

Quantum estimation of unknown parameters

Esteban Martínez-Vargas,¹ Carlos Pineda,² François Leyvraz,³ and Pablo Barberis-Blostein¹

¹*Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, Universidad Nacional Autónoma de México, México D. F. 01000, Mexico*

²*Instituto de Física, Universidad Nacional Autónoma de México, México D. F. 01000, Mexico*

³*Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca 62210, Mexico*

(Received 25 June 2016; published 27 January 2017)

We discuss the problem of finding the best measurement strategy for estimating the value of a quantum system parameter. In general the *optimum* quantum measurement, in the sense that it maximizes the quantum Fisher information and hence allows one to minimize the estimation error, can only be determined if the value of the parameter is already known. A modification of the quantum Van Trees inequality, which gives a lower bound on the error in the estimation of a random parameter, is proposed. The suggested inequality allows us to assert if a particular quantum measurement, together with an appropriate estimator, is optimal. An adaptive strategy to estimate the value of a parameter, based on our modified inequality, is proposed.

DOI: [10.1103/PhysRevA.95.012136](https://doi.org/10.1103/PhysRevA.95.012136)

I. INTRODUCTION

The problem of determining the value of an unknown parameter θ from a set of measurements which depend on θ probabilistically, has a long history [1–3]. When one wants to estimate θ , one, in general, does not measure θ or even a value in one to one correspondence with it [4]. One rather obtains a variable \mathbf{y} , generally vectorial, chosen from a probability distribution that depends on θ . The probability of obtaining \mathbf{y} is written as $p(\mathbf{y}|\theta)$. Let us now assume that we have actually measured \mathbf{y} and we consider a function $\hat{\theta}(\mathbf{y})$, which we call the estimator of θ and provides a—necessarily imperfect—estimation of the value of θ .

We would now like to distinguish between good and bad estimators, and to find out when a given estimator is the best possible. We define the *variance* of the estimator as follows:

$$\sigma^2[\hat{\theta}(\mathbf{y}); \theta] := \int d\mathbf{y} p(\mathbf{y}|\theta) (\hat{\theta}(\mathbf{y}) - \theta)^2. \quad (1)$$

A fundamental result states that for an *unbiased* estimator, that is one which averaged over $p(\mathbf{y}|\theta)$ yields the value θ , the variance is bounded from below by

$$\sigma^2[\hat{\theta}(\mathbf{y}); \theta] \geq \left(\int d\mathbf{y} \Omega[p(\mathbf{y}|\theta)] \right)^{-1} = I(\theta)^{-1}. \quad (2)$$

The operator Ω acts on functions, taking $\alpha(\theta)$ to $\Omega[\alpha(\theta)] = (\partial_\theta \ln \alpha)^2 \alpha$. Here $I(\theta)$ is known as the Fisher information. Equation (2) is the so-called Cramér-Rao inequality [5]. Note that the Fisher information only depends on the value of θ , as well as, of course, on the probability distribution $p(\mathbf{y}|\theta)$ and in no way on the estimator. Thus, if we know the probabilistic model $p(\mathbf{y}|\theta)$, this inequality gives a limit to the attainable variance for any unbiased estimator $\hat{\theta}(\mathbf{y})$. It therefore states, among other things, when an estimator is optimal.

However, the inequality is of limited use if one does not know a way of finding estimators which saturate this bound, at least approximately. Such a procedure exists for a broad class of cases and leads to the so-called maximum likelihood estimator. Note that, if the vector \mathbf{y} consists of n independent measurements of a quantity x having probability distribution $p(x|\theta)$, then (2) immediately leads to a lower bound of $1/n$ for the variance. Therefore, in this case, it is impossible to

obtain estimators that are systematically closer to θ than $n^{-1/2}$. This rate of convergence is a central and very general feature in classical statistics. Using quantum mechanics, one may do better [6], but this will not be the point of this paper.

Now let us consider the same problem for a quantum mechanical system. We consider a density matrix $\rho(\theta)$ depending in a known manner on an unknown parameter θ , which one wishes to determine on the basis of measurements performed on the system; for example, in an interferometry experiment we may want to estimate a phase shift due to the presence of a crystal. The crucial difference with the classical case has to do with the issue of measurement, which is more complex in quantum systems. Indeed, *the choice of the observable to be measured determines which aspects of the quantum system will appear*.

