“© 2005 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Catalyst-assisted Probabilistic Entanglement Transformation

Yuan Feng, Runyao Duan and Mingsheng Ying

***Abstract*—We are concerned with catalyst-assisted probabilistic entanglement transformations. A necessary and sufficient condition is presented under which there exist partial catalysts that can increase the maximal transforming probability of a given entanglement transformation. We also design an algorithm which leads to an efficient method for finding the most economical partial catalysts with minimal dimension. The mathematical structure of catalyst-assisted probabilistic transformation is carefully investigated.**

***Index Terms*—Probabilistic entanglement transformation, majorization, catalysis, catalyst-assisted transformation, partial catalyst.**

I. Introduction

QUANTUM entanglement plays an essential role in quantum information processing [1]. Indeed, it is a necessary resource in quantum cryptography [2], quantum superdense coding [3], and quantum teleportation [4], which are striking tasks in quantum information processing. When entanglement is treated as a type of resource, the study of how to quantify and manipulate it becomes crucial (for a survey of quantum information theory, we refer to [5]). A fruitful research direction is to try to discover the laws that must be obeyed when transforming between different forms of entanglement using only local operations on the separate subsystems and classical communication between them. This kind of transformation is usually abbreviated as LOCC. The communication constraints that characterize LOCC are fundamentally and practically important, since many applications of quantum information processing involve spatially separated parties who must manipulate an entanglement state without performing joint operations.

Suppose two spatially separated parties, Alice and Bob, share a bipartite quantum state $|\psi_1\rangle \in \mathcal{C}^n \otimes \mathcal{C}^n$ with Schmidt decomposition

$$|\psi_1\rangle = \sum_{i=1}^{n} \sqrt{\alpha_i}|i_A\rangle|i_B\rangle,$$

where $\alpha_1 \geq \ldots \geq \alpha_n \geq 0$ are the Schmidt coefficients of $|\psi_1\rangle$ and $\sum_i \alpha_i = 1$. $|i_A\rangle$ and $|i_B\rangle$ are orthonormal bases of Alice's and Bob's systems, respectively. Suppose the parties want to transform this initial state into a desired bipartite state $|\psi_2\rangle$ with Schmidt decomposition

$$|\psi_2\rangle = \sum_{i=1}^{n} \sqrt{\beta_i}|i'_A\rangle|i'_B\rangle,$$

where $\beta_1 \geq \ldots \geq \beta_n \geq 0$, and $\sum_i \beta_i = 1$. The orthonormal bases $|i'_A\rangle$ and $|i_A\rangle$ (also $|i'_B\rangle$ and $|i_B\rangle$) are not necessarily the same. Nielsen [6] proved that Alice and Bob can realize this transformation from $|\psi_1\rangle$ to $|\psi_2\rangle$ by LOCC if and only if

$$\sum_{i=1}^{l} \alpha_i \leq \sum_{i=1}^{l} \beta_i \quad \text{for any} \quad 1 \leq l \leq n,$$

with equality holding when $l = n$, or equivalently, by the theory of majorization [7][8], λ_{ψ_1} is majorized by λ_{ψ_2}, written

$$\lambda_{\psi_1} \prec \lambda_{\psi_2},$$

where the probability vectors λ_{ψ_1} and λ_{ψ_2} denote the Schmidt coefficient vectors of $|\psi_1\rangle$ and $|\psi_2\rangle$, respectively.

Nielsen's work establishes a connection between the theory of majorization in linear algebra and entanglement transformation. Furthermore, since the necessary and sufficient condition mentioned above is very easy to check, it is extremely useful in telling whether one bipartite entangled state can be transformed into another by LOCC. Nielsen's theorem directly implies that there exist incomparable states in the sense that any one cannot be transformed into another only using LOCC. To treat the case of transformations between incomparable states, Vidal [9] generalized Nielsen's work by allowing probabilistic transformations. He found that although deterministic transformation cannot be realized between incomparable bipartite pure states, a probabilistic one is always possible (notice that when multipartite states are considered, this statement does not hold [10]). Furthermore, he gave an explicit expression of the maximal probability of transforming one state to another. To be more specific, let $P(|\psi_1\rangle \to |\psi_2\rangle)$ denote the maximal probability of transforming $|\psi_1\rangle$ into $|\psi_2\rangle$ by LOCC. Then

$$P(|\psi_1\rangle \to |\psi_2\rangle) = \min_{1\leq l\leq n} \frac{\sum_{i=l}^{n} \alpha_i}{\sum_{i=l}^{n} \beta_i} = \min_{1\leq l\leq n} \frac{E_l(\lambda_{\psi_1})}{E_l(\lambda_{\psi_2})},$$

where $E_l(\lambda_{\psi_1})$ denotes $\sum_{i=l}^{n} \alpha_i$. In what follows, we extend this notation to any probability vector.

Another interesting phenomenon discovered by Jonathan and Plenio [11] is that sometimes an entangled state can enable otherwise impossible entanglement transformations without

This work was partly supported by the National Foundation of Natural Sciences of China (Grant Nos: 60496321, 60321002, and 60433050) and by the Key grant Project of Chinese Ministry of Education (Grant No: 10402).

The authors are with the State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing, China 100084. E-mails: feng-y@tsinghua.edu.cn (Yuan Feng), dry02@mails.tsinghua.edu.cn (Runyao Duan) and yingmsh@tsinghua.edu.cn (Mingsheng Ying)

being consumed at all. A simple but well known example is $|\psi_1\rangle \nrightarrow |\psi_2\rangle$ but $|\psi_1\rangle \otimes |\phi\rangle \rightarrow |\psi_2\rangle \otimes |\phi\rangle$, where

$$|\psi_1\rangle = \sqrt{0.4}|00\rangle + \sqrt{0.4}|11\rangle + \sqrt{0.1}|22\rangle + \sqrt{0.1}|33\rangle,$$

$$|\psi_2\rangle = \sqrt{0.5}|00\rangle + \sqrt{0.25}|11\rangle + \sqrt{0.25}|22\rangle,$$

and

$$|\phi\rangle = \sqrt{0.6}|44\rangle + \sqrt{0.4}|55\rangle.$$

The role of the state $|\phi\rangle$ in this transformation is analogous to that of a catalyst in a chemical process. The mathematical structure of this phenomenon, so called catalyst-assisted entanglement transformation, was carefully examined by Daftuar and Klimesh [12]. They found that there does not exist an upper bound on the dimension of catalysts that should be considered, in trying to determine which states can be transformed into a given state. Furthermore, they proved that any nonuniform state, which has at least two nonzero Schmidt coefficients nonequal, can serve as a catalyst for some entanglement transformation. On the other hand, Eisert and Wilkens found that catalysis is also helpful in entanglement transformations for bipartite mixed states [13].

In this paper, we examine the power of catalysis in probabilistic entanglement transformations. We have noticed that in [11] Jonathan and Plenio found that in some cases, an appropriately chosen catalyst can increase the maximal transformation probability of incomparable states. The example they presented is as follows. Let

$$|\psi_1\rangle = \sqrt{0.6}|00\rangle + \sqrt{0.2}|11\rangle + \sqrt{0.2}|22\rangle$$

and

$$|\psi_2\rangle = \sqrt{0.5}|00\rangle + \sqrt{0.4}|11\rangle + \sqrt{0.1}|22\rangle.$$

The maximal probability of transforming $|\psi_1\rangle$ into $|\psi_2\rangle$ under LOCC is 80% while when a catalyst

$$|\phi\rangle = \sqrt{0.65}|33\rangle + \sqrt{0.35}|55\rangle$$

is introduced, the probability can be increased to 90.4%. They also showed that enhancement of the transformation probability is not always possible. However, no further results on this topic were given in their paper.

The main aim of our paper is to study the structure of catalysis as applied to probabilistic entanglement transformations. We give a necessary and sufficient condition for the existence of partial catalysts (quantum states which can increase the maximal transforming probability while not being consumed) for a given entanglement transformation. Rather surprisingly, we find that whether or not the probability can be increased depends only on the minimal Schmidt coefficients of the original state and the target state, provided that the maximal transforming probability is less than 1 (the probability cannot of course be increased if equal to 1). To be specific, a probabilistic transformation has partial catalysts if and only if the maximal transforming probability is less than the minimum of 1 and the ratio of the minimal Schmidt coefficient of the original state to that of the target state. Furthermore, we show that if the maximal probability of a transformation can be increased by some catalyst, then there is a sequence of 2×2 dimensional states that increases the maximal probability of the transformation.

For any given entanglement transformation, we present a systematic way to construct partial catalysts. The catalysts are, however, not economical in general in the sense that they do not necessarily have the minimal dimension among all partial catalysts for this transformation. In fact, the problem of constructing systematically the most economical partial catalysts for any given transformation seems to be hard and remains open. We can, however, give a numerical solution to this problem by solving a series of inequalities.

