Quantum Information and Computation, Vol. 1, No. 1 (2001) 76–93

MAJORIZATION AND THE INTERCONVERSION OF BIPARTITE STATES

Michael A. Nielsen
Center for Quantum Computer Technology, University of Queensland, Queensland 4072, Australia

Guifré Vidal
Institut für Theoretische Physik, Universität Innsbruck, Queensland 4072, Australia

Received May 28, 2001

Majorization is a powerful, easy-to-use and flexible tool which arises frequently in quantum mechanics as a consequence of fundamental connections between unitarity and the majorization relation. Entanglement theory does not escape from its influence. Thus the interconversion of bipartite pure states by means of local manipulations turns out to be ruled to a great extend by majorization relations. This review both introduces some elements of majorization theory and describes recent results on bipartite entanglement transformations, with special emphasis being placed on explaining the connections between these two topics. The latter implies analyzing two other aspects of quantum mechanics similarly influenced by majorization, namely the problem of mixing of quantum states and the characterization of quantum measurement.

Keywords: majorization, quantum information, entanglement

1. Introduction

Majorization was developed to answer the following question: what does it mean to say that one probability distribution is more disordered than another? In the quantum mechanical context, this question becomes: given two quantum states, what does it mean to say that one is more disordered than the other? Majorization gives a means for comparing two probability distributions or two density matrices in an elegant way. It arises surprisingly often in fields such as computer science, economics, and, most important for us, quantum mechanics. On the other hand, since the early days of quantum theory entanglement has been one of its highlights, mainly due to its relation to non-locality. Entanglement is nowadays conceived also as a resource in quantum information theory. As recent results have shown, majorization relations play a remarkable role in the theory of quantum entanglement, and this will be the subject of the present review, which has the following three goals:

- to introduce some of the ideas and results of the theory of majorization;
- to point out the fundamental reasons why connections between majorization and entanglement theory —or, more generally, quantum mechanics— exist;

- to describe some applications of majorization in the context of local transformation of bipartite pure states.

The basic intuition underlying majorization may be understood from the following definition: we say the d-dimensional real vector r is *majorized* by the d-dimensional real vector s, written $r \prec s$, if there exist a set of d-dimensional permutation matrices P_j and a probability distribution $\{p_j\}$ such that

$$r = \sum_j p_j P_j s. \tag{1}$$

That is, r is majorized by s precisely when r can be obtained from s by randomly permuting the components of s, and then averaging over the permutations. At least naively this appears to be a natural and appealing approach to defining the notion that one vector is more disordered than another. This naive appeal is more than justified by the rich mathematical structure arising from this definition.

As a simple example of majorization, suppose the vector s is a probability distribution on d outcomes, that is, the components are non-negative and sum to one. Then it is easy to see that

$$\left(\frac{1}{d}, \ldots, \frac{1}{d}\right) \prec s, \tag{2}$$

since the uniform distribution $(1/d, \ldots, 1/d)$ may be obtained by averaging over permutations $P_\pi s$ of s, where π is chosen uniformly at random from the symmetric group on d elements. This simple example agrees with our intuition that the uniform distribution on d elements is at least as disordered as any other probability distribution over d elements.

What connections are there between majorization and quantum mechanics? The quantum mechanical analogue of a probability distribution is the density operator, so a natural beginning is to define an operator notion of majorization. Letting R and S be d-dimensional Hermitian operators, we define $R \prec S$ if $\lambda(R) \prec \lambda(S)$, where $\lambda(R)$ denotes the vector whose components are the eigenvalues of R, arranged in decreasing order.

Then, the essential reason for the close connection between majorization and quantum mechanics may be appreciated by inspection of two elegant (and closely related) results: *Horn's lemma* and *Uhlmann's theorem.* Horn's lemma states that for vectors r and s, $r \prec s$ if and only if $r_i = \sum_j |u_{ij}|^2 s_j$ for some unitary matrix $u = (u_{ij})$ of complex numbers. Uhlmann's theorem states, in direct analogy with Eq. (1), that $R \prec S$ for Hermitian matrices R and S if and only if there exist unitary matrices U_j and a probability distribution $\{p_j\}$ such that

$$R = \sum_j p_j U_j S U_j^\dagger. \tag{3}$$

The fundamental role of unitarity in quantum mechanics ensures that relations of the type found in Horn's lemma and Uhlmann's theorem arise frequently, and it is this fact what accounts for many of the applications of majorization in quantum mechanics.

Let us consider which particular mechanism make the above results eventually reach the domain of entanglement theory. In the context of local manipulation of a bipartite system AB, our ultimate aim is to be able to answer questions such as: given an entangled pure state ψ, can it be transformed into another state ϕ by means of local operations on the two subsystems A and B aided with classical communication (LOCC for short)? The feasibility of the transformation $\psi \rightarrow \phi$ by means of LOCC turns out to depend on whether a simple majorization relation is satisfied between the reduced density matrix ρ_A of part A of the system for states ψ and ϕ, namely on whether

$$\rho_A^\psi \prec \rho_A^\phi. \tag{4}$$

To understand the origin of this condition, a close inspection of the allowed manipulation of the bipartite system is required. We will see that any LOCC transformation involving only pure states can be implemented by means of a single local measurement on one of the parts of the bipartite system followed by an outcome-dependent unitary operation on the other part. In addition, local unitary operations are known to be irrelevant as far as the entanglement properties of the final state are concerned. We are therefore left with a single local measurement as the only element potentially responsible for condition (4). It comes then as natural to investigate the role of majorization in quantum measurements.

It will be convenient, however, to first study the influence of majorization in the context of mixing of quantum states. This, apart from helping us understand certain restrictions on quantum measurements, will also provide an extra result that can be applied to design local conversion strategies. Thus, we shall make a small detour and analyze some aspects of mixing and measurement in quantum mechanics. After that we shall be prepared to understand why majorization theory, a tool originally developed to study disorder, rules the interconversion of bipartite pure states under local manipulations of the system.

Despite its indisputable usefulness, one might well ask at the outset why we need the notion of majorization when measures of disorder such as the Shannon and von Neumann entropies are already available. Could not these other measures be as useful as majorization? This is a good question. It turns out that the entropic measures arise naturally out of the theory of majorization in a sort of "law of large numbers limit" where a large number of identical systems are considered. Of course, this also has consequences for entanglement transformations. In the so-called asymptotic regime, where many copies of a pure state ψ are transformed into many copies of a pure state ϕ, conversions are possible if and only if a single inequality for the von Neumann entropy of the reduced density matrix is obeyed. The essential point will be that measures such as the entropy are essentially *weaker* than the notion of majorization, and as such do not give as much detailed information as provided by majorization.

