

Maximum-likelihood coherent-state quantum process tomography

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2012 New J. Phys. 14 105021

(<http://iopscience.iop.org/1367-2630/14/10/105021>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 136.159.58.214

The article was downloaded on 22/10/2012 at 16:57

Please note that [terms and conditions apply](#).

Maximum-likelihood coherent-state quantum process tomography

Aamir Anis¹ and A I Lvovsky^{2,3,4}

¹ Indian Institute of Technology, Kharagpur, Kharagpur 721302, West Bengal, India

² Institute for Quantum Information Science, University of Calgary, Calgary, AB T2N 1N4, Canada

³ Russian Quantum Center, 100 Novaya St., Skolkovo, Moscow 143025, Russia

E-mail: lvov@ucalgary.ca

New Journal of Physics **14** (2012) 105021 (16pp)

Received 25 April 2012

Published 19 October 2012

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/14/10/105021

Abstract. Coherent-state quantum process tomography (csQPT) is a method for completely characterizing a quantum-optical ‘black box’ by probing it with coherent states and carrying out homodyne measurements on the output (Lobino *et al* 2008 *Science* **322** 563). We present a technique for csQPT that is fully based on statistical inference, in particular, quantum expectation-maximization. The method relies on the Jamiolkowski isomorphism and iteratively reconstructs the process tensor in the Fock basis directly from the experimental data. This approach permits incorporation of *a priori* constraints into the reconstruction procedure, thereby guaranteeing that the resulting process tensor is physically consistent. Furthermore, our method is easier to implement and requires a narrower range of coherent states than its predecessors. We test its feasibility using simulations on several experimentally relevant processes.

⁴ Author to whom any correspondence should be addressed.



Content from this work may be used under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 licence](https://creativecommons.org/licenses/by-nc-sa/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Contents

1. Introduction	2
2. The method	3
2.1. Iterative process estimation using the Jamiolkowski isomorphism	3
2.2. Practical issues	6
3. Implementation and results	9
4. Summary	13
Acknowledgments	13
Appendix. Proof of a monotonic increase of log-likelihood for the diluted algorithm	13
References	16

1. Introduction

The art of determining states of quantum systems—quantum tomography—relies on carrying out measurements over multiple copies of the state in various bases, followed by reconstruction of the state’s density matrix using suitable algorithms on the procured data. Methods of state tomography can be extended to the quantum version of the ‘black box’ problem [1–3], giving rise to quantum process tomography (QPT). In QPT, measurements on the black box response to a certain set of probe states allow one to predict the effect of that black box on any arbitrary state within a given Hilbert space. QPT emerged in response to ever-increasing demands in the field of quantum information processing, as the assembly of any quantum information processor requires precise knowledge of each of its components [4].

A popular approach to QPT involves determining the output $\mathcal{E}(\hat{\rho}_i)$ for each state of a spanning set $\{\hat{\rho}_i\}$ of the space of density matrices over the Hilbert space of interest. Due to the linearity of quantum processes over its density operators, the output of any arbitrary state $\hat{\rho} = \sum_i c_i \hat{\rho}_i$ can then be found as $\mathcal{E}(\hat{\rho}) = \sum_i c_i \mathcal{E}(\hat{\rho}_i)$.

This approach has recently been extended to the continuous-variable domain of quantum optics [5]. The reconstruction procedure involves probing the process with coherent states, i.e. simple laser pulses. It relies on the ability of the Glauber–Sudarshan P representation to express the density matrix of any quantum state as a linear combination of coherent states’ density matrices. Improvements of the algorithm have been presented in [6]. The algorithm has been tested in an experiment on characterizing quantum-optical memory [7]. Similar principles have recently been used to perform a characterization of quantum optical detectors [8, 9].

This method, known as coherent-state QPT (csQPT), has the advantage of employing only easy-to-prepare coherent states for probing. However, the numerical reconstruction procedures employed in [5, 6] involve an intermediate step of determining the density matrices of the output states $\mathcal{E}(|\alpha\rangle\langle\alpha|)$ for each probe coherent state $|\alpha\rangle$ and their subsequent integration with the P function. This approach requires a multistep calculation and it cannot be guaranteed to yield a process that is physically plausible, i.e. completely positive (CP) and trace non-increasing.

