Optimal estimation of two-qubit pure-state entanglement

Antonio Acín, Rolf Tarrach, and Guifré Vidal
Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain
(Received 3 November 1999; published 16 May 2000)

We present optimal measuring strategies for an estimation of the entanglement of unknown two-qubit pure states and of the degree of mixing of unknown single-qubit mixed states, of which N identical copies are available. The most general measuring strategies are considered in both situations, to conclude in the first case that a local, although collective, measurement suffices to estimate entanglement, a nonlocal property, optimally.

PACS number(s): 03.67.−a, 03.65.Bz

I. INTRODUCTION

Plenty of work has been performed in recent years on optimal quantum measurements, i.e., on measurements which provide the maximum possible information about an unknown quantum-mechanical pure [1–5] or mixed [6] state, of which N identical copies are available. These works focused mainly on a determination of the unknown state as a whole, and consequently any of its properties is also estimated, although maybe not in an optimal way.

On the other hand, recent developments on the field of quantum information theory stressed the importance of the quantum correlations—or entanglement—displayed by some states of composite systems. In the simplest of such composite systems, the two-qubit case, all nonlocal properties of pure states depend upon only one single parameter. Such a nonlocal parameter is the only relevant quantity invariant under local unitary transformations on each qubit, and plays a central role in the quantification and optimal manipulation of entanglement [7–11].

In this work we analyze and solve the problem of optimally estimating the entanglement of an unknown pure state of two qubits. This problem was also independently addressed by Sancho and Huelga in a recent work [12], where only a restricted class of measuring strategies is considered. Here, on the contrary, we will consider most general quantum measurements on N identical copies of the state. Their quality will be assessed through the gain of information they provide about the nonlocal parameter of the state. After presenting and proving the solution, we will conclude that the optimal measuring strategies so defined are not equivalent to the ones used to fully reconstruct the unknown state. As a matter of fact, *all* information about some relative phase of the unknown state turns out to be irreversibly erased as the entanglement is estimated.

An estimation of the degree of mixing of an unknown mixed state is a different but very much related topic that we shall also consider here. For the single-qubit case the amount of mixing is again specified by just one parameter, the modulus of the corresponding Bloch vector, whereas in order to completely specify the state two more parameters, namely, the direction of the Bloch vector, are also required. We shall show that in this case the optimal measuring strategy on any number N of qubits prepared in the same mixed state can be made compatible with an optimal estimation of the direction of its Bloch vector.

Finally, we will show that a possible way of optimally determining the entanglement of an unknown, two-qubit pure state consists precisely of estimating, also optimally, the degree of mixture of any of its two reduced density matrices. Therefore, in this simple bipartite case it turns out that the optimal estimation of a nonlocal parameter can be done through a local measurement.

The paper is structured as follows. Section II is devoted to background material. We introduce a convenient parametrization of two-qubit pure states, and consider their isotropic distribution. We also review some basic aspects on parameter estimation and on quantum measurements. In Sec. III we pose the problem of entanglement estimation on firmer grounds and announce the main result of this paper: its optimal performance. Section IV, which is rather technical and could well be skipped in a first reading, is devoted to a computation of some effective density matrix $\rho^{(N)}(b)$, an object which plays a central role in deriving the optimal strategy for estimating entanglement. In Sec. V the $N=1$, 2, and 3 cases are presented in more detail in order to illustrate the general case. Optimal estimation of the degree of mixing is discussed and solved in Sec. VI, and finally Sec. VII contains a discussion relating estimation of both entanglement and mixing, and some concluding remarks.

II. PRELIMINARIES

Here we will consider a two-party scenario. Alice and Bob will share N copies of a completely unknown two-qubit pure state $|\psi\rangle$, and their aim will be to obtain as much information as possible about its entanglement. The sense in which the state is *unknown*, the mechanisms for *extracting* information from the system, and the scheme for *evaluating* the extracted information will be briefly reviewed in what follows.

A. Homogeneous distribution

All that is initially known about the state of each pair of qubits is that it is pure. This corresponds to the unbiased distribution on the Hilbert space $\mathcal{H}_4=\mathcal{H}_2\otimes\mathcal{H}_2$ of two qubits, that is, to the only probability distribution invariant under arbitrary unitary transformations on $\mathcal{H}_4$. It is convenient to

express the unknown state $|\psi\rangle \in \mathcal{H}_2 \otimes \mathcal{H}_2$, which depends on six parameters, in its Schmidt-like decomposition

$$|\psi\rangle = \sqrt{\frac{1+b}{2}}|\hat{a}\rangle|\hat{b}\rangle + \sqrt{\frac{1-b}{2}}e^{i\alpha}|-\hat{a}\rangle|-\hat{b}\rangle, \quad (1)$$

where the phase $e^{i\alpha}$, which is usually absorbed by one of the kets it goes with, has been left explicit. The nonlocal parameter $b \in [0,1]$ characterizes the entanglement of $|\psi\rangle$. Only for $b=1$ is $|\psi\rangle$ a product state $|\hat{a}\rangle \otimes |\hat{b}\rangle$, and thus unentangled. For $b<1$ the state contains quantum correlations $b=0$ corresponding to a maximally entangled state. Recall that this parameter is the modulus of the Bloch vector of the reduced density matrix ρ_A on Alice's side,

$$\rho_A \equiv tr_B|\psi\rangle\langle\psi| = \frac{1+b}{2}|\hat{a}\rangle\langle\hat{a}| + \frac{1-b}{2}|-\hat{a}\rangle\langle-\hat{a}|, \quad (2)$$

and equivalently for ρ_B. The other four parameters correspond to the two directions $\hat{a}$ and $\hat{b}$ of the Bloch vectors of ρ_A and ρ_B. Then the unbiased distribution of pure states corresponds [13] to the isotropic distribution of $\hat{a}$ in S^2, $\hat{b}$ in S^2, α in S^1, and the quadratic distribution of b in [0,1], which is actually also a flat distribution, as b^2 is just the Jacobian corresponding to going from Cartesian to spherical coordinates:

$$\int_{S^2}\frac{d\hat{a}}{4\pi}\int_{S^2}\frac{d\hat{b}}{4\pi}\int_{S^1}\frac{d\alpha}{2\pi}\int_0^1 db\, 3b^2 = 1. \quad (3)$$