This paper is organized as follows. In Section II, we consider briefly probabilistic entanglement transformations without the aid of catalysis. We first provide a simple connection between probabilistic transformations and deterministic ones, which is helpful in realizing probabilistic transformations since deterministic ones have been well researched. We then examine properties such as monotonicity and continuity of the set of states that can be transformed into a given state by LOCC with a probability not less than a given positive number. Section III is the main body of this paper. We present here a necessary and sufficient condition under which a given probabilistic transformation has partial catalysts. Moreover, the catalysts are systematically constructed. To find the most economical ones, we present first an algorithm to decide whether there exist partial catalysts with a given dimension and find out all suitable ones. Based on this algorithm, the mission of constructing the most economical partial catalysts is achieved by applying the algorithm to state spaces with increasing dimension from 2 (an upper bound on the dimension we should consider can be predetermined because we have constructed a non-economical one). In Section IV, we generalize the result of Daftuar and Klimesh [12] to the set of states that can be transformed into a given state by catalyst-assisted LOCC, with a probability not less than a given positive number. We find that this set shares many properties with the well known set which consists of all states being trumped by the given state (for the latter set, we refer to [12] for details). To be more specific, the generalized set is convex and not closed in general; the dimensions of catalysts that should be considered in trying to determine the states in the set have no upper bound. We further investigate the mathematical structure of this generalized set and find out all the boundary and extreme points. This gives an answer to Nielsen's open problem in the case of deterministic transformations. Finally, a conclusion is drawn and some open problems are presented in Section V.

For simplicity, in what follows we denote a bipartite quantum state by the probability vector of its Schmidt coefficients. This will not cause any confusion because it is well known that the fundamental properties of a bipartite state under LOCC are completely determined by its Schmidt coefficients. Therefore, from now on, we consider only probability vectors instead of bipartite quantum states and always identify a probability vector with the corresponding quantum state. Sometimes we even omit the normalization of positive vectors to be probability ones for the sake of simplicity.

II. Probabilistic Entanglement Transformation

In this section, we discuss the structure of probabilistic entanglement transformations without catalysis. Denote by V^n the set of all n-dimensional probability vectors and let $x, y, \ldots$ range over V^n. Given a positive number $\lambda \leq 1$, let

$$S^\lambda(y) = \{x \in V^n : P(x \to y) \geq \lambda\},$$

which is the set of n-dimensional probability vectors that can be transformed into y by LOCC with the maximal probability not less than λ. When $\lambda = 1$, the set reduces to the well known set $S(y)$ which includes exactly the vectors that can be transformed into y with certainty, or equivalently, that are majorized by y (see [14] for details of $S(y)$).

What we would like to point out first is that there is a simple relationship between probabilistic entanglement transformation and the theory of weak majorization, just like the connection between deterministic transformation and the theory of majorization discovered by Nielsen in [6]. Recall that an n-dimensional positive vector u is called super-majorized [7] by another n-dimensional positive vector v, written $u \prec^\omega v$, if and only if

$$\sum_{i=l}^{n} u_i^\downarrow \geq \sum_{i=l}^{n} v_i^\downarrow$$

for each l in the range 1 through n. Here $u^\downarrow$ denotes the vector obtained by arranging the components of u in nonincreasing order. Notice that the only difference between super-majorization and majorization is the omission of the equality requirement at $l = 1$. It is very easy to check by G. Vidal's formula that $x \in S^\lambda(y)$, or equivalently, $P(x \to y) \geq \lambda$ if and only if the super-majorization relation

$$x \prec^\omega \lambda y$$

holds.

It is well known that there is a close connection between majorization and doubly stochastic matrices [15]. To be specific, for all $x, y \in V^n$, $x \prec y$ if and only if $x = Dy$ for some doubly stochastic matrix D. Here a matrix D is called doubly stochastic if it is positivity preserving and every row and column sums to 1. That is,

$$\forall i, j : D_{ij} \geq 0 \quad \text{and} \quad \forall j : \sum_i D_{ij} = \sum_i D_{ji} = 1.$$

Unfortunately, to the authors' knowledge, super-majorization does not have such a correspondence. In order to make use of the known results about majorization, we must connect probabilistic entanglement transformation and majorization. The following lemma is just for this purpose.

Lemma 1: For $x, y \in V^n$ and $0 \leq \lambda \leq 1$, $x \in S^\lambda(y)$ if and only if $x \in S(y_\lambda)$, where

$$y_\lambda = (1 - \lambda E_2(y), \lambda y_2^\downarrow, \ldots, \lambda y_n^\downarrow). \tag{1}$$

That is, $S^\lambda(y) = S(y_\lambda)$.

Proof: It follows directly from the definitions and we omit the proof here. ∎

In the sequel, we expand the notation y_λ in Eq.(1) to any probability vectors. Another equivalent expression of this lemma is that $x \in S^\lambda(y)$ if and only if $x \prec y_\lambda$. In this paper, we will switch between these two expressions from time to time for convenience. This simple lemma is quite useful because it establishes a relationship between probabilistic transformations and deterministic ones, while the latter have been well researched.

We know that $S(y)$ is just the convex hull of all vectors which may be obtained by permutating the components of y. A direct application of the above lemma is a similar description of the generalized set $S^\lambda(y)$, as the following theorem shows.

Theorem 1: For all $y \in V^n$ and $0 \leq \lambda \leq 1$, $S^\lambda(y)$ is compact and convex. Furthermore, $S^\lambda(y)$ is the convex hull of the vectors in the following set

$$\{Py_\lambda : \ P \ \text{ is any } n \text{ dimesional permutation}\}.$$

Proof: Direct from Lemma 1 and the known structure of $S(y_\lambda)$. ∎

The next theorem shows that the set $S^\lambda(y)$, as a function of λ, is monotonic, and the intersection of all $S^\lambda(y)$, $0 < \lambda < 1$, gives rise to $S(y)$.

Theorem 2: Suppose $y \in V^n$ and $0 \leq \lambda_1 \leq \lambda_2 \leq 1$. Then $S^{\lambda_2}(y) \subseteq S^{\lambda_1}(y)$. Furthermore, we have that

$$S(y) = \bigcap_{0<\lambda<1} S^\lambda(y). \tag{2}$$

Proof: The monotonicity of $S^\lambda(y)$ is obvious from the definition. So

$$S(y) \subseteq \bigcap_{0<\lambda<1} S^\lambda(y)$$

holds. To show

$$S(y) \supseteq \bigcap_{0<\lambda<1} S^\lambda(y),$$

suppose $x \in \bigcap_{0<\lambda<1} S^\lambda(y)$. It follows that $x \prec y_\lambda$, or equivalently,

$$E_l(x) \geq E_l(y_\lambda) = \lambda E_l(y)$$

for all $1 < l \leq n$ and $0 < \lambda < 1$. When λ tends to 1, we have $E_l(x) \geq E_l(y)$ for all $1 < l \leq n$. Thus $x \in S(y)$. This completes the proof. ∎

From the monotonicity of $S^\lambda(y)$ as a function of λ, we can define the notions of limit as follows. We call $S^\lambda(y)$ left continuous at λ if for all nondecreasing sequences $\{\lambda_i : i = 1, 2, \ldots\}$, $\lim_{i\to\infty} \lambda_i = \lambda$ implies that

$$\bigcap_{i=1}^{\infty} S^{\lambda_i}(y) = S^\lambda(y).$$

While $S^\lambda(y)$ is said to be right continuous at λ if for all nonincreasing sequences $\{\lambda_i : i = 1, 2, \ldots\}$, $\lim_{i\to\infty} \lambda_i = \lambda$ implies that

$$\bigcup_{i=1}^{\infty} S^{\lambda_i}(y) = S^\lambda(y).$$

Furthermore, $S^\lambda(y)$ is continuous at λ if it is both left continuous and right continuous. Having these notions, we are able to present the following theorem.

Theorem 3: For all $y \in V^n$, $S^\lambda(y)$ is continuous at any λ when $0 < \lambda < 1$. It is also right continuous at 0 and left continuous at 1.

Proof: Easy to check from the definitions. ∎

We have examined thoroughly probabilistic transformations without catalysis; in the following sections, we will consider catalyst-assisted ones. At the end of this section, we introduce some lemmas that are useful for later discussion.

Lemma 2: Given $y \in V^n$, the function $P(x \to y)$ is concave in x.

Proof: For all $x, x' \in V^n$ and $t \in [0,1]$, we have

$$\begin{aligned} P(tx+(1-t)x' \to y) &= \min_l \frac{E_l(tx+(1-t)x')}{E_l(y)} \\ &\geq \min_l \frac{E_l(tx)+E_l((1-t)x')}{E_l(y)} \\ &\geq tP(x \to y)+(1-t)P(x' \to y). \end{aligned}$$

So $P(x \to y)$ is concave in x. ∎

The next two lemmas consider the properties of the maximal transformation probability $P(x \to y)$ under the operations of direct summation and tensor product on its parameters x and y.

