Local Discrimination of Mixed States

J. Calsamiglia, J. I. de Vicente, R. Muñoz-Tapia, and E. Bagan
Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain
(Received 29 April 2010; published 18 August 2010)

We provide rigorous, efficiently computable and tight bounds on the average error probability of multiple-copy discrimination between qubit mixed states by local operations assisted with classical communication (LOCC). In contrast with the pure-state case, these experimentally feasible protocols perform strictly worse than the general collective ones. Our numerical results indicate that the gap between LOCC and collective error rates persists in the asymptotic limit. In order for LOCC and collective protocols to achieve the same accuracy, the former can require up to twice the number of copies of the latter. Our techniques can be used to bound the power of LOCC strategies in other similar settings, which is still one of the most elusive questions in quantum communication.

DOI: 10.1103/PhysRevLett.105.080504 PACS numbers: 03.67.Hk, 03.65.Ta, 03.67.Mn

Quantum communication and computation tasks involve, broadly speaking, transforming an input state and reading the corresponding output state. One of the most prominent features of quantum mechanics is that, hard as one may try, the readout will be unavoidably imperfect unless the various output states are orthogonal. This has both fundamental and practical implications that lie at the heart of quantum mechanics and its applications. The most elementary scenario where this deep fact manifests itself is what is known as quantum state discrimination. In its simplest form, given one of two equiprobable sources that provide N independent copies of a state ρ_0 or ρ_1 (i.e., $\sigma_0 = \rho_0^{\otimes N}$ or $\sigma_1 = \rho_1^{\otimes N}$), we ask ourselves what is on average the minimum error probability P_e of making a conclusive guess of the identity of the state. State discrimination is an essential primitive for many quantum information tasks, such as quantum cryptography [1], or even quantum algorithms [2]. Moreover, with the remarkable experimental advances in the preparation and measurement of quantum states, it becomes essential to assess the performance of state discrimination protocols.

The error probability for two arbitrary states was given already decades ago by Helmstrom [3], who provided a formal expression for P_e in terms of the trace distance between the two density matrices σ_0 and σ_1. An explicit simple expression can be given for single qubit states ($N = 1$), and only very recently has the asymptotic error rate when $N \to \infty$ for qudits been found through the quantum Chernoff bound [4,5]. For a finite (moderately large) number of qubits, the permutation invariance of the multiple-copy states $\{\sigma_a\}$, of size $\sim 2^N$, enables us to write them in block-diagonal form. Thus, we can numerically compute the error probability in terms of the trace distance between small-sized ($\sim N$) blocks, and the difficulty of the problem (required memory size) becomes polynomial in N despite the exponentially growing dimension of the Hilbert space we are dealing with. This collection of results constitutes a fairly complete theory regarding the optimal (unconstrained- or collective-measurement) multiple-copy discrimination of quantum states.

The picture changes completely when it comes to discrimination of states using local operations (on individual copies) and classical communication (LOCC) instead of general (collective) measurements. This scenario is interesting from the fundamental point of view, as it sheds light on the role of quantum correlations in quantum information tasks, but it is of paramount interest from a practical point of view since it puts under scrutiny the attainability of previous bounds in implementations, where collective measurements are usually unfeasible. For pure states it has been shown [6] that the minimum collective error probability can be attained by a LOCC one-way adaptive protocol consisting in performing a different von Neumann measurement on each copy, where each measurement is chosen according to the outcome(s) of the previous one(s). Furthermore, it is shown that an even simpler fixed local strategy, where the *very same* particular von Neumann measurement is repeated on every copy, though not optimal for finite N, does provide the collective asymptotic error rate as $N \to \infty$. The latter also holds when only one of the states is pure [5].

For mixed states, recently, Hayashi [7] has proven that, as far as the asymptotic error rate is concerned, one-way adaptive strategies are not advantageous over fixed strategies, which do not make use of classical communication between measurements. In the last months Higgins *et al.* [8] have studied theoretically and experimentally the performance of various local strategies including adaptive von Neumann measurements. These adaptive strategies, which are optimized by using dynamic programming techniques [9], outperform the others under consideration, but there still remain the fundamental open questions of whether or not these strategies are the best one can achieve by LOCC (which include generalized local measurements and unlimited communication rounds) and whether or not those can attain the collective bounds.

