Tomography for quantum diagnostics

J Řeháček\textsuperscript{1,3}, D Mogilevtsev\textsuperscript{2} and Z Hradil\textsuperscript{1}

\textsuperscript{1} Department of Optics, Palacky University, 17. listopadu 50, 772 00 Olomouc, Czech Republic
\textsuperscript{2} Institute of Physics, Belarus National Academy of Sciences, F Skarina Ave 68, Minsk 220072, Belarus

E-mail: rehacek@phoenix.inf.upol.cz

New Journal of Physics \textbf{10} (2008) 043022 (12pp)

Received 8 January 2008
Published 14 April 2008

Online at \url{http://www.njp.org/}
doi:10.1088/1367-2630/10/4/043022

Abstract. The quantification of relevant statistical errors is an indispensable but often neglected part of any tomographic scheme used for quantum diagnostic purposes. We introduce a novel resolution measure, which provides ‘error bars’ for any inferred quantity of interest. This is illustrated with an example of the diagnostics of non-classical states based on the value of the reconstructed Wigner function at the origin of the phase space. We show that such diagnostics is meaningful only when a lot of prior information on the measured quantum state is available. Our resolution measure also provides an effective tool for optimization and resolution tuning of tomography schemes.

Quantum information technologies have recorded enormous progress in the last fifteen years. They have developed from the early stage of thought experiments into almost ready-to-use technology today. In view of the many possible applications the question of efficient analysis and diagnostics of quantum systems appears to be crucial. The quantum state is not observable and as such it cannot be measured in the traditional sense of the word. Information encoded in a quantum state may be portrayed by various ways yielding the most complete and detailed picture of the quantum object available. Due to the formal similarities between the quantum estimation and medical non-invasive three-dimensional (3D) imaging, this method is also called quantum tomography.

Almost 20 years ago, Vogel and Risken [1] demonstrated theoretically that the quantum state of light could be reconstructed from homodyne measurement data. The experimental realization followed soon after [2]. Later on, many different methods of quantum tomography were proposed and implemented for various physical systems [3]. Experiments are being permanently improved in order to increase our ability to unravel even the most exquisite and
fragile non-classical effects [4, 5]. Progress has been made not only on the detection side of tomography schemes. Mathematical algorithms have also been improved. The original linear methods [6] based on the inverse Radon transformation are prone to producing artifacts and have other serious drawbacks. For example, the positivity of the reconstructed state required by quantum theory is not guaranteed. This may obviously lead to inconsistent statistical predictions about future events. For such reasons, the simple linear methods are gradually being replaced by statistically motivated methods, for example, by Bayesian [7]–[11] or maximum-likelihood (ML) [12]–[14] tomography methods.

A proper and correct statistical treatment of the data forms an equally important part of the quantum tomography scheme as the detection does, since both these issues are irreducibly intertwined. The following example demonstrates how subtle and counter-intuitive the realm of quantum considerations can be. Fidelity defined for an estimate \( \rho \) of the (pure) true state \( |\psi\rangle \) by their overlap, \( f = \langle \psi | \rho | \psi \rangle \), is, for its simplicity, often used for the quantification of how successful the reconstruction has been. While in most cases fidelity is an adequate measure of ‘similarity’, it may fail in a specific context. Indeed, the fidelity of two coherent states with amplitudes 1 and 0.9 is better than 99%, yet their energies differ by almost 20%. This paradox is a consequence of reducing the complicated mutual relationship of two quantum objects to just a single number. Going back to ML tomography, one cannot believe that the most likely state alone is able to characterize all aspects of the measured signal. Proper statistical interpretation is of paramount importance especially in the realm of quantum world, where the registered data are the only source of information. Prior to measurement the true state is not known and an independent check of the reconstructed state is usually not possible. In this sense it is highly important to tell which observed effects are real and what should be attributed to the reconstruction noise.

