Separability and Entanglement of Composite Quantum Systems

Maciej Lewenstein and Anna Sanpera
CEA/DSM/DRECAM/SPAM, Centre d'Etudes de Saclay, 91191 Gif-Sur-Yvette, France
(Received 22 July 1997)

We provide a constructive algorithm to find the best separable approximation to an arbitrary density matrix of a composite quantum system of finite dimensions. The method leads to a condition of separability and to a measure of entanglement. [S0031-9007(98)05501-X]

PACS numbers: 03.65.Bz, 42.50.Dv, 89.70.+c

Entanglement and nonlocality are some of the most emblematic concepts embodied in quantum mechanics [1]. The nonlocal character of an entangled system is usually manifested in quantum correlations between subsystems that have interacted in the past but are no longer interacting. Furthermore, these concepts play a crucial role in quantum information theory [2].

From a formal point of view, a state of a composite quantum system is called "inseparable" (or "entangled") if it cannot be represented as a tensor product of states of its subsystems. On the contrary, a density matrix ρ describes a *separable* state if it can be expressed as a finite [3] sum of tensor products of its subsystems:

$$\rho_s = \sum_i p_i(\rho_i^A \otimes \rho_i^B \cdots \otimes \rho_i^N); \qquad 1 \geq p_i \geq 0, \quad (1)$$

where $\rho_i^A, \rho_i^B, \ldots, \rho_i^N$ are density matrices describing subsystems $A, B, \ldots, N$, respectively, and $\sum_i p_i = 1$. Thus, separable states are those that can be produced by N distant observers (Alice, Bob, ..., Norberto) that prepare their states $(\rho_i^A, \rho_i^B, \ldots, \rho_i^N)$ independently, following common instructions (p_i) from a source [4]. Let us, for the moment, restrict ourselves to binary composite systems, i.e., $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Using the spectral decompositions of ρ_i^A and ρ_i^B it is easy to rewrite Eq. (1) in the form

$$\rho_s = \sum_\alpha \lambda_\alpha P_\alpha \qquad 1 \geq \lambda_\alpha \geq 0; \qquad \sum_\alpha \lambda_\alpha = 1, \quad (2)$$

where α is a multiindex running over all distinct eigenvectors of the matrices $\rho_i^A \otimes \rho_i^B$, and P_α are projectors onto product states, i.e., $P_\alpha \equiv |e,f\rangle\langle e,f|$ (where $|e\rangle \in \mathcal{H}_A$ and $|f\rangle \in \mathcal{H}_B$). Separable states, ρ_s, are thus mixtures of product states and as such their correlations are purely classical.

The distinction between entangled and separable states is well established for pure states: entangled pure states do always violate Bell inequalities [5]. For mixed states, however, the statistical properties of the mixture can hide the quantum correlations embodied in the system, making thus the distinction between separable and entangled enormously difficult [6,7]. Besides the importance of the subject from a fundamental point of view, this distinction has also important consequences for quantum information theory. Consider, for instance, Werner's family of entangled mixed states [8], which does not violate any kind of Bell inequalities but, nevertheless, can be used for quantum teleportation [9].

Recently, a first step in such distinction has been done by Peres [4] and the Horodecki family [3,10]. They have formulated two necessary conditions to characterize separable density matrices. The first condition [4] states that if a matrix ρ is separable, then its partial transposition (with respect to subsystem A or B) must be a density matrix, i.e., must have non-negative eigenvalues:

$$\rho = \rho_s \Rightarrow \rho^{T_B} = (\rho^{T_A})^* \geq 0. \quad (3)$$

This can be easily grasped from the representation (2) of separable matrices, since the partial transposition with respect to system B amounts to replacing P_α by $P_\alpha^{T_B} = |e,f^*\rangle\langle e,f^*|$, so that evidently

$$\rho^{T_B} = \sum_\alpha \lambda_\alpha |e,f^*\rangle\langle e,f^*| \geq 0. \quad (4)$$

This condition is sufficient to guarantee separability only for composite systems of dimension 2×2 or 2×3.

