Separable approximations of density matrices of composite quantum systems

2001 J. Phys. A: Math. Gen. 34 6919

(http://iopscience.iop.org/0305-4470/34/35/318)

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

J. Phys. A: Math. Gen. **34** (2001) 6919–6937 PII: S0305-4470(01)19362-5

Separable approximations of density matrices of composite quantum systems

Siniša Karnas and Maciej Lewenstein

Institut für Theoretische Physik, Universität Hannover, D-30167 Hannover, Germany

E-mail: karnas@itp.uni-hannover.de and lewen@itp.uni-hannover.de

Received 22 November 2000
Published 24 August 2001
Online at stacks.iop.org/JPhysA/34/6919

Abstract
We investigate optimal separable approximations (decompositions) of states ϱ of bipartite quantum systems A and B of arbitrary dimensions $M \times N$ following the lines of Lewenstein and Sanpera. Such approximations allow to represent in an optimal way any density operator as a sum of a separable state and an entangled state of a certain form. For two-qubit systems ($M = N = 2$) the best separable approximation has the form of a mixture of a separable state and a projector onto a pure entangled state. We formulate a necessary condition that the pure state in the best separable approximation is not maximally entangled. We demonstrate that the weight of the entangled state in the best separable approximation in arbitrary dimensions provides a good entanglement measure. We prove for arbitrary M and N that the best separable approximation corresponds to a mixture of separable and entangled states, both of which are unique. We develop also a theory of optimal separable approximations for states with positive partial transpose (PPT states). Such approximations allow to decompose any density operator with positive partial transpose as a sum of a separable state and an entangled PPT state. We discuss procedures for constructing such decompositions.

PACS numbers: 03.67.Hk, 03.65.Bz, 03.67.−a, 89.70.+c

1. Introduction

The problem of characterization of entangled states of composite quantum systems is one of the fundamental open problems of quantum theory. Entanglement is one of the quantum properties which make quantum mechanics so fascinating: it leads to famous apparent paradoxes [1, 2], and it is of great importance for applications in quantum communication and information processing [3].

In the case of pure states it is easy to check whether a given state is or is not entangled. So far, the answer to this question when applied to quantum mixtures is not known in general. The definition (introduced by Werner [4]) says that a state (in general a mixed state) is entangled when it is not separable. Separable states defined on a Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ are those that can be described as a convex combination of projections onto product states

$$\varrho = \sum_{i=1}^{K} p_i \left|e_A^i, f_B^i\right\rangle\left\langle e_A^i, f_B^i\right| \qquad \sum_i p_i = 1. \tag{1}$$

In finite-dimensional spaces, the number of terms in the sum can be restricted to $K \leqslant \dim(\mathcal{H}_{AB})^2$ (in other words, when the density matrix is separable, then it can be represented in the above form with K terms, where K is not larger than the dimension of the space of linear operators acting in $\mathcal{H}_{AB}$, see [5]).

Several necessary conditions for separability are known: Werner's condition based on the mean value of the, so-called, flipping operator [4], Horodecki's criterion based on α-entropy inequalities [6], and many others [7, 8]. Perhaps, the most important necessary criterion has been formulated by Peres [9] who has demonstrated that the partial transpose ϱ^{T_A} of any separable matrix ϱ defined as $\langle m, \mu|\varrho^{T_A}|n, \nu\rangle = \langle n, \mu|\varrho|m, \nu\rangle$ for any fixed orthonormal product basis $|n, \nu\rangle \equiv |e_n\rangle_A \otimes |e_\nu\rangle_B$ must be positively defined. In the following we will call states with positive partial transpose PPT states. The physical meaning of the PPT property is that for a PPT state the time reversal operation in one subsystem (either Alice's or Bob's) is physically sound [8, 10].

It is worth stressing that the problem of separability is directly related to the theory of positive maps on C^*-algebras [11, 12]. This has been established in [13] in which it was shown, in particular, that for systems of low dimensions ($M \times N \leqslant 6$) the PPT condition is also sufficient for separability. For systems of higher dimensions ($M \times N > 6$) there exist entangled states having the PPT property. The first examples of such were provided by means of the so-called range separability criterion based on analysis of the range of the density matrix [5] (see also [11]). Such states represent bound entanglement, i.e. cannot be distilled [14].

In a recent Letter [15] we have also looked at the range of the entangled density operators in order to formulate an algorithm of optimal decomposition of mixed states into the separable and inseparable parts. Our method of the *best separable approximations* (BSA) was based on subtracting projections on product vectors from a given density matrix in such a way that the remainder remained positively defined. This approach allowed to achieve a variety of very strong results: optimal decompositions with minimal number of terms in the form of mixtures and pseudo-mixtures for 2×2 and 2×3 systems [16], separability criteria for $2 \times N$ systems [17], and in general for $M \times N$ systems (with $M \leqslant N$) [18] for density matrices of low ranks. In particular it was shown that: (i) all PPT states of rank smaller than N are separable; (ii) for generic states such as $r(\varrho) + r(\varrho^{T_A}) \leqslant MN - M - N + 2$ constructive separability criteria were given that reduce the problem to finding roots of some complex polynomials; and (iii) for $2 \times N$ it was shown that for the states invariant under partial transpose with respect to the two-dimensional subsystem, and those that are not 'very different' from their partial transpose are necessarily separable. Very recently, these findings have allowed us to present general schemes of constructing non-decomposable entanglement witnesses (i.e. observables that have a positive mean value on all separable states, and have a negative mean value on a PPT entangled state [19, 20]) and non-decomposable positive maps in arbitrary dimensions, i.e. the maps that cannot be decomposed into a sum of a completely positive map and another completely positive map combined with the transposition [21]. It should be stressed that our approach goes beyond the methods of constructing examples of PPT entangled states and positive maps based on the so-called unextendible product bases [19, 22]. More importantly, we were able to present methods of constructing optimal entanglement witnesses and optimal non-decomposable maps which provide very strong separability criteria [21]. In a series of important papers, Englert and his collaborators have obtained remarkable analytic results concerning the BSA decompositions for 2×2 systems [23]. These results give a new deep insight into the fundamental problem of quantum correlations in two-qubit systems.

All of the above-mentioned applications indicate that the BSA method is very useful. The aim of this paper is to generalize and to complete results of the [15]. We present several results that characterize the BSA decompositions in 2×2 and, in general in $M \times N$ systems. Concerning the two-qubit systems our results are complementary to those of [23]. The plan of the paper is as follows: In section 2 we remind the reader of some basic facts about the optimal and the best separable approximations. In section 3 (using also the results presented in the appendix) we demonstrate a necessary condition that for a two-qubit systems ($M = N = 2$) the best separable approximation has the form of a mixture of a separable state and a projector onto an entangled state which is not *maximally entangled*. In section 4 we remind the reader of the basic facts about entanglement measures; we prove here that the weight of the fully entangled state in the BSA decomposition of $M \times N$ states provides a legitimate entanglement measure. In section 5 we prove that in general for arbitrary M and N the best separable approximation corresponds to a mixture of separable and entangled states, both of which are *uniquely* determined. Finally, in section 6 we formulate the theory of optimal separable approximations for states with positive partial transpose (PPT states). Such approximations allow to represent any density operator with positive partial transpose as a sum of a separable state and an entangled PPT state. Decompositions of this sort play an essential role in the theory of non-decomposable positive maps [21]. We present and discuss efficient numerical procedures of construction of such decompositions.

2. Introduction to BSA

Consider a state ρ acting on $\mathcal{C}^M \otimes \mathcal{C}^N$. Such a state will be called a PPT state if its partial transpose satisfies $\rho^{T_A} \geqslant 0$ (or equivalently $\rho^{T_B} \geqslant 0$). Throughout this paper $K(X)$, $R(X)$, $k(X)$, and $r(X)$ denote the kernel, the range, the dimension of the kernel, and the rank of the operator X, respectively. By $|e^*\rangle$ we will denote the complex conjugated vector of $|e\rangle$ in the basis $|0\rangle_A, |1\rangle_A, \ldots$ in which we perform the partial transposition in Alice's space; that is, if $|e\rangle = \alpha|0\rangle + \beta|1\rangle + \cdots$ then $|e^*\rangle = \alpha^*|0\rangle + \beta^*|1\rangle + \cdots$. Similar notation will be used for vectors in Bob' s space. By $|\hat{e}\rangle$ we denote a vector orthogonal to $|e\rangle$.

In this section we give a short repetition of what we call optimal and the best separability approximations (OSA and BSA respectively). Although the results below have been proven in [15], we repeat them here using the notation of the present work. The idea of BSA is that, because of the fact that a set of separable states is compact, for any density matrix ρ there exist a 'optimal' separable matrix ρ_s^* and 'optimal' $\Lambda \geqslant 0$ such that $\Lambda\rho_s$ can be subtracted from ρ maintaining the positivity of the difference, $\rho - \Lambda\rho_s^* \geqslant 0$. This situation is characterized by the following theorem:

Theorem 1. *For any density matrix ρ (separable, or not) and for any (fixed) countable set V of product vectors belonging to the range of ρ, i.e. $|e_\alpha, f_\alpha\rangle \in R(\rho)$, there exist $\Lambda(V) \geqslant 0$ and a separable matrix*

$$\rho_s^*(V) = \sum_\alpha \Lambda_\alpha P_\alpha \tag{2}$$

where $P_\alpha = |e_\alpha, f_\alpha\rangle\langle e_\alpha, f_\alpha|$, while all $\Lambda_\alpha \geqslant 0$, such that $\delta\rho = \rho - \Lambda\rho_s^* \geqslant 0$, and that $\rho_s^*(V)$ provides the optimal separable approximation (OSA) to ρ since $\mathrm{Tr}(\delta\rho)$ is minimal or, equivalently, Λ is maximal. There exists also the best separable approximation ρ_s^* for which $\Lambda = \max_V \Lambda(V)$. Obviously, $\Lambda(V) \leqslant \Lambda(V')$ when $V' \subset V$.

