Two-way classical communication remarkably improves local distinguishability

Masaki Owari$^{1,2,5}$ and Masahito Hayashi$^{3,4}$

$^1$ Collaborative Institute of Nano Quantum Information Electronics, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku Tokyo 113-0033, Japan
$^2$ Department of Physics, Graduate School of Science, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku Tokyo 113-0033, Japan
$^3$ ERATO-SORST Quantum Computation and Information Project, JST, Hongo 5-28-3, Bunkyo-ku Tokyo 113-0033, Japan
$^4$ Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai, 980-8579, Japan
E-mail: m.owari@imperial.ac.uk, masakiowari@is.s.u-tokyo.ac.jp and hayashi@math.is.tohoku.ac.jp

New Journal of Physics 10 (2008) 013006 (25pp)
Received 23 August 2007
Published 14 January 2008
Online at http://www.njp.org/
doi:10.1088/1367-2630/10/1/013006

Abstract. We analyze the difference in the local distinguishability among the following three restrictions: (i) local operations and only one-way classical communications (one-way LOCC) are permitted; (ii) local operations and two-way classical communications (two-way LOCC) are permitted; and (iii) all separable operations are permitted. We obtain two main results concerning the discrimination between a given bipartite pure state and the completely mixed state with the condition that the given state should be detected perfectly. As the first result, we derive the optimal discrimination protocol for a bipartite pure state in cases (i) and (iii). As the second result, by constructing a concrete two-way local discrimination protocol, it is proven that case (ii) is much better than case (i), i.e. two-way classical communication remarkably improves the local distinguishability in comparison with one-way classical communication at least for a low-dimensional bipartite pure state.

$^5$ Author to whom any correspondence should be addressed.
1. Introduction

Recently, quantum communication has been investigated by many groups as a future technology. Similar to conventional information technology, practical quantum communication technology will require distributed information processing among two or more spatially separated parties. In order to treat this problem, it is necessary to clarify what kind of information processing is possible under respective constraints for permitted operations. In the quantum case, when our quantum system consists of distinct two parties \( A \) and \( B \), we often restrict our operations to local (quantum) operations and classical communications (LOCC) because sending quantum states over long distances is technologically more difficult than sending classical information [1].

Even in this restriction, we can consider the following two formulations: (i) the classical communication is restricted to the direction \( A \) to \( B \) (we can similarly treat the restriction of the opposite direction) and (ii) all parties are allowed to communicate classically with each other as much as they like. Case (i) is called one-way LOCC and case (ii) is called two-way LOCC.

Since, by definition, two-way LOCC apparently includes one-way LOCC, two-way LOCC is always more powerful than one-way LOCC in principle. However, due to the following two reasons, it is not easy to characterize the difference between the performance of the two situations. The first reason is that it is mathematically hard to rigorously evaluate the performance of a given two-way LOCC protocol, because the mathematical description of two-way LOCC is too complicated. The second reason is that for several simple tasks, the performance of two-way LOCC was actually shown to be the same as the performance of one-way LOCC. In fact, there are several settings that have no difference between one-way LOCC and two-way LOCC, e.g. LOCC convertibility of bipartite pure state [2], and Stein’s lemma bound in the simple asymptotic hypothesis testing of the \( n \)-tensor product of identical states [3].

On the other hand, in several settings, two-way LOCC is strictly more powerful than one-way LOCC. For example, the distillable entanglement (the amount of maximally entangled states which can be derived from a given state by LOCC) with two-way classical communication has been proven to be greater than that with one-way classical communication [4]. Also, this
type of comparison has been done by several papers [5]–[7] only on the discrimination among orthogonal states. Although several researchers treated this problem, they did not treat the discrimination among non-orthogonal states. In this paper, we compare the two performances quantitatively on the ‘local discrimination’ among states that are not necessarily orthogonal, whose purpose is discriminating given states by only LOCC with a single copy. In fact, the general discrimination problem is closely related to sending classical information via quantum channels [8] and quantum algorithms [9, 10].

In order to quantify the difference between two-way LOCC and one-way LOCC, in this paper, we concentrate on a simple setting: the local discrimination of the first state $\rho$ on a bipartite system $H$ from the second state $\tilde{\rho}$ under the condition where the first state $\rho$ should be detected perfectly. When both the states $\rho$ and $\tilde{\rho}$ are pure, there is no difference between one-way LOCC and two-way LOCC because any global discrimination protocol can be simulated by one-way LOCC [11, 12]. Surprisingly, as our result, we found that there usually exists non-negligible difference between two restrictions when the second state $\tilde{\rho}$ is the completely mixed state $\rho_{\text{mix}} := I / \dim H$. At first glance, this setting seems specific; however, due to the following six reasons, it is closely related to several research topics. Firstly, this type of analysis produces a bound of the number of perfectly locally distinguishable states. Secondly, as is explained later, there is a relation between the performance of local distinguishability and the amount of entanglement in the case of pure states. Thirdly, this kind of distinguishability is often treated in quantum complexity as a triviality of the coset state [10, 13]. Fourthly, when the second state $\tilde{\rho}$ is close to the completely mixed state $\rho_{\text{mix}}$, we obtain a similar conclusion because the power of our test is continuous concerning the second state. Fifthly, in the community of classical statistics, the problem of discriminating the given two distributions is widely accepted as the fundamental problem of hypothesis testing because the general hypothesis testing problem can be treated by using this type of problem [14]. Sixthly, as was mentioned in the preceding papers [8], hypothesis testing with two candidate states is closely related to quantum channel coding. Hence, it is suitable to treat this kind of local discrimination problem.

In order to analyze this problem in the respective settings, we introduce the minimum error probabilities to detect the complete mixed state $\beta_{\text{mix}}(\rho)$, $\beta_{\text{sep}}(\rho)$ and $\beta_{\leftrightarrow}(\rho)$ by one-way LOCC, two-way LOCC and separable operations, respectively. Indeed, these functions are considered as appropriate measures of the local distinguishability because they give not only the minimum error probability of the above problem, but also the upper bound of the size of locally distinguishable sets in general perfect local discrimination problems [15]. Under this formulation, we first analyze the local distinguishability by means of one-way LOCC and separable operations, and derive the optimal discrimination protocol with one-way LOCC and separable operations; we should note that the minimum error probability $\beta_{\text{sep}}(\rho)$ with separable operations gives a lower bound for the minimum error probability $\beta_{\leftrightarrow}(\rho)$ with two-way LOCC. After that, constructing a concrete two-way local discrimination protocol, we show that two-way classical communication remarkably improves the local distinguishability in comparison with the local discrimination by one-way classical communication at least for a low-(less than five) dimensional bipartite pure state. Indeed, since the power of our test is continuous concerning the first and the second states, our result indicates that two-way classical communication remarkably improves the local distinguishability in a wider class of the first and second states. Moreover, as a by-product, we extend the characterization of local distinguishability by one-way LOCC by Cohen [7] to a set of mixed states.
This paper is organized as follows: in section 2, we introduce the discrimination problem between an arbitrary given state $\rho$ and a completely mixed state $\rho_{\text{mix}}$ on a bipartite system $H$ under the condition that the given state is detected perfectly. Then, we explain another meaning of $\beta_{\rightarrow}(\rho), \beta_{\leftarrow}(\rho)$ and $\beta_{\text{sep}}(\rho)$ from the viewpoint of general local discrimination problems. In section 3, constructing the optimal separable POVM (positive operator valued measure) for the local discrimination, we prove that $D\beta_{\text{sep}}(|\Psi\rangle) - 1$ coincides with the entanglement monotone called robustness of the entanglement for a bipartite pure state, where $D$ is the dimension of the bipartite Hilbert space $\mathcal{H}$. In section 4, we show that the amount $D\beta_{\rightarrow}(|\Psi\rangle)$ with one-way LOCC coincides with the Schmidt rank (the rank of the reduced density matrix) of the states. Also, as a corollary, we extend Cohen’s characterization to a set of mixed states. Finally, in section 5, constructing a concrete three-step two-way LOCC discrimination protocol, we derive an upper bound for $\beta_{\leftarrow}(\rho)$. Calculating this upper bound analytically and also numerically, we show that $\beta_{\leftarrow}(|\Psi\rangle)$ is strictly smaller than $\beta_{\rightarrow}(|\Psi\rangle)$, and moreover, $\beta_{\rightarrow}(\rho)$ and $\beta_{\text{sep}}(\rho)$ give almost the same value for a lower dimensional bipartite pure state; these results can be seen in figures 2–6. As a result, we conclude that two-way classical communication remarkably improves the local distinguishability in comparison with one-way classical communication for a low-dimensional pure state at least in the present problem setting.

2. Local discrimination between an arbitrary state and the completely mixed state

In this paper, we treat the bipartite system $\mathcal{H} := \mathcal{H}_A \otimes \mathcal{H}_B (\dim \mathcal{H} = D)$ composed of two finite-dimensional subsystems $\mathcal{H}_A$ and $\mathcal{H}_B$. In the following sections, we often focus on the case when $\rho$ is pure. In such a case, we assume that the dimension $d$ of $\mathcal{H}_A$ is equal to that of $\mathcal{H}_B$. Note that the given pure state belongs to the composite system of the same-dimensional subsystem. Then, the dimension ($D$) of the Hilbert space $\mathcal{H}$ is equal to $d^2$. In the composite system $\mathcal{H}$, we call a positive operator $T$ with $0 \leq T \leq I$ a one-way LOCC POVM element, where $I$ is an identity operator on $\mathcal{H}$, if the two-valued POVM $\{T, I - T\}$ can be implemented by the one-way LOCC; we also define a two-way LOCC POVM element and a separable POVM element in the same manner by using the two-way LOCC and the separable operations instead of the one-way LOCC, respectively [16]. We write a set of one-way LOCC, two-way LOCC, separable POVM elements, and all (global) POVM elements as $\mathcal{T}_{\rightarrow}, \mathcal{T}_{\leftarrow}, \mathcal{T}_{\text{sep}}$ and $\mathcal{T}_g$. Obviously, they satisfy the relation $\mathcal{T}_{\rightarrow} \subset \mathcal{T}_{\leftarrow} \subset \mathcal{T}_{\text{sep}} \subset \mathcal{T}_g$. We can see that the condition $T \in \mathcal{T}_c$ is equivalent to the condition $I - T \in \mathcal{T}_c$, where $c$ can be either $\rightarrow, \leftrightarrow, \text{sep}$ or $g$.

In this paper, we discuss the comparison of the performance of the local discrimination in the case of the one-way LOCC, the two-way LOCC, and the separable operations. In order to find this difference, although there are many problem settings for the local discrimination, we particularly focus on one of the simplest problem settings as follows: we consider local discrimination of a given arbitrary state $\rho$ and another state $\tilde{\rho}$, and investigate how well we can detect $\tilde{\rho}$ under the additional condition that we do not make any error in detecting $\rho$ when the second state $\tilde{\rho}$ is the completely mixed state $\rho_{\text{mix}} \overset{\text{def}}{=} \frac{1}{D} I_n (D = \dim \mathcal{H})$; namely, by only LOCC, how well we can distinguish a given entangled state $\rho$ from the white noise state $\rho_{\text{mix}}$ without making any error in judging the given state is $\rho_{\text{mix}}$ when the real state is $\rho$.

