Local distinguishability of quantum states in infinite-dimensional systems

2006 J. Phys. A: Math. Gen. 39 3059

(http://iopscience.iop.org/0305-4470/39/12/014)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 3059–3069 doi:10.1088/0305-4470/39/12/014

Local distinguishability of quantum states in infinite-dimensional systems

Yoshiko Ogata

Department of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8914, Japan

Received 4 July 2005, in final form 23 January 2006
Published 8 March 2006
Online at stacks.iop.org/JPhysA/39/3059

Abstract

We investigate local distinguishability of quantum states by use of convex analysis of *joint numerical range* of operators on a Hilbert space. We show that any two orthogonal pure states are distinguishable by local operations and classical communications, even for infinite-dimensional systems. An estimate of the local discrimination probability is also given for some families of more than two pure states.

PACS numbers: 03.65.Db, 03.65.Ud

1. Introduction

Local operations and classical communications (LOCC) are basic operations in quantum information theory. Many interesting studies have arisen from the question, what we can/cannot do using only LOCC. The question is highly non-trivial and difficult to solve due to the lack of simple characterization of LOCC. The necessary and sufficient condition for the deterministic convertibility of one pure state to the other was derived by Nielsen for general bipartite systems in [1]. Furthermore, in [2], Vidal obtained the optimal probability of converting one pure state to the other, non-deterministically. However, when we start thinking of simultaneous convertibility of more than one state, the problem becomes more difficult, because of the fact that the Lo–Popescu theorem [3] is not applicable there. The Lo–Popescu theorem alternates the given LOCC with one-way operation which *depends on* the given states; hence it is not applicable when we consider the operation on more than one state. Instead of considering the general simultaneous convertibility problem, we consider some particular situations, like distinguishing or copying quantum states by LOCC [4, 5], which are already nontrivial interesting questions.

In this paper, we study the distinguishability problem, in bipartite systems. In [4], Walgate *et al* proved that any two orthogonal pure states in finite-dimensional systems are distinguishable. Unfortunately, because of the nature of their proof (that is constructed from the two-dimensional case), this important result has been restricted to finite-dimensional systems so far. As it is indispensable to consider infinite-dimensional systems in the real world, the analogous result in infinite-dimensional systems is desirable. In this paper, we prove the infinite version:

Theorem. *Any two orthogonal pure states in a bipartite system are distinguishable by LOCC, even for infinite-dimensional systems.*

In spite of these simple results for two pure states, it is known that more than two pure states are not always distinguishable by LOCC. It was proved that three Bell states cannot be distinguished with certainty by LOCC and four Bell states cannot, even probabilistically [6]. A set of non-entangled pure states that are not locally distinguishable was introduced in [7]. The probability of the discrimination for the worst case was estimated in [8]. In [9], a family of linearly independent states, given by the generalized Pauli matrices, were shown to be indistinguishable, deterministically or probabilistically. Furthermore, recently, the necessary condition for the perfect LOCC discrimination of general multipartite states was given in terms of the global robustness of entanglement in [10]. In this paper, we give an estimate of discrimination probability for some families of more than two pure states. This result also holds for infinite-dimensional systems.

The key notion of our approach is the *joint numerical range* of operators (see the definition below). We represent pure states in terms of Hilbert Schmidt operators and investigate trace class operators given by them. We see that the convexity of the joint numerical range of these trace class operators restricted to arbitrary sub-Hilbert spaces implies the distinguishability of states. From the fact that the convexity condition holds for a joint numerical range of two self-adjoint operators, the distinguishability of two pure states is proven.

The remainder of the paper is organized in the following way: in section 2, we introduce a representation of a vector in $\mathcal{H}_A \otimes \mathcal{H}_B$ as an operator from $\mathcal{H}_B$ to $\mathcal{H}_A$. And from them, we define the real vector space $\mathcal{K}$. Then we represent our main results in terms of the vector space $\mathcal{K}$. In section 3, we correlate the convexity of the joint numerical range of the basis operators of $\mathcal{K}$. The theorems are proven in section 4.

2. The distinguishability of states

In this section, we introduce a representation of pure states on $\mathcal{H}_A \otimes \mathcal{H}_B$ as operators from $\mathcal{H}_B$ to $\mathcal{H}_A$, and describe our main results in terms of the operator representation. In finite-dimensional systems, the operator representation corresponds to the well-known matrix representation of states, by use of a maximal entangled state. (See for example [11].)

Let $\mathcal{H}_A, \mathcal{H}_B$ be separable (possibly infinite dimensional) Hilbert spaces. Let us fix some orthonormal basis $\{f_i\}$ of $\mathcal{H}_B$. A vector ψ in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be decomposed as

$$\psi = \sum_i \varphi_i \otimes f_i,$$

in general. The vectors φ_i in $\mathcal{H}_A$ satisfy

$$\sum_i \|\varphi_i\|^2 = \|\psi\|^2. \tag{1}$$

Now we define a bounded linear operator X from $\mathcal{H}_B$ to $\mathcal{H}_A$ by

$$X\eta \equiv \sum_i \langle f_i|\eta\rangle \cdot \varphi_i, \qquad \forall \eta \in \mathcal{H}_B. \tag{2}$$

