

All Entangled States are Useful for Channel Discrimination

Marco Piani¹ and John Watrous²

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*

(Received 20 January 2009; published 24 June 2009)

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.

DOI: [10.1103/PhysRevLett.102.250501](https://doi.org/10.1103/PhysRevLett.102.250501)

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

Despite its sometimes counterintuitive 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 with respect to 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 its pairing with another entangled state. For example, in [10] it was proved that for any entangled state, there exists 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 that 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 probe-ancilla quantum state, to apply the given channel to one part of this state—the probe—and finally to measure the resulting bipartite state by a POVM with two outcomes that correspond to predictions of which channel was given.

It is well-known that probe-ancilla entanglement is sometimes useful for channel discrimination. This phenomenon seems to have been identified first by Kitaev [13], who introduced the diamond norm on superoperators to deal with precisely this phenomenon in the context of quantum error correction and fault tolerance [14]. Subsequent work [16–26] 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: 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. 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 $A: \mathcal{X} \rightarrow \mathcal{X}$ is denoted by $L(\mathcal{X})$. We will denote by $\mathbb{1}_Z$ the identity operator on Z and by $\mathbb{1}_{L(Z)}$ the identity superoperator on $L(Z)$. An operator $\rho \in L(\mathcal{X})$ is a density

operator, and represents a state, if it is positive semidefinite ($\rho \geq 0$) and has unit trace ($\text{Tr}(\rho) = 1$). The set of such density operators is denoted $D(\mathcal{X})$. A state $\sigma \in D(\mathcal{X} \otimes \mathcal{Z})$ of a bipartite system is said to be separable if it takes the form $\sigma = \sigma^{\text{sep}} = \sum_i p_i \sigma_{\mathcal{X}}^i \otimes \sigma_{\mathcal{Z}}^i$ for density operators $\{\sigma_{\mathcal{X}}^i\}$ and $\{\sigma_{\mathcal{Z}}^i\}$ on the Hilbert spaces \mathcal{X} and \mathcal{Z} , respectively, and $\{p_i\}$ a probability distribution, and otherwise is entangled. The set of all separable states is denoted $\text{Sep}(\mathcal{X}:\mathcal{Z})$. The trace norm of an operator A is defined as $\|A\|_{\text{tr}} \equiv \text{Tr}\sqrt{A^\dagger A}$ [27]. The trace distance between two states ρ_0 and ρ_1 is $\|\rho_0 - \rho_1\|_{\text{tr}}$.

Channels are particular elements of the set of linear superoperators $T(\mathcal{X}, \mathcal{Y}) \equiv \{\Phi | \Phi: L(\mathcal{X}) \rightarrow L(\mathcal{Y})\}$ that map operators on a Hilbert space \mathcal{X} into operators on a (possibly different) Hilbert space \mathcal{Y} . A superoperator $\Phi \in T(\mathcal{X}, \mathcal{Y})$ is said to be (i) *Hermiticity preserving* if $\Phi[X]^\dagger = \Phi[X^\dagger]$, $\forall X \in L(\mathcal{X})$; (ii) *trace preserving* if $\text{Tr}(\Phi[X]) = \text{Tr}(X)$, $\forall X \in L(\mathcal{X})$; (iii) *trace annihilating* if $\text{Tr}(\Phi[X]) = 0$, $\forall X \in L(\mathcal{X})$; (iv) *positive* if $\Phi[X] \geq 0$ for every positive semidefinite operator $L(\mathcal{X}) \ni X \geq 0$; (v) *completely positive* if $\Phi \otimes \mathbb{1}_{L(\mathbb{C}^n)}$ is positive for all n ; (vi) a *channel* if it is both completely positive and trace preserving; (vii) an *entanglement-breaking channel* if it is a channel that destroys all entanglement: $(\Phi \otimes \mathbb{1}_{L(\mathcal{Z})}) \times [\rho_{\mathcal{X}\mathcal{Z}}] \in \text{Sep}(\mathcal{Y}:\mathcal{Z})$ 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-Jamiołkowski representation [28,29] of a superoperator $\Phi \in T(\mathcal{X}, \mathcal{Y})$ is given by

