

Quasi-inversion of quantum and classical channels in finite dimensions

Fereshte Shahbeigi^{1,2}, Koorosh Sadri¹, Morteza Moradi¹, Karol Życzkowski^{2,3},
and Vahid Karimipour¹

¹Department of Physics, Sharif University of Technology, Tehran, Iran

²Center for Theoretical Physics, Polish Academy of Sciences, 02-668
Warszawa, Poland

³Institute of Theoretical Physics, Uniwersytet Jagielloński, 30-348 Kraków,
Poland

June 14, 2021

Abstract

We introduce the concept of quasi-inverse of quantum and classical channels, prove general properties of these inverses and determine them for a large class of channels acting in an arbitrary finite dimension. Therefore we extend the previous results of [1] to arbitrary dimensional channels and to the classical domain. We demonstrate how application of the proposed scheme can increase on the average the fidelity between a given random pure state and its image transformed by the quantum channel followed by its quasi-inversion.

1 Introduction

It is generally understood that quantum resources make significant improvements over the classical ones in most of the information processing tasks [2]. However, these resources are usually fragile under the noise caused by inevitable interactions with environment which may drastically neutralize the quantum advantage mentioned above. Stated in other words, an open quantum system is usually interacting with an environment, so its dynamics cannot be described by a unitary evolution, $\rho \longrightarrow U\rho U^\dagger$. This unitary dynamics is an idealization which almost never occurs in reality. There are always inevitable and unknown couplings with the environment which destroy the coherence, decrease the purity of the state, and deteriorate information encoded into a quantum system [3, 4]. One of the central results in the quantum theory is that a general non-unitary dynamics of an open quantum system can be characterized by operators acting entirely within the quantum system [5, 6]. The latter dynamics has long been studied and by now there is an extensive literature on the subject. The simplest way to describe a non-unitary dynamics is to use the Kraus form of a channel acting on a density matrix ρ of order d ,

$$\rho \longrightarrow \mathcal{E}(\rho) = \sum_{\alpha} K_{\alpha} \rho K_{\alpha}^{\dagger}, \quad (1)$$

which can be interpreted as a generalization of the unitary dynamics. The standard unitarity condition $U^\dagger U = I$ has been replaced above by the identity resolution, $\sum_{\alpha=1}^r K_{\alpha}^{\dagger} K_{\alpha} = I$. Here

K_α denotes a Kraus operator of size d . The number of these operators, r , may vary, but it is always possible to find representations with $r \leq d^2$. Any map of this class is Completely Positive and Trace-preserving (CTP) and is called a quantum operation or a quantum channel [5, 6]. It captures the effect of errors (noise, decoherence and dissipation) in a quantum system caused by interaction with the environment.

It is easy to see that any unitary evolution can be explicitly inverted by replacing U with U^{-1} , so one can get back the original state by turning the dynamics backward. Even if we set aside practical considerations for reducing the effect of noise and errors, since after all there exist error correcting codes and other methods [7–16] for dealing with these issues, it is a mathematical curiosity to ask, whether a general quantum channel can be inverted too [17, 18]. It is however an established fact that a quantum channel can be exactly inverted only if it is a unitary transformation.

By inversion of a given quantum channel we mean here using another quantum channel, which is physically possible. One may therefore ask, whether a quantum channel can be quasi-inverted, in the sense that another quantum channel \mathcal{E}^{qi} exists, such that $\mathcal{E}^{qi} \circ \mathcal{E}$ is as close to the identity map as possible. This was the approach taken in [1], where the case of qubit channels was studied in detail. As qubit channels are completely characterized and classified in [19, 20], the authors of [1] managed to find the quasi-inverse of all qubit channels. It was shown that the quasi-inverse of every quantum channel, except for a measure zero set, is nothing but a suitably defined unitary map [1]. An explicit formula for deriving this unitary map was also derived.

Slightly related issue was earlier considered by Koenig et al. [21] who analyzed the following problem: given a bipartite quantum state ρ_{AB} one wishes to convert it as closely as possible to a maximally entangled state by applying a quantum channel only on the system B . Due to Jamiolkowski isomorphism any quantum channel acting on a d -dimensional system can be treated as a bipartite state ρ_{AB} of size d^2 – see [22]. Thus the maximal fidelity optimized over all possible channels can be related to the conditional min-entropy of the state ρ_{AB} . A similar approach was later used by Chiribella and Ebler [23], who demonstrated that an unknown unitary channel can be optimally inverted with an average fidelity of $1/d^2$.

Extension of results obtained in [1] for single-qubit channels for higher dimensions is not straightforward, as very little is known concerning the structure of the set of channels acting in dimensions $d \geq 3$. In addition to the quartic increase of the number of parameters with the dimension d , which defies any kind of geometrical picture for these channels, certain important and simplifying theorems which hold for qubit channels do not extend to the higher dimensional case. For instance, for $d = 2$ any unital channel, which leaves the maximally mixed state invariant, $\mathcal{E}(I) = I$ belongs to the class of mixed unitary channels, so it can be represented as a mixture of unitary operations [24]. This property does not hold for general d [25, 26] and already for $d \geq 3$ there exists the unital channel of Landau and Streater, which is not mixed unitary [24], see the recent study [27] for further information on this map. This is also related to another important difference which concerns the extreme points of a convex set, i.e. those points which cannot be written as convex combination of other points of the set. While the extreme points of the convex set of unital qubit channels are unitary maps, this is no longer the case for higher dimensional channels.

All these properties make the study of quasi-inverse of quantum channels a non-trivial task. Nevertheless, we obtain here some general results on higher dimensional channels and their

quasi-inverses and substantiate these results by several examples of general families of d -dimensional channels. We provide upper and lower bounds for the performance of a channel after it is compensated by its quasi-inverse, i.e. $\mathcal{E}^{qi} \circ \mathcal{E}$. In particular, we show that the quasi-inverse of a channel of an arbitrary dimension d need not be a unitary map. We also study a class of self-quasi-inverse quantum channels including the interesting case of Landau-Streater [24].

In the second part of this work we study an analogous problem posed for classical channels, namely for stochastic matrices which map the set of probability vectors into itself. This question, left open in [1], is of a special interest in view of the correspondence between quantum channels and their classical counterparts [28]. We show that several results originally formulated for the quantum domain find their natural parallels in the classical scenario.

The structure of this paper is as follows: In Section 2, we collect the preliminary ingredients, in Section 3 we define the quasi-inverse channel whose general properties and examples are respectively explored in Sections 4 and 5. These studies are extended to the classical domain in Sections 6 and 7. We conclude the paper with a discussion. Some of the detailed calculations and proofs are collected in the Appendices.

2 Preliminaries

Let H_d be a complex d -dimensional Hilbert space for which we choose the computational basis $\{|m\rangle, m = 1, \dots, d\}$. Let $L(H_d)$ be the space of linear operators on this Hilbert space. The state of a d -dimensional quantum system (a qudit) is described by a density matrix ρ which is a positive operator of unit trace acting on this Hilbert space. The space of all density matrices is denoted by $D(H_d)$. This is a convex subset of $L(H_d)$. Any point of this convex set is described by $d^2 - 1$ real parameters. Any linear map $A \in L(H_d)$ can be uniquely mapped to a vector $|A\rangle \in H_d \otimes H_d$ in the form

$$|A\rangle = d(A \otimes I)|\phi^+\rangle, \quad (2)$$

where $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_i |i, i\rangle$ is a maximally entangled state. This is called vectorization of a matrix $A = \sum_{i,j} A_{i,j} |i\rangle\langle j|$ into a vector $|A\rangle = \sum_{i,j} A_{i,j} |i, j\rangle$, with the correspondence between the inner products:

$$\text{Tr}(A^\dagger B) = \langle A|B\rangle. \quad (3)$$

Taking $d^2 - 1$ traceless and Hermitian matrices Γ_i along with identity I as a complete basis, one can write any density matrix as

$$\rho = \frac{1}{d} \left(I + \sum_{i=1}^{d^2-1} r_i \Gamma_i \right) = \frac{1}{d} (I + \mathbf{r} \cdot \mathbf{\Gamma}) \quad (4)$$

We normalize Γ_i matrices to satisfy

$$\text{Tr}(\Gamma_i \Gamma_j) = d(d-1) \delta_{ij}, \quad (5)$$

a concrete choice for this basis is given in Appendix A. We will then have

$$\text{Tr}(\rho^2) = \frac{1}{d} [1 + (d-1) \mathbf{r} \cdot \mathbf{r}]. \quad (6)$$

The $(d^2 - 1)$ -dimensional vector $\mathbf{r} = (x_1, x_2, \dots, x_{d^2-1})^T$ is a real vector called the generalized Bloch vector. In contrast to the qubit case $d = 2$, the convex set of physical states (positive matrices with unit trace) is no longer a unit ball. In fact the geometry of this convex set can be quite complicated and only partial facts are known about low dimensions, i.e. for $d = 3$ [22, 29]. The set of pure states $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle = \sum_{i=1}^d \psi_i |i\rangle$ is a sphere of dimension $2d - 2$ which we denote by S_{2d-2} . On the other hand, any pure state has the property $\text{Tr}(\rho^2) = 1$ which in view of Eqs. (5) and (6) is equivalent to $\mathbf{r} \cdot \mathbf{r} = 1$. This is a sphere of dimension S_{d^2-2} which for $d > 2$ has a higher dimension than the set of pure states. Hence the set of pure states is a subset of this larger sphere. This larger sphere contains other points which are not even states at all, see Appendix A for concrete examples. The necessary and sufficient condition for the Bloch vector to describe a pure quantum state can be found in [30]. The necessary condition for the Bloch vector to produce a general mixed state is provided in [31].

Consider now a quantum channel \mathcal{E} represented by its Kraus operators K_α acting on states of dimension d through Eq. (1). The correspondence (2) leads to the following representation of channels by matrices acting on vectorized form of density matrices,

$$\mathcal{E}(\rho) = \sum_{\alpha} K_{\alpha} \rho K_{\alpha}^{\dagger} \quad \longrightarrow \quad |\mathcal{E}(\rho)\rangle = \Phi_{\mathcal{E}} |\rho\rangle \quad (7)$$

where

$$\Phi_{\mathcal{E}} = \sum_{\alpha} K_{\alpha} \otimes K_{\alpha}^*, \quad (8)$$

is called the superoperator of the map \mathcal{E} [22] in which $*$ denotes complex conjugation. Although a quantum channel has many different sets of Kraus operators, connected by $L_{\beta} = \sum_{\alpha} U_{\alpha,\beta} K_{\alpha}$, where U is a unitary matrix, it is straightforward to see that Φ is unique. It also has the property that

$$\Phi_{\mathcal{E} \circ \mathcal{E}'} = \Phi_{\mathcal{E}} \Phi_{\mathcal{E}'}.$$

The quantum channel \mathcal{E} induces an affine transformation on the generalized Bloch vector \mathbf{r}

$$\mathbf{r} \longrightarrow \mathbf{r}' = M\mathbf{r} + \mathbf{t} \quad (9)$$

where

$$M_{ij} = \frac{1}{d(d-1)} \text{Tr}(\Gamma_i \mathcal{E}(\Gamma_j)) \quad \text{and} \quad t_i = \frac{1}{d(d-1)} \text{Tr}(\Gamma_i \mathcal{E}(I)). \quad (10)$$

The matrix M of order $d^2 - 1$ is called the distortion matrix. Writing the basis $\{I, \Gamma_i\}$ in vectorized form $\{|I\rangle, |\Gamma_i\rangle\}$ which are vectors of dimension d^2 , we have the normalization condition

$$\langle I|I\rangle = d, \quad \langle I|\Gamma_i\rangle = \langle \Gamma_i|I\rangle = 0, \quad \langle \Gamma_i|\Gamma_j\rangle = d(d-1)\delta_{i,j}. \quad (11)$$

The d^2 dimensional vector $|\rho\rangle$ can be written in terms of the normalized basis vectors $|\tilde{I}\rangle = \frac{1}{\sqrt{d}}|I\rangle$ and $|\tilde{\Gamma}_i\rangle = \frac{1}{\sqrt{d(d-1)}}|\Gamma_i\rangle$, in the symbolic form

$$|\rho\rangle = \frac{1}{\sqrt{d}} \begin{pmatrix} 1 \\ \sqrt{d-1} \mathbf{r} \end{pmatrix}, \quad (12)$$

where the first component is the coefficient of $|\tilde{I}\rangle$ and the second component encapsulates the coefficients of $|\tilde{\Gamma}_i\rangle$ as a vector. The quantum channel \mathcal{E} turns this into the vector

$$|\rho'\rangle = \frac{1}{\sqrt{d}} \begin{pmatrix} 1 \\ \sqrt{d-1} M\mathbf{r} + \mathbf{t} \end{pmatrix}. \quad (13)$$

This means that in this basis the superoperator can be written as

$$\Phi_{\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ \sqrt{d-1} \mathbf{t} & M \end{pmatrix}. \quad (14)$$

In the case of unital channels the translation vector vanishes, $\mathbf{t} = 0$. If the entire matrix M also vanishes the corresponding map Φ_* represents the completely depolarizing channel, which sends any state ρ into the maximally mixed state, $\mathcal{E}_*(\rho) = I/d$.

The above form of the map Φ , which represents the evolution of the Bloch vector \mathbf{r} , is also called its *Liouville representation* [32]. On the other hand, it can be interpreted as the Fano form [33] of the bi-partite state representing the map in the Choi-Jamiołkowski isomorphism. Such a state, proportional to the Choi matrix,

$$C_{\mathcal{E}} := d(\mathcal{E} \otimes I)(|\phi^+\rangle\langle\phi^+|) = \sum_{i,j} \mathcal{E}(|i\rangle\langle j|) \otimes |i\rangle\langle j|, \quad (15)$$

forms a positive operator with $\text{Tr}(C_{\mathcal{E}}) = d$. It is related to the superoperator of the channel through the simple relation [22]

$$C_{\mathcal{E}} = \Phi_{\mathcal{E}}^R, \quad (16)$$

where A^R denotes the following reshuffling of the entries of a matrix A of order d^2

$$A_{ij,kl}^R := A_{ik,jl}. \quad (17)$$

2.1 Average fidelity of a channel

The performance of a quantum channel \mathcal{E} can be studied in different ways. Among these we choose the input-output fidelity averaged over a uniform distribution of pure states, $\overline{F}(\mathcal{E})$, and the entanglement fidelity $F_E(\mathcal{E})$ [34]. These are respectively defined as follows:

$$\overline{F}(\mathcal{E}) = \int_{S_{2d-2}} d\psi \langle \psi | \mathcal{E}(|\psi\rangle\langle\psi|) | \psi \rangle, \quad (18)$$

where $\int d\psi = 1$ and $d\psi = d(U\psi)$ for any unitary, and

$$F_E(\mathcal{E}) = \langle \phi^+ | \mathcal{E} \otimes I(|\phi^+\rangle\langle\phi^+|) | \phi^+ \rangle. \quad (19)$$

The average fidelity measures how much the input and output states are similar to each other and the entanglement fidelity measures how much a maximally entangled state is affected if the channel \mathcal{E} acts on one part of this state. As we will see the two quantities are related to each other in a simple way. Below we calculate these quantities by different methods and each method sheds light on these quantities from a different angle.

First consider the average fidelity where it can be written as

$$\begin{aligned} \overline{F}(\mathcal{E}) &= \sum_{\alpha} \int_{S_{2d-2}} d\psi \langle \psi | K_{\alpha} | \psi \rangle \langle \psi | K_{\alpha}^{\dagger} | \psi \rangle \\ &= \sum_{\alpha} \int_{S_{2d-2}} d\psi \langle \psi | K_{\alpha} | \psi \rangle \langle \psi^* | K_{\alpha}^* | \psi^* \rangle = \text{Tr}(L\Phi_{\mathcal{E}}), \end{aligned} \quad (20)$$

where

$$L := \int_{S_{2d-2}} d\psi |\psi\rangle\langle\psi| \otimes |\psi^*\rangle\langle\psi^*|, \quad (21)$$

is the isotropic state and $\Phi_{\mathcal{E}}$ is defined in Eq. (8). Since L is an isotropic operator in the sense that $[L, U \otimes U^*] = 0 \quad \forall U$, it can be written in the form [35],

$$L = \frac{1}{d(d+1)} (I + \sum_{i,j} |i, i\rangle\langle j, j|). \quad (22)$$

Inserting (22) in (20) and using (16) and (17) one finds the following formula for the average fidelity

$$\overline{F}(\mathcal{E}) = \frac{1}{d(d+1)} [d + \text{Tr} \Phi_{\mathcal{E}}]. \quad (23)$$

More details on these calculations will be given in Appendix B. From (8) and (14) we see that this can also be written as $\overline{F}(\mathcal{E}) = \frac{1}{d(d+1)} [d + \sum_{\alpha} |\text{Tr} K_{\alpha}|^2]$ and $\overline{F}(\mathcal{E}) = \frac{1}{d(d+1)} [d + 1 + \text{Tr} M]$.

