

Objective approach to biased tomography schemes

D. Mogilevtsev

Institute of Physics, Belarus National Academy of Sciences, F.Skarina Ave. 68, Minsk 220072, Belarus and Departamento de Física, Universidade Federal de Alagoas Cidade Universitária, 57072-970, Maceió, AL, Brazil

J. Řeháček and Z. Hradil

Department of Optics, Palacky University, 17. listopadu 50, 772 00 Olomouc, Czech Republic

(Received 8 September 2006; published 23 January 2007)

The intrinsic relationship between the maximum-likelihood quantum-state estimation and the representation of the signal is elaborated. A quantum analogy of the transfer function determines the space where a successful reconstruction can be achieved. This provides a tool for reducing the number of dimensions of the observed system based on physical characteristics of the reconstruction scheme rather than some *ad hoc* truncations. The method is illustrated with two examples of practical importance: an optical quantum homodyne tomography and a simple and robust tomography of an optical signal recorded by realistic binary detectors.

DOI: [10.1103/PhysRevA.75.012112](https://doi.org/10.1103/PhysRevA.75.012112)

PACS number(s): 03.65.Wj, 42.50.Lc

I. INTRODUCTION

The development of effective, robust, and reliable methods of quantum state reconstruction is of utmost importance for both fundamental and practical reasons. It is needed for diagnostics of quantum systems—a crucial task in future applications of quantum information technology. Experimentalists may use quantum state reconstruction for analysis of quantum processes, decoherence, and other changes occurring between the input and output ports of the quantum devices. Quantum state reconstruction is a formidable task requiring a sophisticated measurement scheme, efficient mathematical tools, and extensive computational effort for practical realization. At the early stage of quantum state tomography [1,2] the main effort was aimed to answer the question “how should the state be reconstructed?” Consequently, various strategies for collection of the data and theoretical tools for data inversion have been developed. Considerable progress has been achieved on this road. Nowadays tomography is considered as a routine experimental technique. It has been successfully applied to probe the structure of entangled states of light and ions, operations (quantum gates) with entangled states of light and ions, and analysis of the internal angular momentum structure of correlated beams, just to mention a few examples [2].

In spite of recent progress there are several open problems associated with quantum tomography. Until now rather passing attention has been paid to the question “where may the state be found?” For example, in the original quantum tomography scheme a region (field of view) should be chosen in the phase plane where the Wigner function of the state in question should be looked for. The choice of this region seems to be only loosely connected with the reconstruction procedure itself. However, such an *ad hoc* guess may have poor impact on the accuracy of the reconstruction or conversely, it may lead to more consistent results. The latter case may happen when an experimentalist seeks a result close to the *a priori* known true state. Such a tacitly accepted assumption may appear as crucial eliminating the large number of unwanted free parameters. Obviously, such a drawback erodes the notion of a reconstruction scheme as one which is objective.

The design of a tomographic measurement reaching optimal performance of the tomographic scheme is another problem which is addressed in this contribution. Since the answer to this problem depends on the technology available and cannot be therefore exhaustive, we will follow the idea of the original proposal [3] to utilize the imperfections of the single photon detectors for making tomographic measurement on a mode of light. Indeed, such schemes based on linear optics and realistic models of on/off detectors are robust and can be considered optimal from the pragmatic viewpoint of feasibility of detection. In this paper approaches and methods briefly outlined in the recently published Letter [4] are extended, discussed in more detail, and are further developed.

The outline of the paper is as follows. In Sec. II the mathematical details of maximum-likelihood (ML) estimation are discussed. In Sec. III the main issues related to successful reconstruction are summarized, which are advised to follow. This part is designated as a brief manual for the potential users of the suggested reconstruction procedure. The recommended approach is demonstrated on two examples of optical tomographic measurement: homodyne detection in Sec. IV and on/off detection in Sec. V. In Sec. VI the latter scheme is detailed by adopting the ML strategy for reconstructing the Wigner function of the measured state of light. Finally, in Sec. VII, the complete information about the measured system using the recommended approach is inferred.

II. THEORY OF ML ESTIMATION

To make this paper self-contained the derivation of the extremal equation for the ML state along the lines presented in [4] is briefly described.

A. Extremal equation

The measurement procedure is generally described by elements \mathbf{A}_j of a positive-operator-valued measure (POVM), where the index j enumerates particular sets of parameters. At this point, any specific assumptions about the nature of the measured elements \mathbf{A}_j , such as their commutation rela-

tions or group properties, need not be done. In particular, we allow for a measurement, where certain output channels of the measuring apparatus are ignored or are not accessible. Of course, in order to get a unique result of the reconstruction, a sufficiently “rich” structure of the measurement should be anticipated. The concept of the “tomographic completeness” (“informational completeness”) can be introduced. Loosely speaking, a set of measurements is tomographically complete provided that the set contains at least as many independent observations as the number of independent parameters. For example, in a p -dimensional Hilbert subspace the free parameters correspond to the $p^2 - 1$ real numbers parametrizing a p -dimensional density matrix. A tomographically complete set of observations A_j thus comprises a Hermitian operator basis in the subspace.

Since for a realistic tomography only experimentally feasible measurements are available, and since tomography must be able to handle any kind of data, we do not impose other prior criteria for the choice of the set of operators A_j with the exception of the above-mentioned tomographical completeness. The adequacy of each particular scheme, however, depends strongly on the nature of measurements done and as such it can always be verified *a posteriori* following our recommendations.

The reader may be guided by the well-known example of quantum homodyne tomography discussed in detail later on. In homodyne tomography, operators A_j are projectors onto different eigenstates of the rotated quadrature operators. Operators A_j satisfy

$$0 \leq A_j \leq 1, \quad j = 1, \dots, N_p. \quad (1)$$

Then, the probabilities of measurement outcomes are given as

$$p_j = \text{Tr}[A_j \rho], \quad (2)$$

ρ being the quantum state. By conditions (1) the probabilities generated by any state are guaranteed to be non-negative and less than one. In general, the probabilities p_j do not add up to unity, as the operator sum

$$\sum_j A_j = \mathbf{G} \geq 0 \quad (3)$$

may differ from the identity operator. Theoretical probabilities p_j can be sampled experimentally by repeating the measurement on an ensemble of identically prepared systems described by the density matrix ρ . In the course of such a repeated measurement an outcome j occurs N_j times. The aim of tomography is to find the quantum state ρ from the observed data $\{N_j\}$, $j = 1, \dots, N_p$.

The ML scenario hinges on the likelihood functional associated with the statistics of the experiment. In the following analysis, the generic form of likelihood for unnormalized probabilities is adopted

$$\log \mathcal{L} = \sum_j N_j \log \left[\frac{p_j}{\sum_{j'} p_{j'}} \right], \quad (4)$$

where index j runs over all registered data. This functional is to be maximized with respect to ρ .

This form, suggested by Fermi, is sometimes called the extended maximum likelihood principle [5]. There is a simple rationale behind the principle. The counted events are discrete and can always be cast in the generic form of Poissonian statistics with an unknown mean number of counts. Counted data N_j correspond therefore to a sampled Poissonian signal with expectations λp_j , where λ is an unknown mean total number of particles (counts). When the corresponding Poissonian log-likelihood function

$$\log \mathcal{L}_{\text{Poiss}} = \sum_j N_j \log(\lambda p_j) - \lambda \sum_j p_j \quad (5)$$

is maximized with respect to λ , a mutual normalization of probabilities in Eq. (4) is readily obtained.

The extremal equation for the maximum-likely state can be derived in three steps: (i) The positivity of ρ is made explicit by decomposing it as $\rho = \sigma^\dagger \sigma$. (ii) The likelihood (4) is varied with respect to an independent matrix σ using $\delta(\log p_j) / \delta \sigma = A_j \sigma^\dagger / p_j$; and (iii) variation obtained is set equal to zero and multiplied from the right-hand side by σ with the result

$$\mathbf{R} \rho = \mathbf{G} \rho, \quad (6)$$

where

$$\mathbf{R} = \sum_j \frac{j'}{\sum_{j'} N_{j'}} \frac{N_j}{p_j(\rho)} A_j. \quad (7)$$

Notice that while operator \mathbf{G} is defined by Eq. (3), operator \mathbf{R} depends on the particular statistical model (likelihood function). Also notice that operator \mathbf{R} depends on ρ via the state dependent probabilities p_j as indicated in Eq. (7). Equation (6) may be cast in the form of the expectation-maximization (EM) algorithm [6]

$$\mathbf{R}_G \rho_G = \rho_G, \quad (8)$$

where

$$\mathbf{R}_G = \mathbf{G}^{-1/2} \mathbf{R} \mathbf{G}^{-1/2}, \quad \rho_G = \mathbf{G}^{1/2} \rho \mathbf{G}^{1/2}.$$

This extremal equation may be iterated in a fixed orthogonal basis. Keeping the positive semidefinite character of ρ_G [by combining Eq. (6) with its Hermitian conjugate] the $(n + 1)$ th iteration reads

$$\rho_G^{(n+1)} = \mathbf{R}_G^{(n)} \rho_G^{(n)} \mathbf{R}_G^{(n)}, \quad (9)$$

where we defined

$$\mathbf{R}_G^{(n)} = \mathbf{G}^{-1/2} \mathbf{R}(\rho^{(n)}) \mathbf{G}^{-1/2}. \quad (10)$$

Starting with an initial guess, $\rho_G^{(0)}$, the iterative process is repeated until the fixed point is reached. In terms of ρ_G , the desired solution is then given by

$$\rho = \mathbf{G}^{-1/2} \rho_G \mathbf{G}^{-1/2}. \quad (11)$$

Going back to likelihood Eq. (4) it is noted that the operator \mathbf{G} , which comes from the mutual normalization of probabilities, $\sum_j p_j = \text{Tr}[\rho \mathbf{G}]$, also provides a mutual normalization of the original biased observations via the transformation

$$\sum_j \mathbf{G}^{-1/2} \mathbf{A}_j \mathbf{G}^{-1/2} = 1_G. \quad (12)$$

The subscript G , appearing on the right-hand side of Eq. (12) denotes that this identity is valid within the support of the operator \mathbf{G} . Thus any set of tomographic observations can formally be described by complete (normalized) POVM elements $\mathbf{G}^{-1/2} \mathbf{A}_j \mathbf{G}^{-1/2}$.

