PHYSICAL REVIEW A

ATOMIC, MOLECULAR, AND OPTICAL PHYSICS

THIRD SERIES, VOLUME 55, NUMBER 3 **MARCH 1997**

RAPID COMMUNICATIONS

The Rapid Communications section is intended for the accelerated publication of important new results. Since manuscripts submitted to this section are given priority treatment both in the editorial office and in production, authors should explain in their submittal letter why the work justifies this special handling. A Rapid Communication should be no longer than 4 printed pages and must be accompanied by an abstract. Page proofs are sent to authors.

Quantum-state estimation

Z. Hradil*

Atominstitut der Österreichischen Universitäten, Schüttelstrasse 115, A-1020 Wien, Austria

(Received 11 September 1996)

An algorithm for quantum-state estimation based on the maximum-likelihood estimation is proposed. Existing techniques for state reconstruction based on the inversion of measured data are shown to be overestimated since they do not guarantee the positive definiteness of the reconstructed density matrix. [S1050-2947(97)50703-7]

PACS number(s): 03.65.Bz, 02.50.Wp, 42.50.Dv

State reconstruction belongs to the topical problems of contemporary quantum theory. This sophisticated technique is trying to determine the maximum amount of information about the system—its quantum state. Even though the history of the problem may be traced back to the early days of quantum mechanics, quantum optics opened a new era for state reconstruction. A theoretical prediction of Vogel and Risken [1] was closely followed by the experimental realization of the suggested algorithm by Smithey *et al.* [2]. Since that time, many improvements and new techniques have been proposed [3–12], to cite without requirements for completeness at least some titles from the existing literature [13]. Even if the method comes from optics, similar methods such as quantum endoscopy, are currently being used also in atomic physics [14]. Homodyne detection of quadrature operators with varying phases of local oscillators (x_ϕ,ϕ) was used as the measurement technique in the original proposal [1,2]. The algorithm served to determine the Wigner function $W(x,p)$ and also other quasiprobabilities representing the density matrix. Measurement of rotated quadrature operators may also be used for direct evaluation of the coefficient of a density matrix in the number-state representation $\rho_{m,n}$ [4] and for the analysis of multimode fields [7]. Simultaneous measurement of the pair of quadrature operators (x,p) using double homodyne or heterodyne detection directly yields the Q function $Q(\alpha)$ [8]. A surprisingly easy technique was suggested by Wallentowitz and Vogel [9] and by Banaszek and Wodkiewicz [10]. Mixing of the signal and coherent fields with controlled amplitude on the beam splitter may serve for reconstruction of the Wigner function and other distribution functions using the photon counting only. Techniques similar to the quantum-state reconstruction have been suggested for indirect observations of particle number; see, for example, Ref. [15]. Though the techniques are different as far as practical realization is concerned, they all may be comfortably represented by the formalism of generalized measurement [16]. As is well known, any measurement may be described using the probability operator measure (POM), $\hat{\Pi}(\xi)$ being any positively defined resolution of the identity operator $\hat{\Pi}(\xi)\geqslant 0$, $\int d\xi\hat{\Pi}(\xi)=\hat{1}$. The probability distribution of the outcome predicted by quantum theory is

$$w_\rho(\xi)=\mathrm{Tr}[\hat{\rho}\hat{\Pi}(\xi)], \tag{1}$$

where $\hat{\rho}$ is the density matrix of the state. The measured variable ξ represents formally the registered data being in general a multidimensional vector with the components belonging to both the discrete and continuous spectrum, as shown in the above-mentioned examples. The key point of the existing reconstruction techniques—inversion of the relation (1)—represents a nontrivial problem. The solution may be formally written as an analytical identity,

$$W_\rho(\alpha)=\int d\xi K(\alpha,\xi)w_\rho(\xi), \tag{2}$$

*Permanent address: Department of Optics, Palacký University, 17. listopadu 50, 772 07 Olomouc, Czech Republic.

$W_\rho(\alpha)$ being a representation of a density matrix. In order to find the representation $W(\alpha)$ of a density matrix corresponding to an unknown signal, the existing reconstruction techniques apply the relation (2) on the actually detected statistics $w(\xi)$.

