All entangled states are useful for channel discrimination

Marco Piani[1] and John Watrous[2]

[1]*Institute for Quantum Computing & Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada*
[2]*Institute for Quantum Computing & School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada*

We prove that every entangled state is useful as a resource for the problem of minimum-error channel discrimination. More specifically, given a single copy of an arbitrary bipartite entangled state, it holds that there is an instance of a quantum channel discrimination task for which this state allows for a correct discrimination with strictly higher probability than every separable state.

PACS numbers: 03.67.Mn, 03.67.Bg, 03.65.Ud

Despite its sometimes counter-intuitive properties, entanglement has firmly been established as a fundamental resource at the core of quantum information theory. Universal quantum computation is generally believed to be impossible in its absence [1], and it plays a principal role in quantum teleportation [2], superdense coding [3], and the one-way model of quantum computation [4]. The classification of entanglement into different types, depending on its usefulness and properties as a resource, is a major focus in the theory of quantum information. For example, *distillable entanglement* [5] may be processed by means of local operations and classical communication into a nearly pure form that is suitable for high fidelity quantum teleportation, while *bound entanglement* cannot [6]. Other classifications of entangled states, such as those that allow or do not allow superdense coding [7, 8], and those from which private shared-randomness can be extracted [9], have also been studied.

Although entanglement is known to be useful in several quantum information-theoretic settings, there are very few known results that establish the usefulness of *every* entangled state, irrespective of the "quality" of its entanglement and of the dimensionality of its underlying systems. The only prior examples that we are aware of involve a type of activation mechanism, where the usefulness of a given entangled state is based on the joint properties of a composite system when it is paired with another entangled state of a special type. For example, in [10] it was proved that for any entangled state, there exist another entangled state such that the fidelity of conclusive teleportation [11] of the latter is enhanced by the presence of the former. A different property holding for all entangled states that has a similar character was proved in [12].

In this Letter we demonstrate a new way in which every entangled state is useful as a resource: for the task of *channel discrimination*. In this task, two known discrete physical processes (or channels) are fixed, and access to one of them is made available—but it is not known which one it is, and only a single application of the channel is possible. The goal is to determine, with minimal probability of error, which of the two channels was given, assuming for simplicity that the two channels were equally likely. The most general approach to solving an instance of this problem is to prepare a (possibly entangled) bipartite quantum state, to apply the given channel to one part of this state, and finally to measure the resulting state by a POVM with two outcomes that correspond to predictions of which channel was given.

It is well-known that entanglement is sometimes useful for channel discrimination. This phenomenon seems to have been identified first by Kitaev [13], who introduced the *diamond norm* on super-operators to deal with precisely this phenomenon in the context of quantum error correction and fault-tolerance [36]. Subsequent work [14–24] by several researchers further illuminated the usefulness of entanglement in the problem of channel discrimination and related tasks. In these works, the focus has mainly been on identifying classes of channel pairs for which some optimally chosen entangled state either does or does not give an advantage over every possible separable (or nonentangled) state.

In this Letter we reverse this question and suppose that some *arbitrary* entangled state is given, and ask whether the entanglement in this state is useful for channel discrimination. We prove that every bipartite entangled state indeed does provide an advantage for this task, in that there necessarily exists an instance of a channel discrimination problem for which the entangled state allows for a correct discrimination with strictly higher probability than every possible separable state. This holds even for a single copy of the entangled state, regardless of its dimensionality or the quality or type of its entanglement (including, for instance, bound entangled states), and does not require the presence of an auxiliary state that it serves to activate. This fact is proved below after brief discussions of notation, terminology, and background information on the problem of channel discrimination.

Notation and terminology. For a given (finite dimensional) Hilbert space $\mathcal{X}$, the set of linear operators taking the form $A : \mathcal{X} \to \mathcal{X}$ is denoted by $\mathrm{L}(\mathcal{X})$. An operator $\rho \in \mathrm{L}(\mathcal{X})$ is a *density operator*, and represents a *state*, if it is positive semidefinite ($\rho \geq 0$) and has unit trace

($\mathrm{Tr}(\rho) = 1$). The set of such density operators is denoted $\mathrm{D}(\mathcal{X})$.