Our problem can be written down rigorously as follows. We measure an unknown state chosen from two candidates $\{\rho, \tilde{\rho}\}$ by the two-values POVM $\{T, I - T\}$, where $T \in \mathcal{T}_{\rightarrow}, \mathcal{T}_{\leftarrow}, \mathcal{T}_{\text{sep}}$ or $\mathcal{T}_g$; that is, if we get the result corresponding to $T$, then we decide that the unknown state is in $\rho$, and if we get the result corresponding to $I - T$, then we decide that...
the unknown state is in $\tilde{\rho}$. We consider two kinds of error probability as follows: the type 1 error probability $\text{Tr}\rho(I - T)$ and the type 2 error probability $\text{Tr}\tilde{\rho}T$; these are common terms in the field of ‘quantum hypothesis testing’ [3], where these two different error probabilities are treated in an asymmetric way. In this case, the type 1 error probability corresponds to the error probability that the real state is $\rho$ and our decision is $\tilde{\rho}$, and the type 2 error probability corresponds to the error probability that the real state is $\tilde{\rho}$ and our decision is $\rho$. Thus, our problem is to minimize the type 2 error probability $\text{Tr}\tilde{\rho}T$ under the additional condition that the type 1 error probability $\text{Tr}\rho(I - T)$ must be 0. Thus, we focus on the following minimum of the type 2 error probability:

$$\beta_c(\rho\|\tilde{\rho}) := \min\{\text{Tr}(\tilde{\rho}T) | T \in T_c, \text{Tr}\rho T = 1\},$$

(1)

where $c = \to$ (one-way LOCC), $\leftrightarrow$ (two-way LOCC), sep (separable operations) and $g$ (global operations). When both the states $\rho$ and $\tilde{\rho}$ are pure states $|\Phi\rangle$ and $|\Psi\rangle$, this quantity does not depend on whether $c = \to, \leftrightarrow, \text{sep}$ or $g$, and is calculated as

$$\beta_c(|\Phi\rangle\||\Psi\rangle) = |\langle\Phi|\Psi\rangle|^2$$

(2)

for $c = \to, \leftrightarrow, \text{sep}$ and $g$. This is because any discriminating protocol between two pure bipartite states can be simulated by one-way LOCC when we focus only on the distribution of the outcome [11, 12]. In this paper, we focus on the minimum of the type 2 error probability in the case of $\tilde{\rho} = \rho_{\text{mix}}$:

$$\beta_c(\rho) := \beta_c(\rho\|\rho_{\text{mix}}) = \frac{t_c(\rho)}{D},$$

(3)

where $t_c(\rho)$ is defined as

$$t_c(\rho) = \min\{\text{Tr}T | T \in T_c, \text{Tr}\rho T = 1\}$$

(4)

and $D$ is the dimension of the whole system $\mathcal{H}$. That is, $t_c(\rho)$ is in proportion to the minimum of the type 2 error probability $\beta_c(\rho)$ of one-way LOCC, two-way LOCC, separable POVM and global POVM in the case where $c = \to, \leftrightarrow, \text{sep}$ and $g$, respectively. Trivially,

$$t_g(\rho) = \text{rank } \rho.$$ 

(5)

Obviously, $t_c(\rho)$ satisfies the inequality $t_g(\rho) \leq t_{\text{sep}}(\rho) \leq t_{\to}(\rho) \leq t_{\leftrightarrow}(\rho)$; as a matter of course, $\beta_c(\rho)$ also satisfies a similar inequality. Note that by normalizing $\beta_c(\rho)$ as the above equation (3), the resulting function $t_c(\rho)$ is no longer a function depending both on $\rho$ and $\rho_{\text{mix}}$, but is a function depending only on $\rho$.

**Remark 1.** In the quantum information community, many papers treat the Bayesian framework, in which the Bayesian prior distribution is assumed [17]–[19]. However, in the statistics community, a non-Bayesian framework is more widely accepted, in which no Bayesian prior distribution is assumed [14]. This is because it is usually quite difficult to decide the Bayesian prior distribution based on prior knowledge. In order to resolve this difficulty, they often treat the two kinds of error probabilities in an asymmetric way in hypothesis testing without assuming prior distribution because the importance of the two errors are not equal in a usual case, e.g. Neyman–Pearson lemma [14], Stein’s lemma [20] and Hoeffding bound [21]. These quantum cases are treated by several papers [22]–[25]. In this paper, according to a conventional statistics framework, we focus on the error probabilities of the first and second, and minimize the second kind of error probability under the constraint of the first one.
Here, we explain the reason why we choose the above special problem of discrimination of an arbitrary state $\rho$ from the completely mixed state $\rho_{\text{mix}}$ and the reason why we add the above additional condition of perfect detection of $\rho$. As we have mentioned before, the first reason is that this additional condition makes the analysis of the problem significantly easier. Actually, as we will see later in this paper, we can derive the optimal POVM of this restricted local-discrimination problem with respect to each one-way LOCC and separable operations for a bipartite pure state. As a result, we make the difference between one-way LOCC and two-way LOCC clear for our local-discrimination problem; this is our main purpose in this paper. Note that it is generally a hard problem to find an optimal protocol for a local-discrimination problem, and only in very limited situations, optimal local-discrimination protocols are known $[5, 11, 12]$. The second reason is that we can clearly see the relationship between local distinguishability and entanglement of a state in this problem setting. In the previous paper $[15]$, we showed the relationship between local distinguishability of a set of states and an average of the values of entanglement monotones for the states in terms of inequalities. However, in this paper, we will show that the minimum error probability $\beta_c(\rho)$ of our problem is in proportion to entanglement monotones in the case of one-way LOCC ($c = \rightarrow$) and separable operations ($c = \text{sep}$) at least for bipartite pure states except for an unimportant constant factor. The third reason is that the minimum error probability $\beta_c(\rho)$ can give a bound of local distinguishability for a more general local-discrimination problem: suppose that a set of states $\{\rho_i\}_{i=1}^N$ is perfectly locally distinguishable by one-way LOCC ($c = \rightarrow$), two-way LOCC ($c = \leftrightarrow$), or separable ($c = \text{sep}$) POVM. From the result obtained in the previous paper, $t_c(\rho_i)$ (which corresponds to $d(\rho)$) gives an upper bound of $N_c$ as $[15]$, $t_c(\rho_i) = 1/\beta_c(\rho_i)$,

$$N_c \leq D/t_c(\rho_i) = 1/\beta_c(\rho_i), \quad (6)$$

where $t_c(\rho_i)$ and $\beta_c(\rho_i)$ are the average of $\{t_c(\rho_i)\}_{i=1}^N$ and $\{\beta_c(\rho_i)\}_{i=1}^N$, respectively $[15]$. Thus, $\beta_c(\rho)$ can be considered as an appropriate measure of local distinguishability in an original operational sense, and also as a function whose average gives an upper bound for the locally distinguishable sets of states. Therefore, we investigate the difference of local distinguishability of $\rho$ by one-way LOCC POVM, two-way LOCC POVM and separable POVM in terms of $\beta_c(\rho)$ in the following sections.

3. Local discrimination by separable POVM

In this section, we investigate the minimum type 2 error probability $\beta_{\text{sep}}(\rho) = \frac{t_{\text{sep}}(\rho)}{D}$ in terms of separable POVMs, which are given by $\{N_i \otimes M_i\}_i$ with the conditions $\sum_i N_i \otimes M_i = I$, $N_i \geq 0$ and $M_i \geq 0$. The main purpose of this section is to prove the following theorem:

**Theorem 1.** The inequality

$$t_{\text{sep}}(\rho) \geq \max\{(\text{Tr} \sqrt{\rho_A})^2, (\text{Tr} \sqrt{\rho_B})^2\} \quad (7)$$

holds for a bipartite state $\rho$ on $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\rho_A$ and $\rho_B$ are the reduced density matrices of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Any pure state satisfies its equality. In other words, the following inequality concerning the minimum error probability $\beta_{\text{sep}}(\rho)$ holds:

$$\beta_{\text{sep}}(\rho) \geq \frac{1}{D} \max\{(\text{Tr} \sqrt{\rho_A})^2, (\text{Tr} \sqrt{\rho_B})^2\} \quad (8)$$
For a bipartite pure state, the right-hand side of equation (7) is proportional to an entanglement monotone called the global robustness of entanglement $R_g(|\Psi\rangle)$ except for an unimportant constant term [26].

Applying theorem 1 for equation (6), we can immediately derive the following corollary concerning the perfect discrimination of a given set of states in terms of separable operations:

**Corollary 1.** If a set of states $\{\rho_i\}_{i=1}^N$ is perfectly distinguishable by separable operations, then the set of states $\{\rho_i\}_{i=1}^N$ satisfies the following inequality:

$$ N \leq D/\max\{ (\text{Tr}\sqrt{\rho_i^A})^2, (\text{Tr}\sqrt{\rho_i^B})^2 \}, \tag{9} $$

where $\max\{ (\text{Tr}\sqrt{\rho_i^A})^2, (\text{Tr}\sqrt{\rho_i^B})^2 \}$ is the average of $\max\{ (\text{Tr}\sqrt{\rho_i^A})^2, (\text{Tr}\sqrt{\rho_i^B})^2 \}$ for all $1 \leq i \leq N$.

The above inequality is weaker than the inequality (6). However, the inequality (6) is superior to the inequality (6) in terms of the efficiency of the computation; that is, in general, we cannot efficiently compute the bound in equation (6), since the function $t_{\text{sep}}(\rho)$ includes the big variational problem.

### 3.1. Pure states case

First, for a technical reason, we concentrate on the pure states case, and define a set of POVM elements $\mathcal{T}_{\text{sep}}$ by

$$ \mathcal{T}_{\text{sep}} \overset{\text{def}}{=} \left\{ T | T \leq I_{AB}, T = \sum_i N_i \otimes M_i, \forall i, N_i \geq 0, M_i \geq 0 \right\}. $$

$\mathcal{T}_{\text{sep}}$ is a set of POVM elements that can be decomposed into a separable form; we say that a positive linear operator $M$ has a separable form if $M/\text{Tr}M$ is a separable state. Since the definition of $\mathcal{T}_{\text{sep}}$ is equivalent to the definition of $\mathcal{T}_{\text{sep}}$ except the condition $I - T \in \mathcal{T}_{\text{sep}}$, $\mathcal{T}_{\text{sep}}$ is a subset of $\mathcal{T}_{\text{sep}}$. Note that even if $T \in \mathcal{T}_{\text{sep}}$, $I - T$ does not necessarily satisfy $I - T \in \mathcal{T}_{\text{sep}}$; that is, $\mathcal{T}_{\text{sep}}$ does not coincide with $\mathcal{T}_{\text{sep}}$. For example, suppose a set of states $\{|\Psi_i\rangle\}_{i=1}^m \in \mathcal{H}_A \otimes \mathcal{H}_B (m < \dim \mathcal{H})$ is an unextendable product basis, and a POVM $T$ is defined as $T \overset{\text{def}}{=} \sum_i |\Psi_i\rangle \langle \Psi_i|$. Then, $T$ belongs to $\mathcal{T}_{\text{sep}}$, but not to $\mathcal{T}_{\text{sep}}$, since $I - T = I - \sum_i^m |\Psi_i\rangle \langle \Psi_i|$ is in proportion to a (bound) entangled state, and does not have a separable form [27]. Similarly, we can define $t_{\text{sep}}$ as

$$ t_{\text{sep}}(\rho) \overset{\text{def}}{=} \min \left\{ \text{Tr} T | T \in \mathcal{T}_{\text{sep}}, \text{Tr} \rho T = 1 \right\}. \tag{10} $$

By definition, $t_{\text{sep}}(\rho)$ apparently gives a lower bound of $t_{\text{sep}}(\rho) = d^2 \beta_{\text{sep}}(\rho)$; that is, for all $\rho \in S(\mathcal{H})$,

$$ t_{\text{sep}}(\rho) \leq t_{\text{sep}}(\rho) = d^2 \beta_{\text{sep}}(\rho). \tag{11} $$

