New Journal of Physics

The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft **DPG**

IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

PAPER

CrossMark

Coherifying quantum channels

Kamil Korzekwa[1,5] (i), Stanisław Czachórski[2], Zbigniew Puchała[2,3] and Karol Życzkowski[2,4]

[1] Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia
[2] Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland
[3] Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, 44-100 Gliwice, Poland
[4] Center for Theoretical Physics, Polish Academy of Sciences, 02-668 Warszawa, Poland
[5] Author to whom any correspondence should be addressed.
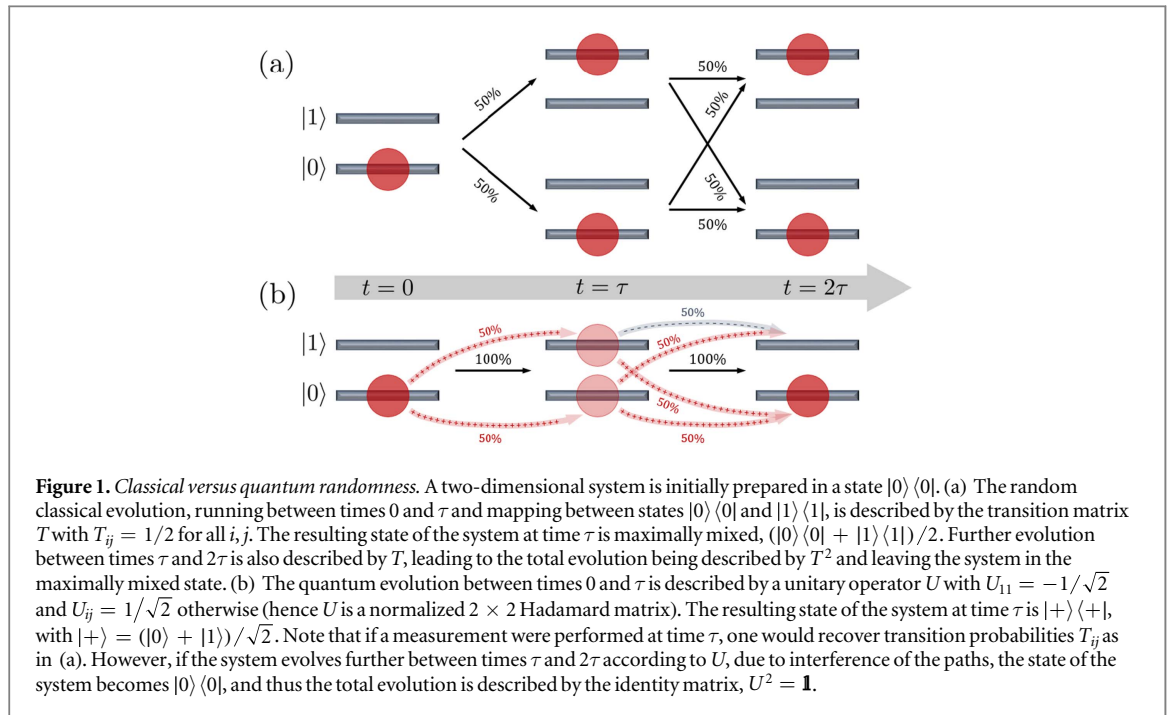
E-mail: kamil.korzekwa@sydney.edu.au

Abstract

Is it always possible to explain random stochastic transitions between states of a finite-dimensional system as arising from the deterministic quantum evolution of the system? If not, then what is the minimal amount of randomness required by quantum theory to explain a given stochastic process? Here, we address this problem by studying possible *coherifications* of a quantum channel Φ, i.e., we look for channels $\Phi^{\mathcal{C}}$ that induce the same classical transitions T, but are 'more coherent'. To quantify the coherence of a channel Φ we measure the coherence of the corresponding Jamiołkowski state J_Φ. We show that the classical transition matrix T can be coherified to reversible unitary dynamics if and only if T is unistochastic. Otherwise the Jamiołkowski state $J_\Phi^{\mathcal{C}}$ of the optimally coherified channel is mixed, and the dynamics must necessarily be irreversible. To assess the extent to which an optimal process $\Phi^{\mathcal{C}}$ is indeterministic we find explicit bounds on the entropy and purity of $J_\Phi^{\mathcal{C}}$, and relate the latter to the unitarity of $\Phi^{\mathcal{C}}$. We also find optimal coherifications for several classes of channels, including all one-qubit channels. Finally, we provide a non-optimal coherification procedure that works for an arbitrary channel Φ and reduces its rank (the minimal number of required Kraus operators) from d^2 to d.

1. Introduction

Random processes are ubiquitous in both classical and quantum physics. However, the nature of randomness in these two regimes differs significantly. On the one hand, classical random evolution is necessarily irreversible. On the other hand, quantum evolution may be completely deterministic (and thus reversible if no measurement is performed), but nevertheless lead to random measurement outcomes of observable A by transforming a system into a coherent superposition of eigenstates of A. When probing the dynamics of the system one can therefore observe the same random transitions, irrespectively of whether the evolution is coherent or incoherent. The question then arises: to what extent an observed random transformation can be explained via the underlying deterministic and coherent process, and how much unavoidable classical randomness must be involved in it?

To formulate this problem more precisely, consider a d-dimensional physical system undergoing some unknown evolution. In order to characterize it, one first measures the system, finding it in some well-defined state j, e.g., an eigenstate of observable A. One then allows the system to evolve for time τ and performs the same measurement again, this time finding the system in state i. By repeating this procedure many times and collecting the statistics of measurement outcomes, one can reconstruct the transition matrix T, with entries T_{ij} describing transition probabilities between states j and i. Now, for a truly random classical process, repeating it (e.g., by letting the system evolve for 2τ instead of τ) leads to the evolution described by T^2. We illustrate this in figure 1(a) for an exemplary two-dimensional system. However, in quantum physics, different transitions (paths) of T can interfere with each other, so that the composition of two processes will generally not be described by a transition matrix T^2. In particular, the compound process can even become fully deterministic, leading to the complete disappearance of the observed randomness (see figure 1(b)). Our question can then be

Figure 1. *Classical versus quantum randomness.* A two-dimensional system is initially prepared in a state $|0\rangle\langle0|$. (a) The random classical evolution, running between times 0 and τ and mapping between states $|0\rangle\langle0|$ and $|1\rangle\langle1|$, is described by the transition matrix T with $T_{ij} = 1/2$ for all i, j. The resulting state of the system at time τ is maximally mixed, $(|0\rangle\langle0| + |1\rangle\langle1|)/2$. Further evolution between times τ and 2τ is also described by T, leading to the total evolution being described by T^2 and leaving the system in the maximally mixed state. (b) The quantum evolution between times 0 and τ is described by a unitary operator U with $U_{11} = -1/\sqrt{2}$ and $U_{ij} = 1/\sqrt{2}$ otherwise (hence U is a normalized 2×2 Hadamard matrix). The resulting state of the system at time τ is $|+\rangle\langle+|$, with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Note that if a measurement were performed at time τ, one would recover transition probabilities T_{ij} as in (a). However, if the system evolves further between times τ and 2τ according to U, due to interference of the paths, the state of the system becomes $|0\rangle\langle0|$, and thus the total evolution is described by the identity matrix, $U^2 = \mathbf{1}$.

rephrased as: what is the optimal *coherification* of the random process described by a classical transition matrix T?

A more formal motivation for our studies comes from the resource-theoretic approach to quantum information. To explain it, let us first consider a simpler and better-known problem: how coherent is a given quantum state ρ of a d-dimensional system, and to what extent can it be transformed into a 'more coherent' state? Note that we do not refer here to the notion of spin-coherent states, which does not depend on the choice of basis [1], but rather to a more recent concept of coherence with respect to a given basis [2], distinguished for instance by the eigenbasis of the system's Hamiltonian. Any state represented by a density matrix that is diagonal in the preferred basis is incoherent in that sense, as it corresponds to a statistical mixture of classical states. On the other hand, a quantum state whose non-diagonal elements, called *coherences*, do not vanish may lead to non-classical effects of quantum interference. However, a generic system-environment interaction leads to the process of *decoherence*, due to which the off-diagonal entries tend to zero and the state becomes classical.

From the perspective of emerging quantum information technologies, coherence can be treated as a resource [3] allowing one to perform tasks impossible otherwise. It is then crucial to assess which quantum states are more valuable, i.e., have more coherence. One is thus confronted with the problem of quantifying coherence [2], which effectively means ordering the set of quantum states according to their coherence properties. Several competing measures of coherence of a quantum state were recently discussed in the literature, e.g., the l_1 norm of the off-diagonal entries of a density matrix or the relative entropy between a state and its decohered version (for a comprehensive review see [4] and references therein). Note that the diagonal density matrices, appearing as a result of the process of decoherence, are situated at the very bottom of this ordering, as they are classical and do not carry any coherence at all.

The problem of quantifying coherence has been taken a step farther by focusing on *cohering power* of quantum channels [5], i.e., studying the degree to which a quantum map can create coherence in an initially incoherent state. One can analyze the maximal or the average gain of coherence, where the average is taken over a suitable set of incoherent states. Such an approach is applicable to unitary transformations [6, 7] and to non-unitary operations [8–10], and in this way quantum channels can be ordered depending on their power to create coherence.

Within this approach, however, quantum states and their coherence are still the central objects of interest. Here, we take an alternative path and make quantum channels themselves the main focus of our study. Employing the well known Jamiołkowski–Choi isomorphism [11, 12], i.e., the fact that every quantum channel Φ is isomorphic to a bipartite quantum state J_Φ [13], we propose to apply the measures of coherence to bipartite states associated with a given channel. This way we can quantitatively investigate the problem posed at the beginning of this section: how coherent can a given random transformation be? More precisely, for a given stochastic transition matrix T we look for the maximally coherent quantum channel Φ^C, which under complete decoherence collapses to T, so that the diagonal parts of both Jamiołkowski states are equal. We first prove that a channel whose classical action is described by a transition matrix T can be coherified to a reversible unitary transformation if and only if T is unistochastic. We then derive general upper and lower bounds on the optimal

coherification of a channel described by a non-unistochastic T. Finally, we construct optimally coherified maps for any one-qubit channel and certain classes of channels acting on higher dimensional systems.

The paper is organized as follows. In section 2 we set the scene by introducing necessary concepts concerning the coherence and mixedness of quantum states and channels, and formulate the optimal coherification problem. General limitations for coherifying quantum channels are then derived and analyzed in section 3, where we also study particular families of maps in detail. In section 4 we discuss physical interpretation of coherence and purity of a channel, and relate these quantities to unitarity [14] and cohering power [5]. Concluding remarks are presented in the final section 5, while some technical results, predominantly concerning channels acting on two- and three-dimensional spaces, are relegated to appendices A–D.

2. Setting the scene

2.1. Coherence and mixedness of quantum states

A state of a finite-dimensional quantum system is described by a density operator ρ acting on a d-dimensional Hilbert space $\mathcal{H}_d$ that is positive, $\rho \geqslant 0$, and normalized by a trace condition, $\mathrm{Tr}(\rho) = 1$. The convex set of all density matrices of size d, denoted by $\mathcal{M}_d$, has $d^2 - 1$ dimensions and contains the $(d - 1)$-dimensional simplex $\mathcal{P}_d$ of normalized probability vectors of length d. By $\boldsymbol{\lambda}(\rho)$ we will denote the probability vector with entries given by the eigenvalues of ρ arranged in a non-increasing order.

A state is pure if $\rho = \rho^2$ (equivalently if $\boldsymbol{\lambda}(\rho) = [1, 0, \ldots, 0]$), so it can be represented by a one-dimensional projector, $\rho = |\psi\rangle \langle\psi|$; and mixed otherwise. Typical measures used to quantify the degree of mixedness of a given state [15, 16] include the von Neumann entropy[6],

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho) = -\sum_i \lambda_i(\rho) \log \lambda_i(\rho), \tag{1a}$$

and purity,

$$\gamma(\rho) = \mathrm{Tr}(\rho^2) = \sum_{i,j} |\rho_{ij}|^2 = \boldsymbol{\lambda}(\rho) \cdot \boldsymbol{\lambda}(\rho). \tag{1b}$$

Note that, since the above measures are directly related to the eigenvalues of ρ, they are unitarily invariant, and thus the mixedness of a quantum state is preserved under unitary dynamics.

On the other hand, in order to study coherence of $\rho \in \mathcal{M}_d$ one first needs to specify a basis with respect to which the coherence is measured [2]. This basis may be distinguished by the problem under study, e.g., within quantum thermodynamics one is mainly concerned with superposition of energy eigenstates [17, 18]. Here, however, we will study the problem in a general quantum information context, and thus we will simply fix an orthonormal basis $\{|i\rangle\}_{i=1}^d$. We say that a state is incoherent, or classical, when it is diagonal in the chosen basis. Classical states can be alternatively represented by a probability distribution $\boldsymbol{p} = \mathrm{diag}(\rho)$, where $\mathrm{diag}(\rho)$ denotes a mapping of a density matrix ρ into a probability vector $\boldsymbol{p} \in \mathcal{P}_d$ with $p_i = \rho_{ii}$. With a slight abuse of the notation we will also write $\rho \in \mathcal{P}_d$ if ρ is diagonal. Before we introduce measures of coherence, let us first define a *completely decohering* quantum channel $\mathcal{D}$,

$$\mathcal{D}(\rho) = \rho^{\mathcal{D}} := \sum_i \langle i| \, \rho |i\rangle \, |i\rangle \, \langle i|. \tag{2}$$

Note that under the action of $\mathcal{D}$ any quantum state ρ undergoes complete decoherence and becomes diagonal in the preferred basis. Thus, $\rho^{\mathcal{D}}$ and the associated probability distribution $\boldsymbol{p} = \mathrm{diag}(\rho)$ can be considered as the classical version of a general quantum state ρ. Notice also that $\mathcal{D}$ is a projector onto $\mathcal{P}_d$ as $\rho^{\mathcal{D}} \in \mathcal{P}_d$ and $\mathcal{D}(\rho^{\mathcal{D}}) = \rho^{\mathcal{D}}$ for all ρ.

The problem of quantifying the amount of coherence present in a state has been addressed in [2], while an earlier work [19] was devoted to quantifying quantum superposition. Two particular measures of coherence that we will focus on in this work are the relative entropy of coherence,
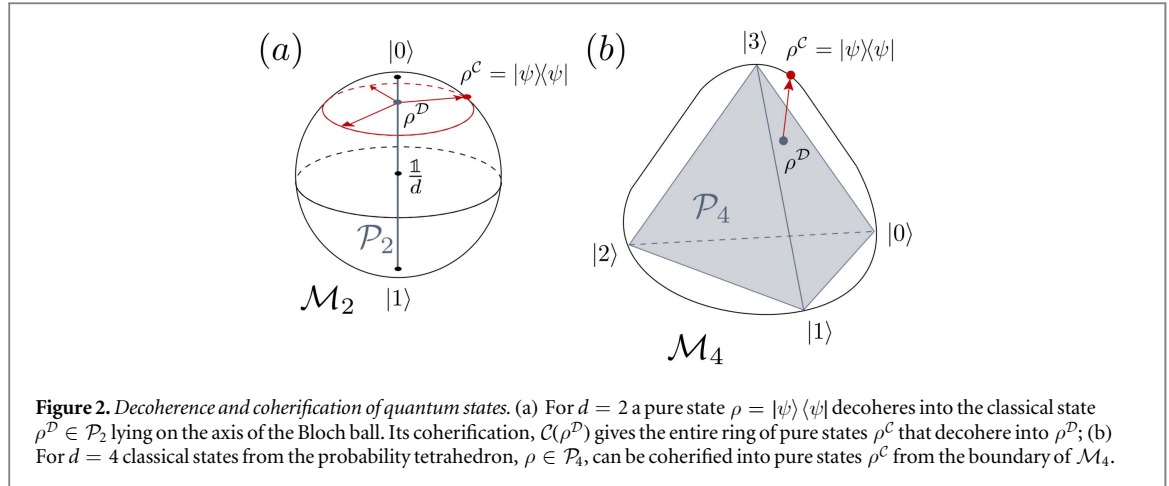
$$\mathcal{C}_e(\rho) := S(\rho||\rho^{\mathcal{D}}) = S(\rho^{\mathcal{D}}) - S(\rho) = S(\boldsymbol{p}) - S(\boldsymbol{\lambda}(\rho)), \tag{3a}$$

and the 2-norm of coherence[7]

$$\mathcal{C}_2(\rho) := \|\rho - \rho^{\mathcal{D}}\|_2 = \gamma(\rho) - \gamma(\rho^{\mathcal{D}}) = \boldsymbol{\lambda}(\rho) \cdot \boldsymbol{\lambda}(\rho) - \boldsymbol{p} \cdot \boldsymbol{p}. \tag{3b}$$

---

[6] Within this work we will use $S(\cdot)$ to denote both the von Neumann entropy of a density matrix, as well as the Shannon entropy of a probability distribution.

