Channel discrimination power of bipartite quantum states

Matteo Caiaffa, Marco Piani
SUPA and Department of Physics, University of Strathclyde, Glasgow G4 0NG, UK

We quantify the usefulness of a bipartite quantum state in the ancilla-assisted channel discrimination of arbitrary quantum channels, formally defining a worst-case-scenario channel discrimination power for bipartite quantum states. We show that such a quantifier is deeply connected with the operator Schmidt decomposition of the state. We compute the channel discrimination power exactly for pure states, and provide upper and lower bounds for general mixed states. We show that highly entangled states can outperform any state that passes the realignment criterion for separability and that the channel discrimination power of a state is bounded by its quantum discord.

I. INTRODUCTION

A quantum channel is the most general deterministic linear transformation a quantum system can undergo, capturing mathematically the notion of physical process and playing the role of basic building block in quantum information processing [1]. Channel discrimination is a fundamental task that falls under the umbrella of quantum metrology [2, 3] and consists in the attempt to tell apart two or more *known* channels; think of the situation where we want to probe the presence or absence of a magnetic field. In the prototypical and simplest case, one of two channels is applied once to a probe, and we try to identify the channel by performing a measurement on the output probe. Channel discrimination is typically performed by tailoring the state of the input probe to the channels to be discriminated. The wrong choice of input might make the probability of correct identification less than optimal or even not better than a random guess.

There can be advantages in channel discrimination by making use of correlations between the probe and a reference ancilla. One possible advantage is that correlations may lead to a probability of success in the discrimination that is higher than what possible without the use of an ancillary system [2, 4–15]. In general, achieving such a higher probability of success requires (i) to tailor the probe-ancilla input state to the specific channels to be discriminated and (ii) input entanglement between probe and ancilla. Another advantage provided by probe-ancilla correlations, on which we focus in this work, is that they may allow to discriminate between an arbitrary pair of known channels, without the need to tailor the input probe-ancilla state to avoid 'being blind' to the difference between the channels. This fact is at the basis of the celebrated Choi-Jamiołkowski isomorphism [16, 17] between linear maps and linear operators, which allows to effectively encode or parametrize quantum transformations, with applications that go from the optimization of protocols in quantum information, to the analysis of rates in quantum communication, all the way to the consideration of the issue of causal order in physics (see, e.g., [18–20]). The use of an ancilla allows one to perform channel tomography – that is, to identify an unknown channel with many uses of the unknown channel – with a fixed input state [21]. Such a feat can be achieved even in the absence of entanglement, and Ref. [21] already identified the Operator Schmidt Rank (OSR; to be defined later) of the probe-ancilla input state as the key property determining whether such state makes ancilla-assisted tomography possible. An equivalent result was independently derived in [22], where the *faithfulnees* of bipartite quantum states was introduced. A bipartite state used in ancilla-assisted channel tomography is faithful if the action of the channel on the probe leads to an output probe-ancilla state that is uniquely associated with the specific channel. Nonetheless, the study of the usefulness of correlations in fixed-input ancilla-assisted channel discrimination and channel tomography has been limited [23]. In this article, we shed light on ancilla-assisted channel discrimination, providing an analysis of how the Operator Schmidt Decomposition (OSD; to be defined later) of the probe-ancilla input state affects the quality of the discrimination. In particular, we introduce a worst-case quantifier for the performance of a probe-ancilla state in channel discrimination, the *Channel Discrimination Power* (CDP). We provide general upper and lower bounds to the CDP of a state in terms of the OSD of the state. We compute the exact CDP of pure states. Remarkably, we show that, while correlated but unentangled states can have non-zero CDP, and allow the discrimination of any pair of channels as long as they have maximal OSR, they cannot have maximal CDP. More in general, we provide a non-trivial bound on the channel discrimination power of any state – entangled or unentangled – that passes the so-called realignment (or computable cross-norm) criterion for separability [24, 25]. Furthermore, we prove that the general quantumness of correlations known as quantum discord [26] provides a bound for the channel discrimination power of a bipartite state.

II. NOTATION AND PRELIMINARIES

We will consider finite-dimensional systems. Hence, each Hilbert space $\mathscr{H}$ will be equivalent to $\mathbb{C}^d$ for some dimension d. The space of linear operators L (equivalent to matrices) on $\mathscr{H}$ will be indicated by $\mathscr{L}(\mathscr{H})$. We will be interested in the p-norms $\|L\|_p := \left(\mathrm{Tr}\left((L^\dagger L)^{\frac{p}{2}}\right)\right)^{\frac{1}{p}}$,

for the values $p = 1, 2, \infty$ [27].

We indicate by d_X the dimension of a system X with Hilbert space $\mathscr{H}_X$. We will focus on bipartite systems AB, and, unless stated otherwise, we will define $d_{\min} = \min\{d_A, d_B\}$. A quantum state on $\mathscr{H}$ corresponds to a density operator ρ belonging to convex subset $\mathscr{D}(\mathscr{H}) \subset \mathscr{L}(\mathscr{H})$ of operators that have unit trace and are positive semidefinite. We indicate by Tr the trace operation, and by Tr_X the partial trace on system X. We denote by ρ_X the (reduced) state of system X. The space $\mathscr{L}(\mathscr{H})$ can be made into a Hilbert space itself by considering the Hilbert-Schmidt inner product $\langle\langle C|D\rangle\rangle := \mathrm{Tr}(C^\dagger D)$. Notice that $\|C\|_2 = \sqrt{\langle\langle C|C\rangle\rangle}$.

The trace distance between two density matrices ρ and σ is defined as $D(\rho,\sigma) := \frac{1}{2}\|\rho - \sigma\|_1$ [1]. Its operational meaning is that of bias in the optimal discrimination of the two states: the probability of correctly identifying the state of a system that is in either the state ρ or σ each with a priori probability 50%, in the single-shot scenario when one is given one copy of the state to measure, is $(1 + D(\rho,\sigma))/2$. The trace distance varies between 0 (for identical states) to 1 (for perfectly distinguishable states, which are mathematically orthogonal, $\langle\langle\rho|\sigma\rangle\rangle = 0$).

A bipartite state ρ_{AB} is unentangled (or separable) if it is the convex combination of product (or uncorrelated) states [28], $\rho_{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B$, with $\{p_i\}$ a probability distribution, so that all the correlations encoded in such a state have an explanation in terms of shared classical randomness. A state is entangled if it is not separable. To decide whether a given state is entangled is a hard problem, but numerous entanglement criteria exist [28].

The physical evolution of a quantum system is formally described in terms of quantum channels [1]. In general, one considers evolutions from an input system X to an output system Y, representing evolution in time or general transfer of information – either in space or in time – from one system to another. Formally, a quantum channel from X to Y is a completely-positive trace-preserving linear map Λ from $\mathscr{L}(\mathscr{H}_X)$ to $\mathscr{L}(\mathscr{H}_Y)$.

In the remainder of this paper it will be convenient to work directly with norms, e.g. $\|X\|_1$, rather than with derived distances, e.g., rather than in terms of the trace distance between two states ρ and σ. It is useful to recall that, for a Hermitian operators $X = X^\dagger$, one has

$$\|X\|_1 = \max_{-\mathbb{1}\leq M\leq \mathbb{1}} |\,\mathrm{Tr}(MX)|.$$

We define the (Hermitian) super-operator 1-norm of an Hermiticity preserving map Γ as [14]

$$\|\Gamma\|_1 = \sup_{X=X^\dagger;\|X\|_1=1} \|\Gamma[X]\|_1.$$

Notice that this is equivalent to

$$\|\Gamma\|_1 = \sup_{X=X^\dagger\neq 0} \frac{\|\Gamma[X]\|_1}{\|X\|_1}.$$

It is immediate to argue by convexity that the best input X can always be taken to be a pure normalized state $|\psi\rangle\langle\psi|$ [14].