The second necessary condition [3] states that if $\rho = \rho_s$, then there exist a set of product vectors $V = \{|e_i, f_i\rangle\}$ that spans $\mathcal{R}(\rho)$ and at the same time $V^{T2} = \{|e_i, f_i^*\rangle\}$ spans $\mathcal{R}(\rho^{T2})$, where $\mathcal{R}(\rho)$ denotes the range of ρ, i.e., the set of all $|\psi\rangle \in \mathcal{H}$ for which $\exists|\phi\rangle \in \mathcal{H}$ such that $|\psi\rangle = \rho|\phi\rangle$. From the representations [2] and [4] we see that if a set of product vectors $\{|e_i, f_i\rangle\}$ spans $\mathcal{R}(\rho)$, it immediately follows that the set of product vectors $\{|e_i, f_i^*\rangle\}$ also spans $\mathcal{R}(\rho^{T_B})$. In general, both conditions are not equivalent. In particular, when the dimension of $\mathcal{R}(\rho)$ is equal to the dimension of $\mathcal{R}(\rho^{T2})$, the second condition may not be sufficient to ensure separability.

Finally, let us point out that for a density matrix which is known to be separable, only if $\dim[\mathcal{H}] \leq 6$ there exist an algorithm for decomposing it according to Eq. (1) [11].

In this Letter we address this last point and provide a constructive way of finding such an algorithm regardless of the (finite) dimension of the composite system. That immediately leads to a necessary condition for separability. Furthermore, we shall demonstrate that any inseparable mixed state in $C^2 \otimes C^2$ can be decomposed in a separable matrix and just a single pure entangled state, providing thus a novel characterization of the "entanglement" of any inseparable state.

The idea behind the algorithm relies on the fact that the set of separable states is compact. Therefore, for any density matrix ρ there exist a "maximal" separable matrix ρ_s^* which can be subtracted from ρ maintaining the positivity of the difference, $\rho - \rho_s^* \geq 0$. Let us express the above idea in a more rigorous way.

Theorem 1.—For any density matrix ρ (separable, or not) and for any set V of product vectors belonging to the range of ρ, i.e., $|e,f\rangle \in \mathcal{R}(\rho)$, there exist a separable (in general not normalized) matrix

$$\rho_s^* = \sum_\alpha \Lambda_\alpha P_\alpha \,, \tag{5}$$

with all $\Lambda_\alpha \geq 0$, such that $\delta\rho = \rho - \rho_s^* \geq 0$, and that ρ_s^* provides the best separable approximation (BSA) to ρ in the sense that the trace $\mathrm{Tr}(\delta\rho)$ is minimal (or, equivalently, $\mathrm{Tr}\,\rho_s^* \leq 1$ is maximal).

The proof of the theorem is simple, and the whole art is, of course, to construct ρ_s^*. Let us consider all separable matrices ρ_s of the form (5) that we can subtract from ρ maintaining the non-negativity of the difference $\delta\rho$. Obviously, the trace of ρ_s must be smaller than one, since $0 \leq \mathrm{Tr}(\delta\rho) = 1 - \mathrm{Tr}\,\rho_s$. The set of such matrices is determined by the set of possible $\Lambda_\alpha \geq 0$ for which $\delta\rho \geq 0$, and $0 \leq \mathrm{Tr}\,\rho_s = \sum_\alpha \Lambda_\alpha \leq 1$. This set is closed (in any reasonable topology). The set of all possible traces of ρ_s is bounded from above, so it must have an upper bound, say $1 - \epsilon$; *ergo* because of the compactness of the set of all ρ_s, there exist a matrix ρ_s^* in this set with the maximal trace, equal to $1 - \epsilon$. That implies that although the matrix $\rho_s^*[V]$ depends on the choice of the set V, and by expanding V we can construct better separable approximations to ρ (i.e., for $V' \supset V$, $\mathrm{Tr}\,\rho_s^*[V'] \geq \mathrm{Tr}\,\rho_s^*[V]$), it is generally sufficient to take $V \subset S$ large enough to obtain already the maximal possible trace $\mathrm{Tr}\,\rho_s^*[V] = \mathrm{Tr}\,\rho_s^*[S]$ [where S is the set of all $|e,f\rangle \in \mathcal{R}(\rho)$]. The latter statement indicates also that although typically the BSA matrix $\rho_s^*[V]$ is not unique, its trace is. Nevertheless, for $C^2 \otimes C^2$ composite systems we shall demonstrate that $\rho_s^*[V]$ is also unique.