A state $\sigma^{\mathrm{sep}} \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$ of a bipartite system is said to be *separable* if it takes the form

$$\sigma^{\mathrm{sep}} = \sum_i p_i \, \sigma^i_{\mathcal{X}} \otimes \sigma^i_{\mathcal{Z}} \tag{1}$$

for density operators $\{\sigma^i_{\mathcal{X}}\}$ and $\{\sigma^i_{\mathcal{Z}}\}$ on the Hilbert spaces $\mathcal{X}$ and $\mathcal{Z}$, respectively, and otherwise is *entangled*. The set of all separable states of the above form (1) is denoted $\mathrm{Sep}(\mathcal{X} : \mathcal{Z})$.

The *trace norm* of an operator A is defined as $\|A\|_{\mathrm{tr}} \equiv \mathrm{Tr}\sqrt{A^\dagger A}$ [37]. The *trace distance* between two states ρ_0 and ρ_1 is $\|\rho_0 - \rho_1\|_{\mathrm{tr}}$.

Channels are particular elements of the set of linear super-operators $\mathrm{T}(\mathcal{X}, \mathcal{Y}) \equiv \{\Phi | \Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})\}$ that map operators on a Hilbert space $\mathcal{X}$ into operators on a (possibly different) Hilbert space $\mathcal{Y}$. A super-operator $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is said to be:

- *Hermiticity-preserving* if $\Phi[X]$ is Hermitian for every Hermitian operator X; or equivalently if $\Phi[X]^\dagger = \Phi[X^\dagger]$ for every operator X;
- *trace-preserving* if $\mathrm{Tr}(\Phi[X]) = \mathrm{Tr}(X)$ for every operator X;
- *trace-annihilating* if $\mathrm{Tr}(\Phi[X]) = 0$ for every operator X;
- *positive* if $\Phi[X] \geq 0$ for every positive semidefinite operator $X \geq 0$;
- *completely positive* if $\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}$ is positive for every Hilbert space $\mathcal{Z}$, where $\mathbb{1}_{\mathrm{L}(\mathcal{Z})}$ denotes the identity super-operator on $\mathrm{L}(\mathcal{Z})$;
- a *channel* if it is both completely positive and trace-preserving;
- an *entanglement-breaking channel* if it is a channel that destroys all entanglement: $\left(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho_{\mathcal{X}\mathcal{Z}}] \in \mathrm{Sep}(\mathcal{Y} : \mathcal{Z})$ for all states $\rho_{\mathcal{X}\mathcal{Z}}$.

A channel describes any physical process which preserves probability, i.e., that happens with certainty.

The *Choi-Jamołkowski representation* [25, 26] of a super-operator $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is given by

$$J(\Phi) = \sum_{1 \leq i,j \leq d_{\mathcal{X}}} \Phi[|i\rangle\langle j|] \otimes |i\rangle\langle j| \; \in \; \mathrm{L}(\mathcal{Y} \otimes \mathcal{X}),$$

where $d_{\mathcal{X}}$ and $\{|1\rangle, \ldots, |d_{\mathcal{X}}\rangle\}$ are the dimension and a fixed orthonormal basis of $\mathcal{X}$, respectively. The mapping $J : \mathrm{T}(\mathcal{X}, \mathcal{Y}) \to \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$ is a linear bijection, which implies that for every operator $A \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$ there exists a unique super-operator $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ such that $J(\Phi) = A$. It holds that a super-operator $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is:

- *Hermiticity-preserving* if and only if $J(\Phi)^\dagger = J(\Phi)$ [27];
- *trace-preserving* if and only if $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$;
- *trace-annihilating* if and only if $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = 0$;
- *completely positive* if and only if $J(\Phi) \geq 0$ [25, 26];
- an *entanglement-breaking channel* if and only if it is a channel and $J(\Phi)/d_{\mathcal{X}} \in \mathrm{Sep}(\mathcal{Y} \otimes \mathcal{X})$ [28].