We present a reconstruction scheme that does away with this intermediate step and reconstructs the process directly from the experimental data using pure statistical inference. The experimental setup is equivalent to that of [5] and is illustrated in figure 1. The process reconstruction algorithm, on the other hand, is entirely different: it relies on the iterative maximum-likelihood approach. Its major advantage is the possibility to incorporate *a priori*

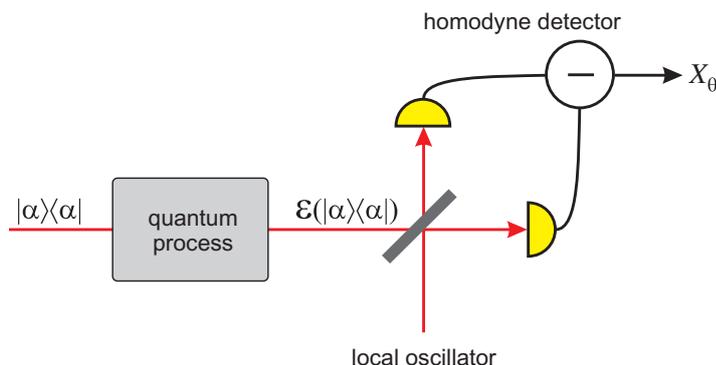


Figure 1. Schematic representation of the experimental setup for performing csQPT.

constraints in the reconstruction procedure in order to ensure physically consistent and meaningful results.

Maximum-likelihood methods have been successfully used in the past for quantum state estimation as well as QPT [4, 10, 11]. However, their role in QPT has been limited to the discrete variable state space. The technique presented in this paper extends the purview of maximum-likelihood QPT to the continuous variable state space, thereby allowing physically consistent quantum process estimation through homodyne tomography experiments [12]. A further advantage of the present technique is the need for a significantly narrower range of coherent states to probe the process as compared to [5, 6]. We test our approach on a number of processes that are relevant to quantum optical information processing: identity, attenuation and photon creation. In doing so, we elaborate a number of recommendations for practical use of the method.

2. The method

2.1. Iterative process estimation using the Jamiolkowski isomorphism

The process reconstruction scheme presented in this paper is based on the application of a maximum-likelihood-based QPT scheme [4, 11] to quadrature measurements in the Hilbert space associated with a harmonic oscillator. Consider a quantum optical process \mathcal{E} acting upon an optical mode prepared in some quantum state $\hat{\rho}_m$. The positivity of density matrices makes it necessary that \mathcal{E} be a CP map, in addition to being trace non-increasing [13]. The output state $\mathcal{E}(\hat{\rho}_m)$ of such a process can be subjected to optical homodyne measurements of its field quadratures $\hat{x}_\theta = \hat{x} \cos \theta + \hat{p} \sin \theta$, where \hat{x} and \hat{p} are the canonical position and momentum operators and θ is the local oscillator phase. For the output of the probe $\hat{\rho}_m$, the probability of detecting a specific quadrature value x for a phase θ is given by

$$p_\theta^m(x) = \text{Tr}[\hat{\Pi}(\theta, x)\mathcal{E}(\hat{\rho}_m)], \quad (1)$$

where $\hat{\Pi}(\theta, x) = |\theta, x\rangle\langle\theta, x|$ is the projector associated with the quadrature eigenstate $|\theta, x\rangle$ and the superscript m on the left-hand side denotes the probe state index. The above expression can be considered as a probability distribution function with \mathcal{E} as the parameter. If one carries out N measurements for each of the M input probe states $\hat{\rho}_m$, obtained as a set of phase and

corresponding quadrature values $\{\theta_{i,m}, x_{i,m}\}$, where $1 \leq i \leq N$ and $1 \leq m \leq M$, one can obtain the log-likelihood functional as

$$\begin{aligned} \mathcal{L}(\mathcal{E}) &= \sum_{m,i} \ln(p_{\theta_{i,m}}^m(x_{i,m})) \\ &= \sum_{m,i} \ln(\text{Tr}[\hat{\Pi}(\theta_{i,m}, x_{i,m})\mathcal{E}(\hat{\rho}_m)]). \end{aligned} \quad (2)$$

This functional is convex over the space of CP maps [14]. The objective of maximum-likelihood estimation is to determine the parameter \mathcal{E}_{est} that is as close to the actual parameter as possible, by maximizing the likelihood functional $\mathcal{L}(\mathcal{E})$ over the space of CP maps

$$\mathcal{E}_{\text{est}} = \arg \max_{\mathcal{E}} \mathcal{L}(\mathcal{E}). \quad (3)$$

This optimization problem is not straightforward and has been handled previously through various methods such as the uphill simplex [15]. However, a more rigorous yet technically simpler approach involves the formulation of an extremal equation that maximizes the log-likelihood functional given in equation (2).

In order to carry out the reconstruction procedure, one needs to first select a certain basis for the representation of the process and the relevant operators. In the Fock (number state) basis, the quantum process can be represented by a rank-4 tensor that relates the density matrix of the input and output states as [5, 6]

$$[\rho_{\text{out}}]_{jk} = \sum_{m,n} \mathcal{E}_{jk}^{mn} [\rho_{\text{in}}]_{mn}, \quad (4)$$

where $\mathcal{E}_{jk}^{mn} = \langle j|\mathcal{E}(|m\rangle\langle n|)|k\rangle$ and $\rho_{mn} = \langle m|\hat{\rho}|n\rangle$. Although the optical Hilbert space is of infinite dimension, in practical process tomography it is truncated to the spanning set of several lowest Fock states, as will be discussed later. Also, the projectors $\hat{\Pi}(\theta, x)$ can be expressed in this basis as

$$\Pi_{mn}(\theta, x) = \langle m|\hat{\Pi}(\theta, x)|n\rangle = \langle m|\theta, x\rangle\langle\theta, x|n\rangle, \quad (5)$$

where the overlap of the quadrature eigenstate with the number state is given as [10, 16]

$$\langle m|\theta, x\rangle = e^{im\theta} \left(\frac{1}{\pi^{1/4}} \right) \frac{H_m(x)}{\sqrt{2^m m!}} e^{-x^2}. \quad (6)$$

