PHYSICAL REVIEW A **71**, 032338 (2005)

Multiple-copy two-state discrimination with individual measurements

A. Acín,[1] E. Bagan,[2] M. Baig,[2] Ll. Masanes,[3,4] and R. Muñoz-Tapia[2]
[1]*ICFO-Institut de Ciències Fotòniques, Jordi Girona 29, Edifici Nexus II, 08034 Barcelona, Spain*
[2]*Grup de Física Teòrica, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*
[3]*School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom*
[4]*Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona, 08028 Barcelona, Spain*
(Received 13 October 2004; published 23 March 2005)

We address the problem of nonorthogonal two-state discrimination when multiple copies of the unknown state are available. We give the optimal strategy when only fixed individual measurements are allowed and show that its error probability saturates the collective (lower) bound asymptotically. We also give the optimal strategy when adaptivity of individual von Neumann measurements is allowed (which requires classical communication) and show that the corresponding error probability is exactly equal to the collective one for any number of copies. We show that this strategy can be regarded as Bayesian updating.

DOI: 10.1103/PhysRevA.71.032338 PACS number(s): 03.67.Hk, 03.65.Ta

I. INTRODUCTION

Measurement is a central tenet of quantum mechanics. As for any sensible theory of nature, it links abstract ideas to reality and makes mathematical concepts truly physical. In contrast to classical measurements, which (ideally) have no demolishing effect whatsoever, in the quantum realm any attempt to acquire information from a system alters it to a degree proportional to the gain of information. Moreover, this gain is limited [1]. Given a single copy of an unknown quantum state it is usually impossible to determine it by performing any conceivable measurement. Nevertheless, if an increasing number of copies of such a state is available, our knowledge of the state can also increase by the use of various measurement strategies, and complete determination can be achieved in the asymptotic limit when the number of copies goes to infinity.

Measurement strategies involving multiple copies of a quantum state fall into two categories: collective and individual (local), depending on whether a single measurement is performed on all the copies as a whole or the strategy consists of individual measurements, each of them performed separately on a single copy. Since the pioneering work of Helstrom [2] and Peres and Wootters [3], it has been repeatedly shown that collective strategies outperform individual ones. This should not come as a surprise, since the latter can be viewed as a subset of the former, which are completely general and unconstrained. Collective measurements, however, are difficult to implement experimentally, and a great deal of effort goes into designing optimal strategies involving only individual measurements. Common examples are quantum tomography [4] and (local) adaptive strategies [5] (where the choice of each individual measurement is based on the outcomes of the previous), the two of them in the context of quantum state estimation. The state of the art of these approaches can only compete with collective strategies in the asymptotic limit.

Many practical applications, however, do not require a full determination of a state. For instance, to asses the security of a key distribution protocol in quantum cryptography [6], one gives full advantage to Eve, the eavesdropper. Hence, one usually assumes she knows the set of possible states that will be used in a secret transmission (e.g., in the B92 protocol [7] this set consists of two nonorthogonal states), and her task is to *discriminate* [8] among them. She can follow two different approaches: use a strategy based on quantum hypothesis testing [2] (unconclusive discrimination), which gives the lowest probability of error, or do unambiguous (or conclusive) discrimination [9]—namely, adopt a strategy that does not tolerate errors.

When the number of copies is greater than 1 (as is the case of a noncompletely attenuated laser pulse which may be split in several identical single-photon states), the discussion above concerning individual versus collective strategies becomes again an issue. In this paper we focus our attention on this situation. To be more concrete, we will consider a hypothesis-testing approach to (nonorthogonal) two-state discrimination under the assumption that we have N identical copies of the transmitted quantum state. We will find the best adaptive strategy—i.e., a particular case of strategies that use local operations and classical communication (LOCC), and we will show that it is optimal regardless of the number of copies, in the sense that its error probability and that of the optimal collective strategy are exactly the same for any N. A similar result was obtained by Brody and Meister [10] for Bayesian updating. Our result could be seen as its extension to general adaptive strategies. However, we will prove the remarkable result that the whole class of adaptive strategies has actually a single element: Bayesian updating. Related work can be found in [11], where the particular case $N=2$ is discussed using the information gain instead of the error probability.