Lemma 3: For all $x, y \in V^n$ and $x', y' \in V^m$,

$$P(x \oplus x' \to y \oplus y') \geq \min\{P(x \to y), P(x' \to y')\},$$

where $\oplus$ means direct summation. In particular, $P(x \oplus c \to y \oplus c) \geq P(x \to y)$ for all c.

Proof: By Vidal's formula for the probability of entanglement transformation, there exists an index l such that $1 \leq l \leq n+m$ and

$$P(x \oplus x' \to y \oplus y') = \frac{E_l(x \oplus x')}{E_l(y \oplus y')}.$$

We assume that $E_l(x \oplus x') = E_{l_x}(x) + E_{l_{x'}}(x')$ for some $l_x \leq n$ and $l_{x'} \leq m$. Notice that $E_l(y \oplus y') \leq E_{l_x}(y) + E_{l_{x'}}(y')$ by definition. It follows that

$$\begin{aligned} P(x \oplus x' \to y \oplus y') &\geq \frac{E_{l_x}(x)+E_{l_{x'}}(x')}{E_{l_x}(y)+E_{l_{x'}}(y')} \\ &\geq \min\{\frac{E_{l_x}(x)}{E_{l_x}(y)}, \frac{E_{l_{x'}}(x')}{E_{l_{x'}}(y')}\}, \end{aligned}$$

where the second inequality follows from the following fact

$$\frac{a+b}{c+d} \geq \frac{b}{d} \Leftrightarrow \frac{a}{c} \geq \frac{b}{d} \quad \text{for any } a,b,c,d \geq 0.$$

Thus $P(x \oplus x' \to y \oplus y') \geq \min\{P(x \to y), P(x' \to y')\}$. ∎

Lemma 4: For all $x, y \in V^n$ and $x', y' \in V^m$,

$$P(x \otimes x' \to y \otimes y') \geq P(x \to y)P(x' \to y'),$$

where $\otimes$ means tensor product. In particular, $P(x \otimes c \to y \otimes c) \geq P(x \to y)$ for all c.

Proof: This result is obvious from the physical meaning of $P(x \to y)$ since the way that separately transforms x into y and x' into y' gives an implementation of transforming $x \otimes x'$ to $y \otimes y'$. The probability of success is the multiplication of the probabilities of those two transformations. That means $P(x \otimes x' \to y \otimes y') \geq P(x \to y)P(x' \to y')$.

We can, however, give a simple pure mathematical proof as follows. Without loss of generality, we assume that the components of x, x', y, and y' are nonincreasingly arranged, respectively. For an arbitrarily fixed integer l satisfying $1 \leq l \leq nm$, let r_i be the smallest index of the components of x' in summands of $E_l(x \otimes x')$ that have the form $x_i x'_j$, where $1 \leq i \leq n$. That is,

$$r_i = \min\{j : \ x_i x'_j \geq (x \otimes x')^\downarrow_l\}. \tag{3}$$

In case of repeated values of components of $x \otimes x'$, we regard the terms with smaller i to be included in the sum first. If the set in the right hand side of Eq.(3) is empty for some i (that is, any term having the form $x_i x'_j$, $1 \leq j \leq m$, does not occur), then let $r_i = m+1$. With these notations, we can arrange the summands of $E_l(x \otimes x')$ as

$$E_l(x \otimes x') = \sum_{i=1}^n x_i \sum_{j=r_i}^m x'_j.$$

By the definition of $P(x' \to y')$, we have $\sum_{j=r_i}^m x'_j \geq P(x' \to y') \sum_{j=r_i}^m y'_j$ for all r_i. Thus

$$E_l(x \otimes x') \geq P(x' \to y') \sum_{i=1}^n x_i \sum_{j=r_i}^m y'_i.$$

Now we rearrange the summands of $\sum_{i=1}^n x_i \sum_{j=r_i}^m y'_i$ such that

$$\sum_{i=1}^n x_i \sum_{j=r_i}^m y'_i = \sum_{j=1}^m y'_i \sum_{i=t_j}^n x_i$$

for some $1 \leq t_1, \ldots, t_m \leq n+1$. By the definition of $P(x \to y)$, we have

$$\sum_{i=t_j}^n x_i \geq P(x \to y) \sum_{i=t_j}^n y_i$$

for all t_j. Thus

$$\begin{aligned} E_l(x \otimes x') &\geq P(x' \to y')P(x \to y) \sum_{j=1}^m y'_i \sum_{i=t_j}^n y_i \\ &\geq P(x' \to y')P(x \to y)E_l(y \otimes y'), \end{aligned}$$

and $P(x \otimes x' \to y \otimes y') \geq P(x \to y)P(x' \to y')$ from the arbitrariness of l. ∎

III. Catalyst-assisted Probabilistic Transformation

The aim of this section is to consider the case of entanglement transformations with the aid of catalysis. First, we present a necessary and sufficient condition for a given probability vector to serve as a partial catalyst for a certain probabilistic transformation.

Without loss of generality, we concentrate on catalysts with nonzero components, since for any probability vector c, c and $c \oplus 0$ have the same catalysis power in the sense that

in any situation, if one serves as a partial catalyst for some transformation, so does the other for the same transformation.

Theorem 4: Suppose x and y are two nonincreasingly arranged n-dimensional probability vectors, and $P(x \to y) < \min\{x_n/y_n, 1\}$. Let

$$L = \{l : 1 < l < n \ \text{ and } \ P(x \to y) = \frac{E_l(x)}{E_l(y)}\}.$$

Then a nonincreasingly arranged k-dimensional probability vector c serves as a partial catalyst for the transformation from x to y, that is,

$$P(x \otimes c \to y \otimes c) > P(x \to y),$$

if and only if for all $r_1, r_2, \ldots, r_k \in L \cup \{n+1\}$ satisfying $r_1 \geq \ldots \geq r_k \neq n+1$, there exist i and j, $1 \leq j < i \leq k$, such that

$$\frac{c_i}{c_j} < \frac{y_{r_j}}{y_{r_i-1}} \qquad \text{or} \qquad \frac{c_i}{c_j} > \frac{y_{r_j-1}}{y_{r_i}}. \tag{4}$$

Here, in order to avoid a too complicated statement, we ignore some extreme cases of Eq.(4); the condition (4) should be understood in the following way: whenever one of the two components of the disjunction in Eq.(4) contains the meaningless term y_{n+1}, it is considered to be violated automatically, and the other component is then required.

Proof: We will prove the theorem by showing that c cannot serve as a partial catalyst for transforming x into y, that is,

$$P(x \otimes c \to y \otimes c) = P(x \to y),$$

if and only if there exist $r_1, r_2, \ldots, r_k \in L \cup \{n+1\}$ satisfying $r_1 \geq \ldots \geq r_k \neq n+1$, such that for all $1 \leq j < i \leq k$,

$$\frac{y_{r_j}}{y_{r_i-1}} \leq \frac{c_i}{c_j} \leq \frac{y_{r_j-1}}{y_{r_i}}, \tag{5}$$

where any constraint containing the meaningless term y_{n+1} is considered to be satisfied automatically.

Notice that for all l, $1 < l \leq nk$, we can arrange the summands of $E_l(x \otimes c)$ such that

$$E_l(x \otimes c) = \sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} x_i,$$

where $1 \leq r_j \leq n+1$. The case $r_j = n+1$ means that any term having the form $c_j x_i$, $1 \leq i \leq n$, does not occur. Without loss of generality, we regard terms with smaller j to be included in the sum first in case of repeated values of components of $x \otimes c$. This assumption guarantees that $r_1 \geq \ldots \geq r_k$. Furthermore, we exclude the possibility of $r_1 = \ldots = r_k = n+1$ from $\sum_j r_j = l$ and $1 < l \leq nk$.

From the definition of $E_l(y \otimes c)$ and $P(x \to y)$, the following inequalities are easy to check:

$$\begin{aligned}
\frac{E_l(x \otimes c)}{E_l(y \otimes c)} &\geq \frac{\sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} x_i}{\sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} y_i} \\
&\geq \frac{P(x \to y)(\sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} y_i)}{\sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} y_i} \\
&= P(x \to y).
\end{aligned}$$

The first equality holds if and only if $E_l(y \otimes c) = \sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} y_i$, while the second equality holds if and only if every r_j is in $L \cup \{n+1\}$. Consequently, we see that $P(x \otimes c \to y \otimes c) = P(x \to y)$ if and only if there exist $r_1, r_2, \ldots, r_k \in L \cup \{n+1\}$ such that

$$E_l(x \otimes c) = \sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} x_i \tag{6}$$

and

$$E_l(y \otimes c) = \sum_{j=1}^{k} c_j \sum_{i=r_j}^{n} y_i \tag{7}$$

for some $1 < l \leq nk$.

