Entanglement is not very useful for estimating multiple phases

Manuel A. Ballester*

Department of Mathematics, University of Utrecht, Box 80010, 3508 TA Utrecht, The Netherlands
(Received 8 April 2004; published 13 September 2004)

The problem of the estimation of multiple phases (or of commuting unitaries) is considered. This is a submodel of the estimation of a completely unknown unitary operation where it has been shown in recent works that there are considerable improvements by using entangled input states and entangled measurements. Here it is shown that, when estimating commuting unitaries, there is practically no advantage in using entangled input states or entangled measurements.

I. INTRODUCTION

A unitary operation is a map that transforms a density operator ρ_0 on $\mathbb{C}^d$ to another density operator $\rho = U\rho_0 U^\dagger$ on $\mathbb{C}^d$, where $U \in \mathrm{SU}(d)$ is a $d \times d$ special unitary matrix. Suppose one is given a device that performs an unknown U. One can learn something about U by learning about how it transforms a known state ρ_0. In order to completely determine a unitary operation one would need to know how it transforms a basis of $\mathbb{C}^d$ plus some linear combinations thereof. This is known as *quantum process tomography* [5]. More precisely, to estimate U one prepares many copies of the necessary input states and performs a measurement on the output that they produce. As a result, some classical data are obtained and from that one can estimate U. This is shown schematically in Fig. 1.

Another approach (used in Refs. [1–4]) is to prepare a bipartite entangled input state ρ_0 on $\mathbb{C}^d \otimes \mathbb{C}^d$ and then use one of the parts as input for U, while nothing is done to the other part, as shown in Fig. 2. The effect of the operation is to transform the state ρ_0 to $(U \otimes \mathbb{1})\rho_0(U^\dagger \otimes \mathbb{1})$. This output state is then measured and estimated, and, since in this case there is a one-to-one relation between the output state and U, one gets an estimate for U as well. The advantages of this method with respect to quantum process tomography are that only one input state is needed, and that there is potentially a better accuracy in the estimation (if nonseparable measurements are used) [4].

In this paper a relatively less difficult problem will be studied, the estimation of unitary operations that commute with one another, that is, only a maximal Abelian subgroup of $\mathrm{SU}(d)$ is considered instead of the whole group. In this case, the number of unknown parameters decreases from $d^2 - 1$ to $d - 1$. This problem has already been addressed in Ref. [6] where it is given the name of "multiple phase estimation" (MPE). One would like to know whether it is also advantageous to use an entangled input in MPE. In what follows the MPE model that uses entanglement (Fig. 2) will be referred to as MPEE and the one that does not use it (Fig. 1) will be referred to as MPEU.

*Electronic address: ballester@math.uu.nl;
URL: http://www.math.uu.nl/people/balleste/

There are two things that need to be optimized here, the input state and the measurement that is to be performed. Therefore one needs a quantitative measure of how good an input state is and of how good a measurement is.

II. PRELIMINARIES

In this section, the necessary concepts of quantum Fisher information (QFI) and Fisher information (FI) will be introduced, and the quantum Cramér-Rao bound (QCRB) will be stated. The QFI and the FI will be used as measures of the performance of an input state and a measurement, respectively. The QCRB relates these two quantities in a nice way.

A. QFI

Suppose that the quantum state density matrix ρ on $\mathbb{C}^d$ is parametrized by $\theta \in \Theta \subset \mathbb{R}^p$ where p is the number of parameters. In our case ρ would be the output state and $p = d - 1$. Define the symmetric logarithmic derivatives (SLDs) $\lambda_1, \ldots, \lambda_p$ as the self-adjoint operators that satisfy

$$\rho_{,i}(\theta) = \partial_{\theta_i}\rho(\theta) = \frac{1}{2}[\rho(\theta)\lambda_i(\theta) + \lambda_i(\theta)\rho(\theta)].$$

For pure states, $\rho = |\psi\rangle\langle\psi|$, they simply are $\lambda_i = 2\rho_{,i}$. The QFI is defined as the $p \times p$ matrix with elements

$$H_{ij}(\theta) = \mathrm{Re}\,\mathrm{tr}[\rho(\theta)\lambda_i(\theta)\lambda_j(\theta)]$$

which for pure states reduces to

$$H_{ij}(\theta) = \mathrm{Re}\langle l_i(\theta)|l_j(\theta)\rangle,$$

where $|l_i(\theta)\rangle = \lambda_i(\theta)|\psi(\theta)\rangle$.