Entanglement fidelity is derived in a simpler way, e.g. by direct expansion of the state $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_i |i, i\rangle$, and hence one finds

$$F_E(\mathcal{E}) = \frac{1}{d} \langle \phi^+ | \sum_{\alpha} K_{\alpha} |i\rangle\langle j| K_{\alpha}^{\dagger} \otimes |i\rangle\langle j| \phi^+ \rangle = \frac{1}{d^2} \sum_{\alpha} |\text{Tr} K_{\alpha}|^2. \quad (24)$$

Therefore, we find the simple relation [36]

$$\overline{F} = \frac{1}{d+1} (1 + d F_E(\mathcal{E})), \quad (25)$$

showing that quantities (18) and (19) are monotone with respect to each other.

3 Definition of the quasi-inverse channel

In the quantum communication context, it is usually the case that Alice (the information source), generates a message state ρ_M from a distribution corresponding to her language alphabet which is known to Bob (the receiver). Let us denote by μ this probability measure over $D(H_d)$, the set of all quantum states. The state is then fed into a quantum channel \mathcal{E} to reach Bob. The received state is denoted by ρ_R . Assuming that the characteristics of the channel are well known to Bob, and before performing any quantum error correction, he may want to pass the received state through some other channel in order to process it to a state more *similar* to the one Alice has sent. This latter channel clearly cannot depend on ρ_M since Bob does not know it. It only depends on the channel \mathcal{E} and the probability measure μ . The second channel is applied to somehow invert the action of \mathcal{E} , therefore, it is natural to call it the quasi-inverse of \mathcal{E} [1] and it is expected that this channel does this inversion in the best possible way, given the constraint of being a CPT map. Therefore, we have the formal definition of the quasi-inverse:

Definition 1. *The quasi-inverse of a channel \mathcal{E} is defined as a CPT map \mathcal{E}^{qi} fulfilling the following condition [1]*

$$\overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) \geq \overline{F}(\mathcal{E}' \circ \mathcal{E}) \quad \forall \mathcal{E}', \quad (26)$$

where \overline{F} is the average of a proper fidelity function between the output state and the input pure state.

Remark 1. *As we will see in the sequel, the quasi-inverse of a quantum channel is unique except when the channel falls on a set of measure zero. In the simplest case of unital qubit channels, in which every channel is unitarily equivalent to a Pauli channel, $\mathcal{E}(\rho) = \sum_{i=0}^3 p_i \sigma_i \rho \sigma_i$, only the channels with two or more equal weights p_i have non-unique quasi-inverses.*

We now prove our first theorem whose validity does not depend on the specific form of the fidelity function, nor on the input state restricted to be pure, rather it depends on two very general properties of the fidelity measure, as described in the proof.

Theorem 1. *For any proper similarity function F , probability measure μ and quantum channel \mathcal{E} , a quasi-inverse \mathcal{E}^{qi} can be found on the boundary of the set of allowed channels.*

Proof. A proper similarity measure F should satisfy the following two conditions: **(a)**

$$F(\rho, \rho') \leq F(\rho, \rho) \quad \forall \rho, \rho' \in D(H_d); \quad (27)$$

the state most similar to a state ρ is the state ρ itself, and **(b)** concavity:

$$F(\rho, \alpha\rho_1 + (1 - \alpha)\rho_2) \geq \alpha F(\rho, \rho_1) + (1 - \alpha)F(\rho, \rho_2). \quad (28)$$

One such measure is the well-known fidelity, but the following argument is independent of the particular form of this measure. We proceed with a proof by contradiction.

Let us denote the set of all possible quantum channels on H_d by \mathcal{C}_d . This is a convex set. Assume the quasi-inverse for the channel $\mathcal{E} \equiv (M, \mathbf{t})$, defined by Definition 1, is a quantum channel \mathcal{E}' in the interior of \mathcal{C}_d , figure (1). Now take the inverse of the above affine map which is given by $(M^{-1}, -M^{-1}\mathbf{t})$. This affine map may not correspond to a legitimate quantum channel and may be outside \mathcal{C}_d . Denote it by \mathcal{E}^{-1} . However for small enough ε , the channel $(1 - \varepsilon)\mathcal{E}' + \varepsilon\mathcal{E}^{-1}$ is a CPT $\in \mathcal{C}_d$. We now note that this channel performs better in quasi-inverting the channel \mathcal{E} since using (27) and (28) we find

$$\begin{aligned} & \int d\mu F(\rho, [(1 - \varepsilon)\mathcal{E}' + \varepsilon\mathcal{E}^{-1}] \circ \mathcal{E}(\rho)) \geq \\ & (1 - \varepsilon) \int d\mu F(\rho, \mathcal{E}' \circ \mathcal{E}(\rho)) + \varepsilon \int d\mu F(\rho, \rho) = \\ & \int d\mu F(\rho, \mathcal{E}' \circ \mathcal{E}(\rho)) + \varepsilon \left(\int d\mu [F(\rho, \rho) - F(\rho, \mathcal{E}' \circ \mathcal{E}(\rho))] \right) \geq \\ & \int d\mu F(\rho, \mathcal{E}' \circ \mathcal{E}(\rho)). \end{aligned}$$

This means we have found a linear path along which the average fidelity increases or stays constant as we go toward the boundary of \mathcal{C}_d . Clearly, the quasi-inverse has to be in the end of such a path and hence at the boundary of \mathcal{C}_d .

To complete the proof, we need to consider the case of singular channels. This corresponds to $\det M = 0$ and therefore has unit co-dimension. It means for any singular channel \mathcal{E} , we may come up with a sequence of nonsingular channels $\mathcal{E}_n \rightarrow \mathcal{E}$. So far we have proved that

$$\forall \mathcal{E}' \in \text{int } \mathcal{C}_d, \quad \exists \mathcal{E}_n^{qi} \in \partial \mathcal{C}_d$$

such that

$$\langle F(\rho, \mathcal{E}_n^{qi} \circ \mathcal{E}_n(\rho)) \rangle \geq \langle F(\rho, \mathcal{E}' \circ \mathcal{E}_n(\rho)) \rangle,$$

where $\langle X \rangle$ denotes the average of X over μ . Using the continuity of F , implied by its concavity, the proposition is proved for the singular channels as well. Note that in this proof we have not assumed any particular form for the fidelity measure, except the two natural properties (27) and (28), neither we have assumed the average fidelity to be defined only for pure states. \square

Hereafter we assume the standard fidelity measure, $F(\rho, \rho') = \left(\text{Tr} \sqrt{\sqrt{\rho} \rho' \sqrt{\rho}} \right)^2$, of Uhlmann and Jozsa [37, 38], and use the average input-output fidelity (18) for the performance of the

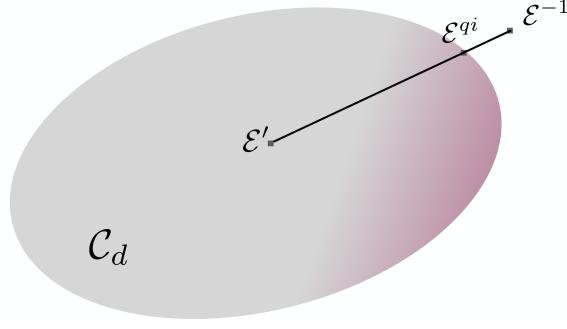


Figure 1: Assume that the requested quasi-inverse of a given channel \mathcal{E} is \mathcal{E}' , located inside the convex set \mathcal{C}_d of all channels. Consider the true inverse map \mathcal{E}^{-1} (defined in the text), which may not belong to \mathcal{C}_d . Then at the line joining \mathcal{E}' and \mathcal{E}^{-1} we always find a legitimate quasi-inverse with a higher average fidelity. Hence the true quasi-inverse \mathcal{E}^{qi} lies on the boundary of the set \mathcal{C}_d of all channels.

channel. In such a case the average value can be compared with the mean fidelity between two random quantum states averaged over the set of all mixed states with an appropriate measure [39].

Under these assumptions, the quasi-inversion is defined as the channel maximizing $\overline{F}(\mathcal{E}' \circ \mathcal{E})$. In view of Eq. (23), the practical method for finding the quasi-inverse is through one of the following maximization problems:

$$\begin{aligned}
 \overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) &= \max_{\Phi'} \frac{1}{d+1} \left(1 + \frac{1}{d} \text{Tr} \Phi' \Phi \right) \\
 &= \max_{\{K'_\beta\}} \frac{1}{d+1} \left(1 + \frac{1}{d} \sum_{\alpha, \beta} |\text{Tr} K'_\beta K_\alpha|^2 \right) \\
 &= \max_{M'} \frac{1}{d} \left(1 + \frac{1}{d+1} \text{Tr} M' M \right).
 \end{aligned} \tag{29}$$

These equations immediately imply that for a given channel, the left and right quasi-inverses are the same, i.e. $\overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) = \overline{F}(\mathcal{E} \circ \mathcal{E}^{qi})$. It is an interesting fact with practical benefit. Either Bob can apply the quasi-inverse after receiving the state or Alice before sending the state to Bob. Furthermore, note that for \mathcal{E} and \mathcal{E}' denoted by (M, \mathbf{t}) and (M', \mathbf{t}') , respectively, their concatenation, $\mathcal{E}' \circ \mathcal{E}$, is represented by $(M'M, M'\mathbf{t} + \mathbf{t}')$. This implies that the translation vector \mathbf{t} , which determines the non-unitality of the channel, plays no role directly in amount of fidelity and fidelity after correction. However, it affects the range of the allowed values of the distortion matrix elements. This is one of the features of quasi-inverse for qubit channels [1], which survives for higher dimensions.

It is worth mentioning that one may find a relation between the input-output fidelity after correction and the conditional min-entropy. The latter quantity for a bipartite state ρ_{AB} is defined as [40]

$$H_{\min}(B|A)_\rho = -\inf_{\sigma_A} \inf_{\lambda} \{ \lambda \in \mathbb{R} \mid \rho_{AB} \leq 2^\lambda (\sigma_A \otimes I) \}, \tag{30}$$

where σ_A is a quantum state. Indeed, it has been proven that [21]

$$2^{-H_{\min}(B|A)_\rho} = d \max_{\mathcal{E}'} \langle \phi^+ | (\mathcal{E}' \otimes I)[\rho_{AB}] | \phi^+ \rangle. \quad (31)$$

Let us assume that ρ_{AB} is the Jamiołkowski state (the normalised form of the Choi matrix (15)) assigned to a quantum channel and denoted by $J_\mathcal{E} = C_\mathcal{E}/d$. Then we get

$$2^{-H_{\min}(B|A)_{J_\mathcal{E}}} = d \max_{\mathcal{E}'} \langle \phi^+ | (\mathcal{E}' \circ \mathcal{E} \otimes I)[|\phi^+\rangle\langle\phi^+|] | \phi^+ \rangle = d \max_{\mathcal{E}'} F_E(\mathcal{E}' \circ \mathcal{E}) = d F_E(\mathcal{E}^{qi} \circ \mathcal{E}). \quad (32)$$

Here $F_E(\mathcal{E}' \circ \mathcal{E})$ is the entanglement fidelity (24) of the composed map $\mathcal{E}' \circ \mathcal{E}$. So the above equation shows that entanglement fidelity after correction (thus input-output fidelity after correction, see (25)) are directly related to the conditional min-entropy of the Choi matrix of the channel.

4 General properties of the quasi-inverse channel

In this section we elaborate on some general properties of the quasi-inverse of quantum channels and discuss the similarities and the crucial differences with the qubit case [1]. Thus far, we have seen in Theorem 1 that for any proper similarity measure the quasi-inverse lies on the boundaries of the set of quantum channels. In what follows, we show that if this similarity measure is linear with respect to quantum channels, we can specify the quasi-inverse in the set of extreme channels, a subset of the boundary points. Recall that a point of a convex set Ω is called extreme if it cannot be written as a convex combination of two other points of Ω . We will use the fact that the set \mathcal{C}_d of quantum channels of dimension d is convex [22] for any d .

Theorem 2. *The quasi-inverse of a quantum channel can always be taken to be an extreme channel.*

Proof. Assume that the quasi-inverse of a channel \mathcal{E} is in the form

$$\mathcal{E}^{qi} = \sum_i \lambda_i \mathcal{E}_i. \quad (33)$$

Let \mathcal{E}_m be the element in the above set for which $\overline{F}(\mathcal{E}_m \circ \mathcal{E}) \geq \overline{F}(\mathcal{E}_i \circ \mathcal{E}) \quad \forall \quad i$. Then using the linearity of the quasi-inverse, we will have

$$\overline{F}(\mathcal{E}_m \circ \mathcal{E}) \geq \overline{F}(\sum_i \lambda_i \mathcal{E}_i \circ \mathcal{E}) = \overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}), \quad (34)$$

which means, according to Definition 1, that the quasi-inverse can always be taken as an extreme channel. \square

The crucial difference between the qubit case and the higher dimensional case is that for qubit channels quasi-inversion is unital and the extreme points of the set of one-qubit unital maps are unitary channels, while this is no longer the case for $d > 2$. Even more than that, not all extreme points of the set of channels are known for $d > 2$, not even for unital channels.

Note that the linearity of $\overline{F}(\mathcal{E}' \circ \mathcal{E})$ over \mathcal{E}' implies that if \mathcal{E}_1^{qi} and \mathcal{E}_2^{qi} are both quasi-inversions of \mathcal{E} , then any convex combination of them, $p\mathcal{E}_1^{qi} + (1-p)\mathcal{E}_2^{qi}$, is the inverse as well. In accordance with the above theorem we arrive at the following result.

Corollary 1. *The quasi-inverse is either unique or an infinite number of them exist where at least two of them are extreme channels.*

An immediate result of Theorem 2 is that the quasi-inversion is not an involution, i.e. $(\mathcal{E}^{qi})^{qi} \neq \mathcal{E}$ for a general \mathcal{E} . It is because a quasi-inverse map is an extreme point. So even if one takes into account the non-uniqueness of quasi-inverse, see Remark 1 and Corollary 1, there are always non-extremal maps which are not quasi-inverse of any other maps.

Proposition 1. *Let \mathcal{E}^{qi} denote the quasi-inverse of \mathcal{E} . Then $(\mathcal{E}^{qi} \circ \mathcal{E})^{qi}$ can be taken to be the identity.*

Proof. The quasi-inverse $(\mathcal{E}^{qi} \circ \mathcal{E})^{qi}$ is the map which maximizes the input-output fidelity

$$\max_{\mathcal{E}''} \overline{F}(\mathcal{E}'' \circ (\mathcal{E}^{qi} \circ \mathcal{E})) \leq \max_{\mathcal{E}'} \overline{F}(\mathcal{E}' \circ \mathcal{E}) = \overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}). \quad (35)$$

This inequality is saturated if we take \mathcal{E}'' equal to the identity map. \square

On the other hand, let $\mathfrak{C}_{\mathcal{E}}$ denote the set of quantum channels for which \mathcal{E} defines the quasi-inverse. We argue that such a set is convex because for any $\mathcal{E}_1, \mathcal{E}_2 \in \mathfrak{C}_{\mathcal{E}}$ we have $p\mathcal{E}_1 + (1-p)\mathcal{E}_2 \in \mathfrak{C}_{\mathcal{E}}$, where $0 \leq p \leq 1$. In this sense, \mathfrak{C}_I is a special convex subset of quantum channels which are not correctable, i.e. the identity map is its quasi-inverse. According to the above proposition, applying the quasi-inversion we actually send a given quantum channel to this special subset since we cannot correct the fidelity afterwards.

As an example consider the tetrahedron of Pauli channels, $\Phi = \sum_i p_i \Phi_i = \sum_{i=0}^3 p_i \sigma_i \otimes \bar{\sigma}_i$ where $\sigma_0 = I$ and other σ_i 's are the Pauli matrices. According to the results of [1], the subset of Pauli channels for which $\forall i \ p_0 \geq p_i$ belongs to \mathfrak{C}_I . This is also a consequence of the result of Example 1 in the next section, since σ_i 's satisfy orthogonality. The set of non-correctable Pauli channels is presented in Fig 2.

