

Convex approximations of quantum channels

Massimiliano F. Sacchi*

*Istituto di Fotonica e Nanotecnologie – CNR, Piazza Leonardo da Vinci 32, I-20133, Milano, Italy
and QUIT Group, Dipartimento di Fisica, Università di Pavia, via A. Bassi 6, I-27100 Pavia, Italy*

Tito Sacchi

Istituto Comprensivo di via Acerbi, via E. Acerbi 21, I-27100 Pavia, Italy

(Received 6 May 2017; published 8 September 2017)

We address the problem of optimally approximating the action of a desired and unavailable quantum channel Φ having at our disposal a single use of a given set of other channels $\{\Psi_i\}$. The problem is recast to look for the least distinguishable channel from Φ among the convex set $\sum_i p_i \Psi_i$, and the corresponding optimal weights $\{p_i\}$ provide the optimal convex mixing of the available channels $\{\Psi_i\}$. For single-qubit channels we study specifically cases where the available convex set corresponds to covariant channels or to Pauli channels, and the desired target map is an arbitrary unitary transformation or a generalized damping channel.

DOI: [10.1103/PhysRevA.96.032311](https://doi.org/10.1103/PhysRevA.96.032311)

I. INTRODUCTION

Quantum channels, or trace-preserving completely positive maps, represent all possible deterministic quantum operations one can perform over a quantum system [1]. It is well known that dilation theorems prove that a quantum channel can be realized by means of a unitary transformation which couples the system to an ancilla with fixed state preparation [2,3]. This “realization” theorem is clearly a powerful theoretical instrument, but in a realistic scenario where the available technology is limited it may give just a poor indication about the best effective experimental realization of a desired quantum channel.

In the same spirit adopted to characterize the universality of fixed quantum gates for quantum computation [1], to study programmable devices to achieve different quantum channels [4–7] or measurements [8,9], and to perform the purification of noisy quantum measurements [10], in this paper we address the problem of optimally approximating the action of a desired and unavailable quantum channel Φ over a Hilbert space \mathbb{H} by an operational approach, when only a given set of quantum channels $\{\Psi_i\}$ for \mathbb{H} is at our disposal for a single use. More specifically, we want to look for the best convex combination among the channels of the given set that most resembles the desired Φ , i.e., that is the least distinguishable from Φ itself. This approach clearly has an immediate experimental application, especially when the quantum operations effectively feasible in a laboratory are limited due to intrinsic restrictions, unavailable technology, or even economic reasons. Further relevance of this approach is due to the fact that a convex sum of quantum channels offers the possibility of performing different experiments followed by postprocessing of experimental data [11,12], when the quantities of interest are linear with respect to quantum operations.

We note that when the target channel is unital and the available set corresponds to all possible unitary transformations, our problem is related to the quantification of the distance

between unital maps and random-unitary channels [13] and to the disproved “quantum Birkoff’s conjecture” [14,15].

II. OPTIMAL CONVEX APPROXIMATION OF QUANTUM CHANNELS

The probability p_{discr} of optimally discriminating between two quantum channels Φ_0 and Φ_1 is quantified by the expression [16]

$$p_{\text{discr}}(\Phi_0, \Phi_1) = \frac{1}{2} + \frac{1}{4} \|\Phi_0 - \Phi_1\|_{\diamond}, \quad (1)$$

where $\|\cdot\|_{\diamond}$ denotes the completely bounded trace norm [17] (or, equivalently, the diamond norm [18]).

By defining the positive Choi operator [19] $R_{\Phi} = (\Phi \otimes I)|\eta\rangle\langle\eta|$, which corresponds to the action of the map Φ over one party of a maximally entangled vector $|\eta\rangle \equiv \sum_{n=1}^d |n\rangle \otimes |n\rangle$ of $\mathbb{H} \otimes \mathbb{H}$, with $d = \dim(\mathbb{H})$, let us recall the identity [16]

$$\begin{aligned} \|\Phi_0 - \Phi_1\|_{\diamond} &= \max_{\text{Tr}[\xi^{\dagger}\xi]=1} \|(I \otimes \xi)(R_{\Phi_0} - R_{\Phi_1})(I \otimes \xi^{\dagger})\|_1, \end{aligned} \quad (2)$$

where $\|A\|_1$ denotes the trace norm of A , namely [20],

$$\|A\|_1 = \text{Tr}\sqrt{A^{\dagger}A} = \sum_i s_i(A), \quad (3)$$

$\{s_i(A)\}$ representing the singular values of A . In the case of Eq. (2), since the operator inside the norm is Hermitian, the singular values just correspond to the absolute value of the eigenvalues. We also note that any operator ξ providing the maximum in Eq. (2) corresponds to an optimal input state $(I \otimes \xi)|\eta\rangle$ for the discrimination [21]. We recall here also the result in Refs. [23] and [24], namely, for arbitrary unitary maps $\mathcal{V} \equiv V(\cdot)V^{\dagger}$ and $\mathcal{Z} \equiv Z(\cdot)Z^{\dagger}$ one has

$$\|\mathcal{V} - \mathcal{Z}\|_{\diamond} = 2\sqrt{1 - r(Z^{\dagger}V)^2}, \quad (4)$$

where $r(Z^{\dagger}V)$ denotes the distance from the origin of the complex plane of the polygon whose vertices stay on the circle of unit radius and correspond to the singular values of the unitary matrix $Z^{\dagger}V$. For $d = 2$, since there are only two eigenvalues, one easily finds by simple geometry that $r(U) = \frac{1}{2}|\text{Tr}[U]|$.