In this Letter, we show how to efficiently compute bounds on LOCC protocols. This enables us to compare such protocols to their collective counterparts and benchmark the performance of particular state discrimination strategies or experiments. Before presenting our analysis, let us note that there has been some recent interest in the related problem of LOCC discrimination of bipartite states (see [10] and references therein), where each of the parties has joint access to a share of all copies. We find that in our scenario, although the states are disentangled, quantum correlations play an essential role, which is, in this sense, a manifestation of nonlocality without entanglement.

At this point we need to go into a more technical discussion. The error probability for the case under consideration can be written as [3]

$$P_e = \tfrac{1}{2}\{1 + \min_{0\le E\le \mathbb{1}} \mathrm{tr}[(\sigma_0 - \sigma_1)E]\}. \tag{1}$$

The matrix E, together with $\mathbb{1} - E$, are the elements of the positive operator valued measure (POVM) that represents mathematically the measuring protocol. Note that they can be taken to be symmetric under permutations of the individual systems (invariant under the action of the symmetric group S_N), because so are σ_0 and σ_1, and can thus be put in block-diagonal form.

We realize that Eq. (1) defines a semidefinite programming (SDP) problem, for which very efficient numerical algorithms have been recently developed [11]. Bounds on LOCC strategies could be obtained if in addition to the positivity constraint $0 \le E \le \mathbb{1}$ on $\{E, \mathbb{1} - E\}$ one would further impose, e.g., positive partial transposition (PPT), i.e., $0 \le E^\Gamma \le \mathbb{1}$ (recall that this condition defines the set of PPT-preserving operations [12], which includes the set of separable operations and, in turn, LOCC operations). The difficulty here is that partial transposition (PT), which we denote by the superscript Γ, does *not* preserve the block-diagonal form of E (just note that Γ breaks permutation invariance). Hence, the size of the matrices one needs to deal with remains $\sim 2^N$, and the error probability cannot be computed but for very small values of N. We next show how to bypass this problem.

The main observation is that for (two) qubit-state discrimination the POVM elements can always be chosen to be PT-invariant, $E^\Gamma = E$ (for any bipartition of the N qubits). This follows from the fact that, with the appropriate choice of basis, one can take the two states ρ_0 and ρ_1 to be real and, thus, symmetric: $\rho_a^T = \rho_a$. Obviously, this implies PT invariance, $\sigma_a^\Gamma = \sigma_a$, for any bipartition. Hence, for a given E that satisfies PPT, the PT-invariant operator $E' = (E + E^\Gamma)/2$, which also satisfies $0 \le E' = E'^\Gamma \le \mathbb{1}$, provides the exact same error probability P_e. Therefore, we can restrict ourselves to PT-invariant operators without any loss of generality. Since such E can be put in block-diagonal form, the sizes of the matrices we have to deal with grow only quadratically in N.

Applying the procedure sketched above requires, nonetheless, finding an efficient parametrization of PT-invariant matrices in block-diagonal form. The first step towards this end is identifying the independent matrix elements in the computational basis $|i_1 i_2 \dots i_N\rangle$. For N qubits, the operator E can be written as

$$E = \sum_{\{i_p\}} \sum_{\{i'_s\}} E^{i'_1 i'_2 \dots i'_N}_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle\langle i'_1 i'_2 \dots i'_N|, \tag{2}$$