State and channel discrimination. The task of channel discrimination is naturally related to the well-studied task of discriminating states [29]. Suppose we are given one of two known states $\rho_0, \rho_1 \in D(\mathcal{X})$, and our goal is to guess which one it is with minimal error probability. A guessing procedure for this task may be described by a two-outcome POVM $\{M_0, M_1\} \subset \mathrm{L}(\mathcal{X})$, $M_0, M_1 \geq 0$, $M_0 + M_1 = \mathbb{1}_{\mathcal{X}}$. The error probability for such a measurement can be expressed as $p_E = 1/2(1 - 1/2\,\mathrm{Tr}[(M_0 - M_1)(\rho_0 - \rho_1)])$, and as quantum states are not perfectly distinguishable in general, this probability of error may be nonzero for every possible measurement. By optimizing over all choices of the measurement one reaches the minimum error probability $p_E = 1/2(1 - 1/2\|\rho_0 - \rho_1\|_{\mathrm{tr}})$ [38].

Now, suppose we want to discriminate two channels $\Phi_0, \Phi_1 \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ with minimal error probability, as discussed above. By "probing" whichever channel was given with a state $\rho \in \mathrm{D}(\mathcal{X})$, we transform the problem into one of discriminating between the states $\Phi_0[\rho]$ and $\Phi_1[\rho]$. Thus, the relevant quantity becomes $\|\Phi_0[\rho] - \Phi_1[\rho]\|_{\mathrm{tr}}$, and the minimal error will be achieved by choosing an optimal input state that minimizes this quantity. In this way we are led to consider the trace distance [39] of two channels $\|\Phi_0 - \Phi_1\|_{\mathrm{tr}} \equiv \max_\rho \|\Phi_0[\rho] - \Phi_1[\rho]\|_{\mathrm{tr}}$. By the convexity of the trace norm, this maximum will be achieved for some pure input state.

As mentioned previously, however, the reduction from channel to state discrimination just described may not always be optimal, for it does not exploited the possibility of feeding the channel with a subsystem of a larger correlated system, and then measuring the resulting output joint system. More precisely, we may consider an input state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, with $\mathcal{Z}$ the Hilbert space of an arbitrary ancillary system, and compare the output states $\left(\Phi_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho]$, for $i = 0, 1$. Thus, the ultimate quantity relevant in minimal-error channel discrimination is actually the diamond norm

$$\|\Phi_0 - \Phi_1\|_\diamond \equiv \sup_{n \geq 1} \|\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_n)} - \Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_n)}\|_{\mathrm{tr}},$$

where $\mathcal{Z}_n$ denotes a Hilbert space with dimension n [40].

By definition, it holds that $\|\Phi_0 - \Phi_1\|_\diamond \geq \|\Phi_0 - \Phi_1\|_{\mathrm{tr}}$, and if it is the case that $\|\Phi_0 - \Phi_1\|_\diamond > \|\Phi_0 - \Phi_1\|_{\mathrm{tr}}$, then it is necessarily because of entanglement. This is due to the fact that the correlations of separable states never helps in the discrimination of channels, as we have

$$\begin{aligned}\left\|\left(\Phi_0\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\sigma^{\mathrm{sep}}]-\left(\Phi_1\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\sigma^{\mathrm{sep}}]\right\|_{\mathrm{tr}}\\ \le \sum_i p_i\|\Phi_0[\sigma^i_{\mathcal{X}}]\otimes\sigma^i_{\mathcal{Z}}-\Phi_1[\sigma^i_{\mathcal{X}}]\otimes\sigma^i_{\mathcal{Z}}\|_{\mathrm{tr}}\\ = \sum_i p_i\|\Phi_0[\sigma^i_{\mathcal{X}}]-\Phi_1[\sigma^i_{\mathcal{X}}]\|_{\mathrm{tr}} \le \|\Phi_0-\Phi_1\|_{\mathrm{tr}}\end{aligned}$$

for every separable state σ^{sep}.

Proof of the main result. To establish our main result, we will connect the characterization of entanglement in terms of positive linear maps with its usefulness for channel discrimination. We will first prove two lemmas, the first being a simplification of Lemma 1 in [30], followed by the main theorem.

Lemma 1. *A state $\rho\in\mathrm{D}(\mathcal{X}\otimes\mathcal{Z})$ is entangled if and only if there exists a positive, trace-preserving super-operator $\Phi\in\mathrm{T}(\mathcal{X},\mathcal{Y})$ such that*

$$\left(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho]\not\geq 0. \tag{2}$$

It suffices to take $\dim\mathcal{Y}\le\dim\mathcal{Z}+1$.