Remark 1. Quite generally one can define the best separable approximations ρ_s of ρ by demanding that $\|\rho - \rho_s\|$ is minimal with respect to some norm in the (Banach) space of operators. Here we minimize $\mathrm{Tr}(\rho - \lambda\rho_s)$ with respect to all ρ_s such that $\rho - \lambda\rho_s \geqslant 0$.

From this theorem it follows then that if any density matrix ρ is separable, then $\Lambda = 1$. Caratheodory's theorem implies then (see discussion in [5]) that there exists a finite set of product vectors $V \subset R(\rho)$ of cardinality $\leqslant r(\rho)^2$ for which the optimal separable approximation to ρ, $\rho_s^*[V]$, is equal to the BSA and $\Lambda = 1$ also. Theorem 1 and Caratheodory's theorem are also true for uncountable families of states V, and appropriate generalizations are discussed in [20, 21].

In order to explain now how the procedure of construction of the matrix ρ_s^* actually works, we introduce two important concepts:

Definition 1. *A non-negative parameter Λ is called* **maximal** *with respect to a (not necessarily normalized) density matrix ρ, and the projection operator $P = |\psi\rangle\langle\psi|$ if $\rho - \Lambda P \geqslant 0$, and for every $\epsilon \geqslant 0$, the matrix $\rho - (\Lambda + \epsilon)P$ is not positive definite.*

This means that Λ determines the maximal contribution of P that can be subtracted from ρ maintaining the non-negativity of the difference. Now we have the following important lemma:

Lemma 1. *Λ is maximal with respect to ρ and $P = |\psi\rangle\langle\psi|$, if: (a) $|\psi\rangle \notin R(\rho)$ then $\Lambda = 0$, and (b) $|\psi\rangle \in R(\rho)$ then*

$$0 \leqslant \Lambda = \frac{1}{\langle\psi|\rho^{-1}|\psi\rangle}. \tag{3}$$

Note that in the case (b) the expression on the RHS of equation (3) makes sense, since $|\psi\rangle \in R(\rho)$, and therefore there exists $|\phi\rangle$ such that $|\psi\rangle = \rho|\phi\rangle$, or equivalently that $\rho^{-1}|\psi\rangle = |\phi\rangle$. Remarkably, this lemma has been used in a completely different context by Jaynes in his works on the foundations of statistical mechanics [24].

Definition 2. *A pair of non-negative (Λ_1, Λ_2) is called* maximal *with respect to ρ and a pair of projection operators $P_1 = |\psi_1\rangle\langle\psi_1|$, $P_2 = |\psi_2\rangle\langle\psi_2|$, if $\rho - \Lambda_1 P_1 - \Lambda_2 P_2 \geqslant 0$, Λ_1 is maximal with respect to $\rho - \Lambda_2 P_2$ and to the projector P_1, Λ_2 is maximal with respect to $\rho - \Lambda_1 P_1$ and to the projector P_2, and the sum $\Lambda_1 + \Lambda_2$ is maximal.*

The condition for the maximality of $\Lambda_1 + \Lambda_2$ is given by the following lemma:

Lemma 2. *A pair (Λ_1, Λ_2) is maximal with respect to ρ and a pair of projectors (P_1, P_2) if:*

- *(a) if $|\psi_1\rangle$, $|\psi_2\rangle$ do not belong to $R(\rho)$ then $\Lambda_1 = \Lambda_2 = 0$;*
- *(b) if $|\psi_1\rangle$ does not belong to $R(\rho)$, while $|\psi_2\rangle \in R(\rho)$ then $\Lambda_1 = 0$, $\Lambda_2 = \langle\psi_2|\rho^{-1}|\psi_2\rangle^{-1}$;*
- *(c) if $|\psi_1\rangle, |\psi_2\rangle \in R(\rho)$ and $\langle\psi_1|\rho^{-1}|\psi_2\rangle = 0$, then $\Lambda_i = \langle\psi_i|\rho^{-1}|\psi_i\rangle$, $i = 1, 2$;*
- *(d) if $|\psi_1\rangle, |\psi_2\rangle \in R(\rho)$ and $\langle\psi_1|\rho^{-1}|\psi_1\rangle, \langle\psi_2|\rho^{-1}|\psi_2\rangle \geqslant |\langle\psi_1|\rho^{-1}|\psi_2\rangle| \neq 0$ then*
$$\Lambda_1 = \left(\langle\psi_2|\rho^{-1}|\psi_2\rangle - \left|\langle\psi_1|\rho^{-1}|\psi_2\rangle\right|\right)/D \tag{4}$$
$$\Lambda_2 = \left(\langle\psi_1|\rho^{-1}|\psi_1\rangle - \left|\langle\psi_2|\rho^{-1}|\psi_1\rangle\right|\right)/D \tag{5}$$
where $D = \langle\psi_1|\rho^{-1}|\psi_1\rangle\langle\psi_2|\rho^{-1}|\psi_2\rangle - |\langle\psi_1|\rho^{-1}|\psi_2\rangle|^2$;
- *(e) finally, if $|\psi_1\rangle, |\psi_2\rangle \in R(\rho)$ and $\langle\psi_1|\rho^{-1}|\psi_1\rangle \geqslant |\langle\psi_1|\rho^{-1}|\psi_2\rangle| \geqslant \langle\psi_2|\rho^{-1}|\psi_2\rangle$, then $\Lambda_1 = \langle\psi_1|\rho^{-1}|\psi_1\rangle^{-1}$, $\Lambda_2 = 0$.*

Note that the Schwarz inequality implies that $D \geqslant 0$. We are in the position now to present the basic BSA theorem:

Theorem 2. *Given the set V of product vectors* $|e_\alpha, f_\alpha\rangle \in R(\rho)$, *the matrix* $\rho_s^* = \sum_\alpha \Lambda_\alpha P_\alpha$ *is the optimal separable approximation* (*OSA*) *of* ρ *if*:

- *all* Λ_α *are maximal with respect to* $\rho_\alpha = \rho - \sum_{\alpha'\neq\alpha} \Lambda_{\alpha'} P_{\alpha'}$, *and to the projector* P_α;
- *all pairs* $(\Lambda_\alpha, \Lambda_\beta)$ *are maximal with respect to* $\rho_{\alpha\beta} = \rho - \sum_{\alpha'\neq\alpha,\beta} \Lambda_{\alpha'} P_{\alpha'}$, *and to the projection operators* (P_α, P_β).

If V is the set of all product vectors in $R(\rho)$ (in general uncountable) then the same theorem holds for the BSA (for the detailed proof see [15, 20, 21]). All information about entanglement is included in the matrix $\delta\rho$. If $\delta\rho$ does not vanish, i.e. if ρ is not separable, the range $R(\delta\rho)$ cannot contain any product vector. The reason is that one can use projectors on product vectors that belong to $R(\delta\rho)$ in order to increase Λ. The rank of the matrix $\delta\rho$ must be smaller, or equal to $(M-1)(N-1)$. This is because the set of all product vectors in the Hilbert space H of dimension $M \times N$ spans a $(N+M-1)$-dimensional manifold, which generically has a non-vanishing intersection with linear subspaces of H of dimension larger than $(N-1)\times(M-1)$. In fact, we have proven rigorously that this is the case for $2 \times N$ systems in [17], and presented some rigorous arguments for the case $M \times N$ in [18].

In particular, for the case of $M = N = 2$, $\delta\rho$ is a simple projector onto an entangled state. For two-qubit systems it is easy to prove that the BSA decomposition is unique and has the form:

$$\rho = \Lambda\rho_s + (1-\Lambda)P_e \qquad \Lambda \in [0, 1] \tag{6}$$

where ρ_s is the normalized density matrix. If it had not been so, we could have another BSA expansion, say $\rho = \Lambda\tilde{\rho}_s + (1-\Lambda)\tilde{P}_e$. But, taking the convex combination of these two decompositions, we obtain another BSA decomposition with the remainder $\delta\rho$ being given by a convex combination of P_e and $\tilde{P}_e$. Such a remainder would have then rank 2, and would necessarily contain product vectors in its range [16]. If this happened, we would be then able to increase the BSA parameter Λ by subtracting from $\delta\rho$ projectors on product vectors in its range. That is, however, impossible since Λ is already maximal. For the case of arbitrary dimensions the OSA and BSA decompositions are also unique. We present the proof of this fact in section 5 of this paper.

3. The BSA reminder in $\mathcal{C}^2 \otimes \mathcal{C}^2$ quantum systems: is it maximally entangled?

We have seen that the BSA reminder in $\mathcal{C}^2 \otimes \mathcal{C}^2$ quantum systems is just given by a projector onto an entangled state $|\psi_e\rangle$. This fact is essential and allows to obtain the BSA decomposition for some states analytically [23]. For many families of states considered by Englert and his collaborators, the BSA remainder consists of a maximally entangled state. Similar conclusions follow from the numerical analysis of [15]. In this section we ask therefore a natural question: under which conditions is, or is not, the BSA remainder maximally entangled? Strictly speaking we present here a necessary condition, that the BSA decomposition for a generic density matrix must fulfil, so that the BSA remainder is not maximally entangled.

We concentrate here on generic quantum states which have the maximal dimension of the range $(r(\rho) = r(\rho^{T_A}) = 4)$. Let us assume that the density matrix ρ has the BSA decomposition

$$\rho = \Lambda\rho_s + (1-\Lambda)P_{\psi_e} \tag{7}$$

so that its partial transposition with respect to Alice's system is $\rho^{T_A} = \Lambda\rho_s^{T_A} + (1-\Lambda)P_{\psi_e}^{T_A}$. When Λ is not equal to 1, ρ is entangled, and ρ^{T_A} must not be positive definite.