As an obvious consequence of Theorem 1, we obtain a necessary and sufficient condition for separability.

Condition 3.—A density matrix ρ is separable iff (if and only if) there exist a set of product vectors $V \subset \mathcal{R}(\rho)$, for which the best separable approximation to ρ, $\rho_s^*[V]$ has the trace 1.

The proof is again simple: The necessity of the cond3 follows directly from (2). From the fact that $\delta\rho = \rho - \rho_s^* \geq 0$, and $\mathrm{Tr}\,\delta\rho = 1 - 1 = 0$, we obtain $\delta\rho \equiv 0$, or equivalently $\rho = \rho_s^*$.

Before we discuss the procedure of construction of the matrix ρ_s^*, let us introduce two concepts which shall play a crucial role in what it follows.

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

The maximal Λ determines thus the maximal contribution of P that can be subtracted from ρ maintaining the non-negativity of the difference. In the following we will apply the above definition to projections onto product vectors, i.e., $|\psi\rangle = |e,f\rangle$. The following lemma characterizes a single maximal Λ completely.

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

$$0 < \Lambda = \frac{1}{\langle\psi|\frac{1}{\rho}|\psi\rangle} \,. \tag{6}$$

Note that in case (b) the expression on the right-hand side of Eq. (6) makes sense, since $|\psi\rangle \in \mathcal{R}(\rho)$, and therefore there exists $|\Psi\rangle \in \mathcal{R}(\rho)$ such that $|\psi\rangle = \rho|\Psi\rangle$. Let us observe that for any $|\phi\rangle$ the Schwartz inequality implies that

$$\langle\phi|P|\phi\rangle = \left|\,\langle\phi|\sqrt{\rho}\,\frac{1}{\sqrt{\rho}}\,|\psi\rangle\,\right|^2 \leq \langle\phi|\rho|\phi\rangle\left\langle\psi\,\left|\,\frac{1}{\rho}\,\right|\,\psi\right\rangle. \tag{7}$$

That proves that for every $|\phi\rangle$, $\langle\phi|\rho - \langle\psi|1/\rho|\psi\rangle^{-1}P|\phi\rangle \geq 0$, i.e., $\rho - \Lambda P \geq 0$. Since on the other hand, $(\rho - \Lambda P)|\Psi\rangle = 0$ for $|\Psi\rangle = \frac{1}{\rho}|\psi\rangle$, thus for every $\epsilon > 0$, $\langle\Psi|[\rho - (\Lambda + \epsilon)P]|\Psi\rangle = -\epsilon\Lambda^{-2} < 0$. This proves that Λ given by expression (6) is indeed maximal.

Definition 2.—A pair of non-negative (Λ_1, Λ_2) is called *maximal* with respect to ρ and a pair of projection operators $P_1 = |\psi_1\rangle\langle\psi_1|$, $P_2 = |\psi_2\rangle\langle\psi_2|$ iff $\rho - \Lambda_1 P_1 - \Lambda_2 P_2 \geq 0$, 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 maximal pair (Λ_1, Λ_2) determines thus the maximal contribution of $\Lambda_1 P_1 + \Lambda_2 P_2$ that can be subtracted from ρ maintaining the non-negativity of the difference, and that has a maximal trace, $\mathrm{Tr}(\Lambda_1 P_1 + \Lambda_2 P_2) = \Lambda_1 + \Lambda_2$.

