Optimization of entanglement witnesses

M. Lewenstein,[1] B. Kraus,[2] J. I. Cirac,[2] and P. Horodecki[3]
[1]*Institute for Theoretical Physics, University of Hannover, D-30167 Hannover, Germany*
[2]*Institute for Theoretical Physics, University of Innsbruck, A-6020 Innsbruck, Austria*
[3]*Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-952 Gdańsk, Poland*
(Received 4 May 2000; published 16 October 2000)

An entanglement witness (EW) is an operator that allows the detection of entangled states. We give necessary and sufficient conditions for such operators to be optimal, i.e., to detect entangled states in an optimal way. We show how to optimize general EW, and then we particularize our results to the nondecomposable ones; the latter are those that can detect positive partial transpose entangled states (PPTES's). We also present a method to systematically construct and optimize this last class of operators based on the existence of ''edge'' PPTES's, i.e., states that violate the range separability criterion [Phys. Lett. A **232**, 333 (1997)] in an extreme manner. This method also permits a systematic construction of nondecomposable positive maps (PM's). Our results lead to a sufficient condition for entanglement in terms of nondecomposable EW's and PM's. Finally, we illustrate our results by constructing optimal EW acting on $H=\mathbb{C}^2\otimes\mathbb{C}^4$. The corresponding PM's constitute examples of PM's with minimal ''qubit'' domains, or—equivalently—minimal Hermitian conjugate codomains.

PACS number(s): 03.67.Hk, 03.65.Bz, 03.65.Ca

I. INTRODUCTION

Quantum entanglement [1,2], which is the essence of many fascinating quantum-mechanical effects [3–6], is a very fragile phenomenon. It is usually very hard to create, maintain, and manipulate entangled states under laboratory conditions. In fact, any system is usually subjected to the effects of external noise and interactions with the environment. These effects turn *pure-state entanglement* into *mixed-state*, or *noisy*, *entanglement*. The separability problem, that is, the characterization of mixed entangled states, is highly nontrivial, and has not been accomplished so far. Even an apparently innocent question—*Is a given state entangled and does it contain quantum correlations, or is it separable, and does not contain any quantum correlations?*—will, in general, be very hard (if not impossible) to answer.

Mathematically, mixed-state entanglement can be described as follows. A density operator $\rho\geqslant 0$, acting on a finite Hilbert space $H=H_A\otimes H_B$ describing the state of two quantum systems A and B, is called separable [7] (or not entangled) if it can be written as a convex combination of product vectors; that is, in the form

$$\rho=\sum_k p_k|e_k,f_k\rangle\langle e_k,f_k|, \tag{1}$$

where $p_k\geqslant 0$, and $|e_k,f_k\rangle\equiv|e_k\rangle_A\otimes|f_k\rangle_B$ are product vectors. Conversely, ρ is nonseparable (or entangled) if it cannot be written in this form. Physically, a state described by a separable (nonseparable) density operator ρ can always (never) be prepared locally. Most applications in quantum information are based on the nonlocal properties of quantum mechanics, [3–6,8] and therefore on nonseparable states. Thus a criterion to determine whether a given density operator is nonseparable, i.e. useful for quantum information purposes, or not is of crucial importance. On the other hand, positive partial transposed entangled states (PPTES's), are objects of special interest since they represent so-called bound entangled states, and therefore provide evidence of irreversibility in quantum information processing [9].

For low-dimensional systems [10,11] there exist operationally simple necessary and sufficient conditions for separability. In fact, in $H=\mathbb{C}^2\otimes\mathbb{C}^2$ and $H=\mathbb{C}^2\otimes\mathbb{C}^3$ the Peres-Horodecki criterion [10,11] establishes that ρ is separable iff its partial transpose is positive. Partial transpose means a transpose with respect to one of the subsystems [12]. For higher-dimensional systems all operators with nonpositive partial transposition are entangled. However, there exist PPTES's [13,14]. Thus the separability problem reduces to finding whether density operators with positive partial transpose are separable or not [15,16].

In the recent years there has been a growing effort in searching for necessary and sufficient separability criteria and checks which would be operationally simple [15,16]. Several necessary [7,17] or sufficient [13,18–21] conditions for separability are known. A particularly interesting necessary condition is given by the so-called *range criterion* [13]. According to this criterion, if a state ρ acting on a finite-dimensional Hilbert space is separable then there must exist a set of product vectors $\{|e_k,f_k\rangle\}$ that spans the range $R(\rho)$, such that the set of partial complex conjugated product states $\{|e_k,f_k^*\rangle\}$ spans the range of the partial transpose of ρ with respect to the second system, i.e., ρ^{T_B}. Among PPTES's that violate this criterion there are particular states with the property that if one subtracts a projector onto a product vector from them, the resulting operator is no longer a PPTES [20,21]. In this sense, these states lie in the *edge* between PPTES's and entangled states with nonpositive partial transposition, and therefore we will call them ''edge'' PPTES's.

An approach involving the analysis of the range of density operators initiated in Ref. [13] turned out to be very fruitful. In particular, it led to an algorithm for the optimal decomposition of mixed states into a separable and an inseparable part [22–24], and to a systematic method of constructing examples of PPTES's using unextendible product bases [14,25]. For low rank operators it has allowed to show that one can reduce the separability problem to the one of determining the roots of certain complex polynomial equations [20,21].

From a different point of view, a very general approach to analyze the separability problem is based on the so-called entanglement witnesses (EW's) and positive maps (PM's) [11]. Entanglement witnesses [25] are operators that detect the presence of entanglement. Starting from these operators one can define PM's [26] that also detect entanglement. An example of a PM is precisely partial transposition [10,27,28]. The importance of EW's stems from the fact that a given operator is separable iff there exists an EW that detects it [11]. Thus, if one was able to construct all possible EW's (or PM's) one would have solved the problem of separability. Unfortunately, it is not known how to construct EW's that detect PPTES's in general. The only result in this direction so far was given in Ref. [25], although some preliminary results exist in the mathematical literature [29]. Starting from a PPTES fulfilling certain properties (related to the existence of unextendible basis of product vectors [14]), it has been shown how to construct an EW (and a corresponding PM) that detects it. Perhaps, one of the most interesting goals regarding the separability problem is to develop a constructive and operational approach using EW's and PM's that allows us to detect mixed entanglement.

In this paper we realize this goal partially: we introduce a powerful technique to construct EW's and PM's that, among other things, allows us to study the separability of certain density operators . In particular, we show how to construct optimal EW's; that is, operators that detect the presence of entanglement in an optimal way. We specifically concentrate on nondecomposable EW's, which are those that detect the presence of PPTES's. Furthermore, we present a way of constructing optimal EW's for edge PPTES's. Our method generalizes the one introduced by Terhal [25] to the case in which there are no unextendible basis of product vectors. When combined with our previous results [20,21] regarding subtracting product vectors from PPTES's, the construction of nondecomposable optimal EW's starting from ''edge'' PPTES's gives rise to a sufficient criterion for nonseparability of general density operators with positive partial transposition. We illustrate our method by constructing optimal EW's that detect some known examples of PPTES's [13] in $H=\mathbb{C}^2\otimes\mathbb{C}^4$. The corresponding PM's constitute the first examples of PM's with minimal ''qubit'' domains, or—equivalently—minimal Hermitian conjugate codomains.

This paper is organized as follows. In Sec. II we review the definition of EW's and fix some notation. In Sec. III we study general EW's. We define optimal witnesses, and find a criterion to decide whether an EW is optimal or not. In Sec. IV we restrict the results of Sec. III to nondecomposable EW's. In particular, we show how to optimize them by subtracting decomposable operators. In Sec. V we give an explicit method to optimize both general and nondecomposable EW's. We also show how to construct nondecomposable EW's, and that this leads to a sufficient criterion of nonseparability. The construction and optimization are based on the use of ''edge'' PPTES's. In Sec. VI we extend our results to positive maps. In Sec. VII we illustrate our methods and results starting from the examples of PPTES's given in Ref. [13]. The paper also contains two appendixes. In Appendix A we describe in detail a method to check whether an EW is optimal or not. In Appendix B we discuss separately some important properties of the edge PPTES's, and show that they provide a canonical decomposition of mixed states with positive partial transpose.

II. DEFINITIONS AND NOTATION

We say that an operator $W=W^\dagger$ acting on $H=H_A\otimes H_B$ is an EW if [11,25]. (I) $\langle e,f|W|e,f\rangle\geqslant 0$ for all product vectors $|e,f\rangle$, (II) it has at least one negative eigenvalue (i.e. is not positive); and (III) $\mathrm{tr}(W)=1$. Property (I) implies that $\langle\rho\rangle_W\equiv\mathrm{tr}(W\rho)\geqslant 0$ for all ρ separable. Thus, if we have $\langle\rho\rangle_W<0$ for some $\rho\geqslant 0$, then ρ is nonseparable. In that case we say that *W detects* ρ. Property (II) implies that every EW detects something, since in particular it detects the projector on the subspace corresponding to the negative eigenvalues of W. Property (III) is just normalization condition that we need in order to compare the action of different EW's [30].

In this paper we will denote the kernel and range of ρ by $K(\rho)$ and $R(\rho)$, respectively. The partial transposition of an operator X will be denoted by X^T [12,31]. On the other hand, we will encounter several kinds of operators (EW's, positive operators, decomposable operators, etc.) and vectors. In order to help to identify the kind of operators and vectors we use, and not to overwhelm the reader by specifying at each point their properties, we will use the following notation: W will denote an EW. P and Q will denote positive operators. Unless specified they will have unit trace $[\mathrm{tr}(P)=\mathrm{tr}(Q)=1]$. D will denote a decomposable operator. That is, $D=aP+bQ^T$, where $a,b\geqslant 0$. Unless stated, all decomposable operators that we use will have unit trace (i.e., $b=1-a$). ρ will denote a positive operator (not necessarily of trace 1). $|e,f\rangle$ will denote product vectors with $|e\rangle\in H_A$ and $|f\rangle\in H_B$. Unless especified, they will be normalized.

III. GENERAL ENTANGLEMENT WITNESSES

In this section we first give some definitions directly related to EW's. Then we introduce the concept of optimal EW's. We derive a criterion to determine when an EW is optimal. This criterion will serve us to find an optimization procedure for these operators.

A. Definitions

Given an EW, W, we define the following. $D_W=\{\rho\geqslant 0$, such that $\langle\rho\rangle_W<0\}$; that is, the set of operators detected by W. Finer: Given two EW's, W_1 and W_2, we say that W_2 is *finer* than W_1, if $D_{W_1}\subseteq D_{W_2}$; that is, if all the operators detected by W_1 are also detected by W_2. Optimal entanglement witness (OEW): We say that W is an OEW if there exists no other EW which is finer. $P_W=\{|e,f\rangle\in H$, such that $\langle e,f|W|e,f\rangle=0\}$; that is, the set of product vectors on which W vanishes. As we will show, these vectors are closely related to the optimality property.

Note the important role that the vectors in P_W play regarding entanglement (for a method to determine P_W in practice, see Appendix A). If we have an EW, W, which detects a given operator ρ, then the operator $\rho' = \rho + \rho_w$, where

$$\rho_w = \sum_k p_k |e_k, f_k\rangle\langle e_k, f_k|, \tag{2}$$

with $p_k \geqslant 0$, and $|e_k, f_k\rangle \in P_W$ is also detected by W. In fact, this means that any operator of the form of Eq. (2) is in the border between separable states and nonseparable states, in the sense that if we add an arbitrarily small amount of ρ to it we obtain a nonseparable state. Thus the structure of the sets P_W characterizes the border between separable and nonseparable states. In fact, from the results of this section it will become clear that we can restrict ourselves to the structure of the set of P_W corresponding to OEW's.

B. Optimal entanglement witnesses

According to Ref. [11] ρ is nonseparable iff there exists an EW which detects it. Obviously, we can restrict ourselves to the study of OEW. For that, we need criteria to determine when an EW is optimal. In this subsection we will derive a necessary and sufficient condition for this to happen (theorem 1 below). In order to do this, we first have to introduce some results that tell us under which conditions an EW is finer than another one.

Lemma 1: Let W_2 be finer than W_1, and

$$\lambda \equiv \inf_{\rho_1 \in D_{W_1}} \left| \frac{\langle \rho_1 \rangle_{W_2}}{\langle \rho_1 \rangle_{W_1}} \right|. \tag{3}$$

Then we have the following. (i) If $\langle \rho \rangle_{W_1} = 0$ then $\langle \rho \rangle_{W_2} \leqslant 0$. (ii) If $\langle \rho \rangle_{W_1} < 0$, then $\langle \rho \rangle_{W_2} \leqslant \langle \rho \rangle_{W_1}$. (iii) If $\langle \rho \rangle_{W_1} > 0$ then $\lambda \langle \rho \rangle_{W_1} \geqslant \langle \rho \rangle_{W_2}$. (iv) $\lambda \geqslant 1$. In particular, $\lambda = 1$ iff $W_1 = W_2$.

Proof: Since W_2 is finer than W_1 we will use the fact that for all $\rho \geqslant 0$ such that $\langle \rho \rangle_{W_1} < 0$ then $\langle \rho \rangle_{W_2} < 0$.

(i) Let us assume that $\langle \rho \rangle_{W_2} > 0$. Then we take any $\rho_1 \in \mathcal{D}_{W_1}$ so that for all $x \geqslant 0$, $0 \leqslant \tilde{\rho}(x) \equiv \rho_1 + x\rho \in \mathcal{D}_{W_1}$. But for sufficiently large x we have that $\langle \tilde{\rho}(x) \rangle_{W_2}$ is positive, which cannot be since then $\rho(x) \notin \mathcal{D}_{W_2}$.