The question, to which extent one can trust the results of quantum tomography, is central to this paper. Firstly, we will show how the errors of the reconstruction schemes can be estimated together with the quantum states. Although some steps in this direction have already been done, we will introduce, for the first time, a simple and operational measure of errors implemented in the context of objective tomography [15]. Several alternative and complementary approaches have recently appeared to address the reliability and resource analysis of quantum-state [14, 16, 17] and quantum-process [18] tomographies. Though the statistical arguments are very similar, the methods differ in practical conclusions. From the Bayesian point of view a quantum state is considered as a multi-dimensional random variable whose posterior distribution is given by the Bayes’ rule. For example, in [11] a ‘Keeping the experimentalist honest’ strategy is elaborated which turns out to be a version of the likelihood principle, namely the (quantum) relative entropy. Such an approach, however, does not yield a simple and tractable recipe that an experimenter could use to handle the intrinsic uncertainties involved in quantum tomography. The formulation of such a recipe is the main goal of the present paper. Adopting the ML approach, we seek for the state, which appears to be most likely from the point of view of the recorded data. The most likely state is found by solving an operator nonlinear equation which is guaranteed to yield an estimate consistent with all the requirements of the quantum theory. Since the most likely solution itself cannot grasp all the complexity of quantum tomography, the width of the likelihood function reflecting the tomographic data will be adopted for further characterization of variance of any inferred variable. In order to demonstrate the utility of the method we will consider explicitly the uncertainty of the value of the Wigner function at the origin of the phase space, whose negativity heralds the nonclassical character of the...
reconstructed state. Finally, we will show that when a thorough analysis of realistic experimental tomography schemes is done along our recipe, some earlier claims of obtaining successful reconstructions of nonclassical states by means of homodyne tomography may appear to be over-optimistic interpretations of the measured data.

Let us consider an ensemble of $N$ quantum systems subject to a generic measurement with $S$ different outcomes/channels, which is described by positive-operator-valued measure (POVM) elements $A_i \geq 0$, $i = 1, \ldots, S$, such that the probability $p_i$ of detecting the $i$th channel is given by $p_i = \text{Tr}(\rho A_i)$. Note that the set of operators $A_i$ may but need not form a POVM in the usual sense: $\sum_i A_i = 1$. In the latter case one may assume that the set $A_i$ is a subset of a POVM whose one or more channels are ignored. Due to the finite size $N$ of the measured ensemble, the observed relative frequency $N_i/N$ of the $i$th outcome will be different from the corresponding theoretical probability $p_i$. In the following we assume that all measurements are ideal and that the above-mentioned statistical noise is the only source of measurement errors.

The likelihood $\mathcal{L}$ of a state $\rho$ is the probability $p(\{N_i\} | \rho)$ that our measurements on state $\rho$ would generate given data $\{N_i\}$. For independent channels, $\mathcal{L}$ is proportional to the product of mutually normalized probabilities $p_i = \text{Tr}(\rho A_i)$ of the corresponding channels [15]:

$$\log \mathcal{L} = \sum_{i=1}^{S} N_i \log(\frac{p_i}{\sum_j p_j}).$$

The goal of quantum tomography is to reconstruct the true unknown state $\rho$ from registered data $\{N_i\}$. Many different approaches to this problem have been proposed. In the asymptotic regime of many detections per channel $N_i \gg 1$ the ML tomography proposed in [13] and adopted to various quantum systems in [4, 5, 14], [19]–[22] becomes the most efficient tomography technique. ML tomography looks for the state $\rho^{\text{ML}}$ for which the likelihood $\mathcal{L}$ is maximized. According to the theory presented in [15] the Hilbert subspace for the reconstruction is delimited by the dominant eigenvalues of the operator $G = \sum_i A_i$. This subspace is sampled more densely and hence is more visible than a subspace where the eigenvalues of $G$ are very small.

However, the result of the quantum tomography cannot be reduced merely to finding the most likely state. What also matters is how much the other states, those being less likely ones, would be consistent with the registered data. In this sense, states lying in the neighborhood of the most likely state should also be taken into account for making future statistical predictions.

