Computing the distance between quantum channels: Usefulness of the Fano representation

Giuliano Benenti[1, 2, *] and Giuliano Strini[3, †]

[1]*CNISM, CNR-INFM, and Center for Nonlinear and Complex Systems, Università degli Studi dell'Insubria, via Valleggio 11, 22100 Como, Italy*
[2]*Istituto Nazionale di Fisica Nucleare, Sezione di Milano, via Celoria 16, 20133 Milano, Italy*
[3]*Dipartimento di Fisica, Università degli Studi di Milano, via Celoria 16, 20133 Milano, Italy*

(Dated: June 24, 2010)

The diamond norm measures the distance between two quantum channels. From an operational vewpoint, this norm measures how well we can distinguish between two channels by applying them to input states of arbitrarily large dimensions. In this paper, we show that the diamond norm can be conveniently and in a physically transparent way computed by means of a Monte-Carlo algorithm based on the Fano representation of quantum states and quantum operations. The effectiveness of this algorithm is illustrated for several single-qubit quantum channels.

PACS numbers: 03.65.Yz, 03.67.-a

I. INTRODUCTION

Quantum information processes in a noisy environment are conveniently described in terms of *quantum channels*, that is, linear, trace preserving, completely positive maps on the set of quantum states [1, 2]. The problem of discriminating quantum channels is of great interest. For instance, knowing the correct noise model might provide useful information to devise efficient error-correcting strategies, both in the fields of quantum communication and quantum computation.

It is therefore natural to consider *distances* between quantum channels, that is to say, we would like to quantify how similarly two channels $\mathcal{E}_1$ and $\mathcal{E}_2$ act on quantum states, or in other words to determine if there are input states ρ on which the two channels produce output states $\rho'_1 = \mathcal{E}_1(\rho)$ and $\rho'_2 = \mathcal{E}_2(\rho)$ that are distinguishable. The trace norm of $\rho'_1 - \rho'_2$ represents how well ρ'_1 and ρ'_2 can be distinguished by a measurement [3]: the more orthogonal two quantum states are, the easier it is to discriminate them. The trace distance of two quantum channels is then obtained after maximizing the trace norm of $\rho'_1 - \rho'_2$ over the input state ρ. However, the trace norm is not a good measure of the distance between quantum channels. Indeed, in general the presence in the input state of entanglement with an ancillary system can help discriminating quantum channels [4–8]. This fact is captured by the *diamond norm* [9, 10]: the trace distance between the overall output states (including the ancillary system) is optimized over all possible input states, including those entangled.

The computation of the diamond norm is not known to be straightforward and only a limited number of algorithms have been proposed [11–13], based on complicated semidefinite programming or convex optimization. On the other hand, analytical solutions are limited to special classes of channels [7, 8]. In this paper, we propose a simple and easily parallelizable Monte-Carlo algorithm based on the Fano representation of quantum states and quantum operations. We show that our algorithm provides reliable results for the case, most significant for present-day implementations of quantum information processing, of single-qubit quantum channels. Furthermore, in the Fano representation quantum operations are described by affine maps whose matrix elements have precise physical meaning: They are directly related to the evolution of the expectation values of the system's polarization measurements [1, 14, 15].

The paper is organized as follows. After reviewing in Sec. II basic definitions of the distance between quantum channels, we discuss in Sec. III two numerical Monte-Carlo strategies for computing the diamond norm. The first one is based on the Kitaev's characterization of the diamond norm. The second one is based on the Fano representation of quantum states and quantum operations. The two methods are compared in Sec. IV for a few physically significant single-qubit quantum channels. Finally, our conclusions are drawn in Sec. V.

II. THE DIAMOND NORM

A. Basic definitions

We consider the following problem: given two quantum channels $\mathcal{E}_1$ and $\mathcal{E}_2$, and a single channel use, chosen uniformly at random from $\{\mathcal{E}_1, \mathcal{E}_2\}$, we wish to maximize the probability of correctly identifying the quantum channel. It seems natural to reformulate the optimization problem into the problem of finding an input state (density matrix) ρ in the Hilbert space $\mathcal{H}$ such that the error probability in the discrimination of the output states $\mathcal{E}_1(\rho)$ and $\mathcal{E}_2(\rho)$ is minimal. In this case, the minimal error

*Electronic address: giuliano.benenti@uninsubria.it
†Electronic address: giuliano.strini@mi.infn.it

probability reads

$$p'_E = \frac{1}{2} - \frac{||\mathcal{E}_1 - \mathcal{E}_2||_1}{4}, \qquad ||\mathcal{E}||_1 \equiv \max_\rho ||\mathcal{E}(\rho)||_1, \tag{1}$$

where $||X||_1 \equiv \mathrm{Tr}\sqrt{X^\dagger X}$ denotes the trace norm.

The superoperator trace distance $||\mathcal{E}_1 - \mathcal{E}_2||_1$ is, however, not a good definition of the distance between two quantum operations. The point is that in general it is possible to exploit quantum entanglement to increase the distinguishability of two quantum channels. In this case, an ancillary Hilbert space $\mathcal{K}$ is introduced, the input state ξ is a density matrix in $\mathcal{K} \otimes \mathcal{H}$, and the quantum operations are trivially extended to $\mathcal{K}$. That is to say, the output states to discriminate are $(I_\mathcal{K} \otimes \mathcal{E}_1)\xi$ and $(I_\mathcal{K} \otimes \mathcal{E}_2)\xi$, where $\mathcal{I}_\mathcal{K}$ is the identity map acting on $\mathcal{K}$. The minimal error probability reads

$$p_E = \frac{1}{2} - \frac{||\mathcal{E}_1 - \mathcal{E}_2||_\diamond}{4}, \qquad ||\mathcal{E}||_\diamond \equiv \max_\xi ||(\mathcal{I}_\mathcal{K} \otimes \mathcal{E})\xi||_1, \tag{2}$$

where $||\mathcal{E}||_\diamond$ denotes the diamond norm of $\mathcal{E}$. It is clear from definition (2) that

$$||\mathcal{E}||_\diamond = ||\mathcal{I}_\mathcal{K} \otimes \mathcal{E}||_1 \geq ||\mathcal{E}||_1 \tag{3}$$

and therefore $p_E \leq p'_E$, so that it can be convenient to use an ancillary system to better discriminate two quantum operations after a single channel use. The two quantum channels $\mathcal{E}_1$ and $\mathcal{E}_2$ become perfectly distinguishable ($p_E = 0$) when their diamond distance $||\mathcal{E}_1 - \mathcal{E}_2||_\diamond = 2$.