B. Interpretation

The closure relation (12) establishes the preferred basis for reconstruction. Due to division by the operator \mathbf{G} in Eq. (11), the reconstruction can be done only in the subspace spanned by the nonzero eigenvalues of \mathbf{G} . The spectrum of \mathbf{G} therefore plays the role of a tomographic transfer function analogous to the transfer function in optical imaging. It quantifies the resolution of the reconstruction scheme. A large eigenvalue of \mathbf{G} indicates that many observations overlap in the corresponding Hilbert subspace and this part of Hilbert space is clearly visible. The Hilbert subspace where the reconstruction was done cannot be chosen freely in a proper statistical analysis. This factor also gives a clue on how to reduce the number of dimensions of some problems, mainly those associated with infinite dimensional systems, simply by restricting the search to the subspace corresponding to the dominant eigenvalues of the operator \mathbf{G} . The result of the reconstruction can be checked easily afterward. If the reconstructed state exhibits dominant contributions from the components with relatively small eigenvalues of \mathbf{G} , the result cannot be trusted.

Naturally, the spectrum of \mathbf{G} is affected by the choice of the measurement apparatus. Its dependence on the true state of the measured system is less obvious. This situation typically arises in the experiments with quantum systems described by continuous variables. For example, in homodyne tomography, random detections of quadrature eigenvectors $|x_1(\theta)\rangle\langle x_1(\theta)|, |x_2(\theta)\rangle\langle x_2(\theta)| \dots$ are accumulated for different phases θ of the local oscillator. Only projectors actually detected enter the sum in Eq. (3). The probability that a particular projector is detected then depends on the true state. For the same reason \mathbf{G} might depend on the number of copies used for tomography, since a larger number of observations makes it more likely to explore rarely happening events. Later on we will illustrate behavior of \mathbf{G} with several examples.

It is intriguing to note that the character of the ML solution (27) is governed by the fundamental laws of quantum

mechanics. Let us present a simplified discussion valid for $\mathbf{G}=1$, the generalization to any other \mathbf{G} is straightforward. Assume that a measurement of \mathbf{A}_j is observed with values of frequencies $f_j = N_j / \sum_j N_j$. Since a total finite number of observations, $\sum_j N_j$, is distributed among $j=1, \dots, M$ channels, the registered data will fluctuate according to a multinomial statistic. The strength of these fluctuations (statistical noise) depends on the unknown true state. Since, for a given state, the fluctuation at two different channels will be different, the registered data cannot be trusted equally. Those channels with higher signal-to-noise ratio should be given larger weighting during the data inversion process than the channels of more noise. Consequently, the optimal data processing requires the knowledge of the true state. However, at the beginning of the reconstruction procedure we know nothing about the true state which governs the observed statistics. That is why the optimal estimation must be nonlinear—the unknown state must be estimated together with the quality of the noisy data. This may be achieved by weighting (renormalizing) each measurement using a coefficient ν_j . Defining the new POVM elements by

$$\mathbf{A}'_j \equiv \nu_j \mathbf{A}_j, \quad (13)$$

which would satisfy the completeness relationship

$$\sum_j \mathbf{A}'_j = 1_\rho, \quad (14)$$

an attempt may be made to match the observed relative frequencies to the new theoretical probabilities as follows:

$$\text{Tr}(\rho \mathbf{A}'_j) = f_j. \quad (15)$$

Expressing unknowns ν_j by means of Eqs. (13) and (15), inserting them into Eq. (14), and multiplying both sides of Eq. (15) by ρ produces the extremal equation

$$R\rho = \rho, \quad (16)$$

which is a special case of Eq. (6) for $\mathbf{G}=1$. Standard linear approaches to quantum tomography which in some special cases may tend to the inverse Radon transformation, also hinge upon the same relations (14) and (15) but their role is *reversed*. In standard tomography, all the measured data are equally trusted, $\nu_j=1$, so the completeness relation (14) is obeyed by definition and the desired state is given by a linear inversion of Eq. (15). Aside from other drawbacks, such a solution does not keep the necessary constraints of quantum mechanics $\rho \geq 0$. In marked contrast, in the optimal ML approach the exact correspondence Eq. (15) between the theory and data is taken as definition and the optimal quantum state is found by solving the completeness relation (14), thus reversing the roles of relations (14) and (15). It should be also noted that the ML solution related to the above-mentioned choice of parameters ν_k always exists due to the convexity of the likelihood function $\log \mathcal{L} = \sum_j f_j \log p_j$, for which Eq. (16) is the extremal equation possessing a unique solution. On the other hand, the existence of a positive semidefinite solution of the linear problem characterized by the conditions $\nu_j=1$ is not guaranteed.

Consider for comparison an alternative form of the likelihood function. Often the noise can be approximated by a

Gaussian probability distribution. The relevant part of the corresponding log-likelihood then reads

$$\log \mathcal{L}_{Gauss} = - \sum_j \frac{(p_j - N_j)^2}{2\sigma_j^2}, \quad (17)$$

where σ_j is the standard deviation of the j th channel given approximately by $\sqrt{N_j}$. This approach is sometimes preferred by experimentalists [7] though there are no good reasons for making such an approximation. Provided we have some prior knowledge about the true state this approach will likely give reasonable results, also for other fits, for example, the least-squares fit. However, when resorting to these methods we do not have any hints telling us where to look for the states since all the links between tomography, performed measurement (expressed by an operator G), and closure relation (12) are lost. Nothing is paid back for this disadvantage since the iteration algorithms are equally involved for alternative choices of likelihoods.

III. MANUAL FOR APPLICATION

Since ML tomography is a rather involved scheme, let us summarize the recommended approach for experimenters in the form of a manual with four steps, which should be followed.

A. Step 1: Design of measurement

The quality of reconstruction relies on the available quantum detection. The choice of measurement depends on the system explored and on the resources available. In our analysis, we will focus on the generic scheme formulated in the infinite dimensional Hilbert space. Here the subspace where the reconstruction can be done should be found simultaneously with the reconstructed object. In the following we will focus on homodyne detection and robust binary (on/off) detection.

B. Step 2: Field of view

Experimental resources should be adopted for the overall characterization of quantum performance. The key role is played by the operator \mathbf{G} defined by the generic relation (3). The spectrum of this positively semidefinite operator defines the efficiency of the tomographic measurement in a particular subspace. Notice that all measurements performed contribute to \mathbf{G} . In particular, in the case of homodyne detection \mathbf{G} will comprise all the observed quadrature eigenvectors. The purpose of this step is to find a field of view with good resolution provided by the particular reconstruction scheme. Good resolution can be expected only in the subspace spanned by those eigenvectors of \mathbf{G} corresponding to significantly nonzero eigenvalues. Similar reasoning can be used to define the number of degrees of freedom of an image [8], for example. The spectrum of \mathbf{G} thus plays a similar role to that of the optical transfer function in classical wave optics. If one wishes to extend the reconstruction to an even larger subspace, data must be acquired from new measurements

introduced to the tomography scheme in order to increase the desired eigenvalues of \mathbf{G} .

C. Step 3: ML reconstruction

After properly choosing the reconstruction space as indicated in Step 2, the reconstruction itself can proceed. Starting from a completely mixed state, $\rho^0 \propto 1$, defined on this subspace the iterative algorithm (8) can be employed to find the optimal state. Notice that the algorithm should not be initialized by a pure state even when we expect the true state to be one (no guarantee of convexity of the likelihood). It is also not advisable to use the Gaussian approximation of likelihood (see the remark above) or rewrite the iterative algorithm directly in terms of Wigner functions (there is no way of imposing non-negativity) even though one can be tempted by its seeming simplicity.

D. Step 4: Error analysis

In the last step the reconstruction should be completed by an error analysis. Notice that the result of the reconstruction is not just the ML state. There are also states in its neighborhood whose likelihoods are still significant that should be considered. Obviously, the rate of decrease of the likelihood with distance from the ML state defines the uncertainty of the reconstruction and puts the “error bars” on the reconstructed density matrix. All predictions based on the reconstructed state must reveal this additional uncertainty. Details associated with error analysis will be given elsewhere.

All these four steps are necessary conditions which are establishing quantum tomography as an objective tool for the analysis and diagnostics of infinite dimensional quantum systems. Indeed, previously reported results of tomographic schemes have considered the space for reconstruction *ad hoc*: But if one knows what the result should be it is not difficult to get it.

IV. QUANTUM HOMODYNE TOMOGRAPHY

The recommended approach to ML tomography is illustrated on the example of homodyne detection, which originally triggered the interests in quantum tomography [1,2]. An iterative algorithm similar to the general scheme described above has already been adopted for the reconstruction of the density matrix in Fock-state space [9,10]. However, the measurement [9,10] was considered to be unbiased $\mathbf{G}=1$, which only approximately holds for the homodyne detection. Also the subspace where the reconstruction was done was fixed *ad hoc* in the neighborhood of the (expected) true state. It is shown in this section how this approach can be improved using the general framework of objective tomography.