If classical communication is not allowed, we show that optimality holds asymptotically for the fixed measurement strategy named the *unanimity vote*, which we also present here.

II. PRELIMINARIES

We will start by reviewing some known facts about two-state discrimination, including a few technical details, which will help us to introduce the notation.

A. One copy

By choosing the appropriate orthonormal basis, any two states $|\psi_0\rangle, |\psi_1\rangle$ (which we will assume to be neither orthogonal nor parallel) can always be written as

$$|\psi_a\rangle = \cos\,\theta|x\rangle + (-1)^a \sin\,\theta|y\rangle, \quad a=0,1, \tag{1}$$

regardless of the dimension of the Hilbert space $\mathcal{H}$ they belong to, where the unit vectors $|x\rangle, |y\rangle$ are the elements of the basis that span the plane $\mathcal{P}$ formed by $|\psi_0\rangle, |\psi_1\rangle$. Now, we ask ourselves what the best measurement for discriminating between $|\psi_0\rangle$ and $|\psi_1\rangle$ is. It can be defined in terms of two orthonormal vectors $\{|\omega_1(0)\rangle, |\omega_1(1)\rangle\}$, which also belong to $\mathcal{P}$, and thus can be written as

$$|\omega_1(a)\rangle = \cos\left(\phi_0 - a\frac{\pi}{2}\right)|x\rangle + \sin\left(\phi_0 - a\frac{\pi}{2}\right)|y\rangle. \tag{2}$$

In our approach, by "best measurement" we mean the measurement that maximizes the probability of discrimination, $P_1 = \Sigma_{a=0}^1 q_a p(\boldsymbol{a}|a) = \Sigma_{a=0}^1 p(\boldsymbol{a}, a)$ (or equivalently the one that minimizes the error probability $\overline{P_1} = 1 - P_1$). Here q_a is the prior probability of $|\psi_a\rangle$ being (secretly) transmitted, $p(\mathbf{0}|0)$ and $p(\mathbf{1}|1)$ are the conditional probabilities of obtaining the outcome 0 or 1 given that the unknown state is $|\psi_0\rangle$ or $|\psi_1\rangle$, respectively, and $p(\mathbf{0},0)$ and $p(\mathbf{1},1)$ are the corresponding joint probabilities. The subindex 1 in $|\omega_1(a)\rangle$ and in the probability of discrimination and error emphasizes that so far we are dealing with just one copy of the unknown state. Throughout this paper, boldfaced random variables will denote the outcomes of our measurement; thus, e.g., $p(1|\mathbf{0})$ is the (*a posteriori*) probability of the transmitted state being $|\psi_1\rangle$ given that the outcome of our measurement is 0. Using elementary quantum mechanics, the conditional probabilities $p(\boldsymbol{a}|b)$ can be computed to be $p(\boldsymbol{a}|b) = |\langle\omega_1(a)|\psi_b\rangle|^2 = \cos^2[\phi_0 - a\pi/2 - (-1)^b\theta]$. The optimal measurement and the corresponding probabilities of discrimination and error are given by

$$\cos 2\phi_0 = \frac{q_0 - q_1}{R_0}\cos 2\theta, \quad \sin 2\phi_0 = \frac{q_0 + q1}{R_0}\sin 2\theta, \tag{3}$$

$$P_1 = \frac{1}{2}(1 + R_0), \quad \overline{P_1} = \frac{1}{2}(1 - R_0), \tag{4}$$

where $R_0 = [(q_0 - q_1)^2 + 4q_0q_1\sin^2 2\theta]^{1/2}$. In terms of the overlap between $|\psi_0\rangle$ and $|\psi_1\rangle$, defined as $c \equiv |\langle\psi_0|\psi_1\rangle| = \cos 2\theta$, the factor R_0 can be written as

$$R_0 = \sqrt{1 - 4q_0q_1c^2}. \tag{5}$$

In the simple case where $q_0 = q_1 = 1/2$, we have $\phi = \pi/4$ and, thus, $|\omega(a)\rangle = \{|x\rangle + (-1)^a|y\rangle\}/\sqrt{2}$, as one would expect.