The paper is divided into three more sections. Section presents and describes some results of majorization theory. This section is purely mathematical in nature, although all the chosen topics play a role in later applications. Section applies the previous results to the analysis of mixing of quantum states and to quantum measurement, which in turn prepares the path for the contents of section , where the interconversion of bipartite entanglement is finally analyzed. There, we shall discuss the criterion for feasibility of the most general pure-state transformations of a bipartite system under LOCC, and shall describe a conversion strategy that requires remarkably little classical communication. The relation between majorization and entanglement monotones, and majorization in asymptotic transformations will be also briefly discussed.

Throughout the paper, specially in the last two sections, we have made historical remarks to the best of our knowledge; our apologies to any researcher inadvertently omitted from citation. We have not given citations for some classic results on majorization as Marshall and Olkin's classic 1979 text[1] comprehensely covers the literature well beyond our competence and available space.

2. Elements of the theory of majorization

In this section we present some results of majorization theory that will be needed later on. A series of theorems are stated and explained, mainly without proof. Any reader seriously interested in majorization is recommended Marshall and Olkin's book[1], which contains a wealth of additional material we have not covered, including many applications of majorization outside physics. Other references on the theory of majorization include Chapters 2 and 3 of the book by Bhatia[2], Ando's[3] survey articles on the subject of majorization, and finally Alberti and Uhlmann's[4] monograph.

2.1. *Alternative definition*

The definition for the majorization relation $r \prec s$ in Eq. (1) in terms of random permutations is satisfying from an intuitive point of view, and is often useful when proving theoretical results, but is rather inconvenient for actual calculations. Given two vectors of numbers r and s, is there some simple procedure to determine whether $r \prec s$? Rather remarkably, such a procedure exists. First, we re-order the components of r and s into decreasing order, writing for example $r^\downarrow = (r_1^\downarrow, \ldots, r_d^\downarrow)$ for the vector whose components are the same as those of r, but ordered so that

$$r_1^\downarrow \geq r_2^\downarrow \geq \ldots \geq r_d^\downarrow. \tag{5}$$

It turns out that $r \prec s$ if and only if

$$\begin{aligned} r_1^\downarrow &\leq s_1^\downarrow \\ r_1^\downarrow + r_2^\downarrow &\leq s_1^\downarrow + s_2^\downarrow \\ &\vdots \\ r_1^\downarrow + \ldots + r_{d-1}^\downarrow &\leq s_1^\downarrow + \ldots + s_{d-1}^\downarrow \\ r_1^\downarrow + \ldots + r_d^\downarrow &= s_1^\downarrow + \ldots + s_d^\downarrow. \end{aligned} \tag{6}$$

Note that we could have actually taken this *calculational* definition from the very beginning, and obtain Eq. (1) as a result. However, this set of inequalities is probably less suggestive than Eq. (1) as far as the nature of the majorization relation is concerned. In any case, we will shortly discuss their equivalence.

Any two vectors x and y such that $x^{\downarrow} = y^{\downarrow}$ are equivalent from the point of view of majorization. We can thus restrict our considerations to "ordered" real vectors $x = x^{\downarrow}$. The relation $\prec$ defines a *partial* order on the set of these vectors. It gives only a *partial* rather than a *total* order, since there are vectors x and y which are *incomparable* in the sense that neither $x \prec y$ nor $y \prec x$. An example of this phenomenon is provided by the vectors $x = (0.5, 0.25, 0.25)$ and $(0.4, 0.4, 0.2)$.

The related notion of super-majorization will also be needed later on. Given any d-dimensional vector r, let $\{r_i^{\uparrow}\}$ denote the elements of r re-ordered into increasing order. Then r is *sub-majorized* by s, written $r \prec^w s$, if $\sum_{j=1}^k r_j^{\uparrow} \geq \sum_{j=1}^k s_j^{\uparrow}$ for each k in the range 1 through d. Notice that the majorization relation can also be rewritten as $\sum_{j=1}^k r_j^{\uparrow} \geq \sum_{j=1}^k s_j^{\uparrow}$, but with the extra constraint $\sum_{j=1}^d r_j^{\uparrow} = \sum_{j=1}^d s_j^{\uparrow}$.

2.2. *Majorization and double stochasticity*

So far we have introduced two different definitions for majorization, one in terms of a probabilistic sum of permutations, Eq. (1), and a second one in terms of a set of inequalities, Eqs. (6). The equivalence between these definitions follows from two important theorems we now turn to. These theorems connect in a notorious way the notion of majorization with doubly stochastic matrices.

A real d by d matrix $D = (D_{ij})$ is *doubly stochastic* if the entries of D are non-negative, and each row and column of D sums to 1. A simple example of a doubly stochastic matrix, the most general of dimensions 2 by 2, is given by

$$D = \begin{bmatrix} t & 1-t \\ 1-t & t \end{bmatrix}, \tag{7}$$

where t is a parameter in the range 0 to 1. As mentioned above, doubly stochastic matrices are closely related to majorization, as reflected in the following theorem.

Theorem 1 *$r \prec s$ if and only if $r = Ds$ for some doubly stochastic matrix D.*

The set of d by d doubly stochastic matrices is convex. Furthermore, it is easily verified that the permutation matrices are extreme points of this set, that is, if P is a permutation matrix, then it is not a convex combination of two distinct doubly stochastic matrices. Birkhoff's theorem asserts that the permutation matrices exhaust the extremal points of the set of doubly stochastic matrices, that is, any doubly stochastic matrix can be written as a convex combination of permutation matrices. In addition, together with Theorem 1 it proves equivalence of the two definitions we have introduced.

Theorem 2 (Birkhoff's theorem) *The set of d by d doubly stochastic matrices is a convex set whose extreme points are the permutation matrices.*

As a simple example of Birkhoff's theorem we can express any 2 by 2 doubly stochastic matrix as a convex combination of permutations:

$$\begin{bmatrix} t & 1-t \\ 1-t & t \end{bmatrix} = t \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (1-t) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{8}$$

In the general d by d case there are $d!$ different permutation matrices. However, Caratheodory's theorem[5] guarantees that a point in an m-dimensional compact convex set may be expressed as a convex combination of at most $m+1$ extremal points of that set. The d by d doubly stochastic matrices form a d^2-2d+1-dimensional set, so an arbitrary doubly stochastic matrix may be expressed as a convex combination of at most d^2-2d+2 permutation matrices.