The $|l\rangle$ vectors have a simple geometric interpretation. Suppose one has a pure state model parametrized by $\theta \in \Theta \subset \mathcal{R}$. For simplicity, the parameter has been taken to be one dimensional. Denote the set of all state vectors by $\mathcal{H} = \{|\psi(\theta)\rangle \,|\, \theta \in \Theta\}$ and the set of all density operators by $\mathcal{M} = \{\rho(\theta) \,|\, \theta \in \Theta\}$. State vectors and density operators are related by the map $\pi: |\psi\rangle \mapsto |\psi\rangle\langle\psi|$, $\mathcal{H} \to \mathcal{M}$. This map is many to one since $\pi(|\psi\rangle) = \pi(e^{i\phi}|\psi\rangle)$ where ϕ is an arbitrary real phase. This means that $\pi^{-1}(\rho_0)$ is a circle in $\mathcal{H}$ and that a curve in $\mathcal{M}$ is mapped through π^{-1} to a tube in $\mathcal{H}$. Conversely, all curves that lie on the surface of the tube, are

FIG. 1. Quantum process tomography.

FIG. 2. Entanglement is used, $\mathbb{C}^d \otimes \mathbb{C}^d$ model.
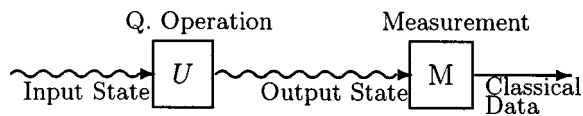
mapped through π to the same curve in $\mathcal{M}$. Then out of all curves $|\psi(\theta)\rangle \in \mathcal{H}$ satisfying $|\psi(0)\rangle = |\psi_0\rangle$ and $\pi(|\psi(\theta)\rangle) = \rho(\theta)$ ($\rho(\theta)$ is a given curve in $\mathcal{M}$) there is a minimal curve $|\widetilde{\psi}(\theta)\rangle$, defined as the one that at every θ goes from the circle $\pi^{-1}[\rho(\theta)]$ to the circle $\pi^{-1}[\rho(\theta + \delta\theta)]$ using the shortest path. $|l(\theta)\rangle$ is a vector pointing in the direction of that shortest path. It can then be calculated as $|l(\theta)\rangle = 2\partial_\theta|\widetilde{\psi}(\theta)\rangle$.

B. FI

Take a positive operator valued measure (POVM) with elements $M_1, \ldots, M_n$. This POVM induces a probability distribution given by $p_\xi(\theta) = \operatorname{tr} \rho(\theta) M_\xi$, the probability to obtain outcome ξ if the parameter has the value θ. The Fisher information for this measurement is defined as the $p \times p$ matrix with elements

$$I_{ij}(M, \theta) = \mathbb{E}_{M,\theta}[\partial_{\theta_i} \ln p_\xi(\theta) \partial_{\theta_j} \ln p_\xi(\theta)].$$

For an estimator $\hat{\theta}$ and a measurement M, locally unbiased at θ_0,[1] the (classical) Cramér-Rao bound is satisfied

$$V(M, \theta_0, \hat{\theta}) \geq I(M, \theta_0)^{-1},$$

i.e., the FI is the smallest variance that a locally unbiased estimator based on this measurement can have. This also means that, if one of the eigenvalues of I is zero, then the variance of the function of the parameters corresponding to that eigenvalue is infinity and therefore cannot be estimated.

If one has N copies of the quantum state and performs the same measurement on each of the copies then the FI of the N copies, I^N, satisfies $I^N(M, \theta) = NI(M, \theta)$ where $I(M, \theta)$ is the FI of one system. It follows that

$$V^N(M, \theta_0, \hat{\theta}) \geq I^N(M, \theta_0)^{-1} = I(M, \theta_0)^{-1}/N.$$

It is a well known fact in mathematical statistics that the maximum likelihood estimator in the limit of large N is asymptotically unbiased and saturates the classical Cramér-Rao bound. Moreover no other reasonable estimator (unbiased or not) can do better.

C. QCRB

The QCRB states that for any measurement M

$$I(M, \theta) \leq H(\theta). \tag{1}$$

In other words, $H(\theta) - I(M, \theta)$ is a positive semidefinite matrix.

