

# Entanglement in channel discrimination with restricted measurements

William Matthews

*Institute for Quantum Computing and Department of Combinatorics and Optimization,  
University of Waterloo, Waterloo, Ontario, Canada*

Marco Piani

*Institute for Quantum Computing and Department of Physics and Astronomy,  
University of Waterloo, Waterloo, Ontario, Canada*

John Watrous

*Institute for Quantum Computing and School of Computer Science,  
University of Waterloo, Waterloo, Ontario, Canada*

We study the power of measurements implementable with local quantum operations and classical communication (or *LOCC measurements* for short) in the setting of quantum channel discrimination. More precisely, we consider discrimination procedures that attempt to identify an unknown channel, chosen uniformly from two known alternatives, that take the following form: (i) the input to the unknown channel is prepared in a possibly entangled state with an ancillary system, (ii) the unknown channel is applied to the input system, and (iii) an LOCC measurement is performed on the output and ancillary systems, resulting in a guess for which of the two channels was given. The restriction of the measurement in such a procedure to be an LOCC measurement is of interest because it isolates the entanglement in the initial input/ancillary systems as a resource in the setting of channel discrimination. We prove that there exist channel discrimination problems for which restricted procedures of this sort can be at either of the two extremes: they may be optimal within the set of all discrimination procedures (and simultaneously outperform all strategies that make no use of entanglement), or they may be no better than unentangled strategies (and simultaneously sub-optimal within the set of all discrimination procedures).

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

## I. INTRODUCTION

Quantum channel discrimination is an interesting problem in the theory of quantum information. In this problem, two known physical processes (or channels) are fixed, and access to one of them is made available—but it is not known which one it is. In the simplest scenario, 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. Several papers, including [1–15], have considered variants of this problem, both from an abstract point of view and a more applied one.

The most general form of a discrimination procedure for a channel discrimination problem of the type described above has the following form: (i) the input to the unknown channel is prepared in a possibly entangled state with an ancillary system, (ii) the unknown channel is applied to the input system, and (iii) a measurement is performed on the output and ancillary systems, resulting in a guess for which of the two channels was given. The primary purpose of this paper is to consider the effect of restricting the measurements in step (iii) to be LOCC (i.e., implementable using only local operations and classical communication with respect to the output/ancillary system splitting).

It is well-known that entanglement between the in-

put and ancillary systems is sometimes advantageous for channel discrimination, in the sense that it may allow for a strictly smaller probability of error to correctly identify the given channel in comparison to the case where there is no entanglement (which turns out to be equivalent to having no ancillary system whatsoever). This phenomenon seems to have been identified first by Kitaev [16], who introduced the *diamond norm* on super-operators to deal with precisely this phenomenon. More recent work, partially represented by the sources cited above, has further illuminated the usefulness of entanglement in the problem of channel discrimination and related tasks.

In [17] the question was reversed, by supposing that some arbitrary entangled state is given and asking whether the entanglement in this state is useful for channel discrimination. It was proved 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 unentangled (or separable) state. One may therefore say that every entangled state is a resource for channel discrimination. However, there is an obvious (potential) way that this result might be improved, which is to prove that every entangled state remains useful in the setting of channel discrimination even when measurements are restricted to be LOCC as sug-

gested above. This would have the effect of isolating the entanglement in the input/ancillary systems as the principal use of entanglement in such a procedure, and therefore as a more fundamentally important resource. We have not been able to determine whether or not such a result holds, and we consider this to be one of the main open problems of interest associated with this work.

Restrictions on operations are standard in entanglement theory [18], where they are introduced both because they are physically motivated (e.g., some initial entanglement between distant labs may be established by means of optical fibers or flying photons, but beyond this only local operations and classical communication may be feasible) and because they make entanglement theory an interesting resource theory, where the specifically quantum properties of entanglement are emphasized. The effect of restrictions on the measurement on the ability to discriminate—either ambiguously or unambiguously—between quantum states has also recently attracted much attention. In particular, the limits of LOCC discrimination have been investigated in [19–34], for instance. Such limits are at the base of the existence of hiding states [34–37], which are orthogonal and therefore perfectly distinguishable by global operations, but hardly distinguishable by LOCC measurements. A systematic approach to investigating the relation between distinguishability of states under various restrictions on measurements was put forward in [38].

In regard to the role of measurements and restrictions on measurements in quantum channel discrimination, the focus of past work has mainly been on the discrimination of bi- and multi-partite unitary operations. It has been proved that with many uses of the unknown unitary, it is possible to perfectly discriminate between any set of unitaries, as long as the parties can apply an LOCC protocol [39, 40]. We stress that the unitaries acts globally on the same parties on which the restrictions are imposed. The issue we consider in this paper—the role of restrictions on the measurements in minimum-error channel discrimination—deals with concepts similar to those present in [39, 40], but with critical differences. Indeed, in our case we consider that (i) only one use of the unknown channel is allowed, and we do not focus on perfect discrimination; and (ii) the channels—possibly unitaries—are applied only to the input, so the input/ancilla evolution is local. In particular, in the case of unitaries it is known that input/ancilla entanglement is useless for minimum-error channel discrimination, so restrictions on the measurements that are based on input/ancilla locality are uninteresting.

The main contribution of the present paper is the identification of instances of channel discrimination problems where procedures restricted to make LOCC measurements are at either of the two extremes: they may be optimal within the set of all discrimination procedures and simultaneously outperform all strategies that make no use of entanglement, or they may be no better than unentangled strategies and simultaneously sub-optimal

within the set of all discrimination procedures. These two possibilities are discussed in Sections III and IV, respectively, which follow Section II that provides some definitions that are useful for describing the examples. The paper concludes with Section V, which discusses some future research directions relating to our work.

## II. DEFINITIONS

Throughout this paper we will use notation and terminology that, for the most part, is standard in the theory of quantum information. For the sake of clarity let us state explicitly that we restrict our attention to finite-dimensional complex Hilbert spaces in this paper, and for any such space  $\mathcal{X}$  we write  $\mathcal{L}(\mathcal{X})$ ,  $\text{Herm}(\mathcal{X})$ ,  $\text{Pos}(\mathcal{X})$  and  $\mathcal{D}(\mathcal{X})$  to denote the sets of all linear operators, Hermitian operators, positive semidefinite operators, and density operators on  $\mathcal{X}$ , respectively. We also write  $\text{Sep}(\mathcal{X} : \mathcal{Y})$  and  $\text{PPT}(\mathcal{X} : \mathcal{Y})$  to denote the sets of all (unnormalized) separable and PPT (i.e., positive partial transpose) operators on a tensor product space  $\mathcal{X} \otimes \mathcal{Y}$ . [45]

When we refer to a *channel* we mean a completely positive, trace-preserving linear mapping of the form

$$\Phi : \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{Y}). \quad (1)$$

Hereafter we will write  $\mathcal{T}(\mathcal{X}, \mathcal{Y})$  to refer to the vector space of all (not necessarily completely positive or trace-preserving) mappings of the form (1).

