Lower bounds for attainable fidelities in entanglement purification

G. Giedke,[1] H. J. Briegel,[1,2] J. I. Cirac,[1] and P. Zoller[1]
[1]*Institut für Theoretische Physik, Universität Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria*
[2]*Departamento de Fisica Aplicada, Universidad de Castilla–La Mancha, 13071 Ciudad Real, Spain*

(Received 22 September 1998)

We derive lower bounds for the attainable fidelity of standard entanglement purification protocols when local operations and measurements are subjected to errors. We introduce an error parameter which measures the distance between the ideal completely positive map describing a purification step and the one in the presence of errors. We derive nonlinear maps for a lower bound of the fidelity at each purification step in terms of this parameter. [S1050-2947(99)01104-X]

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

I. INTRODUCTION

Entanglement purification [1–3] is one of the most important tools in the theory of quantum information and, in particular, in quantum communication. It allows, in principle, creation of maximally entangled states of particles at different locations, even if the channel that connects those locations is noisy [4]. These entangled particles can then be used for faithful teleportation [5] or secure quantum cryptography [6,7].

The basic idea in entanglement purification is to ''distill'' a few N' pairs of particles [quantum bits (qubits), for example, the case which we will consider exclusively in the following] in highly entangled states out of $N \geqslant N'$ pairs in a mixed state with lower fidelity of the entanglement (or, in short, fidelity) using local operations and measurements. This fidelity is defined as the maximum overlap of the density operator of a pair of qubits with a maximal entangled state. If the initial pairs are in a nonseparable state [8,9], then one can obtain asymptotically (in the limit $N \rightarrow \infty$) maximally entangled states [10] provided all local operations and measurements are perfect [2,11]. In practice, there will be errors in both the local operations and measurements. The purpose of this paper is to analyze this problem for the purification protocols introduced in Refs. [1,7]. We are interested in analyzing the conditions under which one can purify in the presence of errors, as well as in the limitations of the purification protocols. In particular, we find a nonlinear map which relates a lower bound for the fidelity at two consecutive steps of the purification protocol, which allows us to derive lower bounds for the reachable fidelity. In order to analyze this problem, we introduce a parameter δ which characterizes the errors. It measures the distance between the ideal operations and measurements and the ones in the presence of errors.

Quantum communication in the presence of errors has been considered previously by Knill and Laflamme [12] in a general context, and by Van Enk *et al.* [13] for a particular experimental setup [14]. The work of Knill and Laflamme introduced ideas of fault-tolerant quantum computation [15] to show that there exists an accuracy threshold for storage of quantum information, which also applies to the case of quantum communication. As shown by Bennett *et al.* [2] one can rephrase this result in terms of entanglement purification with *one-way classical communication*. In Ref. [16], entanglement purification together with a generic error model is used to estimate the possibilities of quantum communication over long distances using quantum repeaters. The employed entanglement purification protocols explicitly utilize *two-way classical communication*, which makes them much more efficient for quantum communication. In the present paper we use purification protocols which utilize *two-way classical communication*, and therefore our error thresholds are much less demanding than those derived from the theory of Knill and Laflamme [12]. On the other hand, we are interested in a rigorous lower bound for the achievable fidelity for arbitrary errors, and not in an estimation [16]. The results and methods developed here can be generalized to derive lower bounds for other interesting problems in which local operations and measurements are imperfect, such as quantum teleportation or quantum cryptography.

This paper is organized as follows. Section II contains a summary of the main results of this paper, and is directed to the reader who is interested neither in the technical details of the definitions of our error parameter, nor in the derivations of the nonlinear maps for the lower bound of the fidelity. In Sec. III we introduce the error parameter δ and derive some properties related to the fact that it is a distance between completely positive linear maps. Finally, in Sec. IV we derive the nonlinear map for the fidelity of entanglement in terms of this distance and sketch its dynamics.

II. SUMMARY OF THE MAIN RESULTS AND DISCUSSION

In the standard scenario of entanglement purification [1], two partners at different locations share N pairs of qubits, each pair being in a state described by a density operator ρ. A purification procedure produces $N' \leqslant N$ pairs in a state ρ' ''closer'' to a maximally entangled state ψ_{me} by only using local operations, local measurements, and classical communication between the partners. More specifically, if we define the fidelity of the entanglement

$$F(\rho) = \max_{\psi_{me}} \langle \psi_{me} | \rho | \psi_{me} \rangle, \tag{1}$$

where the maximization is taken with respect to maximally entangled states ψ_{me}, then $F(\rho') > F(\rho)$. In the following we will call $F(\rho)$ simply fidelity.

It has been shown [10] that if ρ is nonseparable (it cannot be written as a convex combination of factorized density operators [8,9]) then there are purification procedures which obtain $F(\rho')=1$ in the asymptotic limit $N\rightarrow\infty$. In particular, if $F(\rho)>1/2$ one can reach this goal by using the purification procedure devised by Bennett *et al.* [1] and improved by Deutsch *et al.* [7]. It consists of a concatenation of *purification steps* involving two pairs of qubits, which give rise to a single pair with higher fidelity. In all these procedures, one assumes that the local operations and measurements are error free. In a real situation, however, there will be errors due to the coupling to the environment, imprecise apparatus, etc. Although small, they will limit the maximum attainable fidelity and will dictate whether purification is possible or not.

In this section we first briefly review the purification protocol introduced in Refs. [1,7], and define the notation that we will use later on. Then we consider the same procedure in the presence of general errors, and characterize these errors in terms of a single parameter δ, which basically expresses the departure of the purification step in the presence of errors from the ideal one. Next, we express the lowest possible fidelity (worst case) in each purification step as a function of the lowest possible fidelity in the previous step, which leads to a non-linear map. We analyze this map and discuss the conditions required for purification with imperfect means. The properties of our definitions and the technical details are presented in the following sections.

A. Error-free purification protocols

In this subsection we review the two purification procedures introduced in Refs. [1,7]. Subsequently we will refer to them as scheme I and II, respectively. We characterize them in two different ways: first, in terms of a completely positive linear map between the initial density operator and the one after the measurement; secondly, in terms of a nonlinear map relating the diagonal matrix elements of the density operator (in the Bell basis) at each step with the ones in the previous step. In the next subsection we will generalize the first characterization to the case of imperfect operations in order to introduce the parameter describing the errors, and then we will generalize the second characterization to find a lower bound for the fidelity.