This bound is not achievable in general. A theorem due to Matsumoto [7] states that for pure states the bound is achievable at $\theta = \theta_0$ if and only if

$$\operatorname{Im}\langle l_i(\theta_0)|l_j(\theta_0)\rangle = 0. \tag{2}$$

If a model satisfies the above condition, it is said to be quasiclassical at θ_0. Furthermore, if condition (2) holds, there is a measurement with $p+2$ elements that achieves the bound. In fact, any measurement of the type

$$M_\alpha = |b_\alpha\rangle\langle b_\alpha|, \quad \alpha = 1, \ldots, p+1,$$

$$M_{p+2} = 1 - \sum_{\alpha=1}^{m+1} M_\alpha,$$

$$|b_\alpha\rangle = \sum_{\beta=1}^{p+1} o_{\alpha\beta}|m_\beta\rangle,$$

$$|m_k\rangle = \sum_l (H^{-1/2})_{kl}|l_l\rangle, \quad |m_{p+1}\rangle = |\phi\rangle, \tag{3}$$

with o a $(p+1) \times (p+1)$ real orthogonal matrix satisfying $o_{\alpha,p+1} \neq 0$ achieves $I(M, \theta_0) = H(\theta_0)$.

Now one can see why these quantities are a good measure of the performance of input states and measurements. From $V \geq I^{-1}/N \geq H^{-1}/N$ one can see that a good input state is one that achieves an H as large as possible and a good measurement is one that achieves an I as large as possible. Since I and H are matrices the best input state and best measurement can not always be decided unambiguously. This ambiguity will not be completely resolved in the case of choosing an input state. However, in the case of choosing an optimal measurement there will be no ambiguity. In the next section, it will be shown that for every input state there is a measurement on the output state that achieves equality in Eq. (1). Therefore, an optimal measurement is one that achieves equality in the QCRB.

III. MPE IS A QUASICLASSICAL MODEL

It will be shown here that both MPEE and MPEU are quasiclassical everywhere (for all $\theta \in \mathbb{R}^{d-1}$) and for any input state. In particular, this means that only one input state is necessary in MPEU also. In this respect MPE is quite different from the estimation of a completely arbitrary U. Actually, since MPEU can be considered as a special case of MPEE, one needs to show quasiclassicality only for MPEE.

The MPEE model is

---

[1]This means that the expectation of the estimator satisfies $\mathbb{E}_{M,\theta_o}(\hat{\theta}_i) = \theta_{0i}$ and $\partial_{\theta_j} \mathbb{E}_{M,\theta}(\hat{\theta}_i)|_{\theta = \theta_0} = \delta_{ij}$
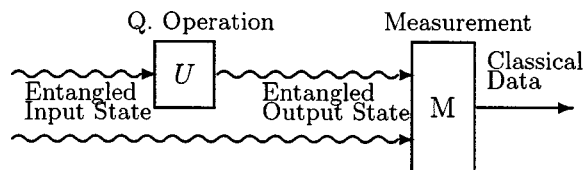
$$\rho(\theta) = U_\theta \rho_0 U_\theta^\dagger, \qquad (4)$$

where

$$U_\theta = \exp\left(i \sum_{m=1}^{d-1} \theta_m T_m \right) \otimes \mathbb{1},$$

ρ_0 on $\mathbb{C}^d \otimes \mathbb{C}^d$ is a pure state density matrix,[2] and $T_1, \ldots, T_{d-1}$ are selfadjoint traceless matrices that commute with one another. They are chosen so that they satisfy an orthonormality condition

$$\mathrm{tr}\, T_m T_n = \delta_{nm}. \qquad (5)$$

The SLDs are

$$\lambda_m = 2\partial_m \rho(\theta) = 2i[T_m \otimes \mathbb{1}, \rho(\theta)]$$

and

$$\mathrm{tr}\, \rho \lambda_m \lambda_n = 4\{\mathrm{tr}\, \rho_0 (T_m T_n \otimes \mathbb{1}) - \mathrm{tr}[\rho_0 (T_m \otimes \mathbb{1})\rho_0 (T_n \otimes \mathbb{1})]\}.$$

It is easy to see that due to the commutativity of the T's this quantity is real and therefore the model is quasiclassical, i.e., it satisfies the condition (2). Therefore for every θ there exists a measurement M (which may depend on θ) such that $I(M, \theta) = H(\theta)$. Furthermore, if one has a large number N of copies, performs the same (optimal) measurement on all N, and calculates the MLE, the mean square error should behave as

