Parameter estimation for mixed states from a single copy

Thomas Konrad,[1] Otfried Gühne,[2] Jürgen Audretsch,[3] and Hans J. Briegel[2,4]
[1]*School of Physics, University of KwaZulu-Natal, Private Bag 54001, Durban 4000, South Africa*
[2]*Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, A-6020 Innsbruck, Austria*
[3]*Fachbereich Physik, Universität Konstanz, Fach M 674, D-78457 Konstanz, Germany*
[4]*Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria*
(Received 13 March 2007; published 1 June 2007)

Given a single copy of a mixed state of the form $\rho=\lambda\rho_1+(1-\lambda)\rho_2$, what is the optimal measurement to estimate the parameter λ if ρ_1 and ρ_2 are known? We present a general strategy to obtain the optimal measurements employing a Bayesian estimator. The measurements are chosen to minimize the deviation between the estimated value and the true value of λ. We explicitly determine the optimal measurements for a general two-dimensional system and for important higher dimensional cases.

DOI: 10.1103/PhysRevA.75.062101 PACS number(s): 03.65.Wj, 03.67.−a, 03.65.Ta

I. INTRODUCTION

The estimation of quantum states is one of the basic primitives in quantum theory. This general task appears in several modifications and situations. When many copies of the state are available, one may try to obtain maximal information on the quantum state by employing state tomography [1], although the effort may become overwhelming if the dimension of the system increases. Furthermore, tomography is clearly not a viable way, if only a few or even only a single copy of the state are available.

Nevertheless, if some *a priori* information about the state is given, there are several estimation problems which can be meaningfully posed and solved even on the single copy level. The most prominent is state discrimination, where one knows that the system is in one of several given states, and one needs to decide in which one. This problem has been discussed in several variations, either one aims for an unambiguous discrimination, or for a discrimination with a minimal error probability [1,2]. Moreover, programmable state discriminators have been proposed [3]. Other problems considered are state estimation from several copies with collective or separable measurements [4], the estimation of certain state parameters (like the time in a unitary evolution) [5], as well as estimation of the state after a generalized measurement when the premeasurement state is unknown [6].

In this paper, we consider the following problem: Let us assume that an apparatus is given, which produces for a given input λ the state

$$\rho(\lambda)=\lambda\rho_1+(1-\lambda)\rho_2, \tag{1}$$

where ρ_1 and ρ_2 are known. Let us further assume that λ is determined by some well characterized random number generator, but we do not know its actual value. The task is now to estimate λ from a single copy of $\rho(\lambda)$.

Such estimation problems can occur in several realistic situations. For instance, one may consider some process of decoherence, where the single copy of $\rho(\lambda)$ describes the state of a single atom coupled to its environment, and the task is to estimate the rate of decoherence. In this case ρ_1 represents an initial state, while ρ_2 is the state of thermal equilibrium, which the system assumes eventually. Another example may be that $\rho(\lambda)$ is the reduced state of some multipartite pure state $|\psi(\lambda)\rangle$, where one tries to estimate λ. We will discuss the decoherence example later in more detail.

For the estimation one has to perform some measurement. The most general measurement is described by a positive operator valued measure (POVM) and we ask for the POVM that minimizes the expected deviation between the true value of λ and the estimated one. Here we take the viewpoint that the value of λ is not exactly known beforehand but we may have information about it in the form of a prior probability distribution. After a measurement this prior probability distribution can be updated according to Bayes theorem. Minimizing the deviation of the estimate from λ expected from the posterior probability distribution leads to the so-called Bayesian estimator (see, for example, [7]).

Bayesian estimators have been successfully applied in state discrimination problems. They are also appropriate in situations where a sequence of consecutive measurements is carried out and the estimate has to be updated after each measurement. This is, for example, the case in real-time monitoring of the dynamics of quantum systems. Using a Bayesian estimator it is possible to monitor oscillations of single qubits in real-time with high accuracy by means of a sequence of weak measurements [8,9]. The Bayesian estimator is also intimately related to the continuous estimation of the wave function of a system with arbitrary-dimensional Hilbert space by means of continuous measurement [10]. Although the search for the best measurement in connection with a Bayesian estimator is performed in this paper for the particular task to estimate the parameter λ in Eq. (1), it may be carried out in an analogous way in different setups.

For the case when ρ is a two-level system, i.e., a qubit, we determine the optimal measurement strategy for arbitrary ρ_1 and ρ_2. Interestingly, for the case of completely unknown λ, it turns out that if ρ_1 and ρ_2 do not have the same purity, the optimal measurement does not commute with $\rho_1-\rho_2$, and hence does not coincide with the intuitive guess which one would make from the Bloch sphere picture. Finally, we discuss how the results from the qubit case can readily be used to solve this problem for important higher dimensional cases.

As already mentioned, similar parameter estimation problems have in generality been studied in Ref. [5] and general conditions on the information which can be obtained about the parameter have been formulated. Our aim, however, is to explicitly construct the optimal measurements. The knowledge of these explicit measurements may further be used to improve schemes for the observation of oscillations with small disturbance [8,9].

Our paper is organized as follows: In Sec. II we pose the problem in mathematical terms and derive a general condition on the optimal POVM. In Sec. III we show how this general condition can already reduce the set of POVMs which have to be considered. In Sec. IV we solve the problem for qubits. We first derive the optimal measurement, if ρ_1 and ρ_2 are pure, then we consider the general case. We also discuss how one can apply the results for estimating decoherence rates. Finally, in Sec. V we derive some results for higher dimensional systems.

II. BAYESIAN ESTIMATOR AND CONDITION FOR OPTIMAL MEASUREMENT

In this paper, we consider the following task. Let us assume that we have a single quantum system with d-dimensional Hilbert space $\mathcal{H}$. Its state is known to be a mixture of two states ρ_1 and ρ_2, i.e.,

$$\rho_\lambda = \lambda\rho_1 + (1-\lambda)\rho_2, \tag{2}$$

where for the parameter λ only some probability distribution is known. What is the measurement which leads to an optimal estimation of the unknown parameter λ?

The estimation of λ has to proceed in two steps: First, a measurement is carried out on the system. The statistics of a general measurement can be characterized by a POVM and throughout this paper we denote by $\mathcal{P}=\{E_m\}$ a POVM with the effects E_m. In a second step the parameter λ is estimated by a number g_m (the estimate) which depends on the result of the measurement m.

