PHYSICAL REVIEW A, VOLUME 61, 010304(R)

Maximum-likelihood estimation of the density matrix

K. Banaszek,[1,2] G. M. D'Ariano,[1] M. G. A. Paris,[1] and M. F. Sacchi[1]
[1]*INFM and Dipartimento di Fisica ''A. Volta,'' Università di Pavia, via A. Bassi 6, I-27100 Pavia, Italy*
[2]*Instytut Fizyki Teoretycznej, Uniwersytet Warszawski, Hoża 69, PL-00-681 Warszawa, Poland*
(Received 15 September 1999; published 8 December 1999)

We present a universal technique for quantum-state estimation based on the maximum-likelihood method. This approach provides a positive-definite estimate for the density matrix from a sequence of measurements performed on identically prepared copies of the system. The method is versatile and can be applied to multimode radiation fields as well as to spin systems. The incorporation of physical constraints, which is natural in the maximum-likelihood strategy, leads to a substantial reduction of statistical errors. Numerical implementation of the method is based on a particular form of the Gauss decomposition for positive-definite Hermitian matrices.

PACS number(s): 03.67.−a, 03.65.Bz

In quantum mechanics, the achievable information on a physical system is encoded into the density matrix $\hat{\varrho}$, which allows one to evaluate all possible expectation values through the Born statistical rule $\langle \hat{O} \rangle = \mathrm{Tr}(\hat{\varrho}\hat{O})$. In order to obtain full information on a quantum system we need to estimate its density matrix. In principle, this can be accomplished by successive measurements on repeated identical preparations of the same system. With a proper choice of the measurements, and after collecting a suitably large number of data, we can arrive at a reliable knowledge of the quantum state of the system.

The problem of inferring the complete quantum state from experimental data has received a lot of attention over the past several years. Physical systems whose quantum state has been fully characterized in recent experiments include now a single light mode [1], a diatomic molecule [2], a trapped ion [3], and an atomic beam [4]. These fascinating advances stimulate further theoretical research in two main directions: on one hand, in implementing effective measurement schemes that connect the density matrix to directly observable quantities, on the other hand, in designing efficient data processing algorithms in a practical experimental setup in order to extract the optimal amount of information on the quantum state. In a laboratory, we always deal with finite ensembles of copies of the measured system [5]. In addition, the process of detection is usually affected by various imperfections. This implies the need of developing novel tools specifically designed to process realistic and finite experimental samples.

In this Rapid Communication we present a general-purpose method for quantum-state estimation based on the maximum-likelihood (ML) approach [6,7]. We consider the statistical treatment of a sample of measurements performed on repeated preparations of a given system. The approach presented in this Rapid Communication is very general: it allows one to extract the information on the quantum state from data collected in a generic scheme, without assuming any specific form of the measurement. Its principle of operation is to find the quantum state that is most likely to generate the observed data. This idea is quantified and implemented using the concept of the likelihood functional.

The ML strategy is an entirely different approach to quantum-state measurement compared to the standard quantum-tomographic techniques [8,9]. In quantum tomography the expectation value of an operator is obtained by averaging a special function (so-called ''pattern function'') of experimental data of a sufficiently complete set of observables—a ''quorum'' of observables. In homodyne tomography the quorum observables are the quadratures of the electromagnetic (e.m.) field for varying phase with respect to the local oscillator. Hence, typically, a matrix element of the quantum state is obtained by averaging its pertaining pattern function over data. This method is very general and efficient; however, in the averaging procedure, the matrix elements are allowed to fluctuate statistically through negative values, with resulting large statistical errors.

In contrast, the ML method estimates the quantum state as a whole. Such a procedure incorporates *a priori* knowledge about relations between elements of the density matrix. This guarantees positivity and normalization of the matrix, with the result of a substantial reduction of statistical errors. These advantages of the ML approach are inevitably related to increased computational complexity of the estimation procedure, which remains a highly nontrivial problem, even if we resort to numerical means. We present in this Rapid Communication an alternative general solution to this problem, which provides an effective numerical algorithm for the ML estimation of the density matrix.