[7] We note that from the resource-theoretic perspective [2] $\mathcal{C}_2$ does not strictly satisfy all desirable requirements for a coherence measure. While it is true that under incoherent CPTP maps the 2-norm of coherence is non-increasing, it can increase on average under selective measurements. However, in our study of coherence of quantum channels, the resource-theoretic constraints have no clear physical meaning, and thus $\mathcal{C}_2$ is a completely legitimate measure of coherence.

Figure 2. *Decoherence and coherification of quantum states.* (a) For $d = 2$ a pure state $\rho = |\psi\rangle\langle\psi|$ decoheres into the classical state $\rho^{\mathcal{D}} \in \mathcal{P}_2$ lying on the axis of the Bloch ball. Its coherification, $\mathcal{C}(\rho^{\mathcal{D}})$ gives the entire ring of pure states $\rho^{\mathcal{C}}$ that decohere into $\rho^{\mathcal{D}}$; (b) For $d = 4$ classical states from the probability tetrahedron, $\rho \in \mathcal{P}_4$, can be coherified into pure states $\rho^{\mathcal{C}}$ from the boundary of $\mathcal{M}_4$.

It is evident that the measures of coherence are directly related to the measures of mixedness. More precisely, the relative entropy of coherence is the difference between the entropy of a classical version of a state and the quantum state itself; and the 2-norm of coherence is the difference between the purity of a quantum state and the purity of its classical version.

Among the family of states with a fixed spectrum, i.e., belonging to a unitary orbit, the measures of mixedness are equal, but the measures of coherence vary significantly. The minimal coherence, equal to zero, is obtained for the diagonal state $U^\dagger \rho U$, where U is the unitary matrix containing the eigenvectors of ρ. The maximal coherence is achieved by the contradiagonal state [20], $\rho^{\mathrm{cont}} = HU^\dagger\rho UH^\dagger$, where H is a Fourier matrix (or, more generally, a complex Hadamard matrix [21]), which is unitary and has entries with the same modulus, $|H_{ij}|^2 = 1/d$. Since all diagonal elements of the contradiagonal state are equal, $\rho_{ii}^{\mathrm{cont}} = 1/d$, one gets $\mathcal{C}_{\mathrm{e}}(\rho^{\mathrm{cont}}) = \log d - S(\rho)$ and $\mathcal{C}_2(\rho^{\mathrm{cont}}) = \gamma(\rho) - 1/d$.

On the other hand, among the family of states with a fixed diagonal, i.e., quantum states that under the action of $\mathcal{D}$ decohere to the same classical state, both mixedness and coherence measures vary. However, they are maximized and minimized by the same states, which can be directly inferred from equations (3a), (3b). The minimum can be obtained by acting with the decohering channel $\mathcal{D}$ (that leaves the diagonal unchanged) on any member of the family, leading to zero coherence. In a similar fashion we can define an optimal *coherifying* transformation $\mathcal{C}$ (which should not be confused with coherence measures) that maps any member of the family into a state that maximizes purity (and thus coherence),

$$\mathcal{C}(\rho) = \rho^{\mathcal{C}} = |\psi\rangle\langle\psi|, \quad \text{such that } \mathrm{diag}(\rho^{\mathcal{C}}) = \mathrm{diag}(\rho). \tag{4}$$

This problem has a simple solution for any mixed state ρ. Identifying its diagonal elements with components of a probability vector $\boldsymbol{p}$, one can write explicitly a family of optimally coherifying transformations

$$\rho \rightarrow \rho^{\mathcal{C}}: \quad \rho_{ij}^{\mathcal{C}} = \sqrt{p_i p_j} \, e^{i(\phi_i - \phi_j)}, \tag{5}$$

where the phases, $\phi_i \in [0, 2\pi)$, are arbitrary. The mixedness of such coherified states is zero and coherence achieves its maximal value, $\mathcal{C}_{\mathrm{e}}(\rho^{\mathcal{C}}) = S(\boldsymbol{p})$ and $\mathcal{C}_2(\rho^{\mathcal{C}}) = 1 - \boldsymbol{p} \cdot \boldsymbol{p}$. In our work we will also refer to non-optimal coherification transformations that map a diagonal state into a state with the same diagonal but some non-zero off-diagonal terms. In figure 2 we illustrate the ideas of decoherence and coherification of quantum states using low-dimensional examples.

Let us emphasize that equation (4) does not describe a realistic physical process, but rather it provides an answer to a legitimate question concerning the possible past of an irreversible quantum dynamics. Such a fictitious coherification can be treated as a kind of a formal inverse of the process of decoherence. More precisely, for a diagonal state $\rho \in \mathcal{P}_d$ we have $\mathcal{D}(\mathcal{C}(\rho)) = \rho$; and for a pure state $|\psi\rangle\langle\psi|$ we have $\mathcal{C}(\mathcal{D}(|\psi\rangle\langle\psi|)) \sim |\psi\rangle\langle\psi|$, where the equivalence is up to phase factors of the off-diagonal terms.

Finally, note that coherification can be compared to the known procedure [15, 22] of *purification* of a quantum state. Any mixed quantum state can be purified at the expense of increasing the dimension of the Hilbert space. More precisely, for any state $\rho \in \mathcal{M}_d$ its purification is given by a pure state $|\psi_{AB}\rangle \in \mathcal{H}_d \otimes \mathcal{H}_d$ of the extended system, such that its partial trace reads $\mathrm{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|) = \rho$. The Schmidt vector of $|\psi_{AB}\rangle$ coincides with the spectrum of ρ, e.g., if the state ρ is maximally mixed, the state $|\psi_{AB}\rangle$ is maximally entangled. Both formal procedures are not unique and they allow one to find possible preimages of ρ with respect to non-invertible physical operations. Namely, purification yields states of an extended system which are transformed into ρ by a partial trace; and coherification of a state ρ provides states of the same size which decay into $\rho^{\mathcal{D}}$ due to decoherence.

2.2. Coherence and mixedness of quantum channels

In this work we generalize the notion of coherence and mixedness of quantum states to quantum channels, i.e., completely positive trace preserving (CPTP) maps acting on density matrices of order d. We will denote by $\mathcal{S}_d$ the set of all quantum channels, also called stochastic maps, acting on $\mathcal{M}_d$. Recall that for any $\Phi \in \mathcal{S}_d$ one can define the associated *Jamiołkowski state* [11], as the image of the extended map acting on a maximally entangled state,

$$J_\Phi = \frac{1}{d}(\Phi \otimes \mathcal{I})|\Omega\rangle\langle\Omega|, \tag{6}$$

with $|\Omega\rangle = \sum_i |ii\rangle$ and $\mathcal{I}$ denoting the identity channel. Note also that the Jamiołkowski state is proportional to the dynamical matrix of Choi [12], so that $J_\Phi = \Phi^R/d$. Here, with a slight abuse of the notation, Φ denotes the representation of the channel as a matrix of size d^2, i.e., a superoperator with entries in the preferred basis given by $\Phi_{ij,kl} = \mathrm{Tr}(|j\rangle\langle i| \Phi(|k\rangle\langle l|))$; and the reshuffling transformation, X^R, exchanges elements of a matrix in such a way that square blocks of size d after reshaping form rows of length d^2, so that $(X_{ij,kl})^R = X_{ik,jl}$—see [13] for further details. Finally, for every quantum channel there exists a Kraus decomposition [15], or the operator-sum representation, of the form

$$\Phi(\cdot) = \sum_i K_i(\cdot)K_i^\dagger, \tag{7}$$

where K_i are called Kraus operators. Due to trace preserving condition these satisfy $\sum_i K_i^\dagger K_i = \mathbf{1}$, where $\mathbf{1}$ denotes the identity matrix of size d.

The condition of $\Phi \in \mathcal{S}_d$ is equivalent to [11]

$$J_\Phi \geqslant 0, \tag{8a}$$

$$\mathrm{Tr}_1(J_\Phi) = \frac{\mathbf{1}}{d}. \tag{8b}$$

These conditions imply that diagonal elements (in the preferred basis) of the Jamiołkowski state correspond to the entries of a $d \times d$ (column-stochastic) transition matrix T,

$$\langle ij| J_\Phi |ij\rangle = \frac{1}{d}T_{ij}, \tag{9}$$

with $T_{ij} \geqslant 0$ and $\sum_i T_{ij} = 1$. The set of stochastic transition matrices of order d will be denoted by $\mathcal{T}_d$. This set of classical maps has $d(d-1)$ dimensions and can be embedded inside the set $\mathcal{S}_d$ of quantum maps with $d^4 - d^2$ dimensions [16]. Since the diagonal of J_Φ (up to a constant $1/d$ factor) is given by the elements of a transition matrix T, we will write $\mathrm{diag}(J_\Phi) = \frac{1}{d}|T\rangle\rangle$, where $|\cdot\rangle\rangle$ denotes the (row-wise) vectorization of a matrix,

$$|T\rangle\rangle = (T \otimes \mathbf{1})|\Omega\rangle, \tag{10}$$

which can be also written as

$$|T\rangle\rangle = [T_{11}, T_{12}, \ldots, T_{1d}, T_{21}, \ldots, T_{dd}]^\top, \tag{11}$$

with $\top$ denoting a transpose. We also define $\langle\langle T|$ by the Hermitian conjugate of the right-hand side of equation (10).

Before we define the coherence of a quantum channel Φ, let us first physically interpret the entries of the corresponding Jamiołkowski state J_Φ. Writing it in the matrix form in the distinguished basis we have the following block form

$$J_\Phi = \frac{1}{d}\begin{bmatrix} D^1 & C^{12} & \ldots & C^{1d} \\ C^{21} & D^2 & \ldots & C^{2d} \\ \vdots & \vdots & \ddots & \vdots \\ C^{d1} & C^{d2} & \ldots & D^d \end{bmatrix} \tag{12}$$

with

$$D^i = \begin{bmatrix} T_{i1} & c_{12}^{ii} & \ldots & c_{1d}^{ii} \\ c_{21}^{ii} & T_{i2} & \ldots & c_{2d}^{ii} \\ \vdots & \vdots & \ddots & \vdots \\ c_{d1}^{ii} & c_{d2}^{ii} & \ldots & T_{id} \end{bmatrix}, \quad C^{ij} = \begin{bmatrix} c_{11}^{ij} & c_{12}^{ij} & \ldots & c_{1d}^{ij} \\ c_{21}^{ij} & c_{22}^{ij} & \ldots & c_{2d}^{ij} \\ \vdots & \vdots & \ddots & \vdots \\ c_{d1}^{ij} & c_{d2}^{ij} & \ldots & c_{dd}^{ij} \end{bmatrix}, \tag{13}$$

where $c_{kl}^{ij} = \langle i| \Phi(|k\rangle\langle l|)|j\rangle$ and formally $T_{ij} = c_{jj}^{ii}$. We thus see that the diagonal elements of D^i describe how initial populations (occupations in the preferred basis) affect the final population of a state i; and the off-diagonal elements of D^i describe how initial coherences affect the final population of a state i. On the other hand, the diagonal elements of C^{ij} describe how initial populations affect the final coherence between states i and j; and the off-diagonal elements of C^{ij} describe how initial coherences affect the final coherence between states i and j.

In analogy with the standard completely decohering map, equation (2), we also define a decohering operation $\mathcal{D}$ which acts on any quantum channels Φ by bringing its corresponding Jamiołkowski state into the diagonal form,

$$\Phi \rightarrow \Phi^{\mathcal{D}}: \quad J_{\Phi^{\mathcal{D}}} = J_{\Phi}^{\mathcal{D}}. \tag{14}$$

Diagonal Jamiołkowski state $J_{\Phi}^{\mathcal{D}} \in \mathcal{M}_{d^2}$ represents the classical map $T \in \mathcal{T}_d$ acting on probability vectors of size d. The action of $\Phi^{\mathcal{D}}$ on any state ρ is first to completely decohere it into $\rho^{\mathcal{D}}$, and then to transform the probability vector $\boldsymbol{p} = \mathrm{diag}(\rho)$ into $T\boldsymbol{p}$, so that the final state is always classical. Therefore, again with a slight abuse of the notation, we will write $\Phi \in \mathcal{T}_d$ and $\Phi \sim T$ if J_{Φ} is diagonal and $\mathrm{diag}(J_{\Phi}) = \frac{1}{d}|T\rangle\rangle$.

As every quantum channel is isomorphic to a density matrix on an extended Hilbert space, $\mathcal{H}_d \otimes \mathcal{H}_d$, it is then natural to apply the standard measures of mixedness and coherence to the Jamiołkowski state J_{Φ} and to characterize in this way the properties of the associated channel. More formally, for any quantum channel Φ acting on quantum states of size d one defines the entropy of a channel [23],

$$S(\Phi) := S(J_{\Phi}), \tag{15a}$$

and the purity of a channel,

$$\gamma(\Phi) := \gamma(J_{\Phi}). \tag{15b}$$

These quantities allow us to introduce

1. *entropic coherence of a channel*,

$$\mathcal{C}_{\mathrm{e}}(\Phi) := \mathcal{C}_{\mathrm{e}}(J_{\Phi}) = S\left(\frac{1}{d}|T\rangle\rangle\right) - S(J_{\Phi}), \tag{16a}$$

and

2. *2-norm coherence of a channel*,

$$\mathcal{C}_2(\Phi) := \mathcal{C}_2(J_{\Phi}) = \gamma(J_{\Phi}) - \frac{1}{d^2}\langle\langle T|T\rangle\rangle = \mathrm{Tr}(J_{\Phi}^2) - \frac{1}{d^2}\mathrm{Tr}(TT^{\dagger}). \tag{16b}$$

Note that $\mathcal{C}_2(\Phi)$ can be decomposed into two terms,

$$\mathcal{C}_2(\Phi) = \mathcal{C}_2^D(\Phi) + \mathcal{C}_2^C(\Phi), \tag{17}$$
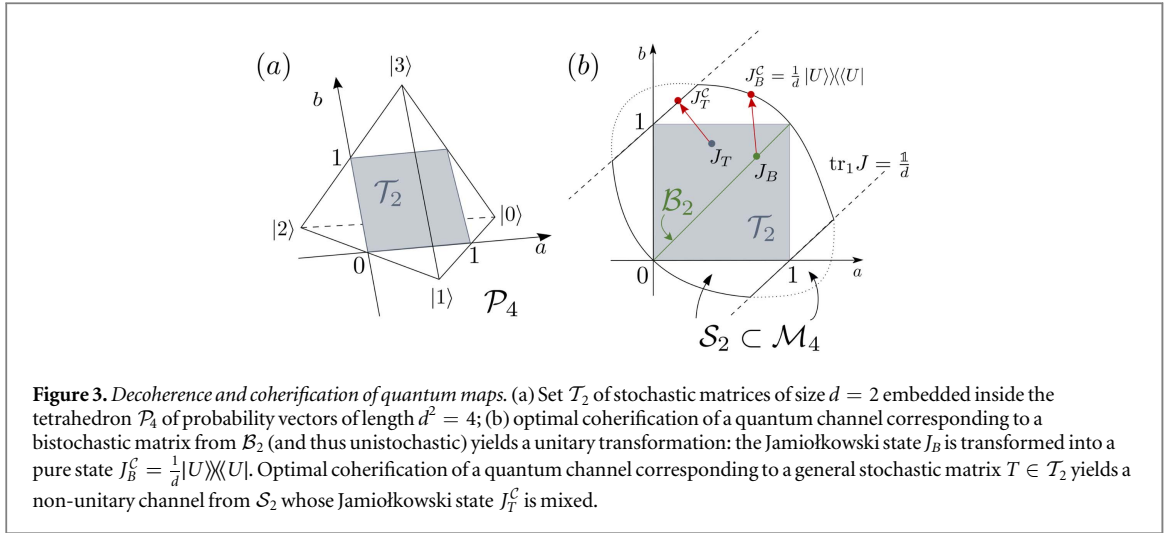
with $\mathcal{C}_2^D(\Phi)$ measuring coherence coming from diagonal blocks D^i and $\mathcal{C}_2^C(\Phi)$ from off-diagonal blocks C^{ij}, i.e.,

$$\mathcal{C}_2^D(\Phi) = \sum_{\substack{i,k,l \\ k \neq l}} |c_{kl}^{ii}|^2, \quad \mathcal{C}_2^C(\Phi) = \sum_{\substack{i,j,k,l \\ i \neq j}} |c_{kl}^{ij}|^2. \tag{18}$$

We now arrive at the central technical problem analyzed in the current work: coherifying quantum channels. Note that, given a fixed diagonal, $\mathrm{diag}(J_{\Phi}) = \frac{1}{d}|T\rangle\rangle$, of the Jamiołkowski state J_{Φ} (equivalently: a transition matrix T specifying the classical action of Φ, i.e., $\Phi^{\mathcal{D}} = T$), one can always find the corresponding coherified pure state by simply employing the optimal coherification recipe, equation (5). In general, however, such a pure state will not satisfy the trace preserving condition, equation (8b). More precisely, this condition is equivalent to $\sum_i D^i = \mathbb{1}$, and thus the choice of the off-diagonal elements of D^i (describing the effect of initial coherences on final populations) is constrained beyond the standard positivity constraint. Hence, for any classical map represented by a stochastic matrix T it is legitimate to ask the following question: what is the corresponding optimally coherified quantum map $\Phi^{\mathcal{C}}$ with the same classical part, such that its coherence is the largest (or its mixedness is the smallest). In figure 3 we illustrate the ideas of decoherence and coherification of quantum channels using one-qubit maps as an example.