Let us first discuss the construction of the optimal estimate g_m for a given outcome m. A criterion for the optimality of the estimate g_m can be formulated as follows: g_m is optimal iff the cost function

$$c(\lambda, g_m) = (\lambda - g_m)^2 \tag{3}$$

is expected to assume its minimum:

$$\mathcal{E}_m((\lambda - g_m)^2) := \int (\lambda - g_m)^2 p(\lambda|m) d\lambda \overset{!}{=} \min. \tag{4}$$

Here the expectation value $\mathcal{E}$ is taken with respect to the posterior probability distribution of λ, which includes the information contained in the occurrence of the measurement result m:

$$p(\lambda|m) = \frac{p(m|\lambda)p(\lambda)}{p(m)}, \tag{5}$$

where $p(m)=\int p(m|\lambda)p(\lambda)d\lambda$ is the probability to obtain the measurement result m averaged over all possible occurring states.

Taking into account the linearity of the expectation value, one can directly verify that the optimal estimator g_m is equal to the expected value of λ:

$$\begin{aligned}\mathcal{E}_m((\lambda - g_m)^2) &= \mathcal{E}_m((\lambda)^2) - 2g_m\mathcal{E}_m(\lambda) + g_m^2 \\ &= (\mathcal{E}_m(\lambda) - g_m)^2 + \mathrm{Var}_m(\lambda),\end{aligned} \tag{6}$$

where

$$\mathrm{Var}_m(\lambda) = \mathcal{E}_m(\lambda^2) - \mathcal{E}_m(\lambda)^2 \tag{7}$$

represents the variance of λ. The right-hand side of Eq. (6) assumes a minimum for $g_m=\mathcal{E}_m(\lambda)$. Such a value g_m is also called a Bayesian estimate.

Having derived the optimal estimate g_m for a certain outcome m we can consider the optimal choice of the POVM. Now, the optimal measurement is the one which leads to the smallest expected costs averaged over all outcomes m:

$$\sum_m p(m)\mathcal{E}_m((\lambda - g_m)^2) \overset{!}{=} \min. \tag{8}$$

Hence the effects $\{E_m\}$ have to minimize the mean variance $\Sigma_m p(m)\mathrm{Var}_m(\lambda)$. Note that this variance, just as the expectation of the costs, is defined with respect to the posterior probability density $p(\lambda|m)$.

For the sake of simplicity we choose in the following an equally distributed prior density $p(\lambda)=1$. We will see later in an example that this does not impose a big restriction. Using $p(m|\lambda)=\mathrm{tr}[E_m\rho_\lambda]$ and the probability to obtain measurement result m,

$$\begin{aligned}p(m) &= \int_0^1 p(m|\lambda)p(\lambda)d\lambda = \int_0^1 \mathrm{tr}\{E_m[\lambda\rho_1 + (1-\lambda)\rho_2]\}d\lambda \\ &= \mathrm{tr}\left[E_m\frac{1}{2}(\rho_1+\rho_2)\right],\end{aligned} \tag{9}$$

we obtain for the posterior probability density in Eq. (5)

$$p(\lambda|m) = \frac{\mathrm{tr}[E_m\rho_\lambda]}{\mathrm{tr}\left[E_m\frac{1}{2}(\rho_1+\rho_2)\right]}. \tag{10}$$

In order to minimize the mean variance of λ we have to first compute the expectation value of λ and λ^2 with respect to the posterior probability density:

$$\begin{aligned}\mathcal{E}_m(\lambda) &= \int_0^1 \lambda p(\lambda|m)d\lambda = \frac{\int_0^1 \lambda\, \mathrm{tr}[E_m\rho_\lambda]d\lambda}{\mathrm{tr}\left[E_m\frac{1}{2}(\rho_1+\rho_2)\right]} \\ &= \frac{\mathrm{tr}\left[E_m\left(\frac{2}{3}\rho_1 + \frac{1}{3}\rho_2\right)\right]}{2\,\mathrm{tr}\left[E_m\frac{1}{2}(\rho_1+\rho_2)\right]}\end{aligned} \tag{11}$$

and in a similar calculation

$$\mathcal{E}_m(\lambda^2) = \int_0^1 \lambda^2 p(\lambda|m) d\lambda = \frac{\mathrm{tr}\left[E_m\left(\frac{3}{4}\rho_1 + \frac{1}{4}\rho_2\right)\right]}{3\,\mathrm{tr}\left[E_m \frac{1}{2}(\rho_1+\rho_2)\right]}.$$

The mean variance thus amounts to [see Eqs. (7) and (9)]:

$$\sum_m p(m)\mathrm{Var}_m(\lambda) = \frac{1}{3} - \frac{1}{4}\sum_m \frac{\left\{\mathrm{tr}\left[E_m\left(\frac{2}{3}\rho_1 + \frac{1}{3}\rho_2\right)\right]\right\}^2}{\mathrm{tr}\left[E_m \frac{1}{2}(\rho_1+\rho_2)\right]}. \tag{12}$$

We can thus summarize the main result of this section.

Proposition 1. The optimal POVM $\mathcal{P}=\{E_m\}$ is the one which maximizes

$$Q(\mathcal{P}) := \frac{1}{4}\sum_m \frac{\left\{\mathrm{tr}\left[E_m\left(\frac{2}{3}\rho_1 + \frac{1}{3}\rho_2\right)\right]\right\}^2}{\mathrm{tr}\left[E_m \frac{1}{2}(\rho_1+\rho_2)\right]}. \tag{13}$$

Note that using the fact that $\frac{2}{3}\rho_1+\frac{1}{3}\rho_2=\frac{1}{2}(\rho_1+\rho_2)+\frac{1}{6}(\rho_1-\rho_2)$ the quantity $Q(\mathcal{P})$ may be rewritten as

$$Q(\mathcal{P}) := \frac{1}{4}\left(1 + \sum_m \frac{\{\mathrm{tr}[E_m(\rho_1-\rho_2)]\}^2}{3\,\mathrm{tr}[E_m(\rho_1+\rho_2)]}\right), \tag{14}$$

which is manifestly invariant under the permutation of ρ_1 and ρ_2. Further note that Eq. (11) allows us to calculate the Bayesian estimate g_m of λ, given a measurement outcome associated with the effect operator E_m, since $g_m=\mathcal{E}_m(\lambda)$.