where all i_p and i'_s ($p, s = 1, 2 \dots N$) are either 0 or 1 and each sum runs over the 2^N possible binary lists (numbers) of N digits. By invoking permutation invariance and Hermiticity, the independent components of E can be chosen to be $E^{11\dots10\dots011\dots10\dots00}_{11\dots11\dots100\dots00\dots00} \equiv \tilde{\mathcal{E}}^{Q,Q'}_R$, where the first R digits in the subscript are ones and the Q (Q') first digits on top of them (the remaining $N - R$ zeros) are also ones. We further impose that $R \le Q + Q'$ and note that $Q \le R$. We next use PT invariance to exchange the last $R - Q$ ones in the subscript with the zeros on top of them by raising (lowering) the corresponding i_p (i'_p). This proves that the PT-invariant matrices we are dealing with have $(N + 1) \times (N + 2)/2$ independent components:

$$E^{11\dots11\dots10\dots00}_{11\dots10\dots00\dots00} \equiv E^q_r, \qquad r \le q, \tag{3}$$

where q and r are the number of ones in the superscript and the subscript, respectively.

We next wish to write E in block-diagonal form. To this end, we map each qubit to a spin 1/2, $|i_p\rangle \to |m_p\rangle = |(-1)^{i_p}/2\rangle$, where m_p is the magnetic number of the pth spin, and change from the uncoupled (computational) basis to the total spin eigenbasis $\{|j, m\rangle\}^j_{m=-j}$, which span the irreducible representations (irreps) of $SU(2)$. In this basis E becomes block-diagonal, and the matrix elements of each block $E^{(j)}$, i.e., $[E^{(j)}]^{m'}_m = \langle j, m|E|j, m'\rangle$, are expressed as linear combinations of the independent parameters $\mathcal{E}^q_r$. We write $[E^{(j)}]^{m'}_m = \sum_{r,q} [\mathcal{M}^{(j)}]^{m'r}_{mq} \mathcal{E}^q_r$, which facilitates the SDP implementation of the optimization. Some comments are in order. For given j and m, the state $|j, m\rangle$ is degenerate. Note, though, that all blocks with the same j are identical, as E is fully symmetric. Therefore, the contribution of $E^{(j)}$ to the error probability will have to be multiplied by the corresponding degeneracy $n_j = \binom{N}{N/2-j}(2j + 1)/(N/2 + j + 1)$. The matrices $\mathcal{M}^{(j)}$ turn out to be (see Ref. [13])

$$[\mathcal{M}^{(j)}]^{m'r}_{mq} = \sum_k [\Delta^{(j)}_k]^{m'}_m \binom{\frac{N}{2} - j}{\frac{q-r+m'-m}{2} - k} \times (-1)^{(q-r+m'-m/2)-k} \delta_{q+r,N-m-m'}, \tag{4}$$

where we have defined

$$[\Delta^{(j)}_k]^{m'}_m = \frac{\sqrt{(j-m)!(j+m)!(j-m')!(j+m')!}}{(j-m-k)!(j+m'-k)!(m-m'+k)!k!} \tag{5}$$

and the sums run over all integer values for which the

factorials make sense. Note that the Wigner d matrices also involve these coefficients [13].

Now that we have a minimal parameterization of the operators that are invariant under permutations and partial transpositions, we can compute the error probability by the following SDP instance:

$$P_e^{\mathrm{PPT}} = \frac{1}{2}\left\{1 + \min_{\{E_r^a\}} \sum_j n_j \mathrm{tr}[(\sigma_0^{(j)} - \sigma_1^{(j)})E^{(j)}]\right\}. \quad (6)$$

Here the minimization is constrained by $0 \leq E^{(j)} \leq \mathbb{1}$, for all possible values of the total spin j, and the matrix blocks $\sigma_a^{(j)}$ are computed to be (see Ref. [13])

$$[\sigma_a^{(j)}]_m^{m'} = \frac{(1-r_a^2)^{(N/2)-j}}{2^N} \sum_k [\Delta_k^{(j)}]_m^{m'} \left[(-1)^a r_a \sin\frac{\theta}{2}\right]^{m-m'+2k} \times \left(1 + r_a \cos\frac{\theta}{2}\right)^{j+m'-k} \left(1 - r_a \cos\frac{\theta}{2}\right)^{j-m-k}, \quad (7)$$

where r_a, often referred to as purity or degree of mixedness, is the length of the Bloch vector $\vec{r}_a$ of the single qubit state ρ_a, and θ is the relative angle between $\vec{r}_0$ and $\vec{r}_1$.

