Optimal estimation of quantum observables

Giacomo Mauro D'Ariano[a)]
QUIT Group, Dipartimento di Fisica, "A. Volta," via Bassi 6, 27100 Pavia, Italy

Vittorio Giovannetti[b)]
NEST-CNR-INFM and Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

Paolo Perinotti[c)]
QUIT Group, Dipartimento di Fisica, "A. Volta," via Bassi 6, 27100 Pavia, Italy

(Received 24 October 2005; accepted 3 January 2006; published online 21 February 2006)

We consider the problem of estimating the ensemble average of an observable on an ensemble of equally prepared identical quantum systems. We show that, among all kinds of measurements performed jointly on the copies, the optimal unbiased estimation is achieved by the usual procedure that consists in performing independent measurements of the observable on each system and averaging the measurement outcomes. © *2006 American Institute of Physics.* [DOI: 10.1063/1.2168122]

I. INTRODUCTION

The astonishing precision of measurements currently available in quantum optics[1] along with the growing demand of quantum devices of the new information technology[2,3] have revived the interest in the theory of quantum measurements.[4] The outcome statistics of a quantum measurement for all possible input states is described by a positive operator valued measure (POVM). The general optimization approach of quantum estimation theory[5] is to maximize over all possible POVM's an appropriate cost function, which depends on the context and on the specific use of the measurement. The output statistics can then be improved by using multiple copies of the same quantum system, all prepared in the same state, and performing a suitable *ensemble measurement* over the copies.

The experimental complexity of ensemble measurements is roughly classified by dividing them into three main categories: (a) *independent*, (b) *separable*, and (c) *entangled* measurements. Category (a) is described by tensor products of independent POVM's; (b) by POVM's with separable elements only; (c) by POVM's where some elements are entangled. Notice that the separability of POVM's generally does not correspond to a physical separability of measuring apparatuses [there exist separable measurements that cannot be performed by separate measuring apparatuses, i.e., by local operations and classical communication (LOCC)], and this classification remains essentially mathematical in nature. However, at least one can say that category (b) contains all *adaptive* measurements (in which the choice of the measuring apparatus on the nth copy depends on the outcomes of previous measurements), whereas category (c) contains those measurements that need quantum interactions between copies, implying that all copies during the measuring time must be at the same physical location, or, otherwise, that a "quantum memory" is available.

Among the three categories of ensemble measurements, the category (c) of entangled POVM's discloses the full exponential growth of the Hilbert space dimension versus the number of copies N for a virtually unlimited optimization of the statistical efficiency of the measurement, with the

[a)]Electronic mail: dariano@unipv.it
[b)]Electronic mail: v.giovannetti@sns.it
[c)]Electronic mail: perinotti@fisicavolta.unipv.it

possibility of largely surpassing the performance of categories (a) and (b).[6–9] Indeed, over the last few years, it has been recognized that entangled measurements are usually more efficient than independent measurements, and the optimal measurement scheme is almost always entangled.[10–13] However, in some situations it has been also shown that asymptotically for $N\to\infty$ an equivalently optimal estimation may be achieved using just independent measurements over the copies.[14–17]

In the above scenario it is natural to ask if the *canonical procedure* of averaging the outcomes of repeated measurements of an observable A over equally prepared systems is the best way of estimating the ensemble average $\langle A\rangle$ of A, or, instead, if a joint entangled measurement over the copies can improve the estimation. As we will see it turns out that the canonical procedure is indeed optimal, however, the derivation of this result is nontrivial, and offers a general warning against easy assumptions and generalizations when evaluating statistical efficiencies of ensemble measurements.

Let us be more precise, and fix precisely the scenario of the quantum estimation. Suppose one has a finite number N of equally prepared distinguishable identical d-dimensional quantum systems, which are described by the state $\rho^{\otimes N}$, and one wants to estimate the ensemble average $\langle A\rangle_\rho \equiv \mathrm{Tr}[\rho A]$ of the observable A. Suppose now that one has unlimited technology at one's disposal, including measuring apparatus that can achieve any desired entangled POVM on all N systems jointly. The question is which is the best measuring apparatus to choose in order to estimate $\langle A\rangle_\rho$ with the minimum statistical error? What we will prove in the present paper is that the best estimation strategy is just the canonical procedure, which consists in averaging the outcomes of repeated measurements of the observable A over the equally prepared quantum systems.

II. PERMUTATIONALLY INVARIANT POLARIZATION IDENTITIES

In the derivation of our main result the following lemma will play a crucial role.

