PHYSICAL REVIEW A **96**, 042325 (2017)

# Optimal convex approximations of quantum states

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 Agostino Bassi 6, I-27100 Pavia, Italy*

(Received 21 July 2017; published 18 October 2017)

We consider the problem of optimally approximating an unavailable quantum state $\rho$ by the convex mixing of states drawn from a set of available states $\{\nu_i\}$. The problem is recast to look for the least distinguishable state from $\rho$ among the convex set $\sum_i p_i \nu_i$, and the corresponding optimal weights $\{p_i\}$ provide the optimal convex mixing. We present the complete solution for the optimal convex approximation of a qubit mixed state when the set of available states comprises the three bases of the Pauli matrices.

DOI: 10.1103/PhysRevA.96.042325

## I. INTRODUCTION

Convex structures are ubiquitous in the realm of quantum mechanics. Density matrices, probability operator-valued measures, and completely positive maps—which represent quantum states, quantum measurements, and quantum channels, respectively—are convex sets. The weights in a convex sum typically represent classical probabilities which have an immediate operational interpretation: They are the weights of the extremal points of the set and may correspond to classical processing.

In Ref. [1], the problem of optimally approximating an unavailable quantum channel by the convex mixing of channels which are supposed to be available was addressed. This operational problem has been recast to the problem of looking for the least distinguishable channel from the target among the convex set of channels constructed by the given set.

Here in this paper we address the analogous problem for quantum states, namely the problem of optimally generating a desired quantum state $\rho$, when only a given set of quantum states $\{\nu_i\}$ is disposable. In this case, we will look for the best convex combination among the states of the given set that mostly resembles the desired $\rho$, i.e., that is the least distinguishable from $\rho$ itself. As for the case of convex approximation of channels, this approach has clearly a prompt experimental application when the effectively available states in a laboratory are limited for intrinsic restrictions, unavailable technology, or even economical reasons. A further relevance of this approach is due to the fact that a convex sum of states offers the possibility of performing different experiments followed by postprocessing of experimental data when the quantities of interest are linear with respect to the input states.

Since the natural measure of distinguishability between quantum states is based on the trace norm [2], we note that our general problem of convex approximation includes the well-studied (and still open) problem of quantifying the coherence of quantum states for the specific case where the available set $\{\nu_i\}$ corresponds to a complete orthogonal basis, via the trace-distance measure of coherence [3–5]. Also, for generic mutipartite state $\rho$ and available set of states given by all product pure states, our problem is equivalent to evaluating the trace norm of entanglement measure [6–9].

## II. CONVEX APPROXIMATION OF QUANTUM STATES

It is well known that the probability $p_{\rm discr}$ of optimally discriminating between two quantum states $\rho_0$ and $\rho_1$ given with equal *a priori* probability is given by [2]

$$p_{\rm discr}(\rho_0,\rho_1) = \tfrac{1}{2} + \tfrac{1}{4}\|\rho_0 - \rho_1\|_1 \;, \qquad (1)$$

where $\|A\|_1$ denotes the trace norm of $A$, namely [10],

$$\|A\|_1 = {\rm Tr}\sqrt{A^\dagger A} = \sum_i s_i(A) \;, \qquad (2)$$

with $\{s_i(A)\}$ representing the singular values of $A$. In the case of Eq. (1), the singular values just correspond to the absolute value of the eigenvalues, since the operator inside the norm is Hermitian. Let us also recall that the optimal measurement for the discrimination is performed by the projectors on the support of the positive and negative parts of the Hermitian operator $\rho_0 - \rho_1$.