Then, we can see that $t_{\text{sep}}(\rho)$ is actually equal to $d(\rho)$, which is defined in theorem 1 of [15] as:

$$ d(\rho) \overset{\text{def}}{=} \min \left\{ \frac{1}{\text{Tr} \rho \omega}, 0 \leq \frac{\omega}{\text{Tr} \rho \omega} \leq 1, \omega \in \text{SEP} \right\}, \tag{12} $$

where SEP is the set of all separable states. We can easily check this fact just by defining $T \overset{\text{def}}{=} \frac{\omega}{\text{Tr} \rho \omega}$; then, $T$ satisfies $0 \leq T \leq I$, $\text{Tr} \rho T = 1$ and $T \in \mathcal{T}_{\text{sep}}$. From theorem 2 of the
previous paper [15], for an arbitrary multipartite pure state $|\Psi\rangle$, $t_{\text{SEP}}(\rho)$ satisfies the following inequality:

$$t_{\text{SEP}}(|\Psi\rangle) = d(|\Psi\rangle) \geq 1 + R_g(|\Psi\rangle),$$

and $R_g(|\Psi\rangle)$ is the global robustness of entanglement [26] defined as:

$$R_g(\rho) \overset{\text{def}}{=} \min \left\{ t \geq 0 \mid \exists \text{ a state } \Delta, \text{ s.t. } \frac{1}{1+t}(\rho + t\Delta) \in \text{SEP} \right\}, \tag{14}$$

For a bipartite pure state, we can know more details of $t_{\text{SEP}}(|\Psi\rangle)$ as follows. First, it was proven that $t_{\text{SEP}}(|\Psi\rangle)$ coincides with the robustness of entanglement $R_g(|\Psi\rangle)$ for an arbitrary pure bipartite state $|\Psi\rangle$ [28]. This fact can be seen by checking that the optimal states of $R_g(\rho)$, which was derived in [26], satisfy the condition of $d(\rho)$; the optimal state of $R_g(\rho)$ is also an optimal state of $d(\rho)$. Moreover, we know that the value of $R_g(|\Psi\rangle)$ is given by the following formula for a bipartite state $|\Psi\rangle$ [26]:

$$R_g(\rho) = \left( \sum_i \sqrt{\lambda_i} \right)^2 - 1,$$

where $\{\lambda_i\}_{i=1}^d$ are the Schmidt coefficients of $|\Psi\rangle$; $|\Psi\rangle$ can be decomposed as $|\Psi\rangle = \sum_i \sqrt{\lambda_i} |e_i\rangle \otimes |f_i\rangle$ by choosing an appropriate orthonormal basis set of the local Hilbert spaces $\{|e_i\rangle\}_{i=1}^d \subset \mathcal{H}_A$ and $\{|f_i\rangle\}_{i=1}^d \subset \mathcal{H}_B$. Thus, we derive

**Lemma 1.** For a bipartite pure state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$.

$$t_{\text{SEP}}(|\Psi\rangle) = d(|\Psi\rangle) = 1 + R_g(|\Psi\rangle) = \left( \sum_i \sqrt{\lambda_i} \right)^2. \tag{15}$$

Although this lemma is a known result [28], as a preparation for the proof of the next theorem, we give a complete proof of equation (15), in which we prove directly the equation $t_{\text{SEP}}(|\Psi\rangle) = (\sum_i \sqrt{\lambda_i})^2$ from the definition of $t_{\text{SEP}}(|\Psi\rangle)$.

**Proof.** This proof is divided into two steps. In the first step, we prove that $(\sum_i \sqrt{\lambda_i})^2$ is the lower bound of $t_{\text{SEP}}(|\Psi\rangle)$. Then, in the second step, we construct POVM element $T$ which attains this lower bound. For convenience, we define $|M_\Psi\rangle = (1/\sqrt{d}) \sum_{i=1}^d |e_i\rangle \otimes |f_i\rangle$ where $\{|e_i\rangle\}_{i=1}^d$ and $\{|f_i\rangle\}_{i=1}^d$ are the Schmidt basis of $|\Psi\rangle$; thus, $|M_\Psi\rangle$ is the maximum entangled state sharing the Schmidt basis with $|\Psi\rangle$. Then, we derive $d(|M_\Psi\rangle^2) = (\sum_i \sqrt{\lambda_i})^2$. As the first step, we prove the following inequality:

$$t_{\text{SEP}}(|\Psi\rangle) = \min \{ \text{Tr}T | 0 \leq T \leq I, T \text{ is sep.} \} \geq \min \{ d(M_\Psi T | M_\Psi) | 0 \leq T \leq I, T \text{ is sep.} \} \geq d(|M_\Psi\rangle^2). \tag{16}$$

To prove the first inequality (16), since both $\text{Tr}T$ and $d(M_\Psi T | M_\Psi)$ are linear for $T$, it is enough to prove only the case that $T$ can be written down as $T = |a\rangle \langle a| \otimes |b\rangle \langle b|$ by using un-normalized vectors $|a\rangle$ and $|b\rangle$. Suppose $|a\rangle = \sum_{i=1}^d \alpha_i |e_i\rangle$ and $|b\rangle = \sum_{i=1}^d \beta_i |f_i\rangle$. Then, using Schwarz’s inequality, we can prove:

$$\text{Tr}T = \left( \sum_i |\alpha_i|^2 \right) \left( \sum_j |\beta_j|^2 \right) \geq \left( \sum_j |\alpha_j\beta_j|^2 \right)^2 = d(M_\Psi T | M_\Psi).$$
For the second inequality (16), since the relations $\langle \Psi | T | \Psi \rangle = 1$ and $T \leq I$ deduce that $| \Psi \rangle$ is an eigenvector of the largest eigenvalue 1 of $T$, we derive $T \geq | \Psi \rangle \langle \Psi |$. Therefore, the inequality $d \langle M_{\Psi} | T | M_{\Psi} \rangle \geq d \langle \langle \langle M_{\Psi} | \Psi \rangle \rangle_{\Psi} \rangle$ is derived by taking the mean value with respect to $| \Psi \rangle$.

As the second step, we construct an example of POVM element $T$ which achieves the lower bound we derived above. Define $T_0$ as $T_0 = | a \rangle \langle a | \otimes | b \rangle \langle b |$, where $| a \rangle = \sum_{i=1}^d (\lambda_i)^{1/4} | e_i \rangle$ and $| b \rangle = \sum_{i=1}^d (\lambda_i)^{1/4} | f_i \rangle$; then, $T_0$ satisfies $\text{Tr} T_0 = (\sum_j \sqrt{\lambda_j})^2$. Moreover, since $P_0 T_0 P_0 = | \Psi \rangle \langle \Psi |$, where $P_0 = \sum_{i=1}^d | e_i \rangle \langle e_i | \otimes | f_i \rangle \langle f_i |$, $T_0$ satisfies $\langle \Psi | T_0 | \Psi \rangle = 1$. Since $T_0$ apparently satisfies $0 \leq T_0$, the inequality $T \leq I$ is the only remaining condition which the optimal POVM element $T$ attaining the equality $\text{Tr} T = t_{\text{sep}}(| \Psi \rangle)$ must satisfy. Since $T_0$ does not generally satisfy the inequality $T_0 \leq I$, we construct a new POVM element $T$ which satisfies $0 \leq T \leq I$ from $T_0$. In order to construct the POVM element $T$ from $T_0$, we use the twirling technique here. We define a family of local unitary operators $U_{-\theta}$ parameterized by $\theta = \{ \theta \}_{i=1}^d$ as follows:

$$U_{-\theta} = \left( \sum_{j=1}^d e^{i\theta_j} | e_j \rangle \langle e_j | \right) \otimes \left( \sum_{k=1}^d e^{-i\theta_k} | f_k \rangle \langle f_k | \right).$$ (17)

Note that $(\mathcal{H}^d, U_{-\theta})$ is a unitary representation of the compact topological group $\widehat{U(1) \times \cdots \times U(1)}$; by means of a unitary representation of a compact topological group, we implement the ‘twirling’ operation (the averaging over parameters) for a state (or POVM) [29]. Then, we define $T$ as the operator which is constructed by twirling $U_{-\theta} T_0 U_{-\theta}^\dagger$ over parameters $\theta = \{ \theta \}_{i=1}^d$. Since by action of the twirling operation, a given state is projected onto the subspace of all invariant elements of the group action [29], we can calculate $T$ as follows:

$$T \overset{\text{def}}{=} \int_0^{2\pi} \cdots \int_0^{2\pi} U_{-\theta} T_0 U_{-\theta}^\dagger \theta_1 \cdots \theta_d$$

$$= \left( \sum_{j=1}^d | e_j \rangle \langle e_j | \otimes | f_j \rangle \langle f_j | \right) T_0 \left( \sum_{j=1}^d | e_j \rangle \langle e_j | \otimes | f_j \rangle \langle f_j | \right)$$

$$+ \sum_{j \neq k} \left( | e_j \rangle \langle e_j | \otimes | f_k \rangle \langle f_k | \right) T_0 \left( | e_j \rangle \langle e_j | \otimes | f_k \rangle \langle f_k | \right)$$

$$= \left( \sum_{i=1}^d \sqrt{\lambda_i} | e_i \rangle \otimes | f_i \rangle \right) \left( \sum_{j=1}^d \sqrt{\lambda_j} | e_j \rangle \otimes | f_j \rangle \right) + \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} | e_i \rangle \langle e_i | \otimes | f_j \rangle \langle f_j |.$$ 

Since $\sqrt{\lambda_i \lambda_j} \leq 1$, $T \leq I$. Moreover, $T$ satisfies $0 \leq T \leq I$, $\langle \Psi | T | \Psi \rangle = 1$, and is in the separable form; we only applied the local unitary $U_{-\theta}$ to the un-normalized product state $T_0$ and, then, took an average over parameters $\theta$. Thus, we derive the inequality $t_{\text{sep}}(| \Psi \rangle) \leq \text{Tr} T = \left( \sum_i \sqrt{\lambda_i} \right)^2$. Since we have already proven the converse inequality, we conclude that $t_{\text{sep}}(| \Psi \rangle) = \left( \sum_i \sqrt{\lambda_i} \right)^2$. □

Finally, by means of lemma 1, we can derive the following theorem, i.e. theorem 1 in the pure states case:

New Journal of Physics 10 (2008) 013006 (http://www.njp.org/)
Theorem 2. For a bipartite pure state $|\Psi\rangle$,
\[
\beta_{\text{sep}}(|\Psi\rangle) = \frac{1}{d^2} t_{\text{sep}}(|\Psi\rangle) = \frac{1}{d^2} \left( 1 + R_{\text{g}}(|\Psi\rangle) \right)
\]
\[
= \frac{1}{d^2} \left( \sum_i \sqrt{\lambda_i} \right)^2 = \frac{1}{d^2} (\text{Tr} \sqrt{\rho})^2 = \frac{1}{d^2} (\text{Tr} \sqrt{\rho_B})^2,
\]
where $\{\lambda_i\}_{i=1}^d$ is the Schmidt coefficients of $|\Psi\rangle$.

Proof. Since by the definition $t_{\text{sep}}(|\Psi\rangle) \leq t_{\text{sep}}(|\Psi\rangle)$, all that we need to prove is that the optimal POVM $T$ for $t_{\text{sep}}(|\Psi\rangle)$ is also the optimal POVM for $t_{\text{sep}}(|\Psi\rangle)$; that is, $I - T$ also has a separable form.