We define the diamond norm of an Hermiticity preserving map Γ as

$$\|\Gamma\|_\diamond := \sup_n \|\Gamma \otimes \mathrm{id}_{\mathbb{C}^n}\|_1,$$

where the supremum is over the dimension of the ancillary space $\mathbb{C}^n$. It is easily argued that one can choose n to be equal to the input dimension of the map Γ [14].

III. OPERATOR SCHMIDT DECOMPOSITION

Any vector state $|\psi\rangle_{AB} \in \mathscr{H}_A \otimes \mathscr{H}_B$ admits a Schmidt decomposition [1]

$$|\psi\rangle_{AB} = \sum_{i=1}^{\mathrm{SR}(\psi)} \sqrt{p_i}\,|a_i\rangle_A \otimes |b_i\rangle_B\,, \tag{1}$$

with $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ some special and $|\psi\rangle$-dependent orthonormal bases for $\mathscr{H}_A$ and $\mathscr{H}_B$, respectively, and $\{\sqrt{p_i}\}$ is a collection of positive numbers that satisfy $\sum_{i=1}^{\mathrm{SR}(\psi)}(\sqrt{p_i})^2 = \sum_{i=1}^{\mathrm{SR}(\psi)} p_i = \langle\psi|\psi\rangle$. Since $\langle\psi|\psi\rangle = 1$, we can think of $\{p_i\}$ as of a probability distribution, whose elements we can imagine ordered, $p_1 \geq p_2 \geq \ldots$, without loss of generality. Here $\mathrm{SR}(\psi)$ denotes the Schmidt rank of $|\psi\rangle_{AB}$, which is the number of non-zero p_i's, and satisfies $\mathrm{SR}(\psi) \leq d_{\min}$. Let ρ_{AB} be a density matrix for the bipartite system AB. We can consider it as a vector in $\mathscr{L}(\mathscr{H}_A \otimes \mathscr{H}_B)$, and hence derive the Operator Schmidt Decomposition (see, e.g., [29, 30] and references therein)

$$\rho_{AB} = \sum_{i=1}^{\mathrm{OSR}(\rho)} r_i A_i \otimes B_i. \tag{2}$$

Here $\mathrm{OSR}(\rho)$ is the number of non-zero Operator Schmidt Coefficients (OSCs) r_i, and $\{A_i\}_{i=1}^{d_A^2}$ and $\{B_i\}_{i=1}^{d_B^2}$ are some (ρ-dependent) orthonormal bases for the spaces $\mathscr{L}(\mathscr{H}_A)$ and $\mathscr{L}(\mathscr{H}_B)$, respectively. The OSR is the minimum number of product terms that need to enter in any product decomposition of ρ_{AB}. Since ρ_{AB} is Hermitian, one can argue that the two orthonormal operator bases in (2) can be (but need not be) taken to be composed of Hermitian operators. The OSCs are the singular values of the correlation matrix $[C_{ij}(\rho_{AB})]_{ij}$, with $C_{ij}(\rho_{AB}) := \langle\langle F_i \otimes G_j|\rho_{AB}\rangle\rangle$, where $\{F_i\}$ and $\{G_j\}$ are arbitrary local orthonormal bases for operators. We will take the OSC to be ordered as $r_1 \geq r_2 \geq \ldots$; they satisfy $\sum_i r_i^2 = \langle\langle\rho|\rho\rangle\rangle = \mathrm{Tr}(\rho^2)$. Notice that $\mathrm{OSR}(\rho_{AB}) \leq d_{\min}^2$, as the vector space $L(\mathscr{H}_A)$ has dimension d_A^2 (similarly for $L(\mathscr{H}_B)$). It is immediate to realize that the SD of a pure state $|\psi\rangle_{AB}$ and the OSD of the corresponding density matrix $|\psi\rangle\langle\psi|_{AB}$ are related: indeed, for a pure state, $r_i = \sqrt{p_k}\sqrt{p_l}$, $A_i = |a_k\rangle\langle a_l|$, and $B_i = |b_k\rangle\langle b_l|$, for $i = (k,l)$ a multi-index.

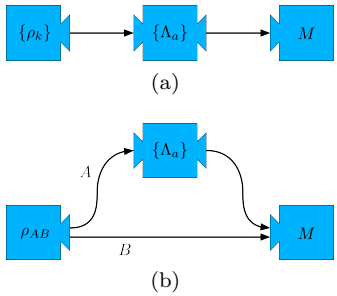

FIG. 1. Two strategies for distinguishing channels. (a) No ancilla is used: a probe undergoes one of many possible quantum evolutions described by channels $\{\Lambda_a\}$, and is later measured (box M). Many different input states $\{\rho_k\}$ are in general needed to discriminate between arbitrary channels, if one cannot tailor the input to the channels. (b) Ancilla-assisted: the probe A is correlated with an ancilla B; the output probe and the ancilla are jointly measured. Depending on the initial probe-ancilla correlations, it might be possible to distinguish between arbitrary evolutions, without modifying the input.

IV. CHANNEL DISCRIMINATION AND CHANNEL TOMOGRAPHY

Channel discrimination is a generalization of state discrimination, where the objects to tell apart are now channels. One can define a physically meaningful notion of distance between two channels Λ_0 and Λ_1 via [14] $D(\Lambda_0, \Lambda_1) := \max_{\rho\in\mathscr{D}(\mathscr{H}_S)} D\left(\Lambda_0[\rho], \Lambda_1[\rho]\right)$, that is by considering the trace distance of the output states of a probe upon acting on the same input state of the probe. One fundamental—and relevant for applications—way in which quantum physics differs from classical physics, is that the distinguishability of two channels, as captured by $D(\Lambda_0, \Lambda_1)$, can be enhanced by the use of entanglement between the input probe and an ancilla [2, 4–15]. One can prove that the best ancilla system can be chosen to be a copy S' of the input probe system S, so that we can define the so-called diamond distance between Λ_0 and Λ_1 as $D_\diamond(\Lambda_0, \Lambda_1) := D(\Lambda_{0,S} \otimes \mathrm{id}_{S'}, \Lambda_{1,S} \otimes \mathrm{id}_{S'})$, where id_X indicates the identity map on system X. The diamond distance formalizes the notion of best possible one-shot distinguishability of two quantum channels.

In general, it is not possible to distinguish arbitrary quantum channels in $\mathscr{T}(\mathscr{H}_X, \mathscr{H}_Y)$ by means of their action on an input state $\rho \in \mathscr{D}(\mathscr{H}_X)$ of the probe alone that is independent of the channels considered [31]. Nonetheless, it is always possible to tell two arbitrary channels in $\mathscr{T}(\mathscr{H}_X, \mathscr{H}_Y)$ apart by 'feeding' them with many different input states ρ_k. As long as $\{\rho_k\}$ constitutes a basis for $\mathscr{L}(\mathscr{H}_X)$, and as long as an arbitrary number of uses of the channel are allowed, one can even perform a tomographic reconstruction of a channel Λ (see Figure 1(a)) [1].

Remarkably, it is possible to perform tomography of the channel, or the non-trivial discrimination of an arbitrary number of channels, even with just a fixed input state, as long as one uses an ancilla: this constitutes the framework of ancilla-assisted channel discrimination and channel tomography (see Figure 1(b)). Ref. [21] proves both theoretically and experimentally that channel tomography is possible also when the state ρ_{AB} of probe A and ancilla B is separable. The key condition that permits channel tomography on A with ρ_{AB} is that $\mathrm{OSR}(\rho_{AB}) = d_A^2$. Indeed, one has $\Lambda_A[\rho_{AB}] = \sum_{i=1}^{\mathrm{OSR}(\rho)} r_i \Lambda[A_i] \otimes B_i$, and, as long as the state has $\mathrm{OSR}(\rho) = d_A^2$, one can reconstruct the action of the map Λ on an arbitrary state $\sigma \in D(\mathscr{H}_A)$ as $\Lambda[\sigma] = \sum_{i=1}^{d_A^2} \frac{1}{r_i} \langle\langle A_i|\sigma\rangle\rangle \,\mathrm{Tr}_B(\mathbb{1}_A \otimes B_{i,B}^\dagger \Lambda_A[\rho_{AB}])$. We improve on this basic observation, by introducing and studying a simple and meaningful measure of merit for the usefulness of a fixed probe-ancilla state in channel discrimination.