B. Several copies: Collective measurements

Let us next suppose that N copies of either $|\psi_0\rangle$ or $|\psi_1\rangle$ are available to us. In full analogy with Eq. (1) we define

$$|\Psi_a\rangle = |\psi_a\rangle^{\otimes N} = \cos\,\Theta|X\rangle + (-1)^a \sin\,\Theta|Y\rangle, \tag{6}$$

where $|X\rangle$ and $|Y\rangle$ belong to a conveniently chosen basis of $\mathcal{H}^{\otimes N}$. In this situation Eqs. (3) and (4) also hold if we replace θ and c with the corresponding uppercased variables Θ and $C = \cos 2\Theta$. In terms of the new basis $\{|X\rangle, |Y\rangle, \ldots\}$, the vectors $|\Omega(a)\rangle (a=0,1)$, which define the measurement on the N copies in full analogy with $|\omega_1(a)\rangle$, are also given by Eq. (2) (uppercasing ω, x, and y). This defines a *collective* measurement, since in general $|\Omega(a)\rangle$ is not a product state. We obviously have $C = |\langle\Psi_0|\Psi_1\rangle| = |\langle\psi_0|\psi_1\rangle^N| = c^N$ and thus conclude that the error probability for this optimal collective measurement is [2]

$$\overline{P_N^{\mathrm{col}}} = \frac{1 - \sqrt{1 - 4q_0q_1c^{2N}}}{2}. \tag{7}$$

Since $c < 1$, in the large-N limit we note that

$$\overline{P_N^{\mathrm{col}}} \simeq q_0q_1c^{2N}. \tag{8}$$

III. INDIVIDUAL MEASUREMENTS

A. Fixed measurements

If we are only allowed to perform the same individual measurement on each of our N copies, one could expect that the lowest probability of error we can achieve is $\overline{P_N^{\mathrm{ind}}} \simeq \eta c^N$, where the constant η is not relevant for the discussion here. This belief may stem from the widespread use of the statistical overlap as a measure of distinguishability; from a statistical analysis of the problem at hand, one concludes that the probability of error is bounded by λ^N, where λ depends on the specific individual measurement we are performing. The statistical overlap is a particularly convenient choice of λ (see below). Optimizing over all possible measurements one finds that $\lambda = c$ for two pure states. This bound is attained by a majority-vote strategy: we perform the best individual measurement, given by Eqs. (2) and (3) on each copy and get N_a times the outcome a. Once the measurement process is complete, we decide in favor of the state $|\psi_a\rangle$ whose corresponding N_a is greatest.

However, there exist tighter bounds for the exponential decrease of the probability of error. The best one is known as the Chernoff bound [12], which for the problem at hand is given by $\lambda = \min_\alpha \Sigma_b p(\boldsymbol{b}|0)^\alpha p(\boldsymbol{b}|1)^{1-\alpha}$, where $0 \leqslant \alpha \leqslant 1$ (the statistical overlap is a particular simplification of this expression obtained by setting $\alpha = 1/2$). We now note that if we assume $q_0 > q_1$, with the choice $|\tilde{\omega}(0)\rangle = |\psi_0\rangle$ and $|\tilde{\omega}(1)\rangle = |\psi_0^\perp\rangle$ for our measurement we have $p(\mathbf{0}|0) = 1, p(\mathbf{1}|0) = 0$, and $p(\mathbf{0}|1) = c^2$ and $\alpha = 0$ gives the absolute minimum (over all measurements and over all values of α) of the sum over b above. Thus, $\overline{P_N^{\mathrm{ind}}} \simeq \eta c^{2N}$, as for collective measurements.

There is a simple strategy that saturates the Chernoff bound: the unanimity vote. Let ourselves perform the measurement defined by $\{|\tilde{\omega}(a)\rangle\}$ on each of our N copies. If we always obtain the outcome 0 ($N_0 = N$), we claim that the unknown state is $|\psi_0\rangle$. However, if we obtain the outcome 1 once or more than once, we decide in favor of $|\psi_1\rangle$.