As we have already shown in the proof of lemma 1, the optimal POVM $T$ for $t_{\text{sep}}(|\Psi\rangle)$ can be written down as
\[
T = \left( \sum_{i=1}^d \sqrt{\lambda_i} |e_i \rangle \otimes |f_i \rangle \right) \left( \sum_{i=1}^d \sqrt{\lambda_i} (e_i | \otimes \langle f_i |) + \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} |e_i \rangle (\langle e_i | \otimes |f_j \rangle (|f_j \rangle |),
\]
where $\{|e_i \rangle \otimes |f_i \rangle\}_{ij}$ and $\{\lambda_i\}_{i=1}^d$ are the Schmidt basis and the Schmidt coefficients corresponding to $|\Psi\rangle = \sum_{i=1}^d \sqrt{\lambda_i} |e_i \rangle$, respectively. Suppose
\[
\mathcal{T}_0 \overset{\text{def}}{=} \frac{1}{2} \sum_{i \neq j} (|a_{ij} \rangle |a_{ij} \rangle \otimes |b_{ij} \rangle |b_{ij} \rangle + \sum_{i \neq j} \lambda_i \lambda_j + (\sqrt{\lambda_i - \sqrt{\lambda_j}})^2 |e_i f_j \langle e_i f_j |,
\]
where $|a_{ij} \rangle$ and $|b_{ij} \rangle$ are defined as $|a_{ij} \rangle \overset{\text{def}}{=} (\lambda_j)^{1/4} |e_i \rangle - (\lambda_i)^{1/4} |e_j \rangle$ and $|b_{ij} \rangle \overset{\text{def}}{=} (\lambda_j)^{1/4} |f_i \rangle + (\lambda_i)^{1/4} |f_j \rangle$ for $i \neq j$, respectively. Then, as is proven in appendix A, the relation
\[
\int_0^{2\pi} \cdots \int_0^{2\pi} U_{\theta} \mathcal{T}_0 U_{\theta}^\dagger \, d\theta_1 \cdots d\theta_d = I - T
\]
holds, where a local unitary operator $U_{\theta}$ is defined as equation (17). By the definition, $\int_0^{2\pi} \cdots \int_0^{2\pi} U_{\theta} \mathcal{T}_0 U_{\theta}^\dagger \, d\theta_1 \cdots d\theta_d$ is apparently a separable POVM element. Therefore, we can conclude the equality $t_{\text{sep}}(|\Psi\rangle) = t_{\text{sep}}(|\Psi\rangle)$ for a bipartite pure state. By means of lemma 1, we derive equation (18).

Thus, $t_{\text{sep}}(|\Psi\rangle)$ is equivalent to $1 + R_{\text{g}}(|\Psi\rangle)$, and the minimum type 2 error probability $\beta_{\text{sep}}(|\Psi\rangle)$ only depends on the global robustness of entanglement $R_{\text{g}}(|\Psi\rangle)$ for a bipartite pure state $|\Psi\rangle$. In this case, the optimal POVM $\{T, I - T\}$ can be derived by using equation (19) as the definition of the POVM element $T$.

We should note that theorem 6 not only gives a way to calculate the minimum type 2 error probability under separable operations $\beta_{\text{sep}}(|\Psi\rangle)$, but this theorem gives a complete relationship between the local distinguishability of a bipartite state under separable operations and the entanglement of the state. In the previous paper [15], it was shown that the global robustness of entanglement $R_{\text{g}}(|\Psi\rangle)$ gives an upper bound for the maximum number of distinguishable states under separable operations. However, the present result shows that $R_{\text{g}}(|\Psi\rangle)$ is nothing but the local distinguishability (against the completely mixed state) itself at least for a bipartite pure
state. In other words, it is shown that robustness of entanglement has rigorously operational meaning for bipartite pure states in terms of the local discrimination from the completely mixed state $\rho_{\text{max}}$.

3.2. Mixed states case

Now, we prove theorem 1 for a general mixed bipartite state.

Proof (Theorem 1). First we prove the inequality $t_{\text{sep}}(\rho) \geq (\text{Tr}\sqrt{\rho_A})^2$. Adding the system $B'$, we choose a purification $|\Phi\rangle$ of $\rho$. In the following, we will prove the inequality $t_{\text{sep}}(\rho) \geq t_{\text{sep}}(|\Phi\rangle)$. If this inequality holds, applying equations (11) and (15), we obtain $t_{\text{sep}}(\rho) \geq t_{\text{sep}}(|\Phi\rangle) = (\text{Tr}\sqrt{\rho_A})^2$.

Define a separable positive operator $T = \sum_i p_i |e_i\rangle \langle e_i| \otimes |f_i\rangle \langle f_i|$, $e_i \in \mathcal{H}_A$, $f_i \in \mathcal{H}_B$, $\|f_i\| = 1$, $\|e_i\| = 1$ such that $0 \leq T \leq I_{AB}$ and $\text{Tr}_B T = 1$. Thus, $\langle \Phi | \sum_i p_i |e_i\rangle \langle e_i| \otimes |f_i\rangle \langle f_i| \otimes I_B |\Phi\rangle = 1$. Now, we focus on the bipartite system $A$ and $BB'$. Then, we choose the state $|f_i\rangle (\|f_i\| = 1)$ on $\mathcal{H}_{BB'}$ such that $\text{Tr}_A |\Phi\rangle \langle \Phi| (|e_i\rangle \langle e_i| \otimes I_{BB'}) = c_i |f_i\rangle \langle f_i|$, where $c_i$ is the normalizing constant. Define the state $|f_i^\prime\rangle (\|f_i^\prime\| = 1)$ on $\mathcal{H}_{BB'}$ by

$$|f_i^\prime\rangle \langle f_i^\prime| := \frac{1}{\langle f_i| P_i |f_i\rangle} P_i |f_i\rangle \langle f_i| P_i \leq P_i,$$

where the projection $P_i$ is defined by $P_i := |f_i\rangle \langle f_i| \otimes I_B$. Since $\langle f_i| P_i |f_i\rangle = \langle f_i| f_i^\prime\rangle \langle f_i^\prime| f_i\rangle$, $\text{Tr}_B (|e_i\rangle \langle e_i| \otimes |f_i^\prime\rangle \langle f_i^\prime| \otimes I_B') = |\langle f_i| f_i^\prime\rangle \langle f_i^\prime| |\langle f_i| f_i^\prime\rangle \langle f_i^\prime| |\langle f_i| f_i^\prime\rangle \langle f_i^\prime| |\langle f_i| f_i^\prime\rangle \langle f_i^\prime| |\rangle\rangle = \langle \Phi | (|e_i\rangle \langle e_i| \otimes |f_i^\prime\rangle \langle f_i^\prime| \otimes I_{BB'}) |\Phi\rangle$.

Thus, the relations

$$T' := \sum_i p_i |e_i\rangle \langle e_i| \otimes |f_i^\prime\rangle \langle f_i^\prime| \leq \sum_i p_i |e_i\rangle \langle e_i| \otimes |f_i\rangle \langle f_i| \otimes I_{BB'} \leq I_{ABB'} (|\Phi\rangle \langle T' | \langle \Phi\rangle) = \text{Tr}_B T = 1$$

hold. Moreover, $T'$ satisfies the equality $\text{Tr} T' = \sum_i p_i = \text{Tr} T$. Thus, the inequality $t_{\text{sep}}(\rho) \geq t_{\text{sep}}(|\Phi\rangle)$ holds. Therefore, the relations $t_{\text{sep}}(\rho) \geq t_{\text{sep}}(\rho) \geq t_{\text{sep}}(|\Phi\rangle) = (\text{Tr}\sqrt{\rho_A})^2$ hold.

Similarly, we can show the inequality $t_{\text{sep}}(\rho) \geq (\text{Tr}\sqrt{\rho_B})^2$. Thus, we obtain (7) in the mixed states case.

4. Local discrimination by one-way LOCC

In this section, we prove the following theorem concerning the local discrimination problem in terms of one-way LOCC in the direction $A \rightarrow B$:

Theorem 3. The inequality

$$t_\rightarrow(\rho) \geq \text{rank} \rho_A$$

holds for a bipartite state $\rho$ on $\mathcal{H}_A \otimes \mathcal{H}_B$. Any maximally correlated state $\rho$ satisfies the equality. In other words, the following inequality concerning the minimum error probability $\beta_\rightarrow(\rho)$ holds:

$$\beta_\rightarrow(\rho) \geq \frac{1}{D} \text{rank} \rho_A.$$
In the above theorem, a maximally correlated state is defined as a state which can be decomposed into the following form:

\[ \rho = \sum_{1 \leq i, j \leq d} \alpha_{ij} |u_i, v_i \rangle \langle u_j, v_j|, \]  

(24)

where \( \{|u_i\rangle\}_{i=1}^d \) and \( \{|v_j\rangle\}_{j=1}^d \) are orthonormal bases of \( \mathcal{H}_A \) and \( \mathcal{H}_B \), respectively [30]; apparently, a pure state is a maximally correlated state. Thus, \( t_-(|\Psi\rangle) = D\beta_-(|\Psi\rangle) \) is equal to the Schmidt rank of a state for a bipartite pure state \( |\Psi\rangle \). In the case when \( \rho \) is a maximally correlated state satisfying equation (24), the optimal way to discriminate between \( \rho \) and the completely mixed state is the following: suppose there are two parties called Alice and Bob. Both Alice and Bob measure their local states \( \mathcal{H}_A \) and \( \mathcal{H}_B \) in the bases \( \{|u_i\rangle\}_{i=1}^d \) and \( \{|v_j\rangle\}_{j=1}^d \), respectively (of course, they only need to detect the support of the local states). Then, Alice informs her measurement result to Bob. Suppose Alice’s result is \( |u_k\rangle \) and Bob’s result is \( |v_l\rangle \). If \( k \) is equal to \( l \), then they judge that the given state is \( \rho \). Otherwise, they judge that the given state is the completely mixed state.

By comparing theorem 1 and theorem 3, we can easily see that if a bipartite pure state \( |\Psi\rangle \) is neither a maximally entangled state nor a product state, then the strict inequality \( \beta_{\text{sep}}(|\Psi\rangle) < \beta_- (|\Psi\rangle) \) holds. Thus, we can conclude from these results that there is a gap between the one-way local distinguishability and the separable local distinguishability for a bipartite pure state at least in the present problem setting.

Applying theorem 3 for equation (6), we can extend Cohen’s characterization [7] concerning the perfect discrimination of a given set of pure states in terms of one-way LOCC to a set of mixed states:

**Corollary 2.** If a set of bipartite states \( \{|\rho_i\rangle\}_{i=1}^N \) on \( \mathcal{H}_A \otimes \mathcal{H}_B \) is perfectly distinguishable by one-way LOCC, then

\[ \sum_{i=1}^N \text{rank} \rho_{iA} \leq D. \]  

(25)

This bound of the size of locally distinguishable sets for one-way LOCC is much stronger than the known bound for separable operations [15].

In preparation for our proof of theorem 3, we see the fact that there are several equivalent representations of the definition of one-way LOCC POVM elements. We start from the following representation which we can see immediately from the definition; that is, in a bipartite system \( \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \), if \( T \in \mathcal{T}_- \), there exist sets of positive operators \( \{M_i\}_i \) and \( \{N^i_j\}_j \) such that

\[ T = \sum_{ij} M_i \otimes N^i_j, \]  

(26)

\[ \sum_i M_i \leq I_B \] and \( \sum_j N^i_j \leq I_A \), where \( \{M_i\}_i \) is the POVM of the local measurement on \( \mathcal{H}_A \) and \( \{N^i_j\}_j \) is the POVM of the local measurement on \( \mathcal{H}_B \) depending on the first measurement result \( i \). Further, redefining \( N^i_j \) as \( N_i = \sum_j N^i_j \), we derive the following equivalence relation:

\[ T \in \mathcal{T}_- \iff \exists \{M_i\}_i \text{ and } \{N_i\}_i \text{ s.t. } \forall i, 0 \leq M_i \leq I_B, \sum_i M_i \leq I_A \text{ and } T = \sum_i M_i \otimes N_i. \]  

(27)
Using this characterization, we obtain the following lemma.