Lemma 1: Any permutationally invariant operator X on $\mathsf{H}^{\otimes N}$ is completely determined by all ensemble averages $\mathrm{Tr}[X\rho^{\otimes N}]$ *on identical equally prepared systems.*

Proof: The statement of the lemma is equivalent to the following logical implication:

$$X \in \mathcal{P}_N(\mathsf{H}), \qquad \rho \in \mathcal{S}(H), \quad \mathrm{Tr}[X\rho^{\otimes N}] = 0 \Rightarrow X = 0, \tag{1}$$

where $\mathcal{S}(H)$ denotes the set of states on H, and $\mathcal{P}_N(\mathsf{H})$ the algebra of permutationally invariant operators on $\mathsf{H}^{\otimes N}$. Indeed, statement (1) is equivalent to the statement that if $\mathrm{Tr}[X\rho^{\otimes N}]=\mathrm{Tr}[Y\rho^{\otimes N}]$ for all states ρ, then $X\equiv Y$.

Consider the following special states of the form

$$\rho_\lambda = \sum_{j=1}^{N} \lambda_j |\psi_j\rangle\langle\psi_j|, \quad \lambda_j > 0, \lambda_j \neq \lambda_i, i \neq j, \tag{2}$$

with $\{\psi_j\}_N$ any set of N unequal states (not necessarily orthogonal). The trace $\mathrm{Tr}[X\rho_\lambda^{\otimes N}]$ is a polynomial in $\Pi_{j=1}^N \lambda_j^{x_j}$, with $\Sigma_{j=1}^N x_j = N$ and $x_j \geqslant 0$ integers. Now, in order to have $\mathrm{Tr}[X\rho_\lambda^{\otimes N}]=0$ for arbitrary ρ_λ, all coefficients of the polynomial must vanish. In particular, the coefficient of $\Pi_{j=1}^N\lambda_j$ is given by

$$\sum_\sigma \langle\psi_1|\ldots\langle\psi_N|\Pi_\sigma X \Pi_\sigma^\dagger|\psi_1\rangle\ldots|\psi_N\rangle \equiv 0, \tag{3}$$

where Π_σ are the permutations of the N systems. By hypothesis we have $\Pi_\sigma X\Pi_\sigma^\dagger = X$, then the vanishing of $\mathrm{Tr}[X\rho_\lambda^{\otimes N}]$ for all states ρ_λ implies

$$\langle\psi_1|\ldots\langle\psi_N|X|\psi_1\rangle\ldots|\psi_N\rangle = 0, \tag{4}$$

for all sets $\{\psi_j\}_N$. If we take $|\psi_k\rangle = \alpha|\phi\rangle + \beta|\phi_\perp\rangle$, by arbitrariness of α and β we have

$$\begin{aligned}\langle\psi_1|\cdots\langle\phi|\cdots\langle\psi_N|X|\psi_1\rangle\cdots|\phi\rangle\cdots|\psi_N\rangle &= \langle\psi_1|\cdots\langle\phi|\cdots\langle\psi_N|X|\psi_1\rangle\cdots|\phi_\perp\rangle\cdots|\psi_N\rangle\\ &= \langle\psi_1|\cdots\langle\phi_\perp|\cdots\langle\psi_N|X|\psi_1\rangle\cdots|\phi\rangle\cdots|\psi_N\rangle\\ &= \langle\psi_1|\cdots\langle\phi_\perp|\cdots\langle\psi_N|X|\psi_1\rangle\cdots|\phi_\perp\rangle\cdots|\psi_N\rangle = 0. \end{aligned}\tag{5}$$

By repeating the same argument for different values of k and choosing ϕ and $\phi_\perp$ as all possible elements of an orthonormal basis $\{\phi_j\}$ we get

$$\langle\phi_{j_1}|\cdots\langle\phi_{j_N}|X|\phi_{k_1}\rangle\cdots|\phi_{k_N}\rangle = 0, \qquad \{j_i\},\{k_i\}. \tag{6}$$

Since all the matrix elements of X on an orthonormal basis are null, one has that $X\equiv 0$. ■

Notice that the proof of the previous lemma contains the following interesting corollary.

Corollary 1: For any operator X on $\mathsf{H}^{\otimes N}$ *the diagonal elements on factorized states completely determine X.*

This is a kind of *factorized* polarization identity for permutation invariant operators.