(ii) We define $\tilde{\rho} = \rho + |\langle \rho \rangle_{W_1}| \mathbb{1} \geqslant 0$. We have that $\langle \tilde{\rho} \rangle_{W_1} = 0$. Using (i) we have that $0 \geqslant \langle \rho \rangle_{W_2} + |\langle \rho \rangle_{W_1}|$.

(iii) We take $\rho_1 \in D_{W_1}$ and define $\tilde{\rho} = \langle \rho \rangle_{W_1} \rho_1 + |\langle \rho_1 \rangle_{W_1}| \rho \geqslant 0$, so that $\langle \tilde{\rho} \rangle_{W_1} = 0$. Using (i) we have $|\langle \rho_1 \rangle_{W_1}| \langle \rho \rangle_{W_2} \leqslant |\langle \rho_1 \rangle_{W_2}| \langle \rho \rangle_{W_1}$. Dividing both sides by $|\langle \rho_1 \rangle_{W_1}| > 0$ and $\langle \rho \rangle_{W_1} > 0$, we obtain

$$\frac{\langle \rho \rangle_{W_2}}{\langle \rho \rangle_{W_1}} \leqslant \left| \frac{\langle \rho_1 \rangle_{W_2}}{\langle \rho_1 \rangle_{W_1}} \right|. \tag{4}$$

Taking the infimum with respect to $\rho_1 \in D_{W_1}$ on the right-hand side of this equation, we obtain the desired result.

(iv) From (ii) it immediately follows that $\lambda \geqslant 1$. On the other hand, we just have to prove that if $\lambda = 1$ then $W_1 = W_2$ (the only if part is trivial). If $\lambda = 1$, using (i) and (iii) we have that $\langle \rho_v \rangle_{W_1} \geqslant \langle \rho_v \rangle_{W_2}$ for all $\rho_v = |e,f\rangle\langle e,f|$ projector on a product vector. Since $\mathrm{tr}(W_1) = \mathrm{tr}(W_2)$ we must have $\mathrm{tr}[(W_1 - W_2)\rho_v] = 0$ for all ρ_v, since we can always find a product basis in which we can take the trace. But now, for any given $\rho \geqslant 0$ we can define $\tilde{\rho}(x) = \rho + x\mathbb{1}$ such that for large enough x, $\tilde{\rho}(x)$ is separable [18]. In that case we have $\langle \tilde{\rho}(x) \rangle_{W_1} = \langle \tilde{\rho}(x) \rangle_{W_2}$ which implies that $\langle \rho \rangle_{W_1} = \langle \rho \rangle_{W_2}$, i.e. $W_1 = W_2$.

Corollary 1: $D_{W_1} = D_{W_2}$ iff $W_1 = W_2$.

Proof: We just have to prove the only if part. For that, we define λ as in Eq. (3). On the other hand, defining

$$\tilde{\lambda} \equiv \inf_{\rho_2 \in D_{W_2}} \left| \frac{\langle \rho_2 \rangle_{W_1}}{\langle \rho_2 \rangle_{W_2}} \right|, \tag{5}$$

we have that $\tilde{\lambda} \geqslant 1$ since W_1 is finer than W_2 [lemma 1(iv)]. Equivalently,

$$1 \geqslant \sup_{\rho_1 \in D_{W_1}} \left| \frac{\langle \rho_1 \rangle_{W_2}}{\langle \rho_1 \rangle_{W_1}} \right| \geqslant \lambda \geqslant 1, \tag{6}$$

where for the last inequality we have used that W_2 is finer than W_1. Now, since $\lambda = 1$ we have that $W_1 = W_2$ according to lemma 1(iv).

Next we introduce one of the basic results of this paper. It basically tells us that one EW is finer than another one if they differ by a positive operator. That is, if we have an EW and we want to find another one which is finer, we have to subtract a positive operator.

Lemma 2: W_2 is finer than W_1 iff there exists a P and $1 > \epsilon \geqslant 0$ such that $W_1 = (1-\epsilon) W_2 + \epsilon P$.

Proof: (If) For all $\rho \in D_{W_1}$ we have that $0 > \langle \rho \rangle_{W_1} = (1-\epsilon)\langle \rho \rangle_{W_2} + \epsilon \langle \rho \rangle_P$ which implies $\langle \rho \rangle_{W_2} < 0$ and therefore $\rho \in D_{W_2}$. (Only if) We define λ as in Eq. (3). Using lemma 1(iv) we have $\lambda \geqslant 1$. First, if $\lambda = 1$ then according to lemma 1(iv) we have $W_1 = W_2$ (i.e., $\epsilon = 0$). For $\lambda > 1$, we define $P = (\lambda - 1)^{-1}(\lambda W_1 - W_2)$ and $\epsilon = 1 - 1/\lambda > 0$. We have that $W_1 = (1-\epsilon) W_2 + \epsilon P$, so that it only remains to be shown that $P \geqslant 0$. But this follows from lemma 1 (i)–(iii), and the definition of λ, $\lambda = \inf_{\rho_1 \in D_{W_1}} |\langle \rho_1 \rangle_{W_2} / \langle \rho_1 \rangle_{W_1}|$.

The previous lemma provides us with a way of determining when an EW is finer than another one. With this result, we are now at the position of fully characterizing OEW.

Theorem 1: W is optimal iff for all P and $\epsilon > 0$, $W' = (1+\epsilon) W - \epsilon P$ is not an EW [does not fulfill (I)].

Proof: (If) According to lemma 2, there is no EW which is finer than W; therefore, W is optimal. (Only if) If W' is an EW, then according to lemma 2 W is not optimal.

The previous theorem tells us that W is optimal iff when we subtract any positive operator from it, the resulting operator is not positive on product vectors. This result is not very practical for two reasons: (1) For a given P it is typically very hard to check whether there exists some $\epsilon>0$ such that $W-\epsilon P$ is positive on all product vectors. (2) It may be difficult to find a particular P that can be subtracted from W among all possible positive operators. In Appendix A we show how to circumvent these two drawbacks in practice: we give a simple criterion to determine when a given P can be subtracted from W. This allows us to determine which are the positive operators which can be subtracted from a given EW.

In the rest of this subsection we will present some simple results related to these two questions. First, it is clear that not every positive operator P can be subtracted from an EW, W. In particular, the following lemma tells us that it must vanish on P_W.

Lemma 3: If $PP_W\neq 0$ then P cannot be subtracted from W.

Proof: There exists some $|e_0,f_0\rangle\in P_W$ such that $\langle e_0,f_0|P|e_0,f_0\rangle>0$. Substituting this product vector into condition (I) for any $W-\epsilon P$, we see that the inequality is not fulfilled for any $\epsilon>0$, i.e. P cannot be subtracted.

Corollary 2: If P_W spans H, then W is optimal.

Note that, as announced at the beginning of this section, the set P_W plays an important role in determining the properties of the separable states which lie on the border with the entangled states. We see here that this set also plays an important role in determining whether an EW is optimal or not.

On the other hand, in order to check whether a given operator P can be subtracted or not from W, one has to check whether there exists some $\epsilon>0$ such that $\langle e,f|W-\epsilon P|e,f\rangle>0$ for all $|e,f\rangle$. The following lemma gives an alternative way to do this. In fact, it gives a necessary and sufficient criterion for an EW to be optimal. For a given $|e\rangle\in H_A$, we will denote by $W_e\equiv\langle e|W|e\rangle$.

Lemma 4: W is optimal iff for all $|\Psi\rangle$ orthogonal to P_W:

$$\epsilon\equiv\inf_{|e\rangle\in H_A}[\langle\Psi|e\rangle W_e^{-1}\langle e|\Psi\rangle]^{-1}=0. \tag{7}$$

Proof: (If) Let us assume that W is not optimal; that is, there exists $W'\neq W$, finer than W. Then, according to lemma 2 we have that there exists $\epsilon_0>0$ and $P\geqslant 0$ such that $W'=(W-\epsilon_0 P)/(1-\epsilon_0)$. Imposing that W' is positive on product vectors (i.e., $W'_e\geqslant 0$ for all $|e\rangle\in H_A$) we obtain $0\leqslant\langle e|W-\epsilon_0 P|e\rangle\leqslant W_e-\epsilon_0\lambda_\Psi\langle e|\Psi\rangle\langle\Psi|e\rangle$, where $|\Psi\rangle$ is any eigenstate of P with nonzero eigenvalue, λ_Ψ. According to Ref. [20], this last operator is positive iff both (i) $\langle e|\Psi\rangle$ is in the range of $\langle e|W|e\rangle$, which imposes that $|\Psi\rangle$ is orthogonal to P_W; and (ii) $\lambda_\Psi\epsilon_0\leqslant[\langle\Psi|e\rangle W_e^{-1}\langle e|\Psi\rangle]^{-1}$, which imposes that $\epsilon\geqslant\lambda_\Psi\epsilon_0>0$ for that given $|\Psi\rangle$. (Only if) Let us assume that there exists some $|\Psi\rangle$ orthogonal to P_W such that $\epsilon>0$. Then using the same arguments one can show that $W'\equiv(W-\epsilon|\Psi\rangle\langle\Psi|)/(1-\epsilon)\neq W$ is an EW. According to lemma 2, W' is finer than W, so that W is not optimal.

C. Decomposable entanglement witnesses

There exists a class of EW which is very simple to characterize, namely, decomposable entanglement witnesses (DEW's) [28]. These are EW's that can be written in the form

$$W=aP+(1-a)Q^T, \tag{8}$$

where $a\in[0,1]$. As it is well known (see Sec. IV), these EW's cannot detect PPTES's. In any case, for the sake of completeness, we will give some simple properties of optimal DEW's.

Theorem 2: Given a DEW, W, if it is optimal then it can be written as $W=Q^T$, where $Q\geqslant 0$ contains no product vector in its range.

Proof: Since W is decomposable, it can be written as $W=aP+(1-a)Q^T$. $W'\propto W-aP$ is also a witness, which according to lemma 2 is finer than W, and therefore W is not optimal. On the other hand, if $|e,f\rangle\in R(Q)$ then for some $\lambda>0$ we have that $W\propto(Q-\lambda|e,f\rangle\langle e,f|)^T$ is finer than W, and therefore this last is not optimal.

This previous result can be slightly generalized as follows

Theorem 2': Given a DEW's, W, if it is optimal then it can be written as $W=Q^T$, where $Q\geqslant 0$ and there is no operator $P\in R(Q)$ such that $P^T\geqslant 0$.

Proof: Is the same as in previous theorem.

Corollary 3: Given a DEW, W, if it is optimal then W^T is not an EW [does not fulfill (II)].

Proof: Using theorem 2 we have that $W=Q^T$ with $Q\geqslant 0$. Then $W^T=Q\geqslant 0$, which does not satisfy property (ii).

IV. NONDECOMPOSABLE ENTANGLEMENT WITNESSES

In Sec. III we were concerned with EW's in general. As mentioned above, when studying separability we just have to consider those EW's that can detect PPTES's. In order to characterize these, one defines nondecomposable witnesses (NDEW's) as EW's which cannot be written in the form of Eq. (8) [28]. This section is devoted to this kind of witness. The importance of NDEW's in order to detect PPTES's is reflected in the following

Theorem 3: An EW is nondecomposable iff it detects PPTES's.

Proof: (If) Let us assume that the EW is decomposable. Then it cannot detect a PPTES, since if $\rho,\rho^T\geqslant 0$ we have $\mathrm{tr}[(aP+(1-a)Q^T)\rho]=a\,\mathrm{tr}(P\rho)+(1-a)\mathrm{tr}(Q\rho^T)\geqslant 0$. (Only if) The set of decomposable witnesses is convex and closed, and W, as a set containing one point, is a closed convex set itself. Thus from Hahn–Banach theorem [32] it follows that there exists an operator ρ such that (i) $\mathrm{tr}[\rho(aP+(1-a)Q^T)]\geqslant 0$ for all $P,Q\geqslant 0$, $a\in[0,1]$; and (ii) $\mathrm{tr}(\rho W)<0$. From (i), taking $a=1$ we infer that $\rho\geqslant 0$; on the other hand, taking $a=0$ we obtain that $\mathrm{tr}[\rho^T Q]\geqslant 0$ for all $Q\geqslant 0$, and therefore $\rho^T\geqslant 0$. Thus, W detects ρ which is a PPTES.

Corollary 4: Given an operator D, it is decomposable iff $\mathrm{tr}(D\rho)\geqslant 0$ for all $\rho,\rho^T\geqslant 0$.

A. Definitions

In this subsection we introduce some definitions which are parallel to those given in Sec. III. Given a NDEW, W, we define $d_W=\{\rho\geqslant 0$, such that $\rho^T\geqslant 0$ and $\langle\rho\rangle_W<0\}$; that is, the set of PPT operators detected by W. Nondecomposable-finer (ND-finer): Given two NDEW's, W_1 and W_2, we say that W_2 is *ND-finer* than W_1, if $d_{W_1} \subseteq d_{W_2}$; that is, if all the operators detected by W_1 are also detected by W_2. Nondecomposable optimal entanglement witness (NDOEW): We say that W is an NDOEW if there exist no other NDEW which is ND-finer. $p_W = \{|e,f\rangle \in H$, such that $\langle e,f|W|e,f\rangle = 0\}$; that is, the product vectors on which W vanishes.

Note again the important role that the vectors in p_W play regarding PPTES's. If we have a NDEW, W, which detects a given PPTES ρ, then the operator $\rho' = \rho + \rho_w$ where ρ_w has the form of Eq. (2) with $p_k \geqslant 0$, and $|e_k, f_k\rangle \in p_W$ also describes a PPTES. Thus any operator of the form of Eq. (2) lies in the border between separable states and PPTES's.