*msacchi@unipv.it

The problem of the optimal convex approximation of a quantum channel is implicitly posed by the following definition.

Definition. The optimal convex approximation of a quantum channel Φ with respect to (w.r.t.) a given set of quantum channels $\{\Psi_i\}$ is given by $\sum_i p_i^{\text{opt}} \Psi_i$, where $\{p_i^{\text{opt}}\}$ denotes the vector of probabilities

$$\{p_i^{\text{opt}}\} = \arg \min_{\{p_i\}} \|\Phi - \sum_i p_i \Psi_i\|_{\diamond}. \quad (5)$$

The effectiveness of the optimal convex approximation is then quantified by the $\{\Psi_i\}$ distance

$$D_{\{\Psi_i\}}(\Phi) \equiv \min_{\{p_i\}} \|\Phi - \sum_i p_i \Psi_i\|_{\diamond}, \quad (6)$$

which provides through Eq. (1) the worst probability of discriminating the desired channel Φ from any of the available channels $\sum_i p_i \Psi_i$. Clearly, our definition of optimal convex approximation can be suitably changed by referring to any other figure of merit that quantifies the distance between quantum operations [25,26].

Since we have in mind an operational approach where the channels in the set $\{\Psi_i\}$ are experimentally available, we always assume that this set contains the identity map \mathcal{I} . We note that the formulation of the diamond norm as a semidefinite program satisfying strong duality [27–29] allows its efficient calculation. Moreover, the convexity of the norm itself allows one to search for the minimum by means of standard software of convex optimization [30,31].

From the convexity of the diamond norm, one obtains the upper bound

$$D_{\{\Psi_i\}}(\Phi) \leq \min_i \|\Phi - \Psi_i\|_{\diamond} = \min_i D_{\Psi_i}(\Phi). \quad (7)$$

On the other hand, since from Eq. (2) one has

$$\|\Phi_0 - \Phi_1\|_{\diamond} \geq \frac{1}{d} \|R_{\Phi_0} - R_{\Phi_1}\|_1, \quad (8)$$

one obtains the lower bound

$$D_{\{\Psi_i\}}(\Phi) \geq \frac{1}{d} \min_{\{p_i\}} \|R_{\Phi} - R_{\sum_i p_i \Psi_i}\|_1. \quad (9)$$

From the unitarily invariance of the diamond norm, note also that for all unitary maps \mathcal{V} and \mathcal{Z} one has the symmetry

$$D_{\{\Psi_i\}}(\Phi) = D_{\{\mathcal{V} \circ \Psi_i \circ \mathcal{Z}\}}(\mathcal{V} \circ \Phi \circ \mathcal{Z}), \quad (10)$$

where \circ denotes the composition of maps. Clearly, if the set itself $\{\Psi_i\}$ is invariant, then

$$D_{\{\Psi_i\}}(\mathcal{V} \circ \Phi \circ \mathcal{Z}) = D_{\{\Psi_i\}}(\Phi), \quad (11)$$

and the probabilities of the optimal convex approximation for $\mathcal{V} \circ \Phi \circ \mathcal{Z}$ are just a permutation of those for Φ . This is the case, for example, when the available channels are unitary maps corresponding to a (projective) representation of some elements of a group.

III. DISTANCE OF A UNITARY MAP FROM COVARIANT CHANNELS

Let us consider the case where the set of available channels is given by $\mathcal{C} = \{\mathcal{I}, \frac{1}{d^2-1}(d \text{Tr}[\cdot]I - \mathcal{I})\}$. The convex hull is

clearly given by the channels

$$\mathcal{C}_p(\rho) = (1-p)\rho + \frac{p}{d^2-1}(d \text{Tr}[\rho]I - \rho), \quad (12)$$

with $p \in [0,1]$. Indeed, Eq. (12) describes all covariant channels for $\text{SU}(d)$, namely, the channels \mathcal{E} satisfying

$$U_g^\dagger \mathcal{E}[U_g \rho U_g^\dagger] U_g = \mathcal{E}(\rho) \quad (13)$$

for all ρ and unitary $U_g \in \text{SU}(d)$ [32]. Note also that Eq. (12) includes all depolarizing channels (for $p \in [0, (d^2-1)/d^2]$). For any orthogonal basis of unitaries $\{V_i\}$ containing the identity $V_0 \equiv I$ (and hence with $\text{Tr}[V_i] = 0$ and $\text{Tr}[V_i^\dagger V_j] = d\delta_{ij}$ for $i, j = 1, \dots, d^2-1$) we have also $\mathcal{C} = \{\mathcal{I}, \frac{1}{d^2-1} \sum_{i=1}^{d^2-1} V_i(\cdot)V_i^\dagger\}$. This means that the optimal convex approximation of a channel Φ with respect to covariant channels can be achieved by the convex mixture of the identity map and equally weighted orthogonal rotations.