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

where $d_{\mathcal{X}}$ and $\{|1\rangle, \dots, |d_{\mathcal{X}}\rangle\}$ are the dimension and a fixed orthonormal basis of \mathcal{X} , respectively. The mapping $J: T(\mathcal{X}, \mathcal{Y}) \rightarrow L(\mathcal{Y} \otimes \mathcal{X})$ is a linear bijection, which implies that for every operator $A \in L(\mathcal{Y} \otimes \mathcal{X})$ there exists a unique superoperator $\Phi \in T(\mathcal{X}, \mathcal{Y})$ such that $J(\Phi) = A$. It holds that a superoperator $\Phi \in T(\mathcal{X}, \mathcal{Y})$ is (i) Hermiticity preserving if and only if $J(\Phi)^\dagger = J(\Phi)$ [30]; (ii) trace preserving if and only if $\text{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$; (iii) trace annihilating if and only if $\text{Tr}_{\mathcal{Y}}(J(\Phi)) = 0$; (iv) completely positive if and only if $J(\Phi) \geq 0$ [28,29]; (v) an entanglement-breaking channel if and only if it is a channel and $J(\Phi)/d_{\mathcal{X}} \in \text{Sep}(\mathcal{Y} \otimes \mathcal{X})$ [31].

State and channel discrimination.—The task of channel discrimination is naturally related to the well-studied task of discriminating states [32]. Suppose we are given one of two known states $\rho_0, \rho_1 \in D(\mathcal{X})$, each with equal *a priori* probability, 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 $\{M_0, M_1\} \subset L(\mathcal{X})$, $M_0, M_1 \geq 0$, $M_0 + M_1 = \mathbb{1}_{\mathcal{X}}$. The error probability for such a measurement can be expressed as $p_E = 1/2(1 - 1/2 \text{Tr}[(M_0 - M_1)(\rho_0 - \rho_1)])$. It may be nonzero for every possible measurement, but by optimizing the measurement one reaches the minimum-error probability $p_E^{\text{min}} = 1/2(1 - 1/2 \|\rho_0 - \rho_1\|_{\text{tr}})$ [33].

Now, suppose we want to discriminate two channels $\Phi_0, \Phi_1 \in T(\mathcal{X}, \mathcal{Y})$ with minimal-error probability, as discussed above. By “probing” whichever channel was given with a state $\rho \in D(\mathcal{X})$, we transform the problem into one of discriminating between the states $\Phi_0[\rho]$ and $\Phi_1[\rho]$. Thus, the relevant quantity becomes $\|\Phi_0[\rho] - \Phi_1[\rho]\|_{\text{tr}}$, 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 [34] of two channels $\|\Phi_0 - \Phi_1\|_{\text{tr}} \equiv \max_{\rho} \|\Phi_0[\rho] - \Phi_1[\rho]\|_{\text{tr}}$. 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 exploit 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 $\rho \in D(\mathcal{X} \otimes \mathcal{Z})$, with \mathcal{Z} the Hilbert space of an arbitrary ancillary system, and compare the output states $(\Phi_i \otimes \mathbb{1}_{L(\mathcal{Z})})[\rho]$, for $i = 0, 1$. Thus, the ultimate quantity relevant in minimal-error channel discrimination is actually the diamond norm [35]:

$$\|\Phi_0 - \Phi_1\|_{\diamond} \equiv \sup_{n \geq 1} \|\Phi_0 \otimes \mathbb{1}_{L(\mathbb{C}^n)} - \Phi_1 \otimes \mathbb{1}_{L(\mathbb{C}^n)}\|_{\text{tr}}.$$