As described in the introduction, this paper concerns the problem of channel discrimination. The specific type of channel discrimination problems we consider are as follows. Two channels  $\Phi_0, \Phi_1 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  are fixed. One of the two channels is selected, uniformly at random, and a single evaluation of this (unknown) channel is made available. The goal is to determine which of the two channels was selected.

A natural, but sometimes sub-optimal, strategy for solving an instance of a channel discrimination problem is to choose a quantum state  $\rho \in \mathcal{D}(\mathcal{X})$ , to apply the unknown channel to  $\rho$ , and to measure the resulting state according to a binary-valued measurement  $\{P_0, P_1\} \subset \text{Pos}(\mathcal{Y})$ . The measurement outcome (0 or 1) is then interpreted as the procedure's guess for which channel was given. The probability that such a procedure correctly identifies the unknown channel is given by

$$\begin{aligned} & \frac{1}{2} \langle P_0, \Phi_0(\rho) \rangle + \frac{1}{2} \langle P_1, \Phi_1(\rho) \rangle \\ &= \frac{1}{2} + \frac{1}{4} \langle P_0 - P_1, \Phi_0(\rho) - \Phi_1(\rho) \rangle \end{aligned}$$

where the inner product is the Hilbert-Schmidt inner product:  $\langle X, Y \rangle = \text{Tr}(X^*Y)$ . Optimizing over all choices of  $\rho \in \mathcal{D}(\mathcal{X})$  and all binary-valued measurements  $\{P_0, P_1\}$  on  $\mathcal{Y}$  yields a correctness probability

$$\frac{1}{2} + \frac{1}{4} \|\Phi_0 - \Phi_1\|_{\text{NE}},$$

where the norm  $\|\Phi\|_{\text{NE}}$  is defined as

$$\|\Phi\|_{\text{NE}} = \max_{\rho \in \mathcal{D}(\mathcal{X})} \|\Phi(\rho)\|_1$$

for every Hermiticity-preserving[46] mapping  $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ . This norm could be extended to arbitrary (non-Hermiticity-preserving) mappings, but it is not necessary for us to consider such extensions in this paper. The subscript NE for this norm is short for “no entanglement,” which refers to the fact that the discrimination procedure has not made use of the possibility that the input system to the unknown channel could have been entangled with an ancillary system.

A more general type of discrimination strategy that does make use of an ancillary system is as follows. A quantum state  $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})$  (for an arbitrary choice of  $\mathcal{Z}$ ) is selected, and the unknown channel is applied to the part of this state corresponding to  $\mathcal{X}$ . A binary-valued measurement  $\{P_0, P_1\} \subset \text{Pos}(\mathcal{Y} \otimes \mathcal{Z})$  is then applied to the resulting state, and (as before) the outcome is interpreted as the procedure’s guess for which channel was given. The probability for such a procedure to correctly identify the unknown channel is

$$\frac{1}{2} + \frac{1}{4} \langle P_0 - P_1, (\Phi_0 \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) - (\Phi_1 \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) \rangle. \quad (2)$$

Optimizing over all choices of  $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})$  and all binary-valued measurements  $\{P_0, P_1\}$  on  $\mathcal{Y} \otimes \mathcal{Z}$ , for any choice of  $\mathcal{Z}$  having dimension at least that of  $\mathcal{X}$ , results in the quantity

$$\frac{1}{2} + \frac{1}{4} \|\Phi_0 - \Phi_1\|_{\diamond},$$

where the *diamond norm*  $\|\cdot\|_{\diamond}$  is defined as

$$\|\Phi\|_{\diamond} = \max_{\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{X})} \|(\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X})})(\rho)\|_1 \quad (3)$$

for every Hermiticity-preserving mappings  $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ . (For general maps, the maximum is taken over all  $X \in \mathcal{L}(\mathcal{X} \otimes \mathcal{X})$  with  $\|X\|_1 \leq 1$ .)

It is known that this more general sort of strategy can give a striking improvement in the probability to correctly discriminate some pairs of channels. It is the *entanglement* between the input and ancillary systems that is responsible for this improvement, when it occurs, because (as observed in [17]) it holds that

$$\max_{\rho \in \text{SepD}(\mathcal{X} : \mathcal{Z})} \|(\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho)\|_1 = \|\Phi\|_{\text{NE}}$$

for every choice of  $\mathcal{Z}$ , where  $\text{SepD}(\mathcal{X} : \mathcal{Z})$  denotes the set of separable density operators on  $\mathcal{X} \otimes \mathcal{Z}$ . In other words, the ancillary system is useless for channel discrimination unless it is entangled with the input system.

With the connection between channel discrimination and the two norms defined above in mind, we define a norm  $\|\Phi\|_{\text{LOCC}}$  by considering a maximization of the expression (2) over those choices of  $P_0$  and  $P_1$  that represent

*LOCC measurements*, as opposed to general measurements. More precisely, for any Hermiticity-preserving mapping  $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  we define

$$\|\Phi\|_{\text{LOCC}} = \max_{\{P_0, P_1\}} \max_{\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})} \langle P_0 - P_1, (\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) \rangle, \quad (4)$$

where the maximization is taken over all LOCC measurements[47]  $\{P_0, P_1\}$  on  $\mathcal{Y} \otimes \mathcal{Z}$ , and where  $\mathcal{Z}$  is any space having dimension at least that of  $\mathcal{X}$ . We also define  $\|\Phi\|_{\text{SEP}}$  and  $\|\Phi\|_{\text{PPT}}$  similarly, where the maximization is over all separable or PPT binary-valued measurements, respectively.[48] In all of these cases, the resulting norm is insensitive to the dimension of  $\mathcal{Z}$ , provided it is at least that of  $\mathcal{X}$ . This follows from the observation that the maximum over  $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})$  for any of these norms is always achieved for a pure state (by a simple convexity argument), and such a state must be supported on a subspace of  $\mathcal{Z}$  having dimension at most that of  $\mathcal{X}$ . Similar to  $\|\Phi\|_{\text{NE}}$ , we do not concern ourselves with extensions of  $\|\Phi\|_{\text{LOCC}}$ ,  $\|\Phi\|_{\text{SEP}}$ , or  $\|\Phi\|_{\text{PPT}}$  to non-Hermiticity-preserving mappings  $\Phi$ .