3. Limitations of coherifying quantum channels

In this section we will investigate the limits to which a given quantum channel $\Phi \in \mathcal{S}_d$, with a prescribed classical action $\Phi^{\mathcal{D}} = T \in \mathcal{T}_d$, can be coherified into an optimal channel $\Phi^{\mathcal{C}}$ with minimal entropy or maximal purity. To characterize potential coherification of a given channel we will simply use both coherence measures introduced above in equations (16a) and (16b). We thus first define optimally coherified channels according to both measures,

Figure 3. *Decoherence and coherification of quantum maps.* (a) Set $\mathcal{T}_2$ of stochastic matrices of size $d = 2$ embedded inside the tetrahedron $\mathcal{P}_4$ of probability vectors of length $d^2 = 4$; (b) optimal coherification of a quantum channel corresponding to a bistochastic matrix from $\mathcal{B}_2$ (and thus unistochastic) yields a unitary transformation: the Jamiołkowski state J_B is transformed into a pure state $J_B^{\mathcal{C}} = \frac{1}{d}|U\rangle\!\rangle\langle\!\langle U|$. Optimal coherification of a quantum channel corresponding to a general stochastic matrix $T \in \mathcal{T}_2$ yields a non-unitary channel from $\mathcal{S}_2$ whose Jamiołkowski state $J_T^{\mathcal{C}}$ is mixed.

$$\Phi^{\mathcal{C}_e} := \operatorname*{argmax}_{\Psi:\, \Psi^{\mathcal{D}} = \Phi^{\mathcal{D}}} \mathcal{C}_e(\Psi), \tag{19a}$$

$$\Phi^{\mathcal{C}_2} := \operatorname*{argmax}_{\Psi:\, \Psi^{\mathcal{D}} = \Phi^{\mathcal{D}}} \mathcal{C}_2(\Psi), \tag{19b}$$

which allows us to define

1. *entropic coherification*,

$$\Delta\mathcal{C}_e(\Phi) := \mathcal{C}_e(\Phi^{\mathcal{C}_e}) - \mathcal{C}_e(\Phi), \tag{20a}$$

and

2. *2-norm coherification*,

$$\Delta\mathcal{C}_2(\Phi) := \mathcal{C}_2(\Phi^{\mathcal{C}_2}) - \mathcal{C}_2(\Phi). \tag{20b}$$

We will be particularly interested in the extremal case when the coherified channel is classical, $\Phi \sim T$. Then, since $\mathcal{C}_e(T) = \mathcal{C}_2(T) = 0$, we have

$$\Delta\mathcal{C}_e(T) = \mathcal{C}_e(\Phi^{\mathcal{C}_e}) = S\!\left(\frac{1}{d}|T\rangle\!\rangle\right) - S(J_\Phi^{\mathcal{C}_e}), \tag{21a}$$

$$\Delta\mathcal{C}_2(T) = \mathcal{C}_2(\Phi^{\mathcal{C}_2}) = \gamma(J_\Phi^{\mathcal{C}_2}) - \frac{1}{d^2}\langle\!\langle T|T\rangle\!\rangle. \tag{21b}$$
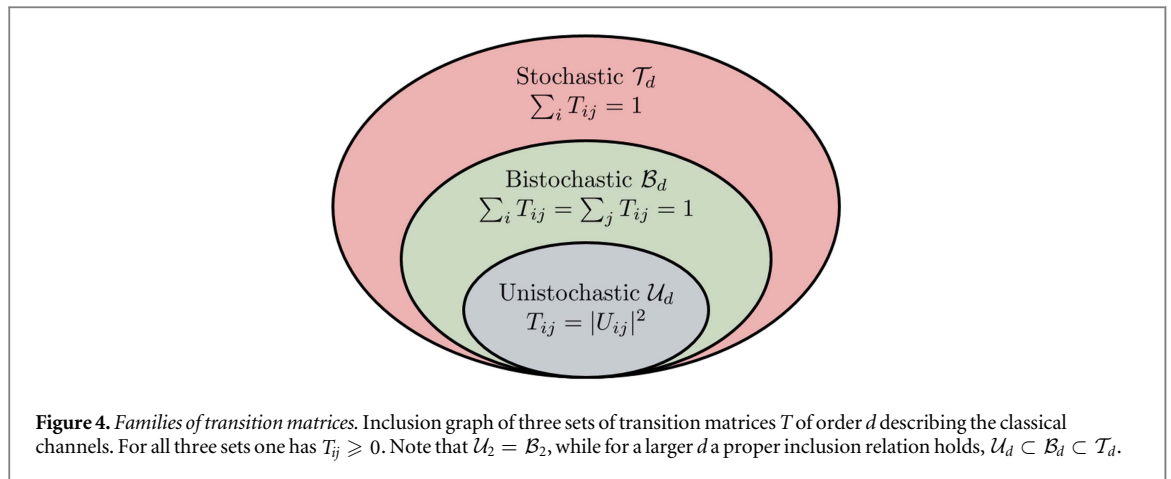
Note that, although we introduced two potentially inequivalent coherification procedures, $\Phi^{\mathcal{C}_e}$ and $\Phi^{\mathcal{C}_2}$ (with corresponding Jamiołkowski states $J_\Phi^{\mathcal{C}_e}$ and $J_\Phi^{\mathcal{C}_2}$), while deriving general bounds affecting both maximization processes, we will simply use $\Phi^{\mathcal{C}}$ (and $J_\Phi^{\mathcal{C}}$).

We now need to point out an important relation linking the classical action of a channel and its Kraus decomposition. Namely, for a channel Φ with a classical action $\Phi^{\mathcal{D}} = T$, the Kraus decomposition of Φ, defined in equation (7), satisfies

$$\sum_i K_i \circ \bar{K}_i = T, \tag{22}$$

with $\circ$ denoting the entry-wise product (also known as Hadamard or Schur product) and $\bar{K}_i$ being the complex conjugate of K_i.

The problem of optimal coherification of a channel naturally splits into three cases, corresponding to three families of transition matrices T presented in figure 4. The biggest family, $\mathcal{T}_d$, consists of all stochastic matrices of size d, i.e., the most general transformations mapping the set of d-dimensional probability vectors into itself. These are defined by $T_{ij} \geqslant 0$ and $\sum_i T_{ij} = 1$. The second family, $\mathcal{B}_d$, is given by the set of bistochastic matrices, which in addition to being stochastic satisfy $\sum_j T_{ij} = 1$. This additional condition encodes the fact that bistochastic matrices map the uniform distribution, $[1/d, \ldots, 1/d]$, into itself. The third analyzed case corresponds to unistochastic matrices $\mathcal{U}_d$, which are such bistochastic matrices T whose entries can be written as $T_{ij} = |U_{ij}|^2$, for some unitary matrix U. Note that this condition, using equation (22), can alternatively be written as $T = U \circ \bar{U}$. While bistochastic matrices form a proper subset of stochastic matrices for all $d \geqslant 2$, the

Figure 4. *Families of transition matrices.* Inclusion graph of three sets of transition matrices T of order d describing the classical channels. For all three sets one has $T_{ij} \geqslant 0$. Note that $\mathcal{U}_2 = \mathcal{B}_2$, while for a larger d a proper inclusion relation holds, $\mathcal{U}_d \subset \mathcal{B}_d \subset \mathcal{T}_d$.

unistochastic matrices form a proper subset of bistochastic matrices only for $d \geqslant 3$, as every bistochastic matrix of order $d = 2$ is unistochastic. Interestingly, the exact boundary of the set of unistochastic matrices is known only for $d \leqslant 3$ [24], while for $d \geqslant 3$ the set of unistochastic matrices is not convex [25].

3.1. Unistochastic matrices and unitary channels

We start our analysis from the smallest family of unistochastic matrices. We will thus consider optimal coherification of a channel for which $\mathrm{diag}(J_\Phi) = \frac{1}{d}|U \circ \bar{U}\rangle\rangle$. We say that a given channel can be completely coherified if there exists $J_\Phi^{\mathcal{C}}$ that is pure, has the same diagonal as J_Φ and still corresponds to a valid channel. From equation (6) it is clear that the Jamiołkowski state is pure if and only if the corresponding map Φ is unitary. This simple observation can then be formalized as follows:

Proposition 1. *A quantum channel Φ, with the corresponding Jamiołkowski state J_Φ, can be completely coherified to a unitary transformation if and only if its classical action is given by a unistochastic matrix.*

Proof. First assume that Φ can be completely coherified. This means that there exists a pure state $J_\Phi^{\mathcal{C}}$ with $\mathrm{diag}(J_\Phi^{\mathcal{C}}) = \mathrm{diag}(J_\Phi)$, and that it corresponds to a valid channel $\Phi^{\mathcal{C}}$. However, pure Jamiołkowski states correspond to unitary channels, so that (comparing equations (6) and (10))

$$J_\Phi^{\mathcal{C}} = \frac{1}{d}|U\rangle\rangle\langle\langle U| = \frac{1}{d}(U \otimes \bar{U})^R. \tag{23}$$

Therefore, the diagonal of $J_\Phi^{\mathcal{C}}$ (and, by assumption, of J_Φ) is given by $\frac{1}{d}|U \circ \bar{U}\rangle\rangle$, which corresponds to a unistochastic matrix. Conversely, assume that the diagonal of J_Φ is described by a unistochastic matrix $U \circ \bar{U}$. Then one can simply choose $J_\Phi^{\mathcal{C}}$ to be a pure state given in equation (23). □

Notice that every non-trivial classical stochastic dynamics is irreversible. However, if it is described by a unistochastic matrix T, one can find a reversible (unitary) quantum channel $\Phi^{\mathcal{C}}$ whose classical action is given by T. On the other hand, if the classical dynamics is not unistochastic, it cannot be completely coherified and made reversible. We will now show such a coherification procedure in action by analyzing some simple examples, and in the following section we will address the limits to which a general stochastic dynamics can be made reversible.

First, consider a transition matrix given by a permutation matrix Π of size d. A quantum channel Φ corresponding to a diagonal Jamiołkowski state with $\mathrm{diag}(J_\Phi) = \frac{1}{d}|\Pi\rangle\rangle$ is a completely decohering channel that permutes diagonal elements. The vector $\frac{1}{d}|\Pi\rangle\rangle$ of length d^2 has d non-zero entries equal to $1/d$, so its Shannon entropy is equal to $\log d$. However, as Π is unistochastic, it can be coherified to a unitary transformation corresponding to the Jamiołkowski state $J_\Phi^{\mathcal{C}} = \frac{1}{d}(\Pi \otimes \Pi)^R$ with zero entropy and $d(d-1)$ off-diagonal entries equal to $1/d$. Thus, the entropic coherification of any classical permutation matrix reads $\Delta\mathcal{C}_{\mathrm{e}}(\Pi) = \log d$, while the 2-norm coherification is equal to $\Delta\mathcal{C}_2(\Pi) = (d-1)/d$. Observe that the optimal coherification of the classical identity matrix $T = \mathbb{1}$ (corresponding to a completely decohering channel $\mathcal{D}$) is indeed the unitary identity quantum channel $\mathcal{I}$, represented by a maximally entangled state $J_\mathcal{I} = \frac{1}{d}|\Omega\rangle\langle\Omega|$.

Let us now move to the other extreme: the uniform van der Waerden matrix W of size d with entries $W_{ij} = 1/d$. A quantum channel Φ corresponding to a diagonal Jamiołkowski state with $\mathrm{diag}(J_\Phi) = \frac{1}{d}|W\rangle\rangle$ is the completely depolarizing channel, which sends any state into the maximally mixed state, $\Phi(\rho) = \mathbb{1}/d$. The vector $\frac{1}{d}|W\rangle\rangle$ has d^2 equal entries, so that its entropy is equal to $2\log d$. However, as for any dimension d there

exists a unitary Fourier matrix F with all entries of the same modulus, the uniform bistochastic matrix is unistochastic, $W = F \circ \bar{F}$. Thus, Φ can be completely coherified to a unitary transformation described by a pure state J_Φ^C of zero entropy. As a result, coherification of the uniform matrix (i.e., completely depolarizing channel) is maximal, with $\Delta C_e(W) = 2 \log d$ and $\Delta C_2(W) = (d^2 - 1)/d^2$.

Finally, we consider a class of quantum channels of an arbitrary dimension that can be completely coherified: a family of Schur product channels [26, 27], defined as

$$\Phi_X : \rho \mapsto \rho \circ X. \tag{24}$$

In the above X is an arbitrary correlation matrix and $\circ$, as before, denotes the entry-wise (Schur) product. The correlation matrix X has ones on the diagonal to assure trace preserving condition, and positivity of X guarantees complete positivity of the map Φ_X. The Choi–Jamiołkowski matrix of this channel is given by

$$J_{\Phi_X} = \frac{1}{d} \sum_{i,j} X_{ij} |i\rangle \langle j| \otimes |i\rangle \langle j|. \tag{25}$$

As $X_{ii} = 1$ for all i, the classical action of Φ_X is given by an identity matrix. Using proposition 1 we see that every Schur product channel can be completely coherified to a single common unitary channel, namely the identity.

3.2. Stochastic matrices and majorization bounds

We now proceed to the analysis of quantum channels whose classical action is given by a general stochastic matrix T that is not bistochastic (the outer shell of the set $\mathcal{T}_d$ presented in figure 4). First, we provide the majorization upper-bound on the spectrum, $\boldsymbol{\lambda}(J_\Phi)$, of all Jamiołkowski states with a given diagonal $\mathrm{diag}(J_\Phi) = \frac{1}{d}|T\rangle\rangle$, i.e., on the spectrum of the optimally coherified state J_Φ^C. This bound allows us to upper-bound any Schur-convex function of the spectrum $\boldsymbol{\lambda}(J_\Phi^C)$ (like the purity $\gamma(J_\Phi^C)$) and lower-bound any Schur-concave function (like entropy $S(J_\Phi^C)$), and thus to bound C_e and C_2 of the optimally coherified channel. This, in turn, is equivalent to bounding the entropic and 2-norm coherifications of a given classical channel. Next, we provide an explicit construction of a particular (non-optimal) coherified Jamiołkowski state $J_\Phi^{C_0}$, which allows us to lower bound coherence measures for the optimally coherified channel. We then illustrate the application of our results by finding optimal coherifications of qubit and qutrit channels, and interpreting their action. Finally, we make a short comment on the coherification of a particular qudit map.

3.2.1. Upper bound for the optimal spectrum

Let us first recall that a probability vector $\boldsymbol{p}$ is said to majorize $\boldsymbol{q}$, which we denote by $\boldsymbol{p} \succ \boldsymbol{q}$, if and only if

$$\sum_{i=1}^{k} p_i^\downarrow \geqslant \sum_{i=1}^{k} q_i^\downarrow, \tag{26}$$

for all $k \in \{1, \ldots, d\}$, where $\boldsymbol{p}^\downarrow$ denotes a probability vector with entries of $\boldsymbol{p}$ arranged in a non-increasing order. We now state the following theorem, which we prove in appendix A (recall that $\boldsymbol{\lambda}(X)$ denotes the eigenvalues of X arranged in a non-increasing order):

Theorem 2. *Given a positive semi-definite matrix J_Φ written in blocks, as in equation (12), we have*

$$\frac{1}{d} \sum_{i=1}^{d} \boldsymbol{\lambda}(D^i) \succ \boldsymbol{\lambda}(J_\Phi), \tag{27}$$

where $d(d - 1)$ zeros are appended to each of the vectors $\boldsymbol{\lambda}(D^i)$, so that their dimension agrees with that of $\boldsymbol{\lambda}(J_\Phi)$.

