Catalytic transformations for bipartite pure states

2007 J. Phys. A: Math. Theor. 40 12185

(http://iopscience.iop.org/1751-8121/40/40/012)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 12185–12212 doi:10.1088/1751-8113/40/40/012

Catalytic transformations for bipartite pure states

S Turgut

Department of Physics, Middle East Technical University, 06531 Ankara, Turkey

E-mail: sturgut@metu.edu.tr

Received 17 July 2007, in final form 3 September 2007
Published 18 September 2007
Online at stacks.iop.org/JPhysA/40/12185

Abstract
Entanglement catalysis is a phenomenon that usually enhances the conversion probability in the transformation of entangled states by the temporary involvement of another entangled state (so-called catalyst), where after the process is completed the catalyst is returned to the same state. For some pairs of bipartite pure entangled states, catalysis enables a transformation with unit probability of success, in which case the respective Schmidt coefficients of the states are said to satisfy the trumping relation, a mathematical relation which is an extension of the majorization relation. This paper provides all necessary and sufficient conditions for the trumping and two other associated relations. Using these conditions, the least upper bound of conversion probabilities using catalysis is also obtained. Moreover, best conversion ratios achievable with catalysis are found for transformations involving many copies of states.

PACS numbers: 03.67.Mn, 03.65.Ud

1. Introduction

A major problem in quantum information theory is to understand the conditions for transforming a given entangled state into another desired state by using only local quantum operations assisted with classical communication (LOCC). Significant development has been achieved for the case of pure bipartite states. Bennett *et al* have established the entropy of entanglement as the sole conversion currency in the asymptotic case, where essentially an infinite number of copies of states are transformed into each other [1]. In this case, conversion is possible as long as the entropy of entanglement decreases with the fractional drop in that quantity and the failure probability can be made desirably small and the fidelity between the final state and the desired target state can be made desirably large.

Away from the asymptotic limit, where a single copy of a given state is to be transformed into another given state, such a simple conversion criterion cannot be found and investigations have unearthed a deep connection of the problem to the mathematical theory of majorization [2]. For setting up the necessary notation, the following definitions need to be introduced first. For two vectors x and y with n real elements, we say that x is *super-majorized* by y

1751-8113/07/4012185+28$30.00

(written $x \prec^w y$), if $F_m(x) \geqslant F_m(y)$ for all $m = 1, 2, \ldots, n$. Here,

$$F_m(x) = x_1^\uparrow + x_2^\uparrow + \cdots + x_m^\uparrow \tag{1}$$

denotes the sum of the smallest m elements of x, where $x^\uparrow$ is the vector x with all elements arranged in a non-decreasing order $\left(x_1^\uparrow \leqslant x_2^\uparrow \leqslant \cdots \leqslant x_n^\uparrow\right)$. If, in addition to these, the two vectors have the same sum ($F_n(x) = F_n(y)$) then we say that x is *majorized* by y (written $x \prec y$).

Given two entangled states in the Schmidt form,

$$|\psi\rangle = \sum_{i=1}^{n} \sqrt{x_i}|i_A \otimes i_B\rangle, \tag{2}$$

$$|\phi\rangle = \sum_{i=1}^{n} \sqrt{y_i}|i'_A \otimes i'_B\rangle, \tag{3}$$

where x and y are the respective Schmidt coefficients $\left(\sum x_i = \sum y_i = 1\right)$, the problem is essentially to determine the probability that $|\phi\rangle$ can be obtained by LOCC starting from the state $|\psi\rangle$. As two entangled states with the same Schmidt coefficients are equivalent under local unitaries, the probability depends only on the Schmidt coefficients and not on the particular local orthonormal bases in which they are expressed. For that reason, the conversion probability of $|\psi\rangle$ into $|\phi\rangle$ will be simply denoted by $P(x \to y)$.

The most important step in the solution of this problem is taken by Nielsen who has shown that $|\psi\rangle$ can be converted into $|\phi\rangle$ with certainty, i.e. $P(x \to y) = 1$, if and only if $x \prec y$ [3]. Subsequently, Vidal has obtained the expression

$$P(x \to y) = \min_{1 \leqslant m \leqslant n} \frac{F_m(x)}{F_m(y)}, \tag{4}$$

for the conversion probability between two arbitrary states [4]. Note that the conversion probability is equal to the largest value of λ such that x is super-majorized by λy, i.e.

$$P(x \to y) = \max\{\lambda : \lambda \geqslant 0,\ x \prec^w \lambda y\}, \tag{5}$$

where λy denotes the vector obtained by multiplying each element of y with λ.

Soon afterwards, Jonathan and Plenio have demonstrated an interesting effect that is termed as catalysis or entanglement-assisted local transformation [5]. If an additional entangled pair (a catalyst) shared by the same parties is involved in the transformation process in such a way that it reappears in the same form at the end, then the conversion probability is improved in some cases. To be explicit, let $|\chi\rangle = \sum_{\ell=1}^{N} \sqrt{c_\ell}|\ell_A \otimes \ell_B\rangle$ be the state of the catalyst which is shared by the same parties. For some cases, even though $|\psi\rangle$ to $|\phi\rangle$ conversion only proceeds with a probability smaller than 1, the state $|\psi\rangle \otimes |\chi\rangle$ can be converted into $|\phi\rangle \otimes |\chi\rangle$ with certainty (in terms of Schmidt coefficients, we have $P(x \otimes c \to y \otimes c) = 1$ and $P(x \to y) < 1$). In such a transformation, the entanglement of the catalyst $|\chi\rangle$ is not consumed, even though it takes part in the transformation. Catalysis is also helpful in probabilistic transformations; in a large number of cases, we have $P(x \otimes c \to y \otimes c) > P(x \to y)$.

Subsequently, a lot of research has been directed to understanding the catalytic transformations [6]. A big impediment in most of these studies is the absence of strong results on majorization and tensor products. One purpose of this paper is to partially remove this obstacle. In doing so, two major problems in entanglement catalysis, which are described below, will be solved.

One problem is to determine when catalysis enables us to carry out a transformation with certainty. Nielsen has suggested the following notation for this purpose: for two n-component vectors x and y we say that x is *trumped* by y (written $x \prec_T y$) when there is a vector c (a catalyst), consisting only of positive numbers, such that $x \otimes c \prec y \otimes c$. Detailed investigations of the mathematical properties of the trumping relation indicate a very rich structure [7]. It is found that, for example, when catalysis is useful for a final state y, then it is possible to find possible initial states which require catalysts having arbitrarily large Schmidt numbers.

The problem at hand can also be equivalently posed as finding all necessary and sufficient conditions for the $x \prec_T y$ relation. Such conditions have been found only for four component states having two component catalysts [8]. Recently, partial solutions to this problem have been obtained by Aubrun and Nechita, who determined the closure of the set of vectors trumped by a fixed final state [9, 10]. But, since the closure is on a space of vectors which have a larger number of components than the final state, the conditions they have found are incomplete.

It is also of some interest to see how far the conversion probability can be improved for cases where catalysis cannot achieve a conversion with certainty. A problem that is investigated in detail is to find cases where catalysis is useful, i.e. $P(x \otimes c \to y \otimes c) > P(x \to y)$ for some c, and to find algorithms for searching possible catalysts for this purpose [11, 12]. The second problem to be solved in this paper is the determination of the least upper bound on conversion probabilities involving a catalyst, i.e. $P_{\text{cat}}(x \to y) = \sup_c P(x \otimes c \to y \otimes c)$, where the supremum is taken over all finite vectors, c, of positive numbers. In other words, for any probability smaller than (in some cases, equal to) this quantity, it is possible to find a catalyst that achieves a transformation with that probability. However, such transformations do not conform fully with the spirit of catalysis. Although the catalyst can be recovered when the transformation is successful, when it fails the catalyst will also be lost, at least partially. The author is not aware of any studies on catalytic conversions where recovery upon failure is also taken into account; apparently this is a very complicated problem. However, the determination of $P_{\text{cat}}(x \to y)$ can be seen as a first step towards such studies as that quantity can be used as an upper bound (but not least) on maximum probability that can be achieved with true catalysis.

That quantity can be expressed in a simple way by introducing another mathematical notation. For two n-element vectors x and y we will say that x is *super-trumped* by y (written $x \prec_T^w y$) when there is a catalyst c, consisting only of positive numbers, such that $x \otimes c \prec^w y \otimes c$. In that case, P_{cat} can be expressed as

$$P_{\text{cat}}(x \to y) = \sup \left\{ \lambda : \lambda \geqslant 0, x \prec_T^w \lambda y \right\}. \tag{6}$$

As a result, finding all necessary and sufficient conditions for $x \prec_T^w y$ will enable us to compute this quantity.

This paper provides the needed necessary and sufficient conditions for both the trumping and super-trumping relations. The organization of the paper is as follows. Section 2 expresses the theorems that provide all of the necessary and sufficient conditions for the relations defined above and an additional relation which is introduced for the sake of completeness. In section 3, two key lemmas are proved. The succeeding three sections are devoted to the proofs of the theorems about these relations. Section 7 discusses some immediate conclusions that can be drawn from these theorems and finally section 8 contains brief conclusions.

2. Catalytic majorization

First, a few statements must be made about the trumping and super-trumping relations. As the applications in the quantum information theory is the main concern, only vectors with non-negative elements will be considered. For comparing two vectors with unequal lengths, zeros should be padded to the shortest vector to make them of equal length. As either $x \prec^w y$ or $x \prec y$ or the associated trumping relations imply that the number of zero elements of x is less than that of y, the corresponding zeros can be dropped from both vectors. As a result, it can be assumed that x does not have any zero elements. This step will be necessary for comparing the conditions given below. Note that, in view of the connection with the entanglement transformations, the vector x in here will be associated with the Schmidt coefficients of the initial state. As a result, x has no zero elements while y may have some which correspond to the fact that the Schmidt number cannot increase in an entanglement transformation (with or without catalysis).

If $x \prec_T^w y$, then it necessarily follows that $\sum x_i \geqslant \sum y_i$. In the special case where the vectors have the same sum $\left(\sum x_i = \sum y_i\right)$, the relation $x \prec_T^w y$ is equivalent to the trumping relation $x \prec_T y$. In other words, the trumping relation is a special case of the super-trumping relation. However, it appears that the two cases for sums (strict inequality and equality) have quite different necessary and sufficient conditions. First, we start with the equality case, which is covered by the following theorem.

Theorem 1. *For two n-element vectors of non-negative numbers x and y such that x has non-zero elements and the vectors are distinct (i.e. $x^\uparrow \neq y^\uparrow$), the relation $x \prec_T y$ is equivalent to the following three strict inequalities:*

$$A_\nu(x) > A_\nu(y), \qquad \forall \nu \in (-\infty, 1), \tag{T1}$$

$$A_\nu(x) < A_\nu(y), \qquad \forall \nu \in (1, +\infty), \tag{T2}$$

$$\sigma(x) > \sigma(y). \tag{T3}$$

Here, σ denotes the function

$$\sigma(x) = -\sum_{i=1}^{n} x_i \ln x_i, \tag{7}$$

which gives the entropy of entanglement in the case the vector x is normalized. However, for checking condition (T3), normalization is not necessary. Moreover, A_ν denotes the νth power mean of vectors, i.e. for x, it is defined as

$$A_\nu(x) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^\nu\right)^{\frac{1}{\nu}}, \tag{8}$$

and similarly for $A_\nu(y)$. This is a bounded, continuous function of ν. It has the limits $A_{-\infty}(x) = x_1^\uparrow$, the minimum element, and $A_{+\infty}(x) = x_n^\uparrow$, the maximum element of x. For the particular value $\nu = 0$, it gives the geometric mean $A_0(x) = \left(\prod x_i\right)^{1/n}$. Note that, if any element of the vector y is zero, then $A_\nu(y) = 0$ for all $\nu \leqslant 0$. In that case, conditions (T1) should only be checked for $0 < \nu < 1$.

By the continuity of the power mean function against ν, the requirement that the vectors x and y have the same sum is included in conditions (T1) and (T2). Moreover, the limits of these inequalities at $\nu = -\infty$ and $\nu = +\infty$ imply that the minimum and maximum elements of the vectors satisfy the respective inequalities $x_1^\uparrow \geqslant y_1^\uparrow$ and $x_n^\uparrow \leqslant y_n^\uparrow$, but these do not have to be strict.