The exact probability of error is straightforward to compute as follows. Let us assume again $q_0 > q_1$. If the unknown state were $|\psi_0\rangle$, we would make no error. If the unknown state were $|\psi_1\rangle$ (it happens with probability q_1), we would give the wrong answer only if $N_0 = N$, which happens with probability c^{2N}. Hence, the probability of error would be $q_1 c^{2N}$. If $q_1 > q_0$, we just exchange the subscripts 0 and 1 everywhere. The error probability is then

$$\overline{P}_N^{\text{ind}} = \min(q_0, q_1) c^{2N}. \tag{9}$$

We note that asymptotically $\overline{P}_N^{\text{ind}}$ may be larger than $\overline{P}_N^{\text{col}}$ only because of the prefactor $\min(q_0, q_1) \geqslant q_0 q_1$, which is not important in most situations. This result has application in the assessment of the security of some quantum cryptographic protocols [13].

B. Adaptive measurements

So far, we have shown that the performance of individual and collective strategies is essentially the same for large ensembles of identical states. We now show that if we are not restricted to perform the same individual measurement on each copy and we use the information we are gathering to optimize these measurement step by step, the overall performance is *exactly* the same as for the optimal collective strategy, regardless of the number of copies of the unknown state. One could reach this conclusion by using the algebraic results in [14] to trade $|\Omega(a)\rangle$ for a set of product states similar to those in Eq. (11) below. We follow here a different approach since we would like to present a constructive procedure within the framework of probability.

We consider the simplest scenario where we perform always von Neumann measurements on each individual copy. The final outcomes are binary sequences or strings of length N—e.g., **011…01**. Let us denote them by $\boldsymbol{x}$. The strategy is designed in such a way that the last outcome (leftmost binary digit in $\boldsymbol{x}$) determines whether our guess is $|\psi_0\rangle$ or $|\psi_1\rangle$. We have

$$P_N^{\text{ad}} = \sum_{\boldsymbol{x} \in \mathcal{L}_{N-1}} \{q_0 p(\mathbf{0}\boldsymbol{x}|0) + q_1 p(\mathbf{1}\boldsymbol{x}|1)\}, \tag{10}$$

where "ad" stands for adaptive, $\mathcal{L}_r$ is the set of binary strings of length r, and $\mathbf{0}\boldsymbol{x}$ and $\mathbf{1}\boldsymbol{x}$ are the strings obtained by appending 0 and 1 respectively, to the left of the string $\boldsymbol{x}$.

Quantum mechanics tells us that the conditional probability of obtaining the set of outcomes $\boldsymbol{x} \in \mathcal{L}_r$ if the initial state were $|\psi_b\rangle$ is $p(\boldsymbol{x}|b) = |\langle \Omega(\boldsymbol{x}) | \psi_b^N \rangle|^2$, where

$$|\Omega(\boldsymbol{x})\rangle = |\omega(\boldsymbol{x}_r)\rangle \otimes |\omega(\boldsymbol{x}_{r-1})\rangle \otimes \cdots \otimes |\omega(\boldsymbol{x}_1)\rangle, \tag{11}$$

$\boldsymbol{x}_k$ is the substring of length $k (0 \leqslant k \leqslant r)$ consisting of the k rightmost digits of $\boldsymbol{x}$, and

$$|\omega(a\boldsymbol{x})\rangle = \cos\left(\phi_{\boldsymbol{x}} - a\frac{\pi}{2}\right)|x\rangle + \sin\left(\phi_{\boldsymbol{x}} - a\frac{\pi}{2}\right)|y\rangle, \tag{12}$$

in analogy with Eq. (2). Note that $\phi_{\boldsymbol{x}}$, the angle that defines the measurement $r+1$, depends only on the list of outcomes, $\boldsymbol{x}$, of the previous r individual measurements. One readily sees that $\Sigma_{\boldsymbol{x} \in \mathcal{L}_r} |\Omega(\boldsymbol{x})\rangle\langle\Omega(\boldsymbol{x})| = \mathbb{1}$ in $\mathcal{H}^{\otimes r}$, which implies that

$$\sum_{\boldsymbol{x}} p(\boldsymbol{x}|b) = 1, \tag{13}$$

as it should be. We start with $r=0$ ($\mathcal{L}_0$ contains only the empty string $\varnothing$) and set $\phi_\varnothing = \phi_0$, as defined in Eq. (3), which gives the optimal measurement for one copy. For $r > 0, \phi_{\boldsymbol{x}}$ will be determined by requiring optimality step by step. We now can write