The condition for an optimal measurement in the case of a general prior probability distribution $p(\lambda)$ can be determined from the corresponding mean variance, which is of similar form as Eq. (12). It can be derived as in Eqs. (9)–(12) with a general $p(\lambda)$ and reads

$$\sum_m p(m)\mathrm{Var}_m(\lambda) = \overline{\lambda^2} - (\bar{\lambda})^2 \sum_m \frac{\left(\mathrm{tr}\left\{E_m\left[\frac{\overline{\lambda^2}}{\bar{\lambda}}\rho_1 + \left(1 - \frac{\overline{\lambda^2}}{\bar{\lambda}}\right)\rho_2\right]\right\}\right)^2}{\mathrm{tr}\{E_m[\bar{\lambda}\rho_1 + (1-\bar{\lambda})\rho_2]\}}, \tag{15}$$

where $\overline{\lambda^n}$ is the nth moment of $p(\lambda)$, i.e., $\overline{\lambda^n}=\int_0^1\lambda^n p(\lambda)d\lambda$. Again, this may be rewritten in a way similar to Eq. (14) which is invariant under permutation of ρ_1 and ρ_2. Although we consider in the following an equidistributed λ, thus $p(\lambda)=1$, we illustrate by an example in Sec. IV that the results we obtain can readily be transcribed for the general case.

III. PROPERTIES OF OPTIMAL MEASUREMENTS

Now we show two facts about the POVM which minimizes the expression of Eq. (12): We first show that this POVM can be chosen to have pure effects, i.e., effects which are of rank one. Then, we show that this POVM must be a so-called extremal POVM.

To show that the POVM which minimizes the expression of Eq. (12) has pure effects E_i, assume that we have a POVM where one effect, say E_1, is not pure. Then we can decompose E_1 as $E_1=E_1^A+E_1^B$ where E_1^A and E_1^B are linearly independent positive operators, which can serve as new effects. The new POVM with the effects $E_1^A, E_1^B, E_2, E_3, \ldots.$ gives the same or a smaller value for the cost function. Indeed, from Eq. (12) it follows that it suffices to show

$$\frac{(Z_1)^2}{N_1} \leqslant \frac{(Z_1^A)^2}{N_1^A} + \frac{(Z_1^B)^2}{N_1^B}, \tag{16}$$

where $Z_\beta^\alpha=\mathrm{tr}\left[E_\beta^\alpha\left(\frac{2}{3}\rho_1+\frac{1}{3}\rho_2\right)\right]$ and $N_\beta^\alpha=\mathrm{tr}\left[E_\beta^\alpha\frac{1}{2}(\rho_1+\rho_2)\right]$. Equation (16) can be straightforwardly verified, using the facts that $Z_1=Z_1^A+Z_1^B$ and $N_1=N_1^A+N_1^B$. Thus for our search for optimal measurement strategies, it suffices to consider POVMs with pure effects.

There are further constraints on the optimal POVM $\mathcal{P}$. These follow from the fact that the set of all POVMs is a convex set. Given two POVMs $\mathcal{P}^{(1)}=\{E_m^{(1)}\}$ and $\mathcal{P}^{(2)}=\{E_m^{(2)}\}$ with K outcomes, the convex combination $\mathcal{P}=p\mathcal{P}^{(1)}+(1-p)\mathcal{P}^{(2)}=\{pE_m^{(1)}+(1-p)E_m^{(2)}\}$ is again a POVM with K outcomes. On the other hand there are POVMs which cannot be expressed as a convex combination of two different POVMs, these are called *extremal* [11]. For our purpose, it is important to note that for convex combinations

$$Q(\mathcal{P}) \leqslant pQ(\mathcal{P}^{(1)}) + (1-p)Q(\mathcal{P}^{(2)}) \tag{17}$$

holds. This follows because it is true for each of the summands of Q. This can be verified, e.g., for the summand with the effect $E_1=pE_1^{(1)}+(1-p)E_1^{(2)}$ by replacing the effects E_1^A and E_1^B in inequality (16) by $pE_1^{(1)}$ and $(1-p)E_1^{(2)}$, respectively. Since a generic convex combination on the right-hand side of Eq. (17) assumes only values less than $\max\{Q(\mathcal{P}^{(1)}), Q(\mathcal{P}^{(2)})\}$, only extremal POVMs can maximize Q and thus minimize the expected costs. We can summarize.

Proposition 2. When maximizing Q in Eq. (13) it suffices to consider extremal POVMs with pure effects.

Indeed, we can require both conditions at the same time, since POVMs with pure effects, which are not extremal, can be written as a convex combination of extremal POVMs with pure effects. The two conditions on the POVMs are *a priori* independent. However, for special cases, certain relations are known. For instance, it has been shown that all extremal POVMs for qubits have pure effects and maximally four outcomes [11]. This fact we will exploit in the following.

IV. OPTIMAL POVM FOR QUBITS

In this section we consider the problem for qubits, that is, ρ_1 and ρ_2 are qubit states. The representation of these states on the Bloch sphere will enable us to determine the optimal POVM. We will first derive some general formulation of the optimization problem in terms of Bloch vectors. Then we will solve the problem for the case where ρ_1 and ρ_2 are pure and finally for general ρ_1 and ρ_2.

We know already that the effects E_m of an optimal measurement are pure, i.e.,

$$E_m = \alpha_m |0_m\rangle\langle 0_m|. \tag{18}$$

Here, the states $|0_m\rangle$ span the whole Hilbert space but they are in general not mutually orthogonal. Since we are dealing with qubits the projectors $|0_m\rangle\langle 0_m|$ can be expressed by means of Bloch vectors $\vec{r}_m = \langle 0_m|\vec{\sigma}|0_m\rangle$:

$$|0_m\rangle\langle 0_m| = \frac{1}{2}(\mathbb{1} + \vec{r}_m \cdot \vec{\sigma}), \tag{19}$$

where the components of the vector $\vec{\sigma}$ are the Pauli operators σ_x, σ_y, and σ_z. Thus the effects read

$$E_m = p_m(\mathbb{1} + \vec{r}_m \cdot \vec{\sigma}) \tag{20}$$

with $p_m := \alpha_m/2$. From the completeness of the effects $\sum_m E_m = \mathbb{1}$ we obtain the following constraints for the weights p_m and the vectors $\vec{r}_m$:

$$\sum_m p_m = 1 \quad \text{and} \sum_m p_m \vec{r}_m = \vec{0}. \tag{21}$$

The optimal POVM maximizes Q in Eq. (13). Expressing the states $\rho_a := (2\rho_1 + \rho_2)/3$ and $\rho_b := (\rho_1 + \rho_2)/2$ in terms of Bloch vectors $\vec{r}_i = \mathrm{tr}[\vec{\sigma}\rho_i]$,

$$\rho_a = \frac{1}{2}(\mathbb{1} + \vec{r}_a \cdot \vec{\sigma}), \quad \rho_b = \frac{1}{2}(\mathbb{1} + \vec{r}_b \cdot \vec{\sigma}), \tag{22}$$