Both purification protocols I and II consist of a sequence of steps in which local operations are applied to two pairs of qubits, followed by a measurement of one of the pairs which is then discarded. Depending on the outcome of the measurement, the other pair is discarded or not. In the latter case the fidelity F_1 of the remaining pair is (on average) larger than that of the original ones. This step is applied to the N pairs obtaining $N_1\leqslant N/2$ pairs of fidelity F_1. Then it is applied to the resulting N_1 pairs obtaining N_2 pairs of fidelity $F_2>F_1$. Continuing in this vein, one can reach asymptotically $F_n\rightarrow 1$ when $n\rightarrow\infty$.

Let us consider a single purification step. It starts out with two pairs 1 and 2 in the state $\rho_{12}=\rho\otimes\rho$, applies the local operations described by the superoperator $\mathcal{U}$, and then measures each of the qubits of the second pair in the basis $\{|0\rangle,|1\rangle\}$. We denote by x the outcome of the measurement: $x=0$ if the qubits are found in the state $|0\rangle_2\equiv|00\rangle_2$; $x=1$ if they are in $|1\rangle_2\equiv|11\rangle_2$; $x=2$ if they are in $|2\rangle_2\equiv|01\rangle_2$; and $x=3$ if they are in $|3\rangle\equiv|10\rangle_2$ (the subscript 2 denotes the second pair). We denote by $\mathcal{P}_x$ $(x=0,\ldots,3)$ the map defined as follows:

$$\mathcal{P}_x(\rho_{12})\equiv {}_2\langle x|\mathcal{U}(\rho_{12})|x\rangle_2. \tag{2}$$

This map is linear and completely positive. The probability of obtaining the outcome x is $p_x(\rho_{12})=\mathrm{tr}[\mathcal{P}_x(\rho_{12})]$. If the outcome is $x=2,3$, then the first pair is discarded and otherwise it is kept. In the latter case, the state of the first pair will be

$$\rho_1'=\frac{\mathcal{P}_0(\rho_{12})+\mathcal{P}_1(\rho_{12})}{p_0(\rho_{12})+p_1(\rho_{12})}. \tag{3}$$

Thus, each (successful) step of the purification protocol is completely characterized by the maps $\mathcal{P}_{0,1}$. (Note that $\mathcal{P}_x$ stand for different maps depending on whether we are discussing scheme I or scheme II.)

On the other hand, if one is only interested in the fidelity at each step, one can use a simpler characterization of each purification step in terms of four real numbers. In the purification protocols I and II, the local operations characterized by $\mathcal{U}$ consist of a bilateral controlled-NOT gate and specific single qubit rotations. In that case, the diagonal elements of the density operator ρ' in the Bell basis only depend on the diagonal elements of the density operator ρ, and therefore each purification step can be characterized by a nonlinear map between these four diagonal matrix elements. We denote by $A_n^i=\langle\phi^i|\rho_n|\phi^i\rangle$, where ρ_n is the density operator of each pair after the nth purification step and $|\phi^i\rangle$ are the elements of the Bell basis $(i=0,1,2,3)$,

$$|\phi^{0,3}\rangle=\frac{1}{\sqrt{2}}(|00\rangle\pm|11\rangle),$$

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

In particular, $A_n^0=F_n$, the entanglement fidelity at each step. For scheme II there is, according to Ref. [7], a simple nonlinear map that relates $\mathbf{A}_{n+1}$ to $\mathbf{A}_n$, namely

$$A_{n+1}^i=\frac{\langle\phi^i|\mathcal{P}_0(\rho_n\otimes\rho_n)+\mathcal{P}_1(\rho_n\otimes\rho_n)|\phi^i\rangle}{\mathrm{tr}[\mathcal{P}_0(\rho_n\otimes\rho_n)+\mathcal{P}_1(\rho_n\otimes\rho_n)]}=:\frac{f^i(\mathbf{A}_n)}{g(\mathbf{A}_n)}, \tag{4}$$

where

$$f^0(\mathbf{A}_n)=(A_n^0)^2+(A_n^1)^2, \tag{5a}$$

$$f^1(\mathbf{A}_n)=2A_n^2A_n^3, \tag{5b}$$

$$f^2(\mathbf{A}_n)=(A_n^2)^2+(A_n^3)^2, \tag{5c}$$

$$f^3(\mathbf{A}_n)=2A_n^0A_n^1, \tag{5d}$$

$$g(\mathbf{A}_n)=(A_n^0+A_n^1)^2+(A_n^2+A_n^3)^2. \tag{5e}$$

The map (4) has a fixed point at $\mathbf{A}=(1,0,0,0)$, which is reached if the initial state has $A_0^0=F>1/2$ [17]. This fact expresses that in the absence of errors, one can use this purification protocol to purify states with $F>1/2$ and reach a fidelity as close to one as we please.

Scheme I [1] is governed by a similar map. The main difference is that at the end of each step the resulting state is brought into Werner form, that is, the three diagonal elements A^1,A^2,A^3 are made equal to $(1-A^0)/3$. Therefore one can concentrate on the first diagonal element, the fidelity A^0, only. The fidelity after the nth purification step is then given by

$$A^0_{n+1}=\frac{f^0(A^0_n,(1-A^0_n)/3)}{g(A^0_n,(1-A^0_n/3)}. \tag{6}$$

Like Eq. (4), this map has an attractive fixed point at $A^0=1$, and all $A^0_0>1/2$ are attracted to it.