It turns out that the diamond norm does not depend on $\mathcal{K}$, provided $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$ [9]. Due to the convexity of the trace norm, it can be shown that the maximum in both Eqs. (1) and (2) is achieved for pure input states [16].

B. Kitaev's characterization of the diamond norm

Kitaev provided a different equivalent characterization of the diamond norm, see, e.g., [9, 10]. Any superoperator (not necessarily completely positive) $\mathcal{E} : L(\mathcal{H}) \to L(\mathcal{H})$, with $L(\mathcal{H})$ space of linear operators mapping $\mathcal{H}$ to itself, can be expressed as

$$\mathcal{E}(X) = \mathrm{Tr}_\mathcal{R}(AXB^\dagger), \tag{4}$$

where $X \in L(\mathcal{H})$, A and B linear operators from $\mathcal{H}$ to $\mathcal{R} \otimes \mathcal{H}$, with $\mathcal{R}$ auxiliary Hilbert space and $\dim(\mathcal{R}) \leq [\dim(\mathcal{H})]$. It is then possible to define completely positive superoperators $\Psi_A, \Psi_B : L(\mathcal{H}) \to L(\mathcal{R})$:

$$\Psi_A(X) = \mathrm{Tr}_\mathcal{H}(AXA^\dagger), \quad \Psi_B(X) = \mathrm{Tr}_\mathcal{H}(BXB^\dagger). \tag{5}$$

Note that the space $\mathcal{H}$ is traced out in the definition of Ψ_A, Ψ_B, rather than the space $\mathcal{R}$. Finally, it turns out that [9, 10]

$$||\mathcal{E}||_\diamond = F_{\max}(\Psi_A, \Psi_B), \tag{6}$$

where $F_{\max}(\Psi_A, \Psi_B)$ is the maximum output fidelity of Ψ_A and Ψ_B, defined as

$$F_{\max}(\Psi_A, \Psi_B) = \max_{\rho_1, \rho_2} F[\Psi_A(\rho_1), \Psi_B(\rho_2)], \tag{7}$$

where ρ_1, ρ_2 are density matrices in $\mathcal{H}$, and the fidelity F is defined as

$$F(\Psi_A, \Psi_B) = \mathrm{Tr}\sqrt{\Psi_A^{1/2}\Psi_B\Psi_A^{1/2}}. \tag{8}$$

Note that Ψ_A, Ψ_B are not density matrices: the conditions $\mathrm{Tr}(\Psi_A) = 1$, $\mathrm{Tr}(\Psi_B) = 1$ are not satisfied.

III. COMPUTING THE DIAMOND NORM

We numerically compute the distance (induced by the diamond norm) between two quantum channels $\mathcal{E}_1$ and $\mathcal{E}_2$ using two Monte-Carlo algorithms. The first one is based on the direct computation of $||\mathcal{E}_1 - \mathcal{E}_2||_\diamond$, with the output states $(\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_1)\xi$ and $(\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_2)\xi$ in Eq. (2) computed from ξ taking advantage of the Fano representation of quantum states and quantum operations. The second Monte-Carlo algorithm uses the Kitaev's representation of the diamond norm to compute the maximum output fidelity $F_{\max}$ of Eq. (7). In the following, we will refer to the two above Monte-Carlo algorithms as F-algorithm and K-algorithm, respectively. For the sake of simplicity we will confine ourselves to one-qubit quantum channels, even though the two algorithms can be easily formulated for two- or many-qubit channels.

A. The F-algorithm

In this section we describe the F-algorithm, which we will use to directly compute the diamond norm (2), with the maximum taken over a large number of randomly chosen input states ξ. A convexity argument shows that it is sufficient to optimize over pure input states $\xi = |\Psi\rangle\langle\Psi|$ [16]. For one-qubit channels, it is enough to add a single ancillary qubit when computing the diamond norm [9]. Therefore, we can write

$$|\Psi\rangle = C_{00}|00\rangle + C_{01}|01\rangle + C_{10}|10\rangle + C_{11}|11\rangle, \tag{9}$$

with

$$\begin{aligned} C_{00} &= \cos\theta_1 \cos\theta_2, \\ C_{01} &= \cos\theta_1 \sin\theta_2 e^{i\phi_1}, \\ C_{10} &= \sin\theta_1 \cos\theta_3 e^{i\phi_2}, \\ C_{11} &= \sin\theta_1 \sin\theta_3 e^{i\phi_3}, \end{aligned} \tag{10}$$

where the angles $\theta_i \in \left[0, \frac{\pi}{2}\right]$ and the phase $\phi_i \in [0, 2\pi]$. Hence, the maximization in the diamond norm is over the 6 real parameters $\theta_1, \theta_2, \theta_3$, and ϕ_1, ϕ_2, ϕ_3. Of course, the number of parameters can be reduced for specific channels when there are symmetries.