B. Optimal nondecomposable entanglement witness

The goal of this section is to find a necessary and sufficient condition for a NDEW to be optimal. We start by proving a similar result to the one given in lemma 1, but for NDEW's:

Lemma 1(b): Let W_2 be ND-finer than W_1,

$$\lambda \equiv \inf_{\rho_1 \in d_{W_1}} \left| \frac{\mathrm{tr}(W_2 \rho_1)}{\mathrm{tr}(W_1 \rho_1)} \right|, \tag{9}$$

and now both $\rho, \rho^T \geqslant 0$. Then we have (i)–(iv) as in lemma 1.

Proof: The proof is basically the same as in lemma 1, and will be omitted here.

Corollary 1(b): Given two NDEW's, $W_{1,2}$, then $d_{W_1} = d_{W_2}$ iff $W_1 = W_2$.

Proof: The proof is basically the same as corollary 1, and will be omitted here.

Lemma 2(b): Given two NDEW's, $W_{1,2}$, W_2 is ND-finer than W_1 iff there exists a decomposable operator D and $1 > \epsilon \geqslant 0$, such that $W_1 = (1-\epsilon) W_2 + \epsilon D$.

Proof: (If) Given any $\rho, \rho^T \geqslant 0$, we have that if $\rho \in d_{W_1}$ then $0 > \langle \rho \rangle_{W_1} = (1-\epsilon)\langle \rho \rangle_{W_2} + \epsilon \langle \rho \rangle_D \geqslant (1-\epsilon)\langle \rho \rangle_{W_2}$, where in the last inequality we have used that $\langle \rho \rangle_D \geqslant 0$ since D is decomposable (see corollary 4). Therefore $\rho \in d_{W_2}$. (Only if) We define λ as in Eq. (9), so that $\lambda \geqslant 1$ according to Lemma 1(b)(iv). If $\lambda = 1$ we have $W_1 = W_2$. If $\lambda > 1$ we define $D = (\lambda - 1)^{-1}(\lambda W_1 - W_2)$ and $\epsilon = 1 - 1/\lambda$. We have that $W_1 = (1-\epsilon) W_2 + \epsilon D$, so that it only remains to be shown that D is decomposable. But from Lemma 1(b)(i)–(iii) and the definition of λ it follows that $\langle \rho \rangle_D \geqslant 0$ for all $\rho, \rho^T \geqslant 0$. Using corollary 4 we then have that D is decomposable.

Now we are able to fully characterize NDOEW's.

Theorem 1(b): Given an NDEW, W, it is *ND* optimal iff for all decomposable operators D and $\epsilon > 0$, $W' = (1+\epsilon) W - \epsilon D$ is not an EW [does not fulfill (I)].

Proof: Same as for Theorem 1.

Theorems 1 and 1(b) allow us to relate OEW's and NDOEW's. In this way we can directly translate the results for general OEW's to NDOEW's. We have

Theorem 4: Given a NDEW, *W*, *W* is a NDOEW iff both W and W^T are OEW's.

Proof: (If) Let us assume that W is not a NDOEW. Then, according to theorem 1(b) there exists $\epsilon > 0$, and a decomposable operator D such that $W' = (1+\epsilon) W - \epsilon D$ is a NDEW. We can write $D = aP + (1-a) Q^T$, with $a \in [0,1]$. If $a \neq 0$, then $W_1 = (1 + a\epsilon) W - a\epsilon P$ fulfills $\langle e,f|W_1|e,f\rangle \geqslant \langle e,f|W'|e,f\rangle \geqslant 0$, and therefore, according to lemma 2, W is not optimal. If $a \neq 1$ then $W_2 = [1 + (1-a)\epsilon] W^T - (1-a)\epsilon Q$ fulfills $\langle e,f|W_2|e,f\rangle \geqslant \langle e,f|(W')^T|e,f\rangle \geqslant 0$, i.e. is an EW, and therefore W^T is not optimal. (Only if) According to theorem 1(b), if W is ND optimal then for all $D = aP + (1-a) Q^T$, with $a \in [0,1]$, and all $\epsilon > 0$ we have that $W' = (1-\epsilon) W - \epsilon D$ does not satisfy (I). Taking $a = 1$ we have for all P and $\epsilon > 0$, $W_1 = (1-\epsilon) W - \epsilon P$ does not fulfill (I), and therefore (theorem 1) W is optimal; analogously, taking $a = 0$ we have that W^T is optimal also.

Corollary 5: W is a NDOEW's iff W^T is a NDOEW.

V. OPTIMIZATION

In this section we give a procedure to optimize EW's, which is based on the results of the previous sections.

A. Optimization of general entanglement witnesses

Our method is based in the following lemma. It tells us how much we can subtract from an EW. Here we will denote $W_e = \langle e|W|e\rangle$ and $P_e = \langle e|P|e\rangle$ where $|e\rangle \in H_A$, by $[\cdots]_{\min}$, the minimum eigenvalue, and by $[\cdots]_{\max}$, the maximum eigenvalue. On the other hand, $X^{-1/2}$ will denote the square root of the pseudoinverse of X [33].

Lemma 5: If there exists some P such that $P P_W = 0$, and

$$\begin{aligned} \lambda_0 &\equiv \inf_{|e\rangle \in H_A} [P_e^{-1/2} W_e P_e^{-1/2}]_{\min} \\ &= (\sup_{|e\rangle \in H_A} [W_e^{-1/2} P_e W_e^{-1/2}]_{\max})^{-1} > 0, \end{aligned} \tag{10}$$

then

$$W'(\lambda) \equiv (W - \lambda P)/(1-\lambda), \tag{11}$$

where $\lambda > 0$ is an EW iff $\lambda \leqslant \lambda_0$.

Proof: Let us find out for which values of $\lambda \geqslant 0$ $W'(\lambda)$ is an EW. We have to impose condition (I), which can be written as $\langle e|W'(\lambda)|e\rangle \geqslant 0$, i.e.,

$$W_e - \lambda P_e \geqslant 0. \tag{12}$$

Multiplying by $P_e^{-1/2}$ on the right and left of this equation, we obtain $P_e^{-1/2} W_e P_e^{-1/2} \geqslant \lambda$, which immediately gives that $\lambda \leqslant \lambda_0$ given in the first part of Eq. (10). On the other hand, multiplying by $W_e^{-1/2}$ on the right and left of Eq. (12), we obtain $W_e^{-1/2} P_e W_e^{-1/2} \leqslant 1/\lambda$, which immediately gives that $\lambda \leqslant \lambda_0$ given in the second equality of Eq. (10).

Lemma 5 provides us with a direct method to optimize EW's by subtracting positive operators for which the elements of P_W are contained in their kernels. The method thus consists of (1) determining P_W; (2) choosing an operator P so that $P P_W = 0$, and determining λ using (10); and (3) subtracting the operator P according to lemma 5 if $\lambda \neq 0$. Continuing in the same vein we will reach an OEW. In Appendix A, we show how to accomplish steps (1) and (2) in practice.

B. Optimization of nondecomposable entanglement witnesses

For NDEW's we have the following generalization of lemma 5:

Lemma 5(b): Given a NDEW, W, if there exists some decomposable operator D such that $Dp_W=0$ and

$$\lambda_0 \equiv \inf_{|e\rangle \in H_A} [D_e^{-1/2} W_e D_e^{-1/2}]_{\min} = \Big(\sup_{|e\rangle \in H_A} [W_e^{-1/2} D_e W_e^{-1/2}]_{\max}\Big)^{-1} > 0, \tag{13}$$

then

$$W'(\lambda) \equiv (W - \lambda D)/(1-\lambda), \tag{14}$$

where $\lambda > 0$ is a NDEW iff $\lambda \leq \lambda_0$.

Proof: Same as for lemma 5.

With the help of lemma 5(b), we can optimize NDEW's by subtracting decomposable operators as follows: (1) determining p_W and p_{W^T}; (2) choosing P and Q so that $Pp_W=0$ and $Qp_{W^T}=0$, building $D=aP+(1-a)Q^T$ with $a\in[0,1]$, and determining λ_0 using (13); and, (3) subtracting the operator D according to Lemma 5(b) if $\lambda_0 \neq 0$.

C. Detectors of ''edge'' PPTES

In the previous subsections we have have given two optimization procedures. In both of them, starting from a general EW one can obtain one which is optimal (or ND optimal). It may well happen that the EW found in this way is nondecomposable, even though the original one was decomposable. To check this one simply has to use corollary 3, that is, check whether W^T is an EW or not. In case it is, then the OEW, W, is nondecomposable. However, nothing guarantees that the final EW is nondecomposable if the original one is not. In this subsection we describe a general method to construct NDEW's using the optimization procedures introduced earlier. This method generalizes the one presented in Ref. [25].

We are going to use the results presented in Refs. [20,21]. There we already used and discussed ''edge'' PPTES's, though without naming them. Let us now introduce the following definition:

Definition (see Ref. [20]): A PPTES δ is an edge PPTES if, for all product vectors $|e,f\rangle$ and $\epsilon>0$, $\delta-\epsilon|e,f\rangle\langle e,f|$ is not a PPTES.

This definition implies that the edge states lie on the boundary between PPTES's and entangled states with nonpositive partial transposes. In this subsection we will show how, out of an edge PPTES, we can construct a NDOEW that detects it. As we mentioned in Sec. I, edge PPTES's are of special importance. In particular, they allow one to provide a canonical form to write PPTES's in arbitrary Hilbert spaces. For these reasons, some properties of the edge PPTES's are discussed in Appendix B.

In order to check whether a PPTES is an edge PPTES we can use a range criterion [13] (also see Ref. [20]). That is, δ is an edge PPTES iff for all $|e,f\rangle \in R(\delta)$, $|e,f^*\rangle \notin R(\delta^{T_B})$.

Let δ be an edge PPTES, and let us denote the projector onto $K(\delta)$ by P_1 and the projector onto $K(\delta^T)$ by Q_1. We define

$$W_\delta = a(P_1 + Q_1^T), \tag{15}$$

where $a=1/\mathrm{tr}(P_1+Q_1)$. Let us also define

$$\epsilon_1 \equiv \inf_{|e,f\rangle} \langle e,f|W_\delta|e,f\rangle. \tag{16}$$

Then we have the following lemma.

Lemma 6: Given an edge PPTES δ, then $W_1 \propto W_\delta - \epsilon_1 \mathbb{1}$ is a NDEW, where ϵ_1 and W_δ are defined in Eqs. (16) and (15), respectively.

Proof: We have that $\langle e,f|W_\delta|e,f\rangle = a(\langle e,f|P_1|e,f\rangle + \langle e,f^*|Q_1|e,f^*\rangle) \geq 0$. This quantity is zero iff $\langle e,f|P_1|e,f\rangle = \langle e,f^*|Q_1|e,f^*\rangle = 0$. But this is not possible since δ is an edge PPTES. Thus $\langle e,f|W_\delta|e,f\rangle > 0$ for all $|e,f\rangle$. Defining ϵ_1 as in Eq. (16), and taking into account that $\langle e,f|W_\delta|e,f\rangle$ is a continuous function of (the coefficients of) $|e,f\rangle$, and that the set in which we are taken the infimum is compact, we obtain $\epsilon_1 > 0$. Then we obviously have that W_1 fulfills properties (I) and (III). On the other hand, $\langle \delta \rangle_{W_1} \propto a(\langle \delta \rangle_{P_1} + \langle \delta^T \rangle_{Q_1}) - \epsilon_1 < 0$, since $P_1\delta = Q_1\delta^T = 0$. Thus W_1 detects a PPTES, and therefore, according to theorem 3, is nondecomposable.

Note that lemma 6 provides an important generalization of the method of Terhal [25], based on the use of unextendible product bases [14]. Our method works in Hilbert spaces of arbitrary dimensions, and in particular when $\dim(H_A)=2$ [in $(2\times N)$-dimensional systems] for which unextendible product basis do not exist. By combining lemma 6 and the optimization procedure introduced earlier, we obtain a way of creating NDOEW's. Once we have W_1 we find p_{W_1} and $p_{W_1^T}$. We denote the projector operators orthogonal to these two sets by P_2 and Q_2 , respectively,

$$\epsilon_2 = \inf_{|e,f\rangle} \frac{\langle e,f|W_1|e,f\rangle}{\langle e,f|P_2+Q_2^T|e,f\rangle}, \tag{17}$$

and $W_2 \propto W_1 - \epsilon_2(P_2+Q_2^T)$. According to lemma 2(b) we have that W_2 is ND-finer than W_1. Now we can define p_{W_2}, $p_{W_2^T}$, P_3, Q_3, and W_3 in the same way, and continue in this vein until for some k, $\epsilon_k=0$. If W_k is not yet optimal, we still have to find other projectors such that we can optimize as explained in the previous subsections.

In Sec. VII we illustrate this method with a family of edge PPTES's from Ref. [13]. In fact, as we will mention in that section, we have checked that the optimization method typically works as well by starting with three random vectors, and following a similar procedure to the one indicated here. This means that in our construction method we do not need in practice to start from an edge PPTES.

D. Sufficient condition for PPTES's

In this subsection we use the results derived in Sec. IV C to construct a sufficient criterion for nonseparability of

PPTES's. As shown in Refs. [20,21], given an operator $\rho \geqslant 0$, with $\rho^T \geqslant 0$, we can always decompose it in the form

$$\rho = \rho_s + \delta, \tag{18}$$

where ρ_s is separable and δ is an edge PPTES. More details concerning this decomposition, and in particular its canonical optimal form are presented in Appendix B. In this section we use this decomposition together with the following.