B. Characterization of errors

In practice, while performing the purification protocols, errors will occur, both in the local operation and in the measurements. The imperfections in the local operations can be accounted for by substituting the action of the superoperator $\mathcal{U}$ in Eq. (2) by the action of some other completely positive, trace-preserving linear map. The errors in the measurements will be related to the following fact: in practice, the outcomes $x=0,1$ will be ultimately attributed to the presence/absence of clicks in some kind of detectors. Due to imperfections, the projection operators (or, more generally, POVMs) corresponding to those clicks are not exactly the same as the ideal ones [see Eq. (2)]. Consequently, the probabilities of the outcomes $x=0,1$ as well as the state remaining after the measurement will differ from the ideal ones. In general, we can describe both these erroneous operations and measurements in terms of a single completely positive linear map $\tilde{\mathcal{P}}_x$ which does not necessarily preserve the trace (we will use tildes in the case in which there are errors). That is, if the two pairs are initially in the state $\rho_{12}=\rho\otimes\rho$, a purification step yields the outcome x with a probability $\tilde{p}_x(\rho_{12})=\mathrm{tr}[\tilde{\mathcal{P}}_x(\rho_{12})]$. The state of the pair after the measurement is

$$\tilde{\rho}_1'=\frac{\tilde{\mathcal{P}}_0(\rho_{12})+\tilde{\mathcal{P}}_1(\rho_{12})}{\tilde{p}_0(\rho_{12})+\tilde{p}_1(\rho_{12})}. \tag{7}$$

Thus, as before, the maps $\tilde{\mathcal{P}}_{0,1}$ completely characterize each purification step.

We characterize the errors by a single parameter as follows:

$$\delta := \max_{x=0,1} d(\mathcal{P}_x,\tilde{\mathcal{P}}_x), \tag{8}$$

where $d(\mathcal{P},\tilde{\mathcal{P}})$ denotes a distance between $\mathcal{P}$ and $\tilde{\mathcal{P}}$. The explicit form of this distance is given in Eq. (13) below. We emphasize that for a given set-up, one can (in principle) perform local measurements to completely characterize $\tilde{\mathcal{P}}_x$, and therefore obtain the value of δ experimentally [18,19]. The error parameter δ has a clear physical meaning since it measures the distance between the ideal process and the erroneous one. We would like to remark here that due to the fact that there are measurements and postselection involved in the process, we have to work with maps $\mathcal{P}_x$ that do not preserve the trace. In Sec. III we discuss why it is advantageous to use those maps instead of trace-preserving maps.

Some remarks concerning the adopted description of errors are in order: We envision $\mathcal{P}$ as the reduced dynamics of the two entangled pairs coupled to some environment [20]. In taking the imperfect system dynamics to be completely positive we do (as discussed in [20]) essentially assume that there is *no initial entanglement* between the system and any environment to which it might be coupled during gate operations. There may be, however, initial entanglement of the system with another environment that is not affected by the gate operations. As in the error-free purification schemes [1,7] we also assume the two pairs that participate in a purification step to be disentangled from each other.

C. Purification with imperfect means

Once we have defined a parameter that characterizes the errors at each purification step, we can analyze the whole purification procedure [1,7] in the nonideal case. In order to do that, we define $\tilde{A}^i_n=\langle\phi^i|\tilde{\rho}_n|\phi^i\rangle$ where $\tilde{\rho}_n$ is the density operator after the nth purification step. We are particularly interested in the fidelity at each step $\tilde{A}^0_n=\tilde{F}_n$. In Sec. IV we show that for suitable initial conditions $\mathbf{A}_0$ and error parameter δ,

$$\tilde{A}^0_n\geqslant a_n,\ \ \tilde{A}^1_n\leqslant b_n\ \ (n=1,2,\dots), \tag{9}$$

where

$$a_{n+1}=\frac{a_n^2+b_n^2-2\delta}{(a_n+b_n)^2+(1-a_n-b_n)^2+2\delta}, \tag{10a}$$

$$b_{n+1}=\frac{(1-a_n)^2/2+2\delta}{a_n^2+(1-a_n)^2-2\delta}, \tag{10b}$$

and $a_0=\tilde{A}^0_0$, $b_0=\tilde{A}^1_0$. For scheme I only the fidelity A^0_n and therefore the bound (10a) with b_n replaced by $(1-a_n)/3$ is relevant.

Equations (10) define a nonlinear map that can be iterated to yield a lower bound for the attainable fidelity $\tilde{F}_\infty\geqslant a_\infty$ which depends on the value of δ. In the following we will analyze the map (10).

Let us first concentrate on the fixed points (a_f,b_f) of this map, and consider in particular scheme II. In Fig. 1 (solid line) we have plotted a_f as a function of the error parameter δ. For small values of $\delta\lesssim 0.01$ there are three fixed points. The ones with largest and the smallest value of a_f are attractive, whereas the intermediate one is a saddle point attractive in one direction and repulsive in the others. For larger values of δ, only the smallest one survives. This means that for the appropriate initial values of a_0 and b_0 if $\delta\lesssim 0.01$ one increases the fidelity using the purification protocol II to a value larger than the one given by the right wing of the appropriate curve of Fig. 1. For example, for $\delta\simeq 0.005$ one can obtain a fidelity $F\gtrsim 0.95$.

Now, let us analyze for which initial conditions (a_0,b_0) the map converges to the fixed point with the largest a_f, i.e., for which the protocol achieves purification. In Fig. 2 we

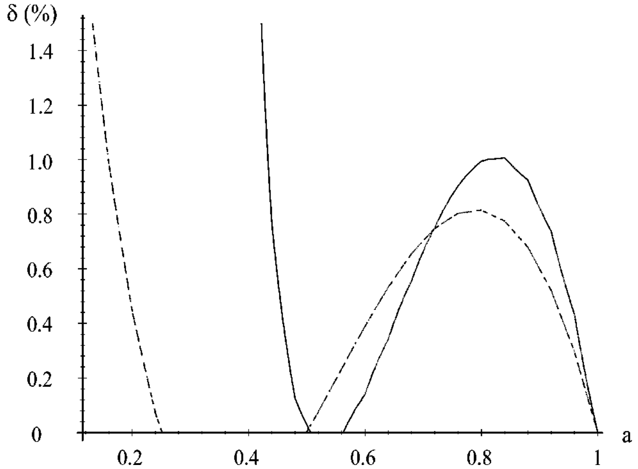

FIG. 1. The fixed points of the nonlinear map: the intersections of a horizontal line at δ with the plotted curve give the a coordinates of the fixed points for scheme I (broken) and scheme II (solid).

have plotted in the (a,b) parameter space the curve (separatrix) between the stable regions for several values of δ $(\delta_k = 0.002k,\ k=0,1,\ldots,5)$. For any initial value (a_0,b_0) lying to the right of each curve, the map will converge to the corresponding fixed point (asterisks in the plot). For $\delta = 0.006$ $(k=3$ in the plot), for example, one can purify from values of $a_0 \gtrsim 0.69$ up to values of $F \geq a_f \approx 0.94$; for $\delta = 0.002$, one can reach $F \gtrsim 0.98$ starting from $a_0 \lesssim 0.61$. The results show that the error threshold for purification is much less restrictive than the one for quantum computation [12].