It is clear that

$$\|\Phi\|_{\text{NE}} \leq \|\Phi\|_{\text{LOCC}} \leq \|\Phi\|_{\text{SEP}} \leq \|\Phi\|_{\text{PPT}} \leq \|\Phi\|_{\diamond} \quad (5)$$

and the main contribution of this paper is to provide examples of channels  $\Phi_0$  and  $\Phi_1$  for which the mapping  $\Phi = \Phi_0 - \Phi_1$  causes various choices of the inequalities in (5) to become either equalities or strict inequalities. We are, in particular, interested in the relationship between  $\|\Phi_0 - \Phi_1\|_{\text{LOCC}}$  and the two norms  $\|\Phi_0 - \Phi_1\|_{\text{NE}}$  and  $\|\Phi_0 - \Phi_1\|_{\diamond}$  for different choices of channels  $\Phi_0$  and  $\Phi_1$ . This relationship addresses the question raised in the introduction, which is whether entanglement between the input and auxiliary systems remains a useful resource for channel discrimination when entangled measurements are disallowed. The two sections that follow show that sometimes entanglement is still useful in this sense and sometimes it is not.

Since we focus on families of measurements defined with respect to locality, and the input states are always optimized, three operational minimum-error channel discrimination scenarios corresponding to the three norms  $\|\cdot\|_{\text{NE}}$ ,  $\|\cdot\|_{\diamond}$  and  $\|\cdot\|_{\text{LOCC}}$  can be depicted as in Figure 1, Figure 2, and Figure 3, respectively. The channels go from Alice to Bob, as in a communication scenario, but the task is not that of transmitting a classical or a quantum message from Alice to Bob having at disposal many uses of the channel. Rather, they want to discriminate between two possible channels with minimum probability of error, having at their disposal only one use of the channel.

In the NE norm case, Alice feeds the channels with a probe that is not correlated with any other subsystems held by either herself or Bob, and Bob measures the output probe, guessing which channel was applied. In the diamond-norm case, at the beginning Alice holds

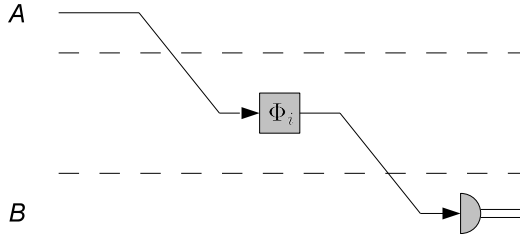


FIG. 1: Operational scenario for the NE norm. The unknown channel  $\Phi_i$  goes from Alice to Bob, and it is not in the hands of either of them. Bob performs the final measurement. In this case, Alice's input to the channel is uncorrelated with any other system in the hands of Alice or Bob.

the probe and Bob the ancilla, which are in whatever needed pre-distributed entangled state; Alice sends the probe down the channel, and Bob can jointly measure the output probe and the ancilla.[49] Finally, in the LOCC-

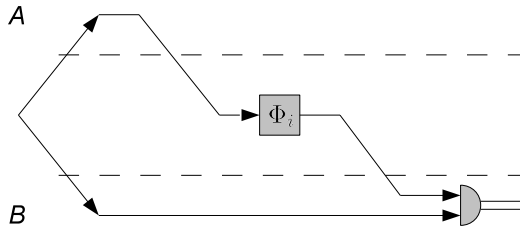


FIG. 2: Operational scenario for the diamond norm. Alice and Bob share an initial—possibly entangled—probe/ancilla state which is optimal for channel discrimination. Alice sends the probe down the channel and Bob proceeds to the measurement of the output probe/ancilla state.

norm case, we can imagine that Alice locally creates an entangled probe/ancilla state, feeding the channel with the probe and keeping the ancilla. The measurement is then performed by LOCC on the output probe held by Bob and the ancilla held by Alice. We remark that this

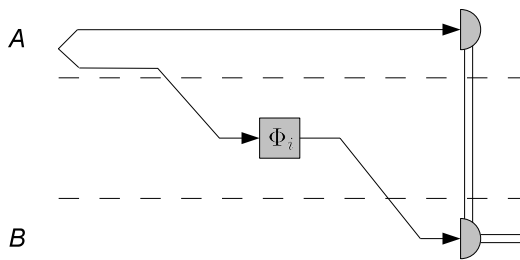


FIG. 3: Operational scenario for the LOCC norm. Alice create an entangled probe/ancilla state, keeps the ancilla, and sends the probe down the channel. Alice and Bob can then proceed to an LOCC measurement of the output probe/ancilla state.

scenario—where the unknown channel  $\Phi_i$  goes from Alice to Bob, and it is not in the hands of either of them—could be thought as corresponding to a practical situ-

ation like that of, e.g., an optical fiber underneath the sea. In particular, the restriction to LOCC output measurement would be the result of the non-availability of long-distance pre-established entanglement—or of some other quantum channel to be used to send also the ancilla from Alice to Bob.

### III. OPTIMAL DISCRIMINATION PROCEDURES WITH LOCC MEASUREMENTS

In this section we provide examples of channels  $\Phi_0$  and  $\Phi_1$  for which

$$\|\Phi_0 - \Phi_1\|_{\text{NE}} < \|\Phi_0 - \Phi_1\|_{\text{LOCC}} = \|\Phi_0 - \Phi_1\|_{\diamond}. \quad (6)$$

Thus, entanglement between the input and ancillary systems may still be useful for channel discrimination when the measurement is LOCC—and indeed an LOCC measurement may even be optimal in this situation.

The first collection of examples we provide achieves a limited gap between the quantities  $\|\Phi_0 - \Phi_1\|_{\text{NE}}$  and  $\|\Phi_0 - \Phi_1\|_{\text{LOCC}}$ , but has the advantage of being simple to describe. The second type of example achieves a large gap, but also requires that the output dimension of the channels be very large. The third example is specifically for input qubits, and while being similar to the second class of examples, we include it because of its simplicity: on the one hand, it is possible to provide exact values for the various norms; on the other hand, it could be experimentally implemented.