Let us first observe:

Lemma 3. *If ρ acting in $\mathcal{C}^2 \otimes \mathcal{C}^2$ has the BSA decomposition $\rho = \Lambda\rho_s + (1-\Lambda)P_{\psi_e}$, then $r(\rho_s^{T_A}) \leqslant 3$.*

Proof. Had the range of $\rho_s^{T_A}$ been full, one could always replace $1-\Lambda$ by $(1-\Lambda-\epsilon)$, keeping $\Lambda\rho_s^{T_A} + \epsilon P_{\psi_-}^{T_A}$ positive definite, while $\rho_s' = \rho_s + \epsilon P_{\psi_-}$ separable. □

The fact that the rank of ρ^{T_A} is not full implies that $\exists|v\rangle$, such that $\rho^{T_A}|v\rangle = 0$. Since $P\psi^{T_A}$ has three positive eigenvalues and one negative eigenvalue [16], where the eigenvector corresponding to a negative eigenvalue in a conveniently chosen basis can be written as

$$\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = |\nu_-\rangle$$

then $\langle v|\psi_-\rangle \neq 0$. If it was not the case, one could also replace $1-\Lambda$ by $(1-\Lambda-\epsilon)$, keeping $\Lambda_{\rho_s}^{T_A} + \epsilon P_{\psi_-}^{T_A}$ positive.

Let us now discuss the optimization procedure, that sometimes allow to increase Λ in the decomposition (7). A given decomposition of such a form is optimal if it cannot be optimized. It will turn out that the optimization strategy works only provided ψ_e is not maximally entangled. The necessary condition, that the BSA remainder is not maximally entangled, is that the decomposition cannot be optimized in the sense formulated below. Our aim is to formulate this necessary condition in an explicit form in this section.

Optimization procedure. Let us observe that we can always write

$$|\psi_e\rangle = N_1|e_1, f_1\rangle + N_2|e_2, f_2\rangle$$

for any basis $|e_1\rangle$, $|e_2\rangle$, where $\langle e_1|e_1\rangle = \langle e_2|e_2\rangle = 1$, but $\langle e_1|e_2\rangle$ does not have to be zero. Let $|\hat{e}_1\rangle$, $|\hat{e}_2\rangle$ denote the basis bi-orthogonal to $|e_1\rangle$, $|e_2\rangle$; we obtain then

$$\langle\hat{e}_1|\psi_e\rangle = N_2\langle\hat{e}_1|e_2\rangle|f_2\rangle$$
$$\langle\hat{e}_2|\psi_e\rangle = N_1\langle\hat{e}_2|e_1\rangle|f_1\rangle.$$

Requiring that $\langle f_1|f_2\rangle = \langle f_2|f_2\rangle = 1$, the above equations allow to determine uniquely N_1, N_2, $|f_1\rangle$ and $|f_2\rangle$. Without losing generality, we may assume $N_1 \geqslant N_2$. Let us introduce

$$|\psi_e(\alpha)\rangle = \frac{1}{N(\alpha)}\left(\alpha N_1|e_1, f_1\rangle + \frac{1}{\alpha}N_2|e_2, f_2\rangle\right)$$

where

$$N(\alpha)^2 = \alpha^2 N_1^2 + \frac{1}{\alpha^2}N_2^2 + 2N_1N_2\mathrm{Re}(\langle e_1|e_2\rangle\langle f_1|f_2\rangle).$$

We can now rewrite the BSA projector

$$P_{\psi_e} = N(\alpha)^2 P_{\psi_e(\alpha)} + (1-\alpha^2)P_{e_1f_1} + \left(1 - \frac{1}{\alpha^2}\right)P_{e_2f_2}. \tag{8}$$

We would like to replace the projector P_{ψ_e} by the expression (8), and in this way improve the BSA decomposition. To this aim we require that $N(\alpha)^2 \leqslant 1$ which implies that $\alpha^2N_1^2 + \frac{1}{\alpha^2}N_2^2 \leqslant N_1^2 + N_2^2$. Defining now $x \equiv N_2^2/N_1^2$, we see that $N(\alpha)^2 < 1$ provided $x < \alpha^2 < 1$. That is only possible if $N_1 \neq N_2$. The latter conditions can be fulfilled if ψ_e is not maximally entangled, as described in the following lemma:

Lemma 4. *Iff $|\psi_e\rangle = N_1|e_1, f_1\rangle + N_2|e_2, f_2\rangle$, where $\langle e_1|e_1\rangle = \langle e_2|e_2\rangle = 1$, then $N_1 = N_2$ if ψ_e is maximally entangled.*

Proof. Let us consider a basis in which $|\psi_e\rangle = a|00\rangle + \sqrt{1-a^2}|11\rangle$, and assume a general form of

$$|\hat{e}_1\rangle = \begin{pmatrix} \sqrt{p} \\ \sqrt{1-p}\mathrm{e}^{\mathrm{i}\varphi} \end{pmatrix} \qquad |\hat{e}_2\rangle = \begin{pmatrix} \sqrt{p'} \\ \sqrt{1-p'}\mathrm{e}^{\mathrm{i}\varphi'} \end{pmatrix}.$$

In the basis considered we can easily calculate that

$$\langle\hat{e}_1|\psi_e\rangle = a\sqrt{p}|0\rangle + \sqrt{1-a^2}\sqrt{1-p}|1\rangle\mathrm{e}^{-\mathrm{i}\varphi} \tag{9}$$

$$\langle\hat{e}_2|\psi_e\rangle = a\sqrt{p'}|0\rangle + \sqrt{1-a^2}\sqrt{1-p'}|1\rangle\mathrm{e}^{-\mathrm{i}\varphi'} \tag{10}$$

so that

$$N_2^2|\langle\hat{e}_1|e_2\rangle|^2 = a^2p + (1-a^2)(1-p) \tag{11}$$

$$N_1^2|\langle\hat{e}_2|e_1\rangle|^2 = a^2p' + (1-a^2)(1-p'). \tag{12}$$

Note that $|\langle\hat{e}_1|e_2\rangle|^2 = |\langle\hat{e}_2|e_1\rangle|^2$, so that indeed $N_1^2 = N_2^2$, if $a^2 = \frac{1}{2}$, that is when the state $|\psi_e\rangle$ is maximally entangled. □

Now we can easily prove:

Lemma 5. *If ρ has the BSA decomposition* (7), *then either ψ_e is maximally entangled, or* $r(\rho_s) = 3$.

Proof. Suppose that $r(\rho_s) = 4$. If ψ_e is not maximally entangled, the optimization procedure allows to optimize the decomposition by taking $\alpha^2 < 1$, but very close to 1. We can indeed improve BSA for ρ, provided we can subtract $\frac{1-\alpha^2}{\alpha^2}P_{e_2^* f_2}$ from $\Lambda\rho_s^{T_A}$. This means that $|e_2^*, f_2\rangle$ must belong to the range $R(\rho_s^{T_A})$. That in turn requires that if $|v\rangle = |\hat{e}_1^*, h_1\rangle + |\hat{e}_2^*, h_2\rangle$, we then need $\langle h_1|f_2\rangle = 0$, or in other words

$$\langle v|e_2^*\rangle\langle\hat{e}_1|\psi_e\rangle = 0. \tag{13}$$

It is easy to see that this equation has many solutions: for example, take $|e_2\rangle = |\hat{e}_1\rangle$ and $|\hat{e}_1\rangle$ proportional to $\begin{pmatrix} 1 \\ \alpha \end{pmatrix} = |0\rangle + \alpha|1\rangle$, then the above equation implies that $[\langle v|0\rangle + \alpha^*\langle v|1\rangle]$ $[\langle 0|\psi_e\rangle + \alpha^*\langle 1|\psi_e\rangle] = 0$, which is a quadratic equation for α^* which obviously has solutions for $|e_2\rangle \neq |\hat{e}_1\rangle$. We conclude that either $r(\rho_s) = 3$, or $N_1 = N_2$. The latter can occur if and only if $|\psi_e\rangle$ is fully entangled. □

Therefore we have to consider the case $r(\rho_s) = r(\rho_s^{t_A}) = 3$. From the results presented in appendix A we know that there exists such a one-dimensional family of product states $|e_2(\delta), f_2(\delta)\rangle$, where δ is real, such that $|e_2(\delta), f_2(\delta)\rangle \in R(\rho_s)$ and $|e_2^*(\delta), f_2(\delta)\rangle \in R\left(\rho_s^{T_A}\right)$ is satisfied.

Now we are in the situation where we can explicitly check whether the vector $|\psi_e\rangle$ in the BSA remainder can be non-maximally entangled. If $|\psi_e\rangle$ is given and we have $|e_2, f_2\rangle = |e(\delta), f(\delta)\rangle$ for a given ρ_s, then we can calculate $|f_1\rangle$ and $|e_1\rangle$ by

$$|f_1\rangle = \frac{\langle\hat{e}_2|\psi_e\rangle}{|\langle\hat{e}_2|\psi_e\rangle|} \tag{14}$$

$$|e_1\rangle = \frac{\langle\hat{f}_2|\psi_e\rangle}{|\langle\hat{f}_2|\psi_e\rangle|} \tag{15}$$

and from $\langle f_1|f_1\rangle = 1$, we obtain $|N_1| = \frac{|\langle\hat{e}_2|\psi_e\rangle|}{|\langle\hat{e}_2|e_1\rangle|}$. Since we know now $|e_1\rangle, |f_1\rangle$, we can also easily calculate $|N_2| = \frac{|\langle\hat{e}_1|\psi_e\rangle|}{|\langle\hat{e}_1|e_2\rangle|}$.

We see that the coefficient N_1 and N_2 can be explicitly constructed from ρ_s and $|\psi_e\rangle$. We obtain therefore the main result of this section:

Theorem 3. *If a generic* ($r(\rho) = r(\rho^{T_A}) = 4$) *state* ρ *in* $\mathcal{C}^2 \otimes \mathcal{C}^2$ *has the BSA decomposition* $\rho = \Lambda\rho_s + (1-\Lambda)P_{\psi_e}$, *then either* ψ_e *is maximally entangled, or* $r(\rho_s) = r(\rho_s^{T_A}) = 3$, *and for any expansion of* $|\psi_e\rangle = N_1|e_1, f_1\rangle + N_1|e_2, f_2\rangle$, *such that* $|e_2, f_2\rangle \in R(\rho_s)$ *and* $|e_2^*, f_2\rangle \in R(\rho_s^{T_A})$ *holds, it must follow that* $N_1 < N_2$.