III. DISTANCE BETWEEN TWO POSITIVE MAPS

We denote by H a finite dimensional complex Hilbert space and by $L(H)$ the complex Banach space of linear operators $A{:}H \to H$ with the trace norm $||A|| = \mathrm{tr}(|A^{\dagger}A|^{1/2}) \equiv \mathrm{tr}(|A|)$ (as usual, $|A| \equiv |A^{\dagger}A|^{1/2}$). We denote by $C(H) \subset L(H)$ the convex set of positive linear operators ρ acting on H with $||\rho|| \leq 1$, and by $P(H,H')$ the set of completely positive linear maps $\mathcal{P}{:}C(H) \to C(H')$ fulfilling

$$||\mathcal{P}(\rho)|| \leq ||\rho||. \tag{11}$$

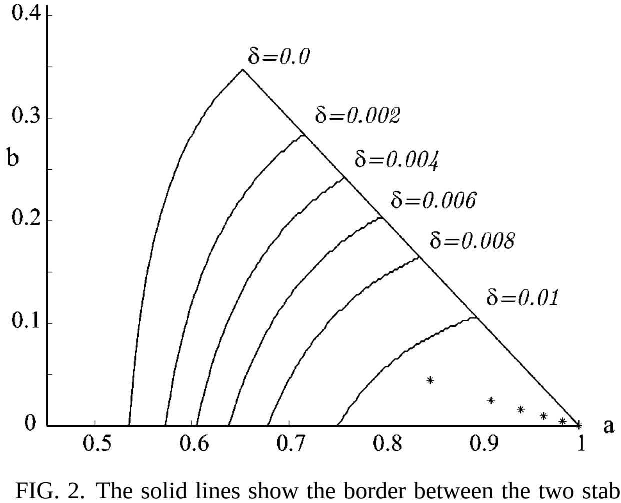

FIG. 2. The solid lines show the border between the two stable sets (the separatrix) for six values of δ. The asterisks show the corresponding (δ increasing from right to left) upper fixed points.

For positive operators, the trace norm simply coincides with the trace, and therefore Eq. (11) is equivalent to

$$\mathrm{tr}[\mathcal{P}(\rho)] \leq \mathrm{tr}(\rho) \leq 1. \tag{12}$$

Given two completely positive maps $\mathcal{P}, \tilde{\mathcal{P}} \in P(H,H')$, we define their distance

$$d(\mathcal{P},\tilde{\mathcal{P}}) = \max_{\rho\ \in C(H)} ||\mathcal{P}(\rho) - \tilde{\mathcal{P}}(\rho)||. \tag{13}$$

It is straightforward to show that d is indeed a distance by using the fact that the trace norm is a norm.

With this definition, we can characterize the errors by using the parameter δ as defined in Eq. (8). The motivation for this definition with respect to other possible definitions is that it easily gives lower bounds even for physical processes where there are measurements and postselection (as it is in the case of entanglement purification, cf. next section), i.e., when the map describing the physical process is not trace preserving. On the other hand (although we will not use this property here), it allows one to easily bound the distance between processes which are composed of several individual processes in terms of the distances between the individual processes themselves (see next subsection).

One can define other distances between trace preserving maps: for example, one can consider the map $\tilde{\mathcal{P}}'$ that transforms $\rho_{12} \to \rho_1'$, where ρ_1' is given in Eq. (7) in terms of the linear maps $\tilde{\mathcal{P}}_{0,1}$. This new map, although trace preserving, is nonlinear. If one defines distances between $\tilde{\mathcal{P}}'$ and the corresponding (trace-preserving) ideal map $\mathcal{P}'$, problems related to the nonlinearity arise: for example, it can happen that while the distance δ between the linear maps $\mathcal{P},\tilde{\mathcal{P}}$ is very small, the similarly defined distance between the nonlinear maps $\mathcal{P}',\tilde{\mathcal{P}}'$ is of the order of 1, which makes the definition useless to derive bounds. The reason is that low probability processes get ‘‘magnified’’ by the normalization and then dominate the maximization used to define the distance.

One can still define other error parameters to find sharper bounds to the fidelity in entanglement purification. However, by increasing the number of parameters one does not gain too much and the bounds become more complicated to analyze. On the other hand, $d(\mathcal{P}\otimes 1,\tilde{\mathcal{P}}\otimes 1) \neq d(\mathcal{P},\tilde{\mathcal{P}})$ [19], which would allow us to use d in processes for which the system in which we perform operations and measurements is entangled with another system, without having to include the other system in the error analysis. This may be useful, for example, in quantum computation where operations are performed on single qubits that are entangled with many other qubits. In that case, one can define other distances, as it is done in Ref. [19]. In any case, in quantum communication if we can bound the fidelity when the system is not entangled, we can automatically derive a bound for the entanglement fidelity [12,4].

A. Properties of d

In this subsection we derive some properties of the distance d introduced above. Given $\mathcal{P},\tilde{\mathcal{P}}\in P(H,H')$ we have the following.