Next, we note that for Jamiołkowski states, due to equation (8b), we have $\sum_i D^i = \mathbf{1}$. This results in the maximal eigenvalue of each D^i to be upper-bounded by 1, as $D^i \leqslant \mathbf{1}$. Consider now a $d \times d$ stochastic matrix T that describes the diagonal of the Jamiołkowski state J_Φ. For every row of T we write the sum over its columns as $\sum_j T_{ij} = n_i + a_i$, with n_i being an integer and $a_i \in [0,1)$. We then define the following set of vectors:

$$\boldsymbol{s}^{(i)}(T) = [\underbrace{1, 1, \ldots, 1}_{n_i \text{ times}}, a_i, 0, \ldots, 0]. \tag{28}$$

Using theorem 2 and the fact that eigenvalues of D^i are upper-bounded by 1 we obtain the following majorization bound:

$$\boldsymbol{\mu}^\succ(T) := \frac{1}{d} \sum_{i=1}^{d} \boldsymbol{s}^{(i)}(T) \succ \boldsymbol{\lambda}(J_\Phi). \tag{29}$$

Since the above majorization bound holds for all J_Φ with a fixed diagonal, in particular it also bounds the optimally coherified channel, $\boldsymbol{\mu}^\succ(T) \succ \boldsymbol{\lambda}(J_\Phi^C)$. This can then be translated into upper-bounds on entropic coherence of Φ^{C_e} and 2-norm coherence of Φ^{C_2} [so also, due to equations (21a) and (21b), on entropic and

2-norm coherifications of a classical channel $\Phi \sim T$]:

$$\mathcal{C}_e(\Phi^{\mathcal{C}_e}) \leqslant S\left(\frac{1}{d}|T\rangle\rangle\right) - S(\boldsymbol{\mu}^{\succ}(T)), \tag{30a}$$

$$\mathcal{C}_2(\Phi^{\mathcal{C}_2}) \leqslant \boldsymbol{\mu}^{\succ}(T) \cdot \boldsymbol{\mu}^{\succ}(T) - \frac{1}{d^2}\mathrm{Tr}(TT^{\dagger}). \tag{30b}$$

To illustrate the application of the bound, let us consider the following transition matrix:

$$T = \begin{bmatrix} 0.7 & 0.2 & 0.6 \\ 0.1 & 0.6 & 0.4 \\ 0.2 & 0.2 & 0.0 \end{bmatrix}, \tag{31}$$

for which the vectors from equation (28) read

$$\boldsymbol{s}^{(1)}(T) = [1, 0.5, 0], \quad \boldsymbol{s}^{(2)}(T) = [1, 0.1, 0], \quad \boldsymbol{s}^{(3)}(T) = [0.4, 0, 0].$$

The bound, equation (29), tells us then that $[0.8, 0.2, 0, \ldots, 0] \succ \boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}})$.

Let us also notice that the bound becomes trivial for bistochastic matrices, as in this case $\boldsymbol{s}^{(i)} = [1, 0, \ldots, 0]$ for all i. This, however, was to be expected, as otherwise one could differentiate between unistochastic matrices and bistochastic matrices that are not unistochastic, a problem that is known to be hard and was solved only for $d \leqslant 3$ [25]. We will come back to this problem in section 3.3.

3.2.2. Lower bound for the optimal spectrum

We will now present a particular (non-optimal) coherification procedure $\mathcal{C}_0$ that can be applied to all quantum channels, irrespectively of their classical action T. The coherence of channel $\Phi^{\mathcal{C}_0}$ coherified in such a way can thus be used as a lower-bound on the optimally coherified channel $\Phi^{\mathcal{C}}$, i.e., $\mathcal{C}_e(\Phi^{\mathcal{C}_e}) \geqslant \mathcal{C}_e(\Phi^{\mathcal{C}_0})$ and $\mathcal{C}_2(\Phi^{\mathcal{C}_2}) \geqslant \mathcal{C}_2(\Phi^{\mathcal{C}_0})$.

We start by reminding that the constraint not allowing one to completely coherify a channel is the TP condition, equation (8b), which means that $\sum_i D^i = \mathbb{1}$. One can then choose all D^i to be diagonal and try to coherify the channel only by modifying C^{ij} matrices. Note that the eigenvalues of D^i are then given by the entries of T, $\boldsymbol{\lambda}(D^i) = \boldsymbol{r}^{(i)}$, where the vectors $\boldsymbol{r}^{(i)}$ are defined by the rows of T arranged in a non-increasing order,

$$\boldsymbol{r}^{(i)} = [T_{i1}, \ldots, T_{id}]^{\downarrow}. \tag{32}$$

From theorem 2 we thus have that

$$\frac{1}{d}\sum_{i=1}^{d} \boldsymbol{r}^{(i)} \succ \boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}_0}). \tag{33}$$

The above majorization bound can be saturated by $J_{\Phi}^{\mathcal{C}_0}$ with the following choice of non-zero elements of C^{ij}. For each block D^i find the maximal diagonal element, $r_1^{(i)}$, and set the off-diagonal elements between them (elements of C^{ij}) to the maximal value allowed by the CP condition, i.e., $[r_1^{(i)}r_1^{(j)}]^{\frac{1}{2}}$. Then, repeat the procedure for the nth largest eigenvalues of D^i, $r_n^{(i)}$, with $n \in \{2, \ldots, d\}$. The structure of the resulting Jamiołkowski state $J_{\Phi}^{\mathcal{C}_0}$ is illustrated in figure 5 for $d = 3$. The spectrum of $J_{\Phi}^{\mathcal{C}_0}$ is given by
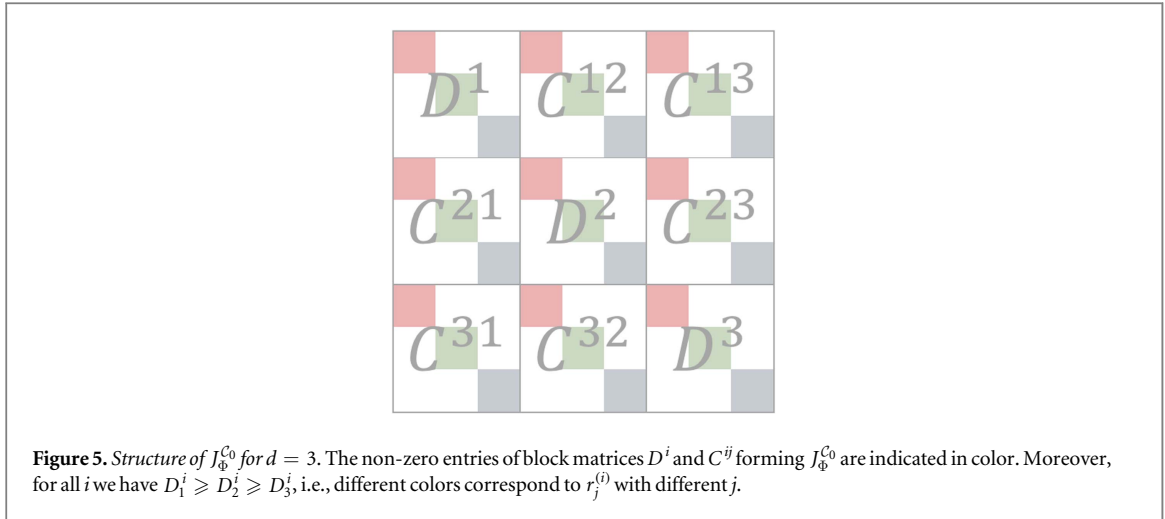
$$\boldsymbol{\mu}^{\backprime}(T) := \frac{1}{d}\sum_{i=1}^{d} \boldsymbol{r}^{(i)} = \boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}_0}), \tag{34}$$

which is also the optimal spectrum for the Jamiołkowski state with a fixed diagonal, when we additionally assume no coherence in its diagonal blocks D^i.

Let us now analyze the action of a channel coherified according to $\mathcal{C}_0$. We first note that a classical channel $\Phi = \Phi^{\mathcal{D}}$ has the following Kraus decomposition:

$$\Phi(\cdot) = \sum_{i,j}^{d} K_{ij}(\cdot)K_{ij}^{\dagger}, \quad K_{ij} = \sqrt{T_{ij}}|i\rangle\langle j|, \tag{35}$$

so that the minimal number of Kraus operators is equal to the number of non-zero entries of a stochastic matrix T (in general d^2). On the other hand, we know that, by construction, $J_{\Phi}^{\mathcal{C}_0}$ is equal to the sum over at most d projectors, so that the number of the corresponding Kraus operators will be smaller or equal to d. More precisely, one can obtain ith Kraus operator directly from the T matrix: in every row of T leave only the ith largest entry, replace it with its square root, and set all other entries to zero. For example, given the transition matrix from equation (31), we get

Figure 5. *Structure of $J_\Phi^{\mathcal{C}_0}$ for $d = 3$. The non-zero entries of block matrices D^i and C^{ij} forming $J_\Phi^{\mathcal{C}_0}$ are indicated in color. Moreover, for all i we have $D_1^i \geqslant D_2^i \geqslant D_3^i$, i.e., different colors correspond to $r_j^{(i)}$ with different j.*

$$K_1 = \begin{bmatrix} \sqrt{0.7} & 0 & 0 \\ 0 & \sqrt{0.6} & 0 \\ \sqrt{0.2} & 0 & 0 \end{bmatrix}, \quad K_2 = \begin{bmatrix} 0 & 0 & \sqrt{0.6} \\ 0 & 0 & \sqrt{0.4} \\ 0 & \sqrt{0.2} & 0 \end{bmatrix}, \quad K_3 = \begin{bmatrix} 0 & \sqrt{0.2} & 0 \\ \sqrt{0.1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We thus see that it is always possible to coherify a channel Φ, so that the number of Kraus operators (the rank of the Jamiołkowski state $J_\Phi^{\mathcal{C}_0}$) realizing a given classical transformation T (with $\mathrm{diag}(J_\Phi^{\mathcal{C}_0}) = \frac{1}{d}|T\rangle\rangle$) decreases from d^2 to d. Physically, we can interpret $\mathcal{C}_0$ as replacing d^2 classical processes (transitions from state i to j) into a classical mixture over d quantum processes, where each quantum process describes a coherent superposition of d classical transitions, each to a different final state. We also note that there exist stochastic matrices T for which one cannot reduce the number of Kraus operators below d. These are given by transitions that move all populations to a fixed i_0th state, i.e., T with all entries in the i_0th row equal to one and all other entries being zero. The D^i matrices of the corresponding Jamiołkowski state are all vanishing, except for $D^{i_0} = \mathbb{1}$, which cannot be coherified due to the condition $\sum_i D^i = \mathbb{1}$.

Finally, since it is always possible to coherify a channel so that the spectrum of its Jamiołkowski state is given by $\boldsymbol{\mu}^{\smallfrown}(T)$, one gets the following lower-bounds

$$\mathcal{C}_e(\Phi^{\mathcal{C}_e}) \geqslant S\left(\frac{1}{d}|T\rangle\rangle\right) - S(\boldsymbol{\mu}^{\smallfrown}(T)), \tag{36a}$$

$$\mathcal{C}_2(\Phi^{\mathcal{C}_2}) \geqslant \boldsymbol{\mu}^{\smallfrown}(T) \cdot \boldsymbol{\mu}^{\smallfrown}(T) - \frac{1}{d^2}\,\mathrm{Tr}(TT^\dagger). \tag{36b}$$

For comparison with the upper-bounds presented in the previous section, equations (30a) and (30b), we note that for the exemplary transition matrix used there, equation (31), we have

$$\boldsymbol{\mu}^{\smallfrown}(T) = [0.5, 0.4, 0.1, 0, \ldots, 0].$$

3.2.3. Qubits

Having described the general bounds on possible coherifications of quantum channels acting on arbitrary d-dimensional spaces, we now want to focus on a particular case of $d = 2$. The classical action of a general qubit channel is given by a 2×2 transition matrix T described by two real parameters,
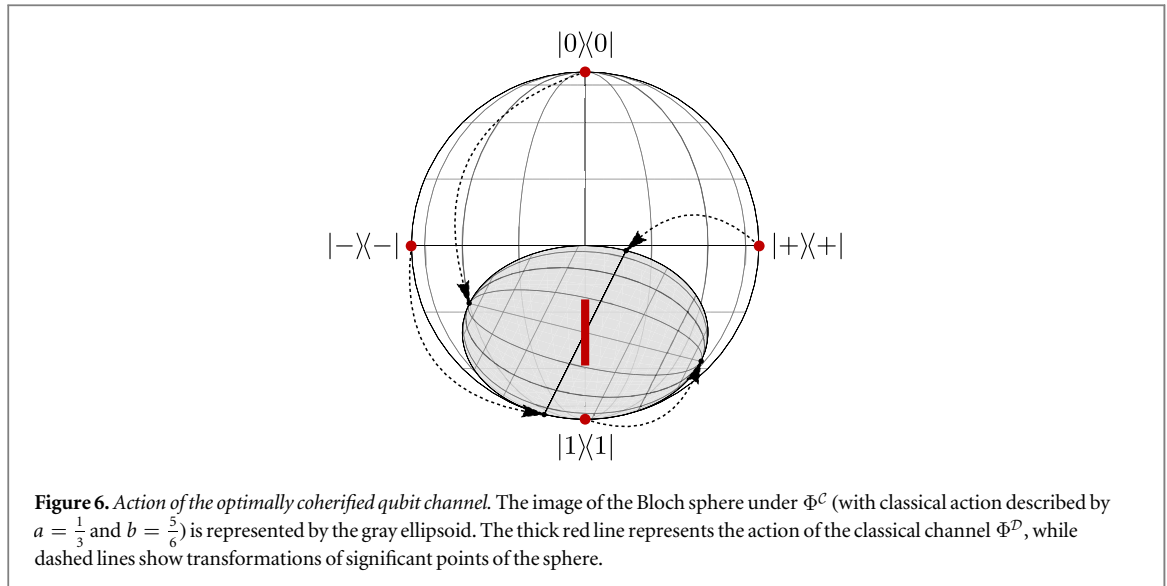
$$T = \begin{bmatrix} a & 1-b \\ 1-a & b \end{bmatrix} =: \begin{bmatrix} a & \tilde{b} \\ \tilde{a} & b \end{bmatrix}, \tag{37}$$

with $\tilde{x} := 1 - x$. We will only focus on the case when $a \leqslant b$, as the results for the case $a > b$ are analogous. Namely, one only needs to exchange a with b in all expressions, and transform all matrices X by replacing X_{kl} with $X_{\tilde{k}\tilde{l}}$. The details of the derivations can be found in appendix B.

For $a \leqslant b$, equation (29) tells us that the spectrum of the Jamiołkowski state J_Φ, corresponding to any channel with classical action specified by T, is bounded in the following way,

$$\boldsymbol{\mu}^{\smallsmile}(T) = \frac{1}{2}[1 + a + \tilde{b}, b - a, 0, \ldots, 0] \succ \boldsymbol{\lambda}(J_\Phi). \tag{38}$$

This bound can be in fact saturated, i.e., there exists $J_\Phi^{\mathcal{C}}$ such that $\boldsymbol{\lambda}(J_\Phi^{\mathcal{C}}) = \boldsymbol{\mu}^{\smallsmile}(T)$. Note that, since the majorization bound is saturated, both coherification procedures (maximizing entropic and 2-norm coherence) coincide, and are simply denoted by $\mathcal{C}$. To express the Kraus operators of the corresponding optimally coherified

Figure 6. *Action of the optimally coherified qubit channel.* The image of the Bloch sphere under $\Phi^{\mathcal{C}}$ (with classical action described by $a = \frac{1}{3}$ and $b = \frac{5}{6}$) is represented by the gray ellipsoid. The thick red line represents the action of the classical channel $\Phi^{\mathcal{D}}$, while dashed lines show transformations of significant points of the sphere.

channel $\Phi^{\mathcal{C}}$, let us first introduce a unitary

$$U = \frac{1}{\sqrt{a + \tilde{b}}} \begin{bmatrix} \sqrt{a} & -\sqrt{\tilde{b}} \\ \sqrt{\tilde{b}} & \sqrt{a} \end{bmatrix}, \tag{39}$$

and a decaying channel $\Psi(\cdot) = L_1(\cdot)L_1^{\dagger} + L_2(\cdot)L_2^{\dagger}$ with

$$L_1 = \begin{bmatrix} \sqrt{a + \tilde{b}} & 0 \\ 0 & 1 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 & 0 \\ \sqrt{b - a} & 0 \end{bmatrix}. \tag{40}$$

Then, $\Phi^{\mathcal{C}}(\cdot) = K_1(\cdot)K_1^{\dagger} + K_2(\cdot)K_2^{\dagger}$ with

$$K_1 = L_1 U^{\dagger} = \begin{bmatrix} \sqrt{a} & \sqrt{\tilde{b}} \\ -\sqrt{\frac{\tilde{b}}{a + \tilde{b}}} & \sqrt{\frac{a}{a + \tilde{b}}} \end{bmatrix}, \quad K_2 = L_2 U^{\dagger} = \begin{bmatrix} 0 & 0 \\ \sqrt{\tilde{a} - \frac{\tilde{b}}{a + \tilde{b}}} & \sqrt{b - \frac{a}{a + \tilde{b}}} \end{bmatrix}. \tag{41}$$

It is straightforward to verify that the classical action of the resulting channel is given by T (using equation (22)), as well as that the spectrum of the Jamiołkowski state is optimal (by checking that $\mathrm{Tr}(K_i K_i^{\dagger}) = \mu_i^{>}(T)$). Let us also explicitly emphasize that the optimally coherified channel is given by the composition of a unitary process and a decaying channel, $\Phi^{\mathcal{C}}(\cdot) = \Psi[U^{\dagger}(\cdot)U]$. As a consequence there exists a pure state $U|1\rangle$ that is mapped by $\Phi^{\mathcal{C}}$ to a pure state $|1\rangle$, and therefore the minimum output entropy of $\Phi^{\mathcal{C}}$ is zero. We illustrate the action of $\Phi^{\mathcal{C}}$ on a Bloch sphere for a particular choice of T in figure 6.

Finally, let us apply the notion of coherification to contribute to the studies on geometry of the set $\mathcal{S}_2$ of one-qubit stochastic maps initiated in [28].