Proof. The proof is obvious using the lemmas of this section, and the optimization procedure. If there exist $|e_2(\delta), f_2(\delta)\rangle$ such that $N_1 > N_2$, the optimization procedure can be applied, which contradicts the optimality of the BSA. □

4. Entanglement measures

Before we turn to the main results of this paper let us also remind the reader in this section of some basic facts about entanglement measures and their properties.

Once one has the physical picture of entanglement as a resource, one needs to formulate this concept mathematically. One way leads through a definition of non-entangled, i.e. separable, states as discussed in previous sections. Another possibility is to try to quantify the amount of entanglement for a given mixed state. The latter approach is realized by defining entanglement measures [25], and by specifying physical properties which the entanglement measure should have. There are several versions of definitions of the entanglement measures; here we follow the approach of Plenio and Verdal [26]:

Definition 3. *Let* ρ *be a quantum state acting in a Hilbert space* $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, *then the function* $E(\rho) \mapsto R$ *is called an* entanglement measure *if it satisfies*:

1. $E(\rho) = 0$, *if* ρ *is separable*
2. *Local unitary operations leave* $E(\rho)$ *invariant, i.e.* $E(\rho) = E(U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger)$;
3. *Let* $\sum_i A_i A_i^\dagger \otimes B_i B_i^\dagger = 1$ *be some complete local measurement* (*i.e. local positive operator-valued map* (*POVM*)), *then*
$$E(\rho) \geqslant \sum_i \mathrm{Tr}(\rho_i) E(\rho_i/\mathrm{Tr}(\rho_i)) \tag{16}$$
where $\rho_i := A_i \otimes B_i \rho A_i^\dagger \otimes B_i^\dagger$. *This property means that entanglement measure cannot increase in the mean under local operations.*
4. *For pure states the measure of entanglement should reduce to the entropy of entanglement, which is defined as the von Neumann entropy of the reduced density matrix,* $\rho_A = \mathrm{Tr}_B\rho$ (*or, alternatively,* $\rho_B = \mathrm{Tr}_A\rho$)
$$E(\rho) := -\mathrm{Tr}(\rho_A \ln \rho_A); \tag{17}$$
5. *The entanglement measure should be additive which means that*
$$E(\rho_1 \otimes \rho_2) = E(\rho_1) + E(\rho_2). \tag{18}$$

It should be pointed out that the necessity of the last two conditions is still disputed in the literature [27, 28], and therefore we will just concentrate on the first three conditions. Notice, that in equation (3) it may happen that $E(\rho_i/\mathrm{tr}(\rho_i)) \geqslant E(\rho)$.

To complete this section, let us list some of the most widely used entanglement measures. Typically, they fulfil some, but not all, of conditions 1–5 of definition 3.

1. *Entanglement of formation* [25] is defined as
$$E_F := \min \sum_i p_i S(\rho_A^i) \tag{19}$$

where $S(\rho_A) := -\text{Tr}(\rho_A \ln \rho_A)$ is the von Neumann entropy and the minimum is taken over all the possible realizations of the state, $\rho = \sum_i |\psi_i\rangle\langle\psi_i|$, where $\rho_A^i = \text{Tr}_B(|\psi_i\rangle\langle\psi_i|)$. Notice that in the case where ρ is a pure state ($\rho = |\psi\rangle\langle\psi|$), the von Neumann entropy of the reduced density matrix is an entanglement measure. The physical meaning of the formation measure is the minimal amount of pure state entanglement needed to create the given entangled state. Calculation of E_F for a given state is a very difficult task. Remarkably, Wootters has derived the analytic formula for E_F for an arbitrary two-qubit state [29].

2. *Relative entropy entanglement measure* [26] is defined as

$$E(\rho) := \min_{\rho_s} E(\rho\|\rho_s) \tag{20}$$

where the minimum is taken over all separable states ρ_s and $E(\rho\|\rho_s)$ is the relative entropy, which is given by the expression

$$E(\rho\|\rho_s) := \text{Tr}(\rho(\ln\rho - \ln\rho_s)). \tag{21}$$

3. *Bures entanglement measure* [25] is defined as

$$E(\rho) := \min_{\rho_s}(2 - 2\sqrt{F(\rho, \rho_s)}) \tag{22}$$

where $F(\rho, \rho_s)$ is Uhlmann's fidelity $F(\rho, \rho_s) := (\text{Tr}(\sqrt{\sqrt{\rho}\rho_s\sqrt{\rho}}))^2$. This entanglement measure does not fulfil the last two conditions of definition 3.

In recent years a very promising approach has been initiated by Vidal who has shown that more parameters (the so-called entanglement monotones) are required in order to quantify completely the non-local character of bipartite pure states [28].

5. The BSA entanglement measure

Let us now investigate how the local POVM's influence a given BSA decomposition. To this aim we consider a POVM of the form of $\sum_i V_i V_i^\dagger = 1$, $V_i = A_i \otimes B_i$. After the ith result is obtained in the measurement we obtain the following density matrix:

$$\begin{aligned}\rho_i &:= \frac{V_i\rho V_i^\dagger}{\text{Tr}(V_i\rho V_i^\dagger)}\\ &= \Lambda\frac{\text{Tr}(V_i\rho_s V_i^\dagger)}{\text{Tr}(V_i\rho V_i^\dagger)}\sum_\alpha \frac{\Lambda_\alpha \text{Tr}(V_i P_\alpha V_i^\dagger)}{\text{Tr}(V_i\rho_s V_i^\dagger)}\frac{V_i P_\alpha V_i^\dagger}{\text{Tr}(V_i P_\alpha V_i^\dagger)}\\ &\quad + (1-\Lambda)\left(\frac{\text{Tr}(V_i\delta\rho V_i^\dagger)}{\text{Tr}(V_i\rho V_i^\dagger)}\right)\left(\frac{V_i\delta\rho V_i^\dagger}{\text{Tr}(V_i\delta\rho V_i^\dagger)}\right).\end{aligned}$$

Defining now

$$\begin{aligned}\Lambda_i &:= \Lambda\frac{\text{Tr}(V_i\rho_s V_i^\dagger)}{\text{Tr}(V_i\rho V_i^\dagger)}\\ \Lambda_{i\alpha} &:= \Lambda_\alpha\frac{\text{Tr}(V_i P_\alpha V_i^\dagger)}{\text{Tr}(V_i\rho_s V_i^\dagger)}\\ P_{i\alpha} &:= \frac{V_i P_\alpha V_i^\dagger}{\text{Tr}(V_i P_\alpha V_i^\dagger)}\\ \delta\rho_i &:= \frac{V_i\delta\rho V_i^\dagger}{\text{Tr}(V_i\delta\rho V_i^\dagger)}\end{aligned}$$

we rewrite the result as

$$V_i \rho V_i^\dagger \to \rho_i = \Lambda_i \sum_\alpha \Lambda_{i\alpha} P_{i\alpha} + (1 - \Lambda_i)\delta\rho_i.$$

We observe that

$$1 - \Lambda = \sum_i (1 - \Lambda_i \mathrm{Tr}(V_i \rho V_i^\dagger)) \tag{23}$$

holds. Since for the BSA decomposition of ρ_i the inequality

$$\Lambda_{BSA_i} \geqslant \Lambda_i \tag{24}$$

holds, we get from (23) that

$$1 - \Lambda \geqslant \sum_i (1 - \Lambda_{BSA_i} \mathrm{Tr}(V_i \rho V_i^\dagger)). \tag{25}$$

The result (25) allows to prove the following property:

Property 1. The BSA entanglement measure

$$E(\rho) = 1 - \Lambda_{BSA}(\rho)$$

fulfils the properties 1–3 of definition 3.

Proof.

1. If ρ is separable, i.e. $\rho = \rho_s$, then $\Lambda = 1$, and $E(\rho) = 1 - \Lambda = 0$.
2. If $\tilde{\rho} = U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger$ then obviously $E(\tilde{\rho}) \geqslant 1 - \Lambda = E(\rho)$, and *vice versa*, since we can invert $U_A \otimes U_B$. That means that $E(\rho)$ is invariant with respect to local unitary transformations.
3. Finally, if we apply a local POVM, we obtain

$$E(\rho) = 1 - \Lambda \geqslant \sum_i (1 - \Lambda_{BSA_i} \mathrm{Tr}(V_i \rho V_i^\dagger))$$
$$\geqslant \sum_i E(\rho_i) \mathrm{Tr}(V_i \rho V_i^\dagger)$$

where $\rho_i = V_i \rho V_i^\dagger / \mathrm{Tr}(V_i \rho V_i^\dagger)$. This follows from (25). □

It is worth noticing that the above argument holds for the Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$ of arbitrary dimensions.

6. The uniqueness of the BSA

In this section we turn back to the general case and present a proof that the BSA in any Hilbert space is unique. To this aim we prove first a lemma, and then the major result.

Lemma 6. *Let a Hermitian density matrix ρ have a decomposition of the form $\rho = \Lambda\rho_s + (1 - \Lambda)\delta\rho$, where ρ_s is the separable part which has the structure $\rho_s = \Lambda \sum_{\alpha=1}^n \Lambda_\alpha P_\alpha$, with P_α being the projection operators onto the product states $|e_\alpha, f_\alpha\rangle$ and $\sum_{\alpha=1}^n \Lambda_\alpha = 1$. Then the set of $\{\Lambda_\alpha\}$, which are maximal with respect to the density matrix ρ and the set of* *the projection operators* $\{P_\alpha\}$, *form a manifold which generically has a dimension* $n-1$ *and is determined by the following equation*:

$$
\begin{aligned}
1 - \sum_{i}^{n} \Lambda_i D_i + \sum_{i<j}^{n} \Lambda_i \Lambda_j D_{ij} - \sum_{i<j<k}^{n} \Lambda_i \Lambda_j \Lambda_k D_{ijk} + \cdots \\
+ (-1)^m \sum_{i_1<i_2<\cdots<i_m} \Lambda_{i_1} \Lambda_{i_2} \cdots \Lambda_{i_m} D_{i_1 i_2 \cdots i_m} + \cdots \\
+ (-1)^n \Lambda_1 \Lambda_2 \cdots \Lambda_n D_{12\cdots n} = 0
\end{aligned}
\tag{26}
$$

where the set of $\{D_{i_1 i_2 \cdots i_m}\}$ *are the subdeterminants* (*minors*) *of the matrix* D, *which is defined as*

$$
D = \begin{pmatrix}
\langle\psi_1|\rho^{-1}|\psi_1\rangle & \langle\psi_1|\rho^{-1}|\psi_2\rangle & \cdots & \langle\psi_1|\rho^{-1}|\psi_n\rangle \\
\langle\psi_2|\rho^{-1}|\psi_1\rangle & \langle\psi_2|\rho^{-1}|\psi_2\rangle & \cdots & \langle\psi_2|\rho^{-1}|\psi_n\rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle\psi_n|\rho^{-1}|\psi_1\rangle & \langle\psi_1|\rho^{-1}|\psi_2\rangle & \cdots & \langle\psi_n|\rho^{-1}|\psi_n\rangle
\end{pmatrix}
$$

and where by $\{|\psi_i\rangle\}$ *we denote for shortness the product vectors which are building the projection operators* $P_i \equiv |\psi_i\rangle\langle\psi_i|$.