Let us now study the case of qubits, where the target map Φ is a unitary transformation, which, up to a global phase, can be parameterized as

$$U(\alpha, \beta, \gamma) = \begin{pmatrix} \cos \alpha e^{i\beta} & \sin \alpha e^{i\delta} \\ -\sin \alpha e^{-i\delta} & \cos \alpha e^{-i\beta} \end{pmatrix}, \quad (14)$$

with $\alpha \in [0, \pi/2]$ and $\beta \in [0, 2\pi]$, and $\delta \in [0, 2\pi]$. Denote, as usual, the Pauli matrices as $\sigma_0 = I$, $\sigma_1 = \sigma_x$, $\sigma_2 = \sigma_y$, and $\sigma_3 = \sigma_z$. The convex hull of the set $\{\mathcal{I}, \frac{1}{3} \sum_{i=1}^3 \sigma_i(\cdot)\sigma_i\}$ provides all covariant channels as in Eq. (12), for $d = 2$. Of course, the convex approximation of the map $\mathcal{U}(\alpha, \beta, \delta) = U(\alpha, \beta, \gamma)(\cdot)U^\dagger(\alpha, \beta, \gamma)$ is rather poor, since the set is highly constrained. However, we can give here a complete analytical solution, and the physical interpretation of the result is crystalline and exemplary for more intricate situations.

Then, for qubits, the *covariance distance* of the unitary map $\mathcal{U}(\alpha, \beta, \delta)$ is given by

$$D_{\mathcal{C}}[\mathcal{U}(\alpha, \beta, \delta)] \equiv \min_{p \in [0,1]} \|\mathcal{U}(\alpha, \beta, \delta) - \mathcal{C}_p\|_{\diamond}. \quad (15)$$

Since the difference of the Choi operators $R_{\mathcal{U}}$ and $R_{\mathcal{C}_p}$ can be diagonalized over orthonormal Bell states [16], one obtains

$$\begin{aligned} & \|\mathcal{U}(\alpha, \beta, \delta) - \mathcal{C}_p\|_{\diamond} \\ &= \frac{1}{2} \|R_{\mathcal{U}} - R_{\mathcal{C}_p}\|_1 \\ &= \frac{2}{3} p + \sqrt{\frac{16}{9} p^2 + \left(1 - \frac{4}{3} p\right) D_I[\mathcal{U}(\alpha, \beta, \delta)]^2}, \end{aligned} \quad (16)$$

where $D_I[\mathcal{U}(\alpha, \beta, \delta)] = \|\mathcal{U}(\alpha, \beta, \delta) - \mathcal{I}\|_{\diamond} = 2(1 - \cos^2 \alpha \cos^2 \beta)^{1/2}$. Equation (16) provides via Eq. (1) the minimum-error discrimination between arbitrary unitary channels and covariant channels. No maximization is present in Eq. (16) because any maximally entangled state can always be used as an input to achieve the optimal discrimination, thus allowing one to choose $\xi = \frac{1}{\sqrt{2}}$ in Eq. (2).

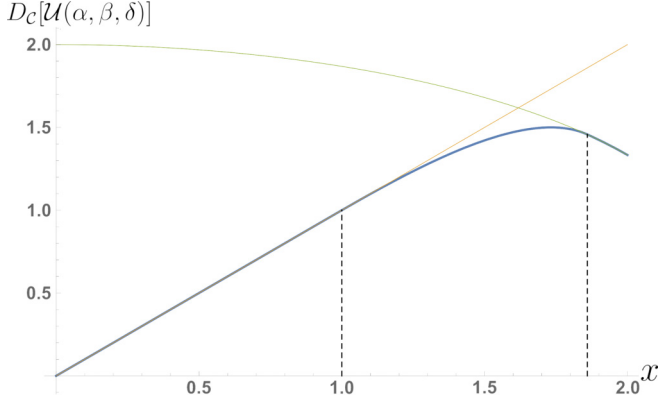


FIG. 1. Optimal convex approximation of a unitary map $\mathcal{U}(\alpha, \beta, \delta)$ for qubits w.r.t. the set including the identity map and the map $\frac{1}{3} \sum_i \sigma_i(\cdot) \sigma_i$. The problem is solved by finding the closest covariant channel to the unitary map in the diamond norm, as in Eq. (15). The covariance distance $D_C[\mathcal{U}(\alpha, \beta, \delta)]$ is a piecewise function of the diamond norm $x \equiv D_I[\mathcal{U}(\alpha, \beta, \delta)] = \|\mathcal{U}(\alpha, \beta, \delta) - \mathcal{I}\|_\diamond$, as given in Eq. (17).