2.3. *Double stochasticity and unitarity*

Let us now build the bridge to quantum mechanics. As discussed in the introduction, *Horn's lemma* connects unitary matrices and majorization through doubly stochastic matrices. Suppose $U=(U_{ij})$ is a unitary matrix. Then the matrix $D_{ij}\equiv|U_{ij}|^2$ is doubly stochastic since the rows and columns of U are unit vector; a doubly stochastic matrix that can be written in this way is called *unitary-stochastic.*

Theorem 3 (Horn's lemma) *$r\prec s$ if and only if a unitary-stochastic D exists such that $r=Ds$.*

D is said to be *ortho-stochastic* if, in addition, U is orthogonal, that is, a real unitary matrix. Although not all doubly stochastic matrices are unitary-stochastic, the following theorem shows that, from the point of view of majorization, ortho-stochastic matrices are actually all we need consider:

Theorem 4 *Suppose $r\prec s$. Then an ortho-stochastic D exists such that $r=Ds$.*

2.4. *Operator majorization*

The notion of majorization can be extended to hermitian operators by focussing on the eigenvalues of the operators R and S to be compared. $R\prec S$ means, by definition, that $\lambda(R)\prec\lambda(S)$. Then Uhlmann's theorem, in analogy with Eq. (1) —or, equivalently, Theorem 1—, states the following.

Theorem 5 (Uhlmann) *$A\prec B$ if and only if there exist unitary matrices U_i and probabilities p_i such that*

$$A=\sum_i p_iU_iBU_i^\dagger. \tag{9}$$

Uhlmann's theorem clearly illustrates the idea that the hermitian operator A (eventually, a density matrix) is more random than B, since A can be obtained by independently applying to B unitary operations $\{U_i\}$, and mixing the resulting operators $U_iBU_i^\dagger$ according to the probabilities $\{p_i\}$. This transformation is known a unitary mixing.

As an elementary consequence of Horn's lemma we have *Ky Fan's maximum principle*, which says that for any Hermitian matrix A, the sum of the k largest eigenvalues of A is the maximum value of $\mathrm{tr}(AP)$, where the maximum is taken over all k-dimensional projectors P.

Theorem 6 (Ky Fan's maximum principle)

$$\sum_{j=1}^{k} \lambda_j(A) = \max_P tr(AP). \tag{10}$$

Proof. Choosing P to be the projector onto the space spanned by the k eigenvectors of A with the k largest eigenvalues results in $\mathrm{tr}(AP) = \sum_{j=1}^{k} \lambda_j(A)$. We only need to show that $\mathrm{tr}(AP) \leq \sum_{j=1}^{k} \lambda_j(A)$ for any k-dimensional projector P. Let $|e_1\rangle, \ldots, |e_d\rangle$ be an orthonormal basis chosen so that $P = \sum_{j=1}^{k} |e_k\rangle\langle e_k|$. Let $|f_1\rangle, \ldots, |f_d\rangle$ be an orthonormal set of eigenvectors for A. Then

$\langle e_j|A|e_j\rangle = \sum_{k=1}^{d} |u_{jk}|^2 \lambda_k(A)$, where $u_{jk} \equiv \langle e_j|f_k\rangle$ is unitary. By Horn's lemma it follows that $(\langle e_j|A|e_j\rangle) \prec \lambda(A)$, which implies that $\mathrm{tr}(AP) = \sum_{j=1}^{k} \langle e_j|A|e_j\rangle \leq \sum_{j=1}^{k} \lambda_j(A)$. $\square$.

Ky Fan's maximum principle implies that for Hermitian matrices A and B,

$$\lambda(A + B) \prec \lambda(A) + \lambda(B). \tag{11}$$

To prove this very useful result, choose a k-dimensional projector P such that $\sum_{j=1}^{k} \lambda_j(A+B) = \mathrm{tr}((A + B)P) = \mathrm{tr}(AP) + \mathrm{tr}(BP) \leq \sum_{j=1}^{k} \lambda_j(A) + \sum_{j=1}^{k} \lambda_j(B)$.

2.5. *Majorization and entropies*

The Shannon entropy $H(x) \equiv -\sum_i x_i \log x_i$ of a probability distribution x satisfies that $x \prec y \Rightarrow -H(x) < -H(y)$. This is a particular case of a more general result, which we state in the following weak form.

Theorem 7 *$x \prec y \Rightarrow F(x) < F(y)$, where $F(x) \equiv \sum_i f(x_i)$, for any convex function $f : \mathbf{R} \to \mathbf{R}$.*

Proof. Suppose $x \prec y$ and $f(\cdot)$ is a convex function. Then $x = \sum_i p_i P_i y$ for some set of probabilities p_i and permutation matrices P_i. $F(\cdot)$ is a sum of convex functions, and thus convex, so $F(x) \leq \sum_i p_i F(P_i y)$. But $F(\cdot)$ is manifestly permutation invariant, so $F(P_i y) = F(y)$, and thus $F(x) \leq \sum_i p_i F(y) = F(y)$, as required. $\square$.

This result can be extended to the domain of operator functionals [6].

Theorem 8 *$\rho \prec \sigma \Rightarrow F(\rho) \leq F(\sigma)$, where $F(\rho) \equiv \sum_{i=1}^{d} f(\lambda_i)$ and λ_i are the eigenvalues of ρ, for any convex functions $f : \mathbf{R} \to \mathbf{R}$.*

Proof. Suppose $\rho \prec \sigma$ and $f(\cdot)$ is a convex function. Then $\rho = \sum_i p_i U_i \sigma U_i^\dagger$ for some set of probabilities p_i and unitary matrices U_i. $F(\cdot)$ is a sum of convex functions, and thus convex, so $F(\rho) \leq \sum_i p_i F(U_i \sigma U_i \dagger)$. But $F(\cdot)$ is manifestly unitarily invariant, so $F(U_i \sigma U_i \dagger) = F(\sigma)$, and thus $F(\rho) \leq \sum_i p_i F(\sigma) = F(\sigma)$, as required. $\square$.

In particular it follows that the von Neumann entropy $S(\rho) = H(\lambda(\rho))$ also obeys $\rho \prec \sigma \Rightarrow -S(\rho) < -S(\sigma)$. Thus, if one probability distribution or one density operator are more disordered than another in the sense of majorization, then they are also so according to the Shannon or the von Neumann entropies, respectively. As the two previous theorems show, there are as a matter of fact many other functions that also preserve the majorization relation. Any such function, called Schur-convex, could in a sense be used as a measure of order. Notice, however, that the majorization relation is a stronger notion of disorder, giving more information than any Shur-convex function. As it is well known, the reason why the Shannon and the von Neumann entropies are so important in practice is that they properly quantify the order in some limiting conditions, namely when many copies of a system are considered. For the situations in which we are interested in this paper, majorization and not entropy is the most convenient tool to be used.