B. General measurements and information gain

The parties are thus provided with N copies of a pure state $|\psi\rangle$ as in Eq. (1), i.e., with the state $|\psi\rangle^{\otimes N}$, and our aim is to construct the most informative measurement on the collective, $2N$-qubit system for the estimation of the parameter b. The optimality criterion to be used is based on the Kullback or mutual information $K[f',f]$ [14], a functional of two probability distributions f' and f that is interpreted as the gain of information in replacing the latter distribution with the former one [15]. In our case, for instance, the prior, unbiased density function for the parameter b is given by Eq. (3), so we have $f(b)=3b^2$. A generic measurement, allowing for the most general manipulation of the system, is represented by a resolution of the identity by means of a set of positive operators:

$$\sum_k M^{(k)} = I. \quad (4)$$

After the above positive operator valued measurement (POVM) has been performed, giving the outcome k with probability $\mathrm{tr}(M^{(k)}\rho^{\otimes N})$, where $\rho=|\psi\rangle\langle\psi|$, we compute the posterior density function for b, $f(b|k)$, through the Bayes formula

$$f_k(b) \equiv f(b|k) = \frac{p(k|b)f(b)}{p(k)}, \quad (5)$$

where $p(k)$ is given by

$$p(k) = \int_0^1 db f(b) p(k|b), \quad (6)$$

and the conditional probability of obtaining outcome k when the state's nonlocal parameter has value b, $p(k|b)$ will be shown later. The gain of information resulting from obtaining the outcome k after the measurement is quantified by the Kullback information corresponding to the prior and posterior probability density functions:

$$K[f_k,f] = \int db f(b|k)\ln\left(\frac{f(b|k)}{f(b)}\right). \quad (7)$$

This expression has to be averaged over all the possible outcomes of the measurement, so that the expected gain of information reads

$$\overline{K}[f_k,f] = \sum_k p(k)K[f_k,f], \quad (8)$$

using Eq. (5), this expression can be written as

$$\overline{K}[f_k,f] = \sum_k \int db f(b)p(k|b)\ln\left(\frac{p(k|b)}{p(k)}\right). \quad (9)$$

Let us note here that the value of $K[f_k,f]$ in Eq. (7) would remain unchanged if we decided to characterize the entanglement of $|\psi\rangle$ by another parameter $b=h(b)$ [where $h(b)$ is any bijective function of the original parameter b]. Consequently, the gain of information we compute for b also applies to any of the measures of entanglement so far proposed, such as the entanglement of formation [7],

$$-\sqrt{\frac{1+b}{2}}\log_2\sqrt{\frac{1+b}{2}} - \sqrt{\frac{1-b}{2}}\log_2\sqrt{\frac{1-b}{2}}, \quad (10)$$

for the asymptotic regime, or the monotone [10]

$$\sqrt{\frac{1-b}{2}} \quad (11)$$

for the single-copy case.

III. OPTIMAL MEASUREMENTS FOR ENTANGLEMENT ESTIMATION

We are looking for a measurement of the form of Eq. (4), such that the expected gain of information [Eq. (9)] is maximized. Here and in Sec. V we will present and explain such optimal measurements, whereas their explicit construction is mainly contained in Sec. IV.

A. Local and global strategies

Before we proceed we comment on four classes of measurements Alice and Bob may consider in order to learn about b [12]:

(i) *Local* measurements on only, say, Alice's side, i.e., on the N qubits supporting the local state $\rho_A^{\otimes N}$, would be the most restrictive class of the hierarchy.

(ii) *Uncorrelated bilocal* measurement, in which each party measures their local N-qubit part independently, is one type of intermediate strategy.

(ii) *Classically correlated bilocal* measurement, with classical communication between Alice and Bob, is a less restrictive intermediate strategy.

(iv) *Global* measurements on the $2N$ qubits constitute the most general case.

Global measurements are in principle the most informative ones. But as the parameter b, which quantifies the entanglement of $|\psi\rangle$, also completely quantifies the mixing of ρ_A (and ρ_B), it could well happen that local measurements, or bilocal measurements on the two parties, optimal for the determination of the mixing, are as informative as the global ones with respect to entanglement. In fact, in reducing $|\psi\rangle\langle\psi|$ to $\rho_A\otimes\rho_B$ only the relative phase α is lost, and the dependence on directions $\hat{a}$ and $\hat{b}$ and on the entanglement b is preserved. We have found the optimal global and local measurement of b. The results obtained following the two strategies are the same, as we will discuss in Sec. VII, so all the extractable information about the entanglement is preserved under the partial trace operation, and the four classes considered above turn out to be equivalent for entanglement estimation.

B. Effective mixed state

Note that all the dependence on the measuring strategy (4) in Eq. (9) is contained in the probability $p(k|b)$ of outcome k conditioned on the entanglement of the state being some given b,

$$p(k|b)=\int_{S^2}\frac{d\hat{a}}{4\pi}\int_{S^2}\frac{d\hat{b}}{4\pi}\int_{S^1}\frac{d\alpha}{2\pi}\mathrm{tr}(M^{(k)}\rho^{\otimes N}), \tag{12}$$

where the sum over the rest of the parameters reflects the fact that we are only interested in the entanglement. This expression can also be written as

$$p(k|b)=\mathrm{tr}[M^{(k)}\rho^{(N)}(b)], \tag{13}$$

where the mixed state $\rho^{(N)}(b)$ is

$$\rho^{(N)}(b)\equiv\int_{S^2}\frac{d\hat{a}}{4\pi}\int_{S^2}\frac{d\hat{b}}{4\pi}\int_{S^1}\frac{d\alpha}{2\pi}|\psi\rangle\langle\psi|^{\otimes N}. \tag{14}$$

Equation (13) allows for an alternative interpretation to our problem: a $2N$-qubit mixed state $\rho^{(N)}(b)$ is drawn randomly with prior probability distribution $f(b)=3b^2$, and we want to determine it by estimating b.