The minimization over p of Eq. (16) gives for the covariance distance the piecewise function

$$D_C[\mathcal{U}(\alpha, \beta, \delta)] = \begin{cases} x, & 0 \leq x \leq 1, \\ \frac{1}{4}[x^2 + \sqrt{3x^2(4-x^2)}], & 1 \leq x \leq \left(\frac{15+\sqrt{33}}{6}\right)^{1/2}, \\ \frac{1}{3}(2 + \sqrt{16-3x^2}), & \left(\frac{15+\sqrt{33}}{6}\right)^{1/2} \leq x \leq 2, \end{cases} \quad (17)$$

where, for brevity, we posed $x \equiv D_I[\mathcal{U}(\alpha, \beta, \delta)]$. The minimum is achieved for $p = 0$, $p = \frac{1}{8}[3x^2 - \sqrt{3x^2(4-x^2)}]$, and $p = 1$ in the three respective pieces. Being just a function of $D_I[\mathcal{U}(\alpha, \beta, \delta)]$, the covariance distance $D_C[\mathcal{U}(\alpha, \beta, \delta)]$ is independent of the parameter δ .

In Fig. 1 the covariance distance $D_C[\mathcal{U}(\alpha, \beta, \delta)]$ for unitary maps is plotted vs the diamond norm $D_I[\mathcal{U}(\alpha, \beta, \delta)]$. It is quite easy to physically interpret the result: as long as the unitary $U(\alpha, \beta, \delta)$ is close enough to the identity (i.e., $x \leq 1$) nothing can be done to approximate its action, whereas, for $U(\alpha, \beta, \delta)$ sufficiently far from the identity, the optimal convex approximation is an equally weighted rotation by the three Pauli matrices. In fact, in these two cases the upper bound in Eq. (7) is saturated with equality. Finally, between these two situations, one has to suitably weight the two previous strategies with probability $p = \frac{1}{8}[3x^2 - \sqrt{3x^2(4-x^2)}]$.

IV. PAULI DISTANCE OF A UNITARY MAP

For a set given by the identity map $\mathcal{V}_0 \equiv \mathcal{I}$ and $(d^2 - 1)$ unitary maps \mathcal{V}_i , with corresponding traceless and orthogonal unitary operators, the optimal convex approximation of a quantum channel Φ corresponds to the closest generalized Pauli channel, namely, to a channel of the form $\sum_{i=0}^{d^2-1} p_i V_i \rho V_i^\dagger$, which provides the minimum *Pauli distance*:

$$D_{\mathcal{P}}(\Phi) \equiv D_{\{\mathcal{V}_i\}}(\Phi). \quad (18)$$

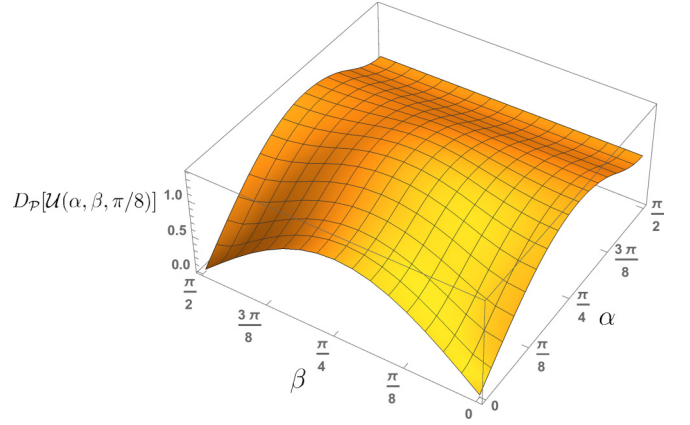


FIG. 2. Optimal convex approximation of a unitary map $\mathcal{U}(\alpha, \beta, \delta)$ w.r.t. the set including the identity map and the three rotations by the three Pauli matrices. The problem is solved by finding the closest Pauli channel to the unitary map in the diamond norm. The solution provides the Pauli distance $D_{\mathcal{P}}[\mathcal{U}(\alpha, \beta, \delta)]$, here plotted vs α and β , for a fixed value of δ , namely, $\delta \equiv \pi/8$.

Let us study in more detail the case of qubit channels and consider Φ as the unitary map $\mathcal{U}(\alpha, \beta, \gamma)$, using again the parametrization of Eq. (14). Exploiting the invariance properties of the set of Pauli matrices, according to Eq. (11), a number of symmetry relations for the Pauli distance can be derived, which can be summarized as follows [33]:

$$\begin{aligned} D_{\mathcal{P}}[\mathcal{U}(\alpha, \beta, \delta)] &= D_{\mathcal{P}}\left[\mathcal{U}\left(\alpha, \frac{\pi}{2} \pm \beta, \delta\right)\right] \\ &= D_{\mathcal{P}}\left[\mathcal{U}\left(\alpha, \beta, \frac{\pi}{2} \pm \delta\right)\right] \\ &= D_{\mathcal{P}}\left[\mathcal{U}\left(\frac{\pi}{2} - \alpha, \delta, \beta\right)\right]. \end{aligned} \quad (19)$$

For specific unitaries $U(\alpha, \beta, \delta)$ we can find exact results for the optimal convex approximation:

(i) for $\beta = \delta = 0$ one has

$$D_{\mathcal{P}}[\mathcal{U}(\alpha, 0, 0)] = |\sin 2\alpha|, \quad (20)$$

with pertaining optimal weights given by $\{p_i^{\text{opt}}\} = \{\cos^2 \alpha, 0, \sin^2 \alpha, 0\}$;

(ii) for $\alpha = 0$ one has

$$D_{\mathcal{P}}[\mathcal{U}(0, \beta, \delta)] = |\sin 2\beta|, \quad (21)$$

with $\{p_i^{\text{opt}}\} = \{\cos^2 \beta, 0, 0, \sin^2 \beta\}$;

(iii) for $\alpha = \pi/2$ one has

$$D_{\mathcal{P}}[\mathcal{U}(\pi/2, \beta, \delta)] = |\sin 2\delta|, \quad (22)$$

with $\{p_i^{\text{opt}}\} = \{0, \sin^2 \delta, \cos^2 \delta, 0\}$.