3. Characterization of mixing and measurements in quantum mechanics

The connection between majorization and quantum mechanics seems to have first been pointed out by Uhlmann[9]. Some of Uhlmann's results were later generalized by Wehrl to the infinite dimensional case[10]. In this section we will briefly describe some of the known applications of majorization theory in the contexts of quantum state mixing and of quantum measurements. For a more complete description see Nielsen's paper [11]. The reason for presenting these results here is that they will be very useful in the next section when discussing bipartite pure state entanglement transformations, since they constitute the link between entanglement theory and majorization.

3.1. *Mixing*

The *mixing problem* in quantum mechanics is the following: given a density matrix ρ, characterize the class of probability distributions p_j and density matrices ρ_j such that $\rho = \sum_j p_j \rho_j$. We start presenting a theorem that classifies all the different ways a given density matrix may be represented by an ensemble. An ensemble of pure states $\{p_i, \psi_i\}$ is said to realize the density matrix ρ if $\rho = \sum_i p_i |\psi_i\rangle\langle\psi|$.

Theorem 9 (Ensemble classification theorem [12]**)** *The ensembles* $\{p_i, |\psi_i\rangle\}$ *and* $\{q_i, |\phi_j\rangle\}$ *realize the same density matrix if and only if*

$$\sqrt{p_i}|\psi_i\rangle = \sum_j u_{ij} \sqrt{q_j} |\phi_j\rangle \tag{12}$$

where u_{ij} *is some unitary matrix and we pad the smaller ensemble with entries having probability zero to ensure that the two ensembles have the same number of elements.*

Combining the ensemble classification theorem with Horn's lemma gives the following theorem classifying the set of probability distributions consistent with a given density matrix.

Theorem 10 [16] *Suppose* ρ *is a density matrix. Let* p_i *be a probability distribution. Then there exist normalized quantum states* $|\psi_i\rangle$ *such that*

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{13}$$

if and only if $(p_i) \prec \lambda_\rho$, *where* λ_ρ *is the vector of eigenvalues of* ρ.

Proof. Suppose there is a set of states $|\psi_i\rangle$ such that $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$. Multiplying (12) by its adjoint gives $p_i = \sum_{jk} u_{ik}^* u_{ij}\lambda_j\delta_{jk}$, which simplifies to

$$p_i = \sum_j |u_{ij}|^2\lambda_j. \tag{14}$$

Since $D_{ij} \equiv |u_{ij}|^2$, we have $(p_i) = D\lambda$ for doubly stochastic D, and thus $(p_i) \prec \lambda$. Conversely, if $(p_i) \prec \lambda$ then we can find orthogonal (and thus unitary) u such that (14) is satisfied. Defining the states $|\psi_i\rangle$ by Eq. (12) gives the result. □.

Later on we will use the following corollary [11]: suppose $|\psi\rangle$ is a pure state of a composite system AB with Schmidt decomposition

$$|\psi\rangle = \sum_i \sqrt{p_i}|i_A\rangle|i_B\rangle. \tag{15}$$

Then given a probability distribution q_i there exists an orthonormal basis $|i'_B\rangle$ for system B and corresponding states $|\psi_i\rangle$ of system A such that

$$|\psi\rangle = \sum_i \sqrt{q_i}|\psi_i\rangle|i'_B\rangle \tag{16}$$

if and only if $(q_i) \prec (p_i)$.

Finally, the next theorem relates the spectrum of ρ, $\lambda(\rho)$, with that of density matrices ρ_i realizing it.

Theorem 11 *Let ρ be a density matrix, p_j a probability distribution, and ρ_j states such that $\rho = \sum_j p_j\rho_j$. Then the following constraint must be obeyed:*

$$\lambda(\rho) \prec \sum_j p_j\lambda(\rho_j) \tag{17}$$

$$\tag{18}$$

Proof. It follows immediately from Eq. (11). □.

3.2. *Measurement*

In quantum mechanics an *efficient* measurement —one where all information about outcomes is kept— is characterized by a single set of operators $\{F_i\}$ satisfying $\sum_i F_i^\dagger F_i = I$. It takes, with probability $p_i \equiv \mathrm{tr}\ F_i^\dagger F_i\rho$, a system originally in state ρ into the state $\rho_i \equiv (F_i^\dagger \rho F_i)/p_i$. Notice that an efficient measurement always maps pure states into pure states. Here, again, we would like to characterize the ensemble $\{p_i, \rho_i\}$ that can be produced from ρ through an efficient measurement. The next two theorems set restrictions on the probability vectors $\lambda(\rho)$, $\lambda(\rho'_i)$ and the probability distribution $\{p_i\}$.

Theorem 12 [11,18] *Suppose an efficient measurement $\{F_i\}$ transforms ρ into ρ'_i with probability p_i. Then the following constraint is obeyed:*

$$\lambda(\rho) \prec \sum_j p_j\lambda(\rho'_i). \tag{19}$$

Proof. (Adapted from theorem 2 in [8] applied to the entanglement monotones of [19]) Suppose the measurement $\{F_i\}$ is performed on part A of a bipartite system in a pure state $|\psi\rangle$ such that $\rho = \rho_A \equiv \mathrm{tr}_B|\psi\rangle\langle\psi|$. The resulting states are $|\psi_i\rangle \equiv F_i \otimes I_B|\psi\rangle/\sqrt{p_i}$, with $\rho'_i = \mathrm{tr}_B|\psi_i\rangle\langle\psi_i|$. Define $\rho_{B,i} \equiv \mathrm{tr}_A|\psi_i\rangle\langle\psi_i|$. Notice that, as expected in order to prevent faster than light signalling from A to B,

$$\rho_B = \mathrm{tr}_A|\psi\rangle\langle\psi| = \mathrm{tr}_A(\sum_i F_i^\dagger F_i \otimes I_B|\psi\rangle\langle\psi|) = \sum_i p_i \rho_{B,i}. \tag{20}$$

But we already known from theorem 11 on mixing that this implies

$$\lambda(\rho_B) \prec \sum_i p_i \lambda(\rho_{B,i}), \tag{21}$$

which is equivalent to Eq. (19) because, for any bipartite pure state, the reduced density matrices ρ_A and ρ_B have equivalent spectra, that is $\lambda(\rho_A) = \lambda(\rho_B)$. □.

