

Separable Measurement Estimation of Density Matrices and its Fidelity Gap with Collective Protocols

E. Bagan,¹ M. A. Ballester,² R. D. Gill,^{3,4} R. Muñoz-Tapia,¹ and O. Romero-Isart¹

¹*Grup de Física Teòrica & IFAE, Facultat de Ciències, Edifici Cn, Universitat Autònoma de Barcelona, 08193 Bellaterra Barcelona, Spain*

²*Centrum voor Wiskunde en Informatica, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands*

³*Department of Mathematics, University of Leiden, Box 9512, 2300 RA Leiden, The Netherlands*

⁴*EURANDOM, Post Office Box 513-5600 MB Eindhoven, The Netherlands*

(Received 7 April 2006; published 25 September 2006)

We show that there exists a gap between the performance of separable and collective measurements in the qubit mixed-state estimation that persists in the large sample limit. We characterize the gap with sharp asymptotic bounds on mean fidelity. We present an adaptive protocol that attains the separable measurement bound. This protocol uses von Neumann measurements and can be easily implemented with current technology.

DOI: [10.1103/PhysRevLett.97.130501](https://doi.org/10.1103/PhysRevLett.97.130501)

PACS numbers: 03.67.Mn, 03.65.Ta, 03.65.Wj

Collective measurements provide the largest amount of information that can be retrieved from a multipartite quantum system. However, they rarely offer much advantage over separable (also known as *unentangled*) measurements as far as state discrimination [1] or estimation based on large samples of identical copies is concerned. To be more precise, as the sample size N goes to infinity the (mean-squared) error in the (optimal) estimate often vanishes at the same rate for both collective and separable measurements, despite the fact that the former are fully general whereas the latter are much more constrained. Examples of this abound in the literature. They include estimation of qubit pure states [2] (separate individual measurements suffice in this case), estimation of spectrum of a qudit density matrix [3] (for qubits see Ref. [4]), discrimination of two nonorthogonal multipartite pure states [5], multiple-copy 2-state discrimination [6], the problem of distinguishing the states of the double-trine ensemble [1,7], etc. (In these discrimination examples N needs not be asymptotically large.)

In this Letter, we address the opposite, less common situation, but for a very important example. (I) We present a state-estimation scenario where separable measurements are outperformed by collective measurements *even* in the large sample limit. (II) We give a protocol based solely on local operations and classical communication (a so-called LOCC protocol) whose mean-squared error (fidelity) attains the lower (upper) separable measurement bound. To the best of our knowledge this is the first time a complete example of (I) and (II) is given. The example which we provide is of great practical relevance: it concerns the estimation of a qubit density matrix $\rho = (\mathbb{1} + \vec{r} \cdot \vec{\sigma})/2$, $|\vec{r}| \equiv r \leq 1$ (the components of $\vec{\sigma}$ are the three Pauli matrices) given N identical copies, i.e., assuming we are given the state $\rho^{\otimes N}$. The fact that the protocol uses von Neumann measurements and just one-step adaptivity adds greatly to its practical interest, since it can be imple-

mented with present-day technology. Similar protocols have been proposed earlier in, e.g., Refs. [8–10], but they have not been studied from the point of view of average fidelity. Our protocol is delicately tuned to attain the asymptotic maximum separable measurement fidelity.

The estimation of qubit mixed states already showed some other puzzling anomalies. In Ref. [11] it was proved that the average error using (local) tomography vanishes as $N^{-3/4}$, which conflicts with the behavior N^{-1} expected on statistical grounds, as well as being much worse than the optimal collective results [12]. A closer look at this problem reveals that this strange power law is intimately connected to the particular but very natural choice of prior distribution used there. This will be explained, after introducing our notation.