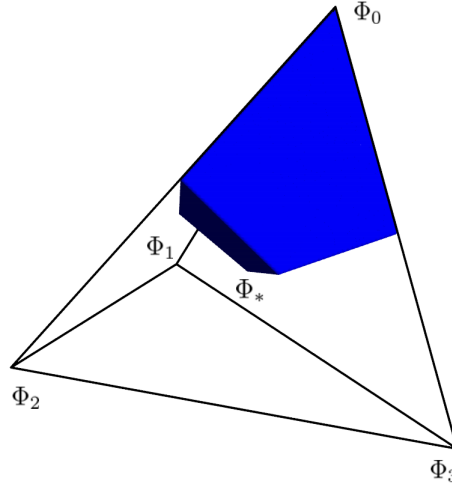


Figure 2: The set of Pauli channels. The blue region shows \mathfrak{C}_I , the convex subset of the channels whose quasi-inversion is the identity map. Φ_* is the center of the tetrahedron defined by $\frac{1}{4} \sum_{i=0}^3 \Phi_i$.

Two unitarily equivalent channels are defined as \mathcal{E}_1 and \mathcal{E}_2 such that $\mathcal{E}_2 = \mathcal{E}_U \circ \mathcal{E}_1 \circ \mathcal{E}_V$ where \mathcal{E}_U and \mathcal{E}_V are two unitary maps. We use Φ_1 and Φ_2 to denote \mathcal{E}_1 and \mathcal{E}_2 . Unitary maps, \mathcal{E}_U

and \mathcal{E}_V , are represented by Φ_U and Φ_V . Hence,

$$\begin{aligned}\text{Tr}(\Phi_2^{qi}\Phi_2) &= \max_{\Phi''} \text{Tr}(\Phi''\Phi_2) = \max_{\Phi''} \text{Tr}(\Phi''\Phi_U\Phi_1\Phi_V) \\ &= \max_{\Phi''} \text{Tr}(\Phi_V\Phi''\Phi_U\Phi_1) = \max_{\Phi'} \text{Tr}(\Phi'\Phi_1) = \text{Tr}(\Phi_1^{qi}\Phi_1),\end{aligned}\quad (36)$$

where we have used the fact that the set of all quantum channels, \mathcal{C}_d , is invariant under unitary transformations. This relation proves the quasi-inverse channels of \mathcal{E}_1 and \mathcal{E}_2 are also unitarily equivalents, $\mathcal{E}_2^{qi} = \mathcal{E}_{V^{-1}} \circ \mathcal{E}_1^{qi} \circ \mathcal{E}_{U^{-1}}$, and they can reach the same amount of fidelity after correction. This fact expands the result of [1] related to unitarily equivalent channels of the form $\mathcal{E}_2 = \mathcal{E}_U \circ \mathcal{E}_1 \circ \mathcal{E}_{U^{-1}}$.

Here a crucial difference between the qubit channels and higher dimensional channels shows up. In the case of qubits, a complete characterization of qubit channels exists and it is known that any qubit channel has the decomposition $\mathcal{E} = \mathcal{E}_U \circ \mathcal{E}_c \circ \mathcal{E}_V$, where \mathcal{E}_c is a canonical map with diagonal distortion matrix $M_c = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$. The signed singular values [22] of the matrix M_c confined inside a tetrahedron Δ whose extreme points are unitary operations $\rho \rightarrow \sigma_i \rho \sigma_i^\dagger$ ($i = 0, 1, 2, 3$). Therefore the task of finding quasi-inverse of any qubit channel is considerably easy compared with higher dimensional channels where such a canonical decomposition does not exist and even if there was, with a presumably diagonal matrix M , we were faced with a highly complex characterization of the vector λ . It is known that the structure of the convex set of higher dimensional channels is far more complex than that of a simple tetrahedron, and in particular it is known that the extreme points of this set are not necessarily unitary channels. A well-known counter example is the Landau-Streater channel [24] which will be discussed in Section 5. Let us now put general bounds on the improved average fidelity.

Theorem 3. *The average input-output fidelity of a channel after correction has the following upper and lower bounds:*

$$\frac{f+1}{d+1} \leq \overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) \leq \frac{p_m+1}{d+1}, \quad (37)$$

where $f = \max_{|\beta\rangle} \langle \beta | C_{\mathcal{E}} | \beta \rangle$ is the fully entangled fraction of the Choi matrix $C_{\mathcal{E}}$ of the channel \mathcal{E} and $|\beta\rangle$ denotes a maximally entangled state, while p_m is the maximal eigenvalue of $C_{\mathcal{E}}$.

Before proceeding with the proof let us mention in view of Eq. (25) and the definition of fully entangled fraction, the lower bound in above equation is actually an upper bound for the input-output fidelity before we correct it with quasi-inverse map, $\overline{F}(\mathcal{E})$.

Proof. To prove the upper bound, we note that

$$\begin{aligned}\text{Tr}(\Phi_{\mathcal{E}'}\Phi_{\mathcal{E}}) &= \sum (\Phi_{\mathcal{E}'}^*)_{ij,kl} (\Phi_{\mathcal{E}})_{kl,ij} = \sum (C_{\mathcal{E}'})_{ik,jl} (C_{\mathcal{E}})_{ki,lj} \\ &= \sum (SC_{\mathcal{E}'}S)_{ki,lj} (C_{\mathcal{E}})_{ki,lj} = \text{Tr}(\tilde{C}_{\mathcal{E}'}C_{\mathcal{E}}) \\ &\leq p_m \text{Tr}(\tilde{C}_{\mathcal{E}'}) = p_m \text{Tr}(C_{\mathcal{E}'}) = dp_m,\end{aligned}\quad (38)$$

where $S = \sum_{i,j} |i,j\rangle\langle j,i|$ is the swap operator, $\tilde{C}_{\mathcal{E}'} = (SC_{\mathcal{E}'}S)^T$, and p_m is the largest eigenvalue of the Choi matrix of the channel \mathcal{E} . In writing these equations we have used Eqs. (16) and (17), and the fact that $C_{\mathcal{E}}$ and $\tilde{C}_{\mathcal{E}'}$ are Hermitian and positive matrices. The above inequality leads to the following upper bound for the improved average fidelity

$$\overline{F}(\mathcal{E}) \leq \frac{p_m+1}{d+1}. \quad (39)$$

In order to obtain a lower bound, we note that in view of (26)

$$\text{Tr}(\tilde{C}_{\mathcal{E}'} C_{\mathcal{E}}) = \text{Tr}(\Phi_{\mathcal{E}'} \Phi_{\mathcal{E}}) \leq \text{Tr}(\Phi_{\mathcal{E}}^{qi} \Phi_{\mathcal{E}}) \quad \forall C_{\mathcal{E}'}, \quad (40)$$

and choose for $\tilde{C}_{\mathcal{E}'}$ to be equal to $d|\beta\rangle\langle\beta|$, where $|\beta\rangle$ is a maximally entangled state. This gives the lower bound. \square

One of the main differences between the qubit case and the higher dimensional channels is that the singular values of M in the qubit case are always less than or equal to one. As a result, the Bloch vector cannot be stretched by applying the distortion matrix M in the qubit case, while as we will show in an explicit example, this is not necessarily the case for higher dimensional channels.

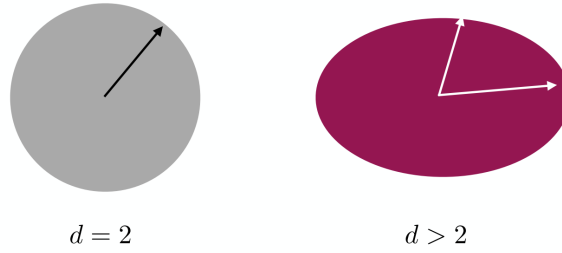


Figure 3: (color online) Left: In $d = 2$ every Bloch vector \mathbf{r} is rotated or shrunk when it is acted on by the distortion matrix of a qubit channel. Right: for $d > 2$, the distortion matrix of a quantum channel can also stretch the Bloch vector.

There are certain channels whose distortion matrix M in Eq. (14) can stretch the generalized Bloch vector \mathbf{r} . This is due to the non-spherical shape of the space of quantum states in higher dimensions, see Fig. 3. We will elaborate on this point and its consequences in Appendix C.

Proposition 2. *The quasi-inverse of the tensor product of quantum channels is the tensor product of the quasi-inverses, i.e. if $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2$ then $\mathcal{E}^{qi} = \mathcal{E}_1^{qi} \otimes \mathcal{E}_2^{qi}$.*

Proof. Let $\rho_{AA'BB'} = \rho_{AB} \otimes \rho_{A'B'}$, then one can show the min-entropy (30) is additive [21], i.e. $H_{\min}(BB'|AA') = H_{\min}(B|A) + H_{\min}(B'|A')$. The proof of the proposition then becomes straightforward using Eq. (32) and the fact that for the tensor product of quantum channels, the Jamiołkowski state is in the tensor product shape. \square

Applying this proposition we show that for N copies of a given quantum channel in a multipartite setting

$$\max_{\mathcal{E}'} \bar{F}(\mathcal{E}' \circ \mathcal{E}^{\otimes N}) = \bar{F}((\mathcal{E}^{qi})^{\otimes N} \circ \mathcal{E}^{\otimes N}) \leq (\bar{F}(\mathcal{E}^{qi} \circ \mathcal{E}))^N. \quad (41)$$

To prove this equation we note that for any quantum channel \mathcal{Q}

$$\text{Tr} \Phi_{\mathcal{Q}^{\otimes N}} = (\text{Tr} \Phi_{\mathcal{Q}})^N. \quad (42)$$

Let us apply $\text{Tr}(\Phi_{\mathcal{Q}}) = x$ for simplicity of notation. In view of Eq. (23), we prove whenever $1 \leq x \leq d^2$ the following inequality holds:

$$\bar{F}(\mathcal{Q}^{\otimes N}) = \frac{d^N + x^N}{d^N(d^N + 1)} \leq \left(\frac{d + x}{d(d + 1)} \right)^N = (\bar{F}(\mathcal{Q}))^N. \quad (43)$$

To prove this relation we note that both sides are increasing functions of x . When $x = 1$ or $x = d^2$ both sides are equal. Comparing their derivatives with respect to x at $x = 1$, we see the right hand side function grows faster at $x = 1$ which proves the inequality (43). To prove Eq. (41), it remains to show that $1 \leq \text{Tr}(\Phi_{\mathcal{E}^{qi} \circ \mathcal{E}}) \leq d^2$ for the composition of any quantum channel and its quasi-inverse. The later is, however, a consequence of Eq. (37) and the fact that fully entangled fraction of the Choi matrix is greater than $1/d$.

5 Examples

Taking into account the upper and lower bounds given in (37) and the general theorems of the previous section, in this section we will consider a few classes of examples, as the optimization problem (29) cannot be solved analytically in the general case. In the case of single-qubit systems a complete classification of quantum channels [19, 20] leads to an explicit description of their quasi-inverse which can be unitary [1]. For higher dimensional channels, the quasi-inverse may not necessarily be a unitary map. Thus identifying the quasi-inverse is related to finding the extreme points of the set \mathcal{C}_d of quantum channels, which remains an open problem.

Example 1 (Mixed unitary channels with orthogonal unitaries). A mixed unitary channel is defined [41, 42] by

$$\mathcal{E}(\rho) = \sum_{\alpha=1}^r q_{\alpha} V_{\alpha} \rho V_{\alpha}^{\dagger}, \quad (44)$$

where $\{V_{\alpha}\}_{\alpha=1}^r$ is an arbitrary set of r unitary transformations. We restrict ourselves to the case where the unitaries are orthogonal with respect to the Hilbert-Schmidt scalar product, $\text{Tr}(V_{\alpha}^{\dagger} V_{\beta}) = d \delta_{\alpha, \beta}$. In analogy to the construction of an approximate time reversal proposed in [17], the quasi-inverse of \mathcal{E} is then the unitary channel $\mathcal{V}_m^{\dagger} : \rho \rightarrow V_m^{\dagger} \rho V_m$, where V_m corresponds to the largest weight q_m in the mixture (44). To see this note that the Choi matrix of this channel reads

$$C_{\mathcal{E}} = \sum_{\alpha} q_{\alpha} |V_{\alpha}\rangle \langle V_{\alpha}|, \quad (45)$$

where we have used the correspondence (2). In view of the orthogonality of the vectors $|V_{\alpha}\rangle$, this is then the spectral decomposition of the Choi matrix $C_{\mathcal{E}}$ with eigenvalues equal to $p_{\alpha} = dq_{\alpha}$ (note that $|V_{\alpha}\rangle$ is not normalized). Let $q_m = p_m/d$ be the largest of coefficients in (44). In view of the diagonal nature of the Choi matrix, p_m is the largest eigenvalue of the Choi matrix. Moreover, by taking $|\beta\rangle = |V_m\rangle$, one sees that $f = p_m$. So the upper and lower bounds of (37) coincide and we find the quasi-inverse is the unitary map induced by V_m^{\dagger} , i.e. $\mathcal{E}^{qi} = \mathcal{E}_{V_m^{\dagger}}$, with

$$\overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) = \frac{p_m + 1}{d + 1}. \quad (46)$$

Let us emphasize again that this result is valid only for mixture of unitary maps corresponding to unitary matrices mutually orthogonal in sense of the Hilbert-Schmidt scalar product. If this assumption is not satisfied the quasi-inverse is not the inverse of one of the unitaries. As the Example 5 shows.

Example 2 (Uniform mixture of orthogonal conjugations). In the following two examples, we will bring some quantum channels which are self-inverse.

Theorem 4. *Let \mathcal{E} be a unital channel obtained by the uniform mixture of conjugation (not necessarily unitary ones) by matrices which are orthogonal to each other. Such a channel is*

specified by

$$\mathcal{E}(\rho) = \sum_{\alpha=1}^q X_{\alpha} \rho X_{\alpha}^{\dagger}, \quad (47)$$

where

$$\sum_{\alpha=1}^q X_{\alpha}^{\dagger} X_{\alpha} = \sum_{\alpha=1}^q X_{\alpha} X_{\alpha}^{\dagger} = I_d, \quad \text{and} \quad \text{Tr}(X_{\alpha}^{\dagger} X_{\beta}) = \frac{d}{q} \delta_{\alpha\beta}. \quad (48)$$

We now show that the quasi-inverse of this map is given by its dual, i.e. $\mathcal{E}^{qi}(\rho) = \sum X_{\alpha}^{\dagger} \rho X_{\alpha}$.

Proof. The Choi matrix of the channel \mathcal{E} (47) is

$$C_{\mathcal{E}} = \sum_{\alpha} |X_{\alpha}\rangle\langle X_{\alpha}|, \quad (49)$$

where $|X_{\alpha}\rangle$ is based on the correspondence of Eq. (2) and it fulfills $\langle X_{\alpha}|X_{\beta}\rangle = \frac{d}{q} \delta_{\alpha\beta}$. Thus we find that the above equation is indeed the spectral decomposition of the degenerated Choi matrix with the q -fold degenerated largest eigenvalue equal to $p_m = \frac{d}{q}$. The superoperator of this channel is given by

$$\Phi_{\mathcal{E}} = \sum_{\alpha=1}^q X_{\alpha} \otimes X_{\alpha}^*. \quad (50)$$

Note that since \mathcal{E} is assumed to be unital, $\Phi_{\mathcal{E}}^{\dagger}$ is also a valid quantum channel corresponding to the dual of \mathcal{E} . Composing $\Phi_{\mathcal{E}}$ and $\Phi_{\mathcal{E}}^{\dagger}$, we get

$$\text{Tr}(\Phi_{\mathcal{E}}^{\dagger} \Phi_{\mathcal{E}}) = \sum_{\alpha, \beta=1}^q |\text{Tr}(X_{\alpha}^{\dagger} X_{\beta})|^2 = \sum_{\alpha=1}^q \left(\frac{d}{q}\right)^2 = \frac{d^2}{q} = dp_m. \quad (51)$$

So by such a composition the upper bound of Eq. (37) is obtained, which completes the proof. \square

The fidelity after correction for the channel (47) then reads $\overline{F}(\mathcal{E}^{\dagger} \circ \mathcal{E}) = \frac{d+q}{q(d+1)}$, while for the case with $\text{Tr}(X_{\alpha}) = 0$ we have $\overline{F}(\mathcal{E}) = \frac{1}{d+1}$ before applying quasi-inversion which admits significant improvement specially in higher dimensions and when q is not large. Note that for $q = 1$ the average fidelity after correction is 1, as it is expected. Moreover, it is obvious that if the operators X_{α} are Hermitian, \mathcal{E} is its own quasi-inverse. Two explicit examples for this case are provided in what follows.