Let $\mathcal{E}_1$ and $\mathcal{E}_2$ denote the two single-qubit superoperators we would like to distinguish. The output states $\xi_1' \equiv (\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_1)\xi$ and $\xi_2' \equiv (\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_2)\xi$ can be conveniently computed using the Fano representation. Any two-qubit state can be written in the Fano form as follows [17–19]:

$$\xi = \frac{1}{4} \sum_{\alpha,\beta=x,y,z,I} R_{\alpha\beta} \sigma_\alpha \otimes \sigma_\beta, \tag{11}$$

where σ_x, σ_y, and σ_z are the Pauli matrices, $\sigma_I \equiv \mathbb{1}$, and

$$R_{\alpha\beta} = \mathrm{Tr}[(\sigma_\alpha \otimes \sigma_\beta)\xi]. \tag{12}$$

Note that the normalization condition $\mathrm{Tr}(\xi) = 1$ implies $R_{II} = 1$. Moreover, the coefficients $R_{\alpha\beta}$ are real due to the hermicity of ξ. Eqs. (11) and (12) allow us to go from the standard representation for the density matrix (in the $\{|0\rangle \equiv |00\rangle, |1\rangle \equiv |01\rangle, |2\rangle \equiv |10\rangle, |3\rangle \equiv |11\rangle\}$ basis) to the Fano representation, and vice versa. It is convenient to write the coefficients $R_{\alpha\beta}$ as a column vector,

$$\mathbf{R} = [R_{xx}, R_{xy}, R_{xz}, R_{xI}, R_{yx}, R_{yy}, R_{yz}, R_{yI}, R_{zx}, R_{zy}, R_{zz}, R_{zI}, R_{Ix}, R_{Iy}, R_{Iz}, R_{II}]^T. \tag{13}$$

Then the quantum operations $\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_1$ and $\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_2$ map, in the Fano representation, $\mathbf{R}$ into $\mathbf{R_1'} = \mathcal{M}_1^{(2)}\mathbf{R}$ and $\mathbf{R_2'} = \mathcal{M}_2^{(2)}\mathbf{R}$, respectively, where $\mathcal{M}_1^{(2)}$ and $\mathcal{M}_2^{(2)}$ are affine transformation matrices. Such matrices have a simple block structure:

$$\mathcal{M}_i^{(2)} = \mathcal{I}^{(1)} \otimes \mathcal{M}_i^{(1)}, \tag{14}$$

with $\mathcal{I}^{(1)}$ and $\mathcal{M}_i^{(1)}$ 4×4 affine transformation matrices corresponding to the quantum operations $\mathcal{I}_\mathcal{K}$ and $\mathcal{E}_i$ (of course, $\mathcal{I}^{(1)}$ is the identity matrix). Matrices $\mathcal{M}_i^{(1)}$ directly determines the transformation, induced by $\mathcal{E}_i$, of the single-qubit Bloch-sphere coordinates (x, y, x). Given the Fano representation for a single qubit, $\rho = \frac{1}{2}\sum_\alpha r_\alpha \sigma_\alpha$, then $\mathbf{r} = [x, y, z, 1]^T$ and $\mathbf{r_i'} \equiv \mathcal{M}_i^{(1)}\mathbf{r}$. We point out that, while one could compute ξ_i' from the Kraus representation of the superoperator $\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_i$, the advantage of the Fano representation is that the matrix elements of $\mathcal{M}_i$ are directly related to the transformation of the expectation values of the system's polarization measurements [1, 14, 15].

Finally, we compute the trace distance between ξ_1' and ξ_2' as

$$||\xi_1' - \xi_2'||_1 = \sum_k |\lambda_k|, \tag{15}$$

where $\lambda_1, ..., \lambda_4$ are the eigenvalues of $\xi_1' - \xi_2'$.

B. The K-algorithm

We consider a special and unnormalized state in the extended Hilbert space $\mathcal{K} \otimes \mathcal{H}$:

$$|\alpha\rangle = \sum_j |j_\mathcal{K}\rangle|j_\mathcal{H}\rangle, \tag{16}$$

where $\dim(\mathcal{K}) = \dim(\mathcal{H})$ and $\{|j_\mathcal{K}\rangle\}$, $\{|j_\mathcal{H}\rangle\}$ are orthonormal bases for $\mathcal{K}$, $\mathcal{H}$. The state $|\alpha\rangle$ is, up to a normalization factor, a maximally entangled state in $\mathcal{K} \otimes \mathcal{H}$.

We define an operator σ on $\mathcal{K} \otimes \mathcal{H}$:

$$\sigma = (\mathcal{I}_\mathcal{K} \otimes \mathcal{E})(|\alpha\rangle\langle\alpha|), \tag{17}$$

where $\mathcal{E} = \mathcal{E}_1 - \mathcal{E}_2$ is the difference of two quantum operations but is not a quantum operation itself. That is, $\mathcal{E}$ is linear but not trace preserving and completely positive.

Using the singular value decomposition of σ, we can write

$$\sigma = \sum_{i=1}^{M} |u^{(i)}\rangle\langle v^{(i)}|, \tag{18}$$

where M is the rank of σ and $|u^{(i)}\rangle$, $|v^{(i)}\rangle$ are vectors in $\mathcal{K} \otimes \mathcal{H}$. The operator σ completely specifies $\mathcal{E}$ and can be exploited to express $\mathcal{E}$ as

$$\mathcal{E}(X) = \sum_{i=1}^{M} A^{(i)} X B^{(i)\dagger}. \tag{19}$$

The operators $A^{(i)}$ and $B^{(i)}$ can be derived by generalizing the construction of the operator-sum representation for quantum operations (see, for instance, Sec. 8.2.4 in Ref. [2]). Given a generic state $|\psi\rangle = \sum_j \psi_j |j_\mathcal{H}\rangle$ in $\mathcal{H}$, we define a corresponding state in $\mathcal{K}$: $|\tilde{\psi}\rangle = \sum_j \psi_j^\star |j_\mathcal{K}\rangle$. Next, we define

$$A^{(i)} = \langle\tilde{\psi}|u^{(i)}\rangle. \tag{20}$$

For instance, in the single-qubit case $|u^{(i)}\rangle = [u_1^{(i)}, u_2^{(i)}, u_3^{(i)}, u_4^{(i)}]^T$ and we obtain

$$A^{(i)} = \begin{bmatrix} u_1^{(i)} & u_3^{(i)} \\ u_2^{(i)} & u_4^{(i)} \end{bmatrix}. \tag{21}$$

Similarly, we define

$$B^{(i)} = \langle\tilde{\psi}|v^{(i)}\rangle. \tag{22}$$

Finally, it can be checked that with the above defined operators $A^{(i)}$, $B^{(i)}$ we can express $\mathcal{E}$ by means of Eq. (19).