Our aim is to maximize the mean fidelity $F = (1 + \Delta)/2$, where

$$\Delta = \sum_{\chi} \int dn dr w(r) \mathbf{r} \cdot \mathbf{R}_{\chi} \text{Tr}[O_{\chi} \rho^{\otimes N}]. \quad (1)$$

In writing (1) we have used that (a) the fidelity between ρ and its estimate $\rho_{\chi} = (\mathbb{1} + \vec{R}_{\chi} \cdot \vec{\sigma})/2$ can be cast in the form $f(\rho, \rho_{\chi}) = (\text{Tr} \sqrt{\sqrt{\rho} \rho_{\chi} \sqrt{\rho}})^2 = (1 + \mathbf{r} \cdot \mathbf{R}_{\chi})/2$, where the (Euclidean) 4-dimensional vector \mathbf{r} (similarly \mathbf{R}_{χ}) is defined as $\mathbf{r} = (\sqrt{1 - r^2}, \vec{r})$; (b) a generalized measurement on $\rho^{\otimes N}$ is represented by a positive operator valued measure, $\mathcal{O} = \{O_{\chi}\}$; (c) the prior probability distribution of ρ is isotropic (dn is the rotationally invariant measure on \mathbb{S}^{d-1} , where $d = 3, 2$ depending, respectively, on whether we deal with the entire Bloch sphere or only with its equatorial plane); and (d) the purity (i.e., r) prior distribution has probability density $w(r)$.

This Letter focuses on measurements for which $O_{\chi} = \bigotimes_{k=1}^N O_{\chi_k}^{(k)}$ (so that $\chi = \chi_N \chi_{N-1} \dots \chi_1$ is a string of outcomes, each of them associated with one of the copies of

ρ). By definition, separable measurements are those whose components are convex combinations of these “local” operators. All LOCC measurements are separable, but not vice versa [13].

It is worth mentioning that, although we do not stick to any particular purity prior, many arguments favor the choice of the Bures distribution,

$$w_{\text{Bures}}(r) = \frac{2\Gamma(\frac{d+1}{2})}{\sqrt{\pi}\Gamma(\frac{d}{2})} \frac{r^{d-1}}{\sqrt{1-r^2}}. \quad (2)$$

Notice that (2) is precisely the volume element induced by the distance $(1/2)\arccos f(\rho, \rho + d\rho)$ [14]. It is monotonically decreasing under coarse graining [15]. It has been argued that it corresponds to maximal randomness of the signal states and hence describes an ensemble for which one has minimal prior knowledge [16].

Let us summarize some known facts about the optimal (collective) protocols. For asymptotically large samples the maximum value of the fidelity reads [12] (see also Ref. [9])

$$F_{d=3}^{\max} = 1 - \frac{3 + 2\bar{r}}{4N} + o(1/N), \quad (3)$$

where \bar{r} stands for the average of the purity with respect to its prior (on which very mild regularity conditions need be assumed), i.e., $\bar{r} \equiv \int_0^1 dr w(r)r$. For the Bures prior, Eq. (2), we have $\bar{r} = 2/(3\pi)$. Similarly [12],

$$F_{d=2}^{\max} = 1 - \frac{1}{2N} + o(1/N), \quad (4)$$

which is independent of $w(r)$. These asymptotic bounds were computed assuming no restriction upon the type of measurements used in the protocols. Although they were shown to be attained by a specific collective measurement, one cannot rule out that separable measurements can also attain the bounds (especially after our introductory remarks about the frequent asymptotic optimality of local protocols).

To get a hint for this problem, we recall some results [11] concerning tomography, which is a standard scheme for quantum state estimation in the laboratory at present. In the simplest approach one performs measurements of the projections of $\vec{\sigma}$ along d fixed orthogonal directions (one does not make use of classical communication), each of them on a corresponding fraction of the sample of size N/d . For the Bures prior the best data processing leads to

$$F_d^{\text{tom}} = 1 - \frac{\xi_d}{N^{3/4}} + o(N^{-3/4}), \quad 0 < \xi_2 \leq \xi_3, \quad (5)$$

where the specific value of the constant ξ_d is irrelevant to this discussion. This asymptotic behavior contrasts drastically with (3) and (4), and might lead us to suspect that there exists no local estimation protocol for which $1 - F \sim N^{-1}$ if $w(r)$ is the Bures prior (2). We show below that this is not so.