#### A. Examples based on flagged Kraus operators

Let  $\mathcal{X} = \mathbb{C}^d$  for any desired choice of  $d \geq 2$ , and let  $\mathcal{Y} = \mathbb{C}^n \otimes \mathcal{X}$  for  $n \geq 1$ . Consider a choice of channels  $\Phi_0, \Phi_1 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  defined as

$$\Phi_a(\rho) = \sum_{j=1}^n |j\rangle\langle j| \otimes A_{a,j} \rho A_{a,j}^*,$$

for some selection of operators  $\{A_{a,j} : a \in \{0,1\}, 1 \leq j \leq n\} \subset \mathcal{L}(\mathcal{X})$  satisfying

$$\sum_{j=1}^n A_{0,j}^* A_{0,j} = \sum_{j=1}^n A_{1,j}^* A_{1,j} = \mathbb{1}_{\mathcal{X}}.$$

A sufficient condition for  $\Phi_0$  and  $\Phi_1$  to be perfectly distinguishable using an LOCC measurement discrimination procedure (i.e.,  $\|\Phi_0 - \Phi_1\|_{\text{LOCC}} = 2$ ) is that  $\langle A_{0,j}, A_{1,j} \rangle = 0$  for all  $j = 1, \dots, n$ . This observation follows from the well-known theorem of [19] stating that every fixed pair of orthogonal pure states can be perfectly distinguished by an LOCC measurement. In particular, if the maximally entangled state

$$\frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle|i\rangle$$

between the input and an ancillary system is selected, then the resulting state produced by  $\Phi_a$  has the form

$$\frac{1}{d} \sum_{j=1}^n |j\rangle\langle j| \otimes |\psi_{a,j}\rangle\langle\psi_{a,j}|$$

for vectors  $\{|\psi_{a,j}\rangle\}$  satisfying  $\langle\psi_{0,j}|\psi_{1,j}\rangle = \langle A_{0,j}, A_{1,k}\rangle = 0$  for every choice of  $j = 1, \dots, n$ . An LOCC measurement that first measures  $j$ , then implements the corresponding measurement of [19] to distinguish  $|\psi_{0,j}\rangle$  and  $|\psi_{1,j}\rangle$  succeeds in discriminating  $\Phi_0$  and  $\Phi_1$  without error.

On the other hand, it holds that

$$\|\Phi_0 - \Phi_1\|_{\text{NE}} = \max_{\rho \in \mathcal{D}(\mathcal{X})} \sum_{j=1}^n \|A_{0,j}\rho A_{0,j}^* - A_{1,j}\rho A_{1,j}^*\|_1. \quad (7)$$

So, to obtain examples of channel pairs for which (6) holds, it suffices to select a collection of operators  $\{A_{a,j}\}$  so that

1.  $\langle A_{0,j}, A_{1,j} \rangle = 0$  for each  $j = 1, \dots, n$ , and
2. the expression in (7) is smaller than 2.

This may be accomplished, for instance, by setting  $n = d^2 - 1$ ,  $A_{0,j} = \mathbb{1}/\sqrt{n}$ , and  $A_{1,j} = W_j/\sqrt{n}$ , for  $j = 1, \dots, n$  and where  $W_1, \dots, W_n$  is any orthonormal collection of traceless unitary operators, such as the non-identity *discrete Weyl* (or *generalized Pauli*) operators. We then have

$$\begin{aligned} \|\Phi_0 - \Phi_1\|_{\text{NE}} &= \max_{|\psi\rangle} \frac{2}{d^2 - 1} \sum_{j=1}^{d^2-1} \sqrt{1 - |\langle\psi|W_j|\psi\rangle|^2} \\ &\leq 2\sqrt{\frac{d}{d+1}} \end{aligned}$$

where the inequality is due to the concavity of the square root (with equality if and only if Zauner's conjecture [41, 42] holds for dimension  $d$ ), and we used the fact that  $\rho + \sum_{j=1}^{d^2-1} W_j \rho W_j^* = d \text{Tr}(\rho) \mathbb{1}_{\mathcal{X}}$

### B. Examples based on random binary-valued measurements

Let  $d$  be an even positive integer, let  $\mathcal{X} = \mathbb{C}^d$  and  $\mathcal{Y} = \mathbb{C}^2$ , and define channels  $\Psi_0, \Psi_1 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  as

$$\begin{aligned} \Psi_0(\rho) &= \text{Tr}(\Pi_0 \rho) |0\rangle\langle 0| + \text{Tr}(\Pi_1 \rho) |1\rangle\langle 1| \\ \Psi_1(\rho) &= \text{Tr}(\Pi_1 \rho) |0\rangle\langle 0| + \text{Tr}(\Pi_0 \rho) |1\rangle\langle 1|, \end{aligned}$$

where

$$\Pi_0 = \sum_{j=1}^{d/2} |j\rangle\langle j| \quad \text{and} \quad \Pi_1 = \sum_{j=d/2+1}^d |j\rangle\langle j|.$$

Now suppose that  $U_1, \dots, U_N \in \mathcal{U}(\mathcal{X})$  is a collection of unitary operators, and define channels  $\Phi_0, \Phi_1 \in \mathcal{T}(\mathcal{X}, \mathbb{C}^N \otimes \mathcal{Y})$  as follows:

$$\begin{aligned} \Phi_0(\rho) &= \frac{1}{N} \sum_{j=1}^N |j\rangle\langle j| \otimes \Psi_0(U_j \rho U_j^*) \\ \Phi_1(\rho) &= \frac{1}{N} \sum_{j=1}^N |j\rangle\langle j| \otimes \Psi_1(U_j \rho U_j^*). \end{aligned}$$

A representation of the action of the channels is provided in Figure 4.

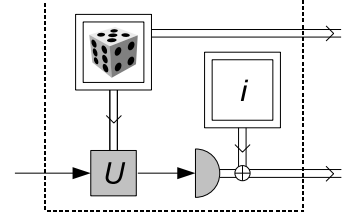


FIG. 4: Structure of some exemplary pairs of channels for which the LOCC norm is equal to the diamond norm, and strictly greater than the NE norm. Single lines indicate quantum systems or messages; double lines classical ones. The channel chooses at random a unitary among a discrete set. An incomplete projective measurement in the computational basis with two possible outcomes is then performed. The result of the measurement is further flipped or not, depending on the channel (on the value of  $i$  in the figure). The output of the channel is classical, meaning that it is diagonal in a fixed computational basis: a register containing the index of the chosen unitary, and a bit containing the (possibly flipped) result of the measurement. A full description of the action of the channels can be found in the main text.

For any choice of unitary operators  $U_1, \dots, U_N$ , the channels  $\Phi_0$  and  $\Phi_1$  can be perfectly distinguished by an LOCC discrimination procedure as follows.

1. The given channel  $\Phi_a$  is evaluated on half of the canonical maximally entangled state

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle|i\rangle.$$

Let us suppose that the resulting state is stored in the registers  $A_1$ ,  $A_2$  and  $B$ , where the vector spaces corresponding to these registers are given by  $\mathbb{C}^N$ ,  $\mathcal{Y}$ , and  $\mathcal{X}$ , respectively; and where Alice holds  $A_1$  and  $A_2$ , and Bob holds  $B$ .