With the selected basis, we proceed to formulate a numerical procedure for the reconstruction of the quantum process. For a concise mathematical visualization, we resort to the Jamiołkowski isomorphism between linear CP maps \mathcal{E} from operators on the Hilbert space \mathcal{H} to the space \mathcal{K} and positive semidefinite operators \hat{E} on the Hilbert space $\mathcal{H} \otimes \mathcal{K}$. The explicit relation between \hat{E} and \mathcal{E} is given as [14]

$$\hat{E} = \sum_{m,n,j,k} \mathcal{E}_{jk}^{mn} |m\rangle\langle n| \otimes |j\rangle\langle k|. \quad (7)$$

With the definition in equation (7), the output $\hat{\rho}_{\text{out}} \in \mathcal{K}$ of a process \mathcal{E} for an input $\hat{\rho}_{\text{in}} \in \mathcal{H}$ is

$$\hat{\rho}_{\text{out}} = \mathcal{E}(\hat{\rho}_{\text{in}}) = \text{Tr}_{\mathcal{H}} \left[\hat{E} \hat{\rho}_{\text{in}}^T \otimes \hat{I}_{\mathcal{K}} \right], \quad (8)$$

where T denotes transposition in the number basis. In addition, one must apply the trace-preservation condition ($\text{Tr}[\hat{\rho}_{\text{out}}] = \text{Tr}[\hat{\rho}_{\text{in}}]$) over the process \mathcal{E} , which yields

$$\text{Tr}_{\mathcal{K}}[\hat{E}] = \hat{I}_{\mathcal{H}}. \quad (9)$$

The reconstruction procedure can be extended to encompass trace non-preserving processes, as will be shown subsequently. The problem has thus reduced to the determination of $(\dim \mathcal{H} \dim \mathcal{K})^2$ parameters subject to $\dim \mathcal{H}^2$ constraints. When the input and output Hilbert spaces are identical, this amounts to evaluating $\dim \mathcal{H}^4 - \dim \mathcal{H}^2$ free parameters.

For the process output of the input probe state $\hat{\rho}_m$, the probability of reading a quadrature value x for a given local oscillator phase θ can be obtained by substituting (8) into equation (1) to obtain

$$p_{\theta}^m(x) = \text{Tr}[\hat{E} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta, x)]. \quad (10)$$

Operator \hat{E} should then maximize a constrained log-likelihood functional in order to stand as the most likely quantum process that has the set of outcomes $\{\theta_{i,m}, x_{i,m}\}$ for the input probes $\{\hat{\rho}_m\}$. The relevant log-likelihood functional is given as

$$\begin{aligned} \mathcal{L}(\hat{E}) &= \sum_{m,i} \ln(p_{\theta_{i,m}}^m(x_{i,m})) - \text{Tr}[\hat{\Lambda} \hat{E}] \\ &= \sum_{m,i} \ln(\text{Tr}[\hat{E} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta_{i,m}, x_{i,m})]) - \text{Tr}[\hat{\Lambda} \hat{E}], \end{aligned} \quad (11)$$

where $\hat{\Lambda} = \hat{\lambda} \otimes \hat{I}_{\mathcal{K}}$ and $\hat{\lambda}$ is the Hermitian matrix of Lagrange multipliers that incorporates the trace preservation condition (9). Again, $\theta_{i,m}$ and $x_{i,m}$ belong to the set of quadrature data for the m th probe state given by $\{\theta_{i,m}, x_{i,m}\}$. An extremal equation can be obtained by varying equation (11) with respect to \hat{E} :

$$\delta \mathcal{L}(\hat{E}) = \mathcal{L}(\hat{E} + \delta \hat{E}) - \mathcal{L}(\hat{E}) = 0, \quad (12)$$

which gives

$$\text{Tr} \left[\left(\sum_{m,i} \frac{1}{p_{\theta_{i,m}}^m(x_{i,m})} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta_{i,m}, x_{i,m}) - \hat{\Lambda} \right) \delta \hat{E} \right] = 0. \quad (13)$$

This holds for all $\delta \hat{E}$, so that the expression in parentheses can be equated to zero and one has

$$\hat{E} = \hat{\Lambda}^{-1} \hat{R} \hat{E}, \quad (14)$$

where

$$\hat{R} = \sum_{m,i} \frac{1}{p_{\theta_{i,m}}^m(x_{i,m})} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta_{i,m}, x_{i,m}). \quad (15)$$