V. CHANNEL DISCRIMINATION POWER

For any quantum state $\rho_{AB} \in \mathscr{D}(\mathscr{H}_A \otimes \mathscr{H}_B)$, we define the channel discrimination power (CDP) of ρ_{AB} on A (and similarly on B) as

$$\mathrm{CDP}_A(\rho_{AB}) := \inf_{\Lambda_0,\Lambda_1} \frac{D(\Lambda_{0,A}[\rho_{AB}], \Lambda_{1,A}[\rho_{AB}])}{D_\diamond(\Lambda_0, \Lambda_1).} \tag{3}$$

The infimum is taken over all pairs Λ_0, Λ_1 of quantum channels with input in $\mathscr{L}(\mathscr{H}_A)$, and we have used the notation $\Lambda_{i,A} := \Lambda_{i,A} \otimes \mathrm{id}_B$. The parameter $\mathrm{CDP}_A(\rho_{AB})$ captures how suitable ρ_{AB} is for ancilla-assisted channel discrimination as compared with the optimal distinguishability of those two channels, in a worst-case scenario approach.

In the following we report a number of results about the channel discrimination power. As notation goes, we will indicate the difference of two channels Λ_0 and Λ_1 as $\Delta = \Lambda_0 - \Lambda_1$. The channel discrimination power can then be expressed as

$$\mathrm{CDP}_A(\rho_{AB}) = \inf_{\Delta} \frac{\|\Delta_A \otimes \mathrm{id}_B[\rho_{AB}]\|_1}{\|\Delta\|_\diamond}.$$

VI. BASIC PROPERTIES OF THE CDP

One can easily prove that $\mathrm{CDP}_A(\rho_{AB})$ is continuous in its argument. To show that, first notice that by definition of (Hermitian) super-operator 1-norm we have the following.

Proposition 1. *Let Γ be any Hermiticity preserving map, and X_{AB} Hermitian. Then*

$$\|\Gamma_A \otimes \mathrm{id}_B[X_{AB}]\|_1 \leq \|\Gamma\|_\diamond \|X_{AB}\|_1 .$$

Proposition 2. $\mathrm{CDP}_A(\rho)$ *is continuous:*

$$|\mathrm{CDP}_A(\rho_{AB}) - \mathrm{CDP}_A(\sigma_{AB})| \leq \|\rho_{AB} - \sigma_{AB}\|_1 ,$$

for any two states ρ_{AB} and σ_{AB}.

Proof. Because of the triangle inequality and Proposition 1, one has

$$\begin{aligned}\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1 &= \|\Delta\otimes\mathrm{id}[\sigma_{AB}]+\Delta\otimes\mathrm{id}[\rho_{AB}-\sigma_{AB}]\|_1\\ &\leq \|\Delta\otimes\mathrm{id}[\sigma_{AB}]\|_1+\|\Delta\|_\diamond\|\rho_{AB}-\sigma_{AB}\|_1\,,\end{aligned}$$

that is

$$\frac{\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1-\|\Delta\otimes\mathrm{id}[\sigma_{AB}]\|_1}{\|\Delta\|_\diamond}\leq\|\rho_{AB}-\sigma_{AB}\|_1\,.$$

The thesis follows immediately. □

Proposition 3. $\mathrm{CDP}_A(\rho)$ *is monotone under local channels on* B*:*

$$\mathrm{CDP}_A(\mathrm{id}_A\otimes\Lambda_B[\rho_{AB}])\leq\mathrm{CDP}_A(\rho_{AB}). \tag{4}$$

Proof. This comes directly from the monotonicity of the trace norm of Hermitian operators under channels, i.e. from $\|\Lambda[X]\|_1\leq\|X\|_1$. One has

$$\begin{aligned}&\|\Delta_A\otimes\mathrm{id}_B[\mathrm{id}_A\otimes\Lambda_B[\rho_{AB}]]\|_1\\ &\qquad=\|\mathrm{id}_A\otimes\Lambda_B[\Delta_A\otimes\mathrm{id}_B[\rho_{AB}]]\|_1\\ &\qquad\leq\|\Delta_A\otimes\mathrm{id}_B[\rho_{AB}]\|_1\,,\end{aligned} \tag{5}$$

for any $\Delta=\Lambda_0-\Lambda_1$, and the thesis follows. □

Proposition 4. *The channel discrimination power* CDP_A *is invariant under local unitaries on* A*.*

Proof. For any map Λ on A and any unitary U on A we can consider the map $\Lambda'[\cdot]=\Lambda[U^\dagger\cdot U]$ such that $(\Lambda_A\otimes\mathrm{id}_B)[\rho_{AB}]=(\Lambda'_A\otimes\mathrm{id}_B)[U_A\rho_{AB}U_A^\dagger]$. Given the freedom in the minimization through which CDP_A is defined, the claim follows immediately. □

Notice that Proposition 3 immediately implies that, for fixed dimension of A, the CDP assumes maximal value for pure states, as any bipartite state ρ_{AB} can be seen as the reduced state of a pure state $\psi_{ABB'}$, with B' a purifying system, and BB' considered together as one ancilla. Furthermore, this fact together with Proposition 4 imply that the CDP of a pure state only depends on its Schmidt coefficients.

VII. CDP FOR PURE STATES

We find that for pure states the CDP can be computed exactly. We will need the following lemma, which is a slight generalization of observations in, e.g., Ref. [47].

Lemma 1. *Let* $|\psi\rangle_{AA'}=\sum_{k=1}^d\sqrt{p_k}\,|a_k\rangle_A\otimes|b_k\rangle_{A'}$ *be a pure state with* $d=d_A=d_{A'}$*, and the Schmidt coefficients ordered as* $p_1\geq p_2\geq\ldots\geq p_d$*. Then*

$$p_d\,\|\Delta\|_\diamond\leq\|\Delta\otimes\mathrm{id}[|\psi\rangle\langle\psi|]\|_1\,.$$

Proof. We use the fact that any pure state $|\psi\rangle_{AA'}$ can be expressed as

$$|\psi\rangle_{AA'}=(\mathbb{1}\otimes C)\,|\tilde{\psi}^+\rangle_{AA'}\,, \tag{6}$$

with $|\tilde{\psi}^+\rangle_{AA'}=\sum_{k=1}^d|k\rangle_A\otimes|k\rangle_{A'}$, and $C=\sum_{l=1}^d\sqrt{p_l}\,|b_l\rangle\langle a_l^*|$, where $|a_l^*\rangle$ is the basis state whose coefficients in the basis $|k\rangle$ are the complex conjugates of those of $|a_l\rangle$. Notice that the singular values of C coincide with the Schmidt coefficients of $|\psi\rangle$, and the fact that $|\psi\rangle$ is normalized implies $\|C\|_2=1$, hence $\|C\|_\infty\leq1$.