We will compute $p(k|b)$ on a basis that diagonalizes $\rho^{(N)}(b)$, which will crucially turn out to be independent of b. Let us denote the positive eigenvalues of $\rho^{(N)}(b)$ by $\lambda_1(b),\ldots,\lambda_m(b)$, and their multiplicity by $n_1,\ldots,n_m$. From the normalization of Eq. (14) the relation $\Sigma_{j=1}^m n_j\lambda_j=1$ follows. The sum $n\equiv\Sigma_j n_j$ of multiplicities of (nonvanishing) eigenvalues equals the dimension of the space which supports $|\psi\rangle\langle\psi|^{\otimes N}$. This is the symmetric subspace of $\mathcal{H}_4^{\otimes N}$, and thus [5]

$$n=\frac{(N+3)!}{3!N!}=\frac{(N+3)(N+2)(N+1)}{6}. \tag{15}$$

With this notation Eq. (13) reads

$$p(k|b)=\lambda_1(b)\sum_{i=1}^{n_1}M_{ii}^{(k)}+\lambda_2(b)\sum_{i=n_1+1}^{n_1+n_2}M_{ii}^{(k)}+\cdots+\lambda_m(b)\sum_{i=n-n_m+1}^{n}M_{ii}^{(k)}\equiv\sum_{j=1}^{m}\lambda_j(b)q_j^{(k)}. \tag{16}$$

By substituting this expression into Eq. (9), and using the inequality [16]

$$(x_1+x_2)\ln\left(\frac{x_1+x_2}{y_1+y_2}\right)\leqslant x_1\ln\left(\frac{x_1}{y_1}\right)+x_2\ln\left(\frac{x_2}{y_2}\right), \tag{17}$$

where $x_i,y_i\geqslant 0$, along with the fact that the POVM is a resolution of the identity in the symmetric subspace of $\mathcal{H}_4^{\otimes N}$, i.e. $\Sigma_k q_j^{(k)}=n_j$, it follows that the average gain of information is bounded by

$$\overline{K}[f_k,f]\leqslant\int db f(b)\sum_{j=1}^{m}n_j\lambda_j(b)\ln\left(\frac{\lambda_j(b)}{\int db f(b)\lambda_j(b)}\right). \tag{18}$$

C. Minimal most informative measuring strategy

Bound (18) can be minimally saturated through a measurement with m outcomes, where each $M^{(k)}$ is the n_k-dimensional projector over the subspace corresponding to the eigenvalue λ_k of $\rho^{(N)}(b)$, then having $p(k|b)=n_k\lambda_k(b)$. Therefore, the construction of the optimal measurement can be readily performed after the computation of the spectral decomposition of state (14), and this is done for an arbitrary N in Sec. IV. For a more detailed account of the $N=1$, 2, and 3 cases, see Sec. V, where also the gain of information up to $N=80$ has been computed explicitly.

Note also that there are other ways measuring strategies that can be evaluated and, consequently, there is not a unique notion of optimality. For instance, in Refs. [1–6] a guess for the unknown state is made depending on the outcome of the measurement, and then both guessed and unknown states are compared using the fidelity. It can be proved, following Ref. [16], that the optimal measurements presented here, the most informative ones, are also optimal if we decide, alternatively, on a fidelitylike figure of merit satisfying some very general conditions [19].

IV. COMPUTATION OF $\rho^{(N)}$

It has been shown that the spectrum of $\rho^{(N)}(b)$ determines the maximal gain of information about b, whereas its eigenprojectors lead to the corresponding measuring strategy. Our

next step will be the computation of the spectral decomposition of this effective mixed state.

Let us rewrite the generic two-qubit pure state [Eq. (1)] as

$$|\psi\rangle = U_A \otimes U_B (c_+ |+\rangle_A \otimes |+\rangle_B + c_- |-\rangle_A \otimes |-\rangle_B) \equiv U_A \otimes U_B |\psi(b)\rangle, \tag{19}$$

where $c_+ \equiv \sqrt{(1+b)/2}$, $c_- \equiv \sqrt{(1-b)/2}$, the single-qubit pure states $|+\rangle_A$ and $|-\rangle_A$ ($|+\rangle_B$ and $|-\rangle_B$) constitute an orthonormal basis in Alice's (Bob's) part (corresponding to some fixed direction in the Bloch sphere), U_A and U_B are unitary transformations in each single-qubit space, and $|\psi(b)\rangle$ is a reference state.

The state $\rho^{(N)}(b)$ corresponds then to a Haar integral over the group $\mathrm{SU}(2)\times\mathrm{SU}(2)$, since it can be expressed as

$$\rho^{(N)}(b) = \int_{g\in G} dg [D(g) M(b) D(g)^\dagger]^{\otimes N}, \tag{20}$$

where the index g denotes the elements of the group $G = \mathrm{SU}(2)\times\mathrm{SU}(2)$, $D(g) = \mathrm{U}_A \otimes \mathrm{U}_B$ is a $\frac{1}{2}\times\frac{1}{2}$ irreducible representation (irrep) of this group and $M(b) = |\psi(b)\rangle\langle\psi(b)|$.