(1) We can restrict the maximization in Eq. (13) to one dimensional projectors, i.e.,

$$d(\mathcal{P},\tilde{\mathcal{P}})=\max_{\psi\in H,\||\psi\rangle\|=1}\|\mathcal{P}(|\psi\rangle\langle\psi|)-\tilde{\mathcal{P}}(|\psi\rangle\langle\psi|)\|. \tag{14}$$

Proof: We just have to prove that the distance as given in Eq. (14) is always larger than or equal to the one given in Eq. (13), since the converse is clearly true. For any $\rho\in C(H)$ we write $\rho=\Sigma P_i|\phi_i\rangle\langle\phi_i|$ with $\Sigma_i P_i\leqslant 1$ and ψ_i normalized states of H. Using the linearity of $\mathcal{P}$ and $\tilde{\mathcal{P}}$ and that $\|\Sigma_i P_i A_i\|\leqslant\max_i\|A_i\|$, we find that $\|\mathcal{P}(\rho)-\tilde{\mathcal{P}}(\rho)\|\leqslant\max_i\|\mathcal{P}(|\phi_i\rangle\langle\phi_i|)-\tilde{\mathcal{P}}(|\phi_i\rangle\langle\phi_i|)\|$. Taking the maximum with respect to ρ in this inequality completes the proof.

(2) For all $\rho\in C(H)$ and $\phi\in H$ (normalized state) we have

$$\langle\phi|\mathcal{P}(\rho)|\phi\rangle-d(\mathcal{P},\tilde{\mathcal{P}})\leqslant\langle\phi|\tilde{\mathcal{P}}(\rho)|\phi\rangle\leqslant\langle\phi|\mathcal{P}(\rho)|\phi\rangle+d(\mathcal{P},\tilde{\mathcal{P}}), \tag{15a}$$

$$\mathrm{tr}[\mathcal{P}(\rho)]-d(\mathcal{P},\tilde{\mathcal{P}})\leqslant\mathrm{tr}[\tilde{\mathcal{P}}(\rho)]\leqslant\mathrm{tr}[\mathcal{P}(\rho)]+d(\mathcal{P},\tilde{\mathcal{P}}). \tag{15b}$$

Proof: For Eq. (15a) we use

$$|\langle\phi|\mathcal{P}(\rho)-\tilde{\mathcal{P}}(\rho)|\phi\rangle|\leqslant\|\mathcal{P}(\rho)-\tilde{\mathcal{P}}(\rho)\|\leqslant d(\mathcal{P},\tilde{\mathcal{P}}), \tag{16}$$

whereas for Eq. (15b) we use

$$|\mathrm{tr}[\mathcal{P}(\rho)-\tilde{\mathcal{P}}(\rho)]|\leqslant\mathrm{tr}[|\mathcal{P}(\rho)-\tilde{\mathcal{P}}(\rho)|]=d(\mathcal{P},\tilde{\mathcal{P}}). \tag{17}$$

Next, we give a property that allows one to bound the distance when one applies sequential maps. This may be useful when one has a concatenation of processes.

(3) Given $\mathcal{P}\in P(H',H'')$ and $\mathcal{Q}\in P(H,H')$, we define $\mathcal{P}\circ\mathcal{Q}\in P(H,H'')$ according to $(\mathcal{P}\circ\mathcal{Q})(\rho)=\mathcal{P}[\mathcal{Q}(\rho)]$. Then, we have

$$d(\mathcal{P}\circ\mathcal{Q},\tilde{\mathcal{P}}\circ\tilde{\mathcal{Q}})\leqslant d(\mathcal{P},\tilde{\mathcal{P}})+d(\mathcal{Q},\tilde{\mathcal{Q}}). \tag{18}$$

Proof: Using the properties of a distance, we have

$$d(\mathcal{P}\circ\mathcal{Q},\tilde{\mathcal{P}}\circ\tilde{\mathcal{Q}})\leqslant d(\mathcal{P}\circ\mathcal{Q},\mathcal{P}\circ\tilde{\mathcal{Q}})+d(\mathcal{P}\circ\tilde{\mathcal{Q}},\tilde{\mathcal{P}}\circ\tilde{\mathcal{Q}}). \tag{19}$$

On the one hand, we have

$$\begin{aligned}d(\mathcal{P}\circ\tilde{\mathcal{Q}},\tilde{\mathcal{P}}\circ\tilde{\mathcal{Q}})&=\max_{\rho\in C(H)}\|\mathcal{P}[\tilde{\mathcal{Q}}(\rho)]-\tilde{\mathcal{P}}[\tilde{\mathcal{Q}}(\rho)]\|\\&\leqslant\max_{\rho'\in C(H')}\|\mathcal{P}(\rho')]-\tilde{\mathcal{P}}(\rho')\|=d(\mathcal{P},\tilde{\mathcal{P}}),\end{aligned} \tag{20}$$

where we have used Eq. (11) for $\tilde{\mathcal{Q}}$. On the other hand,

$$\begin{aligned}d(\mathcal{P}\circ\mathcal{Q},\mathcal{P}\circ\tilde{\mathcal{Q}})&=\max_{\rho\in C(H)}\|\mathcal{P}[\mathcal{Q}(\rho)]-\mathcal{P}[\tilde{\mathcal{Q}}(\rho)]\|\\&=\max_{\rho\in C(H)}\|\mathcal{P}[\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)]\|.\end{aligned} \tag{21}$$

Now, since $\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)$ is self-adjoint, we can substitute in this last equation its spectral decomposition

$$\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)=\sum_{\phi}|\phi\rangle\langle\phi|\langle\phi|\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)|\phi\rangle \tag{22}$$

obtaining

$$\begin{aligned}d(\mathcal{P}\circ\mathcal{Q},\mathcal{P}\circ\tilde{\mathcal{Q}})&=\max_{\rho\in C(H)}\sum_{\phi}|\langle\phi|\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)|\phi\rangle|\\&\quad\times\|\mathcal{P}(|\phi\rangle\langle\phi|)\|\\&\leqslant\max_{\rho\in C(H)}\sum_{\phi}|\langle\phi|\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)|\phi\rangle|\end{aligned} \tag{23}$$

$$=\max_{\rho\in C(H)}\|\mathcal{Q}(\rho)-\tilde{\mathcal{Q}}(\rho)\|=d(\mathcal{Q},\tilde{\mathcal{Q}}), \tag{24}$$

which completes the proof.