But let us first provide an explanation of the behavior shown in (5), which can be traced back to the contribution to F from states near the surface of the Bloch sphere (almost pure states). Let \mathcal{S}_ε be the outer shell of the Bloch sphere of thickness ε , i.e., $\mathcal{S}_\varepsilon = \{\vec{r}: 1 - \varepsilon < r \leq 1\}$. As an extreme case, consider the fidelity $f(\rho, \rho')$ between two states of \mathcal{S}_ε whose Bloch vectors point in the same direction while one is of length $r = 1$, the other of length $r' = 1 - \varepsilon$. We find that $1 - f(\rho, \rho') \sim \varepsilon/2$ [instead of $1 - f(\rho, \rho') \sim \varepsilon^2$, which holds for states in the complement of \mathcal{S}_ε whose Bloch vectors are a distance ε apart]. For a signal state not in the direction of one of the measurement axes, however we process the tomographic data χ we cannot hope to reconstruct the location and, in particular, the *length* of the Bloch vector to an accuracy better than $N^{-1/2}$. Thus $1 - f(\rho, \rho_\chi) \sim N^{-1/2}$ when $\rho \in \mathcal{S}_{N^{-1/2}}$. Integrating (2) from $r = 1 - \varepsilon$ to $r = 1$, we see that the signal state has a probability $p \sim \varepsilon^{1/2}$ of being in \mathcal{S}_ε . In particular, $\rho \in \mathcal{S}_{N^{-1/2}}$ with probability $p \sim N^{-1/4}$. In this case, the best processing of the data can at most result in $1 - f(\rho, \rho_\chi) \sim N^{-1/2}$. The (dominant) contribution to $1 - F$ from this increasingly thin outer shell of signal states, $\mathcal{S}_{N^{-1/2}}$, is therefore of order $[1 - f(\rho, \rho_\chi)] \times p \sim N^{-3/4}$.

Let us now move on to the central part of our work and prove the existence of a gap between the asymptotic fidelities of separable and collective protocols. This, in turn, provides an upper bound for the LOCC fidelity. To this end, we recall some results concerning quantum statistical inference theory (a comprehensive summary of general results can be found in Ref. [12]). Hereafter we write simply f for $f(\rho, \rho_\chi)$. From classical statistical arguments, the average of the fidelity over the outcomes can be expected to be

$$\langle f \rangle_\chi = 1 - \frac{1}{4N} \text{Tr}\{H(\theta)[I(\theta, \mathcal{O})]^{-1}\} + o(1/N), \quad (6)$$

where $\theta = (r, \theta, \phi)$ (i.e., the standard spherical coordinates), $H(\theta) = \text{diag}[1/(1-r^2), r^2, r^2 \sin^2 \theta]$ is the quantum Fisher information matrix, and $I(\theta, \mathcal{O})$ is related to the (“classical”) Fisher information matrix (FI) $I_N(\theta, \mathcal{O})$ corresponding to a measurement \mathcal{O} on $\rho^{\otimes N}$ through the equation $I(\theta, \mathcal{O}) = \lim_{N \rightarrow \infty} I_N(\theta, \mathcal{O})/N$. (For states on the equatorial plane, $d = 2$, we just drop the ϕ entry in the above expressions.) Up to a constant factor, H is the Riemannian metric corresponding to the fidelity [17]. The FI plays a similar role with respect to the classical fidelity (overlap) between probability distributions. The inverse of H is a lower bound to the inverse of the FI which is a lower bound for the variance of “reasonable” estimators of θ [18].

If one restricts oneself to separable measurements, the following bound holds:

$$\text{Tr}\{[H(\theta)]^{-1}I(\theta, \mathcal{O})\} \leq 1, \quad (7)$$

as proved in Ref. [8]. It follows straightforwardly that

$$\text{Tr}\{H(\theta)[I(\theta, \mathcal{O})]^{-1}\} \geq d^2. \quad (8)$$

Equations (6) and (8) suggest that for any separable measurement scheme the following bound should apply:

$$\lim_{N \rightarrow \infty} N[1 - \langle f \rangle_\chi] \geq \frac{d^2}{4}. \quad (9)$$

One could, moreover, hope that the bound remains true after averaging with respect to any prior.