In what follows, we derive the conditions presented in Eq.(5) from Eqs.(6) and (7). In fact, Eq.(7) means that $\max_{1 \leq i \leq k}\{y_{r_i} c_i\} \leq \min_{1 \leq i \leq k}\{y_{r_i-1} c_i\}$, or equivalently,

$$\frac{y_{r_j}}{y_{r_i-1}} \leq \frac{c_i}{c_j} \leq \frac{y_{r_j-1}}{y_{r_i}} \tag{8}$$

for all $i, j = 1, 2, \ldots, k$ and $i > j$. The special case of $r_i = n+1$ or $r_j = n+1$ can be included in Eq.(8) by simply assuming that the constraints in Eq.(8) containing the meaningless term y_{n+1} are automatically satisfied. Analogously, we can show that Eq.(6) is equivalent to

$$\frac{x_{r_j}}{x_{r_i-1}} \leq \frac{c_i}{c_j} \leq \frac{x_{r_j-1}}{x_{r_i}} \tag{9}$$

for all $i > j$. Notice that for all $r_i, r_j \in L$,

$$\frac{x_{r_j}}{y_{r_j}} \leq P(x \to y) \leq \frac{x_{r_i-1}}{y_{r_i-1}}.$$

It follows that the constraints in Eq.(9) can be derived from those in Eq.(8). That completes our proof. ■

Intuitively, if we decompose $x \otimes c$ and $y \otimes c$ as

$$x \otimes c = c_1 x \oplus \ldots \oplus c_k x$$

and

$$y \otimes c = c_1 y \oplus \ldots \oplus c_k y,$$

respectively, then when the conditions in Eq.(8) are satisfied for some $r_1, r_2, \ldots, r_k \in L \cup \{n+1\}$, we have $c_i x_{r_i} \leq c_j x_{r_j-1}$ and $c_j x_{r_j} \leq c_i x_{r_i-1}$ for all $1 \leq i, j \leq n$. So the

smallest $k(n+1)-\sum_{i=1}^{k} r_i$ components of $x\otimes c$ are exactly the components of the form $x_i c_j$, where $1 \le j \le k$ and $r_j \le i \le n$. A similar argument holds for $y\otimes c$. It follows that when we take $l=\sum_{i=1}^{k} r_i - k + 1$, then $E_l(x\otimes c) = P(x\to y)E_l(y\otimes c)$ and thus $P(x\otimes c\to y\otimes c)=P(x\to y)$.

To our surprise, the constraints presented in Eq.(4) for the probability vector c to serve as a partial catalyst for transforming x into y are almost irrelevant to x. The only effect of x is to determine the index set L.

Corollary 1: Let x, y, c, and L be as in the above theorem. If

$$P(x\otimes c\to y\otimes c) > P(x\to y),$$

then

$$\frac{c_k}{c_{k-1}} > \max\{\frac{y_n}{y_l} : l\in L\}. \tag{10}$$

Proof: For any $l\in L$, take $r_1=\ldots=r_{k-1}=n+1$ and $r_k=l$. Noticing that the constraints having the term y_{n+1} are violated automatically, we can reduce Eq.(4) to the condition that

$$\frac{c_k}{c_j} > \frac{y_n}{y_l} \tag{11}$$

for some $j<k$. So

$$\frac{c_k}{c_{k-1}} \ge \frac{c_k}{c_j} > \frac{y_n}{y_l},$$

and the corollary holds from the arbitrariness of l. ■

Intuitively, Corollary 1 shows that the difference between the smallest two components of a partial catalyst cannot be too large. The following corollary, on the other hand, shows that the difference between the smallest and the largest components of a partial catalyst cannot be too small. Recall that a uniform state is a state which has equal nonzero Schmidt coefficients.

Corollary 2: Let x, y, c, and L be as in the above theorem. If

$$P(x\otimes c\to y\otimes c) > P(x\to y),$$

then

$$\frac{c_k}{c_1} < \min\{\frac{y_l}{y_{l-1}} : l\in L\}. \tag{12}$$

In particular, any uniform state cannot serve as a partial catalyst for any probabilistic transformation.

Proof: For any $l\in L$, let $r_1=\ldots=r_k=l$. Then the conditions in Eq.(4) become

$$\frac{c_i}{c_j} < \frac{y_l}{y_{l-1}} \tag{13}$$

for some $i>j$. Noticing that $\frac{c_k}{c_1}\le\frac{c_i}{c_j}$ for all $1\le j<i\le k$ and l is taken arbitrarily from L, we complete the proof. ■

A special and perhaps more interesting case of Theorem 4 is when L has only one element, that is, $L=\{l\}$ for some $1<l<n$. In this case, we have the following corollary.

Corollary 3: Let x, y, and L be as in Theorem 4. If $L=\{l\}$ for some $1<l<n$, and

$$\frac{y_n}{y_l} < \frac{y_l}{y_{l-1}}, \tag{14}$$

then any nonincreasingly arranged k-dimensional probability vector c with

$$\frac{y_n}{y_l} < \frac{c_k}{c_t} < \frac{y_l}{y_{l-1}} \tag{15}$$

for some $t<k$ serves as a partial catalyst for transforming x into y.

Proof: Since $L=\{l\}$ contains one element, any choice of $r_1, r_2, \ldots, r_k \in L\cup\{n+1\}$ satisfying $r_1\ge\ldots\ge r_k\ne n+1$ has the form

$$r_1=\ldots=r_h=n+1,\quad r_{h+1}=\ldots=r_k=l$$

for some $0\le h<k$. Take $i=k$ and $j=t$. In what follows we show that under this choice of i, j, Eq.(15) implies Eq.(4) in Theorem 4 and thus c is a partial catalyst for transforming x into y.

In fact, if $t\le h$, then $r_i=l$ and $r_j=n+1$. Thus

$$\frac{y_{r_j-1}}{y_{r_i}} = \frac{y_n}{y_l} < \frac{c_k}{c_t} = \frac{c_i}{c_j}.$$

Similarly, if $t>h$ then $r_i=r_j=l$ and

$$\frac{y_{r_j}}{y_{r_i-1}} = \frac{y_l}{y_{l-1}} > \frac{c_k}{c_t} = \frac{c_i}{c_j}.$$

Thus the conditions in Eq.(4) holds and that completes our proof. ■

When 2-dimensional catalysts are considered, Eq.(15) reduces to

$$\frac{y_n}{y_l} < \frac{c_2}{c_1} < \frac{y_l}{y_{l-1}}. \tag{16}$$

Furthermore, it is easy to check that Eq.(16) is indeed a necessary and sufficient condition for a 2-dimensional probability vetor c to be a partial catalyst. Thus Eq.(14) also becomes a necessary and sufficient one to guarantee the transformation from x to y has 2-dimensional partial catalysts.

From Theorem 4 we can derive a necessary and sufficient condition for when a given probabilistic transformation has partial catalysts.

Theorem 5: Suppose $x, y\in V^n$ and the components of x and y are nonincreasingly arranged, respectively. Then the probabilistic transformation from x to y has partial catalysts if and only if

$$P(x\to y) < \min\{x_n/y_n, 1\}.$$

Proof: The 'only if' part is easy and we omit the details here. The proof of 'if' part is as follows.

We abbreviate $P(x\to y)$ to P for simplicity in this proof. Let l_{min} and l_{max} be the minimal element and maximal element in L, respectively. Then $y_n < y_{l_{max}}$ since otherwise

$$\frac{E_{l_{max}}(x)}{E_{l_{max}}(y)} \ge \frac{(n-l_{max}+1)x_n}{(n-l_{max}+1)y_n} = \frac{x_n}{y_n} > P,$$

which contradicts the assumption that $l_{max}\in L$. Now let α be a real number such that $y_n/y_{l_{max}} < \alpha < 1$ and k a positive integer such that $\alpha^{k-1} < y_{l_{max}}/y_{l_{min}-1}$. In what follows, we prove that

$$c=(1,\alpha,\alpha^2,\ldots,\alpha^{k-1})$$

will serve as a partial catalyst for the transformation from x to y, that is,

$$P(x \otimes c \to y \otimes c) > P(x \to y).$$

Here we omit the normalization of c for simplicity.

For all $r_1, r_2, \ldots, r_k \in L \cup \{n+1\}$ satisfying $r_1 \geq \ldots \geq r_k \neq n+1$. There are two cases to consider.

Case 1. $r_1 \in L$. In this case, let $i = k$ and $j = 1$. Then

$$\frac{c_i}{c_j} = \alpha^{k-1} < \frac{y_{l_{max}}}{y_{l_{min}-1}} \leq \frac{y_{r_1}}{y_{r_k-1}} = \frac{y_{r_j}}{y_{r_i-1}}.$$

Case 2. $r_1 = n+1$. In this case, denote by m, $1 \leq m \leq k-1$, the (unique) integer such that $r_m = n+1$ but $r_{m+1} \in L$. Now let $i = m+1$ and $j = m$. Then

$$\frac{c_i}{c_j} = \alpha > \frac{y_n}{y_{l_{max}}} \geq \frac{y_n}{y_{r_{m+1}}} = \frac{y_{r_j-1}}{y_{r_i}}.$$

In either case, the constraints in Eq.(4) are satisfied. So from Theorem 4 we know that the present theorem holds. ■

To illustrate the utility of Theorem 5, let us see an example from [11] (it has been presented in Introduction). Let $x = (0.6, 0.2, 0.2)$ and $y = (0.5, 0.4, 0.1)$. We show how to construct a partial catalyst by the above theorem. It is easy to check that $L = \{2\}$ and

$$\frac{y_3}{y_2} = 0.25 < 0.8 = \frac{y_2}{y_1}.$$

So we can take $k = 2$ and any two dimensional nonnormalized vector $c = (1, \alpha)$ for $0.25 < \alpha < 0.8$ serves as a partial catalyst for transforming x into y. Furthermore, from the remark behind Corollary 3, $\{c = (1, \alpha) : 0.25 < \alpha < 0.8\}$ is exactly the set of all two dimensional partial catalysts for this transformation.

