

# Characterizing entanglement

Cite as: J. Math. Phys. **43**, 4237 (2002); <https://doi.org/10.1063/1.1494474>

Submitted: 09 May 2002 • Accepted: 16 May 2002 • Published Online: 20 August 2002

Dagmar Bruß



View Online



Export Citation

## ARTICLES YOU MAY BE INTERESTED IN

[Entangled quantum systems and the Schmidt decomposition](#)

American Journal of Physics **63**, 415 (1995); <https://doi.org/10.1119/1.17904>

[Quantum Computation and Quantum Information](#)

American Journal of Physics **70**, 558 (2002); <https://doi.org/10.1119/1.1463744>

[Quantum Entanglement: A Modern Perspective](#)

Physics Today **56**, 46 (2003); <https://doi.org/10.1063/1.1580049>

Journal of  
Mathematical Physics

Receive the latest **research updates**

Sign up for Alerts

# Characterizing entanglement

Dagmar Bruß<sup>a)</sup>

*Institut für Theoretische Physik, Universität Hannover,  
Appelstr. 2, D-30167 Hannover, Germany*

(Received 9 May 2002; accepted for publication 16 May 2002)

Quantum entanglement is at the heart of many tasks in quantum information. Apart from simple cases (low dimensions, few particles, pure states), however, the mathematical structure of entanglement is not yet fully understood. This tutorial is an introduction to our present knowledge about how to decide whether a given state is separable or entangled, how to characterize entanglement via witness operators, how to classify entangled states according to their usefulness (i.e., distillability), and how to quantify entanglement with appropriate measures.

[DOI: 10.1063/1.1494474]

## I. INTRODUCTION

The role of entanglement in quantum information processing is manifold. Rather than considering entanglement as a mystery, like in the early years of quantum mechanics, it is nowadays viewed as a resource for certain tasks that can be performed faster or in a more secure way than classically. This genuinely new aspect of quantum properties has launched intensive experimental efforts to create entangled states and theoretical efforts to understand the mathematical structure of entanglement. This tutorial presents the status regarding our understanding of entanglement.

As the number of articles about entanglement has enormously increased during the last 10 years, it is almost impossible to give a complete overview, and this is not the purpose of this article. It will rather introduce the reader to the established knowledge and some important tools in this field, and discuss some of the questions that remain open at present.

Throughout this tutorial, we will mostly consider entanglement of just two parties, unless stated otherwise explicitly. Most concepts can be explained best with bipartite systems; some of them could then be generalized to more parties in an evident way, for others the situation changes completely for more than two parties. Not many results are known for multipartite systems. In Sec. VD some reasons for this will become clear from studying tripartite states.

Various aspects of entanglement have recently been summarized in the following review articles, which were partially used as a source for this tutorial: the “primer”<sup>1</sup> aims at introducing the nonexpert reader to the problem of separability and distillability of quantum states. The Horodecki family discusses entanglement in the context of quantum communication,<sup>2</sup> where the distillability properties of a given state are important. B. Terhal summarizes the use of witness operators for detecting entanglement in Ref. 3. Various entanglement measures are presented in the context of the theorem of their uniqueness in Ref. 4. Other reviews on theoretical and experimental aspects of entanglement can be found in the first issues of the newly launched journal QIC.<sup>5</sup>

This article is organized as follows: Section II presents various possible answers to the question “what is entanglement?,” thus shedding light upon different facets of quantum correlations. In Sec. III several criteria are introduced that allow us to distinguish separable from entangled states. In Sec. IV we discuss the possibility to distill entanglement, and give a distillability criterion. Finally, Sec. V concerns attempts to quantify entanglement via entanglement measures. Some important measures are defined and their properties are discussed. The classification of

---

<sup>a)</sup>Electronic mail: bruss@itp.uni-hannover.de

entangled states according to their Schmidt number is introduced, and a generalization to tripartite states is included.

## II. WHAT IS ENTANGLEMENT?

It is nearly 70 years ago that Erwin Schrödinger gave the name “Verschränkung” to a correlation of quantum nature.<sup>6</sup> In colloquial German for nonphysicists this term is only used in the sense of “folding the arms.” It was then rather loosely translated to “entanglement,” with more inspiring connotations.

Over the decades the meaning of the word “entanglement” has changed its flavor. The following list is an attempt to sketch the attitude towards entanglement of various important persons in the fields of foundations of quantum physics and later in quantum information theory. These statements are not quotations, unless indicated explicitly.

*Einstein/Podolsky/Rosen:* An entangled wavefunction does not describe the physical reality in a complete way.

*E. Schrödinger:* For an entangled state “the best possible knowledge of the whole does *not* include the best possible knowledge of its parts.”<sup>6</sup>

Entanglement is ...

*J. Bell:* ...a correlation that is stronger than any classical correlation.

*D. Mermin:* ...a correlation that contradicts the theory of elements of reality.

*A. Peres:* “...a trick that quantum magicians use to produce phenomena that cannot be imitated by classical magicians.”<sup>7</sup>

*C. Bennett:* ...a resource that enables quantum teleportation.

*P. Shor:* ...a global structure of the wavefunction that allows for faster algorithms.

*A. Ekert:* ...a tool for secure communication.

*Horodecki family:* ...the need for first applications of positive maps in physics.

Our view of the nature of entanglement may continue to be modified during the coming years.

## III. GIVEN A QUANTUM STATE, IS IT SEPARABLE OR ENTANGLED?

In this section we will summarize *operational* and *nonoperational* criteria that allow us to classify a given state as separable or entangled. Here the word “operational” is used in the sense of “user-friendly:” an operational criterion is a recipe that can be applied to an explicit density matrix  $\varrho$ , giving some immediate answer like “ $\varrho$  is entangled,” or “ $\varrho$  is separable,” or “this criterion is not strong enough to decide whether  $\varrho$  is separable or entangled.”