We can now give explicit expressions for Ψ_A and Ψ_B in Eq. (5):

$$[\Psi_A(X)]_{ij} = \sum_{\alpha,m,n} A_{\alpha m}^{(i)} X_{mn} A_{\alpha n}^{(j)\star}, \tag{23}$$

$$[\Psi_B(X)]_{ij} = \sum_{\alpha,m,n} B^{(i)}_{\alpha m} X_{mn} B^{(j)\star}_{\alpha n}, \tag{24}$$

with $1 \leq i, j \leq M$. Therefore, Ψ_A and Ψ_B are $M \times M$ matrices. In the single-qubit case, $M \leq 4$ and to compute the diamond norm through Eq. (7) we need to calculate eigenvalues and eigenvectors of matrices of size M. A simpler but less efficient decomposition of $\mathcal{E}$, Ψ_A, Ψ_B is provided in Appendix A.

For single-qubit channels the optimization (7) is over 6 real parameters, 3 for ρ_1 and 3 for ρ_2 (for instance, the Bloch-sphere coordinates of ρ_1 and ρ_2). As discussed in Sec. III B, the same number of parameters are needed in the F-method. However, the F-method has computational advantages in that only the eigenvalues of the 4×4 matrix $\xi'_1 - \xi'_2$ are required, while in the K-method we need to evaluate both eigenvalues and eigenvectors of matrices in general of the same size ($M = 4$). Moreover, the singular-values decomposition of matrix σ is needed. Besides computational advantages, the F-method is physically more transparent, as it is based on affine maps, whose matrix elements have physical meaning, being directly related to the transformation of the expectation values of the system's polarization measurements [1, 14, 15].

IV. EXAMPLES FOR SINGLE-QUBIT QUANTUM CHANNELS

In this section, we illustrate the working of the F- and K-methods for the case, most significant for present-day implementations, of single-qubit quantum channels.

A. Pauli channels

We start by considering the case of Pauli channels,

$$\mathcal{E}_i(\rho) = \sum_{\alpha=x,y,z,I} (q_\alpha)_i \sigma_\alpha \rho \sigma_\alpha, \quad \sum_\alpha (q_\alpha)_i = 1, \quad i = 1, 2, \tag{25}$$

for which the diamond norm can be evaluated analytically [7]:

$$||\mathcal{E}||_\diamond = ||\mathcal{E}_1 - \mathcal{E}_2||_\diamond = \sum_\alpha |(q_\alpha)_1 - (q_\alpha)_2|, \tag{26}$$

this value of the diamond norm being achieved for maximally entangled input states. The Pauli-channel case will serve as a testing ground for the F- and K-algorithms and help us develop a physical and geometrical intuition.

Let us focus on a couple of significant examples. We first consider the bit-flip and the phase-flip channels:

$$\begin{aligned} \mathcal{E}_1(\rho) &= \frac{1+c_1}{2}\rho + \frac{1-c_1}{2}\sigma_x \rho \sigma_x, \\ \mathcal{E}_2(\rho) &= \frac{1+c_2}{2}\rho + \frac{1-c_2}{2}\sigma_z \rho \sigma_z, \end{aligned} \tag{27}$$

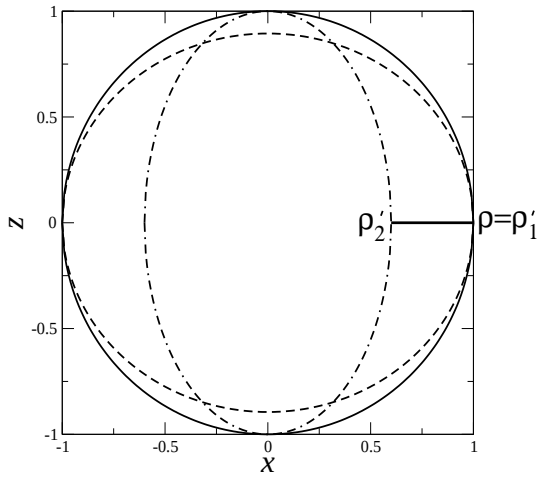

FIG. 1: Schematic drawing of the trace (and diamond) distance distance between the bit-flip and the phase-flip channels.

with $0 \leq c_1, c_2 \leq 1$. We then readily obtain from Eq. (26) that

$$||\mathcal{E}||_\diamond = ||\mathcal{E}_1 - \mathcal{E}_2||_\diamond = \max\{1 - c_1, 1 - c_2\}. \tag{28}$$

For this example, the diamond norm coincides with the trace norm, $||\mathcal{E}||_\diamond = ||\mathcal{E}||_1$, and therefore has a simple geometrical interpretation: For a single qubit the trace distance between two single-qubit states is equal to the Euclidean distance between them on the Bloch ball [2]. Superoperators $\mathcal{E}_1$ and $\mathcal{E}_2$ map the Bloch sphere into an ellipsoid with x (for the bit-flip channel) and z (for the phase-flip channel) as symmetry axis. If we call (x, y, z), (x'_1, y'_1, z'_1), and (x'_2, y'_2, z'_2), the initial Bloch-sphere coordinates and the new coordinates after application of quantum operations $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively, we obtain

$$\begin{aligned} &x'_1 = x, \;\; y'_1 = c_1 y, \;\; z'_1 = c_1 z, \\ &x'_2 = c_2 x, \;\; y'_2 = c_2 y, \;\; z'_2 = z. \end{aligned} \tag{29}$$

The geometrical meaning of the trace norm for the present example is clear from Fig. 1: the length of the line segment $\overline{\rho\rho'_2} = \overline{\rho'_1\rho'_2}$ is the trace (and the diamond) distance $||\mathcal{E}_1 - \mathcal{E}_2||_1 = ||\mathcal{E}_1 - \mathcal{E}_2||_\diamond$ (note that in this figure $1 - c_2 > 1 - c_1$).