Proof. In [31] it was proved that a state $\rho\in\mathrm{D}(\mathcal{X}\otimes\mathcal{Z})$ is entangled if and only if there exists a positive super-operator $\Omega\in\mathrm{T}(\mathcal{X},\mathcal{Z})$ such that $(\Omega\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]\not\geq 0$. The main issue that must be addressed is that the super-operator Ω may not, in general, be trace-preserving.

Let us define $\lambda(\Omega)=\max_\rho\mathrm{Tr}(\Omega[\rho])$, where the maximum is over all density operators $\rho\in\mathrm{D}(\mathcal{X})$, and consider the normalized map $\hat\Omega\equiv\Omega/\lambda(\Omega)$. By construction, this super-operator satisfies $\mathrm{Tr}(X)\ge\mathrm{Tr}(\hat\Omega[X])$ for all $X\ge 0$, and so the map $\hat\Omega_{\mathrm{TP}}\in\mathrm{T}(\mathcal{X},\mathcal{Z}\oplus\mathbb{C})$ defined as $\hat\Omega_{\mathrm{TP}}[X]=\hat\Omega[X]+(\mathrm{Tr}(X)-\mathrm{Tr}(\hat\Omega[X]))\,|0\rangle\langle 0|$, where $|0\rangle$ is a normalized vector orthogonal to $\mathcal{Z}$, is also positive and satisfies $(\hat\Omega_{\mathrm{TP}}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]\not\geq 0$. Moreover $\hat\Omega_{\mathrm{TP}}$ is trace-preserving, so that by taking $\Phi=\hat\Omega_{\mathrm{TP}}$ and $\mathcal{Y}=\mathcal{Z}\oplus\mathbb{C}$ the proof is complete. □

Remark 1. Any positive and trace-preserving super-operator Φ that satisfies the condition (2) must necessarily satisfy the following properties:

1. $\|(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]\|_{\mathrm{tr}}=1+2\sum_{i:r_i<0}|r_i|>1$, where r_i are the eigenvalues of $(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]$;
2. $\left(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\sigma^{\mathrm{sep}}]\ge 0$, and $\|(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\sigma]\|_{\mathrm{tr}}=1$ for every separable state $\sigma\in\mathrm{Sep}(\mathcal{X}:\mathcal{Z})$.

Lemma 2. *Let $\Phi\in\mathrm{T}(\mathcal{X},\mathcal{Y})$ be a Hermiticity preserving, trace-annihilating super-operator. Then there exist channels $\Psi_0,\Psi_1\in\mathrm{T}(\mathcal{X},\mathcal{Y})$ and a scalar $c_\Phi>0$ such that $c_\Phi\Phi=\Psi_0-\Psi_1$.*

Proof. Given that Φ is Hermiticity-preserving and trace-annihilating, it holds that its Choi-Jamiołkowski representation $J(\Phi)$ is Hermitian and satisfies $\mathrm{Tr}_{\mathcal{Y}}J(\Phi)=0$. Let $J(\Phi)=P_0-P_1$ be a Jordan decomposition of $J(\Phi)$ (meaning that $P_0,P_1\ge 0$ and $\mathrm{Tr}(P_0P_1)=0$), and note that $\mathrm{Tr}_{\mathcal{Y}}P_0=\mathrm{Tr}_{\mathcal{Y}}P_1=Q\ge 0$. Take $c_\Phi=1/\|Q\|$, so that $c_\Phi Q\le\mathbb{1}_{\mathcal{X}}$. Next, consider any positive operator $\xi\in L(\mathcal{Y}\otimes\mathcal{X})$ such that $\mathrm{Tr}_{\mathcal{Y}}\xi_{\mathcal{Y}\mathcal{X}}=\mathbb{1}_{\mathrm{L}(\mathcal{X})}-c_\Phi Q$ [41], and let $\Psi_0,\Psi_1\in\mathrm{T}(\mathcal{X},\mathcal{Y})$ be the unique super-operators for which $J(\Psi_i)=c_\Phi P_i+\xi$ for $i=0,1$. We have $J(\Psi_i)\ge 0$ and $\mathrm{Tr}_{\mathcal{Y}}(J(\Psi_i))=c_\Phi Q+\mathbb{1}_{\mathrm{L}(\mathcal{X})}-c_\Psi Q=\mathbb{1}_{\mathcal{X}}$, therefore Ψ_0,Ψ_1 are channels. Moreover, $J(\Psi_0)-J(\Psi_1)=c_\Phi(P_0-P_1)=c_\Phi J(\Phi)$, therefore $\Psi_0-\Psi_1=c_\Phi\Phi$. □