But, first of all, we need a mathematical definition for entanglement versus separability. This is very simple for *pure states*: a pure state  $|\psi\rangle$  is called *separable* iff it can be written as  $|\psi\rangle = |a\rangle \otimes |b\rangle$ , otherwise it is *entangled*. (Remember that throughout most of this article we talk about bipartite entanglement. In some cases the generalization to more particles is straightforward, like here.) An example for a pure separable state is  $|\psi\rangle = |00\rangle$ ; examples for pure entangled states are the *Bell states*

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad (1)$$

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \quad (2)$$

A *mixed state* is called separable if it can be prepared by the two parties (which are traditionally called Alice and Bob) in a “classical” way, that is, by agreeing over the phone on the *local* preparation of states. A density matrix that has been created in this way can only contain classical correlations. Mathematically this means: a mixed state  $\varrho$  is called *separable* iff it can be written as<sup>8</sup>

$$\varrho = \sum_i p_i |a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i|, \quad (3)$$

otherwise it is *entangled*. Here the coefficients  $p_i$  are probabilities, i.e.,  $0 \leq p_i \leq 1$  and  $\sum_i p_i = 1$ . Note that in general  $\langle a_i | a_j \rangle \neq \delta_{ij}$ , and also Bob’s states need not be orthogonal. This decomposition is not unique. An example for a mixed separable state that contains classical correlations, but no quantum correlations, is  $\varrho = 1/2(|00\rangle\langle 00| + |11\rangle\langle 11|)$ . An example for a mixed entangled state is a Werner state, an admixture of a Bell state as in (1) or (2) to the identity:  $\varrho_W = (1-p)/4 \mathbb{1} + p|\Phi^+\rangle\langle\Phi^+|$  with  $\frac{1}{3} < p \leq 1$ . The lower limit of  $p$  for  $\varrho_W$  to be entangled can be easily found with the operational criteria discussed below. In fact, any density matrix that is “close enough” to the identity is separable.<sup>9</sup>

Finding a decomposition as in (3) for a given  $\varrho$ , or proving that it does not exist, is a nontrivial task which has been solved explicitly only for a few cases. Therefore, this simple-looking definition of separability is by no means “user-friendly,” and we are in demand of criteria that are easier to test.

### A. Operational separability criteria

Here we present some separability criteria that are easy to check in an explicit case. In the following we will assume that  $\varrho \in \mathcal{H}_A \otimes \mathcal{H}_B$  with  $\dim \mathcal{H}_A = M$  and  $\dim \mathcal{H}_B = N \geq M$ , without loss of generality.

For *pure states* there is a very simple necessary and sufficient criterion for separability, the **Schmidt decomposition**. A pure state has Schmidt rank  $r \leq M$  if it can be decomposed as the biorthogonal sum

$$|\psi^r\rangle = \sum_{i=1}^r a_i |e_i\rangle |f_i\rangle, \quad (4)$$

with  $a_i > 0$  and  $\sum_{i=1}^r a_i^2 = 1$ , where  $\langle e_i | e_j \rangle = \delta_{ij} = \langle f_i | f_j \rangle$ . Note that  $a_i^2$  are the eigenvalues of the reduced density matrices, and therefore the Schmidt rank is easy to compute. A given pure state  $|\psi\rangle$  is separable iff  $r = 1$ .

For *mixed states* the situation is less simple. There are several operational separability criteria for this case. Here they are ordered in decreasing strength, i.e., the last criterion fails to detect an entangled state as entangled in more cases than the previous ones:

(1) **Peres-Horodecki criterion (positive partial transpose):**<sup>10,11</sup> The partial transpose of a composite density matrix is given by transposing only one of the subsystems. Thus, the entries of a density matrix that is partially transposed with respect to Alice are given by

$$(\varrho^{TA})_{m\mu, n\nu} = \varrho_{n\mu, m\nu}, \quad (5)$$

where Latin indices are referring to Alice’s subsystem and Greek ones to Bob’s subsystem. As any separable state can be decomposed according to (3), its partial transpose is given by

$$\varrho_{\text{sep}}^{T_A} = \sum_i p_i (|a_i\rangle\langle a_i|)^T \otimes |b_i\rangle\langle b_i|. \quad (6)$$

Since the  $(|a_i\rangle\langle a_i|)^T$  are again valid density matrices for Alice, one finds immediately that  $\varrho_{\text{sep}}^{T_A} \geq 0$ . The same holds for partial transposition with respect to Bob (or any other party for multipartite systems). In conclusion, the partial transpose of a separable state  $\varrho$  with respect to any subsystem is positive.<sup>10</sup> (In our terminology a positive operator has positive or vanishing eigenvalues—more precisely it should be called positive semidefinite. The expectation value of a positive operator with any state is positive or zero.)

It was shown in Ref. 11 for bipartite systems that the converse (i.e., if  $\varrho^{T_A} \geq 0$ , then  $\varrho$  is separable) is true only for low-dimensional systems, namely for composite states of dimension  $2 \times 2$  and  $2 \times 3$ . In this case the positivity of the partial transpose (PPT) is a necessary and sufficient condition for separability. For higher dimensions it is only necessary, and the existence of entangled PPT states has been shown<sup>12</sup>—these states have been called *bound entangled states*, as their entanglement does not seem to be “useful,” as explained in Sec. IV.

(2) **Reduction criterion:**<sup>13</sup> According to the reduction criterion, if  $\varrho$  is separable, then

$$\varrho_A \otimes \mathbb{1} - \varrho \geq 0 \quad \text{and} \quad \mathbb{1} \otimes \varrho_B - \varrho \geq 0, \quad (7)$$

where  $\varrho_A$  is Alice’s reduced density matrix, and  $\varrho_B$  is Bob’s. In order to understand why the positivity of the left-hand sides in (7) is a separability criterion, one has to note that they correspond to the application of the positive map  $\Lambda(\sigma) = (\text{Tr} \sigma) \mathbb{1} - \sigma$  to Bob’s subsystem, or to Alice’s subsystem. (The important role of positive maps will be discussed in the next subsection, Sec. III B.) A positive map applied to one subsystem of a separable state preserves the properties of a density matrix—therefore the resulting density matrix has to remain positive.

Like the partial transpose criterion, the reduction criterion is a necessary and sufficient separability condition only for dimensions  $2 \times 2$  and  $2 \times 3$ , and a necessary condition otherwise.

(3) **Majorization criterion:**<sup>14</sup> The majorization criterion says that if a state  $\varrho$  is separable, then

$$\lambda_{\varrho}^{\downarrow} < \lambda_{\varrho_A}^{\downarrow} \quad \text{and} \quad \lambda_{\varrho}^{\downarrow} < \lambda_{\varrho_B}^{\downarrow} \quad (8)$$

has to be fulfilled. Here  $\lambda_{\varrho}^{\downarrow}$  denotes the vector consisting of the eigenvalues of  $\varrho$ , in decreasing order, and a vector  $x^{\downarrow}$  is *majorized* by a vector  $y^{\downarrow}$ , denoted as  $x^{\downarrow} < y^{\downarrow}$ , when  $\sum_{j=1}^k x_j^{\downarrow} \leq \sum_{j=1}^k y_j^{\downarrow}$  holds for  $k=1, \dots, d-1$ , and the equality holds for  $k=d$ , with  $d$  being the dimension of the vector. Zeros are appended to the vectors  $\lambda_{\varrho_{A,B}}^{\downarrow}$  in (8), in order to make their dimension equal to the one of  $\lambda_{\varrho}^{\downarrow}$ .