Owing to Hermiticity, one may rewrite equation (14) as $\hat{E} = \hat{E} \hat{R} \hat{\Lambda}^{-1}$. Using this, along with equation (14), we arrive at

$$\hat{E} = \hat{\Lambda}^{-1} \hat{R} \hat{E} \hat{R} \hat{\Lambda}^{-1}. \quad (16)$$

$\hat{\Lambda}$ can be determined by substituting the expression for \hat{E} in equation (16) into the trace-preservation condition (9):

$$\hat{\lambda} = \left(\text{Tr}_{\mathcal{K}}[\hat{R} \hat{E} \hat{R}] \right)^{1/2}. \quad (17)$$

Equations (16) and (17) can be solved numerically through iterations, starting from an unbiased initial \hat{E} , such as $\hat{E}^{(0)} = \hat{I}_{\mathcal{H} \otimes \mathcal{K}} / (\dim \mathcal{K})$. At each step of the iterations, the positive semi-definiteness of \hat{E} is ensured and the constraint $\text{Tr}_{\mathcal{K}}[\hat{E}] = \hat{I}_{\mathcal{H}}$ is satisfied.

Quantum processes may also be probabilistic, in which case the trace of the input quantum state is not preserved. The probability of occurrence of a probabilistic quantum process is given by

$$p_{\text{success}} = \text{Tr}[\mathcal{E}(\hat{\rho})]. \quad (18)$$

The reconstruction of probabilistic quantum processes can be viewed as a reconstruction of a trace-preserving, deterministic CP map $\tilde{\mathcal{E}}$ if the failure of the process is taken to be a measurement event associated with the projection operator $\hat{\Pi}_{\emptyset}$ onto a fictitious state $|\emptyset\rangle$ [14]. In order to analyze such a process, one can extend the Hilbert space to form $\mathcal{K}_{\text{total}} = \mathcal{K} \oplus \mathcal{K}_{\text{fail}}$, where $\mathcal{K}_{\text{fail}}$ is spanned by the single state $|\emptyset\rangle$. The original set of projectors $\hat{\Pi}_{\theta}(x)$ for each θ is augmented by adding $\hat{\Pi}_{\emptyset}$, so that the new set of projectors satisfies the closure relation over $\mathcal{K}_{\text{total}}$, i.e. $\forall \theta \int \hat{\Pi}_{\theta}(x) dx + \hat{\Pi}_{\emptyset} = I$. Subsequently, the likelihood functional, with the extended trace-preserving map $\tilde{\mathcal{E}}$ as parameter, can be rewritten as

$$\mathcal{L}(\tilde{\mathcal{E}}) = \sum_{m,i} [g_m \ln(p_{\theta_{i,m}}^m(x_{i,m})) + (1 - g_m) \ln(p_{\emptyset}^m)] - \text{Tr}[\hat{\Lambda} \hat{\tilde{E}}], \quad (19)$$

where g_m is the fraction of successful events over total events, which can be determined experimentally. The extremal equation would then contain a modified operator \hat{R} given by

$$\hat{R} = \sum_{m,i} \left[\frac{g_m}{p_{\theta_{i,m}}^m(x_{i,m})} \hat{\rho}_m^T \otimes \hat{\Pi}(\theta_{i,m}, x_{i,m}) + \frac{1 - g_m}{p_{\emptyset}^m} \hat{\rho}_m^T \otimes \hat{\Pi}_{\emptyset} \right]. \quad (20)$$

Iterations can now be performed with the new \hat{R} to obtain the trace-preserving process tensor $\hat{\tilde{E}}$. The actual process tensor \hat{E} is obtained by taking the projection of the estimated tensor $\hat{\tilde{E}}$ onto the subspace $\mathcal{H} \otimes \mathcal{K}$.

Our analysis so far did not specify which states were to be used as probes; the only requirement is that these states compose a spanning set in the space of density matrices. In csQPT, the role of probe states is played by coherent states [5]. The density operator of an arbitrary state can be written as a linear combination of coherent-state density operators using the optical equivalence theorem:

$$\hat{\rho} = \int P_{\hat{\rho}}(\alpha) |\alpha\rangle \langle \alpha| d^2\alpha, \quad (21)$$

where $P_{\hat{\rho}}(\alpha)$ is the Glauber–Sudarshan P function of state $\hat{\rho}$. Using the linearity of quantum processes with respect to density matrices, the process output is then given by

$$\mathcal{E}(\hat{\rho}) = \int P_{\hat{\rho}} \mathcal{E}(|\alpha\rangle \langle \alpha|) d^2\alpha. \quad (22)$$

Therefore, if the response of the quantum system to all coherent states is known, the output of any arbitrary unknown quantum state can be computed. In other words, measurements on the set of responses $\mathcal{E}(|\alpha_m\rangle \langle \alpha_m|)$ for coherent states $|\alpha_m\rangle$ provide tomographically complete information about the quantum process.