As argued above, P_e^{PPT} provides a lower bound to the error probability attainable by the most general LOCC strategy, which includes weak generalized local measurements interlaced with an unlimited number of classical communication rounds. In what follows we will compare this bound to the error probability of the optimal collective strategy P_e^{col} and to that of the following two LOCC strategies: (i) *repeated* strategy, where the very same two-outcome measurement is performed on every copy. The error probability P_e^{rep} is obtained by minimizing over the azimuthal angle Θ that specifies the unit Bloch vector of the two measurement projectors (as this vector can be chosen to lie in the plane spanned by $\vec{r}_0$ and $\vec{r}_1$). The corresponding asymptotic error rate can also be obtained from the classical Chernoff bound [5]. (ii) *Adaptive* strategy, where copies are measured sequentially and the choice of the azimuthal angle Θ_s, corresponding to the projective measurement on the sth copy, depends on the outcomes obtained upon measuring the preceding $s-1$ copies; i.e., it makes use of one-way communication. If the number of available copies N is known beforehand, it is possible to find the optimal adaptive strategy very efficiently by using dynamic programming [8], as detailed in Ref. [13].

Figure 1 shows the error probability of discrimination between two states with equal purity ($r_0 = r_1 = 0.8$) and $\theta = \pi/2$ for the various strategies and bounds discussed above. We notice that there is a significant gap between the collective strategy and the LOCC lower bound (P_e^{PPT}). In addition, we note that the error probability for repeated and adaptive strategies falls almost on top of the LOCC lower bound. This shows that P_e^{PPT} is a very tight bound and that it can be taken as a good estimate of the minimal LOCC error probability for most practical purposes. The figure clearly shows that, as expected, the error probability falls exponentially with N: $P_e \sim e^{-CN}$. By fitting the data in the figure to an error rate of the form [14] $C = C_0 + C_1 \log N/N + C_2/N$ for $25 \leq N \leq 35$, we can obtain its asymptotic value C_0 for the various strategies. For the collective and repeated strategies the results agree up to the third significant digit with the analytical results provided by the quantum and classical Chernoff bounds, respectively [5]. More interestingly, within numerical accuracy the fits indicate that the gap between collective and LOCC error rates persists in the asymptotic limit. The results are also consistent with a convergence of the asymptotic error rates for the two LOCC strategies and that of the LOCC lower bound (P_e^{PPT}), although here, due to the already small differences, it is harder to exclude the existence of a (tiny) nonvanishing gap.

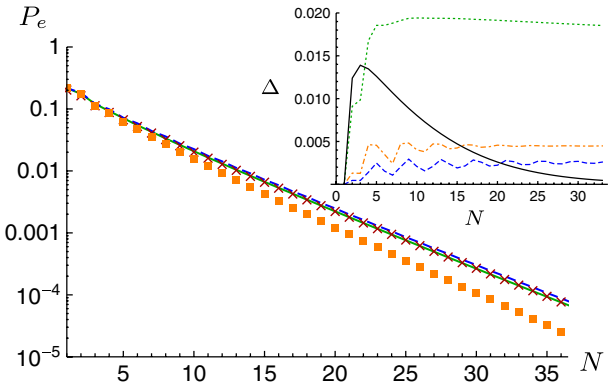

FIG. 1 (color online). LOCC lower bound P_e^{PPT} (solid line) and error probability for collective (squares), adaptive (crosses), and repeated (dashed line) strategies for $r_0 = r_1 = 0.8$ and $\theta = \pi/2$. Inset: Gap vs N for $\theta = \pi/2$, $r_0 = 0.8$ and $r_1 = 1$ (solid line), $r_1 = 0.9$ (dotted line), $r_1 = 0.5$ (dotted-dashed line), and $r_1 = 0.4$ (dashed line).