From (1), the sum in (2) absolutely converges in norm of $\mathcal{H}_B$, and we obtain $\|X\| \leqslant \|\psi\|$. Then the vector ψ is represented as

$$\psi = \sum_i \varphi_i \otimes f_i = \sum_i (X f_i) \otimes f_i.$$

The bounded operator X^*X on $\mathcal{H}_B$ satisfies

$$\mathrm{Tr}\, X^*X = \sum_i \|\varphi_i\|^2 = \|\psi\|^2 < \infty, \tag{3}$$

i.e., X^*X is a trace class operator on $\mathcal{H}_B$. By operating $1 \otimes |f_i\rangle\langle f_i|$ on ψ, we see that X is the unique operator such that $\psi = \sum_i X f_i \otimes f_i$. On the other hand, for any bounded linear operator X from $\mathcal{H}_B$ to $\mathcal{H}_A$ satisfying $\mathrm{Tr}\, X^*X < \infty$, there exists a unique vector $\sum_i X f_i \otimes f_i$. Hence we obtain the following one-to-one correspondence:

$$\psi \in \mathcal{H}_A \otimes \mathcal{H}_B \quad \Leftrightarrow \quad X \in B(\mathcal{H}_B, \mathcal{H}_A), \quad s.t. \quad \mathrm{Tr}\, X^*X < \infty,$$

through the relation

$$\psi = \sum_i (X f_i) \otimes f_i. \tag{4}$$

Here $B(\mathcal{H}_B, \mathcal{H}_A)$ indicates the set of bounded operators from $\mathcal{H}_B$ to $\mathcal{H}_A$.

Now let us consider a set of orthonormal M vectors $\psi_1, \ldots, \psi_M$ in $\mathcal{H}_A \otimes \mathcal{H}_B$. We can associate each ψ_l with an operator X_l through (4). As in (3), $X_m^* X_l$ are trace class operators on $\mathcal{H}_B$ for all $1 \leqslant m, l \leqslant M$ and satisfy

$$\mathrm{Tr}\, X_m^* X_l = \langle \psi_m, \psi_l \rangle = \delta_{m,l}, \qquad 1 \leqslant m, \quad l \leqslant M. \tag{5}$$

Let $\mathcal{K}$ be the real linear subspace of trace class self-adjoint operators on $\mathcal{H}_B$ spanned by operators $\{X_m^* X_l + X_l^* X_m, \mathrm{i}(X_m^* X_l - X_l^* X_m)\}_{m \neq l}$. Let N be the dimension of $\mathcal{K}$ and $(A_1, \ldots, A_N)$ an arbitrary basis of $\mathcal{K}$. The dimension N is bounded as $N \leqslant M(M-1)$. Because each $X_m^* X_l$ satisfies (5), we have

$$\mathrm{Tr}\, A_i = 0, \qquad i = 1, \ldots, N. \tag{6}$$

We will call $\mathcal{K}$ the real vector space of trace class self-adjoint operators associated with $\psi_1, \ldots, \psi_M$.

Now we are ready to state our main results. In this paper, we show the following theorems:

Theorem 2.1. *Let $\mathcal{H}_A, \mathcal{H}_B$ be (possibly infinite dimensional) separable Hilbert spaces. Let $\psi_1, \ldots, \psi_M$ be a set of orthogonal pure states in $\mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{K}$ the associated real vector space of trace class self-adjoint operators on $\mathcal{H}_B$. Then if the dimension of $\mathcal{K}$ is 2, the states $\psi_1, \ldots, \psi_M$ are distinguishable by LOCC with certainty. In particular, any pair of orthogonal pure states ψ_1, ψ_2 are distinguishable by LOCC with certainty.*

Theorem 2.2. *Let $\mathcal{H}_A, \mathcal{H}_B$ be (possibly infinite dimensional) separable Hilbert spaces. Let $\psi_1, \ldots, \psi_M$ be a set of orthogonal pure states in $\mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{K}$ the associated real vector space of trace class self-adjoint operators on $\mathcal{H}_B$. Suppose that the dimension of $\mathcal{K}$ is 3. Then $\psi_1, \ldots, \psi_M$ are distinguishable by conclusive LOCC protocol with probability P_d such that*

$$P_d \geqslant 1 - \max_{1 \leqslant l \leqslant M} \left(\sum_{k=1}^{2} p_k^l \right).$$

Here, p_k^l represents the kth Schmidt coefficient of ψ_l, ordered in the decreasing order.

Remark 2.3. The last statement of theorem 2.1 is the extension of [4] to an infinite-dimensional system. Applying the argument in [4], we can extend the result to multipartite systems: any two orthogonal pure states in multipartite systems are distinguishable by LOCC even in infinite-dimensional systems.

Remark 2.4. In [12], Virmani *et al* showed that any two (even non-orthogonal) multipartite pure states in finite-dimensional systems can be optimally distinguished using only LOCC. It was derived using the result of the orthogonal case in [4]. The argument there can be applied to our infinite-dimensional case. Therefore, any two bipartite pure states can be optimally distinguished using only LOCC, even for infinite-dimensional systems.