Example 2.1 (Landau-Streater (LS) channel). Consider the Landau-Streater channel [24],

$$\mathcal{E}(\rho) = \frac{1}{j(j+1)} (J_1 \rho J_1 + J_2 \rho J_2 + J_3 \rho J_3), \quad (52)$$

where J_i are the Hermitian generators of $SU(2)$ in its irreducible representation in dimension $d = 2j+1$ and they satisfy $\text{Tr}(J_m^{\dagger} J_n) = \frac{1}{3} j(j+1)(2j+1) \delta_{mn}$. It is clear then that LS channel is a special case of Eq. (47) with $q = 3$. It is worth mentioning that the Landau-Streater channel (52) is an extreme point of \mathcal{C}_d when $d \geq 3$ [24]. However for $d = 2$, the Landau-Streater channel is not an extreme point of \mathcal{C}_2 , so in view of Theorem 2 there exist several quasi-inverse channels.

Example 2.2 (Generalized Landau-Streater channel). Let G be a Lie group with dimension n , with Lie algebra generators A_{α} where $1 \leq \alpha \leq n$. Let D_{μ} be an irreducible unitary representation of the Lie algebra. We define the generalized Landau-Streater channel as

$$\mathcal{E}_{\mu}(\rho) = \frac{1}{c_{\mu}} \sum_{\alpha=1}^n D_{\mu}(A_{\alpha}) \rho D_{\mu}(A_{\alpha})^{\dagger}, \quad (53)$$

where c_μ is the value of the second Casimir operator in this representation

$$c_\mu I = \sum_{\alpha=1}^n D_\mu(A_\alpha)^2. \quad (54)$$

The generators can be made orthogonal so that

$$\text{Tr}[D_\mu(A_\alpha)D_\mu(A_\beta)] = \frac{c_\mu d_\mu}{n}. \quad (55)$$

Such a channel satisfies the assumption of Theorem 4 and is hence its own inverse.

Example 3 (The transverse-depolarizing and depolarizing channel). The transverse-depolarizing channel is defined as $\mathcal{E}_w^{td}(\rho) = (1-w)\rho^T + \frac{w}{d}\text{Tr}(\rho)I_d$, where $\frac{d}{d+1} \leq w \leq \frac{d}{d-1}$ to satisfy complete positivity. The superoperator of this channel is given by:

$$\Phi_w^{td} = (1-w)S + w|\phi^+\rangle\langle\phi^+|, \quad (56)$$

and the Choi matrix is equal to

$$C_w^{td} = (1-w)S + \frac{w}{d}I \otimes I. \quad (57)$$

S is the swap operator introduced in the proof of Theorem 3. It is a Hermitian unitary so its eigenvalues are ± 1 . Thus, the largest eigenvalue of C_w^{td} is

$$\begin{cases} p_m = 1 - w + \frac{w}{d}, & \frac{d}{d+1} \leq w \leq 1, \\ p_m = w - 1 + \frac{w}{d}, & 1 \leq w \leq \frac{d}{d-1}. \end{cases} \quad (58)$$

Let $\mathcal{E}_\pm(\rho) = \frac{1}{d \pm 1}(\text{Tr}(\rho)I_d \pm \rho^T)$. The channel \mathcal{E}_+ is the transverse-depolarizing channel with $w = \frac{d}{d+1}$ and \mathcal{E}_- , also called Werner-Holevo channel [43], is equal to \mathcal{E}_w^{td} for $w = \frac{d}{d-1}$. Indeed, any transverse-depolarizing channel is a convex combination of \mathcal{E}_+ and \mathcal{E}_- . Now it is straightforward to see $\overline{F}(\mathcal{E}_+ \circ \mathcal{E}_w^{td})$ saturates the upper bound of Eq. (37) when $\frac{d}{d+1} \leq w \leq 1$, so \mathcal{E}_+ is the quasi-inverse on this domain. However, we find in this region the equality $\overline{F}(\mathcal{E}_+ \circ \mathcal{E}_w^{td}) = \overline{F}(\mathcal{E}_w^{td})$ holds. This implies that for this range of w the transverse-depolarizing channel is non-correctable. On the other hand, for $1 \leq w \leq \frac{d}{d-1}$, the upper bound of Eq. (37) is achievable by $\mathcal{E}_- \circ \mathcal{E}_w^{td}$ confirming that \mathcal{E}_- is the quasi-inverse in this interval and it can indeed improve the fidelity by $\Delta \overline{F} = \frac{2(w-1)}{d+1}$. The parameter space of the transverse-depolarizing channel is shown in Fig. 4.

We can also consider the depolarizing channel defined as $\mathcal{E}_q^d = (1-q)\rho + \frac{q}{d}\text{Tr}(\rho)I_d$, where q is a probability. The superoperator and the Choi matrix are given by

$$\Phi_q^d = (1-q)I \otimes I + q|\phi^+\rangle\langle\phi^+|, \quad C_q^d = d(1-q)|\phi^+\rangle\langle\phi^+| + \frac{q}{d}I \otimes I. \quad (59)$$

It is now obvious that the largest eigenvalue of the Choi matrix is given by $p_m = d(1-q) + \frac{q}{d}$ which is actually equal to $\text{Tr}(\Phi_q^d)/d$. This implies that quasi-inverse is the identity map. So the depolarizing channel is not correctable.

Example 4 (Covariant Channels). A quantum channel $\mathcal{E} : L(H_d) \longrightarrow L(H_d)$ is called *covariant* with respect to a group G , if the following property holds:

$$\mathcal{E}(U(g)\rho U^\dagger(g)) = V(g)\mathcal{E}(\rho)V^\dagger(g), \quad \forall g \in G. \quad (60)$$

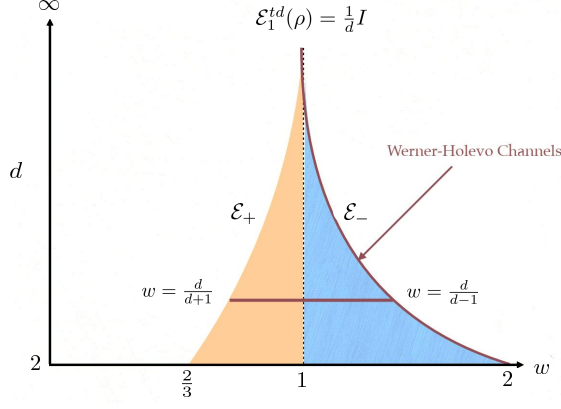


Figure 4: (Color Online) A rough sketch of the parameter space of the transverse-depolarizing channel (56). The colored region defines the valid range of the parameters for this map to be complete positive. The orange part belongs to the set of non-correctable maps, also can be quasi-inverted by \mathcal{E}_+ , the boundary of the orange part. The blue part is the set of correctable channels, invertible by \mathcal{E}_- , the boundary of the blue part. Fixing d , any transverse-depolarizing channel is a convex combination of \mathcal{E}_+ and \mathcal{E}_- .

in which $U(g)$ and $V(g)$ are two not necessarily equivalent representations of g . From the vectorization (2), we find

$$\Phi_{\mathcal{E}} U(g) \otimes U^*(g) = V(g) \otimes V^*(g) \Phi_{\mathcal{E}}. \quad (61)$$

This property has implications for the quasi-inverse. To see this we note that

$$\begin{aligned} \mathcal{E}^{qi} &= \operatorname{argmax}_{\mathcal{E}'} \operatorname{Tr}(\Phi_{\mathcal{E}'} \Phi_{\mathcal{E}}) = \operatorname{argmax}_{\mathcal{E}'} \operatorname{Tr}(\Phi_{\mathcal{E}'} (V^\dagger(g) \otimes V^T(g)) \Phi_{\mathcal{E}} (U(g) \otimes U^*(g))) \\ &= \operatorname{argmax}_{\mathcal{E}'} \left((U(g) \otimes U^*(g)) \Phi_{\mathcal{E}'} (V^\dagger(g) \otimes V^T(g)) \Phi_{\mathcal{E}} \right), \end{aligned} \quad (62)$$

from which we obtain

$$\Phi_{\mathcal{E}^{qi}} V(g) \otimes V^*(g) = U(g) \otimes U^*(g) \Phi_{\mathcal{E}^{qi}}. \quad (63)$$

Equivalently this means that

$$\mathcal{E}^{qi}(V(g) \rho V^\dagger(g)) = U(g) \mathcal{E}^{qi}(\rho) U^\dagger(g), \quad \forall g \in G. \quad (64)$$

Therefore the quasi-inverse of a covariant channel is also covariant except that the order of the two representations of the group is reversed.

Example 5 (mixed unitary channels with commuting unitaries). In contrast to the qubit case where a classification of quantum channels facilitates the study of various aspects of them including their quasi-inverses, for higher dimensional channels, many aspects do not easily yield an analytical treatment. In this section we pose a simple problem whose solution, as we will see, is quite nontrivial and yet instructive. We have seen in Example 1, that the quasi-inverse of a mixed unitary channel of the form

$$\mathcal{E}(\rho) = \sum_k p_k U_k \rho U_k^\dagger \quad (65)$$

when U_k 's are orthogonal to each other, is the unitary channel $\mathcal{E}^{qi}(\rho) = U_{max}^\dagger \rho U_{max}$, where U_{max} is the unitary corresponding to the maximum probability p_{max} in (65). When the

unitaries are not orthogonal to each other, then we know the complete answer only for qubit channels [1]. For higher dimensional channels we can tackle the simplified version of the problem, if all unitary matrices commute, so that they are diagonal in a certain basis. Under such assumptions we will prove that for qutrit channels the quasi-inverse is a unitary map. Our analysis also reveals certain facts about higher dimensional channels which may be of interest in their own right. At the end of this section we will consider a concrete case for $d = 3$ and the reader can follow the general arguments here by looking at that special case.

We aim to find the quasi-inverse of the channel in (65) when $[U_k, U_l] = 0 \quad \forall k$. In the basis in which all the unitaries are diagonal, $U_k = \text{diag}(e^{i\theta_1^{(k)}}, e^{i\theta_2^{(k)}}, \dots, e^{i\theta_d^{(k)}})$, the superoperator of the channel reads

$$\Phi = \sum_{k=1}^M p_k U_k \otimes U_k^* = \sum_{i,j} \sum_{k=1}^M p_k e^{i(\theta_i^{(k)} - \theta_j^{(k)})} |i, j\rangle\langle i, j| = \sum_{i,j} w_{ij} |i, j\rangle\langle i, j|, \quad (66)$$

where

$$w_{ij} = \langle e^{i\theta_{ij}} \rangle := \sum_k p_k e^{i(\theta_i^{(k)} - \theta_j^{(k)})}. \quad (67)$$

Let G be the unitary group that consists of all diagonal unitary matrices in this basis.¹ Since U_k 's commute with any diagonal matrix, the channel $\mathcal{E} = \{U_k\}$ is G -covariant.

$$U \mathcal{E}(\rho) U^\dagger = \mathcal{E}(U \rho U^\dagger) \quad \forall \rho, \text{ and } \forall U \in G. \quad (68)$$

According to property (64), this implies that the quasi-inverse of this channel is also G -covariant. Expressed in terms of the superoperators, this means that the superoperator $\Phi_{\mathcal{E}}$ must commute with the superoperator of all unitary maps $\rho \rightarrow U \rho U^\dagger$, where $U \in G$. The superoperator of the latter is of the diagonal form

$$\Phi_U = \sum_{mn} e^{i(\phi_m - \phi_n)} |m, n\rangle\langle m, n|. \quad (69)$$

Let the superoperator of the quasi-inverse be given by

$$\Phi_{\mathcal{E}^{qi}} = \sum \Phi_{ij,kl} |ij\rangle\langle kl|. \quad (70)$$

Equation (68) now restricts the form of this superoperator to the following simple form

$$\Phi_{\mathcal{E}^{qi}} = \sum_{i,j} q_{ij} |ij\rangle\langle ij| + r_{ij} |ii\rangle\langle jj|. \quad (71)$$

The Choi-matrix of the quasi-inverse is obtained by reshuffling the entries of the superoperator, see (16), which amounts to

$$C_{\mathcal{E}^{qi}} = \sum_{i,j} r_{ij} |ij\rangle\langle ij| + q_{ij} |ii\rangle\langle jj|. \quad (72)$$

According to $\overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) = \frac{1}{d+1} \left(1 + \frac{1}{d} \text{Tr} \Phi_{\mathcal{E}^{qi}} \Phi_{\mathcal{E}}\right)$, the quasi-inverse is the channel which maximizes the following quantity

$$F := \text{Tr}(\Phi_{\mathcal{E}^{qi}} \Phi_{\mathcal{E}}) = \sum_{i=1}^d (q_{ii} + r_{ii}) + \sum_{i \neq j=1}^d q_{ij} w_{ij}. \quad (73)$$

¹This is also known as the phase group.

Here we have used the equality $w_{ii} = 1 \forall i$, subject to the condition that $C_{\mathcal{E}^{qi}}$ in (72) designates the Choi matrix of a legitimate quantum channel, i.e. it is a positive matrix with partial trace equal to the identity

$$C_{\mathcal{E}^{qi}} \geq 0 \quad \text{and} \quad \text{Tr}_1 C_{\mathcal{E}^{qi}} = I_d. \quad (74)$$

In view of the block-diagonal form of the Choi matrix, it turns out that its eigenvalues are of the form $\{r_{i \neq j}\} \cup \text{Eigenvalues of } Y$, where Y is the d -dimensional matrix

$$Y = \sum_{i,j=1}^d q_{ij} |i\rangle\langle j| + \sum_i r_{ii} |i\rangle\langle i|. \quad (75)$$

The second condition in (74) leads to the following set of equalities

$$q_{ii} + \sum_j r_{ij} = 1 \quad \forall i, \quad (76)$$

which can be rewritten as

$$q_{ii} + r_{ii} + \sum_{j \neq i} r_{ij} = 1 \quad \forall i. \quad (77)$$

Note that $r_{i \neq j}$, being the eigenvalues of the Choi matrix are non-negative. Therefore in order to maximize the right hand side of (73), we can take all of them to be zero, reducing (77) to

$$q_{ii} + r_{ii} = 1 \quad \forall i, \quad (78)$$

which further simplifies the expression (73) and reduces our problem to maximization of the expression

$$F = d + \sum_{i \neq j=1}^d q_{ij} w_{ij} \equiv \text{Tr}(X^T W), \quad (79)$$

subject to the positivity of the following matrix

$$X = I + \sum_{i \neq j} q_{ij} |i, j\rangle\langle i, j|. \quad (80)$$

The set of all matrices X , denoted by Ω is a convex set. Therefore the linear function F takes its maximum at the extreme points of the set Ω .

In genera, the problem of finding the extreme points of d -dimensional channels is a difficult and rather non-trivial. It is only known that the extreme points of the set of 2 dimensional unital channels are unitary maps. Here we show that for channels defined by the Choi matrix (72), the extreme points are unitary maps if $d \leq 3$. We show also a stronger result: for any dimension d the rank of any extreme point of this set is less than \sqrt{d} . To this end, we first need to clarify a few definitions and a lemma. In what follows, \mathcal{S} is a vector space and $\Omega \subset \mathcal{S}$ is a convex subset of \mathcal{S} .

A basic property of extreme points of a convex set Ω is depicted in figure (5). In this figure X_0 and X'_0 are extreme points while Y_0 is not. We note that any line (no matter how short) passing through an extreme point like X_0 or X'_0 contains points which do not belong to Ω , while for the non-extreme point Y_0 , there exists a sufficiently short line (namely the one lying on the edge) which lies entirely in Ω . We present this more formally in the following statement.

Definition 2. Let Y_0 be a non-extreme point of Ω . Then there is a $Z \in \mathcal{S}$, and $\epsilon > 0$, such that $Y_0 + t(Z - Y_0) \in \Omega$ for all $t \in (-\epsilon, \epsilon)$. We call Z a witness of non-extremality of Y_0 .

This definition implies its equivalent form.