In the inset in Fig. 1, we plot the gap between the collective error rate and that of the LOCC lower bound (P_e^{PPT}), i.e., $\Delta = C^{\mathrm{col}} - C^{\mathrm{PPT}} = -\frac{1}{N}\log\frac{P_e^{\mathrm{col}}}{P_e^{\mathrm{PPT}}}$. We notice that the gap reaches its asymptotic value already for a small number of copies. This is so for all values of r_1 but for $r_1 = 1$ (solid line), for which Δ, after growing to a maximum at $N \approx 4$, decreases to zero as it should, according to Ref. [5]. There, as mentioned at the beginning of this Letter, it is shown that when one of the states, say, ρ_1, is pure, the collective error rate is asymptotically attainable by a repeated strategy or, more precisely, by one consisting in performing the measurement defined by $E = \rho_1$ on each copy. The unknown state is claimed to be ρ_1 if the N outcomes of the measurements correspond to E (none to $\mathbb{1} - E$), and it is claimed to be ρ_0 otherwise (*unanimity vote*). The asymptotic error rate of this strategy attains the upper bound $C_0 \leq -\log F(\rho_0, \rho_1)$, where $F(\rho_0, \rho_1) = (\mathrm{tr}|\sqrt{\rho_0}\sqrt{\rho_1}|)^2$ is the fidelity [5]. For the collective strategy it also holds that $-(1/2)\log F(\rho_0, \rho_1) \leq C_0^{\mathrm{col}}$.

Figure 2 shows the error rate $C = -(1/N)\log P_e$, for two equally mixed states, $\theta = \pi/2$ and $N = 25$, as one varies their degree of mixedness r. We identify four pa-

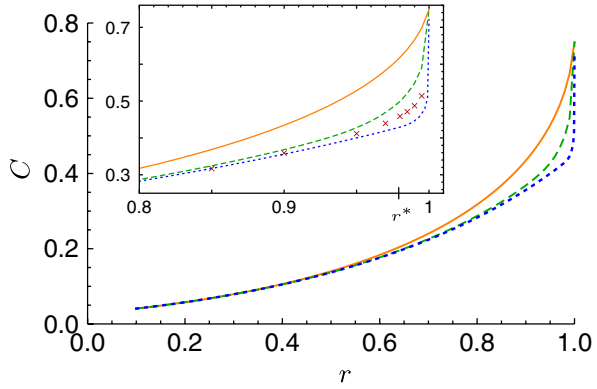

FIG. 2 (color online). Error rate vs $r = r_0 = r_1$ for $N = 25$ and $\theta = \pi/2$, and for collective (solid line), LOCC (PPT bound) (dashed line), and repeated (dotted line) strategies. The inset also shows the rate for the adaptive strategy (crosses).

rameter regions in this plot: (i) For very mixed states ($r \lesssim 0.5$), collective and repeated local strategies have essentially the same performance. (ii) As the purity increases ($0.5 \lesssim r \lesssim 0.8$), the collective strategy starts to outperform the LOCC one, but the repeated strategy nearly attains the LOCC lower bound P_e^{PPT} (upper bound on the error rate). The measurements of this repeated strategy have $\Theta = \pi/2$, which means that their Bloch vector is proportional to $\vec{r}_0 - \vec{r}_1$ [15]. The decision is taken by "*majority vote*"; i.e., the most frequent outcome determines the decision. (iii) At very high purities ($0.8 \lesssim r \leq r^*$), the three LOCC curves start to split. (iv) At purities larger than a critical one, r^*, all LOCC curves start to rapidly converge to the collective one. At $r = r^*$ the measurement angle of the repeated strategy starts to change from $\Theta = \pi/2$ towards $\Theta = \pi/4$ as $r \to 1$ (for arbitrary θ, one has $\Theta \to \theta/2$; i.e., the Bloch vector of the measurements goes to either $\vec{r}_0$ or $\vec{r}_1$) and the decision rule gradually shifts from a majority vote to the unanimity vote described above. In the asymptotic limit $N \to \infty$, r^* can be computed to arbitrary accuracy as a solution of a transcendental equation; for $\theta = \pi/2$ one has $r^* \approx .9819$. For $r < r^*$ the error rate of the repeated strategy, C_0^{rep}, saturates the fidelity lower bound introduced earlier: $C_0^{\mathrm{rep}} = -(1/2)\log F(\rho_0, \rho_1)$.