Lemma 2.—A pair (Λ_1, Λ_2) is maximal with respect to ρ and a pair of projectors (P_1, P_2) iff (a) if $|\psi_1\rangle$, $|\psi_2\rangle$ do not belong to $\mathcal{R}(\rho)$ then $\Lambda_1 = \Lambda_2 = 0$; (b) if $|\psi_1\rangle$ does not belong to $\mathcal{R}(\rho)$, while $|\psi_2\rangle \in \mathcal{R}(\rho)$ then $\Lambda_1 = 0$, $\Lambda_2 = \langle\psi_2|1/\rho|\psi_2\rangle^{-1}$; (c) if $|\psi_1\rangle$, $|\psi_2\rangle \in \mathcal{R}(\rho)$ and $\langle\psi_1|1/\rho|\psi_2\rangle = 0$ then $\Lambda_i = \langle\psi_i|1/\rho|\psi_i\rangle$, $i = 1,2$; (d) finally, if $|\psi_1\rangle$, $|\psi_2\rangle \in \mathcal{R}(\rho)$ and $\langle\psi_1|1/\rho|\psi_2\rangle \neq 0$ then

$$\Lambda_1 = (\langle\psi_2|1/\rho|\psi_2\rangle - |\langle\psi_1|1/\rho|\psi_2\rangle|)/D \,, \tag{8a}$$

$$\Lambda_2 = (\langle\psi_1|1/\rho|\psi_1\rangle - |\langle\psi_1|1/\rho|\psi_2\rangle|)/D \,, \tag{8b}$$

where $D = \langle\psi_1|1/\rho|\psi_1\rangle\langle\psi_2|1/\rho|\psi_2\rangle - |\langle\psi_1|1/\rho|\psi_2\rangle|^2$.

The proof of (a) and (b) is the same as the proof of Lemma 1. In case (c) observe that

$(\rho - \Lambda_1 P_1)^{-1}|\psi_2\rangle = \rho^{-1}|\psi_2\rangle$, $(\rho - \Lambda_2 P_2)^{-1}|\psi_1\rangle = \rho^{-1}|\psi_1\rangle$, so that maximality of Λ_i implies automatically that $\Lambda_i = \langle\psi_i|\rho^{-1}|\psi_i\rangle$, $i = 1, 2$. Finally, in case (d) we get $(\rho - \Lambda_2 P_2)^{-1}|\psi_1\rangle = \rho^{-1}|\psi_1\rangle + B\rho^{-1}|\psi_2\rangle$, with $B = \Lambda_2\langle\psi_2|1/\rho|\psi_1\rangle/D$. The maximality of Λ_1 assures then automatically the maximality of Λ_2 provided

$$1 - \Lambda_1\langle\psi_1|1/\rho|\psi_1\rangle - \Lambda_2\langle\psi_2|1/\rho|\psi_2\rangle + \Lambda_1\Lambda_2 D = 0. \tag{9}$$

Maximizing the sum $\Lambda_1 + \Lambda_2$ with the constraint (9), we arrive after elementary algebra at Eqs. (8).

We can now formulate the basic theorem of this paper.

Theorem 2.—Given the set V of product vectors $|e, f\rangle \in \mathcal{R}(\rho)$, the matrix $\rho_s^* = \sum_\alpha \Lambda_\alpha P_\alpha$ is the BSA to ρ iff (a) 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; (b) 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).

Let us prove now that maximizing all the pairs $(\Lambda_\alpha, \Lambda_\beta)$ with respect to $\rho_{\alpha\beta} = \rho - \sum_{\alpha'\neq\alpha,\beta}\Lambda_{\alpha'}P_{\alpha'}$, (P_α, P_β) is a necessary and sufficient condition to subtract the maximal separable matrix ρ_s^* from ρ. Obviously, if ρ_s^* is the BSA then all Λ_α, as well as all pairs $(\Lambda_\alpha, \Lambda_\beta)$ must be maximal, since otherwise maximalizing Λ_α, or the sum $\Lambda_\alpha + \Lambda_\beta$ would increase the trace of ρ_s^*, maintaining non-negativity of $\rho - \rho_s^*$.