Lemma 1. *Let X_0 be an extreme point of Ω and let $Z \in \mathcal{S}$ be an arbitrary element in \mathcal{S} such that for all $t \in (-\epsilon, \epsilon)$, $X_0 + t(Z - X_0) \in \mathcal{S}$. Then $Z = X_0$.*

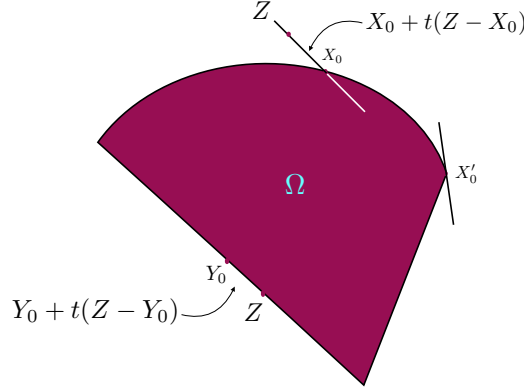


Figure 5: X_0 and X'_0 are extreme points of the convex set Ω . Any short segment passing through these points contains points outside Ω . On the other hand Y_0 is not an extreme point and there always exists a short segment passing through it which lies entirely inside Ω .

Now we can state and prove the following result.

Theorem 5. *Let $\mathbb{V}_m \subset \mathbb{M}_d$ be an m -dimensional subspace of the set of d -dimensional complex matrices and let $\Omega \subset \mathbb{V}_m$ be the cone of semi-definite positive matrices. Then the rank r of any extreme point of Ω is bounded as $r \leq \sqrt{d^2 - m}$.*

Proof. Let Y_0 be any element Ω with rank r . We show that if $r > \sqrt{d^2 - m}$, then we can always find an element $Z \in \Omega$ different from Y_0 such that the sufficiently short line segment $Y_0 + t(Z - Y_0)$ belongs entirely to Ω . This shows that such points cannot be extreme points of Ω . To this end, let us expand Y_0 in its eigenbasis as

$$Y_0 = \sum_{\alpha=1}^r \lambda_{\alpha} |u_{\alpha}\rangle \langle u_{\alpha}|, \quad (81)$$

where $\lambda_{\alpha} > 0 \quad \forall \alpha$. We choose $Z - Y_0 \in \mathbb{M}_d$ in the form

$$Z - Y_0 = \sum_{\alpha, \beta} z_{\alpha\beta} |u_{\alpha}\rangle \langle u_{\beta}|, \quad (82)$$

and let t be so small that $Y_0 + t(Z - Y_0) \geq 0$. The only other requirement that is needed for this matrix to belong to Ω is to satisfy $d^2 - m$ linear homogeneous equations which define the subspace Ω . This is a system of $d^2 - m$ linear homogeneous equations on r^2 variables and if $d^2 - m < r^2$, it has always a non-zero solution. This means that the point Y_0 is not an extreme point of the set Ω . Hence the rank of any extreme point of this set should be less than or equal to $\sqrt{d^2 - m}$. \square

Corollary 2. *As a corollary we find that the extreme points of the set of matrices (80) which is a subset of a $d^2 - d$ dimensional space, have rank $r < \sqrt{d}$. This means that for $d = 3$, the extreme points of the set (80) have unit rank and hence the quasi-inverse of mixed unitary channels with commuting unitaries is a unitary channel. This theorem by itself does not preclude the existence of quasi-inverses which are unitary maps in higher dimensions.*

Consider now the case of $d = 3$. Here after setting $r_{ij} = 0$ and satisfying the constraint (78), the superoperator is a diagonal matrix given by $\Phi_{\mathcal{E}^{qi}} = \text{diag}[1, q_{12}, q_{21}, 1, q_{23}, q_{31}, q_{32}, 1]$ and its Choi matrix is as follows:

$$C_{\mathcal{E}^{qi}} = \begin{pmatrix} 1 & . & . & . & q_{12} & . & . & . & q_{13} \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ q_{21} & . & . & . & 1 & . & . & . & q_{23} \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ q_{31} & . & . & . & q_{32} & . & . & . & 1 \end{pmatrix}. \quad (83)$$

The Hermitian matrices W and X , which appear in Eq. (79), are now 3 dimensional and take the form

$$X = \begin{pmatrix} 1 & q_{12} & q_{13} \\ q_{21} & 1 & q_{23} \\ q_{31} & q_{32} & 1 \end{pmatrix}, \quad \text{and} \quad W = \begin{pmatrix} 1 & w_{12} & w_{13} \\ w_{21} & 1 & w_{23} \\ w_{31} & w_{32} & 1 \end{pmatrix}. \quad (84)$$

Having proved that the matrix X_0 which maximizes the expression (79) is of unit rank, we can write it as $X = |\phi\rangle\langle\phi|$, where $|\phi\rangle = (q_1 \ q_2 \ q_3)^T$ maximizes $\text{Tr}(W|\phi\rangle\langle\phi|)$, with q_i being unimodular complex numbers. Note that the maximum will be smaller than the largest eigenvalue of W , if the corresponding eigenvectors is not built of unimodular entries. Therefore the quasi-inverse is the unitary map $\mathcal{E}^{qi}(\rho) = U\rho U^\dagger$, where the unitary operator U is given by $U = \text{diag}(q_1, q_2, q_3)$. In view of (29) and (79), the final average fidelity becomes

$$\overline{F} = \frac{1}{4} \left(1 + \frac{1}{3} \text{Tr}(\Phi_{\mathcal{E}^{qi}} \Phi_{\mathcal{E}}) \right) = \frac{1}{4} \left(1 + \frac{1}{3} \max_{|\phi\rangle} \langle\phi|W|\phi\rangle \right). \quad (85)$$

As a concrete example, consider a spin-1 particle subject to a magnetic field in the z direction, where the strength of the magnetic field or the exposure time is random. The evolution of the state is given by the following channel

$$\mathcal{E}(\rho) = \int d\tau f(\tau) U(\tau) \rho U^\dagger(\tau), \quad U(\tau) = \begin{pmatrix} e^{i\tau} & & \\ & 1 & \\ & & e^{-i\tau} \end{pmatrix}, \quad (86)$$

for some distribution $f(\tau)$. For this channel we have

$$w_{12} = w_{23} = \langle e^{i\tau} \rangle, \quad w_{13} = \langle e^{2i\tau} \rangle, \quad (87)$$

where the matrix W is Hermitian with unit diagonal entries and the averages are taken with respect to the distribution function $f(\tau)$. The vector $|\phi\rangle$ is then of the form $|\phi\rangle = (e^{i\tau_m} \ 1 \ e^{-i\tau_m})^T$, where τ_m is chosen to maximize $\langle\phi|W|\phi\rangle$ which is

$$\langle\phi|W|\phi\rangle = \text{Re} [2\langle e^{i\tau} \rangle e^{-i\tau_m} + \langle e^{2i\tau} \rangle e^{-2i\tau_m}]. \quad (88)$$

In view of the form of X , the quasi-inverse will be given by

$$V = \begin{pmatrix} e^{i\tau_m} & & \\ & 1 & \\ & & e^{-i\tau_m} \end{pmatrix}, \quad (89)$$

and the average fidelity after application of the quasi-inverse is given by

$$\overline{F} = \frac{3 + \langle (1 + 2 \cos(\tau - \tau_m))^2 \rangle}{12}. \quad (90)$$

Figure 6 shows the average fidelity and the increase in average fidelity for a simple discrete distribution $P(\tau = 0) = 1 - p$ and $P(\tau = \tau_0) = p$.

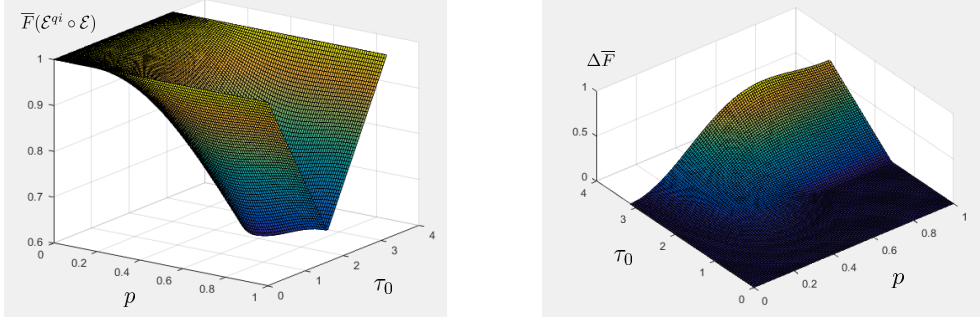


Figure 6: (Left) The average fidelity, \overline{F} , and (Right) the increase of the average fidelity, $\Delta \overline{F}$, of the channel (86) after correction with quasi-inverse (89).

6 Quasi-inversion of classical channels

There are several known parallels between probability distributions and stochastic matrices on the one hand and their quantum counterparts, namely density matrices and quantum channels, on the other hand. For instance, stochastic and bi-stochastic matrices acting on probability vectors form classical analogues of quantum channels and unital quantum channels. Furthermore, the discrete group of permutations is the analog of the continuous group of unitary channels. As the concept of convexity is critical in both domains, it is instructive to analyze, how the notion of quasi-inverse works in the classical setup.

In more concrete terms, the state of classical stochastic system of dimension d is a real vector \mathbf{p} whose entries are non-negative and add up to one. The set of all probability vectors of length d , forms the simplex Δ_d which is a $(d-1)$ -dimensional compact and convex set. For the sake of simplicity, let us use the bra-ket notation here to mention a probability vector. In this sense, let $|i\rangle$ for $i \in \{1, \dots, d\}$ denotes the pure probability vector whose all components are zero but the i -th element which is equal to 1. Therefore, a general mixed probability vector $\mathbf{p} = (p_1, \dots, p_d)^T$ can be stated as a convex combination of $\{|i\rangle\}$, i.e. $|\mathbf{p}\rangle = \sum_i p_i |i\rangle$.

A classical channel is represented by a stochastic transition matrix T of order d with non-negative elements where the sum of all elements in each column is equal to 1. This is the analog of trace-preserving property. The space of stochastic matrices of order d is a $(d^2 - d)$ -dimensional convex and compact set which will be denoted by \mathcal{S}_d .

Making use of the analogy to the quantum case consider the generalized Bloch representation (4) of a diagonal density matrix $\rho_{\text{diag}} = \text{diag}(\mathbf{p})$. Let us order the generators of $SU(d)$ matrices used in (5) in such a way that Γ_i are diagonal for $i = 1, \dots, d-1$. Then any diagonal matrix ρ_{diag} , representing the classical state \mathbf{p} , is represented in the Bloch form (4), where the Bloch vector \mathbf{t}_{cl} has now only $d-1$ components.

Using such a Bloch representation for point \mathbf{p} of the probability simplex Δ_d one can represent the action of an arbitrary stochastic matrix in the form [44],

$$T = \begin{pmatrix} 1 & 0 \\ \sqrt{d-1} \mathbf{t}_{cl} & M_{cl} \end{pmatrix}. \quad (91)$$

Note that this representation mimics the Liouville form (14) of a quantum operation, with the only difference that the classical distortion matrix M_{cl} forms a $(d-1)$ dimensional truncation of the quantum distortion matrix M of order $d^2 - 1$, while the classical translation vector \mathbf{t}_{cl} consists of $d-1$ components of the original translation vector \mathbf{t} of size $d^2 - 1$. Hence in the Bloch basis the classical transition matrix T forms a block of a matrix representing a quantum operation $\Phi_{\mathcal{E}}$, which decoheres to it [28].

Let us mention explicitly two distinguished classical maps. The first is the permutation map P_{α} where $P_{\alpha}|i\rangle = |\alpha_i\rangle$, where $1 \leq i \leq d$, while α_i denote permutations of $\{1, 2, \dots, d\}$. The next one is the flat, van der Weerden matrix, denoted by T_* , which sends all input states to the uniform state. It implies $(T_*)_{ij} = 1/d$ for any $1 \leq i, j \leq d$. As a result, the summation of elements of the columns of an assumed matrix A is equal to 1 if and only if $T_* A = T_*$. Accordingly, if A is an invertible matrix, the summation of elements on the columns of A^{-1} is also equal to 1.

Finally we note the fidelity of two probability vectors $|\mathbf{p}\rangle$ and $|\mathbf{q}\rangle$ which is defined as

$$\langle \mathbf{p} | \mathbf{q} \rangle := \left(\sum_{i=1}^d \sqrt{p_i q_i} \right)^2. \quad (92)$$

Proceeding in the same way that we did for the quantum case, here we should define the average fidelity of a classical channel to be the fidelity of an output state with the input pure state averaged over all input states. Thus we define

$$\overline{\mathcal{F}}(T) := \frac{1}{d} \sum_{i=1}^d F(|i\rangle, T|i\rangle) = \frac{1}{d} \text{Tr}(T). \quad (93)$$

Under these assumptions, the quasi-inversion is the channel (the stochastic matrix) increasing Eq. (93) as much as possible:

$$\text{Tr}(T^{qi} T) \geq \text{Tr}(T' T), \quad \forall T' \in \mathcal{S}_d. \quad (94)$$

To emphasize even further similarity to the quantum case, consider the Bloch representation (91) of the classical map T involving its distortion matrix M_{cl} . Then the average fidelity of

the corrected classical transformation $T^{qi}T$ can be expressed as the maximum over the set of allowed classical distortion matrices,

$$\overline{\mathcal{F}}(T^{qi}T) = \max_{M'_{cl}} \frac{1}{d} (1 + \text{Tr} M'_{cl} M_{cl}), \quad (95)$$

which is in analogy to Eq. (29). However, the difference in prefactor with respect to Eq. (29) is due to the averaging over the set of classical probabilities of a dimension smaller than the set of quantum states.

All the arguments of Theorem 1, based on the linearity of the fidelity function and its two properties (27) and (28) are also valid here and therefore Theorems 1 and 2 and the corollary 1 hold true also if we replace the quantum channels with classical ones. In particular, the fact that the quasi-inverse of a quantum channel can be taken to correspond to an extreme point is very important, since compared with the quantum case, we have a much better knowledge of the convex set of classical channels and its extreme points. In the general case, our basic theorem is the following:

Theorem 6. *Let the stochastic matrix T be such that in each row i , the element in the a_i -th column be the maximum. Then its quasi-inverse is found by replacing that single element by 1 and setting all the other elements in that row equal to zero and then transposing the matrix.*

Proof. Let the matrix T be written as $T = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n)^T$ where \mathbf{t}_i^T denotes the i -th row as a vector. Denote its quasi-inverse as a matrix $T^{qi} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$, where \mathbf{x}_i is the i -th column as a vector. The aim is to maximize $\text{Tr}(TT^{qi}) = \sum_i \mathbf{t}_i \cdot \mathbf{x}_i$. We can maximize this sum if we maximize each inner product $\mathbf{t}_i \cdot \mathbf{x}_i$ independently. Since none of the components of the vectors \mathbf{x}_i can be larger than one, the maximization is achieved if we choose each $\mathbf{x}_i = (0, 0, \dots, 1, 0, 0)$ such that 1 stands in the position of maximum component of \mathbf{t}_i . This proves the theorem. \square

For example assume a general two dimensional stochastic matrix parameterized as

$$T = \begin{pmatrix} 1-x & y \\ x & 1-y \end{pmatrix}, \quad 0 \leq x, y \leq 1. \quad (96)$$

In this case, the quasi-inverse is either I , σ_x , or a mixture of these two for $x+y < 1$, $x+y > 1$ and $x+y = 1$, respectively. This is depicted in Fig. 7. In higher dimensions, however, more options are possible.

Example 6. *As examples, the quasi-inverse of the stochastic matrices*

$$T_1 = \frac{1}{8} \begin{pmatrix} 2 & 6 & 2 \\ 4 & 1 & 1 \\ 2 & 1 & 5 \end{pmatrix}, \quad \text{and} \quad T_2 = \frac{1}{48} \begin{pmatrix} 12 & 6 & 3 \\ 4 & 36 & 42 \\ 32 & 6 & 3 \end{pmatrix} \quad (97)$$

are given by

$$T_1^{qi} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad T_2^{qi} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \quad (98)$$

The second example shows that the quasi-inverse of a classical map is not necessarily a permutation. In the examples above, the average fidelities of the stochastic maps T_1 and T_2 increase from $1/3$ and $5/18$ to $35/72$ and $43/72$, respectively.