Let us adopt a fixed Hermitian operator basis $\Omega_i$, $i = 0, \ldots, M$, such that $\Omega_0$ is a unity operator, $\text{Tr}(\Omega_i) = 0$, $i > 1$ and $\text{Tr}(\Omega_i \Omega_j) = \delta_{ij}$, where $M = d^2 - 1$ denotes the number of independent real parameters of $d$-dimensional density matrix. Using this representation, the ML estimate $\rho^{\text{ML}}$ and a generic quantum state $\rho$ read $\rho^{\text{ML}} = \Omega_0 / d + \sum_i \rho_i^{\text{ML}} \Omega_i$, and $\rho = \Omega_0 / d + \sum_i \rho_i \Omega_i$, respectively. By denoting $r_i = \rho_i - \rho_i^{\text{ML}}$ and $i = 1, \ldots, M$ and expanding the log-likelihood around the ML solution one finds that the posterior distribution of quantum states close to ML solution can be approximated by a multi-mode normal distribution (see the appendix),

$$P_\rho(r) = (2\pi)^{-M/2} (\text{det} F)^{1/2} \exp \left\{ -\frac{1}{2} r \cdot F r \right\},$$

(1)

$F$ being the Fisher information matrix,

$$F_{jk} = N^2 \sum_i \frac{1}{N_i} \frac{\partial}{\partial r_j} \left[ \frac{p_i}{P} \right] \frac{\partial}{\partial r_k} \left[ \frac{p_i}{P} \right],$$

and $P = \sum_i p_i$. In this sense the ML reconstruction is given by a family of states described by the posterior distribution. Any prediction based on the result of reconstruction should encompass this additional source of uncertainty.
Let us consider a prediction of the expectation value $z = \text{Tr}[\rho^\text{ML}Z]$ of an observable $Z = \sum_n z_n \Omega_n$ based on the tomographic reconstruction. For example, the expectation value of the (pure) true state $Z = |\psi_\text{true}\rangle \langle \psi_\text{true}|$ yields the fidelity of the reconstruction, and the parity operator having Fock representation $Z = \sum_{n=0}^\infty (-1)^n |n\rangle \langle n|$ provides the value $W(0,0)$ of the Wigner function [6] at the origin of the phase space, to give just two examples of practical importance. Fluctuations of the reconstructed state, which is now considered to be a random variable distributed according to the posterior distribution equation (1), are reflected in the fluctuations of the predicted expectation values of $Z$. They are characterized by their mean value and variance

$$
\langle z \rangle = \text{Tr}[\rho^\text{ML}Z], \quad \langle (\Delta z)^2 \rangle = z \cdot F^{-1}z.
$$

Using this formula one can place ‘error bars’ on any quantity inferred from the reconstructed state; this is the central result of our paper. Let us stress that $\Delta z$ is not only a quantitative measure of our uncertainty about the observable $Z$ but often its knowledge is crucial for the correct interpretation of results. In experiments done so far, only the mean values were adopted for diagnostic purposes. For example, a negative value of the reconstructed Wigner function at the origin was taken as a sign of nonclassical behavior. Such a statement is, however, valid only for diagnostic purposes. For example, a negative value of the reconstructed Wigner function at the origin was taken as a sign of nonclassical behavior. Such a statement is, however, valid only for diagnostic purposes. For example, a negative value of the reconstructed Wigner function at the origin was taken as a sign of nonclassical behavior. Such a statement is, however, valid only for diagnostic purposes.

Hence, an efficient quantum-state diagnostics is not possible without efficient control of the relevant statistical errors. In addition, the errors of different quantities are not independent. This can be shown by deriving uncertainty relation analogous to the famous Heisenberg principle, which holds for any tomographic scheme. Due to the Cauchy–Schwarz inequality the errors of a pair of generic observables $A$ and $B$ fulfill the relation

$$
(a \cdot F^{-1}a)(b \cdot F^{-1}b) \geq |a \cdot F^{-1}b|^2.
$$

That is why tomography schemes should be optimized with respect to the inferred variable of interest. In principle, our resolution measure (2) can be used for finding the optimal tomography scheme, which would minimize the error of a particular variable. Such a complicated multi-dimensional optimization is, however, beyond the scope of the present paper.