Note that in all the above specific examples the optimal vector of probabilities has just two nonzero elements. More generally, however, the optimal convex approximation requires three or even all four σ_i operations. For generic values of α, β, δ one can look for a numerical solution. As an example, in Fig. 2 we present the result of the optimal convex approximation of the unitary maps $\mathcal{U}(\alpha, \beta, \pi/8)$. The unitary maps which are worst approximated correspond to $\alpha = \beta = \delta = \frac{\pi}{4}$ [along with those related by the symmetries in Eqs. (19)], and their

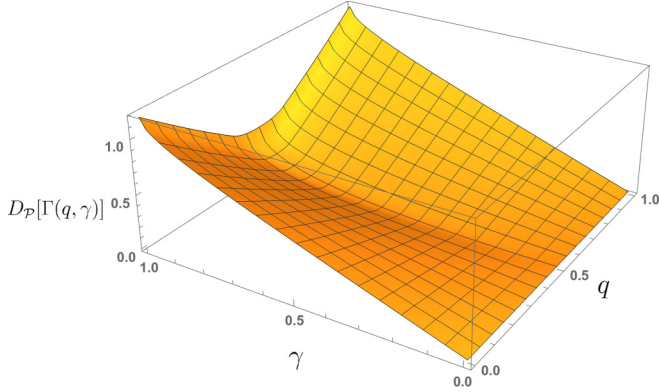


FIG. 3. Optimal convex approximation of a generalized damping channel $\Gamma(q, \gamma)$ w.r.t. the set including the identity map and the three rotations by the three Pauli matrices. The problem is solved by finding the closest Pauli channel to $\Gamma(q, \gamma)$ in the diamond norm. The solution provides the Pauli distance $D_{\mathcal{P}}[\Gamma(q, \gamma)]$, here plotted vs q and γ .

Pauli distance equals $\frac{3}{2}$ and is achieved for equal weights $p_i = \frac{1}{4}$, namely, by the completely depolarizing channel.

V. PAULI DISTANCE OF A GENERALIZED DAMPING CHANNEL

A generalized damping channel $\Gamma(q, \gamma)$ for qubits is described by the completely positive map

$$\Gamma(q, \gamma)[\rho] = q(A_\gamma \rho A_\gamma + C_\gamma \rho C_\gamma^\dagger) + (1 - q)(B_\gamma \rho B_\gamma + C_\gamma^\dagger \rho C_\gamma), \quad (23)$$

where $A_\gamma = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}$, $B_\gamma = \begin{pmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{pmatrix}$, and $C_\gamma = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}$, with $0 \leq q, \gamma \leq 1$. This channel is a mixture of an amplitude damping channel ($q = 1$) and an amplitude amplification channel ($q = 0$), and thus q plays the role of the temperature.

Let us look for the optimal convex approximation of $\Gamma(q, \gamma)$ w.r.t. the set of Pauli matrices. From Eq. (11) we note that the identity $\sigma_x \circ \Gamma(q, \gamma) \circ \sigma_x = \Gamma(1 - q, \gamma)$ provides the symmetry relation $D_{\mathcal{P}}[\Gamma(q, \gamma)] = D_{\mathcal{P}}[\Gamma(1 - q, \gamma)]$.

We have numerically solved the problem, and the results are plotted in Fig. 3, where we show the Pauli distance of $D_{\mathcal{P}}[\Gamma(q, \gamma)]$ vs q and γ . The pertaining weights $\{p_i^{\text{opt}}\}$ of the optimal Pauli approximation are of the form $\{1 - 2p, p, p, 0\}$. On one hand, one has $p_1^{\text{opt}} = p_2^{\text{opt}}$, namely, the rotations by σ_x and σ_y are equally weighted. In fact, this condition guarantees that the convex approximation enjoys the same covariance property of $\Gamma(q, \gamma)$, i.e.,

$$\mathcal{V}(\phi) \circ \Gamma(q, \gamma) = \Gamma(q, \gamma) \circ \mathcal{V}(\phi) \quad \forall \phi, \quad (24)$$

where $\mathcal{V}(\phi) = e^{i\phi\sigma_z}(\cdot)e^{-i\phi\sigma_z}$ denotes the rotation map around σ_z . On the other hand, the additional result $p_3^{\text{opt}} = 0$ stems from the fact that any phase rotation by σ_z would make the resulting Pauli channel more distinguishable from $\Gamma(q, \gamma)$, which instead preserves the phase of the off-diagonal matrix elements of quantum states. We have numerical evidence that the optimal probability satisfies $p_1^{\text{opt}} \geq \frac{\gamma}{4}$, along with the