**Lemma 2.** A one-way LOCC POVM element $T = \sum_{ij} M_i \otimes N^i_j \in \mathcal{T}$, satisfies $\text{Tr} \rho T = 1$ if and only if $\text{Tr} (\rho_A \sum_i M_i) = 1$ and $\text{Tr} (\rho_{B,M_i} \sum_j N^j_j) = 1$ for all $i$, where $\rho_A \界定 = \text{Tr}_B \rho$ and $\rho_{B,M_i} \界定 = \text{Tr}_A \rho M_i \otimes I_B / \text{Tr} \rho M_i \otimes I_B$.

**Proof.** We can calculate $\text{Tr} \rho T$ as follows:

$$
\text{Tr} \rho T = \sum_{ij} \text{Tr} \rho M_i \otimes N^i_j = \sum_{ij} \text{Tr} (\text{Tr}_A \rho (M_i \otimes I_B)) N^i_j
$$

$$
= \sum_i \text{Tr} \rho_A M_i \cdot \text{Tr} \rho_{B,M_i} \left( \sum_j N^j_j \right) = 1.
$$

Since $\sum_i \text{Tr} \rho_A M_i \leq 1$ and $\text{Tr} \rho_{B,M_i} \left( \sum_j N^j_j \right) \leq 1$ for all $i$, we derive $\sum_i \text{Tr} \rho_A M_i = 1$ and $\text{Tr} \rho_{B,M_i} \left( \sum_j N^j_j \right) = 1$. The opposite direction is trivial. \qed

Now, we prove theorem 3 using the above lemma.

**Proof (Theorem 3).** In order to detect a state perfectly, we need to detect the reduced density operator of the local system $A$, $\rho_A$, as well as that of the other local system $B$, $\rho_{B,M_i}$, perfectly in each step. Thus, we can assume that $N^i_j$ is a projection on $B$ without loss of generality. Hence, $\text{Tr} T = \sum_i \text{Tr} M_i \cdot \text{Tr} N^i_j \geq \sum_i \text{Tr} M_i$. Since we have to detect the reduced density operator of the local system $A$, $\rho_A$, we obtain $\text{Tr} \sum_i M_i \rho_A = 1$, i.e. (22). When the state $\rho$ is a maximally mixed state $\sum_{1 \leq i,j \leq d} a_{i,j} |u_i, v_i\rangle \langle u_j, v_j|$, the reduced density $\rho_A$ is $\sum_{i=1}^d a_{i,i} |u_i\rangle \langle u_i|$. Thus, rank $\rho_A = d$. In this case, we can perfectly detect this state by the one-way LOCC test $\sum_{i=1}^d |u_i\rangle \langle u_i| \otimes |v_i\rangle \langle v_i|$. \qed

We should note the following fact: although a maximally correlated state satisfies the equality of equation (22), the converse is not necessarily true. Even if $v_i$ is not orthogonal, we can perfectly detect this state by the one-way LOCC test $\sum_{i=1}^d |u_i\rangle \langle u_i| \otimes |v_i\rangle \langle v_i|$. When the rank of the state $\sum_{1 \leq i,j \leq d} a_{i,j} |v_j| v_i\rangle |u_i\rangle \langle u_j|$ is $d$, the rank of $\rho_A$ is $d$. That is, this gives a counterexample of the converse.

### 5. Local discrimination by two-way LOCC

So far, we have calculated the minimum error probability of the local discrimination problem for one-way LOCC $\beta_- (|\Psi\rangle)$ and separable operations $\beta_{sep} (|\Psi\rangle)$. In this section, we focus on discrimination protocols by two-way LOCC. Since two-way LOCC is mathematically complicated, it is difficult to derive the minimum two-way LOCC discrimination protocol, and as a result, it is difficult to derive the exact value of $\beta_- (|\Psi\rangle)$. However, in order to show the difference in the efficiency of one-way and two-way local discrimination protocols (which is actually the main purpose of this paper), it is enough to find the upper bound of the two-way error probability $\beta_- (|\Psi\rangle)$. Thus, we concentrate on deriving an upper-bound of $\beta_- \界定$ by constructing a concrete two-way LOCC discrimination protocol. For simplicity, we only
treat three-step LOCC discrimination protocols on a bipartite system, which are in the simplest

class of genuine two-way LOCC protocols. As a result, we show that even three-step LOCC

protocols can discriminate a given state from the completely mixed state much better than by

one-way (that is, two-step) LOCC protocols.

We can generally describe a three-step LOCC protocol to discriminate a pure state $|\Psi\rangle$ from

$\rho_{\text{mix}} = \frac{I}{d^2}$ on a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$ without making any error in detecting $|\Psi\rangle$ as follows: suppose there are two parties called Alice and Bob. Firstly, Alice performs a POVM $\{M_i\}_i$ on her system $\mathcal{H}_A$, and sends the measurement result $i$ to Bob. Secondly, depending on $i$, Bob performs a POVM $\{N_j\}_j$ on his system $\mathcal{H}_B$, and sends the measurement result $j$ to Alice. If the given state is $\rho$, by easy calculation, we can check that Alice’s state after this step is

$$\rho_{A}^{ij} = \frac{\sqrt{M_i} \sqrt{\rho_A} N_j^T \sqrt{\rho_A} \sqrt{M_i}}{\mathrm{Tr} \left( \sqrt{M_i} \sqrt{\rho_A} N_j^T \sqrt{\rho_A} \right)}, \quad (28)$$

where $\rho_A^{ij} \overset{\text{def}}{=} \mathrm{Tr}_B \langle \Psi | \langle \Psi | = 1$, and the transposition is taken in the Schmidt basis of $|\Psi\rangle$. Thus, in order not to make an error in detecting the above state, finally, Alice should make a measurement in $\{\sigma_A^{ij} > 0\}$, $I_A - \{\sigma_A^{ij} > 0\}$, where $\{\sigma_A^{ij} > 0\}$ is a projection operator on to the support of $\sigma_A^{ij}$ (the subspace spanned by eigenvectors corresponding to non-zero eigenvalue of $\sigma_A^{ij}$), and $I_A$ is an identity operator in $\mathcal{H}_A$. Then, if she detects $\{\sigma_A^{ij} > 0\}$, she judges that the first state was $|\Psi\rangle$, and if she detects $I_A - \{\sigma_A^{ij} > 0\}$, she judges that the first state was $\rho_{\text{mix}}$. Suppose $\{T, I_A - T\}$ is the POVM corresponding to the above local discrimination protocol, where $T$ corresponds to $|\Psi\rangle$ and $I - T$ corresponds to $\rho_{\text{mix}}$. Then, we can check that the whole POVM $\{T, I - T\}$ can be written down as follows:

$$T = \sum_i \sum_j \left( \sqrt{M_i} \left\{ \sigma_A^{ij} > 0 \right\} \sqrt{M_i} \right) \otimes N_j^i, \quad (29)$$

where $\sigma_A^{ij}$ is defined by equation (28), and all $M_i$ and $N_j^i$ are positive operators satisfying $\sum_i M_i = I_A$ and $\sum_j N_j^i = I_B$. We can also check that $T$ defined by equation (29) satisfies

$$\langle \Psi | T^j | \Psi \rangle = \sum_{ij} \langle \Psi | \left( \sqrt{M_i} \left\{ \sigma_A^{ij} > 0 \right\} \sqrt{M_i} \right) \otimes N_j^i | \Psi \rangle$$

$$= d \sum_{ij} \langle \Phi^+ | \left( \sqrt{\rho_A} \otimes I_B \right) \left( \sqrt{M_i} \left\{ \sigma_A^{ij} > 0 \right\} \sqrt{M_i} \right) \otimes \left( \sqrt{\rho_A} \otimes I_B \right) | \Phi^+ \rangle$$

$$= d \sum_{ij} \langle \Phi^+ | \left( \sqrt{N_j^T} \sqrt{\rho_A} \sqrt{M_i} \left\{ \sigma_A^{ij} > 0 \right\} \sqrt{M_i} \sqrt{\rho_A} \sqrt{N_j^T} \right) \otimes I_B | \Phi^+ \rangle$$

$$= \sum_{ij} \mathrm{Tr} \sqrt{N_j^T} \sqrt{\rho_A} \sqrt{M_i} \left\{ \sigma_A^{ij} > 0 \right\} \sqrt{M_i} \sqrt{\rho_A} \sqrt{N_j^T}$$

$$= \sum_{ij} \mathrm{Tr} \sqrt{M_i} \sqrt{\rho_A} N_j^T \sqrt{\rho_A} \sqrt{M_i} \left\{ \sigma_A^{ij} > 0 \right\}$$

$$= \sum_{ij} (\mathrm{Tr} M_i \sqrt{\rho_A} N_j^T \sqrt{\rho_A}) \cdot \left( \mathrm{Tr} \sigma_A^{ij} \left\{ \sigma_A^{ij} > 0 \right\} \right)$$
\[
\sum_{ij} \text{Tr} \left( N^i_j \left( \sqrt{\rho_A M_i} \sqrt{\rho_A} \right)^T \right) = \sum_{ij} \text{Tr} \rho_A M_i = 1,
\]

where \(|\Phi^+\rangle\) is the maximally entangled state sharing the Schmidt basis with \(|\Psi\rangle\), and the transposition \(T\) is always taken in the Schmidt basis of \(|\Psi\rangle\). In the second line of the above equalities, we used the equality \(|\Psi\rangle = \sqrt{\rho_A} \otimes I_B |\Phi^+\rangle\). In the third line, we used the equalities \(I_A \otimes X |\Phi^+\rangle = X^T \otimes I_B |\Phi^+\rangle\), which is valid for an arbitrary operator \(X\). In the sixth line, we used equation (28).

The above three-step LOCC protocol is general enough. However, it is too complicated to optimize \(\text{Tr}T\) over all choices of POVM \(\{M_i\}_i\) and \(\{N^i_j\}_j\). In this section, our aim is only to find a good (not necessary optimal) two-way LOCC protocol by which we can discriminate a state from the completely mixed state better than by any one-way LOCC protocols. Thus, to make the problem simpler, we make the following assumptions on Alice’s POVM from the completely mixed state.

In other words, an orthonormal set of states \(\{\xi^i_j\}_{ij=1}^{\text{rank} M_i}\) of an orthonormal set of eigen vectors of

\[
M_i = \sum_{k=1}^d \delta_{ki} |k\rangle \langle k|,
\]

where \(\text{rank} M_i = i\), and the coefficients \(\delta_{ki} \geq 0\) satisfy \(\sum_{i=1}^d \delta_{ki} = 1\) for all \(k\). Moreover, we assume that \(\{N^i_j\}_j\) is a von Neumann measurement corresponding to a mutually unbiased basis [11, 31] \(\{\xi^i_j\}_{ij=1}^{\text{rank} M_i}\) of an orthonormal set of eigen vectors of

\[
\omega_B \overset{\text{def}}{=} \frac{\sqrt{\rho_B} M_i^T \sqrt{\rho_B}}{\text{Tr} \sqrt{\rho_B} M_i^T \sqrt{\rho_B}}
\]

corresponding to non-zero eigenvalues; that is, \(\{\xi^i_j\}_{ij=1}^{\text{rank} M_i}\) only spans

\[
\text{rank} \frac{\sqrt{\rho_B} M_i^T \sqrt{\rho_B}}{\text{Tr} \sqrt{\rho_B} M_i^T \sqrt{\rho_B}}.
\]

In other words, an orthonormal set of states \(\{\xi^i_j\}_{ij=1}^{\text{rank} M_i}\) satisfies

\[
\left| \frac{\sqrt{\rho_B} M_i^T \sqrt{\rho_B}}{\text{Tr} \sqrt{\rho_B} M_i^T \sqrt{\rho_B}} \right| \langle \xi^i_j | = \frac{1}{\text{rank} M_i}.
\]