Theorem 1. *A state $\rho\in\mathrm{D}(\mathcal{X}\otimes\mathcal{Z})$ is entangled if and only if there exist channels $\Psi_0,\Psi_1\in\mathrm{T}(\mathcal{X},\mathcal{Y})$ such that*

$$\left\|\left(\Psi_0\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho]-\left(\Psi_1\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho]\right\|_{\mathrm{tr}}>\|\Psi_0-\Psi_1\|_{\mathrm{tr}}.$$

It suffices to take $\dim\mathcal{Y}\le\dim\mathcal{Z}+2$.

Proof. We have already argued that if ρ allows, for some choice of channels Ψ_0,Ψ_1, a discrimination better than the one corresponding to $\|\Psi_0-\Psi_1\|_{\mathrm{tr}}$, then ρ must be entangled. On the other hand, if ρ is entangled, then by Lemma 1 there exists a positive, trace-preserving super-operator $\Phi\in\mathrm{T}(\mathcal{X},\mathcal{W})$ such that $\|(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]\|_{\mathrm{tr}}=1+2\sum_{i:r_i<0}|r_i|>1$, where r_i's are the eigenvalues of $(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]$. Let us define a new map $\Phi_{\mathrm{TA}}\in\mathrm{T}(\mathcal{X},\mathcal{W}\oplus\mathbb{C})$ as $\Phi_{\mathrm{TA}}[X]=\Phi[X]-\mathrm{Tr}(X)|0\rangle\langle 0|$, where $|0\rangle$ is a normalized vector orthogonal to $\mathcal{W}$. By construction, Φ_{TA} is Hermiticity-preserving and trace-annihilating. By Lemma 2, there exists a scalar $c_{\Phi_{\mathrm{TA}}}$ such that $c_{\Phi_{\mathrm{TA}}}\Phi_{\mathrm{TA}}=\Psi_0-\Psi_1$ for two channels $\Psi_0,\Psi_1\in\mathrm{T}(\mathcal{X},\mathcal{W}\oplus\mathbb{C})$.

Now, for a generic state $\tau\in\mathrm{D}(\mathcal{X}\otimes\mathcal{Z})$, one finds

$$\begin{aligned}&\|\left(\Psi_0\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\tau]-\left(\Psi_1\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\tau]\|_{\mathrm{tr}}\\&=c_{\Phi_{\mathrm{TA}}}\|(\Phi_{\mathrm{TA}}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\tau]\|_{\mathrm{tr}}\\&=c_{\Phi_{\mathrm{TA}}}\|(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\tau]-|0\rangle\langle 0|\otimes\mathrm{Tr}_{\mathcal{X}}(\tau)\|_{\mathrm{tr}}\\&=c_{\Phi_{\mathrm{TA}}}(1+\|(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\tau]\|_{\mathrm{tr}})\\&=2c_{\Phi_{\mathrm{TA}}}(1+\sum_{i:t_i<0}|t_i|),\end{aligned}$$

where t_i are the eigenvalues of $(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\tau]$. For every separable state $\tau=\sigma^{\mathrm{sep}}\in\mathrm{Sep}(\mathcal{X}:\mathcal{Z})$ we obtain

$$\|(\Psi_0\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\sigma^{\mathrm{sep}}]-(\Psi_1\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\sigma^{\mathrm{sep}}]\|_{\mathrm{tr}}=2c_{\Phi_{\mathrm{TA}}},$$

so that $\|\Psi_0-\Psi_1\|_{\mathrm{tr}}=2c_{\Phi_{\mathrm{TA}}}$. Thus,

$$\|(\Psi_0\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]-(\Psi_1\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]\|_{\mathrm{tr}}-\|\Psi_0-\Psi_1\|_{\mathrm{tr}}=2c_{\Phi_{\mathrm{TA}}}\sum_{i:r_i<0}|r_i|>0. \tag{3}$$

According to Lemma 1 it is sufficient to have $\dim\mathcal{W}\le\dim\mathcal{Z}+1$. Taking $\mathcal{Y}=\mathcal{W}\oplus\mathbb{C}$ shows that it is sufficient to have $\dim\mathcal{Y}\le\dim\mathcal{Z}+2$, and completes the proof. □

In regard to the type of channels that allow entangled states to give improved discrimination, one has the following interesting corollary.