Suppose now x is just as above while $y = (0.5, 0.3, 0.2)$. Then $L = \{2\}$ but $y_3/y_2 = 2/3 > y_2/y_1 = 0.6$. So any two-dimensional state cannot serve as a partial catalyst for the probabilistic transformation from x to y. We can, however, construct a higher dimensional partial catalyst from Theorem 5 as follows. First, take a real number $\alpha > y_3/y_2 = 2/3$. In order not to make k too large, we should take α as small as possible. For example, $\alpha = 0.67$. Then from the constraint $\alpha^{k-1} < y_2/y_1 = 0.6$ in the theorem, we have $k \geq 3$. Thus the state $c = (1, \alpha, \alpha^2)$ can serve as a partial catalyst for transforming x into y.

It is worth noting that the catalyst c presented in the proof of the above theorem can be replaced by a sequence of 2-dimensional vectors. To see this, notice that the only constraint on the dimension k of the catalyst c is $\alpha^{k-1} < y_{l_{max}}/y_{l_{min}-1}$, that is, for all sufficiently large k,

$$c = (1, \alpha, \alpha^2, \ldots, \alpha^{k-1})$$

is an appropriate partial catalyst which can increase the maximal probability of transforming x into y. In particular, take $k = 2^{m+1} - 1$ for some positive integer m. From the simple fact that

$$(1, \alpha, \alpha^2, \ldots, \alpha^{2^{m+1}-1}) = (1, \alpha) \otimes (1, \alpha^2) \otimes \ldots \otimes (1, \alpha^{2^m}),$$

the effect of the catalyst in the left hand side can be implemented by the sequence of 2-dimensional catalysts listed in the right hand side. From this observation, a potential 'catalyst bank' need only prepare sufficiently many 2-dimensional catalysts in order to help probabilistic transformation.

We state the arguments above as the following theorem.

Theorem 6: The set V^2 constitutes a complete set of partial catalysts for all probabilistic entanglement transformations. That is, for all positive n and $x, y \in V^n$, if $P(x \otimes c \to y \otimes c) > P(x \to y)$ for some c, then there exists a sequence of probability vectors $c_1, \ldots, c_m \in V^2$ such that

$$P(x \otimes c_1 \otimes \ldots \otimes c_m \to y \otimes c_1 \otimes \ldots \otimes c_m) > P(x \to y).$$

We have presented a necessary and sufficient condition under which a given entanglement transformation has partial catalysts. Furthermore, the proof process constructs real catalysts. The constructed catalysts are, however, not very economical in the sense that they are usually not with the minimal dimension among all the probability vectors which can serve as partial catalysts for this transformation. In what follows, we show how to construct economical ones.

First, from Theorem 4, we can design an efficient algorithm to decide whether a probabilistic transformation has partial catalysts with a given dimension.

Theorem 7: Suppose x, y are two n-dimensional probability vectors and $P(x \to y) < \min\{x_n^{\downarrow}/y_n^{\downarrow}, 1\}$. Let k be a given positive integer. Then the problem of whether there exists a k-dimensional partial catalyst c for transforming x into y can be decided in polynomial time about n.

Proof: Without loss of generality, we assume that the components of x, y, and c are respectively arranged non-increasingly. Notice that from the proof of Theorem 4 (see Eq.(8)), the necessary and sufficient condition under which c can increase the maximal probability of transforming x into y can be reexpressed as, for all $r_1, r_2, \ldots, r_k \in L \cup \{n+1\}$ satisfying $r_1 \geq \ldots \geq r_k \neq n+1$,

$$\max_{1 \leq i \leq k} \{y_{r_i} c_i\} > \min_{1 \leq i \leq k} \{y_{r_i - 1} c_i\}.$$

This condition leads to the following algorithm to decide whether a k-dimensional partial catalyst exists for transforming x into y:

1. Calculate $P(x \to y)$ and determine the set $L = \{l : 1 < l < n \text{ and } P(x \to y) = \frac{E_l(x)}{E_l(y)}\}$.

2. For all k positive integers $r_1, r_2, \ldots, r_k$ chosen from $L \cup \{n+1\}$, if $r_1 \geq \ldots \geq r_k \neq n+1$, then solve the following inequality about $c_1, \ldots, c_k$:

$$\max_{1 \leq i \leq k} \{y_{r_i} c_i\} > \min_{1 \leq i \leq k} \{y_{r_i - 1} c_i\}. \tag{17}$$

Then there exists a k-dimensional partial catalyst c if and only if the intersection of the solution areas of Eqs. in (17) is not empty when $r_1, r_2, \ldots, r_k$ range over $L \cup \{n+1\}$ but satisfy $r_1 \geq \ldots \geq r_k \neq n+1$. Notice that the solution area of Eq.(17) for a given sequence $r_1, r_2, \ldots, r_k$ is just the union of those of the following k^2 inequalities:

$$y_{r_i} c_i > y_{r_j - 1} c_j,$$

which can be solved in polynomial time of k. Furthermore, the number of choices of $r_1, r_2, \ldots, r_k$ is less than $(\#L + 1)^k$, where $\#L$ denotes the number of elements in L and obviously, $\#L < n - 1$. So the algorithm presented above runs in $O(k)(\#L+1)^k = O(n^k)$ time, which is a polynomial of n when k is treated as a constant. ■

Notice that in [16], Sun *et al* have presented a polynomial time algorithm to decide whether a given entanglement transformation has k-dimensional catalysts, that is, k-dimensional partial catalysts which can increase the maximal transforming probability to 1. So a little modification of Sun's algorithm can also be used to determine whether or not a k-dimensional partial catalyst exists. What we would like to point out is that the complexity of Sun's algorithm is $O(n^{2k})$ while our algorithm presented in Theorem 7 is $O((\#L)^k)$. Although they are both exponential of k and in the worst case $\#L = n - 2$, in practice our algorithm is more efficient since $\#L$ is generally much less than n.

Theorem 7 and Theorem 5 together give a method for finding out the most economical catalysts for a given entanglement transformation as follows. First, we use Theorem 5 to decide whether partial catalysts exist for this transformation and an upper bound m on the dimensions of the most economical ones can also be derived. Second, for $k = 2, 3, \ldots, m$ we use the algorithm presented in Theorem 7 to decide whether there exist k-dimensional partial catalyst. Moreover, from the algorithm presented in Theorem 7, the most economical partial catalysts can be constructed explicitly.

IV. Structure of Catalyst-assisted Probabilistic Transformation

In this section, we investigate the mathematical structure of catalyst-assisted probabilistic entanglement transformation. Given a probability vector $y \in V^n$ and $0 \leq \lambda \leq 1$, denote by $T^\lambda(y)$ the set of probability vectors which, with the aid of some catalyst, can be transformed into y with a probability not less than λ, that is

$$T^\lambda(y) = \{x \in V^n : P(x \otimes c \to y \otimes c) \geq \lambda \text{ for some probability vector } c\}.$$

The special case of $\lambda = 1$ corresponds to $T(y)$, which is just the set of all probability vectors that can be transformed deterministically into y by catalyst-assisted LOCC (for the definition, we refer to [12] or [17]). It is easy to check that the set $T^\lambda(y)$, as a function of λ, is monotonic.

Recall that $S^\lambda(y)$ is just equal to $S(y_\lambda)$ from Lemma 1. One may wonder if $T^\lambda(y) = T(y_\lambda)$, or if there exists a simple connection between catalyst-assisted probabilistic transformation and catalyst-assisted deterministic one. If so, then all the known properties of $T(y_\lambda)$ can be used to give simple proofs of those of $T^\lambda(y)$. In fact, we can prove the following.

Lemma 5: For $x, y \in V^n$, if $x \in T(y_\lambda)$, then $x \in T^\lambda(y)$. That is, $T(y_\lambda) \subseteq T^\lambda(y)$.