Intuitively, we know that quantum measurements acquire (rather than lose) information about the system being measured. The above theorem just made this intuition mathematically precise: the eigenvalues of the initial state are more disordered than the average eigenvalues of the post-measurement state. A type of converse to this result (see below) also holds: provided an equation similar to (19) is fulfilled —with some additional technical restrictions—, it is possible to find a quantum measurement which gives the post-measurement state ρ'_i with probability p_i when performed with ρ as the initial state. Thus majorization provides a natural language to express sharp fundamental constraints on the ability of quantum measurements to acquire information about a quantum system. Since the entanglement of a bipartite state is related to how mixed the corresponding reduced density matrices are, and in order to modify their spectra we need to locally make efficient measurements, we can already start suspecting that the constraints regulating information acquisition in efficient measurements may be exactly the same as the ones ruling bipartite pure state entanglement manipulation.

Theorem 13 [11] *Suppose ρ is a density matrix with vector of eigenvalues λ, and σ_i are density matrices with vectors of eigenvalues λ_i. Suppose p_i are probabilities such that*

$$\lambda \prec \sum_i p_i \lambda_i \tag{22}$$

Then there exist matrices $\{E_{ij}\}$ and a probability distribution p_{ij} such that

$$\sum_{ij} E_{ij}^\dagger E_{ij} = I, \quad E_{ij}\rho E_{ij}^\dagger = p_{ij}\sigma_i, \quad \sum_j p_{ij} = p_i. \tag{23}$$

Proof. By Birkhoff theorem, $\lambda \prec \sum_i p_i \lambda_i$ implies that there exist permutation matrices P_j and probabilities q_j such that

$$\lambda = \sum_{ij} p_i q_j P_j \lambda_i. \tag{24}$$

Without loss of generality we may assume that ρ and σ_i are all diagonal in the same basis, with non-increasing diagonal entries, since if this is not the case then it is an easy matter to append unitary matrices to the measurement matrices to obtain the correct transformation. With this convention, we define matrices E_{ij} by

$$E_{ij}\sqrt{\rho} \equiv \sqrt{p_i q_j}\sqrt{\sigma_i}P_j^\dagger. \tag{25}$$

For simplicity sake, we will assume that ρ is invertible. Note that we have

$$\sqrt{\rho}\left(\sum_{ij} E_{ij}^\dagger E_{ij}\right)\sqrt{\rho} \quad = \quad \sum_{ij} p_i q_j P_j \sigma_i P_j^\dagger. \tag{26}$$

Comparing with (24) we see that the right-hand side of the last equation is just ρ and thus

$$\sqrt{\rho}\left(\sum_{ij} E_{ij}^\dagger E_{ij}\right)\sqrt{\rho} = \rho, \tag{27}$$

from which we deduce that $\sum_{ij} E_{ij}^\dagger E_{ij} = I$. Furthermore, from the definition (25) it follows that

$$E_{ij}\rho E_{ij}^\dagger \quad = \quad p_i q_j \sigma_i, \tag{28}$$

and thus upon performing a measurement defined by the measurement matrices $\{E_{ij}\}$ the result (i, j) occurs with probability $p_{ij} = p_i q_j$, $\sum_j p_{ij} = p_i$, and the post-measurement state is σ_i. $\square$.

4. Interconversion of bipartite states under LOCC

Entanglement transformations have been intensively studied in recent years. This active area of research has been so far specially successful in the case of transformations of bipartite systems, where, for instance, it has revealed the structure of entangled pure states [7,19,20,21,22]. It has also led to physically motivated criteria for the quantification of the entanglement properties of pure [23] and mixed [24] multipartite states. In addition these studies generate useful tools to address other relevant issues in quantum information theory, given the predominant role entanglement plays in the whole field. A good example is provided by the bounds on the quantum communication complexity of some distributed computations derived by estimating the amount of entanglement necessarily produced while performing the computation [25]. In general, any task involving distant parties and using up entangled states as a resource, such as quantum teleportation, benefits from a better understanding of entanglement.

So far in this paper we have introduced some results of majorization theory and have applied them both in the context of mixing of quantum states and in that of quantum measurements. We move now to the main aim of this review, the analysis of entanglement transformations from the point of view of majorization. First we describe and justify the necessary and sufficient conditions for a pure-state entanglement transformation of a bipartite system to be feasible under LOCC. Such conditions naturally endow the set of bipartite pure states with a majorization-based structure. They also lead to uncovering the surprising effect of entanglement catalysis. Then, we will be concerned with explicit protocols that locally perform a given transformation. In particular, we will describe a local strategy requiring very little communication between the parts of the system. We will move to describe the frame set by the theory of entanglement monotones, that applies to mixed states and to systems with more than two parties. Finally, we will comment on asymptotic transformations.

4.1. *Pure-state transformations under local operations and classical communication*

Consider a composite quantum system consisting of two spatially separated parts, A and B. Part A is controlled by Alice, who can perform the most general transformation on it. Similarly, Bob is in charge of part B, that he can manipulate arbitrarily. In addition, Alice and Bob can talk to each other. Under these circumstances, only a restricted set of transformations of the system AB are feasible, the so-called LOCC transformations, after local operations and classical communication. We are then interested in the characterization of the states Alice and Bob can take the system into, starting from a given pure state ψ and by means of LOCC.

The following theorems provide us with the answer when the final states are also pure. Let $\lambda(\psi) = \lambda(\rho_A^\psi)$ denote the vector of decreasingly ordered eigenvalues λ_i of the reduced density matrix ρ_A^ψ, or equivalently, of the Schmidt coefficients of ψ,

$$|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}|i_A\rangle|i_B\rangle, \tag{29}$$

a state of two d-level systems. Then the first theorem, due to Nielsen [7], refers to a local deterministic conversion, that is, a local conversion achieving the final state with certainty. It shows that the feasibility of the transformation depends on whether a majorization relation between the Schmidt coefficients of the initial and final states is obeyed.

Theorem 14 *State ψ can be converted into state ϕ by means of LOCC if, and only if,*

$$\lambda(\psi) \prec \lambda(\phi). \tag{30}$$

Notice that since $\lambda(\psi)$ corresponds to the spectrum of the reduced density matrix $\rho_A^\psi \equiv \mathrm{tr}_B|\psi\rangle\langle\psi|$, this condition is equivalent to Eq. (4).