Example 2.5. Let $\{|e_k\rangle\}_{k=1}^{\infty}$, $\{|f_k\rangle\}_{k=1}^{\infty}$ be the orthonormal basis of $\mathcal{H}_A, \mathcal{H}_B$, respectively. Then for orthogonal three states $|\psi_1\rangle, \ldots, |\psi_3\rangle$ given by

$$|\psi_1\rangle = \sum_k c_k |e_{2k}\rangle \otimes |f_{2k}\rangle,$$

$$|\psi_2\rangle = \sum_k d_k |e_{2k+1}\rangle \otimes |f_{2k+1}\rangle,$$

$$|\psi_3\rangle = \sum_k d_k |e_{2k}\rangle \otimes |f_{2k+1}\rangle + c_k |e_{2k+1}\rangle \otimes |f_{2k}\rangle,$$

the dimension of $\mathcal{K}$ is 2. Therefore, by theorem 2.1 these states are deterministically distinguishable by LOCC.

Example 2.6. Let us consider three orthogonal simultaneously Schmidt decomposable pure states [13] with real coefficients

$$|\psi_i\rangle = \sum_k c_k^i |e_k\rangle \otimes |f_k\rangle, \qquad c_k^i \in \mathbb{R}, \quad c_1^i \geqslant c_2^i \geqslant \cdots, \quad i = 1, 2, 3.$$

The dimension of the associated real vector space $\mathcal{K}$ is then less than 3, because the imaginary parts of the associated operators are 0. Therefore, from theorem 2.2, they are distinguishable with probability larger than

$$1 - \max_{i=1,2,3} \left(\left|c_1^i\right|^2 + \left|c_2^i\right|^2 \right).$$

For example, the following vectors are distinguishable with probability more than $(2 - 2r + r^2)/(2(1+r))$, for r small enough:

$$|\psi_1\rangle = \sqrt{\frac{r(2-r)}{2(r+1)}} \left[\sqrt{\frac{2r}{r(2-r)}} |e_0\rangle \otimes |f_0\rangle + \sum_{k=0}^{\infty} (1-r)^k (|e_{3k+1}\rangle \otimes |f_{3k+1}\rangle - |e_{3k+2}\rangle \otimes |f_{3k+2}\rangle) \right]$$

$$|\psi_2\rangle = \sqrt{\frac{r(2-r)}{3+2r+2r^2}} \left[-\sqrt{\frac{2r}{r(2-r)}} |e_0\rangle \otimes |f_0\rangle + \sum_{k=0}^{\infty} (1-r)^k ((1+r)|e_{3k+1}\rangle \otimes |f_{3k+1}\rangle + (1-r)|e_{3k+2}\rangle \otimes |f_{3k+2}\rangle + |e_{3k+3}\rangle \otimes |f_{3k+3}\rangle) \right]$$

$$|\psi_3\rangle = \sqrt{\frac{r(2-r)}{6}} \left[\sum_{k=0}^{\infty} (1-r)^k \left(|e_{3k+1}\rangle \otimes |f_{3k+1}\rangle + |e_{3k+2}\rangle \otimes |f_{3k+2}\rangle - 2|e_{3k+3}\rangle \otimes |f_{3k+3}\rangle \right) \right].$$

Here, $\{e_k\}_{k=0}^{\infty} \{f_k\}_{k=0}^{\infty}$ are CONS of Alice and Bob, respectively.

3. Distinguishability and joint numerical range

In this section, we introduce the key notion of our proof, *joint numerical range*. Then we state how it is related to the distinguishability problem.

Let $(A_1, \ldots, A_N)$ be bounded self-adjoint operators on a Hilbert space $\mathcal{H}$. A subset of $\mathbb{R}^N$ given by

$$\{(\langle z, A_1 z\rangle, \langle z, A_2 z\rangle, \ldots, \langle z, A_N z\rangle), z \in \mathcal{H}, \|z\| = 1\} \subset \mathbb{R}^N$$

is called the *joint numerical range* (or generalized numerical range) of $(A_1, \ldots, A_N)$. Furthermore, for an orthogonal projection P on $\mathcal{H}$, we will call the set

$$C_P(A_1, \ldots, A_N) \equiv \{(\langle z, A_1 z\rangle, \langle z, A_2 z\rangle, \ldots, \langle z, A_N z\rangle); z \in P\mathcal{H}, \|z\| = 1\} \subset \mathbb{R}^N,$$

the joint numerical range of $(A_1, \ldots, A_N)$ restricted to the sub-Hilbert space $P\mathcal{H}$. The distinguishability of states is related to the convexity of joint numerical ranges as follows:

Proposition 3.1. *Let $\psi_1, \ldots, \psi_M$ be a set of orthogonal pure states in $\mathcal{H}_A \otimes \mathcal{H}_B$, and $\mathcal{K}$ the associated real vector space of trace class self-adjoint operators on $\mathcal{H}_B$. Let $(A_1, \ldots, A_N)$ be a basis of $\mathcal{K}$. Suppose that for any projection P on $\mathcal{H}_B$, $C_P(A_1, \ldots, A_N)$ is convex. Then the states $\psi_1, \ldots, \psi_M$ are distinguishable by LOCC with certainty.*