Proof. Let us first remark that generically the matrix D does not have a block structure. If the matrix D consists k diagonal n_k-dimensional blocks, then not only is equation (26) fulfilled, but also the k corresponding equations for the blocks, so that the corresponding manifold has the dimension n_k, and is a Cartesian product of k manifolds of dimension $n_k - 1$. In the following we will concentrate on the generic case.

The proof of the lemma goes with induction. First we prove it for $n = 2$ and we get

$$1 - \Lambda_1 D_1 - \Lambda_2 D_2 + \Lambda_1 \Lambda_2 D_{12} = 0$$

or for $n = 3$ where we get

$$1 - \Lambda_1 D_1 - \Lambda_2 D_2 - \Lambda_3 D_3 + \Lambda_1 \Lambda_2 D_{12} + \Lambda_1 \Lambda_3 D_{13} + +\Lambda_2 \Lambda_3 D_{23} - \Lambda_1 \Lambda_2 \Lambda_3 D_{123} = 0.$$

Now, let us assume that the lemma is true for n, and show that it must also be true for $n+1$. Let ρ have the decomposition $\rho = \Lambda \rho_s + (1-\Lambda)\delta\rho$, with

$$\rho_s = \Lambda \sum_{\alpha=1}^{n+1} \Lambda_\alpha P_\alpha .$$

The lemma holds for the matrix $\tilde{\rho} = \rho - \Lambda_{n+1}|\psi_{n+1}\rangle\langle\psi_{n+1}|$ so that the first n coefficients Λ_α fulfil equation (26) with coefficients D calculated as above with the substitution $\rho^{-1} \rightarrow \tilde{\rho}^{-1} = (\rho - \Lambda_{n+1}|\psi_{n+1}\rangle\langle\psi_{n+1}|)^{-1}$. The latter inverse can be calculated using a power series expansion in the projector $\Lambda_{n+1}|\psi_{n+1}\rangle\langle\psi_{n+1}|$. The result is

$$
\begin{aligned}
(\rho - \Lambda_{n+1}|\psi_{n+1}\rangle\langle\psi_{n+1}|)^{-1}|\psi_i\rangle &= \rho^{-1}|\psi_i\rangle \\
&\quad + \frac{\Lambda_{n+1}\langle\psi_{n+1}|\rho^{-1}|\psi_i\rangle\langle\psi_i|\rho^{-1}|\psi_{n+1}\rangle}{1 - \Lambda_{n+1}\langle\psi_{n+1}|\rho^{-1}|\psi_{n+1}\rangle}\rho^{-1}|\psi_{n+1}\rangle .
\end{aligned}
$$

Inserting the above result into equations defining the surface for the first n Λ's we get, after tedious, but elementary algebraic calculation

$$
\begin{aligned}
1 - \sum_{i}^{n} \Lambda_i D_i + \sum_{i<j}^{n} \Lambda_i \Lambda_j D_{ij} - \sum_{i<j<k}^{n} \Lambda_i \Lambda_j \Lambda_k D_{ijk} + \cdots \\
+ (-1)^m \sum_{i_1<i_2<\cdots<i_m} \Lambda_{i_1} \Lambda_{i_2} \cdots \Lambda_{i_m} D_{i_1 i_2 \cdots i_m} + \cdots \\
+ (-1)^n \sum_{i_1<i_2<\cdots<i_n} \Lambda_{i_1} \Lambda_{i_2} \cdots \Lambda_{i_n} D_{i_1 i_2 \cdots i_n} \\
+ (-1)^{n+1} \Lambda_1 \Lambda_2 \cdots \Lambda_{n+1} D_{12\cdots n+1} = 0
\end{aligned}
$$

which proves the lemma for $n + 1$. □

Note that, in particular, if the decomposition discussed in the above lemma is the BSA, then the corresponding Λ's fulfil equation (26). This observation allows us to prove the uniqueness of the BSA in arbitrary dimensions. It is important to note that the surface defined by equation (26) can be considered for arbitrary Λ's, not necessarily positive! This surface is strictly convex and divides the space of all Λ's into two sets: a convex set of those sets of $\{\Lambda\text{'s}\}$ which have the property that $\rho - \Lambda \sum_{\alpha=1}^{n+1} \Lambda_\alpha P_\alpha$ is positive definite, and a concave set for which the latter matrix is not positive definite. If this surface contains a part of a hyperplane (linear subspace), it must contain the whole hyperplane, since it is defined by the polynomial equation (26). This observation is essential to prove the uniqueness of the expansion.

Lemma 7 (The uniqueness of the BSA). *Any density matrix ρ has a unique decomposition $\rho = \Lambda\rho_s + (1 - \Lambda)\delta\rho$, where ρ_s is a separable density matrix, $\delta\rho$ is a inseparable matrix with no product vectors in its range, and Λ is maximal.*

Proof. The proof the lemma goes by assuming the decomposition is not unique; then there must exist at least two BSA decompositions, $\rho = \Lambda\rho_{s1} + (1 - \Lambda)\delta\rho_1$ and $\rho = \Lambda\rho_{s2} + (1 - \Lambda)\delta\rho_2$, with the same maximal Λ. Now, any convex combination of these two BSA decompositions is also a BSA decomposition,

$$
\begin{aligned}
\rho &= \epsilon\rho_{s1} - (1-\epsilon)\rho_{s2} + \epsilon\delta\rho_1 + (1-\epsilon)\delta\rho_2 \\
&= \sum_i (\epsilon\Lambda\Lambda_{1i} - (1-\epsilon)\Lambda\Lambda_{2i}) P_i + (\epsilon\delta\rho_1 - (1-\epsilon)\delta\rho_2) \\
&\equiv \rho_s(\epsilon) + \delta\rho(\epsilon)
\end{aligned}
$$

where $\epsilon \in [0, 1]$. The part of the one-dimensional hyper plane (line) $\epsilon\Lambda_{1i} - (1 - \epsilon)\Lambda_{2i}$ for $\epsilon \in [0, 1]$ lies on the surface (26).

From the form of the surface it follows that the whole line $\epsilon\Lambda_{1i} - (1 - \epsilon)\Lambda_{2i}$ for all ϵ lies on that surface. This cannot be, since for some $\epsilon \notin [0, 1]$, and $\delta\rho_1 \neq \delta\rho_2$, $\delta\rho(\epsilon)$ must become non-positive definite. This is easy to see since for $\epsilon \to \pm\infty$, $\delta\rho(\epsilon) \propto \delta\rho_1 - \delta\rho_2$, and the latter matrix is non-zero and has the trace zero, so that it has to have eigenvalues of opposite signs. This is thus a contradiction with the assumption made at the beginning, *ergo* the BSA decomposition must be unique. □

7. The PPT BSA

In this section we discuss in detail the generalization of the BSA approach for PPT states used in [17, 18].

Theorem 4. *Let ρ be an arbitrary PPT state. For any countable set $V = \{P_i = |e_i, f_i\rangle\langle e_i, f_i|\}$, such that $|e_i, f_i\rangle \in R(\rho)$ and $|e_i^*, f_i\rangle \in R(\rho^{T_A})$, there exists the best separable* *approximation of* ρ *in the form*

$$\rho = \Lambda\rho_s + (1 - \Lambda)\delta\rho \tag{27}$$

where $\rho_s = \sum_i \Lambda_i P_i$ *is a separable state,* Λ *is maximal, and both* $\delta\rho \geqslant 0$, *and* $\delta\rho^{T_A} \geqslant 0$. *We call such a decomposition a PPT BSA if it preserves the PPT of the remainder* $\delta\rho$ *and*

$$\Lambda_{PPT} \equiv \max_{\mathrm{V}}(\mathrm{Tr}(\rho_s[V])). \tag{28}$$

Proof. Let us consider the set of all separable matrices $\rho_s = \sum_i \lambda_i |e_i, f_i\rangle\langle e_i, f_i|$, where $|e_i, f_i\rangle \in V$, $\rho - \rho_s \geqslant 0$ and $\rho^{T_A} - \rho_s^{T_A} \geqslant 0$. This set of ρ's form a convex and bounded set, which means that this set is compact. Because of the compactness there must exist a separable matrix ρ_s which has maximal trace $\Lambda = \mathrm{Tr}(\rho_2[V])$. By expanding V we will finally get the maximal PPT contribution. □

Let us analyze the PPT BSA decomposition in more detail. All information about the PPT entanglement is included in the PPT BSA parameter Λ and $\delta\rho$. If the PPT BSA remainder $\delta\rho$ does not vanish, then there exists no product vector $|e, f\rangle$, such that $|e, f\rangle \in R(\delta\rho)$ and simultaneously $|e^*, f\rangle \in R(\delta\rho^{T_A})$ is satisfied. This means that the PPT state $\delta\rho$ is entangled.