2.2. Practical issues

We now proceed to discussing a few practical issues arising in the implementation of the above algorithm of csQPT. The first issue is associated with infinite dimension of the optical Hilbert

space. In practical implementation of csQPT, the process tensor is reconstructed for a subspace $\mathbb{H}(n_{\max})$ of the Hilbert space spanned by Fock states up to a certain cut-off value, n_{\max} . The choice of n_{\max} is correlated with the maximum amplitude α_{\max} of the set of coherent probe states. Given a data set with a specific α_{\max} , the choice of n_{\max} depends on many factors, in particular, the process itself (see the online supplementary material of [5]).

For the iterative cycle, the cut-off value must be chosen sufficiently high so that $\mathbb{H}(n_{\max})$ accommodates all of the coherent probe states and the associated output states. Otherwise, the probe states and the quadrature data will be inadequately represented by $\mathbb{H}(n_{\max})$. This will lead to inaccurate reconstruction of the process tensor; we refer to this phenomenon as *truncation errors*.

For physically realistic processes, we expect the fractions of $|\alpha_m\rangle$ and $\mathcal{E}(|\alpha_m\rangle\langle\alpha_m|)$ that lie outside the reconstruction subspace to vanish as n_{\max} tends to infinity. Hence, for a given α_{\max} , it is possible to choose a value of n_{\max} such that the associated truncation errors are arbitrarily low [17].

However, a high cut-off value may give rise to another class of inaccuracies, which we call *data insufficiency errors*. If the overlap of a given Fock state $|n\rangle$ with all of the $|\alpha_m\rangle$ is low, so is the contribution of $|n\rangle$ to the log-likelihood functional, and hence the available data will not provide sufficient information about the effect of the process on $|n\rangle$. In contrast to the truncation errors, the data insufficiency errors grow with n_{\max} , but only apply to the process tensor elements associated with high input photon numbers.

Therefore, the following dual-step procedure for the choice of the cutoff may be necessary. The initial value of n_{\max} must be sufficiently high to ensure the absence of truncation errors. Subsequently, after the iterative cycle has been completed, we choose a secondary cut-off value, $n'_{\max} \leq n_{\max}$, and remove all the process tensor elements containing indices above n'_{\max} . The choice of n'_{\max} can be determined by calculating the statistical errors associated with each process tensor element—similar to the error estimations for state tomography [14, 18–20]. However, further research is required to determine statistical errors for QPT and establish a concrete bound for n'_{\max} . In the next section, we illustrate the effect of the chosen subspace dimension on the process reconstruction through various simulations.

As in the case of state tomography [10], our algorithm permits automatic correction for optical losses and inefficient detectors in the process tensor reconstruction. In order to account for non-unitary efficiency η , the projection operators are replaced by

$$\hat{\Pi}_\eta(\theta, x) = \sum_{m,n,j,k} B_{n+k,n}(\eta) B_{m+k,m}(\eta) \langle m | \hat{\Pi}(\theta, x) | n \rangle \times |m+k\rangle \langle n+k|, \quad (23)$$

where $B_{n+k,n} = \left[\binom{n+k}{k} \eta^n (1-\eta)^k \right]^{1/2}$. Substituting this into equation (15) and performing the iterations generates the original process tensor pertaining to the case of ideal detection.

Many physically relevant processes are phase-invariant: applying an optical phase shift to the input state results in the same shift to the output. Mathematically, such processes satisfy the following relation [6, 16]:

$$\mathcal{E}[\hat{U}(\phi)\hat{\rho}\hat{U}^\dagger(\phi)] = \hat{U}(\phi)\mathcal{E}(\hat{\rho})\hat{U}^\dagger(\phi). \quad (24)$$

In this case, further simplifications can be made. If the action of the process on a coherent state $|\alpha\rangle$ is known, so is the outcome for $|\alpha e^{i\phi}\rangle$. Therefore, one needs to only carry out measurements for input coherent states with amplitudes on the positive real axis. When condition (24) is applied in the Fock basis, the elements of the process tensor \mathcal{E}_{jk}^{mn} for a phase-invariant

process vanish except when $m - n = j - k$. This condition is incorporated into the probability distribution (10) as

$$p_{\theta}^m(x) = \text{Tr}[\mathcal{M}(\hat{E})\hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta, x)], \quad (25)$$

where \mathcal{M} denotes a masking operation over \hat{E} . If $\hat{\Pi}_m = |m\rangle\langle m|$ denotes a projection operator in the number basis, then $\mathcal{M}(\hat{E})$ can be expressed as

$$\mathcal{M}(\hat{E}) = \sum_{m,n,j,k} \delta_{m-n,j-k} (\hat{\Pi}_m \otimes \hat{\Pi}_j) \hat{E} (\hat{\Pi}_n \otimes \hat{\Pi}_k). \quad (26)$$