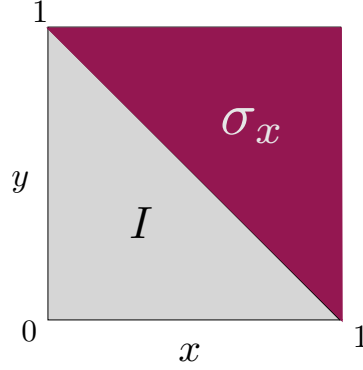


Figure 7: The set of all stochastic matrices of dimension 2. Transitions located in the gray region cannot be corrected in sense of input-output fidelity. For other stochastic maps one can apply σ_x to modify the fidelity.

Corollary 3. If T denotes a stochastic matrix for which $\forall i, j : T_{ii} \geq T_{ij}$, then its quasi-inverse is I_d which means it is not possible to increase the average fidelity.

As it is seen, these quasi-inverses are at extreme points of the space of stochastic matrices. To have quasi-inverses which are not necessarily at the extreme points, we should consider the case where there are more than one maximum entry in each row. In this case if we follow the argument leading to Theorem 6, we see the quasi-inverse can be the convex combination of all quasi-inverses which we construct when we consider only one of these elements. The next example illustrates this point.

Example 7. *The stochastic matrix*

$$T = \frac{1}{24} \begin{pmatrix} 8 & 3 & 8 \\ 12 & 6 & 6 \\ 4 & 15 & 10 \end{pmatrix} \quad (99)$$

has as its quasi-inverse

$$T^{qi} = \lambda \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & 0 & 1 \\ 1 - \lambda & 0 & 0 \end{pmatrix}. \quad (100)$$

Here the average fidelity increases from the value $\frac{1}{3}$ to $\frac{35}{72}$ which is expectedly independent from λ .

Theorem 7. *Among the stochastic matrices with a unique quasi-inverse, only symmetric permutations are their own quasi-inverse.*

Proof. Uniqueness of the quasi-inverse suggests that in each row of T there exists a single entry which is strictly larger than other elements in the row. Hence, we have exactly one non-zero array (which is equal to 1) in each column of T^{qi} . Suppose that the classical channel T is a self-quasi-inverse stochastic matrix, i.e. $T^{qi} = T$. This equality implies the existence of exactly one non-zero element equal to unity in each row of T^{qi} . Because we would have had more than one leading value in a row of T which violates the uniqueness of T^{qi} , otherwise. So T^{qi} and consequently T are the same permutation matrix. However, a permutation matrix has a real inverse which is equal to its transpose. So we have $T^{qi} = T = T^T$ suggesting that T is a symmetric permutation matrix. \square

6.1 On the commutativity of super-decoherence and quasi-inverse.

Now that we have discussed the quasi-inverses of both quantum and classical channels, a natural question is whether or not through super-decoherence [28] of quasi-inverse of a quantum channel \mathcal{E} , the quasi-inverse of a classical channel $T_{\mathcal{E}}$ can be obtained. In other words, we want to see if the action of taking quasi-inverse commutes with super-decoherence. As we will see below, in general the answer is negative. For convenience we first remind the concept of super-decoherence [28].

The decoherence channel removes off-diagonal elements of density matrices and sends any quantum state ρ into its diagonal $\mathcal{D}(\rho) = \rho_d = \sum \rho_{ii}|i\rangle\langle i|$, i.e. a classical state. Defining an analogous process in the space of quantum channels, one may extract a classical map, a stochastic matrix, from any quantum channel. This process is called super-decoherence, noted by Δ , to emphasize that it acts on quantum channels and not states. For a quantum channel \mathcal{E} , the assigned classical transition matrix obtained by super-decoherence is defined by $T_{\mathcal{E}} = \Delta(\mathcal{E})$ where

$$(T_{\mathcal{E}})_{ij} = \langle i|\mathcal{E}(|j\rangle\langle j|)|i\rangle = (\Phi_{\mathcal{E}})_{ii,jj} = (C_{\mathcal{E}})_{ij,ij}. \quad (101)$$

The last equality above shows that this stochastic matrix is actually gained by decohering the Choi matrix and clarifies why it is called super-decoherence. It is straightforward to see stochasticity of $T_{\mathcal{E}}$ is guaranteed by the fact that \mathcal{E} is a positive and trace preserving map. This is related to the fact that the classical distortion matrix M_{cl} of size $(d-1)$ used in (91) forms a block of the quantum distortion matrix M of size d^2-1 present in the Liouville form (14) of any corresponding quantum operation [44].

Moreover, one can show that if the channel \mathcal{E} is described by the set of Kraus operators $\{K_{\alpha}\}$, then $T_{\mathcal{E}} = \sum K_{\alpha} \odot K_{\alpha}^*$ where \odot defines Hadamard (entry-wise) product. Through this relation it is easy to see by super-decohering a unital channel we get a bistochastic matrix, while a unistochastic matrix is obtained if the input channel is unitary.

However, a simple counter example shows that the answer to the question posed at the beginning of this subsection is negative. Consider the case of a single-qubit channel, where the quasi-inverse of a channel is in general a unitary map, $U = e^{i\theta \mathbf{n} \cdot \boldsymbol{\sigma}}$, where θ and \mathbf{n} depend on the channel. On the other hand, as shown in Fig. 7, the quasi-inverse of any 2-dimensional classical map is either identity I or the permutation σ_x , which arise due to super-decoherence of a small subset of the unitary channels [45].

7 Further results and examples

Although Theorem 6 gives a complete prescription for finding the quasi-inverse of any stochastic matrix, it is instructive to consider a few special classes. These examples illustrate further the parallels between classical and quantum notions of maps and their quasi-inverses. The first result is the analog of unitarily equivalent quantum maps.

Consider two transition matrices T_1 and T_2 which are related by two arbitrary permutations P_{α} and P_{β} in the following way $T_2 = P_{\alpha} \circ T_1 \circ P_{\beta}$. Let us call them permutationally equivalent stochastic matrices. Pursuing the same approach adopted in obtaining Eq. (36), and noting that the set of stochastic matrices, \mathcal{S}_d , is invariant under permutations, we get

$$T_2^{qi} = P_{\beta}^{-1} \circ T_1^{qi} \circ P_{\alpha}^{-1}. \quad (102)$$

Moreover, T_1 and T_2 have the same amount of fidelity after correction. The next example illustrates the connection with its quantum analog, Example 1.

Example 8 (Convex combination of orthogonal permutations). A bi-stochastic matrix T is a stochastic matrix with the extra property that the sum of entries of each row equals unity. It is a well-known result due to Birkhoff's theorem that any such matrix can be written as a convex combination of permutations. For dimension d , there are $d!$ such permutations which form the extreme points of the convex set of these matrices, conventionally called the Birkhoff polytope. Note however that a bi-stochastic matrix has $(d-1)^2$ independent parameters and hence the convex decomposition of an arbitrary bi-stochastic matrix in terms of $d!$ permutations is not unique. Consider now a special class of bi-stochastic matrices which are convex combination of orthogonal permutations. These permutations are defined by the property that

$$\text{Tr}(P_m^{-1} P_n) = d\delta_{n,m}. \quad (103)$$

Equation (103) indicates that

$$\forall k \quad P_m|k\rangle \neq P_n|k\rangle, \quad \text{if } m \neq n. \quad (104)$$

This means that there are at most d orthogonal permutations in the group S_d of all permutations of d objects. Furthermore, it implies that

$$\langle j | \sum_m P_m | k \rangle = 0, 1 \quad \forall j, k. \quad (105)$$

Here the sum runs over all the permutations in the convex combination, which may be a subset of all the d orthogonal permutations. A simple example consists of the set of permutations of the form $\{P_m = X^m, \quad m = 0, \dots, d-1\}$, where $X = \sum_{n=0}^{d-1} |n+1\rangle\langle n|$ is the full-cycle permutation, or any subset thereof. By definition we have $P_m|k\rangle = |k+m\rangle$ which clearly satisfies (105) as $\sum_m \langle j | P_m | k \rangle = \sum_m \delta_{j,k+m} = 0, 1$. In this example all the permutations commute with each other. As an example consisting of non-commuting but orthogonal permutations consider the following:

$$P_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad (106)$$

with matrix representations

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \quad (107)$$

One can see that P_1 and P_2 are orthogonal while P_1 and P_3 are not. Also one can see that

$$P_1 + P_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad (108)$$

satisfying the condition (105) while

$$P_1 + P_3 = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad (109)$$

violates (105). Furthermore one can check that in the group S_3 with generators $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ and $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, each of the two sets of even and odd permutations, respectively given by $(I, \sigma_1\sigma_2, \sigma_2\sigma_1)$ and $(\sigma_1, \sigma_2, \sigma_1\sigma_2\sigma_1)$ comprise orthogonal permutations.

Consider now a bi-stochastic matrix of the form

$$B = \sum \lambda_m P_m, \quad (110)$$

where $\{P_m\}$ are orthogonal. In what follows we will show the quasi-inverse of any such bi-stochastic matrix is $P_{m_0}^{-1}$ where m_0 refers to the index of the greatest coefficient λ_m . To see this we note that

$$\begin{aligned} \overline{\mathcal{F}}(T \circ B) &= \frac{\text{Tr}(TB)}{d} = \frac{\sum_m \lambda_m \text{Tr}(TP_m)}{d} \\ &\leq \frac{\lambda_{m_0} \text{Tr}(T \sum_m P_m)}{d} = \frac{\lambda_{m_0} \sum_{k,j} \langle k|T|j \rangle \langle j| \sum_m P_m |k \rangle}{d}. \end{aligned} \quad (111)$$

Using Eq. (105) for orthogonal permutations, we find

$$\overline{\mathcal{F}}(T \circ B) \leq \frac{\lambda_{m_0} \sum_{k,j} \langle k|T|j \rangle}{d} = \lambda_{m_0}. \quad (112)$$

where we used the fact that for every stochastic matrix, the sum of its elements is equal to the dimension d . Now, it is straightforward to see $T^{q^i} = P_{m_0}^{-1}$ recovers the average fidelity of T to this upper bound.

Thus the quasi-inverse of convex combination of orthogonal permutations is the inverse of the single permutation which has the largest share in the convex combination. This is however not the case if the permutations are not orthogonal. This is shown in the next example.

Example 9 (Convex combination of non-orthogonal permutations). Consider the following permutations

$$P_1 = I \oplus I \oplus I, \quad P_2 = \sigma_x \oplus \sigma_x \oplus I, \quad P_3 = \sigma_x \oplus I \oplus \sigma_x, \quad (113)$$

where $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. These permutations are obviously non-orthogonal in the sense that $\text{Tr}(P_i P_j^T) \neq 0$. Consider now the following convex combination

$$P = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3, \quad (114)$$

where $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and all $\lambda_i < \frac{1}{2}$. In explicit form this permutation matrix is given by

$$P = \begin{pmatrix} \lambda_1 & \lambda_2 + \lambda_3 & \cdot & \cdot & \cdot & \cdot \\ \lambda_2 + \lambda_3 & \lambda_1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \lambda_1 + \lambda_3 & \lambda_2 & \cdot & \cdot \\ \cdot & \cdot & \lambda_2 & \lambda_1 + \lambda_3 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \lambda_1 + \lambda_2 & \lambda_3 \\ \cdot & \cdot & \cdot & \cdot & \lambda_3 & \lambda_1 + \lambda_2 \end{pmatrix}. \quad (115)$$

Since $\lambda_i < \frac{1}{2}$ for all i , we find $\lambda_i + \lambda_j > \frac{1}{2}$ for all $(i \neq j)$ (due to the requirement that $\lambda_1 + \lambda_2 + \lambda_3 = 1$). Thus according to Theorem 6, the quasi-inverse of P is given by replacing the largest entry of each row with unity and then replacing the resulting matrix, hence $P^{q^i} = \sigma_x \oplus I \oplus I$ which is not equal to inverse of any of the permutations P_1, P_2 or P_3 .

8 Concluding remarks

We have extended the concept of quasi-inversion of qubit channels [1] to quantum channels in arbitrary dimensions and to the classical domain, i.e. Markov processes in discrete time. In both cases, a quasi-inverse is a map which when combined with the original channel increases the average input-output fidelity in an optimal way. While the complete classification of qubit channels [19, 20], makes a complete characterization of their quasi-inverse possible, in higher dimensions the lack of such classification makes the problem a highly non-trivial one. The most notable difference is that in the qubit case, the extreme points of all unital channels are the unitary maps while in higher dimensions this is not the case any more and no general theorem is known on extreme points. Therefore in this paper, we have established certain general theorems on the nature of the quasi-inverse, and have provided certain bounds on the average fidelity after quasi-inversion. Applying these general results to some concrete cases, we have found in examples 1 through 5 the quasi-inverse for a large class of quantum channels. Moreover, we have done a parallel analysis for the classical channels, represented by stochastic matrices.

As shown in the Appendix D, we have also obtained exact expressions for the average input-output fidelity for classical channels in any dimension d and have shown that the quasi-inversion increases this quantity from d^{-1} to, approximately, $d^{-1/2}$. In the quantum case, we analyzed numerically in Appendix E the improvement of the average fidelity of a random channel after correction with quasi-inversion and after applications of the best possible unitary evolution – see Fig. 10. If the dimension of the systems increases, the non-unitarity of quasi-inversion becomes larger. In other words, for a generic channel acting in higher dimensions, its quasi-inverse is usually non-unitary.

It is noteworthy to mention that by applying quasi-inversion, we aim to get as close as possible to the identity map in the sense of the maximizing average fidelity and so by the reduction of the average Bures distance. We do not expect the notion of quasi-inversion to improve some other properties of channels, for which the identity map is not the most distinguished channel. There are other figures of merit, like the average output purity of a channel

$$\bar{P}(\mathcal{E}) := \int d\psi \operatorname{Tr}[(\mathcal{E}(|\psi\rangle\langle\psi|))^2], \quad (116)$$

or the unitarity of a channel [46]

$$\bar{u}(\mathcal{E}) := \frac{d}{d-1} \int d\psi \operatorname{Tr}\left[\mathcal{E}(|\psi\rangle\langle\psi|) - \mathcal{E}\left(\frac{I}{d}\right)\right]^2, \quad (117)$$

which can also be studied in the same way, leading to a different version of the quasi-inverse, both in the quantum and the classical domain. We are aware of examples which show that these notions of quasi-inverse are different. That is, a quasi-inverse which increases the average fidelity can increase, decrease or keep constant the unitarity of a quantum channel. One can also study the effect of quasi-inversion on the cohering power of quantum channels as defined in [47–49].

Finally, an interesting by-product of our study is Appendix C, where we have constructed special types of channels with affine maps (M, \mathbf{t}) , where the distortion matrix M can stretch the Bloch vectors. Such channels are possible only in dimensions higher than two. These are of course different from channels of the form $\mathcal{E}(\rho) = |\psi\rangle\langle\psi|$, where the distortion matrix M vanishes and the stretching is due only to the translation vector \mathbf{t} .

We hope that this study can be pursued in different directions, i.e. in obtaining more information about the extreme points of the space of quantum channels, and hence the quasi-inverse of larger classes of channels, those channels which do not have unique quasi-inverses, and those which are their own quasi-inverse. And also more importantly in finding connections with the recovery maps [50–53].

It is a pleasure to thank Seyed Javad Akhtarshenas, Erik Aurell, Giulio Chiribella, Sergey Filippov, Kamil Korzekwa, and Lukasz Pawela for several discussions and helpful remarks. This research was partially supported by the grant number G98024071 from Iran National Science Foundation. Financial support by Narodowe Centrum Nauki under the grant number DEC-2015/18/A/ST2/00274 and by the Foundation for Polish Science under the Team-Net NTQC project is gratefully acknowledged.