As a further example, we discuss a special instance of the channels considered in Ref. [7]:

$$\begin{aligned} \mathcal{E}_1(\rho) &= \frac{1}{2}\rho + \frac{1}{4}\sigma_x \rho \sigma_x + \frac{1}{4}\sigma_y \rho \sigma_y, \\ \mathcal{E}_2(\rho) &= \sigma_z \rho \sigma_z. \end{aligned} \tag{30}$$

In this case,

$$\begin{aligned} &x'_1 = \frac{1}{2}x, \;\; y'_1 = \frac{1}{2}y, \;\; z'_1 = 0, \\ &x'_2 = -x, \;\; y'_2 = -y, \;\; z'_2 = z. \end{aligned} \tag{31}$$

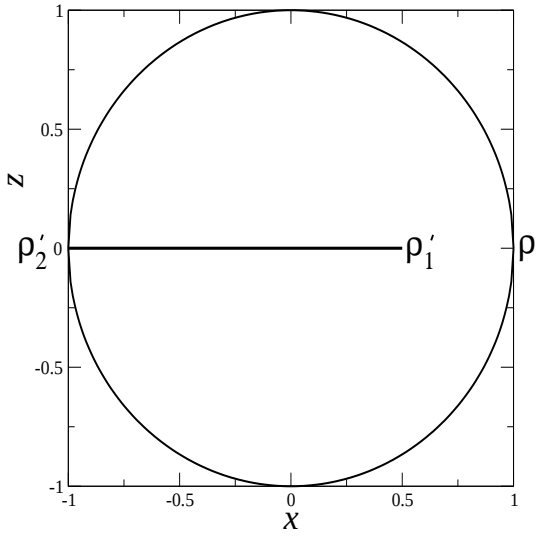

FIG. 2: Schematic drawing of the trace distance distance between the two Pauli channels of Eq. (30).

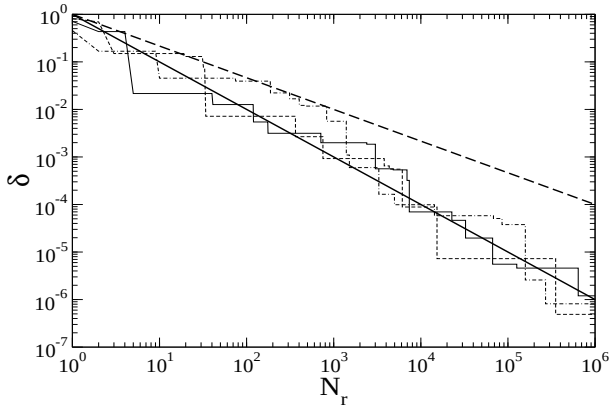

FIG. 3: Error δ in evaluating the diamond distance between channels (30) by means of the F-algorithm after N_r randomly chosen initial conditions. Three different runs are shown, together with the $1/N_r$ (full line) and $(1/N_r)^{2/3}$ (dashed line) dependences.

The trace norm $||\mathcal{E}||_1 = \frac{3}{2}$, as shown in Fig. 2, where $||\mathcal{E}||_1$ is given by the length of the segment $\overline{\rho_1'\rho_2'}$. On the other hand, the two channels $\mathcal{E}_1$ and $\mathcal{E}_2$ are perfectly distinguishable, as we readily obtain from Eq. (26) that $||\mathcal{E}||_\diamond = 2$, this value being achieved by means of maximally entangled input states. Therefore, in this example $||\mathcal{E}||_\diamond > ||\mathcal{E}||_1$, that is, entangled input states improve the distinguishability of the two channels.

Channels (30) are very convenient to illustrate the convergence properties of the F- and K-algorithms. Let us first consider the F-algorithm. If the maximum of $||\xi_1' - \xi_2'||_1$, with $\xi_i' = (\mathcal{I}_\mathcal{K} \otimes \mathcal{E}_i)\xi$ $(i = 1, 2)$ is taken over N_r random initial conditions, then the obtained value differs from the diamond norm $||\mathcal{E}||_\diamond$ by an amount $\delta(N_r)$ which must converge to zero when $N_r \to \infty$. To get convergence to $||\mathcal{E}||_\diamond$ for channels (30) it is enough to optimize over real initial conditions. Numerical results, shown in Fig. 3 for a few runs, are consistent with $\delta(N_r) \sim 1/N_r$. A rough argument can be used to explain the $1/N_r$ dependence. Assuming a smooth quadratic dependence of the distance $D(\xi) \equiv ||\xi_1' - \xi_2'||_1$ on the parameters for optimization [the angles θ_i in (10), with $\theta_1 \in \left[0, \frac{\pi}{2}\right]$, $\theta_2, \theta_3 \in [0, 2\pi)$] for ξ around the value ξ_0 optimizing D, then $|D(\xi) - D(\xi_0)| \sim \delta$ when the distance between ξ and ξ_0 is of the order of $\sqrt{\delta}$. The number N_r of randomly distributed initial conditions typically requested to get a point satisfying $||\xi - \xi_0||_1 < \sqrt{\delta}$ is of the order of $(1/\sqrt{\delta})^{n_p}$, where n_p is the number of parameters for optimization. Therefore, $\delta(N_r) \sim (1/N_r)^{2/n_p}$. In the example of Fig. 3, $n_p = 3$ leading to $\delta(N_r) \sim (1/N_r)^{2/3}$. On the other hand, numerical data exhibit a $1/N_r$ dependence. This fact has a simple explanation: the maximum is achieved for Bell states, which are invariant under rotations. Hence, the maximum distance is obtained not on a single point but on a curve and the number N_r of initial conditions requested to get a point within a trace distance smaller than $\sqrt{\delta}$ from this curve is of the order of $(1/\sqrt{\delta})^{n_p - 1}$, thus leading to $\delta(N_r) \sim 1/N_r$.