The POVM describing the ideal quantum homodyne tomography measurement consists of projectors to the eigenstates of quadrature operators [1]:

$$\mathbf{A}(\gamma) = |\gamma\rangle\langle\gamma|, \quad (18)$$

whose components in the Fock basis $|n\rangle$ are given by

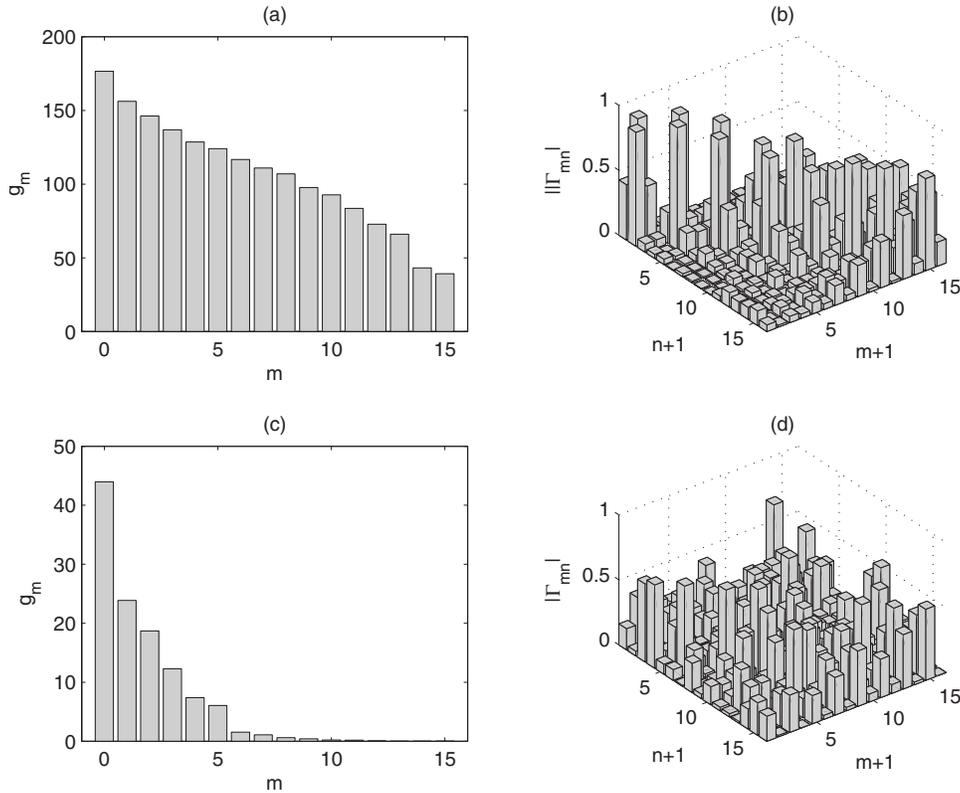


FIG. 1. Operator \mathbf{G} in homodyne tomography; (a,c) eigenvalues of \mathbf{G} ; (b,d) absolute values of its eigenvector elements computed in Fock basis and truncated at $N_{tr}=15$; m and n label the eigenvectors and their components, respectively. The γ points describing the measured projections are taken from an equidistant polar mesh, whose parameters are as follows: (a,b) 16 amplitudes $\text{abs}(\gamma) \in [0.01, 3] \times 16$ phases $\text{arg}(\gamma) \in [0, 2\pi]$ and (c,d) 4 amplitudes $\times 4$ phases from the same intervals.

$$\langle n|\gamma\rangle = \left(\frac{2}{\pi}\right)^{1/4} \frac{H_n(\sqrt{2}|\gamma|)}{\sqrt{2^n n!}} \exp\{-|\gamma|^2 + in \arg(\gamma)\}.$$

For illustration, we adopt here a measurement set and reconstruction subspace close to those used in Ref. [9], where the reconstruction of a state with the average number of photons less than one was considered. We take the same region near the origin of the phase plane, see Fig. 1, using the same truncation number $N_{tr}=15$. Let us examine the effect of several data configurations. Figures 1(c) and 1(d) correspond to a sparse data set—just $N_p=16$ different projections are taken. Clearly, when the set of measured γ -points is very small, the subspace spanned by the dominant eigenvectors of \mathbf{G} is small compared to N_{tr} , and the effective field of view does not extend up to N_{tr} photons. Increasing the number of γ -points in the chosen region to $N_p=256$, the resolution of the tomography can be improved, see Figs. 1(a) and 1(b), and the eigenvalues of the operator \mathbf{G} get closer to Fock states, as seen by comparing Figs. 1(b) and 1(d). Any state lying in this subspace can be successfully reconstructed with the help of the ML technique. Notice also that no data binning is needed here, raw data consisting of detected quadrature values can also be processed. One should, however, remember that since the \mathbf{G} operator corresponding to a homodyne tomography is not proportional to the identity operator, the correct algorithm (9) should be used instead of the simplified version adopted in Refs. [9,10].

Increasing further the number of γ -points, the \mathbf{G} operator is approaching the identity operator on the given subspace. This is because the integral taken over all projectors (18) results in the identity operator. This tendency can be seen in

Fig. 2(a), where the relative differences between the first few eigenvalues of \mathbf{G} are considerably reduced. Consequently the simplified algorithm assuming $\mathbf{G}=1$ becomes a good approximation of the correct algorithm (9) on such a small subspace. This also explains why the reconstruction done in Refs. [9,10] gave reasonable results. However, this argument would not apply if reconstructions of states with larger photon numbers were attempted. In that case even a measurement consisting of $\sim 10^5$ projections (γ -points) as used in Ref. [9] may be insufficient for establishing the closure relation $\mathbf{G}=1$ so the correct general algorithm (9) should be adopted.

This analysis illustrates how a measurement can be tuned to the required performance. Any measurement brings some information so loosely speaking any measurement can be considered informationally complete in some subspace. This feature is demonstrated in Fig. 2(c). Here, the data collected is from only the first quadrant of the phase plane. Such observations seem to be “incomplete” from the point of view of the standard inverse Radon transformation since projections from the full $(0, \pi)$ interval are needed for linear inversion. However, even such an imperfect observation may be informationally complete on a properly chosen subspace of the infinite dimensional Hilbert space and the ML algorithm can be used to extract this information.

A simulation of the homodyne tomography of a coherent state based on the quadrature measurements of Fig. 1(a) is shown in Fig. 3. For each of N_p different projections γ_j , the corresponding theoretical probability p_j has been calculated using Eqs. (2) and (18). This probability was then sampled N_r times in order to simulate N_r experimental runs. As a result, γ_j was detected $N_j \leq N_r$ times. Due to the probabilistic

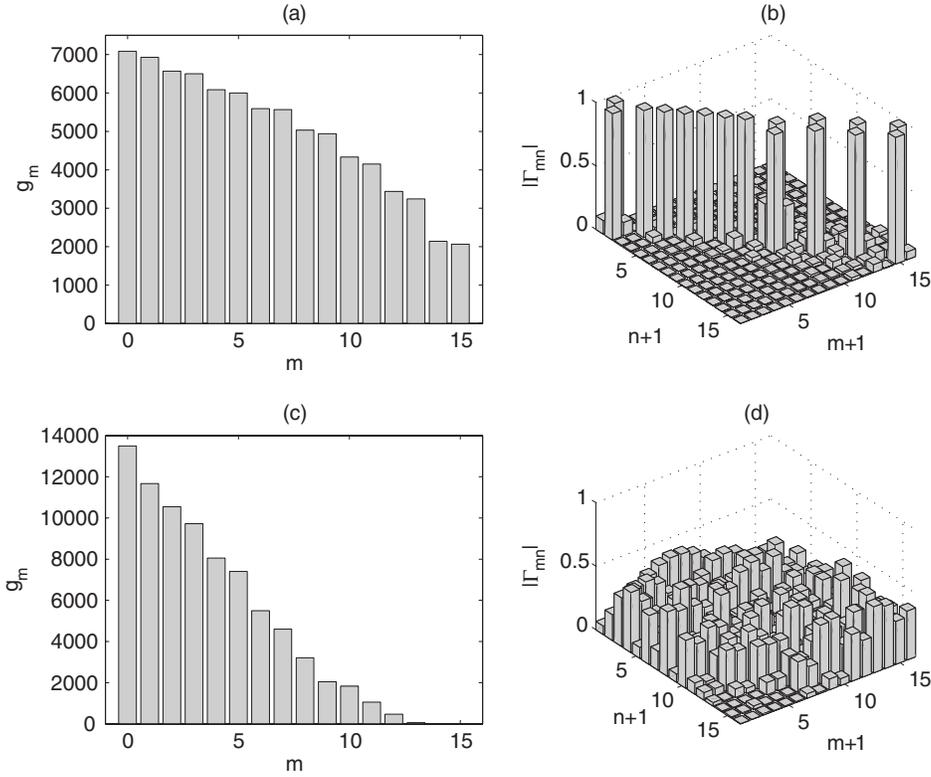


FIG. 2. Operator \mathbf{G} in homodyne tomography; for description, see Fig. 1. The γ form a polar mesh: (a,b) 121 amplitudes $\text{abs}(\gamma) \in [0.01, 3] \times 101$ phases $\text{arg}(\gamma) \in [0, 2\pi]$ and (c,d) 121 amplitudes $\text{abs}(\gamma) \in [0.01, 3] \times 101$ phases $\text{arg}(\gamma) \in [0, \pi/2]$.

nature of quantum theory, N_j in general differs from the mean value $\bar{N}_j = p_j N_r$. The set of generated data was then processed using the iterative ML procedure of Eq. (9). The confidence intervals on the reconstructed density matrix elements can be provided by calculating the variances

$$\sigma(\rho_{mn}) = [F(\rho_{mn}) N_{mes}]^{-1/2}, \quad (19)$$

where $N_{mes} = N_p N_r$ is the total number of measurements, and the Fisher information F can be defined for the real part of the density matrix elements as follows [11]:

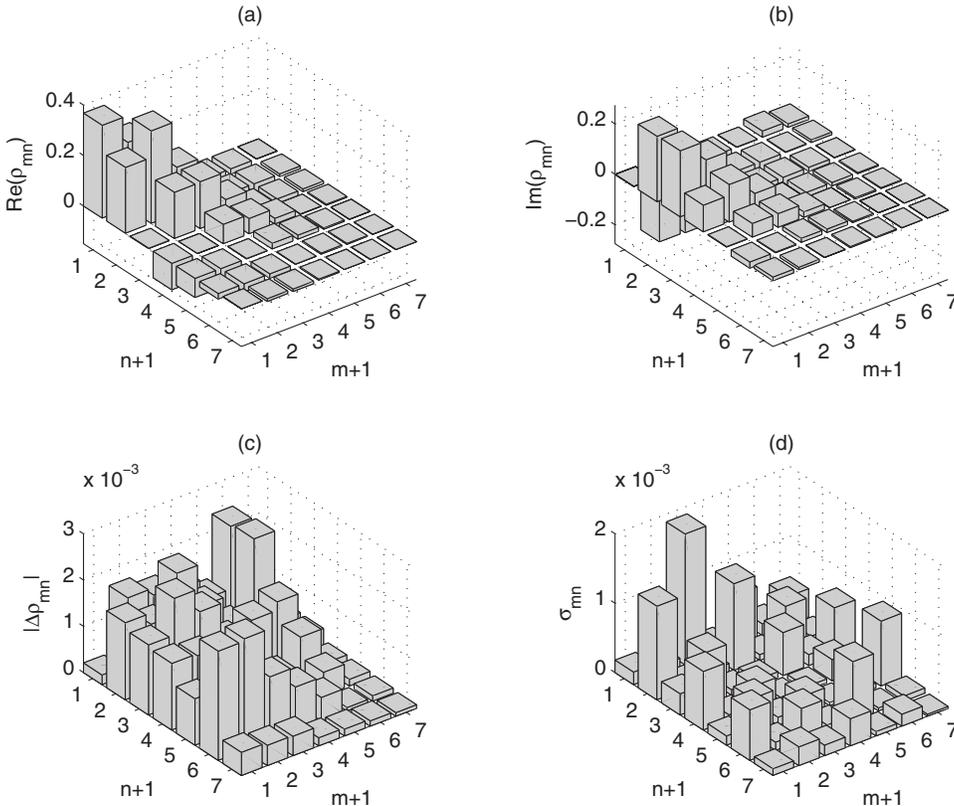


FIG. 3. Simulated homodyne tomography of a coherent state of amplitude $\alpha = \exp(i\pi/4)$; (a) real and (b) imaginary parts of the reconstructed density matrix; (c) absolute values of the differences between the reconstructed and the true density matrix elements; and (d) variances of real (imaginary) parts of the reconstructed density matrix elements in the region $m \leq n$ ($n > m$) as estimated via Eqs. (19) and (20). The reconstruction was done using $N_{it} = 10^3$ iterations of the ML iterative procedure based on Eq. (8) for homodyne measurements [Eq. (18)]. Amplitudes γ are taken as for Fig. 1(a); the number of experimental runs is $N_r = 10^2$; Fock space is truncated at $N_{tr} = 15$.

$$F(\text{Re}[\rho_{mn}]) = \sum_j \frac{j'}{p_j} \left[\frac{\partial}{\partial \text{Re}(\rho_{mn})} \frac{p_j}{\sum_{j'} p_{j'}} \right]^2 \quad (20)$$

and similarly for the imaginary part of ρ with Re replaced by Im.

Immediately a significant quantitative advantage of the ML algorithm over methods relying on a tomographically complete measurement can be seen. Indeed, for Fig. 3 we have used data collected in just $N_p=256$ points in the γ -plane. As confirmed by Fig. 3 it is possible to achieve very good accuracy of the reconstruction with the total number of measurements of the order of $N_{mes}=10^5$ although the corresponding operator \mathbf{G} is quite far from the identity operator, see Fig. 1(c).

V. ON/OFF DETECTION

Any realistic reconstruction scheme must be robust with respect to noise. In real experiments the presence of noise is unavoidable due to losses and because detectors are not ideal. It is well-known [2] that the standard tomography via the inverse Radon transformation fails when the detector's efficiency is below 50%. Although the same conclusion does not hold for the ML scheme developed above, the presence of losses is obviously detrimental to the accuracy of a reconstruction. However, the very presence of losses can be turned into an advantage and used for reconstruction purposes. As shown in Ref. [12], complete characterization of the input optical signal can be achieved using nonideal photon counters. Here it is enough to measure the probabilities of no-counts provided the quantum efficiency of such a detector can be controlled.

From the practical point of view such a possibility is very convenient. Practically any photodetector used in quantum optics can be implemented as a binary one, i.e., a detector discriminating between the presence and absence of the signal only. Apart from the avalanche photodiodes being the most natural detectors for the task, one can also use photomultiplier tubes, hybrid photodetectors, and cryogenic thermal detectors [13]. High quantum efficiencies are not needed for this purpose. Further, ideal 100% efficient detectors are of no use in such a reconstruction scheme. They would be sensitive just to the contribution of the vacuum component. The only requirement for a successful reconstruction is our ability to discriminate between the signal counts and spurious "dark" counts. When this is not possible, the noise coming from dark counts can still be incorporated into the reconstruction scheme, provided a proper physical model of the detection process is available [14].

The first reconstruction scheme based on binary detectors was developed and analyzed for inference of photon number distribution [3]. The scheme is extremely simple: just a set of binary detectors is used with different quantum efficiencies, which can be modeled by a single detector preceded by a beam splitter with a variable beam-splitting ratio to mix the input signal with the vacuum field. The scheme is fast, efficient, and robust. It can provide an adequate estimation of

diagonal elements of the signal's density matrix (photon number distribution) even when the set of different efficiencies used is smaller than the number of reconstructed parameters. The scheme also works for very low detector efficiencies of the order of a few percent [14]. In addition, this scheme has an important advantage compared to the homodyne tomography, which is the other important technique for determining the photon number distribution: it is not an interferometric technique, so there is no need to mode match the signal with the probe coherent field of the local oscillator. The scheme was realized experimentally in [15]. A more advanced setup based on a multichannel fiber loop detector was developed and experimentally verified earlier [16].

A generalization of the schemes exists that allows a reconstruction of the nondiagonal elements of the signal density matrix. It was noticed that diagonal elements of a coherently shifted signal contain full information about the signal density matrix. To extract this information it is necessary to perform shifts with appropriately chosen phases and amplitudes [17]. Such a scheme was exploited in proposals [18,19]. In the former scheme [18], the generalized distribution function may be recovered from the measured photocount statistics, whereas in the latter scheme [19] the Wigner function can be probed directly. This reconstruction technique has also been implemented experimentally [20]. It should be noted, however, that these schemes of quantum state reconstruction have their drawbacks. Since the schemes are based on linear inversion they would typically give non-physical results. This is due to the *a priori* constraints that any quantum state has to satisfy, namely the non-negativity of a density matrix $\rho \geq 0$, which is not guaranteed by the linear methods. This is also a drawback of the proposal representing the Wigner function of the signal state by projectors into the shifted Fock states. While it seems to be intractable to implement the condition of positive semidefiniteness in Wigner representation, it can be done in the general formalism adopting the maximum-likelihood estimation.

Reconstruction of a signal state from binary detections using the ML approach has been briefly discussed in Ref. [4]. The idea is to mix the signal not only with the vacuum field but also with a class of probe states. Coherent states, which can be produced easily in real experiments can be used for this purpose. In the following text the method is discussed in detail and its robustness with respect to the measurement errors is investigated.

The probability of registering no counts on a detector of quantum efficiency ν_c is given by the well-known Mandel formula [21]:

$$p = \langle : \exp\{-\nu_c \mathbf{c}^\dagger \mathbf{c}\} : \rangle, \quad (21)$$

where \mathbf{c}^\dagger and \mathbf{c} denote creation and annihilation operators of the output mode, and $::$ denotes normal ordering. For simplicity, it is assumed here that in the absence of the signal, the detector does not produce any clicks. It is also assumed that a beam splitter transforms input modes \mathbf{a} and \mathbf{b} in the following way:

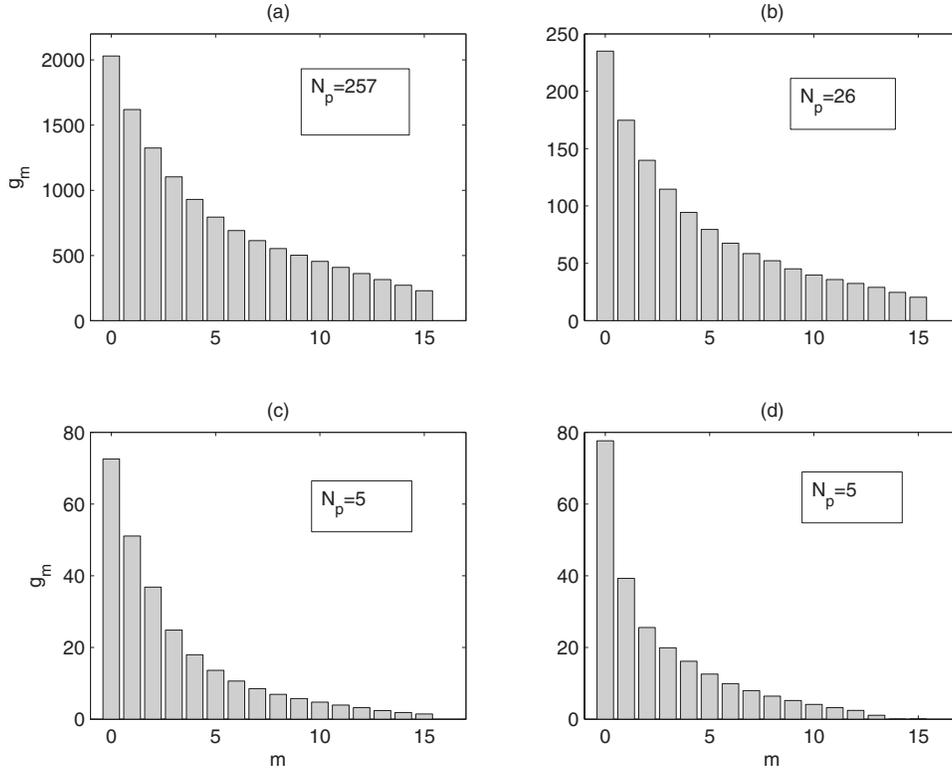


FIG. 4. Spectrum of \mathbf{G} operators corresponding to four different binary detection schemes. Probe amplitudes γ are as follows: (a) 16 equidistant amplitudes $\text{abs}(\gamma) \in [0.01, 3] \times 16$ equidistant phases $\text{arg}(\gamma) \in [0, 1.99\pi]$ plus the point in the center of coordinates; (b) the same as (a) but only 4 amplitudes $\times 4$ phases; (c) amplitudes $\text{abs}(\gamma) = \{0.01, 1\} \times \text{phases } \text{arg}(\gamma) = \{0, 0.1\}$ plus the point in the center of coordinates; and (d) amplitudes $\text{abs}(\gamma) = \{0.999, 1.001\} \times \text{phases } \text{arg}(\gamma) = \{0.1, 0.1001\}$ plus an additional point $\text{abs}(\gamma) = 1, \text{arg}(\gamma) = 0$. In all panels, 20 equidistant values of the detector efficiencies were chosen from the interval $\eta \in [0.1, 0.9]$.