As mentioned above, for asymptotically large N repeated and adaptive error rates coincide [7]. The tiny gap between the corresponding curves in Fig. 2 is explained by the relatively small number of copies ($N = 25$) used in the plot. It is plausible, and consistent with our data, that also the gap between the LOCC error rate bound and C_0^{rep} vanishes as $N \to \infty$.

Note that the smaller the error rate, the more copies we need to achieve the same error probability. The ratio $f = C^{\mathrm{col}}/C^{\mathrm{PPT}}$ tells us that we need fN copies in order for the best LOCC strategy to discriminate with the accuracy of the collective one. The general features of f as a function of r can be immediately grasped from Fig. 2. For very mixed states, LOCC and collective strategies require a similar number of copies ($f \approx 1$) to discriminate with the same error probability, but as the states become more pure the LOCC strategies demand an increasing number of copies ($f \leq 2$, where equality corresponds to the ratio of the two fidelity bounds). In the limit of very pure states ($r \gtrsim r^*$), the factor f drops back down again to one. The factor f attains its maximum value when the states are nearly parallel ($\theta \sim 0$) and nearly pure (but strictly mixed).

In summary, we have lower-bounded the error probability of LOCC discrimination between two qubit mixed states. Our results indicate an error rate gap between the best LOCC and collective discrimination protocols that persists as the number of copies goes to infinity. This gap takes its largest value in the region of nearly pure, but strictly mixed, states. Excluding this region, there are no significant differences in performance between the simplest (repeated) and optimal LOCC strategies.

We thank E. Ronco and G. Via for their contributions in the earlier stages of this work. We acknowledge financial support from the Spanish MICINN, through the Ramón y Cajal program (J. C.), FIS2008-01236, and QOIT (CONSOLIDER2006-00019), from the Generalitat de Catalunya CIRIT, 2009SGR-0985, and from Alianza 4 Universidades (J. I. d. V.).

[1] A. Acin *et al.*, Phys. Rev. A **73**, 012327 (2006); see also the review N. Gisin *et al.*, Rev. Mod. Phys. **74**, 145 (2002).
[2] D. Bacon *et al.*, in *Proceedings of the 46th IEEE FOCS 2005* (IEEE, New York, 2005), pp. 469–478.
[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[4] K. Audenaert *et al.*, Phys. Rev. Lett. **98**, 160501 (2007); M. Nussbaum and A. Szkola, Ann. Stat. **37**, 1040 (2009).
[5] J. Calsamiglia *et al.*, Phys. Rev. A **77**, 032311 (2008).
[6] D. Brody and B. Meister, Phys. Rev. Lett. **76**, 1 (1996); A. Acin *et al.*, Phys. Rev. A **71**, 032338 (2005).
[7] M. Hayashi, IEEE Trans. Inf. Theory **55**, 3807 (2009).
[8] B. L. Higgins *et al.*, Phys. Rev. Lett. **103**, 220503 (2009).
[9] G. L. Nemhauser, *Introduction to Dynamic Programming* (Wiley, New York, 1966).
[10] M. Hayashi *et al.*, Phys. Rev. Lett. **96**, 040501 (2006).
[11] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996); J. Löfberg, Yalmip, http://control.ee.ethz.ch/~joloef/yalmip.php.
[12] E. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001)
[13] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.105.080504 for the derivation of (4), (5), and (7) and the use of dynamic programming for adaptive discrimination.
[14] These terms appear in the classical (commuting) and pure-state limits and provide accurate fits in general cases.
[15] This holds for N even, but it is only an approximation for N odd, becoming exact as $N \to \infty$.