Apart from how ingeniously the individual inversions have been done, some problems are caused by application of this treatment in quantum theory. In particular, the algorithm may give a density matrix only for such measured probability distributions, which are given *exactly* by the relation (1). Deviations between actually detected $w(\xi)$ and the true statistics $w_\rho(\xi)$ spoil the positivity of the reconstructed density matrix. There are at least the following imperfections of the detected statistics $w(\xi)$, which should always be taken into account: (i) the sampling error caused by the limited number of available scanned positions of continuous variable at which the measurement was done; (ii) the counting error caused by the limited set of available data counted at each position. For example, in Ref. [2] the former one is caused by the division of the quadrature x_ϕ into 64 bins and the phase into 27 values, whereas the latter one is caused by the detection of the quadrature x_ϕ at each bin. Other errors, such as imperfections of detectors or external noises, may appear in practice as well. In the quantum case, the algorithm based on the inversion provides a result, but does not guarantee the positive definiteness of the reconstructed density matrix [17]. In the example of Ref. [2], the positive definiteness of the reconstructed matrix has not been checked explicitly, but can be judged according to the papers [4,5,18]. Here the negative part of the photocount distribution indicates the spoiling of positive-definiteness. Even if there were a connection between the dimension of Hilbert space where this happens and the number of phases [18], a rigorous way to treat the positive definiteness within reconstruction has not yet been suggested. The goal of the existing techniques is the estimation of the error bars of the coefficients parametrizing an unknown density matrix [19] involving also the possible nonpositive parts. Ordinary reconstruction techniques, therefore, determine the coefficients of a particular representation of the density matrix as a fluctuating variable rather than a positive-definite density matrix itself. As pointed out by Jones [20] in his Ref. [12], the failure of similar methods is the rule rather than the exception in the case where the measured data underdetermine the solution. Instead of inversion of the detected data, a technique motivated by quantum information theory [16,21] and by phase-shift estimation [22] will be suggested in this paper. Previous reconstruction techniques will be embedded into the common scheme based on the maximum-likelihood estimation.

Many parameters characterizing the quantum state should be estimated in state reconstruction. As pointed out by Helstrom [16], this may be done by restricting the dimension of Hilbert space, and accepting some residual uncertainty. Similarly, Jones [20] investigated the fundamental limitations of quantum-state measurement using Bayesian methodology. On the contrary, realistic measurements, such as those in the existing techniques, will be anticipated here. Assuming the repeated (or multiple) measurement performed on the n copies of the system, the output of the observation may be parametrized by the set of states (projectors) formally denoted as $|y_1\rangle, \ldots, |y_n\rangle$, repetition of a particular outcome being allowed. Pure states represent here the case of sharp measurement, whereas unsharp measurement involving the finite resolution should be represented by an appropriate POM. Since formal considerations are valid for both these cases, the notation of sharp measurement will be kept in the following for the sake of simplicity. Maximum-likelihood estimation ascribes to such a measurement the state $\hat{\rho}$ maximizing the likelihood functional

$$\mathcal{L}(\hat{\rho}) = \prod_i^n \langle y_i|\hat{\rho}|y_i\rangle. \tag{3}$$

The aim of this contribution is to find this state and to clarify the fluctuations of such a prediction. As the mathematical tool, the inequality between the geometric and arithmetic averages of non-negative numbers q_i will be used, $(\Pi_i^n q_i)^{1/n} \leqslant (1/n)\,\Sigma_i{}^n q_i$. The equality is achieved if and only if all the numbers q_i are equal. The variables will be formally replaced by $q_i = x_i/a_i$, where $x_i \geqslant 0$ are positive and $a_i > 0$ are auxiliary positive nonzero numbers. In the following the n-dimensional vectors will be denoted, in boldface, by $\mathbf{a}$, $\mathbf{x}$, $\mathbf{y}$, etc. Assume now that the numbers q_i are chosen from the given set of values so that the value q_i appears k_i times in the collection of n data. Hence k_i represents the frequency, $f_i = k_i/n$ being the relative frequency $\Sigma_i' f_i = 1$. Parametrization explicitly revealing the frequency will be denoted by an upper prime in sums and products, indicating that the index runs over a spectrum of different values. Without loss of generality the variable $\mathbf{x}$ may be interpreted as probability $\Sigma_i' x_i = 1$, since the normalization may always be involved in auxiliary variables $\mathbf{a}$. The relation, known as Jensen's inequality [23], then reads

$$\prod_i{}' \left[\frac{x_i}{a_i}\right]^{f_i} \leqslant \sum_i{}' f_i \frac{x_i}{a_i}. \tag{4}$$