Lemma 7: Given a nonseparable operator $\rho = \rho_s + \delta$, where $\rho_s \geqslant 0$ is separable, then for all EW's, W, such that $\langle \rho \rangle_W < 0$ we have that $\langle \delta \rangle_W < 0$.

Proof: Obvious from the definition of EW.

Lemma 7 tells us that if ρ is nonseparable, then there must exist some EW that detects both δ and ρ. Actually, it is clear that there must exist an OEW with that property. In particular, if $\rho^T \geqslant 0$, it must be a NDOEW. In Sec. IV C we showed how to build these out of edge PPTES's. Thus, given ρ we can always decompose it in the form of Eq. (18), construct an OEW that detects δ, and check whether it detects ρ. In that case, we will have that ρ is nonseparable. Thus this provides a sufficient criterion for nonseparability.

We stress the fact that for PPTES's only a special class of states, namely, the class of edge PPTES's, is responsible for the entanglement properties. In fact, one should stress that many of the examples of PPTES's discussed so far in the literature belong to the class of edge PPTES's: the $2\otimes 4$ family from Ref. [13], the $n\otimes n$ states obtained via unextendible product basis construction [14], the $3\otimes 3$ states obtained via the chessboard method [34](b), and projections of continuous variable PPTES onto finite-dimensional subspaces [34](c).

VI. POSITIVE MAPS

It is known that PM's allow for necessary and sufficient conditions for separability (or, equivalently, entanglement) of bipartite mixed states [11]. Positive maps have been also applied in the context of distillation of entanglement [35] and information theoretic analyses of separability [36]. In this section we will use the isomorphism between operators and linear maps to extend the properties derived for witnesses to PM's [26]. We will first review some of the definitions and properties of linear maps.

Let us consider a linear map $\mathcal{E}: B(H_A) \to B(H_C)$. We say that $\mathcal{E}$ is positive if for all $Y \in B(H_A)$ positive, $\mathcal{E}(Y) \geqslant 0$. One can extend a linear map as follows. Given $\mathcal{E}: B(H_A) \to B(H_C)$, we define its extension $\mathcal{E}\otimes 1_B : B(H_A)\otimes B(H_B) \to B(H_C)\otimes B(H_B)$ according to $\mathcal{E}\otimes 1_C(\Sigma_i Y_i \otimes Z_i) = \Sigma_i \mathcal{E}(Y_i)\otimes Z_i$, where $Y_i \in B(H_A)$ and $Z_i \in B(H_B)$. A linear map is completely positive if all extensions are positive. The classification and characterization of positive (but not completely positive) maps is an open question (see, e.g., Refs. [28,29]).

An example of a positive (but not completely positive) map is transposition (in a given basis O_A); that is, the map $\mathcal{E}_T$ such that $\mathcal{E}_T(Y) = Y^T$. The corresponding extension is the partial transposition [12]. A map $\mathcal{E}$ is called decomposable if it can be written as $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 \cdot \mathcal{E}_T$, where $\mathcal{E}_{1,2}$ are completely positive.

One can relate linear maps with linear operators in the following way. We will assume $d_A \equiv \dim(H_A) \leqslant \dim(H_C)$, but one can otherwise exchange H_A by H_C in what follows. Given $X \in B(H_A \otimes H_C)$ and an orthonormal basis $O_A = \{|k\rangle\}_{k=1}^{d_A}$ in H_A, we define the linear map $\mathcal{E}_X : B(H_A) \to B(H_C)$ according to

$$\mathcal{E}(Y) = \mathrm{tr}_A(X^{T_A} Y) \tag{19}$$

for all $Y \in B(H_A)$, where tr_A denotes the trace in H_A, and the partial transpose is taken in the basis O_A. Similarly, given a linear map we can always find an operator X such that Eq. (19) is fulfilled. For instance, if we choose $T = (|\Psi\rangle\langle\Psi|)^{T_A}$, where

$$|\Psi\rangle = \sum_{k=1}^{d_A} |k\rangle_A \otimes |k\rangle_C, \tag{20}$$

then the corresponding map $\mathcal{E}_T$ is precisely the transposition in the basis O_A.

Given a linear map $\mathcal{E}_X$, one can easily show the following relations: (a) $\mathcal{E}_X$ is completely positive iff $X \geqslant 0$; (b) $\mathcal{E}_X$ is positive but not completely positive iff X is an EW [except for the normalization condition (III)]; and (c) $\mathcal{E}_X$ is decomposable iff X is decomposable. Thus the problem of studying and classifying PM's is very much related to the problem of EW's. Furthermore, PM's can be also used to detect entanglement [11]. Let us consider the extension $\overline{\mathcal{E}_X} \equiv \mathcal{E}_X \otimes 1 : B(H_A)\otimes B(H_B) \to B(H_C)\otimes B(H_B)$, where we take $d_B \equiv \dim(H_B) = \dim(H_C)$. Then we have that, given $\rho \in B(H_A \otimes H_B)$,

$$\langle \rho \rangle_{X^+} = \langle \Psi | \overline{\mathcal{E}_X}(\rho) | \Psi \rangle, \tag{21}$$

where

$$|\Psi\rangle = \sum_{k=1}^{d_B} |k\rangle_C \otimes |k\rangle_B. \tag{22}$$

Thus, if an EW, W, detects ρ, then $\overline{\mathcal{E}_W}(\rho)$ is a nonpositive operator. Consequently, $\rho \geqslant 0$ is entangled iff there exists a PM such that acting on ρ gives a nonpositive operator. In that case we say that the PM "detects" ρ. Actually, PM's are "more efficient" in detecting entanglement than EW's. The reason is that it may happen that $\overline{\mathcal{E}_X}(\rho)$ is nonpositive, but still $\langle \rho \rangle_{X^+} \geqslant 0$.

It is convenient to define finer and optimal PM's as it is for EW's. That is, given two PM's, $\mathcal{E}_{1,2}$, we say that $\mathcal{E}_2$ is finer than $\mathcal{E}_1$ if it detects more. We say that a PM, $\mathcal{E}$, is optimal if there exists no PM that is finer. In the same way we can define ND-finer and ND-optimal. The results presented in the previous sections can be directly translated to PM given the following fact.

Lemma 8: If W_2 is finer (ND-finer) than W_1 then $\mathcal{E}_{W_2}$ is finer (ND-finer) than $\mathcal{E}_{W_1}$.

Proof: Using lemma 2 we can write $W_1 = (1-\epsilon) W_2 + \epsilon P$. According to Eq. (19) we have that $\mathcal{E}_{W_1} = (1-\epsilon)\mathcal{E}_{W_2}$

$+\epsilon\mathcal{E}_P$. Since $\mathcal{E}_P(\rho)\geqslant 0$ for all $\rho\geqslant 0$, we have that $\mathcal{E}_{W_2}$ is finer than $\mathcal{E}_{W_1}$. Using lemma 2(b), we can also prove that it is ND-finer.

From this lemma it follows that optimizing EW's implies optimizing PM's. In fact, the constructions that we gave in Sec. V C can be viewed as ways of constructing nondecomposable PM's. In fact, since the method works for $\dim(H_A)=2$, the resulting PM $\mathcal{E}:B(H_A)\to B(H_C)$ has a minimal ''qubit'' domain, or—equivalently—minimal Hermitian conjugate codomain. To our knowledge, our method is the first that permits one to construct nondecomposable PM's with these characteristics.

VII. ILLUSTRATION

In this section we explicitly give construct a NDOEW out of edge PPTES. We use, as a starting point, the family of PPTES's introduced in Ref. [13]).

A. Family of ''edge'' PPTES's

We consider $H_A=\mathbb{C}^2$ and $H_B=\mathbb{C}^4$, and denote $\{|k\rangle\}_{k=0}^{d_\alpha}$ ($\alpha=A,B$) an orthonormal basis in these spaces, respectively. Most of the time we will write the operators in those bases; that is, as matrices. For operators acting in $H_A\otimes H_B$ we will always use the order $\{|0,0\rangle,|0,1\rangle,\ldots,|1,0\rangle,|1,1\rangle,\ldots\}$. On the other hand, all partial transposes will be taken with respect to H_B.

We consider the following family of positive operators [13]

$$\rho_b=\frac{1}{7b+1}\begin{pmatrix} b&0&0&0&0&b&0&0\\ 0&b&0&0&0&0&b&0\\ 0&0&b&0&0&0&0&b\\ 0&0&0&b&0&0&0&0\\ 0&0&0&0&\frac{1+b}{2}&0&0&\frac{\sqrt{1-b^2}}{2}\\ b&0&0&0&0&b&0&0\\ 0&b&0&0&0&0&b&0\\ 0&0&b&0&\frac{\sqrt{1-b^2}}{2}&0&0&\frac{1+b}{2} \end{pmatrix}, \tag{23}$$

where $b\in[0,1]$. For $b=0$ and 1 those states are separable, whereas for $0<b<1$, ρ_b is an edge PPTES. This can be shown by checking directly that they violate the range criterion of Ref. [13], i.e., the definition given in Sec. IV C.

If we take the partial transpose in the basis $\{|k\rangle\}$, the density operators ρ_b have the property that $\rho_b^T=U_B\rho_bU_B^\dagger$ with $U_B=(\sigma_x)_{03}\oplus(\sigma_x)_{12}$. Here, the subscript ij denotes the subspace, $\mathcal{H}_{Bij}\subset\mathcal{H}_B$ spanned by $\{|i\rangle,|j\rangle\}$, and σ_x is one of the Pauli operators. Note that U_B is a real unitary operator acting only on H_B. This immediately implies that

$$\tilde{\rho}_b^T=\tilde{\rho}_b, \tag{24}$$

where $\tilde{\rho}_b=V_B\rho_bV_B^\dagger$ and $V_B=1\sqrt{2}[(\mathbb{1}+i\sigma_x)_{03}\oplus(\mathbb{1}+i\sigma_x)_{12}]$. We will use property (24) to simplify the problem of constructing the NDOEW. Thus we will concentrate from now on the operators $\tilde{\rho}_b$ [37]. Obviously, $\tilde{\rho}_b$ is an edge PPTES for $1>b>0$.

The projector onto the kernel of $\tilde{\rho}_b$, P_1, is invariant under the transformation $T_{AB}=T_A\otimes T_B$, where

$$T_A=\begin{pmatrix}1&0\\0&e^{i2\pi/3}\end{pmatrix},$$

$$T_B=\begin{pmatrix}1&0&0&0\\0&\cos(2\pi/3)&-\sin(2\pi/3)&0\\0&\sin(2\pi/3)&\cos(2\pi/3)&0\\0&0&0&1\end{pmatrix}. \tag{25}$$

Note that T_B is a real matrix. Later on we will need its eigenstates with real coefficients; they are $|0\rangle\pm|3\rangle$. Note also that $T_{AB}^3=\mathbb{1}$.

B. Construction of NDEW's

We use now the methods developed in Sec. V to obtain a NDOEW starting from $\tilde{\rho}_b$. That is, we define $W_b=P_1+P_1^T$, where P_1 is the projector onto $K(\tilde{\rho}_b)=K(\tilde{\rho}_b^T)$. Our procedure consists of first subtracting the identity to obtain $W_1=W_b-\epsilon_1\mathbb{1}$. Then we subtract $P_2+Q_2^T$, $P_3+Q_3^T$, etc. In the nth step, we have

$$W_n=W_{n-1}-\epsilon_n(P_n+Q_n^T), \tag{26}$$

where P_n (Q_n) is the projector orthogonal to the space spanned by $P_{W_{n-1}}$ ($P_{W_{n-1}^T}$). We will use the symmetries of $\tilde{\rho}_b$ to better understand the structure of W_n.

(a) $W_n=W_n^T$. We can prove this by induction. First, it is clear that $W_1=W_1^T$. Let us now assume that $W_{n-1}=W_{n-1}^T$. Then we show that $W_n=W_n^T$. For this, we just have to show that the subspace spanned by $P_{W_{n-1}}$ is the same as the one spanned by $P_{W_{n-1}^T}$, so that $Q_n=P_n$. But this is clear since $W_{n-1}=W_{n-1}^T$.

(b) $T_{AB}W_nT_{AB}^\dagger=W_n$. We prove this by induction. First, for $W_1=P_1+P_1^T-\epsilon_1\mathbb{1}$ we have that $T_{AB}W_1T_{AB}^\dagger=T_{AB}P_1T_{AB}^\dagger+T_{AB}P_1^TT_{AB}^\dagger-\epsilon_1\mathbb{1}=W_1$, since $T_{AB}P_1^TT_{AB}^\dagger=(T_{AB}P_1T_{AB}^\dagger)^T$ (given the fact that T_B is real) and P_1 is invariant under T_{AB}. Then, let us assume that $T_{AB}W_{n-1}T_{AB}^\dagger=W_{n-1}$. In order to show that $T_{AB}W_nT_{AB}^\dagger=W_n$ we just have to show that P_n is invariant under T_{AB}, or, equivalently, that the subspace spanned by $P_{W_{n-1}}$ is invariant under T_{AB}. But this follows immediately from the fact that $T_{AB}W_{n-1}T_{AB}^\dagger=W_{n-1}$.