A well-known result in group representation theory following from Schur's lemma, the so-called orthogonality lemma, will be useful in the calculation of this integral. Consider a matrix $A^{\alpha\beta}(B)$ given by

$$A^{\alpha\beta}(B) = \int_{g\in G} dg D^\alpha(g) B D^{\beta\dagger}(g), \tag{21}$$

where D^α and D^β are two unitary irreps of the group G. Then we have the following,

Lemma 1 (orthogonality lemma):

$$A^{\alpha\beta}(B) = a(B)\,\delta^{\alpha\beta} I, \tag{22}$$

so $A^{\alpha\beta}(B)$ is zero if the two representations are inequivalent, and proportional to the identity if the two representations are equivalent.

In order to benefit from this lemma we identify B with $M(b)^{\otimes N} = |\psi(b)\rangle\langle\psi(b)|^{\otimes N}$ and then consider the relevant irreps of $\mathrm{SU}(2)\times\mathrm{SU}(2)$ borne by the N-fold tensor product of the $\frac{1}{2}\times\frac{1}{2}$ irrep of the group. These representations are the support of the state $|\psi(b)\rangle^{\otimes N}$, and our next task is to recognize them.

The state $|\psi(b)\rangle^{\otimes N}$ can be expanded as

$$\begin{aligned}|\psi(b)\rangle^{\otimes N} = {} & c_+^N |++\cdots++\rangle_A \otimes |\cdot\rangle_B, \\ & + c_+^{N-1} c_- (|++\cdots+-\rangle_A \otimes |\cdot\rangle_B + \cdots \\ & + |-+\cdots++\rangle_A \otimes |\cdot\rangle_B), \\ & + c_+^{N-2} c_-^2 (|+\cdots+--\rangle_A \otimes |\cdot\rangle_B + \cdots \\ & + |--+\cdots+\rangle_A \otimes |\cdot\rangle_B), \\ & + c_+^{N-3} c_-^3 (\quad) + \cdots + c_+ c_-^{N-1} (\quad), \\ & + c_-^N |--\cdots--\rangle_A \otimes |\cdot\rangle_B,\end{aligned} \tag{23}$$

where $|\cdot\rangle_B$ means that we have exactly the same vector in the second subsystem. Notice that in the expression above all the elements of the product basis $\{|u_i\rangle\}$ of the local spaces $\mathcal{H}_2^{\otimes N}$ of Alice's and Bob's N qubits—i.e., $|u_1\rangle = |++\cdots++\rangle$, $|u_2\rangle = |++\cdots+-\rangle$, $\ldots$, $|u_{2^N}\rangle = |--\cdots--\rangle$ — appear in the form $|u_i\rangle_A \otimes |u_i\rangle_B$. Notice, in addition, that if we denote by m_T the sum of the third spin component of all spinors in each ket—i.e., for instance $m_T(|+++\rangle) = 3/2$, $m_T(|++-\rangle) = 1/2$, $m_T(|-+-\rangle) = -1/2$, $\cdots$–, the terms multiplied by the same combination of the factors c_+ and c_- have the same m_T in A and B. State (23) can thus also be expressed as

$$\begin{aligned}|\psi(b)\rangle^{\otimes N} = {} & c_+^N \sum_{i;m_T=N/2} |u_i\rangle_A \otimes |u_i\rangle_B \\ & + c_+^{N-1} c_- \sum_{i;m_T=(N/2)-1} |u_i\rangle_A \otimes |u_i\rangle_B + \cdots \\ & + c_-^N \sum_{i;m_T=-N/2} |u_i\rangle_A \otimes |u_i\rangle_B .\end{aligned} \tag{24}$$

We now move from the local spin basis $\{|u_i\rangle_A\}$ to the coupled one $\{|v_i\rangle_A\}$ in Alice's N qubits, and we also do the same in Bob's. The following lemma, that can be easily checked, will be useful here.

Lemma 2: Let $\{|e_i\rangle\}$ and $\{|f_i\rangle\}$ be two orthonormal basis in $\mathcal{C}^l$, related by an orthogonal transformation O, so that $|e_i\rangle = \Sigma_j O_{ij} |f_j\rangle$, with $O^* = O$, and $O^{-1} = O^\dagger$. Then,

$$\sum_{i=1}^{l} |e_i\rangle \otimes |e_i\rangle = \sum_{i=1}^{l} |f_i\rangle \otimes |f_i\rangle. \tag{25}$$

Now, note that the unitary transformation relating the local basis and the coupled one is real (since all the Clebsch-Gordan coefficients are real), and that there is a conservation rule for the total third spin component (i.e., the Clebsch-Gordan coefficients that couple two states with third component m_1 and m_2 to a coupled state with third component m are proportional to δ_{m,m_1+m_2}). Then Eq. (24) can be reexpressed, using the previous two facts and lemma 2, in the coupled basis as

$$\begin{aligned}|\psi(b)\rangle^{\otimes N} = {} & c_+^N \sum_{i;m_T=N/2} |v_i\rangle_A \otimes |v_i\rangle_B \\ & + c_+^{N-1} c_- \sum_{i;m_T=(N/2)-1} |v_i\rangle_A \otimes |v_i\rangle_B + \cdots \\ & + c_-^N \sum_{i;m_T=-N/2} |v_i\rangle_A \otimes |v_i\rangle_B\end{aligned} \tag{26}$$

(see the examples in Sec. V for more details). We note that the symmetry between the terms in A and in B allows us to derive Eq. (26) from Eq. (24).