$$\mathbf{c} = \mathbf{a} \cos(\alpha) + \mathbf{b} \sin(\alpha). \quad (22)$$

Theoretically, the choice of probes to make reconstruction through a set of recorded probabilities p is rather wide. A set of arbitrary states with nonzero off-diagonal elements in the Fock-state representation would be sufficient [3,12]. For the sake of feasibility we will limit our choice of probe states b to coherent states of amplitudes β . As it is shown in the Appendix, averaging over the coherent probe mode \mathbf{b} in Eq. (21) gives

$$p = \sum_{n=0} (1-\nu)^n \langle n | \mathbf{D}^\dagger(\gamma) \rho \mathbf{D}(\gamma) | n \rangle, \quad (23)$$

where

$$\nu = \nu_c \cos^2(\alpha), \quad \gamma = -\beta \tan(\alpha), \quad (24)$$

and

$$\mathbf{D}(\gamma) = \exp\{\gamma \mathbf{a}^\dagger - \gamma^* \mathbf{a}\}$$

is the coherent shift operator; $|n\rangle$ denotes a Fock state of signal mode \mathbf{a} . Using the operator notation

$$\mathbf{R}_{n,\gamma} = \mathbf{D}(\gamma) |n\rangle \langle n| \mathbf{D}^\dagger(\gamma), \quad (25)$$

we find that no-count probabilities (23) are generated by the following POVM elements:

$$\mathbf{A}_{\nu,\gamma} = \sum_n (1-\nu)^n \mathbf{R}_{n,\gamma} \quad (26)$$

and, defining a collective index $j = \{\nu, \gamma\}$, the counted probabilities coincide with those in Eq. (2).

Now let us go back to our main goal and illustrate the process of finding and optimizing the subspace available for

the reconstruction described in Sec. III. Figure 4 shows how a suitable choice of γ -points can be achieved. Obviously, the amount of data used in Fig. 4(a) as compared to Fig. 4(b) is excessive for the reconstruction. It would be enough to use a smaller number of different phase shifts. On the other hand, when the number of points is too small, or when they are chosen in an inappropriate way, the eigenvalues of \mathbf{G} differ strongly thus making reconstruction unreliable. We have already seen the analogous behavior of homodyne tomography in Sec. IV. For example, the last eigenvalue in Fig. 4(d) is only of the order of 10^{-3} .

So far we have been specifying the field of view of a given reconstruction technique by means of inspecting eigenvalues of operator \mathbf{G} (Step 2). In the next section we will move to Step 3 considering particular realizations of the reconstruction procedure described above.

VI. POINT-BY-POINT RECONSTRUCTION

Before implementing the complete reconstruction described in Sec. II, it is interesting to take a closer look at Eq. (23). Consider a set of measurements with different quantum efficiencies ν but with a fixed value of γ . Such measurements are compatible,

$$[\mathbf{A}_{\nu,\gamma} \mathbf{A}_{\mu,\gamma}] = 0.$$

This follows from Eq. (26) since quantum states $|n_\gamma\rangle = \mathbf{D}(\gamma) |n\rangle$ are obtained by a unitary transformation of the orthonormal Fock-state basis and as such they form an orthonormal basis,

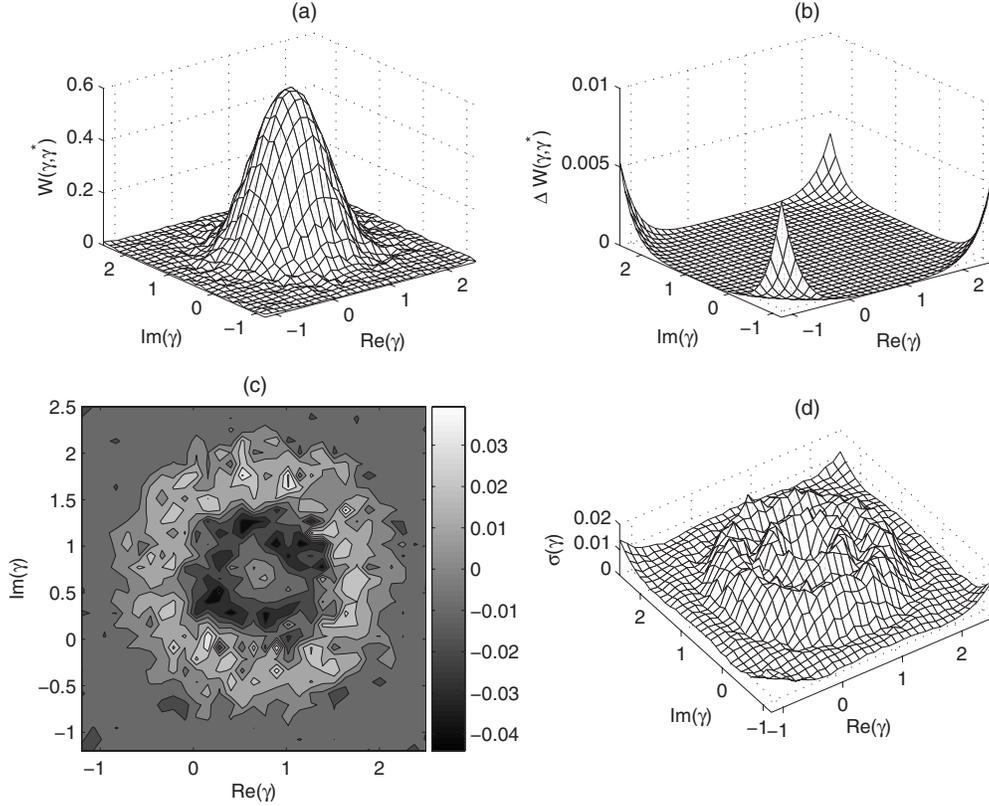


FIG. 5. Point-by-point reconstruction of a signal coherent state of amplitude $\alpha = \exp\{i\pi/4\}$; (a) the reconstructed Wigner function; (b) the difference between Wigner functions of the exact and truncated true states; (c) the difference between Wigner functions of the true and reconstructed states; and (d) the variances $\sigma(\gamma, \gamma^*)$. The Wigner function was reconstructed at $N_p = 2500$ points of the phase plane using $N_{it} = 10^3$ iterations of the EM algorithm per point; $N_r = 10^4$; $N_{tr} = 12$; 30 different values of detector efficiencies were used equidistantly distributed in the interval $[0.1, 0.9]$.

$$\langle n_\gamma | m_\gamma \rangle = \langle n | \mathbf{D}^\dagger(\gamma) \mathbf{D}(\gamma) | m \rangle = \delta_{nm},$$

where δ_{nm} is the Kronecker symbol. Hence all POVM elements $A_{\nu, \gamma}$ corresponding to the same value of γ take diagonal form in the representation $|n_\gamma\rangle$. Now let us introduce the diagonal elements of ρ in this basis,

$$R_n(\gamma) = \langle n | \mathbf{D}^\dagger(\gamma) \rho \mathbf{D}(\gamma) | n \rangle.$$

As follows from Eq. (23), measurements with a fixed value of γ yield information about the corresponding elements $R_n(\gamma)$. Reconstructing the values of $R_n(\gamma)$ from this data is thus formally equivalent to reconstructing the diagonal elements of ρ from the binary detections [3,14,22]. In this case, due to the mentioned mutual compatibility of the observations $A_{\nu_1, \gamma}, A_{\nu_2, \gamma}, \dots$, the algorithm (6) and (7) simplifies to

$$R_n^{(k+1)}(\gamma) = w_n^{(k)}(\gamma) \sum_{j=0}^{M-1} \frac{(1 - \nu_j)^n p_j^{\text{exp}}}{f_j p_j^{(k)}}, \quad (27)$$

where

$$f_j = \sum_{n=0}^{N-1} (1 - \nu_j)^n,$$

p_j^{exp} are the experimentally measured relative frequencies of detecting no click on the detector, and $p_j^{(k)}$ is the left-hand

side of Eq. (23) calculated using the result of the k th iteration. Procedure (27) guarantees non-negativity of the reconstructed $R_n(\gamma)$. To ensure a unit sum of the reconstructed values of $R_n(\gamma)$, one should either renormalize them or multiply the right-hand part of Eq. (27) by $\sum_j p_j^{(k)} / \sum_k p_k^{\text{exp}}$.

Having found $R_n(\gamma)$ a value of the signal Wigner function at point γ can be calculated [17]:

$$W(\gamma^*, \gamma) = \frac{2}{\pi} \sum_{n=0}^{N-1} (-1)^n R_n(\gamma). \quad (28)$$

The algorithm (27) together with the formula (28) gives a practically realizable and seemingly simple way of doing the quantum state tomography based on binary detections.