Since the trace operation is invariant under cyclic rearrangements of the operators and the Kronecker delta is invariant under transposition of indices, the probability distribution (25), and consequently, the expression for the operator \hat{R} in equation (15) change to

$$p_{\theta}^m(x) = \text{Tr}[\hat{E} \mathcal{M}(\hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta, x))], \quad (27)$$

$$\hat{R} = \sum_{m,i} \frac{1}{p_{\theta_{i,m}}^m(x_{i,m})} \mathcal{M}(\hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(\theta_{i,m}, x_{i,m})). \quad (28)$$

When the above relations are used, the elements of \mathcal{E}_{jk}^{mn} , for which $m - n \neq j - k$, vanish, resulting in the incorporation of the phase invariance condition.

In some cases, the value of the log-likelihood oscillates before converging to the maximum owing to overshoots. Stabilization can be achieved using the diluted algorithm that slows down but guarantees convergence [21]. The operator \hat{R} , in that case, is modified to a weighted sum of itself and the identity operator as

$$\hat{R}' = \mu \hat{R} + (1 - \mu) \hat{I}, \quad (29)$$

where $0 \leq \mu \leq 1$. As the value of μ decreases, the algorithm becomes more and more dilute, resulting in increased stability but a reduced rate of convergence. In addition, a monotonic increase of the likelihood is guaranteed for small values of μ (see the [appendix](#)). One may try to find the optimal value of μ that maximizes the increase in likelihood at the cost of increased computational complexity. Gradually varying the value of μ during the iterations might be justified for some processes.

The number of quadrature measurements for each probe state typically is in the range of tens of thousands. With multiple probe states, the iteration cycle may require significant computation time. In order to speed up the computation, binning of the data points in the quadrature and phase axes might be useful. A suitable step size is chosen for each axis as a trade-off between the desired computational time and the quantization error. For each $\mathcal{E}(|\alpha_m\rangle\langle\alpha_m|)$, quadrature data points are then clubbed into bins with centers $\{\theta_{u,m}, x_{v,m}\}$. With this modification, the log-likelihood functional (2) now reads as

$$\mathcal{L}(\mathcal{E}) = \sum_{m,u,v} h_{m;u,v} \ln[p_{\theta_{v,m}}^m(x_{u,m})], \quad (30)$$

where $h_{m;u,v}$ denotes the number of data points in the bin with the center $(\theta_{u,m}, x_{v,m})$. Ideally, one must obtain the POVM associated with the bin center as a function of all the positive-operator valued measures (POVMs) lying in the bin. However, given a small size of the bin, this element

can be approximated by projection onto the quadrature value at the center of the bin. Similarly, the operator \hat{R} in equation (15) can be rewritten as

$$\hat{R} = \sum_{m,u,v} \frac{h_{m:u,v}}{p_{\theta_{v,m}}^m(x_{u,m})} \hat{\rho}_m^T \otimes \hat{\Pi}(\theta_{v,m}, x_{u,m}). \quad (31)$$

For further speedup, one can compute \hat{R} in a parallel fashion on different threads owing to the absence of interdependence in the summation procedure.

In practical experiments on probabilistic processes, the frequency of successful events can be low: $g_m \ll 1$. In this case, the process tensor elements of interest (i.e. those related to \mathcal{K}) will be small and thus suffer from increased relative error. This issue can be resolved by rescaling the values of all g_m by the same factor for all probe states, keeping in mind the requirement that $g_m < 1$ for all m . Physically relevant elements of the process tensor will then rescale by the same factor, reducing the relative error.

3. Implementation and results

In order to test the algorithm, we have implemented it using Matlab and studied the reconstruction of a few quantum processes using simulated data. Theoretical process tensors of identity, attenuation and photon creation [6] were used to find the marginal probability distribution functions for various probe states using equation (10). From the marginals, we generated synthetic experimental data through Monte Carlo simulations.

Each process was applied to four coherent probe states with α_m ranging from 0 to 0.9375 in steps of 0.3125. For each input probe state, the output state data set consisted of 100 000 phase and quadrature points $\{\theta, x\}$. This set of data was subjected to the reconstruction method described. The iterations were halted when the change in process tensor elements was insignificant over a large number of iterations. However, a better approach would be to set a threshold for the increase of the log-likelihood [22].

The result obtained by running the reconstruction technique is a four-dimensional (4D) process tensor whose diagonal elements have a simple interpretation. For a given quantum process \mathcal{E} , the diagonal element \mathcal{E}_{kk}^{mm} denotes the probability that the output contains k photons when the process is subjected to m input photons. A comparison between the diagonal elements of the theoretical and the reconstructed process tensor is made in figure 2 and reveals the close match between the two.