Q reads

$$Q = \frac{1}{4} \sum_m p_m \frac{(1 + \vec{r}_m \cdot \vec{r}_a)^2}{1 + \vec{r}_m \cdot \vec{r}_b}. \tag{23}$$

The quantity Q can be further simplified by expanding it in powers of the difference vector $\Delta\vec{r} = \vec{r}_a - \vec{r}_b$,

$$\begin{aligned} Q &= \frac{1}{4} \sum_m p_m \frac{[1 + \vec{r}_m \cdot (\vec{r}_b + \Delta\vec{r})]^2}{1 + \vec{r}_m \cdot \vec{r}_b} \\ &= \frac{1}{4} \sum_m p_m \left(1 + \vec{r}_m \cdot (\vec{r}_b + 2\Delta\vec{r}) + \frac{(\vec{r}_m \cdot \Delta\vec{r})^2}{1 + \vec{r}_m \cdot \vec{r}_b} \right) \\ &= \frac{1}{4} \left(1 + \sum_m p_m \frac{(\vec{r}_m \cdot \Delta\vec{r})^2}{1 + \vec{r}_m \cdot \vec{r}_b} \right), \end{aligned} \tag{24}$$

where we have employed conditions (21) to obtain the last line from the second. Starting with this representation of Q, we can further reduce the set of interesting POVMs for our task.

Proposition 3. When determining the optimal POVM for qubits it suffices to consider POVMs with pure effects and maximally three outcomes where the vectors $\vec{r}_m$ lie in the plane spanned by $\vec{r}_a, \vec{r}_b$, and $\vec{0}$ in the Bloch sphere.

Proof. In fact the effects E_m of any qubit-POVM $\mathcal{P}$ can be represented by means of Bloch vectors $\vec{r}_m$:

$$E_m = p_m(\mathbb{1} + \vec{r}_m \cdot \vec{\sigma}), \tag{25}$$

where the p_m, $\vec{r}_m$ satisfy condition (21). Positivity of E_m implies $0 \leq \|\vec{r}_m\| \leq 1$ with $\|\vec{r}_m\| = 1$ iff E_m is pure.

Now, let us assume without restriction of generality that the plane spanned by $\vec{r}_a, \vec{r}_b$ and $\vec{0}$ is the x-z plane. Starting from the vectors $\vec{r}_m$ which describe the effects E_m of the POVM $\mathcal{P}$ we consider their projection $\vec{q}_m = \vec{r}_m|_{xz}$ on the x-z plane. This results in a new POVM $\tilde{\mathcal{P}}$ with the effects

$$\tilde{E}_m = p_m(\mathbb{1} + \vec{q}_m \cdot \vec{\sigma}). \tag{26}$$

This POVM possesses in general not only pure effects, since the $\vec{q}_m$ are in the generic case not normalized. However, Eqs. (23) and (24) are still valid, which implies that $Q(\mathcal{P}) = Q(\tilde{\mathcal{P}})$. To proceed, we can now decompose the nonpure effects $\tilde{E}_m$ into pure ones. This can, for example, be done by means of spectral decomposition:

$$\tilde{E}_m = p_m^A \frac{\mathbb{1} + \hat{\vec{q}}_m \cdot \vec{\sigma}}{2} + p_m^B \frac{\mathbb{1} - \hat{\vec{q}}_m \cdot \vec{\sigma}}{2} =: \tilde{E}_m^A + \tilde{E}_m^B, \tag{27}$$

where the eigenvalues are given by $p_m^{A/B} = p_m(1 \pm \|\vec{q}_m\|)$. Please observe that the Bloch vectors $\pm\hat{\vec{q}}_m := \pm\vec{q}_m/\|\vec{q}_m\|$ of the pure effects $\tilde{E}_m^{A/B}$ point into the x-z plane. Thus it suffices to consider POVMs with pure effects (but maybe many outcomes) in the x-z plane. Finally, we can choose extremal POVMs. If the POVM with many pure effects in the x-z plane is not extremal, we can write it as a convex combination of extremal ones, which obviously have to have pure effects in the x-z plane again. The point is, that it is known that for extremal POVMs with four outcomes the projectors cannot lie in one plane [11]. This proves the claim. □

In the following we study the problem of optimal measurements beginning with the simplest case, the optimal von Neumann measurement for pure states ρ_1 and ρ_2. Then we attempt to show that this von Neumann measurement leads to the least mean variance compared to all generalized qubit measurements. We increase the level of generality by studying the case of mixed states ρ_1 and ρ_2.

A. Optimal PVM for mixtures of pure qubit states

Let us first consider pure states ρ_1 and ρ_2 of a qubit, and find the optimal von Neumann measurement, i.e., the PVM $\{P_+ = |\psi\rangle\langle\psi|,\ P_- = \mathbb{1} - |\psi\rangle\langle\psi|\}$ for which the expected variance of λ is least. The projectors $\{P_+, P_-\}$ can be expressed by means of normalized Bloch vectors $\vec{r}_+ = \langle\psi|\vec{\sigma}|\psi\rangle \equiv \vec{r}$ and $\vec{r}_- = -\vec{r}$,