$$D_{\mathcal{P}}[\Gamma(0.7, \gamma)]$$

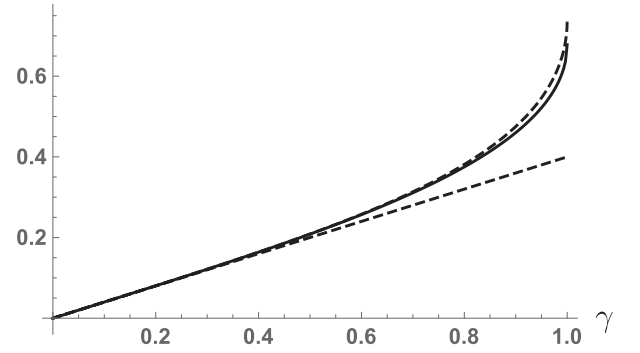


FIG. 4. Optimal convex approximation of a generalized damping channel $\Gamma(0.7, \gamma)$ vs the parameter γ , with respect to the set of Pauli channels. The Pauli distance $D_{\mathcal{P}}[\Gamma(0.7, \gamma)]$ is plotted as the solid line; the upper and lower bounds given by Eq. (25), by the dashed lines.

following bounds for the Pauli distance:

$$\gamma|1 - 2q| \leq D_{\mathcal{P}}[\Gamma(q, \gamma)] \leq \frac{1}{2}[\gamma|1 - 2q| + f(q, \gamma)], \quad (25)$$

with the function $f(q, \gamma)$ given in Eq. (A6) in the Appendix, where in fact these bounds are proved.

In Fig. 4 we plot the Pauli distance for the channel $\Gamma[0.7, \gamma]$ vs the parameter γ , along with the upper and lower bounds. It is apparent that these bounds are tighter for decreasing values of γ , and the upper bound is indeed very good.

VI. CONCLUSIONS

Let us conclude our paper with the following observations. Imagine that we want to approximate N parallel uses of a map Φ acting on an (unknown) N -partite quantum state, and we have at our disposal a set of maps $\{\Psi_i\}$ with which we can act independently on each subsystem. The optimal convex approximation in this case provides the distance $D_{\{\otimes_{j=1}^N \Psi_{i_j}\}}(\Phi^{\otimes N})$. Since obviously the convex hull of $\{\otimes_{j=1}^N \Psi_{i_j}\}$ contains all the N -fold tensor products $\otimes_{j=1}^N (\sum_i p_{i_j} \Psi_{i_j})$, one has

$$\begin{aligned} D_{\{\otimes_{j=1}^N \Psi_{i_j}\}}(\Phi^{\otimes N}) &\leq \min_{\{p_{i_j}\}} \left\| \Phi^{\otimes N} - \otimes_{j=1}^N \left(\sum_i p_{i_j} \Psi_{i_j} \right) \right\|_{\diamond} \\ &\leq \left\| \Phi^{\otimes N} - \left(\sum_i p_i^{\text{opt}} \Psi_i \right)^{\otimes N} \right\|_{\diamond}, \end{aligned} \quad (26)$$

where $\{p_i^{\text{opt}}\}$ denotes the vector of probabilities pertaining to the optimal convex approximation of a single copy of the target map Φ . The interesting fact is that generally both inequalities in Eq. (26) can be strict, and a simple explicit example is provided in the following.

Consider the unitary map \mathcal{U} corresponding to the phase rotation

$$U = \begin{pmatrix} e^{i\frac{\pi}{6}} & 0 \\ 0 & e^{-i\frac{\pi}{6}} \end{pmatrix} \quad (27)$$

and its convex approximation w.r.t. $\{\mathcal{I}, \mathcal{Z}\}$, where $\mathcal{Z} = \sigma_z(\cdot)\sigma_z$. Thus, according to Eq. (21), the optimal convex approximation

is the closest dephasing channel $\Psi(p) = p\mathcal{I} + (1-p)\mathcal{Z}$ to \mathcal{U} , and one has

$$D_{\{\mathcal{I}, \mathcal{Z}\}}(\mathcal{U}) = \frac{\sqrt{3}}{2}, \quad (28)$$

with corresponding optimal weight $p^{\text{opt}} = \frac{3}{4}$. The diamond norm for the two-fold tensor product can be evaluated as

$$\|\mathcal{U}^{\otimes 2} - \Psi(3/4)^{\otimes 2}\|_{\diamond} \simeq 1.314. \quad (29)$$