We must therefore first define clearly what we mean by measurement. We shall take a somewhat more general definition than that commonly used in textbooks: We define a set of (not necessarily commuting) positive operators $\{E_\xi\}$ indexed by a parameter ξ to be a *positive operator-valued measure* (POVM) if it satisfies $\sum_\xi E_\xi = 1$. The outcome of applying such a POVM to a given density matrix ρ is one of the parameter values ξ , with the probability $p(\xi) = \text{Tr}(\rho E_\xi)$. Note that, if $\{E_\xi\}$ is a set of commuting projectors satisfying the aforementioned identity, the procedure reduces to the traditional quantum mechanical prescription. Advantages of considering POVM's are described in [7].

Once a POVM has been determined, the problem is reduced to a classical one: If the state of the system is given by $\rho(\theta)$ and the POVM is given by $\{E_\xi\}$, the probabilities of obtaining ξ , given the value of θ , are given by

$$p(\xi|\theta) = \text{Tr}[\rho(\theta)E_\xi]. \quad (3)$$

It might now seem that everything is solved and reduced to finding a POVM which maximizes the Fisher information of $p(\xi|\theta)$. The difficulty that arises is the following: This optimal POVM generally depends on the actual value of θ [8], which is, however, always assumed to be unknown. *Indeed, the very problem one is trying to solve is that of estimating θ .*

II. THE VAN TREES INFORMATION

So how can one solve this conundrum? If several copies of the system are available, solutions of this conundrum can be found in the limit where the quantum Fisher information does not depend on the value of the parameter [9] or using adaptive measurements [10]. These strategies depend on (i) having access to several copies of the system and, for the adaptive measurement solution, (ii) the possibility of changing, prior to the measurement of each copy, the POVM to be used. What is the solution if (i) and/or (ii) are not satisfied?

We propose the following: Let us start by introducing a probability distribution $\lambda(\theta)$, according to which we assume the values of θ are distributed. We do not, of course, necessarily have such information, but we surely have some idea of the range of values the parameter θ is liable to take, and we may at least incorporate such knowledge into $\lambda(\theta)$. Instead of the variance defined in (2), we define the quality of the estimator through the quantity

$$\sigma^2[\hat{\theta}(\mathbf{y}); \lambda] := \int d\mathbf{y} d\theta \lambda(\theta) p(\mathbf{y}|\theta) (\hat{\theta}(\mathbf{y}) - \theta)^2, \quad (4)$$

corresponding to the estimator variance *averaged* over the distribution of θ . In such cases, a generalization of the Cramér-Rao inequality has been derived, the so-called Van Trees inequality [11]:

$$\sigma^2[\hat{\theta}(\mathbf{y}); \lambda] \geq Z(\lambda)^{-1}, \quad (5)$$

with

$$Z(\lambda) \equiv \int d\mathbf{y} d\theta \{\Omega[p(\mathbf{y}|\theta)] + \Omega[\lambda(\theta)]\} p(\mathbf{y}|\theta) \lambda(\theta). \quad (6)$$

In words, the lower bound on the average variance is the inverse of the average over the joint distribution of \mathbf{y} and θ of the sum of the Fisher information of $p(\mathbf{y}|\theta)$ and $\lambda(\theta)$. We will call $Z(\lambda)$ the generalized Fisher information [12]. The Cramér-Rao bound (2) can be obtained from (6) when the prior probability distribution $\lambda(\theta)$ is constant and the Fisher information does not depend on the parameter. If $\lambda(\theta)$ is a Dirac delta, $Z(\lambda)$ diverges and the error for the optimal strategy is zero. This makes sense because the prior knowledge of the parameter gives complete knowledge of it.