Proposition 3.2. *Let $\psi_1, \ldots, \psi_M$ be a set of orthonormal pure states in $\mathcal{H}_A \otimes \mathcal{H}_B$, and $\mathcal{K}$ the associated real vector space of trace class self-adjoint operators on $\mathcal{H}_B$. Let $(A_1, \ldots, A_N)$ be a basis of $\mathcal{K}$. Suppose that for any projection P of $\mathcal{H}_B$ with dimension larger than N_p, $C_P(A_1, \ldots, A_N)$ is convex. Then the states $\psi_1, \ldots, \psi_M$ are distinguishable by LOCC with the probability P_d such that*

$$P_d \geqslant 1 - \max_{1 \leqslant l \leqslant M} \left(\sum_{k=1}^{N_p} p_k^l \right).$$

Here, p_k^l represents the kth Schmidt coefficient of ψ_l, ordered in the decreasing order.

Theorems 2.1 and 2.2 are derived as corollaries of these propositions.

Remark 3.3. In general, joint numerical range is not convex. For example, it is easy to see that the joint numerical range of the Pauli operators on $\mathbb{C}^2$ is not convex (see [14]).

Propositions 3.1, 3.2 will be proved in section 4. Here, we explain the proof of proposition 3.1 briefly, for the finite N-dimensional case, in order to give readers the essence of the proof: as operators $\{X_l\}$ satisfy the condition (5), for an orthonormal basis $\{e_k\}_k$, we have

$$\sum_{k=1}^{D} \langle e_k, X_l^* X_m e_k\rangle = 0,$$

for all l, m, which is equivalent to

$$\frac{1}{D} \sum_{k=1}^{D} (\langle e_k, A_1 e_k\rangle, \ldots, \langle e_k, A_N e_k\rangle) = 0.$$

Note that the left-hand side of this equation is the convex combination of elements in the joint numerical range. Therefore, if the convexity assumption in proposition 3.1 holds, 0 is in the joint numerical range. In other words, there exists a normalized vector $|g_1\rangle$ such that

$$(\langle g_1, A_1 g_1\rangle, \ldots, \langle g_1, A_N g_1\rangle) = 0. \tag{7}$$

This is equivalent to having

$$\langle g_1, X_l^* X_m g_1\rangle = 0,$$

for all l, m. Let us consider the restriction of $\{A_i\}$ to the sub-Hilbert space orthogonal to $|g_1\rangle$. We can repeat the same argument as above and obtain a normalized vector $|g_2\rangle$ orthogonal to $|g_1\rangle$, such that

$$\langle g_2, X_l^* X_m g_2\rangle = 0.$$

Repeating this argument, we obtain an orthonormal basis $\{g_k\}_{k=1}^D$ such that

$$\langle g_k, X_l^* X_m g_k\rangle = 0, \qquad k = 1, \ldots, D.$$

From this, we have a decomposition

$$\psi_l = \sum_k \xi_k^l \otimes \bar{g}_k, \qquad l = 1, \ldots, M.$$

where $|\bar{g}_k\rangle$ is the complex conjugation of $|g_k\rangle$ and $\left\{\left|\xi_k^m\right\rangle\right\}$ satisfying

$$\left\langle \xi_k^l \middle| \xi_k^m \right\rangle = 0 \qquad \forall l \neq m, \quad \forall k.$$

Hence as in [4], the states are deterministically distinguishable.

The proof of the infinite-dimensional case is basically the same. The main difference is that $\mathrm{Tr}\, X_l^* X_m$ cannot be regarded as scalar multiplication of the convex combination of elements in the joint numerical range. ($D = \infty$ for this case.) The point is that we still can show the existence of a vector satisfying (7). This is proven in lemma 4.2 in the next section.

4. Proof of theorems 2.1 and 2.2

First we prove proposition 3.1. The proof consists of four steps:

Step 1. First, we show that if $\mathcal{H}_B$ has an orthonormal basis $\{g_k\}$ such that $\langle g_k, A_i g_k\rangle = 0$ for all $i = 1, \ldots, N$ and k, then, $\psi_1, \ldots, \psi_N$ are distinguishable by LOCC (lemma 4.1).

Step 2. Second, using convex analysis, we show that if the joint numerical range of $(A_1, \ldots, A_N)$ is convex, there exists at least one vector $z \in \mathcal{H}_B$ such that $\langle z, A_i z\rangle = 0$ for all $i = 1, \ldots, N$ (lemma 4.2).

Step 3. Third, using lemma 4.2, we show the existence of the orthonormal basis satisfying the desired condition in *step 1* (lemma 4.4).

Step 4. Finally, combining the results of *step 1* and *step 3*, we obtain proposition 3.1.

Now let us start the proof. First we show the following lemma:

Lemma 4.1. *Let $(A_1, \ldots, A_N)$ be a basis of $\mathcal{K}$ associated with $\psi_1, \ldots, \psi_M$. Suppose that there exists an orthonormal basis $\{g_k\}$ of $\mathcal{H}_B$ such that*

$$\langle g_k, A_i g_k\rangle = 0, \qquad \forall k, \quad i = 1 \ldots N. \tag{8}$$

Then the states $\psi_1, \ldots, \psi_M$ are distinguishable by LOCC.