Conditions (T1)–(T3) could have been expressed using power sums, ℓ_ν norms or Renyi entropies. But, the well-defined behavior of the power means at $\nu = 0$ is the main reason why they are preferred (the associated strict inequality will be used in the proofs). However, it should be noted that all of these three conditions can be expressed simply as the positivity of the expression

$$\Delta R_\nu = \frac{1}{\nu - 1} \ln \frac{A_\nu(y)}{A_\nu(x)} = \frac{\sigma_\nu(x) - \sigma_\nu(y)}{\nu} \tag{9}$$

for all finite values of ν, where σ_ν represents the Renyi entropy

$$\sigma_\nu(x) = \frac{1}{1 - \nu} \ln \left(\sum_{i=1}^{n} x_i^\nu \right), \tag{10}$$

where it is understood that an appropriate limit should be taken for ΔR_ν at $\nu = 0$. Moreover, the definition of the Renyi entropy is extended to negative powers ν, but some care is needed for such cases as the number of components n plays a crucial role. If y has some zero components, then $\sigma_\nu(y)$ should be taken as $+\infty$ for all $\nu \leqslant 0$.

Power means also have the following valuable feature: as a function of the vector x, $A_\nu(x)$ is convex for $\nu > 1$ and concave for $\nu < 1$. The simplest way to see this is to compute the second derivative

$$\left. \frac{\mathrm{d}^2}{\mathrm{d}t^2} A_\nu(x + tz) \right|_{t=0} = (\nu - 1) A_\nu(x) (\langle u^2 \rangle - \langle u \rangle^2), \tag{11}$$

where $u_i = z_i / x_i$ and the averages are taken with weight factors x_i^ν. If x is constrained to a subspace of real vectors with a fixed sum, then $A_\nu(x)$ is strictly convex or concave in the associated regions. In view of conditions (T1)–(T3), this property of the power mean function conforms well with the convexity of the set of vectors trumped by a fixed vector y.

Conditions (T1)–(T3) imply that the statement $x^{\otimes k} \prec_T y^{\otimes k}$ for any integer k is equivalent to $x \prec_T y$. Some of the results in [13] about the connection between the multiple-copy entanglement transformation, a related phenomenon discovered by Badyopadhyay *et al* [14], and the trumping relation can then be easily understood. In other words, if k copies of a state with coefficients x can be transformed into k copies of another state with coefficients y, either with or without catalysis, then x must be trumped by y.

Next, we consider the super-trumping relation, $x \prec_T^w y$, for the case where the vectors have different sums. The necessary and sufficient conditions for this case appear to be only the first set of strict inequalities.

Theorem 2. *For two n-element vectors x and y of non-negative numbers such that x has only positive elements and* $\sum x_i > \sum y_i$*, the relation* $x \prec_T^w y$ *is equivalent to the conditions (T1).*

Again, all inequalities in (T1) are strict. The end point $\nu = -\infty$ is not included in these conditions, but from the limit, it can be found that a non-strict inequality, i.e. $x_1^\uparrow \geqslant y_1^\uparrow$, is satisfied at that point. The other end point $\nu = 1$ has the strict inequality, $A_1(x) > A_1(y)$, but this case is covered by the assumptions of the theorem.

Although it will not be used for the applications in entanglement transformations, a third theorem related to the sub-majorization relation must be given for the sake of completeness. For two vectors x and y of n elements, we say that x is *sub-majorized* by y (written $x \prec_w y$) if $x_1^\downarrow + x_2^\downarrow + \cdots + x_m^\downarrow \leqslant y_1^\downarrow + y_2^\downarrow + \cdots + y_m^\downarrow$ for all $m = 1, 2, \ldots, n$. Here, $x^\downarrow$ represents the vector x with all elements arranged in a non-increasing order. Similarly, we will say that x is *sub-trumped* by y (written $x \prec_{wT} y$) when there is a catalyst c such that $x \otimes c \prec_w y \otimes c$. In such a case, we have $\sum x_i \leqslant \sum y_i$ with the equality case being equivalent to $x \prec_T y$. The following theorem covers the strict inequality case.

Theorem 3. *For two n-element vectors x and y of non-negative numbers such that* $\sum x_i < \sum y_i$*, the relation* $x \prec_{wT} y$ *is equivalent to conditions (T2).*

Similar comments apply for the inequalities at the end points of (T2). The $\nu = +\infty$ limit gives $x_1^{\downarrow} \leqslant y_1^{\downarrow}$, a non-strict inequality, and at $\nu = 1$ we have $A_1(x) < A_1(y)$, which is covered by the assumptions. Note that in this case, any of the vectors x or y might have more zeros than the other.

Although the trumping relation is a special case of the super-trumping and sub-trumping relations, the necessary and sufficient conditions for the trumping relation cannot be obtained by simply extrapolating those for the other two relations. The condition for the positive drop in the entropy of entanglement (T3) is independent of the other two conditions (T1) and (T2); it cannot be derived starting from these. Perhaps the simplest counterexample is the following:

$$y = \tfrac{1}{96}(u_8 \oplus 4u_{18} \oplus 16u_1), \tag{12}$$

$$x = \tfrac{1}{96}(2u_{20} \oplus 8u_7), \tag{13}$$

where u_m denotes a uniform vector of m elements consisting of only 1's and $\oplus$ denotes the concatenation of the vectors. Here, x and y are normalized vectors consisting of $n = 27$ elements. It can easily be checked that x and y satisfy both of the conditions (T1) and (T2), but condition (T3) is not satisfied as they have equal values for the entropy function, $\sigma(x) = \sigma(y) = (17/6)\ln 2 + \ln 3$. As a result, the pair x and y does not satisfy any of the trumping relations, i.e. $x \not\prec_T y$, $x \not\prec_T^w y$ and $x \not\prec_{wT} y$.

The proofs of all of these three theorems are lengthy and will be done in separate sections. The following properties of the majorization and trumping relations will be used occasionally in the proofs [2].

(1) All of the majorization relations $\prec$, $\prec^w$ and $\prec_w$, and the corresponding trumping relations $\prec_T$, $\prec_T^w$ and $\prec_{wT}$ are partial orders on vectors with n-elements (up to equivalence under permutation of vector elements).

(2) If two vectors satisfy a particular majorization relation, then the corresponding trumping relation will also be satisfied (with any vector being a possible catalyst), e.g., $x \prec^w y$ implies $x \prec_T^w y$.

(3) If $\bar{x}$ and x are vectors such that $\bar{x}_i \geqslant x_i$ for all i, then it necessarily follows that $\bar{x} \prec^w x$ and $x \prec_w \bar{x}$.

(4) For any vector x, we define the *characteristic function*

$$H_x(t) = \sum_{i=1}^{n} (t - x_i)^+, \tag{14}$$

where $(\alpha)^+ = \max(\alpha, 0)$ denotes the positive-part function. For non-negative vectors x and y, the relation $x \prec^w y$ can be equivalently stated as

$$H_x(t) \leqslant H_y(t), \qquad \forall t \geqslant 0. \tag{15}$$

Moreover, if the vectors are distinct, i.e. $x^{\uparrow} \neq y^{\uparrow}$, then the difference $H_y(t) - H_x(t)$ is strictly positive for some interval in t. If the vectors have the same sum, $\sum x_i = \sum y_i$, then (15) is equivalent to $x \prec y$.

(5) A different characteristic function, $H_x'(t) = \sum (x_i - t)^+$, has to be used for the sub-majorization relation. When x and y are non-negative, the relation $x \prec_w y$ is equivalent to

$$H_x'(t) \leqslant H_y'(t), \qquad \forall t \geqslant 0. \tag{16}$$

(6) If $x \prec y$, then for any convex function f, we have

$$\sum_{i=1}^{n} f(x_i) \leqslant \sum_{i=1}^{n} f(y_i). \tag{17}$$

Moreover, if $x^{\uparrow} \neq y^{\uparrow}$ and f is strictly convex, then the inequality above is strict.

(7) For the tensor-product of two vectors, we have $H_{x\otimes c} = \sum_{\ell} c_\ell H_x(t/c_\ell)$.

(8) Concatenation or removal of identical numbers to or from two given vectors does not change the majorization or trumping relation between them (e.g., if z is a vector of positive numbers, then the statement $x \prec y$ is equivalent to $x \oplus z \prec y \oplus z$, and similarly for all the other relations). Moreover, all of the inequalities in (T1)–(T3) are also unchanged under such concatenation and removals. This enables us to supply the proofs of these theorems only for the special case where x and y have no common elements and then generalize it to arbitrary vectors. Such an assumption will be made whenever it is convenient.

3. Two key lemmas

The proofs of the theorems stated in the previous section need the solution of the following problem. Let $\gamma(s) = \sum_{m=0}^{N} \gamma_m s^m$ be a real polynomial where some of the coefficients γ_m might be negative. The problem is to express γ as a ratio of two polynomials as $\gamma(s) = b(s)/a(s)$, where a and b have non-negative coefficients. Additionally, it is also required that the coefficients of a are integers. It is obvious that if that problem has a solution, then γ should have no positive root. The following lemma shows that this condition is also sufficient.

Lemma 1. *Let $\gamma(s)$ be a polynomial such that $\gamma(s) > 0$ for all $s > 0$. Then,*

(a) $\gamma(s)$ can be expressed as $\gamma(s) = b(s)/a(s)$, where $a(s)$ and $b(s)$ are polynomials with non-negative coefficients;

(b) moreover, $a(s)$ can be chosen as a polynomial with integer coefficients.

Proof. The proof of part (a) will first be given for a quadratic polynomial with complex roots. The general case follows easily once this is completed.

Consider $\gamma(s) = 1 - 2\xi s + \lambda s^2$, where $\lambda > \xi^2$. Obviously, for $\xi \leqslant 0$ there is nothing to be proven, so consider $\xi > 0$ for the following. Let N be an integer sufficiently large so that

$$\frac{1}{4}\left(\frac{(2N)!}{N!^2}\right)^{\frac{1}{N}} \geqslant \frac{\xi^2}{\lambda}. \tag{18}$$

Such an N can always be found as the left-hand side has limit 1 as $N \to \infty$ and the right-hand side is strictly less than 1. In that case, the polynomials

$$a(s) = \sum_{k=0}^{2N-1} (1 + \lambda s^2)^k (2\xi s)^{2N-1-k}, \tag{19}$$

$$b(s) = (1 + \lambda s^2)^{2N} - (2\xi s)^{2N}, \tag{20}$$

satisfy the desired properties. All coefficients of a are obviously non-negative. That is true for b as well, since the coefficient of s^{2N} is $\lambda^N (2N)!/N!^2 - (2\xi)^{2N}$ which is also non-negative by the special choice of N.

Now, consider an arbitrary polynomial γ which satisfies the conditions of the lemma. Express γ as a product of its real irreducible factors as

$$\gamma(s) = As^r \prod_i (1 - \zeta_i s) \prod_i \left(1 - 2\xi_i s + \left(\xi_i^2 + \eta_i^2\right)s^2\right), \tag{21}$$

where $A > 0, r\ (\geqslant 0)$ is the multiplicity of a possible root at 0, $1/\zeta_i$ are the real roots and $(\xi_i \pm \mathrm{i}\eta_i)^{-1}$ are the complex roots of γ. Since γ has no positive root, $\zeta_i < 0$ for all i. As a result, the product of linear factors of γ is a polynomial with non-negative coefficients. Finally, each quadratic factor with complex roots can be written as a ratio of polynomials with non-negative coefficients. As the product of such polynomials has non-negative coefficients, the statement in part (a) follows.

For the proof of part (b), it will be assumed without loss of generality that γ has no root at 0. In that case, $a(s)$ and $b(s)$ can be chosen to have non-zero constant terms. Before passing on to the proof, it will first be shown that $b(s)$ can be chosen such that all of its coefficients are strictly positive. Consider a degree m solution for $b(s)$, i.e. $b(s) = \sum_{k=0}^{m} b_k s^k$ where $b_0 > 0, b_m > 0$ and $b_k \geqslant 0$ for all $1 \leqslant k < m$. Let $e(s) = 1 + s + \cdots + s^{m-1}$. Then $e(s)b(s)$ is a polynomial with degree $2m - 1$ and all of its $2m$ coefficients are positive. Moreover, the polynomials $e(s)b(s)$ and $e(s)a(s)$ satisfy the conditions of part (a). This shows that the polynomial $b(s)$ can be chosen to have non-zero coefficients.