In this form it represents a remarkably powerful relation since the equality sign may be achieved for an arbitrary probability $\mathbf{x} = \mathbf{a}$. For example, the Gibbs inequality [21] follows as a special case choosing the parameters $a_i = f_i$, since the inequality (4) may be rewritten as $-\Sigma_i' f_i \ln(f_i/x_i) \leqslant 0$. These formal manipulations are tightly connected to the maximization of the likelihood function. Using the definition

$$x_i = \langle y_i|\hat{\rho}|y_i\rangle, \tag{5}$$

a_i being a subject of further consideration, the likelihood functional may be simply estimated as

$$(\mathcal{L}(\hat{\rho}))^{1/n} = \prod_i{}' (\langle y_i|\hat{\rho}|y_i\rangle)^{f_i} \leqslant \prod_j{}' a_j^{f_j} \mathrm{Tr}\{\hat{\rho}\hat{R}(\mathbf{y},\mathbf{a})\}. \tag{6}$$

The operator $\hat{R}$ is given, in general, by nonorthogonal decomposition as

$$\hat{R}(\mathbf{y},\mathbf{a}) = \sum_i{}' \frac{f_i}{a_i} |y_i\rangle\langle y_i|. \tag{7}$$

Relation (6) simply follows from the definition (3) and from the inequality (4). Further treatment is distinguished by the following specifications of auxiliary parameters $\mathbf{a}$.

Reconstructions of the wave function. Condition $a_i = f_i$ tends to considerable simplifications. Since the measurement need not be complete, $\hat{R}(\mathbf{y},\mathbf{a}=\mathbf{f}) \leq \hat{1}$, the right-hand side of the relation (6) reads

$$(6) = \prod_j' f_j^{f_j} \mathrm{Tr}\left\{ \hat{\rho} \sum_i' |y_i\rangle\langle y_i| \right\} \leq \prod_j' f_j^{f_j}. \qquad (8)$$

This represents a state-independent upper bound. The necessary condition for the equality sign in Eq. (6) is given by the conditions $\langle y_i|\hat{\rho}|y_i\rangle / a_i = \mathrm{const}$ for any i, whereas the equality sign appears in relation (8) for complete measurements. These relations, together with the normalization of relative frequencies, tend to the necessary condition for searched state $\hat{\rho}$,

$$\langle y_i|\hat{\rho}|y_i\rangle = f_i. \qquad (9)$$

This is merely the experimental counterpart of the relation (1) and hence the starting point of reconstruction based on inversion. The relation (9) may be simply inverted in the case of orthogonal measurements, which may be considered as complete on the given subspace, tending to the solution

$$\hat{\rho}_f = \sum_i' f_i |y_i\rangle\langle y_i|. \qquad (10)$$

Unfortunately, such measurements do not reveal information about the full density matrix since the nondiagonal elements are lost, as, for example, in the case of particle number measurement. Techniques dealing with orthogonal measurements are therefore not suitable for full state reconstruction, which should be based on the usage of nonorthogonal states. On the other hand, in these cases the completeness and the existence of a solution of Eq. (9) cannot be guaranteed. Quantum analogy of the Gibbs inequality corresponds to an overestimated upper bound and tends to the conditions imposed by reconstruction techniques.

Maximum-likelihood estimation. The problems with the existence of a state achieving the upper bound descend obviously from the fixing of the auxiliary parameters **a.** The remedy is to keep them free as a subject of further optimization. For any positively defined operator $\hat{R} = \Sigma_i \lambda_i |r_i\rangle\langle r_i|$, $\lambda_i \geq 0$, and any density operator $\hat{\rho}$, the simple lemma holds

$$\mathrm{Tr}(\hat{\rho}\hat{R}) \leq \max_i \lambda_i. \qquad (11)$$

The inequality sign is achieved for the density matrix corresponding to the spectral projector of operator $\hat{R}$ with maximal eigenvalue. Using this lemma, the estimation of the right-hand side of the inequality (6) then reads

$$(6) \leq \lambda(\mathbf{y},\mathbf{a}) \prod_i' a_i^{f_i}, \qquad (12)$$

where $\lambda(\mathbf{y},\mathbf{a})$ denotes formally the maximal eigenvalue of the operator $\hat{R}(\mathbf{y},\mathbf{a})$ with the corresponding eigenvector $|\psi(\mathbf{y},\mathbf{a})\rangle$. Equality signs in the chain of inequalities are achieved simultaneously if and only if

$$\frac{|\langle y_i|\psi(\mathbf{y},\mathbf{a})\rangle|^2}{a_i} = \mathrm{const}, \qquad (13)$$

independently on the index i. Finally, the maximum-likelihood estimation determines the desired state as $|\psi(\mathbf{y},\mathbf{a})\rangle$, where vector **a** is given by the solution of the set of nonlinear equations (13). The uncertainty of such a quantum-state estimation may be, according to the Bayesian formulation [20], characterized by the likelihood functional (3). Since the interpretation of the probability distribution on the space of states is rather complicated, the uncertainty of the prediction may be involved in an alternative way. The measured data are fluctuating according to the distribution function $P(\mathbf{y})$ depending on the true state of the system. Fluctuations of quantum-state estimates may be represented by the sum of independent contributions

$$\hat{\rho}_{\mathrm{MLE}} = \langle |\psi(\mathbf{y})\rangle\langle\psi(\mathbf{y})| \rangle_{\mathbf{y}} = \int d\mathbf{y} P(\mathbf{y}) |\psi(\mathbf{y})\rangle\langle\psi(\mathbf{y})|. \qquad (14)$$