Note that \(\omega_B\) is Bob’s state after Alice’s first measurement in the case where the given state is \(|\Psi\rangle\), and thus, Bob only needs to detect the subspace \(\text{Ran} \omega_B\) in this case. That is, Bob’s POVM consists of \(\{\xi^i_j\}_{ij=1}^{\text{rank} M_i}\) and \(I_B - \sum_{j=1}^{\text{rank} M_i} |\xi^i_j \rangle \langle \xi^i_j|\); if Bob derives the measurement
result corresponding to \( I_B - \sum_{j=1}^{\text{rank} M_i} |\xi_j^i\rangle \langle \xi_j^i| \), then he judges that the given state is \( \rho_{mix} \). We should also note that due to Bob’s mutually unbiased measurement, our three-step protocol cannot be reduced to a two-step one-way LOCC protocol. If all Bob’s POVMs are commutative with the eigen basis of \( \omega_B \), the whole protocol can be reduced to a one-way LOCC protocol; however, \( \omega_B \) never commutes the projection on to the mutually unbiased basis of the eigen basis of \( \omega_B \). Finally, under the above assumptions, we can write down the trace of the whole POVM element TrT as follows:

\[
\text{Tr} T = \text{Tr} \left( \sum_{i=1}^{d} \sum_{j=1}^{\text{rank} M_i} \left( \sqrt{M_i} \left\{ \sigma^{ij}_X > 0 \right\} \sqrt{M_i} \right) \otimes N_j \right)
\]

\[
= \sum_{i=1}^{d} \sum_{j=1}^{\text{rank} M_i} \text{Tr} \left( \left( \frac{\sqrt{M_i} \sqrt{\rho_A} \langle |\xi_j^i\rangle \langle \xi_j^i| |\xi_j^i\rangle T \sqrt{\rho_A} \sqrt{M_i}}{|\xi_j^i\rangle \sqrt{\rho_A} M_i T \sqrt{\rho_A} |\xi_j^i\rangle} \right) \sqrt{M_i} \right) \otimes |\xi_j^i\rangle \langle \xi_j^i|) \]

We have

\[
= \sum_{i=1}^{d} \sum_{j=1}^{\text{rank} M_i} \frac{|\xi_j^i\rangle \sqrt{\rho_A} (M_i T)^2 \sqrt{\rho_A} |\xi_j^i\rangle}{|\xi_j^i\rangle \sqrt{\rho_A} M_i T \sqrt{\rho_A} |\xi_j^i\rangle} \frac{\text{Tr} \sqrt{\rho_A} (M_i T)^2 \sqrt{\rho_A}}{\text{Tr} \sqrt{\rho_A} M_i T \sqrt{\rho_A}} \sum_{i=1}^{d} \lambda_i \delta_{ki}.
\]

In the second line of the above equalities, we used equation (28) (the definition of \( \sigma^{ij}_X \)) and the equality

\[
\left\{ \frac{\sqrt{M_i} \sqrt{\rho_A} \langle |\xi_j^i\rangle \langle \xi_j^i| |\xi_j^i\rangle T \sqrt{\rho_A} \sqrt{M_i}}{|\xi_j^i\rangle \sqrt{\rho_A} M_i T \sqrt{\rho_A} |\xi_j^i\rangle} > 0 \right\} = \frac{\sqrt{M_i} \sqrt{\rho_A} \langle |\xi_j^i\rangle \langle \xi_j^i| |\xi_j^i\rangle T \sqrt{\rho_A} \sqrt{M_i}}{|\xi_j^i\rangle \sqrt{\rho_A} M_i T \sqrt{\rho_A} |\xi_j^i\rangle}.
\]

In the fourth line of the above equalities, we used the relation \( \rho_A = \rho_B \) and the condition of mutually unbiased basis equation (31). Therefore, our problem is reduced to the optimization of \( \sum_{i=1}^{d} \delta_{ki} \) over \( \{\delta_j\}_k \) subject to the constraints \( \delta_{ki} \geq 0 \) and \( \sum_{i=1}^{d} \delta_{ki} = 1 \). In other words, we can summarize the above discussion in the form of the following lemma.

**Lemma 3.** For a bipartite pure state \( |\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \), \( \beta_{\times}(|\Psi\rangle) \) satisfies the following inequality:

\[
\beta_{\times}(|\Psi\rangle) \leq \beta_{\ast}(|\Psi\rangle)
\]

\[
\text{def} \frac{1}{d^2} \min_{\{\delta_k\}_{k=1}^{d}} \left\{ \sum_{i=1}^{d} i \cdot \sum_{k=1}^{d} \lambda_k \delta_{ki} \left| \forall k, \sum_{i=1}^{d} \delta_{ki} \geq 0, \text{ and } \forall k, \sum_{i=1}^{d} \delta_{ki} = 1 \right. \right\}, \tag{32}
\]

where \( \lambda_k \) is the Schmidt coefficients of \( |\Psi\rangle \) and satisfies \( \lambda_k \geq \lambda_{k+1} \) for all \( k \), and the indices \( (k, i) \) are in the range of the triangle region \( 1 \leq k \leq i \leq d \).

Then, the above inequality can be written as \( \beta_{\times}(|\Psi\rangle) \leq \beta_{\ast}(|\Psi\rangle) \). Together with the results of the above sections, we derive the following inequalities related to the minimum type 2 error...
probability for a bipartite pure state:
\[ \beta_{g}(\ket{\Psi}) = \frac{1}{d^2} \leq \beta_{\text{sep}}(\ket{\Psi}) \]
\[ = \frac{1}{d^2} \left( \sum_{i} \sqrt{\lambda_i} \right)^2 \leq \beta_{\leftrightarrow}(\ket{\Psi}) \leq \beta_{\tilde{\leftrightarrow}}(\ket{\Psi}) \leq \frac{1}{d^2} \text{rank}_{\text{Tr}}(\ket{\Psi} \bra{\Psi}) \]
\[ = \beta_{\rightarrow}(\ket{\Psi}). \] (33)

For a two-qubit system, we can analytically calculate the exact value of the upper bound \( \beta_{\tilde{\leftrightarrow}}(\ket{\Psi}) \), and can derive the following lemma.

**Lemma 4.** In a two-qubit system,
\[ \beta_{\tilde{\leftrightarrow}}(\ket{\Psi}) = \frac{1}{2} \left( 1 - \sqrt{2\lambda} \right)^2, \] (34)
where \{1 - \lambda, \lambda\} is the Schmidt coefficient of \( \ket{\Psi} \) satisfying \( 1 \leq \lambda \leq \frac{1}{2} \).

**Proof.** Without losing generality, we can write a bipartite state as \( \ket{\Psi} = \sqrt{1 - \lambda} \ket{00} + \sqrt{\lambda} \ket{11} \).
Then, by a straightforward calculation, we derive
\[
\beta_{\tilde{\leftrightarrow}}(\ket{\Psi}) = \frac{1}{4} \min_{0 \leq \delta \leq 1} \left\{ \frac{2 - \delta \{ (1 - 2\lambda) + \delta (1 - \lambda) \}}{1 - \delta (1 - \lambda)} \right\}, \] (35)
where we substitute \( \lambda_1 = 1 - \lambda, \lambda_2 = \lambda, \delta_{11} = \delta, \delta_{12} = 1 - \delta \) and \( \delta_{22} = 1 \) into equation (32).

Suppose
\[ t'_{\tilde{\leftrightarrow}}(\lambda, \delta) \overset{\text{def}}{=} 2 - \frac{\delta \{ (1 - 2\lambda) + \delta (1 - \lambda) \}}{1 - \delta (1 - \lambda)}. \]
Then, we can calculate the derivative of \( t'_{\tilde{\leftrightarrow}}(\lambda, \delta) \) as
\[ \frac{\partial t'_{\tilde{\leftrightarrow}}(\lambda, \delta)}{\partial \delta} = - \left\{ (1 - \lambda)\delta - \left( 1 - \sqrt{2\lambda} \right) \right\} \left\{ (1 - \lambda)\delta - \left( 1 + \sqrt{2\lambda} \right) \right\}, \]
for fixed \( 0 \leq \lambda \leq \frac{1}{2} \). Thus, under the condition \( 0 \leq \delta \leq 1 \), \( t'_{\tilde{\leftrightarrow}}(\lambda, \delta) \) attains its minimum when
\[ \delta = \frac{1 - \sqrt{2\lambda}}{1 - \lambda}. \] Therefore, we derive
\[ \beta_{\tilde{\leftrightarrow}}(\ket{\Psi}) = \frac{1}{4} \min_{0 \leq \delta \leq 1} t'_{\tilde{\leftrightarrow}}(\lambda, \delta) = \frac{1}{2} - \frac{(1 - \sqrt{2\lambda})^2}{4 (1 - \lambda)}. \]

Therefore, for a two-qubit state \( \ket{\Psi_{\lambda}} = \sqrt{1 - \lambda} \ket{00} + \sqrt{\lambda} \ket{11} \), the inequality (33) can be reduced as follows:
\[ \beta_{g}(\ket{\Psi}) = \frac{1}{4} \leq \beta_{\text{sep}}(\ket{\Psi}) \]
\[ = \frac{1}{4} + \frac{1}{2} \sqrt{\lambda (1 - \lambda)} \leq \beta_{\leftrightarrow}(\ket{\Psi}) \leq \frac{1}{2} - \frac{(1 - \sqrt{2\lambda})^2}{4 (1 - \lambda)} \leq \frac{1}{2} \]
\[ = \beta_{\rightarrow}(\ket{\Psi}), \]
where the equality of the last inequality holds, if and only if the state is a product state or a maximally entangled state. We present the graph of these bounds in figure 1. From this figure, we can see that there is a big gap between $\beta_-$ $(|\Psi\rangle)$ and $\beta_+$ $(|\Psi\rangle)$ and the difference between $\beta_-(|\Psi\rangle)$ and $\beta_{sep}(|\Psi\rangle)$ is (if the difference exists) relatively small. Thus, for any non-maximally entangled pure states, there is a gap between the one-way and two-way local distinguishability at least for two-qubit systems in terms of $\beta_{sep}(|\Psi\rangle)$.

In a system with a dimension of local systems $d \geq 3$, the optimization in the definition of $\beta_{sep}(|\Psi\rangle)$ (equation (32)) is too complicated to be solved by an analytical calculation anymore. Thus, we numerically calculate the right-hand side of equation (32) for a $\mathbb{C}^3 \otimes \mathbb{C}^3$ (two-qubit) system and a $\mathbb{C}^4 \otimes \mathbb{C}^4$ system. For a $\mathbb{C}^3 \otimes \mathbb{C}^3$ system, we calculate equation (32) for three different one-parameter families of pure states:

1. $|\Psi_\lambda\rangle = \sqrt{1-2\lambda}|11\rangle + \sqrt{\lambda}|22\rangle + \sqrt{\lambda}|33\rangle$, $(0 \leq \lambda \leq \frac{1}{3})$: in this case, $\beta_{g}(|\Psi_\lambda\rangle) = \frac{1}{9}$, $\beta_{sep}(|\Psi_\lambda\rangle) = \frac{1}{9} (\sqrt{1-2\lambda} + 2\sqrt{\lambda})^2$ and $\beta_-(|\Psi_\lambda\rangle) = \frac{1}{3}$. We give the results of a numerical calculation of $\beta_{sep}(|\Psi_\lambda\rangle)$ in figure 2.

2. $|\Psi_\lambda\rangle = \sqrt{1-3\lambda}|11\rangle + 2\sqrt{2\lambda}|22\rangle + \sqrt{\lambda}|33\rangle$, $(0 \leq \lambda \leq \frac{1}{3})$: in this case, $\beta_{g}(|\Psi_\lambda\rangle) = \frac{1}{9}$, $\beta_{sep}(|\Psi_\lambda\rangle) = \frac{1}{9} (\sqrt{1-3\lambda} + (1 + \sqrt{2})\sqrt{\lambda})^2$ and $\beta_-(|\Psi_\lambda\rangle) = \frac{1}{3}$. We give the results of a numerical calculation of $\beta_{sep}(|\Psi_\lambda\rangle)$ in figure 3.