Proof. In order to investigate distinguishability, we look for a suitable decomposition of the states. Let us decompose the vectors $\psi_1, \ldots, \psi_M$ with respect to an orthonornal basis $\{e_k\}$ of $\mathcal{H}_B$:

$$\psi_l = \sum_k \xi_k^l \otimes e_k, \qquad l = 1, \ldots, M. \tag{9}$$

If the orthogonal conditions

$$\left\langle \xi_k^l \middle| \xi_k^m \right\rangle = 0 \qquad \forall l \neq m, \quad \forall k \tag{10}$$

hold, then Alice and Bob can distinguish these states by LOCC [4]. (Note that this orthogonality condition does not hold in general.) Therefore, it suffices to show the existence of this decomposition.

Let $\{f_i\}$ be the orthonormal basis fixed in section 2. (Recall that we defined the operators X_ls in terms of $\{f_i\}$.) We define an antilinear operator $J : \mathcal{H}_B \to \mathcal{H}_B$ to be the complex conjugation with respect to $\{f_i\}$:

$$J \sum_i \alpha_i f_i \equiv \sum_i \bar{\alpha_i} f_i.$$

As J is an antilinear isometry, $\{J g_k\}$ is an orthonormal basis of $\mathcal{H}_B$. Therefore, we can decompose $\psi_1, \ldots, \psi_M$ with respect to $\{J g_k\}$:

$$\psi_l = \sum_k \xi_k^l \otimes J g_k. \tag{11}$$

We show that for each k, $\left\{\xi_k^1, \ldots, \xi_k^M\right\}$ are mutually orthogonal.

Let us decompose ψ_l with respect to $\{f_i\}$:

$$\psi_l = \sum_i \varphi_i^l \otimes f_i. \tag{12}$$

Comparing (11) and (12), we obtain

$$\xi_k^l = \sum_i \varphi_i^l \langle J g_k, f_i \rangle = \sum_i \varphi_i^l \langle f_i, g_k \rangle = X_l g_k.$$

As $(A_, \ldots, A_N)$ is a basis of $\mathcal{K}$, the assumption (8) implies

$$\left\langle \xi_k^l, \xi_k^m \right\rangle = \langle X_l g_k, X_m g_k \rangle = 0 \qquad \forall l \neq m, \quad \forall k.$$

Hence for each k, $\left\{\xi_k^1, \ldots, \xi_k^M\right\}$ are mutually orthogonal.

Thus (11) takes the form of (9), with the orthogonality condition (10). Therefore, we can distinguish $\psi_1, \ldots, \psi_M$ by LOCC with certainty. □

Next we show the following lemma which holds on a general Hilbert space $\mathcal{H}$:

Lemma 4.2. *Let $(A_1, \ldots, A_N)$ be a set of trace class self-adjoint operators on a Hilbert space $\mathcal{H}$ such that* $\operatorname{Tr} A_i = 0$ *for each $1 \leqslant i \leqslant N$. Suppose that the joint numerical range of $(A_1, \ldots, A_N)$ is a convex subset of $\mathbb{R}^N$. Then there exists a vector $z \in \mathcal{H}$ with $\|z\| = 1$ such that*

$$\langle z, A_i z \rangle = 0, \qquad i = 1, \ldots, N.$$

Proof. Before starting the proof, we review some basic facts from convex analysis [15]. Let $x_1, \ldots, x_k$ be elements in $\mathbb{R}^N$. An element $\sum_{i=1}^k \alpha_i x_i$ with real coefficients α_i satisfying $\sum_{i=1}^k \alpha_i = 1$ is called an affine combination of $x_1, \ldots, x_k$. An affine manifold in $\mathbb{R}^N$ is a set containing all its affine combinations. Let S be a nonempty subset of $\mathbb{R}^N$. The affine hull of S is defined to be the smallest affine manifold containing S. We denote the affine hull of S by $\operatorname{aff} S$. In other words, $\operatorname{aff} S$ is the affine manifold generated by S. As easily seen, it is a closed plane parallel to a linear subspace in $\mathbb{R}^N$. Its dimension may be lower than N in general. The relative interior of S, $\operatorname{ri} S$, is the interior of S with respect to the topology relative to $\operatorname{aff} S$. In other words,

$$\operatorname{ri} S \equiv \{x \in S; \exists\, \varepsilon > 0 \ s.t.\ B(x, \varepsilon) \cap \operatorname{aff} S \subset S\}.$$

Here, $B(x, \varepsilon)$ is a ball of radius ε, centred at x. The following fact is known: □

Lemma 4.3. *Let C be a nonempty convex subset of $\mathbb{R}^N$. Then for any point x_0 in* $\mathrm{aff}C\backslash\mathrm{ri}C$, *there exists a non-zero vector $s \in \mathbb{R}^N$ parallel to* $\mathrm{aff}C$, *such that*

$$\langle\langle s, x - x_0\rangle\rangle \geqslant 0, \qquad \forall x \in C.$$

Here $\langle\langle , \rangle\rangle$ is the inner product of $\mathbb{R}^N$:

$$\langle\langle s, x\rangle\rangle \equiv \sum_{i=1}^{N} s_i \cdot x_i.$$

Now we are ready to prove lemma 4.2. The claim is equivalent to saying that 0 is included in the joint numerical range of the operators $(A_1, \ldots, A_N)$. We denote the joint numerical range by C_1:

$$C_1 \equiv \{(\langle z, A_1 z\rangle, \langle z, A_2 z\rangle, \ldots, \langle z, A_N z\rangle) \in \mathbb{R}^N, z \in \mathcal{H}, \|z\| = 1\}.$$