For the proof of part (b), suppose that $b(s)$ is a degree m polynomial with positive coefficients and let $\beta = \min_{0\leqslant k\leqslant m} b_k$ be the minimum of those. Let

$$\epsilon = \frac{\beta}{\sum_k |\gamma_k|}, \tag{22}$$

where γ_k are the coefficients of the polynomial $\gamma(s)$. Define a new polynomial $\bar{a}(s)$ such that it has the same degree as $a(s)$ and its coefficients are chosen from rational numbers such that

$$|\bar{a}_k - a_k| \leqslant \epsilon, \qquad k = 0, 1, \ldots, N, \tag{23}$$

where $\bar{a}_k$ and a_k are the coefficients of $\bar{a}(s)$ and $a(s)$ respectively. As the rational numbers are dense, this can always be done. If $\bar{a}(s)\gamma(s) = \bar{b}(s)$, then the coefficients of $\bar{b}(s)$ satisfy

$$\bar{b}_k - b_k = \sum_{\ell} (\bar{a}_\ell - a_\ell)\gamma_{k-\ell} \geqslant -\epsilon \sum_{\ell} |\gamma_\ell| = -\beta. \tag{24}$$

Therefore, $\bar{b}_k \geqslant b_k - \beta \geqslant 0$, i.e. $\bar{b}(s)$ has non-negative coefficients as desired. Multiplying $\bar{a}(s)$ by the common denominator of its coefficients gives a polynomial with integer coefficients. □

Our next job is to generalize lemma 1 to cases where the polynomial γ has a positive root. In that case, $a(s)$ and $b(s)$ cannot be polynomials and should be expressed as infinite series. Although the relation $a(s)\gamma(s) = b(s)$ could be satisfied as formal power series, for the purposes of this paper, these two series will be required to have a sufficiently large radius of convergence. For example, it can be required that $a(R)$ and $b(R)$ are convergent where R is a positive number. The following lemma gives the solution to this problem.

Lemma 2. *Let $R > 0$ and $\gamma(s)$ be a polynomial such that $\gamma(s) > 0$ for all $s \in (0, R]$. Then,*

(a) it is possible to find two non-zero power series, $a(s) = \sum_{m=0}^{\infty} a_m s^m$ and $b(s) = \sum_{m=0}^{\infty} b_m s^m$, such that (i) $a_m \geqslant 0$ and $b_m \geqslant 0$ for all m, (ii) $a(R)$ and $b(R)$ are convergent and (iii) $\gamma(s) = b(s)/a(s)$;

(b) the series $a(s)$ can be chosen such that all series coefficients, a_m, are rational numbers;

(c) moreover, if $R > 1$, then the series $a(s)$ can be chosen such that (b) is satisfied and the value $a(1)$ is a rational number.

Proof. For part (a), express γ as a product of its real irreducible factors as in (21). It is sufficient to show that each linear factor with a positive root can be expressed as a ratio of two series satisfying the desired conditions. For this purpose, consider $\gamma_i(s) = 1 - \zeta_i s$, where $\zeta_i > 0$.

As γ has no root in $(0, R]$, we necessarily have $\zeta_i R < 1$ and therefore $\gamma_i(s) = b_i(s)/a_i(s)$ where $b_i(s) = 1$ and

$$a_i(s) = \gamma_i(s)^{-1} = \sum_{m=0}^{\infty} \zeta_i^m s^m, \tag{25}$$

which is convergent at $s = R$ and has positive coefficients. As the product of power series convergent at R is also convergent at R, and the rest of the factors in (21) are covered by lemma 1, we reach to the desired result.

For the rest of the lemma, it is sufficient to prove part (c), as a special case of the proof will give part (b). First, note that if the series $a(s)$ and $b(s)$ are a solution of the problem, then for any series $e(s)$ with non-negative coefficients, $e(s)a(s)$ and $e(s)b(s)$ are also a solution. It will be shown that $e(s)$ can be chosen such that all requirements in part (c) are satisfied. Let $a(s)$ be an arbitrary solution such that $a_0 = 1$. Let λ be a positive number such that $e^{\lambda}a(1)$ is a rational number and let $g_n = \lambda^n/n!$. The series coefficients e_m will be defined as follows: choose $e_0 = 1$ and, successively for all $m > 1$, choose e_m in the interval $[G_m, G_m + g_m]$ where

$$G_m = \sum_{\ell=0}^{m-1} g_\ell - e_\ell, \tag{26}$$

in such a way that

$$e_m + \sum_{\ell=1}^{m} a_\ell e_{m-\ell} = \bar{a}_m \tag{27}$$

is a rational number. As the interval has a finite width, this can always be done. From the upper bound on e_{m-1}, it can be seen that $G_m \geqslant 0$; as a result, each e_m is non-negative. Also, from the lower bound on e_{m-1}, it can be seen that $G_m \leqslant g_{m-1}$, from which we deduce that $e_m \leqslant g_m + g_{m-1}$. As a result, the series $e(s) = \sum_{m=0}^{\infty} e_m s^m$ has infinite radius of convergence. Finally, note that

$$\sum_{\ell=0}^{m-1} g_\ell \leqslant \sum_{\ell=0}^{m} e_\ell \leqslant \sum_{\ell=0}^{m} g_\ell, \tag{28}$$

and therefore $e(1)$ converges to e^{λ}. In conclusion, $\bar{a}(s) = e(s)a(s)$ is convergent at $s = R$, has non-negative rational coefficients and $\bar{a}(1) = e^{\lambda}a(1)$ is rational. □

4. Super-trumping relation

We start with the proof of theorem 2, which gives the necessary and sufficient conditions for the super-trumping relation $x \prec_T^w y$ for the case $\sum x_i > \sum y_i$. First, we prove that conditions (T1) are necessary.

Suppose x and y are vectors satisfying the conditions of theorem 2, such that $x \prec_T^w y$. In that case, there is a catalyst c with positive elements such that $x \otimes c \prec^w y \otimes c$ and therefore

$$\Delta(t) = H_{y\otimes c}(t) - H_{x\otimes c}(t) \tag{29}$$

is non-negative for all $t \geqslant 0$. Note that for $t > c_{\max} \max\left(x_n^{\uparrow}, y_n^{\uparrow}\right)$, the function $\Delta(t)$ has the constant value $\left(\sum x_i - \sum y_i\right) \sum c_\ell$. For that reason, the integral

$$I_\nu = \int_0^{\infty} \Delta(t) t^{\nu-2}\, \mathrm{d}t \tag{30}$$

is convergent at $t = \infty$ for all values of $\nu < 1$. For the convergence of the integral at $t = 0$, two cases must be distinguished. (i) If y has no zero elements, then $\Delta(t) = 0$ for a sufficiently small t and the integral is convergent at $t = 0$ for any ν. (ii) If y has zero entries, then $\Delta(t) \propto t$ near $t = 0$ and therefore the integral is convergent only for $0 < \nu < 1$, but this is sufficient for us as (T1) is automatically satisfied for all $\nu \leqslant 0$ in that special case. Finally, strict positivity of $\Delta(t)$ in some interval implies that I_ν is strictly positive. Since the integral is

$$I_\nu = \begin{cases} \dfrac{1}{\nu(1-\nu)} \left(\displaystyle\sum_{i=1}^{n} x_i^\nu - y_i^\nu \right) \displaystyle\sum_\ell c_\ell^\nu & \nu \neq 0 \\ \left(\ln \prod x_i \Big/ \prod y_i \right) \left(\displaystyle\sum_\ell 1 \right) & \nu = 0 \end{cases} \tag{31}$$

investigating $\nu < 0$, $\nu = 0$ and $\nu > 0$ cases separately, it can be seen that (T1) are satisfied.

The proof of sufficiency of the inequalities (T1) for the super-trumping relation will be completed in three steps, where in each step a special case is investigated. The last two steps will rely on the proof completed in the previous steps. The first case deals with a very special situation where both vectors can be expressed as integer powers of a common number.

Case A. y has strictly positive elements such that $y_i = K\omega^{\alpha_i}$ and $x_i = K\omega^{\beta_i}$ for some integers α_i and β_i and for some numbers $K > 0$ and $\omega > 1$, respectively.

Proof. Without loss of generality, it is assumed that x and y have no common elements and they are arranged in a non-decreasing order, i.e. $x = x^\uparrow$ and $y = y^\uparrow$. Redefine K such that $\alpha_1 = 0$ (as a result, $\alpha_i \geqslant 0$ for all i) and then set $K = 1$ by dividing each vector by a common number. Note that the $\nu \to -\infty$ limit of (T1) gives $x_1 \geqslant y_1$. As x and y have no common elements, we have $\beta_i > 0$ for all i. Let the polynomial $\Gamma(s)$ be defined as

$$\Gamma(s) = \sum_{i=1}^{n} (s^{\alpha_i} - s^{\beta_i}) = \sum_k \Gamma_k s^k, \tag{32}$$

and let $\gamma(s) = \Gamma(s)/(1-s)$. Since $\Gamma(1) = 0$, $\gamma(s)$ is also a polynomial. We will first show that γ is strictly positive for $s \in [0, \omega]$. The inequality (T1) at $\nu = 0$ implies that $\gamma(1) = \sum_{i=1}^{n} (\beta_i - \alpha_i)$ is strictly positive. Next, let $s = \omega^\nu$ where ν is any value in $(-\infty, 1]$ excluding $\nu = 0$. In that case, we have

$$\gamma(s) = \frac{1}{1-\omega^\nu} \sum_{i=1}^{n} \left(y_i^\nu - x_i^\nu \right). \tag{33}$$

Investigating the cases $\nu < 0$ and $\nu > 0$ separately, one finds that $\gamma(s) > 0$. Finally, $\gamma(0) = \Gamma(0) > 0$.

By lemma 2 of the previous section, there exists two (possibly infinite) series $a(s)$ and $b(s)$ which are convergent at $s = \omega$ and have non-negative series coefficients. Moreover, $a(s)$ can be chosen in such a way that all of its coefficients and $a(1)$ are rational numbers. As $\gamma(0) > 0$, a_0 and b_0 can be chosen non-zero. The relationship $a(s)\Gamma(s) = (1-s)b(s)$ implies that $\sum_{k=0}^{m} a_k \Gamma_{m-k} = b_m - b_{m-1}$, where we define $b_{-1} = 0$ for simplicity.

Let $\bar{h}(t) = \sum_{m=0}^{\infty} a_m (t - \omega^m)^+$, a function which is a sum of a finite number of terms for any fixed t. Let

$$\begin{aligned} \bar{\delta}(t) &= \sum_{i=1}^{n} y_i \bar{h}\left(\frac{t}{y_i}\right) - x_i \bar{h}\left(\frac{t}{x_i}\right) = \sum_k \Gamma_k \omega^k \bar{h}(t\omega^{-k}), \\ &= \sum_{m=0}^{\infty} (b_m - b_{m-1})(t - \omega^m)^+. \end{aligned} \tag{34}$$

It can be shown that $\bar{\delta}(t) \geqslant 0$ for all $t \geqslant 0$, but better lower bounds can be placed as follows (i) For $t \leqslant \omega$, we have $\bar{\delta}(t) = b_0(t-1)^+ \geqslant 0$. (ii) For $t \geqslant \omega$, there is an integer $N \geqslant 1$ such that $\omega^N \leqslant t \leqslant \omega^{N+1}$, and we have

$$\bar{\delta}(t) = b_N(t-\omega^N) + (\omega - 1)\sum_{m=0}^{N-1} b_m\omega^m \geqslant (\omega-1)b_0, \tag{35}$$

i.e. a strictly positive lower bound.

The basic idea in here is that the catalyst vector c should be constructed from the powers of ω such that a_m is the 'relative frequency' of ω^m. In that case, $\bar{h}(t)$ is the (unnormalized) characteristic function of that vector. However, as $a(s)$ is a possibly infinite series, this does not define a valid catalyst with a finite number of Schmidt coefficients. For that reason, the series $a(s)$ should be somehow terminated for finding a valid catalyst. The procedure for such a truncation is detailed below.

Let $\epsilon = (\omega-1)b_0\big/\left(\sum_k |\Gamma_k|\omega^k\right)$. Since $a(\omega) < \infty$, we can find an integer M ($\geqslant 1$) such that $\sum_{m=M}^{\infty} a_m\omega^m < \epsilon/2$. Define $A = \sum_{m=M}^{\infty} a_m$. This is a rational number and satisfies the inequality $A\omega^M < \epsilon/2$. Consider the function

$$h(t) = \sum_{m=0}^{M-1} a_m(t-\omega^m)^+ + A(t-\omega^M)^+. \tag{36}$$

The following bounds can be placed on $|\bar{h}(t) - h(t)|$: (i) if $t \leqslant \omega^M$, we have $h(t) = \bar{h}(t)$; (ii) if $t \geqslant \omega^M$, there is an $N \geqslant M$ such that $\omega^N \leqslant t \leqslant \omega^{N+1}$ and

$$\begin{aligned} \left|\bar{h}(t) - h(t)\right| &= \left| A\omega^M - \sum_{m=M}^{N} a_m\omega^m - \sum_{m=N+1}^{\infty} a_m t \right| \\ &\leqslant A\omega^M + \sum_{m=M}^{\infty} a_m\omega^m < \epsilon. \end{aligned} \tag{37}$$

As a result, the following function

$$\begin{aligned} \delta(t) &= \sum_{i=1}^{n} y_i h\left(\frac{t}{y_i}\right) - x_i h\left(\frac{t}{x_i}\right) = \sum_k \Gamma_k\omega^k h(t\omega^{-k}), \\ &= \bar{\delta}(t) + \sum_k \Gamma_k\omega^k(h(t\omega^{-k}) - \bar{h}(t\omega^{-k})) \end{aligned} \tag{38}$$

is non-negative everywhere since (i) for $t \leqslant \omega$ we have $\delta(t) = \bar{\delta}(t) \geqslant 0$ and (ii) for $t \geqslant \omega$ we have $\delta(t) > \bar{\delta}(t) - \sum_k |\Gamma_k|\omega^k\epsilon = \bar{\delta}(t) - (\omega-1)b_0 \geqslant 0$.