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

*Definition.* The optimal convex approximation of a quantum state $\rho$ with respect to (w.r.t.) a given set of quantum states $\{\nu_i\}$ is given by $\sum_i p_i^{\rm opt} \nu_i$, where $\{p_i^{\rm opt}\}$ denotes the vector of probabilities

$$\left\{p_i^{\rm opt}\right\} = \arg \min_{\{p_i\}} \left\| \rho - \sum_i p_i \nu_i \right\|_1 . \qquad (3)$$

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

$$D_{\{\nu_i\}}(\rho) \equiv \min_{\{p_i\}} \left\| \rho - \sum_i p_i \nu_i \right\|_1 , \qquad (4)$$

which provides the worst probability of discriminating the desired state $\rho$ from any of the available states $\sum_i p_i \nu_i$. 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 states (e.g., a decreasing function of the fidelity).

We notice that the formulation of the trace norm as a semidefinite program [11] allows its efficient calculation.

*msacchi@unipv.it

 

Moreover, the convexity of the norm itself allows one to search the minimum by means of standard software of convex optimization [12,13].

From the convexity of the trace norm, it follows the upper bound

$$D_{\{\nu_i\}}(\rho) \leqslant \min_i \|\rho - \nu_i\|_1 = \min_i D_{\nu_i}(\rho)\ . \tag{5}$$

Notice also that for all unitary operators $U$, from the unitary invariance of the trace norm, one has the symmetry

$$D_{\{\nu_i\}}(\rho) = D_{\{U\nu_i U^\dagger\}}(U\rho U^\dagger)\ . \tag{6}$$

Clearly, if the set itself $\{\nu_i\}$ is invariant, then

$$D_{\{\nu_i\}}(U\rho U^\dagger) = D_{\{\nu_i\}}(\rho)\ , \tag{7}$$

and the probabilities of the optimal convex approximation for $U\rho U^\dagger$ are just a permutation of those for $\rho$. This is the case, for example, when the set of the available quantum states is covariant w.r.t. a (projective) unitary representation of a group.

## III. PAULI DISTANCE OF QUBIT STATES

In the following, we provide the complete analytical solution for the optimal convex approximation of an arbitrary mixed qubit state, when the available set of states is given by the eigenvectors of the three Pauli matrices.

Let us first consider the simpler case where the set of available states is an orthogonal basis. Without loss of generality, let us identify such basis as the eigenstates $\mathsf{B}_1 = \{|0\rangle,|1\rangle\}$ of $\sigma_z$-Pauli matrix and parametrize the target qubit state $\rho$ as

$$\rho = \begin{pmatrix} 1-a & k\sqrt{a(1-a)}e^{-i\phi} \\ k\sqrt{a(1-a)}e^{i\phi} & a \end{pmatrix}, \tag{8}$$

with $a \in [0,1]$, $\phi \in [0,2\pi]$, and $k \in [0,1]$. A straightforward calculation provides the optimal convex approximation of $\rho$ as the diagonal matrix

$$\rho_d = \begin{pmatrix} 1-a & 0 \\ 0 & a \end{pmatrix}. \tag{9}$$

Clearly, the optimal weights are given by $p_0^{\text{opt}} = 1-a$ and $p_1^{\text{opt}} = a$, and the approximation is quantified by the $\mathsf{B}_1$ distance

$$D_{\{|0\rangle,|1\rangle\}}(\rho) = 2k\sqrt{a(1-a)}\ . \tag{10}$$

This result also corresponds to the trace-distance measure of coherence for the state $\rho$ referred to the $\sigma_z$ eigenstates [14]. Better approximations can be obviously obtained when a larger set of states is available.

Let us consider now the set containing the eigenstates of all Pauli matrices, namely

$$\mathsf{B}_3 = \Big\{|0\rangle,|1\rangle,|2\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle),|3\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle-|1\rangle),$$
$$|4\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle),|5\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)\Big\}. \tag{11}$$