3. $|\Psi_\lambda\rangle = \sqrt{1-4\lambda}|11\rangle + 3\sqrt{2\lambda}|22\rangle + \sqrt{\lambda}|33\rangle$, $(0 \leq \lambda \leq \frac{1}{3})$: in this case, $\beta_{g}(|\Psi_\lambda\rangle) = \frac{1}{9}$, $\beta_{sep}(|\Psi_\lambda\rangle) = \frac{1}{9} (\sqrt{1-4\lambda} + (1 + \sqrt{3})\sqrt{\lambda})^2$ and $\beta_-(|\Psi_\lambda\rangle) = \frac{1}{3}$. We give the results of a numerical calculation of $\beta_{sep}(|\Psi_\lambda\rangle)$ in figure 4.

From figures 2–4, we can confirm that the shapes of the graphs of $\beta_{sep}(|\Psi_\lambda\rangle)$ and $\beta_{sep}(|\Psi_\lambda\rangle)$ hardly depend on the choice of a one-parameter family $|\Psi_\lambda\rangle$ in $\mathbb{C}^3 \otimes \mathbb{C}^3$. For a $\mathbb{C}^4 \otimes \mathbb{C}^4$ system,

\footnote{In a strict sense, we can show that there exists an analytical solution for the optimization problem in equation (33) by means of a Lagrange multiplier. However, even for a $3 \times 3$ dimensional system, the general solution is too complicated and too ugly to write here.}

Figure 1. The bound as a function of $\lambda$ (the Schmidt coefficient of $|\Psi\rangle$). Thin line: $\beta_{sep}(|\Psi\rangle)$; broken line: $\beta_{\geq}(|\Psi\rangle)$; (an upper bound of $\beta_{\geq}(|\Psi\rangle)$); thick line: $\beta_{-}(|\Psi\rangle)$, thin broken line: $\beta_{g}(|\Psi\rangle)$.
Figure 2. The bound as a function of $\lambda$ for a family of states $|\Psi_\lambda\rangle = \sqrt{1 - 2\lambda} |11\rangle + \sqrt{\lambda} |22\rangle + \sqrt{\lambda} |33\rangle$. Thick broken line: results of a numerical calculation of $\beta_{\tilde{\leftrightarrow}}(|\Psi_\lambda\rangle)$ (a lower bound of $\beta_{\leftrightarrow}(|\Psi_\lambda\rangle)$); thin line: $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$; thick line: $\beta_{\to}(|\Psi_\lambda\rangle)$; thin broken line: $\beta_{\to}(|\Psi_\lambda\rangle)$.

Figure 3. The bound as a function of $\lambda$ for a family of states $|\Psi_\lambda\rangle = \sqrt{1 - 3\lambda} |11\rangle + \sqrt{2\lambda} |22\rangle + \sqrt{\lambda} |33\rangle$. Thick broken line: results of a numerical calculation of $\beta_{\tilde{\leftrightarrow}}(|\Psi_\lambda\rangle)$ (a lower bound of $\beta_{\leftrightarrow}(|\Psi_\lambda\rangle)$); thin line: $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$; thick line: $\beta_{\to}(|\Psi_\lambda\rangle)$; thin broken line: $\beta_{\to}(|\Psi_\lambda\rangle)$. 
we calculate equation (32) for two different one-parameter families of pure states:

1. $|\Psi_\lambda\rangle = \sqrt{1-3\lambda}|11\rangle + \sqrt{3\lambda}|22\rangle + \sqrt{\lambda}|33\rangle + |44\rangle$, ($0 \leq \lambda \leq \frac{1}{3}$): in this case, $\beta_{\text{sep}}(|\Psi_\lambda\rangle) = \frac{1}{16}$, $\beta_{\text{sep}}(|\Psi_\lambda\rangle) = \frac{1}{16}(\sqrt{1-3\lambda} + 3\sqrt{\lambda})^2$ and $\beta_{\text{lim}}(|\Psi_\lambda\rangle) = \frac{1}{4}$. We give the results of a numerical calculation of $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$ in figure 5.

2. $|\Psi_\lambda\rangle = \sqrt{1-(9/2)\lambda}|11\rangle + \sqrt{2\lambda}|22\rangle + \sqrt{(3/2)\lambda}|33\rangle + \sqrt{\lambda}|44\rangle$, ($0 \leq \lambda \leq \frac{2}{11}$): in this case, $\beta_{\text{sep}}(|\Psi_\lambda\rangle) = \frac{1}{16}$, $\beta_{\text{sep}}(|\Psi_\lambda\rangle) = \frac{1}{16}(\sqrt{1-(9/2)\lambda} + (1 + \sqrt{(3/2) + \sqrt{2}})\sqrt{\lambda})^2$ and $\beta_{\text{lim}}(|\Psi_\lambda\rangle) = \frac{1}{4}$. We give the results of a numerical calculation of $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$ in figure 6.

From figures 5 and 6, we can confirm that the shapes of the graphs of $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$, and $\beta_{\text{lim}}(|\Psi_\lambda\rangle)$ hardly depend on the choice of a one-parameter family $|\Psi_\lambda\rangle$ in $\mathbb{C}^4 \otimes \mathbb{C}^4$ as well as in $\mathbb{C}^3 \otimes \mathbb{C}^3$.

Note that, for the all above families of states, we choose a parameter $\lambda$ so that $|\Psi_\lambda\rangle$ can be converted to $|\Psi_\lambda\rangle$ by LOCC for all $\lambda \geq \lambda'$, and $|\Psi_0\rangle$ is a product state; that is, in a naive sense, the degree of entanglement increases monotonically when $\lambda$ increases. From figures 2–6, as well as for a two-qubit system (figure 1), we can see that there is always a big gap between $\beta_{\text{lim}}(|\Psi\rangle)$ and $\beta_{\text{sep}}(|\Psi\rangle)$ and the difference between $\beta_{\text{lim}}(|\Psi\rangle)$ and $\beta_{\text{sep}}(|\Psi\rangle)$ is (if the difference exists) relatively small for $\mathbb{C}^3 \otimes \mathbb{C}^3$ and $\mathbb{C}^4 \otimes \mathbb{C}^4$ systems. Moreover, since the shape of graph corresponding to $\beta_{\text{lim}}(|\Psi\rangle)$ seems not to change depending on the dimension of a system, we may guess that, for any non-maximally entangled pure states (even in a high dimensional system), there is a gap between the one-way and two-way local distinguishability in terms of $\beta_{\text{lim}}(|\Psi\rangle)$. That is, the two-way classical communication remarkably improves the local distinguishability compared to local discrimination by one-way classical communication at least for bipartite pure states.

Figure 4. The bound as a function of $\lambda$ for a family of states $|\Psi_\lambda\rangle = \sqrt{1-4\lambda}|11\rangle + \sqrt{3\lambda}|22\rangle + \sqrt{\lambda}|33\rangle$. Thick broken line: results of a numerical calculation of $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$ (a lower bound of $\beta_{\text{lim}}(|\Psi_\lambda\rangle)$); thin line: $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$; thick line: $\beta_{\text{lim}}(|\Psi_\lambda\rangle)$; thin broken line: $\beta_{\text{lim}}(|\Psi_\lambda\rangle)$.

Figure 5. Coefficients: $(1-4\lambda, 3\lambda, \lambda)$.
Figure 5. The bound as a function of $\lambda$ for a family of states $|\Psi_\lambda\rangle = \sqrt{1-3\lambda} |11\rangle + \sqrt{\lambda} (|22\rangle + |33\rangle + |44\rangle)$. Thick broken line: results of a numerical calculation of $\beta_{\leftrightarrow}(|\Psi_\lambda\rangle)$ (a lower bound of $\beta_{\leftrightarrow}(|\Psi_\lambda\rangle)$; thin line: $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$; thick line: $\beta_{\rightarrow}(|\Psi_\lambda\rangle)$; thin broken line: $\beta_{\text{g}}(|\Psi_\lambda\rangle)$.

Figure 6. The bound as a function of $\lambda$ for a family of states $|\Psi_\lambda\rangle = |\Psi_\lambda\rangle = \sqrt{1-(9/2)\lambda} |11\rangle + \sqrt{2\lambda} |22\rangle + \sqrt{(3/2)\lambda} |33\rangle + \sqrt{\lambda} |44\rangle$. Thick broken line: results of a numerical calculation of $\beta_{\leftrightarrow}(|\Psi_\lambda\rangle)$ (a lower bound of $\beta_{\leftrightarrow}(|\Psi_\lambda\rangle)$; thin line: $\beta_{\text{sep}}(|\Psi_\lambda\rangle)$; thick line: $\beta_{\rightarrow}(|\Psi_\lambda\rangle)$; thin broken line: $\beta_{\text{g}}(|\Psi_\lambda\rangle)$.
6. Conclusion

In this paper, in order to clarify the difference between two-way LOCC and one-way LOCC on local discrimination problems, we concentrated on the local discrimination of a given bipartite state from the completely mixed state \( \rho_{\text{mix}} \) under the condition where the given state should be detected perfectly, while previous research [11, 12] treated the same problem between two bipartite pure states. We defined \( \beta_-(\rho), \beta_+(\rho) \) and \( \beta_{\text{sep}}(\rho) \) as the minimum error probability in detecting the completely mixed state by one-way LOCC, two-way LOCC and separable operation, respectively, under the condition that a given state \( \rho \) is detected perfectly. Firstly, in section 3, for separable operations, we showed that the minimum error probability \( t_{\text{sep}}(\rho) \) coincides with an entanglement measure called the global robustness of entanglement for a bipartite pure state except for an unimportant constant term. Secondly, in section 4, for one-way LOCC, we showed that the minimum error probability \( \beta_-(\rho) \) coincides with the Schmidt rank for a bipartite pure state except for an unimportant constant term. Finally, in section 5, by constructing a concrete three-step two-way LOCC discrimination protocol, we derived an upper bound for the minimum error probability \( \beta_+(\rho) \) for a bipartite pure state. By calculating this upper bound analytically and also numerically, we showed that \( \beta_+(\rho) \) is strictly smaller than \( \beta_-(\rho) \), and moreover, \( \beta_+(\rho) \) and \( \beta_{\text{sep}}(\rho) \) give almost the same value for a lower dimensional bipartite pure state; these results can be seen in figures 2–6. As a result, although there is no difference between one-way LOCC and two-way LOCC concerning local discrimination between two bipartite pure states [11, 12], we conclude that two-way classical communication remarkably improves the local distinguishability in comparison with one-way classical communication for a low-dimensional pure state, at least in the present problem setting. Due to our quantitative comparison, from the continuity of the second kinds of error probabilities, a similar result should hold when the second state \( \tilde{\rho} \) belongs to the neighborhood of the completely mixed state. Further, we are preparing a forthcoming paper concerning this kind of problem for the multi-partite case [32].

Acknowledgments

MO thanks Professor Mio Murao, Dr Damian Markham and Dr Shashank Virmani for helpful comments and discussions. MO is grateful to Professor Hiroshi Imai for his support and encouragement. We also appreciate the reviewers’ comments. This work was supported by Special Coordination Funds for Promoting Science and Technology and a MEXT Grant-in-Aid for Scientific Research on Priority, Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI), no. 18079014.