With regard to the K-algorithm, we have observed in the example of Eq. (30) the same $1/N_r$ convergence to the expected asymptotic value $||\mathcal{E}||_\diamond = 2$. However, the cost per initial condition in the K-algorithm is much larger than in the F-algorithm, in agreement with the general discussion of Sec. III B. Furthermore, the physical meaning of the F-method is much more transparent. For instance, for the Pauli channels (25) the affine transformation matrix $\mathcal{M}_i^{(1)}$ such that $\mathbf{R_i'} = (\mathcal{I}^{(1)} \otimes \mathcal{M}_i^{(1)})\mathbf{R}$ has a simple diagonal structure:

$$\mathcal{M}_i^{(1)} = \mathrm{diag}[(c_x)_i, (c_y)_i, (c_z)_i, 1], \tag{32}$$

where

$$\begin{aligned} (c_x)_i &= 1 - 2[(q_y)_i + (q_z)_i], \\ (c_y)_i &= 1 - 2[(q_z)_i + (q_x)_i], \\ (c_z)_i &= 1 - 2[(q_x)_i + (q_y)_i]. \end{aligned} \tag{33}$$

Matrix $\mathcal{M}_i^{(1)}$ simply accounts for the transformation, induced by $\mathcal{E}_i$, of the polarization measurements for the system qubit: $\alpha_i' = (c_\alpha)_i \alpha$, with $\alpha = x, y, z$ Bloch-sphere coordinates.

We have checked the computational advantages of the F-method also for all the other examples discussed in this paper. For this reason in what follows we shall focus on this method only.

B. Nonunital channels

In this section, we consider nonunital channels, that is, channels that do not preserve the identity. Therefore, in

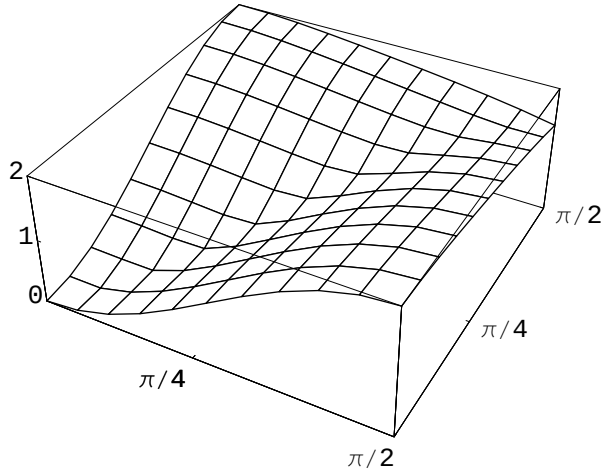

FIG. 4: Trace norm $||\mathcal{E}||_1 = ||\mathcal{E}_1 - \mathcal{E}_2||_1$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are displacement channels along the $+x$- and $+z$-axis of the Bloch sphere.

contrast to the Pauli channels considered in Sec. IV A, single-qubit nonunital channels displace the center of the Bloch sphere. Such channel are physically relevant in the description, for instance, of energy dissipation in open quantum systems, the simplest case being the amplitude damping channel [1, 2].

As a first illustrative example, we compute the distance between two channels $\mathcal{E}_1$ and $\mathcal{E}_2$ corresponding to displacements of the Bloch sphere along the $+x$- and $+z$-direction [14]. In the Fano representation the affine transformation matrices $\mathcal{M}_i^{(1)}$ corresponding to maps $\mathcal{E}_i$ have a simple structure:

$$\mathcal{M}_1^{(1)} = \begin{bmatrix} C_x^2 & 0 & 0 & S_x^2 \\ 0 & C_x & 0 & 0 \\ 0 & 0 & C_x & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{34}$$

$$\mathcal{M}_2^{(1)} = \begin{bmatrix} C_z & 0 & 0 & 0 \\ 0 & C_z & 0 & 0 \\ 0 & 0 & C_z^2 & S_z^2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{35}$$

where we have used the shorthand notation $C_x \equiv \cos\theta_x$ and $C_z \equiv \cos\theta_z$. The two channels depend parametrically on $\theta_x, \theta_z \in \left[0, \frac{\pi}{2}\right]$. The limiting cases $\theta_x = 0$ and $\theta_z = 0$ correspond to $\mathcal{E}_1 = \mathcal{I}$ and $\mathcal{E}_2 = \mathcal{I}$, respectively. Moreover, for $\theta_x = \frac{\pi}{2}$ quantum operation $\mathcal{E}_1$ maps the Bloch ball onto the single point $(x_1' = 1, y_1' = 0, z_1' = 0)$; for $\theta_z = \frac{\pi}{2}$ the mapping operated by $\mathcal{E}_2$ is onto the north pole of the Bloch sphere, $(x_2' = 0, y_2' = 0, z_2' = 1)$.

The numerically computed trace distance $||\mathcal{E}_1 - \mathcal{E}_2||_1$ is shown in Fig. 4, as a function of the parameters θ_x and θ_z. We gathered numerical evidence that for such channels the diamond distance equals the trace distance.

It is interesting to examine the analytical solutions for two limiting cases: (i) $\theta_x = 0$ (the case $\theta_z = 0$ is analogous) and (ii) $\theta_x = \theta_z = \frac{\pi}{2}$. In the first case, $\mathcal{E}_1$ maps the Bloch sphere into an ellipsoid with z as symmetry axis [14], while $\mathcal{E}_2 = \mathcal{I}$. The trace norm is given by the length of the line segment $\overline{\rho\rho_2'} = \overline{\rho_1'\rho_2'}$ shown in Fig. 5 (left),

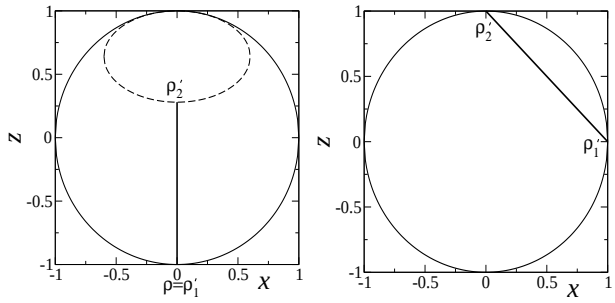

FIG. 5: Bloch-sphere visualization of the trace norm $||\mathcal{E}||_1$ in the limiting cases $\theta_x = 0$ (left) and $\theta_x = \theta_z = \frac{\pi}{2}$ (right).