Proof: Suppose $x \in T(y_\lambda)$. By definition, there exists a probability vector c such that $x \otimes c \prec y_\lambda \otimes c$. Noticing that

$$y_\lambda \otimes c = (1 - \lambda E_2(y))c \oplus \lambda y_2 c \oplus \ldots \oplus \lambda y_n c \prec (y \otimes c)_\lambda,$$

we have $x \otimes c \prec (y \otimes c)_\lambda$, which implies that $x \in T^\lambda(y)$. ■

But unfortunately, $T(y_\lambda) \neq T^\lambda(y)$. Moreover, we can show informally that for all $z \in V^n$, $T(z) \neq T^\lambda(y)$ as follows. From Corollary 5 in [12], the boundary points of $T(z)$ is the set $\{x \in T(z) : x_1^\downarrow = z_1^\downarrow \text{ or } x_n^\downarrow = z_n^\downarrow\}$, while from Theorem 9 in this section, the boundary points of $T^\lambda(y)$ is the set $\{x \in T^\lambda(y) : x_n^\downarrow = \lambda y_n^\downarrow\}$. It is obvious that these two sets cannot be equal. So $T(z) \neq T^\lambda(y)$. Since $T^0(y) = V^n$ and $T^1(y) = T(y)$, we assume from now on that $0 < \lambda < 1$. The next two lemmas are useful for latter discussion.

Lemma 6: Let $x, y \in V^n$ and $0 < \lambda < 1$. Suppose x and y can be decomposed into two parts respectively as

$$x = x' \oplus \lambda z \text{ and } y = y' \oplus z$$

for some z such that $z_1^\downarrow < y_1^\downarrow$. Let μ denote the sum of the components of z. Then $x \in S^\lambda(y)$ if and only if

$$\frac{x'}{1 - \mu\lambda} \in S^{\lambda'}(\frac{y'}{1 - \mu}),$$

where

$$\lambda' = \frac{\lambda - \mu\lambda}{1 - \mu\lambda}. \tag{18}$$

Proof: Without loss of generality, we can assume that

$$y' = (y_{i_1}, y_{i_2}, \ldots, y_{i_m}),$$

and the components of y' have been arranged nonincreasingly. If $x \in S^\lambda(y)$, then $x \prec y_\lambda$ from Lemma 1. Furthermore, we have $y_\lambda = y'' \oplus \lambda z$ and then

$$\frac{x'}{1 - \mu\lambda} \prec \frac{y''}{1 - \mu\lambda},$$

where

$$y'' = (1 - \lambda E_2(y), \lambda y_{i_2}, \ldots, \lambda y_{i_m}).$$

A simple calculation shows that

$$\frac{y''}{1 - \mu\lambda} = (\frac{y'}{1 - \mu})_{\lambda'},$$

so the 'only if' part of the lemma holds.

The 'if' part can be proved by simply retracing the arguments above. ■

Lemma 7: Let x, y, λ, μ, and λ' be as in Lemma 6. Then $x \in T^\lambda(y)$ if and only if

$$\frac{x'}{1 - \mu\lambda} \in T^{\lambda'}(\frac{y'}{1 - \mu}).$$

Proof: Notice that $x \in T^\lambda(y)$ if and only if there exists a probability vector c such that $x \otimes c \in S^\lambda(y \otimes c)$. The current lemma is just a simple application of Lemma 6. ■

The following theorem shows that some properties of $T(y)$ such as convexity and containing corresponding no-catalysis case as a subset are also shared by $T^\lambda(y)$.

Theorem 8: For all $y \in V^n$ and $0 < \lambda < 1$,

1) $T^\lambda(y)$ is convex;

2) $S^\lambda(y) \subseteq T^\lambda(y)$;

3) suppose $x \in T^\lambda(y)$ and the components of x and y are nonincreasingly arranged respectively. If $x \neq y_\lambda$, then $x_d/y_d > \lambda$, where d is the maximal index of the components such that $x_d/y_d \neq \lambda$.

Proof: 1) Suppose $x, x' \in T^\lambda(y)$. By definition, there exist probability vectors c and c' such that $P(x\otimes c \to y\otimes c) \geq \lambda$ and $P(x' \otimes c' \to y \otimes c') \geq \lambda$. For all t, $0 \leq t \leq 1$, we have

$$P((tx + (1-t)x') \otimes \widetilde{c} \to y \otimes \widetilde{c})$$

$$\geq tP(x \otimes \widetilde{c} \to y \otimes \widetilde{c}) + (1-t)P(x' \otimes \widetilde{c} \to y \otimes \widetilde{c})$$

$$\geq tP(x \otimes c \to y \otimes c) + (1-t)P(x' \otimes c' \to y \otimes c') \geq \lambda$$

from Lemma 2 and Lemma 4, where $\widetilde{c} = c \otimes c'$.

2) It is obvious from Lemma 4.

3) Suppose $x \in T^\lambda(y)$. Decompose x as $x = x' \oplus x''$ where

$$x' = (x_1, x_2, \ldots, x_d) \quad \text{and} \quad x'' = (x_{d+1}, \ldots, x_n).$$

The vector y is decomposed analogously. From the definition of d, we have $x'' = \lambda y''$. Applying Lemma 7, we have

$$\frac{x'}{1-\mu\lambda} \in T^{\lambda'}(\frac{y'}{1-\mu}), \tag{19}$$

where μ is the sum of the components of y'' and λ' is defined by Eq.(18). We can then deduce that $x_d/y_d > \lambda$ since otherwise (notice that $x_d/y_d \neq \lambda$ by assumption)

$$P(\frac{x'}{1-\mu\lambda} \otimes c \to \frac{y'}{1-\mu} \otimes c) \quad \leq \frac{(1-\mu)x'_d c_k}{(1-\mu\lambda)y'_d c_k}$$

$$= \frac{(1-\mu)x_d}{(1-\mu\lambda)y_d} < \lambda'$$

for any probability vector c, where k is the dimension of c. This contradicts Eq.(19). ∎

Notice that in 3), no constraints on the largest components of x and y are needed for $x \in T^\lambda(y)$, in contrast to the necessary condition that $x_1 \leq y_1$ of $x \in T(y)$. This is due to the asymmetry of roles of the largest and the smallest components in determining the maximal transforming probability.

Lemma 8: Let $y \in V^n$ and $0 < \lambda < 1$. For all $x \in S^\lambda(y)$, if $P(x \to y) < x_n^\downarrow / y_n^\downarrow$, then x is in the interior of $T^\lambda(y)$.

Proof: Suppose $x \in S^\lambda(y)$ and $P(x \to y) < x_n^\downarrow/y_n^\downarrow$. If $P(x \to y) = 1$, then x is in the interior of $S^\lambda(y)$ for $\lambda < 1$. Thus x is also an interior point of $T^\lambda(y)$ since $S^\lambda(y) \subseteq T^\lambda(y)$. When $P(x \to y) < 1$, by Theorem 5, there exists a partial catalyst c such that $P(x\otimes c \to y \otimes c) > P(x \to y) \geq \lambda$. Thus x is an interior point of $T^\lambda(y)$. ∎

Theorem 9: Let $y \in V^n$, $0 < \lambda < 1$ and $x \in T^\lambda(y)$. Then x is on the boundary of $T^\lambda(y)$ if and only if $x_n^\downarrow/y_n^\downarrow = \lambda$.

Proof: Notice first that $x \in T^\lambda(y)$ implies $x_n^\downarrow/y_n^\downarrow \geq \lambda$ from Theorem 8. Suppose now that $x_n^\downarrow/y_n^\downarrow = \lambda$. For all $\epsilon > 0$, consider the probability vector

$$x' = (x_1^\downarrow, \ldots, x_{n-2}^\downarrow, x_{n-1}^\downarrow + \epsilon, x_n^\downarrow - \epsilon).$$

It is easy to check that

$$P(x' \otimes c \to y \otimes c) \leq \frac{(x'\otimes c)^\downarrow_{nk}}{(y\otimes c)^\downarrow_{nk}} = \frac{x_n^\downarrow - \epsilon}{y_n^\downarrow} < \frac{x_n^\downarrow}{y_n^\downarrow} = \lambda$$

for any probability vector c, where k is the dimension of c. Thus $x' \notin T^\lambda(y)$. It follows that x is a boundary point of $T^\lambda(y)$.

Conversely, suppose $x_n^\downarrow/y_n^\downarrow > \lambda$. By the definition of $x \in T^\lambda(y)$, there exists a probability vector c such that $P(x\otimes c \to y\otimes c) \geq \lambda$. If the inequality holds strictly, then x is of course in the interior of $T^\lambda(y)$ by the continuum of $P(x\otimes c \to y\otimes c)$ on x. So we need only consider the case of $P(x\otimes c \to y\otimes c) = \lambda$. In this case we have

$$P(x \otimes c \to y \otimes c) = \lambda < \frac{x_n^\downarrow}{y_n^\downarrow} = \frac{(x\otimes c)^\downarrow_{nk}}{(y\otimes c)^\downarrow_{nk}},$$

where k is the dimension of c. Thus $x \otimes c$ is an interior point of $T^\lambda(y \otimes c)$ from Lemma 8. On the other hand, since the function $x \mapsto x \otimes c$ is continuous, it follows that x is in the interior of the set $\{x : x \otimes c \in T^\lambda(y \otimes c)\}$, which is just a subset of $T^\lambda(y)$. So x is in the interior of $T^\lambda(y)$. ∎

Recall that the boundary point set of $T(y)$ is $\{x \in T(y) : x_1^\downarrow = y_1^\downarrow \text{ or } x_n^\downarrow = y_n^\downarrow\}$. Once again, the asymmetry of roles of the largest and the smallest components in determining the maximal transforming probability makes the boundary point set of $T^\lambda(y)$ different from that of $T(y)$.