The task of estimating the density matrix has been approached also using the least-squares inversion [10,11], based on the assumption of experimental Gaussian noise. In contrast, our estimation strategy will be derived from the exact statistical description of raw results. Furthermore, though the least-squares inversion allows one to reduce the statistical uncertainty by appropriate regularization, it does not guarantee in principle that the reconstructed density matrix is positive definite. So far, the positivity constraints have been applied only in the much simpler problem of reconstructing the photon-number distribution [7,10]. Here, we will demonstrate how to implement in numerical calculations the complete set of physical constraints on the density matrix.

We start with the derivation of the likelihood functional $\mathcal{L}(\hat{\varrho})$, which links the raw experimental results with the object to be reconstructed, i.e., the density matrix. The physical situation we have in mind is an experiment consisting of N

measurements performed on identically prepared copies of a given system. Quantum mechanically, each measurement is described by a positive operator-valued measure (POVM). We shall denote by $\hat{\mathcal{F}}_i$ the POVM corresponding to the particular measurement and the observed outcome of the ith run. The likelihood functional $\mathcal{L}(\hat{\varrho})$ describes the probability of obtaining the set of outcomes for a given density matrix $\hat{\varrho}$. For measurements performed on repeated preparations of the system, it is given by the product

$$\mathcal{L}(\hat{\varrho}) = \prod_{i=1}^{N} \mathrm{Tr}(\hat{\varrho}\hat{\mathcal{F}}_i). \tag{1}$$

After the experiment is performed, the operators $\hat{\mathcal{F}}_i$ are determined by the outcomes of the measurements. The unknown element of the above expression, which we want to infer from our data, is the density matrix describing the measured ensemble. The general estimation strategy of the ML technique is to maximize the likelihood functional over the set of the density matrices. Several properties of the likelihood functional are easily found, if we restrict ourselves to finite-dimensional Hilbert spaces. In this case, it can be easily proved that $\mathcal{L}(\hat{\varrho})$ is a concave function defined on a convex and closed set of density matrices. Therefore, its maximum is achieved either on a single isolated point, or on a convex subset of density matrices. In the latter case, the experimental data are insufficient to provide a unique estimate for the density matrix using the ML strategy. On the other hand, the existence of a single maximum allows us to assign unambiguously the ML estimate for the density matrix. This estimate satisfies all the physical constraints, such as normalization and positivity.

The ML estimation of the quantum state, despite its elegant general formulation, presents a highly nontrivial constrained optimization problem, even if we resort to purely numerical means. The central difficulty lies in the appropriate parametrization of the set of all density matrices. The parameter space should be of the minimum dimension in order to preserve the maximum of the likelihood function as a single isolated point. Additionally, the expression of quantum expectation values in terms of this parametrization should enable fast evaluation of the likelihood function, as this step is performed many times in the course of numerical maximization.

Here, we introduce a parametrization of the set of density matrices that provides an efficient algorithm for maximization of the likelihood function. We represent the density matrix in the form

$$\hat{\varrho} = \hat{T}^{\dagger}\hat{T}, \tag{2}$$

which automatically guarantees that $\hat{\varrho}$ is positive and Hermitian. The remaining condition of unit trace $\mathrm{Tr}\,\hat{\varrho} = 1$ will be taken into account using the method of Lagrange multipliers. In order to achieve the minimal parametrization, we assume that $\hat{T}$ is a complex lower triangular matrix, with real elements on the diagonal. This form of $\hat{T}$ is motivated by the Cholesky decomposition known in numerical analysis [12] for the arbitrary non-negative Hermitian matrix. For an M-dimensional Hilbert space, the number of real parameters in the matrix $\hat{T}$ is $M + 2M(M-1)/2 = M^2$, which equals the number of independent real parameters for a Hermitian matrix. This confirms that our parametrization is minimal, up to the unit trace condition.