Let us illustrate the meaning and relevance of our main result equation (2) with an example of optical homodyne tomography [6]. Homodyne measurements consist of projections into the eigenvectors of rotated quadrature operators $Q(\theta) = x \cos \theta + p \sin \theta$. Denoting the phase of the local oscillator $\theta$ together with the detected eigenvalue $q$ by a single number $\gamma = q \exp(i\theta)$, the corresponding POVM elements read [6]

$$
A_\gamma = \sum_{m,n,k=0} f_{mnk}(\nu,\gamma) |m+k\rangle \langle n+k|,
$$

where $\nu$ is the overall detection efficiency of the scheme,

$$
f_{mnk}(\nu,\gamma) = \nu^{(m+n)/2} |\gamma|^m |\gamma|^n (1 - \nu)^k \sqrt{C_n^m C_{nk}^m C_{mk}^m},
$$

$C_n^m = m! / n!(m - n)!$ and the projections of the quadrature operator eigenstates $|\gamma\rangle$ on the Fock states $|n\rangle$ are given as

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

Although the resolution measure in equation (2) has been derived for the ML estimate, it is rather general. In the asymptotic regime of many detections per channel, where the approximation equation (1) holds, all reasonable reconstruction methods yield very similar results [17].

*New Journal of Physics* 10 (2008) 043022 (http://www.njp.org/)
Figure 1. Homodyne tomography of a noised ‘Schrödinger kitten’ state. Panel (a) shows the real part of a typical reconstruction $\rho_{\text{ML}}^{nm}$ in a 8D Fock subspace. Panel (b) shows the corresponding Wigner function.

Our method of error estimation has been applied to a simulated homodyne tomography of a noisy ‘Schrödinger kitten’ state $\rho = x|0\rangle\langle 0| + (1-x)|\alpha_{\text{odd}}\rangle\langle \alpha_{\text{odd}}|$, which is an almost equal mixture of the odd coherent state $|\alpha_{\text{odd}}\rangle = \mathcal{N}(|\alpha\rangle - |\alpha\rangle)$ of amplitude $\alpha = 0.6$ with the vacuum component, described by the mixing ratio $x = 0.4975$. States of this type can be prepared experimentally, e.g. by subtracting a photon from a squeezed vacuum state [4, 5, 23]. With the parameters as given above the Wigner representation takes a slightly negative value at the origin, $W(0, 0) = -0.0157$, which is an indicator of a nonclassical behavior. We have deliberately chosen a state close to the classical/nonclassical border in order to demonstrate how delicate the quantum-state diagnostics is. We ask whether the nonclassicality of this true state can be revealed by standard homodyne tomography.

To make our simulations as realistic as possible, similar parameters of the homodyne detection were used as those found in [4, 5]: the overall detection efficiency was set at $\nu = 0.8$ and six different phases of the local oscillator (phase cuts) were used uniformly distributed in the $[0, \pi]$ phase interval. For each phase cut, 20,000 random samples were generated. They were grouped into 64 bins before passing them to the reconstruction routine.

Figure 1 shows the standard interpretation of quantum homodyne tomography of a noisy Schrödinger kitten state. Panel (a) shows a typical ML reconstruction obtained in 8D Fock subspace. Notice that the Wigner function of the reconstructed state depicted in panel (b) is slightly negative at the origin, which is a sign of nonclassicality. One may be tempted to take this result as a proof of the nonclassical character of the measured state. Let us analyze in detail to what extent this interpretation is real or spurious.

To check the consistency of the result presented in figure 1, the simulation was repeated one thousand times. From each set of simulated data three reconstructions were always done in Fock subspace truncated at 5, 8 and 18 photons. These artificial cut-offs may correspond to different amounts of prior information that are available to the experimenter. From each reconstruction the value of $W(0, 0)$ was obtained and its standard deviation was calculated according to equation (2). Resulting histograms of $W(0, 0)$ and $\Delta W(0, 0)$ are shown in figure 2.