A direct and rigorous proof of the desired result can be given using exactly the same arguments as in Appendix H of Ref. [12]. Alternatively, it follows from a general theorem proved in Ref. [19]. Either way, we have the inequality

$$\lim_{N \rightarrow \infty} N(1 - F_d^{\text{sep}}) \geq d^2/4 > \lim_{N \rightarrow \infty} N(1 - F_d^{\text{max}}) \quad (10)$$

where the second, strict, inequality, which follows from (3) and (4), proves the existence of a gap between the two asymptotic optimal fidelities.

Within the so-called pointwise approach to quantum state estimation, the existence of a gap between optimal collective and separable measurements on multiparameter problems has been known for some time; see Hayashi and Matsumoto [9] and their references. In this approach, one compares the pointwise rate of convergence (i.e., at each fixed θ), with respect to mean square error, of estimators satisfying regularity conditions (e.g., “asymptotically locally unbiased”). Differing efficiencies in this approach suggest, but *do not prove*, that a corresponding gap exists when we compare *average* (with respect to a prior) *fidelity* of *arbitrary* estimators.

We are now in a position to state precisely and prove our main result: there exists a (LOCC) one-step adaptive protocol that saturates the separable measurement bound (10). The protocol which, taking inspiration from the Gill-Massar approach [8], makes use of adaptivity (and thus of classical communication) only once, is as follows. In a first step, we spend a vanishing fraction, $N^\alpha \equiv N_0$ ($1/2 < \alpha < 1$), of copies of ρ to get a rough estimate of \vec{n} (θ and ϕ), to which we refer as \vec{n}_0 . To this purpose we may use, e.g., tomography.

In a second step we use tomography again on the remaining $N - N_0 \equiv N_1 d$ copies of ρ , but now we measure the projection of $\vec{\sigma}$ along \vec{n}_0 and along $d - 1$ other orthogonal axis in the plane normal to \vec{n}_0 . In the following we refer to these axis as \vec{z} , \vec{x} , and \vec{y} respectively; they define a spatial reference frame related to the original one through a known rotation. The outcomes of this second step can be written as $\chi = (\chi_x, \chi_y, \chi_z) \equiv (2\alpha_x - 1, 2\alpha_y - 1, 2\alpha_z - 1)$, where α_x is the relative frequency of pluses (+) obtained in the N_1 measurements of $\vec{x} \cdot \vec{\sigma}$ (α_y and α_z are defined similarly). The estimate of \vec{r} is given by $\vec{R}_\chi = R_\chi \vec{n}_\chi$, where we have defined

$$R_\chi = \chi_z, \quad (11)$$

$$\vec{n}_\chi = \vec{x} \sin \hat{\theta} \cos \hat{\phi} + \vec{y} \sin \hat{\theta} \sin \hat{\phi} + \vec{z} \cos \hat{\theta}, \quad (12)$$

and

$$\sin \hat{\theta} = \frac{\sqrt{\chi_x^2 + \chi_y^2}}{R_\chi}, \quad \tan \hat{\phi} = \frac{\chi_y}{\chi_x}. \quad (13)$$

(For $d = 2$ we drop the y component of χ and set $\hat{\phi} = 0$.) This protocol is similar to the one used in Ref. [4], where one was only interested in estimating the purity. The main difference is that, in purity estimation, after the first step one measures the rest of the copies along the estimated direction. In the case studied here, however, part of the copies are used to refine the estimate of the direction. The other main difference is in the model: purity estimation is a one-parameter model, which essentially behaves as a classical problem, and LOCC protocols do attain the optimal (collective) asymptotic accuracy. In contrast, the estimation of the whole density matrix is not classical and collective measurements can and do provide an advantage.