Thus, for a separable state the ordered vector of eigenvalues for the whole density matrix is majorized by the ones of the reduced density matrices. This is was summarized by Nielsen and Kempe as “Separable states are more disordered globally than locally.”<sup>14</sup> Note that the spectra of a density matrix and its reduced density matrices do not allow us to distinguish separable and entangled states. The majorization criterion is only a necessary, not a sufficient condition for separability.

The logical ordering of the separability criteria introduced in this section is as follows:

- (i) *dimension  $2 \times 2$  and  $2 \times 3$ :*  $\varrho$  is separable  $\Leftrightarrow$  satisfies PPT  $\Leftrightarrow$  satisfies reduction criterion  $\Rightarrow$  satisfies majorization criterion, and
- (ii) *higher dimensions:*  $\varrho$  is separable  $\Rightarrow$  satisfies PPT  $\Rightarrow$  satisfies reduction criterion  $\Rightarrow$  satisfies majorization criterion

## B. Nonoperational separability criteria

In this subsection we will discuss two nonoperational separability criteria. Both are necessary and sufficient criteria for any bipartite system. They bear the major problem, however, that they do not provide us with a simple procedure to check the separability properties of a given state. This will become clear from their description.

(1) **Positive maps:** It was shown in Ref. 11 that  $\varrho$  is separable iff for any positive map  $\Lambda$

$$(\mathbb{1} \otimes \Lambda)(\varrho) \geq 0 \quad (9)$$

holds.

A positive map is a map that takes positive operators to positive operators. A positive map  $\Lambda$  is called *completely positive* (CP) if any extension to a larger Hilbert space, i.e.,  $\mathbb{1}_x \otimes \Lambda$ , is a positive map. Here  $x$  denotes the dimension of the extension and is arbitrary. It is clear from Eq. (9) that, for the purpose of finding separability criteria, only those maps are interesting which are positive, but *not* CP, as a CP map will fulfill (9) for any given  $\varrho$ .

In the previous section we have already studied two examples for positive maps that are not CP and the extensions of which provide separability criteria: the transpose and the map  $\Lambda(\sigma) = (\text{Tr} \sigma) \mathbb{1} - \sigma$ . There we have already explained the reason why (9) has to hold for separable states: it is the possibility to decompose a separable state into a sum of tensor products according to (3). Applying a positive map to one of the subsystems will keep each term positive, and therefore also their sum.

Note that the problem about the nonoperational criterion of positive maps lies in the little word “any” just before (9): we do not have a complete characterization of the set of all positive maps.

(2) **Entanglement witnesses:** The criterion of the so-called entanglement witnesses was given in Ref. 11 and studied in Ref. 15:

*A density matrix  $\varrho$  is entangled iff there exists a Hermitian operator  $\mathcal{W}$  with  $\text{Tr}(\mathcal{W}\varrho) < 0$  and  $\text{Tr}(\mathcal{W}\varrho_{\text{sep}}) \geq 0$  for any separable state  $\varrho_{\text{sep}}$ .*

We say that the witness  $\mathcal{W}$  “detects” the entanglement of  $\varrho$ .

**Correspondence between (1) and (2):** These two criteria are not independent—there is the Jamiołkowski isomorphism<sup>16</sup> that provides us with a correspondence between an entanglement witness and a positive map,

$$\mathcal{W} = (\mathbb{1} \otimes \Lambda)(P_+), \quad (10)$$

where  $P_+ = (1/M) (\sum_{i=1}^M |ii\rangle)(\sum_{j=1}^M \langle jj|)$  is the projector onto the maximally entangled state.

Let us for the rest of this subsection pursue the concept of entanglement witnesses. What seems at first sight to be a rather abstract theorem will prove to be a powerful tool, as it allows us to answer explicit questions about entanglement properties of certain states (see, e.g., Sec. V E).

The existence of entanglement witnesses is a consequence of the Hahn–Banach theorem, which states the following:

*Let  $S$  be a convex, compact set, and let  $\varrho \notin S$ . Then there exists a hyper-plane that separates  $\varrho$  from  $S$ .*

This fact is illustrated in Fig. 1: the bigger “egg”-shaped set symbolizes the convex, compact set of all density matrices. The smaller one stands for the separable density matrices, and is a convex, compact subset of the bigger one. The state  $\varrho$  is entangled and therefore  $\notin S$ . The dotted line sketches the hyper-plane that separates  $\varrho$  from  $S$ , and is given by those  $\sigma$  that fulfill  $\text{Tr}(\mathcal{W}\sigma) = 0$ .

It is helpful to realize that  $\text{Tr}(\mathcal{W}\varrho)$  defines a scalar product. Let us for a moment look at a scalar product which is familiar to everybody, the scalar product of two vectors of unit length with common origin. Let us call them  $\vec{w}$  and  $\vec{r}$ . Their scalar product  $\vec{w} \cdot \vec{r}$  is equal to  $\cos \alpha$ , where  $\alpha$  is



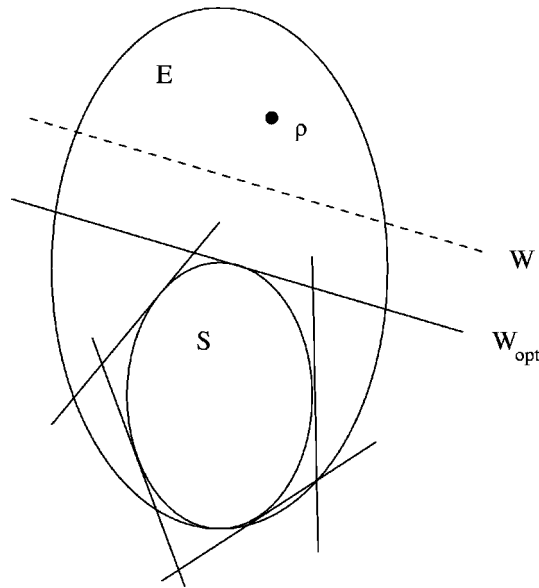


FIG. 1. Intuitive picture of entanglement witnesses and their optimization.

the relative angle between the two vectors. A fixed  $\vec{w}$  defines a certain plane—the one to which  $\vec{w}$  is orthogonal. Vectors  $\vec{r}$  from this plane have a vanishing scalar product with  $\vec{w}$ . All vectors  $\vec{r}$  that are “on one side” of this plane have a positive scalar product with  $\vec{w}$ ; all vectors  $\vec{r}$  “on the other side” a negative one—due to the properties of the cosine function. The scalar product  $\text{Tr}(\mathcal{W}\varrho)$  has the same property: all density operators on one side of the hyper-plane lead to a positive outcome; the ones on the other side to a negative one.