For a quantum mechanical system described by a density matrix $\rho(\theta)$ depending in a known manner on an unknown parameter, we follow an approach analog to what was done above. Using a POVM we construct a probabilistic model through (3), then we can use (5) to find the optimal estimator. The POVM that maximizes $Z(\lambda)$ —the optimal POVM—together with the estimator that saturates Eq. (5) minimizes the error (4); they will be called the optimal measurement strategy. Let $\mathcal{E} = \{E_\xi\}$ represent the set of all POVMs acting on the quantum system. If we define the quantum Van Trees information as

$$\begin{aligned} Z_Q(\lambda) &= \max_{\{E_\xi\}} \int d\xi d\theta \{\Omega[p_E(\xi|\theta)] + \Omega[\lambda(\theta)]\} p_E(\xi|\theta) \lambda(\theta) \\ &= \max_{\mathcal{E}} Z(\lambda), \end{aligned} \quad (7)$$

a Cramér-Rao-type equation for the quantum case can be written as

$$\sigma^2[\hat{\theta}(\xi); \lambda] \geq [Z_Q(\lambda)]^{-1}. \quad (8)$$

The POVM and estimator that saturates it constitute the optimal strategy for estimating θ .

Inequality (8) resembles the quantum Van Trees inequality, which can be stated defining the generalized quantum Fisher information,

$$V_Q(\lambda) = \int d\xi d\theta \max_{\{E_\xi\}} \{\Omega[p_E(\xi|\theta)] + \Omega[\lambda(\theta)]\} p_E(\xi|\theta) \lambda(\theta),$$

and takes the simple form [12],

$$\sigma^2[\hat{\theta}(\xi); \lambda] \geq (V_Q(\lambda))^{-1}. \quad (9)$$

Notice that the maximization over all POVMs is taken inside the integral whereas in (7) it is taken outside the integral. Since the POVM that maximizes the Fisher information depends on the value of the parameter to estimate, there is not, in general, a single POVM that saturates the quantum Van Trees inequality; in this case the inequality is useless to find the best POVM for parameter estimation. By construction $V_Q \geq Z_Q$. When $V_Q > Z_Q$, the Van Trees inequality predicts smaller errors for the optimal measurement strategy; it can then be argued that the Van Trees inequality (9) gives better results, but as shown above, the POVM that saturates it does not exist and *no strategy exists to achieve the smaller error predicted by (9)*.

If only one measurement over a quantum state $\rho(\theta)$ is allowed, and we codify what we know about the parameter in the probability distribution $\lambda(\theta)$, the inequality (8) tells us that in order to minimize the error (4) we should use the quantum measurement that maximizes $Z(\lambda)$. This POVM does not depend on the parameter we want to estimate and thus, *this way of deciding how to measure the quantum state solves the conundrum discussed above*. In our proposal the optimal measurement strategy depends on what we already know about the parameter, i.e., the *a priori* probability distribution, and not on what will be estimated *after* the measurement. Once the optimal POVM has been found, the problem reduces to a classical one and the mean of the posterior probability distribution minimizes the risk (4) [13]. Note that Van Trees inequalities do not require unbiased estimators, and one thus expects them to provide better error bounds than the usual Cramér-Rao inequalities [14]. This fact underlines the importance of Eq. (8).

III. MEASURING SEVERAL COPIES OF THE SYSTEM

Suppose we have a machine that generates n copies of the quantum state $\rho(\theta)$ and we want to estimate θ with the smallest error. We consider three different strategies to minimize the error in the estimation, each of them depending on the type of measurement that can be done on the n copies. The conundrum described in the introduction does not appear because the POVMs that minimize the error in each of the strategies to be described do not depend on the value of the parameter we want to estimate.

A. Any possible measurement on the n copies

If we can make any possible quantum measurement—including collective ones—over the n copies, the POVM that maximizes the generalized Fisher information, see Eq. (7),

corresponds to the measurement that gives the smallest error given the prior information. This is the best case scenario.

B. Independent measurement on each copy

Let us think of a more plausible situation: The states are provided in such a way that collective measurements over two or more copies are not possible. For example, the machine gives the states sequentially and the time it needs to create the next copy is larger than the decoherence time of the state, so that there are never two or more copies of the same state available to make a collective measurement; then only n independent measurements, each on one of the n copies of the system, can be performed. If the same measurement is going to be performed in each copy, we can use the additive property of the Fisher information for independent measurements to calculate the quantum Van Trees information for n independent measurements performed with the same POVM,

$$Z_Q^I(\lambda) = \max_{\{E_\xi\}} \int d\xi d\theta \{n \times \Omega[p_E(\xi|\theta)] + \Omega[\lambda(\theta)]\} p_E(\xi|\theta) \lambda(\theta). \quad (10)$$

The maximization is done over all the POVMs acting on one copy of the system; the number of measurements performed appears as a factor multiplying the first summand of the integral in the previous equation. The Cramér-Rao bound reduces to

$$\sigma^2[\hat{\theta}(\xi); \lambda] \geq [Z_Q^I(\lambda)]^{-1}. \quad (11)$$

For the case of performing the same measurement over several copies of the system, the risk (4) is minimized by the measurement that maximizes the integral in (10).