Let us now have a closer look into Eq. (26). The term with coefficient c_+^N corresponds simply to the state with a total spin j maximal in both Alice's and Bob's subsystem

(i.e., $j_A=j_B=N/2$) and also maximal third spin component m, namely, $m_A=m_B=N/2$. We can thus write, with the notation $|^{j_A}m_A\rangle_A\otimes|^{j_B}m_B\rangle_B$, $|v_1\rangle\equiv|v_1\rangle_A\otimes|v_1\rangle_B=|^{N/2}N/2\rangle_A\otimes|^{N/2}N/2\rangle_B$. This state belongs to a $N/2\otimes N/2$ irrep of the group $\mathrm{SU}(2)\times\mathrm{SU}(2)$. The coefficient $c_+^{N-1}c_-$ corresponds to all states with $m_A=m_B=(N/2)-1$. Apart from $|v_2\rangle\equiv|^{N/2}(N/2)-1\rangle_A\otimes|^{N/2}(N/2)-1\rangle_B$, which again belongs to the previous $N/2\otimes N/2$ irrep, the remaining $N-1$ kets, $|v_3\rangle\cdots|v_{N+1}\rangle$ have $j_A=j_B=(N/2)-1$, and thus belong to $N-1$ different (but equivalent)

$$\left(\frac{N}{2}-1\right)\otimes\left(\frac{N}{2}-1\right)$$

irreps of the group. But since only the linear combination $|v_3\rangle+\cdots+|v_{N+1}\rangle$ appears, the relevant irrep is just the symmetric combination of the latter $N-1$ ones, which we will denote by

$$\left\{\left(\frac{N}{2}-1\right)\otimes\left(\frac{N}{2}-1\right)\right\}_{sym},$$

and which no longer decomposes as the product of two irreps of $SU(2)$. The same applies for

$$\left(\frac{N}{2}-2\right)\otimes\left(\frac{N}{2}-2\right)$$

irreps, and so on.

Thus, the space which supports the initial state can be decomposed in terms of irreps of $\mathrm{SU}(2)\times\mathrm{SU}(2)$ as

$$\frac{N}{2}\otimes\frac{N}{2}\oplus\left\{\left(\frac{N}{2}-1\right)\otimes\left(\frac{N}{2}-1\right)\right\}_{sym}\oplus\cdots\oplus\left\{\frac{N\bmod 2}{2}\otimes\frac{N\bmod 2}{2}\right\}_{sym},\tag{27}$$

where $N \bmod 2$ is equal to 1 for odd N and equal to zero for even N. It can be checked that this result agrees dimensionally with formula (15).

The decomposition shown above in terms of the relevant irreps of the group $\mathrm{SU}(2)\times\mathrm{SU}(2)$, together with the orthogonality lemma, can be used to solve the integral in Eq. (20). As we have argued, when plugging Eq. (26) into Eq. (20) the cross terms corresponding to inequivalent representations—such as $|v_1\rangle(\langle v_3|+\cdots+\langle v_{N+1}|)$—vanish as we integrate, while the terms within the same representation—such as $|v_1\rangle\langle v_1|$—lead to a contribution proportional to the identity in the subspace associated with the representation. So the state $\rho^{(N)}(b)$ is equal to

$$\rho^{(N)}(b)=\lambda_1(b)I_{N/2\otimes N/2}+\lambda_2(b)I_{\{[(N/2)-1]\otimes[(N/2)-1]\}_{sym}}+\cdots+\lambda_m(b)I_{\{[(N\bmod 2)/2]\otimes[(N\bmod 2)/2]\}_{sym}}.\tag{28}$$

This is the spectral decomposition we are looking for, where $\{\lambda_j\}$ are the entanglement dependent eigenvalues of $\rho^{(N)}(b)$, the trace of the identities giving the corresponding multiplicities $\{n_j\}$. It is important to notice that, as it was mentioned before, the eigenspaces are independent of b.

The calculation of $n_j\lambda_j$ can now be readily performed from Eq. (26) by computing the trace of the projection of $|\psi(b)\rangle^N$ into each relevant irrep. The determination of the spectrum of $\rho^{(N)}(b)$ completes, as we have shown, the construction of the optimal measurement for the estimation of the entanglement. In Sec. V some examples are studied in order to clarify the implementation of the procedure.

V. SOME EXAMPLES: THE $N=1,2,3$ CASES AND BEYOND

In this section we will apply the procedure described above to obtain the optimal estimation of b when one, two, and three identical copies of the initial state are at our disposal.

A. $N=1$

The simplest case, $N=1$, is now straightforward. The state written as in Eq. (19) belongs to the $\frac{1}{2}\otimes\frac{1}{2}$ irrep of $\mathrm{SU}(2)\times\mathrm{SU}(2)$. From Eq. (20) we have, using the orthogonality lemma as in Eq. (28),

$$\rho^{(1)}(b)=\int dg D(g)M(b)D(g)^\dagger=\lambda_1(b)I.\tag{29}$$

The eigenvalue $\lambda_1(b)=\frac{1}{4}$ is obtained by taking the trace in the expression above. The probability $p(k|b)$ [see Eq. (13)] is independent of b, so that $p(k)=p(k|b)$ and the average Kullback information [Eq. (9)] vanishes. Consequently, no information whatsoever can be obtained about the entanglement of a completely unknown pure state if only one copy is at our disposal.

B. $N=2$

For the $N=2$ case the initial state has the form, from Eqs. (23) or (24),

$$|\psi(b)\rangle^{\otimes 2}=c_+^2|++\rangle_A\otimes|\cdot\rangle_B+c_+c_-(|+-\rangle_A\otimes|\cdot\rangle_B+|-+\rangle_A\otimes|\cdot\rangle_B)+c_-^2|--\rangle_A\otimes|\cdot\rangle_B,\tag{30}$$

Now, using lemma 2 and the conservation law mentioned above for the Clebsch-Gordan coefficients [cf. Eq. (26)], we can rewrite the state as

$$|\psi(b)\rangle^{\otimes 2}=c_+^2|^1 1\rangle_A\otimes|\cdot\rangle_B+c_+c_-(|^1 0\rangle_A\otimes|\cdot\rangle_B+|^0 0\rangle_A\otimes|\cdot\rangle_B)+c_-^2|^1-1\rangle_A\otimes|\cdot\rangle_B,\tag{31}$$

where for each party the coupled basis is related to the local one by means of an orthogonal transformation, as usual,

$$|^1 1\rangle=|++\rangle,\quad |^1-1\rangle=|--\rangle,\qquad |^1 0\rangle=\frac{1}{\sqrt{2}}(|+-\rangle+|-+\rangle),\tag{32}$$

$$|^0 0\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle).$$

The state $|\psi(b)\rangle^{\otimes 2}$ in Eq. (31) is supported then in the $1\otimes 1$ and the $0\otimes 0$ irreps of $SU(2)\times SU(2)$, and now the application of lemma 1 gives for $\rho^{(2)}(b)$:

$$\rho^{(2)}(b) = \lambda_1(b) I_{1\otimes 1} + \lambda_2(b) I_{0\otimes 0}. \tag{33}$$