(4) Finally, we show that the distance d stems from a norm, which may be useful to derive some other properties. First, let us enlarge the set $C(H)$ so that it becomes a Banach space. The simplest way is to define $S(H)=\mathrm{lin}_R\{C(H)\}$, that is, the set of operators that can be written as a (finite) linear combination of positive operators with real coefficients. The real Banach space $S(H)\subset L(H)$ is simply the space of self-adjoint operators acting on H. In the same way, we can enlarge the set $P(H,H')$. First, given a map $\mathcal{P}\in P(H,H')$ we define $\hat{\mathcal{P}}:S(H)\rightarrow S(H)$ by using the linearity of $\mathcal{P}$ [that is, if $S(H)\ni A=\Sigma_i\lambda_i\rho_i$ with $\rho_i\in C(H)$, we define $\mathcal{P}(A)=\Sigma_i\lambda_i\mathcal{P}(\rho_i)$]. Then, we define $Q(H,H')=\mathrm{lin}_R\{P(H,H')\}$, which is a real vector space. Using the operator norm

$$\|\mathcal{P}\|_{\mathrm{op}}=\max_{A\in S(H)\|A\|\leqslant 1}\|\mathcal{P}(A)\|, \tag{25}$$

it becomes a real Banach space. With this definition we have

$$d(\mathcal{P},\tilde{\mathcal{P}})=\|\mathcal{P}-\tilde{\mathcal{P}}\|_{\mathrm{op}}. \tag{26}$$

Proof: We show that the distance given in Eq. (26) is smaller than or equal to the one defined in Eq. (13), since the converse is obviously true since $C(H)\subset S(H)$. For any $A\in S(H)$ with $\|A\|\leqslant 1$ we can write $A=\Sigma_i\lambda_i|\phi\rangle\langle\phi|$, where $\Sigma_i|\lambda_i|\leqslant 1$. Now, arguing as in the proof of the property (1), we obtain that $\|\mathcal{P}(A)-\tilde{\mathcal{P}}(A)\|\leqslant\max_\phi\|\mathcal{P}(|\phi\rangle\langle\phi|)-\tilde{\mathcal{P}}(|\phi\rangle\langle\phi|)\|$. Taking the maximum over all possible $A\in S(H)$ we complete the proof.

The distance d is not unrelated to other quantities used in the literature to characterize erroneous operations. Typically, given one of the other quantities, one can bound d (and vice versa within the respective domains of applicability). Specifically this is true for the minimum fidelity, the error amplitude [12], and the generic error model [16]. The diamond norm introduced in [19] is a generalization of the distance used here and particularly useful to discuss operations on systems that are strongly entangled with other systems.

IV. NONLINEAR MAP FOR ENTANGLEMENT PURIFICATION

In this section we derive the nonlinear map (10) for the bounds of the diagonal matrix elements in the Bell basis of the density operator after each step of the purification process. As above, let $\tilde{A}_n^i = \langle \phi^i | \tilde{\rho}_n | \phi^i \rangle$, $i=0\ldots 3$. Analogous to Eq. (4), we have

$$\tilde{A}_{n+1}^i = \frac{\langle \phi^i | \tilde{\mathcal{P}}_0(\tilde{\rho}_n \otimes \tilde{\rho}_n) + \tilde{\mathcal{P}}_1(\tilde{\rho}_n \otimes \tilde{\rho}_n) | \phi^i \rangle}{\mathrm{tr}[\tilde{\mathcal{P}}_0(\tilde{\rho}_n \otimes \tilde{\rho}_n) + \tilde{\mathcal{P}}_1(\tilde{\rho}_n \otimes \tilde{\rho}_n)]}. \tag{27}$$

Using Eq. (10) we have that

$$\frac{f^i(\tilde{\mathbf{A}}_n) - 2\delta}{g(\tilde{\mathbf{A}}_n) + 2\delta} \leqslant \tilde{A}_{n+1}^i \leqslant \frac{f^i(\tilde{\mathbf{A}}_n) + 2\delta}{g(\tilde{\mathbf{A}}_n) - 2\delta}, \tag{28}$$

where f^i and g are defined in Eq. (5). In the following subsections we will discuss the two purification schemes separately in detail.

A. Scheme I

As stated above for scheme I we can use Eq. (6) instead of f^0 and forget about the other three diagonal elements. This gives

$$\tilde{A}_{n+1}^0 \geqslant \frac{(\tilde{A}_n^0)^2 + [(1-\tilde{A}_n^0)/3]^2 - 2\delta}{[\tilde{A}_n^0 + (1-\tilde{A}_n^0/3]^2 + [1-\tilde{A}_n^0 + (1-\tilde{A}_n^0)/3]^2 + 2\delta}. \tag{29}$$

Now we observe that the right hand side of Eq. (29) is monotonically increasing with $\tilde{A}_n^0$ for all $\tilde{A}_n^0 \geqslant 1/8$. Therefore replacing $\tilde{A}_n^0$ by $\frac{1}{8} \leqslant a_n \leqslant \tilde{A}_n^0$ in Eq. (29) yields a lower bound for $\tilde{A}_{n+1}^0$. Since the interval [1/8,1] is mapped into itself by the left hand side of Eq. (29) we arrive at the dynamical system defined by $a_0 = A_0^0$ and

$$a_{n+1} = \frac{a_n^2 + [(1-a_n)/3]^2 - 2\delta}{[a_n + (1-a_n)/3]^2 + [1-a_n-(1-a_n)/3]^2 + 2\delta}. \tag{30}$$

For every n the value of a_n is a lower bound of the fidelity after n purification steps.

In the case $\delta=0$ the original map of Bennett *et al.* is recovered. The three fixed points of that map at $a_l(\delta) \approx 0.25$, $a_i(\delta) \approx 0.5$, and $a_u(\delta) \approx 1$ survive even for nonzero δ and are given by the roots of the cubic polynomial

$$x^3 - \frac{7}{4}x^2 + \left[\frac{7}{8} + \frac{9}{4}\delta\right]x - \left[\frac{1}{8} - \frac{9}{4}\delta\right].$$

They are plotted as a function of δ in Fig. 1 (broken line). For $\delta \geqslant 0.008$ only the lower fixpoint survives.

The upper and lower fixpoints are attractive, while the intermediate is repulsive. Consequently even an imperfectly implemented scheme I allows us to purify ensembles with initial fidelity $F_{\mathrm{in}} > a_i(\delta)$ up to a fidelity $F_{\mathrm{out}} \geqslant a_u(\delta)$, provided that $\delta \leqslant 0.008$.