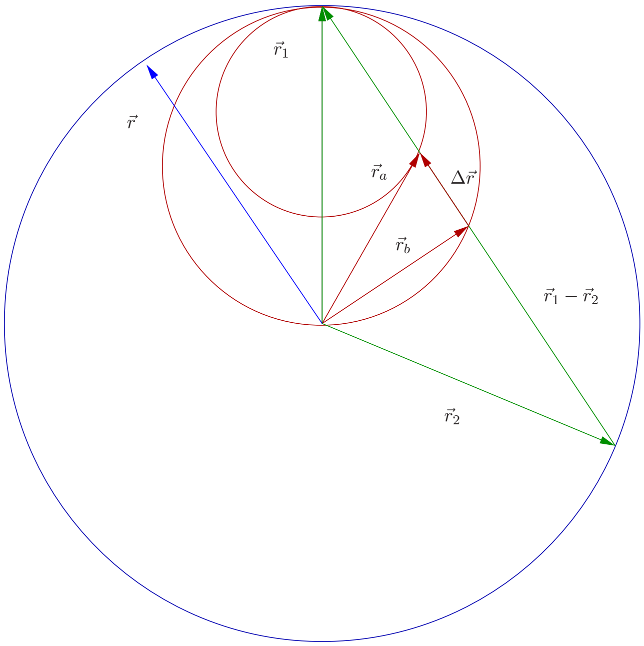

FIG. 1. (Color online) Schematic figure of the Bloch representation for the case where ρ_1 and ρ_2 are pure. The Bloch vectors $\vec{r}_b$ and $\Delta\vec{r} \propto \vec{r}_1 - \vec{r}_2$ are orthogonal because of the circle of Thales. $\vec{r}$ denotes the direction of the considered measurement. In the calculation, it turns out that for the optimal measurement $\vec{r}$ is parallel to $\vec{r}_1 - \vec{r}_2$. See text for further details.

$$P_\pm = \frac{1}{2}(\mathbb{1} \pm \vec{r}\vec{\sigma}). \tag{28}$$

This implies that the weights which appear in the parametrization of the pure effects (20) are given by $p_\pm = 1/2$.

Inserting the weights $p_\pm$ and the Bloch vectors $\vec{r}_\pm$ into Q as given by Eq. (24) yields

$$Q = \frac{1}{4}\left(1 + \frac{(\vec{r}\cdot\Delta\vec{r})^2}{1 - (\vec{r}\cdot\vec{r}_b)^2}\right). \tag{29}$$

Without losing generality we assume that ρ_1 and ρ_2 are represented by the Bloch vectors $\vec{r}_1 = (0,0,1)$ and $\vec{r}_2 = (\sin\theta, 0, \cos\theta)$, respectively. In this case $\vec{r}_b = (\vec{r}_1 + \vec{r}_2)/2 = [\sin\theta\vec{e}_x + (1+\cos\theta)\vec{e}_z]/2$ and $\Delta\vec{r} = [-\sin\theta\vec{e}_x + (1-\cos\theta)\vec{e}_z]/6$ are perpendicular, see Fig. 1. Because of this orthogonality Q in Eq. (29) reduces to

$$Q = \frac{1}{4}\left(1 + \frac{(\Delta r)^2\cos^2\alpha}{1 - r_b^2\sin^2\alpha}\right), \tag{30}$$

where α is the angle between $\vec{r}$ and $\Delta\vec{r}$, $\Delta r = \|\Delta\vec{r}\|$ and $r_b = \|\vec{r}_b\| \leqslant 1$. Q assumes its maximum value

$$Q_{\max} = \frac{1}{4}[1 + (\Delta r)^2] \tag{31}$$

at $\alpha = k\pi$ with $k = 0,1,2,\ldots$, i.e., if the Bloch vectors $r_\pm$ of the projectors point into the same (opposite) direction as $\Delta\vec{r}$. This is the case for measurements of nondegenerated observables O which commute with $\rho_1 - \rho_2$.

It remains to show that this von Neumann measurement is also optimal among all possible POVM. This we will do in the next section, where we solve the problem for general ρ_1, ρ_2.

B. Mixtures of mixed states

Which is the optimal measurement if ρ_1 and ρ_2 are mixed qubit states? In order to answer this question, we first determine the optimal projector-valued measure and then show that this leads also to a maximal value of Q for all POVMs. According to Eq. (29), Q for a given PVM can be written as

$$Q(\alpha) = \frac{1}{4}\left(1 + \frac{(\Delta r)^2\cos^2(\alpha)}{1 - r_b^2\cos^2\beta}\right), \tag{32}$$

where α is the angle between the Bloch vector $\vec{r}$ which represents one projector of the PVM and $\Delta\vec{r}$ and β is the angle between $\vec{r}$ and $\vec{r}_b$. Introducing the angle γ between $\Delta\vec{r}$ and $\vec{r}_b$ and taking into account that $\beta = \alpha + \gamma$, one can directly verify that Q assumes its maximum at

$$\alpha_0 = \pm\arccos\left[\frac{\cos\gamma}{\sqrt{\frac{r_b^2}{2}(r_b^2 - 2)[1 - \cos(2\gamma)] + 1}}\right] - \gamma$$
$$+ k\pi \quad \text{with} \quad k = 0,1,\ldots, \tag{33}$$

where the positive sign is for $0 \leqslant \gamma < \pi$, and the negative sign has to be taken for $-\pi \leqslant \gamma < 0$ (as it is in Fig. 2). If both states were pure (as in the previous section), we had $\gamma = \pi/2$ and hence $\alpha_0 = k\pi$. For mixed states, however, this is in general not the case, hence the projectors of the optimal PVM measurement do not correspond to the eigenvectors of $\rho_1 - \rho_2$ (see Fig. 2). Only if both states ρ_1 and ρ_2 have the same purity [that is, $\mathrm{tr}(\rho_1^2) = \mathrm{tr}(\rho_2^2)$ which implies $r_1 = r_2 \Leftrightarrow \gamma = 0$] the best measurement is again that of a nontrivial observable O which commutes with $\rho_1 - \rho_2$.

The deviation from the pure state case $\alpha = k\pi$ is shown in Fig. 3. Physically, this deviation may be explained as follows. If ρ_2 is relatively mixed compared to ρ_1, then no ob-

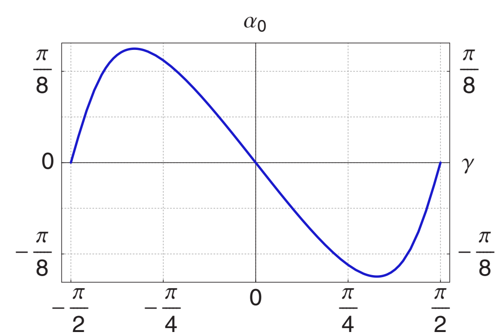

FIG. 3. (Color online) Deviation of optimal measurement direction $\vec{r}$ from direction of $\vec{r}_1 - \vec{r}_2$ specified by the angle α_0 as a function of angle γ between $\Delta\vec{r}$ and $\vec{r}_b$ for $r_b = 0.8$. For states ρ_1, ρ_2 with the same entropy ($\gamma = \pm\pi/2$) the deviation is zero. If, say, state ρ_1 corresponds to less entropy $\vec{r}$ is tilted towards the Bloch vector of ρ_1. For example, compare the case depicted in Fig. 2, where $\pi/2 < \gamma < 0$ and thus $0 < \alpha_0 < \pi/2$, indicating the named tilting.