We just need to pick up the contributions of Eq. (31) to each irrep, that is the trace of the corresponding projections, to find that

$$n_1\lambda_1(b) = (c_+^4 + c_+^2 c_-^2 + c_-^4) = \frac{3+b^2}{4},$$
$$n_2\lambda_2(b) = c_+^2 c_-^2 = \frac{1-b^2}{4}. \tag{34}$$

The optimal measurement [see Eq. (18)] then consists of two projectors onto the $1\otimes 1$ and $0\otimes 0$ irreps of $SU(2)\otimes SU(2)$, with probabilities $p(1|b) = n_1\lambda_1(b) = (3+b^2)/4$ and $p(2|b) = n_2\lambda_2(b) = (1-b^2)/4$, and from them $p(1) = \frac{9}{10}$ and $p(2) = \frac{1}{10}$. Finally the gain of information can be computed, using Eq. (9), and it gives $\overline{K} = 0.0375$ bits.

C. $N=3$

The last case we want to discuss is $N=3$. Starting now from Eq. (26), we have

$$\begin{aligned}|\psi(b)\rangle^{\otimes 3} = {} & c_+^3 |^{3/2}\tfrac{3}{2}\rangle_A \otimes |\cdot\rangle_B + c_+^2 c_- (|^{3/2}\tfrac{1}{2}\rangle_A \otimes |\cdot\rangle_B + |^{1/2}\tfrac{1}{2}\rangle_A \\ & \otimes |\cdot\rangle_B + |^{1/2'}\tfrac{1}{2}\rangle_A \otimes |\cdot\rangle_B) + c_+ c_-^2 (|^{3/2}-\tfrac{1}{2}\rangle_A \otimes |\cdot\rangle_B \\ & + |^{1/2}-\tfrac{1}{2}\rangle_A \otimes |\cdot\rangle_B + |^{1/2'}-\tfrac{1}{2}\rangle_A \otimes |\cdot\rangle_B) \\ & + c_-^3 |^{3/2}-\tfrac{3}{2}\rangle_A \otimes |\cdot\rangle_B,\end{aligned} \tag{35}$$

we observe that only contributions to the $\frac{3}{2}\otimes\frac{3}{2}$ and to two different $\frac{1}{2}\otimes\frac{1}{2}$ irreps of $SU(2)\times SU(2)$ appear. Notice, in addition, that since in this expansion the contributions to $\frac{1}{2}\otimes\frac{1}{2}$ and to $\frac{1}{2}'\otimes\frac{1}{2}'$ only appear in a symmetric linear combination (i.e., $|^{1/2}\tfrac{1}{2}\rangle_A\otimes|\cdot\rangle_B + |^{1/2'}\tfrac{1}{2}\rangle_A\otimes|\cdot\rangle_B$ and $|^{1/2}-\tfrac{1}{2}\rangle_A \otimes|\cdot\rangle_B + |^{1/2'}-\tfrac{1}{2}\rangle_A\otimes|\cdot\rangle_B$), the relevant irreps is precisely a symmetric combination of the two latter ones, $\{\frac{1}{2}\otimes\frac{1}{2}\}_{sym}$. The orthogonality lemma gives now

$$\rho^{(3)}(b) = \lambda_1(b) I_{3/2\otimes 3/2} + \lambda_2(b) I_{\{1/2\otimes 1/2\}_{sym}}. \tag{36}$$

Finally, by collecting the traces of each projection of Eq. (35) onto each irrep, we obtain

$$n_1\lambda_1(b) = (c_+^6 + c_+^4 c_-^2 + c_+^2 c_-^4 + c_-^6) = \frac{1+b^2}{2},$$
$$n_2\lambda_2(b) = 2(c_+^4 c_-^2 + c_+^2 c_-^4) = \frac{1-b^2}{2}, \tag{37}$$

TABLE I. Average gain of information $\overline{K}$ about b given N copies of the state $|\psi\rangle$.

N	$\overline{K}$
1	0
2	0.03751
3	0.08397
4	0.13259
5	0.18059
10	0.39245
20	0.69639
40	1.07422
60	1.32005
80	1.50261

and thus the optimal measurement is composed by 16-dimensional and four-dimensional projectors into the two irreps shown above, the corresponding probabilities being $p(1|b) = (1+b^2)/2$ and $p(2|b) = (1-b^2)/2$. From these, $p(1) = \frac{4}{5}$ and $p(2) = \frac{1}{5}$, and the gain of information is of 0.084 bits.

D. $N>3$

We have applied the same, general procedure to obtain the gain of information up to $N=80$, as reported in Table I and Fig. 1. We observe a logarithmic asymptotic dependence of the gain of information on the number N of available copies of $|\psi\rangle$, which reads

$$\overline{K} \approx 0.44 \log_2 N \tag{38}$$

bits of information on b.