$$P_N^{\text{ad}} = \sum_{a=0}^{1} \sum_{\boldsymbol{x} \in \mathcal{L}_{N-1}} p(\boldsymbol{x}, a) |\langle \omega(a\boldsymbol{x}) | \psi_a \rangle|^2, \tag{14}$$

where $p(\boldsymbol{x}, a)$ is the joint probabilities of $|\psi_a\rangle$ being transmitted *and* we obtaining the (partial) outcome list $\boldsymbol{x}$. Namely, $p(\boldsymbol{x}, a) = q_a p(\boldsymbol{x}|a) = q_a \Pi_{s=1}^r |\langle \omega(\boldsymbol{x}_s) | \psi_a \rangle|^2$ (assuming $\boldsymbol{x} \in \mathcal{L}_r$). Equation (14) can be written in terms of the angles θ and $\phi_{\boldsymbol{x}}$ using Eqs. (1) and (12). Maximizing over $\phi_{\boldsymbol{x}}$, we obtain

$$\cos 2\phi_{\boldsymbol{x}} = \frac{p(\boldsymbol{x},0) - p(\boldsymbol{x},1)}{R(\boldsymbol{x})} c, \tag{15}$$

where

$$R(\boldsymbol{x}) = \sqrt{[p(\boldsymbol{x},0) + p(\boldsymbol{x},1)]^2 - 4p(\boldsymbol{x},0)p(\boldsymbol{x},1)c^2}, \tag{16}$$

and we also have

$$\sin 2\phi_{\boldsymbol{x}} = \frac{p(\boldsymbol{x},0) + p(\boldsymbol{x},1)}{R(\boldsymbol{x})} \sin 2\theta. \tag{17}$$

Substituting back in Eq. (14) we obtain

$$P_N^{\text{ad}} = \frac{1}{2} + \frac{1}{2} \sum_{x \in \mathcal{L}_{N-1}} R(\mathbf{x}), \tag{18}$$

where we have used that $\Sigma_{\boldsymbol{x}} p(\mathbf{x}, a) = q_a$, which follows from Eq. (13). Equations (15), (16), and (18) are analogous to Eqs. (3) and (4). Actually, the latter can be seen as a particular case of the former if we define $p(\varnothing, a) = q_a$ (this definition is sensible, since the empty binary string means that no measurement has yet been performed).

Having set up this framework, one can prove our main result. Namely, that this adaptive strategy gives exactly the same error probability as the optimal collective one for any N. A straightforward calculation yields

$$p(a\boldsymbol{x}, b) = \frac{p(\boldsymbol{x}, b)}{2} \left\{ 1 + (-1)^{a+b} \times \frac{p(\boldsymbol{x}, b) + (1 - 2c^2) p(\boldsymbol{x}, b \oplus 1)}{R(\boldsymbol{x})} \right\}, \tag{19}$$

where $\oplus$ stands for sum mod 2, and one can prove by induction the relation

$$q_0 q_1 c^{2r} [p(\boldsymbol{x},0) + p(\boldsymbol{x},1)]^2 - p(\boldsymbol{x},0) p(\boldsymbol{x},1) = 0, \tag{20}$$

for $\boldsymbol{x} \in \mathcal{L}_r$, which is obviously satisfied for $r=0$.

Using this relation in Eq. (16) and recalling again that $\Sigma_{\boldsymbol{x}} p(\boldsymbol{x}, a) = q_a$, we finally have the result $\overline{P}_N^{\text{ad}} = \overline{P}_N^{\text{col}}$, where $\overline{P}_N^{\text{col}}$ is the exact error probability (7).

It is not difficult to show that

$$\cos 2\phi_{\boldsymbol{x}} = (-1)^{i_r} c \sqrt{\frac{1-4q_0q_1c^{2r}}{1-4q_0q_1c^{2r+2}}}, \tag{21}$$

where i_r is the leftmost digit in $\boldsymbol{x} \in \mathcal{L}_r$ and we have used that $\mathrm{sgn}[p(\boldsymbol{x},0)-p(\boldsymbol{x},1)]=(-1)^{i_r}$ [note that $\mathrm{sgn}(q_0-q_1)=(-1)^{i_0}$].