The next theorem tells us when catalysis is helpful for probabilistic transformation with destination state y by giving a necessary and sufficient condition for $T^\lambda(y) = S^\lambda(y)$.

Theorem 10: Let $y \in V^n$ be a nonincreasingly arranged probability vector and $0 < \lambda < 1$. Then $T^\lambda(y) = S^\lambda(y)$ if and only if $y_2 = y_n$.

Proof: If $y_2 = y_n$, then for any nonincreasingly arranged $x \in V^n$ and any integer l satisfying $1 < l \leq n$,

$$\frac{E_l(x)}{E_l(y)} = \frac{\sum_{i=l}^n x_i}{(n-l+1)y_n} \geq \frac{(n-l+1)x_n}{(n-l+1)y_n} = \frac{x_n}{y_n}.$$

Thus $P(x \to y) = \min\{x_n/y_n, 1\}$ and for any vector c, $P(x\otimes c \to y \otimes c) = P(x \to y)$. It follows that $T^\lambda(y) = S^\lambda(y)$.

Conversely, suppose $y_2 > y_n$. Let $m > 1$ be the maximal index of the components of y which are not equal to y_n, that is, $y_m > y_{m+1} = \ldots = y_n$. Then we have $E_m(y) > (n - m + 1)y_n$. Let μ be a real number such that

$$\lambda < \mu < \min\{\frac{\lambda E_m(y)}{(n-m+1)y_n}, 1\},$$

and define a probability vector

$$x = (y_1 + \frac{1-\lambda}{m-1}E_m(y), \ldots, y_{m-1} + \frac{1-\lambda}{m-1}E_m(y),$$

$$\lambda E_m(y) - \mu E_{m+1}(y), \mu y_{m+1}, \ldots, \mu y_n).$$

It is a little tedious but very easy to check that the components of x have been nonincreasingly arranged and $P(x \to y) = \lambda < \mu = x_n/y_n$. By Theorem 9, x is an interior point of $T^\lambda(y)$. That completes our proof that $S^\lambda(y) \neq T^\lambda(y)$ since x is obviously on the boundary of $S^\lambda(y)$. ∎

Compared with the conditions of $T(y) = S(y)$ presented in [12], our condition in Theorem 10 is more simple and even a little surprising since it depends only on whether or not $y_2 = y_n$ (totally irrelevant to the value of λ).

In what follows, we consider the interesting question of whether or not there exists a bound on the dimension of partial catalysts that we should consider to help transforming some vector x into a given probability vector y. This is in fact a generalization of the problem considered for $T(y)$ in [12]. We will give a negative answer to this question by showing that in general $T^\lambda(y) \neq T_k^\lambda(y)$ for all positive k, where

$$T_k^\lambda(y) = \{x \in V^n : P(x \otimes c \to y \otimes c) \geq \lambda \ \text{ for some } c \in V^k\}$$

is defined to be the set of probability vectors which, with the aid of some k-dimensional catalyst vector, can be transformed into y with the maximal probability not less than λ.

Lemma 9: For all $y \in V^n$ and any positive integer k, $T_k^\lambda(y)$ is a closed set.

Proof: Suppose $x^1, x^2, \ldots$ is an arbitrary vector sequence in $T_k^\lambda(y)$ that converges to x. By the definition of $T_k^\lambda(y)$, there exists a catalyst sequence $c^1, c^2, \ldots$ in V^k such that $P(x^i \otimes c^i \to y \otimes c^i) \geq \lambda$ for $i = 1, 2, \ldots$. Notice that V^k is a compact set in R^k. There exists a convergent subsequence $c^{i_1}, c^{i_2}, \ldots$ of $c^1, c^2, \ldots$ that converges to, say, $c \in V^k$. Then we can deduce that $P(x \otimes c \to y \otimes c) \geq \lambda$ and so $x \in T_k^\lambda(y)$ from the fact that the function $P(x \otimes c \to y \otimes c)$ is continuous on the parameters x and c. ∎

Notice that in [12], a similar lemma about $T_k(y)$ was presented but the proof there was a little complex. The proof technique of the above lemma can be used to give a simpler one.

Theorem 11: Let $y \in V^n$ be a nonincreasingly arranged probability vector, $y_2 > y_n$, and $0 < \lambda < 1$. Then $T^\lambda(y) \neq T_k^\lambda(y)$ for all positive k.

Proof: We will complete the proof by showing that under the assumptions in this theorem, $T^\lambda(y)$ is not a closed set. Then from Lemma 9, we have $T^\lambda(y) \neq T_k^\lambda(y)$ for all positive k.

Let $e = (1/n, \ldots, 1/n)$ be the uniform vector in V^n. Notice that $e \in T^\lambda(y)$ and from Theorem 9, e is also an interior point of $T^\lambda(y)$. Denote by m the maximal index of the components of y which are not equal to y_n, that is,

$$y_m > y_{m+1} = \ldots = y_n. \tag{20}$$

Since $y_2 > y_n$, we have $1 < m < n$. Let $\mu \in (0, \lambda)$ and define a probability vector

$$x = (y_1 + \frac{1-\mu}{m-1}E_m(y), \ldots, y_{m-1} + \frac{1-\mu}{m-1}E_m(y), \mu y_m, \ldots, \mu y_n).$$

It is easy to check that the components of x are arranged nonincreasingly and $P(x \to y) = x_n/y_n = \mu < \lambda$. So $x \notin T^\lambda(y)$. If $T^\lambda(y)$ is closed, then the closed set $\{tx + (1-t)e : 0 \leq t \leq 1\}$ must intersect $T^\lambda(y)$ at a boundary point of $T^\lambda(y)$, say, $x' = t'x + (1-t')e$. Thus $x'_n/y_n = \lambda$ from Theorem 9. On the other hand, from Eq.(20) and noticing that $0 < t' < 1$, we have

$$\frac{x'_n}{y_n} = t'\mu + \frac{1-t'}{ny_n} > t'\mu + \frac{1-t'}{ny_m} = \frac{x'_m}{y_m},$$

and so

$$\frac{x'_m}{y_m} < \frac{x'_{m+1}}{y_{m+1}} = \ldots = \frac{x'_n}{y_n} = \lambda,$$

which contradicts 3) of Theorem 8. ∎

Notice that when $y_2 = y_n$, we have $T^\lambda(y) = S^\lambda(y)$ from Theorem 10, and so $T^\lambda(y) = T_k^\lambda(y)$ for any positive integer k since

$$T^\lambda(y) = S^\lambda(y) \subseteq T_k^\lambda(y) \subseteq T^\lambda(y).$$

In order to describe $T^\lambda(y)$ more deeply, we examine its extreme points in what follows.

The following lemma describes what kind of perturbations will not remove a point from $T^\lambda(y)$, even when the point is on the boundary. This lemma can also be regarded as a generalization of Corollary 5 in [12].

Lemma 10: Suppose $0 < \lambda < 1$, and $x, y \in V^n$ be two nonincreasingly arranged n-dimensional probability vectors satisfying $x \in T^\lambda(y)$ but $x \neq y_\lambda$. Let d be the maximal index of the components such that $x_d > \lambda y_d$. Then any sufficiently small perturbation will not remove x from $T^\lambda(y)$, provided that the perturbation does not affect the components $x_{d+1}, \ldots, x_n$. (Notice that if $d = n$, then x is an interior point of $T^\lambda(y)$ and the result here is just a simple property of interior points.)

Proof: To begin with, let us decompose x as $x = x' \oplus x''$ where

$$x' = (x_1, x_2, \ldots, x_d) \quad \text{and} \quad x'' = (x_{d+1}, \ldots, x_n).$$

Similarly, y can be decomposed as $y = y' \oplus y''$. Applying Lemma 7, we have

$$x \in T^\lambda(y) \quad \Leftrightarrow \quad \widetilde{x} \in T^{\lambda'}(\widetilde{y}), \tag{21}$$

where

$$\widetilde{x} = \frac{x'}{1-\mu\lambda} \ \text{ and } \ \widetilde{y} = \frac{y'}{1-\mu},$$

and μ is the sum of the components of y'' and λ' is defined by Eq.(18). Since $x_d > \lambda y_d$, we have

$$\widetilde{x}_d = \frac{x_d}{1-\mu\lambda} > \lambda' \frac{y_d}{1-\mu} = \lambda' \widetilde{y}_d,$$

and so $\widetilde{x}$ is an interior point of $T^\lambda(\widetilde{y})$ from Theorem 9. Notice that any perturbation to x which does not affect the components $x_{d+1}, \ldots, x_n$ is also a perturbation to $\widetilde{x}$ and vice versa. Therefore the lemma follows. ∎

Using this lemma, we can easily find out all the extreme points of $T^\lambda(y)$ for a given $y \in V^n$. It is rather surprising that $T^\lambda(y)$ and $S^\lambda(y)$ share the same extreme points, although they satisfy very different properties and also $S^\lambda(y) \subsetneqq T^\lambda(y)$ in general. Notice that $T^\lambda(y)$ is not closed in general by the proof of Theorem 11, so we cannot determine the whole set $T^\lambda(y)$ only by its extreme points.