As shown in figure 1 a typical Wigner function reconstructed from the simulated data does take negative values. Now suppose the experimenter would repeat the experiment several times.
Figure 2. Diagnostics of the Wigner function at the origin based on homodyne data. Panels on the left show the histograms of 1000 reconstructed values of $W(0,0)$. Panels on the right show the corresponding histograms of uncertainties $\Delta W(0,0)$ calculated using equation (2). The reconstructions were done in 5D (upper panels), 8D (lower panels) and 18D (bottom panels) Fock subspaces. The bottom set of panels suggests that the homodyne tomography is not able to analyze unambiguously the nonclassical character of the measured state.

to see how much the reconstructed values of $W(0,0)$ would fluctuate. The left panels of figure 2 indicate that this would even strengthen his/her impression that the true state is nonclassical. However, the histograms of theoretical uncertainties $\Delta W(0,0)$ do not support such an optimistic view. Based on the calculated uncertainties, the estimated values of $W(0,0)$ can be summarized as follows:

$$W(0,0) = \begin{cases} -0.015 \pm 0.015, & d = 5, \\ -0.015 \pm 0.020, & d = 8, \\ -0.015 \pm 0.245, & d = 18. \end{cases}$$ (5)

Notice that the estimation errors strongly depend on the cut-off dimensions $d$. Without prior knowledge about the true state the reconstruction space (field of view) is specified by the measurement. By elementary logic, in order to make sure that no important components of the true state are left out, one should adopt the largest subspace of the total infinitely dimensional Hilbert space that may have contributed to the measured signal. Such a selection of the field of view of quantum tomography for the above mentioned set of homodyne measurements is demonstrated in figure 3. Here the spectrum of operator $G = \sum_i A_i$ is plotted for the original homodyne scheme measuring six phase cuts as well as for schemes with the number of phase cuts.
Figure 3. Spectrum of the operator $G$. The field of view of the tomography scheme is defined by the dominant eigenvalues $g_m$ of the operator $G$. It is obvious, that this space becomes smaller as the number of phase cuts is reduced from 6 (light gray bars) down to 4 (dark gray bars) or even 2 (full bars).

cuts reduced down to four or just two. Notice that the eigenspace spanned by the significant eigenvectors of operator $G$ is rather big—its dimension is around 80 for the six phase cuts registered. All states from this large subspace may have contributed to the data measured. Keeping in mind that the set of homodyne measurements in this example is sufficiently rich to explore the Fock space up to several tens of photons, the uncertainties of equation (5) nullify any nonclassical hypothesis.

One may wonder, why the fluctuations of $W(0, 0)$ in repeated simulations/experiments tend to underestimate the true uncertainty of $\Delta W(0, 0)$ as given by equation (2). There is a simple rationale behind the explanation. In high-dimensional Hilbert spaces, there are many states with $W(0, 0) > 0$ which can give rise to observations very similar to those generated by a nonclassical state in our example. These events appear rarely but may have a deep impact on the resolution. This may resemble the situation of a single photon diffraction at a rectangular slit. The width of the theoretical intensity distribution, $I(x) \propto \text{sinc}^2 x$ is infinite but the experimenter can never establish this simple fact from a finite data set: the experimentally sampled uncertainty grows with the size of data acquired. Similarly, it may be difficult to measure quantities, where significant contributions are due to very improbable events. For instance, consider a mixture of the vacuum state and a highly excited Fock state, $\rho = (1 - 1/n)|0\rangle\langle 0| + 1/n|n\rangle\langle n|$, where $n \gg 1$. Notice that the mean number of photons $\langle n \rangle = 1$ is independent of $n$, while the probability of observing no photons goes to unity for large $n$: $\lim_{n \to \infty} \langle n | \rho | n \rangle = 0$. For concreteness let us choose $n = 10^6$, for which the probability of observing nonzero signal is $p = 10^{-6}$. Hence even a long series of thousands of repetition of the photon-number measurement would, most probably, yield zero outcomes, from which the wrong result $\langle n | \rho | n \rangle = 0$ would have been calculated. Loosely speaking: even though many particles have already been observed the most important piece of data is still to come. In this respect, the tomography of high-dimensional objects is similar. Only in the unrealistic limit of a very large number of repetitions...
would the histogram of results start to reveal the real uncertainty of the estimated quantity. The important message is the following: the accuracy of the estimated quantity cannot be determined directly from the fluctuations of data, it must be calculated from equation (2).