In numerical calculations, it is convenient to replace the likelihood functional by its natural logarithm, which of course does not change the location of the maximum. Thus the function subjected to numerical maximization is given by

$$L(\hat{T}) = \sum_{i=1}^{N} \ln \mathrm{Tr}(\hat{T}^{\dagger}\hat{T}\hat{\mathcal{F}}_i) - \lambda\, \mathrm{Tr}(\hat{T}^{\dagger}\hat{T}), \tag{3}$$

where λ is a Lagrange multiplier accounting for normalization of $\hat{\varrho}$ that equals the total number of measurements N [13]. This formulation of the maximization problem allows one to apply standard numerical procedures for searching the maximum over the M^2 real parameters of the matrix $\hat{T}$. The examples presented below use the downhill simplex method [15].

Our first example is the application of the ML estimation in quantum homodyne tomography of a single-mode radiation field [8], which is so far the most successful method in measuring nonclassical states of light [1,14]. The experimental apparatus used in this technique is the homodyne detector. The realistic, imperfect homodyne measurement is described by the positive operator-valued measure

$$\hat{\mathcal{H}}(x;\varphi) = \frac{1}{\sqrt{\pi(1-\eta)}} \exp\left(-\frac{(x - \sqrt{\eta}\hat{x}_{\varphi})^2}{1-\eta} \right), \tag{4}$$

where η is the detector efficiency and $\hat{x}_{\varphi}$ is the quadrature operator, depending on the externally adjustable local oscillator (LO) phase φ [16].

After repeating the measurement N times, we obtain a set of pairs $(x_i;\varphi_i)$ consisting of the outcome x_i and the LO phase φ_i for the ith run, where $i = 1,\ldots,N$. The log-likelihood functional is given by Eq. (3) with $\hat{\mathcal{F}}_i \equiv \hat{\mathcal{H}}(x_i;\varphi_i)$. Of course, for a light mode it is necessary to truncate the Hilbert space to a finite-dimensional basis. We shall assume that the highest Fock state has $M-1$ photons, i.e., that the dimension of the truncated Hilbert space is M. For the expectation $\mathrm{Tr}[\hat{T}^{\dagger}\hat{T}\hat{\mathcal{H}}(x;\varphi)]$ it is necessary to use an expression that is explicitly positive, in order to protect the algorithm against the occurrence of small negative numerical arguments of the logarithm function. A simple derivation yields

$$\mathrm{Tr}[\hat{T}^{\dagger}\hat{T}\hat{\mathcal{H}}(x;\varphi)] = \sum_{k=0}^{M-1} \sum_{j=0}^{k} \left| \sum_{n=0}^{k-j} \langle k|\hat{T}|n+j\rangle B_{n+j,n} \langle n|x\rangle e^{in\varphi} \right|^2, \tag{5}$$

where $B_{n+j,n} = [\binom{n+j}{n} \eta^n (1-\eta)^j]^{1/2}$ and $\langle n|x\rangle = H_n(x)\exp(-x^2/2)/\sqrt{2^n n!\,\pi^{1/2}}$ are eigenstates of the harmonic oscillator in the position representation—$H_n(x)$ being the nth Hermite polynomial.

We have applied the ML technique to reconstruct the density matrix in the Fock basis from Monte Carlo simulated homodyne statistics. Figure 1 depicts the matrix elements of

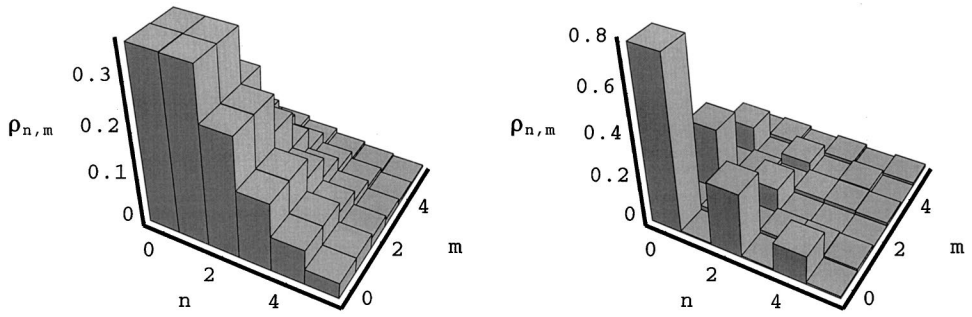

FIG. 1. ML reconstruction of the density matrix of a single-mode radiation field. On the left are the matrix elements obtained for a coherent state with $\langle \hat{a}^\dagger \hat{a} \rangle = 1$ photon; on the right for a squeezed vacuum with $\langle \hat{a}^\dagger \hat{a} \rangle = 0.5$ photons. In both cases the ML technique has been applied to a sample of 50 000 simulated homodyne data and for quantum efficiency $\eta = 80\%$.