Theorem 12: For all $y \in V^n$, the set of extreme points of $T^\lambda(y)$ is the same as that of $S^\lambda(y)$. That is, it is just the set

$$\{Py_\lambda : \ P \ \text{ is any } n \text{ dimesional permutation}\}.$$

Proof: First, we prove that any extreme point of $S^\lambda(y)$ is an extreme point of $T^\lambda(y)$. To prove this, we need only show

that y_λ is an extreme point of $T^\lambda(y)$. Suppose the components of y have been arranged nonincreasingly and

$$y_\lambda = ty' + (1-t)y'' \tag{22}$$

is a convex combination of y' and y'' in $T^\lambda(y)$, where $0 < t < 1$. From 3) of Theorem 8 we have

$$y'_n \geq y'^{\downarrow}_n \geq \lambda y_n$$

and

$$y''_n \geq y''^{\downarrow}_n \geq \lambda y_n$$

(notice that y' and y'' are not necessarily nonincreasingly arranged). But from Eq.(22) we have $ty'_n + (1-t)y''_n = \lambda y_n$. It follows that

$$y'_n = y''_n = \lambda y_n, \tag{23}$$

and furthermore, y'_n and y''_n are the smallest components of y' and y'', respectively. Now using 3) of Theorem 8 again, we can deduce that

$$y'_{n-1} = y''_{n-1} = \lambda y_{n-1}$$

from Eqs.(23) and (22). And again, y'_{n-1} and y''_{n-1} are the second smallest components of y' and y'', respectively. Repeating the above arguments, we finally come to the conclusion that

$$y' = y'' = y_\lambda.$$

Thus y_λ is an extreme point of $T^\lambda(y)$ as claimed.

Now suppose x is an extreme point of $T^\lambda(y)$ and x is nonincreasingly arranged. If it is not an extreme point of $S^\lambda(y)$, then $x \neq y_\lambda$ and so there exists a positive integer d, $1 < d \leq n$, such that $x_d > \lambda y_d$. By Lemma 10, we can find a sufficiently small but positive ϵ such that $x' \in T^\lambda(y)$ and $x'' \in T^\lambda(y)$, where

$$x' = (x_1, \ldots, x_{d-2}, x_{d-1} - \epsilon, x_d + \epsilon, x_{d+1}, \ldots, x_n)$$

and

$$x'' = (x_1, \ldots, x_{d-2}, x_{d-1} + \epsilon, x_d - \epsilon, x_{d+1}, \ldots, x_n).$$

Obviously $x = (x' + x'')/2$, which contradicts our assumption that x is an extreme point of $T^\lambda(y)$. That completes our proof. ∎

What we would like to point out here is that a similar argument on $T(y)$ instead of $T^\lambda(y)$ in the above theorem can lead to a solution to the open problem Nielsen proposed in his lecture notes on the theory of majorization and its applications in quantum information theory [14]. More specifically, $T(y)$ has a discrete set, but not a continuum as Nielsen conjectured, of extreme points (in fact, the extreme points of $T(y)$ are just $\{Py$: where P is any n-dimensional permutation$\}$).

V. Conclusion and Open problems

In this paper, we investigate carefully the power of catalysis in probabilistic entanglement transformations by LOCC. We give a necessary and sufficient condition for when some appropriately chosen catalyst can be helpful in increasing the maximal probability for a given probabilistic transformation. An efficient algorithm is presented to decide whether partial catalysts with a given dimension exist for a certain probabilistic transformation, which leads to a method for constructing the most economical partial catalysts with minimal dimension. We also study the set of states that can be transformed by catalyst-assisted LOCC into a given state with the maximal probability not less than a given positive number. We prove that this set shares many properties with the well known set consisting of all vectors being trumped by the given state. More mathematical structure of catalyst-assisted probabilistic transformation is also considered.

The emphasis of this paper is to determine when the maximal transforming probability can be increased in the presence of partial catalysts and how to construct appropriate ones. The amount of the probability increased is, however, not considered. So an open problem and also an important direction for further study is to determine the maximum of the transforming probability that can be reached with the aid of partial catalysts. This is also a generalization of deterministic catalysis since the problem of the existence of catalysts for deterministic transformations is equivalent to the problem of the existence of partial catalysts which can increase the transforming probability to 1 for probabilistic transformations.

At the end of Section IV, we have determined all the extreme points of $T^\lambda(y)$. However, we still know little about the geometric structure of $T^\lambda(y)$. The main reason is that $T^\lambda(y)$ is not closed in general and so it is not the convex hull of its extreme points. How to determine the accumulation points outside $T^\lambda(y)$ is really a hard problem and remains open. We believe that $\overline{T}^\lambda(y)$, the closure of $T^\lambda(y)$, has a continuum of extreme points, just as Nielsen conjectured.

Acknowledgment

The authors are grateful to the two referees. Their detailed comments and suggestions have greatly improved the quality of the paper.

References

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[2] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing", *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, Bangalore, India, 1984, pp. 175-179.

[3] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states", *Phys. Rev. Lett.*, vol. 69, pp. 2881-2884, Nov. 1992.

[4] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels", *Phys. Rev. Lett.*, vol. 70, pp. 1895-1899, Mar. 1993.

[5] C. H. Bennett and P. W. Shor, "Quantum Information Theory", *IEEE Trans. Inform. Theory*, vol. 44, pp. 2724-2742, Oct. 1998.

[6] M. A. Nielsen, "Conditions for a class of entanglement transformations", *Phys. Rev. Lett.*, vol. 83, pp. 436-439, Jul. 1999.

[7] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

[8] P. Alberti and A. Uhlmann, *Stochasticity and Partial Order: Doubly Stochastic Maps and Unitary Mixing*, Dordrecht, Boston, 1982.

[9] G. Vidal, "Entanglement of Pure States for a Single Copy", *Phys. Rev. Lett.*, vol. 83, pp. 1046-1049, Aug. 1999.

[10] W. Duer, G. Vidal, J. I. Cirac., Three qubits can be entangled in two inequivalent ways, *Phys. Rev. A*, vol. 62, pp. 062314, Nov. 2000.

[11] D. Jonathan and M. B. Plenio, "Entanglement-Assisted Local Manipulation of Pure Quantum States", *Phys. Rev. Lett.*, vol. 83, pp. 3566-3569, Oct. 1999.
[12] S. Daftuar and M. Klimesh, "Mathematical structure of entanglement catalysis", *Phys. Rev. A*, vol. 64, pp. 042314, Sep. 2001.
[13] J. Eisert,M. Wilkens, Catalysis of Entanglement Manipulation for Mixed States, *Phys. Rev. Lett.*, vol. 85, pp. 437-440, Jul. 2000.
[14] M. A. Nielsen, Majorization and its applications to quantum information theory. [Online]. Available: http://www.qinfo.org/people/nielsen /info/maj.ps.
[15] G. H. Hardy, J. E. Littlewood and G. Polya, *Inqualities*, Cambridge university press, Cambridge, 1952.
[16] X. M. Sun, R. Y. Duan and M. S. Ying, The Existence of Quantum Entanglement Catalysts. [Online]. Available: http://www.arXiv.org/ abs/quant-ph/0311133.
[17] R. Y. Duan, Y. Feng, X. Li, and M. S. Ying, Multiple-Copy Entanglement Transformation and Entanglement Catalysis. [Online]. Available: http://www.arXiv.org/abs/quant-ph/0404148.

Yuan Feng received the B.S. degree from Department of Mathematics, Tsinghua University in 1999, and the Ph. D degree in Computer Science from Department of Computer Science and Technology, Tsinghua University in 2004. His current research is focusing on Quantum Information and Quantum Computation.

Runyao Duan is a PhD candidate in Department of Computer Science and Technology, Tsinghua University, China. He received the B.S. degree in Computer Science from Tsinghua University in 2002. His research interests are in Quantum Computation and Quantum Information Theory.

Mingsheng Ying graduated from Department of Mathematics, Fuzhou Teachers College in 1981. He is curently Cheung Kong Chair Professor at State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing, China. His research interests are formal methods, foundations of artificial intelligence, quantum information, and fuzzy logic. He has publishe more than 50 papers in various international journals. Also, he is the author of the book "Topology in Process Calculus: Approximate Correctness and Infinite Evolution of Concurrent Program" (Springer-Verlag, New York, 2001).