The following generalization of theorem 14, due to Vidal[19], considers local conversions that succeed with some probability p at transforming the initial state into the wished final state. Notice that the necessary and sufficient condition is given in terms of a supermajorization relation.

Theorem 15 *State ψ can be conclusively converted into state ϕ with probability p by means of LOCC if, and only if,*

$$\lambda(\psi) \prec^w p\lambda(\phi). \tag{31}$$

This result explicitly characterizes then the maximal probability in the conversion of ψ to ϕ when the transformation can not be performed with certainty. Such probability had been previously presented in [26] for the case where the final state is maximally entangled.

Finally, the most general transformation involving only pure states, that is, from ψ to one element of the ensemble $\{p_j, \psi_j\}$, is addressed in this theorem by Jonathan and Plenio[20]. Again, the feasibility of the transformation by local means depends on a majorization relation, this time between the Schmidt coefficients of the initial state and an average of the Schmidt coefficients of the final states.

Theorem 16 *The probabilistic transformation $\psi \rightarrow \{p_j, \psi_j\}$ can be accomplished using only LOCC if, and only if*

$$\lambda(\psi) \prec \sum_j p_j \lambda(\psi_j). \tag{32}$$

It is quite remarkable that the feasibility of a pure-state transformation under LOCC depends only on a system of d inequalities. Notice that the set of protocols based on LOCC is of very difficult characterization, since a local protocol may consist of arbitrarily many rounds of communication and local operations. One key result to understand this simple solution is due to Lo and Popescu[26]. They exploited the symmetry (up to local unitary operations) of any bipartite pure state under exchange of parts A and B to show that the most general LOCC-based protocol involving only pure states can be replaced with another protocol simply consisting of an *efficient* measurement on any one of the subsystems, say A, followed by local unitary operations on B.

Let us justify only theorem 16, since theorems 14 and 15 follow as particular cases. The feasibility by LOCC of a transformation depends, thus, only on whether an efficient measurement on A exists such that transforms the Schmidt coefficients $\lambda(\psi)$ of the initial state ψ into those of the final states ψ_j (recall that a local unitary operation on B cannot modify the Schmidt coefficients of the final states). That is, on whether a measurement can transform the mixed state $\rho_A(\psi)$, with spectrum $\lambda(\psi)$, into the ensemble $\{p_j, \rho_A^{\psi_j}\}$, with corresponding spectra $\lambda(\psi_j)$. Then, theorem 12 on quantum measurements implies that condition (32) must be fulfilled for such a measurement on A to exists, whereas its partial converse, theorem 13, ensures that when condition (32) is fulfilled, then the wished measurement exists. Notice, moreover, that the proof of theorem 13 is constructive, displaying and explicit measurement that we can apply on part A of the system and supplement with convenient, outcome-dependent unitary operations on part B to readily obtain a local protocol for the transformation $\psi \rightarrow \{p_j, \psi_j\}$.

4.2. *Local order, incomparable states and entanglement catalysis.*

In a similar way as Uhlmann's theorem justifies regarding a density operator ρ as more mixed than σ when $\lambda(\rho) \prec \lambda(\sigma)$, theorem 14 says that we can regard state ψ as being more entangled than state ϕ when $\lambda(\psi) \prec \lambda(\phi)$. The idea is that whatever nonlocal resources state ϕ may contain, they are also contained in ψ at least in the same amount. This is so because state ϕ can be obtained from ψ by means of a transformation that, by definition, can not enhance non-local properties.

The degree of entanglement of a bipartite pure state ψ is directly related to the degree of disorder of the corresponding reduced density matrix ρ_A. A pure state transformation through LOCC of a bipartite system AB is possible if, and only if, it implies an increase in the order of part A (equivalently, of part B). There are other properties that entangled states inherit from majorization. Recall that the later defines only a partial order. Thus, we have couples of pure states whose entanglement is incomparable, that is, neither of the states can be converted into the other by means of LOCC. And example of this are states ψ and ϕ with Schmidt coefficients given by $\lambda(\psi) = (0.5, 0.25, 0.25)$ and $\lambda(\phi) = (0.4, 0.4, 0.2)$.

A surprising result of entanglement transformations is the so called entanglement catalysis. As shown by Jonathan and Plenio [21], there are incomparable pure states ψ and ϕ such that the transformation $\psi \rightarrow \phi$ cannot be accomplished, but there is a catalyzing state τ such that $\psi \otimes \tau \rightarrow \phi \otimes \tau$ can be accomplished. Here state τ is necessary to transform ψ into ϕ by LOCC, but is not used up during the process. As an explicit example of this , consider the states

$$|\psi\rangle = \sqrt{0.4}|00\rangle + \sqrt{0.4}|11\rangle + \sqrt{0.1}|22\rangle + \sqrt{0.1}|33\rangle \quad (33)$$

$$|\phi\rangle = \sqrt{0.5}|00\rangle + \sqrt{0.25}|11\rangle + \sqrt{0.25}|11\rangle \quad (34)$$

$$|\tau\rangle = \sqrt{0.6}|00\rangle + \sqrt{0.4}|11\rangle. \quad (35)$$

It is then easy to check that $\lambda(\psi \otimes l) \prec \lambda(\phi \otimes l)$ while condition $\lambda(\psi) \prec \lambda(\phi)$ is not fulfilled, and thus the transformation of ψ into ϕ is only possible in presence of τ.

4.3. *Explicit conversion strategies and their classical communication cost*

Let us discuss now two explicit protocols for a LOCC deterministic conversion between pure states. We shall point out in both cases how much classical communication is required. The classical communication cost of a protocol is an important issue in some circumstances. Suppose the conversion is needed, for instance, while computing a distributed function. If the aim of using entanglement is to reduce the communication complexity of the function, it is important to have available conversion protocols requiring little classical communication. The study of the classical communication cost in entanglement transformations was initiated in [27], in the context of asymptotic transformations.

We already pointed out that the proof of theorem 13 on quantum measurements provides us with a protocol to transform bipartite pure states. This first construction goes as follows. If a deterministic conversion $\psi \rightarrow \phi$ is to be possible by means of LOCC, then we have that $\lambda(\psi) \prec \lambda(\phi)$, which by theorem 1 implies that $\lambda(\psi) = D\lambda(\phi)$ for some doubly stochastic matrix D. In turn, Birkhoff theorem says that D decomposes into n random permutations, $D = \sum_j p_j P_j$, and Caratheodory's theorem ensures that a decomposition exists with $n = d^2 - 2d + 2$. The measurement operators $\{F_j\}$, given by $F_j \equiv q_j^{1/2} (\rho_A^{\phi})^{1/2} P_j^{\dagger} (\rho_A^{\psi})^{-1/2}$, define then a n outcome measurement for Alice that, when supplemented with extra, outcome-dependent local unitary operations, deterministically transforms ψ into ϕ. This protocol requires sending about $2 \log_2 d$ bits from Bob to Alice, corresponding to communicating which of the n measurement outcomes Alice has obtained.