A tiny improvement is found by looking for the closest tensor product of dephasing channels as

$$\min_{q_1, q_2 \in [0,1]} \|\mathcal{U}^{\otimes 2} - \Psi(q_1) \otimes \Psi(q_2)\|_{\diamond} \simeq 1.312, \quad (30)$$

with optimal weights $q_1^{\text{opt}} = q_2^{\text{opt}} \simeq 0.77$. By allowing correlations, a larger improvement is obtained by the optimal convex approximation, and one has

$$\begin{aligned} & D_{\{\mathcal{I} \otimes \mathcal{I}, \mathcal{Z} \otimes \mathcal{I}, \mathcal{I} \otimes \mathcal{Z}, \mathcal{Z} \otimes \mathcal{Z}\}}(\mathcal{U}^{\otimes 2}) \\ &= \min_{\{p_{ij}\}} \|\mathcal{U}^{\otimes 2} - (p_{00}\mathcal{I} \otimes \mathcal{I} + p_{10}\mathcal{Z} \otimes \mathcal{I} \\ &\quad + p_{01}\mathcal{I} \otimes \mathcal{Z} + p_{11}\mathcal{Z} \otimes \mathcal{Z})\|_{\diamond} \\ &\simeq 1.281, \end{aligned} \quad (31)$$

where the optimal weights are given by $p_{00}^{\text{opt}} \simeq 0.60$, $p_{10}^{\text{opt}} = p_{01}^{\text{opt}} \simeq 0.20$, and $p_{11}^{\text{opt}} = 0$. The resulting optimal convex approximation is obviously a channel with correlated uses.

The first inequality in Eq. (26) comes from the fact that the introduction of correlations in the approximating map can be beneficial even if the target map is indeed the product of independent maps (as happens, for example, in the optimal cloning of quantum states [34]). The second inequality is due to the fact that the distance $\|\Phi_0 - \Phi_1\|_{\diamond}$ quantifying the distinguishability of two channels is not additive or multiplicative when considering multiple copies, namely, we clearly only know that $\|\Phi_0 - \Phi_1\|_{\diamond} \leq \|\Phi_0^{\otimes N} - \Phi_1^{\otimes N}\|_{\diamond}$. This also implies that we do not have a direct expression for the scaling with N of the distance between a quantum channel and its convex approximations. The results related to the quantum Chernoff bound for quantum states [35,36], suitably generalized to the case of quantum channels might be useful

for a systematic study of the scaling of the optimal convex approximations with the number of uses.

APPENDIX: PROOF OF THE BOUNDS IN EQ. (25)

Let us consider the use of input state $|0\rangle$ or $|1\rangle$ for discriminating a generalized amplitude damping channel $\Gamma(q, \gamma)$ from a Pauli channel $\mathcal{P} = \sum_{i=0}^3 p_i \sigma_i(\cdot) \sigma_i$. Then the diamond norm is bounded as

$$\|\Gamma(q, \gamma) - \mathcal{P}\|_{\diamond} \geq \|(\Gamma(q, \gamma) - \mathcal{P})|u\rangle\langle u|\|_1 \quad (A1)$$

for $u = 0, 1$. A straightforward calculation gives

$$\|(\Gamma(q, \gamma) - \mathcal{P})|0\rangle\langle 0|\|_1 = 2|\gamma(1-q) - (p_1 + p_2)|, \quad (A2)$$

$$\|(\Gamma(q, \gamma) - \mathcal{P})|1\rangle\langle 1|\|_1 = 2|\gamma q - (p_1 + p_2)|. \quad (A3)$$

Then the Pauli distance $D_{\mathcal{P}}[\Gamma(q, \gamma)]$ of the damping channel can be bounded as

$$\begin{aligned} & D_{\mathcal{P}}[\Gamma(q, \gamma)] \\ &\equiv \min_{\{p_i\}} \|\Gamma(q, \gamma) - \mathcal{P}\|_{\diamond} \\ &\geq \min_{\{p_i\}} \max\{2|\gamma(1-q) - (p_1 + p_2)|, 2|\gamma q - (p_1 + p_2)|\} \\ &= \gamma|1 - 2q|, \end{aligned}$$

where the minimum is achieved for $p_1 = p_2 = \frac{\gamma}{4}$. This proves the lower bound in Eq. (25).

The upper bound can be simply obtained by choosing the Pauli channel \mathcal{P}_{γ} with $p_0 = 1 - \frac{\gamma}{2}$, $p_1 = p_2 = \frac{\gamma}{4}$, and $p_3 = 0$. Then, obviously,

$$D_{\mathcal{P}}[\Gamma(q, \gamma)] \leq \|\Gamma(q, \gamma) - \mathcal{P}_{\gamma}\|_{\diamond}. \quad (A4)$$

The diamond norm in Eq. (A4) can be explicitly evaluated as

$$\|\Gamma(q, \gamma) - \mathcal{P}_{\gamma}\|_{\diamond} = \frac{1}{2}[\gamma|1 - 2q| + f(q, \gamma)], \quad (A5)$$

with

$$\begin{aligned} & f(q, \gamma) \\ &= \{8(1-\gamma) - 4(2-q)\sqrt{1-\gamma} + \gamma^2[2 - 4q(1-q)]\}^{1/2}. \end{aligned} \quad (A6)$$

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
 [2] K. Kraus, *States, Effects and Operations*, Lecture Notes in Physics, Vol. 190 (Springer, Berlin, 1983).
 [3] F. Buscemi, G. M. D'Ariano, and M. F. Sacchi, *Phys. Rev. A* **68**, 042113 (2003).
 [4] M. A. Nielsen and I. L. Chuang, *Phys. Rev. Lett.* **79**, 321 (1997).
 [5] G. Vidal, L. Masanes, and J. I. Cirac, *Phys. Rev. Lett.* **88**, 047905 (2002).
 [6] M. Hillery, M. Ziman, and V. Bužek, *Phys. Rev. A* **73**, 022345 (2006).