To prove the inverse, assume that the total number of α's is K, and that ρ_s^* has all pairs of Λ's maximal. Consider matrices $\rho_s = \sum_\alpha \lambda_\alpha P_\alpha$ in the vicinity of ρ_s^*, for which all individual λ_α are maximal, i.e., ρ_s belong to the boundary of the set Z of all separable matrices such that $\rho - \rho_s \geq 0$; λ_α's lie thus on a $(K-1)$-dimensional manifold, defined through a constraint,

$$f(\lambda_1, \ldots, \lambda_K) = 0. \tag{10}$$

Maximality of $(\Lambda_\alpha, \Lambda_\beta)$ implies that $(\lambda_\alpha + \lambda_\beta)$ has a maximum at $\lambda_{\alpha,\beta} = \Lambda_{\alpha,\beta}$ under the constraint (10), and for all $\gamma \neq \alpha, \beta$; $\lambda_\gamma = \Lambda_\gamma$ which implies $(\partial f/\partial\lambda_\alpha|_{\lambda=\Lambda}) = (\partial f/\partial\lambda_\beta|_{\lambda=\Lambda})$. Using this identity for a sufficient number of pairs we get that $(\partial f/\partial\lambda_\alpha|_{\lambda=\Lambda}) = \text{const}$ for all α. That is equivalent to the fact that the gradient of $\mathrm{Tr}(\rho_s)$ under the constraint (10) vanishes for $\rho_s = \rho_s^*$. The trace of ρ_s has thus either a local maximum, or a minimum, or a saddle point at $\lambda = \Lambda$. The two latter possibilities cannot occur, since the trace is maximal with respect to all pairs of λ's, and since the set Z is *convex* (i.e., if $\rho_s, \rho_s' \in Z$ then $\epsilon\rho_s + (1-\epsilon)p_s' \in Z$ for every $0 \leq \epsilon \leq 1$). For the same reason of convexity, the local maximum at ρ_s^* must be a global one, i.e., there cannot exist two matrices ρ_s^* and $\tilde\rho_s^*$, which both provide local maxima of the trace, and have $\mathrm{Tr}\,\rho_s^* \neq \mathrm{Tr}\,\tilde\rho_s^*$; *ergo* ρ_s^* is the BSA, and any other matrix $\tilde\rho_s^*$ which has all pairs of Λ's maximal, must have the same trace as ρ_s^*.

In any case, we have shown that any density matrix ρ of composite system $\mathcal{H}$ can be decomposed according to $\rho = \rho_s^* + \delta\rho$, where ρ_s^* is a separable matrix (in general not normalized) with maximal trace. Let us analyze such decomposition in more detail. All the information concerning "inseparability" is included in the matrix $\delta\rho$. If it does not vanish, i.e., if ρ is not separable, its range $\mathcal{R}(\delta\rho)$ cannot contain any product vector. The fact that the range of $\delta\rho$ does not contain any product vector restricts also the number of linearly independent entangled states that it can contain. Note that the set of all product vectors in the Hilbert space H of dimension $N \times M$ spans a $(N + M - 1)$-dimensional manifold. Thus, a generic linear subspace of $\mathcal{H}$ of dimension larger than $(N-1)\times(M-1)$ contains product vectors. The above statement implies that the dimension of $\mathcal{R}(\delta\rho)$ is $\leq (N-1)\times(M-1)$; in particular for $N = M = 2$, $\delta\rho$ is a simple projector onto one entangled state.