In Fig. 5 one can see an example of a reconstruction of a coherent signal state with the help of the procedure described above. For this particular case the procedure turns out to be fast and accurate. Accurate reconstruction was achieved with only $N_{it} = 10^3$ iterations of the EM algorithm (27) for $N_r = 10^4$ experimental runs at each point in the phase plane. This impression is further strengthened by looking at how the total error of the reconstruction propagates. The average distance between the exact and reconstructed Wigner functions

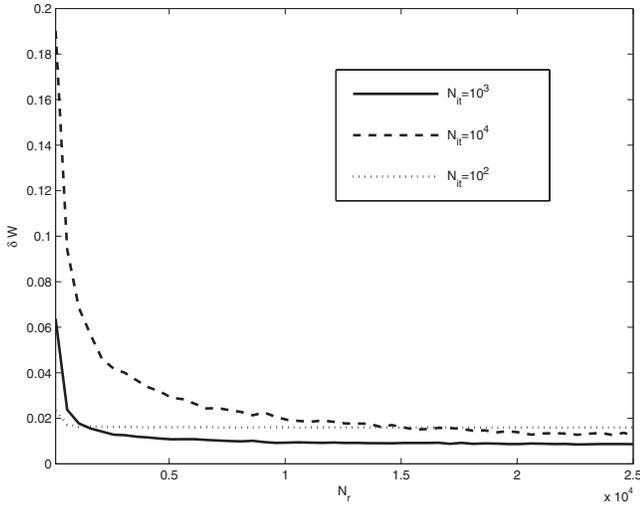


FIG. 6. Reconstruction errors δW defined by Eq. (29) for different numbers of iterations N_{it} of the ML algorithm in dependence on the number of experiment runs N_r ; for all curves $N_p=2500$ and the following region on the phase plane was taken: $\text{Re}(\gamma), \text{Im}(\gamma) \in [-1.2, 2.5]$. Other parameters are as in Fig. 5.

$$\delta W = \frac{1}{N_p} \sum_{\forall \gamma} |W_{\text{exact}}(\gamma^*, \gamma) - W(\gamma^*, \gamma)| \quad (29)$$

is used as a measure of error, where the summation is taken over all points on the phase plane where the estimation was made; N_p stands for the number of points on the phase plane. Figure 6 shows that for a given number of performed iterations N_{it} , the accuracy of the reconstruction improves with the number of runs N_r . This is expected: longer measuring times translate to less statistical noise and more accurate results. However, the improvement is negligible at larger values of N_r . This is a sign of robustness. When the iterations are continued, the dependence of the residual error on N_r becomes more pronounced. Notice also that increasing the number of iterations above a certain limit, $N_{it} \gg N_r$, might even increase the error δW . This has already been observed in [14] in connection with a photon number reconstruction. Naturally, this does not mean that the algorithm (27) does not converge but rather occurs because the error measure (29) and likelihood are not monotonic functions of each other in the neighborhood of the ML state. It can be concluded that in practice it is reasonable to stop the iterative reconstruction procedure after making $N_{it} \approx N_r$ iterations.

So, one can see that a point-by-point reconstruction of the Wigner function is feasible, robust, and a simple scheme. However, a closer look reveals that this simplicity is rather deceiving. In fact, the scheme is far from being optimal. Above we have been estimating a continuous quantity, namely the Wigner function. In the recommended approach outlined in Sec. III the same measurement may be adopted for the estimation of the full density matrix using an appropriately chosen discrete representation. It is confirmed in the next section that this is indeed a more efficient approach.

There are also other problems with implementing the procedure (27). First of all one needs to specify the dimension

N_{tr} at which the Hilbert space is truncated. Then, one has to choose some initial values of $R_n^{(0)}(\gamma)$. Once N_{tr} is chosen, a choice of the initial distribution $R_n^{(0)}(\gamma)$ is not a problem. It is noted in the work [14], the choice of $R_n^{(0)}(\gamma) \neq 0$ has only a weak effect on the convergence for any fixed γ . In particular, one might take the uniform distribution as the initial one, as was done in our example of Fig. 5. However, an appropriate choice of N_{tr} is far from being trivial. Actually, one needs to investigate the convergence for every point γ on the phase plane. Additionally for any N_{tr} one can find regions on the phase plane where this very N would be insufficient for the accurate reconstruction because the average number of photons of the shifted state can be arbitrarily large for large γ 's,

$$\sum_n n R_n(\gamma) |_{|\gamma| \rightarrow \infty} \rightarrow \infty.$$

This is clearly seen in Fig. 5(b). Such truncation errors further increase the errors of the reconstruction procedure, see Fig. 5(d). Moreover, for different points on the phase plane the rate of convergence of the sum (28) might differ strongly. In regions where the Wigner function changes steeply, more iterations and measurements are needed to achieve the same precision. For instance, the variances depicted in Fig. 5(d) are smaller near the peak of $W(\gamma, \gamma^*)$. An explanation can be found in formula (28): in the region of such a steep change one needs to find several comparable $R_n(\gamma)$ with high precision, whereas the behavior of the Wigner function near $\gamma = \alpha$ is defined mostly by $R_0(\gamma)$.

One should also keep in mind that even very small deviations from the true Wigner function, which will inevitably arise in the process of reconstruction, will likely make it nonphysical. Such a Wigner function would not correspond to any physical, positive definite density matrix. This is because operators $\mathbf{R}_{n,\gamma}$ do not commute for different γ 's, so noisy measurements may give inconsistent results. This is confirmed by the inset of Fig. 7—going back from the Wigner function to density matrix using Glauber's formula [23]

$$\rho = 2 \int d^2 \gamma (-1)^n W(\gamma^*, \gamma) D(2\gamma), \quad (30)$$

some diagonal elements of the reconstructed matrix are found to be slightly negative despite the fact that $R_n(0)$ estimated via the formula (27) is strictly non-negative.

VII. ALL-POINTS RECONSTRUCTION SCHEME

Here we consider an optimal reconstruction in which all collected data are used for the reconstruction of the signal density matrix according to the recipe given in Sec. III. We refer to it as an “all-points” reconstruction scheme to distinguish it from the “point-by-point” scheme discussed in the preceding section.

According to the prescription of Sec. III, after choosing a measurement one should also determine what this measurement can reveal. In all-points reconstruction the question of a proper choice of the reconstruction subspace can be answered in quite a simple way. Indeed, let us work out a

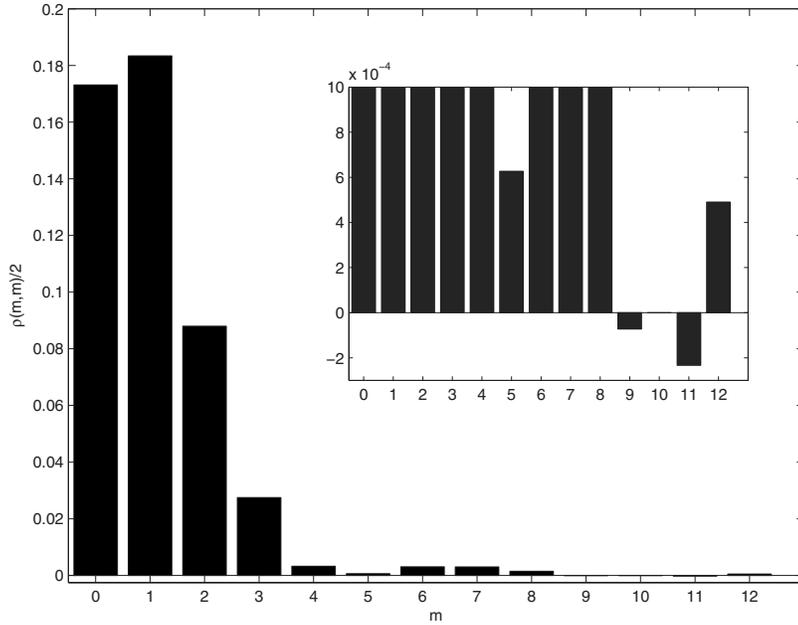


FIG. 7. The diagonal elements of the reconstructed density matrix (30); parameters are as in Fig. 5.

particular example of the measurements choice, illustrated in Fig. 8. Already for a modest number of different γ 's and detector efficiencies, the eigenvectors of the closure operator \mathbf{G} are very close to Fock states. So, by inspecting the eigenvalues of \mathbf{G} one can easily estimate the largest subspace of the Fock space visible to the chosen measurement scheme. An obvious way to do that is to first perform measurements for different detector efficiencies and keep $\gamma=0$. As we have demonstrated before, such measurements are sufficient for reconstructing the diagonal elements of the signal density matrix. The smallest truncation number N_{tr} sufficient to confine the signal can be found, e.g., by checking the conver-

gence with respect to increasing N_{tr} , see also [14]. Then, if the eigenvalues of the operator \mathbf{G} in the Fock subspace $n \leq N_{tr}$ are comparable, its eigenvectors being close to Fock states, one can conclude that the observed data will be sufficient to yield a faithful reconstruction.