References

- [1] V. Karimipour, F. Benatti, and R. Floreanini, “Quasi-inversion of qubit channels,” *Phys. Rev. A* **101**, 032109 (2020).
- [2] M. A. Nielsen, and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, 2000).
- [3] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, “Completely positive dynamical semigroups of N -level systems,” *J. Math. Phys.* **17**, 821 (1976).
- [4] G. Lindblad, “On the generators of quantum dynamical semigroups,” *Commun. Math. Phys.* **48**, 119 (1976).
- [5] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra Appl.* **10**, 285–290 (1975).
- [6] K. Kraus, “General state changes in quantum theory,” *Ann. Physics* **64**, 311–335 (1971).
- [7] E. Knill, “Quantum computing with realistically noisy devices,” *Nature* **434**, 39 (2004).
- [8] D. A. Lidar, I. L. Chuang, and K. B. Whaley, “Decoherence-free subspaces for quantum computation,” *Phys. Rev. Lett.* **81**, 2594 (1998).
- [9] A. Beige, D. Braun, B. Tregenna, and P. L. Knight, “Quantum computing using dissipation to remain in a decoherence-free subspace,” *Phys. Rev. Lett.* **85**, 1762 (2000).
- [10] P. G. Kwiat, A. J. Berglund, J. B. Altepeter, and A. G. White, “Experimental verification of decoherence-free subspaces,” *Science* **290**, 498 (2000).
- [11] N. Katz et al., “Reversal of the weak measurement of a quantum state in a superconducting phase qubit,” *Phys. Rev. Lett.* **101**, 200401 (2008).
- [12] Q. Q. Sun, M. Al-Amri, and M. S. Zubairy, “Reversing the weak measurement of an arbitrary field with finite photon number,” *Phys. Rev. A* **80**, 033838 (2009).
- [13] Q. Q. Sun, M. Al-Amri, L. Davidovich, and M. S. Zubairy, “Reversing entanglement change by a weak measurement,” *Phys. Rev. A* **82**, 052323 (2010).
- [14] Y. S. Kim, J. C. Lee, O. Kwon, and Y. H. Kim, “Protecting entanglement from decoherence using weak measurement and quantum measurement reversal,” *Nat. Phys.* **8**, 117 (2012).

- [15] Y. Aharonov, D. Z. Albert, and L. Vaidman, “How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100,” *Phys. Rev. Lett.* **60**, 1351 (1988).
- [16] G. J. Pryde, J. L. O’Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, “Measurement of quantum weak values of photon polarization,” *Phys. Rev. Lett.* **94**, 220405 (2005).
- [17] E. Aurell, J. Zakrzewski, and K. Życzkowski, “Time reversals of irreversible quantum maps,” *J. Phys. A* **48**, 38FT01 (2015).
- [18] G. Chiribella, E. Aurell, and K. Życzkowski, “Symmetries of quantum evolutions,” arXiv:2101.04962, *Phys. Rev. Research*, in press.
- [19] M. B. Ruskai, S. Szarek, and E. Werner, “An analysis of completely-positive trace-preserving maps on M_2 ,” *Linear Algebra and its Applications* **347**, 159–187 (2002).
- [20] A. Fujiwara, and P. Algoet, “One-to-one parametrization of quantum channels,” *Phys. Rev. A* **59**, 3290–3294 (1999).
- [21] R. Koenig, R. Renner, and C. Schaffner, “The operational meaning of min- and max-entropy,” *IEEE Trans. Inf. Th.* **55**, 4337 (2009).
- [22] I. Bengtsson, and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, II ed., Cambridge University Press, 2017.
- [23] G. Chiribella and D. Ebler, “Optimal quantum networks and one-shot entropies,” *New J. Phys.* **18**, 093053 (2016).
- [24] L. J. Landau, and R.F. Streater, “On Birkhoff’s theorem for doubly stochastic completely positive maps of matrix algebras,” *Lin. Alg. Appl.* **193**, 107 (1993).
- [25] S. L. Tregub, “Bistochastic operators on finite-dimensional von Neumann algebras,” *Soviet Math.* **30**, 105 (1986).
- [26] B. Kümmerner, and H. Maassen, “The essentially commutative dilations of dynamical semigroups on M_n ,” *Commun. Math. Phys.* **109**, 1 (1987).
- [27] S. N. Filippov, and K. V. Kuzhamuratova, “Quantum informational properties of the Landau–Streater channel,” *J. Math. Phys.* **60**, 042202 (2019).
- [28] K. Korzekwa, S. Czachórski, Z. Puchała, and K. Życzkowski, “Coherifying quantum channels,” *New J. Phys.* **20**, 043028 (2018).
- [29] S. K. Goyal, B. N. Simon, R. Singh, and S. Simon, “Geometry of the generalized Bloch sphere for qutrits,” *J. Phys. A: Math. Theor.* **49**, 165203 (2016).
- [30] L. Jakóbczyk, and M. Siennicki, “Geometry of Bloch vectors in two-qubit system,” *Phys. Lett. A* **286**, 383 (2001).
- [31] G. Kimura, “The Bloch vector for N -level systems,” *Phys. Lett. A* **314**, 339 (2003).
- [32] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki, “Robust extraction of tomographic information via randomized benchmarking,” *Phys. Rev. X* **4**, 011050 (2014).
- [33] U. Fano, “Pairs of two-level systems,” *Rev. Mod. Phys.* **55**, 855 (1983).
- [34] B. Schumacher, and M. A. Nielsen, “Quantum data processing and error correction,” *Phys. Rev. A* **54**, 2629 (1996).

- [35] M. Horodecki, and P. Horodecki, “Reduction criterion of separability and limits for a class of distillation protocols,” *Phys. Rev. A* **59**, 4206 (1999).
- [36] M. Horodecki, P. Horodecki, and R. Horodecki, “General teleportation channel, singlet fraction, and quasidistillation,” *Phys. Rev. A* **60**, 1888 (1999).
- [37] A. Uhlmann, “The “transition probability” in the state space of a $*$ -algebra,” *Rep. Math. Phys.* **9**, 273 (1976).
- [38] R. Jozsa, “Fidelity for mixed quantum states,” *J. Mod. Opt.* **41**, 2315 (1994).
- [39] K. Życzkowski, and H.-J. Sommers, “Average fidelity between random quantum states,” *Phys. Rev. A* **71**, 032313 (2005).
- [40] R. Renner, “Security of quantum key distribution,” Ph.D. dissertation, ETH Zurich, arXiv:0512258 (2005).
- [41] R. Alicki, and K. Lendi, *Quantum Dynamical Semigroups and Applications* Springer-Verlag Berlin, 2007.
- [42] K. Audenaert, and S. Scheel, “On random unitary channels,” *New J. Phys.* **10**, 023011 (2008).
- [43] R. F. Werner, and A. S. Holevo, “Counterexample to an additivity conjecture for output purity of quantum channels,” *J. Math. Phys.* **43**, 4353 (2002).
- [44] R. Kukulski, I. Nechita, L. Pawela, Z. Puchała, and K. Życzkowski, “Generating random quantum channels,” *J. Math. Phys.* **62**, 062201 (2021).
- [45] K. Korzekwa, S. Czachórski, Z. Puchała, and K. Życzkowski, “Distinguishing classically indistinguishable states and channels,” *J. Phys. A* **52**, 475303 (2019).
- [46] J. J. Wallman, C. Granade, R. Harper, and S. T. Flammia, “Estimating the coherence of noise,” *New J. Phys.* **17**, 113020 (2015).
- [47] A. Streltsov, G. Adesso, and M. B. Plenio, “Colloquium: Quantum coherence as a resource,” *Rev. Mod. Phys.* **89**, 041003 (2017).
- [48] A. Mani, and V. Karimipour, “Cohering and decohering power of quantum channels,” *Phys. Rev. A* **92**, 032331 (2015).
- [49] P. Zanardi, G. Styliaris, and L. C. Venuti, “Coherence-generating power of quantum unitary maps and beyond,” *Phys. Rev. A* **95**, 052306 (2017).
- [50] O. Fawzi and R. Renner, “Quantum conditional mutual information and approximate Markov chains,” *Comm. Math. Phys.* **340**, 575–611 (2015).
- [51] M. Berta, M. Lemm, and M. M. Wilde, “Monotonicity of quantum relative entropy and recoverability,” *Quant. Inf. Compu.* **15**, 1333-1354 (2015).
- [52] M. Berta, and M. Tomamichel, “The fidelity of recovery is multiplicative,” *IEEE Transactions on Information Theory* **62**, 1758-1763 (2016).
- [53] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, “Universal recovery maps and approximate sufficiency of quantum relative entropy,” In *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, pp. 2494–2498 (2016).
- [54] R. Bhatia, *Positive Definite Matrices*, Princeton University press, 2007.

A An explicit basis for $L(H_d)$

In this appendix we choose an explicit representation for the matrices Γ_i belonging to $L(H_d)$. Let $E_{ij} := |i\rangle\langle j|$ be the standard basis of matrices, then the set

$$H_k = \sqrt{\frac{d(d-1)}{k(k+1)}} \left(\sum_{i=1}^k E_{ii} - k E_{k+1,k+1} \right) \quad k = 1, \dots, d-1 \quad (118)$$

and

$$X_{ij} = \sqrt{\frac{d(d-1)}{2}} (E_{ij} + E_{ji}) \quad Y_{ij} = \sqrt{\frac{d(d-1)}{2}} (-iE_{ij} + iE_{ji}) \quad (119)$$

form a basis of Hermitian matrices with the proper normalization in (5). They are nothing but the standard Gell-Mann matrices, properly normalized to satisfy (5). To work in parallel with the Bloch representation of any classical probability vector \mathbf{p} it is convenient to order the matrix basis in such a way that the $d-1$ diagonal matrices appear first, $\Gamma_k = H_k$, for $k = 1, \dots, d-1$ – see Section 6.

A density matrix can be written as $\rho = \frac{1}{d}(I + \mathbf{r} \cdot \mathbf{\Gamma})$, where \mathbf{r} is the generalized Bloch vector and $\mathbf{\Gamma}$ is a vector constructed by the elements Γ_i . We note, as stated after Eq. (6), that any pure state corresponds to unit vector \mathbf{r} . However the converse is not true. In fact one can easily verify that a matrix like $\frac{1}{d}(I - H_{d-1}) = |d\rangle\langle d|$ is a pure state while a state like $\frac{1}{d}(I + H_{d-1})$ is not a state at all, since it has negative eigenvalues. Both correspond to antipodal points on the sphere S_{d^2-2} .

B Some details on the calculations of average input-output fidelity

As shown in Eq. (20), the input-output fidelity is given by $\overline{F}(\mathcal{E}) = \text{Tr}(L\Phi_{\mathcal{E}})$, where L is given by

$$\begin{aligned} L &= \int d\psi |\psi\rangle\langle\psi| \otimes |\psi^*\rangle\langle\psi^*| \\ &= \int dU (U \otimes U^*) |00\rangle\langle 00| (U \otimes U^*)^\dagger. \end{aligned} \quad (120)$$

The last line shows $L = T(|00\rangle\langle 00|)$ is an isotropic state obtained by twirling the state $|00\rangle$. In general for any state ρ , its twirling gives [35]:

$$\begin{aligned} T(\rho) &= \int dU (U \otimes U^*) \rho (U \otimes U^*)^\dagger \\ &= \frac{d^2}{d^2-1} \left((1-f_\rho) \frac{I}{d^2} + (f_\rho - \frac{1}{d^2}) P_+ \right), \end{aligned} \quad (121)$$

where $f_\rho = \text{Tr}(P_+ \rho)$ and $P_+ = |\phi_+\rangle\langle\phi_+|$ is the maximally entangled state. So for the state $|00\rangle\langle 00|$ of a d -dimensional system, one has:

$$f_{|00\rangle} = \text{Tr}(P_+ |00\rangle\langle 00|) = \frac{1}{d}, \quad (122)$$

and

$$L = T(|00\rangle\langle 00|) = \frac{1}{d(d+1)} I + \frac{1}{(d+1)} P_+, \quad (123)$$

which coincides with Eq. (22) in the text.

One can also write the average fidelity (20) in the form

$$\overline{F}(\mathcal{E}) = \text{Tr}(\Lambda \Psi_{\mathcal{E}}), \quad (124)$$

where $\Psi_{\mathcal{E}} = \sum_{\alpha} K_{\alpha} \otimes K_{\alpha}^{\dagger}$ and

$$\Lambda = \int d\psi |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|$$

is a symmetric Werner state

$$\Lambda = \frac{1}{d(d+1)}(I + P_{sym}) = \frac{1}{d(d+1)}(I + \sum_{i,j} |i,j\rangle\langle j,i|). \quad (125)$$

This will also lead to the same result as in (23).

Finally let us calculate the average fidelity in yet another way, by expressing the pure states as

$$|\psi\rangle\langle\psi| = \frac{1}{d}(I + \mathbf{n} \cdot \mathbf{\Gamma}). \quad (126)$$

From

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \frac{1}{d}(I + (M\mathbf{n} + \mathbf{t}) \cdot \mathbf{\Gamma}). \quad (127)$$

and Eq. (126), one finds

$$\langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle = \frac{1}{d}(1 + (d-1)\mathbf{n} \cdot (M\mathbf{n} + \mathbf{t})). \quad (128)$$

If we now assume a uniform distribution of pure states on the sphere S_{d^2-2} (which is of course not dense) and hence use the relations

$$\int d\mathbf{n} n_i = 0, \quad \int d\mathbf{n} n_i n_j = \frac{1}{(d^2-1)}\delta_{ij} \quad (129)$$

we arrive at

$$\overline{F}(\mathcal{E}) = \frac{1}{d}\left(1 + \frac{1}{d+1}\text{Tr}M\right), \quad (130)$$

which coincides with the result (23) which we obtained by integrating over the invariant volume $d\psi$ of all pure states. Here we have not used any specific measure of volume over the sphere S_{d^2-2} and only have assumed that pure states are distributed symmetrically (albeit in a parse way) on this sphere.

C Quantum channels which stretch the Bloch vector

In two dimension, to satisfy the positivity condition, the affine matrix M of any positive map, and thus any quantum channel, fulfills $M^T M \leq I$. As a result, the Bloch vector corresponding to a qubit state is always shrunk when M acts on it. Here we show that this is no longer the case in higher dimensions. This is one of the strange or un-expected properties of higher dimensional channels which makes the study of their quasi-inverses, among other things, difficult. As describing this class of quantum maps is not straightforward we construct

an example of such a channel in this Appendix.

Let H_d be a d -dimensional Hilbert space and let P_1, P_2, Q_1 and Q_2 be orthogonal projectors of rank d_1, d_2, m_1 and m_2 respectively, with $(d = d_1 + d_2 = m_1 + m_2)$, i.e.

$$P_1 + P_2 = Q_1 + Q_2 = I, \quad P_1 P_2 = Q_1 Q_2 = 0.$$

The following operators are Hermitian and traceless

$$A := \frac{m_2 Q_1 - m_1 Q_2}{\sqrt{m_1 m_2}}, \quad B = \frac{d_2 P_1 - d_1 P_2}{\sqrt{d_1 d_2}}, \quad (131)$$

and we have

$$\text{Tr}(AB) = 0, \quad \text{Tr}(A^2) = d, \quad \text{Tr}(B^2) = d. \quad (132)$$

Consider now the following measure-and-prepare CPT map

$$\mathcal{E}(\rho) = \frac{1}{d_1} \text{Tr}(Q_1 \rho) P_1 + \frac{1}{d_2} \text{Tr}(Q_2 \rho) P_2. \quad (133)$$

By expanding the projectors as

$$\begin{aligned} P_1 &= \sum_{i=1}^{d_1} |i\rangle\langle i|, & P_2 &= \sum_{i=d_1+1}^d |i\rangle\langle i|, \\ Q_1 &= \sum_{\alpha=1}^{m_1} |\varphi_\alpha\rangle\langle \varphi_\alpha|, & Q_2 &= \sum_{\alpha=m_1+1}^d |\varphi_\alpha\rangle\langle \varphi_\alpha|. \end{aligned} \quad (134)$$

a set of Kraus operators for this map is given by

$$K_{i,\alpha} := \frac{1}{\sqrt{d_1}} |i\rangle\langle \varphi_\alpha| \quad i = 1, \dots, d_1, \quad \alpha = 1, \dots, m_1 \quad (135)$$

and

$$L_{i,\alpha} := \frac{1}{\sqrt{d_2}} |i\rangle\langle \varphi_\alpha| \quad i = d_1 + 1, \dots, d, \quad \alpha = m_1 + 1, \dots, d. \quad (136)$$

Note that

$$\sum_{i,\alpha} K_{i,\alpha}^\dagger K_{i,\alpha} + \sum_{i,\alpha} L_{i,\alpha}^\dagger L_{i,\alpha} = I_d, \quad (137)$$

but

$$\sum_{i,\alpha} K_{i,\alpha} K_{i,\alpha}^\dagger + \sum_{i,\alpha} L_{i,\alpha} L_{i,\alpha}^\dagger = \frac{m_1}{d_1} P_1 + \frac{m_2}{d_2} P_2, \quad (138)$$

hence the channel is non-unital, unless $m_1 = d_1$ and $m_2 = d_2$.