Let $\mathcal{N}$ be a sufficiently large integer so that all of $\mathcal{N}a_0, \mathcal{N}a_1, \ldots, \mathcal{N}a_{M-1}, \mathcal{N}A$ are integers. Schmidt coefficients of the catalyst vector c will be chosen as ω^m, repeated $\mathcal{N}a_m$ times (for $0 \leqslant m \leqslant M-1$), and as ω^M, repeated $\mathcal{N}A$ times. Then $H_c(t) = \mathcal{N}h(t)$ is the characteristic function of c and the non-negativity of $\delta(t)$ is equivalent to $x \otimes c \prec^w y \otimes c$. This proves our assertion that $x \prec_T^w y$. □

Case B. y has strictly positive elements.

Proof. Without loss of generality assume that x and y have no common elements, in which case the inequalities (T1) imply that $x_1^\uparrow > y_1^\uparrow$. Let, $\theta = \min_{\nu\in[-\infty,1]} A_\nu(x)/A_\nu(y)$. Since the end points are included, the minimum exists and therefore $\theta > 1$. Let $\omega = \theta^{1/3}$ and define two n-element vectors $\bar{x}$ and $\bar{y}$ as $\bar{y}_i = \omega^{\alpha_i}$ and $\bar{x}_i = \omega^{\beta_i}$, respectively, where

$$\alpha_i = \left]\frac{\ln y_i}{\ln \omega}\right[, \qquad \beta_i = \left[\frac{\ln x_i}{\ln \omega}\right], \tag{39}$$

$[t]$ is the largest integer smaller than t and $]t[$ is the smallest integer greater than t. Using $]t[-1 < t \leqslant]t[$ and $[t] \leqslant t < [t]+1$, we get

$$\frac{\bar{y}_i}{\omega} < y_i \leqslant \bar{y}_i, \qquad \bar{x}_i \leqslant x_i < \omega \bar{x}_i. \tag{40}$$

Then for any $\nu \in [-\infty, 1]$, we have

$$\frac{A_\nu(\bar{x})}{A_\nu(\bar{y})} > \frac{1}{\omega^2}\frac{A_\nu(x)}{A_\nu(y)} \geqslant \omega > 1. \tag{41}$$

As a result, $\bar{x}$ and $\bar{y}$ fulfills the conditions of case A, and therefore $\bar{x} \prec_T^w \bar{y}$. Finally, the inequalities (40) imply $x \prec^w \bar{x}$ and $\bar{y} \prec^w y$. All of these prove our assertion that $x \prec_T^w y$. □

Case C. y has zero elements.

The proof will be carried out by replacing all zero elements of y with a small value ϵ in such a way that this case is reduced to case B. Suppose that $y = y^\uparrow$ and it has exactly m entries equal to 0 ($0 < m < n$), i.e. $y_1 = \cdots = y_m = 0 < y_{m+1} \leqslant \cdots \leqslant y_n$. Note that the inequalities (T1) are automatically satisfied for $\nu \leqslant 0$. Using the premise that (T1) are satisfied for $\nu \in (0, 1]$, we can deduce that the function

$$J_\nu = \left(\frac{\sum_{i=1}^n x_i^\nu - \sum_{i=m+1}^n y_i^\nu}{m}\right)^{\frac{1}{\nu}} \tag{42}$$

is strictly positive for all $\nu \in (0, 1]$. Moreover, it has a positive limit $J_0 = \left(\prod x / \prod_{i=m+1}^n y_i\right)^{1/m}$ at the end point $\nu = 0$. As a result, $J_{\min} = \min_{\nu\in[0,1]} J_\nu$ exists and is non-zero as the minimum is taken over a compact interval. Let ϵ be a positive number such that

$$\epsilon < \min\left(J_{\min}, y_n\left(\frac{x_1}{y_n}\right)^{\frac{n}{m}}\right), \tag{43}$$

and define a new vector $\bar{y}$ as $\bar{y}_1 = \cdots = \bar{y}_m = \epsilon$ and $\bar{y}_i = y_i$ for all $i > m$. It is obvious that $\bar{y} \prec^w y$. Showing that $x \prec_T^w \bar{y}$ will complete the proof. For this purpose, we look at the power means. (i) For $\nu \in (0, 1]$, it is trivial to check that $J_\nu > \epsilon$ is equivalent to $A_\nu(x) > A_\nu(\bar{y})$. (ii) For $\nu = 0$, we have

$$\frac{A_0(x)}{A_0(\bar{y})} = \frac{\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}}{\left(\epsilon^m \prod_{i=m+1}^n y_i\right)^{\frac{1}{n}}} \geqslant \frac{x_1}{y_n}\left(\frac{y_n}{\epsilon}\right)^{\frac{m}{n}} > 1. \tag{44}$$

(iii) For $\nu < 0$, we use Bernoulli's inequality, which states that $\alpha^r - 1 \geqslant r(\alpha - 1)$ for all $r \geqslant 1$ and $\alpha > 0$, as follows:

$$m\left(\epsilon^\nu - y_n^\nu\right) > m y_n^\nu \left(\left(\frac{x_1}{y_n}\right)^{\nu\frac{n}{m}} - 1\right) \tag{45}$$

$$\geqslant m y_n^\nu \frac{n}{m}\left(\left(\frac{x_1}{y_n}\right)^{\nu} - 1\right) \tag{46}$$

$$= n\left(x_1^\nu - y_n^\nu\right), \tag{47}$$

which implies that

$$\sum_{i=1}^{n} \bar{y}_i^{\nu} = m\epsilon^{\nu} + \sum_{i=m+1}^{n} y_i^{\nu} \geqslant m\epsilon^{\nu} + (n-m)y_n^{\nu} \tag{48}$$

$$> nx_1^{\nu} \geqslant \sum_{i=1}^{n} x_i^{\nu}. \tag{49}$$

The result $A_\nu(x) > A_\nu(\bar{y})$ follows from here. As power mean inequalities are satisfied for all $\nu \in (-\infty, 1]$, we have $x \prec_T^w \bar{y}$ by the result in case B, which completes the proof. □

5. Sub-trumping relation

This section contains a proof of theorem 3, which gives the necessary and sufficient conditions for the sub-trumping relation. The proof is quite similar but simpler than the proof of theorem 2, as zero entries do not create any difficulty in here.

Proof of necessity. Let x and y be vectors satisfying the conditions of theorem 3 such that $x \prec_{wT} y$. It will be shown that the strict inequalities (T2) are satisfied. Since there is a catalyst c with positive entries such that $x \otimes c \prec_w y \otimes c$, the function

$$\Delta'(t) = H'_{y\otimes c}(t) - H'_{x\otimes c}(t) \tag{50}$$

is non-negative. Moreover, $\Delta'(t) = 0$ for all $t \geqslant y_1^{\downarrow}$. At $t = 0$, it has the strictly positive limit $\Delta'(0) = \left(\sum y_i - \sum x_i\right) \sum c_\ell$. For that reason, the integral

$$I_\nu = \int_0^\infty \Delta'(t) t^{\nu-2}\, \mathrm{d}t \tag{51}$$

is convergent for all $\nu > 1$ and is strictly positive. Since the integral is

$$I_\nu = \frac{1}{\nu(\nu-1)} \left(\sum_{i=1}^{n} y_i^{\nu} - x_i^{\nu} \right) \sum_\ell c_\ell^{\nu}, \tag{52}$$

all inequalities in (T2) follow. □

Proof of sufficiency. This will be done in two steps, the first one dealing with the special situation where the non-zero elements of both vectors can be expressed as integer powers of a common number.

Case A. All non-zero entries of x and y are such that $y_i = K\omega^{\alpha_i}$ and $x_i = K\omega^{\beta_i}$ for some integers α_i and β_i and for some numbers $K > 0$ and $\omega < 1$ respectively.

Proof. Either vector might have a zero element, but this will not cause any difficulty. Without loss of generality, it is assumed that x and y have no common elements and they are arranged in a non-increasing order, i.e. $x = x^{\downarrow}$ and $y = y^{\downarrow}$. Redefine K such that $\alpha_1 = 0$ (as a result, $\alpha_i \geqslant 0$ for all i) and then set $K = 1$ by dividing each vector by a common number. Note that the $\nu \to +\infty$ limit of (T1) gives $x_1 < y_1$, and therefore we have $\beta_i > 0$ for all i. Let the polynomial $\Gamma(s)$ be defined as

$$\Gamma(s) = {\sum_i}' s^{\alpha_i} - {\sum_i}' s^{\beta_i} = \sum_k \Gamma_k s^k, \tag{53}$$

where the primes denote that the sum over i should be done only for non-zero entries. Note that $\Gamma(0) > 0$ and all inequalities in (T2) are equivalent to $\Gamma(s) > 0$ for $s \in (0, \omega]$.

By lemma 2 (b) of section 3, we can find two power series $a(s)$ and $b(s)$ with non-negative coefficients such that $a(s)\Gamma(s) = b(s)$, both series are convergent at $s = \omega$, both have non-zero constant terms and $a(s)$ has rational coefficients. Using the coefficients of $a(s)$, the following function is defined:

$$\bar{h}'(t) = \sum_{m=0}^{\infty} a_m(\omega^m - t)^+. \tag{54}$$

Note that, for any finite t, the expression above consists of a sum of a finite number of positive terms. Let

$$\bar{\delta}'(t) = \sum_i{}' y_i \bar{h}'\left(\frac{t}{y_i}\right) - \sum_i{}' x_i \bar{h}'\left(\frac{t}{x_i}\right), \tag{55}$$

$$= \sum_k \Gamma_k \omega^k \bar{h}'(t\omega^{-k}), \tag{56}$$

$$= \sum_{m=0}^{\infty} b_m(\omega^m - t)^+. \tag{57}$$

It is obvious that $\bar{\delta}'(t)$ is non-negative everywhere. For our purposes, it is better to use the strictly positive lower bound $\bar{\delta}'(t) \geqslant b_0(1-\omega)$ for $t \leqslant \omega$.

To obtain a valid catalyst, the series $a(s)$ can be truncated as follows. Let $\epsilon = b_0(1-\omega)\big/\left(\sum_k |\Gamma_k|\omega^k\right)$. Since $a(\omega) < \infty$, we can find an integer $M(\geqslant 1)$ such that $\sum_{m=M}^{\infty} a_m\omega^m < \epsilon$. Consider the function

$$h'(t) = \sum_{m=0}^{M-1} a_m(\omega^m - t)^+. \tag{58}$$

It is obvious that $h'(t) \leqslant \bar{h}'(t)$. For $t \geqslant \omega^M$, these functions are equal. For any $t \leqslant \omega^M$, we can find an integer $N(\geqslant M)$ such that $\omega^{N+1} < t \leqslant \omega^N$, for which case we have

$$\bar{h}'(t) - h'(t) = \sum_{m=M}^{N} a_m(\omega^m - t) \tag{59}$$

$$< \sum_{m=M}^{N} a_m\omega^m \leqslant \epsilon. \tag{60}$$

In other words, these two functions do not deviate much from each other.

Finally, consider the following function:

$$\delta'(t) = \sum_i{}' y_i h'\left(\frac{t}{y_i}\right) - \sum_i{}' x_i h'\left(\frac{t}{x_i}\right),$$

$$= \sum_k \Gamma_k \omega^k h'(t\omega^{-k}), \tag{61}$$

$$= \bar{\delta}'(t) - \sum_k \Gamma_k \omega^k (\bar{h}'(t\omega^{-k}) - h'(t\omega^{-k})). \tag{62}$$

This function is non-negative everywhere since (i) for $t \geqslant \omega$ we have $\delta'(t) = \bar{\delta}'(t) \geqslant 0$ and (ii) for $t \leqslant \omega$ we have $\delta'(t) \geqslant \bar{\delta}'(t) - \epsilon \sum_k |\Gamma_k|\omega^k \geqslant 0$.