Since the $\mathsf{B}_3$ distance (or, equivalently, *Pauli distance*) is invariant for the state transformations $\rho(a,k,\phi) \to \rho(1-a,k,\phi)$ and $\rho(a,k,n\pi/2 \pm \phi) \to \rho(a,k,\phi)$ (with integer $n$), we can restrict the study to the case $a \in [0,1/2]$ and $\phi \in [0,\pi/2]$. One can immediately find a large set of density matrices which indeed correspond to a convex mixing of the six states of $\mathsf{B}_3$. In fact, we can rewrite $\rho$ in Eq. (8) as follows:

$$\begin{aligned}\rho = &(1-2a)|0\rangle\langle 0| \\ &+ 2k\sqrt{a(1-a)}(\cos\phi|2\rangle\langle 2| + \sin\phi|4\rangle\langle 4|) \\ &+ [a - k\sqrt{a(1-a)}(\cos\phi+\sin\phi)]I,\end{aligned} \tag{12}$$

where $I$ denotes the two-dimensional identity matrix. It follows that there is a threshold value for the coherence parameter $k$ under which $D_{\mathsf{B}_3}(\rho) = 0$, namely for

$$k \leqslant k_{th} \equiv \frac{a}{\sqrt{a(1-a)}(\cos\phi+\sin\phi)}\ . \tag{13}$$

The pertaining weights that provide such an exact convex decomposition can be chosen as follows [15]:

$$\begin{aligned} p_0 &= 1 - a - k\sqrt{a(1-a)}(\cos\phi+\sin\phi), \\ p_1 &= a - k\sqrt{a(1-a)}(\cos\phi+\sin\phi), \\ p_2 &= 2k\sqrt{a(1-a)}\cos\phi, \\ p_4 &= 2k\sqrt{a(1-a)}\sin\phi, \\ p_3 &= p_5 = 0. \end{aligned} \tag{14}$$

In terms of the expectation values $\langle\sigma_\alpha\rangle = \text{Tr}[\rho\sigma_\alpha]$, with $\alpha = x,y,z$ and $\rho$ as in Eq. (8), notice the identities

$$\begin{aligned} \langle\sigma_x\rangle &= 2k\sqrt{a(1-a)}\cos\phi\ , \\ \langle\sigma_y\rangle &= 2k\sqrt{a(1-a)}\sin\phi\ , \\ \langle\sigma_z\rangle &= 1-2a. \end{aligned} \tag{15}$$

Thus, the condition $k \leqslant k_{th}$ in Eq. (13) can be rewritten more transparently as

$$\langle\sigma_x\rangle + \langle\sigma_y\rangle + \langle\sigma_z\rangle \leqslant 1\ . \tag{16}$$

With the help of symbolic computation, by imposing a vanishing value to the gradient of $\|\rho - \sum_{i=0}^5 p_i|i\rangle\langle i|\|_1$ with respect to the probabilities $\{p_i\}$, one can obtain the complete analytical solution for the optimal convex approximation of $\rho$ when $k > k_{th}$, and hence $D_{\mathsf{B}_3}(\rho) > 0$. Explicitly, one obtains the following three cases:

(i) For $k_{th} < k \leqslant \frac{a}{\sqrt{a(1-a)}}$, or $k > \frac{a}{\sqrt{a(1-a)}}$ and $\phi \in [\phi_{th},\pi/2-\phi_{th}]$, with

$$\phi_{th} = 2\arctan\left[\frac{\sqrt{5k^2a(1-a)-a^2} - 2k\sqrt{a(1-a)}}{a+k\sqrt{a(1-a)}}\right], \tag{17}$$

the optimal convex approximation has Pauli distance

$$D_{\mathsf{B}_3}(\rho) = \frac{2}{\sqrt{3}}\sqrt{a(1-a)(1+\sin 2\phi)}(k-k_{th})\ , \tag{18}$$

with pertaining optimal weights

$$\begin{aligned} p_0 &= 1 - \tfrac{4}{3}a - \tfrac{2}{3}k\sqrt{a(1-a)}(\cos\phi+\sin\phi), \\ p_2 &= \tfrac{2}{3}[a + k\sqrt{a(1-a)}(2\cos\phi - \sin\phi)], \\ p_4 &= \tfrac{2}{3}[a + k\sqrt{a(1-a)}(2\sin\phi - \cos\phi)], \\ p_1 &= p_3 = p_5 = 0. \end{aligned} \tag{19}$$