As an immediate consequence, we obtain that any density matrix ρ in $C^2 \otimes C^2$ has a unique decomposition in the form

$$\rho = \lambda\rho_s + (1-\lambda)P_e; \qquad \lambda \in [0, 1], \tag{11}$$

where ρ_s is a separable density matrix (normalized), P_e denotes a single pure entangled projector ($P_e \equiv |\Psi_e\rangle\langle\Psi_e|$), and λ is maximal. Any other decomposition of the form $\rho = \tilde\lambda\tilde\rho_s + (1-\tilde\lambda)\tilde P_e$ with $\tilde\lambda \in [0, 1]$ such that $\tilde\rho_s \neq \rho_s$ necessarily implies that $\tilde\lambda < \lambda$. If not, that is, if $\lambda = \tilde\lambda$ for $\tilde\rho_s \neq \rho_s$, it follows from Ref. [11] that for $P_e \neq \tilde P_e$, we can always find projectors onto product states in the plane formed by P_e and $\tilde P_e$ and therefore increase λ, which is impossible since λ is already maximal.

The decomposition given by expression (11) leads straightforwardly to an unambiguous measure of the entanglement for any mixed state ρ (in $C^2 \otimes C^2$):

$$E(\rho) = (1-\lambda)E(|\Psi_e\rangle), \tag{12}$$

where $E(|\Psi_e\rangle)$ is the entanglement of its pure state expressed in terms of the von Neumann entropy of the reduced density matrix of either of its subsystems [12]:

$$E(|\Psi_e\rangle) = -\mathrm{Tr}\,\rho_A \log_2 \rho_A \equiv -\mathrm{Tr}\,\rho_B \log_2 \rho_B, \tag{13}$$

where $\rho_{\{A,B\}} = \mathrm{Tr}_{\{B,A\}}\rho$. This measure of entanglement is clearly independent of any purification or formation procedure [12,13].

Let us illustrate with an example the ideas stressed in the paper. Consider a pair of spin-$\frac{1}{2}$ particles in an impure state consisting of a fraction x of the singlet and a mixture in equal proportions of the singlet and the triplet [8]. This state is described, in the computational basis, by the density matrix

$$\rho_w(x) = \begin{pmatrix} \frac{1-x}{4} & 0 & 0 & 0 \\ 0 & \frac{1+x}{4} & -\frac{x}{2} & 0 \\ 0 & -\frac{x}{2} & \frac{1+x}{4} & 0 \\ 0 & 0 & 0 & \frac{1-x}{4} \end{pmatrix}; \qquad 0 < x < 1. \tag{14}$$

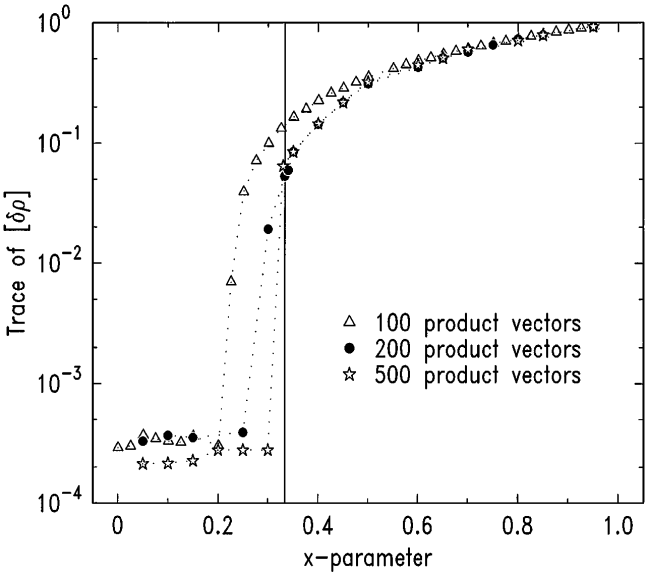

FIG. 1. The best separable approximation to a Werner state ρ_w. We plot the value $\mathrm{Tr}(\delta\rho)$ for the matrix ρ_w. The vertical line indicates the separability border [Eq. (3)]. (The numerical precision of the algorithm is set to 10^{-4}, so that $\mathrm{Tr}(\delta\rho)$ must be $\geq 10^{-4}$).

For this case Eq. (3) is sufficient to ensure separability: ρ_w is separable if $x \leq 1/3$ and inseparable otherwise. Nevertheless, we use our procedure to check the separability and to obtain the decomposition of ρ given by Eq. (11) for different values of x.