Starting from property (a), it follows that the vectors $|e,f\rangle\in P_{W_n}$ will have $|f\rangle$ real (unless we have degeneracies). This can be seen by noting that those vectors minimize $\langle e,f|W_n|e,f\rangle$; defining $W_e\equiv\langle e|W_n|e\rangle$, we have that $W_e^T=W_e=W_e^\dagger$ is symmetric, and therefore the eigenstate corre-

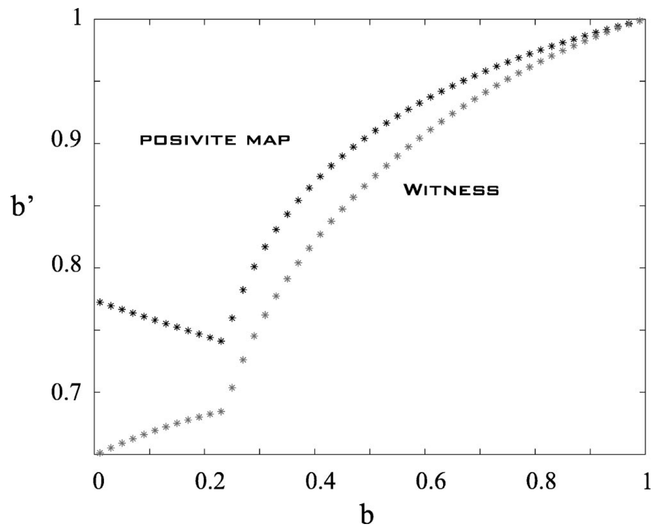

FIG. 1. Values of b' for which if $\tilde{b} \leqslant b'$, $\tilde{\rho}_{\tilde{b}}$ is detected by the witness and the positive map created starting from $\tilde{\rho}_b$.

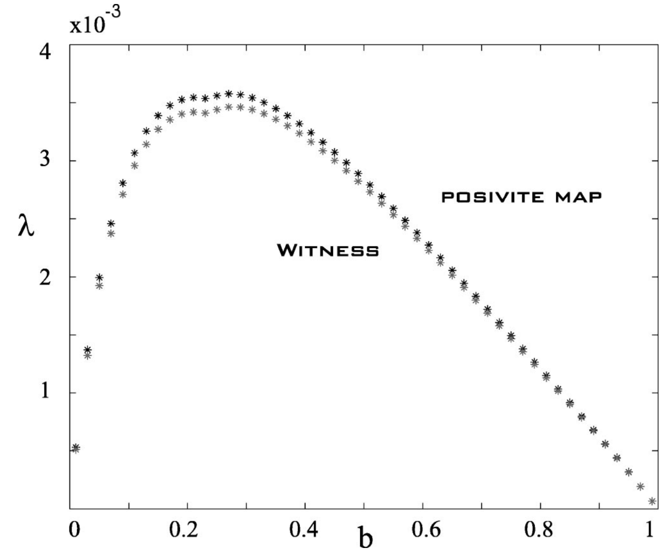

FIG. 2. Maximum λ such that $\tilde{\rho}_b + \lambda \mathbb{1}$ is still detected by the witness, and the positive map created starting from $\tilde{\rho}_b$.

sponding to its minimum eigenvalue can be chosen to be real. On the other hand, starting from the property (b), it follows that if $|e,f\rangle \in P_{W_n}$, then $T_{AB}^{\dagger}|e,f\rangle, T_{AB}^{\dagger 2}|e,f\rangle \in P_{W_n}$. According to this, we will typically have two kinds of product vectors in P_{W_n}.

(1) $|e,f\rangle$ is an eigenstate of $T_{AB}^{\dagger}$ with $|f\rangle$ real: There are only four possible product vectors which fulfill these conditions: $\{|0\rangle,|1\rangle\} \otimes \{|0\rangle + |3\rangle, |0\rangle - |3\rangle\}$.

(2) $|e,f\rangle$ is not an eigenstate of $T_{AB}^{\dagger}$: Then we will also have $T_{AB}^{\dagger}|e,f\rangle$ and $(T_{AB}^{\dagger})^2|e,f\rangle \in P_W$.

We have carried out this procedure for $\tilde{\rho}_b$ and found NDOEW's for each b. We find that for optimal EW's we have two vectors of type (1) and six of type (2). In total we find eight product vectors in P_W, which span the whole Hilbert space, and therefore the corresponding EW's are optimal (see corollary 2). This means that any operator of type (2) with $|e_k, f_k\rangle \in P_W$, the product vectors we have found, and $p_k > 0$ will be a full range separable density operator that lies on the boundary between separable and PPTES's. To our knowledge, this constitutes the first example of these operators [38]. We have also created PM's corresponding to NDOEW's, which we believe are the first examples of nondecomposable PM's with minimal ''qubit'' domains, or—equivalently—minimal Hermitian conjugate codomains.

In Fig. 1 we show for which b' $\tilde{\rho}_{b'}$ is still detected by the NDOEW created out of $\tilde{\rho}_b$. We find that for a given b, the optimal witness that we create detects all $\tilde{\rho}_{\tilde{b}}$ for $\tilde{b} \leqslant b'$. Thus, in the figure we plot b' as a function of b. As explained above, the corresponding positive map detects more than the witness itself. In this figure one can also see how much is detected by the positive map.

Obviously, the witnesses that we create do not only detect the density operators $\tilde{\rho}_b$. For instance, one can check how much one can add to the identity of a certain $\tilde{\rho}_b$ but still keep the state entangled, that is, for which λ the witness still detects $\tilde{\rho}_b + \lambda \mathbb{1}$. This is shown in Fig. 2.

Finally, let us note that using numerical calculations we have observed that if one starts with a random projector P of rank 3, and optimizes the decomposable operator $W \equiv P + P^{T_B}$ in the same way as the one described here, then one will end up with a NDOEW $\tilde{W}$, where $p_{\tilde{W}}$ is complete. This means that in order to create NDOEW one does not need to know in practice an edge PPTES. In another words, optimization itself is a way to reach nondecomposableness.

C. Analytical procedure

In this subsection we will present an analytical way to create NDEW's. Furthermore we will give an example of such a witness, which detects ρ_b for all $b \in (0,1)$. From Fig. 1 we see that the witness which detects the most is the one we created out of $\tilde{\rho}_b$, where b is very close to 1. We will work with the original ρ_b [Eq. (23)].

We consider two Hermitian operators A and B, with A positive on product vectors, i.e., $\langle e,f|A|e,f\rangle \geqslant 0$, whereas B does not have to. As before we denote by P_A (P_B) the (not necessarily complete) set of product vectors on which A (B) vanishes. We require that for all $|e,f\rangle \in P_A$, $\langle e,f|B|e,f\rangle \geqslant 0$. Then we define $W(x) \equiv 1/x(A + xB)$ for any real x. So we have the following lemma

Lemma 9: If $\lim_{x\to 0}\langle \rho \rangle_{W(x)} < 0$, then ρ is entangled.

Proof: We prove that $\lim_{x\to 0}\langle e,f|W(x)|e,f\rangle \geqslant 0$. This implies that if ρ is separable, then $\lim_{x\to 0}\langle \rho \rangle_{W(x} \geqslant 0$. Let us therefore distinguish two cases: (i) if $|e,f\rangle \in P_A$, then we have that $\lim_{x\to 0}\langle e,f|W(x)|e,f\rangle = \langle e,f|B|e,f\rangle$, which is, per assumption, positive. (ii) $|e,f\rangle \notin P_A$; then we have $\lim_{x\to 0}\langle e,f|W(x)|e,f\rangle = \lim_{x\to 0}(a/x) + b$, where $a = \langle e,f|A|e,f\rangle > 0$ and $b = \langle e,f|B|e,f\rangle$. Thus this limit tends to infinity, which proves the statement.

Note that $W(x)$ is not an EW, since it is not necessarily positive on product vectors. However, one can make it positive by adding the identity operator to convert it into an EW.

Corollary 6: Given any $x_0>0$, then $W(x_0)\equiv(1/x_0)(A+x_0B)+\lambda_{x_0}\mathbb{1}$, where $\lambda_{x_0}=-\min_{|e,f\rangle}\langle e,f|1/x_0(A+x_0B)|e,f\rangle$ is an EW.

Let us now illustrate how we can use lemma 9 to detect all the states ρ_b. We define

$$A=\begin{pmatrix} 0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&-2&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&-2&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0 \end{pmatrix}, \tag{27}$$

$$B=\begin{pmatrix} 1&0&0&1&0&-2&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&-2\\ 1&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&-1\\ -2&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&-2&0&-1&0&0&1 \end{pmatrix}. \tag{28}$$

One can easily show that $A=|\psi\rangle\langle\psi|+(|\phi\rangle\langle\phi|)^{T_B}$, where $|\psi\rangle=|01\rangle-|12\rangle$ and $|\phi\rangle=|02\rangle-|11\rangle$. Thus this operator is positive on product vectors, since it is decomposable. Let us now use unnormalized states in order to present the set of product vectors on which A vanishes, i.e., P_A. $P_A=P_{A_1}\cup P_{A_2}$, where $P_{A_1}=\{(|0\rangle+\alpha|1\rangle)\otimes(x|0\rangle+y|3\rangle)\forall\alpha,x,y\}$ and $P_{A_2}=\{(|0\rangle+e^{i\Phi}|1\rangle)\otimes[x|0\rangle+y|3\rangle+z(|1\rangle+e^{-i\Phi}|2\rangle]\forall\Phi,x,y,z\}$. Operator B has to be positive on those product vectors, i.e., $\forall|e,f\rangle\in P_A$, $\langle e,f|B|e,f\rangle\geq 0$. In order to show that this is indeed like that let us distinguish the two cases $|e,f\rangle\in P_{A_1}$ and $|e,f\rangle\in P_{A_2}$. In the first case we have that

$$\langle e|B|e\rangle=\begin{pmatrix} 1+|\alpha|^2 & -2\alpha & 0 & 1-|\alpha|^2\\ -2\alpha^* & 1+|\alpha|^2 & 0 & 0\\ 0&0&1+|\alpha|^2&-2\alpha\\ 1-|\alpha|^2&0&-2\alpha^*&1+|\alpha|^2 \end{pmatrix}, \tag{29}$$

and so $\langle e,f|B|e,f\rangle=|x+y|^2+|\alpha|^2|x-y|^2\geq 0$. If $|e,f\rangle\in P_{A_2}$ then

$$\langle e|B|e\rangle=\begin{pmatrix} 2&-2e^{i\Phi}&0&0\\ -2e^{-i\Phi}&2&0&0\\ 0&0&2&-2e^{i\Phi}\\ 0&0&-2e^{-i\Phi}&2 \end{pmatrix}, \tag{30}$$

which is a positive operator, and so $\langle e,f|B|e,f\rangle\geq 0$. So those two operators A and B fulfill all the required properties. Furthermore, one can show that $\langle\rho_b\rangle_A=0$ and $\langle\rho_b\rangle_B<0$ for all $0<b<1$. Thus we have that $\lim_{x\to 0}\langle W(x)\rho_b\rangle<0$ for all $0<b<1$, where we defined $W(x)=(1/x)(A+xB)$.

As mentioned above, we can use now $W(x)$ in order to create other PPTES's just by adding product vectors on which $W(x)$ vanishes. To find the product vectors we can add, all we need to do is to determine the intersection between P_A and P_B. Since $P_B=\{(|0\rangle+e^{i\phi}|1\rangle)\otimes[a(|0\rangle+e^{-i\phi}|1\rangle+b(|2\rangle+e^{-i\phi}|3\rangle)]\forall\phi,a,b\}$ we have that $S\equiv P_A\cap P_B=P_{A_2}\cap P_B=\{(|0\rangle+e^{i\phi}|1\rangle)\otimes(|0\rangle+e^{-i\phi}|1\rangle+e^{-i2\phi}|2\rangle+e^{-i3\phi}|3\rangle)\forall\phi\}$. Note that S spans a five dimensional subspace, and that the orthogonal subspace is spanned by the vectors $\{-|02\rangle+|13\rangle,-|01\rangle+|12\rangle,-|00\rangle+|11\rangle\}$.

VIII. CONCLUSIONS

Entanglement witnesses allow us to study the separability properties of density operators. We have defined OEW's, which are those that detect entanglement in an optimal way. We have given necessary and sufficient conditions for an EW to be optimal, and we have shown a way to construct them. We have also concentrated on NDEW's, which are those that detect PPTES's. We have extended the definitions of optimality and the optimization procedure to those EW's. It turns out that one can optimize NDEW by subtracting decomposable operators. We have also given an explicit method to construct NDEW's starting from "edge" PPTES's. We have also mentioned that this method works by starting out from random operators. We have extended our techniques to PM's, and therefore given a method to systematically construct nondecomposable positive maps. We have illustrated our methods with a family of edge PPTES acting on $\mathbb{C}^2\otimes\mathbb{C}^4$. The corresponding PM's constitute the first examples of PM's with minimal "qubit" domains, or—equivalently—minimal Hermitian conjugate codo mains. We have also constructed examples of separable states of full range that lie on the boundary between separable states and PPTES's. These states can be used for experimental realization of PPTES's [38].

In this paper we have also introduced the edge PPTES's, which violate the range criterion of separability. As shown in Appendix B, edge PPTES's allow us to construct a canonical form of PPTES's in Hilbert spaces of arbitrary dimensions. They also allow us to give a sufficient condition for nonseparability which applies to operators with positive partial transposes. It is based on the fact that among all PM's (or EW's) only the subset $\{\Lambda_{edge}\}$ of those PM's that detect edge PPTES's are needed to study the separability of PPTES's. This opens many interesting questions. Is it possible that in the set $\{\Lambda_{edge}\}$ there is some map that is globally finer than the transposition? In another words, is there a map detecting the entanglement of *all* the states with nonpositive partial transposes? What is the minimal subset of $\{\Lambda_{edge}\}$ providing such condition? Is it finite?