$$V^N(M, \theta, \hat{\theta}_{MLE}) = H(\theta)^{-1}/N + o(1/N). \qquad (6)$$

IV. OPTIMAL INPUT STATE

It is now clear from Eq. (6) that one needs to choose the input state so that the QFI is "large." Suppose there is an input state that has a QFI that is larger than or equal to the QFI of any other input state. In that case one can unambiguously choose that state as the optimal one. Unfortunately, in our case there is not such a state. Furthermore, there are situations in which one has two input states ρ_1 and ρ_2 with QFI H_1 and H_2, respectively, which satisfy neither $H_1 \leq H_2$ nor $H_1 \geq H_2$. This is resolved by maximizing a quantity like $\mathrm{Tr}\, GH$, with G a real positive semidefinite matrix. By doing this one assigns relative weights to the mean square errors of the different parameters. Furthermore, achieving maximum $\mathrm{Tr}\, GH$ ensures that no other input state can have a larger QFI. In this paper a particular choice is made: all parameters are given the same importance, i.e., the input state will be chosen so that it maximizes $\mathrm{Tr}\, H$. With this particular choice it is possible to obtain nice analytic results and, since the trace of the QFI is parametrization invariant, the results obtained will not depend on the chosen parametrization.

A. MPEE

Since the self-adjoint matrices $T_1, \ldots, T_{d-1}$ commute with one another, there is a basis where all of them are diagonal

$\{|1\rangle, \ldots, |d\rangle\}$ (this basis is considered to be known). From this point, all calculations will be made in this basis.

The input state is $\rho_0 = |\Psi_0\rangle\langle\Psi_0|$, and $|\Psi_0\rangle$ can be expanded as $|\Psi_0\rangle = \Sigma_{kl} R_{kl}|kl\rangle$. The partial trace is then $\mathrm{tr}_B |\Psi_0\rangle\langle\Psi_0| = RR^\dagger$, and since $RR^\dagger$ is self-adjoint and has trace 1, it can be written as

$$RR^\dagger = \frac{\mathbb{1}}{d} + \sum_{\alpha=1}^{d^2-1} t_\alpha T_\alpha,$$

where, in general, the last sum includes all generators of the Lie algebra $su(d)$. Then the QFI is

$$\begin{aligned} H_{mn} &= 4[\langle\Psi_0|(T_m T_n \otimes \mathbb{1})|\Psi_0\rangle \\ &\quad - \langle\Psi_0|(T_m \otimes \mathbb{1})|\Psi_0\rangle\langle\Psi_0|(T_n \otimes \mathbb{1})|\Psi_0\rangle] \\ &= 4[\mathrm{tr}(RR^\dagger T_m T_n) - (\mathrm{tr}\, RR^\dagger T_m)(\mathrm{tr}\, RR^\dagger T_n)] \\ &= 4[\mathrm{tr}(RR^\dagger T_m T_n) - t_m t_n] \qquad (7) \end{aligned}$$

and its trace is

$$\mathrm{Tr}\, H = 4\left[\mathrm{tr}\left(RR^\dagger \sum_{m=1}^{d-1} T_m^2 \right) - \sum_{m=1}^{d-1} t_m^2 \right].$$

The commuting T's can be written as $T_m = \Sigma_{k=1}^d c_{mk}|k\rangle\langle k|$. The tracelessness condition implies $\Sigma_{k=1}^d c_{mk} = 0$, while the orthonormality condition (5) implies $\Sigma_{k=1}^d c_{mk} c_{nk} = \delta_{mn}$. These two together lead to

$$\sum_{m=1}^{d-1} c_{mk} c_{ml} = \delta_{kl} - \frac{1}{d}$$

and then to

$$\sum_{m=1}^{d-1} T_m^2 = \sum_{k=1}^d \sum_{m=1}^{d-1} c_{mk}^2 |k\rangle\langle k| = \frac{d-1}{d}\mathbb{1}.$$

Substituting this in the equation for the trace, one gets

$$\mathrm{Tr}\, H = 4\left[\frac{d-1}{d} - \sum_{m=1}^{d-1} t_m^2 \right],$$

and one immediately sees that $\mathrm{tr}\, H$ is maximal if and only if $t_m = 0$, $m = 1, \ldots, d-1$. Of course the rest of the t's can be anything (as long as $RR^\dagger$ remains positive).

Concluding, any pure input state ρ_0 satisfying

$$\mathrm{tr}\, \rho_0(T_m \otimes \mathbb{1}) = 0, \forall \quad m = 1, \ldots, d-1,$$

achieves a QFI $H_{mn} = 4\, \delta_{mn}/d$ which has the maximum possible trace among all QFIs. In particular, maximally entangled states satisfy this condition.

In the full $SU(d)$ model one obtains a similar result, and the maximum is attained at and only at the maximally entangled state; this will be shown in the Appendix.

B. MPEU

It will be shown here that for every entangled input state $|\Psi_0\rangle = \Sigma_{kl} R_{kl}|kl\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ and every value of the

---

[2]ρ_0 can be taken to be pure since it was shown in [8] that the QFI is convex.