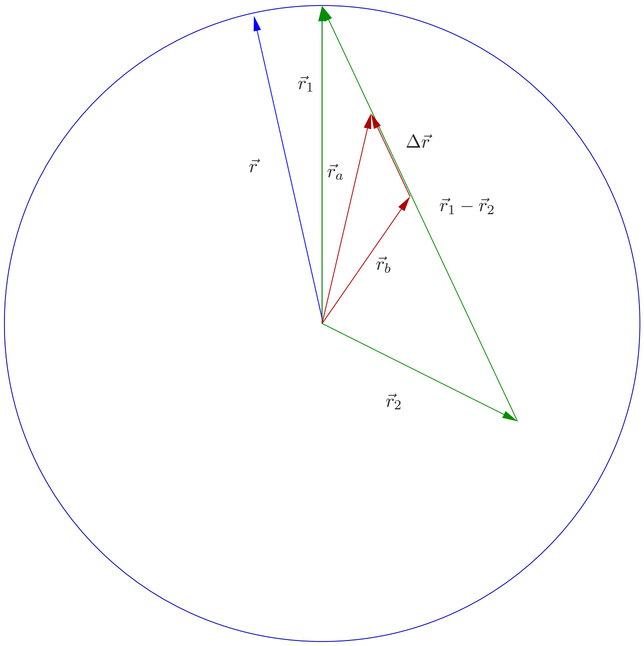

FIG. 2. (Color online) Schematic figure of the Bloch representation for the case where ρ_1 is pure, but ρ_2 is not. Now, the optimal measurement is a measurement in the direction of $\vec{r}$ which is tilted towards $\vec{r}_1$.

servable does efficiently resolve the value of λ for small λ. This is simply due to the fact that the possible states with a small λ are close to the maximally mixed state. So it is favorable to measure in the direction of ρ_1 which gives at least a high resolution for high values of λ.

Finally, we have to show that the optimal projection measurement is also the optimal POVM measurement. To do so, we prove that the Q-value for all POVMs is less than or equal to its maximum value $Q_{\max}^{\mathrm{PVM}}$ for PVMs. For a general qubit-POVM [see Eq. (24)], Q is given by

$$Q = \frac{1}{4}\left(1 + \sum_m p_m \frac{(\Delta r)^2 \cos^2 \alpha_m}{1 + r_b \cos(\alpha_m + \gamma)}\right), \tag{34}$$

whereas the maximum Q for projection measurements is of the form

$$Q_{\max}^{\mathrm{PVM}} = \frac{1}{4}\left(1 + \frac{(\Delta r)^2 \cos^2 \alpha_0}{1 - r_b^2 \cos^2(\alpha_0 + \gamma)}\right) \tag{35}$$

with α_0 given by Eq. (33). We can formulate.

Proposition 4. For qubits POVMs cannot improve the optimal projective measurement, with α_0 given in Eq. (33). That is, we have for all POVMs $Q \leq Q_{\max}^{\mathrm{PVM}}$.

Proof. We introduce the positive function

$$\alpha \to f(\alpha) = \frac{\cos^2(\alpha)}{1 - r_b^2 \cos^2(\alpha + \gamma)}, \tag{36}$$

then the claim can be expressed as follows:

$$Q \leq Q_{\max}^{\mathrm{PVM}} \Leftrightarrow \sum_m p_m f(\alpha_m)(1 - r_b \cos \beta_m) \leq f(\alpha_0) \Leftrightarrow \sum_m p_m \frac{f(\alpha_m)}{f(\alpha_0)}(1 - r_b \cos \beta_m) \leq 1, \tag{37}$$

where $\beta_m := \alpha_m + \gamma$. Since $f(\alpha) \leq f(\alpha_0)$ for all $\alpha \in [0, 2\pi]$, it follows that

$$\sum_m p_m \frac{f(\alpha_m)}{f(\alpha_0)}(1 - r_b \cos \beta_m) \leq \sum_m p_m (1 - r_b \cos \beta_m) = 1, \tag{38}$$

which proves the claim. The right-hand side is equal to unity because of the constraints on weights p_m and the Bloch vectors $\vec{r}_m$ of the effects (21), i.e., $\Sigma_m p_m = 1$ and $\Sigma_m p_m \cos \beta_m = 0$.□

C. Decoherence as an example

As an example, let us discuss the following problem. We consider a two-level atom under the influence of decoherence. The task is to estimate the decay rate from one copy of the state. The evolution of the density matrix may be described by the master equation

$$\frac{\partial}{\partial t}\rho = \frac{1}{i\hbar}[H, \rho] + \mathcal{L}\rho, \tag{39}$$

where $H = \hbar\omega\sigma_z/2$ is the Hamiltonian of the atom, and the incoherent evolution is of the Lindblad form [12]

$$\mathcal{L}\rho = -\frac{B}{2}(1 - s)(\sigma^+\sigma^-\rho + \rho\sigma^+\sigma^- - 2\sigma^-\rho\sigma^+) - \frac{B}{2}s(\sigma^-\sigma^+\rho + \rho\sigma^-\sigma^+ - 2\sigma^+\rho\sigma^-) - \frac{2C - B}{4}(\rho - \sigma_z\rho\sigma_z). \tag{40}$$

Here, $s = \langle 0|\rho_\infty|0\rangle$ denotes the population of the exited state in thermal equilibrium, $\sigma^\pm = \sigma_x \pm i\sigma_y$, and B and C are the decay rates of the expectation values of σ_z and $\sigma^\pm$.

It can be straightforwardly shown that for the case $C = B$ the time dependent density matrix is given by

$$\rho(t) = U\tilde{\rho}(t)U^\dagger,$$

$$\tilde{\rho}(t) = e^{-Bt}\rho(0) + (1 - e^{-Bt})\begin{pmatrix} s & 0 \\ 0 & 1 - s \end{pmatrix}, \tag{41}$$

with $U = e^{-iHt/\hbar}$. Note that this form is independent from the starting density matrix $\rho(0)$. Now we can ask: assuming we know s, t, and $\rho(0)$, how can we estimate the decay parameter B from a single copy of $\rho(t)$?