[7] X.-C. Yao, J. Fiurášek, H. Lu, W.-B. Gao, Y.-A. Chen, Z.-B. Chen, and J.-W. Pan, *Phys. Rev. Lett.* **105**, 120402 (2010).
 [8] J. Fiurášek, M. Dušek, and R. Filip, *Phys. Rev. Lett.* **89**, 190401 (2002).
 [9] G. M. D'Ariano and P. Perinotti, *Phys. Rev. Lett.* **94**, 090401 (2005).
 [10] M. Dall'Arno, G. M. D'Ariano, and M. F. Sacchi, *Phys. Rev. A* **82**, 042315 (2010).
 [11] N. K. Bernardes, Á. Cuevas, A. Orioux, C. H. Monken, P. Mataloni, F. Sciarrino, and M. F. Santos, *Sci. Rep.* **5**, 17520 (2015).
 [12] Á. Cuevas, M. Proietti, M. A. Ciampini, S. Duranti, P. Mataloni, M. F. Sacchi, and C. Macchiavello, *Phys. Rev. Lett.* **119**, 100502 (2017).

[13] N. Yu, R. Duan, and Q. Xu, *IEEE Trans. Info. Theor.* **63**, 1299 (2017).

[14] B. Kümmerer and H. Maassen, *Commun. Math. Phys.* **109**, 1 (1987).

[15] U. Haagerup and M. Musat, *Commun. Math. Phys.* **303**, 5051 (2011).

[16] M. F. Sacchi, *Phys. Rev. A* **71**, 062340 (2005).

[17] V. I. Paulsen, *Completely Bounded Maps and Dilations* (Longman Scientific and Technical, New York, 1986).

[18] D. Aharonov, A. Kitaev, and N. Nisan, in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing, Dallas, TX, USA, May 24–26, 1998* (Association for Computing Machinery, New York, 1998), p. 20.

[19] M.-D. Choi, *Linear Alg. Appl.* **10**, 285 (1975).

[20] See, for example, R. Bhatia, *Matrix Analysis*, Springer Graduate Texts in Mathematics, Vol. 169 (Springer, New York, 1996).

[21] From the unitarily invariance of the trace norm and the polar decomposition of ξ , the matrix ξ can be taken as a positive matrix P . Moreover, for numerical maximization, positivity can be constrained by means of the Choleski decomposition, namely, $P = T^\dagger T$, with T the upper diagonal with positive diagonal entries. For qubits, a useful parametrization is also

$$P = \begin{pmatrix} \cos \theta_1 \cos \theta_2 & \frac{1}{\sqrt{2}} \sin \theta_2 e^{i\phi} \\ \frac{1}{\sqrt{2}} \sin \theta_2 e^{-i\phi} & \sin \theta_1 \cos \theta_2 \end{pmatrix}, \quad (\text{A7})$$

with $\theta_1 \in [0, \pi/2]$, $\theta_2 \in [0, \arctan \sqrt{(\sin 2\theta_1)}]$, and $\phi \in [0, 2\pi]$. Finally, the formulation of the trace norm as a semidefinite program [22] allows its efficient computation.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004).

[23] A. Acín, *Phys. Rev. Lett.* **87**, 177901 (2001).

[24] G. M. D’Ariano, P. L. Presti, and M. G. A. Paris, *Phys. Rev. Lett.* **87**, 270404 (2001).

[25] A. Gilchrist, N. K. Langford, and M. A. Nielsen, *Phys. Rev. A* **71**, 062310 (2005).

[26] V. P. Belavkin, G. M. D’Ariano, and M. Raginsky, *J. Math. Phys.* **46**, 062106 (2005).

[27] J. Watrous, *Theory Comput.* **5**, 117 (2009).

[28] A. Ben-Aroya and A. Ta-Shma, *Quantum Info. Comput.* **10**, 77 (2010).

[29] J. Watrous, *Chicago J. Theor. Comp. Sci.* **2013**(11), 8 (2013).

[30] M. C. Grant and S. P. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*; <http://cvxr.com/cvx> (2017).

[31] M. C. Grant and S. P. Boyd, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. D. Blondel, S. P. Boyd, and H. Kimura (Springer-Verlag, Berlin, 2008), p. 95.

[32] This can be easily proven since, in terms of the Choi operator $\mathcal{R}_\mathcal{E}$, Eq. (13) is equivalent to $[U_g \otimes U_g^*, \mathcal{R}_\mathcal{E}] = 0$. The application of Shur’s lemma for unitary irreducible representations forces channel \mathcal{E} to the form as in Eq. (12).

[33] The symmetry relations are easily obtained by the fact that the set of Pauli matrix is a closed algebra, along with the covariance property $\{\sigma_i\} = \{S\sigma_i S\} = \{S\sigma_i S^\dagger\}$ up to global phases, with $S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$.

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

[35] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. I. Masanes, A. Acin, and F. Verstraete, *Phys. Rev. Lett.* **98**, 160501 (2007).

[36] M. Nussbaum and A. Szkoła, *Ann. Stat.* **37**, 1040 (2009).