We introduce now, just like in the first version of the BSA, a procedure for constructing the matrix ρ_s. But before we do this let us define some basic concepts for that:

Definition 4. *A non-negative parameter* Λ *is called* PPT maximal *with respect to a positive PPT operator* ρ, *and a projection operator* $P = |\psi\rangle\langle\psi| \in V$ *if* $\rho - \Lambda P \geqslant 0$, $\rho^{T_A} - \Lambda\rho^{T_A} \geqslant 0$, *and for every* $\epsilon \geqslant 0$, *the matrix* $\rho - (\Lambda + \epsilon)P$ *is not a PPT state.*

This means that Λ is the maximal contribution of P that can be subtracted from ρ by maintaining the PPT of the difference. Now let us introduce the following lemma:

Lemma 8. Λ *is PPT maximal with respect to* ρ *and* $P = |\, e, f\,\rangle\langle\, e, f\,|$ *iff*:

- if $|e, f\rangle \notin R(\rho)$ and $|e^*, f\rangle \notin R(\rho^{T_A})$, or $|e, f\rangle \notin R(\rho)$ and $|e^*, f\rangle \in R(\rho^{T_A})$ or $|e, f\rangle \in R(\rho)$ and $|e^*, f\rangle \notin R(\rho^{T_A})$ then $\Lambda = 0$;
- if $|e, f\rangle \in R(\rho)$ and $|e^*, f\rangle \in R(\rho^{T_A})$ then

$$\Lambda = \min\left(\left(\langle e, f|\frac{1}{\rho}|e, f\rangle\right)^{-1}, \left(\langle e^*, f|\frac{1}{\rho^{T_A}}|e^*, f\rangle\right)^{-1}\right). \tag{29}$$

Proof. From lemma (1) we know that $\Lambda = \left(\langle e, f|\frac{1}{\rho}|e, f\rangle\right)^{-1}$ is the maximal contribution to ρ and $\tilde{\Lambda} = \left(\langle e^*, f|\frac{1}{\rho^{T_A}}|e^*, f\rangle\right)^{-1}$ is the maximal contribution to ρ^{T_A}. In order to maximize and keep the PPT of the difference we have to take the minimum of Λ and $\tilde{\Lambda}$. □

Definition 5. *A pair of non-negative* (Λ_1, Λ_2) *is called* maximal *with respect to* ρ *and a pair of projection operators* $P_1 = |e_1, f_1\rangle\langle e_1, f_1|$ *and* $P_2 = |e_2, f_2\rangle\langle e_2, f_2|$ *iff*

- $\rho - \Lambda_1 P_1 - \Lambda_2 P_2 \geqslant 0$ *and* $(\rho - \Lambda_1 P_1 - \Lambda_2 P_2)^{t_A} \geqslant 0$,
- Λ_1 *is PPT maximal with respect to* $\rho - \Lambda_2 P_2$,
- Λ_2 *is PPT maximal with respect to* $\rho - \Lambda_1 P_1$, *and*
- $\Lambda_1 + \Lambda_2$ *is maximal.*

The conditions for PPT maximizing of pairs $P_1 = |e_1, f_1\rangle\langle e_1, f_1|$ and $P_2 = |e_2, f_2\rangle\langle e_2, f_2|$ are described in appendix B.

Let us now prove that for a given countable set V of product vectors we can obtain the optimal PPT separable approximation by maximizing all pairs of product vectors in V. But before we do this, we have to define the PPT BSA manifold:

Definition 6. *Let the equation* $F(\lambda_1,\ldots,\lambda_K)=0$ *(or* $\lambda_1=f_1(\lambda_2,\ldots,\lambda_k)$*) describe the BSA manifold with respect to* ρ*, and* $\tilde{F}(\Lambda_1,\ldots,\lambda_K)=0$ *(or* $\lambda_1=\tilde{f}_1(\lambda_2,\ldots,\Lambda_k)$*) for* ρ^{T_A}*. Without losing generality in order to obtain the manifold which preserves the PPT of the difference* $(\rho-\rho_s)$ *we have to define*

$$\lambda_1=\min(\lambda_1=f_1(\lambda_2,\ldots,\lambda_K),\lambda_1=\tilde{f}_1(\lambda_2,\ldots,\lambda_K)\\ \equiv\bar{f}_1(\lambda_2,\ldots,\lambda_K). \tag{30}$$

The implicit form will then be given by $\bar{F}(\lambda_1,\ldots,\lambda_K)=0$.

Notice that the PPT BSA manifold is continuous and almost everywhere differentiable.

Theorem 5. *Given the set V of product vectors* $|e_i,f_i\rangle\in R(\rho)$ *where also* $|e_i^*,f_i\rangle\in R(\rho^{T_A})$*, then the matrix* $\tilde{\rho}_s=\sum_{i=1}\Lambda_iP_i$ *is the optimal PPT separable approximation of* ρ *if*:

- *all* Λ_i *are PPT maximal with respect to* $\rho_i=\rho-\sum_{i'\neq i}\Lambda_{i'}P_{i'}$*, and to the projector* P_i;
- *all pairs* (Λ_i,Λ_j) *are PPT maximal with respect to* $\rho_{ij}=\rho-\sum_{i'\neq i,j}\Lambda_{i'}P_{i'}$*, and to the projectors* (P_i,P_j).

Proof. If $\tilde{\rho}_s$ is a PPT BSA decomposition then all Λ_i, as well as all pairs (Λ_i,Λ_j) must be PPT maximal (otherwise maximizing Λ_i would increase the trace of $\tilde{\rho}_s$).

To prove the inverse, consider matrices $\rho_s=\sum_i\lambda_iP_i$ for which all individual λ_i are PPT maximal. This means that ρ_s belongs to the boundary of the set Z of all separable matrices such that $\rho-\rho_s\geqslant 0$ and $(\rho-\rho_s)^{t_A}\geqslant 0$. This boundary is the PPT BSA manifold:

$$\bar{F}(\lambda_1,\ldots,\lambda_K)=0. \tag{31}$$

The manifold (31) can be written as a function $\lambda_i=f_i(\{\lambda_j\}_{j\neq i})$, depending on which side of the manifold we are. Let $\rho_s^m=\sum_i\Lambda_iP_i$ be the separable matrix for which all pairs of Λ's are PPT maximal. The maximum of (Λ_i,Λ_j) then implies that

$$\frac{\partial}{\partial\lambda_i}(\lambda_i+f_j)|_{\lambda=\Lambda}=\frac{\partial}{\partial\lambda_i}\left(\sum_{i'\neq j}\lambda_{i'}+f_j\right)\Bigg|_{\lambda=\Lambda}\leqslant 0 \tag{32}$$

for all sides of the manifold $\bar{F}=0$ and i,j. This means that ρ_s^m is either a local maximum or a saddle point (not necessarily the same derivative in every direction of $\lambda=\Lambda$). Now we have the same situation just like in the original version of the BSA. The latter possibility cannot occur, since the set Z is *convex* (i.e. if $\rho_s,\rho_s'\in Z$ then $\epsilon\rho_s+(1-\epsilon)\rho_s'\in Z$ for every $0\leqslant\epsilon\leqslant 1$). Since (32) also describes a convex set it can for sure not be a saddle point. The same argument holds also for the local minimum. And finally the local maximum must be also a global one, because on a convex set there cannot exist two of them. This means that $\tilde{\rho}_s=\rho_s^m$. □

At the end of this section we would like to stress that the PPT BSA can be straightforwardly generalized to multicomposite systems.

8. Conclusion

In this paper we have presented several novel results concerning the BSA decompositions of density matrices of composite quantum systems. General results concern the uniqueness of the BSA decompositions, the existence of the BSA entanglement mass, and efficient methods of construction of the BSA decomposition for PPT states. More specific results for two-qubit systems deal with the necessary condition that the projector onto a nonmaximally entangled state provides the remainder in the BSA decomposition. There are several open questions concerning the BSA decompositions in higher-dimensional iHlbert spaces: what is the structure of the remainder in such a case; how to parametrize the remainders (the so-called edge states [21] in the case of PPT BSA)? The physical interpretation of the BSA entanglement mass is not known so far. In the case of 2×2 space, we hope that our results, together with remarkable analytic results of Englert and his colleagues [23], will bring us closer to the challenging goal of analytic construction of the BSA decomposition for the arbitrary two-quibit density matrix.

After our completion of this paper, Wellens and Kuś [30] have been able to construct analytically the BSA for generic states in the 2×2 case, and to relate the BSA entanglement measure to the so-called Wootters concurrence [11].

Acknowledgments

This work has been supported by Deutsche Forschungsgemeinschaft (SFB 407 and Schwerpunkt 'Quanteninformationsverarbeitung') and by the IST Programm 'EQUIP'. We thank D Bruss, J I Cirac, B-G Englert, P Horodecki, B Kraus, M Neilsen, A Sanpera, R Werner and M Wilkens for fruitful discussions.

Appendix A. Product vectors in the range

In this appendix we prove some lemmas that have been used in section 4. Both the results as well as the proofs are very much parallel to that used by Woronowicz [11].

Lemma 9. *If ρ is a density matrix in a 2×2 space having a positive partial transpose and $r(\rho) = r(\rho^{T_A}) = 3$, then there exists a product vector $|e, f\rangle \in R(\rho)$ such that $|e^*, f\rangle \in R(\rho^{T_A})$.*

Proof. Let there be given a density matrix $\rho = \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix}$ (A and C are invertible, because otherwise we would have product vectors in the kernel [17], and the existence of $|e, f\rangle$ would follow from the results of [17]). Now, we choose the basis in $\mathcal{H}_A$ to be

$$\left\{ \frac{1}{\sqrt{1+|\alpha|^2}} \begin{pmatrix} 1 \\ \alpha \end{pmatrix}, \frac{1}{\sqrt{1+|\alpha|^2}} \begin{pmatrix} -\alpha^* \\ 1 \end{pmatrix} \right\}.$$

In this new basis we obtain that

$$B(\alpha^*) = \frac{1}{\sqrt{1+\|\alpha\|^2}} (1 - \alpha^*) \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix} \begin{pmatrix} 1 \\ \alpha^* \end{pmatrix}$$

is a function of α^* only. This means that we can choose α such that $\det B(\alpha^*) = \det B(\alpha) = 0$. Choosing such an α, we get $r(\mathrm{B}) = r(B^*) = 1$.