$$||\mathcal{E}||_1 = 2\sin^2\theta_z. \tag{36}$$

It is interesting to remark that, in contrast to the Pauli channels, the optimal input state is not a maximally entangled input state. In the limiting case (i) the channels are strictly better discriminated by means of an appropriate separable input state, i.e., the south pole of the Bloch sphere (see the left plot of Fig. 5)) rather than by Bell states. Indeed, given the maximally entangled input state

$$\xi = |\Psi\rangle\langle\Psi|, \quad |\Psi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \tag{37}$$

we obtain

$$\xi_1' - \xi_2' = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & C_z - 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & S_z^2 & 0 \\ C_z - 1 & 0 & 0 & -S_z^2 \end{bmatrix}. \tag{38}$$

The trace distance between ξ_1' and ξ_2' is then computed by means of Eq. (15):

$$\begin{aligned} ||\xi_1' - \xi_2'||_1 = \frac{S_z^2}{2} + \frac{\left|S_z^2 + \sqrt{S_z^4 + 4(1-C_z)^2}\right|}{4} \\ + \frac{\left|S_z^2 - \sqrt{S_z^4 + 4(1-C_z)^2}\right|}{4}. \end{aligned} \tag{39}$$

As shown in Fig. 6, $||\xi_1' - \xi_2'||_1 < ||\mathcal{E}||_1$ for any $C_z > 0$.

In case (ii), for any initial state ρ, the Bloch coordinates of ρ_1' and ρ_2' are given by $(1, 0, 0)$ and $(0, 0, 1)$. The trace norm is given by the distance between these two points, $||\mathcal{E}||_1 = \sqrt{2}$ [see Fig. 5 (right)]. In this case, given an input Bell state or any other two-qubit input state ξ, the trace distance $||\xi_1' - \xi_2'||_1 = \sqrt{2} = ||\mathcal{E}||_1$. Thus, there is no advantage in using an ancillary system.

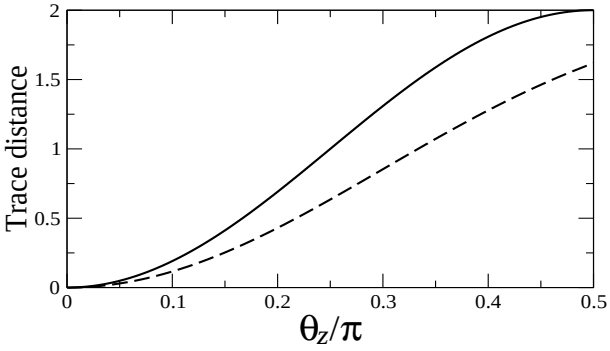

FIG. 6: Comparison between $||\mathcal{E}_1 - \mathcal{E}_2||_1$ (upper curve) and $||\xi_1' - \xi_2'||_1$ (lower curve) for $\theta_x = \frac{\pi}{2}$.

As a further example, we compute the distance between the depolarizing channel $\mathcal{E}_1$ and the nonunital channel $\mathcal{E}_2$ corresponding to the displacement along the $+z$-axis of the Bloch sphere. The depolarizing channel belongs to the class of Pauli channels (25), with $q_I = 1-p$ and $q_x = q_y = q_z = \frac{p}{3}$. This channel contracts the Bloch sphere by a factor $\left(1 - \frac{4}{3}p\right)$, with $0 \leq p \leq \frac{3}{4}$.

Fig. 7 shows the numerically computed $||\mathcal{E}||_1$ and $||\mathcal{E}||_\diamond$ as well as their difference $||\mathcal{E}||_\diamond - ||\mathcal{E}||_1$. The use of an ancillary qubit improves the distinguishability of the two channels. However, we remark again that maximally entangled input states can be detrimental. For instance, in the limiting case $p = 0$, $\theta_z = \frac{\pi}{2}$ we obtain from Eqs. (36) and (39) $||\mathcal{E}||_1 = 2 > ||\xi_1' - \xi_2'||_1 = \frac{1+\sqrt{5}}{2}$. A clear advantage is instead seen in another limiting case, $p = \frac{3}{4}$, $\theta_z = 0$. The fully depolarizing channel maps each point of the Bloch ball onto its center, so that the trace distance is given by the radius of the Bloch sphere, $||\mathcal{E}_1 - \mathcal{E}_2||_1 = 1$. On the other hand, by means of Eq. (26) we obtain $||\mathcal{E}_1 - \mathcal{E}_2||_\diamond = \frac{3}{2}$.

V. CONCLUSIONS

We have shown that the distance between two quantum channels can be conveniently computed by means of a Monte-Carlo algorithm based on the Fano representation. The effectiveness of this algorithm is illustrated in the case, most relevant for present-day implementations of quantum information processing, of single-qubit quantum channels. A main computational advantage of this algorithms is that it is easily parallelizable. Furthermore, being based on the Fano representation, is enlights the physical meaning of the involved quantum channels: the matrix elements of the affine map representing a quantum channel directly account for the evolution of the expectation values of the system's polarization measurements. More generally, we believe that the Fano representation provides a computationally convenient and physically trasparent representation of quantum noise.