2. Alice measures  $A_1$  with respect to the standard basis and transmits the result to Bob. Upon receiving a value  $j \in \{1, \dots, N\}$ , Bob applies the operation  $\overline{U}_j$  to  $B$ , where the complex conjugation is taken with respect to the standard basis. The state of the pair  $(A_2, B)$  at this point is given by  $(\Psi_a \otimes \mathbb{1})(|\psi\rangle\langle\psi|)$ .
3. Bob applies  $\Psi_0$  to  $B$  and measures the resulting qubit with respect to the standard basis. Alice measures her

qubit  $A_2$  with respect to the standard basis as well, and Alice and Bob compare their measurements. If they agree then  $a = 0$  and otherwise  $a = 1$ .

Thus, it holds that  $\|\Phi_0 - \Phi_1\|_{\text{LOCC}} = 2$ .

Now we will argue that for some choice of unitary operators  $U_1, \dots, U_N$  it holds that

$$\|\Phi_0 - \Phi_1\|_{\text{NE}} = \sqrt{\frac{2}{\pi d}} + O\left(\frac{1}{d}\right).$$

This is done by taking  $N$  to be large and selecting  $U_1, \dots, U_N$  independently according to the normalized Haar measure on  $U(\mathcal{X})$ . It holds that

$$\begin{aligned} & \|\Phi_0 - \Phi_1\|_{\text{NE}} \\ &= \max_{|\psi\rangle} \frac{2}{N} \sum_{k=1}^N \left| \sum_{j=1}^{d/2} |\langle\psi|U_k|j\rangle|^2 - \sum_{j=d/2+1}^d |\langle\psi|U_k|j\rangle|^2 \right|, \end{aligned} \quad (8)$$

which, in the limit of large  $N$ , approaches

$$2 \max_{|\psi\rangle} \int \left| \sum_{j=1}^{d/2} |\langle\psi|U_j|j\rangle|^2 - \sum_{j=d/2+1}^d |\langle\psi|U_j|j\rangle|^2 \right| d\mu(U)$$

for  $\mu$  denoting the Haar measure on  $U(\mathcal{X})$ , normalized so that  $\mu(U(\mathcal{X})) = 1$ . A formula for this type of integral is given in Appendix B of [38], which in this case evaluates to

$$\frac{1}{d} \sum_{j=1}^{d/2} 2^{-2j} \binom{2j}{j} = \sqrt{\frac{2}{\pi d}} + O\left(\frac{1}{d}\right)$$

as claimed.

### C. Simple pairs of channels with qubit input

Consider a special case of the exemplary class presented in Section III B, with qubit input dimension  $d = 2$ . In particular, take the number of random unitaries  $N = 2$ , with  $U_1 = \mathbb{1}$  and  $U_2 = (\sigma_x + \sigma_z)/\sqrt{2}$ . Then Eq. (8) becomes

$$\|\Phi_0 - \Phi_1\|_{\text{NE}} = \max_{|\psi\rangle} \sum_{k=x,z} |\langle\psi|\sigma_k|\psi\rangle| = \sqrt{2}$$

where the maximum is realized by a pure state in the  $xz$  plane of the Bloch sphere.

This qubit-input example can also be computed exactly for  $N = 3$ , with  $U_1 = \mathbb{1}$ ,  $U_2 = (\sigma_x + \sigma_z)/\sqrt{2}$  and  $U_3 = (\sigma_y + \sigma_z)/\sqrt{2}$ . Then one has  $\|\Phi_0 - \Phi_1\|_{\text{NE}} = 2/\sqrt{3}$ , with a slightly bigger gap between the LOCC and diamond norms, still both equal to 2, and the NE norm.

## IV. SUB-OPTIMALITY OF LOCC MEASUREMENTS

In this section we show that there are cases where restricting measurements to be LOCC in channel discrimination procedures eliminates the benefits of entanglement altogether. That is, there exist channels  $\Phi_0$  and  $\Phi_1$  for which

$$\|\Phi_0 - \Phi_1\|_{\text{NE}} = \|\Phi_0 - \Phi_1\|_{\text{LOCC}} < \|\Phi_0 - \Phi_1\|_{\diamond} \quad (9)$$

Three sets of examples are discussed.

### A. Channels with one-qubit output

For any pair of channels whose outputs correspond to a single qubit, discrimination strategies using LOCC measurements can be no better than strategies that use no entanglement. This follows from the following theorem.

**Theorem 1.** *For any choice of channels  $\Phi_0, \Phi_1 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  it holds that*

$$\|\Phi_0 - \Phi_1\|_{\text{SEP}} \leq \frac{\dim(\mathcal{Y})}{2} \|\Phi_0 - \Phi_1\|_{\text{NE}}.$$

*Proof.* Let  $\Phi = \Phi_0 - \Phi_1$ , and let  $\mathcal{Z} = \mathbb{C}^n$  for an arbitrary choice of  $n \geq 1$ . Also let  $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})$  be a density operator and let  $\{Q_1 \otimes R_1, \dots, Q_N \otimes R_N\}$  represent an arbitrary product measurement over  $\mathcal{Y} \otimes \mathcal{Z}$ , so that  $Q_1, \dots, Q_N \in \text{Pos}(\mathcal{Y})$  and  $R_1, \dots, R_N \in \text{Pos}(\mathcal{Z})$  satisfy

$$\sum_{j=1}^N Q_j \otimes R_j = \mathbb{1}_{\mathcal{Y} \otimes \mathcal{Z}}.$$

There is no loss of generality in assuming that each operator  $Q_j$  is a density operator, because each  $R_j$  can be re-scaled appropriately.