VI. OPTIMAL ESTIMATION OF MIXING

So far we have considered the most general measurement involving the whole space $(\mathcal{H}_2\otimes\mathcal{H}_2)^{\otimes N}$ of N copies of a two-qubit pure state. Now we are going to study optimal *local* measurements for the estimation of its entanglement. Alice will perform a collective measurement over the N cop-

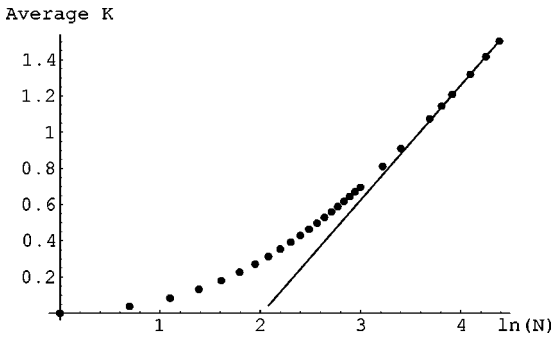

FIG. 1. Average gain of information $\overline{K}$ about b given N copies of the state $|\psi\rangle$. The points represent the results obtained by the described optimal measurement, while the line shows the asymptotic behavior.

ies of the state ρ_A in Eq. (2) at her disposal in order to estimate the parameter b. Consequently, we are also studying optimal strategies for estimating the degree of mixing of a single-qubit mixed state, when N copies are available.

In order to study the latter with more generality we will consider a generic prior distribution $f(b)$ for the degree of mixing while keeping an isotropic distribution in the Bloch vector direction $\hat{a}$ of the unknown mixed state, with

$$\int_{S^2} \frac{d\hat{a}}{4\pi} \int_0^1 db\, f(b) = 1. \tag{39}$$

A general measurement on the local composite system supporting the state $\rho_A^{\otimes N}$ consists of a resolution of the identity in the corresponding Hilbert space $\mathcal{H}_2^{\otimes N}$ by means of positive operators $M^{(k)}$. The gain of information is as in Eq. (9), where now

$$p(k|b) = \mathrm{tr}[M^{(k)} \rho_A^{(N)}(b)], \tag{40}$$

so that we need to compute the effective mixed state

$$\rho_A^{(N)}(b) \equiv \int_{g \in G} dg [D(g) \rho_A(b) D(g)^\dagger]^{\otimes N}, \tag{41}$$

where the integral is performed over the group $G = \mathrm{SU}(2)$ and a single copy of the mixed state

$$\rho_A = U_A \rho_A(b) U_A^\dagger \tag{42}$$

has been expressed, as before, in terms of a reference state $\rho_A(b) \equiv (c_+^2 |+\rangle\langle+| + c_-^2 |-\rangle\langle-|)$ and a unitary transformation U_A. The procedure to be followed is analogous to the previous one, the spectral decomposition of the state (41), allowing us to build the optimal measurement.

The density matrix $\rho_A(b)^{\otimes N}$ can be written—by using a straightforward modification of lemma 2 and the mentioned properties of the Clebsh-Gordan coefficients—in terms of the coupled basis $\{|v_i\rangle_A\}$ as

$$\begin{aligned} \rho_A(b)^{\otimes N} = {} & c_+^{2N} \sum_{i; m_T = N/2} |v_i\rangle\langle v_i|_A \\ & + c_+^{2(N-1)} c_-^2 \sum_{i; m_T = (N/2)-1} |v_i\rangle\langle v_i|_A + \cdots \\ & + c_-^{2N} \sum_{i; m_T = -(N/2)} |v_i\rangle\langle v_i|_A . \end{aligned} \tag{43}$$

Notice that the important role played before by the symmetry between the kets in A and B [cf. Eq. (26)] is now played by the symmetry between the terms in the bra and in the ket. However we see that now there are no cross-terms between inequivalent irreps of SU(2), and that equivalent irreps, such as the $N-1$ copies of the $[(N/2)-1]$ irrep, obtain equal but independent contributions. The space $\mathcal{H}_2^{\otimes N}$, decomposed in terms of irreps of SU(2) is (see also Refs. [6] and [17])

$$\begin{aligned} \mathcal{H}_2^{\otimes N} = {} & \frac{N}{2} \oplus \left(\frac{N}{2} - 1\right) \oplus \cdots \oplus \left(\frac{N}{2} - 1\right) \\ & \oplus \cdots \oplus \frac{N \bmod 2}{2} \oplus \cdots \oplus \frac{N \bmod 2}{2}. \end{aligned} \tag{44}$$

The spectral decomposition of $\rho_A^{(N)}(b)$ is determined by application of the orthogonality lemma. Since equivalent irreps receive always the same contributions in the decomposition (43), the corresponding eigenvalues are equal, so that Eq. (41) reads

$$\begin{aligned} \rho_A^{(N)}(b) = {} & \lambda_1^L(b) I_{N/2} + \lambda_2^L(b) (I_{(N/2)-1} + \cdots + I_{(N/2)-1}) + \cdots \\ & + \lambda_m^L(b) (I_{(N \bmod 2)/2} + \cdots + I_{(N \bmod 2)/2}). \end{aligned} \tag{45}$$

This is, of course, simply what remains from Eq. (28) when Bob's subsystem is traced out, and we have included the whole derivation only for completeness.

Equations (16)–(18) still hold, and therefore the optimal measurement for the degree of mixing b corresponds, for any isotropic distribution, to projections onto each of the subspaces associated with the eigenvalues $\{\lambda_k^L\}$. The gain of information is then given by the right-hand side of Eq. (18). Notice that both the number of outcomes and the corresponding probabilities $p(k|b) = n_k^L \lambda_k^L(b)$ are equal to the ones obtained before for entanglement estimation. In particular, it follows that there is no way to learn about the degree of mixture of an unknown mixed state if only one copy is available.

VII. DISCUSSION AND CONCLUSIONS

In this work we have presented an optimal strategy for the estimation of the entanglement of two-qubit pure states, when N copies are available. Such optimal measurement is also minimal, in the sense that it consists of the minimum number of outcomes, namely, $N/2+1$-$(N+1)/2$ outcomes for the even-odd-N-copy case. Most of the corresponding projectors are of dimension greater than 1, and of course any further decomposition of them can be used in principle to obtain, simultaneously, some additional information about other properties of the unknown state, although our optimal POVM is not compatible with projecting onto states of the form $|\psi_i\rangle^{\otimes N}$ as optimal POVM for state determination are [2–5], and they are thus less powerful for that purpose.