the density operator as obtained for a coherent state and a squeezed vacuum, respectively. Remarkably, only 50 000 homodyne data have been used for quantum efficiency at photodetectors $\eta = 80\%$. The truncation dimension (which depends only on the available computational resources—typically $M = 10$ on an Intel 686, 512 Mb RAM, running REDHAT Linux 5.2) is not a crucial parameter, since more excited states can be reconstructed by adaptive techniques [18].

Since statistical aspects of standard quantum homodyne tomography have been thoroughly studied [17], this gives us an opportunity to compare it with the ML estimation. In the tomographic approach, statistical errors are known to grow rapidly with decreasing efficiency η of the detector. In contrast, the elements of the density matrix reconstructed using the ML approach remain bounded, as the whole matrix must satisfy positivity and normalization constraints. This results in much smaller statistical errors. As a comparison one could see that the same precision of the reconstructions in Fig. 1 could be achieved using $10^7 - 10^8$ data samples with the conventional quantum tomography of Ref. [8]. On the other hand, in order to find numerically the ML estimate we need to set *a priori* the cutoff parameter for the photon number, and its value is limited by increasing computation time.

Another relevant example is the reconstruction of the quantum state of the two-mode field using single-LO homodyning [19]. Here, the full joint density matrix can be measured by scanning the quadratures of all possible linear combinations of modes. For two modes the measured quadrature operator is given by $\hat{x}_{\theta\psi_0\psi_1} = (\hat{a}e^{-i\psi_0}\cos\theta + \hat{b}e^{-i\psi_1}\sin\theta + \text{H.c.})/\sqrt{2}$, where $(\theta, \psi_0, \psi_1) \in S^2 \times [0, 2\pi]$, S^2 being the Poincaré sphere and one phase ranging between 0 and 2π. In each run these parameters are chosen randomly. The POVM describing the measurement is given by the right-hand side of Eq. (4), with $\hat{x}_\varphi$ replaced by $\hat{x}_{\theta\psi_0\psi_1}$, and thus the probability distribution is written as

$$\mathrm{Tr}[\hat{T}^\dagger \hat{T} \hat{\mathcal{H}}(x; \theta, \psi_0, \psi_1)] = \sum_{\substack{k_1, k_2 \\ j, n_2}} \left| \sum_{\substack{m_1, m_2 \\ n_1}} \langle k_1 k_2 | \hat{T} | m_1 m_2 \rangle \times \langle m_1 m_2 | \hat{U}^\dagger(\theta, \psi_0, \psi_1) | n_1 + j, n_2 \rangle B_{n_1+j, n_1} \langle n_1 | x \rangle \right|^2 . \tag{6}$$

We have simulated an experiment for the two orthogonal states $|\Psi_1\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and $|\Psi_2\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. We reconstructed the density matrix in the two-mode Fock basis using the ML technique. The results are depicted in Fig. 2.

Finally, we mention that the ML procedure can be applied also for reconstructing the density matrix of spin systems. For example, let us consider N repeated preparations of a pair of spin-1/2 particles. The particles are shared by two parties. In each run, the parties select randomly and independently from each other a direction along which they perform spin measurement. The obtained result is described by the joint projection operator (spin coherent states [20]) $\hat{\mathcal{F}}_i = |\Omega_i^A, \Omega_i^B\rangle\langle\Omega_i^A, \Omega_i^B|$, where Ω_i^A and Ω_i^B are the vectors on the Bloch sphere corresponding to the outcomes of the ith