By assumption, C_1 is a nonempty convex subset of $\mathbb{R}^N$. Let $\{e_k\}$ be an arbitrary orthonormal basis of $\mathcal{H}$. By the definition of C_1,

$$x_k \equiv (\langle e_k, A_1 e_k\rangle, \ldots, \langle e_k, A_N e_k\rangle)$$

is an element of C_1 for each k.

The finite-dimensional case $\mathcal{H} = \mathbb{C}^n$ is immediate. By the convexity of C_1, we obtain

$$0 = \frac{1}{n}(\mathrm{Tr}\, A_1, \ldots, \mathrm{Tr}\, A_N) = \frac{1}{n}\sum_{k=1}^{n}(\langle e_k, A_1 e_k\rangle, \ldots, \langle e_k, A_N e_k\rangle) \in C_1.$$

Below we prove the infinite-dimensional case.

First we observe that 0 is included in the closure of C_1. In particular, 0 is in $\mathrm{aff}C_1$. To see this, note that for all $l \in \mathbb{N}$, we have

$$\frac{1}{l}\sum_{k=1}^{l}(\langle e_k, A_1 e_k\rangle, \ldots, \langle e_k, A_N e_k\rangle) \in C_1.$$

As A_i is a trace class operator, the sum $\sum_{k=1}^{\infty}\langle e_k, A_i e_k\rangle$ converges absolutely. By taking $l \to \infty$ limit, we obtain

$$0 = \lim_{l\to\infty}\frac{1}{l}\sum_{k=1}^{l}(\langle e_k, A_1 e_k\rangle, \ldots, \langle e_k, A_N e_k\rangle) \in \overline{C_1} \subset \mathrm{aff}C_1.$$

Hence 0 is in $\mathrm{aff}C_1$.

Second, we show that 0 is actually in $\mathrm{ri}C_1$. To prove this, assume 0 is not included in $\mathrm{ri}C_1$. Then it is an element of $\mathrm{aff}C_1\backslash\mathrm{ri}C_1$. As C_1 is a nonempty convex set, from lemma 4.3, there exists a non-zero vector $s = (s_1, \ldots, s_N) \in \mathbb{R}^N$ parallel to $\mathrm{aff}C_1$, such that

$$\langle\langle s, x\rangle\rangle \geqslant 0, \qquad \forall x \in C_1.$$

As $x_k \in C_1$, we have

$$\langle\langle s, x_k\rangle\rangle \geqslant 0, \tag{13}$$

for all k. On the other hand, we have

$$\sum_{k=1}^{\infty}\langle\langle s, x_k\rangle\rangle = \sum_{i=1}^{N} s_i \sum_{k=1}^{\infty} \cdot \langle e_k, A_i e_k\rangle = \sum_{i=1}^{N} s_i \cdot \mathrm{Tr}\, A_i = 0. \tag{14}$$

From (13) and (14), we obtain $\langle\langle s, x_k\rangle\rangle = 0$ for all k. As the orthonormal basis $\{e_k\}$ can be taken arbitrarily, we obtain

$$\langle\langle s, x\rangle\rangle = 0, \qquad \forall x \in C_1.$$

As s is a non-zero vector parallel to $\mathrm{aff}C_1$, this means that C_1 is included in some affine manifold that is strictly smaller than $\mathrm{aff}C_1$. This contradicts the definition of $\mathrm{aff}C_1$. (Recall that $\mathrm{aff}C_1$ is the smallest affine manifold including C_1.) Therefore, we obtain $0 \in \mathrm{ri}C_1$. In particular, $0 \in C_1$ and this completes the proof.

Using lemma 4.2, we obtain the following lemma:

Lemma 4.4. *Let* $(A_1, \ldots, A_N)$ *be a set of trace class self-adjoint operators on a Hilbert space* $\mathcal{H}$ *such that* $\mathrm{Tr}\, A_i = 0$ *for each* $1 \leqslant i \leqslant N$. *Suppose that for every orthogonal projection P on* $\mathcal{H}$, $\mathcal{C}_\mathcal{P}(\mathcal{A}_\infty, \ldots, \mathcal{A}_\mathcal{N})$ *is convex. Then there exists an orthonormal basis* $\{g_k\}$ *of* $\mathcal{H}$, *such that*

$$\langle g_k, A_i g_k\rangle = 0, \qquad \forall i = 1, \ldots N, \quad \forall k.$$

Proof. We will say that a set of vectors Z in $\mathcal{H}$ satisfies *property** if it satisfies the following conditions:

*Property**

(i) Z is a set of mutually orthogonal unit vectors of $\mathcal{H}$.
(ii) $\langle z, A_i z\rangle = 0$, $i = 1, \ldots, N$ for all $z \in Z$.