By definition, it holds that $\|\Phi_0 - \Phi_1\|_{\diamond} \geq \|\Phi_0 - \Phi_1\|_{\text{tr}}$, and if it is the case that $\|\Phi_0 - \Phi_1\|_{\diamond} > \|\Phi_0 - \Phi_1\|_{\text{tr}}$, then it is necessarily because of entanglement. Indeed, the correlations of separable states never help in the discrimination of channels, as for every separable state $\sigma^{\text{sep}} \in \text{Sep}(\mathcal{X} \otimes \mathcal{Z})$ we have

$$\|(\Phi_0 \otimes \mathbb{1}_{L(\mathcal{Z})})[\sigma^{\text{sep}}] - (\Phi_1 \otimes \mathbb{1}_{L(\mathcal{Z})})[\sigma^{\text{sep}}]\|_{\text{tr}} \leq \sum_i p_i \|(\Phi_0 - \Phi_1)[\sigma_{\mathcal{X}}^i] \otimes \sigma_{\mathcal{Z}}^i\|_{\text{tr}} = \sum_i p_i \|\Phi_0[\sigma_{\mathcal{X}}^i] - \Phi_1[\sigma_{\mathcal{X}}^i]\|_{\text{tr}} \leq \|\Phi_0 - \Phi_1\|_{\text{tr}}.$$

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 begin with a lemma that can be considered an improvement of Lemma 1 in [36]: the well-known characterization of entanglement by positive maps proved in [37] continues to hold if the extra constraint of trace-

preservation is placed on the positive maps. The improvement of the following lemma lies in a significantly simpler proof and in a better bound on the output dimension of the positive maps.

Lemma 1.—A state $\rho \in D(\mathcal{X} \otimes \mathcal{Z})$ is entangled if and only if there exists a positive, trace-preserving superoperator $\Phi_{\text{TP}} \in T(\mathcal{X}, \mathcal{Y})$ such that

$$(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)}[\rho]) \neq 0. \quad (1)$$

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

Proof.—In [37] it was proved that a state $\rho \in D(\mathcal{X} \otimes Z)$ is entangled if and only if there exists a positive superoperator $\Omega \in T(\mathcal{X}, Z)$ such that $(\Omega \otimes \mathbb{1}_{L(Z)}[\rho]) \neq 0$. The main issue that must be addressed is that the superoperator Ω may not, in general, be trace preserving.

Let us define $\lambda(\Omega) \equiv \max_{\rho} \text{Tr}(\Omega[\rho])$, where the maximum is over all density operators $\rho \in D(\mathcal{X})$, and consider the normalized map $\hat{\Omega} \equiv \Omega/\lambda(\Omega)$. By construction, this superoperator satisfies $\text{Tr}(X) \geq \text{Tr}(\hat{\Omega}[X])$ for all $X \geq 0$, and so the map $\Phi_{\text{TP}} \in T(\mathcal{X}, Z \oplus \mathbb{C})$ defined as $\Phi_{\text{TP}}[X] \equiv \hat{\Omega}[X] + (\text{Tr}(X) - \text{Tr}(\hat{\Omega}[X]))|0\rangle\langle 0|$, where $|0\rangle$ is a normalized vector orthogonal to Z , is also positive and satisfies $(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)}[\rho]) \neq 0$. By taking $\mathcal{Y} = Z \oplus \mathbb{C}$ and noticing that Φ_{TP} is trace-preserving, the proof is complete. ■

It is helpful to note at this point that any positive and trace-preserving superoperator $\Phi_{\text{TP}} \in T(\mathcal{X}, \mathcal{Y})$ allows one to define, for every state $\rho \in D(\mathcal{X} \otimes Z)$, a generalized negativity [38,39] parameter as [40]

$$N_{\Phi_{\text{TP}}}(\rho) \equiv \frac{\|(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)}[\rho])\|_{\text{tr}} - 1}{2} = \sum_{i: r_i < 0} |r_i|,$$

where $\{r_i\}$ is the set of eigenvalues of $(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)}[\rho])$. Of course, $(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)}[\sigma^{\text{sep}}]) \geq 0$ and $N_{\Phi_{\text{TP}}}(\sigma^{\text{sep}}) = 0$, for every separable state $\sigma^{\text{sep}} \in \text{Sep}(\mathcal{X}:Z)$, while $N_{\Phi_{\text{TP}}}(\rho) > 0$ if ρ is entangled and detected as in (1).