We immediately realize that the actual dependence of the individual measurement $r+1$ on previous outcomes is extremely simple: it is just a function of the rth outcome—i.e., of i_r—rather than a function of the whole binary sequence $\boldsymbol{x}$. In this sense, the optimal one step adaptive scheme is "Markovian." It is thus convenient to change the notation and define $\phi_r \equiv \phi_{\boldsymbol{x}}, |\omega_{r+1}(a)\rangle = |\omega(a\boldsymbol{x})\rangle$, for $\boldsymbol{x} \in \mathcal{L}_r$. Equation (12) becomes

$$|\omega_{r+1}(a)\rangle = \cos\left(\phi_r - a\frac{\pi}{2}\right)|x\rangle + \sin\left(\phi_r - a\frac{\pi}{2}\right)|y\rangle, \tag{22}$$

where the subscript $r+1$ refers to the measurement on copy $r+1$ and $a=0$, 1 is the corresponding outcome. Equation (2) is a particular case of this equation.

C. Bayesian updating interpretation

Finally, we would like to show that the adaptive strategy we have presented has a natural interpretation as Bayesian updating (we refer to [10] for an alternative point of view). This, along with the results of the previous section, proves that Bayesian updating is the *unique* solution to the recursion relations (19) that define the best adaptive strategy.

Note that our knowledge of the system, which changes after each measurement, is encoded in the *a posteriori* probabilities of $|\psi_a\rangle$ being the unknown state *given that* a specific outcome has occurred when performing the measurement on, say, the rth copy. We will show below that these *a posteriori* probabilities can be identified with P_r^{ad} and $\overline{P}_r^{\mathrm{ad}}$. Assuming this for the time being, we might be tempted to take a Bayesian point of view and use P_r^{ad} to update our prior probabilities for the next measurement. Hereafter, we drop the superscript "ad" to further simplify the notation.

Suppose we have got the first copy of the unknown state. Our optimal measurement will be defined by ϕ_0 in Eq. (3). If we obtain the outcome $i_1=0$, we will update our priors using the rule $q_0 \to p(0|\mathbf{0})=P_1$, and we will use again Eq. (3) to optimize the measurement on the second copy [similarly, if the first outcome is $i_1=1$, we will view $p(1|\mathbf{1})=P_1$ as our prior q_1 for the second measurement]. Hence, the second measurement is defined by $\cos 2\phi_1=(-1)^{i_1}c|P_1-\overline{P}_1|(1-4P_1\overline{P}_1c^2)^{-1/2}$, and we obtain that the discrimination (error) probability after the second measurement is $P_2=[1+(1-4P_1\overline{P}_1)^{1/2}]/2$ **(**$\overline{P}_2=[1-(1-4P_1\overline{P}_1)^{1/2}]/2$**)**. This updating of the prior probabilities can be carried out step by step until we run out of copies. At step r we will have

$$\cos 2\phi_r = (-1)^{i_r}\frac{|P_r-\overline{P}_r|}{R_r}c, \tag{23}$$

where by analogy with $R(\boldsymbol{x})$, we have defined $R_r=(1-4P_r\overline{P}_rc^2)^{1/2}$, and we obtain

$$P_{r+1} = (1+R_r)/2. \tag{24}$$

This leads to the recursion relation

$$R_{r+1} = \sqrt{1-(1-R_r^2)c^2}, \tag{25}$$

whose solution can readily be seen to be $R_r=[1-4q_0q_1c^{2r+2}]^{1/2}$, and we again find that $\overline{P}_N^{\mathrm{ad}}=\overline{P}_N^{\mathrm{col}}$.

We still need to show that the *a posteriori* probabilities indeed coincide with P_r. It suffices to prove it for the case $r=1$, where this statement amounts to $P_1=p(0|\mathbf{0})=p(1|\mathbf{1})$. This result follows from the obvious formula

$$P_1 = p(0|\mathbf{0})p(\mathbf{0}) + p(1|\mathbf{1})p(\mathbf{1}), \tag{26}$$

where $p(\boldsymbol{b})$ is the probability of obtaining the outcome b, *if* the "detailed balance" relation

$$p(0|\mathbf{0}) = p(1|\mathbf{1}) \tag{27}$$

holds for the optimal scheme. Let us prove this is the case.