We examine now an alternative protocol[28] that requires only $\log_2 d$ bits of communication. It is based on exploiting the corollary of theorem 10 and constitutes yet another instance of application of majorization theory to entanglement transformations. Notice that if

$$|\psi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}|i_A\rangle|i_B\rangle \tag{36}$$

can be transformed into $|\phi\rangle$ by LOCC, that is, $\lambda(\psi) \prec \lambda(\phi)$, then we can write, up to local unitary operations,

$$|\phi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}|\phi_i\rangle|i_B\rangle, \tag{37}$$

that is, with the same coefficients $\sqrt{\lambda_i}$ as ψ, but using a set of (possibly) non-orthogonal normalized vectors $\{|\phi_i\rangle\}$ in part A.

Then the operator $F_j \equiv \sum_i \omega^{ij}|\phi_i\rangle\langle i_A|$, $\omega \equiv \exp(i2\pi/d)$, applied on part A transforms ψ into $|\phi^j\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}\omega^{ij}|\phi_i\rangle|i_B\rangle$, which becomes ϕ after a diagonal unitary operation on part B that removes the phases w^{ij}. To see that $\sum_j F_j^\dagger F_j = I$, that is, that $\{F_j\}$ define indeed a d-outcome measurement, just regard the F_j as components of a vector $\vec{F}$ and notice that

$$\vec{F} = \begin{pmatrix} F_1 \\ \vdots \\ F_d \end{pmatrix} = U \begin{pmatrix} |\phi_1\rangle\langle 1_A| \\ \vdots \\ |\phi_d\rangle\langle|d_A| \end{pmatrix} \equiv U\vec{v}, \tag{38}$$

where U is a unitary matrix with entries $u_{ij} = \omega^{ij}$ and vector $\vec{v}$ satisfies $\vec{v}^\dagger \cdot \vec{v} = I_A$. Now, $\sum_j F_j^\dagger F_j = \vec{F}^\dagger \cdot \vec{F} = (\vec{v}^\dagger U^\dagger) \cdot (U\vec{v}) = \vec{v}^\dagger \cdot \vec{v} = I_A$. Since this measurement has only d outcomes, Alice needs to communicate $\log_2 d$ bits of information to Bob so that he knows which unitary operation he is to apply on part B.

4.4. *Majorization versus entanglement monotones*

A fruitful approach to entanglement transformations, not necessarily restricted to the bipartite pure-state case, has been the identification and study of those quantities that have a monotonic behaviour under LOCC, the so called *entanglement monotones*[8]. In a similar way as conservation laws, such as that of energy and of linear and angular momenta, imply constraints in the evolution of a dynamical system, entanglement monotones are functions of the state of a composite system that set constraints LOCC transformations must satisfy.

Here we have been concerned with transformations involving only pure states. Suppose that the transformation $\psi \to \{p_j, \psi_j\}$ can be accomplished by means of LOCC. A function μ is then a (non-increasing) entanglement monotone if, for any such initial and final states, it fulfills

$$\mu(\psi) \geq \sum_j p_j \mu(\psi_j). \tag{39}$$

Let us now suppose $\mu(\psi)$ has been shown to be an entanglement monotone. Obviously, if a given pure-state transformation $\tau \to \{q_j, \tau_j\}$ results in an increase of μ, i.e. if condition (39) is not fulfilled for τ and $\{q_j, \tau_j\}$, then that transformation cannot be implemented using only LOCC. Thus, for each entanglement monotone μ that we identify, we obtain a constraint the transformation $\tau \to \{q_j, \tau_j\}$ has to fulfill if it is to be accomplished by means of LOCC. A remarkable achievement would then be to eventually identify a complete set of entanglement monotones, that is, a set $\{\mu_1, \mu_2, ...\}$ such that, given any transformation, we could assess whether it can be accomplished locally by simply checking whether none of these functions increases.

In a series of contributions by Nielsen [7], Vidal [19] and Jonathan and Plenio [20] a complete set of monotones for bipartite pure states has been identified. Consider the ordered Schmidt coefficients $\lambda_1^\downarrow \geq ... \geq \lambda_d^\downarrow \geq 0$ of a state ψ and, for each $l = 1, ..., d$, define the entanglement monotones $E_l(\psi)$ as

$$E_l(\psi) \equiv \sum_{i=l}^{n} \lambda_i^\downarrow. \tag{40}$$

Then theorem 16 can be rephrased in terms of the entanglement monotones E_l as

Theorem 17 *The pure state transformation $\psi \to \{p_j, \psi_j\}$ can be accomplished using only LOCC if, and only if,*

$$E_l(\psi) \geq \sum_j p_j E_l(\psi_j), \qquad l = 1, ..., d. \tag{41}$$

That is, a necessary and sufficient condition for the transformation $\psi \to \{p_j, \psi_j\}$ to be possible by means of LOCC is that none of the monotones E_l, $l = 1, ..., d$, increases on average during the transformation. Clearly, each of the inequalities in (41) is just one of the majorization inequalities in of Eq. (32). Nevertheless, approaching entanglement transformations from the viewpoint of entanglement monotones has the advantage that, contrary to majorization, the analysis is not restricted to bipartite pure states.

In this more general setting several generalizations of the previous results have been reported. These involve approximate transformations[29], conversion of mixed states[30] and transformations in multipartite systems[31]. Even a criterion for separability has been found in terms of majorization relations[32].

4.5. *Asymptotic transformations*

Entanglement transformations have also been studied in the so called asymptotic regime[23], where instead of a single copy of the state ψ, a large number N of copies is collectively manipulated, using LOCC, to obtain a large number of copies of ϕ. Asymptotic transformations fall beyond the scope of the present review paper. Nevertheless, we find it instructive to conclude it by making the connection between the conditions ruling non-asymptotic and asymptotic entanglement transformations.