Let us prove that the fidelity of the protocol above attains the separable bound (10). The accuracy with which \vec{n} is estimated in step one can be quantified by the average of $C = \cos \Theta$ over the N_0 outcomes of this first measurement, where Θ is the angle between \vec{n} and its rough guess \vec{n}_0 . One has

$$\langle C \rangle_0 = 1 - \frac{\eta_d(r)}{N_0} + o(1/N_0), \quad (14)$$

where $\eta_3(r) = 3(1/r^2 - 1/5)$ and $\eta_2(r) = (1/r^2 - 1/4)$. A shorthand notation similar to that in (6) and (14) will be used below to denote other averages. That is, $\langle \cdots \rangle_{\vec{r}}$ will stand for the average over the prior distribution $d\rho = dr w(r) dn$. Likewise, the average fidelity (after step two) can be written as $F = \langle f \rangle_{\chi_0 \vec{r}} \equiv (1 + \langle \delta \rangle_{\chi_0 \vec{r}})/2$. Note that the frequencies α_x , α_y , and α_z are binomially distributed as $\alpha_{x,y} \sim \text{Bin}[N_1, (1 + r n_{x,y})/2]$ and $\alpha_z \sim \text{Bin}[N_1, (1 + r C)/2]$, where the components of \vec{n} are referred to the rotated reference frame. Hence, for *large* N_1 (N), the components of χ are close to normally distributed; $\chi_{x,y} \sim N[r n_{x,y}, N_1^{-1/2}(1 - r^2 n_{x,y}^2)^{1/2}]$ and similarly $\chi_z \sim N[r C, N_1^{-1/2}(1 - r^2 C^2)^{1/2}]$.

To compute the asymptotic form of $\Delta = \langle \delta \rangle_{\chi_0 \vec{r}}$, we note that $\delta = \mathbf{r} \cdot \mathbf{R}_\chi = \vec{r} \cdot \vec{R}_\chi + (1 - r^2)^{1/2}(1 - R_\chi^2)^{1/2} \equiv \delta^V + \delta^S$, as can be read off from Eq. (1), and make in $\langle \delta^V \rangle_{\chi_0}$ (no average over the prior) the approximation $R_\chi \approx rC$, along with the substitutions $\sin \hat{\theta} \cos \hat{\phi} = \chi_x/R_\chi$ and $\sin \hat{\theta} \sin \hat{\phi} = \chi_y/R_\chi$. Retaining only terms up to order $\hat{\theta}^2$ (on average, $\hat{\theta}$ is small for large N_1) we have

$$\langle \delta^V \rangle_{\chi_0} \approx r^2 \left\langle C^2 - \frac{\langle \chi_x^2 + \chi_y^2 \rangle_\chi}{2r^2} + \frac{1}{r} \sum_{i=x,y} n_i \langle \chi_i \rangle_\chi \right\rangle_0. \quad (15)$$

(For $d = 2$ we just drop χ_y .)

The average $\langle \chi_x^2 + \chi_y^2 \rangle_\chi$ (or just $\langle \chi_x^2 \rangle_\chi$ if $d = 2$) can be computed trivially recalling that $\chi_{x,y}$ are almost normally distributed random variables. The resulting expression can be written as

$$\langle \delta^V \rangle_{\chi_0} = \frac{r^2}{2} \langle (1 - C)^2 \rangle_0 + r^2 \langle C \rangle_0 - \frac{d-1}{2N_1} + o(1/N_1), \quad (16)$$

where we have used the relation $n_x^2 + n_y^2 = 1 - C^2$ ($n_x^2 = 1 - C^2$ if $d = 2$).

We now observe that $r^2 \langle C \rangle_0$ can be approximated by $\langle rR_\chi \rangle_{\chi_0}$ and that the term $\langle (1 - C)^2 \rangle_0$ is of order N_0^{-2} [see Eq. (14)]. The latter is thus subdominant if $\alpha > 1/2$ and can be dropped. Therefore we obtain (up to the order we are interested in)

$$\Delta = \langle rR_\chi + (1 - r^2)^{1/2} (1 - R_\chi^2)^{1/2} \rangle_{\chi_0} - \frac{d-1}{2N_1} + o(1/N). \quad (17)$$

This is a gratifying result because the term in brackets can be recognized as the average fidelity used in the purity estimation problem discussed in Ref. [4], and we just need to borrow the asymptotic expression obtained there: $F^{\text{purity}} = 1 - 1/(2N_1) + o(1/N_1)$. We have $\Delta = 1 - d/(2N_1) + o(1/N_1) = 1 - d^2/[2(N - N_0)] + o(1/N)$, and therefore

$$F_d = 1 - \frac{d^2}{4N} + o(1/N), \quad (18)$$

which is the separable measurement bound (10).