This intuitive picture of entanglement witnesses also helps to understand how they can be optimized:<sup>17</sup> performing a parallel transport of the hyper-plane such that it becomes tangent to the set of separable states means that the corresponding optimized witness  $\mathcal{W}_{\text{opt}}$  detects more entangled states than before. This is also indicated in Fig. 1.

In order to completely characterize the set  $S$  one would in principle need infinitely many witnesses, unless the shape of  $S$  is a polytope. This is not known nowadays. But several witnesses can already give a good approximation of the set of separable states. Methods to construct entanglement witnesses in a canonical way have been provided in Ref. 17.

#### IV. GIVEN AN ENTANGLED STATE, IS THE ENTANGLEMENT *USEFUL*?

In an ideal experiment an initially prepared maximally entangled state would remain maximally entangled. In reality, the resource of entanglement is very fragile, due to interaction with the environment. As entanglement is the foundation of many quantum information processing tasks, it would therefore be desirable to concentrate nonmaximal entanglement. A central question is: given several copies of a nonmaximally entangled state, is there a process that allows us to locally “distill” its entanglement, i.e., to retrieve a maximally entangled state?

This is a straightforward motivation to study entanglement distillation. In addition, this concept is useful in any other quantum communication task: assume that Alice wants to send a quantum message to Bob. As a simple illustration she will send a polarized photon along an optical fiber. Interaction with the environment disturbs the original quantum state—in general one has to deal with “noisy channels.” An initial pure state  $|\psi\rangle$  will arrive as some disturbed mixed state  $\varrho$ . This situation is sketched in Figs. 2(a) and 2(b).

In the wide field of error correction one deals with this problem by “repairing the state,” i.e., via encoding, finding the error and then restoring the original state. The idea of distillation of

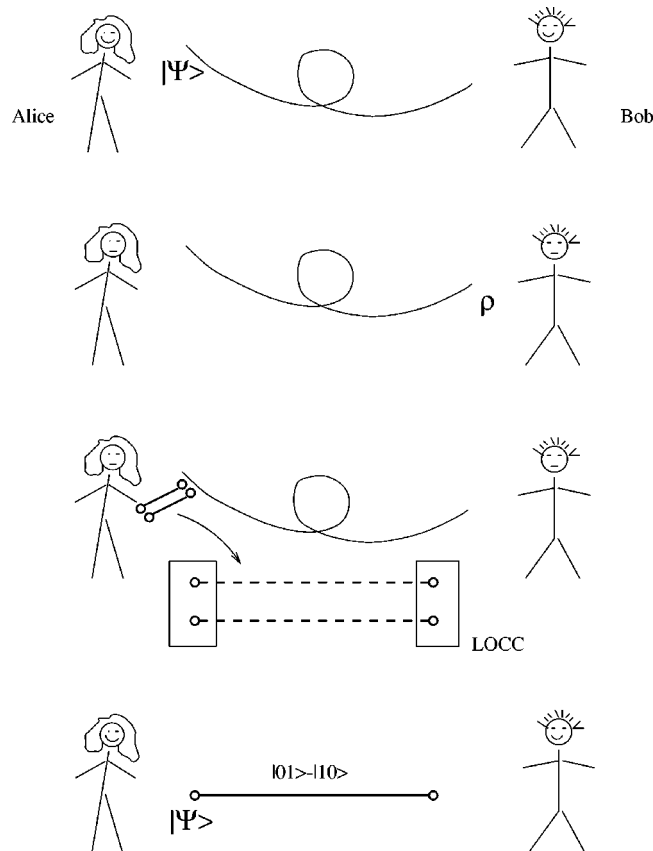


FIG. 2. Providing a noiseless channel via distillation: (a) Alice wants to send the message  $|\psi\rangle$  to Bob. (b) Bob receives  $\rho$  instead, as the channel is noisy. (c) Alice sends one subsystem of a maximally entangled state through the noisy channel to Bob, and repeats this with a second pair. They employ a distillation protocol. (d) Alice and Bob have created a maximally entangled singlet which they can use as a noiseless teleportation channel.

nonmaximally entangled states pursues a different path, by “providing a noiseless channel,” as explained below.

Figure 2(c) visualizes this concept: Alice sends one subsystem of a maximally entangled state through the noisy channel to Bob. The resulting state will not be maximally entangled and mixed, due to the noise. She repeats this with a second pair or more pairs. Alice and Bob then operate locally on their respective qubits and communicate classically (LOCC=local operations and classical communication), thus employing a distillation protocol. One explicit protocol will be introduced below. Thus they create a maximally entangled state as indicated in Fig. 2(d). This state can now be used as a noiseless channel via teleportation.

In summary, it is an essential question to ask: Given an entangled density matrix  $\rho$ , can its entanglement be *distilled*?

### A. A distillation protocol

The following distillation protocol for nonmaximally entangled mixed states was proposed in Ref. 18. It is designed for the case that Alice and Bob share a supply of many identical entangled bipartite systems of qubits. They can always convert them by local operations to the isotropic state  $\rho_{iso} = (1-p)/4 \mathbb{1} + p|\Phi^+\rangle\langle\Phi^+|$ , with  $\frac{1}{3} < p \leq 1$ .

In the first step Alice and Bob use two  $\rho$ 's, as illustrated in Fig. 3. Each of them applies a local CNOT-gate to his/her two qubits. The action of this gate is given by  $U_{\text{CNOT}}|a_1\rangle|a_2\rangle = |a_1\rangle|(a_1 + a_2) \bmod 2\rangle$ .



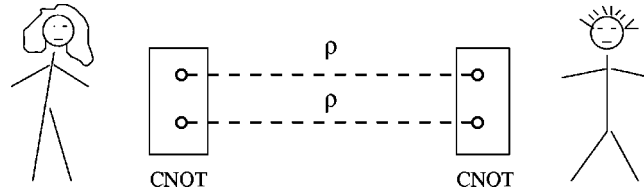


FIG. 3. First step in the distillation protocol: applying local CNOT's.