An interesting particular case is when the initial state is a product state, i.e., $b=1$. It can be seen that in this situation we have only an outcome corresponding to the space of maximum spin, since $n_1 \lambda_1(1) = 1$. Therefore, if the outcome k, with $k>1$, is obtained, we can be assured that the state is entangled.

In Sec. VI we were also concerned with the optimal estimation of the degree of mixing. Our optimal measurement, again minimal, can be used, for instance, to quantify the degree of purity of states created by a preparation device whose polarization direction we ignore. Our strategy is actually complementary to the one aiming at optimally revealing the direction of polarization of the state [1]. As a matter of

fact, the optimal POVM we obtained is just a coarse graining of the one obtained in Ref. [6] for optimal estimation of mixed states, which turned out also to reach the optimal standards of direction estimation obtained in Ref. [1]. Consequently, the direction and modulus of the Bloch vector of an unknown mixed state can be optimally estimated simultaneously. Note that this is not a frequent situation. If, instead, we would like to estimate the x, y, and z components of the Bloch vector independently, we would have obtained incompatible optimal strategies (consider, e.g., the $N=1$ case, where an optimal measurement for the component of the Bloch vector along direction $\hat{n}$ consists of a two outcome measurement projecting on that direction).

Finally, we can argue that *bilocal* measurements, either *uncorrelated* or *classically correlated*, do not imply any improvement of the simpler, *local* ones for entanglement estimation. Once we obtain an outcome from Alice's local measurement, we can compute Bob's effective state, and it is clear from Eq. (28) that his outcome will be the same as Alice's, so that no extra information on b will be obtained. We have also seen that the optimal global measurement on $|\psi\rangle^{\otimes N}$ is perfectly mimicked by a local one on $\rho_A^{\otimes N}$ (or $\rho_B^{\otimes N}$), so that actually all four classes of measurements considered in Sec. III A are equivalent. In fact, with hindsight, one can understand this result: local measurements are performed on the reduced density matrix, which is obtained by a partial trace over the other subsystem. This operation erases the information contained in the parameters α and $\hat{b}$ of Eq. (1). On the other hand, the global measurement can be interpreted as being performed on the effective density matrix of Eq. (14), where the same parameters have been integrated over. This operation erases the information contained in them as well.

It would be challenging to address the same question for bipartite mixed states, and for systems shared by more than two parties. Note that in none of these cases is optimal estimation of the nonlocal parameters possible by means of local (or even uncorrelated bilocal) measuring strategies. This is the case for mixed states because any given reduced density matrix ρ_A may correspond to infinitely many mixed states ρ, with different degrees of entanglement, so that not even in the limit $N\to\infty$ can the entanglement of ρ be properly inferred from $\rho_A^{\otimes N}$. The mere existence of hidden nonlocal parameters [18]—that is, of entanglement parameters that are erased during the partial trace operation—also prevents uncorrelated local strategies from being optimal for estimation of pure-state tripartite entanglement. To conclude, two-qubit pure-state entanglement, a quantum nonlocal property, can be optimally estimated by means of local, but collective, measurements.

ACKNOWLEDGMENTS

We thank Susana Huelga for reactivating our interest in this problem and for interesting discussions, and J. I. Latorre for helping us with the computation of the values of Fig. 1. G.V. acknowledges CIRIT Grant No. 1997FI-00068 PG. A. A. acknowledges a grant from MEC. Financial support from CICYT Contract No. AEN98-0431 and CIRIT Contract No. 1998SGR-00026 are also aknowledged. This work was partially elaborated during the ''Complexity, Computation and the Physics of Information'' workshop of the Isaac Newton Institute. The authors thank the Institute and the European Science Foundation for support during this period.

[1] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[2] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
[3] R. Derka, V. Buzek, and A. K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998); e-print quant-ph/9707028.
[4] J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. Lett. **81**, 1351 (1998); e-print quant-ph/9803066.
[5] A. Acín, J. I. Latorre, and P. Pascual, e-print quant-ph/9904056 [Phys. Rev. A (to be published)].
[6] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. A **60**, 126 (1999); e-print quant-ph/9812068.
[7] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[8] H.-K. Lo and S. Popescu, e-print quant-ph/9707038.
[9] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[10] G. Vidal, Phys. Rev. Lett. **83**, 1046 (1999).
[11] D. Jonathan and M. B. Plenio, Phys. Rev. Lett. **83**, 1455 (1999).
[12] The problem of optimally estimating the entanglement of two-qubit pure states was recently analyzed by J. M. G. Sancho and S. F. Huelga, preprint, quant-ph/9910041. In their work they considered strategies for the N-copy case that measured only on one copy of the unknown state at a time. Their work and ours can be thus regarded as complementary.
[13] M. J. W. Hall, Phys. Lett. A **242**, 123 (1998); e-print quant-ph/9802052.
[14] S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).
[15] A. Hobson, J. Stat. Phys. **1**, 383 (1969).
[16] R. Tarrach and G. Vidal, e-print quant-ph/9907098 [Phys. Rev. A (to be published)].
[17] J. I. Cirac, A. K. Ekert, and C. Macchiavello, Phys. Rev. Lett. **82**, 4344 (1999); e-print quant-ph/9812075.
[18] J. Kempe, Phys. Rev. A **A60**, 910 (1999).
[19] More specifically, the most informative measurements presented in this work are also optimal with respect to a fidelity-guided scheme if the quality of the guesses is evaluated through any *concave* fidelity function $F(b-b_k)$—where b is the unknown parameter and b_k is the guess made after outcome k—that reasonably takes its maximum for $b_k=b$, i.e., $F((x+x')/2)\geq[F(x)+F(x')]/2$ and $F(0)\geq F(x\in[-1,1])$.