The cornerstone of all these difficulties lies in the indirect character of a tomographic measurement. Objective quantum diagnostics is only possible with ‘focused’ observations that are sensitive only to states from a reasonably small subspace, so that the number of free parameters is limited. Curiously enough, it has been shown above that the seminal example of homodyne tomography [1] which started the boom of reconstruction techniques some 20 years ago does not meet this criterion. Here projections are done into strongly nonclassical states—the eigenvectors of rotated quadrature operators—which are not bounded in the phase space. As a consequence the homodyne measurement is not able to determine quantum states in the objective (i.e. data based only) way unless some prior information about the state is available.

As seen above the choice of a proper dimension of the reconstruction space is vital for successful diagnostics of nonclassical states. There are two concurring tendencies for the choice of this dimension. When the reconstruction space is low-dimensional, the reconstruction noise is kept low, however there may not be enough free parameters left for fitting of a possibly high-dimensional true state. In the case of the high-dimensional reconstruction space, the danger of missing important components of the true state is smaller, however the reconstruction errors may easily exceed acceptable levels as has been demonstrated above in the case of homodyne tomography.

A partial remedy for this dilemma is to introduce a dimension-dependent penalization into the fitting procedure. Suppose the true state is contained in a subspace of dimension $d_T$ spanned by the reconstruction basis $|n\rangle$: $\langle m|\rho|n\rangle \approx 0, \ m, n > d_T$. Since the measured data are noisy, the data fit monotonically improves with the dimension $d$ of the reconstruction subspace until a perfect fit is obtained. However, going from $d$ to $d + 1$, this improvement is much larger when $d < d_T$ compared to $d \geq d_T$. Loosely speaking, in the former case a substantial improvement of the fit is achieved by taking into account new components of the true state not contained in the $d$-dimensional subspace. In the latter case all important components of the true state are already contained in the $d$-dimensional subspace and the extra free parameters enable (slightly) better fitting of data noise. Such an increase of the fitting function can be subtracted from it to yield a new fitting function that increases only up to $d_T$.

In the framework of ML tomography, some additional dimension-dependent terms, such as the volume of the parameter space, are used as constraints for maximizing the likelihood giving rise to various information criteria such as the Akaike [17, 24], Schwarz [25] or modified-Schwarz [26] criteria. Considering the statistics of quantum measurements, the modified-Schwarz information [26],

$$I_{MS} = \log \mathcal{L}(\rho) - \frac{M}{2} \log S + \frac{M}{2} \log(2\pi) - \frac{1}{2} \log \det F,$$

seems to introduce penalization into quantum tomography in the most natural way. The first two terms on the right-hand side of equation (6) comprise Schwarz information [25] and the last two terms correspond to the normalization of the likelihood function.

A plot of the modified-Schwarz information as a function of $d$ is given in figure 4. The true state and the simulated measurements are the same as in figures 1 and 2. Notice that $I_{MS}$ exhibits the maximum for relatively low dimensions in agreement with the low photon numbers of the true state. At least in this particular case, the optimal dimension given by $I_{MS}$ seems to be small enough yet sufficient for fitting the true state. One may be tempted to use $I_{MS}$ or similar
quantities as a pragmatic tool for interpreting the results of quantum tomography. However, there seems to be a significant difference between the cut-offs imposed by the eigenvalues of operator $G$ and optimal dimension obtained from $I_{\text{MS}}$. The former one is given by the nature of the measurement: it does not make sense to do the predictions about state components which have not been measured. The tomography scheme is simply insensitive to this part of the Hilbert space. The latter condition is more of a statistical nature: higher dimensions of the model are not needed for fitting measured data. However, this does not necessarily mean that the true state has no components there. More research is needed to settle this important problem. As for now, the strategy recommended for quantum tomography is to stay on the safe side and do ‘well resolving’ tomography matched to the family of observed quantum states, that is measurements with a sufficiently localized spectrum of the $G$ operator. If this is not feasible or even not possible, the modified-Schwarz information may be used to keep the dimension of the reconstruction space reasonably small. In that case there will always remain a danger that some important parts of the measured signal are thrown away.