Proposition 3. *Coherification of any classical one-qubit stochastic map, specified by a matrix T from equation (37), yields a channel which is extremal in the set $\mathcal{S}_2$.*

Proof. In the bistochastic case $a = b$, we obtain a unitary channel, which is extremal. In other cases, without loss of generality, we may assume that $c = b - a > 0$. The products of the Kraus operators from equation (40), corresponding to a decaying channel, read

$$L_1^{\dagger} L_1 = \begin{bmatrix} 1 - c & 0 \\ 0 & 1 \end{bmatrix}, \quad L_1^{\dagger} L_2 = \begin{bmatrix} 0 & 0 \\ \sqrt{c} & 0 \end{bmatrix}, \quad L_2^{\dagger} L_1 = \begin{bmatrix} 0 & \sqrt{c} \\ 0 & 0 \end{bmatrix}, \quad L_2^{\dagger} L_2 = \begin{bmatrix} c & 0 \\ 0 & 0 \end{bmatrix}. \tag{42}$$

Now, by direct inspection, we see that the above matrices form a linearly independent set. Thus, invoking the theorem of Choi [12], the channel described by two Kraus operators L_1 and L_2 is extremal. To see that $\Phi^{\mathcal{C}}$ is extremal, we note that the additional unitary matrix applied to Kraus operators will not introduce a linear dependence of the set defined in equation (42).

3.2.4. Qutrits

We will now illustrate how our results can be applied beyond the simplest qubit scenario, by using them to find optimally coherified qutrit channels $\Phi^{\mathcal{C}}$. Again, the coherification procedure $\mathcal{C}$ will optimize both considered coherence measures simultaneously. The classical action of a general qutrit channel is given by a 3×3 transition matrix T described by six real parameters. Here, we will consider three families of such matrices, each parametrized by three real numbers:

$$T \in \left\{ \begin{bmatrix} 0 & a & b \\ c & 0 & \tilde{b} \\ \tilde{c} & \tilde{a} & 0 \end{bmatrix}, \begin{bmatrix} a & b & 0 \\ 0 & 0 & c \\ \tilde{a} & \tilde{b} & \tilde{c} \end{bmatrix}, \begin{bmatrix} a & b & c \\ \tilde{a} & \tilde{b} & \tilde{c} \\ 0 & 0 & 0 \end{bmatrix} \right\}, \tag{43}$$

with $\tilde{x} := 1 - x$. We will refer to the above as cyclic matrices, single-row matrices and double-row matrices, respectively. For each family of T we will provide the optimal spectrum of the Jamiołkowski state (which yields tight bounds on $\mathcal{C}_e$ and $\mathcal{C}_2$ via equations (30a), (30b)), as well as the Kraus decomposition of the optimally coherified channel $\Phi^{\mathcal{C}}$. The details of the derivations can be found in appendix C.

For cyclic matrices $\Phi^{\mathcal{C}}$ is given by $\Phi^{\mathcal{C}_0}$, i.e., the optimal coherification procedure is given by the one we defined in section 3.2.2. Introducing

$$\mu = \max(a, b) + \max(c, \tilde{b}) + \max(\tilde{c}, \tilde{a}). \tag{44}$$

we thus get

$$\boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}}) = \left[\frac{\mu}{3}, 1 - \frac{\mu}{3}, 0, \dots, 0 \right]. \tag{45}$$

The Kraus operators can be obtained by using the procedure described in section 3.2.2, e.g., for $a \geqslant b$, $\tilde{b} \geqslant c$ and $\tilde{c} \geqslant \tilde{a}$ we get

$$K_1 = \begin{bmatrix} 0 & \sqrt{a} & 0 \\ 0 & 0 & \sqrt{\tilde{b}} \\ \sqrt{\tilde{c}} & 0 & 0 \end{bmatrix}, \ K_2 = \begin{bmatrix} 0 & 0 & \sqrt{b} \\ \sqrt{c} & 0 & 0 \\ 0 & \sqrt{\tilde{a}} & 0 \end{bmatrix}. \tag{46}$$

For single-row transition matrices there are three separate cases depending on parameters a, b and c. If $a + b \leqslant 1$ then the optimal spectrum is given by

$$\boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}}) = \frac{1}{3}[1 + a + b + c, \ \max(\tilde{a} - b, \tilde{c}), \ \min(\tilde{a} - b, \tilde{c}), 0, \dots, 0], \tag{47}$$

and the Kraus decomposition of the corresponding optimally coherified channel is given by

$$K_1 = \begin{bmatrix} \sqrt{a} & \sqrt{b} & 0 \\ 0 & 0 & \sqrt{c} \\ \sqrt{\frac{b}{a+b}} & -\sqrt{\frac{a}{a+b}} & 0 \end{bmatrix}, \ K_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{\tilde{c}} \end{bmatrix}, K_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sqrt{\tilde{a} - \frac{b}{a+b}} & \sqrt{\tilde{b} - \frac{a}{a+b}} & 0 \end{bmatrix}. \tag{48}$$

If $a + b \geqslant 1$ the number of non-zero elements of the optimal spectrum (so also of the Kraus operators) reduces from three to two. The optimal spectrum is again given by equation (45), but this time with

$$\mu = 1 + \max(\tilde{a} + \tilde{b}, \tilde{c}) + c. \tag{49}$$

If $\tilde{c} \geqslant \tilde{a} + \tilde{b}$ we thus get $\boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}}) = [2/3, 1/3, \dots, 0]$, i.e., the optimal spectrum is constant for all parameters satisfying $a + b \geqslant 1 + c$, and the Kraus operators of the optimal map are given by:
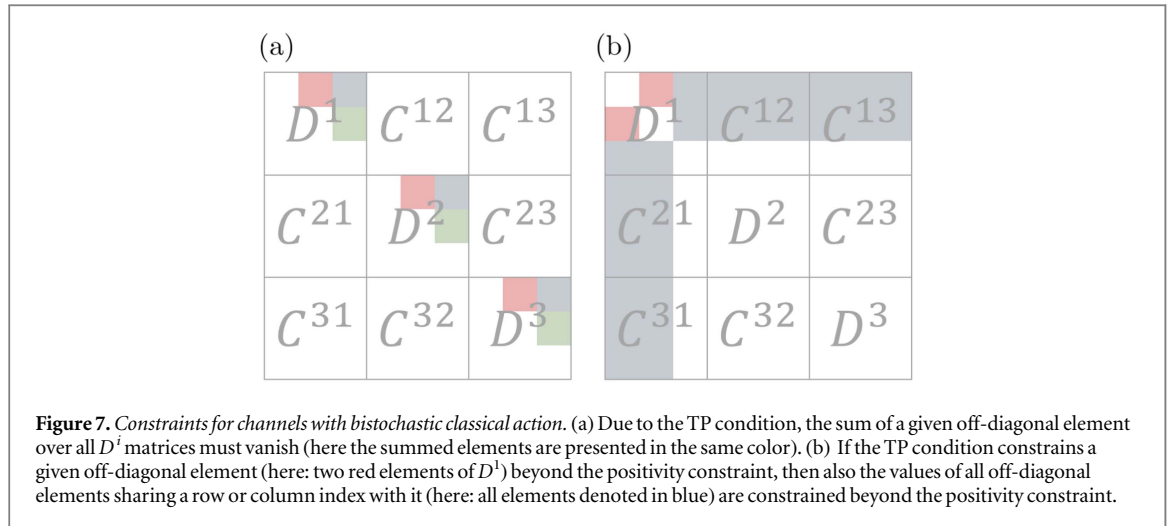
$$K_1 = \begin{bmatrix} \sqrt{\frac{\tilde{b}}{\tilde{a}+\tilde{b}}} & \sqrt{\frac{\tilde{a}}{\tilde{a}+\tilde{b}}} & 0 \\ 0 & 0 & \sqrt{c} \\ 0 & 0 & \sqrt{\tilde{c}} \end{bmatrix}, K_2 = \begin{bmatrix} \sqrt{a - \frac{\tilde{b}}{\tilde{a}+\tilde{b}}} & -\sqrt{b - \frac{\tilde{a}}{\tilde{a}+\tilde{b}}} & 0 \\ 0 & 0 & 0 \\ \sqrt{\tilde{a}} & -\sqrt{\tilde{b}} & 0 \end{bmatrix}. \tag{50}$$

On the other hand, if $\tilde{c} \leqslant \tilde{a} + \tilde{b}$ then there is a slight change in the Kraus decomposition of the optimal map. Namely, the last rows of K_1 and K_2 in equation (50) are swapped.

For double-row matrices there are again three separate cases depending on the value of $s := a + b + c$. These are specified by $s \in [0, 1]$, $s \in (1, 2)$ and $s \in [2, 3]$, and the optimal spectra are then given by

$$\boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}}) = \frac{1}{3}[1 + s, 1, 1 - s, 0, \dots, 0], \tag{51a}$$

$$\boldsymbol{\lambda}(J_{\Phi}^{\mathcal{C}}) = \frac{1}{3}[2, 1, 0, \dots, 0], \tag{51b}$$

Figure 7. *Constraints for channels with bistochastic classical action.* (a) Due to the TP condition, the sum of a given off-diagonal element over all D^i matrices must vanish (here the summed elements are presented in the same color). (b) If the TP condition constrains a given off-diagonal element (here: two red elements of D^1) beyond the positivity constraint, then also the values of all off-diagonal elements sharing a row or column index with it (here: all elements denoted in blue) are constrained beyond the positivity constraint.

$$\boldsymbol{\lambda}(J_\Phi^{\mathcal{C}}) = \frac{1}{3}[4 - s, 1, s - 2, 0, \dots, 0], \tag{51c}$$

respectively. Due to the lack of concise expressions, we provide Kraus decompositions of the resulting optimally coherified channels in appendix C.

3.2.5. Qudits
Finally, we want to make a short comment about a special family of channels in the general d-dimensional case. Consider a *completely contracting channel* Ψ_σ, which sends any initial state into a single point, $\Psi_\sigma(\rho) = \sigma$. The corresponding Jamiołkowski state has a product structure and reads $J_{\Psi_\sigma} = \sigma \otimes \mathbf{1}/d$ [29]. The output state can be coherified to a pure state $\sigma^{\mathcal{C}} = |\psi\rangle\langle\psi|$ by the standard procedure given in equation (5). Hence the contracting channel Ψ_σ, can be coherified to a channel contracting into a pure state with $J_{\Psi_{|\psi\rangle}} = |\psi\rangle\langle\psi| \otimes \mathbf{1}/d$ and zero output entropy. Note that this coherification procedure increases the entropic coherence of a channel by $S(\sigma)$. Notice also, that for a mixed state σ such a procedure is not optimal, as can be immediately seen by recalling the result presented in section 3.1, where we showed that $\Psi_{\mathbf{1}/d}$ can be completely coherified.

3.3. Bistochastic matrices and polygon constraints
We now proceed to the analysis of quantum channels whose classical action is described by bistochastic matrices that are not unistochastic (the middle shell of the graph presented in figure 4). On the one hand, due to proposition 1, we know that these cannot be completely coherified. On the other, our majorization result derived in section 3.2.1 yields a trivial bound for bistochastic matrices. Moreover, a non-trivial constraint for all bistochastic matrices could serve as a witness of unistochasticity, and thus it is unlikely that such a concise bound can be found [25]. Therefore, here we will present an approach that allows one to obtain limitations on possible coherifications of quantum channels with classical action described by a particular subset of $\mathcal{B}_d$.

We start by noting that due to the TP condition, equation (8b), for every $k \neq l$ we have $\sum_i D_{kl}^i = 0$ (see figure 7(a)). This, via the polygon inequality, implies that

$$|D_{kl}^i| \leqslant \sum_{j \neq i} |D_{kl}^j|. \tag{52}$$

Recalling that matrices D^i are all positive, we have

$$|D_{kl}^i| \leqslant \sqrt{D_{kk}^i D_{ll}^i} = \sqrt{T_{ik} T_{il}}. \tag{53}$$

Combining the above two equations we arrive at

$$|D_{kl}^i| \leqslant \sum_{j \neq i} \sqrt{T_{jk} T_{jl}}. \tag{54}$$

We thus see that the maximum value of $|D_{kl}^i|$ allowed by CP condition is $\sqrt{T_{ik} T_{il}}$, whilst the TP condition restricts it via equation (54). Therefore, if for some i, k, l we have

$$\sqrt{T_{ik} T_{il}} > \sum_{j \neq i} \sqrt{T_{jk} T_{jl}}, \tag{55}$$

then $|D_{kl}^i|$ is constrained beyond the positivity condition and we know that the resulting Jamiołkowski state cannot be pure, so that the corresponding channel cannot be completely coherified. More precisely, for every i,

k, l such that $T_{ik} T_{il} \neq 0$, we introduce

$$\sqrt{\alpha_{kl}^i} := \min\left(\frac{\sum_{j \neq i} \sqrt{T_{jk} T_{jl}}}{\sqrt{T_{ik} T_{il}}}, 1\right), \tag{56}$$

which describes the maximum fraction of the coherence (between states k and l of a matrix D^i) that could be achieved if there was no TP constraint.

Now, whenever $\alpha_{kl}^i < 1$ (i.e., $|D_{kl}^i|$ must be smaller than necessary for complete coherification), other off-diagonal elements of the Jamiołkowski state J_Φ also become constrained beyond the positivity condition. Before we prove this and explain how it restricts the coherification of quantum channels, let us first comment on the $\alpha_{kl}^i < 1$ condition. First of all, there exist $T \in \mathcal{B}_d$ that are neither unistochastic, nor they satisfy this condition for any i, k, l. Thus, the presented bounds will, in general, work only for a subset of quantum channels with bistochastic classical action. However, as for $d = 3$ a bistochastic matrix is either unistochastic or $\alpha_{kl}^i < 1$ for some i, k, l [25], we will obtain non-trivial bounds for all qutrit channels. Further improvements would require finding a clearer separation between the sets $\mathcal{U}_d$ and $\mathcal{B}_d$.

We start by showing how $\alpha_{kl}^i < 1$ can be used to constrain the purity of the optimally coherified channel. Note that Sylvester's criterion states that $J_\Phi \geqslant 0$ implies that all 3×3 submatrices of J_Φ must have positive determinant. In particular, it means that for a part of a matrix $d \cdot J_\Phi$ containing T_{ik}, T_{il} and any other T_{jm} we have

$$\det \begin{bmatrix} T_{ik} & D_{kl}^i & a \\ D_{kl}^{i*} & T_{il} & b \\ a^* & b^* & T_{jm} \end{bmatrix} \geqslant 0. \tag{57}$$

Since we know that the maximum value of $|D_{kl}^i|^2$ is upper-bounded by $\alpha_{kl}^i T_{ik} T_{il}$, the above equation constrains all off-diagonal elements a, b of J_Φ sharing a row or column with T_{ik} or T_{il} (see figure 7(b)). This results in the following bound on the purity of the optimally coherified Jamiołkowski state $J_\Phi^{\mathcal{C}_2}$ (see appendix D for details):

$$\gamma(J_\Phi^{\mathcal{C}_2}) \leqslant 1 - \Delta_1 - \Delta_2, \tag{58}$$

with

$$\Delta_1 = \frac{2}{d^2} T_{ik} T_{il} (1 - \alpha_{kl}^i), \qquad \Delta_2 = \frac{d - T_{ik} - T_{il}}{d^2} (T_{ik} + T_{il} - \beta_{kl}^i), \tag{59}$$

and

$$\beta_{kl}^i := \sqrt{(T_{ik} - T_{il})^2 + 4\alpha_{kl}^i T_{ik} T_{il}}. \tag{60}$$

The purity deficits, Δ_1 and Δ_2, add up for every i, k, l for which equation (55) holds (however, care needs to be taken not to count twice the same terms). We illustrate this bound on the purity of the optimally coherified channel in figure 8 for an exemplary case of a family of qutrit maps.