The claim is trivial if $p_d=0$. If $p_d>0$, then C is invertible, and we can express any other state $|\phi\rangle_{AA'}=(\mathbb{1}\otimes D)\,|\tilde{\psi}^+\rangle_{AA'}$ as

$$|\phi\rangle_{AA'}=(\mathbb{1}\otimes DC^{-1})\,|\psi\rangle_{AA'}\,.$$

Let $|\phi\rangle_{AA'}$ be the state that achieves the diamond norm $\|\Delta\|_\diamond$, that is $\|\Delta\|_\diamond=\|\Delta_A\otimes\mathrm{id}_{A'}[|\phi\rangle\langle\phi|_{AA'}]\|_1$. Then

$$\begin{aligned}&\|\Delta\|_\diamond\\ &=\|\Delta_A\otimes\mathrm{id}_{A'}[|\phi\rangle\langle\phi|_{AA'}]\|_1\\ &=\|(\mathbb{1}\otimes DC^{-1})\big(\Delta_A\otimes\mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\big)(\mathbb{1}\otimes DC^{-1})^\dagger\|_1\\ &\leq\|\mathbb{1}\otimes DC^{-1}\|_\infty^2\|\Delta_A\otimes\mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\|_1\\ &\leq\|D\|_\infty^2\|C^{-1}\|_\infty^2\|\Delta_A\otimes\mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\|_1\\ &=p_d^{-1}\|\Delta_A\otimes\mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\|_1,\end{aligned}$$

where in the first inequality we have used Hölder's inequality, $|\operatorname{Tr}(XY)|\leq\|X\|_\infty\|Y\|_1$, twice. For the last line, just observe that the largest singular value of C^{-1} is the reciprocal of the smallest singular value of C. □

Theorem 1. *Let* $|\psi\rangle_{AB}$ *be a pure state with Schmidt decomposition as in* (1)*. Then, if* $d_{\min}=d_A=d_B$*,* $\mathrm{CDP}_A(\psi_{AB})=\mathrm{CDP}_B(\psi_{AB})=p_{d_{\min}}$*, while, if* $d_{\min}=d_A<d_B$*,* $\mathrm{CDP}_A(\psi_{AB})=p_{d_{\min}}$ *and* $\mathrm{CDP}_B(\psi_{AB})=0$*.*

Notice that it might be that $p_{d_{\min}}=0$, in which case both $\mathrm{CDP}_A(\psi_{AB})$ and $\mathrm{CDP}_B(\psi_{AB})$ vanish. We remark that $p_{d_{\min}}$ is a quantifier of the entanglement of $|\psi\rangle_{AB}$ [33]. Having already established that CDP_A is maximal for pure states, we find that it achieves its maximum, $1/d_A$, for maximally entangled states, e.g., for $|\psi^+\rangle_{AB}=\frac{1}{\sqrt{d_A}}\sum_{i=1}^{d_A}|i\rangle_A\,|i\rangle_B$. Note that it is reasonable that the maximum of the channel discrimination power, being defined as in Eq. (3), decreases with d_A, since the number of parameters describing an arbitrary channel with input in A increases with the size of A.

Proof. (of Theorem 1) Lemma 1 implies immediately $\mathrm{CDP}_A(|\psi\rangle\langle\psi|)\geq p_{d_A}$. We will prove the inequality in the other direction, that is, $\mathrm{CDP}_A(|\psi\rangle\langle\psi|)\leq p_{d_A}$, by constructing a pair of perfectly distinguishable channels that are hard to distinguish by means of $|\psi\rangle$. We observe that, because in the case of pure states CDP_A only depends on

the Schmidt coefficients, we can assume $|a_k\rangle = |b_k\rangle = |k\rangle$, without loss of generality. Let us introduce the channels

$$\begin{aligned}\Lambda_0[X] &= \mathrm{Tr}[PX]\,|2\rangle\langle 2| + \mathrm{Tr}[(\mathbb{1}-P)X]\,|0\rangle\langle 0|\,,\\ \Lambda_1[X] &= \mathrm{Tr}[PX]\,|2\rangle\langle 2| + \mathrm{Tr}[(\mathbb{1}-P)X]\,|1\rangle\langle 1|\,,\end{aligned} \tag{7}$$

with $P = \sum_{i=1}^{d_A-1} |i\rangle\langle i|$ and $\mathbb{1} - P = |d_A\rangle\langle d_A|$. Then, $\Delta[X] = \langle d_A|\,X\,|d_A\rangle\,(|0\rangle\langle 0| - |1\rangle\langle 1|)$. It is clear by their definition that the two channels are perfectly distinguishable, even without the use of an ancilla, since

$$\Lambda_0[|d_A\rangle\langle d_A|] = |0\rangle\langle 0|\,, \quad \Lambda_1[|d_A\rangle\langle d_A|] = |1\rangle\langle 1|\,,$$

so that $\|\Lambda_0 - \Lambda_1\|_\diamond = \|\Lambda_0 - \Lambda_1\|_1 = 2$. On the other hand,

$$\begin{aligned}&\|(\Lambda_0 - \Lambda_1)\otimes \mathrm{id}\,|\psi\rangle\langle\psi|\|_1\\ &\quad= \|(|0\rangle\langle 0| - |1\rangle\langle 1|)\otimes \mathrm{Tr}_A(|d_A\rangle\langle d_A|_A\,|\psi\rangle\langle\psi|_{AB})\|_1\\ &\quad= p_{d_A}\,\|(|0\rangle\langle 0| - |1\rangle\langle 1|)\otimes |d_A\rangle\langle d_A|)\|_1\\ &\quad= 2p_{d_A}.\end{aligned}$$

Thus, we have proven that it must be $\mathrm{CDP}_A(|\psi\rangle\langle\psi|) \leq p_{d_A}$. □

We now show that the CDP attains its maximum for maximally entangles states.

Corollary 1. *The channel discrimination power* CDP_A *is maximal for maximally entangled states, for which it is equal to* $1/d_A$.

Proof. As observed after Proposition 4, the maximum of the channel discrimination power is achieved by pure states. On the other hand, Theorem 1 tells us that the CDP of a pure state is equivalent to the (square) of the smallest Schmidt coefficient. The latter cannot be bigger than $1/d_A$, which is achieved for a maximally entangled state. □

VIII. BOUNDS FOR MIXED STATES

We now present general bounds for the CDP.

Theorem 2. *Let* ρ_{AB} *have an OSD as in Eq.* (2)*, with* $\{A_i\}$, $\{B_i\}$ *Hermitian orthonormal bases for* $\mathscr{L}(\mathscr{H}_A)$ *and* $\mathscr{L}(\mathscr{H}_B)$*, respectively. Then*

$$\frac{r_{d_A^2}}{d_A^{5/2}} \leq \mathrm{CDP}_A(\rho_{AB}) \leq \min_i \left\{ r_i \frac{\|B_i\|_1}{\|A_i\|_\infty}\right\} \leq r_{d_A^2}\sqrt{d_A d_B}. \tag{8}$$

Proof. We first prove $r_{d_A^2}/d_A^{5/2} \leq \mathrm{CDP}_A(\rho_{AB})$.

We start by finding a lower bound for the numerator in the definition of the $\mathrm{CDP}_A(\rho_{AB})$. First, observe that

$$\begin{aligned}&\|\Delta\otimes \mathrm{id}[\rho_{AB}]\|_1\\ &\quad= \left\|\sum_i r_i \Delta(A_i)\otimes B_i\right\|_1\\ &\quad= \max_{-\mathbb{1}\leq M_{AB}\leq \mathbb{1}} \left|\mathrm{Tr}\left(M_{AB}\sum_i r_i\Delta(A_i)\otimes B_i\right)\right|\\ &\quad\geq \max_{\substack{-\mathbb{1}\leq M_A\leq \mathbb{1}\\ -\mathbb{1}\leq M_B\leq \mathbb{1}}} \left|\mathrm{Tr}\left(M_A\otimes M_B\sum_i r_i\Delta(A_i)\otimes B_i\right)\right|\\ &\quad\geq \max_i\left\{r_i\frac{\|\Delta[A_i]\|_1}{\|B_i\|_\infty}\right\}.\\ &\quad\geq r_{d_A^2}\max_i \|\Delta[A_i]\|_1\end{aligned}$$

The first inequality is due to restricting the class of operators M_{AB} to be product. The second inequality is due to further choosing M_A such that $\|\Delta[A_k]\|_1 = |\,\mathrm{Tr}(M_A\Delta[A_k])|$ and $M_B = B_k/\|B_k\|_\infty$, with k the index such that the maximum over i in the last line is achieved. Notice that, because of the orthonormality of the $B_i's$, this choice for M_B selects only one term in the OSD of ρ_{AB}. The last inequality is due to the fact that $\|B_i\|_\infty \leq \|B_i\|_2 = 1$, and that $r_i \geq r_{d_A^2}$ by assumption.