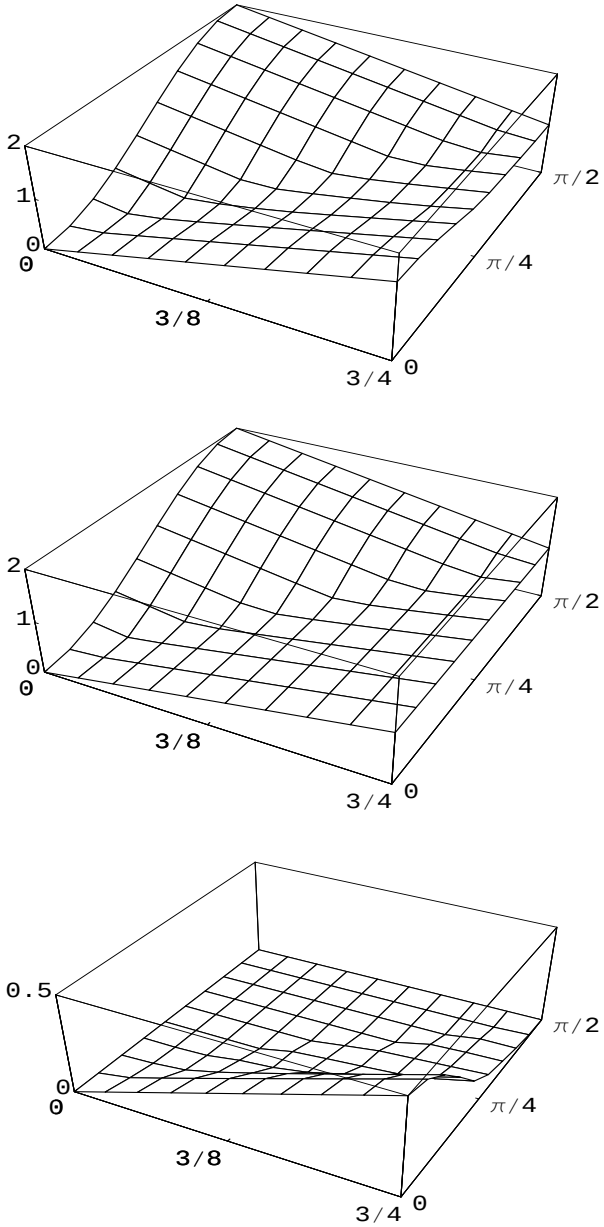

FIG. 7: From top to bottom: trace norm $||\mathcal{E}||_1 = ||\mathcal{E}_1 - \mathcal{E}_2||_1$, diamond norm $||\mathcal{E}||_\diamond$, and their difference $||\mathcal{E}||_\diamond - ||\mathcal{E}||_1$; $\mathcal{E}_1$ is the depolarizing channel, $\mathcal{E}_2$ the displacement along the $+z$-axis of the Bloch sphere.

APPENDIX A: ALTERNATIVE DECOMPOSITION OF $\mathcal{E}$

Given a superoperator $\mathcal{E} = \mathcal{E}_1 - \mathcal{E}_2$, with $\mathcal{E}_1$, $\mathcal{E}_2$ quantum operations, we start from the Kraus representation [1, 2] of $\mathcal{E}_1$ and $\mathcal{E}_2$:

$$\mathcal{E}_1(X) = \sum_{k=1}^{M_1} E_k X E_k^\dagger, \quad \mathcal{E}_2(X) = \sum_{k=1}^{M_2} F_k X F_k^\dagger, \qquad \text{(A1)}$$

and define new operators:

$$\tilde{A}^{(2k-1)} = \frac{1}{\sqrt{2}}(F_k + G_k), \quad \tilde{A}^{(2k)} = \frac{1}{\sqrt{2}}(F_k - G_k),$$
$$\tilde{B}^{(2k-1)} = \tilde{A}^{(2k)}, \quad \tilde{B}^{(2k)} = \tilde{A}^{(2k-1)}, \tag{A2}$$

where $k = 1, ..., \overline{M} \equiv \max(M_1, M_2)$. Note that, if $M_1 > M_2$, we set $G_k = 0$ for $k = M_2 + 1, ..., M_1$; vice-versa, if $M_1 < M_2$, $F_k = 0$ for $k = M_1 + 1, ..., M_2$. It is easy to see that

$$\mathcal{E}(X) = \sum_{i=1}^{2\overline{M}} \tilde{A}^{(i)} X \tilde{B}^{(i)\dagger}. \tag{A3}$$

In contrast to (19), the present decomposition of $\mathcal{E}(X)$ is simpler, in that no singular value decomposition is required, but less efficient. Indeed, the maximum number of terms in (A3) is twice that of decomposition (19). This implies that, if Ψ_A, Ψ_B are expressed in terms of operators $\tilde{A}^{(i)}$, $\tilde{B}^{(i)}$ rather than $A^{(i)}$, $B^{(i)}$, to evaluate $F(\Psi_A, \Psi_B)$ we need to compute eigenvalues and eigenvectors of matrices of size $2\overline{M}$. In the single-qubit case, typically $2\overline{M} = 8$.

ACKNOWLEDGMENTS

We thank Massimiliano Sacchi for interesting comments on our work.

[1] G. Benenti, G. Casati, and G. Strini, *Principles of Quantum Computation and Information*, Vol. I: Basic concepts (World Scientific, Singapore, 2004); Vol. II: Basic tools and special topics (World Scientific, Singapore, 2007).

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[4] A. M. Childs, J. Preskill, and J. Renes, J. Mod. Opt. **47**, 155 (2000).

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

[6] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).

[7] M. F. Sacchi, Phys. Rev. A **71**, 062340 (2005); *ibid.* **72**, 014305 (2005).

[8] G. M. D'Ariano, M. F. Sacchi, and J. Kahn, Phys. Rev. A **72**, 052302 (2005).

[9] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, vol. 47 of Graduate Studies in Mathematics (American Mathematical Society, Providence, Rhode Island, 2002), Sec. 11.

[10] J. Watrous, *Advanced Topics in Quantum Information Processing*, chap. 22, Lecture notes (2004), at http://www.cs.uwaterloo.ca/∼watrous/lecture-notes/701/.

[11] N. Johnston, D. W. Kribs, and V. I. Paulsen, Quantum Inf. Comput. **9**, 16 (2009).

[12] J. Watrous, Theory of Computing **5**, 11 (2009).

[13] A. Ben-Aroya and A. Ta-Shma, Quantum Inf. Comput. **10**, 87 (2010).

[14] G. Benenti, S. Felloni, and G. Strini, Eur. Phys. J. D **38**, 389 (2006).

[15] G. Benenti and G. Strini, Phys. Rev. A **80**, 022318 (2009).

[16] B. Rosgen and J. Watrous, in Proceedings of the 20th Annual Conference on Computational Complexity, pages 344-354 (2005), preprint arXiv:cs/0407056.

[17] U. Fano, Rev. Mod. Phys. **29**, 74 (1957); *ibid.* **55**, 855 (1983).

[18] F. T. Hioe and J. H. Eberly, Phys. Rev. Lett. **47**, 838 (1981).

[19] J. Schlienz and G. Mahler, Phys. Rev. A **52**, 4396 (1995).