C. Adaptive measurements

Let's assume now a setup similar to the previous one but allowing the use of the outcomes of previous measurements to choose how to measure the next copy. We model this situation by choosing n different individual quantum measurements, one for each copy of the state. The quality of this adaptive estimation process is given by (4) with $\mathbf{y} = (\xi_1, \xi_2, \dots, \xi_n)$ the outcomes of measuring n copies of the quantum state $\rho(\theta)$. We now discuss possible ways to choose this quantum measurement.

In [10] the following adaptive scheme is proposed: First, arbitrarily guess a value for the parameter θ . Let us call it θ_1 . Then, use the quantum Fisher information to choose the POVM optimized for θ_1 , make a measurement with outcome ξ , and estimate θ using the maximum likelihood function. We call the second estimation θ_2 . Then, repeat the procedure using the POVM optimized for θ_2 . The procedure is repeated n times. It is shown that in the limit of n going to infinity the procedure saturates the Cramér-Rao inequality and is, in this sense, optimal. Nevertheless, preparing several copies of a quantum state and measuring them can be expensive, difficult, or time consuming; when we are in this situation strategies that are designed to get smaller estimation errors when n is small are of interest.

We propose an adaptive method similar to the adaptive scheme proposed in [10] but using the quantum Van Trees information instead of the quantum Fisher information: We

use the *a priori* knowledge distribution $\lambda(\theta)$ to obtain a POVM $\{E_\xi\}$ which maximizes $Z(\lambda)$. A measurement is carried out with this POVM obtaining ξ as its result with a probability $p(\xi|\theta)$ given by (3). Then, we use the Bayes rule to obtain a new probability distribution,

$$\lambda_1(\theta) = \frac{p(\xi|\theta)\lambda(\theta)}{\int p(\xi|\theta')\lambda(\theta')d\theta'}. \quad (12)$$

Afterwards, we calculate $Z_Q(\lambda)$ using $\lambda_1(\theta)$ as our new *a priori* knowledge distribution. A new POVM is obtained and the process can be iterated from here on n times. Our method has the advantage that all the information we know about the parameter enters in the determination of the optimal POVM. When the number of measurements goes to infinity both methods predict the same result. We now show an example where for a small number of measurements our approach is better than the one in [10].

IV. EXAMPLE

We study the quantum Van Trees information and apply the adaptive estimation process outlined above for the problem of phase estimation with an initial coherent state [4]. Specifically, we consider estimating θ from the state defined by

$$|\alpha(\theta)\rangle = e^{i\hat{n}\theta} |\alpha\rangle, \quad (13)$$

where $|\alpha\rangle$ is the coherent state corresponding to the complex parameter α . The *a priori* distribution $\lambda(\theta)$ is assumed to be a Gaussian centered at 0 and of width σ (we ignore the effect of periodicity in θ). We shall present an analytical approach, which involves a restriction over the considered POVMs, and a numerical one, which will provide an independent check to validate the analytical approach.

The problem of maximizing Fisher information $I(\theta)$ for the phase estimation problem, with pure states, has been solved in [4]. The corresponding POVM belongs to the family characterized by the set of operators,

$$\tilde{\mathcal{E}} = \{|\alpha(\epsilon)\rangle\langle\alpha(\epsilon)|, 1 - |\alpha(\epsilon)\rangle\langle\alpha(\epsilon)|\}_{\epsilon \in [0, 2\pi)}. \quad (14)$$

Here $|\alpha(\epsilon)\rangle$ is the state defined by (13). In particular, the POVM that maximizes I is obtained setting ϵ to the actual value of the parameter to estimate. This leads to a value of $4|\alpha|^2$ for the Fisher information, independent of θ , which in turn means that $V_Q = 4|\alpha|^2 + \sigma^{-2}$.