III. THE MAIN RESULT

Let us now come back to the original problem of determining the optimal measurement for estimating the ensemble average of an observable. Consider a generic joint POVM $P(r)$ on $\mathsf{H}^{\otimes N}$, with outcome r providing an estimate of the expectation $\langle A\rangle_\rho$ of the observable A on N identical systems all in the same state ρ. Clearly, one has $\lambda_m \leqslant r \leqslant \lambda_M$, with λ_m and λ_M minimum and maximum eigenvalues of A, respectively. The POVM $P(r)$ provides an estimate of the expectation $\langle A\rangle_\rho$ if the conditional probability $p(r|\rho)$ of estimating expectation value r for actual value $\mathrm{Tr}[A\rho]$ is expressed via the Born rule as follows:

$$p(r|\rho)\mathrm{d}r = \mathrm{Tr}[P(r)\rho^{\otimes N}]\mathrm{d}r. \tag{7}$$

Since the state $\rho^{\otimes N}$ is permutation invariant, we can consider permutation invariant POVM's. Indeed, using invariance of $\rho^{\otimes N}$ under permutations, one has

$$p(r|\rho)\mathrm{d}r = \frac{1}{N!}\sum_\sigma \mathrm{Tr}[\Pi_\sigma\rho^{\otimes N}\Pi_\sigma^\dagger P(r)]\mathrm{d}r = \mathrm{Tr}\left[\rho^{\otimes N}\frac{1}{N!}\sum_\sigma (\Pi_\sigma^\dagger P(r)\Pi_\sigma)\right]\mathrm{d}r = \mathrm{Tr}[\Pi'(r)\rho^{\otimes N}]\mathrm{d}r, \tag{8}$$

where the POVM

$$P'(r) \equiv \frac{1}{N!}\sum_\sigma \Pi_\sigma^\dagger P(r)\Pi_\sigma \tag{9}$$

is permutation invariant by construction. This means that for any POVM there is a permutation invariant one giving the same probability distributions for all states $\rho^{\otimes N}$. Therefore, without loss of generality, in the following we can assume that $P(r)$ is permutation invariant. We will consider now the case in which the POVM is *unbiased*, that is the averaging over r coincides with the value to be estimated. Mathematically this means that for all states ρ the following identity holds:

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, rp(r|\rho) = \mathrm{Tr}[A\rho]. \tag{10}$$

The statistical error in the estimate is given by the rms of the probability distribution

$$\epsilon_N(A) \doteq \left[\int_{\lambda_m}^{\lambda_M} \mathrm{d}r(r-\langle A\rangle_\rho)^2 p(r|\rho)\right]^{1/2}, \tag{11}$$

which for unbiased estimation equals

$$\epsilon_N(A) \doteq \left[\int_{\lambda_m}^{\lambda_M} \mathrm{d}r[r^2 p(r|\rho)] - \langle A \rangle_\rho^2 \right]^{1/2}. \tag{12}$$

Since the only part which depends on the POVM is the conditional probability $p(r|\rho)$, the optimization of the error resorts to minimize the quantity

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, r^2 \,\mathrm{Tr}[P(r)\rho^{\otimes N}], \tag{13}$$

with the constraints

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, P(r) = I, \tag{14}$$

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, r\, \mathrm{Tr}[P(r)\rho^{\otimes N}] = \langle A \rangle_\rho. \tag{15}$$

Using the following identity:

$$\langle A \rangle_\rho = \mathrm{Tr}\left[\rho^{\otimes N} \frac{1}{N!} \sum_\sigma \Pi_\sigma (A \otimes I^{\otimes(N-1)}) \Pi_\sigma^\dagger \right] = \mathrm{Tr}\left[\rho^{\otimes N} \frac{1}{N} \sum_{k=1}^N A^{(k)} \right] \tag{16}$$

with $A^{(k)} \doteq I^{\otimes(k-1)} \otimes A \otimes I^{\otimes N-k}$, by virtue of Lemma 1 we can recast Eq. (15) as follows:

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, rP(r) = \frac{1}{N} \sum_{k=1}^N A^{(k)} \doteq \Theta. \tag{17}$$

The operator $\Delta \geqslant 0$ defined as

$$\Delta \doteq \int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, r^2 P(r), \tag{18}$$

allows to reexpress the statistical error as follows:

$$\epsilon_N(A)^2 = \mathrm{Tr}[\Delta \rho^{\otimes N}] - \langle A \rangle_\rho^2. \tag{19}$$

In the representation in which Δ is diagonal, the constraints (14) and (15) become

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, P(r)_{lk} = \delta_{lk}, \tag{20}$$

$$\int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, rP(r)_{lk} = \Theta_{lk}, \tag{21}$$

whereas the error (12) becomes

$$\epsilon_N(A)^2 = \sum_n (\rho^{\otimes N})_{nn} \int_{\lambda_m}^{\lambda_M} \mathrm{d}r\, r^2 P(r)_{nn} - \langle A \rangle_\rho^2. \tag{22}$$