By Zorn's lemma, there exists a maximal set of orthonormal vectors $\{g_k\}$ in $\mathcal{H}$ which satisfies the *property**. It suffices to show that $\{g_k\}$ is complete.

Suppose that $\{g_k\}$ is not complete in $\mathcal{H}$, and let P be the orthogonal projection onto the sub-Hilbert space spanned by $\{g_k\}$. From *property**, we have

$$\mathrm{Tr}\, P A_i P = \sum_k \langle g_k, A_i g_k\rangle = 0, \qquad i = 1, \ldots N.$$

Let $\bar{P} = 1 - P$. Now we regard $(\bar{P} A_1 \bar{P}, \ldots, \bar{P} A_N \bar{P})$ as self-adjoint trace class operators on the Hilbert space $\bar{P}\mathcal{H}$ such that

$$\mathrm{Tr}_{\bar{P}\mathcal{H}}\, (\bar{P} A_i \bar{P}) = \mathrm{Tr}\,(A_i) - \mathrm{Tr}\,(P A_i P) = 0, \qquad i = 1, \ldots N.$$

By the assumption, the joint numerical range of $(\bar{P} A_1 \bar{P}, \ldots, \bar{P} A_N \bar{P})$ on $\bar{P}\mathcal{H}$ is convex. Thus, applying lemma 4.2, there exists a unit vector $z \in \bar{P}\mathcal{H}$ such that $\langle z, A_i z\rangle = 0$ for all $i = 1, \ldots, N$. As z is orthogonal to all g_k, the set $\{z\} \cup \{g_k\}$ satisfies the *property**, and is strictly larger than $\{g_k\}$. This contradicts the maximality of $\{g_k\}$. Therefore, $\{g_k\}$ is complete. □

Now, let us complete the proof of proposition 3.1. The basis of $\mathcal{K}$, $(\mathcal{A}_\infty, \ldots, \mathcal{A}_\mathcal{N})$ are trace class self-adjoint operators satisfying $\mathrm{Tr}\, A_i = 0, i = 1, \ldots, N$ (6). Therefore if $C_P(A_1, \ldots, A_N)$ is a convex subset of $\mathbb{R}^N$ for any orthogonal projection P on $\mathcal{H}_B$, there exists an orthonormal basis $\{g_k\}$ of $\mathcal{H}_B$ such that $\langle g_k, A_i g_k\rangle = 0$, for all $i = 1, \ldots N$ and k, from lemma 4.4. By lemma 4.1, this concludes that $\psi_1, \ldots, \psi_M$ are distinguishable by LOCC.

Proposition 3.2 can be shown in the same way. We have the following lemma:

Lemma 4.5. *Let* $(A_1, \ldots, A_N)$ *be a set of trace class self-adjoint operators on a Hilbert space* $\mathcal{H}$ *such that* $\mathrm{Tr}\, A_i = 0$ *for each* $1 \leqslant i \leqslant N$. *Suppose that for every orthogonal projection P on* $\mathcal{H}$ *with dimension larger than* N_p, $C_P(A_1, \ldots, A_N)$ *is convex. Then there exists an orthonormal basis* $\{g_k\}$ *of* $\mathcal{H}$, *such that*

$$\langle g_k, A_i g_k\rangle = 0, \qquad i = 1, \ldots N, \quad \forall k > N_p.$$

Proof. The same as the proof of lemma 3.6. We can find a set of orthonormal vectors satisfying *property**, such that the dimension of its complementary subspace is N_p. □

Decomposing each ψ_l with respect to $\{Jg_k\}$, we obtain

$$\psi_l = \sum_k \xi_k^l \otimes Jg_k, \tag{15}$$

with

$$\langle \xi_k^l | \xi_k^m \rangle = 0 \qquad \forall l \neq m \qquad \forall k > N_p, \tag{16}$$

then $\psi_1, \ldots, \psi_M$ can be distinguished by conclusive LOCC protocol, probabilistically: Bob carries out the measurement $\{|Jg_k\rangle\langle Jg_k|\}$. We call it an error if Bob obtains $|e_k\rangle$ for $k \leqslant N_p$. Otherwise, Alice sequentially carries out the projective measurement given by $\{|xi_k^l\rangle\}_l$, and distinguishes the states. This conclusive protocol distinguishes the states with probability P_d, lower bounded as

$$P_d \geqslant 1 - \max_{1 \leqslant l \leqslant M} \sum_{k=1}^{N_p} \|\xi_k^l\|^2. \tag{17}$$

By the argument in the proof of lemma 4.1, (15) takes the form of (9) with the orthogonality condition. Therefore, for the protocol in the introduction, the probability that the error occurs is $\sum_{k=1}^{N_p} \|\xi_k^l\|^2$ when $\psi = \psi_l$. It is bounded from above as follows:

$$\sum_{k=1}^{N_p} \|\xi_k^l\|^2 = \sum_{k=1}^{N_p} \|(1 \otimes |Jg_k\rangle\langle Jg_k|)\, \psi_l\|^2$$
$$\leqslant \sup \left\{ \sum_{k=1}^{N_p} \|(1 \otimes |z_k\rangle\langle z_k|)\, \psi_l\|^2, \{z_k\}_{k=1}^{N_p} : \text{orthonormal set of } \mathcal{H}_B \right\} = \sum_{k=1}^{N_p} p_k^l.$$

Here, p_k^l is the kth Schmidt coefficient of ψ_l, ordered in the decreasing order. Therefore, $\psi_1, \ldots, \psi_M$

$$P_d \geqslant 1 - \max_{1 \leqslant l \leqslant M} \left(\sum_{k=1}^{N_p} \|\xi_k^l\|^2 \right) \geqslant 1 - \max_{1 \leqslant l \leqslant M} \left(\sum_{k=1}^{N_p} p_k^l \right),$$

and we obtain proposition 3.2.