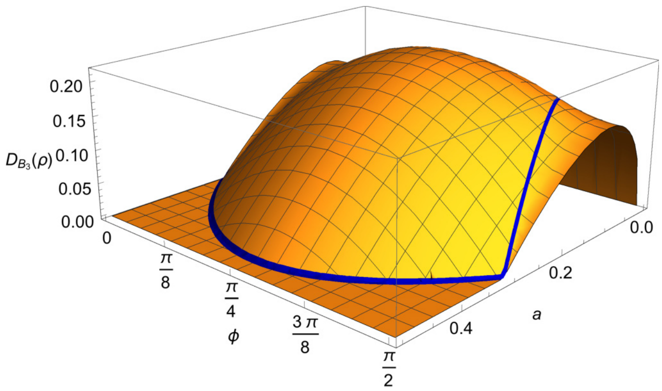


FIG. 1. Optimal convex approximation of a qubit mixed state $\rho$ w.r.t. the set $B_3$ of the eigenstates of the three Pauli matrices. The Pauli distance $D_{B_3}(\rho)$ is here plotted vs the target state parameters $a$ and $\phi$, for fixed value of the parameter $k = \frac{2}{3}$ [see the parametrization of $\rho$ in Eq. (8)]. The plotted surface is a piecewise function obtained from Eqs. (18), (20), and (22) in their region of definition. According to Eq. (13), the minimal trace distance vanishes in the region where $\frac{a}{\sqrt{a(1-a)}(\cos\phi+\sin\phi)} \geqslant k = \frac{2}{3}$.

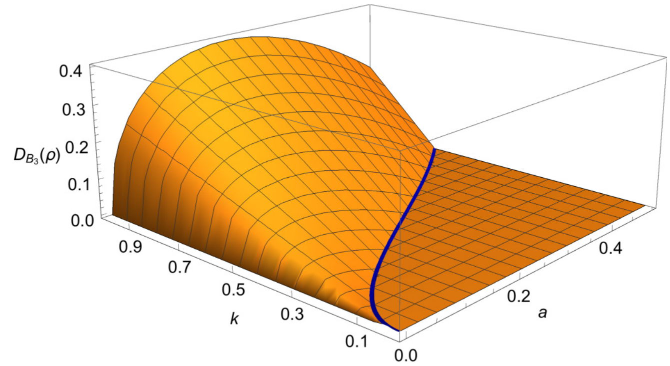


FIG. 2. Optimal convex approximation of a qubit mixed state $\rho$ w.r.t. the set $B_3$ of the eigenstates of the three Pauli matrices. The Pauli distance $D_{B_3}(\rho)$ is here plotted vs the target state parameters $a$ and $k$, for fixed value of the phase $\phi = \frac{\pi}{3}$. According to Eq. (13), the minimal trace distance vanishes for $k \leqslant \frac{2a}{(\sqrt{3}+1)\sqrt{a(1-a)}}$.

(ii) For $k > \frac{a}{\sqrt{a(1-a)}}$ and $\phi \in [0,\phi_{th}]$, the optimal convex approximation has Pauli distance

$$D_{B_3}(\rho) = \{2a[a - 2k\sqrt{a(1-a)}\cos\phi + k^2(1-a)(2-\cos^2\phi)]\}^{1/2}, \quad (20)$$

with optimal weights

$$\begin{aligned} p_0 &= 1 - a - k\sqrt{a(1-a)}\cos\phi, \\ p_2 &= a + k\sqrt{a(1-a)}\cos\phi, \\ p_1 &= p_3 = p_4 = p_5 = 0. \end{aligned} \quad (21)$$