The next step is to perform a non-unitary, but invertible local transformation $\rho \to I_A \otimes \frac{1}{\sqrt{C}} \rho I_A \otimes \frac{1}{\sqrt{C}}$, and redefine $A \to \frac{1}{\sqrt{C}} A \frac{1}{\sqrt{C}}$, $B \to \frac{1}{\sqrt{C}} B \frac{1}{\sqrt{C}}$. After that, the new matrix is given by $\rho = \begin{pmatrix} A & B \\ B^\dagger & I \end{pmatrix}$. Now, we use our assumption that $r(\rho) = 3$, from which it follows that $A = BB^\dagger + \lambda P$, where P is a projector on some vector $|\psi\rangle$. The assumption that also $(r\rho^{T_A}) = 3$, leads us to $A = B^\dagger B + \tilde{\lambda}\tilde{P}$, where $\tilde{P}$ is a projector on some other vector $|\tilde{\psi}\rangle$. This leads us to $BB^\dagger + \lambda P = B^\dagger B + \tilde{\lambda}\tilde{P}$, and since $\mathrm{tr}(BB^\dagger - B^\dagger B) = 0$, we get that $\lambda = \tilde{\lambda}$. What is the necessary condition now for $\begin{pmatrix} |f\rangle \\ z|f\rangle \end{pmatrix} \in r(\rho)$ and $\begin{pmatrix} |f\rangle \\ z^*|f\rangle \end{pmatrix} \in r(\rho^{T_A})$? This condition means nothing else than that there exist two vectors, say $\binom{|h\rangle}{|g\rangle}$ and $\binom{|\tilde{h}\rangle}{|\tilde{g}\rangle}$, such that

$$\begin{pmatrix} BB^\dagger + \lambda P & B \\ B^\dagger & I \end{pmatrix}\begin{pmatrix} |h\rangle \\ |g\rangle \end{pmatrix} = \begin{pmatrix} |f\rangle \\ z|f\rangle \end{pmatrix} \tag{A1}$$

$$\begin{pmatrix} B^\dagger B + \lambda \tilde{P} & B^\dagger \\ B & I \end{pmatrix}\begin{pmatrix} |\tilde{h}\rangle \\ |\tilde{g}\rangle \end{pmatrix} = \begin{pmatrix} |f\rangle \\ z^*|f\rangle \end{pmatrix} \tag{A2}$$

from which we get the equation

$$\frac{1}{1-zB}|\psi\rangle = \eta \frac{1}{1-z^*B^\dagger}|\tilde{\psi}\rangle \tag{A3}$$

with some complex η. In order to prove our lemma we must show that there exists a solution for (A3). The trick is now to describe the right-hand side of equation (A3) as a complex conjugate of the left-hand side, so that we can construct a solution explicitly.

We will show now that equation (A3) can indeed be transformed into the form

$$\frac{1}{1-zB}|\psi\rangle = \sigma_x \eta \frac{1}{1-z^*B^*}|\psi^*\rangle \tag{A4}$$

where σ_x is the Pauli matrix. Defining $\frac{1}{1-zB}|\psi\rangle = \binom{v_1}{v_2}$, we must have that $v_1 = \eta e^{i\phi} v_2^*$ and $v_2 = \eta e^{i\phi} v_1^*$. This equation has a solution if $v_1 = v e^{i\theta}$ and $v_2 = v e^{i\theta+\delta}$, where $\|v_1\| = \|v_2\| = v$. Let us take now an arbitrary δ and require $\binom{1}{e^{i\delta}} \sim \frac{1}{1-zB}|\psi\rangle$, which means that

$$(e^{i\delta}, -1)\frac{1}{1-zB}|\psi\rangle = 0 \tag{A5}$$

must hold. Obviously, this equation has not only one solution, but an infinite family of solutions for every δ.

Let us now prove that equation (A4) indeed holds. First we choose a basis $|\psi_1\rangle$, $|\psi_2\rangle$ such that $B^\dagger B - BB^\dagger = \begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix}$. Therefore we have that $\lambda(P - \tilde{P}) = \begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix}$. Since the overall phases of $|\psi\rangle$ and $|\tilde{\psi}\rangle$ are irrelevant, we parameterize $|\psi\rangle$ and $|\tilde{\psi}\rangle$ in our new basis as

$$|\psi\rangle = \begin{pmatrix} \sqrt{p} \\ \sqrt{1-p}e^{i\phi} \end{pmatrix} \qquad |\tilde{\psi}\rangle = \begin{pmatrix} \sqrt{1-\tilde{p}} \\ \sqrt{\tilde{p}}e^{i\tilde{\phi}} \end{pmatrix}.$$

This parameterization yields $\tilde{p} = p, \tilde{\phi} = \phi$ and $\Lambda = \lambda(1-2p)$. We observe now that for $r(B) = 1$, there exists always a unitary K such that $KBK^\dagger = B^T$. From this it trivially follows, of course, that $(K^\dagger)^T B^T K^T = B$, and therefore $(K^\dagger)^T KBK^\dagger K^T = B$, from which then $BU = UB$, where $U = K^\dagger K^T$.[1]

Now we will prove that $K = e^{i\varphi_0}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Let $M = BB^\dagger - B^\dagger B = \lambda(\tilde{P} - P)$ (note that $M = M^*$ in our basis). Then we have $KMK^\dagger = B^T B^* - B^* B^\dagger = B^*(B^T)^* - (B^\dagger)^* B^* = -M^* = -M$. Therefore $M = \lambda(K|\psi\rangle\langle\psi|K^\dagger - K|\tilde{\psi}\rangle\langle\psi K^\dagger)$, and for the vectors $|\psi\rangle$, $|\tilde{\psi}\rangle$ we get

$$K|\psi\rangle = \begin{pmatrix} e^{i\varphi_1}\sqrt{1-p} \\ e^{i\varphi_1}\sqrt{p}e^{i\phi} \end{pmatrix}$$

$$K|\tilde{\psi}\rangle = \begin{pmatrix} e^{i\varphi_2}\sqrt{p} \\ e^{i\varphi_2}\sqrt{1-p}e^{i\phi} \end{pmatrix}.$$

[1] In general, there exists always K, such that $KBK^* = B^T$. The existence of such K is sufficient to prove lemma 9 without the assumption $r(B) = 1$.

This implies $K = \begin{pmatrix} 0 & e^{i\theta_1} \\ e^{i\theta_2} & 0 \end{pmatrix}$ and therefore $\theta_2 = \varphi_1 + \phi, \theta_1 + \phi = \varphi_1, \varphi_2 = \theta_1 + \phi$ and $\varphi_2 + \phi = \theta_2$. But, if $\theta_1 \neq \theta_2$ then $U = \begin{pmatrix} e^{i(\theta_1-\theta_2)} & 0 \\ 0 & e^{i(\theta_1-\theta_2)} \end{pmatrix}$. U will commute with B, if B is diagonal in the chosen basis. But then $BB^\dagger - B^\dagger B = 0$, from which follows that $|\psi\rangle \sim |\tilde{\psi}\rangle$, and thus $\begin{pmatrix} \psi \\ 0 \end{pmatrix}$ in the range of ρ which proves the lemma. This means that $\theta_1 = \theta_2$, and $K = e^{i\varphi_0}\sigma_x$. Since the overall phases of K are irrelevant, we can assume that $K = \sigma_x$. This proves, however, (A4) which consequently proves the lemma too.

The reader may think now that we have finished the proof of the lemma, but remember that at the beginning of the proof we have made a non-unitary local operation. What we must do now is to transform back the density matrix ρ, and check if our results still hold. Let us see what happens after the inverse transformation:

$$\rho = \begin{pmatrix} \sqrt{C}BB^\dagger\sqrt{C} + \lambda\sqrt{C}P\sqrt{C} & \sqrt{C}B\sqrt{C} \\ \sqrt{C}B^\dagger\sqrt{C} & C \end{pmatrix}.$$

Demanding that $\begin{pmatrix} |f\rangle \\ z|f\rangle \end{pmatrix} \in R(\rho)$ and $\begin{pmatrix} |f\rangle \\ z^*|f\rangle \end{pmatrix} \in R(\rho^{T_A})$ leads to the following conditions:

$$\frac{1}{1 - \sqrt{C}B\frac{1}{\sqrt{C}}z}\sqrt{C}|\psi\rangle = \eta\frac{1}{1 - \sqrt{C}B^\dagger\frac{1}{\sqrt{C}}z^*}\sqrt{C}|\tilde{\psi}\rangle$$

$$\sqrt{C}(1 - f(z)B)|\psi\rangle = \sqrt{C}\eta(1 - f^*(z)B^\dagger)|\tilde{\psi}\rangle$$

$$(1 - f(z)B)|\psi\rangle = \eta(1 - f^*(z)B^\dagger)\sigma_x|\psi\rangle.$$

We see that the equations are equivalent after the rescaling, so that the lemma holds. □

The proof of the above lemma allows to parameterize the set of all product vectors $|e(\delta), f(\delta)\rangle$, which satisfy the condition $|e(\delta), f(\delta)\rangle \in R(\rho_s)$ and $|e(\delta)^*, f(\delta)\rangle \in R\left(\rho_s^{T_A}\right)$, by a one-dimensional real parameter δ. This will be used in section 3.

Appendix B. PPT pair maximizing

In this appendix we explain how to PPT maximize a pair of product projectors $|\psi_1\rangle\langle\psi_1| = |e_1, f_1\rangle\langle e_1, f_1|, |\psi_2\rangle\langle\psi_2| = |e_1, f_1\rangle\langle e_1, f_1|)$.