Let us assume that we know that at least $B \in [0, B^{\max}]$. Then the state under consideration is, in the rotating frame, of the form

$$\rho(\lambda) = \lambda\rho_1 + (1 - \lambda)\rho_2, \quad \lambda = e^{-Bt} \in [e^{-tB^{\max}}, 1]. \tag{42}$$

While $B \in [0, B^{\max}]$ is equidistributed [i.e., $p(B) = 1/B^{\max}$ for all B], the parameter λ is not. Its probability density is given by

$$q(\lambda) = \frac{1}{\lambda t B^{\max}}, \tag{43}$$

as can be checked by direct calculation. Employing the density $q(\lambda)$ we can calculate the mean variance using Eq. (15):

$$\sum_m p(m)\mathrm{Var}_m(\lambda) = \frac{1-e^{-2tB^{\max}}}{2tB^{\max}} - x^2 \sum_m \frac{\left(\mathrm{tr}\left\{E_m\left[\left(1-\frac{xtB^{\max}}{2}\right)\rho_1 + \left(\frac{xtB^{\max}}{2}\right)\rho_2\right]\right\}\right)^2}{\mathrm{tr}\{E_m[x\rho_1 + (1-x)\rho_2]\}};$$

$$x = \frac{1-e^{-tB^{\max}}}{tB^{\max}}. \tag{44}$$

The remaining optimization problem is essentially the same as in the previous case. The only difference is that the weights of ρ_1 and ρ_2 in the previous case [(2/3,1/3) and (1/2,1/2)] are now replaced by more complicated expressions. Due to that, the discussion from above for the case where ρ_1 and ρ_2 are pure, is no longer valid, since the orthogonality of $\vec{r}_b$ and $\Delta\vec{r}$ is not guaranteed anymore. However, the solution of the general case is directly applicable, the only difference is the new definition of $\vec{r}_a$ and $\vec{r}_b$. Hence the calculations of Sec. IV can be applied, and finally Eq. (33) solves the problem of the estimation of the decay parameter.

V. OPTIMAL MEASUREMENT FOR HIGHER DIMENSIONAL SYSTEMS

Let us now discuss the case of higher dimensional systems. That is, we consider a single copy of the state $\rho_\lambda = \lambda\rho_1 + (1-\lambda)\rho_2$ where the ρ_i are density matrices acting on a d-dimensional complex Hilbert space $\mathcal{H}_d$.

A general solution for this case is quite complicated. However, for many important cases the problem can be solved as follows. Let us choose d^2 operators G_i, $i=0,\ldots,d^2-1$ such that they form an orthonormal basis of the operator space. This means they are Hermitian and fulfill $\mathrm{tr}(G_iG_j)=\delta_{ij}$. We can choose them in such a way that ρ_1 as well as ρ_2 can be written as linear combinations of $G_0=\mathbb{1}/\sqrt{d}$, G_1, and G_2. Any Hermitian operator with trace one can be written as

$$O = \frac{1}{d}\left(\mathbb{1} + \sqrt{d^2-d}\sum_{i=1}^{d^2-1} r_i G_i\right) \tag{45}$$

and if O denotes a pure state we have $\mathrm{Tr}(O^2)=1$ which is equivalent to $\Sigma_i r_i^2 = 1$.

Now one can apply the results of the qubit case: First, one can argue as in Proposition 3 that the effects of the optimal POVM are linear combinations of G_0, G_1, and G_2. Then, the previous section allows one to compute the optimal α_0 and the corresponding vectors. Note that the normalizations $1/d$ and $\sqrt{d^2-d}$ in Eq. (45) are chosen in such a way that all the formulas of the qubit case can be applied without modification.

The drawback of this ansatz is that the obtained optimal vectors are not guaranteed to correspond to valid POVMs. Namely, in contrast to qubits, for higher dimensional systems the condition $\Sigma_i r_i^2=1$ does only imply that $\mathrm{tr}(O^2)=1$, but not that O is positive. The conditions for positivity for Bloch vectors in higher dimensional systems are quite involved [13,14]. However, if the resulting solution yields positive effects, the obtained solution is clearly the optimal one. We will discuss now important examples when this is the case.

Let us first consider the case when ρ_1 and ρ_2 have support on the same two-dimensional subspace of $\mathcal{H}_d$. This is, for instance, the case if $\rho_1=|\psi_1\rangle\langle\psi_1|$ and $\rho_2=|\psi_2\rangle\langle\psi_2|$ are pure states. Then, by choosing an appropriate basis of the two-dimensional subspace, we can assume that ρ_1 and ρ_2 are real. Hence G_1 and G_2 can be chosen as the Pauli matrices on the subspace, and the solution of the two qubit case can directly be applied. Moreover, if $\rho_1=|\psi_1\rangle\langle\psi_1|$ and $\rho_2=|\psi_2\rangle\langle\psi_2|$ the optimal measurement consists of a von Neumann measurement of an observable commuting with $|\psi_1\rangle\langle\psi_1|-|\psi_2\rangle\langle\psi_2|$.

The other important case occurs if the first state is a pure state, $\rho_1=|\psi\rangle\langle\psi|$, mixed with white noise $\rho_2=\mathbb{1}/d$. Then we have to choose $G_1=(\mathbb{1}-d|\psi\rangle\langle\psi|)/\sqrt{d^2-1}$ and G_2 arbitrary. The optimal measurement is then a von Neumann-Lüders measurement with two effects: $P_1=|\psi\rangle\langle\psi|$ and $P_2=\mathbb{1}-|\psi\rangle\langle\psi|$.