We have presented a reconstruction scheme for the diagnostics of quantum objects. An objective approach to tomography hinges upon the identification of the proper subspace where the reconstruction can be done. We have given a simple and operational recipe for placing error bars on any quantity inferred from a tomography measurement. We have applied this new resolution measure to quantum-optical homodyne tomography and shown that some nonclassical aspects of quantum states, such as the negativity of the Wigner function at the origin may be undetectable with the present technology unless some prior information on the measured system is available. Our measure may be adopted for designing optimized tomography schemes with resolution tuned to a particular purpose.

**Figure 4.** Negative modified Schwarz information equation (6) of ML states is shown as a function of the Fock space cut-off dimension $d$. The circle, square and triangle symbols correspond to three different sets of simulated homodyne data.
Acknowledgments

We acknowledge support from the Research Project MSM 6198959213 of the Czech Ministry of Education, grant no. 202/06/0307 of the Czech Grant Agency and Project COVQIAL FP6-511004 of the EU (JR and ZH), and from projects BRFFI of Belarus and CNPq of Brazil (DM).

Appendix. Derivation of the posterior distribution

A derivation of the posterior distribution equation (1) is given here for the special case of $G \equiv \sum_i A_i = 1$. The generalization to $G \neq 1$ is straightforward.

We start by expanding the log-likelihood function in the neighborhood of the maximum-likely state $\rho_{\text{ML}}$,

$$\log L(r) = \log L(0) + \frac{1}{2} \sum_{ij} F_{jk} r_j r_k,$$

(A.1)

where we defined

$$F_{jk} = -\frac{\partial^2 \log L(0)}{\partial r_j \partial r_k}$$

(A.2)

and used the extremal condition for the ML state: $\partial \log L(0)/\partial r_j = 0$. Since different outcomes of an ideal POVM measurement are independent, the corresponding statistics is multinomial

$$\log L = \sum_{i=1}^S N_i \log p_i,$$

which gives

$$F_{jk} = \sum_{i=1}^S \frac{N_i \partial p_i \partial p_i}{p_i^2 \partial r_j \partial r_k} - \sum_{i=1}^S \frac{N_i \partial^2 p_i}{p_i \partial r_j \partial r_k},$$

(A.3)

This expression can be further manipulated. For statistically significant sampling (many detections per channel) it holds that $N_i \approx N p_i$ and the first term simplifies to

$$\sum_{i=1}^S \frac{N_i \partial p_i \partial p_i}{p_i^2 \partial r_j \partial r_k} \approx N^2 \sum_{i=1}^S \frac{\partial p_i \partial p_i}{N_i \partial r_j \partial r_k},$$

(A.4)

while the second term can be neglected since

$$\sum_{i=1}^S \frac{N_i \partial^2 p_i}{p_i \partial r_j \partial r_k} \approx N \frac{\partial^2}{\partial r_j \partial r_k} \sum_{i=1}^S p_i = 0. $$

(A.5)

On substituting $F$ into equation (A.1) and normalizing the resulting likelihood function to unity the posterior distribution equation (1) is established. Here a remark seems to be in the order concerning the parametrization of quantum states used in this derivation. Even though some $r$-vectors may correspond to nonphysical nonpositive ‘states’, this has no consequences for the validity of the final result within the approximations made. In the asymptotic regime of many detections per channel we are interested in, the likelihood function becomes localized inside the set of physical density matrices and the likelihood of nonpositive ‘states’ gets vanishingly small. This is true for all measured true states except for a family of true states comprising the boundary of physical states (pure states for example). This family is however a set of measure zero and hence can be excluded from practical considerations.
It is worth mentioning that matrix $F$ can be interpreted as a Fisher information matrix for a generic quantum tomography measurement. The Fisher information matrix is defined in statistics by the following expression \[(A.6)\]

\[
F_{jk} = \left\langle \frac{\partial \log L}{\partial r_j} \frac{\partial \log L}{\partial r_k} \right\rangle,
\]

where $\langle \cdots \rangle$ denotes averaging over data. For multinomial statistics of a generic quantum measurement $F$ assumes the form \[(A.7)\]

\[
F_{jk} = \sum_{i,i'=1}^{s} \frac{1}{p_i p_{i'}} \frac{\partial p_{i'}}{\partial r_j} \frac{\partial p_i}{\partial r_k} \langle N_i N_{i'} \rangle.
\]