In the next step, both Alice and Bob do a measurement on their second qubit, as shown in Fig. 4. They only keep the first density matrix, which had changed to some  $\varrho'$ , if their outcomes are identical. Otherwise the two pairs have to be discarded.

This process has increased the overlap of the new density matrix  $\varrho'$  with the maximally entangled state. Thus both entanglement and purity are enhanced. The new fidelity is defined as  $F' = \langle \Phi^+ | \varrho' | \Phi^+ \rangle$  and is plotted in Fig. 5, as a function of the original fidelity  $F = \langle \Phi^+ | \varrho | \Phi^+ \rangle = (1+3p)/4$ . This procedure is then repeated with new pairs of the higher fidelity. In this way the entanglement is increased in successive steps, finally being maximal when enough original pairs were available. A distillation protocol is successful, i.e., enhances the entanglement, whenever the curve for the new fidelity lies above the line  $F' = F$  (dashed line in Fig. 5).

### B. Which states can be distilled?

For the general reasons discussed at the beginning of Sec. IV it is a very fundamental question to ask: which entangled states can be distilled? For two-qubit states the answer was given in Ref. 19: *all* entangled two-qubit states are distillable.

In general, this question is unsolved, however. A necessary and sufficient criterion for distillability of a given  $\varrho$  was proved in Ref. 20:

*The state  $\varrho$  is distillable iff there exists  $|\psi^2\rangle = a_1|e_1\rangle|f_1\rangle + a_2|e_2\rangle|f_2\rangle$  such that  $\langle \psi^2 | (\varrho^{T_A})^{\otimes n} | \psi^2 \rangle < 0$  for some  $n$ .*

In other words, if for a certain number  $n$  of copies the partial transpose of the total state has a negative expectation value with some vector of Schmidt rank 2, then  $\varrho$  can be distilled (one says:  $\varrho$  is  $n$ -distillable), and vice versa. From this theorem it follows immediately that a state with a positive partial transpose cannot be distilled: if  $\varrho^{T_A} \geq 0$ , then  $(\varrho^{T_A})^{\otimes n} \geq 0$ , and thus PPT states are undistillable. As mentioned above, entangled PPT states are therefore called “bound entangled.”

It is an open question whether the reverse of the statement “PPT states are undistillable” is also true, i.e., if a state is undistillable, does it have to be PPT? For the dimension  $2 \times N$  this is indeed true.<sup>21</sup> For higher dimensions there is a strong conjecture that it is false, i.e., that there are undistillable states with a nonpositive partial transpose (NPT). A family of such states in dimension  $N \times N$  was discussed in Refs. 21 and 22. This family consists of a convex combination of projectors onto the symmetric and the antisymmetric subspace, where the relative weight of these two contributions is the only free parameter. Depending on the value of this parameter, the state is PPT or NPT. The undistillability of NPT states with a certain constant finite range of this param-

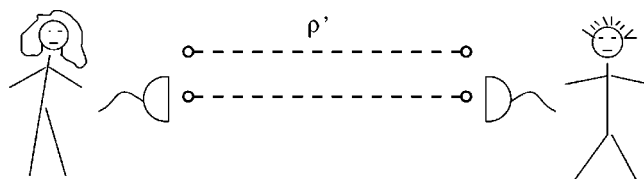


FIG. 4. Second step in the distillation protocol: measurement.

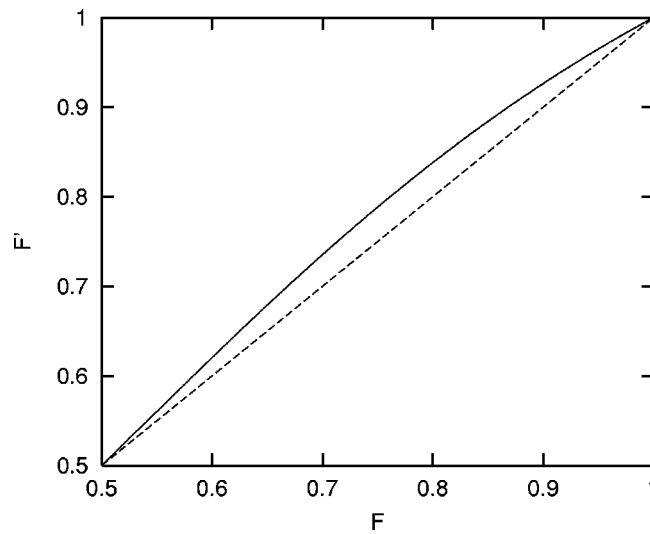


FIG. 5. New fidelity after one distillation step, as a function of the fidelity in the previous step.

eter was shown numerically for up to three copies in the case  $N=3$ . The problem in finding a rigorous answer to the question of distillability for these NPT states lies in the fact that one has to consider the limit  $n \rightarrow \infty$ .

Our present understanding of how the set of all states is decomposed into separable, entangled undistillable and distillable states is summarized in Fig. 6.

## V. GIVEN AN ENTANGLED STATE, *HOW MUCH* IS IT ENTANGLED?

So far we have shown that it is still an open question how to *qualify* a given state as separable versus entangled or undistillable versus distillable. How can we hope to answer the question of *quantifying* the amount of entanglement of a given state? It is not surprising that there is no simple answer to that. We will summarize the requirements for a good entanglement measure, and introduce the reader to some important entanglement measures, without making the attempt to discuss all existing entanglement measures.

In Sec. VC we will explain the concept of Schmidt witnesses, which us allows to classify entangled states in classes according to their Schmidt number. In the final two subsections we will study composite systems of three qubits, showing that the question of quantification of entanglement has to be reformulated in that case.

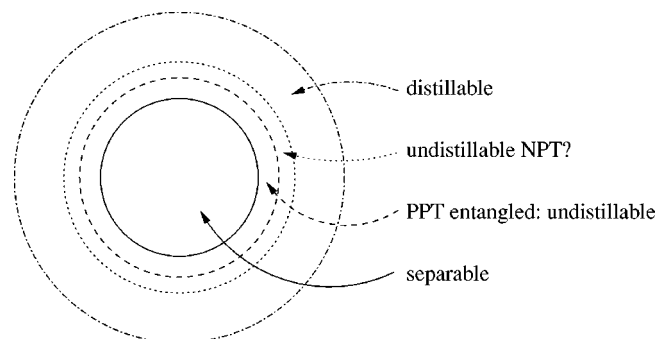


FIG. 6. Decomposition of the set of all states into distillable and undistillable states.