Alternatively, one can use the fact that off-diagonal terms of D^i are constrained beyond the positivity condition to bound $\boldsymbol{\lambda}(D^i)$, and then use theorem 2 to obtain a non-trivial majorization bound on the eigenvalues of the Jamiołkowski state. In appendix D we show for example that

$$[1 - \mu^i, \mu^i, 0, \ldots, 0] \succ \boldsymbol{\lambda}(D^i), \tag{61}$$

with

$$\mu^i = \frac{1}{2}(T_{ik_i} + T_{il_i} - \beta_{k_i l_i}^i), \tag{62}$$

where k_i and l_i are indices for which $\alpha_{k_i l_i}^i$ is minimized (so that we obtain non-trivial majorization bounds on the spectra of D^i, whenever $\alpha_{kl}^i < 1$ for some k and l). This, in turn, means that

$$\left[1 - \frac{1}{d}\sum_i \mu^i, \frac{1}{d}\sum_i \mu^i, 0, \ldots, 0\right] \succ \boldsymbol{\lambda}(J_\Phi^{\mathcal{C}}), \tag{63}$$

which can be used to obtain bounds on $\mathcal{C}_e$ and $\mathcal{C}_2$ via equations (30a), (30b). As an example consider a quantum channel with classical action given by

$$T = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{64}$$

Figure 8. *Purity bound.* The upper-bound on purity γ of the family of optimally coherified qutrit channels with the classical action given by a bistochastic transition matrix $T = \sum_{i=1}^{3} q_i \Pi^i$, with $\sum_i q_i = 1$ and Π being a cyclic permutation matrix. Any unistochastic matrix $T \in \mathcal{U}_3$ can be completely coherified, so that $\gamma = 1$.
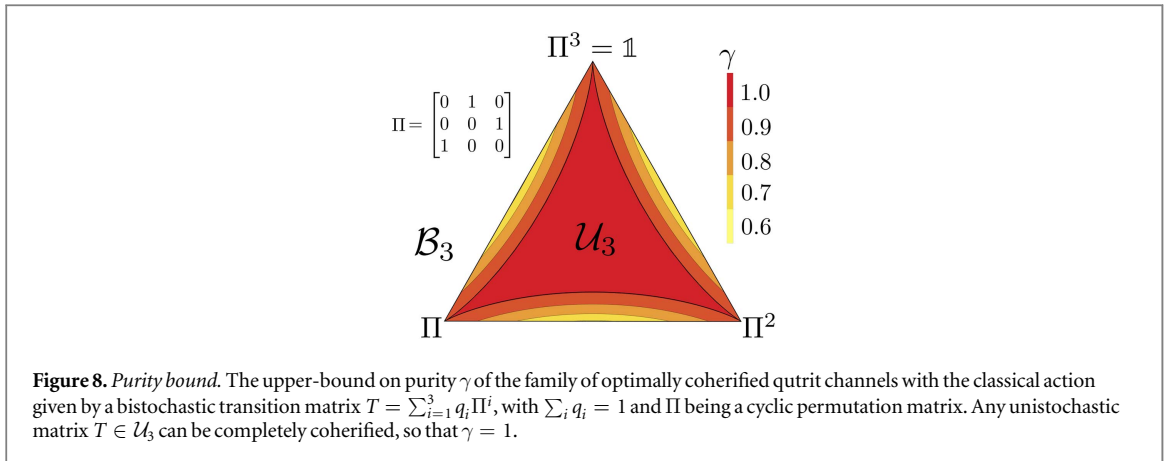
corresponding to the middle-point between Π and Π^2 in figure 8. One then gets

$$\alpha_{23}^1 = \alpha_{13}^2 = \alpha_{12}^3 = 0, \tag{65}$$

resulting in $\mu^i = 1/2$, so that the spectrum of the optimally coherified Jamiołkowski state is majorized by $[1/2, 1/2, 0, \ldots, 0]$.

Finally, let us note that, in particular cases, the tools introduced in section 3.2 can also be used to find limitations on coherifying channels with $T \in \mathcal{B}_d$. As an example consider the family of qutrit channels described by cyclic matrices (first entry of equation (43)) with $a = \tilde{b} = \tilde{c}$. The matrix T is then bistochastic and the spectrum of the optimally coherified Jamiołkowski state is given by $[a, 1 - a, 0, \ldots, 0]$ (for $a \geqslant 1/2$) or $[\tilde{a}, 1 - \tilde{a}, 0, \ldots, 0]$ (for $a < 1/2$). This shows that the majorization bound, equation (63), applied to the channel with T specified by equation (64) is tight.

4. Physical interpretation

We started this paper asking about the extent to which a given random transformation can be explained via the underlying deterministic and coherent process. Now, being equipped with formal bounds limiting possible coherifications of quantum channels, we will try to address this initial question. We will also provide interpretation of the purity of a channel by relating it to the notions of *unitarity* and average output purity. Finally, we will comment on the links and differences between our approach to the study of coherence of quantum channels, and the ones existing in the literature.

Let us start by recalling that the evolution of a pure quantum state $|\psi\rangle$ under the action of a channel Φ, described by

$$\Phi(|\psi\rangle \langle\psi|) = \sum_i K_i |\psi\rangle \langle\psi| K_i^\dagger, \tag{66}$$

can be interpreted as an incoherent (probabilistic) mixture of different pure state transformations,

$$|\psi\rangle \xrightarrow{\Phi} \frac{1}{\sqrt{q_i}} K_i |\psi\rangle \text{ with probability } q_i, \tag{67}$$

where

$$q_i = \text{Tr}(K_i |\psi\rangle \langle\psi| K_i^\dagger), \tag{68}$$

and we refer to the canonical Kraus form in which all Kraus operators K_i are mutually orthogonal, as they are obtained by reshaping eigenvectors of the Choi–Jamiołkowski matrix. Each independent path, described by K_i and being chosen with probability q_i, describes a coherent evolution, as it preserves the ability of a state to interfere (it maps a pure state to a pure state). Thus, the probability distribution over different paths, $\boldsymbol{q}$, can be seen as describing the incoherent randomness associated with Φ. Note, however, that the probability of evolving along a given path depends on the initial state of the system $|\psi\rangle$. In order to achieve a state-independent statement characterizing a given quantum channel Φ, one can then focus on the average probability of choosing a given path. Introducing the average over (Haar distributed) pure states,

$$\langle \cdot \rangle_\psi = \int d\psi (\cdot), \tag{69}$$

we see that the probability describing which path is chosen is on average[8]

$$\langle q_i \rangle_\psi = \mathrm{Tr}(K_i \langle |\psi\rangle\langle\psi| \rangle_\psi K_i^\dagger) = \mathrm{Tr}(K_i K_i^\dagger) = \lambda_i(J_\Phi), \tag{70}$$

where $\boldsymbol{\lambda}(J_\Phi)$ denotes, as usual, the eigenvalues of the Jamiołkowski state J_Φ corresponding to Φ.

We thus see that the incoherent randomness of the evolution coming from the random choice of different paths is (on average) described by the spectrum of J_Φ. The extent to which a quantum channel with a given classical action T can be coherified tells us how coherent the underlying evolution, leading to transitions described by T, can be. On one extreme, we have unistochastic transitions that can be completely coherified and, therefore, explained by a single deterministic path (unitary dynamics). On the other hand, the majorization upper-bounds on $\boldsymbol{\lambda}(J_\Phi)$ that we derived in section 3, yield lower-bounds on the randomness of path distribution of the underlying process necessary to induce classical transformation T. Moreover, our majorization lower-bound provides a particular coherent explanation of every classical process T (decreasing the randomness of path distribution) and, in particular, shows that all transitions T can be explained with at most d paths.

Let us now focus on a particular measure describing the randomness of the path distribution, namely on the purity of a channel $\gamma(\Phi)$. One could be tempted to think that the bigger $\gamma(\Phi)$ is, the purer the average output purity,

$$\langle \gamma_\Phi \rangle := \langle \gamma(\Phi(|\psi\rangle\langle\psi|)) \rangle_\psi, \tag{71}$$

will be. Although the two notions are related, as we will shortly see, they are not in direct correspondence. As an illustrative example consider two quantum channels,

$$\Phi_1(\cdot) = U(\cdot)U^\dagger \;\; \text{with} \;\; U = \frac{1}{\sqrt{d}} \sum_{k,l} \mathrm{e}^{\frac{2\pi i k l}{d}} |k\rangle\langle l|, \tag{72a}$$

$$\Phi_2(\cdot) = |\psi_+\rangle\langle\psi_+| \;\; \text{with} \;\; |\psi_+\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle, \tag{72b}$$

with classical action given by the van der Waerden matrix $T = W$ with flat entries $W_{ij} = 1/d$, which maps every probability distribution to a uniform one. Both channels have the same average output purity equal to one; but for a reversible Φ_1 we have $\gamma(\Phi_1) = 1$, while for an irreversible Φ_2 we get $\gamma(\Phi_1) = 1/d$. This suggests that the purity of a channel is somehow related to reversibility of the process, which leads us to the concept of unitarity. It was originally introduced in [14] to measure the departure of a channel from the unitary dynamics, and for trace-preserving channels is defined by:

$$u(\Phi) := \frac{d}{d-1} \Big[\langle\gamma_\Phi\rangle - \gamma\Big(\Phi\Big(\frac{\mathbf{1}}{d}\Big)\Big) \Big]. \tag{73}$$

We note in the passing that one can also relate unitarity to the variance of the random variable $X = \Phi(|\psi\rangle\langle\psi|)$:

$$u(\Phi) = \mathrm{Tr}(\langle X^2 \rangle_\psi - \langle X \rangle_\psi^2) = \mathrm{Tr}(\mathrm{Var}_\psi(X)). \tag{74}$$

For our exemplary channels we see that $u(\Phi_1) = 1$ and $u(\Phi_2) = 0$, in accordance with the purity of the channel and capturing the fact that a completely irreversible process is as far as possible from a unitary transformation.

We will now formally relate $\gamma(\Phi)$, $\langle\gamma_\Phi\rangle$ and $u(\Phi)$. The authors of [14] showed that

$$u(\Phi) = \frac{d}{d^2-1} \Big[d\gamma(\Phi) - \gamma\Big(\Phi\Big(\frac{\mathbf{1}}{d}\Big)\Big) \Big], \tag{75}$$

which, using the definition of unitarity, directly leads to the general expression for the average output purity derived by Cappellini [30],

$$\langle\gamma_\Phi\rangle = \frac{d}{d+1} \Big[\gamma(\Phi) + \gamma\Big(\Phi\Big(\frac{\mathbf{1}}{d}\Big)\Big) \Big]. \tag{76}$$

We thus see that both average output purity and unitarity are proportional to the purity of a channel corrected by a term describing the purity of the transformed maximally mixed state. Moreover, for large d, $u(\Phi)$ actually approaches $\gamma(\Phi)$. Now, by noting that the minimal value of purity for a d-dimensional system is $1/d$, we obtain the following inequalities[9]:

$$u(\Phi) \leqslant \frac{d^2}{d^2-1} \Big[\gamma(\Phi) - \frac{1}{d^2} \Big], \tag{77a}$$

---

[8] We note that the average here could actually be taken over all states, pure and mixed. However, in order to be consistent with the averaging process used in the definition of unitarity, we restrict the average to pure states only.

[9] Note that by restricting to unital channels these inequalities actually become equalities.

$$\langle \gamma_\Phi \rangle \geqslant \frac{d}{d+1}\left[\gamma(\Phi) + \frac{1}{d}\right]. \tag{77b}$$

Using our majorization and purity upper-bounds we can thus upper-bound the optimal unitarity of a channel with a given classical action T. On the other hand, using the majorization lower-bound, we can lower-bound the average output purity of such an optimal channel. The above bounds can actually be tightened by noting that

$$\gamma\left(\Phi\left(\frac{\mathbf{1}}{d}\right)\right) \geqslant \gamma\left(D\left(\Phi\left(\frac{\mathbf{1}}{d}\right)\right)\right) = \sum_i \left(\frac{1}{d}\sum_j T_{ij}\right)^2. \tag{78}$$

Let us also mention that the optimally coherified qubit channel $\Phi^{\mathcal{C}}$, with Kraus operators specified by equation (41), not only maximizes purity, but also minimizes the output purity for the maximally mixed state (as it saturates the bound in equation (78)). Therefore, it maximizes unitarity among all qubit channels with the same classical action.

Finally, we would like to relate the work presented here to studies on *cohering power P* [5] and *coherence generating power* $\tilde{P}$ [7, 8] of quantum channels. These notions were introduced within the framework of resource theory of coherence [2] and measure the ability of a channel to transform initially incoherent state to a coherent one. More formally they are defined by

$$P(\Phi) = \max_{\rho \in \mathcal{P}} \mathcal{C}_x(\Phi(\rho)), \tag{79a}$$

$$\tilde{P}(\Phi) = \langle \mathcal{C}_2(\Phi(\rho)) \rangle_{\mathcal{P}}, \tag{79b}$$

where $\mathcal{P}$ denotes the set of incoherent states ($\rho \in \mathcal{P}$ if and only if $\langle i| \rho |j\rangle = 0$ for $i \neq j$), $\langle \cdot \rangle_{\mathcal{P}}$ denotes the average over all incoherent states, and $\mathcal{C}_x$ is any measure of coherence for states, e.g., the relative entropy of coherence $\mathcal{C}_e$. Since both definitions involve only the action of Φ on incoherent states, we see that the only relevant parameters (defining the values of P and $\tilde{P}$) are given by the diagonals of D^i and C^{ij}, which are not constrained by the TP condition. Hence, given a fixed classical action T of Φ (so fixed diagonals of D^i), we can choose the diagonals of C^{ij} to be maximal possible (constrained only by complete positivity condition, $J_\Phi \geqslant 0$):

$$C_{kk}^{ij} = \sqrt{T_{ik}T_{jk}}, \tag{80}$$

and set all other matrix elements of J_Φ to zero. This way we will obtain a channel that maximizes both P and $\tilde{P}$ among all channels with a fixed classical action T. The action of such an optimal map is given by

$$\Phi(\cdot) = \Phi(\mathcal{D}(\cdot)) = \sum_j \langle j| \cdot |j\rangle |\psi_j\rangle\langle\psi_j|, \tag{81}$$

with

$$|\psi_j\rangle = \sum_i \sqrt{T_{ij}}|i\rangle. \tag{82}$$

We note that the above channels that maximize P and $\tilde{P}$ for a fixed T do not coincide with optimally coherified channels studied in this work. The reason for this is that the latter optimization depends on all coherence terms, whereas the former one only on the ones lying on the diagonal of C^{ij}. This emphasizes the main difference between our channel-oriented approach (when one focuses on the properties of the channel itself, specifically how close it is to a unitary evolution) and the state-oriented approach used in the studies of cohering power and coherence generating power (when one focuses on the properties of the output states for a restricted set of input states).

5. Conclusions

Any classical state of size d, represented by a diagonal density matrix $\rho = \rho^{\mathcal{D}}$, can be coherified to a pure state $\rho^{\mathcal{C}}$ with maximal coherence, which is transformed back into ρ by decoherence (see figure 2). In a similar way, one can try to coherify a quantum operation Φ represented by the corresponding Jamiołkowski state J_Φ. However, due to the trace preserving condition, the problem of coherifying a quantum channel has a much richer structure.

In this work we posed a general question: how to coherify a given classical map, represented by a stochastic transition matrix T, in an optimal way? Physically, this can be understood as looking for the most coherent (deterministic) underlying quantum evolution that can explain the observed random transformation T. Mathematically, among all quantum channels that decohere to T we looked for the one whose Jamiołkowski state has maximal coherence (as measured by entropic and 2-norm coherence). We demonstrated that the complete coherification to a (reversible) unitary channel is possible if and only if T is unistochastic, as schematically visualized in figure 3. To capture the limitations of possible coherifications of non-unistochastic maps we derived explicit bounds for the purity and entropy of the optimally coherified channel. Furthermore,

we provided an explicit coherification procedure that allows one to lower-bound the coherence of the optimal channel, and solved the optimal coherification problem for several classes of channels, including all one-qubit channels.

Studying possible coherifications of quantum channels can also shed some light on the structure and geometry of the set $\mathcal{S}_d$ of quantum operations [28]. For $d = 2$ the set of pure quantum states (the Bloch sphere) can be obtained by coherifying the set $\mathcal{P}_2$ of one-bit classical states. Analogously, the square $\mathcal{T}_2$ of classical stochastic matrices forms a skeleton of the larger set $\mathcal{S}_2$ of one-qubit quantum operations. Any unistochastic matrix $B \in \mathcal{B}_2 = \mathcal{U}_2$ can be coherified into a quantum unitary transformation, corresponding to a pure Jamiołkowski state $J_B^{\mathcal{C}}$. Furthermore, we have demonstrated that any classical transition matrix $T \in \mathcal{T}_2$ can be coherified to an optimal quantum channel, corresponding to a mixed state $J_T^{\mathcal{C}}$, that is an extremal point of $\mathcal{S}_2$. One would then like to check under what conditions a similar statement holds for higher dimensions, i.e., when the optimally coherified channels are extremal and have vanishing minimum output entropy.

Besides this problem concerning the geometry of $\mathcal{S}_d$, there are also other open questions that we would like to conclude this paper with. One could ask whether the optimally coherified channels are unique up to a unitary equivalence, i.e., can one find two channels whose Jamiołkowski states are not connected via unitary, and which maximize a given coherence measure among all Jamiołkowski states with a fixed diagonal? Furthermore, the expressions that lower- and upper-bound possible coherifications can definitely be improved, especially for bistochastic matrices. In this special case, exploring the boundary between unistochastic and bistochastic maps could be beneficial. Moreover, one might pursue a statistical approach and ask a question concerning a possible degree of coherification of a random stochastic matrix, or a generic quantum channel [31]. Last but not least, it would be very interesting to establish a closer connection between coherification approach to quantum channels, pursued in this work and based on the coherence of the corresponding Jamiołkowski states, with the earlier notion of the coherence power, related to the increase of coherence of selected quantum states by the action of a channel [5, 6, 8].

Note: Shortly after our work appeared on arXiv, another preprint studying the coherence of quantum channels was posted there [32].

Acknowledgments

Appendix A. Proof of theorem 2

We will make use of the following known results (see lemma 3.4 of [33] and equation (2.5) of [34]):

Lemma 4. *For every positive semi-definite matrix written in blocks we have the following decomposition*

$$\begin{bmatrix} A & X \\ X^\dagger & B \end{bmatrix} = U \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} U^\dagger + V \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix} V^\dagger, \tag{A1}$$

for some unitary operators U and V.