$(d-1)$-dimensional parameter θ there is an input state $|\psi_0\rangle \in \mathbb{C}^d$ that achieves the same QFI everywhere. This means that in this model, unlike in the full SU(d) model, there is no improvement in the accuracy of the estimation by using entangled inputs.

The model is now

$$\rho(\theta) = U_\theta \rho_0 U_\theta^\dagger,$$

where

$$U_\theta = \exp\left(i \sum_{m=1}^{d-1} \theta_m T_m \right)$$

and $\rho_0 = |\psi_0\rangle\langle\psi_0|$ is a pure state density matrix. Of course this model is also quasiclassical, and by a calculation identical to the one made in the $\mathbb{C}^d \otimes \mathbb{C}^d$ case one gets that the QFI is

$$H_{mn} = 4[\langle\psi_0|T_m T_n|\psi_0\rangle - \langle\psi_0|T_m|\psi_0\rangle\langle\psi_0|T_n|\psi_0\rangle].$$

It is not difficult to see that any input state of the form

$$|\psi_0\rangle = \sum_{k=1}^{d} \sqrt{\langle k|RR^\dagger|k\rangle} e^{i\phi_k}|k\rangle,$$

where the ϕ's are arbitrary phases, achieves

$$H_{mn} = 4[\langle\Psi_0|(T_m T_n \otimes \mathbb{1})|\Psi_0\rangle - \langle\Psi_0|(T_m \otimes \mathbb{1})|\Psi_0\rangle$$
$$\times \langle\Psi_0|(T_n \otimes \mathbb{1})|\Psi_0\rangle],$$

the QFI when the input is the entangled state $|\Psi_0\rangle$.

In particular, any input state of the form

$$|\psi_0\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} e^{i\phi_k}|k\rangle \tag{8}$$

achieves the maximum trace of the QFI. A state of this form (with all the ϕ's set to zero) was used in Ref. [6].

V. OPTIMAL MEASUREMENT

A. MPEU

This model (and actually also the MPEE) has the property that $I(\theta, U_\theta M U_\theta^\dagger) = I(0, M)$, which is easy to prove; the only necessary ingredient is that $U_\theta^\dagger \partial_{\theta_k} U_\theta = \partial_{\theta_k} U_\theta|_{\theta=0}$. Therefore, if one has an optimal measurement M at $\theta=0$, then the measurement $U_\theta M U_\theta^\dagger$ will be optimal at θ.

One can now use the recipe given by Eq. (3) to find an optimal measurement at $\theta=0$. The optimal input state[3] given by Eq. (8) is used. At the origin one has

$$|l_n\rangle = \frac{2i}{\sqrt{d}} \sum_{k=1}^{d} c_{nk} e^{i\phi_k}|k\rangle,$$

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} e^{i\phi_k}|k\rangle.$$

From these vectors one can form an orthonormal set

---

[3]That is, the one that achieves maximum trace of the QFI.

$$|m_n\rangle = i \sum_{k=1}^{d} c_{nk} e^{i\phi_k}|k\rangle, \quad n = 1, \dots, d-1,$$

$$|m_d\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} e^{i\phi_k}|k\rangle$$

and with the choice $o_{kl} = \delta_{kl} - 2/d$ one gets the set of orthonormalized states onto which the measurement elements will project,

$$|b_k\rangle = |m_k\rangle - \frac{2}{d} \sum_{l=1}^{d} |m_l\rangle, \quad k = 1, \dots, d. \tag{9}$$

The optimal measurement at $\theta=0$ has elements $M_k = |b_k\rangle\langle b_k|$ and, since the above vectors form an orthonormal basis of $\mathbb{C}^d$, they satisfy $\Sigma_{k=1}^{d} M_k = \mathbb{1}$. The optimal measurement at θ has elements $U_\theta|b_k\rangle\langle b_k|U_\theta^\dagger$.

Note that the above choice of the orthogonal matrix o works for $d \geq 3$; for $d=2$ another matrix must be chosen. The $d=2$ case will be treated in the following example.