To calculate the quantum Van Trees information, the maximization in Eq. (7) should be done over all POVMs. However, to achieve analytic results, we shall restrict the optimization to the family Eq. (14), inspired by the similarity of the Fisher information problem and the current one. Coincidence with the numeric method will justify this massive simplification.

If we define $p(\theta, \epsilon) = |\langle\alpha(\theta)|\alpha(\epsilon)\rangle|^2$, the conditional probability distribution reads $p(1|\theta) = p(\theta, \epsilon)$ and $p(2|\theta) = 1 - p(\theta, \epsilon)$. The Fisher information for a POVM with parameter ϵ is then

$$F(\theta, \epsilon) = \frac{1}{p(\theta, \epsilon)(1 - p(\theta, \epsilon))} \left[\frac{dp(\theta, \epsilon)}{d\theta} \right]^2. \quad (15)$$

Notice that θ is the parameter we want to estimate and ϵ is the label that enumerates different POVMs. We restrict the

maximization in (7) to the subset $\tilde{\mathcal{E}}$ and obtain

$$Z_Q^{\tilde{\mathcal{E}}} \equiv \max_{\tilde{\mathcal{E}}} \left[\int \frac{1}{p(\theta, \epsilon)(1 - p(\theta, \epsilon))} \left[\frac{dp(\theta, \epsilon)}{d\theta} \right]^2 \lambda(\theta) d\theta \right] + \int \left(\frac{\partial \ln \lambda(\theta)}{\partial \theta} \right)^2 \lambda(\theta) d\theta. \quad (16)$$

Since the maximization is done over a proper subset of all POVMs, $Z_Q^{\tilde{\mathcal{E}}} \leq Z_Q$. With the additional assumption $|\alpha|^2 \ll 1$ we find, after a lengthy but straightforward calculation [15],

$$Z_Q^{\tilde{\mathcal{E}}} \approx 2|\alpha|^2(e^{-\frac{\sigma^2}{2}} + 1) + \frac{1}{\sigma^2}. \quad (17)$$

The first term of the sum coincides with the optimized Fisher information when $\sigma \rightarrow 0$, as expected. For large σ , $Z_Q^{\tilde{\mathcal{E}}}$ is smaller than the quantum Fisher information, as we are estimating a parameter about which we have little information.

To numerically estimate Z_Q we first truncate the infinite dimensional Hilbert space, in which Eq. (13) lives, to a finite dimension, larger than $|\alpha|$. Next, we recall that any POVM is equivalent to a projective measurement in a larger Hilbert space [7]. We shall vary the dimension of such an enlarged Hilbert space until the value of the Van Trees information no longer changes up to the first two significant digits.

Notice that the columns of any unitary matrix defines an orthonormal basis, whose elements define a projective measurement (up to arbitrary values of the measurement outcome). We select the aforementioned unitary matrix from the Gaussian unitary ensemble [16], which guarantees a uniform exploration of all orthonormal bases. Such an ensemble is composed of all unitary matrices, and is weighted by its Haar measure. That way we obtained a POVM in the original space, and thus a particular value of the integral within Eq. (7). Repeating the procedure many times, we approach the value that maximizes $Z(\lambda)$. We add two remarks. First, the value of the dimension of the truncated Hilbert space was varied, for fixed α , until the correction was smaller than could be noticed by visual inspection of the plot. Second, a downhill method to optimize the orthonormal basis in the enlarged Hilbert space, to fine tune the POVM, was used. However, this procedure only provided a small advantage, suggesting that the landscape of this very large dimensional space had no small and deep depressions.

In Fig. 1 it can be seen that our analytical approximation Eq. (16) gives very good agreement with the numerical simulation, suggesting that indeed $Z_Q^{\tilde{\mathcal{E}}} \approx Z_Q$. In that figure we also show that the quantum Van Trees information is significantly smaller than V_Q ; *this means that there does not exist a single POVM that saturates the quantum Van Trees inequality* (9), and thus an optimal strategy must come from the optimization as performed in Z_Q and not V_Q .