Finally, let us consider the implications of the our results for the very interesting problem of locality of PPTES. There is a conjecture [39] that those states can be local in the sense that they admit a local hidden variable (LHV) model for any set of possible local measurements. The problem is not trivial given the fact that it may be important to take into account the role of sequential measurements and the possible existence of many copies. Quite recently it was shown that PPTES satisfy Bell-type inequalities introduced by Mermin [40]. It is not difficult to convince oneself that the set of states admitting a LHV model for *any* fixed type of measurements is a convex set. Furthermore, extending the reasoning from Ref. [7], it is easy to see that the set of separable states admits LHV models for any possible set of measurements. Hence, taking into account the results of this paper it follows that in order to prove or disprove the locality of PPTES's it is enough to study only edge PPTES's.

Note that edge states typically have very small rank (the minimal rank is four in $3\otimes 3$ systems, see Ref. [21]). There have been no examples of LHV models for states of low rank, so far. Thus, perhaps, completely new techniques will be needed to study this problem. In this case the most symmetric PPTES's provided recently [34](c) seem to be the best suitable for the first test.

ACKNOWLEDGMENTS

This work was supported in part by the Deutsche Forschungsgemeinschaft (SFB 407 and Schwerpunkt ''Quanteninformationsverarbeitung''), the DAAD, the Austrian Science Foundation (SFB ''control and measurement of coherent quantum systems''), the ESF PESC Program on Quantum Information, TMR network Contract No. ERB–FMRX–CT96–0087, the IST Program EQUIP, and the Institute for Quantum Information GmbH.

APPENDIX A: OPTIMALITY OF EW'S

In this appendix we study necessary and sufficient conditions for an EW to be optimal. According to theorem 1 of Sec. III, an EW, W, is optimal iff no positive operator can be subtracted from W while retaining property (I). This condition can be reexpressed in terms of the infimum of some scalar products in lemma 4. This infimum is, in general, difficult to calculate (at least analytically). In this section we will give a different method to determine whether an EW is optimal or not. This method will turn out to be very simple for the case in which $\dim(H_A)=2$. The idea is to find the conditions such that a given operator $P\geqslant 0$ can or cannot be subtracted from an EW. This will automatically give us a criterion to determine when W is optimal.

In all this appendix we will use that, given an EW, W, and an operator $P\geqslant 0$, we say that P cannot be subtracted from W if for all $\lambda>0$, $W-\lambda P$ does not fulfill (I). In other words, there exist $|e(\lambda)\rangle \in H_A$ and $|f(\lambda)\rangle \in H_B$ such that

$$\langle e(\lambda),f(\lambda)|(W-\lambda P)|e(\lambda),f(\lambda)\rangle<0. \tag{A1}$$

Note that $\langle e(\lambda),f(\lambda)|P|e(\lambda),f(\lambda)\rangle$ must be strictly positive, so that Eq. (A1) can be expressed as

$$\lim_{\lambda\to 0}\frac{\langle e(\lambda),f(\lambda)|W|e(\lambda),f(\lambda)\rangle}{\langle e(\lambda),f(\lambda)|P|e(\lambda),f(\lambda)\rangle}=0. \tag{A2}$$

In Appendix A 1 we will introduce some definitions and notation. In Appendix A 2 we give a method to determine the set of product vectors P_W, on which W vanishes. In Appendix A 3 we find a necessary and sufficient condition under which an operator cannot be subtracted from an EW. We will see that there must exist a vector $|e_0,f_0\rangle \in P_W$, some other vectors $|e_1\rangle$ and $|f_1\rangle$, and certain phases $\phi_{e,f}$ and θ such that some quantity is zero. In Appendix A 1 we will see that the problem can be reduced to finding only the vectors $|e_{0,1}\rangle$ and $|f_{0,1}\rangle$. Finally, we will show that if $\dim(H_A)=2$ we just have to find $|e_0\rangle$ and $|f_0\rangle$, which is very simple.

1. Definitions and notation

In order to prove the results of this appendix in a compact and readable form we have made an extensive numbers of definitions. We will always denote by $|e_0,f_0\rangle$ a product vector in P_W, and by $|e_1\rangle \in H_A$ and $|f_1\rangle \in H_B$ two vectors orthogonal to $|e_0\rangle$ and $|f_0\rangle$, respectively. We will use the notation

$$W_{i,j}^{k,l}=\langle e_i,f_j|W|e_k,f_l\rangle \quad (i,j,k,l=0,1), \tag{A3}$$

and we will write

$$W_{1,0}^{0,1}=|W_{1,0}^{0,1}|e^{i\phi_0}, \tag{A4a}$$

$$W_{0,0}^{1,1}=|W_{0,0}^{1,1}|e^{i\phi_1}. \tag{A4b}$$

We will also define the following operators:

$$w_{i,j}^{e}\equiv\langle e_i|W|e_j\rangle, \tag{A5a}$$

$$w_{i,j}^{f}\equiv\langle f_i|W|f_j\rangle. \tag{A5b}$$

The following vectors will be used in the context of Eq. (A2):

$$|e(\epsilon)\rangle=\frac{1}{\sqrt{1+|\cos(\theta)\epsilon|^2}}[|e_0\rangle+\epsilon\cos(\theta)e^{i\phi_e}|e_1\rangle], \tag{A6a}$$

$$|f(\epsilon)\rangle=\frac{1}{\sqrt{1+|\sin(\theta)\epsilon|^2}}[|f_0\rangle+\epsilon\sin(\theta)e^{i\phi_f}|f_1\rangle], \tag{A6b}$$

where ϵ is a real number, and $\phi_{e,f}\in[0,\pi)$ and $\theta\in[0,\pi/2]$ are certain constants. Given a product vector $|e(\epsilon),f(\epsilon)\rangle$ and an operator W, we will expand $\langle e(\epsilon),f(\epsilon)|W|e(\epsilon),f(\epsilon)\rangle$ by collecting terms with the same powers in ϵ; that is, except for a normalization constant,

$$\langle e(\epsilon),f(\epsilon)|W|e(\epsilon),f(\epsilon)\rangle\propto\sum_{i=1}^{4}\epsilon^i A_i(W), \tag{A7}$$

where

$$A_0(W)=W^{0,0}_{0,0}, \tag{A8a}$$

$$A_1(W)=2\,\mathrm{Re}[\cos(\theta)e^{i\phi_e}W^{1,0}_{0,0}+\sin(\theta)e^{i\phi_f}W^{0,1}_{0,0}], \tag{A8b}$$

$$A_2(W)=\cos^2(\theta)W^{1,0}_{1,0}+\sin^2(\theta)W^{0,1}_{0,1}+2\sin(\theta)\cos(\theta)\times\mathrm{Re}[e^{-i(\phi_e-\phi_f)}W^{0,1}_{1,0}+e^{i(\phi_e+\phi_f)}W^{1,1}_{0,0}], \tag{A8c}$$

$$A_3(W)=2\sin(\theta)\cos(\theta)\times\mathrm{Re}[\cos(\theta)e^{i\phi_f}W^{1,1}_{1,0}+\sin(\theta)e^{i\phi_e}W^{1,1}_{0,1}], \tag{A8d}$$

$$A_4(W)=\sin^2(\theta)\cos^2(\theta)W^{1,1}_{1,1}. \tag{A8e}$$

On the other hand, we will define

$$|\Psi_{0,1}\rangle\equiv\sin(\theta)e^{i\phi_f}|e_0,f_1\rangle+\cos(\theta)e^{i\phi_e}|e_1,f_0\rangle. \tag{A9}$$

Finally, the following quantity will play an important role in determining whether there exist vectors and parameters for which [Eq. (A2)]

$$X(W)\equiv W^{1,0}_{1,0}W^{0,1}_{0,1}-(|W^{0,1}_{1,0}|+|W^{1,1}_{0,0}|)^2. \tag{A10}$$

2. Determining P_W

As stated in lemma 3, not every positive operator P can be subtracted from an EW, W; it must vanish on P_W. Thus, in order to choose P one has to know the set P_W. In this subsection we give a method to determine it. We start by characterizing the vectors in P_W:

Lemma A1: Given an operator W satisfying (I), then $|e_0,f_0\rangle\in P_W$ iff

$$\langle e_0|W|e_0\rangle|f_0\rangle=0, \tag{A11a}$$

$$\langle f_0|W|f_0\rangle|e_0\rangle=0. \tag{A11b}$$

Proof: (If) We just apply $\langle f_0|$ to Eq. (A11a). (Only if) Since W fulfills (I) then $W_{e_0}\equiv\langle e_0|W|e_0\rangle$ must be positive. Thus, $\langle f_0|W_{e_0}|f_0\rangle=0$ implies Eq. (A11a). In the same way we obtain Eq. (A11b).

In practice, for a given W the set P_W can be found as follows. Due to the fact that W is an EW we have that for any $|e\rangle\in H_A$, $W_e\equiv\langle e|W|e\rangle$ must be a positive operator (i.e., $\langle f|W_e|f\rangle\geq 0$ for all $|f\rangle\in H_B$). Thus, the determinant $\det(W_e)\geq 0$. According to lemma A1, this determinant is zero iff there exists some $|f_0\rangle\in H_B$ such that $\langle f_0|W_{e_0}|f_0\rangle=0$, i.e., if $|e_0,f_0\rangle\in P_W$. That is, the determinant as a function of $|e\rangle$ has a minimum (which is zero) at $|e_0\rangle$. We can use this fact to find $|e_0\rangle$. Then, we can easily obtain $|f_0\rangle$ via Eq. (A11a). We can expand an unnormalized state $|e\rangle$ in an orthonormal basis $\{|k\rangle\}$ as

$$|e\rangle=\sum_{k=1}^{\dim(H_A)}c_k|k\rangle, \tag{A12}$$

and impose that the corresponding determinant is zero. This gives us a polynomial equation for the coefficients c_k, i.e.,

$$P(c_k,c_k^*)=0. \tag{A13}$$

We also impose that, given the fact that the determinant is a minimum,

$$\frac{\partial}{\partial c_k}P(c_k,c_k^*)=\frac{\partial}{\partial c_k^*}P(c_k,c_k^*)=0, \tag{A14}$$

which also give a set of polynomial equations. These equations can be solved using the method mentioned in Ref. [20].

3. Necessary and sufficient conditions for subtracting an operator

In this subsection we give a necessary and sufficient condition for an operator P to be subtractable from an EW. We start out by giving some properties of the coefficients $A(W)$ defined above (A8).

Lemma A2: Given W satisfying (I) and $|e_0,f_0\rangle\in P_W$, then for all $|e_1\rangle\in H_A$ and $|f_1\rangle\in H_B$ we have (i) $A_0(W)=A_1(W)=0$. (ii) $A_2(W)\geq 0$. (iii) If $A_2(W)=0$, then $A_3(W)=0$.

Proof: (i) It is a direct consequence from lemma A1. In order to prove (ii) and (iii) we use the fact that W satisfies (I). We define $|e(\epsilon)\rangle$ and $|f(\epsilon)\rangle$ as in Eq. (A6). We impose that $\langle e(\epsilon),f(\epsilon)|W|e(\epsilon),f(\epsilon)\rangle\geq 0$. Using expansion (A7) and taking (i) into account, we have $A(\epsilon)\equiv A_2(W)+\epsilon A_3(W)+\epsilon^2A_4(W)\geq 0$ for all ϵ. This automatically implies (ii), since otherwise for sufficiently small ϵ we would have $A(\epsilon)<0$. It also implies (iii), since if $A_3(W)<0$ ($A_3(W)>0$) then for sufficiently small $\epsilon>0$ ($\epsilon<0$) we would have $A(\epsilon)<0$.

Now, we are at the position of giving a necessary and sufficient condition under which an operator cannot be subtracted from an EW:

Lemma A3: Given P fulfilling $PP_W=0$, it cannot be subtracted from W iff there exists $|e_0,f_0\rangle\in P_W$, $|e_1\rangle\perp|e_0\rangle$, $|f_1\rangle\perp|f_0\rangle$, $\phi_{e,f}$, and θ such that $A_2(W)=0$ but $A_2(P)\neq 0$.

Proof: (If) We define $|e(\lambda)\rangle$ and $|f(\lambda)\rangle$ as in Eq. (A6). Using lemma A2(i), we have $A_0(W)=A_0(P)=A_1(W)=A_1(P)=0$. Using lemma A2(iii), we have that $A_3(W)=0$. Thus we can write limit (A2) as

$$\lim_{\lambda\to 0}\frac{\lambda^2A_4(W)}{A_2(P)+\lambda A_3(P)+\lambda^2A_4(P)}, \tag{A15}$$

which obviously tends to zero given that $A_2(P)\neq 0$. (Only if) There exist two normalized vectors $|\tilde{e}(\lambda)\rangle$ and $|\tilde{f}(\lambda)\rangle$ (continuous functions of λ) fulfilling Eq. (A2). Taking the limit $\lambda\to 0$ in this expression we have that $\langle\tilde{e}(0),\tilde{f}(0)|W|\tilde{e}(0),\tilde{f}(0)\rangle=0$, and therefore $|e_0,f_0\rangle\equiv|\tilde{e}(0),\tilde{f}(0)\rangle\in P_W$. This means that we can always choose $|\tilde{e}(\lambda)\rangle=|e[\epsilon(\lambda)]\rangle$ and $|\tilde{f}(\lambda)\rangle=|f[\epsilon(\lambda)]\rangle$ given in Eq. (A6), where $|e_1\rangle\perp|e_0\rangle$ and $|f_1\rangle\perp|f_0\rangle$ are two normalized vectors, $\lim_{\lambda\to 0}\epsilon(\lambda)=0$, and $\langle e(\epsilon),f(\epsilon)|P|e(\epsilon),f(\epsilon)\rangle\neq 0$. We use Eq. (A6) to expand the numerator and denominator of Eq. (A2) as in Eq. (A7). According to lemma A2(i), we have that $A_0(W)=A_0(P)=A_1(W)=A_1(P)=0$. Thus we must have

$$\lim_{\epsilon\to 0}\frac{A_2(W)+\epsilon A_3(W)+\epsilon^2 A_4(W)}{A_2(P)+\epsilon A_3(P)+\epsilon^2 A_4(P)}=0. \tag{A16}$$

This implies $A_2(W)=0$ and $A_2(P)\neq 0$. Note that if both $A_2(W)=A_2(P)=0$ then, according to lemma A2(iii) we have that $A_3(W)=A_3(P)=0$, so that Eq. (A15) would require $A_4(W)/A_4(P)=0$. But this cannot be, since $A_4(W)=0$ would imply that $|e(\epsilon),f(\epsilon)\rangle\in P_W$, and therefore $\langle e(\epsilon),f(\epsilon)|P|e(\epsilon),f(\epsilon)\rangle=0$.