Example V1 ($d=2$). The orthonormal set formed from the input state and $|l\rangle$ is

$$|m_1\rangle = \frac{i}{\sqrt{2}}(|0\rangle - |1\rangle),$$

$$|m_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle). \tag{10}$$

These vectors are then rotated to obtain

$$|b_1\rangle = \cos\eta|m_1\rangle - \sin\eta|m_2\rangle = \frac{i}{\sqrt{2}}(e^{i\eta}|0\rangle - e^{-i\eta}|1\rangle),$$

$$|b_2\rangle = \sin\eta|m_1\rangle + \cos\eta|m_2\rangle = \frac{1}{\sqrt{2}}(e^{i\eta}|0\rangle + e^{-i\eta}|1\rangle),$$

$$\tag{11}$$

where η must satisfy $\sin\eta \neq 0$ and $\cos\eta \neq 0$. The measurement elements are $|b_1\rangle\langle b_1|$ and $|b_2\rangle\langle b_2|$. The FI of this measurement at the origin is equal to the QFI (equal to 2) as expected. Furthermore, this equality happens to hold everywhere and not only at the origin. This feature is very useful in practice, it means that the optimal measurement does not depend on the actual (unknown) value of the parameter and therefore an adaptive scheme is not necessary. Whether it is possible to find measurements with this characteristic for $d>2$ is an open problem.

B. MPEE

An optimal measurement in this case can also be derived using the recipe given by Eq. (3). In general, such a measurement is a joint measurement on the two output systems. One could ask whether it is possible to achieve the bound with local measurements and classical communication. In what follows, it will be shown that this is indeed possible.

The input state is taken to be the maximally entangled state $\Sigma_{k=1}^{d}|kk\rangle/\sqrt{d}$. Alice measures the system coming out of U and Bob measures the other one. The strategy is the following. Bob performs the von Neumann measurement $B_k = |w_k\rangle\langle w_k|$ with $\Sigma_k B_k = 1$ on his system where

$$|w_k\rangle = \frac{1}{\sqrt{d}} \sum_{l=1}^{d} \exp\left(\frac{2\pi kl}{d}i\right)|l\rangle, \quad k = 1, \dots, d.$$

He obtains outcome k with probability $1/d$; he then phones Alice and tells her the outcome of his measurement. The net result of this is that Bob prepares the state

$$\frac{1}{\sqrt{d}} \sum_{l=1}^{d} \exp\left(-\frac{2\pi kl}{d}i\right)|l\rangle$$

at the input of U. It is easy to recognize this state as one of the optimal states in the $\mathbb{C}^d$ case. Now Alice can perform the measurement A_{kl} with $\Sigma_l A_{kl} = 1$, which is the optimal measurement described above for the $\mathbb{C}^d$ case, and where the arbitrary phases are now fixed to $\phi_l = 2\pi k/d$. It is not difficult to check that this measurement indeed achieves equality in the QCRB at $\theta = 0$. The measurement on $\mathbb{C}^d \otimes \mathbb{C}^d$ is then $\Sigma_{kl} A_{kl} \otimes B_k = 1 \otimes 1$ and the measurement $\Sigma_{kl} U_\theta A_{kl} U_\theta^\dagger \otimes B_k = 1 \otimes 1$ is optimal at θ.

If this is applied to the $d = 2$ case and one uses the θ-independent optimal measurement derived for MPEU, a θ-independent optimal measurement is obtained for MPEE.

VI. ASYMPTOTIC FIDELITY

From the results obtained previously, one can infer the asymptotic behavior of the average fidelity. Indeed, in Ref. [9], it was established that the fidelity between nearby states is given by[4]

$$\mathcal{F}(\theta, \theta + \delta\theta) = (\mathrm{tr}\sqrt{\sqrt{\rho(\theta)}\rho(\theta + \delta\theta)\sqrt{\rho(\theta)}})^2$$

$$= 1 - \sum_{\alpha,\beta=1}^{p} \frac{H(\theta)_{\alpha\beta}}{4}\delta\theta_\alpha\delta\theta_\beta + o(\delta\theta^2), \quad (12)$$

where p is the number of parameters and H is the QFI. In the case studied here, this fidelity would be between output states.

Denote by $\hat{\theta}_\xi$ the guess for θ if the outcome of the measurement was ξ, then the fidelity, averaged over all possible outcomes, is

$$F(\theta, \hat{\theta}) = \sum_\xi \mathrm{tr}[\rho(\theta)M_\xi]\mathcal{F}(\theta, \hat{\theta}_\xi)$$

$$= 1 - \frac{\mathrm{tr}\,H(\theta)V(M, \theta, \hat{\theta})}{4} + o(\delta\theta^2), \quad (13)$$

where $V(M, \theta, \hat{\theta})_{\alpha\beta} = \Sigma_\xi \mathrm{tr}[\rho(\theta)M_\xi](\hat{\theta}_{\xi\alpha} - \theta_\alpha)(\hat{\theta}_{\xi\beta} - \theta_\beta)$ is the mean square error of measurement M and estimator $\hat{\theta}$.