Now we discuss the adaptive method proposed in this paper to estimate the parameter, assuming that one has several copies of $\rho(\theta)$. We shall compare our results with the method proposed in [10]. For both methods there are 2^n possible outcomes after n binary measurement; its probability distribution $p(y|\theta)$ and the estimator depend on the method used.

We assume that the *a priori* probability distribution is flat, $\lambda_0(\theta) = 1/2\pi$. This is the worst case scenario: We only know

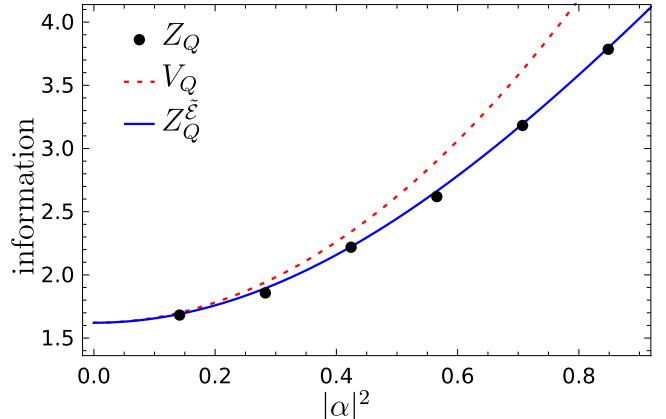


FIG. 1. The blue continuous curve is our analytical approximation to the quantum Van Trees information for phase estimation as a function of the coherent state $|\alpha\rangle$. The *a priori* probability distribution is a Gaussian centered in zero with width $\sigma = \pi/4$. The black dots are calculated numerically. The red dashed curve is the generalized quantum information.

that the parameter to be measured is an angle between 0 and 2π . For each possible value of the parameter to be estimated, θ_r , we simulate both adaptive methods. After n measurements and assuming that the optimal estimator is used, the smallest error for the quantum Fisher adaptive method for θ_r is $\sigma_n^2(\theta_r) = [I_n(\theta_r)]^{-1}$, where $I_n(\theta_r)$ is the Fisher information for the probability distribution of the 2^n possible outcomes. Then, the mean error for the quantum Fisher information adaptive method is

$$\overline{\sigma^2}_{\text{Fisher}}(n) = \int_{-\pi}^{\pi} \lambda_0(\theta) \sigma_n^2(\theta) d\theta_r = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{I_n(\theta)} d\theta_r. \quad (18)$$

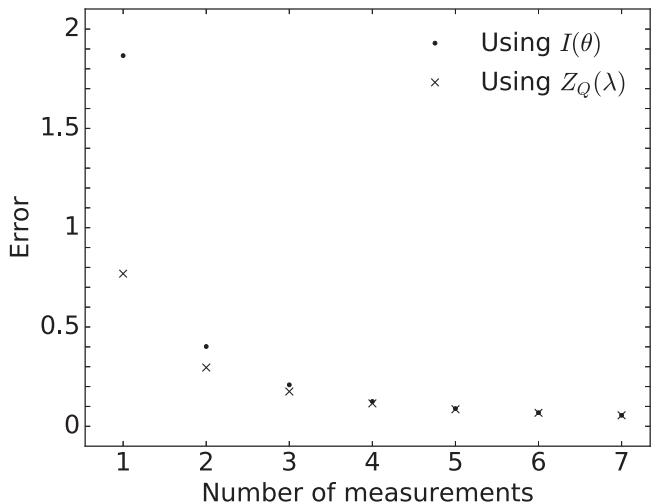


FIG. 2. Comparison of the smallest mean error predicted by the Fisher information [Eq. (18)], with the smallest mean error predicted by the quantum Van Trees information, for two adaptive quantum estimation schemes. The points use the quantum Fisher information as a tool to choose the POVM to be used in the next measurement. The crosses use the quantum Van Trees information to choose the POVM to be used in the next measurement; with this scheme we obtain smaller errors.

After n measurements the smallest error for the quantum Van Trees information adaptive method is $\sigma_{\text{VanTrees}}^2(n) = 1/Z_{Q_n}(\lambda_{n-1})$, where λ_{n-1} is the prior distribution probability for the parameter to estimate after $n-1$ measurements, and $Z_{Q_n}(\lambda_{n-1})$ is the quantum Van Trees information for that prior probability distribution.