Using Bayes formula we can cast Eq. (27) as

$$\frac{|\langle\omega_1(0)|\psi_0\rangle|^2 q_0}{p(\mathbf{0})} = \frac{|\langle\omega_1(1)|\psi_1\rangle|^2 q_1}{p(\mathbf{1})}. \tag{28}$$

We further note that the probabilities of obtaining the outcome a can simply be written as $p(\boldsymbol{a})=\Sigma_b|\langle\omega_1(a)|\psi_b\rangle|^2q_b$. Therefore, Eqs. (27) and (28) are equivalent to

$$\frac{|\langle\omega_1(0)|\psi_1\rangle|^2 q_1}{|\langle\omega_1(0)|\psi_0\rangle|^2 q_0} = \frac{|\langle\omega_1(1)|\psi_0\rangle|^2 q_0}{|\langle\omega_1(1)|\psi_1\rangle|^2 q_1}. \tag{29}$$

This, in terms, is equivalent to

$$(q_0-q_1)\sin 2\phi\cos 2\theta = (q_0+q_1)\cos 2\phi \sin 2\theta, \tag{30}$$

which obviously holds for the optimal strategy [see Eq. (3)] and concludes the proof.

IV. CONCLUDING REMARKS

In summary, multiple-copy two-state discrimination strategies based on individual measurements can be as good as the best collective ones. For fixed measurements, this statement holds only asymptotically. By relaxing this constrain and allowing Bayesian updating, which is arguably the simplest, easiest to implement, adaptive strategy, the statement holds for *any* finite number of copies. Furthermore, our approach provides very simple recursion relations [e.g., Eqs. (23)–(25)] or even closed-form expressions [e.g. Eq. (21); recall the change of notation $\phi_r=\phi_{\boldsymbol{x}}$] for the angles ϕ_r defining the optimal von Neumann measurements and the discrimination and error probabilities.

Finally, we would like to point out that the general adaptive setup of Sec. III B, where measurements are allowed to depend on histories or lists of outcomes (rather than just the very last outcome), has a unique solution which can be regarded as Bayesian updating. Despite all this generality, the optimal solution is as simple as can be.

ACKNOWLEDGMENTS

We acknowledge financial support from Spanish Ministry of Science and Technology Project No. BFM2002-02588, a "Ramón y Cajal" grant, UB Grant No. 2002FI-00373, CIRIT Project No. SGR-00185, and QUPRODIS working group EEC Contract No. IST-2001-38877.

[1] A. S. Holevo, *Probabilistic and Statiscal Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[3] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991).

[4] A. G. White *et al.*, Phys. Rev. Lett. **83**, 3103 (1999); D. F. V. James *et al.*, *ibid.* **64**, 052312 (2001); R. T. Thew *et al.*, *ibid.* **66**, 012303 (2002); J. B Altepeter *et al.*, Phys. Rev. Lett. **90**, 193601 (2003).

[5] D. G. Fischer, S. H. Kienle, and M. Freyberger, Phys. Rev. A **61** 032306 (2000); R. D. Gill and S. Massar, *ibid.* **61**, 042312 (2000); Th. Hannemann *et al.*, *ibid.* **65**, 050303(R) (2002); E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. Lett. **89**, 277904 (2002).

[6] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[7] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[8] A. Chefles, Contemp. Phys. **41**, 401 (2000).

[9] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).

[10] D. Brody and B. Meister. Phys. Rev. Lett. **76**, 1 (1996).

[11] M. Ban, K. Yamazaki, and O. Hirota, Phys. Rev. A **55**, 22 (1997).

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications (Wiley, New York, 1991).

[13] A. Acín *et al.*, e-print quant-ph/0411092.

[14] J. Walgate *et al.*, Phys. Rev. Lett. **85**, 4972 (2000); S. Virmani *et al.*, Phys. Lett. A **228**, 62 (2001).