Lemma 5. *For $\boldsymbol{\lambda}(A)$ denoting the vector of eigenvalues of A arranged in a decreasing order we have*

$$\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B) \succ \boldsymbol{\lambda}(A + B). \tag{A2}$$

We are now ready to present the proof of theorem 2.

Proof. First, using lemma 4 iteratively one gets:

$$d \cdot J_\Phi = U_1 \begin{bmatrix} D^1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} U_1^\dagger + U_2 \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & D^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} U_2^\dagger + \dots + U_d \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D^d \end{bmatrix} U_d^\dagger. \tag{A3}$$

Then, using unitary invariance of the spectrum, the fact that $\boldsymbol{\lambda}_1 \succ \boldsymbol{\lambda}_1'$ and $\boldsymbol{\lambda}_2 \succ \boldsymbol{\lambda}_2'$ induces $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 \succ \boldsymbol{\lambda}_1' + \boldsymbol{\lambda}_2'$, and iteratively applying lemma 5 one arrives at equation (27). □

Appendix B. Coherifying qubit channels

Before we proceed to deriving the results presented in the main text, let us first recall the following fact. Given two channels, Φ_1 and Φ_2, whose Jamiołkowski states, J_{Φ_1} and J_{Φ_2}, are connected via a local unitary acting on the second subsystem, $J_{\Phi_2} = (\mathbb{1} \otimes \bar{U}) J_{\Phi_1} (\mathbb{1} \otimes \bar{U})^\dagger$, their Kraus decomposition satisfies

$$\Phi_1(\cdot) = \sum_i K_i(\cdot) K_i^\dagger, \quad \Phi_2(\cdot) = \sum_i K_i U^\dagger(\cdot) U K_i^\dagger. \tag{B1}$$

Now, using the block structure of the Jamiołkowski state, equation (12), and taking into account the TP condition, $\sum_i D^i = \mathbb{1}$, for a general qubit channel, we get:

$$J_\Phi = \frac{1}{2} \begin{bmatrix} D^1 & C^{12} \\ C^{21} & \mathbb{1} - D^1 \end{bmatrix}, \quad D^1 = \begin{bmatrix} a & c \\ c^* & \tilde{b} \end{bmatrix}. \tag{B2}$$

Consider a unitary U diagonalizing D^1, i.e., $U D^1 U^\dagger = \mathrm{diag}(\lambda_1, \lambda_2)$. Now, the same unitary will obviously also diagonalize $\mathbb{1} - D^1$. Therefore, a 4×4 unitary $V = \mathbb{1} \otimes U$ diagonalizes the D^i blocks of the Jamiołkowski state J_Φ:

$$V J_\Phi V^\dagger = \frac{1}{2} \begin{bmatrix} \lambda_1 & 0 & & \\ 0 & \lambda_2 & & \\ & & \tilde{\lambda}_1 & 0 \\ & & 0 & \tilde{\lambda}_2 \end{bmatrix}, \tag{B3}$$

where blank spaces mean arbitrary entries, $0 \leqslant \lambda_i \leqslant 1$ and $\lambda_1 + \lambda_2 = a + \tilde{b}$.

For $a \leqslant b$ we may obtain the optimally coherified state $J_\Phi^{\mathcal{C}}$ (with the spectrum saturating the bound given by $\boldsymbol{\mu}^\frown(T)$ from equation (38)) in the following way. We choose $\lambda_1 = a + \tilde{b}$ (resulting in $\lambda_2 = 0$), and set the off-diagonal element between λ_1 and $\tilde{\lambda}_2 = 1$ to the maximal value allowed by positivity, i.e., $\sqrt{a + \tilde{b}}$. As a result, $V J_\Phi^{\mathcal{C}} V^\dagger$ becomes a projector on two orthogonal pure states, which in turn means that the corresponding map is given by the decaying Kraus operators L_1, L_2 from equation (40). Since $V = \mathbb{1} \otimes U$, we can use equation (B1), to find the Kraus of decomposition of $\Phi^{\mathcal{C}}$ given by $L_i U^\top$. Finally, U is defined by $U D^1 U^\dagger = \mathrm{diag}(a + \tilde{b}, 0)$, which is exactly the unitary given in equation (39) (note that, since U is real, we have $U^\dagger = U^\top$)

Similarly, for $a \geqslant b$ we may choose $\lambda_1 = 1$ (resulting in $\lambda_2 = a - b$), and set the off-diagonal element between λ_1 and $\tilde{\lambda}_2 = \tilde{a} + b$ to the maximal value allowed by positivity, i.e., $\sqrt{\tilde{a} + b}$. As described in the main text, this then leads to the same results as in $a \leqslant b$ case, just with a and b exchanged, as well as with all 2×2 matrices X transformed by replacing X_{kl} with $X_{\tilde{k}\tilde{l}}$.

Appendix C. Coherifyng qutrit channels

C.1. Cyclic matrices

The general form of D^i matrices is as follows:

$$D^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & a & x \\ 0 & x^* & b \end{bmatrix}, \quad D^2 = \begin{bmatrix} c & 0 & y \\ 0 & 0 & 0 \\ y^* & 0 & \tilde{b} \end{bmatrix}, \quad D^3 = \begin{bmatrix} \tilde{c} & z & 0 \\ z^* & \tilde{a} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, in order to satisfy the TP condition, $\sum_i D^i = \mathbb{1}$, we need $x = y = z = 0$. Hence, the Jamiołkowski state J_Φ can be recast in the following form (note that columns and rows number 1, 5 and 9, composed only of zeros, have been removed):

$$J_\Phi = \frac{1}{3} \begin{bmatrix} a & 0 & x_1 & x_2 & x_3 & x_4 \\ 0 & b & y_1 & y_2 & y_3 & y_4 \\ x_1^* & y_1^* & c & 0 & x_5 & x_6 \\ x_2^* & y_2^* & 0 & \tilde{b} & y_5 & y_6 \\ x_3^* & y_3^* & x_5^* & y_5^* & \tilde{c} & 0 \\ x_4^* & y_4^* & x_6^* & y_6^* & 0 & \tilde{a} \end{bmatrix}. \tag{C1}$$

Now, using theorem 2, we get:

$$\boldsymbol{\mu}_{\mathrm{opt}} := \frac{1}{3} [\mu, 3 - \mu, 0, \ldots, 0] \succ \boldsymbol{\lambda}(J_\Phi), \tag{C2}$$

where

$$\mu = \max(a, b) + \max(c, \tilde{b}) + \max(\tilde{c}, \tilde{b}). \tag{C3}$$

Moreover, one can construct optimally coherified matrix $J_\Phi^{\mathcal{C}}$ such that $\boldsymbol{\lambda}(J_\Phi^{\mathcal{C}}) = \boldsymbol{\mu}_{\mathrm{opt}}$. To do this one simply needs to group together the maximal/minimal terms of each 2×2 matrix and set the corresponding off-diagonal terms to the maximal values allowed by the positivity constraint. For example, if $a \geqslant b, \tilde{b} \geqslant c$ and $\tilde{c} \geqslant \tilde{a}$, one chooses

$$x_2 = \sqrt{a\tilde{b}}, \quad x_3 = \sqrt{a\tilde{c}}, \quad y_5 = \sqrt{\tilde{b}\tilde{c}},$$
$$y_1 = \sqrt{bc}, \quad y_4 = \sqrt{\tilde{a}b}, \quad x_6 = \sqrt{\tilde{a}c},$$

and sets the rest of off-diagonal terms to zero. Note that this is exactly the construction introduced in section 3.2.2 and illustrated in figure 5.

C.2. Single-row matrices

The general form of D^i matrices is as follows:

$$D^1 = \begin{bmatrix} a & x & 0 \\ x^* & b & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad D^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{bmatrix}, \quad D^3 = \begin{bmatrix} \tilde{a} & x' & y \\ x'^* & \tilde{b} & z \\ y^* & z^* & \tilde{c} \end{bmatrix}.$$

Clearly, in order to satisfy the TP condition we need $x' = -x$ and $y = z = 0$. We now note that a 3×3 unitary matrix U diagonalizing D^1,

$$U = \begin{bmatrix} U_{11} & U_{12} & 0 \\ U_{21} & U_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad UD^1U^\dagger = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{C4}$$

also diagonalizes D^2 (by keeping it unchanged) and D^3. Therefore, a 9×9 unitary $V = \mathbf{1} \otimes U$ diagonalizes the D^i blocks of the Jamiołkowski state J_Φ:

$$VJ_\Phi V^\dagger = \frac{1}{3} \begin{bmatrix} \lambda_1 & 0 & 0 & & & & & & \\ 0 & \lambda_2 & 0 & & & & & & \\ 0 & 0 & 0 & & & & & & \\ & & & 0 & 0 & 0 & & & \\ & & & 0 & 0 & 0 & & & \\ & & & 0 & 0 & c & & & \\ & & & & & & \tilde{\lambda}_1 & 0 & 0 \\ & & & & & & 0 & \tilde{\lambda}_2 & 0 \\ & & & & & & 0 & 0 & \tilde{c} \end{bmatrix}, \tag{C5}$$

where blank spaces mean arbitrary entries and $\lambda_1 + \lambda_2 = a + b$.

Without loss of generality let us assume that $\lambda_1 \geqslant \lambda_2$. Then, using theorem 2, we get $\boldsymbol{\mu}(\lambda_1) \succ \boldsymbol{\lambda}(J_\Phi)$ with $\boldsymbol{\mu}^\downarrow(\lambda_1)$ given by

$$\frac{1}{3}[\lambda_1 + \max(\tilde{\lambda}_2, \tilde{c}) + c, \lambda_2 + \mathrm{med}(\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{c}), \min(\tilde{\lambda}_1, \tilde{c}), 0, \ldots, 0],$$

where med denotes the second largest element of the set. Note also that, for fixed a, b, c, the vector $\boldsymbol{\mu}(\lambda_1)$ is just a function of λ_1, since $\lambda_2 = a + b - \lambda_1$. Now, maximizing λ_1 maximizes both $\mu_1^\downarrow$ and $\mu_1^\downarrow + \mu_2^\downarrow$ (recall that $\mu_1^\downarrow + \mu_2^\downarrow + \mu_3^\downarrow$ is constant and equal to 1), and so $\boldsymbol{\mu}(x) \succ \boldsymbol{\mu}(y)$ for $x \geqslant y$. In order to find the optimal $\boldsymbol{\mu}_{\mathrm{opt}}$ (optimal meaning that for all λ_1 we have $\boldsymbol{\mu}_{\mathrm{opt}} \succ \boldsymbol{\mu}(\lambda_1)$) we thus need to maximize λ_1. Recalling that we have two constraints, $0 \leqslant \lambda_1 \leqslant a + b$ and $\lambda_1 \leqslant 1$, we arrive at two cases.

For $a + b \leqslant 1$ the maximal (and thus optimal) value of λ_1 is $a + b$, which also results in $\lambda_2 = 0$. The optimal bounding vector is then given by

$$\boldsymbol{\mu}_{\mathrm{opt}} = \frac{1}{3}[1 + a + b + c, \max(\tilde{a} - b, \tilde{c}), \min(\tilde{a} - b, \tilde{c}), 0, \ldots, 0]. \tag{C6}$$

Moreover, one can construct the Jamiołkowski state $J_\Phi^{\mathcal{C}}$ that saturates this optimal bound, i.e., $\boldsymbol{\lambda}(J_\Phi^{\mathcal{C}}) = \boldsymbol{\mu}_{\mathrm{opt}}$. This can be achieved, again, by setting the adequate off-diagonal terms in equation (C5) to the maximal possible value allowed by the positivity condition. More precisely, we group $\lambda_1 = a + b, c$ and $\tilde{\lambda}_2 = 1$ together, leaving the remaining two terms, $\tilde{\lambda}_1 = \tilde{a} - b$ and $\tilde{c}$, ungrouped. As a result, $VJ_\Phi^{\mathcal{C}}V^\dagger$ becomes a projector on three orthogonal pure states, which in turn means that the corresponding map is given by the following Kraus operators:

$$K_1 = \begin{bmatrix} \sqrt{a+b} & 0 & 0 \\ 0 & 0 & \sqrt{c} \\ 0 & 1 & 0 \end{bmatrix}, \quad K_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sqrt{\tilde{a}-b} & 0 & 0 \end{bmatrix}, \quad K_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{\tilde{c}} \end{bmatrix}.$$

Finally, using equation (B1) we conclude that the Kraus operators corresponding to the optimally coherified channel (with Jamiołkowski state J_Φ^C) are given by $K_i U^\top$ with K_i as above and U defined by equation (C4) with $\lambda_1 = a + b$ and $\lambda_2 = 0$, i.e.,

$$U = \frac{1}{\sqrt{a+b}} \begin{bmatrix} \sqrt{a} & \sqrt{b} & 0 \\ \sqrt{b} & -\sqrt{a} & 0 \\ 0 & 0 & \sqrt{a+b} \end{bmatrix}. \tag{C7}$$

For $a + b \geqslant 1$ the maximal (and thus optimal) value of λ_1 is 1, which also results in $\lambda_2 = a - \tilde{b}$. The optimal bounding vector is then given by

$$\boldsymbol{\mu}_{\mathrm{opt}} = \frac{1}{3}[\mu, \, 3-\mu, \, 0, \, \dots, 0], \quad \mu = 1 + \max(\tilde{a}+\tilde{b}, \tilde{c}) + c. \tag{C8}$$

The bound $\boldsymbol{\mu}_{\mathrm{opt}} \succ \boldsymbol{\lambda}(J_\Phi)$ can be saturated in a usual way—by proper grouping of diagonal elements and setting the corresponding off-diagonal elements to the maximal value allowed by the positivity condition. If $\tilde{c} \geqslant \tilde{a} + \tilde{b}$ then we group together $\lambda_1 = 1, c$ and $\tilde{c}$, with $\lambda_2 = a - \tilde{b}$ and $\tilde{\lambda}_2 = \tilde{a} + \tilde{b}$ forming the other group; otherwise we group together λ_1, c and λ_2, with λ_2 and $\tilde{c}$ forming the other group. In the former case the resulting Kraus operators of the optimally coherified channel read

$$K_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \sqrt{c} \\ 0 & 0 & \sqrt{\tilde{c}} \end{bmatrix} U^\top, \quad K_2 = \begin{bmatrix} 0 & \sqrt{a-\tilde{b}} & 0 \\ 0 & 0 & 0 \\ 0 & \sqrt{\tilde{a}+\tilde{b}} & 0 \end{bmatrix} U^\top,$$

and in the latter case they read

$$K_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \sqrt{c} \\ 0 & \sqrt{\tilde{a}+\tilde{b}} & 0 \end{bmatrix} U^\top, \quad K_2 = \begin{bmatrix} 0 & \sqrt{a-\tilde{b}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{\tilde{c}} \end{bmatrix} U^\top,$$

with U in both cases defined by equation (C4) with $\lambda_1 = 1$ and $\lambda_2 = a - \tilde{b}$, i.e.,

$$U = \frac{1}{\sqrt{\tilde{a}+\tilde{b}}} \begin{bmatrix} \sqrt{\tilde{b}} & \sqrt{\tilde{a}} & 0 \\ \sqrt{\tilde{a}} & -\sqrt{\tilde{b}} & 0 \\ 0 & 0 & \sqrt{\tilde{a}+\tilde{b}} \end{bmatrix}. \tag{C9}$$

C.3. Double-row matrices

The general form of D^i matrices is as follows:

$$D^1 = \begin{bmatrix} a & x & y \\ x^* & b & z \\ y^* & z^* & c \end{bmatrix}, D^2 = \begin{bmatrix} \tilde{a} & -x & -y \\ -x^* & \tilde{b} & -z \\ -y^* & -z^* & \tilde{c} \end{bmatrix}, D^3 = 0.$$

We now note that a 3×3 unitary matrix U diagonalizing D^1,

$$U = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix}, \quad U D^1 U^\dagger = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \tag{C10}$$

also diagonalizes D^2 and D^3 (by keeping it unchanged). Therefore, a 9×9 unitary $V = \mathbf{1} \otimes U$ diagonalizes the D^i blocks of the Jamiołkowski state J_Φ:

$$
V J_\Phi V^\dagger = \frac{1}{3}
\begin{bmatrix}
\lambda_1 & 0 & 0 & & & & & & \\
0 & \lambda_2 & 0 & & & & & & \\
0 & 0 & \lambda_3 & & & & & & \\
& & & \tilde{\lambda}_1 & 0 & 0 & & & \\
& & & 0 & \tilde{\lambda}_2 & 0 & & & \\
& & & 0 & 0 & \tilde{\lambda}_3 & & & \\
& & & & & & 0 & 0 & 0 \\
& & & & & & 0 & 0 & 0 \\
& & & & & & 0 & 0 & 0
\end{bmatrix},
\tag{C11}
$$

where blank spaces mean arbitrary entries and $\lambda_1 + \lambda_2 + \lambda_3 = a + b + c$. To shorten the notation we define $s := a + b + c$.