The process of photon creation \hat{a}^\dagger deserves additional discussion because it corresponds to a non-unitary, trace non-preserving operator. Therefore, in experimental practice it can only be implemented probabilistically. The optical mode containing the target state $|\psi\rangle$ is directed into the signal channel of a parametric down-conversion setup. The state of the down-conversion output in the signal (s) and idler (i) channels can then be written as $|\psi\rangle_s |0\rangle_i + g(\hat{a}^\dagger |\psi\rangle)_s |1\rangle_i$, where g is the down-conversion amplitude. Detection of a photon in the idler channel projects the signal state onto $\hat{a}^\dagger |\psi\rangle$, thereby heralding a photon addition event [23]. For coherent-state input $|\psi\rangle = |\alpha\rangle$, the event probability, corresponding to the quantity g_m in equation (19), is proportional to $|g|^2(1 + |\alpha|^2)$.

We take the value $g^2 = 0.1$ during simulations to ensure that success probabilities remain less than 1 for the probe states selected. This makes the photon creation process trace non-increasing and thus physical. Note that the process tensor reported in figure 2(f) has been

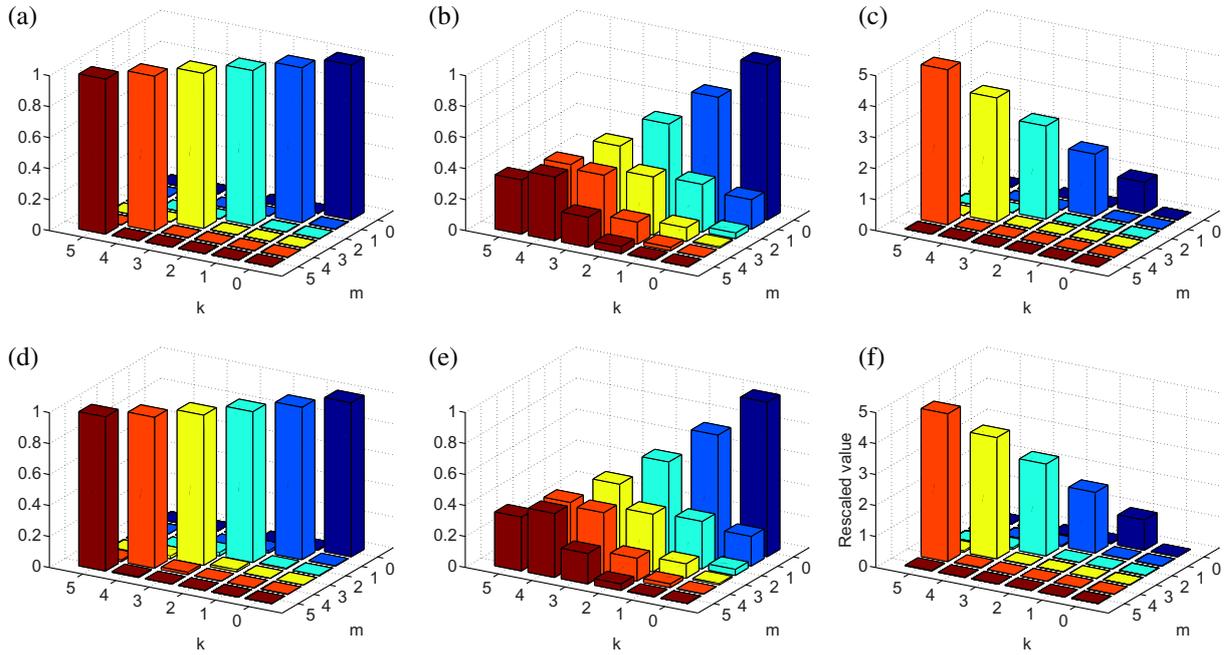


Figure 2. Comparison of the diagonal values (\mathcal{E}_{kk}^{mm}) of theoretical and reconstructed process tensors. Panels (a)–(c) show process tensors for theoretical identity, attenuation (by a factor of 0.9) and photon creation processes, respectively. Panels (d)–(f) show the process tensors for the corresponding reconstructed processes using the algorithm presented in this paper. The photon creation process tensor has been scaled to match the theoretical one (i.e. $g = 1$).

normalized by dividing by g^2 , so that its scale matches that of the process tensor for the photon creation operator \hat{a}^\dagger given in figure 2(c).

The iterative reconstruction of photon creation exhibited relatively poor convergence. Diluted iterations (29) were required in the beginning in order to curb oscillations. However, as the iterations progressed, the rate of increase of the likelihood value became extremely low. We circumvented this issue by implementing the successive over-relaxation technique. Setting μ in equation (29) to slightly over 1 while iterating accelerated the increase in likelihood. As soon as a decrease in the likelihood value was registered due to an overshoot, μ was reset to 1, and then slowly increased again after stabilization. This procedure was applied multiple times until a fair amount of convergence was observed. In a loose sense, the over-relaxation method employs linear extrapolation by selecting a tensor that lies on the line joining the current iterate and the next iterate but is beyond the latter by a fraction. If the iterations happen to proceed in the direction of maximum-likelihood gradient, it allows faster convergence by inducing greater leaps. Additionally, it may also help in escaping limit cycles encountered during the iterations.