Corollary 1. *A state* $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$ *is entangled if and only if there exist entanglement-breaking channels* $\Psi_0, \Psi_1 \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ *such that*

$$\left\| \left(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho] - \left(\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho] \right\|_{\mathrm{tr}} > \|\Psi_0 - \Psi_1\|_{\mathrm{tr}}.$$

Proof. Generalizing the result of [21], we observe that if an entangled state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$ increases the distinguishability of two channels $\Psi_0, \Psi_1 \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then it also increases the distinguishability of two entanglement breaking channels of the form $\Xi_i^p = p\Psi_i + (1-p)\Omega$, for $i = 0, 1$. Here $p \in [0,1]$ and $\Omega \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is the totally depolarizing channel $\Omega[X] = (\mathrm{Tr}(X)/d_{\mathcal{Y}})\mathbb{1}_{\mathcal{Y}}$.

For sufficiently small $p > 0$, the channels Ξ_i^p are entanglement breaking, as their Choi-Jamiołkowski representations are separable by the existence of a ball of positive radius containing only separable states around the maximally mixed state [32]. It holds that

$$\begin{aligned}\left\| \left(\Xi_0^p \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho] - \left(\Xi_1^p \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right)[\rho] \right\|_{\mathrm{tr}} \\ = p \|(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho] - (\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})[\rho]\|_{\mathrm{tr}}\end{aligned}$$

and $\|\Xi_0^p - \Xi_1^p\|_{\mathrm{tr}} = p\|\Psi_0 - \Psi_1\|_{\mathrm{tr}}$, therefore the state ρ enhances the distinguishability of these channels for all choices of $p > 0$. □

Example. The steps in the proof of Theorem 1 are constructive. In particular, while the value of the enhancement in distinguishability depends on the particular state, the channels that are better distinguished by means of the states depend exclusively on the map.

The most well-known example of a positive linear map that detects entanglement is transposition $T : L(\mathcal{X}) \ni X \mapsto X^T \in L(\mathcal{X})$ with respect to some fixed basis of $\mathcal{X}$ [31, 33]. For transposition one finds $c_{T_{\mathrm{TA}}} = 2/(d_{\mathcal{X}}+1)$, and channels $\Psi_0^T, \Psi_1^T \in T(\mathcal{X}, \mathcal{X} \oplus \mathbb{C})$, $\Psi_0^T : X \mapsto \frac{1}{d_{\mathcal{X}}+1}((\mathrm{Tr}\, X)\mathbb{1} + X^T)$, $\Psi_1^T : X \mapsto \frac{1}{d_{\mathcal{X}}+1}((\mathrm{Tr}\, X)(\mathbb{1} + 2|0\rangle\langle 0|) + X^T)$, with $|0\rangle$ a normalized vector orthogonal to $\mathcal{X}$. Thus, for any state $\rho \in D(\mathcal{X} \otimes \mathcal{Z})$, we obtain $\|\Psi_0^T \otimes \mathbb{1}_{L(\mathcal{Z})}[\rho] - \Psi_1^T \otimes \mathbb{1}_{L(\mathcal{Z})}[\rho]\|_{\mathrm{tr}} - \|\Psi_0^T - \Psi_1^T\|_{\mathrm{tr}} = \frac{4}{d_{\mathcal{X}}+1} N(\rho)$, with $N(\rho) = \frac{\|T \otimes \mathbb{1}(\rho)\|_{\mathrm{tr}} - 1}{2}$ the negativity of ρ [32, 34].

Conclusions. We have proved that any entangled state is useful to distinguish some pair of channels strictly better than what is possible by means of a separable state in the minimum-error, single-shot scenario. This pair of channels may be taken to be arbitrarily noisy and able to destroy the entanglement of any arbitrary strongly entangled state. One may consider this result as a physically meaningful interpretation of the characterization of entangled states by means of positive but not completely positive linear maps [31].

MP thanks C. E. Mora, P. Stelmachovic and M. Ziman for inspiring discussions, and acknowledges support from NSERC and Ontario Centres of Excellence. JW acknowledges support from NSERC and the Canadian Institute for Advanced Research (CIFAR).