From the above relations and the fact that $P_1 + P_2 = Q_1 + Q_2 = I$, we find

$$P_1 = \frac{1}{d} (d_1 I + \sqrt{d_1 d_2} B) \quad P_2 = \frac{1}{d} (d_2 I - \sqrt{d_1 d_2} B) \quad (139)$$

and

$$Q_1 = \frac{1}{d} (m_1 I + \sqrt{m_1 m_2} A) \quad Q_2 = \frac{1}{d} (m_2 I - \sqrt{m_1 m_2} A) \quad (140)$$

Consider now a state

$$\rho = \frac{1}{d} (1 + xA) \quad (141)$$

where x represents the Bloch vector of this state, since $x = \text{Tr}(\rho A)$. We will then find

$$\mathcal{E}(\rho) = \frac{1}{d}(I + x'B), \quad (142)$$

where

$$x' = \sqrt{\frac{m_1 m_2}{d_1 d_2}} x + \frac{m_1 d_2 - m_2 d_1}{d \sqrt{d_1 d_2}}. \quad (143)$$

Straightforward calculation completes the reasoning. Hence we find that if $m_1 m_2 > d_1 d_2$, then the Bloch vector can be stretched. It is interesting to note that the inhomogeneous translation does not compensate the stretching of x , rather it enhances it. Obviously, this is not possible in dimensions $d = 2, 3$ but it is possible in dimensions $d \geq 4$, where for example we can take $(m_1, m_2) = (2, 2)$ and $(d_1, d_2) = (1, 3)$. Moreover we see that if the channel is unital, then $x' = x$.

To see how the existence of these channels may affect quasi-inversion, let us denote by \mathcal{M}_s a subset of quantum channels whose distortion matrices M_s can only shrink the Bloch vector, i.e. $\mathcal{M}_s = \{\mathcal{E}_s | M_s^T M_s \leq I\}$. Let us highlight two remarks related to the set \mathcal{M}_s . First, according to Russo-Dye theorem any linear positive map obtains its norm at the identity [54]. Thus, singular values of any unital map are less than or equal to one which implies all unital maps belong to \mathcal{M}_s . Moreover, for any M_s included in the set \mathcal{M}_s it is always possible to find unitary operators U_1 and U_2 such that $M_s = (U_1 + U_2)/2$. For now, let us assume that quasi-inverse for a quantum channel \mathcal{E} belongs to \mathcal{M}_s . In that case, one has (29)

$$\begin{aligned} \overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) &= \max_{\mathcal{E}'} \overline{F}(\mathcal{E}' \circ \mathcal{E}) = \max_{\mathcal{E}_s \in \mathcal{M}_s} \overline{F}(\mathcal{E}_s \circ \mathcal{E}) \\ &= \max_{M_s} \frac{1}{d} \left(1 + \frac{1}{d+1} \text{Tr} M_s M \right) \\ &\leq \max_{U_1, U_2} \frac{1}{d} \left(1 + \frac{1}{2(d+1)} \text{Tr}[(U_1 + U_2)M] \right) \\ &= \frac{1}{d} \left(1 + \frac{1}{d+1} \text{Tr}|M| \right). \end{aligned} \quad (144)$$

This inequality imposes an upper bound on the corrected fidelity whenever quasi-inversion lies in the set \mathcal{M}_s . An example of which is a qubit channel where not only quasi-inversion is a unital and unitary map in \mathcal{M}_s , but also the entire set of qubit channels are contained in \mathcal{M}_s . However, in higher dimensions \mathcal{M}_s is a nontrivial subset of quantum channels as our above example shows. It is a question then whether quasi-inversion of any quantum channel belongs to the set \mathcal{M}_s . In what follows, we show with an example it is possible to exceed the bound in (144), which implies a negative answer to the question. Moreover, as the bound (144) is valid when quasi-inversion is a unital map, one infers in higher dimensions, against qubit channel, quasi-inverse is not necessarily unital.

Before presenting the example, we mention explicitly taking $\Gamma_1 = \sqrt{d-1}A$ and $\Gamma_2 = \sqrt{d-1}B$, see Eq. (131), along with the identity and $d^2 - 3$ other Hermitian and traceless operators orthogonal to A and B as the set of basis, the distortion matrix M and the translation vector t of the channel \mathcal{E} in the beginning of this section are given by their entries as

$$M_{ij} = \sqrt{\frac{m_1 m_2}{d_1 d_2}} \delta_{i2} \delta_{j1}, \quad t_i = \frac{m_1 d_2 - m_2 d_1}{d \sqrt{d-1} \sqrt{d_1 d_2}} \delta_{i2}. \quad (145)$$

Now assume $\lambda > 0$ is sufficiently small so the d -dimensional map \mathcal{E}' described by the following affine parameters is a quantum channel

$$M'_{ij} = \lambda \delta_{i1} \delta_{j2}, \quad t'_i = 0. \quad (146)$$

Note that $\text{Tr}(MM') = \sqrt{\frac{m_1 m_2}{d_1 d_2}} \lambda$. This amount is certainly larger than $\text{Tr}|M'| = \lambda$ if $m_1 m_2 > d_1 d_2$, i.e. if the affine matrix M can stretch the Bloch vector and \mathcal{E} is not contained in \mathcal{M}_s . So for the channel \mathcal{E}' specified in Eq. (146) we have through Eq. (29)

$$\bar{F}(\mathcal{E}'^{qi} \circ \mathcal{E}') \geq \frac{1}{d} \left[1 + \frac{\lambda}{d+1} \max \sqrt{\frac{m_1 m_2}{d_1 d_2}} \right]. \quad (147)$$

For an even d one has $\max \sqrt{\frac{m_1 m_2}{d_1 d_2}} = \frac{d}{2\sqrt{d-1}}$, while $\max \sqrt{\frac{m_1 m_2}{d_1 d_2}} = \frac{\sqrt{d+1}}{2}$ for an odd d . In any case, assuming $d > 3$ we get

$$\bar{F}(\mathcal{E}'^{qi} \circ \mathcal{E}') \geq \frac{1}{d} \left[1 + \frac{\lambda}{2\sqrt{d+1}} \right] > \frac{1}{d} \left[1 + \frac{\lambda}{d+1} \right] = \frac{1}{d} \left[1 + \frac{\text{Tr}|M'|}{d+1} \right]. \quad (148)$$

D Statistical Properties of Classical Channels and their quasi-inverses

In this section we discuss some of the properties of classical channels and their inverses. A random classical channel is a stochastic matrix T in which the columns are independently picked from the uniform ensemble over the probabilistic simplex

$$\Delta_d \equiv \{ \mathbf{p} \in \mathbb{R}^d \mid p_i \geq 0; \sum_i p_i = 1 \} \quad (149)$$

Let the Probability Distribution Function (PDF) for any component p_i be denoted by $f_d(x)$, that is $\mathbb{P}[x \leq p_i \leq x + dx] = f_d(x)dx \quad \forall i$. Noting as an example figure (8) for the $d = 3$ case, we see that this PDF is proportional to the area of the narrow slab on the triangle which defines

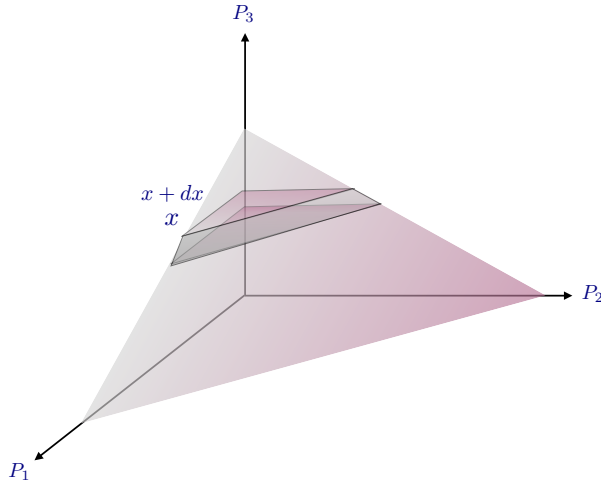


Figure 8: A probability simplex in 3 dimensions. Each point on the large triangle represents a valid probability vector.

For the general case, this function is given by

$$\phi_d(x) \equiv (d-1)(1-x)^{d-2}; \quad x \in [0, 1]. \quad (150)$$

It is also useful to have an expression for the corresponding CDF

$$\Phi_d(x) = \mathbb{P}[p_i \leq x] = \int_0^x \phi_d(t) dt = 1 - (1-x)^{d-1}. \quad (151)$$

The average fidelity for any classical channel T is given by $\overline{\mathcal{F}}(T) \equiv \frac{1}{d} \text{Tr} T$. If we now take the average over the uniform ensemble of all stochastic matrices, we find

$$\langle \overline{\mathcal{F}} \rangle = \frac{1}{d} \sum_i \langle T_{ii} \rangle = \int_0^1 (d-1)(1-x)^{d-2} x dx = \frac{1}{d}, \quad (152)$$

where we have used the statistical independence of the columns of T in the uniform ensemble. Moreover we can also find the variance of this average fidelity which turns out to be

$$\text{Var } \overline{\mathcal{F}} = \frac{1}{d^2} \sum_i \text{Var } T_{ii} = \frac{d-1}{d^3(d+1)}. \quad (153)$$

The interesting question is how much this average fidelity increases for classical channels when we apply the quasi-inversion. To find this we note that the average fidelity after quasi-inversion is

$$\overline{\mathcal{F}}(T^{qi} \circ T) = \frac{1}{d} \sum_i \max_j T_{ij}, \quad (154)$$

that is the average fidelity after quasi-inversion is the average of the maximum element in each row of T . Assuming that the largest element of two different rows do not occur on the same column (which can happen only for a subset of measure zero in the ensemble), the ensemble average of the improved average fidelity becomes

$$\langle \overline{\mathcal{F}}(T^{qi} \circ T) \rangle = \langle p_m \rangle, \quad (155)$$

where p_m is the largest element in a single probability vector chosen uniformly. To find the average of this quantity for the uniform ensemble, we invoke the proposition that only on a set of measure zero, the maximum values of two different rows may occur on the same column. Therefore we first find the following Cumulative Distribution Function (CDF), i.e. the probability that the maximum values in all columns are less than x

$$\Psi_d(x) := \mathbb{P}[p_m \leq x, \forall j] = \Phi_d(x)^d = [1 - (1-x)^{d-1}]^d$$

this leads to the following PDF, i.e. the probability that the maximum value is between x and $x + dx$,

$$\psi_d(x) = \frac{d\Psi_d(x)}{dx} = d(d-1)[1 - (1-x)^{d-1}]^{d-1} (1-x)^{d-2}$$

Now it is easy to see that

$$\begin{aligned} \langle \overline{\mathcal{F}}(T^{qi} \circ T) \rangle = \langle p_m \rangle &= 1 - \langle 1 - p_m \rangle = 1 - d(d-1) \int_0^1 dx [x(1-x^{d-1})]^{d-1} \\ &= 1 + \sum_{n=0}^d \frac{(-1)^n (d-1)n}{(d-1)n+1} \binom{d}{n}. \end{aligned} \quad (156)$$

Figure 9 shows the average fidelity after quasi-inversion versus dimension. The curve fits the equation

$$\langle \overline{\mathcal{F}}(T^{qi} \circ T) \rangle = cd^x \sim d^{-1/2}, \quad (157)$$

with $c \approx 1.055$ and the exponent $x \approx -0.5042$. Therefore, the average fidelity of a classical channel appears to be increasing due to quasi-inversion from $1/d$ to $1/\sqrt{d}$.

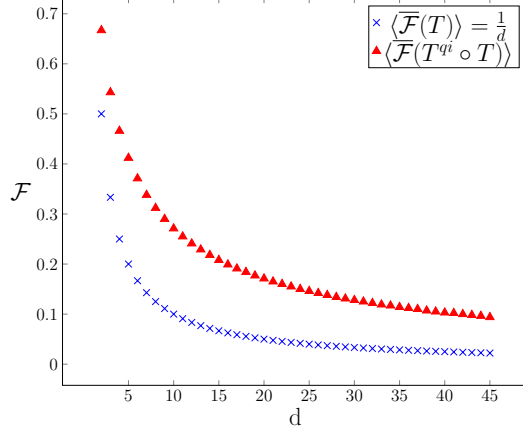


Figure 9: The average fidelity of a classical channel before $\langle \overline{\mathcal{F}}(T) \rangle$ (blue cross) and after $\langle \overline{\mathcal{F}}(T^{qi} \circ T) \rangle$ (red triangle) it is compensated by the quasi-inversion.

E A note on statistical Properties of Quantum Channels and their quasi-inverses

Investigating the same problem in the set of quantum channels seems a highly difficult task since in quantum case the quasi-inversion is not generally known. Nonetheless, numerical analysis is possible in lower dimensions and it shows improvement in amount of fidelity for a typical channel. To see that, one may notice for random quantum channels distributed uniformly in the set of quantum channels [44]

$$\langle \Phi_{\mathcal{E}} \rangle = \Phi_*, \quad (158)$$

where Φ_* is the maximally depolarizing channel, see Section 2. This implies $\langle \text{Tr} \Phi_{\mathcal{E}} \rangle = 1$. Thus, the average amount of input-output fidelity over random channels distributed uniformly in d dimension is given by (23)

$$\langle \overline{F}(\mathcal{E}) \rangle = \frac{1}{d}. \quad (159)$$

To get the fidelity after correction averaged over the set of uniformly distributed quantum channels, we applied numerically searching for the quasi-inverse. The sketch of our method is based on the observation that any quantum channel $\mathcal{E} : D(H_d) \longrightarrow D(H_d)$ should satisfy a continuous family of inequalities

$$\langle \phi | (\mathcal{E} \otimes I)(|\psi\rangle\langle\psi|) | \phi \rangle \geq 0 \quad |\phi\rangle, |\psi\rangle \in H_d \otimes H_d. \quad (160)$$

To approximate the quasi-inverse of a channel numerically, we first approximate the set of CP operators by only including a finite subset of such constraints obtained by random entangled states in $H_d \otimes H_d$. Then the quasi-inverse will be calculated as the answer to a finite linear programming problem using the simplex method or other efficient algorithms. The results for dimensions $d = 2, 3, 4, 5$ are presented in Fig. 10 and they confirm improvement in the average fidelity after correction for random channels.

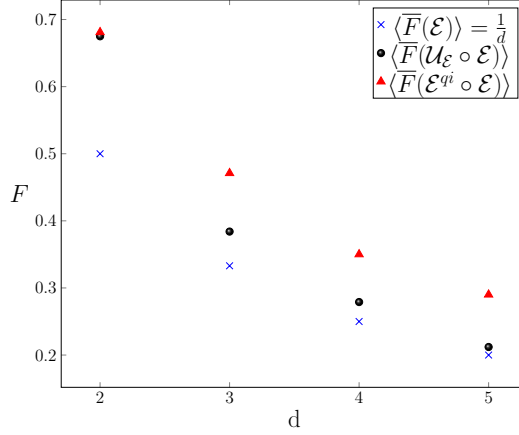


Figure 10: The amount of input-output fidelity $\langle \overline{F}(\mathcal{E}) \rangle$ (blue cross), fidelity after correction with the best possible unitary $\langle \overline{F}(\mathcal{U}_{\mathcal{E}} \circ \mathcal{E}) \rangle$ (black circle), and fidelity after correction with quasi-inversion $\langle \overline{F}(\mathcal{E}^{qi} \circ \mathcal{E}) \rangle$ (red triangle) averaged over the set of random channels distributed uniformly for $d = 2, \dots, 5$.

One may also think about applying a unitary evolution to correct the input-output fidelity. However, our numerical results show that the best unitary quantum channel, which will be denoted by $\mathcal{U}_{\mathcal{E}}$, cannot modify the fidelity in the best possible way, see Fig. 10 for a comparison. Indeed, we can take the purity of Jamiolkowski state as a signature of unitarity of a given channel \mathcal{E} , i.e. this quantity is equal to 1 if and only if the channel is unitary and any deviation from 1 shows non-unitarity. Adopting such a function, one finds unitarity for the quasi-inverse of random channels averaged over the set of uniformly distributed CPT maps is equal to 0.99 ± 0.01 for $d = 3$, 0.68 ± 0.05 for $d = 4$, and 0.52 ± 0.05 for $d = 5$. This interesting fact counter-intuitively shows in higher dimensions for a typical channel quasi inversion is almost a non-unitary channel. However, if we measure unitality by $1 - |\mathbf{t}|^2$ (\mathbf{t} is the translation vector, see (9)) averaged over the set of random channels, we see for $d = 3, 4, 5$ it is respectively given by: 0.98 ± 0.01 , 0.97 ± 0.01 , and 0.97 ± 0.01 , which confirms that the quasi-inverse of a typical channel is close to be unital.