The maximally entangled state can be expressed as $|\psi^+\rangle\langle\psi^+| = \frac{1}{d_A}\sum_{i=1}^{d_A^2} C_i\otimes C_i^*$ for any orthonormal operator basis $\{C_k\}\subset\mathscr{L}(\mathscr{H}_A)$, in particular for the one appearing in the OSD of ρ_{AB}. Thus, using Lemma 1,

$$\begin{aligned}\|\Delta\|_\diamond &\leq d_A\left\|\Delta\otimes \mathrm{id}[|\psi^+\rangle\langle\psi^+|]\right\|_1\\ &= d_A\left\|\frac{1}{d_A}\sum_i \Delta\left[A_i\right]\otimes A_i^*\right\|_1\\ &\leq \sum_i \|\Delta\left[A_i\right]\|_1\|A_i^*\|_1\\ &\leq d_A^{5/2}\max_i\|\Delta[A_i]\|_1\,,\end{aligned}$$

having used the triangle inequality, the fact that there are d_A^2 terms in the sum, and that $\|A_i^*\|_1 = \|A_i\|_1 \leq \sqrt{d_A}\|A_i\|_2 = \sqrt{d_A}$. Thus, combining the above,

$$\mathrm{CDP}_A(\rho_{AB}) = \inf_\Delta \frac{\|\Delta\otimes \mathrm{id}[\rho_{AB}]\|_1}{\|\Delta\|_\diamond} \geq \frac{r_{d_A^2}}{d_A^{5/2}},$$

which completes the first part of the theorem.

We now show how to upper bound the CDP. To do that, let us consider the following channels:

$$\Lambda_i[X] = \mathrm{Tr}(X)\frac{\mathbb{1}}{d_A} + \epsilon\,\mathrm{Tr}(A_l X)Y_i,$$

for $i = 0, 1$, with traceless Hermitian operators Y_0 and Y_1, and where A_l is the local basis operator of the OSD of ρ_{AB} corresponding to the lth OSC r_l. Such

maps are trace-preserving by construction, and completely positive for ϵ small enough, e.g. for $\epsilon \leq 1/(d_A\|A_l\|_\infty\|\max\{\|Y_0\|_\infty, \|Y_1\|_\infty\})$. Then,

$$\Delta[X] = \epsilon\,\mathrm{Tr}(A_l X)(Y_0 - Y_1),$$

and

$$\begin{aligned}\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1 &= \epsilon\left\|\sum_i r_i\,\mathrm{Tr}(A_l A_i)(Y_0-Y_1)\otimes B_i\right\|_1\\ &= \epsilon\,\|r_l(Y_0-Y_1)\otimes B_l\|_1\\ &= r_l\epsilon\,\|Y_0-Y_1\|_1\,\|B_l\|_1\,. \end{aligned}\tag{9}$$

On the other hand, we claim that

$$\|\Delta\|_\diamond = \epsilon\,\|Y_0-Y_1\|_1\,\|A_l\|_\infty\,. \tag{10}$$

Before proving such claim, let us notice that Eqs. (9) and (10) complete the proof of the theorem. Indeed, by recalling the definition of the CDP and using Eqs. (9) and (10), one gets

$$\mathrm{CDP}_A(\rho_{AB}) \leq r_l\frac{\|B_l\|_1}{\|A_l\|_\infty},$$

for any l, that is

$$\mathrm{CDP}_A(\rho_{AB}) \leq \min_i\left\{r_i\frac{\|B_i\|_1}{\|A_i\|_\infty}\right\}.$$

We observe that the right-hand side can be itself upper bounded:

$$\begin{aligned}\min_i\left\{r_i\frac{\|B_i\|_1}{\|A_i\|_\infty}\right\} &\leq r_{d_A^2}\frac{\|B_d\|_1}{\|A_d\|_\infty}\\ &\leq r_{d_A^2}\frac{d_B^{1/2}\,\|B_d\|_2}{d_A^{-1/2}\,\|A_d\|_2}\\ &= r_{d_A^2}(d_A d_B)^{1/2},\end{aligned}$$

where we have used properties of the p-norms in the second inequality.

We now prove Eq. (10). To do so, let us consider an arbitrary

$$\begin{aligned}|\psi\rangle &= \sum_i\sqrt{p_i}\,|a_i\rangle\,|b_i\rangle\\ &= (\mathbb{1}\otimes C)\left|\tilde\psi^+\right\rangle\end{aligned}$$

where $\|C\|_2 = 1$ for $|\psi\rangle$ to be normalized (see the proof of Lemma 1). Notice that

$$\begin{aligned}&\|\Delta\otimes\mathrm{id}[|\psi\rangle\langle\psi|]\|_1\\ &\quad= \|(\mathbb{1}\otimes C)(\Delta\otimes\mathrm{id}[|\tilde\psi^+\rangle\langle\tilde\psi^+|])(\mathbb{1}\otimes C)^\dagger\|_1\\ &\quad= \|\epsilon(Y_0-Y_1)\otimes CA_l^TC^\dagger\|_1\\ &\quad= \epsilon\|Y_0-Y_1\|_1\|CA_l^TC^\dagger\|_1.\end{aligned}$$

Thus, it is sufficient to prove that, for a given $X = X^\dagger$,

$$\max_{\|C\|_2=1}\|CXC^\dagger\|_1 = \|X\|_\infty.$$

Notice that $\|X\|_\infty = \|X^T\|_\infty$.

Let $|x\rangle$ be the eigenvector of X corresponding to the largest eigenvalue (in modulus) $\|X\|_\infty$. Choosing $C = |x\rangle\langle x|$ we have $\|CXC^\dagger\|_1 = \|\,|x\rangle\langle x|\,X\,|x\rangle\langle x|\,\|_1 = \|X\|_\infty$, thus $\max_{\|C\|_2=1}\|CXC^\dagger\|_1 \geq \|X\|_\infty$.

To prove the other direction it is enough to prove that

$$\|CXC^\dagger\|_1 \leq \|X\|_\infty\,\mathrm{Tr}(C^\dagger C) = \|X\|_\infty\|C\|_2^2 = \|X\|_\infty,$$

for $X = X^\dagger$ and C satisfying $\|C\|_2 = 1$. The inequality can be seen as a trivial consequence of the fact that, for any vector $|\psi\rangle$, one has

$$\begin{aligned}|\,\langle\psi|\,CXC^\dagger\,|\psi\rangle\,| &= |\,\langle\psi|\,C(X_+ - X_-)C^\dagger\,|\psi\rangle\,|\\ &\leq \langle\psi|\,C(X_+ + X_-)C^\dagger\,|\psi\rangle\\ &\leq \|X\|_\infty\,\langle\psi|\,CC^\dagger\,|\psi\rangle\,,\end{aligned}$$

where we have used that any Hermitian matrix can be expressed as the difference of two positive semidefinite matrices with orthogonal support,

$$X = X^+ - X^-,$$

with $X^\pm \geq 0$, $X^+X^- = X^-X^+ = 0$, and that

$$X^+ + X^- \leq \|X\|_\infty\mathbb{1}.$$

We have proved the claim in Eq. (10), hence the theorem.

□

The bounds above are not tight in general, as proven by the results about pure states. Nonetheless, they capture quantitatively, rather than purely qualitatively, the fact that the necessary and sufficient condition for ρ_{AB} to always enable ancilla-assisted discrimination and tomography of an arbitrary channel on A is that $\mathrm{OSR}(\rho) = d_A^2$.

IX. BOUND FOR SEPARABLE STATES

We recall that mixed unentangled states may have maximal OSR, that is $\mathrm{OSR}(\rho_{AB}) = d_A^2$, so that, according to Eq. (8), they have non-zero CDP. This is the case, for example, of isotropic states, considered more in detail below.