Now, given that we assume that  $Q_1, \dots, Q_N$  are density operators, it holds that

$$\sum_{j=1}^N R_j = \dim(\mathcal{Y}) \mathbb{1}_{\mathcal{Z}}.$$

Our goal will be to establish an upper bound on the quantity

$$\sum_{j=1}^N |\langle Q_j \otimes R_j, (\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) \rangle|. \quad (10)$$

To this end let us write  $X_j = \text{Tr}_{\mathcal{Z}}[(\mathbb{1}_{\mathcal{Y}} \otimes R_j)\rho]$ , so that

$$\begin{aligned} |\langle Q_j \otimes R_j, (\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) \rangle| &= |\langle Q_j, \Phi(X_j) \rangle| \\ &\leq \frac{1}{2} \|\Phi\|_{\text{NE}} \text{Tr}(X_j) \end{aligned}$$

for each  $j = 1, \dots, N$ , where the inequality follows from the observations that each  $\Phi(X_j)$  is traceless and that  $Q_j \leq \mathbb{1}_{\mathcal{Y}}$ . It follows that

$$\begin{aligned} & \sum_{j=1}^N |\langle Q_j \otimes R_j, (\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) \rangle| \\ & \leq \frac{1}{2} \|\Phi\|_{\text{NE}} \sum_{j=1}^N \text{Tr}(X_j) = \frac{\dim(\mathcal{Y})}{2} \|\Phi\|_{\text{NE}}. \end{aligned} \quad (11)$$

Finally, we note that the quantity  $\|\Phi\|_{\text{SEP}}$  is given by the supremum of the expression (10) over all choices of  $n$  and all product measurements  $\{Q_1 \otimes R_1, \dots, Q_N \otimes R_N\}$ , which completes the proof.  $\square$

**Corollary 1.** *Suppose  $\Phi_0, \Phi_1 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  are channels for  $\dim(\mathcal{Y}) = 2$ . Then  $\|\Phi\|_{\text{NE}} = \|\Phi\|_{\text{SEP}}$ .*

As there exist channels  $\Phi_0$  and  $\Phi_1$  having single-qubit outputs for which  $\|\Phi_0 - \Phi_1\|_{\text{NE}} < \|\Phi_0 - \Phi_1\|_{\diamond}$ , we have that (9) holds for these channels.

## B. Werner-Holevo channels

The Werner-Holevo channels  $\Phi_0$  and  $\Phi_1$  are defined for any dimension  $d \geq 2$  as

$$\begin{aligned} \Phi_0(X) &= \frac{1}{d+1} (\text{Tr}(X)\mathbb{1} + X^\top), \\ \Phi_1(X) &= \frac{1}{d-1} (\text{Tr}(X)\mathbb{1} - X^\top), \end{aligned}$$

where transposition is taken with respect to the standard basis [43]. It holds that

$$\|\Phi_0 - \Phi_1\|_{\text{NE}} = \frac{4}{d+1},$$

whereas  $\|\Phi_0 - \Phi_1\|_{\diamond} = 2$ ; the channels are almost indistinguishable for large  $d$  without the use of entanglement, but are perfectly distinguishable with entanglement.

We will now prove that for this choice of channels, LOCC measurements render entanglement useless for their discrimination. Indeed, even PPT measurements have this property:

$$\|\Phi_0 - \Phi_1\|_{\text{LOCC}} = \|\Phi_0 - \Phi_1\|_{\text{PPT}} = \frac{4}{d+1}.$$

To prove that this is so, it suffices to prove that for every space  $\mathcal{Z}$  with  $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$ , every unit vector  $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Z}$ , and every PPT measurement  $\{P_0, P_1\}$  on  $\mathcal{Y} \otimes \mathcal{Z}$ , it holds that

$$\langle P_0 - P_1, ((\Phi_0 - \Phi_1) \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|) \rangle \leq \frac{4}{d+1}.$$

Given that  $((\Phi_0 - \Phi_1) \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|)$  is traceless, and that  $P_0 = \mathbb{1} - P_1$  for every choice of a measurement  $\{P_0, P_1\}$ , it suffices to prove

$$\langle P_1, ((\Phi_1 - \Phi_0) \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|) \rangle \leq \frac{2}{d+1} \quad (12)$$

for every unit vector  $|\psi\rangle$  and every PPT operator  $P_1$  satisfying  $0 \leq P_1 \leq \mathbb{1}$ . (It may be the case that  $P_0 = \mathbb{1} - P_1$  fails to be PPT for some choices of a PPT operator  $P_1$ , but the above bound will hold nevertheless—we do not require the assumption that  $P_0$  is PPT to establish the bound.)

To prove the bound (12), note first that any unit vector  $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Z}$  may be written

$$|\psi\rangle = (\mathbb{1}_{\mathcal{X}} \otimes A) \sum_{j=1}^d |j\rangle|j\rangle$$

for some choice of a linear mapping  $A : \mathcal{X} \rightarrow \mathcal{Z}$  satisfying  $\|A\|_2 \stackrel{\text{def}}{=} \sqrt{\langle A, A \rangle} = 1$ . Thus, for any mapping  $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  it holds that

$$(\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|) = (\mathbb{1}_{\mathcal{Y}} \otimes A)J(\Phi)(\mathbb{1}_{\mathcal{Y}} \otimes A)^*,$$

where

$$J(\Phi) = \sum_{1 \leq j, k \leq d} \Phi(|j\rangle\langle k|) \otimes |j\rangle\langle k|$$

is the *Choi-Jamiołkowski representation* of  $\Phi$ . For the mapping  $\Phi = \Phi_1 - \Phi_0$ , it holds that

$$J(\Phi_1 - \Phi_0) = \frac{2}{d-1}R - \frac{2}{d+1}S,$$

for  $R$  and  $S$  denoting the projections onto the anti-symmetric and symmetric subspaces of  $\mathcal{Y} \otimes \mathcal{X} = \mathbb{C}^d \otimes \mathbb{C}^d$ , respectively. Thus,

$$\begin{aligned} & \langle P_1, ((\Phi_1 - \Phi_0) \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(|\psi\rangle\langle\psi|) \rangle \\ &= \left\langle (\mathbb{1}_{\mathcal{Y}} \otimes A^*)P_1(\mathbb{1}_{\mathcal{Y}} \otimes A), \frac{2}{d-1}R - \frac{2}{d+1}S \right\rangle. \end{aligned}$$

It is not difficult to prove that  $\langle Q, R - S \rangle \leq 0$  for every PPT operator  $Q$ . Given that  $P_1$  is PPT, so too is the operator  $(\mathbb{1}_{\mathcal{Y}} \otimes A^*)P_1(\mathbb{1}_{\mathcal{Y}} \otimes A)$ , and therefore

$$\begin{aligned} & \left\langle (\mathbb{1}_{\mathcal{Y}} \otimes A^*)P_1(\mathbb{1}_{\mathcal{Y}} \otimes A), \frac{2}{d-1}R - \frac{2}{d+1}S \right\rangle \\ & \leq \frac{4}{d^2-1} \langle (\mathbb{1}_{\mathcal{Y}} \otimes A^*)P_1(\mathbb{1}_{\mathcal{Y}} \otimes A), R \rangle \\ & \leq \frac{4}{d^2-1} \langle \mathbb{1}_{\mathcal{Y}} \otimes A^*A, R \rangle. \end{aligned}$$

As  $A^*A$  is a density operator and  $\text{Tr}_{\mathcal{Y}}(R) = \frac{d-1}{2}\mathbb{1}_{\mathcal{Y}}$ , it follows that

$$\frac{4}{d^2-1} \langle \mathbb{1}_{\mathcal{Y}} \otimes A^*A, R \rangle = \frac{2}{d+1},$$

which establishes the claimed bound (12).