It is interesting to note that coherent shifts depicted in Fig. 8 correspond to coherent states with the average number of photons less than unity. In other words, one does not need to implement large coherent shifts in order to reconstruct even signal density components corresponding to Fock states with a high number of photons $n \gg 1$. This observation is important for optimizing the measurement procedure. "Intuition"

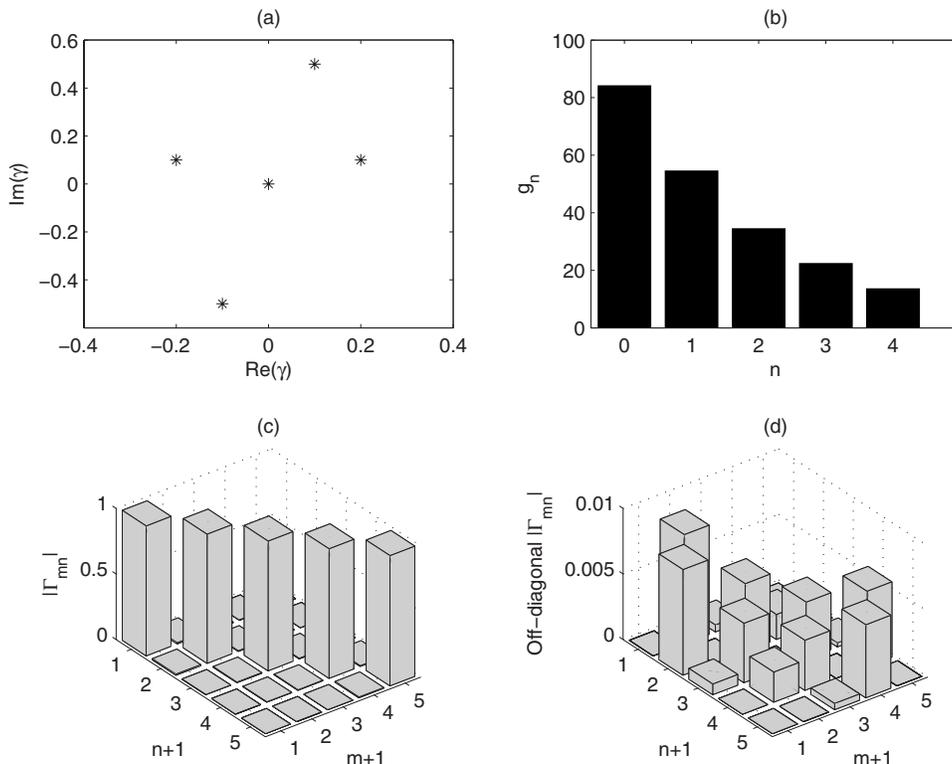


FIG. 8. Operator \mathbf{G} in all-points scheme; (a) γ points on the phase plane taken for the reconstruction; (b) spectrum of \mathbf{G} ; (c, d) absolute values of its eigenvector elements, the diagonal elements are zeroed in (d) for better clarity. Fock-state is truncated at $N_{tr}=15$ photons; m and n label the eigenvectors and their components, respectively. The following five different measurements are used: $\text{Re}(\gamma)=-0.2, -0.1, 0, 0.1, 0.2$; $\text{Im}(\gamma)=0.1, -0.5, 0, 0.5, 0.1$; 20 equidistantly distributed detector efficiencies in the interval $[0.1, 0.9]$ are taken.

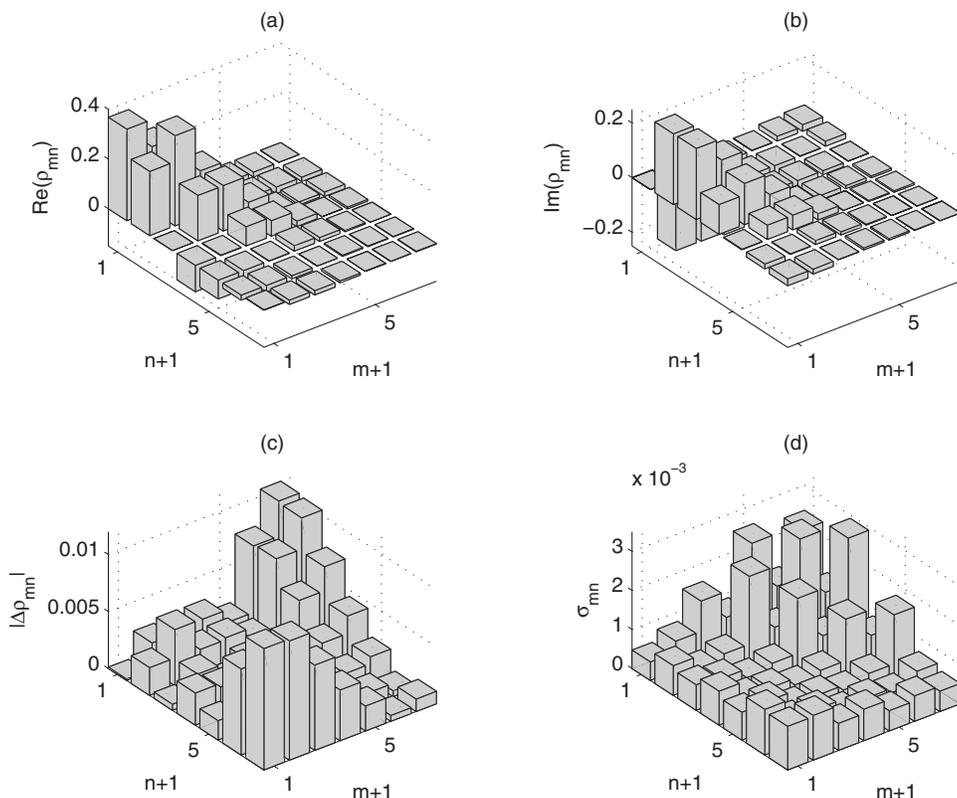


FIG. 9. All-points reconstruction of a coherent state of amplitude $\alpha = \exp\{i\pi/4\}$; (a) real part and (b) imaginary part of the reconstructed density matrix; (c) absolute values of the difference between the reconstructed and true density matrices; and (d) the corresponding variances of real (imaginary) parts of the reconstructed density matrix elements are shown in regions $m \leq n$ ($n > m$) as estimated via Eq. (20). Probe γ -points are taken as for Fig. 4(b); $N_{tr} = 15$; $N_{ii} = 10^3$, and $N_r = 10^4$.

would suggest the use of probe states with a mean number of photons comparable to the expected mean number of photons in the signal mode. Then one would have to use a rather large cutoff N_{tr} to contain such highly populated states. This is not necessary. The all-points scheme does not require strong perturbations of the signal via probe states with large photon content.

To illustrate this, let us consider a reconstruction of the coherent state with the amplitude $\exp\{i\pi/4\}$ taken as an example in the previous section, which is summarized in Fig. 9. One can see that with a modest number of measurements and iterations of the ML algorithm good results of the reconstruction procedure can be obtained. The all-points reconstruction scheme is far more economical than the point-by-point scheme considered previously, one needs much less variation of the probe field to make the reconstruction. Furthermore, since it is not necessary to use probe states with large mean numbers of photons, the truncation error is significantly smaller. Figure 10 shows that the number of iterations of the ML algorithm required to achieve the optimal state is rather small. Here fidelity defined by

$$f = \text{Tr}\{\rho_{rec}\rho\} \quad (31)$$

was used as a measure of the reconstruction errors. It quantifies a distance of the true pure state ρ from the reconstructed one ρ_{rec} . Notice approximately the same number of iterations is needed for reaching optimal fidelity no matter what is the size of the processed data. The fidelity practically does not improve with further iterations. It may also be noticed that the maximal fidelity obtained scales with the number of experimental runs as $(1-f) \propto 1/\sqrt{N_r}$. In this simula-

tion, an $N_{tr} \times N_{tr}$ identity matrix has been used as the initial guess. The choice of the starting point for the EM algorithm has little effect on its convergence as long as we choose it to be strictly positive. On the contrary, for different signal states the convergence of the ML algorithm can be very different. Let us consider, for example, the following superposition of the vacuum and two-photon states:

$$|\phi\rangle = (|0\rangle + \exp\{0.5i\}|2\rangle)/\sqrt{2}. \quad (32)$$

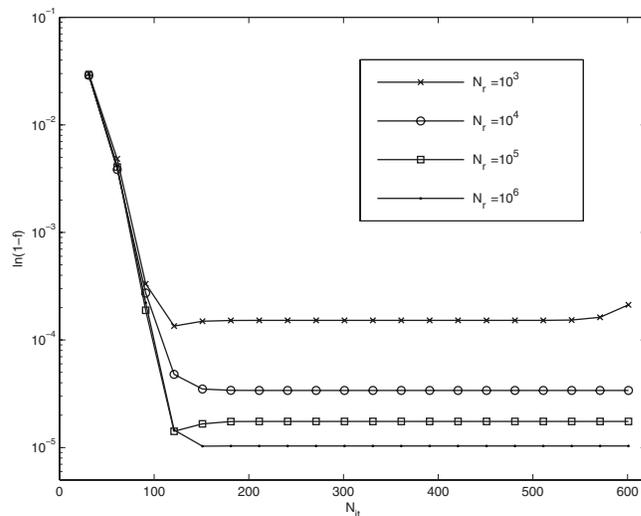


FIG. 10. Fidelities of the reconstructed coherent state of amplitude $\alpha = \exp\{i\pi/4\}$ in dependence on the number of iterations of the ML algorithm for several different numbers of experimental runs N_r . Other parameters are as for Fig. 9. Fidelity f is defined by Eq. (31).

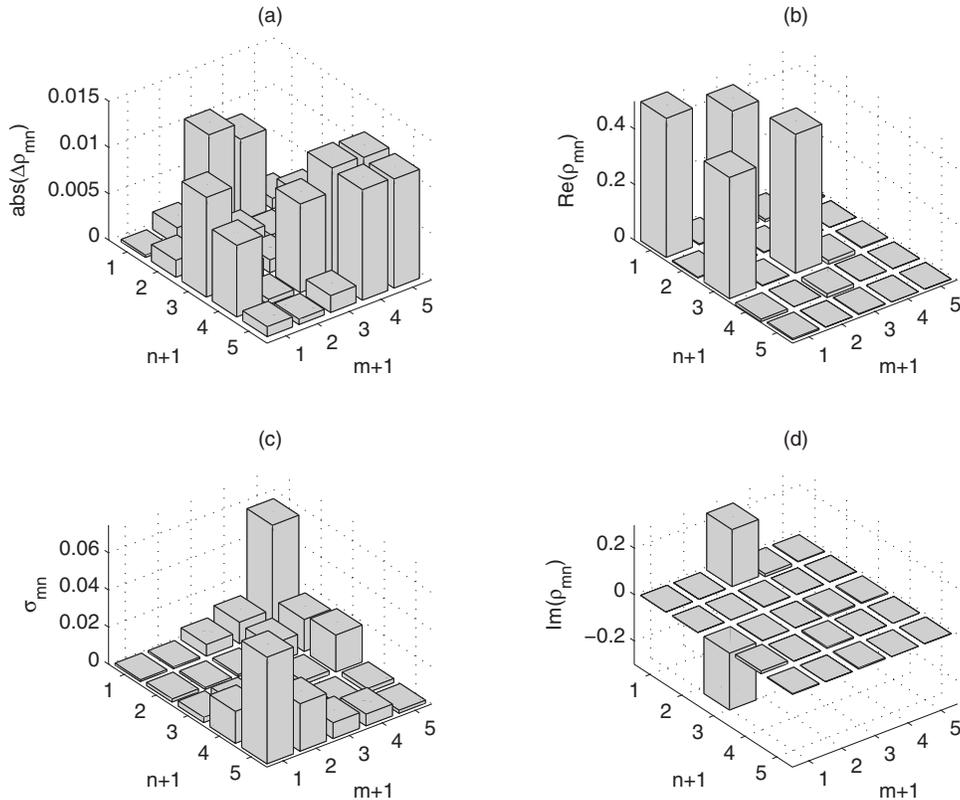


FIG. 11. All-points reconstruction of the state Eq. (32); (a) absolute values of the differences between the exact and reconstructed density matrices; (b) real and (d) imaginary parts of the reconstructed matrix obtained using 10^6 iteration of the ML algorithm; and (c) the variances of the real (imaginary) parts of the reconstructed density matrix are shown in regions $n \leq m$ ($n > m$) as given by Eq. (20). The measurements are that of Fig. 8; $N_r = 10^7$.