We now focus on the case $d_A = d_B = d$. As we have seen, CDP can be as high as $1/d$. We prove that such a value cannot be achieved by states passing the realignment (or computable cross-norm) criterion for separability, i.e., such that its OSCs satisfy $\sum_i r_i \leq 1$, in the same way those of a separable state necessarily do [24, 25]. The proof makes use of the following bounds, which characterize the correlations present in a state in terms of its purity, and may be of independent interest.

Lemma 2. *For any ρ_{AB} and any product state $\sigma_A \otimes \sigma_B$, one has $\sum_{i\geq 2} r_i^2(\rho_{AB}) = \mathrm{Tr}(\rho^2) - r_1^2 \leq \|\rho_{AB} - \sigma_A \otimes \sigma_B\|_2^2$.*

Proof. We recall that the OSCs $r_i(\rho_{AB})$ are the singular values of the correlation matrix $[C_{ij}(\rho_{AB})]_{ij}$, with

$$C_{ij}(\rho_{AB}) := \langle\langle F_i \otimes G_j | \rho_{AB} \rangle\rangle,$$

where $\{F_i\}$ and $\{G_j\}$ are arbitrary local orthonormal bases for operators. We will use that, for any two matrices M and N, with ordered singular values $\sigma_i(M)$ and $\sigma_i(N)$, respectively, it holds (see Corollary 7.3.5 in [27]),

$$\sum_i (\sigma_i(M) - \sigma_i(N))^2 \leq \|M - N\|_2^2.$$

Notice that $r_i(\sigma_A \otimes \sigma_B) = 0$, for $i \geq 2$. Thus,

$$\begin{aligned}
\sum_{i\geq 2} r_i^2(\rho_{AB}) &= \sum_{i\geq 2} (r_i(\rho_{AB}) - r_i(\sigma_A \otimes \sigma_B))^2 \\
&\leq \sum_i (r_i(\rho_{AB}) - r_i(\sigma_A \otimes \sigma_B))^2 \\
&\leq \|C(\rho_{AB}) - C(\sigma_A \otimes \sigma_B)\|_2^2 \\
&= \|C(\rho_{AB} - \sigma_A \otimes \sigma_B)\|_2^2 \\
&= \|\rho_{AB} - \sigma_A \otimes \sigma_B\|_2^2,
\end{aligned}$$

having used that $\|C(X)\|_2 = \|X\|_2$ for any X. □

Proposition 5. *For any state ρ_{AB} on $\mathbb{C}^d \otimes \mathbb{C}^d$, the smallest operator Schmidt coefficient obeys*

$$r_{d^2} \leq \sqrt{\mathrm{Tr}(\rho^2) - \frac{1}{d^2}}.$$

Proof. Immediate, by using Lemma 2 in the case $\sigma_A \otimes \sigma_B = \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d}$, and the fact that

$$\begin{aligned}
\left\| \rho_{AB} - \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d} \right\|_2^2 &= \mathrm{Tr}\left(\left(\rho_{AB} - \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d} \right)^2 \right) \\
&= \mathrm{Tr}(\rho^2) - \frac{1}{d^2}.
\end{aligned}$$

□

Applying these bounds, we obtain the following.

Theorem 3. *If the OSCs of ρ_{AB} satisfy $\sum_i r_i \leq 1$, then $r_d \leq r_{\mathrm{CN}}$ with*

$$r_{\mathrm{CN}} = \frac{d(d^2-1) - \sqrt{d^2-1}}{d(d^2-1)^2 + d^3} < \frac{1}{d^2}.$$

Proof. We want to find the maximal value r_{d^2} can assume under the condition

$$\sum_i r_i \leq 1. \tag{11}$$

We notice that Proposition 5 implies that the OSCs of every state respect

$$r_{d^2}^2 \leq \sum_i r_i^2 - \frac{1}{d^2} \tag{12}$$

(recall that $\mathrm{Tr}(\rho^2) = \sum_i r_i^2$). We want aim to find the maximum of r_{d^2} under conditions (11) and (12), irrespectively of the physicality of the choice coefficient—as long as they respect (11) and (12). Notice that, by definition, $r_i \geq 0$, and $r_1 \geq r_2 \geq \ldots \geq r_d^2$.

It it clear that the maximum r_{d^2} will be found for the condition (11) being satisfied with equality, since, if the left-hand side of (11) was smaller than 1, then we could increase all the OSCs, including r_{d^2}, to make it equal to 1. Moreover, for fixed r_{d^2}, the largest value of $\sum_i r_i^2$ is achieved for $r_2 = r_3 = \ldots = r_{d^2} = r$ and $r_1 = 1 - r$. This is due to the fact that $\sum_i r_i^2$ is Schur convex. Thus, we can find the maximal r_{d^2} compatible with the constraints, by finding the largest r such that

$$r^2 \leq (d^2-1)r^2 + (1 - (d^2-1)r)^2 - \frac{1}{d^2}.$$

One finds that such a value is given by

$$r_{CN} = \frac{d(d^2-1) - \sqrt{d^2-1}}{d(d^2-1)^2 + d^3} < \frac{1}{d^2}.$$

□

By combining Theorem 3 with Theorem 2 we prove that, if the OSCs of ρ_{AB} satisfy $\sum_i r_i \leq 1$, then $\mathrm{CDP}(\rho_{AB}) \leq r_{\mathrm{CN}} d < 1/d$. We remark that the realignment criterion for separability is satisfied by all separable states, and by many (weakly) entangled states [24, 25, 28].

X. RELATION WITH DISCORD

As we have just seen, entanglement is needed to achieve the maximal possible CDP. Nonetheless, separable states can have non-vanishing CDP, when they have maximal OSR. As pointed out in Ref. [34], this is not possible for states that do not exhibit quantum discord. A bipartite state is classical on A if it can be expressed as $\rho_{AB} = \sum_i p_i |a_i\rangle\langle a_i|_A \otimes \rho_i^B$, for some orthonormal basis $\{|a_i\rangle\}$, and manifestly has OSR $\leq d_A$. States that are not classical on A are said to possess quantum discord [26, 35, 36] and may be detected as discordant by looking at their OSR [34, 37]. All entangled states necessarily possess discord, but also unentangled states can. Discord plays a basic role in quantum information processing, being linked, e.g., to the impossibility of local broadcasting of correlations and information [38], to quantum data hiding [39], to quantum data locking [40], to entanglement distribution [41, 42], to quantum cryptography [44], to quantum metrology [43]. Here we shed light on the role of discord in the latter. To do so, it will

be convenient to first study the behaviour of the CDP under the action of maps that reduce the OSR.

Theorem 4. *We have*

$$\mathrm{CDP}_A(\rho_{AB}) \leq \min_{\substack{\Lambda \ s.t. \\ \mathrm{OSR}(\Lambda\otimes\mathrm{id}[\rho_{AB}])<d_A^2}} \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1$$

where the minimization is over all channels that acting on A reduce the OSR of ρ_{AB} to less than maximal.