For each given set V, we first construct the matrix

$$\rho_s^*[V] = \sum_V \Lambda_\alpha P_\alpha \tag{15}$$

with the Λ' maximized pairwise, according to the definitions [14]. When the numerical convergence has been achieved we obtain $\delta\rho = (\rho_w - \rho_s^*[V])$ and compute its trace. Typically, we observe that (a) only very few projectors P_α of each set V contribute to the matrix $\rho_s^*[V]$, and (b) if the set V is large enough (i.e., >300), the results become independent of the chosen set.

The results are presented in Fig. 1, for a set of 100, 200, and 500 P_α projectors randomly chosen. Each point represents the corresponding value of $\mathrm{Tr}(\delta\rho)$ for a given $\rho_w(x)$. The vertical line indicates the condition of separability, derived from Eq. (3). For $x \leq 1/3$, $\mathrm{Tr}(\delta\rho) = 0$ indicating that ρ_w is separable. At $x \sim 1/3$, a clear "phase transition" occurs, and the value $\mathrm{Tr}(\delta\rho) \neq 0$, indicating thus the nonseparable character of the state. Therefore, our numerical results reproduce accurately the conditions of separability derived from Eq. (3).

Let us now analyze the inseparability properties of ρ_w. The matrix $\delta\rho$ when it does not vanish, i.e., for $x > 1/3$, corresponds to the projector onto the maximally entangled singlet $|\Psi^-\rangle = 1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. Thus, a Werner state of the type ρ_w can always be decomposed as

$$\rho_w(x) = \lambda(x)\rho_s + [1 - \lambda(x)]|\Psi^-\rangle\langle\Psi^-| \tag{16}$$

with $\lambda = 1$ for $x \leq 1/3 (\Longleftrightarrow \rho_w = \rho_s)$, and $0 \leq \lambda < 1$ for $x > 1/3$. A measure of the entanglement of ρ_w is, therefore, naturally provided by the value of the corresponding λ, i.e., $E[\rho_w(x)] = [1 - \lambda(x)]$ ebits, since the singlet has a value of entanglement of 1 ebit (see [Eq. (13)]. This measure does not coincide with other measures of the entanglement of ρ_w [12,13]. A further analysis of this entanglement measure will be presented elsewhere.

Summarizing, we have presented a method to construct the best separable approximation to an arbitrary density matrix of a composite quantum system (of arbitrary dimensions). The method provides a necessary condition for separability of a density matrix. Furthermore, for composite systems of $\dim[\mathcal{H}] = 4$, it also provides us with an unambiguous measure of the entanglement of its nonseparable states.

It is a pleasure to thank A. Barenco, A. Peres, P. Horodecki, S. Popescu, and R. Werner for very useful discussions. A. S. acknowledges the Elsag-Baily-I.S.I. Foundation meeting on quantum computation 1997 and financial support from European Community.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] See, for instance, C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wooters, Phys. Rev. A **54**, 3824 (1996), and references therein.
[3] P. Horodecki, Phys. Lett. A **232**, 333 (1997).
[4] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[5] J. S. Bell, Physics **1**, 195 (1964).
[6] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
[7] R. Horodecki and M. Horodecki, Phys. Rev. A **54**, 1838 (1996).
[8] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[9] S. Popescu, Phys. Rev. Lett. **72**, 797 (1994); C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[10] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 8 (1996).
[11] A. Sanpera, R. Tarrach, and G. Vidal, e-print quant-ph/9707041 (to be published); W. K. Wootters, e-print quant-ph/970909.
[12] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[13] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).
[14] The solution of Eqs. (8) requires us to solve inverse equations of the type $(|\phi\rangle = \rho^{-1}|\psi\rangle)$ several times. From a numerical point of view, the convergence of these kinds of problems is often ill defined. To circumvent this difficulty, we maximize each pair of projection operators (P_α, P_β) by a trial-and-error method. Although this slows down the calculation, the program takes a few minutes to calculate BSA for a given set of 100 product vectors.