Let $\mathcal{N}$ be a sufficiently large integer so that all of $\mathcal{N}a_0, \mathcal{N}a_1, \ldots, \mathcal{N}a_{M-1}$ are integers. Schmidt coefficients of the catalyst vector c will be chosen as ω^m, repeated $\mathcal{N}a_m$ times. Then $H_c'(t) = \mathcal{N}h'(t)$ and the non-negativity of $\delta'(t)$ is equivalent to $x \otimes c \prec_w y \otimes c$. This proves our assertion that $x \prec_{wT} y$. □

Case B. The general case.

Proof. Without loss of generality assume that x and y have no common elements, in which case the inequalities (T2) imply that $x_1^\downarrow < y_1^\downarrow$. Let $\theta = \max_{\nu\in[1,+\infty]} A_\nu(x)/A_\nu(y)$. Since the end points are included, the minimum exists and therefore $\theta < 1$. Let $\omega = \theta^{1/3}$ and define the following integers for all non-zero entries of x and y:

$$\alpha_i = \left]\frac{\ln y_i}{\ln\omega}\right[, \qquad \beta_i = \left[\frac{\ln x_i}{\ln\omega}\right]. \tag{63}$$

Two n-element vectors $\bar{x}$ and $\bar{y}$ will be defined as $\bar{y}_i = \omega^{\alpha_i}$ and $\bar{x}_i = \omega^{\beta_i}$, respectively. If a particular element of original vectors is zero, say y_i, then the corresponding element $\bar{y}_i$ will be chosen as zero. A similar choice will be made for x and $\bar{x}$. In that case, the following inequalities hold for each element:

$$\omega y_i \leqslant \bar{y}_i \leqslant y_i, \qquad x_i \leqslant \bar{x}_i \leqslant \frac{1}{\omega} x_i. \tag{64}$$

From these inequalities, it can be shown that $A_\nu(\bar{x}) \leqslant \omega A_\nu(\bar{y}) < A_\nu(\bar{y})$ for all $\nu \in [1, +\infty]$ and the proof given in case A enables us to conclude that $\bar{x} \prec_{wT} \bar{y}$. Finally, the inequalities above imply that $x \prec_w \bar{x}$ and $\bar{y} \prec_w y$ which prove the assertion that $x \prec_{wT} y$. □

6. Trumping relation

This section contains the proof of theorem 1, which gives all necessary and sufficient conditions for the trumping relation. The main difficulty in this case is the requirement that the two vectors that are to be related need to have the same sum. As a result, in order to surmount this particular difficulty, details of the proofs given in this section become more complicated.

Proof of the necessity of conditions (T1)–(T3) for the trumping relation is trivial. Given that there is a catalyst c so that we have $x \otimes c \prec y \otimes c$, the strict inequality of (17) can be used for the following strictly convex functions: $f(t) = t^\nu$ for $\nu > 1$ and $\nu < 0$, $f(t) = -t^\nu$ for $0 < \nu < 1$, $f(t) = -\ln t$ and $f(t) = t\ln t$. All inequalities (T1)–(T3) follow from these.

The proof of sufficiency will follow along the same lines as the proof of the super-trumping relation given in section 4, where three special cases will be considered separately.

Case A. y has strictly positive elements such that $y_i = K\omega^{\alpha_i}$ and $x_i = K\omega^{\beta_i}$ for some integers α_i and β_i and for some numbers $K > 0$ and $\omega > 1$ respectively.

Proof. Without loss of generality, it is assumed that x and y have no common elements and they are arranged in a non-decreasing order. The smallest of the exponents is α_1 which can be set equal to 0 by a redefinition of K. Finally, both x and y can be divided by K which is equivalent to setting $K = 1$. As a result, in here, it is not required that the vectors are normalized (i.e. they do not add up to 1). Since $\alpha_1 = 0$, all other exponents satisfy $\alpha_i \geqslant 0$ and $\beta_i > 0$.

Let the polynomial $\Gamma(s)$ be defined as

$$\Gamma(s) = \sum_{i=1}^{n}(s^{\alpha_i} - s^{\beta_i}) = \sum_k \Gamma_k s^k. \tag{65}$$

First, note that $\Gamma(s)$ has simple roots at $s = 1$ and $s = \omega$. This can be simply seen by evaluating its derivative at these points,

$$\Gamma'(1) = \sum_{i=1}^{n} (\alpha_i - \beta_i) < 0, \tag{66}$$

$$\Gamma'(\omega) = \frac{\sigma(x) - \sigma(y)}{\ln \omega} > 0, \tag{67}$$

where the former strict inequality follows from (T1) at $\nu = 0$ and the latter follows from (T3). The fact that x and y are not normalized does not invalidate the latter inequality.

Therefore, $\gamma(s) = \Gamma(s)/((1-s)(1-s/\omega))$ is a polynomial. It can be seen that $\gamma(0) = \Gamma(0) > 0$. Moreover, we will show that $\gamma(s)$ has no positive root. For this purpose, let $s = \omega^\nu$ where ν is any real number ($\nu = 0$ and $\nu = 1$ can be excluded if desired). Then

$$\gamma(\omega^\nu) = \frac{1}{(1-\omega^\nu)(1-\omega^{\nu-1})} \sum_{i=1}^{n} \left(y_i^\nu - x_i^\nu\right), \tag{68}$$

which can be seen to be strictly positive by virtue of (T1) and (T2) for all values of ν. (For $\nu = 0$ and $\nu = 1$, we have seen above that $\gamma(s)$ has no root at 1 and ω).

By lemma 1 of section 3, there are two polynomials $a(s)$ and $b(s)$ with non-negative coefficients such that $a(s)\gamma(s) = b(s)$ and $a(s)$ has integral coefficients. The constant coefficients $a(0)$ and $b(0)$ will also be chosen to be non-zero. In terms of Γ, the relation can be expressed as

$$a(s)\Gamma(s) = (1-s)(1-s/\omega)b(s). \tag{69}$$

Let $a(s)$ has degree N. The catalyst vector c will be chosen from the numbers ω^k which are repeated a_k times ($k = 0, 1, \ldots, N$). In that case, the characteristic function of c is

$$H_c(t) = \sum_{k=0}^{N} a_k (t - \omega^k)^+. \tag{70}$$

We would like to show that the function

$$\Delta(t) = H_{y\otimes c}(t) - H_{x\otimes c}(t) \tag{71}$$

$$= \sum_{i=1}^{n} y_i H_c(t/y_i) - x_i H_c(t/x_i) \tag{72}$$

$$= \sum_{\ell} \Gamma_\ell \omega^\ell H_c(t\omega^{-\ell}) \tag{73}$$

$$= \sum_{k,\ell} a_k \Gamma_\ell (t - \omega^{k+\ell})^+ \tag{74}$$

is non-negative for all $t \geqslant 0$. First, note that

$$\Delta(t) = \sum_{m=0}^{M+1} (f_m - f_{m-1})(t - \omega^m)^+, \tag{75}$$

where $f(s) = (1 - s/\omega)b(s)$, f_m are the coefficients of the polynomial $f(s)$ and we have chosen $f_{-1} = 0$ for simplicity. Here, M is the degree of f ($M + 1$ is the degree of $a(s)\Gamma(s)$). Since $\Delta(t)$ is a piecewise linear function, for showing its positivity, it is sufficient to look at its value at the turning points and at the $t = 0$ and $t = \infty$ limits. First, note that $\Delta(t) = 0$ for $t \leqslant 1$ and $\Delta(t)$ is constant for $t \geqslant \omega^{M+1}$. As a result, we only need to check the values of $\Delta(t)$ at $t = \omega, \omega^2, \ldots, \omega^{M+1}$. For any $1 \leqslant k \leqslant M+1$,

$$\Delta(\omega^k) = \sum_{m=0}^{k-1}(f_m - f_{m-1})(\omega^k - \omega^m) \tag{76}$$

$$= (\omega - 1)\sum_{m=0}^{k-1}(f_m - f_{m-1})\sum_{p=m}^{k-1}\omega^p \tag{77}$$

$$= (\omega - 1)\sum_{p=0}^{k-1}\omega^p\sum_{m=0}^{p}(f_m - f_{m-1}) \tag{78}$$

$$= (\omega - 1)\sum_{p=0}^{k-1} f_p\omega^p. \tag{79}$$

Finally, as $f(s) = (1 - s/\omega)b(s)$, the coefficients of these polynomials satisfy

$$f_p = b_p - \frac{b_{p-1}}{\omega}, \tag{80}$$

where $b_{-1} = 0$, which leads to

$$\Delta(\omega^k) = (\omega - 1)b_{k-1}\omega^{k-1} \geqslant 0. \tag{81}$$

This completes the proof of $\Delta(t) \geqslant 0$ for all $t \geqslant 0$. It also shows that $x \otimes c \prec y \otimes c$. Therefore, $x \prec_T y$. □

Before passing on to the next case, it is necessary to state another theorem that shows the stability of the inequalities (T1)–(T3) against small variations in x and y when these have different minimum and maximum entries. Although it is possible to generalize to vectors which have zero entries, the stability will be shown only for vectors having non-zero elements. For this purpose, it will also be appropriate to measure the distance between two vectors by the deviation of the ratio of the corresponding elements from 1. For two vectors x and $\bar{x}$ which has no zero elements, the distance between them is defined as

$$D(x; \bar{x}) = \max_i \left| \ln \frac{x_i}{\bar{x}_i} \right|. \tag{82}$$

The following theorem expresses the stability of conditions (T1)–(T3).

Theorem 4. *Let x and y be n-element vectors formed from positive numbers such that $x_1^{\uparrow} > y_1^{\uparrow}$ and $x_n^{\uparrow} < y_n^{\uparrow}$. If x and y satisfy the inequalities (T1)–(T3), then there is a positive number ϵ such that whenever $D(x; \bar{x}) \leqslant \epsilon$ and $D(y; \bar{y}) \leqslant \epsilon$, and $\sum \bar{x}_i = \sum \bar{y}_i = \sum x_i$, the vectors $\bar{x}$ and $\bar{y}$ satisfy the same strict inequalities.*

The proof of theorem 4 is postponed to the appendix. This result will be used in the proof of the next case.

Case B. y has strictly positive elements.

Before proceeding with the proof, it will first be argued that giving the proof for the special case where y has a single maximum element is sufficient. For this purpose, consider the case where y has more than one maximum element. Let $y = y^{\uparrow}$ and suppose that the last

$k+1$ entries of y are equal ($k \geqslant 1$), i.e. $y_{n-k-1} < y_{n-k} = \cdots = y_n$. The inequalities (T1) and (T2) imply that $k < n-1$, i.e. y is not a uniform vector. Consider y^ϵ, a vector which is also a function of a parameter ϵ, which is defined as follows:

$$y_i^\epsilon = \begin{cases} y_i & i < n-k-1 \\ y_{n-k+1} + k\epsilon & i = n-k-1 \\ y_n - \epsilon & n-k \leqslant i \leqslant n-1 \\ y_n & i = n. \end{cases} \tag{83}$$

It can be seen that for all $0 < \epsilon \leqslant (y_n - y_{n-k-1})/(k+1)$, we have $y^\epsilon \prec y$ and y^ϵ has a single maximum entry. Moreover, y^ϵ converges to y as ϵ goes to zero. Theorem 4 then ensures that there is a sufficiently small non-zero ϵ such that x and y^ϵ satisfy all inequalities (T1)–(T3). As a result, proving that $x \prec_T y^\epsilon$ will also show that $x \prec_T y$.

As a result, it can be assumed without loss of generality that y has a single maximum entry. It will also be assumed that x and y are normalized $\left(\sum x_i = \sum y_i = 1\right)$ and they are arranged in a non-decreasing order ($x = x^\uparrow$, $y = y^\uparrow$). The proof will be carried out by choosing two new vectors $\bar{x}$ and $\bar{y}$ which are sufficiently near to x and y respectively such that theorem 4 can be invoked, and it will be made sure that $\bar{x}$ and $\bar{y}$ satisfy the conditions considered in case A. Let $H = \sigma(x) - \sigma(y) > 0$ be the entropy difference of these vectors and let $L = \left|\ln y_1^\uparrow\right|$. Note that the logarithm of all elements are bounded by L, i.e. $|\ln y_i| \leqslant L$ and $|\ln x_i| \leqslant L$. Let ϵ_0 be a positive number such that whenever $D(x,\bar{x}) \leqslant \epsilon_0$ and $D(y,\bar{y}) \leqslant \epsilon_0$, the vectors $\bar{x}$ and $\bar{y}$ satisfy all the inequalities in (T1)–(T3). A positive number ϵ is chosen such that

$$\epsilon < \min\left(\frac{\epsilon_0}{2}, \frac{1}{8n}, \frac{1}{n^2}, \frac{H}{96nL}, \frac{1}{2}\ln\left(\frac{y_n}{y_{n-1}}\right)\right). \tag{84}$$

It necessarily follows that $\epsilon < L$, an inequality that will be used below.