Using the well-known expression for the correlation of multinomially distributed data $\langle N_i N_{i'} \rangle = \langle N_i \rangle \delta_{ii'} = N p_i \delta_{ii'}$, we get \[(A.8)\]

\[
F_{jk} = \sum_{i=1}^{s} \frac{\langle N_i \rangle \partial p_i}{p_i^2} \frac{\partial p_i}{\partial r_j} \frac{\partial p_i}{\partial r_k} \approx N^2 \sum_{i=1}^{s} \frac{1}{N_i} \frac{\partial p_i}{\partial r_j} \frac{\partial p_i}{\partial r_k},
\]

which coincides with our equation \((A.4)\).

References

[1] Vogel K and Risken H 1989 Phys. Rev. A \textbf{40} 2847
[2] Smity D T, Beck M, Raymer M G and Faridani A 1993 Phys. Rev. Lett. \textbf{70} 1244
[3] Paris M G A and Řeháček J (ed) 2004 Quantum States Estimation (Lecture Notes in Physics vol 649) (Berlin: Springer)
[4] Ourjoumtsev A, Tualle-Brouri R and Grangier P 2006 Phys. Rev. Lett. \textbf{96} 213601
[5] Neergaard-Nielsen J S, Nielsen B M, Hettich C, Moelmer K and Polzik E S 2006 Phys. Rev. Lett. \textbf{97} 083604
[6] Leonhardt U 1997 Measuring the Quantum State of Light (Cambridge: Cambridge University Press)
[7] Bernardo J M and Smith A F M 1994 Bayesian Theory (Chichester: Wiley)
[8] Bužek V and Derka R 2001 Quantum observations Coherence and Statistics of Photons and Atoms ed J Peina (New York: Wiley) p 198
[9] Caves C M, Fuchs C A and Schack R 2002 J. Math. Phys. \textbf{43} 4537
[10] Blume-Kohout R 2006 Optimal, reliable estimation of quantum states Preprint quant-ph/0611080
[11] Blume-Kohout R and Hayden P 2006 Accurate quantum state estimation via ‘Keeping the experimentalists honest’ Preprint quant-ph/0603116
[12] Helstrom C W 1976 Quantum Detection and Estimation Theory (New York: Academic)
[13] Hradil Z 1997 Phys. Rev. A \textbf{55} R1561
[14] Banaszek K, D’Ariano G M, Paris M G A and Sacchi M F 1999 Phys. Rev. A \textbf{61} 010304
[15] Hradil Z, Mogilevtsev D and Řeháček J 2006 Phys. Rev. Lett. \textbf{96} 230401
[16] de Burgh M D, Langford N K, Doherty A C and Gilchrist A 2007 Choice of measurement sets in qubit tomography Preprint 0706.3756
[17] Usami K, Nambu Y, Tsuda Y, Matsumoto K and Nakamura K 2003 Phys. Rev. A \textbf{68} 022314
[18] Mohseni M, Rezakhani A T and Lidar D A 2008 Phys. Rev. A \textbf{77} 032322
[19] James D F V, Kiwary P G, Munro W J and White A G 2001 Phys. Rev. A \textbf{64} 052312
[20] Lvovsky A I 2004 J. Opt. B: Quantum Semiclass. Opt. \textbf{6} S556
[21] Babichev S A, Brezger B and Lvovsky A I 2004 Phys. Rev. Lett. 92 047903
[22] Zambra G et al 2005 Phys. Rev. Lett. 95 063602
[23] Wakui K, Takahashi H, Furusawa A and Sasaki M 2007 Opt. Express 15 3568
[24] Akaike H 1974 IEEE Trans. Auto. Control 19 716
[25] Schwarz G 1978 Ann. Stat. 6 461
[26] Konishi S, Ando T and Imoto S 2004 Biometrika 91 27