## C. The channels of [17]

Finally, we prove that channels  $\Psi_0$  and  $\Psi_1$  constructed in Theorem 1 of [17] satisfy

$$\|\Psi_0 - \Psi_1\|_{\text{NE}} = \|\Psi_0 - \Psi_1\|_{\text{SEP}} < \|\Psi_0 - \Psi_1\|_{\diamond}.$$

Channels constructed in this way demonstrate that every entangled state is useful for channel discrimination, provided that the final measurement is unrestricted—but by the above relationship we see that the advantage is lost when the final measurements are restricted to be separable.

Let us first recall that channels  $\Psi_0$  and  $\Psi_1$  constructed in [17] satisfy  $\Psi_0 - \Psi_1 = \alpha\Phi$  for some  $\alpha > 0$  and  $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y} \oplus \mathbb{C})$  being a mapping defined as

$$\Phi(X) = \begin{pmatrix} \Xi(X) & 0 \\ 0 & -\text{Tr}(X) \end{pmatrix}$$

for  $\Xi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  being positive (but not completely positive) and trace-preserving. (The mapping  $\Phi$  is therefore *trace-annihilating*, which is a necessary and sufficient condition for the expression  $\Psi_0 - \Psi_1 = \alpha\Phi$  to hold for channels  $\Psi_0$  and  $\Psi_1$  and a scalar  $\alpha > 0$ .)

Now, it is not difficult to see that  $\|\Phi\|_{\text{SEP}} = \|\Xi\|_{\text{SEP}} + 1$ . Specifically, the inequality  $\|\Phi\|_{\text{SEP}} \leq \|\Xi\|_{\text{SEP}} + 1$  follows from the triangle equality, while the reverse inequality  $\|\Phi\|_{\text{SEP}} \geq \|\Xi\|_{\text{SEP}} + 1$  is obtained by considering the separable measurement  $\{Q_0, Q_1\} \subset \text{Sep}((\mathcal{Y} \oplus \mathbb{C}) : \mathcal{X})$  defined as

$$Q_0 = \begin{pmatrix} P_0 & 0 \\ 0 & 0_{\mathcal{X}} \end{pmatrix} \quad \text{and} \quad Q_1 = \begin{pmatrix} P_1 & 0 \\ 0 & \mathbb{1}_{\mathcal{X}} \end{pmatrix}$$

for  $\{P_0, P_1\} \subset \text{Sep}(\mathcal{Y} : \mathcal{X})$  being an optimal separable measurement in the definition of  $\|\Xi\|_{\text{SEP}}$ .

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

$$\begin{aligned} \|(\Xi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho)\|_{\text{SEP}} &= \langle P_0 - P_1, (\Xi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho) \rangle \\ &= \langle (\Xi^* \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(P_0) - (\Xi^* \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(P_1), \rho \rangle \end{aligned}$$

for some choice of a separable measurement  $\{P_0, P_1\}$ . Given that  $\Xi$  is positive and trace-preserving we have that the adjoint mapping  $\Xi^*$  is positive and unital, and given that  $\{P_0, P_1\}$  is separable it therefore follows that  $\{(\Xi^* \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(P_0), (\Xi^* \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(P_1)\}$  is also a (separable) measurement. The value  $\|(\Xi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(\rho)\|_{\text{SEP}}$  is therefore equal to 1 (where the equality follows by choosing the trivial measurement  $P_0 = \mathbb{1}$  and  $P_1 = 0$ , for instance).

Consequently, we have that  $\|\Psi_0 - \Psi_1\|_{\text{SEP}} = 2\alpha$ . As is proved in [17], it holds that

$$2\alpha = \|\Psi_0 - \Psi_1\|_{\text{NE}} < \|\Psi_0 - \Psi_1\|_{\diamond},$$

and so we have  $\|\Psi_0 - \Psi_1\|_{\text{NE}} = \|\Psi_0 - \Psi_1\|_{\text{SEP}} < \|\Psi_0 - \Psi_1\|_{\diamond}$  as claimed.

## V. CONCLUSIONS

We have introduced a new norm  $\|\cdot\|_{\text{LOCC}}$  on Hermiticity-preserving maps, representing a “weakened” version of the diamond norm. This norm is motivated by

an interest in the usefulness of entanglement for channel discrimination, and in particular whether entanglement in the input/ancillary systems remains useful when the final measurement is implementable by local operations and classical communication alone. We provided examples of pairs of channels where input entanglement is useful only if general measurements are available, while there are other cases where the advantage of using input entanglement is fully maintained in spite of the LOCC restriction on measurements.

In the context of bipartite state discrimination, one may study the gap in the probability of error associated to discrimination protocols that are or are not restricted (e.g., to be LOCC) [38]. This gap has been shown to depend on the total dimension of the underlying systems. In the present paper we showed two results that regard dimensions. One is that a restriction on the output measurement makes input entanglement useless if the output dimension is too small (one qubit), and the second is a dimensional lower bound on the gap that can be achieved between the diamond norm and the LOCC norm, where the relevant dimension is in this case the input one. One may consider in more detail what is the role of the input and output dimensions for the usefulness of entanglement when measurements are restricted.

In [17] it was proved that, for any entangled state, there 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. The result was obtained by resorting to the use of the fundamental characterization of entanglement in terms of linear maps [44]. The actual construction used in [17] completely fails when the output measurement is of the separable type. Indeed, in our view the main open problem concerning channel discrimination with LOCC measurements is whether the results of [17] can be extended to LOCC-restricted measurements. While we know that there are entangled input states useful for channel discrimination even with constrained output measurements, the possibility is left open that some entangled states become useless in this setting. If this is the case, it would be interesting to understand what kind of features make some states remain useful and some not.

An issue that has been the focus of a large part of the literature on the role of entanglement in channel discrimination is the question of what channels are better distinguished by the use of entanglement. Similarly, one can ask what channels that are better discriminated by the use of an entangled ancilla are still (or are not any more) better discriminated when we impose constraints on the measurements. The results presented in this work show that both situations occur. Is it possible to find large classes of channels for which input entanglement stays or fails to stay a useful resource?



## Acknowledgement

The authors would like to thank F. G. S. L. Brandão and T. Ito for discussions. WM acknowledges the sup-

port of the NSERC and QuantumWorks, MP acknowledges support from NSERC, QuantumWorks, and Ontario Centres of Excellence, and JW acknowledges support from NSERC, CIFAR, and QuantumWorks.