This density matrix shows how closely the maximum likelihood method allows us to estimate an unknown state hidden in the measured statistics $P(\mathbf{y})$. Unfortunately, the proposed method is rather complicated and examples of reconstructions specified above should be solved separately, case by case. Considerable technical difficulties may be caused, for example, by possible degeneracy of operator $\hat{R}$ reflecting the structure of performed quantum measurement. This particular question is beyond the scope of this contribution and represents an advanced program for further reinterpretation of existing reconstruction techniques.

A developed technique may be illustrated on simple but theoretically worthwhile examples. Quantum-state reconstruction after the measurement of a Hermitian operator with an orthogonal spectrum is the simplest problem. The solution corresponds to the application of the Gibbs inequality, since the relation (9) may be solved in this case. The quantum state is then reconstructed after each measurement by the density matrix (10). This is a consequence of the possible degeneracy of the operator $\hat{R}$ mentioned above. The treatment based on the Gibbs inequality is overestimated in general. Provided that Eq. (9) is fulfilled in some special cases, then the solution should coincide with the prediction of the maximum-likelihood estimation.

The cases of strongly underdetermined data are also simple, when the state is estimated after single detection $n=1$. Assume for concreteness the standard ''measurement of Q function'' corresponding to the detection of coherent states $|y\rangle = e^{y\hat{a}^\dagger - y^*\hat{a}}|0\rangle$. If the value y is detected, the system is with the highest likelihood just in the state $|y\rangle$. Provided that system *was in a coherent state* $|\alpha\rangle$, the output fluctuates as $|\langle\alpha|y\rangle|^2/\pi$. Estimation after single detection then yields the density matrix of superposition of coherent signal α and the thermal noise [24] with the mean number of particles equal to 1,

$$\hat{\rho}_{\mathrm{MLE}} = \frac{1}{\pi} \int d^2 y e^{-|y-\alpha|^2} |y\rangle\langle y|.$$

The difference between the true state and its estimation is negligible in the case of classical fields, but considerable in the quantum domain.

Estimating the quantum state after multiple detection of coherent states, the matrix $\hat{R}$ should be diagonalized. Using the assumption for eigenstates as $|\varphi\rangle = \Sigma_i' V_i |y_i\rangle$, linear equations for desired coefficients V_i and eigenvalues λ follow as

$$\frac{f_k}{a_k} \sum_i{}' \; V_i C_{ki} = \lambda V_k, \qquad (15)$$

where $C_{ki} = C_{ik}^* = \langle y_k | y_i \rangle, C_{ii} = 1$. This solution determines the coefficients **a** according to the relation (13) as $|\Sigma_i' V_i C_{ki}|^2 / a_k =$ const for any index k. Let us illustrate this strategy in the case of double detection $n=2$ yielding the values y_1 and y_2. Parameters are given as $f_1 = f_2 = 1/2$ and without loss of generality $a_1 = 1$, $a_1/a_2 = x$. The secular equation for the maximal eigenvalue λ reads $\lambda^2 - (1+x)\lambda + x - x|C_{12}|^2 = 0$, yielding easily solutions for the maximal eigenvalue and its eigenvector. Equations (13) impose the single condition as $|C_{12}|\lambda = \sqrt{x}(\lambda - 1 + |C_{12}|^2)$. This nonlinear system of equations may be easily solved yielding an expected solution, such as $\lambda = 1 + |C_{12}|, x = 1$. The projector is given by the normalized Schrödinger-cat-like state,

$$|\varphi\rangle = \frac{1}{\sqrt{2(1+|C_{12}|)}} \, (e^{i \arg C_{12}} |y_1\rangle + |y_2\rangle).$$

The density matrix ''reconstructing'' the coherent state is then given as

$$\hat{\rho}_{\mathrm{MLE}} = \frac{1}{\pi^2} \int d^2 y_1 d^2 y_2 e^{-|y_1 - \alpha|^2 - |y_2 - \alpha|^2} |\varphi\rangle\langle\varphi|.$$

The proposed method describes easily the cases in which the data seem to be underdetermined. There is also a strong effort to apply the developed technique to the case of large data sets estimating properly the quantum state in the cases of realistic measurements.