TABLE I. Properties of entanglement measures.

	$E_C$	$E_F$	$E_R$	$E_D$
continuity	?	✓	✓	?
additivity	✓	?	no <sup>a</sup>	✓
convexity	✓	✓	✓	no (?) <sup>b</sup>

<sup>a</sup>Reference 24.<sup>b</sup>Reference 25.

### A. Requirements for entanglement measures

A good entanglement measure  $E$  has to fulfill several requirements. However, it is still an open question whether all of these conditions are indeed necessary. In fact, some of the entanglement measures that are introduced below do not fulfill the whole list of properties (see Table I).

- (1) If  $\varrho$  is *separable*, then  $E(\varrho)=0$ .
- (2) *Normalization*: The entanglement of a maximally entangled state of two  $d$ -dimensional systems is given by

$$E(P_+^d)=\log d. \quad (11)$$

- (3) *No increase under LOCC*: Applying local operations to  $\varrho$  and classically communicating cannot increase the entanglement of  $\varrho$ , i.e.,

$$E(\Lambda_{\text{LOCC}}(\varrho))\leq E(\varrho). \quad (12)$$

- (4) *Continuity*: In the limit of vanishing distance between two density matrices the difference between their entanglement should tend to zero, i.e.,

$$E(\varrho)-E(\sigma)\rightarrow 0 \quad \text{for } \|\varrho-\sigma\|\rightarrow 0. \quad (13)$$

- (5) *Additivity*: A certain number  $n$  of identical copies of the state  $\varrho$  should contain  $n$  times the entanglement of one copy,

$$E(\varrho^{\otimes n})=nE(\varrho). \quad (14)$$

- (6) *Subadditivity*: The entanglement of the tensor product of two states  $\varrho$  and  $\sigma$  should not be larger than the sum of the entanglement of each of the states,

$$E(\varrho\otimes\sigma)\leq E(\varrho)+E(\sigma). \quad (15)$$

- (7) *Convexity*: The entanglement measure should be a convex function, i.e.,

$$E(\lambda\varrho+(1-\lambda)\sigma)\leq\lambda E(\varrho)+(1-\lambda)E(\sigma) \quad (16)$$

for  $0<\lambda<1$ .

### B. Some important entanglement measures

For a pure bipartite state  $|\psi\rangle$  a good entanglement measure is the von Neumann entropy of its reduced density matrix,  $S(\varrho_{\text{red}})=-\text{Tr}(\varrho_{\text{red}}\log\varrho_{\text{red}})$ . For mixed states there is no unique entanglement measure, but all entanglement measures should coincide on pure bipartite states and be equal to the von Neumann entropy of the reduced density matrix (uniqueness theorem).<sup>5</sup> Some important entanglement measures are defined as follows:

*Entanglement cost*: The entanglement cost tells us how expensive it is to create an entangled state  $\varrho$ , i.e., what is the ratio of the number of maximally entangled input states  $|\Phi^+\rangle$  over the produced output states  $\varrho$ , minimized over all LOCC operations. In the limit of infinitely many outputs this reads

$$E_C(\varrho) = \inf_{\{\Lambda_{\text{LOCC}}\}} \lim_{n_{\varrho} \rightarrow \infty} \frac{n_{|\Phi^+\rangle}^{\text{in}}}{n_{\varrho}^{\text{out}}}. \quad (17)$$

*Entanglement of formation:* Any state  $\varrho$  can be decomposed as a convex combination of projectors onto pure states,  $\varrho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ . The entanglement of formation is the averaged von Neumann entropy of the reduced density matrices of the pure states  $|\psi_i\rangle$ , minimized over all possible decompositions,

$$E_F(\varrho) = \inf_{\{\text{dec}\}} \sum_i p_i S(\varrho_{i,\text{red}}). \quad (18)$$

*Relative entropy of entanglement:* The relative entropy can be seen intuitively as the “distance” of the entangled  $\varrho$  to the closest separable state  $\sigma$ , although it is not a distance in the mathematical sense,

$$E_R(\varrho) = \inf_{\sigma \in S} \text{tr}[\varrho(\log \varrho - \log \sigma)]. \quad (19)$$

*Distillable entanglement:* The distillable entanglement tells us how much entanglement we can extract from an entangled state  $\varrho$ , i.e., what is the ratio of the number of maximally entangled output states  $|\Phi^+\rangle$  over the needed input states  $\varrho$ , maximized over all LOCC operations. In the limit of infinitely many inputs this reads

$$E_D(\varrho) = \sup_{\{\Lambda_{\text{LOCC}}\}} \lim_{n_{\varrho} \rightarrow \infty} \frac{n_{|\Phi^+\rangle}^{\text{out}}}{n_{\varrho}^{\text{in}}}. \quad (20)$$

There are some known relations between these entanglement measures: the distillable entanglement (entanglement cost) is a lower (upper) bound for *any* entanglement measure, i.e.,  $E_D(\varrho) \leq E(\varrho) \leq E_C(\varrho)$ . For any bound entangled state  $E_D(\varrho) < E_C(\varrho)$  holds, but there is also an example for a free entangled state, i.e., a distillable state, with the same property.<sup>23</sup> It is conjectured that the entanglement of formation and the entanglement cost are identical, i.e.,  $E_F(\varrho) \stackrel{?}{=} E_C(\varrho)$ .

Some known and unknown properties of the entanglement measures discussed above are given in Table I.<sup>24,25</sup>

### C. Schmidt witnesses

A slightly different question from “how much entangled is a state  $\varrho$ ?” can be addressed via the generalization of entanglement witnesses to so-called Schmidt witnesses. They give an answer to the question “how many degrees of freedom are entangled in  $\varrho$ ?” This corresponds to a finer classification of entangled states.

The Schmidt rank for pure bipartite states, as defined in Eq. (4), was generalized to the so-called Schmidt number  $k$  for mixed bipartite states in Ref. 26. For a given decomposition of  $\varrho$  into a convex combination of projectors onto pure states, let us call the highest occurring Schmidt rank  $r_{\text{max}}$ . The Schmidt number is the minimization of this highest Schmidt rank over all possible decompositions:

$$\varrho = \sum_i p_i |\Psi_i^{r_i}\rangle\langle\Psi_i^{r_i}|, \quad k = \min_{\{\text{dec}\}}(r_{\text{max}}). \quad (21)$$

The Schmidt number cannot be higher than  $M$ , the smaller of the dimensions of the two subsystems.