The results of the all-points reconstruction procedure can be seen in Fig. 11. We have used only five different points on the phase plane depicted in Fig. 8(a), and $N_r = 10^7$ simulated experimental runs. Therefore the number of the probe field amplitudes needed to get a tomographically complete set of measurements is quite small, too. We can safely limit ourselves to a modest number of Fock-state vectors spanning the reconstruction subspace. In spite of the lower dimensionality of this problem compared to the previous example of the signal coherent state, finding the ML solution numerically appears to be much more difficult. To reach an acceptable accuracy of the reconstruction ($f > 99\%$) in this case, one has to make more than 10^5 iterations.

VIII. CONCLUSION

In this paper we have developed quantum tomography as an objective approach defining a relation between a set of measurements chosen for the reconstruction and a possible result of the reconstruction procedure. It has been demonstrated that choosing *a priori* the measurement set one might severely limit the space available for the reconstruction, or might increase unnecessarily the data acquisition time needed for getting a faithful reconstruction. With help of the ML reconstruction procedure we demonstrated how the subspace available for the reconstruction could be found; moreover, we suggested a way to build an orthogonal basis in such a subspace. This method was demonstrated first on a simple example of a quantum optical tomography scheme.

After that attention was paid to a reconstruction scheme based on binary detectors capable of only distinguishing be-

tween the presence and absence of photons. We proved here that a complete reconstruction of the quantum state with such detectors could be made. The general objective ML estimation was adapted to such a tomography scheme. Two possible scenarios of reconstruction using binary detectors were considered: the point-by-point and the all-points schemes. The former allows the finding of a value of a signal state Wigner function at any point on the phase plane. While being very simple and methodologically similar to the reconstruction of the photon number [3], it was shown to be far from being optimal. Moreover, in the presence of noise it may often fail to provide a physically sound result, e.g., by yielding a nonpositive “density matrix.” The all-points reconstruction method is free from such shortcomings. The implementation of the complete formalism developed in the paper for the all-points scheme was given: first we found the subspace required for a successful reconstruction and determined the appropriate amplitudes of the probe field and built an orthogonal basis in the reconstruction subspace. Subsequently we used the iterative ML algorithm and estimated errors with the help of the Fisher information matrix. The suggested scheme is simple, robust, and sufficiently effective to become a successful competitor to the conventional linear quantum tomography schemes such as the inverse Radon transformation.

ACKNOWLEDGMENTS

We would like to thank A.I. Lvovsky and M.G.A. Paris for helpful discussions. The authors acknowledge the support from Research Project MSM 6198959213 of the Czech Min-

istry of Education, EU project COVAQIAL FP6-511004, Grant No. 202/06/0307 of the Czech Grant Agency (J.R. and Z.H.), and from BRFFI of Belarus and CNPq of Brazil (D.M).

APPENDIX

The derivation of the detection formula, Eq. (23), is given here for the sake of completeness. We start by expanding the normally ordered exponential operator in the quantum averaging, Eq. (21), into a power series,

$$p = \text{Tr}_{a,b} \left\{ (\rho_a \otimes |\beta\rangle_b \langle\beta|) \sum_l \frac{\mathbf{c}^{\dagger l} \mathbf{c}^l}{l!} (-\nu_c)^l \right\}, \quad (\text{A1})$$

where the lower indexes a , b , and c denote the respective modes, ρ_a is the input state of mode a and $|\beta\rangle_b$ is the coherent probe state of amplitude β . Expressing the output variable \mathbf{c} in terms of the input variables \mathbf{a} and \mathbf{b} using the input-output relation Eq. (22), tracing over mode b , and introducing the parameters (24), we get

$$p = \sum_l \frac{(-\bar{\nu})^l}{l!} \text{Tr} \{ \rho (\mathbf{a}^\dagger - \gamma^*)^l (\mathbf{a} - \gamma)^l \}, \quad (\text{A2})$$

where the mode index a has been dropped. Recalling the basic property of the coherent shift operator, $\mathbf{D}(\alpha) \mathbf{a} \mathbf{D}^\dagger(\alpha) = \mathbf{a} - \alpha$, and using the cyclic invariance of the trace operation, this expression further simplifies to

$$p = \sum_l \frac{(-\bar{\nu})^l}{l!} \text{Tr} \{ \mathbf{D}^\dagger(\gamma) \rho \mathbf{D}(\gamma) \mathbf{a}^{\dagger l} \mathbf{a}^l \}. \quad (\text{A3})$$

The indicated tracing over mode a is most conveniently done in the computational basis of Fock states $|n\rangle$. Since $\mathbf{a}^l |n\rangle = 0$ when $l > n$, the summation over l stops at n ,

$$p = \sum_{n=0}^{\infty} \sum_{l=0}^n \frac{n!}{l!(n-l)!} (-\bar{\nu})^l \langle n | \mathbf{D}^\dagger(\gamma) \rho \mathbf{D}(\gamma) | n \rangle. \quad (\text{A4})$$

Recognizing the binomial factor, the detection formula, Eq. (23), is finally arrived at.

-
- [1] K. Vogel and H. Risken, *Phys. Rev. A* **40**, 2847 (1989).
 [2] M. Raymer and M. Beck, in *Quantum States Estimation*, edited by M. G. A. Paris and J. Rehacek, Lecture Notes in Physics Vol. 649 (Springer, Heidelberg, 2004).
 [3] D. Mogilevtsev, *Opt. Commun.* **156**, 307 (1998); D. Mogilevtsev, *Acta Phys. Slov.* **49**, 743 (1999).
 [4] Z. Hradil, D. Mogilevtsev, and J. Rehacek, *Phys. Rev. Lett.* **96**, 230401 (2006).
 [5] R. Barlow, *Statistics* (Wiley, New York, 1989), p. 90.
 [6] A. P. Dempster, N. M. Laird, and D. B. Rubin, *J. R. Stat. Soc. Ser. B (Methodol.)* **39**, 1 (1977); R. A. Boyles, *ibid.* **45**, 47 (1983); Y. Vardi and D. Lee, *ibid.* **55**, 569 (1993).
 [7] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Phys. Rev. A* **64**, 052312 (2001).
 [8] G. Toraldo di Francia, *J. Opt. Soc. Am.* **59**, 799 (1969).
 [9] A. I. Lvovsky, *J. Opt. B: Quantum Semiclassical Opt.* **6**, S556 (2004); A. I. Lvovsky and M. G. Raymer, e-print arXiv:quant-ph/05110444 (unpublished).
 [10] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier, *Phys. Rev. Lett.* **96**, 213601 (2006); *Science* **312**, 83 (2006); J. S. Neergaard-Nielsen, B. M. Nielsen, C. Hettich, K. Moelmer, E. S. Polzik, *Phys. Rev. Lett.* **97**, 083604 (2006); S. A. Babichev, B. Brezger, and A. I. Lvovsky, *ibid.* **92**, 047903 (2004); S. A. Babichev, J. Appel, and A. I. Lvovsky, *ibid.* **92**, 193601 (2004).
 [11] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).
 [12] D. Mogilevtsev, Z. Hradil, and J. Perina, *Quantum Semiclass. Opt.* **10**, 345 (1998).
 [13] E. Hergert, *Single Photon Detector Workshop* (NIST, Gaithersburg, 2003).
 [14] A. R. Rossi, S. Olivares, and M. G. A. Paris, *Phys. Rev. A* **70**, 055801 (2005); A. R. Rossi and M. G. A. Paris, *Eur. Phys. J. D* **32**, 223 (2005).
 [15] G. Zambra, A. Andreoni, M. Bondani, M. Gramegna, M. Genovese, G. Brida, A. Rossi, and M. G. A. Paris, *Phys. Rev. Lett.* **95**, 063602 (2005).
 [16] J. Rehacek, Z. Hradil, O. Haderka, J. Perina, Jr., and M. Hamar, *Phys. Rev. A* **67**, 061801(R) (2003); O. Haderka, M. Hamar, and J. Perina, Jr., *Eur. Phys. J. D* **28**, 149 (2004).
 [17] A. Royer, *Phys. Rev. Lett.* **55**, 2745 (1985); H. Moya-Cessa and P. L. Knight, *Phys. Rev. A* **48**, 2479 (1993).
 [18] S. Wallentowitz and W. Vogel, *Phys. Rev. A* **53**, 4528 (1996).
 [19] K. Banaszek and K. Wodkiewicz, *Phys. Rev. Lett.* **76**, 4344 (1996).
 [20] K. Banaszek, C. Radzewicz, K. Wodkiewicz, and J. S. Krasinski, *Phys. Rev. A* **60**, 674 (1999).
 [21] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, England, 1995).
 [22] K. Banaszek, *Phys. Rev. A* **59**, 4797 (1999).
 [23] K. E. Cahill and R. J. Glauber, *Phys. Rev.* **177**, 1857 (1969); **177**, 1882 (1969).