We will define α_i and β_i to be some rational approximations to numbers $\ln y_i$ and $\ln x_i$, respectively. Let ϕ_i and θ_i represent the deviation of these rational approximations from the true values,

$$\alpha_i = \ln y_i + \phi_i, \tag{85}$$

$$\beta_i = \ln x_i + \theta_i. \tag{86}$$

As the rational numbers are dense, these deviations can be chosen essentially arbitrarily. But, for our purposes, they are going to be chosen as

$$\frac{\epsilon}{2n} \leqslant \phi_i \leqslant \frac{\epsilon}{n} \qquad \text{for} \quad 1 \leqslant i \leqslant n-1, \tag{87}$$

$$\left|\sum_{i=1}^{n} y_i \phi_i\right| \leqslant \epsilon^2. \tag{88}$$

In other words, the rational approximations α_i for all elements except the last one are to be chosen such that the corresponding deviations ϕ_i are positive and small, but they are also required to be sufficiently far away from zero. The last element is an exception. In that case, α_n has to be chosen as a rational number so that this time the sum in (88) is made very small. In that case, ϕ_n does not need to be positive. Note that conditions (87) and (88) provide n separate intervals to choose α_i from. As rational numbers are dense, all of α_i can be chosen as rational numbers. Similarly, we define β_i and the corresponding deviations θ_i such that

$$-\frac{\epsilon}{n} \leqslant \theta_i \leqslant -\frac{\epsilon}{2n} \qquad \text{for} \quad 1 \leqslant i \leqslant n-1, \tag{89}$$

$$\left|\sum_{i=1}^{n} x_i\theta_i\right| \leqslant \epsilon^2, \tag{90}$$

where the deviations for the first $n-1$ elements are chosen this time to be negative. Similar comments apply in here.

Below, however, we will need a uniform bound on all of the deviations. For this purpose, note the following bound on ϕ_n:

$$y_n|\phi_n| \leqslant \epsilon^2 + \sum_{i=1}^{n-1} y_i|\phi_i| \leqslant \epsilon^2 + (1-y_n)\frac{\epsilon}{n} \tag{91}$$

$$|\phi_n| \leqslant \frac{\epsilon^2}{y_n} + \left(\frac{1}{y_n} - 1\right)\frac{\epsilon}{n} \tag{92}$$

$$\leqslant n\epsilon^2 + (n-1)\frac{\epsilon}{n} \leqslant \epsilon, \tag{93}$$

where we have used the fact that $y_n \geqslant 1/n$ for the maximum element of y. Therefore, the following uniform bounds can be placed on all deviations:

$$|\phi_i| \leqslant \epsilon, \quad |\theta_i| \leqslant \epsilon \qquad \text{for} \quad i = 1, 2, \ldots, n, \tag{94}$$

where the bounds on θ_i follow by a similar analysis. For most of the following, we will use these uniform bounds. The stricter bounds given in (87) and (89) will only be necessary at the very end. The following bounds on the rational approximations will be occasionally used: $|\alpha_i| \leqslant |\ln y_i| + |\phi_i| \leqslant L + \epsilon \leqslant 2L$ and similarly $|\beta_i| \leqslant 2L$.

Consider the following function:

$$F(\lambda) = \sum_{i=1}^{n} (\mathrm{e}^{\lambda\alpha_i} - \mathrm{e}^{\lambda\beta_i}). \tag{95}$$

Our first job is to establish that this function has a root near 1, i.e. there is a number λ_0, which is very close to 1 such that $F(\lambda_0) = 0$. Once this problem is solved, the two new vectors $\bar{x}$ and $\bar{y}$ can be defined as

$$\bar{x}_i = \frac{\mathrm{e}^{\lambda_0\beta_i}}{Z_0}, \tag{96}$$

$$\bar{y}_i = \frac{\mathrm{e}^{\lambda_0\alpha_i}}{Z_0}, \tag{97}$$

where $Z_0 = \sum_{i=1}^{n} \mathrm{e}^{\lambda_0\alpha_i} = \sum_{i=1}^{n} \mathrm{e}^{\lambda_0\beta_i}$. In that case, both $\bar{x}$ and $\bar{y}$ are normalized vectors. However, in order to reach to the final conclusion, we also need to place bounds on the deviation of both λ_0 and Z_0 from 1. Therefore, the following analysis of bounds is needed.

First, we must show that $F(\lambda)$ has a root somewhere near 1. For this purpose, we look at the value of $F(1)$. By using the following inequalities satisfied by the exponential function, $1 + t \leqslant e^t \leqslant 1 + t + t^2$ for all $|t| \leqslant 1$, the following bounds can be placed on the first term of $F(1)$:

$$\sum_{i=1}^{n} \mathrm{e}^{\alpha_i} = \sum_{i=1}^{n} y_i\,\mathrm{e}^{\phi_i} \tag{98}$$

$$\geqslant \sum_{i=1}^{n} y_i(1+\phi_i) \geqslant 1 - \epsilon^2, \tag{99}$$

$$\sum_{i=1}^{n} \mathrm{e}^{\alpha_i} \leqslant \sum_{i=1}^{n} y_i\left(1+\phi_i+\phi_i^2\right) \leqslant 1+2\epsilon^2. \tag{100}$$

The same bounds can also be placed for the second term as well, which lead to

$$|F(1)| \leqslant 3\epsilon^2, \tag{101}$$

a very small quantity, which indicates that a root is very close to 1.

However, to verify that there is a root around 1 and to place a bound on the deviation of the root from 1, we must make sure that the derivative $F'(\lambda)$ does not rapidly go to zero around $\lambda = 1$. For this purpose, a lower bound will be placed on the derivative for $|\lambda - 1| \leqslant \epsilon/L$. First, note that

$$\sum_{i=1}^{n} \alpha_i\, \mathrm{e}^{\lambda\alpha_i} = -\sigma(y) + \sum_{i=1}^{n} y_i\phi_i \tag{102}$$

$$+ \sum_{i=1}^{n} y_i\alpha_i\left(\mathrm{e}^{(\lambda-1)\ln y_i+\lambda\phi_i} - 1\right), \tag{103}$$

and the argument of the exponential is small as

$$|(\lambda-1)\ln y_i + \lambda\phi_i| \leqslant \frac{\epsilon}{L}L + \left(1+\frac{\epsilon}{L}\right)\epsilon \leqslant 3\epsilon. \tag{104}$$

Now, using $|e^t - 1| \leqslant |t| + t^2 \leqslant 2|t|$ for all $|t| \leqslant 1$, we can find the following lower bound on the expression above:

$$\sum_{i=1}^{n} \alpha_i\, \mathrm{e}^{\lambda\alpha_i} \geqslant -\sigma(y) - \epsilon^2 - 2L\cdot 6\epsilon \tag{105}$$

$$\geqslant -\sigma(y) - 13L\epsilon \tag{106}$$

A similar analysis for the second term of $F'(\lambda)$ gives

$$\sum_{i=1}^{n} \beta_i\, \mathrm{e}^{\lambda\beta_i} \leqslant -\sigma(x) + 13L\epsilon. \tag{107}$$

Both of these give the following lower bound on the derivative $F'(\lambda)$ for $|\lambda - 1| \leqslant \epsilon/L$:

$$F'(\lambda) \geqslant H - 26L\epsilon > \tfrac{1}{2}H. \tag{108}$$

By using the lower bound given above, it is possible to see that $F(1+\epsilon/L)$ is positive and $F(1-\epsilon/L)$ is negative. This guarantees the presence of the root in the specified interval. But, this interval is too large for our purposes, and we need to find a better bound on the place of the root. Using $F(\lambda_0) = 0$, we can get

$$-F(1) = \int_1^{\lambda_0} F'(\lambda)\, \mathrm{d}\lambda, \tag{109}$$

$$|F(1)| \geqslant |\lambda_0 - 1|\frac{H}{2}, \tag{110}$$

$$|\lambda_0 - 1| \leqslant \frac{2|F(1)|}{H} \leqslant \frac{6\epsilon^2}{H}. \tag{111}$$

In other words, the root is very close to the value 1.

One final bound, this time a bound on $\ln Z_0$, will be needed. For this, we first note that

$$Z_0 = \sum_{i=1}^{n} \mathrm{e}^{\alpha_i}\mathrm{e}^{(\lambda_0-1)\alpha_i} \leqslant \left(\sum_{i=1}^{n} \mathrm{e}^{\alpha_i}\right)\mathrm{e}^{+2|\lambda_0-1|L}, \tag{112}$$

and a similar analysis for the lower bound gives

$$|\ln Z_0| \leqslant \left|\ln\left(\sum_{i=1}^{n} \mathrm{e}^{\alpha_i}\right)\right| + 2|\lambda_0 - 1|L. \tag{113}$$

Finally, (99) and (100) give

$$\left|\ln\left(\sum_{i=1}^{n} \mathrm{e}^{\alpha_i}\right)\right| \leqslant 2\epsilon^2, \tag{114}$$

where we have used the fact that $t-1 \geqslant \ln t \geqslant (t-1)/t$. As a result, we get

$$|\ln Z_0| \leqslant \left(2 + \frac{12L}{H}\right)\epsilon^2. \tag{115}$$

Now, it is possible to show that the vectors $\bar{x}$ and $\bar{y}$ satisfy all the required properties to complete the proof. First, we will show that x is majorized by $\bar{x}$. For this reason, we will look at the ratio $x_i/\bar{x}_i$ for $i = 1, 2, \ldots, n-1$, i.e. for all elements except the last one. Here, we will make use of the upper bounds given in (89) as

$$\ln\frac{x_i}{\bar{x}_i} = -\theta_i + (1-\lambda_0)\beta_i + \ln Z_0 \tag{116}$$

$$\geqslant \frac{\epsilon}{2n} - \left(2 + \frac{24L}{H}\right)\epsilon^2 \geqslant 0, \tag{117}$$

where the last inequality can be obtained simply by inspecting (84). In other words, we have $x_i \geqslant \bar{x}_i$ for all $i < n$. The conclusion $x \prec \bar{x}$ then follows.

Showing that $\bar{y}$ is majorized by y is a little more involved. First, we note

$$\ln\frac{\bar{y}_i}{y_i} = \phi_i + (\lambda_0 - 1)\alpha_i - \ln Z_0 \tag{118}$$

$$\geqslant \frac{\epsilon}{2n} - \left(2 + \frac{24L}{H}\right)\epsilon^2 \geqslant 0; \tag{119}$$

in other words $\bar{y}_i \geqslant y_i$ for all $i < n$. Next, note that $\bar{y}_n$ is the maximum element of $\bar{y}$ as for any $i < n$, we have $\bar{y}_n/\bar{y}_i \geqslant \mathrm{e}^{-2\epsilon} y_n/y_{n-1} > 1$. Suppose that $\bar{y}$ can be put in a non-decreasing order as $\bar{y}_{i_1} \leqslant \bar{y}_{i_2} \leqslant \cdots \leqslant \bar{y}_{i_{n-1}} < \bar{y}_n$. As a result, for any $m < n$, we have

$$F_m(\bar{y}) = \bar{y}_{i_1} + \cdots + \bar{y}_{i_m} \geqslant y_{i_1} + \cdots + y_{i_m} \geqslant F_m(y), \tag{120}$$

which shows that $\bar{y} \prec y$.

Finally, we have

$$D(x;\bar{x}) = \max_i |\theta_i - (1-\lambda_0)\beta_i - \ln Z_0| \tag{121}$$

$$\leqslant \epsilon + \left(2 + \frac{24L}{H}\right)\epsilon^2 < \epsilon_0, \tag{122}$$

and similarly $D(y;\bar{y}) < \epsilon_0$. Therefore, the inequalities (T1)–(T3) are also satisfied by $\bar{x}$ and $\bar{y}$. It is easy to see that $\bar{x}$ and $\bar{y}$ satisfy the conditions of case A. The number ω is given as

$\exp(\lambda_0/\mathcal{N})$, where $\mathcal{N}$ is the common denominator of the rational numbers α_i and β_i. As a result, the conclusion $\bar{x} \prec_T \bar{y}$ follows. Combined with $x \prec \bar{x}$ and $\bar{y} \prec y$, it leads to the desired result $x \prec_T y$. □

Case C. y has zero components.