Proof of theorems 2.1 and 2.2. Now we apply the known results about joint numerical range to propositions 3.1, 3.2 and derive theorems 2.1 and 2.2. For $N = 2$ case, the following theorem is known [16]: □

Theorem 4.6. *For any bounded self-adjoint operators T_1, T_2 on a separable Hilbert space $\mathcal{H}$, the set*

$$\{(\langle z, T_1 z\rangle, \langle z, T_2 z\rangle) \in \mathbb{R}^2, z \in \mathcal{H}, \|z\| = 1\}$$

is a convex subset of $\mathbb{R}^2$.

This is called the Toeplitz Hausdorff theorem. By this theorem, $C_P(A_1, A_2)$ is a convex subset of $\mathbb{R}^2$ for any projection P on $\mathcal{H}_B$. Therefore, applying proposition 3.1, we obtain theorem 2.1. The last statement comes from the fact $N \leqslant M(M-1) = 2$ for $M = 2$.

On the other hand, for $N = 3$, the next theorem is known [17, 18].

Theorem 4.7. *Let $\mathcal{H}$ be a separable Hilbert space with* $\dim \mathcal{H} \geqslant 3$. *Then for any self-adjoint operators T_1, T_2, T_3 in $\mathcal{H}$, the set*

$$\{(\langle z, T_1 z\rangle, \langle z, T_2 z\rangle, \langle z, T_3 z\rangle) \in \mathbb{R}^3, z \in \mathcal{H}, \|z\| = 1\}$$

is a convex subset of $\mathbb{R}^3$.

By this theorem, $C_P(A_1, A_2, A_3)$ is a convex subset of $\mathbb{R}^3$ for any projection P on $\mathcal{H}_B$ with dimension larger than 2. Therefore, applying proposition 3.2, we obtain theorem 2.2.

Acknowledgments

This work was supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

References

[1] Nielsen M A 1999 Conditions for a class of entanglement transformations *Phys. Rev. Lett.* **83** 436–9
[2] Vidal G 1999 Entanglement of pure states for a single copy *Phys. Rev. Lett.* **83** 1046–9
[3] Lo H-K and Popescu S 2001 Concentrating entanglement by local actions: beyond mean values *Phys. Rev.* A **63** 022301
[4] Walgate J, Short A J, Hardy L and Vedral V 2000 Local distinguishability of multipartite orthogonal quantum states *Phys. Rev. Lett.* **85** 4972–5
[5] Anselmi E, Chefles A and Plenio M B 2004 Local copying of orthogonal entangled quantum states *New J. Phys.* **6** 164
[6] Ghosh S, Kar G, Roy A, Sen(De) A and Sen U 2001 Distinguishability of bell states *Phys. Rev. Lett.* **87** 277902
[7] Bennett C H, DiVincenzo D P, Fuchs C A, Mor T, Rains E, Shor P W, Smolin J A and Wooters W K 1999 Quantum nonlocality without entanglement *Phys. Rev.* A **59** 1070-91
[8] Nathanson M 2005 Distinguishing bipartite orthogonal states using lOCC: best and worst cases *J. Math. Phys.* **46** 062103
[9] Fan H 2004 Distinguishability and indistinguishability by LOCC *Phys. Rev. Lett.* **92** 177905
[10] Hayashi M, Markham D, Murao M, Owari M and Virmani S 2005 LOCC State discrimination and multipartite entanglement measures *Preprint* quant-ph/0506170
[11] Keyl M, Schlingemann D and Werner R F 2003 Infinitely entangled states *Quantum Inf. Comput.* **3** 281–306
[12] Virmani S, Sacchi M F, Plenio M B and Markham D 2001 Optimal local discriminations of two multipartite pure states *Phys. Lett.* A **288** 62–8
[13] Hiroshima T and Hayashi M 2004 Finding a maximally correlated state—simultaneous Schmidt decomposition of bipartite pure states *Phys. Rev.* A **70** 030302
[14] Gutkin E, Jonckheere E A and Karow M 2004 Convexity of the joint numerical range: topological and differential geometric viewpoints *Linear Algebr. Appl.* **376** 143–71
[15] Hiriart-Urruty J-B and Lemaréchal C 2001 *Fundamentals of Convex Analysis* (Berlin: Springer)
[16] Halmos P 1982 *A Hilbert Space Problem Book* (New York: Springer)
[17] Binding P 1985 Hermitian forms and the fibration of spheres *Proc. Am. Math. Soc.* **94** 581–4
[18] Fan M K H and Tits A L 1987 On the generalized numerical range *Linear Multilinear Algebr.* **21** 313–20