As we know from the BSA, the BSA manifold for ρ and $(|\psi_1\rangle\langle\psi_1| = |e_1, f_1\rangle\langle e_1, f_1|$, $|\psi_2\rangle\langle\psi_2| = |e_1, f_1\rangle\langle e_1, f_1|)$ is given by

$$F(\Lambda_1, \Lambda_2) \equiv 1 - \Lambda_1 D_1^0 - \Lambda_2 D_2^0 - \Lambda_1\Lambda_2 D^0 = 0 \tag{B1}$$

where

$$D_1^0 = \left\langle e_1, f_1|\rho^{-1}|e_1, f_1\right\rangle$$

$$D_2^0 = \left\langle e_2, f_2|\rho^{-1}|e_2, f_2\right\rangle$$

and

$$D^0 = \left\langle e_1, f_1|\rho^{-1}|e_1, f_1\right\rangle\left\langle e_2, f_2|\rho^{-1}|e_2, f_2\right\rangle - \left\|\left\langle e_1, f_1|\rho^{-1}|e_2, f_2\right\rangle\right\|^2.$$

But we also have to consider the BSA manifold for ρ^{T_A}. This one is given by

$$\tilde{F}(\Lambda_1, \Lambda_2) \equiv 1 - \Lambda_1 D_1^1 - \Lambda_2 D_2^1 - \Lambda_1\Lambda_2 D^1 = 0 \tag{B2}$$

where

$$D_1^1 = \left\langle e_1^*, f_1\left|\left(\rho^{t_A}\right)^{-1}\right|e_1^*, f_1\right\rangle$$

$$D_2^1 = \left\langle e_2^*, f_2\left|(\rho^{t_A})^{-1}\right|e_2^*, f_2\right\rangle$$

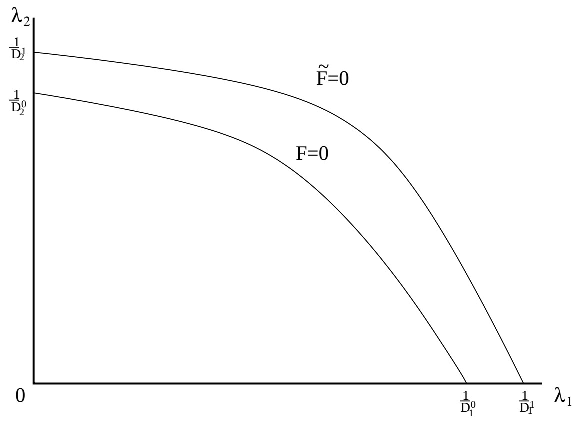

Figure 1. The manifold $F = 0$ is under $\tilde{F} = 0$.

and

$$D^1 = \langle e_1^*, f_1 | (\rho^{t_A})^{-1} | e_1^*, f_1 \rangle \langle e_2^*, f_2 | (\rho^{t_A})^{-1} | e_2^*, f_2 \rangle - \| \langle e_1^* - f_1 | (\rho^{t_A})^{-1} | e_2^*, f_2 \rangle \|^2.$$

Now we have to consider two basic cases which can occur.

Case 1. One of the BSA manifolds is under the other manifold. Without losing generality we assume that this is $F = 0$. Then, we have the situation as in figure 1. In this case we have to take the maximum on the manifold $F = 0$. From lemma 2 we know explicitly the condition for that. Of course, we are also including in case 1 that there can be an overlap at one endpoint (i.e. if $1/D_1^0 = 1/D_1^1$).

Case 2. The manifolds have a cross section point between $0 < \Lambda_1 \leqslant \max(1/D_1^0, 1/D_1^1)$. Without losing generality we assume that this describes figure 2. Now we can see from figure 2 how the PPT BSA manifold $\overline{F} = 0$ is constructed, and why it is not differentiable everywhere.

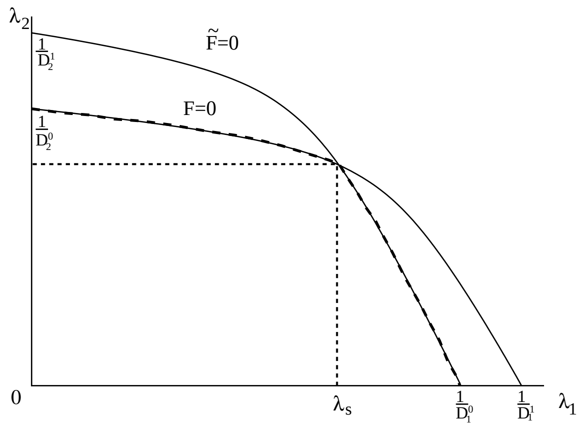

Figure 2. The manifolds have a cross section point λ_s.

Let us denote by Λ_m the maxima of the manifold $F = 0$ and also $\tilde{\Lambda}_m$ as the maxima of $\tilde{F} = 0$. Now we can have the following situations:

- If $\Lambda_m < \lambda_s$ and $\tilde{\Lambda}_m < \lambda_s$ then one has to take $\Lambda_{\max} = \Lambda_m$;
- If $\Lambda_m > \lambda_s$ and $\tilde{\Lambda}_m > \lambda_s$ then one has to take $\Lambda_{\max} = \tilde{\Lambda}_m$;

- If $\tilde{\Lambda}_m > \lambda_s$ and $\Lambda_m < \lambda_s$ then one has to take $\Lambda_{\max} = \lambda_s$;
- Both maxima are in λ_s, so that $\Lambda_{\max} = \Lambda_s$;
- The case where $\tilde{\Lambda}_m > \lambda_s$ and $\Lambda_m > \lambda_s$ cannot occur.

References

[1] Einstein A, Podolsky B and Rosen N 1935 *Phys. Rev.* **47** 777
[2] Schrödinger E 1935 *Proc. Camb. Phil. Soc.* **31** 555
[3] Ekert A 1991 *Phys. Rev. Lett.* **67** 661
Bennett C H and Wiesner S J 1992 *Phys. Rev. Lett.* **69** 2881
Bennett C, Brassard G, Crepeau C, Jozsa R, Peres A and Wootters W K 1993 *Phys. Rev. Lett.* **70** 1895
[4] Werner R 1989 *Phys. Rev.* A **40** 4277
[5] Horodecki P 1997 *Phys. Lett.* A **232** 333
[6] Horodecki R, Horodecki P and Horodecki M 1996 *Phys. Lett.* A **230** 377
[7] For a review, see:
Horodecki M, Horodecki P and Horodecki R 2000 *Quantum Information—An Introduction to Basic Theoretical Concepts and Experiments* (*Springer Tracts in Modern Physics*) ed G Alber *et al* (Berlin: Springer) at press
[8] For a primer, see:
Lewenstein M, Bruß D, Cirac J I, Kraus B, Kuś M, Samsonowicz J, Sanpera A and Tarrach R *J. Mod. Phys.* **47** 2481
(Lewenstein M, Bruß D, Cirac J I, Kraus B, Kuś M, Samsonowicz J, Sanpera A and Tarrach R 2000 *Preprint* quant-ph/0006064)
[9] Peres A 1996 *Phys. Rev. Lett.* **76** 1413
[10] Sanpera A, Tarrach R and Vidal G 1997 *Preprint* quant-ph/9707041
Busch P and Lahti P 1997 *Found. Phys. Lett.* **10** 113
[11] Woronowicz S L 1976 *Rep. Math. Phys.* **10** 165
see also:
Strömer E 1963 *Acta Math.* **110** 233
Choi M D 1975 *Linear Algear Appl.* **10** 285
Choi M D 1982 *Proc. Symp. Pure Math.* **38** 583
[12] Woronowicz S L 1976 *Commun. Math. Phys.* **51** 243
Kruszyński P and Woronowicz S L 1979 *Lett. Math. Phys.* **3** 319
[13] Horodecki M, Horodecki P and Horodecki R 1996 *Phys. Lett.* A **223** 1
[14] Horodecki M, Horodecki P and Horodecki R 1998 *Phys. Rev. Lett.* **80** 5239
[15] Lewenstein M and Sanpera A 1998 *Phys. Rev. Lett.* **80** 2261
[16] Sanpera A, Tarrach R and Vidal G 1998 *Phys. Rev.* A **58** 826
[17] Kraus B, Cirac J I, Karnas S and Lewenstein M 2000 *Phys. Rev.* A **61** 062302
(Kraus B, Cirac J I, Karnas S and Lewenstein M 1999 *Preprint* quant-ph/9912010)
[18] Horodecki P, Lewenstein M, Vidal G and Cirac J I 2000 *Phys. Rev.* A **62** 032310
(Horodecki P, Lewenstein M, Vidal G and Cirac J I 2000 *Preprint* quant-ph/0002089)
[19] Terhal B 1998 *Preprint* quant-ph/9810091
[20] Lewenstein M, Kraus B, Horodecki P and Cirac J I 2000 *Phys. Rev.* A at press
(Lewenstein M, Kraus B, Horodecki P and Cirac J I 2000 *Preprint* quant-ph/0005112)
[21] Lewenstein M, Kraus B, Cirac J I and Horodecki P 2000 *Phys. Rev.* A **62** 052310
[22] Bennett C H, DiVincenzo D P, Mor T, Shor P W, Smolin J A and Terhal B M 1999 *Phys. Rev. Lett.* **83** 3081
DiVincenzo D P, Mor T, Shor P W, Smolin J A and Terhal B M 1999 *Preprint* quant-ph/9908070
Bennett C H, DiVincenzo D P, Fuchs Ch A, Mor T, Rains E, Shor P W, Smolin J A and Wootters W K 1998 *Preprint* quant-ph/9804053
see also:
Horodecki R, Horodecki M and Horodecki P 1998 *Preprint* quant-ph/9811004
[23] Englert B G and Metwally N 2000 *J. Mod. Opt.* **47** 2221
(Englert B G and Metwally N 2000 *Preprint* quant-phy/0007053)
[24] Jaynes E T 1957 *Phys. Rev.* **106** 620
Jaynes E T 1957 *Phys. Rev.* **108** 171
[25] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 *Phys. Rev.* A **54** 3824
[26] Vedral V and Plenio M B 1998 *Phys. Rev.* A **57** 3
[27] Vidal G 1999 *Phys. Rev.* A **59** 141
[28] Vidal G 2000 *J. Mod. Opt.* **47** 355
[29] Wootters W K 1998 *Phys. Rev. Lett.* **80** 2245
[30] Wellens T and Kuś M 2001 *Preprint* quant-ph/0104098