Without loss of generality, it is supposed that x and y are normalized, are arranged in a non-decreasing order and have no common elements. Let y have m zeros, i.e. $y_1 = y_2 = \cdots = y_m = 0 < y_{m+1} \leqslant \cdots \leqslant y_n$. Let z^ϵ be a vector defined as follows:

$$z_i^\epsilon = \epsilon \qquad \text{for} \quad i = 1, 2, \ldots, m, \tag{123}$$

$$z_i^\epsilon = (1 - m\epsilon) y_i \qquad \text{for} \quad i = m + 1, \ldots, n, \tag{124}$$

where ϵ is a non-negative parameter. We will only be interested in the values of ϵ in the range $0 \leqslant \epsilon \leqslant \left(y_{m+1}^{-1} + m\right)^{-1}$, where z^ϵ is arranged in an increasing order. It is easy to see that all such vectors are related to each other by the majorization relation, i.e. if $\epsilon_A > \epsilon_B$ then $z^{\epsilon_A} \prec z^{\epsilon_B}$. As $z^0 = y$, we have $z^\epsilon \prec y$ for all values of ϵ in the range considered.

Our job is to show that if ϵ is sufficiently small, then x and z^ϵ satisfy the inequalities (T1)–(T3). This is a straightforward but laborious procedure which is detailed below. For this purpose, different intervals of ν values will be considered separately and for each interval, the existence of a separate upper bound for ϵ will be provided.

(a) *For* $\nu \leqslant 0$. The quantity $\epsilon_1 = y_n (x_1/y_n)^{n/m}$ is a possible upper bound for this range. Let $\epsilon < \epsilon_1$. For the special case $\nu = 0$, we have

$$\frac{A_0(x)}{A_0(z^\epsilon)} = \left(\frac{\prod_{i=1}^n x_i}{\epsilon^m (1 - m\epsilon)^{n-m} \prod_{i=m+1}^n y_i} \right)^{\frac{1}{n}} \tag{125}$$

$$\geqslant \frac{x_1}{y_n} \left(\frac{y_n}{\epsilon} \right)^{\frac{m}{n}} > 1. \tag{126}$$

For all negative values of ν, we make use of Bernoulli's inequality again to reach $m\left(\epsilon^\nu - y_n^\nu\right) > n\left(x_1^\nu - y_n^\nu\right)$. This then leads to

$$\sum_{i=1}^n \left(z_i^\epsilon\right)^\nu = m\epsilon^\nu + (1 - m\epsilon)^\nu \sum_{i=m+1}^n y_i^\nu \tag{127}$$

$$> m\epsilon^\nu + (n - m) y_n^\nu > n x_1^\nu \geqslant \sum_{i=1}^n x_i^\nu. \tag{128}$$

As a result, we conclude that $A_\nu(x) > A_\nu(z^\epsilon)$ for all $\nu \leqslant 0$ whenever $\epsilon < \epsilon_1$.

(b) *For* $0 < \nu \leqslant 1/2$. The function

$$J_\nu = \left(\frac{\sum_{i=1}^n x_i^\nu - \sum_{i=m+1}^n y_i^\nu}{m} \right)^{\frac{1}{\nu}} \tag{129}$$

is strictly positive in the interval $(0, 1/2]$ and moreover it has a strictly positive limit at $\nu = 0$. Therefore, $\epsilon_2 = \min_{\nu \in [0,1/2]} J_\nu$ is a positive number. If $\epsilon < \epsilon_2$, we have

$$\sum_{i=1}^n x_i^\nu > m\epsilon^\nu + \sum_{i=m+1}^n y_i^\nu > \sum_{i=1}^n \left(z_i^\epsilon\right)^\nu, \tag{130}$$

which leads to $A_\nu(x) > A_\nu(z^\epsilon)$ in this interval.

(c) *For* $2 \leqslant \nu$. Let K be defined as

$$K = \max_{\nu\in[2,\infty]} \frac{A_\nu(x)}{A_\nu(y)}, \tag{131}$$

which is a positive number such that $K < 1$. Note that, as x and y have no common elements, the ratio above at $\nu = +\infty$ gives $x_n^\uparrow / y_n^\uparrow$ which is smaller than 1. Let $\epsilon_3 = (1-K)/m$. Then, for any $\epsilon < \epsilon_3$ and for all $\nu \geqslant 2$ we have

$$\sum_{i=1}^{n} \left(z_i^\epsilon\right)^\nu > (1-m\epsilon)^\nu \sum_{i=m+1}^{n} y_i^\nu \tag{132}$$

$$> K^\nu \sum_{i=m+1}^{n} y_i^\nu > \sum_{i=1}^{n} x_i^\nu. \tag{133}$$

This shows the desired inequality, $A_\nu(z^\epsilon) > A_\nu(x)$.

(d) *For* $1/2 \leqslant \nu \leqslant 2$. Let

$$\Delta R_\nu = \frac{1}{\nu-1} \ln \frac{A_\nu(y)}{A_\nu(x)}. \tag{134}$$

The inequalities (T1)–(T3) imply that ΔR_ν is a strictly positive continuous function in the interval considered. Therefore, the minimum $M = \min_{\nu\in[1/2,2]} \Delta R_\nu$ is a positive number. Let

$$\Delta R_\nu(\epsilon) = \frac{1}{\nu-1} \ln \frac{A_\nu(z^\epsilon)}{A_\nu(x)}. \tag{135}$$

Since all vectors z^ϵ are related to each other by the majorization relation, for any $\epsilon_A > \epsilon_B$ we have $\Delta R_\nu(\epsilon_A) \leqslant \Delta R_\nu(\epsilon_B)$ for all ν. In other words, as ϵ decreases, the function $\Delta R_\nu(\epsilon)$ monotonically increases. Finally, we note that $\Delta R_\nu(\epsilon)$ converges pointwise to ΔR_ν as ϵ goes to zero. At this point, we invoke Dini's theorem, which states that a sequence of monotonically increasing, continuous and pointwise convergent functions on a compact space are uniformly convergent. Therefore, there is a positive number ϵ_4 such that whenever $\epsilon < \epsilon_4$, we have $\Delta R_\nu(\epsilon) > M/2$.

For such values of ϵ, the inequalities (T1) and (T2) are satisfied for all $\nu \in [1/2, 2]$. Moreover, the inequality (T3) is also satisfied, as $\Delta R_1(\epsilon) = \sigma(x) - \sigma(z^\epsilon) > M/2 > 0$.

As a result, if $\epsilon < \min(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$, then the vectors x and z^ϵ satisfies all the inequalities (T1)–(T3). The proof of case B enables us to conclude that $x \prec_T z^\epsilon$. Finally, by $z^\epsilon \prec y$ we reach to the desired result $x \prec_T y$. □

7. Discussion

In this section, some immediate implications of the proven theorems on the catalytic transformations of entangled states are discussed. For simplicity, the entangled states, such as (2)–(3), will be represented only by their respective Schmidt coefficients. For this reason, the vectors x and y that appear in this section are normalized vectors of non-negative numbers $\left(\sum x_i = \sum y_i = 1\right)$. We first start with probabilistic transformations.

7.1. Probabilistic conversion

The least upper bound on the conversion probabilities $P(x \otimes c \to y \otimes c)$ can be computed using (6) and the conditions given in theorem 2 as

$$P_{\text{cat}}(x \to y) = \min_{\nu\in[-\infty,1]} \frac{A_\nu(x)}{A_\nu(y)}, \tag{136}$$

where the minimum is used by the inclusion of the end points. Although the minimization is over a continuous variable, it is possible to compute $P_{\text{cat}}(x \to y)$ to any desired accuracy.

For any probability p with $p < P_{\text{cat}}(x \to y)$, it is possible to find a catalyst such that the state x can be converted into y with a success probability p. The case $p = P_{\text{cat}}(x \to y)$, i.e. the attainability of the upper bound has to be investigated separately by using theorems 1 and 2. If there is a ν in the interval $(-\infty, 1)$ which gives the minimum of (136), then $P_{\text{cat}}(x \to y)$ cannot be achieved by any catalyst c. This is always the case when $P_{\text{cat}}(x \to y) < \min\left(1, x_1^{\uparrow}/y_1^{\uparrow}\right)$.

On the other hand, if the minimum of (136) occurs only at the end point $\nu = -\infty$, in which case we have $P_{\text{cat}}(x \to y) = x_1^{\uparrow}\big/y_1^{\uparrow} < 1$, then there is a catalyst that achieves that maximum probability. It is possible to find non-trivial examples of this case, where catalysis improves the conversion probability, i.e. $P_{\text{cat}}(x \to y) > P(x \to y)$.

Finally, if $P_{\text{cat}}(x \to y) = 1$, then the upper bound can be reached by a catalyst if and only if $x \prec_T y$. An interesting situation occurs if this is not the case, i.e. when $x \not\prec_T y$. For such pairs of states, it is possible to find catalysts that achieve any conversion probability p with $p < 1$, but there will always be a possibility of failure, in which case, the catalyst may also be destroyed. If the trumping relation is violated because of an inversion of some inequality in (T2), i.e. if there is a $\nu > 1$ such that $A_\nu(x) > A_\nu(y)$, then the reason for such a behavior can be understood by using entanglement monotones [15]. Let $\|x\|_\nu = n^{1/\nu} A_\nu(x)$ be the ℓ_ν norm of the vector x. Then $-\|x\|_\nu$ is an entanglement monotone. This monotone increases in the transformation of $x \otimes c$ into $y \otimes c$ for any catalyst c, and therefore there should always be a failure which significantly decreases the monotone. This monotone can also be used to find minimal resources that the catalyst c should have if it achieves a probability p as follows:

$$\|c\|_\nu \leqslant \frac{1 - p}{\|x\|_\nu - p\|y\|_\nu}. \tag{137}$$

As each component of c has to be smaller than the bound given, the number of components of c diverges inversely proportional to $1 - p$ as the probability value approaches to 1, i.e. larger and larger resources are needed to get closer to the upper bound.

7.2. *Closure of $T(y)$*

It is also of some interest to investigate the vectors in the closure of

$$T(y) = \{x : x \prec_T y\}, \tag{138}$$

which is the set of vectors trumped by y. Using theorem 1, it is possible to prove a conjecture which is attributed to Nielsen: $x \in \overline{T(y)}$ if and only if conditions (T1)–(T2) are satisfied with non-strict inequalities, i.e. $A_\nu(x) \geqslant A_\nu(y)$ for $\nu < 1$ and $A_\nu(x) \leqslant A_\nu(y)$ for $\nu > 1$ (the third, $\sigma(x) \geqslant \sigma(y)$, follows from these). In fact, if these conditions are satisfied (excluding the trivial cases of maximally entangled and non-entangled states), it is possible to find different vectors $\bar{x}$ very near to x such that $\bar{x} \prec x$ and $\bar{y}$ very near to y such that $y \prec \bar{y}$. Strict inequalities (T1)–(T3) for the pair $(\bar{x}, x)$ then show that the pair $(\bar{x}, y)$ satisfies the same inequalities which lead to $\bar{x} \prec_T y$. This shows that any neighborhood of x contains a vector trumped by y. The same argument can also be used to show that $x \prec_T \bar{y}$, i.e. any neighborhood of y contains a vector which trumps x.

If $x \in \overline{T(y)}$ but $x \not\prec_T y$, it means that no catalyst can achieve the conversion of x into y with probability 1, but it is possible to find a sequence of catalysts (with growingly large Schmidt numbers) such that the conversion probability is made to approach 1. However, this case has a significantly different aspect than the other pairs of states which have $P_{\text{cat}} = 1$. Namely, it is possible to find states $\bar{y}$ very near to y such that x to $\bar{y}$ conversion is possible with certainty. In other words, high fidelity conversion is possible without running the risk of losing the catalyst.