B. Scheme II

Scheme II converges faster than scheme I and can tolerate somewhat larger errors, but the analysis becomes significantly more complicated, since all four diagonal elements of the density matrix come into play. Using Eq. (28) we have

$$\tilde{A}_{n+1}^0 \geqslant \frac{(\tilde{A}_n^0)^2 + (\tilde{A}_n^1)^2 - 2\delta}{(\tilde{A}_n^0 + \tilde{A}_n^1)^2 + (\tilde{A}_n^2 + \tilde{A}_n^3)^2 + 2\delta}, \tag{31a}$$

$$\tilde{A}_{n+1}^1 \leqslant \frac{2\tilde{A}_n^2 A_n^3 + 2\delta}{(\tilde{A}_n^0 + \tilde{A}_n^1)^2 + (\tilde{A}_n^2 + \tilde{A}_n^3)^2 - 2\delta}. \tag{31b}$$

To proceed the same way as in the preceding subsection we need again a monotonicity property of the right hand sides of Eqs. (31) so that we can replace the values $\tilde{A}_n^i$ (which are typically not known, since their exact value depends on the unkown errors in $\tilde{\mathcal{P}}$) by lower or upper bounds, respectively.

Using $\Sigma_i \tilde{A}_n^i = 1$ we can express the right hand side of Eq. (31a) in terms of $\tilde{A}_n^0, \tilde{A}_n^1$ only. It is straightforward to check that the resulting expression is monotonically increasing in $\tilde{A}_n^0$ and monotonically decreasing in $\tilde{A}_n^1$ for all $(\tilde{A}_n^0, \tilde{A}_n^1)$ fulfilling

$$\tilde{A}_n^0 \geqslant \frac{1}{2} + \frac{3\delta}{1-2\delta} \quad \text{and} \quad \tilde{A}_n^1 \leqslant 0.5. \tag{32}$$

Thus, provided that $\tilde{A}_n^0 \geqslant a_n$, $\tilde{A}_n^1 \leqslant b_n$, and (a_n, b_n) fulfill the condition (32), then a_{n+1} as given in Eq. (10a) is a lower bound for $\tilde{A}_{n+1}^0$.

It remains to justify Eq. (10b). Starting from Eq. (31b) we can this time express the right hand side only in terms of $\alpha_n = \tilde{A}_n^2 + \tilde{A}_n^3$ and $\beta_n = \tilde{A}_n^2 - \tilde{A}_n^3$ using the normalization condition

$$\tilde{A}_{n+1}^1 \leqslant \frac{\frac{1}{2}(\alpha_n^2 - \beta_n^2) + 2\delta}{\alpha_n^2 + (1-\alpha_n)^2 - 2\delta}.$$

Now it is easy to check that the right hand side of this inequality is monotonically increasing in α_n (for fixed β_n) and takes (for fixed α_n) its maximum at $\beta_n = 0$, where we use the fact that $\alpha_n \leqslant 1 - \tilde{A}_n^0$ and $\tilde{A}_n^0 \geqslant 0.5$. Since $\alpha_n = \tilde{A}_n^2 + \tilde{A}_n^3 \leqslant 1 - \tilde{A}_n^0 \leqslant 1 - a_n$ we arrive at Eq. (10b) by replacing $\beta_n \to 0$ and $\alpha_n \to 1 - a_n$.

The discrete dynamical system defined by the map (10) has for $0 \leqslant \delta \lesssim 0.01$ three fixpoints with a coordinate around $a_l \approx 0.5$, $a_i \approx 0.6$, $a_u \approx 1$. Figure 1 (solid line) shows them as a function of δ. For $\delta > 0.01$ only the lower fixpoint survives. The exact a values are given by the real roots of a polynomial of seventh degree or equivalently by the intersections of the curves $b_{n+1}(a)$ and

$$b_{\text{fix}}(a) = -a + \sqrt{a - \left(1 + \frac{3}{2a-1}\right)\delta}, \tag{33}$$

the latter of which is defined by $a_{n+1}\ (a_n, b_{\text{fix}}(a_n)) = a_n$. The corresponding b coordinates are $b_{n+1}(a_x)$, where $x = l,i,u$.

As in the previous case the upper and lower fixpoints are attractive, while the intermediate one is now a saddle point, attractive in one direction and repulsive in the others. Now essentially the same argument as in the preceding subsection applies: points between intermediate and upper fixed points are purified to a final fidelity $F_{\text{out}} \geqslant a_u$. There are, however, two complications: first, the eventual fate of a point (a,b) depends on both a and b. Second, we need to make sure that the conditions (32) are fulfilled in every step of the iteration, otherwise it is no longer valid to interpret (a_n, b_n) as bounds of the actual values $(\tilde{A}_n^0, \tilde{A}_n^1)$. For both of these complications we have been unable to find complete analytical answers. Therefore we first give the numerical results before mentioning partial analytical solutions.

Numerical calculations show that the physically meaningful set $\{(a,b): 0 \leqslant a \leqslant 1, 0 \leqslant b \leqslant 1-a\}$ is divided in two parts by a curve passing through the intermediate fixed point, the separatrix (see Fig. 2). Points to the right of that curve converge to the upper fixed point, points to the left towards the lower one. Moreover, all points to the right satisfy the conditions (32) and so do the orbits of all these points. For all ensembles described by density matrices with diagonal elements A_0^0, A_0^1 in that region, a_n, b_n as defined in Eq. (10) provide lower and upper bounds for the respective fidelities after n purification steps. For initial values to the left of the separatrix our approach allows no statement. The case $\delta = 0$ in Fig. 2 indicates how many ''good'' points our worst-case consideration misses: as shown in [17] the exact border of the set of purifiable points in the (a,b) plane is given by the straight line $a = 0.5$.

For a subset of the points to the right of the separatrix it is easy to *prove* convergence: All the points (a,b) fulfilling $a \geqslant a_i$, $b \leqslant b_i$, and $a + b \leqslant 1$ converge to the upper fixed point P_u (except for P_i, of course).