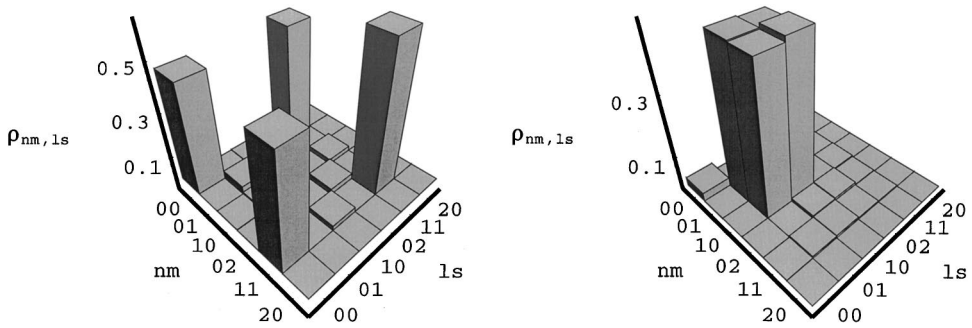

FIG. 2. ML reconstruction of the density matrix of a two-mode radiation field. On the left are the matrix elements obtained for the state $|\Psi_1\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$; on the right for $|\Psi_2\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. For $|\Psi_1\rangle$ we used 100 000 simulated homodyne data and $\eta = 80\%$; for $|\Psi_2\rangle$ we used 20 000 data and $\eta = 90\%$.

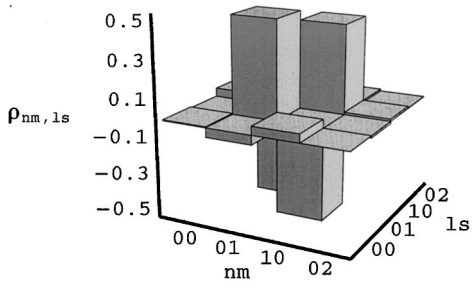

FIG. 3. ML reconstruction of the density matrix of a pair of spin-1/2 particles in the singlet state. The particles are shared by two parties. In each run, the parties select randomly and independently from each other a direction along which they perform spin measurement. The matrix elements have been obtained by a sample of 500 simulated data.

run, and the indices A and B refer to the two particles. As in the previous examples, it is convenient to use an expression for the quantum expectation value $\mathrm{Tr}(\hat{T}^\dagger\hat{T}\hat{\mathcal{F}}_i)$ that is explicitly positive. The suitable form is

$$\mathrm{Tr}(\hat{T}^\dagger\hat{T}\hat{\mathcal{F}}_i)=\sum_\mu |\langle\mu|\hat{T}|\Omega_i^A,\Omega_i^B\rangle|^2, \tag{7}$$

where $|\mu\rangle$ is an orthonormal basis in the Hilbert space of the two particles. The result of a simulated experiment with only 500 data for the reconstruction of the density matrix of the singlet state is shown in Fig. 3.

We conclude this Rapid Communication with a brief discussion of the statistical uncertainty of the ML estimate. The likelihood function can be formally regarded as a probability distribution on the parameter space. In our case, this space is spanned by M^2 real parameters that constitute the triangular matrix $\hat{T}$. We shall denote these parameters in the vector form as **t**. The formal distribution is given, up to the normalization constant, by $\delta[\mathrm{Tr}(\hat{T}^\dagger\hat{T})-1]\exp L(\hat{T})$. In the limit of the large number of measurements, $\exp L(\hat{T})$ takes the form of the Gaussian [21], with the quadratic form in the exponent given by the matrix $G=-\partial^2 L/\partial\mathbf{t}\partial\mathbf{t}'$. Furthermore, the constraint $\mathrm{Tr}(\hat{T}^\dagger\hat{T})=1$ means locally orthogonality to the gradient $\mathbf{u}=\partial\,\mathrm{Tr}(\hat{T}^\dagger\hat{T})/\partial\mathbf{t}$. The covariance matrix for the parameters **t** is consequently given by [22]

$$V=G^{-1}-G^{-1}\mathbf{u}\mathbf{u}^T G^{-1}/\mathbf{u}^T G^{-1}\mathbf{u}. \tag{8}$$

With this result, we can estimate errors for the density matrix using simply the propagation law applied to Eq. (2).