Next we will prove a lemma that relates a Hermiticity-preserving, trace-annihilating superoperator—an apparently abstract object—to the existence of two channels.

Lemma 2.—Let $\Phi_{\text{TA}} \in T(\mathcal{X}, \mathcal{Y})$ be a Hermiticity-preserving, trace-annihilating superoperator. Then there exist channels $\Psi_0, \Psi_1 \in T(\mathcal{X}, \mathcal{Y})$ and a scalar $c_{\Phi_{\text{TA}}} > 0$ such that $c_{\Phi_{\text{TA}}} \Phi_{\text{TA}} = \Psi_0 - \Psi_1$.

$$\begin{aligned} \|((\Psi_0 - \Psi_1) \otimes \mathbb{1}_{L(Z)})[\tau]\|_{\text{tr}} &= c_{\Phi_{\text{TA}}} \|(\Phi_{\text{TA}} \otimes \mathbb{1}_{L(Z)})[\tau]\|_{\text{tr}} = c_{\Phi_{\text{TA}}} \|(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)})[\tau] - |0\rangle\langle 0| \otimes \text{Tr}_{\mathcal{X}}(\tau)\|_{\text{tr}} \\ &= c_{\Phi_{\text{TA}}} (1 + \|(\Phi_{\text{TP}} \otimes \mathbb{1}_{L(Z)})[\tau]\|_{\text{tr}}) = 2c_{\Phi_{\text{TA}}} (1 + N_{\Phi_{\text{TP}}}(\tau)). \end{aligned}$$

For every separable state $\sigma^{\text{sep}} \in \text{Sep}(\mathcal{X}:Z)$ we obtain $\|((\Psi_0 - \Psi_1) \otimes \mathbb{1}_{L(Z)})[\sigma^{\text{sep}}]\|_{\text{tr}} = 2c_{\Phi_{\text{TA}}}$, and therefore $\|\Psi_0 - \Psi_1\|_{\text{tr}} = 2c_{\Phi_{\text{TA}}}$. Thus,

$$\|((\Psi_0 - \Psi_1) \otimes \mathbb{1}_{L(Z)}[\rho])\|_{\text{tr}} - \|\Psi_0 - \Psi_1\|_{\text{tr}} = 2c_{\Phi_{\text{TA}}} N_{\Phi_{\text{TP}}}(\rho) > 0.$$

According to Lemma 1 it is sufficient to have $\dim \mathcal{W} \leq \dim Z + 1$. Taking $\mathcal{Y} = \mathcal{W} \oplus \mathbb{C}$ shows that it is sufficient to have $\dim \mathcal{Y} \leq \dim Z + 2$, 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 $\rho \in D(\mathcal{X} \otimes Z)$ is entangled if and only if there exist entanglement-breaking channels $\Psi_0, \Psi_1 \in T(\mathcal{X}, \mathcal{Y})$ such that

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

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

We are now ready for the proof of the main theorem, which will rely on the careful definition of a trace-annihilating map—starting from a trace-preserving map as in Lemma 1—and on the application of Lemma 2.