Proof: The proof proceeds in four steps. The main tool is the monotonic dependence of a_{n+1}, b_{n+1} on a and b. [It is easily checked by calculation that the coordinates of the intermediate fixed point satisfy the conditions (32) for all δ so that monotonicity holds.]

(i) Consider (a,b) in the set enclosed by the two curves $b_{n+1}(a)$ and $b_{\text{fix}}(a)$ [Eq. (33), cf. Fig. 3]. For these points, we have for all n

$$a_{n+1} \geqslant a_n \quad \text{and} \quad b_{n+1} \leqslant b_n .$$

Since a_n and b_n are bounded by the coordinates of the upper and intermediate fixpoints, they form monotonical, bounded sequences and therefore converge. Since a_n increases and b_n decreases, they converge towards (a_u, b_u).

(ii) Similarly it is seen that all points $(a \geqslant a_u, b \leqslant b_u)$ do converge to the fixed point ''from above.''

(iii) Now, consider a point $X = (a, b \leqslant b_u)$ below the curve $b_{n+1}(a)$.

Let us call a point (a,b) *better* than (a',b'), if $a \geqslant a'$ and $b \leqslant b'$. Monotonicity implies that if (a,b) better than (a',b') then this will also be true for the images of these points after one iteration of the dynamical system.

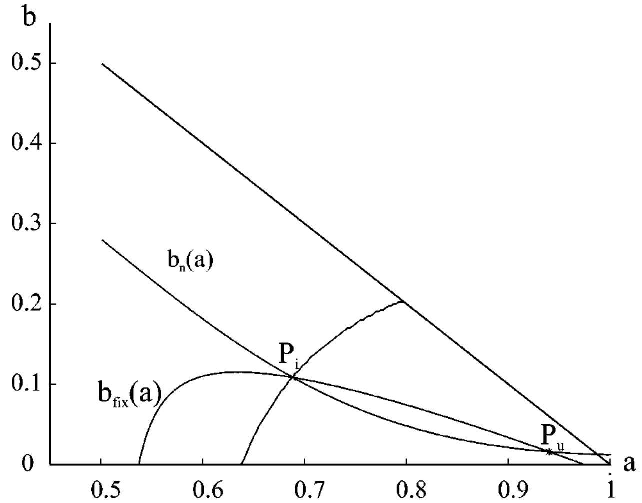

FIG. 3. For $\delta = 0.006$ the curves b_n (10b) and b_{fix} (33) are plotted. Their intersections are fixed points of the dynamical system.

Now compare X with $X' = (a' = a, b')$ between the curves but with the same a as X, and with $X'' = (a'' \geqslant a_u, b'' = b)$. Clearly, X is better than X' but worse than X''. Since both X' and X'' converge towards the upper fixpoint, so does X.

(iv) A similar argument applies, if we compare a point $Y = (a, b > b_{\text{fix}}(a))$ with $Y' = (a' < a, b' = b)$ between the curves and $Y'' = (a'' = a, b'' \leqslant b)$ below the curves: the primed points converge to the upper fixpoint, and thus (a,b)—being better than Y' and worse than Y''—does so, too. This completes the proof.

V. SUMMARY

The entanglement purification protocols [1,7] in the presence of errors in gate operations and measurements have been investigated. The errors are quantified by a single parameter derived from the trace norm. We have shown that these protocols allow us to increase the fidelity of the entanglement even if implemented with imperfect quantum gates and measurements, as long as the errors are below a threshold of the order 1%. We derived a nonlinear map to calculate a lower bound for the fidelity after n purification steps. Polynomials are given, a root of which gives a lower bound for the asymptotically attainable fidelity.

The methods and definitions introduced in this work can be applied to other interesting problems in quantum information, like teleportation or quantum cryptography. Furthermore, they can be used to analyze other purification protocols which, under certain circumstances, are more efficient than the ones studied here (see, for example, Refs. [1,2]).

ACKNOWLEDGMENTS

This work was supported in part by the Österreichischer Fonds zur Förderung der wissenschaftlichen Forschung and by the European TMR Network No. ERB-FMRX-CT96-0087. G.G. thanks Wolfgang Dür for useful discussions. Part of this work was completed during the 1998 Elsag-Bailey–I.S.I. Foundation research meeting on quantum computation.

[1] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[2] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[3] N. Gisin, Phys. Lett. A **210**, 151 (1996).

[4] B. Schumacher, Phys. Rev. A **54**, 2614 (1996); B. Schumacher and M. D. Westmoreland, *ibid.* **56**, 131 (1997).

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

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

[7] D. Deutsch, A. Ekert, R. Josza, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).

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

[9] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[10] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **78**, 574 (1997).

[11] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).

[12] E. Knill and R. Laflamme, e-print quant-ph/9608012; E. Knill, R. Laflamme, and W. Zurek, Proc. R. Soc. London, Ser. A **454**, 365 (1998).

[13] S. Van Enk, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 4293 (1997); Science **279**, 205 (1998).

[14] J. I. Cirac, P. Zoller, J. H. Kimble, and H. Mabuchi, and Phys. Rev. Lett. **78**, 3221 (1997).

[15] P. Shor, SIAM J. Comput. **26**, 1484 (1997); e-print quant-ph/9605011; A. M. Steane, Phys. Rev. Lett. **78**, 2252 (1997); D. Gottesman, e-print quant-ph/970229.

[16] H. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **81**, 5932 (1998); W. Dür, H. Briegel, J. I. Cirac, and P. Zoller, Phys. Rev. A **59**, 169 (1999).

[17] C. Macchiavello, Phys. Lett. A **246**, 385 (1998).

[18] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997); I. Chuang and M. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[19] D. Aharonov, A. Kitaev, and N. Nisan, e-print quant-ph/9806029.

[20] As shown in P. Pechukas, Phys. Rev. Lett. **73**, 1060 (1994), reduced dynamics in general need not be completely positive (not even positive) on the whole system space. In the case of weak coupling between system and environment, short memory of the environment and time coarse graining, the description of the reduced dynamics by a completely positive map is justified even in the case of initial entanglement [see A. Royer, Phys. Rev. Lett. **77**, 3272 (1996)]. We thank D. Lidar for pointing out these references.