From Eqs. (20) and (21) it follows that the diagonal elements $P(r)_{nn}$ are probability densities in r over $[\lambda_m, \lambda_M]$, with average Θ_{nn}. Denoting the variance of $P(r)_{nn}$ by σ_n^2, we can write

$$\epsilon_N(A)^2 = \sum_n (\rho^{\otimes N})_{nn}(\sigma_n^2 + \Theta_{nn}^2) - \langle A \rangle_\rho^2. \tag{23}$$

Therefore, $\epsilon_N(A)^2$ is minimized by taking $\sigma_n^2=0$, corresponding to $P(r)_{nn} \equiv \delta(\Theta_{nn}-r)$. This implies that the outcomes of the optimal POVM are actually discrete, corresponding to $r_n=\Theta_{nn}$. In this discrete version, the POVM has $P(r_n)_{nm}=\delta_{nm}$ [which also implies that $\Theta_{nm}=\delta_{nm}\Theta_{nn}$ via Eq. (21)], that is $P(r_n)$ is projection valued on the nth eigenvector of Δ (when it happens that $\Theta_{nn}=\Theta_{mm}$ for some $m \neq n$, then the projector has rank equal to the number of equal diagonal elements). We have finally

$$\epsilon_N(A)^2 = \sum_n (\rho^{\otimes N})_{nn}\Theta_{nn}^2 - \langle A \rangle_\rho^2. \tag{24}$$

Moreover, we have

$$I = \sum_n P(r_n), \tag{25}$$

$$\Theta = \sum_n \Theta_{nn} P(r_n), \tag{26}$$

$$\Delta = \sum_n \Theta_{nn}^2 P(r_n). \tag{27}$$

Since optimization makes Θ and Δ jointly diagonal, one has $\Sigma_n \Theta_{nn}^2 (\rho^{\otimes N})_{nn} = \mathrm{Tr}[\Theta^2 \rho^{\otimes N}]$, and using Eqs. (17) and (24) we can write the following expression for the minimal error:

$$\epsilon_N(A)^2 = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathrm{Tr}[A^{(i)} A^{(j)} \rho^{\otimes N}] - \langle A \rangle_\rho^2. \tag{28}$$

Notice that the sum in the first term contains N terms with $i=j$ equal to $\mathrm{Tr}[A^2\rho]$ and $N(N-1)$ with $i \neq j$ equal to $\langle A \rangle_\rho^2$, resulting in

$$\epsilon_N(A) = \sqrt{\frac{\langle A^2 \rangle_\rho - \langle A \rangle_\rho^2}{N}}, \tag{29}$$

that is the optimal error equals the statistical error occurring when measuring A separately on all the identical quantum systems in the state ρ, and then averaging. Indeed, the optimal POVM coincides with the spectral resolution of $\Theta=(1/N)\Sigma_n A^{(n)}$ on $\mathsf{H}^{\otimes N}$.

ACKNOWLEDGMENTS

Two of the authors (G.M.D.) and (P.P.) acknowledge financial support by the EC under the program ATESIT (Contract No. IST-2000-29681), and by MIUR under programs Cofinanziamento 2005 and FIRB 2001. One of the authors (V.G.) acknowledges financial support by MIUR through Cofinanziamento 2003 and by "Quantum Information" research programme of Centro di Ricerca Matematica "Ennio De Giorgi" of Scuola Normale Superiore.

[1] *Quantum Optics, Experimental Gravity, and Measurement Theory*, edited by P. Meystre and M. O. Scully (Plenum, New York, London, 1983).
[2] *Introduction to Quantum Computation and Information*, edited by H.-K. Lo, S. Popescu, and T. Spiller (World Scientific, Singapore, 1998).
[3] I. L. Chuang and M. A. Nielsen, *Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, 2000).
[4] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, Lecture Notes in Physics Vol. 2 (Springer-Verlag, Telos, 1996).
[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[6] R. Derka, V. Bužek, and A. K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).
[7] A. Chefles, Phys. Rev. A **64**, 062305 (2001).
[8] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).
[9] A. S. Holevo, quant-ph/0307225.
[10] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. Lett. **93**, 180503 (2004).
[11] G. M. D'Ariano, C. Macchiavello, and P. Perinotti, Phys. Rev. A **72**, 042327 (2005).
[12] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. A **70**, 062105 (2004).
[13] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. A **72**, 042108 (2005).
[14] E. Bagan, M. A. Ballester, R. Muñoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **95**, 110504 (2005).
[15] M. A. Ballester, to appear in J. Phys. A.
[16] M. A. Ballester, Phys. Rev. A **70**, 032310 (2004).
[17] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).