Without loss of generality we may assume $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$, so that $\tilde{\lambda}_1 \leqslant \tilde{\lambda}_2 \leqslant \tilde{\lambda}_3$. Then, using theorem 2, we have

$$
\boldsymbol{\mu}(\lambda_1, \lambda_3) = \frac{1}{3}[\lambda_1 + \tilde{\lambda}_3, 1, \lambda_3 + \tilde{\lambda}_1, 0, \ldots, 0] \succ \boldsymbol{\lambda}(J_\Phi).
\tag{C12}
$$

Now, we observe that $\boldsymbol{\mu}(x, y) \succ \boldsymbol{\mu}(x', y')$ for $x \geqslant x'$ and $y \leqslant y'$. We thus aim at maximizing the largest eigenvalue of D^1 while minimizing its smallest eigenvalue. Again, noting that we are constrained by $0 \leqslant \lambda_i \leqslant s$ and $\lambda_i \leqslant 1$ we arrive at three distinct cases dependent on the value of s.

For $s \leqslant 1$ the maximal (and thus optimal) value of λ_1 is s, which also results in $\lambda_2 = \lambda_3 = 0$. The optimal bounding vector is then given by

$$
\boldsymbol{\mu}_{\mathrm{opt}} = \frac{1}{3}[1 + s, 1, 1 - s, 0, \ldots, 0].
\tag{C13}
$$

The above optimal spectrum can be realized by the Jamiołkowski state $J_\Phi^{\mathcal{C}}$ by simply setting in equation (C11) the off-diagonal terms between $\lambda_1 = s$ and $\tilde{\lambda}_2 = 1$ (or $\tilde{\lambda}_3 = 1$) to $\sqrt{s}$. Recalling the relation between the Kraus operators corresponding to Jamiołkowski states connected via a local unitary, equation (B1), we find that the Kraus decomposition of the optimally coherified channel is given by:

$$
K_1 = \begin{bmatrix} \sqrt{s} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top, \quad
K_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} U^\top, \quad
K_3 = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{1-s} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top.
$$

with U being a unitary such that

$$
U = \frac{1}{\sqrt{s}} \begin{bmatrix} \sqrt{a} & \times & \times \\ \sqrt{b} & \times & \times \\ \sqrt{c} & \times & \times \end{bmatrix},
\tag{C14}
$$

and $\times$ denoting arbitrary entries as long as U stays unitary, e.g.,

$$
U = \begin{bmatrix}
\sqrt{\dfrac{a}{s}} & -\sqrt{\dfrac{b}{a+b}} & -\sqrt{\dfrac{ac}{(a+b)s}} \\[8pt]
\sqrt{\dfrac{b}{s}} & \sqrt{\dfrac{a}{a+b}} & -\sqrt{\dfrac{bc}{(a+b)s}} \\[8pt]
\sqrt{\dfrac{c}{s}} & 0 & \dfrac{a+b}{\sqrt{(a+b)s}}
\end{bmatrix}.
$$

Also note that the position of 1 in matrices describing K_1 and K_2 can be exchanged.

For $s \geqslant 2$ the optimal values are $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = s - 2$. The optimal bounding vector is then given by

$$
\boldsymbol{\mu}_{\mathrm{opt}} = \frac{1}{3}[4 - s, 1, s - 2, 0, \ldots, 0],
\tag{C15}
$$

which can be achieved by the coherified Jamiołkowski state in an analogous way to the first case. This leads to the following decomposition of $\Phi^{\mathcal{C}}$ into the set of Kraus operators,

$$
K_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \sqrt{3-s} \\ 0 & 0 & 0 \end{bmatrix} U^\top, \quad
K_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top, \quad
K_3 = \begin{bmatrix} 0 & 0 & \sqrt{s-2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top,
$$

with U being a unitary such that

$$
U = \frac{1}{\sqrt{3-s}} \begin{bmatrix} \times & \times & \sqrt{\tilde{a}} \\ \times & \times & \sqrt{\tilde{b}} \\ \times & \times & \sqrt{\tilde{c}} \end{bmatrix}.
\tag{C16}
$$

and × denoting arbitrary entries as long as U stays unitary, e.g.,

$$U = \begin{bmatrix} -\sqrt{\dfrac{\tilde{a}\tilde{c}}{(\tilde{a}+\tilde{b})(3-s)}} & -\sqrt{\dfrac{\tilde{b}}{\tilde{a}+\tilde{b}}} & \sqrt{\dfrac{\tilde{a}}{3-s}} \\[2mm] -\sqrt{\dfrac{\tilde{b}\tilde{c}}{(\tilde{a}+\tilde{b})(3-s)}} & \sqrt{\dfrac{\tilde{a}}{\tilde{a}+\tilde{b}}} & \sqrt{\dfrac{\tilde{b}}{3-s}} \\[2mm] \dfrac{\tilde{a}+\tilde{b}}{\sqrt{(\tilde{a}+\tilde{b})(3-s)}} & 0 & \sqrt{\dfrac{\tilde{c}}{3-s}} \end{bmatrix}. \tag{C17}$$

Again, we note that the position of 1 in matrices describing K_1 and K_2 can be exchanged.

Finally, for $1 < s < 2$ the optimal values are $\lambda_1 = 1$, $\lambda_2 = s - 1$ and $\lambda_3 = 0$. The optimal bounding vector is then given by

$$\boldsymbol{\mu}_{\mathrm{opt}} = \frac{1}{3}[2, 1, 0, \ldots, 0]. \tag{C18}$$

The above spectrum can be achieved by the optimally coherified Jamiołkowski state $J_\Phi^{\mathcal{C}}$ by setting the off-diagonal terms in equation (C11) appropriately. More precisely, one chooses the term between $\lambda_1 = 1$ and $\tilde{\lambda}_3 = 1$ to be 1, and the term between $\lambda_2 = s - 1$ and $\tilde{\lambda}_2 = 2 - s$ to be $\sqrt{(s-1)(2-s)}$. The resulting Kraus operators are given by

$$K_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} U^\top, \quad K_2 = \begin{bmatrix} 0 & \sqrt{s-1} & 0 \\ 0 & \sqrt{2-s} & 0 \\ 0 & 0 & 0 \end{bmatrix} U^\top, \tag{C19}$$

with

$$U = \begin{bmatrix} \sqrt{\dfrac{\tilde{b}}{\tilde{a}+\tilde{b}}} & \sqrt{\dfrac{\tilde{a}(a-\tilde{b})}{(\tilde{a}+\tilde{b})(s-1)}} & \sqrt{\dfrac{\tilde{a}c}{(\tilde{a}+\tilde{b})(s-1)}} \\[2mm] -\sqrt{\dfrac{\tilde{a}}{\tilde{a}+\tilde{b}}} & \sqrt{\dfrac{\tilde{b}(a-\tilde{b})}{(\tilde{a}+\tilde{b})(s-1)}} & \sqrt{\dfrac{\tilde{b}c}{(\tilde{a}+\tilde{b})(s-1)}} \\[2mm] 0 & \sqrt{\dfrac{c}{s-1}} & -\sqrt{\dfrac{a-\tilde{b}}{s-1}} \end{bmatrix} \tag{C20}$$

if $a + b \geqslant 1$ and

$$U = \begin{bmatrix} \sqrt{\dfrac{a\tilde{c}}{(a+b)(2-s)}} & \sqrt{\dfrac{a(\tilde{a}-b)}{(a+b)(2-s)}} & \sqrt{\dfrac{b}{a+b}} \\[2mm] \sqrt{\dfrac{b\tilde{c}}{(a+b)(2-s)}} & \sqrt{\dfrac{b(\tilde{a}-b)}{(a+b)(2-s)}} & -\sqrt{\dfrac{a}{a+b}} \\[2mm] -\sqrt{\dfrac{\tilde{a}-b}{2-s}} & \sqrt{\dfrac{\tilde{c}}{2-s}} & 0 \end{bmatrix} \tag{C21}$$

if $a + b \leqslant 1$.

Appendix D. Polygon constraints

First, we derive the expression for the purity bound, equation (58). The expression for Δ_1, equation (59), comes directly from the fact that $|D_{kl}^i|^2 \leqslant \alpha_{kl}^i T_{ik} T_{il}$. To obtain Δ_2, equation (59), let us start by parametrizing the matrix from equation (57), i.e., the 3×3 submatrix of J_Φ, in the following way

$$A := \frac{1}{d} \begin{bmatrix} T_{ik} & x\sqrt{T_{ik}T_{il}} & y\sqrt{T_{ik}T_{jm}} \\[2mm] x\sqrt{T_{ik}T_{il}} & T_{il} & z\sqrt{T_{il}T_{jm}} \\[2mm] y\sqrt{T_{ik}T_{jm}} & z\sqrt{T_{il}T_{jm}} & T_{jm} \end{bmatrix}, \tag{D1}$$

with $0 \leqslant x, y, z \leqslant 1$. We then have $\det A \geqslant 0$ if and only if

$$1 + 2xyz - x^2 - y^2 - z^2 \geqslant 0. \tag{D2}$$

Now, our aim is to upper-bound the squared moduli of the off-diagonal terms of A (for fixed T), given the above constraint and the fact that $x \leqslant \alpha_{kl}^i$ for some $\alpha_{kl}^i < 1$. First, assume that x is fixed, so that effectively we want to find the maximum of $T_{ik}y^2 + T_{il}z^2$(in fact, the optimal choice is to maximize x, i.e., set $x = \alpha_{kl}^i$). It is straightforward to check that it is achieved at the boundary of the constraint, i.e., when equation (D2) becomes an equality. One can then solve for y, substitute it to $T_{ik}y^2 + T_{il}z^2$, and find the maximum of the resulting expression over z. This leads to the following bound on the off-diagonal terms of A:

$$|A_{13}|^2 + |A_{23}|^2 \leqslant \frac{T_{jm}}{2d^2}(T_{ik} + T_{il} + \beta_{kl}^i), \tag{D3}$$

with β_{kl}^i defined in equation (60). As in order to achieve unit purity one needs $|A_{13}|^2 + |A_{23}|^2 = T_{jm}(T_{ik} + T_{il})/d^2$, the above bound leads to the following deficit of purity:

$$\delta_{kl}^i := \frac{T_{jm}}{2d^2}(T_{ik} + T_{il} - \beta_{kl}^i). \tag{D4}$$

Finally, the above deficit adds up for every choice of T_{jm} not equal to T_{ik} or T_{il} (i.e., for all off-diagonal elements sharing row or column with T_{ik} or T_{il} in figure 7(b)), so that using $\sum_{i,j} T_{ij} = d$, we finally arrive at equation (59).

We now proceed to the proof of the majorization bound, equation (61). Note that, using lemma 4 from appendix A, we can rewrite D^i (up to permutations) as

$$D^i = \begin{bmatrix} A & X \\ X^\dagger & B \end{bmatrix} = U \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} U^\dagger + V \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix} V^\dagger, \tag{D5}$$

with

$$A = \begin{bmatrix} T_{ik} & \sqrt{\alpha_{kl}^i T_{ik} T_{il}} \\ \sqrt{\alpha_{kl}^i T_{ik} T_{il}} & T_{il} \end{bmatrix}, \tag{D6}$$

and U, V being unitary. The eigenvalues of A are given by

$$\boldsymbol{\lambda}(A) = \frac{1}{2}[T_{ik} + T_{il} + \beta_{kl}^i, \, T_{ik} + T_{il} - \beta_{kl}^i, \, 0, \, \ldots, \, 0], \tag{D7}$$

whereas the largest eigenvalue of B is constrained by

$$\lambda_1(B) \leqslant \mathrm{Tr}(B) = \mathrm{Tr}(D^i) - T_{ik} - T_{il} = 1 - T_{ik} - T_{il}, \tag{D8}$$

where we used the fact that T is bistochastic. Thus, using lemma 5 and choosing $k = k_i$ and $l = l_i$ minimizing α_{kl}^i, we arrive at equation (61).

Note that the above construction can be easily generalized to cases where for a given D^i there are many pairs (k, l) for which $\alpha_{kl}^i < 1$. Instead of a 2×2 matrix A, one simply chooses it to contain all off-diagonal elements of D^i that are constrained beyond the positivity condition, finds its eigenvalues, and obtains a tighter bound using lemma 5 again.

ORCID iDs

Kamil Korzekwa ⬤ https://orcid.org/0000-0002-0683-5469

References

[1] Zhang W-M, Feng D H and Gilmore R 1990 Coherent states: theory and some applications *Rev. Mod. Phys.* **62** 867

[2] Baumgratz T, Cramer M and Plenio M B 2014 Quantifying coherence *Phys. Rev. Lett.* **113** 140401

[3] Brandão F G S L and Gour G 2015 Reversible framework for quantum resource theories *Phys. Rev. Lett.* **115** 070503

[4] Streltsov A, Adesso G and Plenio M B 2016 Quantum coherence as a resource *Rev. Mod. Phys.* **89** 041003

[5] Mani A and Karimipour V 2015 Cohering and decohering power of quantum channels *Phys. Rev. A* **92** 032331

[6] García-Díaz M, Egloff D and Plenio M B 2016 A note on coherence power of *n*-dimensional unitary operators *Quant. Inf. Comput.* **16** 1282

[7] Zanardi P, Styliaris G and Venuti L C 2017 Coherence-generating power of quantum unitary maps and beyond *Phys. Rev. A* **95** 052306

[8] Zanardi P, Styliaris G and Venuti L C 2017 Measures of coherence-generating power for quantum unital operations *Phys. Rev. A* **95** 052307

[9] Xi Z, Hu M, Li Y and Fan H 2016 Entropic characterization of coherence in quantum evolutions arXiv:1510.06473

[10] Bu K, Kumar A, Zhang L and Wu J 2017 Cohering power of quantum operations *Phys. Lett. A* **381** 1670–6

[11] Jamiołkowski A 1972 Linear transformations which preserve trace and positive semidefiniteness of operators *Rep. Math. Phys.* **3** 275–8

[12] Choi M-D 1975 Completely positive linear maps on complex matrices *Linear Algebr. Appl.* **10** 285–90

[13] Życzkowski K and Bengtsson I 2004 On duality between quantum maps and quantum states *Open Syst. Inf. Dyn.* **11** 3–42

[14] Wallman J, Granade C, Harper R and Flammia S T 2015 Estimating the coherence of noise *New J. Phys.* **17** 113020

[15] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)

[16] Bengtsson I and Życzkowski K 2017 *Geometry of Quantum States* 2nd edn (Cambridge: Cambridge University Press)

[17] Lostaglio M, Jennings D and Rudolph T 2015 Description of quantum coherence in thermodynamic processes requires constraints beyond free energy *Nat. Commun.* **6** 6383

[18] Lostaglio M, Korzekwa K, Jennings D and Rudolph T 2015 Quantum coherence, time-translation symmetry, and thermodynamics *Phys. Rev. X* **5** 021001

[19] Åberg J 2006 Quantifying superposition arXiv:quant-ph/0612146

[20] Lakshminarayan A, Puchała Z and Życzkowski K 2014 Diagonal unitary entangling gates and contradiagonal quantum states *Phys. Rev. A* **90** 032303

[21] Tadej W and Życzkowski K 2006 A concise guide to complex Hadamard matrices *Open Syst. Inf. Dyn.* **13** 133–77

[22] Kleinmann M, Kampermann H, Meyer T and Bruß D 2006 Physical purification of quantum states *Phys. Rev.* A **73** 062309
[23] Roga W, Fannes M and Życzkowski K 2011 Entropic characterization of quantum operations *Int. J. Quantum Inf.* **9** 1031–45
[24] Dunkl C and Życzkowski K 2009 Volume of the set of unistochastic matrices of order 3 and the mean Jarlskog invariant *J. Math. Phys.* **50** 123521
[25] Bengtsson I, Ericsson Å, Kuś W, Tadej W and Życzkowski K 2005 Birkhoff's polytope and unistochastic matrices, $n = 3$ and $n = 4$ *Commun. Math. Phys.* **259** 307–24
[26] Li C-K and Woerdeman H J 1997 Special classes of positive and completely positive maps *Linear Algebr. Appl.* **255** 247–58
[27] Levick J, Kribs D W and Pereira R 2017 Quantum privacy and Schur product channels arXiv:1709.01752
[28] Ruskai M B, Szarek S and Werner E 2002 An analysis of completely-positive trace-preserving maps on M_2 *Linear Algebr. Appl.* **347** 159–87
[29] Życzkowski K 2008 Quartic quantum theory: an extension of the standard quantum mechanics *J. Phys. A: Math. Theor.* **41** 355302
[30] Cappellini V 2007 private communication
[31] Bruzda W, Cappellini V, Sommers H-J and Życzkowski K 2009 Random quantum operations *Phys. Lett.* A **373** 320
[32] Datta C, Sazim S, Pati A K and Agrawal P 2017 Coherence of quantum channels arXiv:1710.05015
[33] Bourin J-C and Lee E-Y 2012 Unitary orbits of Hermitian operators with convex or concave functions *Bull. London Math. Soc.* **44** 1085–102
[34] Ando T 1994 Majorizations and inequalities in matrix theory *Linear Algebr. Appl.* **199** 17–67