Therefore, using Eq. (6) one gets

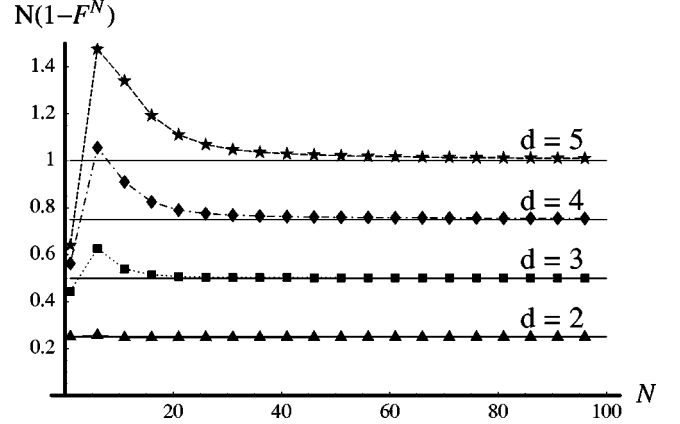[4]Actually, this fidelity is the square of the fidelity used in Ref. [9].

FIG. 3. The points are $N(1-F^N)$ as a function of N for $d = 2, \dots, 5$, where F^N is the optimal fidelity obtained in Ref. [6]. The continuous lines are at the value $(d-1)/4$ for $d = 2, \dots, 5$, (0.25, 0.5, 0.75, and 1, respectively.)

$$\lim_{N\to\infty} N[1 - F(\theta, \hat{\theta})^N] = \frac{\mathrm{Tr}\,H(\theta)H(\theta)^{-1}}{4} = \frac{d-1}{4}. \quad (14)$$

This result can be compared with the optimal fidelities obtained in Ref. [6]. These optimal fidelites were also averaged with respect to a uniform prior distribution on θ. However, since the result obtained here does not depend on θ, its average with respect to any prior will be itself. The comparison is shown in Fig. 3.

One can easily see that for large N the optimal fidelity of Ref. [6] agrees with the result obtained here. Actually, this can also be proved analytically but the proof will not be shown here.

VII. CONCLUSIONS

Two models have been compared, the model of estimating commuting unitaries with and without the use of entangled inputs (MPEE and MPEU, respectively).

It has been shown that the quantum Crámer-Rao bound is achievable in both MPEE and MPEU. It has also been shown that any quantum Fisher information matrix that can be attained in MPEE can also be achieved in MPEU. These two facts imply that an entangled input state is unnecessary. A condition for attaining maximal trace of the QFI has been derived.

In the MPEE it has also been shown that there is a separable measurement that achieves equality in the QCRB.

In the $d = 2$ case, measurements that are optimal everywhere have been found in both MPEU and MPEE. This is a useful feature in practice since this means that an adaptive scheme would not be necessary in this case. However, it is unclear whether it is possible to find measurements with this characteristic in general. It is also an open question whether entanglement could prove itself useful in this respect (for $d > 2$).

These facts show that entanglement is, at best, not as useful in estimating commuting unitaries as in the estimation of a completely unknown unitary.

APPENDIX: GENERALIZATION OF THE RESULT IN SECTION IV TO THE FULL MODEL

In this appendix, a result similar to that of Sec. IV will be proved in the model that includes the whole of $SU(d)$ and not only a commuting subgroup, i.e., the model considered in Ref. [4].

Denote by $H^{\rho_0}(\theta)$ the QFI at θ if the input state is ρ_0, and by $\widetilde{H}$ the QFI when the input state is a maximally entangled state; in what follows the dependence on θ will be omitted. Then the following inequality holds for any input state ρ_0:

$$\mathrm{Tr}(\widetilde{H}^{-1}H^{\rho_0}) \leqslant d^2 - 1, \qquad (A1)$$

and equality is attained if and only if ρ_0 is a maximally entangled state. This will be proved in what follows.

Notice that this trace is parametrization invariant, and not just $\mathrm{Tr}\, H^\rho$ as in Sec. IV; in that case $\widetilde{H}$ was proportional to the identity so there is no contradiction.