Proof. It holds

$$\begin{aligned}\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1 &\leq \|\Delta\otimes\mathrm{id}[\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]]\|_1 \\ &\quad + \|(\Delta\circ\Lambda)\otimes\mathrm{id}[\rho_{AB}]\|_1 \\ &\leq \|\Delta\|_\diamond \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1 \\ &\quad + \|\Delta\otimes\mathrm{id}[\Lambda\otimes\mathrm{id}[\rho_{AB}]]\|_1 ,\end{aligned}$$

having used Proposition 1. Then,

$$\begin{aligned}\inf_\Delta \frac{\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1}{\|\Delta\|_\diamond} &\leq \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1 \\ &\quad + \inf_\Delta \frac{\|\Delta\otimes\mathrm{id}[\Lambda\otimes\mathrm{id}[\rho_{AB}]]\|_1}{\|\Delta\|_\diamond} \\ &= \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1 ,\end{aligned}$$

where we have used that the CDP of $\Lambda\otimes\mathrm{id}[\rho_{AB}]$ (the second term on the right-hand side of the inequality) vanishes under the assumption $\mathrm{OSR}\left(\Lambda\otimes\mathrm{id}\left[\rho_{AB}\right]\right) < d_A^2$. The claim then follows. □

As a particular example involving the last theorem, let $\Pi[X] = \sum_{i=1}^d |i\rangle\langle i| X |i\rangle\langle i|$ be the channel which dephases in an arbitrary basis. Then

$$\mathrm{CDP}_A(\rho_{AB}) \ \leq \min_{\Pi_A\otimes\mathrm{id}_B} \|\rho_{AB} - \Pi_A\otimes\mathrm{id}_B[\rho_{AB}]\|_1 .$$

In the light of the last theorem, we find that

$$\begin{aligned}\mathrm{CDP}_A(\rho_{AB}) &\leq \min_{\substack{\Lambda_A \text{ s.t.} \\ OSR(\Lambda_A[\rho_{AB}])<d_A^2}} 2D(\rho_{AB}, \Lambda_A[\rho_{AB}]) \\ &\leq \min_{\Pi_A} 2D(\rho_{AB}, \Pi_A[\rho_{AB}]).\end{aligned}$$

The first minimization is over channels that reduce the OSR of ρ_{AB} to less than maximal. The second minimization is over projective measurements $\Pi[L] = \sum_i |a_i\rangle\langle a_i| L |a_i\rangle\langle a_i|$, for a choice of basis $\{|a_i\rangle\}$ to be optimized over. The quantity on the second line is a known geometric discord quantifier [45]. Thus, we see that the bipartite state ρ_{AB} must be contain a large amount of discord in order for ρ_{AB} to be useful in one-shot, worst-case ancilla-assisted channel discrimination.

XI. EXAMPLES

As an example that goes beyond pure states, we consider the class of isotropic states [46]

$$\rho_{\mathrm{iso}}(p) = (1-p)\frac{\mathbb{1}_{AB}}{d^2} + p\left|\psi^+\right\rangle\left\langle\psi^+\right|_{AB} , \qquad (13)$$

where $0 \leq p \leq 1$ and $|\psi^+\rangle$ is the standard maximally entangled state. This is a paradigmatic class of noisy states that interpolates between an uncorrelated state (for $p = 0$) and a maximally entangled state (for $p = 1$). Isotropic states are separable for $0 \leq p \leq \frac{1}{d+1}$ and entangled for $\frac{1}{d+1} < p \leq 1$. This is also the class of states used in Ref. [21] in the context of ancilla-assisted channel tomography, where it was already observed that this class of states enables channel tomography as soon as $p > 0$. It is known and immediate to check that

$$\left|\psi^+\right\rangle\left\langle\psi^+\right| = \frac{1}{d}\sum_{k=1}^{d^2} A_k\otimes A_k^*$$

for any orthonormal operator basis $\{A_k\}$, with complex conjugation taken in the local Schmidt basis of the maximally entangled state. We can choose $A_1 = \frac{\mathbb{1}}{\sqrt{d}}$, and find immediately

$$\rho_{\mathrm{iso}}(p) = \frac{1}{d}\frac{\mathbb{1}}{\sqrt{d}}\otimes\frac{\mathbb{1}}{\sqrt{d}} + \frac{p}{d}\sum_{k=2}^{d^2} A_k\otimes A_k^*, \qquad (14)$$

where $\{A_k\}$ is any collection of d^2-1 traceless orthonormal operators. Thus, the OSCs of $\rho_{\mathrm{iso}}(p)$ are evidently $(1/d, p/d, \ldots, p/d)$. Notice that $r_{d^2} = p/d$, so that the general bounds (8) become $p/d^{7/2} \leq \mathrm{CDP}(\rho_{\mathrm{iso}}(p)) \leq p$; we are able to prove the following bounds, which reproduce the correct value for CDP in the limit in which the isotropic states become maximally entangled.

Theorem 5. *For the isotropic state it holds*

$$\frac{p}{d+1-p} \leq \mathrm{CDP}_A(\rho_{iso}(p)) \leq \min\left\{2\frac{p}{d}, \frac{1}{d}\right\}. \qquad (15)$$

Proof. We start by proving the upper bound. That $\mathrm{CDP}_A(\rho_{\mathrm{iso}}(p)) \leq 1/d$ can be straightforwardly be verified by using the same two maps (7) that were used to prove the upper bound for pure states. In order to prove $\mathrm{CDP}_A(\rho_{\mathrm{iso}}(p)) \leq 2p/d$, we will use the bound $\mathrm{CDP}_A(\rho_{AB}) \leq \min_i \left\{r_i \frac{\|B_i\|_1}{\|A_i\|_\infty}\right\}$ from Theorem 2, exploiting the freedom in choosing the decomposition (14). E.g., we can choose $A_2 = (|1\rangle\langle 2| + |2\rangle\langle 1|)/\sqrt{2}$, with $B_2 = A_2^* = A_2$, so that $\|A_2\|_\infty = 1/\sqrt{2}$ and $\|B_2\|_1 = \sqrt{2}$. Thus,

$$\mathrm{CDP}_A(\rho_{\mathrm{iso}}(p)) \leq r_2\frac{\|B_2\|_1}{\|A_2\|_\infty} = \frac{p}{d}\frac{\|A_2\|_1}{\|A_2\|_\infty} = \frac{p}{d}2.$$

For the lower bound, we generalize the approach of Lemma 1.

Given two arbitrary channels, let $|\psi\rangle\langle\psi|$ be optimal for the diamond norm of their difference, i.e.

$$\|\Delta\|_\diamond = \sup_\rho \|\Delta \otimes \mathrm{id}[\rho]\|_1 = \|\Delta \otimes \mathrm{id}[|\psi\rangle\langle\psi|]\|_1$$

and let us consider C such that

$$|\psi\rangle_{AA'} = (\mathbb{1} \otimes C) \left|\tilde{\psi}^+\right\rangle_{AA'}$$

Notice that $\mathrm{Tr}_A(|\psi\rangle\langle\psi|) = CC^\dagger$, with $CC^\dagger \geq 0$ a normalized state.

Let us define the state

$$\begin{aligned}\sigma(p) &:= (1-p)\frac{\mathbb{1}}{d} \otimes CC^\dagger + p\,|\psi\rangle\langle\psi| \\ &= d(\mathbb{1} \otimes C)\left[(1-p)\frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d} + p|\psi^+\rangle\langle\psi^+|\right](\mathbb{1} \otimes C^\dagger) \\ &= d(\mathbb{1} \otimes C)\ \rho_{\mathrm{iso}}(p)\ (\mathbb{1} \otimes C^\dagger).\end{aligned}$$

Then,

$$|\psi\rangle\langle\psi| = \frac{1}{p}\left[\sigma(p) - (1-p)\frac{\mathbb{1}}{d} \otimes CC^\dagger\right],$$

and

$$\begin{aligned}\|\Delta\|_\diamond &= \|\Delta \otimes \mathrm{id}[|\psi\rangle\langle\psi|]\|_1 \\ &= \left\|\frac{1}{p}\left[\Delta \otimes \mathrm{id}[\sigma(p)] - (1-p)\Delta\left[\frac{\mathbb{1}}{d}\right] \otimes CC^\dagger\right]\right\|_1 \\ &\leq \frac{1}{p}\|\Delta \otimes \mathrm{id}[\sigma(p)]\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 \\ &= \frac{d}{p}\left\|(\mathbb{1} \otimes C)\ \Delta \otimes \mathrm{id}[\rho_{\mathrm{iso}}(p)]\ (\mathbb{1} \otimes C^\dagger)\right\|_1 \\ &\quad + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 \\ &\leq \frac{d}{p}\|C\|_\infty^2\,\|\Delta \otimes \mathrm{id}[\rho_{\mathrm{iso}}(p)]\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 \\ &\leq \frac{d}{p}\|\Delta \otimes \mathrm{id}[\rho_{\mathrm{iso}}(p)\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1.\end{aligned}$$