(iii) For $k > \frac{a}{\sqrt{a(1-a)}}$ and $\phi \in [\pi/2 - \phi_{th}, \pi/2]$, the optimal convex approximation has Pauli distance

$$D_{B_3}(\rho) = \{2a[a - 2k\sqrt{a(1-a)}\sin\phi + k^2(1-a)(2-\sin^2\phi)]\}^{1/2}, \quad (22)$$

with optimal weights

$$\begin{aligned} p_0 &= 1 - a - k\sqrt{a(1-a)}\sin\phi, \\ p_4 &= a + k\sqrt{a(1-a)}\sin\phi, \\ p_1 &= p_2 = p_3 = p_5 = 0. \end{aligned} \quad (23)$$

Notice that the exact convex decomposition (when $k \leqslant k_{th}$) involves four states, whereas the optimal convex approximation corresponds to a mixture of three states in case (i) and just two states in cases (ii) and (iii).

In Figs. 1 and 2, respectively, we plot the results for the optimal convex approximation of $\rho$ versus parameters $a$ and $\phi$ with fixed value of the parameter $k = \frac{2}{3}$, and versus parameters $a$ and $k$ with fixed value of the phase parameter $\phi = \frac{\pi}{3}$.

## IV. CONCLUSIONS

Let us conclude our paper with the following observations. Imagine that we want to approximate $N$ copies of the state, namely $\rho^{\otimes N}$, and we have at disposal a set of single-copy states $\{\nu_i\}$. The optimal convex approximation in this case provides the distance $D_{\{\otimes_{j=1}^N \nu_{i_j}\}}(\rho^{\otimes N})$. Since the convex hull of $\{\otimes_{j=1}^N \nu_{i_j}\}$ contains all the $N$-fold tensor products $\otimes_{j=1}^N(\sum_i p_{i_j}\nu_i)$, one has

$$\begin{aligned} D_{\{\otimes_{j=1}^N \nu_{i_j}\}}(\rho^{\otimes N}) &\leqslant \min_{\{p_{i_j}\}} \left\| \rho^{\otimes N} - \otimes_{j=1}^N \left( \sum_i p_{i_j}\nu_i \right) \right\|_1 \\ &\leqslant \left\| \rho^{\otimes N} - \left( \sum_i p_i^{\rm opt}\nu_i \right)^{\otimes N} \right\|_1, \end{aligned} \quad (24)$$

where $\{p_i^{\rm opt}\}$ denotes the vector of probabilities pertaining to the optimal convex approximation of a single copy of the state $\rho$.

We notice that one can find strict inequalities in both lines of Eq. (24). The first inequality arises because the presence of correlations in the convex approximation can be beneficial even if the target state is indeed the product of independent states (as it occurs, for example, in the optimal cloning of quantum states [16], where the copies are correlated). In fact, it is also known that correlations limit the extractable information [17–20], and here indeed we want to minimize the probability of discriminability. The second inequality stems from the fact that the distance for a convex optimization has no additive and/or multiplicative property with respect to the tensor product. Hence, even looking for a tensor-product state for the optimal convex approximation, the corresponding optimal weights will be in general different from those pertaining to the optimal convex approximation of a single copy. This also implies that we do not have an exact expression for the scaling with $N$ of the distance between a quantum state and its convex approximation. In a systematic study of the scaling of the optimal convex approximation with the number of copies, the results related to the quantum Chernoff bound [21,22] might be very useful.

A specific example where Eq. (24) is satisfied with two strict inequalities is the following. Consider the pure qubit state $|\psi\rangle = \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle$. Its optimal convex approximation with respect to the basis $\{|0\rangle,|1\rangle\}$ is achieved by $p_0^{\rm opt}=3/4$ and $p_1^{\rm opt}=1/4$, with corresponding distance $D_{\{|0\rangle,|1\rangle\}}(|\psi\rangle\langle\psi|)=\frac{\sqrt{3}}{2}$. The two-copy trace distance can be evaluated as

$$\left\||\psi\rangle\langle\psi|^{\otimes 2}-\left(\tfrac{3}{4}|0\rangle\langle 0|+\tfrac{1}{4}|1\rangle\langle 1|\right)^{\otimes 2}\right\|_1\simeq 1.299. \quad (25)$$