The model is again described by Eq. (4) but now U is

$$U_\theta = V_\theta \otimes 1$$

where $V_\theta = \exp(i\sum_{\alpha=1}^{d^2-1}\theta_\alpha T_\alpha)$. As before, the T's are traceless self-adjoint matrices chosen so that $\mathrm{tr}(T_\alpha T_\beta) = \delta_{\alpha\beta}$. The input state ρ_0 is chosen to be pure because of the convexity of the QFI [8]. The SLDs are $\lambda_\alpha(\theta) = 2\rho_{,\alpha}(\theta)$, where $\rho_{,\alpha}(\theta)$ means the partial derivative of $\rho(\theta)$ with respect to θ_α. The matrix elements of H^{ρ_0} are

$$H^{\rho_0}_{\alpha\beta} = \mathrm{Re}\,\mathrm{tr}[\rho\lambda_\alpha\lambda_\beta] = 4\,\mathrm{Re}\,\mathrm{tr}[\rho\rho_{,\alpha}\rho_{,\beta}].$$

Since ρ_0 is pure it can be written as $\rho_0 = |\psi_0\rangle\langle\psi_0|$ and $|\psi_0\rangle = \sum_{kl}R_{kl}|kl\rangle$. $H^{\rho_0}_{\alpha\beta}(\theta)$ can then be calculated to be

$$H^{\rho_0}_{\alpha\beta} = 4\,\mathrm{Re}[\mathrm{tr}(RR^\dagger V^\dagger_{,\alpha}V_{,\beta}) + \mathrm{tr}(RR^\dagger V^\dagger V_{,\alpha})\mathrm{tr}(RR^\dagger V^\dagger V_{,\beta})].$$

Denote $S_\alpha = -iV^\dagger V_{,\alpha}$; then

$$H^{\rho_0}_{\alpha\beta} = 4\,\mathrm{Re}[\mathrm{tr}(RR^\dagger S_\alpha S_\beta) - \mathrm{tr}(RR^\dagger S_\alpha)\mathrm{tr}(RR^\dagger S_\beta)].$$

Note that $S_\alpha \in su(d)$. Substituting $RR^\dagger = 1/d$ in the expression for H^{ρ_0}, one gets

$$\widetilde{H}_{\alpha\beta} = \frac{4}{d}\mathrm{tr}(S_\alpha S_\beta).$$

The matrices $S_1, \ldots, S_{d^2-1}$ can be orthonormalized,

$$\mathrm{tr}\left[\left(\frac{2}{\sqrt{d}}\sum_\mu \widetilde{H}^{-1/2}_{\alpha\mu}S_\mu\right)\left(\frac{2}{\sqrt{d}}\sum_\nu \widetilde{H}^{-1/2}_{\beta\nu}S_\nu\right)\right] = \delta_{\alpha\beta}.$$

The operator

$$\sum_\alpha \left(\frac{2}{\sqrt{d}}\sum_\mu \widetilde{H}^{-1/2}_{\alpha\mu}S_\mu\right)\left(\frac{2}{\sqrt{d}}\sum_\nu \widetilde{H}^{-1/2}_{\alpha\nu}S_\nu\right) = \frac{4}{d}\sum_{\mu\nu}\widetilde{H}^{-1}_{\mu\nu}S_\mu S_\nu$$

is a Casimir operator and therefore proportional to the identity. The proportionality factor can be found by taking the trace, and finally one gets

$$\sum_{\mu\nu}\widetilde{H}^{-1}_{\mu\nu}S_\mu S_\nu = \frac{d^2-1}{4}1.$$

The wanted trace is

$$\mathrm{Tr}(\widetilde{H}^{-1}H^{\rho_0}) = d^2 - 1 - \sum_{\alpha\beta}\widetilde{H}^{-1}_{\alpha\beta}\mathrm{tr}(RR^\dagger S_\alpha)\mathrm{tr}(RR^\dagger S_\beta).$$

This quantity is always less than or equal to $d^2 - 1$ and, furthermore, this value is attained if and only if $\mathrm{tr}(RR^\dagger S_\alpha) = 0$ for all $\alpha = 1, \ldots, d^2 - 1$, which implies that $RR^\dagger = 1/d$, i.e., ρ_0 is maximally entangled. In particular, this implies that there is no input state ρ_0 for which $H^{\rho_0} \geqslant \widetilde{H}$.

[1] A. Fujiwara, Phys. Rev. A **65**, 012316 (2001).

[2] F. De Martini, A. Mazzei, M. Ricci, and G. M. D'Ariano, Phys. Rev. A **67**, 062307 (2003).

[3] A. Acín, E. Jané, and G. Vidal, Phys. Rev. A **64**, 050302 (2001).

[4] M. A. Ballester, Phys. Rev. A **69**, 022303 (2004).

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).

[6] C. Macchiavello, Phys. Rev. A **67**, 062302 (2003).

[7] K. Matsumoto, J. Phys. A **35**, 3111 (2002).

[8] A. Fujiwara, Phys. Rev. A **63**, 042304 (2001).

[9] M. Hübner, Phys. Lett. A **163**, 239 (1992).