Even if the problem of positive definiteness used for motivation may seem trivial, it has far-reaching consequences. Since a quantum state comprises the maximum possible information about the system, its proper description is of fundamental interest. The method based on maximum likelihood addresses the state reconstruction in close analogy to the ordinary methods, completing rather than denying them in the following way: For measured data, whenever the ordinary reconstruction schemes provide a positive matrix, the solution coincides with the maximum-likelihood estimation. The predictions will differ only if the ordinary scheme yields a matrix that is not positively defined.

Unfortunately, the proposed technique is nonlinear and the evaluation of realistic data for existing state reconstructions represents a nontrivial problem. Since the algorithm never admits the existence of ''negative probabilities,'' the conjecture concerning accuracy may be formulated as the following statement: *Enforcing positiveness will enhance the uncertainty of state estimation*, in comparison to the ordinary techniques. These and other questions deserve further attention, since an information-theoretic approach allows us to analyze the state reconstruction free of any additional assumptions.

I am grateful to Professor H. Rauch for the hospitality of Atominstitut der Österreichischen Universitäten and to T. Opatrný for discussions concerning state reconstruction. This work was supported by the East-West Program of the Austrian Academy of Sciences and by a grant from the Czech Ministry of Education, No. VS 96028.

[1] K. Vogel and H. Risken, Phys. Rev. A **40**, 2847 (1989).
[2] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993).
[3] U. Leonhardt, and H. Paul, J. Mod. Opt. **41**, 1427 (1994).
[4] G. M. D' Ariano, C. Macchiavello, and M. G. A. Paris, Phys. Rev. A **50**, 4298 (1994).
[5] G. M. D' Ariano, U. Leonhardt, and H. Paul, Phys. Rev. A **52**, R1801 (1995).
[6] T. Richter and A. Wünsche, Phys. Rev. A **53**, R1974 (1996).
[7] H. Kühn, D.-G. Welsch, and W. Vogel, Phys. Rev. A **51**, 4240 (1995).
[8] M. G. A. Paris, Phys. Rev. A **53**, 2658 (1996).
[9] S. Wallentowitz and W. Vogel, Phys. Rev. A **53**, 4528 (1996).
[10] K. Banaszek and K. Wódkiewicz, Phys. Rev. Lett. **76**, 4344 (1996).
[11] U. Leonhardt and M. G. Raymer, Phys. Rev. Lett. **76**, 1985 (1996).
[12] H. Paul, P. Törma, T. Kiss, and I. Jex, Phys. Rev. Lett. **76**, 2464 (1996).
[13] Besides these, there are still other methods based on the principle of maximum entropy. An unknown state is approximated on the class of density matrices specified by moments of a certain measurable operator, rather than reconstructed. This technique is beyond the scope of this contribution; see V. Bužek, G. Adam, and G. Drobný, Phys. Rev. A **54**, 804 (1996).
[14] S. Wallentowitz and W. Vogel, Phys. Rev. Lett. **75**, 2932 (1996).
[15] M. Munroe, D. Boggavarapu, M. E. Anderson, and M. G. Raymer, Phys. Rev. A **52**, R924 (1995).
[16] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[17] Positive definiteness of $\hat{\rho}$, i.e., the property $\langle x|\hat{\rho}|x\rangle \geqslant 0$ for any state $|x\rangle$, is very complicated for testing. It should not be confused with the positivity of the Wigner function, which need not be positive.
[18] U. Leonhardt, and M. Munroe, Phys. Rev. A **54**, 3682 (1996).
[19] U. Leonhardt, M. Munroe, T. Kiss, Th. Richter, and M. G. Raymer, Opt. Commun. **127** , 144 (1996).
[20] K. R. W. Jones, Phys. Rev. A **50**, 3682 (1994).
[21] C. M. Caves and P. D. Drummond, Rev. Mod. Phys. **66**, 481 (1994).
[22] A. S. Lane, S. L. Braunstein, and C. M. Caves, Phys. Rev. A **47**, 1667 (1993); Z. Hradil, R. Myška, J. Peřina, M. Zawisky, Y. Hasegawa, and H. Rauch, Phys. Rev. Lett. **76**, 4295 (1996).
[23] This inequality was used in quantum estimation, in for example, C. A. Fuchs, Ph.D. thesis, University of New Mexico (unpublished).
[24] J. Peřina, *Coherence of Light* (Kluwer Academic Publ., Dordrecht, 1985).