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

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

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

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 $\|\Psi_0 - \Psi_1\|_{\text{tr}}$, then ρ must be entangled. On the other hand, if ρ is entangled, then by Lemma 1 there exists a positive, trace-preserving superoperator $\Phi_{\text{TP}} \in T(\mathcal{X}, \mathcal{W})$ such that $N_{\Phi_{\text{TP}}}(\rho) > 0$. Let us define a new map $\Phi_{\text{TA}} \in T(\mathcal{X}, \mathcal{W} \oplus \mathbb{C})$ as $\Phi_{\text{TA}}[X] \equiv \Phi_{\text{TP}}[X] - \text{Tr}(X)|0\rangle\langle 0|$, 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 $c_{\Phi_{\text{TA}}}$ such that $c_{\Phi_{\text{TA}}} \Phi_{\text{TA}} = \Psi_0 - \Psi_1$ for two channels $\Psi_0, \Psi_1 \in T(\mathcal{X}, \mathcal{W} \oplus \mathbb{C})$.

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

Proof.—Generalizing the result of [23], we observe that if an entangled state $\rho \in D(\mathcal{X} \otimes Z)$ increases the distinguishability of two channels $\Psi_0, \Psi_1 \in T(\mathcal{X}, \mathcal{Y})$, 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 $\Omega \in T(\mathcal{X}, \mathcal{Y})$ is the totally depolarizing channel $\Omega[X] = (\text{Tr}(X)/d_{\mathcal{Y}})\mathbb{1}_{\mathcal{Y}}$.

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 containing

only separable states around the maximally mixed state [38]. It holds that $\|((\Xi_0^p - \Xi_1^p) \otimes \mathbb{1}_{L(Z)})[\rho]\|_{\text{tr}} = p\|((\Psi_0 - \Psi_1) \otimes \mathbb{1}_{L(Z)})[\rho]\|_{\text{tr}}$ and $\|\Xi_0^p - \Xi_1^p\|_{\text{tr}} = p\|\Psi_0 - \Psi_1\|_{\text{tr}}$; therefore, the state ρ enhances the distinguishability of channels Ξ_0^p, Ξ_1^p 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 positive map Φ_{TP} . We further remark that, for any entangled state, there exist tools to find a positive map that detects the state as entangled [42]. Unfortunately, it is not likely that this can efficiently be done [43,44].

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} [37,45]. For transposition one finds $c_T = 2/(d_{\mathcal{X}} + 1)$, and channels $\Psi_0, \Psi_1 \in T(\mathcal{X}, \mathcal{X} \otimes \mathbb{C})$, $\Psi_0: X \mapsto \frac{1}{d_{\mathcal{X}}+1}((\text{Tr}X)\mathbb{1}_{\mathcal{X}} + X^T)$, $\Psi_1: X \mapsto \frac{1}{d_{\mathcal{X}}+1}((\text{Tr}X)(\mathbb{1}_{\mathcal{X}} + 2|0\rangle\langle 0|) - X^T)$, with $|0\rangle$ a normalized vector orthogonal to \mathcal{X} . Thus, for any state $\rho \in D(\mathcal{X} \otimes \mathcal{Z})$, we obtain $\|(\Psi_0 \otimes \mathbb{1}_{L(\mathcal{Z})})[\rho] - (\Psi_1 \otimes \mathbb{1}_{L(\mathcal{Z})})[\rho]\|_{\text{tr}} - \|\Psi_0 - \Psi_1\|_{\text{tr}} = \frac{4}{d_{\mathcal{X}}+1}N_T(\rho)$, with $N_T(\rho)$ the standard negativity of ρ [38,39].

Conclusions.—We have proved that any entangled state is useful to distinguish some pair of (entanglement-breaking) channels strictly better than what is possible by means of a separable state in the minimum-error, single-shot scenario. 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 [37]. We expect that our result will stimulate further investigations on the role of entanglement in the discrimination of physical processes.

M. P. thanks C. E. Mora, P. Stelmachovic and M. Ziman for inspiring discussions. We acknowledge support from NSERC, CIFAR, QuantumWorks, and Ontario Centres of Excellence.