In Fig. 2 we compare $\overline{\sigma^2}_{\text{Fisher}}(n)$ with $\sigma_{\text{VanTrees}}^2(n)$; it can be seen that the average estimation error is smaller using the quantum Van Trees information adaptive method than the quantum Fisher adaptive method. As expected, both methods tend to give the same error as the number of measurements increases.

If the adaptive step cannot be implemented, the same measurement is performed in all the copies. Using the additive property for the Fisher information for independent measurements (see Sec. III B), we get that $\overline{\sigma^2}_{\text{Fisher}}(n) \gtrsim 1.9/n$ and $\sigma_{\text{VanTrees}}^2(n) \gtrsim 0.8/n$; for both cases the error scales the same with respect to the number of measurements, but if we use inequality (8) to choose the POVM to be implemented, the error in the estimation can be smaller by more than a factor of two.

V. CONCLUSIONS

Assume we have a quantum state that depends on an unknown parameter chosen from a known distribution. A central problem in quantum metrology is to determine as

accurately as possible the parameter from measurements on the state. It is customary to use the quantum Fisher information to find the optimal measurement, which, in general, depends on the unknown value of the parameter we want to estimate. Using an inequality proposed in this paper this problem is solved. The inequality bounds the error in determining the parameter and depends on its prior distribution. This bound can thus be used to find a quantum measurement whose results applied to the appropriate estimator gives the minimum error. The most important application of this approach consists of determining the optimal way to use whatever *a priori* information is available in the best possible way. This is particularly important if we can only perform one, or a very small, number of measurements. We propose an adaptive quantum estimation scheme, based on the inequality, that can be used when several copies of the system are available but collective measurements are not possible.

ACKNOWLEDGMENTS

We acknowledge support from UNAM-PAPIIT Grants No. IN111015, No. IN114014, and No. IN103714, and CONACyT Grant No. 254515. We also acknowledge support from the high performance computer laboratory (LUCAR) at IIMAS, as well as useful discussions with Carlton M. Caves, Mankei Tsang, and Luiz Davidovich.

-
- [1] C. Helstrom, *Phys. Lett. A* **25**, 101 (1967).
 - [2] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
 - [3] V. Giovannetti, S. Lloyd, and L. Maccone, *Nat. Photon.* **5**, 222 (2011).
 - [4] B. M. Escher, R. L. Matos Filho, and L. Davidovich, *Braz. J. Phys.* **41**, 229 (2011).
 - [5] R. A. Fisher, *Phil. Trans. R. Soc. Lond. A* **222**, 309 (1922); H. Cramér, *Mathematical Methods of Statistics*, *Princeton Landmarks in Mathematics and Physics* (Princeton University Press, Princeton, 1945); C. R. Rao, *Bull. Calcutta Math. Soc.* **37**, 81 (1945); *Resonance - J. Sci. Edu.* **20**, 76 (2015).
 - [6] M. J. Holland and K. Burnett, *Phys. Rev. Lett.* **71**, 1355 (1993).
 - [7] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, New York, 2011).
 - [8] O. E. Barndorff-Nielsen and R. D. Gill, *J. Phys. A: Math. Gen.* **33**, 4481 (2000).
 - [9] M. Jarzyna and R. Demkowicz-Dobrzaski, *New J. Phys.* **17**, 013010 (2015).
 - [10] A. Fujiwara, *J. Phys. A: Math. Gen.* **39**, 12489 (2006).
 - [11] R. D. Gill and B. Y. Levit, *Bernoulli* **1**, 59 (1995).
 - [12] M. G. A. Paris, *Int. J. Quantum. Inform.* **07**, 125 (2009).
 - [13] R. Blume-Kohout, *New J. Phys.* **12**, 043034 (2010).
 - [14] M. Tsang, [arXiv:1605.03799](https://arxiv.org/abs/1605.03799).
 - [15] E. Martinez, C. Pineda, F. Leyvraz, and P. Barberis-Blostein (unpublished).
 - [16] R. Balian, *Il Nuovo Cimento B* **57**, 183 (1968).