Appendix. Proof of equation (20)

Now, we prove equation (20), which is used in the proof of theorem 2. Suppose

\[
P \overset{\text{def}}{=} \frac{1}{2} \sum_{i \neq j} |\alpha_{ij}\rangle \langle \alpha_{ij}| \otimes |\overline{\beta}_{ij}\rangle \langle \overline{\beta}_{ij}|,
\]

where

\[
\alpha_{ij} = \sqrt{\sum_{j \neq i} |\langle \alpha_{ij}| \alpha_{ij}\rangle|^2},
\]

\[
\beta_{ij} = \frac{1}{2} \sum_{j \neq i} |\langle \beta_{ij}| \beta_{ij}\rangle|^2.
\]

\[\text{New Journal of Physics 10 (2008) 013006 (http://www.njp.org/)}\]
and
\[
Q \overset{\text{def}}{=} \sum_{i \neq j} \left\{ \sum_{k \neq i, j} \lambda_k + (\sqrt{\lambda_i} - \sqrt{\lambda_j})^2 \right\} |e_i f_j\rangle \langle e_i f_j|;
\]
that is, \( \mathcal{T}_0 = P + Q \). Then, by applying a twirling operation over \( U^{-\theta} \), we derive the following equality:
\[
\int_0^{2\pi} \cdots \int_0^{2\pi} U^{-\theta} P U^\dagger_{-\theta} \, d\theta_1 \cdots d\theta_d = \left( \sum_{j'=1}^d |e_{j'}\rangle \langle e_{j'}| \otimes |f_{j'}\rangle \langle f_{j'}| \right) P \left( \sum_{j'=1}^d |e_{j'}\rangle \langle e_{j'}| \otimes |f_{j'}\rangle \langle f_{j'}| \right)
+ \sum_{i' \neq k} \langle e_{i'} | e_{i'} | \otimes | f_{k'} \rangle \langle f_{k'} | P \langle e_{i'} | e_{i'} | \otimes | f_{k'} \rangle \langle f_{k'} |).
\]
This equality can be proven as follows: the action of a twirling operation (group-averaging) over a unitary representation of a compact topological group is equal to the action of the projection on to the subspace of all invariant elements under the group action [29]. For the action of \( U^{-\theta} \) and \( U^\dagger_{-\theta} \), the subspace (of operator-space \( XX(\mathcal{H}) \)) consisting of all the invariant elements is spanned by the operators \( \{ |e_{j} f_{k}\rangle \langle e_{j} f_{k}|\}_{j \neq k} \) and \( \{ |e_{j} f_{k}\rangle \langle e_{k} f_{k}|\}_{ij} \). Therefore, we can easily see the above equation. For \( i \neq j \) and \( i' \neq k' \), we have
\[
\left( \sum_{j'=1}^d |e_{j'}\rangle \langle e_{j'}| \otimes |f_{j'}\rangle \langle f_{j'}| \right) |\overline{a}_{ij}\rangle |\overline{b}_{ij}\rangle = \sqrt{\lambda_j} |e_i f_i\rangle - \sqrt{\lambda_i} |e_j f_j\rangle
\]
\[
\langle e_{i'} | e_{i'} | \otimes | f_{k'} \rangle \langle f_{k'} | |\overline{a}_{ij}\rangle |\overline{b}_{ij}\rangle = \delta_{i', i} \delta_{k', j} (\lambda_j \lambda_i)^{\frac{1}{2}} |e_i f_i\rangle - \delta_{i', j} \delta_{k', i} (\lambda_j \lambda_i)^{\frac{1}{2}} |e_j f_j\rangle.
\]
Since
\[
\left( \sqrt{\lambda_j} |e_i f_i\rangle - \sqrt{\lambda_i} |e_j f_j\rangle \right) \left( \sqrt{\lambda_j} |e_i f_i\rangle - \sqrt{\lambda_i} |e_j f_j\rangle \right) = \lambda_j |e_i f_i\rangle |e_i f_i\rangle + \lambda_i |e_j f_j\rangle |e_j f_j\rangle - \sqrt{\lambda_j} \sqrt{\lambda_i} |e_i f_i\rangle |e_j f_j\rangle - \sqrt{\lambda_i} \sqrt{\lambda_j} |e_j f_j\rangle |e_i f_i\rangle,
\]
we obtain
\[
\int_0^{2\pi} \cdots \int_0^{2\pi} U^{-\theta} P U^\dagger_{-\theta} \, d\theta_1 \cdots d\theta_d
\]
\[
= \sum_{i \neq j} \left( \lambda_j |e_i f_i\rangle \langle e_i f_i| + \lambda_i |e_j f_j\rangle \langle e_j f_j| - \sqrt{\lambda_j} \sqrt{\lambda_i} |e_i f_i\rangle \langle e_j f_j| - \sqrt{\lambda_i} \sqrt{\lambda_j} |e_j f_j\rangle \langle e_i f_i| \right)
+ \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} |e_i f_j\rangle \langle e_i f_j|.
\]
In the same way, we can also show that the equality \( \int_0^{2\pi} \cdots \int_0^{2\pi} U^{-\theta} P \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} |e_i f_j\rangle \langle e_i f_j| \cdot \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} |e_i f_j\rangle \langle e_i f_j| = Q; Q \)
is invariant under the twirling operation.
Finally, we can calculate \( \int_0^{2\pi} \cdots \int_0^{2\pi} U \frac{\partial}{\partial \theta} T U^\dagger \frac{\partial}{\partial \theta} \, d\theta_1 \cdots d\theta_d \) as follows:

\[
\int_0^{2\pi} \cdots \int_0^{2\pi} U \frac{\partial}{\partial \theta} T U^\dagger \frac{\partial}{\partial \theta} \, d\theta_1 \cdots d\theta_d = \int_0^{2\pi} \cdots \int_0^{2\pi} U \frac{\partial}{\partial \theta} (P + Q) U^\dagger \frac{\partial}{\partial \theta} \, d\theta_1 \cdots d\theta_d
\]

\[
= \left( \sum_i |e_i f_i \rangle \langle e_i f_i| \right) - \left( \sum_{i=1}^d \sqrt{\lambda_i} |e_i f_i \rangle \langle e_i f_i| \right) + \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} |e_i f_j \rangle \langle e_i f_j| \\
+ \sum_{i \neq j} \left\{ \sum_{k \neq i, j} \lambda_k + (\sqrt{\lambda_i} - \sqrt{\lambda_j})^2 \right\} |e_i f_j \rangle \langle e_i f_j| \\
= \left( \sum_i |e_i f_i \rangle \langle e_i f_i| \right) + \left( \sum_{i \neq j} |e_i f_i \rangle \langle e_i f_i| \right) - \left( \sum_{i=1}^d \sqrt{\lambda_i} |e_i f_i \rangle \langle e_i f_i| \right) \left( \sum_{i=1}^d \sqrt{\lambda_i} |e_i f_i \rangle \langle e_i f_i| \right) \\
- \sum_{i \neq j} \sqrt{\lambda_i \lambda_j} |e_i f_j \rangle \langle e_i f_j| \\
= I - T,
\]

which proves equation (20).

References

[1] Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 Phys. Rev. A 53 2046–52
Nielsen M A 1999 Phys. Rev. Lett. 83 436
Vidal G 1999 Phys. Rev. Lett. 83 1046
[2] Lo H K and Popescu S 2001 Phys. Rev. A 63 022301
[3] Hayashi M 2006 Quantum Information: An Introduction (Berlin: Springer)
[4] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Phys. Rev. A 54 3824
[5] Bennett C H, DiVincenzo D P, Fuchs C A, Mor T, Rains E, Shor P W, Smolin J A and Wootters W K 1999 Phys. Rev. A 50 1070–91
Ghosh S, Kar G, Roy A and Sarkar D 2004 Phys. Rev. A 70 022304
Fan H 2004 Phys. Rev. Lett. 92 177905
Nathanson M 2005 J. Math. Phys. 46 062103
[6] Groisman B and Vaidman L 2001 J. Phys. A: Math. Gen. 34 6881
[7] Cohen S M 2007 Phys. Rev. A 75 052313
[8] Holevo A S 1998 IEEE Trans. Inform. Theory 44 269
Buranshev M V and Holevo A S 1998 Problems Inform. Transmission 34 97
Schumacher B and Westmoreland M D 1997 Phys. Rev. A 56 131
Hayashi M and Nagaoka H 2003 IEEE Trans. Inform. Theory 49 1753
Ogawa T and Nagaoka H 2007 IEEE Trans. Inform. Theory 53 2261–6
Winter A 2005 Preprint quant-ph/0507045
[9] Ambainis A 1999 Preprint quant-ph/9902053
[10] Hayashi M, Kawachi A and Kobayashi H 2008 Quantum Inform. Comput. at press (Preprint quant-ph/0604174)
[11] Virmani S, Sacchi M, Plenio M B and Markham D 2001 Phys. Lett. A 288 62
[12] Walgate J, Short A J, Hardy L and Vedral V 2000 Phys. Rev. Lett. 85 4972

New Journal of Physics 10 (2008) 013006 (http://www.njp.org/)
[13] Alagic G, Moore C and Russell A 2005 Preprint quant-ph/0511054
Grigni M, Schulman L J, Vazirani M and Vazirani U 2004 Combinatorica 24 137
Hallgren S 2002 Proc 34th Annu. ACM Symp. on Theory of Computing p 653
Hallgren S 2005 Proc. 37th Annu. ACM Symp. on Theory of Computing p 468
Kempe J and Shalev A 2005 Proc. 16th ACM-SIAM Symp. on Discrete Algorithms p 1118
Moore C, Rockmore D, Russell A and Schulman L J 2004 Proc. 15th ACM-SIAM Symp. on Discrete Algorithms p 1106
Moore C and Russell A 2005 Preprint quant-ph/0501177

[14] Lehmann E L 1986 Testing Statistical Hypotheses 2nd edn (New York: Wiley)

[15] Hayashi M, Markham D, Murao M, Owari M and Virmani S 2006 Phys. Rev. Lett. 96 040501

[16] Rains E M 2001 IEEE Trans. Inform. Theory 47 2921
Donald M J, Horodecki M and Rudolph O 2002 J. Math. Phys. 43 4252

[17] Audenaert K M R, Calsamiglia J, Masanes L, Munoz-Tapia R, Acin A, Bagan E and Verstraete F 2007 Phys. Rev. Lett. 98 160501

[18] Nussbaum M and Szoka A 2006 A lower bound of Chernoff type for symmetric quantum hypothesis testing Preprint quant-ph/0607216

[19] Massar S and Popescu S 1995 Optimal extraction of information from finite quantum ensembles Phys. Rev. Lett. 74 1259

[20] Bucklew J A 1990 Large Deviation Techniques in Decision, Simulation and Estimation (New York: Wiley)

[21] Hoeffding W 1965 Asymptotically optimal test for multinomial distributions Ann. Math. Stat. 36 369–400

[22] Hiai F and Petz D 1991 The proper formula for relative entropy and its asymptotics in quantum probability Commun. Math. Phys. 143 99–114

[23] Ogawa T and Nagaoka H 2000 Strong converse and Stein’s lemma in quantum hypothesis testing IEEE Trans. Inform. Theory 46 2428–33

[24] Hayashi M 2008 Error exponent in asymmetric quantum hypothesis testing and its application to classical–quantum channel coding Phys. Rev. A at press (Preprint quant-ph/0611013)

[25] Nagaoka H 2006 The converse part of the theorem for quantum Hoeffding bound Preprint quant-ph/0611289

[26] Vidal G and Tarrach R 1999 Phys. Rev. A 59 141–55

[27] Bennett C H, DiVincenzo D P, Mor T, Shor P W, Smolin J A and Terhal B M 1999 Phys. Rev. Lett. 82 5385

[28] Virginani S and Markham D 2004 private communication

[29] Hayashi M, Markham D, Murao M, Owari M and Virmani S 2008 Entanglement and group symmetries: stabilizer, symmetric and antisymmetric states Phys. Rev. A at press (Preprint 0710.1056)

[30] Rains E M 2001 IEEE Trans. Inform. Theory 47 2921

[31] Wootters W K and Fields B C 1989 Ann. Phys. 191 363

[32] Owari M and Hayashi M in preparation