Finally, we show in the next lemma that condition $A_2(P)=0$ is equivalent to having a certain vector in the kernel of P. We will use the vector $|\Psi_{0,1}\rangle$ defined in Eq. (A9).

Lemma A4: Given a positive operator P and a set of vectors $|e_0,f_0\rangle\in K(P)$, $|e_1\rangle\perp|e_0\rangle$ $|f_1\rangle\perp|f_0\rangle$ and parameters $\phi_{e,f}$ and θ, then $A_2(P)=0$ iff $|\Psi_{0,1}\rangle\in K(P)$.

Proof: Since $P\geqslant 0$ and $|e_0,f_0\rangle\in K(P)$, we have $P_{1,1}^{0,0}=0$. Then we can write $A_2(P)=\langle\Psi_{0,1}|P|\Psi_{0,1}\rangle$, with $|\Psi\rangle$ is defined in Eq. (A9), from which it is obvious that $A_2(P)=0$ iff $|\Psi_{0,1}\rangle\in K(P)$.

4. Necessary and sufficient conditions for $A_2(W)=0$

The previous lemmas tell us that we cannot subtract a given operator P provided we can find some vectors and parameters such that $A_2(W)=0$. The task of finding these vectors is difficult, in general. Here we will give a way to check whether these vectors exist. As before, we will denote by $|e_0,f_0\rangle$ a vector in P_W, and two vectors orthogonal to the first two by $|e_1\rangle$ and $|f_1\rangle$. The quantity $X(W)$ defined in Eq. (A10) will play an important role in determining whether there exist vectors and parameters for which $A_2(W)=0$. In this subsection, we will always have to choose the phases $\phi_{e,f}$ that minimize $A_2(W)$. That is,

$$e^{-i(\phi_e-\phi_f-\phi_0)}=-1,\quad e^{i(\phi_e+\phi_f+\phi_1)}=-1. \tag{A17}$$

We will denote $\tilde{A}_2(W)$ the value of $A_2(W)$ for this particular choice of phases. We have

$$\tilde{A}_2(W)=\cos^2(\theta)W_{1,0}^{1,0}+\sin^2(\theta)W_{0,1}^{0,1} -2\sin(\theta)\cos(\theta)\sqrt{W_{1,0}^{1,0}W_{0,1}^{0,1}-X(W)}, \tag{A18}$$

where we have used Eq. (A10).

Let us start showing that $X(W)$ is positive. We will use this property later on to reexpress condition $A_2(W)=0$ in terms of one that is simpler to check.

Lemma A5: $X(W)\geqslant 0$.

Proof: This follows from the fact that $A_2(W)\geqslant 0$ for all values of $\phi_{e,f}$. In particular, $\tilde{A}_2(W)\geqslant 0$, which according to Eq. (A18) implies $X(W)\geqslant 0$.

The next lemma shows that we just have to check whether $X(W)=0$ if we want to see if there exist parameters $\phi_{e,f}$ and θ such that $A_2(W)=0$. This first condition is therefore much more useful than the last one.

Lemma A6: $X(W)=0$ iff there exist $\phi_{e,f}^0$ and θ^0, such that $A_2(W)=0$.

Proof: (If) Given the phase $\theta=\theta^0$ we have that $0=A_2(W)\geqslant\tilde{A}_2(W)$. Thus $\tilde{A}_2(W)=0$. According to Eq. (A18) we can have two cases: (a) $\theta_0\neq 0,\pi/2$. In this case it is obvious that $X(W)=0$. (b) $\theta_0=0,\pi/2$. In the first (second) case we must have $W_{1,0}^{1,0}=0$ ($W_{0,1}^{0,1}=0$). But this implies that $W_{0,0}^{1,1}=W_{0,1}^{1,0}=0$, since otherwise we could always find some other value of θ such that $\tilde{A}_2(W)<0$. Then, $X(W)=0$. (Only if) We choose $\phi_{e,f}$ as in Eq. (A17). For this value, according to Eq. (A18) we have

$$\tilde{A}_2(W)=[\cos(\theta)\sqrt{W_{1,0}^{1,0}}-\sin(\theta)\sqrt{W_{0,1}^{0,1}}]^2, \tag{A19}$$

which can always be zero for some particular value of θ.

Note that according to the proof of lemma A6, if $W_{1,0}^{1,0}=0$ then $A_2(W)=0$ only for $\theta=0$. But in that case one can easily check that the vector $|e(\lambda),f(\lambda)\rangle\in P_W$ [see Eq. (A6)] which cannot be. Similarly, we conclude that $W_{1,0}^{1,0}\neq 0$ if we want $A_2(W)=0$. Thus from now on we will assume that both $W_{1,0}^{1,0}$ and $W_{1,0}^{1,0}$ are not zero.

5. Optimality test

We can now state the steps to check whether an EW, W, can be optimized or not. (1) For each $|e_0,f_0\rangle\in P_W$ we must check whether there exist $|e_1\rangle\perp|e_0\rangle$ and $|f_1\rangle\perp|f_0\rangle$ such that $X(W)=0$. Let us denote by $|e_{0,1}^{(i)}\rangle$ and $|f_{0,1}^{(i)}\rangle$ the set of vectors fulfilling that. (2) For each of these vectors, we have to find the corresponding values of $\phi_{e,f}^{(i)}$ by using Eq. (A17) and of $\theta^{(i)}$ by imposing that $\tilde{A}_2(W)=0$ in Eq. (A19). (3) Construct $|\Psi^{(i)}\rangle$ according to Eq. (A9). (4) See whether the space spanned by P_W and $\{|\Psi^{(i)}\rangle\}$ is equal to $H_A\otimes H_B$. If it is, then W is optimal. If it is not, we can always find some $|\psi\rangle$ orthogonal to that subspace that can be subtracted from W.

6. Necessary and sufficient conditions for $X(W)=0$

The hard part of the procedure outlined before to see whether an EW is optimal is step (1), namely, to find $|e_1\rangle$ and $|f_1\rangle$ such that $X(W)=0$. We start out by giving a necessary and sufficient condition for $X(W)=0$.

Lemma A7: Given $|e_0,f_0\rangle\in P_W$, and $|e_1\rangle\perp|e_0\rangle$ and $|f_1\rangle\perp|f_0\rangle$, then $X(W)=0$ iff

$$w_{0,0}^{e}|f_1\rangle=-\sqrt{\frac{W_{0,1}^{0,1}}{W_{1,0}^{1,0}}}e^{-i\phi_f}(e^{-i\phi_e}w_{1,0}^{e}+e^{i\phi_e}w_{0,1}^{e})|f_0\rangle, \tag{A20a}$$

$$w_{0,0}^{f}|e_1\rangle=-\sqrt{\frac{W_{0,1}^{0,1}}{W_{1,0}^{1,0}}}e^{-i\phi_e}(e^{-i\phi_f}w_{1,0}^{f}+e^{i\phi_f}w_{0,1}^{f})|e_0\rangle, \tag{A20b}$$

where $\phi_{e,f}$ are given in Eq. (A17).

Proof: (If) We multiply by $\langle f_1|$ [Eq. (A20a)], and take the square of the absolute value of the result. We obtain

$$W^{1,0}_{1,0}W^{0,1}_{0,1}=|e^{-i(\phi_e+\phi_f)}W^{0,0}_{1,1}+e^{i(\phi_e-\phi_f)}W^{1,0}_{0,1}|^2 \leq(|W^{1,1}_{0,0}|+|W^{0,1}_{1,0}|)^2. \quad \text{(A21)}$$

Using lemma A5 we conclude that $X(W)=0$. (Only if) Since $X(W)=0$ and according to lemma A5, $X(W)\geq 0$, then $X(W)$ must be a minimum with respect to $|e_1\rangle$ and $|f_1\rangle$. Taking the derivatives of $X(W)$ with respect to these two vectors and imposing that they vanish, one obtains Eq. (A20).

Equations (A20) are particularly useful if the dimension of one of the Hilbert spaces is 2. Without loss of generality, let us assume that $\dim(H_A)=2$. In that case we can choose $|e_1\rangle$ as the one that is orthogonal to $|e_0\rangle$ (with an arbitrary choice of the global phase). The determination of ϕ_e can be done as follows. Using Eq. (A20), we write

$$\sqrt{\frac{W^{1,0}_{1,0}}{W^{0,1}_{0,1}}}e^{i\phi_f}|f_1\rangle=-\frac{1}{w^e_{0,0}}(e^{-i\phi_e}w^e_{1,0}+e^{i\phi_e}w^e_{0,1})|f_0\rangle, \quad \text{(A22)}$$

where $1/w^e_{0,0}$ denotes the pseudoinverse [33]. We can use this expression to impose

$$W^{0,1}_{1,0}e^{-i(\phi_e-\phi_f)},W^{1,1}_{0,0}e^{i(\phi_e+\phi_f)}<0, \quad \text{(A23)}$$

i.e. they are negative real numbers. We obtain that

$$e^{-i2\phi_e}\langle f_0|w^e_{1,0}\frac{1}{w^e_{0,0}}w^e_{1,0}|f_0\rangle<0, \quad \text{(A24)}$$

so that we determine ϕ_e. With these results, we can prove the following necessary and sufficient condition for $X(W)=0$ when $\dim(H_A)=2$.

Lemma A8: If $\dim(H_A)=2$, given $|e_0,f_0\rangle\in P_W$, then there exists $|e_1,f_1\rangle$ such that $X(W)=0$ iff

$$\langle f_0|\left[w^e_{1,1}-w^e_{0,1}\frac{1}{w^e_{0,0}}w^e_{1,0}-w^e_{1,0}\frac{1}{w^e_{0,0}}w^e_{0,1}\right]|f_0\rangle =2\left|\langle f_0|w^e_{0,1}\frac{1}{w^e_{0,0}}w^e_{0,1}|f_0\rangle\right|. \quad \text{(A25)}$$

Proof: (If) We define

$$|f_1\rangle=-\frac{1}{w^e_{0,0}}(e^{-i\phi_e}w^e_{1,0}+e^{i\phi_e}w^e_{0,1})|f_0\rangle, \quad \text{(A26)}$$

where ϕ_e is determined by condition (A24). Using this expression to calculate $X(W)$ one finds that indeed $X(W)=0$. (Only if) Using lemma A7 we can write $|f_1\rangle$ as in Eq. (A22) so that the phases $\phi_{e,f}$ ensure that Eq. (A23) is fulfilled. Substituting $|f_1\rangle$ into the equation $X(W)=0$, one finds Eq. (A25).

In summary, for a given $|e_0,f_0\rangle\in P_W$, in order to find whether there exist $|e_1,f_1\rangle$ such that $X(W)=0$ we just have to check condition (A25). If it is fulfilled, we can easily find $|f_1\rangle$ and the phases $\phi_{e,f}$ using Eqs. (A23) and (A22).

APPENDIX B: CANONICAL FORMS OF PPTES'S

The concept of edge PPTES's seems to play a very special role in the characterization of PPTES's. In particular, in view of the criterion given in Sec. V D, which is based on the fact that any density operator ρ can be decomposed into a separable part and an edge PPTES [Eq. (18)]. Among all the possible decompositions there might be one for which the trace of the separable part is maximal. When it exists, such a decomposition was termed positive partial transpose best separable approximation (PPT BSA) to ρ [21]. It extended the idea of the BSA introduced in Refs. [23,22] to the case of PPTES's, which were based on the method of diminishing the range of ρ by subtracting product vectors from its range, while keeping the remainder and, at the same time, its partial transpose, positive [23,22,20,21]. In this appendix we formalize the results regarding the existence and properties of the PPT BSA. In particular, the proofs presented in the quoted papers were restricted to the case in which there exists a finite or, at most, countable number of projectors on product vectors that can be subtracted from ρ. We will extend them below to continuous families of product vectors. This appendix is written in a self-contained way, and can be read independently of the body of the paper.

We denote by Γ_ρ the set of projectors on product vectors $\{|e_\alpha,f_\alpha\rangle\langle e_\alpha,f_\alpha|\}$ such that $|e_\alpha,f_\alpha\rangle\in R(\rho)$ and $|e_\alpha,f^*_\alpha\rangle\in R(\rho^{T_B})$. In Ref. [21] we showed that if Γ_ρ is finite then there exists an optimal decomposition (the PPT BSA) $\rho=(1-p)\rho_{sep}+p\delta$, where δ is an edge PPTES, and p is minimal. Note that the PPT BSA involves a state δ which violates the range criterion in a rather special way, i.e., with the additional requirement that Γ_ρ is a finite set. It can happen that there is an uncountable family of product vectors depending on a continuous parameter that can be used for subtracting projectors. In the following we will show that in such case the above result is valid. In order to consider the case of continuous families of product vectors we first prove the following lemma

Lemma B1: Let ρ will be a PPTES defined on a Hilbert space $\mathcal{H}$, $\dim\mathcal{H}<\infty$. Then the set of product vectors Γ_ρ is compact.