---

[1] R. Jozsa and N. Linden, Proceedings of the Royal Society A **459**, 2011 (2003).
[2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Physical Review Letters **70**, 1895 (1993).
[3] C. H. Bennett and S. J. Wiesner, Physical Review Letters **69**, 2881 (1992).
[4] R. Raussendorf and H. J. Briegel, Physical Review Letters **86**, 5188 (2001).
[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Physical Review Letters **76**, 722 (1996).
[6] M. Horodecki, P. Horodecki, and R. Horodecki, Physical Review Letters **80**, 5239 (1998).
[7] D. Bruss, G. M. D'Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen, Physical Review Letters **93**, 210501 (2004).
[8] M. Horodecki and M. Piani, *Constraints on many-to-one multipartite dense coding*, arXiv.org e-Print quant-ph/0701134 (2007).
[9] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Physical Review Letters **94**, 160502 (2005).
[10] L. Masanes, Physical Review Letters **96**, 150501 (2006).
[11] P. Horodecki, M. Horodecki, and R. Horodecki, Physical Review Letters **82**, 1056 (1999).
[12] L. Masanes, Y.-C. Liang, and A. C. Doherty, Physical Review Letters **100**, 090403 (2008).
[13] A. Kitaev, Russian Mathematical Surveys **52**, 1191 (1997).
[14] A. Childs, J. Preskill, and J. Renes, Journal of Modern Optics **47**, 155 (2000).
[15] G. M. D'Ariano, P. LoPresti, and M. G. A. Paris, Physical Review Letters **87**, 270404 (2001).
[16] A. Acín, Physical Review Letters **87**, 177901 (2001).
[17] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
[18] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Physical Review A **71**, 062310 (2005).
[19] B. Rosgen and J. Watrous, in *Proceedings of the 20th Annual Conference on Computational Complexity* (2005), pp. 344–354.
[20] M. F. Sacchi, Physical Review A **71**, 062340 (2005).
[21] M. F. Sacchi, Physical Review A **72**, 014305 (2005).
[22] S. Lloyd, Science **321**, 1463 (2008).
[23] B. Rosgen, in *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science* (2008), pp. 597–608.
[24] J. Watrous, Quantum Information and Computation **8**, 819 (2008).
[25] A. Jamiołkowski, Reports on Mathematical Physics **3**, 275 (1972).
[26] M.-D. Choi, Linear Algebra and Its Applications **10**, 285 (1975).
[27] J. de Pillis, Pacific Journal of Mathematics **23**, 129 (1967).
[28] M. Horodecki, P. Shor, and M. Ruskai, Reviews in Mathematical Physics **15**, 629 (2003).
[29] C. Helstrom, Journal of Statistical Physics **1**, 231 (1969).
[30] M. Horodecki, P. Horodecki, and R. Horodecki, Open Systems and Information Dynamics **13**, 103 (2006).
[31] M. Horodecki, P. Horodecki, and R. Horodecki, Physics

Letters A **223**, 1 (1996).

[32] K. Zyczkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Physical Review A **58**, 883 (1998).

[33] A. Peres, Physical Review Letters **77**, 1413 (1996).

[34] G. Vidal and R. F. Werner, Physical Review A **65**, 032314 (2002).

[35] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, 2002).

[36] The diamond norm turns out to be essentially equivalent to a norm known as the *completely bounded norm*, which is an object of study in the theory of operator algebras [35].

[37] When $A = \sum_i a_i |i\rangle\langle i|$ is Hermitian, the trace norm coincides with the sum of the absolute values of the eigenvalues $\|A\|_{\mathrm{tr}} = \sum_i |a_i|$.

[38] This is achieved by choosing M_0 and M_1 to be the projectors on the positive and negative subspaces of $\rho_0 - \rho_1$.

[39] This norm is different from the super-operator norm that is induced by the trace norm, which is sometimes also denoted $\|\cdot\|_{\mathrm{tr}}$.

[40] The supremum is always achieved for $n \leq \dim(\mathcal{X})$.

[41] One possible canonical choice, uniquely defined up to an isometry on $\mathcal{Y}$, is a not normalized purification $\xi = |\xi\rangle\langle\xi|$, $|\xi\rangle \in \mathcal{Y} \otimes \mathcal{X}$.