Summarizing, we have developed a universal maximum likelihood algorithm for estimating the density matrix. With respect to conventional quantum tomography this method has the great advantage of needing much smaller experimental samples, making experiments with low data rates now feasible, albeit with a truncation of the Hilbert space dimension. We have shown that the method is general and the algorithm has a solid methodological background, its reliability being confirmed in a number of Monte Carlo simulations.

We would like to thank Zdenek Hradil for interesting discussions. This work has been cosponsored by MURST under the project ''Amplificazione e Rivelazione di Radiazione Quantistica.'' K.B. is supported by INFM and by KBN Grant No. 2P03B 089 16. M.F.S. is supported by INFM PAIS 1999.

[1] D. T. Smithey *et al.*, Phys. Rev. Lett. **70**, 1244 (1993); G. Breitenbach *et al.*, Nature (London) **387**, 471 (1997).
[2] T. J. Dunn *et al.*, Phys. Rev. Lett. **74**, 884 (1995).
[3] D. Leibfried *et al.*, Phys. Rev. Lett. **77**, 4281 (1996).
[4] Ch. Kurtsiefer *et al.*, Nature (London) **386**, 150 (1997).
[5] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995); R. Derka *et al.*, *ibid.* **80**, 1571 (1998).
[6] Z. Hradil, Phys. Rev. A **55**, R1561 (1997); in the context of phase measurement, see S. L. Braunstein *et al.*, Phys. Rev. Lett. **69**, 2153 (1992).
[7] K. Banaszek, Phys. Rev. A **57**, 5013 (1998).
[8] G. M. D'Ariano *et al.*, Phys. Rev. A **52**, R1801 (1995). For a review, see G. M. D'Ariano, in *Quantum Optics and the Spectroscopy of Solids*, edited by T. Hakioğlu *et al.* (Kluwer, Dordrecht, 1997), p. 139. The quantum tomographic method has been recently generalized to arbitrary dynamical systems [G. M. D'Ariano (unpublished)].
[9] K. Banaszek, J. Mod. Opt. **46**, 675 (1999).
[10] S. M. Tan, J. Mod. Opt. **44**, 2233 (1997).
[11] T. Opatrný *et al.*, Phys. Rev. A **56**, 1788 (1997).
[12] A. S. Householder, *The Theory of Matrices in Numerical Analysis* (Blaisdell, New York, 1964), Sec. 5.2.
[13] Writing $\hat{\varrho}$ in terms of its eigenvectors $|\psi_\mu\rangle$ as $\hat{\varrho}=\Sigma_\mu y_\mu^2|\psi_\mu\rangle\langle\psi_\mu|$, with real y_μ, the maximum-likelihood condition $\partial L/\partial y_\nu=0$ reads $\lambda y_\nu=\Sigma_{i=1}^N[y_\nu\langle\psi_\nu|\hat{\mathcal{F}}_i|\psi_\nu\rangle/\mathrm{Tr}(\hat{\varrho}\hat{\mathcal{F}}_i)]$, which, after multiplication by y_ν and summation over ν, yields $\lambda=N$.
[14] G. M. D'Ariano, P. Kumar, and M. F. Sacchi, Phys. Rev. A **59**, 826 (1999).
[15] W. H. Press *et al.*, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, 1992), Sec. 10.4.
[16] Equation (4) holds for strong LO, a condition usually well fulfilled in current experiments. See M. G. Raymer, J. Opt. Soc. Am. B **12**, 1801 (1995).
[17] G. M. D'Ariano *et al.*, Quantum Semiclassic. Opt. **9**, 929 (1997); G. M. D'Ariano and M. G. A. Paris, Phys. Lett. A **233**, 49 (1997).
[18] G. M. D'Ariano and M. G. A. Paris (unpublished).
[19] G. M. D'Ariano, M. F. Sacchi, and P. Kumar, Phys. Rev. A (to be published).
[20] F. T. Arecchi *et al.*, Phys. Rev. A **6**, 2211 (1972).
[21] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).
[22] W. T. Eadie *et al.*, *Statistical Methods in Experimental Physics* (North-Holland, Amsterdam, 1971), Sec. 8.3.4.