Proof: Obviously Γ_ρ is a bounded set in finite-dimensional space, so it is enough to show that it is closed. Consider any sequence $|g_n,h_n\rangle\to|\phi\rangle$, $|g_n,h_n\rangle\in R(\rho)$, $|g_n,h^*_n\rangle\in R(\rho^{T_B})$. The limit vector must (i) respect the condition of orthogonality to $K(\rho)$ [i.e. they must belong to $R(\rho)$], (ii) belong to the sphere (i.e., set of all vectors $|\phi\rangle$ with $||\phi||=1$); and (iii) must be a product state, because if it was entangled then its distance from the compact set of product pure states [13] defined as $\min_{|e,f\rangle}|||\phi\rangle-|e,f\rangle||$ would be nonzero, which is obviously impossible. We conclude thus $|\phi\rangle=|g,h\rangle\in R(\rho)$ for some $|g\rangle$, and $|h\rangle$, which implies (up to irrelevant phase factors) that $|g_n\rangle\to|g\rangle$ and $|h_n\rangle\to|h\rangle$. We have (again up to an irrelevant external phase factor)

$|g_n, h_n^*\rangle \rightarrow |g, h^*\rangle$. The latter must belong to $R(\rho^{T_B})$ as any element of the corresponding sequence is orthogonal to $K(\rho^{T_B})$.

Let us now prove the following general lemma, which is a generalization of one theorem from Ref. [22]:

Lemma B2: Let the PPTES ρ be defined on a finite-dimensional Hilbert space. Consider the set Σ_ρ consisting of the trivial zero operator plus all unnormalized states $\tilde{\rho}$(tr$\tilde{\rho}$ $\leqslant 1$) such that $\tilde{\delta} \equiv \rho - \tilde{\rho}$ is positive and has positive partial transpose. Then one can find $\hat{\rho} \in \Sigma_\rho$, such that with tr($\hat{\rho}$) $\leqslant 1$ is optimal in the sense that (i) the trace of $\hat{\delta} \equiv \rho - \hat{\rho}$ is minimal with respect to all separable $\tilde{\rho}$'s, leading to positive partial transpose $\tilde{\delta}$'s; and (ii) the state $\delta = \hat{\delta}/\mathrm{tr}(\hat{\delta})$ is an edge PPTES.

Proof: To prove the existence of $\hat{\rho} \in \Sigma_\rho$, we just have to show that Σ_ρ is compact. This can be done by showing that Σ_ρ is a closed subset of another compact set, namely, $C = \mathrm{conv}\{\Gamma_\rho \cup \mathbf{0}\}$. The latter set C is compact as it is a convex hull of the compact set $\{\Gamma_\rho \cup \mathbf{0}\}$ in a finite-dimensional space.

Note first that $\Sigma_\rho \subset C$. Indeed, by virtue of $\tilde{\delta} \geqslant 0$ any nonzero $\tilde{\rho}$ cannot have any vector in its range not belonging to $R(\rho)$. Analogously $R(\tilde{\rho}^{T_B}) \subset R(\rho^{T_B})$. Hence, according to the properties of the ranges of density operators in general [13], $\tilde{\rho}$ must be a convex combination of vectors from Γ_ρ, and as such it belongs to C. Let us show that Σ_ρ is closed. This follows immediately from the fact that Σ_ρ is a cross section (performed over any projections P, and Q) of the sets $\Sigma_{\rho,P}^1 \equiv \{\tilde{\rho}: f_{P,\rho}(\tilde{\rho}) \equiv \mathrm{tr}(P\rho - P\tilde{\rho}) \geqslant 0\}$ and $\Sigma_{\rho,Q}^2 = \{\tilde{\rho}: g_{Q,\rho}(\tilde{\rho}) \equiv \mathrm{tr}(Q^{T_B}\rho - Q^{T_B}\tilde{\rho}) \geqslant 0\}$. Since the functions $f_{P,\rho}, g_{Q,\rho}$ are continuous, all the sets participating in the cross section are closed. Now the cross section of closed sets is again a closed one.

Now consider statement (ii). Since $\delta, \delta^{T_B} \geqslant 0$, we always have $\delta = \beta P_{R(\delta)} + A$ and $\delta^{T_B} = \beta' P_{R(\delta^{T_B})} + A'$ with $\beta, \beta' > 0$, some positive operators A and A' (here P_X denotes a projector onto the subspace $X \subset \mathcal{H}$). Then if, contrary to (ii), there were any $|e,f\rangle \in R(\delta)$ such that $|e,f^*\rangle \in R(\delta^{T_B})$, then the new operator $\hat{\rho}^* = \hat{\rho} + \gamma |e,f\rangle\langle e,f|$, $\gamma = \min[\beta, \beta']$ would fulfill that $\hat{\delta}^* = \rho - \hat{\rho}^*$ is PPTES, and would contradict optimality with respect to (i).

Let us remark that if we give up the condition regarding the positivity of $\tilde{\delta}^{T_B}$, then we obtain a modified statement (ii) where state δ has no product vectors in its range. This is nothing but the BSA of Ref. [22], extended here rigorously to the states ρ having uncountable set of product vectors in $R(\rho)$.

From the lemma B2 we obtain the following characterization of PPTES's, which can be regarded to be among the main results of this appendix, since it provides *a canonical form of PPTES's*:

Proposition: If the state ρ is a PPTES, then it is a convex combination

$$\rho = (1-p)\rho_{sep} + p\,\delta \tag{B1}$$

of some normalized separable ρ_{sep} and an normalized edge PPTES δ. In the above decomposition the weight p is minimal [i.e., there does not exist a decomposition of type (B1) with a smaller p].

The above proposition means, in particular, that *edge PPTES's are responsible for PPT-type entanglement.*

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[2] E. Schrödinger, Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[3] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[4] C.H. Bennett and S.J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[5] C. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[6] See A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996); A. Steane, Rep. Prog. Phys. **61**, 117 (1998); A. Barenco, Contemp. Physics **37**, 375 (1996).

[7] R.F. Werner, Phys. Rev. A **40**, 4277 (1989).

[8] An excellent introduction to the problematics of quantum correlation and entanglement is provided by A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1995).

[9] See P. Horodecki, M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **82**, 1056 (1999).

[10] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[11] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 8 (1996).

[12] Given an operator X and an orthonormal basis $\{|k\rangle\}_{k=1}^N \in H_B$, one defines the partial transpose of X with respect to B in that basis as follows:

$$X^{T_B} = \sum_{k,k'=1}^N |k'\rangle_B\langle k|X|k'\rangle_B\langle k|. \tag{B2}$$

One can analogously define the partial transpose of X with respect to A in a given basis, X^{T_A}. Partial transposition fulfills the following useful property:

$$\mathrm{tr}(X^{T_B}Y) = \mathrm{tr}(XY^{T_B}). \tag{B3}$$

We say that a positive operator has a positive partial transposition if $\rho^{T_B} \geqslant 0$. Note that this property is basis independent, and that $\rho^{T_B} \geqslant 0$ iff $\rho^{T_A} \geqslant 0$. The relation between separable and positive partial transpose operators was established in Refs. [10,11]. All separable operators have a positive partial transposition. However, the converse is not true in general. That is, there are positive partial transpose operators which are nonseparable.

[13] P. Horodecki, Phys. Lett. A **232**, 333 (1997).

[14] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999); D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal, quant-ph/9908070; C.H. Bennett, D.P. DiVincenzo, Ch.A.

Fuchs, T. Mor, E. Rains, P.W. Shor, J.A. Smolin, and W.K. Wootters, Phys. Rev. A **59**, 1070 (1999); See also R. Horodecki, M. Horodecki, and P. Horodecki, *ibid.* **60**, 4144 (1999).

[15] For a primer on separability, see M. Lewenstein, D. Bruß, J. I. Cirac, B. Kraus, M. Kuś, J. Samsonowicz, A. Sanpera, and R. Tarrach, in *Proceedings of the Conference ''Quantum Optics Kuhtai 2000,''* edited by F. Ehlotzky and P. L. Knight [J. Mod. Opt. (to be published)].

[16] For an extensive review, see M. Horodecki, P. Horodecki, and R. Horodecki, in *Quantum Information—Basic Concepts and Experiments*, edited by A. Zeilinger, H. Weinfurter, R. Werner, and Th. Beth (Springer, Berlin, 2000).

[17] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **230**, 377 (1996).

[18] K. Zyczkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).

[19] G. Vidal and R. Tarrach, Phys. Rev. A **59**, 141 (1999); S.L. Braunstein, C.M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack, Phys. Rev. Lett. **83**, 1054 (1999).

[20] B. Kraus, J.I. Cirac, S. Karnas, and M. Lewenstein, e-print quant-ph/9912010; Phys. Rev. A (to be published).

[21] P. Horodecki, M. Lewenstein, G. Vidal, and J.I. Cirac, e-print quant-ph/0002089; Phys. Rev. A (to be published).

[22] M. Lewenstein and A. Sanpera, Phys. Rev. Lett. **80**, 2261 (1998).

[23] A. Sanpera, R. Tarrach, and G. Vidal, Phys. Rev. A **58**, 826 (1998); G. Vidal, Ph.D. thesis, Universitat de Barcelona, 1999 (unpublished).

[24] For recent progress on optimal decompositions see B.G. Englert and N. Metwally, quant-ph/9912989; quant-ph/0007053.

[25] B.M. Terhal, e-print quant-ph/9810091; also see Phys. Lett. A **271**, 319 (2000).

[26] A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).

[27] E. Störmer, Acta Math. **110**, 233 (1963).

[28] S.L. Woronowicz, Rev. Mod. Phys. **10**, 165 (1976).

[29] S.L. Woronowicz, Commun. Math. Phys. **51**, 243 (1976); P. Kryszyński and S.L. Woronowicz, Lett. Math. Phys. **3**, 319 (1979); M.D. Choi, Proc. Sympos. Pure Math. **38**, 583 (1982).

[30] This normalization condition expresses the fact that if W fulfills (I) and (II), and λ is a positive number, then λW also fulfills (I) and (II), so that we can always normalize W to have (III). Given an operator satisfying (I) and (II) we can always define $W' = W/\mathrm{tr}(W)$, so that it fulfills both properties as well and detects the same operators as W. This follows from the fact that $\mathrm{tr}(W)$ must be strictly positive since, if we take the trace in an orthonormal basis of product vectors, using (I) we have that $\mathrm{tr}(W) \geqslant 0$. If $\mathrm{tr}(W)=0$ then we would have that $\langle e,f|W|e,f\rangle = 0$ for all product vectors; since we can always construct a complete set of operators $\{|e_i,f_i\rangle\langle e_i,f_i|\}$ out of product vectors in the space of the operators acting on $H_A \otimes H_B$, this immediately implies that $W=0$, which contradicts (II). This set of operators can be constructed as follows. Let us consider a Hilbert space H_A of dimension d_a and an orthonormal basis $\{|k\rangle\}_{k=1}^{d_a}$. We define the states $|\Psi^r_{k,k'}\rangle = (|k\rangle + |k'\rangle)/\sqrt{2}$ and $|\Psi^i_{k,k'}\rangle = (|k\rangle + i|k'\rangle)/\sqrt{2}$ for $k<k'$, and the projectors on these states $A_k = |k\rangle\langle k|$, $A^r_{k,k'} = |\Psi^r_{k,k'}\rangle\langle\Psi^r_{k,k'}|$ and $A^i_{k,k'} = |\Psi^i_{k,k'}\rangle\langle\Psi^i_{k,k'}|$. The set $S_A = \{A_k, A^r_{k,k'}, A^i_{k,k'}, k=1,\ldots,d_a,\ k'=k+1,\ldots,d_a\}$ is complete in the space of operators acting on H_A. This can be easily shown by noting that $|k\rangle\langle k'| = A^r_{k,k'} + iA^r_{k,k'} - (1+i) \times (A_k + A_{k'})/2$ for $k \neq k'$. We can do the same construction for H_B, and obtain a set S_B. The set composed of tensor products of all elements of S_A with S_B is complete in the space of operators acting on $H_A \otimes H_B$.

[31] Whenever it does not matter whether we take the partial transpose of an operator X with respect to the first or second system, we will simply write X^T.

[32] W. Rudin, *Functional Analysis* (McGraw Hill, New York, 1973).

[33] Given a self-adjoint operator X, its pseudoinverse X^{-1} is defined as the one that vanishes in $K(X)$, and so that $XX^{-1} = X^{-1}X$ is the projector in $R(X)$. If X has no zero eigenvalue, the pseudoinverse coincides with the inverse operator.

[34] For more examples of PPTES's also see (a) Ref. [9]; (b) D. Bruß and A. Peres, Phys. Rev. A **61**, 030301(R) (2000); (c) P. Horodecki and M. Lewenstein, e-print quant-ph/0001035; Phys. Rev. Lett. (to be published).

[35] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).

[36] N.J. Cerf, C. Adami, and R.M. Gingrich, Phys. Rev. A **60**, 893 (1999).

[37] Note that if we define a basis in H_B as $\{V_B^\dagger|k\rangle\}$ we have that $\rho_b^T = \rho_b$, where the partial transposition is taken in that basis.

[38] These states can be prepared locally, and can be used to construct PPTES's by mixing them ''weakly'' with entangled states. This provides an interesting possibility of an experimental realization of PPTES's. This suggestion was formulated by A. Weinfurter.

[39] A. Peres, Found. Phys. **29**, 589 (1999).

[40] R.F. Werner and M.M. Wolf, e-print quant-ph/9910063.