7.3. Catalytic conversion ratio

Let x and y be the Schmidt coefficients of some given states. What is the best ratio for M/N if N copies of x is needed to be transformed into M copies of y? The asymptotic procedure of Bennett *et al* achieves the ratio $C_{\text{asymp}} = \sigma(x)/\sigma(y)$, but this procedure is associated with a small probability of failure and a loss in fidelity. It is of some interest to find the best ratio when there is no possibility of failure and the exact final state y is needed to be produced (with fidelity 1). A simple answer can now be found when catalysis is allowed. To simplify the discussion, it is assumed that $n_x^N > n_y^M$, where n_x and n_y are Schmidt numbers of the states x and y respectively. In that case, it is only necessary to check (T1)–(T3) for positive ν values. In conclusion, if

$$\frac{M}{N} < C_{\text{cat}} = \min_{\nu\in[0,\infty]} \frac{\sigma_\nu(x)}{\sigma_\nu(y)}, \tag{139}$$

where σ_ν are the Renyi entropies, then $x^{\otimes N} \prec_T y^{\otimes M}$ and it is possible to convert N copies of x into M copies of y with certainty by using a suitable catalyst. When failure is acceptable, a different conversion ratio can be found. If

$$\frac{M}{N} < C_{\text{cat,prob}} = \min_{\nu\in[0,1]} \frac{\sigma_\nu(x)}{\sigma_\nu(y)} \tag{140}$$

is satisfied, then conversion by a suitable catalyst is possible where the success probability can be made sufficiently close to 1. However, failure for such a procedure implies the loss of entanglement not only of x, but also of the catalyst as well. In both of these cases, the numbers of copies N and M do not need to be large, in contrast to the asymptotic procedure. As, $C_{\text{cat}} \leqslant C_{\text{cat,prob}} \leqslant C_{\text{asymp}}$, a better conversion ratio of the asymptotic transformation might make it more advantageous than the probabilistic catalysis, but a detailed investigation of losses incurred in the event of failure is needed before reaching to a definite conclusion.

There are two special cases of particular interest, i.e. when either the initial or the final state is maximally entangled. By using the fact that the Renyi entropy σ_ν is a non-increasing function of ν, it is possible to compute the catalytic conversion ratio explicitly for both of these cases. When x is maximally entangled and y is partially entangled, then catalysis does not have any advantage in the transformation process. The conversion is possible as long as the Schmidt number does not increase. The expressions above also give the same result, namely $C_{\text{cat}} = C_{\text{cat,prob}} = \ln n_x / \ln n_y$.

For the other case, when it is desired to concentrate the partially entangled state x to maximally entangled y, we have

$$C_{\text{cat}} = \frac{\sigma_\infty(x)}{\ln n_y} = \frac{(-\ln x_{\max})}{\ln n_y}, \tag{141}$$

for the deterministic case, where $x_{\max}$ is the largest component of x. If failure is allowed, then the conversion ratio matches with the asymptotic one, $C_{\text{cat,prob}} = C_{\text{asymp}}$.

8. Conclusions

The necessary and sufficient conditions for the trumping relation and two associated relations have been found. These conditions will be valuable in the investigations of catalytic transformations. Once it is understood that catalysis is possible, the problem of finding a suitable catalyst can in principle be solved by going backwards along the proofs. Although possible solutions of the problem posed in lemma 1 in section 3 can be found by the well-established procedures of linear programming, carrying out the whole procedure for realistic cases might be forbidding, as the degree of the polynomial $\gamma(s)$ and of the sought for polynomial $a(s)$ might be very large. However, the method used in the proof of the lemma can be used to place an upper bound on the degree of $a(s)$ (but not on the Schmidt number). This also suggests a conjecture that the complex roots, ν, of the equation $A_\nu(x) = A_\nu(y)$, and their closeness to the real line could be used for estimating the minimum amount of resources the catalysts should have.

Appendix. Proof of theorem 4

The most troublesome part of the proof of theorem 4 is the neighborhood of $\nu = 1$. This part can be handled with the following theorem.

Theorem 5. *For any positive vector x and any given $\delta > 0$, there is a positive number ϵ such that for any $\bar{x}$ with $\sum_i \bar{x}_i = \sum x_i$, and $D(x;\bar{x}) \leqslant \epsilon$, we have*

$$e^{-\delta|\nu-1|} \leqslant \frac{A_\nu(\bar{x})}{A_\nu(x)} \leqslant e^{\delta|\nu-1|} \qquad \forall \nu \in [1/2, 2]\,. \tag{A.1}$$

Proof of theorem 5. Without loss of generality, it is supposed that $x = x^\uparrow$. Let $S_\nu(x) = \sum_i x_i^\nu$ be the νth power sum and K_ν be defined as

$$K_\nu = \ln \frac{S_\nu(\bar{x})}{S_\nu(x)}. \tag{A.2}$$

Note that $K_1 = 0$. We place the following bound on the absolute value of ν derivative of K_ν:

$$\left|\frac{dK_\nu}{d\nu}\right| = \frac{1}{S_\nu(x)S_\nu(\bar{x})}\left|\sum_{ij} \bar{x}_i^\nu x_j^\nu \ln \frac{\bar{x}_i}{x_j}\right| \tag{A.3}$$

$$\leqslant \epsilon + \frac{1}{S_\nu(x)S_\nu(\bar{x})}\left|\sum_{ij} \bar{x}_i^\nu x_j^\nu \ln \frac{x_i}{x_j}\right|$$

$$\leqslant \epsilon + \frac{1}{S_\nu(x)S_\nu(\bar{x})}\left|\sum_{i>j} \left(\bar{x}_i^\nu x_j^\nu - \bar{x}_j^\nu x_i^\nu\right) \ln \frac{x_i}{x_j}\right| \tag{A.4}$$

Since for $i > j$ we have $x_i \geqslant x_j$, all of the logarithmic terms are non-negative in the expression above. As a result, for any positive ν we have

$$\left|\frac{dK_\nu}{d\nu}\right| \leqslant \epsilon + \frac{e^{\nu\epsilon} - e^{-\nu\epsilon}}{S_\nu(x)S_\nu(\bar{x})} \sum_{i>j} x_i^\nu x_j^\nu \ln \frac{x_i}{x_j} \tag{A.5}$$

$$\leqslant \epsilon + \frac{e^{\nu\epsilon} - e^{-\nu\epsilon}}{2S_\nu(x)S_\nu(\bar{x})} \sum_{i,j} x_i^\nu x_j^\nu \left|\ln \frac{x_i}{x_j}\right| \tag{A.6}$$

$$\leqslant \epsilon + \frac{e^{\nu\epsilon} - e^{-\nu\epsilon}}{2} \frac{S_\nu(x)}{S_\nu(\bar{x})} \ln \frac{x_n}{x_1}. \tag{A.7}$$

Finally, we can apply $S_\nu(\bar{x}) \geqslant \mathrm{e}^{-\nu\epsilon} S_\nu(x)$ to the last line which gives

$$\left|\frac{\mathrm{d}K_\nu}{\mathrm{d}\nu}\right| \leqslant \epsilon + \frac{\mathrm{e}^{2\nu\epsilon} - 1}{2} \ln \frac{x_n}{x_1} \tag{A.8}$$

$$\leqslant \epsilon + \frac{\mathrm{e}^{4\epsilon} - 1}{2} \ln \frac{x_n}{x_1}, \tag{A.9}$$

where the last inequality is valid for all $0 < \nu \leqslant 2$. Note that the right-hand side of the last expression has a zero limit as $\epsilon \to 0$. This enables us to choose the value of ϵ so small that the right-hand side is less than $\delta/2$. In other words, $|\mathrm{d}K_\nu/\mathrm{d}\nu| \leqslant \delta/2$.

Next, we express K_ν as

$$K_\nu = \int_1^\nu \frac{\mathrm{d}K_\nu}{\mathrm{d}\nu}\,\mathrm{d}\nu. \tag{A.10}$$

The inequality above then implies that

$$|K_\nu| \leqslant \tfrac{1}{2}\delta|\nu - 1|. \tag{A.11}$$

Finally, considering only the values of ν in the interval $[1/2, 2]$, we have

$$\left|\ln \frac{A_\nu(\bar{x})}{A_\nu(x)}\right| = \frac{|K_\nu|}{\nu} \leqslant \frac{2\delta|\nu - 1|}{\nu} \leqslant \delta|\nu - 1|, \tag{A.12}$$

which is the desired result. □

Proof of theorem 4. Without loss of generality suppose that x and y are normalized, i.e. $\sum x_i = \sum y_i = 1$. Since the minimum and maximum values of these vectors are different by the assumptions of the theorem, the strict inequalities are valid at the infinities, i.e. $A_{-\infty}(x) > A_{-\infty}(y)$ and $A_\infty(x) < A_\infty(y)$. Let G_ν be defined as

$$G_\nu = \ln \frac{A_\nu(y)}{A_\nu(x)}. \tag{A.13}$$

By the inequalities (T1)–(T3), we have $G_\nu < 0$ for all $\nu < 1$ and $G_\nu > 0$ for all $\nu > 1$. At infinities, G_ν approaches to non-zero limits. Moreover, the derivative of G_ν at $\nu = 1$ is

$$G_1' = \sigma(x) - \sigma(y) > 0. \tag{A.14}$$

Therefore, both of the following quantities are strictly positive:

$$B = \min_{\nu \in [-\infty, 1/2] \cup [2, \infty]} |G_\nu|, \tag{A.15}$$

$$M = \min_{\nu \in [1/2, 2]} \frac{G_\nu}{\nu - 1}. \tag{A.16}$$

By theorem 5, there are numbers ϵ_1 and ϵ_2 such that $D(x; \bar{x}) \leqslant \epsilon_1$ implies that $|\ln A_\nu(\bar{x})/A_\nu(x)| \leqslant M|\nu - 1|/3$ and $D(y; \bar{y}) \leqslant \epsilon_2$ implies that $|\ln A_\nu(\bar{y})/A_\nu(y)| \leqslant M|\nu - 1|/3$. We choose $\epsilon = \min(\epsilon_1, \epsilon_2, B/3)$.

Let $\bar{x}$ and $\bar{y}$ be arbitrary vectors such that $\sum \bar{x}_i = \sum \bar{y}_i = 1$, $D(x; \bar{x}) \leqslant \epsilon$ and $D(y; \bar{y}) \leqslant \epsilon$. Let

$$\bar{G}_\nu = \ln \frac{A_\nu(\bar{y})}{A_\nu(\bar{x})} = G_\nu + \ln \frac{A_\nu(\bar{y})}{A_\nu(x)} + \ln \frac{A_\nu(x)}{A_\nu(\bar{x})}. \tag{A.17}$$

Our purpose is to show that $\bar{G}_\nu$ satisfies the desired properties, i.e. it is negative for $\nu < 1$, positive for $\nu > 1$ and has a simple zero at $\nu = 1$. Note that $D(x; \bar{x}) \leqslant \epsilon$ implies that $|\ln A_\nu(x)/A_\nu(\bar{x})| \leqslant \epsilon$ for all ν.

We consider the following ranges of ν values separately.

(a) For $\nu \leqslant 1/2$, we have $\bar{G}_\nu \leqslant -B + 2\epsilon \leqslant -B/3 < 0$.
(b) For $\nu \geqslant 2$, we have $\bar{G}_\nu \geqslant B - 2\epsilon \geqslant B/3 > 0$.
(c) For $\nu \in [1/2, 2]$, we have

$$\frac{\bar{G}_\nu}{\nu - 1} \geqslant M - \frac{2M}{3} > 0. \tag{A.18}$$

As a result, $\bar{G}_\nu$ satisfies the desired properties in this interval as well.

Finally, for the inequality (T3), we note that for $\nu \in [1/2, 1)$, we have

$$\frac{\bar{G}_\nu}{\nu - 1} \geqslant \frac{M}{3}. \tag{A.19}$$

Taking the $\nu \to 1$ limit gives the derivative of $\bar{G}_\nu$, which is

$$\bar{G}_1' = \sigma(\bar{x}) - \sigma(\bar{y}) \geqslant \frac{M}{3}. \tag{A.20}$$

This completes the proof of theorem 4. □

References

[1] Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 *Phys. Rev.* A **53** 2046
[2] Bhatia R 1997 *Matrix Analysis, Graduate Texts in Mathematics* vol 169 (New York: Springer-Verlag)
[3] Nielsen M A 1999 *Phys. Rev. Lett.* **83** 436
[4] Vidal G 1999 *Phys. Rev. Lett.* **83** 1046
[5] Jonathan D and Plenio M B 1999 *Phys. Rev. Lett.* **83** 3566
[6] Duan R, Ji Z, Feng Y and Ying M 2006 *J. Comput. Sci. Technol.* **21** 776
[7] Daftuar S and Klimesh M 2001 *Phys. Rev.* A **64** 042314
[8] Sun X, Duan R and Ying M 2005 *IEEE Trans. Inf. Theory* **51** 75
[9] Aubrun G and Nechita I 2007 *Preprint* quant-ph/0702153v2
[10] Aubrun G and Nechita I 2007 *Preprint* arXiv:0707.0211v1
[11] Feng Y, Duan R and Ying M 2004 *Phys. Rev.* A **69** 062310
[12] Feng Y, Duan R and Ying M 2005 *IEEE Trans. Inf. Theory* **51** 1090
[13] Duan R, Feng Y, Li X and Ying M 2005 *Phys. Rev.* A **71** 042319
[14] Bandyopadhyay S, Roychowdhury V and Sen U 2002 *Phys. Rev.* A **65** 052315
[15] Vidal G 2000 *J. Mod. Opt.* **47** 355