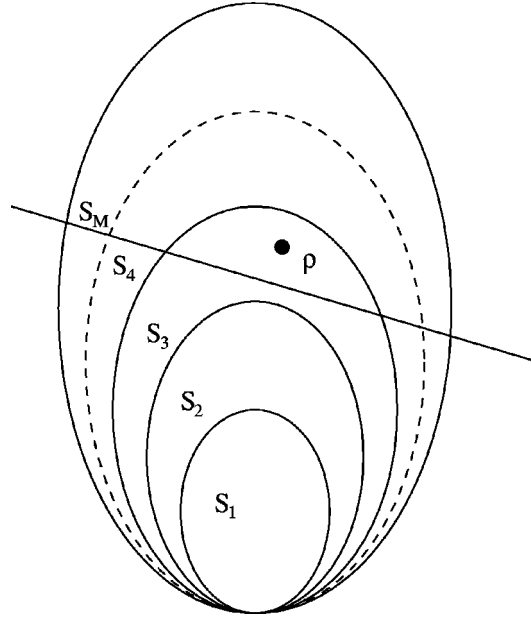


FIG. 7. Schmidt classes and the detection of the Schmidt number by a Schmidt witness.

Entangled states can now be classified according to Schmidt classes:<sup>27</sup> the Schmidt class  $S_k$  is a subset of the set of all density matrices and contains all density matrices with Schmidt number  $\leq k$ . The Schmidt classes are successively embedded into each other, as visualized in Fig. 7:  $S_1 \subset S_2 \subset \dots \subset S_M$ .

A Schmidt witness  $\mathcal{W}_k$  for the Schmidt class  $S_k$  is defined as a straightforward generalization of the entanglement witnesses discussed in Sec. III B. A Schmidt witness  $\mathcal{W}_k$  is a Hermitian operator for which

$$\exists \varrho \in S_k \quad \text{with} \quad \text{Tr}(\mathcal{W}_k \varrho) < 0, \quad (22)$$

$$\forall \varrho_{k-1} \in S_{k-1}: \quad \text{Tr}(\mathcal{W}_k \varrho_{k-1}) \geq 0. \quad (23)$$

Constructive methods for Schmidt witnesses have been shown; they can be optimized in analogy with entanglement witnesses, and they can be used for example as a tool to study properties of bound entangled states.

#### D. “Many” systems: Pure three-qubit states

So far we have only considered bipartite systems. Unfortunately, according to current knowledge one already has to call three subsystems “many.” Can we generalize the concepts that were introduced so far to tripartite states? Again, one can ask whether a given state is separable or entangled. But now one can also specify the kind of entanglement: is it genuine three-particle entanglement or are just two of the three subsystems entangled? Like for bipartite states, one can ask whether a given tripartite state can be distilled. This question will not be addressed here, however. And, finally, does it make sense to ask “how much is a given tripartite state entangled?”

For *pure* three-qubit states the situation is as follows:

A *separable* state can be written as

$$|\psi_S\rangle = |\phi_A\rangle \otimes |\phi_B\rangle \otimes |\phi_C\rangle. \quad (24)$$

A *biseparable* state is a state where only two out of the three systems are entangled, and the third system is a tensor product with the entangled ones, e.g., A–BC:

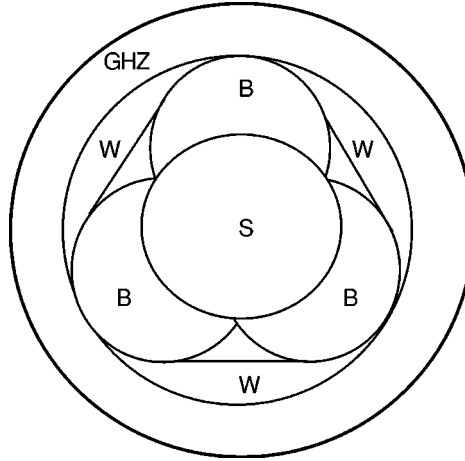


FIG. 8. Schematic structure of the set of all three-qubit states.  $S$ : separable class;  $B$ : biseparable class (convex hull of biseparable states with respect to any partition);  $W$ -class and GHZ-class.

$$|\psi_B\rangle = |\phi_A\rangle \otimes \sum_{i=1}^2 a_i |e_i\rangle |f_i\rangle. \quad (25)$$

The other two possible partitions for biseparable states are  $B-AC$  and  $C-AB$ .

A *three-qubit correlated* state is one with genuine entanglement of all three subsystems. It was shown in Ref. 28 that there exist two classes of *inequivalent* states:

$$|\psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \quad (26)$$

$$|\psi_W\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle). \quad (27)$$

Any three-qubit correlated pure state  $|\psi\rangle$  can be transformed into either  $|\psi_{\text{GHZ}}\rangle$  or  $|\psi_W\rangle$  by local reversible operations  $\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{C}$ .

Therefore, it is not enough to ask whether a given three-qubit state is separable or biseparable or three-party entangled, but also whether a genuinely three-qubit entangled state belongs to the GHZ- or  $W$ -class. For mixed states, the tool of witness operators is again useful for this purpose, as will be discussed in the following subsection.

### E. Classification of mixed three-qubit states

Let us introduce entanglement classes for mixed three-qubit states.<sup>29</sup> A mixed three-qubit state  $\varrho$  can be written as convex combination of pure states: if  $\varrho$  can be decomposed as a sum of projectors onto pure separable states  $|\psi_S\rangle\langle\psi_S|$ , then it belongs to the convex compact set  $S$ . If one needs at least one biseparable state  $|\psi_B\rangle\langle\psi_B|$  in the sum, but no genuine tripartite entanglement, then  $\varrho$  belongs to the class  $B$ , more precisely to  $B \setminus S$ . In the same way we define the  $W$ -class (needs at least one  $W$ -state in the decomposition) and the GHZ-class (at least one GHZ-state needed).

These sets are embedded into each other:  $S \subset B \subset W \subset \text{GHZ}$ . This is schematically shown in Fig. 8. It is important to note that  $W \subset \text{GHZ}$  and not the other way round: otherwise the class GHZ would not be compact, as can be seen by studying the most general form of a  $W$ -versus a GHZ-state, as given in Ref. 30.