Finally, since $\frac{\mathbb{1}}{d} = \mathrm{Tr}_B(\rho_{\mathrm{iso}}(p))$ and the partial trace is a channel, the monotonicity of the trace distance implies

$$\begin{aligned}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 &= \|\Delta_A\left[\mathrm{Tr}_B(\rho_{\mathrm{iso}}(p))\right]\|_1 \\ &= \|\mathrm{Tr}_B\left(\Delta_A[\rho_{\mathrm{iso}}(p)]\right)\|_1 \\ &\leq \|\Delta \otimes \mathrm{id}[\rho_{\mathrm{iso}}(p)]\|_1.\end{aligned}$$

Thus,

$$\|\Delta\|_\diamond \leq \left(\frac{d+1-p}{p}\right)\|\Delta \otimes \mathrm{id}[\rho_{\mathrm{iso}}(p)]\|_1,$$

from which we obtain

$$\mathrm{CDP}_A(\rho_{\mathrm{iso}}(p)) \geq \frac{p}{d+1-p}.$$

□

XII. CONCLUSIONS

Quantum correlations [26, 28] play an important role in several areas of physics, including quantum foundations, condensed-matter physics, quantum information processing, and quantum technologies. In particular, quantum correlations can be exploited in quantum metrology [2, 3]. We focused on the usefulness of quantum correlations for ancilla-assisted channel discrimination with fixed input, introducing the channel discrimination power (CDP) of the input state. We argued that the key factor that dictates the CDP of a state is its smallest operator Schmidt coefficient. We proved that the CDP is maximal for maximally entangled states. This can be seen as an argument to consider the Choi-Jamiołkowski isomorphism [16, 17] as the best possible one-to-one mapping between states and maps. We derived general bounds for the CDP that allowed us to prove that highly entangled states outperform—in the sense of having a larger CDP—all states that pass the so-called realignment criterion of separability [24, 25]. We added to the list of quantum information processing tasks for which the quantum discord provides a bound on the performance: we proved that a disturbance-based discord quantifier bounds the CDP. Several questions remain open, like whether the CDP is equal to the lowest operator-Schmidt-coefficient of the state, and which channels are the hardest to discriminate for a given input state. Finally, while the CDP is defined in terms of optimal probability of channel discrimination, it would be interesting to consider more in general how a probe-ancilla state induces a mapping between a metric on the space of channels and a metric in the space of output probe-ancilla states.

ACKNOWLEDGMENTS

We thank Vern Paulsen for correspondence and John Watrous for discussions. We acknowledge financial support from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Action OPERACQC (Grant Agreement No. 661338), and from the Foundational Questions Institute under the Physics of the Observer Programme (Grant number FQXi-RFP-1601).

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).
[2] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
[3] G. Tóth and I. Apellaniz, Journal of Physics A: Mathematical and Theoretical **47**, 424006 (2014).
[4] G. M. D'Ariano, P. LoPresti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).
[5] A. Y. Kitaev, Russian Mathematical Surveys **52**, 1191 (1997).
[6] A. M. Childs, J. Preskill, and J. Renes, Journal of modern optics **47**, 155 (2000).
[7] A. Acín, Physical review letters **87**, 177901 (2001).
[8] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Physical Review A **71**, 062310 (2005).
[9] B. Rosgen and J. Watrous, in *Computational Complexity, 2005. Proceedings. Twentieth Annual IEEE Conference on* (IEEE, 2005) pp. 344–354.
[10] M. F. Sacchi, Physical Review A **71**, 062340 (2005).
[11] M. F. Sacchi, Physical Review A **72**, 014305 (2005).
[12] S. Lloyd, Science **321**, 1463 (2008).
[13] B. Rosgen, Journal of Mathematical Physics **49**, 102107 (2008).
[14] J. Watrous, Quantum Info. Comput. **5**, 58 (2005).
[15] J. Watrous, Quantum Information & Computation **8**, 819 (2008).
[16] M.-D. Choi, Linear algebra and its applications **10**, 285 (1975).
[17] A. Jamiołkowski, Reports on Mathematical Physics **3**, 275 (1972).
[18] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Physical review letters **101**, 060401 (2008).
[19] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Nature Communications **8**, 15043 EP (2017).
[20] O. Oreshkov, F. Costa, and C. Brukner, arXiv preprint arXiv:1105.4464 (2011).
[21] J. B. Altepeter, D. Branning, E. Jeffrey, T. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, Physical Review Letters **90**, 193601 (2003).
[22] G.M. D'Ariano , P. Lo Presti, Phys. Rev. Lett. **91**, 047902 (2003).
[23] A. Jenčová and M. Plávala, Journal of Mathematical Physics **57**, 122203 (2016), http://dx.doi.org/10.1063/1.4972286.
[24] K. Chen and L.-A. Wu, Quantum Inf. Comput **3**, 193 (2003).
[25] O. Rudolph, Letters in Mathematical Physics **70**, 57 (2004).
[26] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, Rev. Mod. Phys. **84**, 1655 (2012).
[27] R. A. Horn and C. R. Johnson, *Matrix analysis* (Cambridge University Press, 2013).
[28] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[29] P. Aniello and C. Lupo, Open Systems & Information Dynamics **16**, 127 (2009).
[30] C. Lupo, P. Aniello, and A. Scardicchio, Journal of Physics A: Mathematical and Theoretical **41**, 415301 (2008).
[31] For example, consider the case where Λ_0 is the identity channel, so that $\Lambda_0[\sigma] = \sigma$ for all σ, and Λ_1 is the channel with fixed output ρ. Then, obviously, $D(\Lambda_0[\rho], \Lambda_1[\rho]) = 0$, even if the two channels are very different, and even having many copies of $\Lambda_i[\rho]$ we cannot tell the two channels apart.
[32] See Supplementary material, which references also [47] and [48] for proofs.
[33] G. Vidal, Physical Review Letters **83**, 1046 (1999).
[34] B. Dakić, V. Vedral, and Č. Brukner, Physical review letters **105**, 190502 (2010).
[35] L. Henderson and V. Vedral, J. Phys. A: Math. Gen. **34**, 6899 (2001).
[36] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2001).
[37] B. Lanyon, P. Jurcevic, C. Hempel, M. Gessner, V. Vedral, R. Blatt, and C. Roos, Physical review letters **111**, 100504 (2013).
[38] M. Piani, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **100**, 090502 (2008).
[39] M. Piani, V. Narasimhachar, and J. Calsamiglia, New J. Phys. **16**, 113001 (2014).
[40] S. Boixo, L. Aolita, D. Cavalcanti, K. Modi, and A. Winter, International Journal of Quantum Information **9**, 1643 (2011).
[41] T. Chuan, J. Maillard, K. Modi, T. Paterek, M. Paternostro, and M. Piani, Phys. Rev. Lett. **109**, 070501 (2012).
[42] A. Streltsov, H. Kampermann, and D. Bruß, Phys. Rev. Lett. **108**, 250501 (2012).
[43] D. Girolami, A. M. Souza, V. Giovannetti, T. Tufarelli, J. G. Filgueiras, R. S. Sarthour, D. O. Soares-Pinto, I. S. Oliveira, and G. Adesso, Physical Review Letters **112**, 210401 (2014).
[44] S. Pirandola, Scientific reports **4** (2014).
[45] S. Luo, Phys. Rev. A **77**, 022301 (2008).
[46] M. Horodecki and P. Horodecki, Physical Review A **59**, 4206 (1999).
[47] F. G. Brandão, M. Piani, and P. Horodecki, Nature communications **6**, 7908 (2015).
[48] A. Uhlmann, Reports on Mathematical Physics **9**, 273 (1976).