On the other hand, the optimal convex approximation with respect to factorized diagonal states is given by

$$\begin{aligned} &D_{\mathsf{B}_1\otimes\mathsf{B}_1}(|\psi\rangle\langle\psi|^{\otimes 2}) \\ &\quad=\min_{p,q\in[0,1]}\||\psi\rangle\langle\psi|^{\otimes 2}-[p|0\rangle\langle 0|+(1-p)|1\rangle\langle 1|] \\ &\qquad\otimes[q|0\rangle\langle 0|+(1-q)|1\rangle\langle 1|]\|_1\simeq 1.272 \end{aligned} \quad (26)$$

and is achieved for $p^{\rm opt}=q^{\rm opt}\simeq 0.859$. Finally, by allowing correlations between the copies of the convex approximation, one obtains

$$\begin{aligned} &D_{\{|00\rangle,|01\rangle,|10\rangle,|11\rangle\}}(|\psi\rangle\langle\psi|^{\otimes 2}) \\ &\quad=\min_{\{p_{ij}\}}\||\psi\rangle\langle\psi|^{\otimes 2}-(p_{00}|00\rangle\langle 00|+p_{01}|01\rangle\langle 01| \\ &\qquad+p_{10}|10\rangle\langle 10|+p_{11}|11\rangle\langle 11|)\|_1\simeq 1.265, \end{aligned} \quad (27)$$

where the optimal weights are given by $p_{00}^{\rm opt}\simeq 0.712$, $p_{01}^{\rm opt}=p_{10}^{\rm opt}\simeq 0.144$, and $p_{11}^{\rm opt}=0$. Notice that the improvement in the convex approximation of $|\psi\rangle\langle\psi|^{\otimes 2}$ is exclusively due to classical correlations, since obviously no entanglement is present in the approximating state.


[1] M. F. Sacchi and T. Sacchi, Phys. Rev. A **96**, 032311 (2017).
[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[3] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).
[4] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, Phys. Rev. Lett. **115**, 020403 (2015).
[5] S. Rana, P. Parashar, and M. Lewenstein, Phys. Rev. A **93**, 012110 (2016).
[6] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).
[7] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).
[8] M. B. Plenio and S. Virmani, Quantum Inf. Comput. **7**, 1 (2007).
[9] T. R. Bromley, M. Cianciaruso, and G. Adesso, Phys. Rev. Lett. **114**, 210401 (2015).
[10] See, for example, R. Bhatia, *Matrix Analysis*, Springer Graduate Texts in Mathematics Vol. 169 (Springer, New York, 1996).
[11] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004).
[12] M. C. Grant and S. P. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, http://cvxr.com/cvx.
[13] M. C. Grant and S. P. Boyd, in *Recent Advances in Learning and Control*, edited by V. D. Blondel, S. P. Boyd, and H. Kimura, Lecture Notes in Control and Information Sciences (Springer-Verlag, Berlin, 2008), p. 95.
[14] L.-H. Shao, Z. Xi, H. Fan, and Y. Li, Phys. Rev. A **91**, 042120 (2015).
[15] The solution is not unique, since clearly one can equivalently ascribe the weight of the identity matrix in Eq. (12) to any of the three Pauli matrix bases.
[16] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín, Rev. Mod. Phys. **77**, 1225 (2005).
[17] R. Demkowicz-Dobrzański, Phys. Rev. A **71**, 062321 (2005).
[18] J. Bae and A. Acín, Phys. Rev. Lett. **97**, 030402 (2006).
[19] G. M. D'Ariano, R. Demkowicz-Dobrzański, P. Perinotti, and M. F. Sacchi, Phys. Rev. Lett. **99**, 070501 (2007).
[20] G. M. D'Ariano, R. Demkowicz-Dobrzański, P. Perinotti, and M. F. Sacchi, Phys. Rev. A **77**, 032344 (2008).
[21] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. I. Masanes, A. Acín, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[22] M. Nussbaum and A. Szkoła, Ann. Statis. **37**, 1040 (2009).