A. Verifying production of entangled states

The last example is similar to a task often occurring in experiments. Namely, one aims to produce an entangled state $|\psi\rangle$, however, noise is added during the preparation process. This situation may be modeled by writing the actual prepared state as

$$\rho(\lambda) = \lambda|\psi\rangle\langle\psi| + (1-\lambda)\frac{1}{d}\mathbb{1}. \tag{46}$$

Now one would like to know whether the state is entangled or not. Our results deliver now a possible strategy for this decision: one may estimate λ with the methods outlined above and then apply separability criteria to the state $\rho(\lambda_{\mathrm{est}})$. However, since only one copy is available, this does not allow one to detect the entanglement unambiguously. It is interesting to see that this method is different from the standard method for many copies. Then, entanglement witnesses allow the unambiguous detection, since for them a negative mean value guarantees entanglement [15,16]. Indeed, these are different observables: Taking the two-qubit case and $|\psi\rangle = \alpha|01\rangle + \beta|10\rangle$ the optimal witness is $\mathcal{W} = |00\rangle\langle 00| + |01\rangle \times \langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11|$ [17] which is different from the observable which leads to the best estimate of λ.

B. Optimal measurement for commuting states

Another case where we can determine the optimal measurement to estimate the qudit state ρ_λ in Eq. (2) is given when ρ_1 and ρ_2 are commuting states,

$$[\rho_1, \rho_2] = 0. \tag{47}$$

It turns out that measurements leading to the least mean variance are von Neumann measurements of observables O which commute with ρ_1 and ρ_2, that is, they have the same eigenvectors as ρ_1, ρ_2. The proof can be accomplished in two steps. First we prove that we can restrict the search for optimal measurements to the class of POVMs with effects commuting with ρ_1 and ρ_2, then we argue that among these measurements the projection measurements yield the best estimation of ρ_λ.

Proposition 5. Let ρ_1 and ρ_2 be commuting, $[\rho_1, \rho_2] = 0$ and let the POVM $\mathcal{P} = \{E_m\}$ represent a measurement to estimate the state $\rho_\lambda = \lambda\rho_1 + (1-\lambda)\rho_2$. Then there is a projective measurement $\widetilde{\mathcal{P}} = \{F_m\}$ with $[F_m, \rho_1] = 0$ for all m which satisfies $Q(\mathcal{P}) = Q(\widetilde{\mathcal{P}})$.

Proof. The commuting states ρ_1 and ρ_2 are simultaneously diagonalizable,

$$\rho_1 = \sum_i t_i |i\rangle\langle i| \quad \text{and } \rho_2 = \sum_i s_i |i\rangle\langle i|. \tag{48}$$

Thus $Q(\mathcal{P})$, see Eq. (13), can be expressed as

$$Q(\widetilde{\mathcal{P}}) = \frac{1}{4}\sum_m \frac{\left[\sum_i e_{mi}\left(\frac{2}{3}t_i + \frac{1}{3}s_i\right)\right]^2}{\sum_i e_{mi}\frac{1}{2}(\rho_1 + \rho_2)} \tag{49}$$

with $e_{mi} := \langle i|E_m|i\rangle$. Now, the positive operators $F_m := \sum_i e_{mi}|i\rangle\langle i|$ form a POVM, i.e., they satisfy the completeness relation:

$$\sum_m F_m = \sum_{m,i} |i\rangle\langle i|E_m|i\rangle\langle i| = \mathbb{1}. \tag{50}$$

It is easy to see that the commutative POVM $\widetilde{\mathcal{P}} = \{F_m\}$ leads to the same value of Q as $\mathcal{P}$. Hence we have only to consider POVMs which commute with ρ_1 and ρ_2 to find a measurement which maximizes Q. In addition we already learned in Sec. III that an optimal POVM is distinguished by pure effects. Together with the commutativity of the effects it follows that the optimal measurement is a projective one. □

VI. CONCLUSION

In conclusion, we have studied parameter estimation for quantum states, when only one copy of the state is available. For one qubit, we solved the problem by explicitly constructing the measurement that minimizes the deviation between the true value of the parameter and the estimated one, using a Bayesian estimator. Furthermore, we showed how the results from the qubit case can readily be used to solve this problem for important higher dimensional cases.

Our work can be extended into several directions. First, one may look at higher dimensional systems, trying to find general solutions for this case. Here, it would be of great interest to find cases where, contrary to the qubit case, general POVMs allow for a better parameter estimation than von Neumann measurements. For practical purposes, it may also be relevant to develop optimal measurement strategies for several but a finite number of copies.

ACKNOWLEDGMENTS

We thank D. Janzing and A. Scherer for useful discussions. This work was supported by the FWF and the EU (OLAQUI, SCALA, QICS) as well as the Center for Applied Photonics in Konstanz, Germany.

[1] For an overview see C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976); A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982); *Quantum State Estimation*, edited by M. Paris and J. Rehacek (Springer, Berlin, 2004).

[2] A. Chefles, Contemp. Phys. **41**, 401 (2000); P. Raynal, N. Lütkenhaus, and S. J. van Enk, Phys. Rev. A **68**, 022308 (2003); S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, Phys. Rev. Lett. **96**, 070401 (2006).

[3] J. A. Bergou and M. Hillery, Phys. Rev. Lett. **94**, 160501 (2005).

[4] See, e.g., D. G. Fischer, S. H. Kienle, and M. Freyberger, Phys. Rev. A **61**, 032306 (2000); E. Bagan, M. A. Ballester, R. D. Gill, R. Munoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **97**, 130501 (2006), and references therein.

[5] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. (N.Y.) **247**, 135 (1996); see also S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[6] J. Audretsch, L. Diosi, and T. Konrad, Phys. Rev. A **68**, 034302 (2003).

[7] T. Leonard and J. S. Hsu, *Bayesian Methods: An Analysis for Statisticians and Interdisciplinary Researchers* (Cambridge Univ. Press, Cambridge, England, 2001).

[8] J. Audretsch, T. Konrad, and A. Scherer, Phys. Rev. A **63**, 052102 (2001).
[9] J. Audretsch, F. E. Klee, and T. Konrad, Phys. Lett. A **361**, 212 (2007).
[10] L. Diosi, T. Konrad, A. Scherer, and J. Audretsch, J. Phys. A **39**, L575 (2006).
[11] G. M. D'Ariano, P. Lo Presti, and P. Perinotti, J. Phys. A **38**, 5979 (2005).
[12] H.-J. Briegel and B.-G. Englert, Phys. Rev. A **47**, 3311 (1993).
[13] G. Kimura, Phys. Lett. A **314**, 339 (2003).
[14] M. S. Byrd and N. Khaneja, Phys. Rev. A **68**, 062322 (2003).
[15] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[16] B. Terhal, Phys. Lett. A **271**, 319 (2000).
[17] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera, J. Mod. Opt. **50**, 1079 (2003).