Some care regarding the constant α that determines the (vanishing) fraction of copies used in the first step must be taken in the above derivation. This constant must be carefully tailored to the specific choice of the prior $w(r)$, Eq. (2). One can show that near $r = 0$ there appears a term of order $\mathcal{O}(1/N^{5\alpha/2})$ while near $r = 1$ there is a term $\mathcal{O}(1/N^{3\alpha/2})$, which comes from the purity estimation part [4]. Therefore the choice $\alpha > 2/3$ renders both terms subdominant. The optimal value of α is hard to find analytically; however, numerical results suggest that it is close to its lower bound, $2/3$.

Let us summarize. We have analyzed LOCC estimation protocols for qubit mixed states. These are the most relevant arrangements for practical purposes. Using statistical tools we have obtained an asymptotic bound on the fidelity for slightly more general approaches; those that use *separable* measurements. Our specific LOCC protocol attains the separable bound. The rate at which perfect determination can be attained is comparable to that of the completely unrestricted optimal protocol, which involves joint measurements: $1 - F$ goes to zero at rate N^{-1} for both separable and collective approaches. The accuracies, however, exhibit a gap, Eq. (10). The separable measurement bounds do not depend on the prior distribution. In view of the fact

that even optimal processing of standard (fixed) tomography leads to accuracies that go to zero more slowly than N^{-1} for the very natural choice of the Bures prior, it is nontrivial and gratifying to exhibit an experimentally feasible LOCC protocol that saturates the separable bounds (showing they are sharp!) and, in particular, has the N^{-1} rate. Our results can be extended to the distillation of pure states [20]. Also they can be applied to higher dimensional systems; e.g., the pointwise approach, Eqs. (6)–(10), shows that the asymptotic expression of the fidelity also satisfies $1 - F = o(1/N)$.

We thank A. Acín, J. Calsamiglia, and W. K. Wootters for useful discussions. We acknowledge financial support from Spanish Ministry of Science and Technology Projects No. BFM2002-02588, No. FIS2005-01369, CIRIT project No. SGR-00185, Netherlands Organization for Scientific Research NWO Project No. 613.003.047, the European Community projects QUPRODIS Contract No. IST-2001-38877 and RESQ Contract No. IST-2001-37559. This work was done while R.D.G. and M.A.B. were at the Mathematical Institute, University of Utrecht.

-
- [1] W. K. Wootters, quant-ph/0506149.
 - [2] E. Bagan, A. Monras, and R. Muñoz-Tapia, Phys. Rev. A **71**, 062318 (2005).
 - [3] M. A. Ballester, J. Phys. A **39**, 1645 (2006).
 - [4] E. Bagan, M. A. Ballester, R. Muñoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **95**, 110504 (2005).
 - [5] S. Virmani, M. F. Sacchi, M. B. Plenio, and D. Markham, Phys. Lett. A **288**, 62 (2001).
 - [6] A. Acín, E. Bagan, M. Baig, L. Masanes, and R. Muñoz-Tapia, Phys. Rev. A **71**, 032338 (2005).
 - [7] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991).
 - [8] R. D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).
 - [9] M. Hayashi and K. Matsumoto, quant-ph/0411073.
 - [10] *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, edited by M. Hayashi (World Scientific, Singapore, 2005).
 - [11] E. Bagan, M. Baig, R. Muñoz-Tapia, and A. Rodríguez, Phys. Rev. A **69**, 010304 (2004).
 - [12] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia, Phys. Rev. A **73**, 032301 (2006).
 - [13] C. H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999).
 - [14] K. Życzkowski and H. J. Sommers, Phys. Rev. A **71**, 032313 (2005).
 - [15] D. Petz and C. Sudar, J. Math. Phys. (N.Y.) **37**, 2662 (1996).
 - [16] M. J. W. Hall, Phys. Lett. A **242**, 123 (1998).
 - [17] M. Hübner, Phys. Lett. A **163**, 239 (1992).
 - [18] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
 - [19] R. D. Gill, math.ST/0512443 [Ann. Stat. (to be published)].
 - [20] O. Romero-Isart *et al.* (to be published).