- 
- [1] A. Childs, J. Preskill, and J. Renes, *Journal of Modern Optics* **47**, 155 (2000).
  - [2] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, *Physical Review Letters* **87**, 270404 (2001).
  - [3] A. Acin, *Physical Review Letters* **87**, 177901 (2001).
  - [4] R. Duan, Y. Feng, and M. Ying, *Physical Review Letters* **103**, 210501 (2009).
  - [5] V. Giovannetti, S. Lloyd, and L. Maccone, *Science* **306**, 1330 (2004).
  - [6] A. Gilchrist, N. K. Langford, and M. A. Nielsen, *Physical Review A* **71**, 062310 (2005).
  - [7] B. Rosgen and J. Watrous, in *Proceedings of the 20th Annual Conference on Computational Complexity* (2005), pp. 344–354.
  - [8] M. F. Sacchi, *Physical Review A* **71**, 062340 (2005).
  - [9] M. F. Sacchi, *Physical Review A* **72**, 014305 (2005).
  - [10] S. Lloyd, *Science* **321**, 1463 (2008).
  - [11] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, *Phys. Rev. Lett.* **101**, 253601 (2008).
  - [12] B. Rosgen, in *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science* (2008), pp. 597–608.
  - [13] J. Watrous, *Quantum Information and Computation* **8**, 819 (2008).
  - [14] J. H. Shapiro and S. Lloyd, *New Journal of Physics* **11**, 063045 (2009), URL <http://stacks.iop.org/1367-2630/11/i=6/a=063045>.
  - [15] A. R. Usha Devi and A. K. Rajagopal, *Phys. Rev. A* **79**, 062320 (2009).
  - [16] A. Kitaev, *Russian Mathematical Surveys* **52**, 1191 (1997).
  - [17] M. Piani and J. Watrous, *Physical Review Letters* **102**, 250501 (pages 4) (2009).
  - [18] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Reviews of Modern Physics* **81**, 865 (pages 78) (2009).
  - [19] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, *Physical Review Letters* **85**, 4972 (2000).
  - [20] S. Virmani, M. Sacchi, M. Plenio, and D. Markham, *Physics Letters A* **288**, 62 (2001).
  - [21] J. Walgate and L. Hardy, *Physical Review Letters* **89**, 147901 (2002).
  - [22] H. Fan, *Physical Review Letters* **92**, 177905 (2004).
  - [23] M. Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, *Physical Review Letters* **93**, 170503 (2004).
  - [24] A. Cheffles, *Physical Review A* **69**, 050307(R) (2004).
  - [25] S. Ghosh, P. Joag, G. Kar, S. Kunkri, and A. Roy, *Physical Review A* **71**, 012321 (2005).
  - [26] J. Watrous, *Physical Review Letters* **95**, 080505 (2005).
  - [27] M. Nathanson, *Journal of Mathematical Physics* **46**, 062103 (pages 15) (2005).
  - [28] M. Hayashi, K. Matsumoto, and Y. Tsuda, *Journal of Physics A: Mathematical and General* **39**, 14427 (2006).
  - [29] M. Hayashi, D. Markham, M. Murao, M. Owari, and S. Virmani, *Physical Review Letters* **96**, 040501 (pages 4) (2006).
  - [30] H. Fan, *Physical Review A* **75**, 014305 (pages 4) (2007).
  - [31] R. Duan, Y. Feng, Z. Ji, and M. Ying, *Physical Review Letters* **98**, 230502 (2007).
  - [32] M. Owari and M. Hayashi, *New Journal of Physics* **10**, 013006 (2008).
  - [33] S. Bandyopadhyay and J. Walgate, *Journal of Physics A: Mathematical and Theoretical* **42**, 072002 (2009).
  - [34] W. Matthews and A. Winter, *Communications in Mathematical Physics* **285**, 161 (2009).
  - [35] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, *Physical Review Letters* **86**, 5807 (2001).
  - [36] T. Eggeling and R. F. Werner, *Physical Review Letters* **89**, 097905 (2002).
  - [37] D. DiVincenzo, D. Leung, and B. Terhal, *IEEE Transactions on Information Theory* **48**, 580 (2002).
  - [38] W. Matthews, S. Wehner, and A. Winter, *Communications in Mathematical Physics* **291**, 813 (2009).
  - [39] X.-F. Zhou, Y.-S. Zhang, and G.-C. Guo, *Physical Review Letters* **99**, 170401 (pages 4) (2007).
  - [40] R. Duan, Y. Feng, and M. Ying, *Physical Review Letters* **100**, 020503 (pages 4) (2008).
  - [41] G. Zauner, Ph.D. thesis, Universitat Wien (1999).
  - [42] D. Appleby, *Journal of Mathematical Physics* **46**, 052107 (2005).
  - [43] R. Werner and A. Holevo, *Journal of Mathematical Physics* **43**, 4353 (2002).
  - [44] M. Horodecki, P. Horodecki, and R. Horodecki, *Physics Letters A* **223**, 1 (1996).
  - [45] Both  $\text{Sep}(\mathcal{X} : \mathcal{Y})$  and  $\text{PPT}(\mathcal{X} : \mathcal{Y})$  are subsets of  $\text{Pos}(\mathcal{X} \otimes \mathcal{Y})$ . A positive operator is PPT if it remains positive under the action of partial transposition; a positive operator is separable if it can be expressed as  $\sum_i P_i \otimes Q_i$ , with  $P_i \in \text{Pos}(\mathcal{X})$ ,  $Q_i \in \text{Pos}(\mathcal{Y})$ .
  - [46] A mapping  $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  is *Hermiticity-preserving* if and only if  $\Phi(X) \in \text{Herm}(\mathcal{Y})$  for every  $X \in \text{Herm}(\mathcal{X})$ . For any choice of quantum channels  $\Phi_0, \Phi_1 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$  it holds that  $\Phi = \Phi_0 - \Phi_1$  is Hermiticity-preserving.
  - [47] It is not known if the set of LOCC measurements on a bipartite system  $\mathcal{Y} \otimes \mathcal{Z}$ , implementable with a finite number of rounds of communication, is closed. For the sake of simplicity, we will consider any measurement in the closure of the set of LOCC measurements to be an LOCC measurement, so that the maximum over  $\{P_0, P_1\}$  in (4) is a maximization over a compact set.
  - [48] A binary-valued measurement  $\{P_0, P_1\}$  is separable or PPT if both  $P_0$  and  $P_1$  are separable or PPT, respectively.
  - [49] In this picture, this is a local measurement, as both the output probe and the ancilla are in Bob's lab, but it is a global probe/ancilla measurement.