[1] R. Jozsa and N. Linden, Proc. R. Soc. A **459**, 2011 (2003).
 [2] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993).
 [3] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
 [4] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).
 [5] C. H. Bennett *et al.*, Phys. Rev. Lett. **76**, 722 (1996).
 [6] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
 [7] D. Bruss *et al.*, Phys. Rev. Lett. **93**, 210501 (2004).
 [8] M. Horodecki and M. Piani, arXiv:quant-ph/0701134.
 [9] K. Horodecki *et al.*, Phys. Rev. Lett. **94**, 160502 (2005).
 [10] L. Masanes, Phys. Rev. Lett. **96**, 150501 (2006).
 [11] P. Horodecki, M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **82**, 1056 (1999).

[12] L. Masanes, Y.-C. Liang, and A. C. Doherty, Phys. Rev. Lett. **100**, 090403 (2008).
 [13] A. Kitaev, Russ. Math. Surv. **52**, 1191 (1997).
 [14] 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 [15].
 [15] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, England, 2002).
 [16] A. Childs, J. Preskill, and J. Renes, J. Mod. Opt. **47**, 155 (2000).
 [17] G. M. D'Ariano, P. LoPresti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).
 [18] A. Acin, Phys. Rev. Lett. **87**, 177901 (2001).
 [19] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
 [20] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Phys. Rev. A **71**, 062310 (2005).
 [21] B. Rosgen and J. Watrous, in *Proceedings of the 20th Annual Conference on Computational Complexity* (IEEE Computer Society Press, Los Alamitos, California, 2005), p. 344.
 [22] M. F. Sacchi, Phys. Rev. A **71**, 062340 (2005).
 [23] M. F. Sacchi, Phys. Rev. A **72**, 014305 (2005).
 [24] S. Lloyd, Science **321**, 1463 (2008).
 [25] B. Rosgen, in *Proceedings of the 26th International Symposium on Theoretical Aspects of Computer Science* (Schloss Dagstuhl, Wadern, Germany, 2008), p. 597.
 [26] J. Watrous, Quantum Inf. Comput. **8**, 819 (2008).
 [27] 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 $\|A\|_{\text{tr}} = \sum_i |a_i|$.
 [28] A. Jamiolkowski, Rep. Math. Phys. **3**, 275 (1972).
 [29] M.-D. Choi, Linear Algebra Appl. **10**, 285 (1975).
 [30] J. de Pillis, Pac. J. Math. **23**, 129 (1967).
 [31] M. Horodecki, P. Shor, and M. Ruskai, Rev. Math. Phys. **15**, 629 (2003).
 [32] C. Helstrom, J. Stat. Phys. **1**, 231 (1969).
 [33] Achieved by choosing M_0 and M_1 to be the projectors on the positive and negative subspaces of $\rho_0 - \rho_1$.
 [34] This norm is different from the superoperator norm that is induced by the trace norm, sometimes denoted $\|\cdot\|_{\text{tr}}$.
 [35] The supremum is always achieved for $n \leq \dim(\mathcal{X})$.
 [36] M. Horodecki, P. Horodecki, and R. Horodecki, Open Syst. Inf. Dyn. **13**, 103 (2006).
 [37] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
 [38] K. Zyczkowski *et al.*, Phys. Rev. A **58**, 883 (1998).
 [39] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).
 [40] Note that the party on which the superoperator is applied is in general relevant.
 [41] One possible canonical choice, uniquely defined up to an isometry on \mathcal{Y} , is a (non-normalized) purification $\xi = |\xi\rangle\langle\xi|$, $|\xi\rangle \in \mathcal{Y} \otimes \mathcal{X}$, of $\mathbb{1}_{\mathcal{X}} - c_{\Phi_{\text{TA}}} Q$.
 [42] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004).
 [43] L. Gurvits, in *Proceedings of the 35th Annual Symposium on Theory of Computation* (Association for Computing Machinery, New York, NY, 2003), p. 10.
 [44] S. Gharibian, arXiv:0810.4507.
 [45] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).