In analogy to entanglement witnesses and Schmidt witnesses one can construct tripartite witnesses. A GHZ witness  $\mathcal{W}_{\text{GHZ}}$  is an Hermitian operator with  $\text{Tr}(\mathcal{W}_{\text{GHZ}}\varrho) < 0$  for some  $\varrho \in \text{GHZ} \setminus W$ , and  $\text{Tr}(\mathcal{W}_{\text{GHZ}}\varrho_W) \geq 0$  for all  $\varrho_W \in W$ . An example for a GHZ witness is given by

$$\mathcal{W}_{\text{GHZ}} = \frac{3}{4}\mathbb{1} - P_{\text{GHZ}}, \quad (28)$$

where  $P_{\text{GHZ}}$  is the projector onto  $|\psi_{\text{GHZ}}\rangle$ , given in (26). It is straightforward to show that this operator has the desired properties, when one realizes that the maximal squared overlap between a pure  $W$ -state and  $|\psi_{\text{GHZ}}\rangle$  is given by  $\frac{3}{4}$ .

In a similar manner one can define a  $W$ -witness. An example for such a witness that detects a  $W$ -state, but has a positive or vanishing expectation value for all states in  $B$ , is given by

$$\mathcal{W}_W = \frac{2}{3}\mathbb{1} - P_W, \quad (29)$$

where  $P_W$  is the projector onto  $|\psi_W\rangle$ , given in (27), and  $\frac{2}{3}$  is the maximal squared overlap between  $|\psi_W\rangle$  and a pure biseparable state.

Using the witness in (29), one can show that the set of mixed  $W \setminus B$ -states is not of measure zero. This is contrary to the pure case, where  $W$ -states are of measure zero.<sup>28</sup> The idea of the proof is to show that there is a finite ball around a state from the family

$$\varrho = \frac{1-p}{8}\mathbb{1} + pP_W, \quad (30)$$

for a certain given parameter  $p$ , such that the ball is contained in the  $W$ -class. This is an example where the concept of witnesses helps to answer an explicit question about the structure of the set of entangled states.

Another interesting topic in this context is the properties of bound entangled states. Using again the tool of witness operators, there is some evidence that bound entangled three-qubit states cannot be in  $\text{GHZ} \setminus W$ , i.e., they are at most in  $W$ .<sup>29</sup>

By studying mixed three-qubit states we have realized that in this case it is not enough to ask “how much a given state is entangled.” As there are different inequivalent entanglement classes, this question makes sense only within a given class  $W$  or  $\text{GHZ}$ . For more than three qubits the number of inequivalent classes grows fast.<sup>28</sup>

## VI. SUMMARY

The section headings in this tutorial were phrased as questions. Let us summarize some answers.

*What is entanglement?*

There are many possible answers, maybe as many as there are researchers in this field.

*Given a state  $\varrho$ , is it separable or entangled?*

This question is easy to answer for pure states, and for low dimensions ( $2 \times 2$  and  $2 \times 3$ ). It is very difficult to answer otherwise. Several separability criteria have been explained.

*Given an entangled  $\varrho$ , is the entanglement useful?*

We have discussed that states with a positive partial transpose are undistillable, most states with a nonpositive partial transpose (NPT) are distillable; but some NPT states are conjectured to be undistillable.

*Given an entangled  $\varrho$ , how much is it entangled?*

There are several different bipartite entanglement measures which quantify the degree of entanglement. For multipartite systems there are inequivalent entanglement classes, and therefore the above question has to be rephrased accordingly.



## ACKNOWLEDGMENTS

Many thanks to Antonio Acín, Maciej Lewenstein, Asher Peres, Martin Plenio, and Anna Sanpera for enjoyable discussions about entanglement. This work has been supported by DFG (Schwerpunkt “Quanteninformationsverarbeitung”) and the EU IST-Program EQUIP.

- <sup>1</sup>M. Lewenstein, D. Bruß, J. I. Cirac, B. Kraus, M. Kuś, J. Samsonowicz, A. Sanpera, and R. Tarrach, J. Mod. Opt. **47**, 2841 (2000).
- <sup>2</sup>M. Horodecki, P. Horodecki, and R. Horodecki, “Mixed-state entanglement and quantum communication,” in “*Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments?*” Springer Tracts in Modern Physics, Vol. 173, edited by G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. Werner, and A. Zeilinger (Springer-Verlag, New York, 2001).
- <sup>3</sup>B. Terhal, “Detecting quantum entanglement,” quant-ph/0101032.
- <sup>4</sup>M. Donald, M. Horodecki, and O. Rudolph, “The uniqueness theorem for entanglement measures,” quant-ph/0105017.
- <sup>5</sup>Quantum Inf. Comput. **1**(1) (2001).
- <sup>6</sup>E. Schrödinger, Naturwissenschaften **23**, 807 (1935).
- <sup>7</sup>A. Peres, private communication.
- <sup>8</sup>R. Werner, Phys. Rev. A **40**, 4277 (1989).
- <sup>9</sup>K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).
- <sup>10</sup>A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
- <sup>11</sup>M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
- <sup>12</sup>P. Horodecki, Phys. Lett. A **232**, 333 (1997).
- <sup>13</sup>M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
- <sup>14</sup>M. Nielsen and J. Kempe, quant-ph/0011117.
- <sup>15</sup>B. Terhal, Phys. Lett. A **271**, 319 (2000).
- <sup>16</sup>A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).
- <sup>17</sup>M. Lewenstein, B. Kraus, I. J. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000).
- <sup>18</sup>C. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
- <sup>19</sup>M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **78**, 574 (1997).
- <sup>20</sup>M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
- <sup>21</sup>W. Dür, J. I. Cirac, M. Lewenstein, and D. Bruß, Phys. Rev. A **61**, 062313 (2000).
- <sup>22</sup>D. DiVincenzo, P. Shor, J. Smolin, B. Terhal, and A. Thapliyal, Phys. Rev. A **61**, 062312 (2000).
- <sup>23</sup>G. Vidal and J. I. Cirac, quant-ph/0107051.
- <sup>24</sup>K. Vollbrecht and R. Werner, quant-ph/0010095.
- <sup>25</sup>P. Shor, J. Smolin, and B. Terhal, Phys. Rev. Lett. **86**, 2681 (2001).
- <sup>26</sup>B. Terhal and P. Horodecki, Phys. Rev. A **61**, 040301(R) (2000).
- <sup>27</sup>A. Sanpera, D. Bruß, and M. Lewenstein, Phys. Rev. A **63**, 050301(R) (2001).
- <sup>28</sup>W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).
- <sup>29</sup>A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **87**, 040401 (2001).
- <sup>30</sup>A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach, Phys. Rev. Lett. **85**, 1560 (2000).