As Bennett, Bernstein, Popescu and Tapp showed in their pionering contribution[23], in the limit of large N the state $\psi^{\otimes N}$ can be converted into the state $\phi^{\otimes M}$ with arbitrarily high fidelity if, and only if,

$$NS(\rho_A^{\psi}) \geq MS(\rho_A^{\phi}), \tag{42}$$

that is, if and only if the von Neumann entropy of the reduced density matrix for subsystem A (identically for B) is not increased during the transformation. Thus, in this asymptotic limit all inequalities involved in the majorization relation of theorem 14 collapse into a single inequality for the von Neumann entropy. This can be understood by expanding states $\psi^{\otimes N}$ and $\phi^{\otimes M}$ in their Schmidt decompositions and by noticing that for large N and M most of the majorization inequalities of Eqs. (6) become very sensitive to perturbations of the Schmidt coeffecients of $\phi^{\otimes M}$. A "small" modification of the latter is sufficient for the inequalities to be fulfilled, even if they were not fulfilled initially. Only those inequalities setting conditions similar to Eq. (42) turn out to be robust against such modifications, and thus only condition (42) matters in the large N regime, where (asymptotically perfect) approximations to $\phi^{\otimes M}$ are accepted as the output of the transformation.

Conclusions

Majorization was originally developed to make precise the notion that a probability distribution is more mixed than another one. In this paper we have described the important role majorization plays in entanglement theory, where it rules the feasibility of bipartite pure-state transformations. We have made special emphasis on the reasons why such a connection between majorization and entanglement transformations exist. We have seen that in order to transform an entangled pure state of a system AB, a local measurement on, say, part A is required. Such a measurement necessarily increases the order of the reduced density matrix of A, in accordance with the idea that by means of a measurement we learn about the system being measured. This process is also ruled by a majorization relation, and this is what accounts for the tight relation between local disorder and non-local correlations, between majorization and entanglement.

Acknowledgements

The authors thank J.I. Latorre and J.I. Cirac for helpful comments on the manuscript. The material of section proceeds from the notes of a mini-course on majorization and its applications to quantum mechanics given by MAN at the California Institute of Technology during June of 1999. GV acknowledges financial support from the European Community under contract HPMF-CT-1999-00200.

References

1. A. W. Marshall and I. Olkin. *Inequalities: theory of majorization and its applications.* Academic Press, New York, 1979.
2. R. Bathia. *Matrix analysis.* Springer-Verlag, New York, 1997.
3. T. Ando. Majorization, doubly stochastic matrices, and comparison of eigenvalues. *Linear Algebra and Its Applications*, 118:163-248, 1989. T. Ando. Majorizations and inequalities in matrix theory. *Linear Algegra and Its Applications*, 199:17-64, 1994

4. P.M. Alberti and A. Uhlmann. *Stochasticity and partial order: doubly stochastic maps and unitary mixing.* Dordrecht, Boston, 1982.
5. R.T. Rockafeller. *Convex Analysis.* Princeton University Pres, Princeton, 1970.
6. This theorem follows from two results of entanglement theory, namely Nielsen's criterion for feasibility of pure-state conversions in terms of majorization relations (theorem 1 of [7]) and Vidal's characterization of entaglement monotones (theorem 2 of [8]).
7. M. A. Nielsen. Phys. Rev. Lett. 83(2):436-439, 1999.
8. G. Vidal, Journ. Mod. Opt. 47: 355, 2000.
9. A. Uhlmann. Rep. Math. Phys. 1(2):147-159, 1970.
10. A. Wehrl Rep. Math. Phys. 6(1):15-28, 1974. A. Wehrl Rev. Mod. Phys., 50:221, 1978.
11. M. A. Nielsen. Phys. Rev. A, 63(02):022114, 2000.
12. This result has been independently discovered several different times — by Schrödinger[13], Jaynes[14], and by Hughston, Jozsa and Wootters[15].
13. E. Schrödinger. Proc. Cambridge Phil. Soc., 32:446-452, 1936.
14. E. T. Jaynes. Phys. Rev. 108(2):171-190, 1957.
15. L.P. Hughston, R. Jozsa, W. K. Wootters. Phys. Lett. A, 183:14-18, 1993.
16. This result was conjectured and proved in the forward direction by Uhlmann[9], and the proof completed by Nielsen[17]; Ruskai had noted the same result prior to the paper of Nielsen, but did not publish the result. Many elements of the proof are implicit in [15], however they did not explicitly draw the connection with majorization.
17. M. A. Nielsen. Phys. Rev. A, 62:052308, 2000.
18. C. A. Fuchs, K. Jacobs. "Information tradeoff relations for finite-strength quantum measurments", quant-ph/0009101.
19. G. Vidal, Phys. Rev. Lett. 83(5):1046-1049, 1999.
20. D. Jonathan and M. B. Plenio. Phys. Rev. Lett. 83(7):1455-1458, 1999.
21. D. Jonathan and M. B. Plenio. Phys. Rev. Lett. 83, 3566, 1999.
22. L. Hardy. Phys. Rev. A 60(3):1912-1923, 1999.
23. C.H. Bennett, H.J. Bernstein, S. Popescu, B. Schumacher. Phys. Rev A 53, 246, 1996.
24. C.H. Bennett, D.P. Divincenzo, J.A. Smolin, W.K. Wootters. Phys. Rev A 54, 3824, 1996.
25. R. Cleve, W. van Dam, M. A. Nielsen, A. Tapp, "Quantum entanglement and the communication complexity of the inner product function", quant-ph/9708018
26. H.-K. Lo, S. Popescu, "Concentrating entanglement by local operations —beyond mean values", quant-ph/9707038.
27. H.-K. Lo, S. Popescu,. Phys. Rev. Lett. 83, 1459, 1999.
28. G. Vidal and J.I. Cirac, unpublished.
29. G. Vidal, D. Jonathan, M.A. Nielsen, Phys. Rev. A 62, 012304 (2000).
30. G. Vidal, Phys. Rev. A 62, 062315 (2000). P. Hayden, B. M. Terhal, A. Uhlmann, quant-ph/0011095.
31. C.H. Bennett, S. Popescu, D. Rohrlich, J.A. Smolin, A.V. Thapliyal, Phys. Rev. A 63, 012307 (2001). W. Dür, G. Vidal, J.I. Cirac, Phys. Rev. A 62, 062314 (2000). A. Acín, E. Jané, W. Dür, G. Vidal, Phys. Rev. Lett. 85, 4811 (2000). H. Barnum, N. Linden, "Monotones and invariants for multiparticle quantum states", quant-ph/0103155. F. Verstraete, J. Dehaene, B. De Moor, "Normal forms, entanglement monotones and optimal filtration of multipartite quantum systems" quant-ph/0105090.
32. M.A. Nielsen, J. Kempe, "Separable states are more disordered globally than locally", quant-ph/0011117.