We have also tested the reconstruction technique for photon creation in the case of inefficient detection. The output density matrices for the probe states were calculated using the beam splitter model of absorption [16]. With these modified density matrices, we have generated test data using Monte Carlo simulations for $\eta = 0.75$ and 0.55. A comparison of the reconstructed process tensors is given in figure 3.

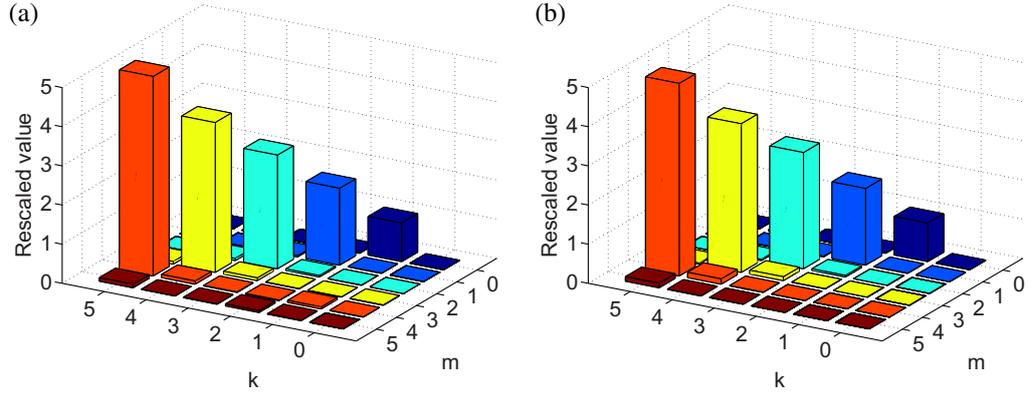


Figure 3. Reconstruction of the photon creation process with correction for inefficiency: (a) $\eta = 0.75$ and (b) $\eta = 0.55$.

Finally, we investigated the effect of the dimension of the subspace of optical Hilbert space chosen for the reconstruction, specifically for the photon creation process. In order to eliminate statistical errors in this reconstruction, we directly used the marginal distributions instead of simulated quadrature data sets to obtain the values of $h_{m;u,v}$ in equations (30) and (31). The performance criterion is taken to be the worst-case fidelity [24] over the input space $\mathbb{H}(n_{\max} - 1)$, defined as

$$\mathcal{F}(\mathcal{E}, \mathcal{E}_{\text{est}}) = \min_{\hat{\rho} \in \mathbb{H}(n_{\max}-1)} \text{Tr} \left(\sqrt{\sqrt{\mathcal{E}(\hat{\rho})} \mathcal{E}_{\text{est}}(\hat{\rho}) \sqrt{\mathcal{E}(\hat{\rho})}} \right), \quad (32)$$

where \mathcal{E} is the actual process tensor and \mathcal{E}_{est} is the estimated process tensor. Note that the photon number cutoff for the fidelity calculation is taken to be $n_{\max} - 1$ to ensure the Hilbert space closure under photon addition. Minimization over $\mathbb{H}(n_{\max} - 1)$ is carried out through a Monte Carlo simulation that involves introducing small random changes in the density matrix $\hat{\rho}$ within $\mathbb{H}(n_{\max} - 1)$ and accepting the change whenever the value of the fidelity decreases.

The solid line in figure 4(a) shows the worst-case fidelity versus n_{\max} for two values of α_{\max} . For each given α_{\max} , the fidelity initially increases with n_{\max} as the truncation effects subside and decreases afterwards due to data insufficiency. The range of n_{\max} over which the process tensor is reconstructed correctly shifts towards the higher photon numbers with increasing α_{\max} owing to the greater contribution of higher photon numbers in probe states of higher amplitudes.

Figures 4(b) and (c) further illustrate the two types of errors associated with the choice of the cut-off point. If n_{\max} is chosen too low (figure 4(b)), truncation errors compromise the entire reconstructed process. If the reconstruction subspace is sufficient to accommodate all the input probe states and associated output states (figure 4(c)), only the process tensor elements associated with high input photon numbers are reconstructed incorrectly. In this case, introducing a secondary cut-off at $n'_{\max} = 5$ is justified.

The dashed lines in figure 4(a) display the advantages of the dual cut-off approach introduced in section 2.2. For example, with $\alpha_{\max} = 0.6$, optimal reconstruction is attained with the initial cut-off point $n_{\max} \geq 6$ and subsequent cropping of the process tensor with $n'_{\max} = 5$. With this approach, the worst-case reconstruction fidelity is higher than for all the cases with $n'_{\max} = n_{\max}$. Note, however, that in most of the examples we studied, the dual cut-off method offers only a small advantage and may not be justified in practical csQPT.

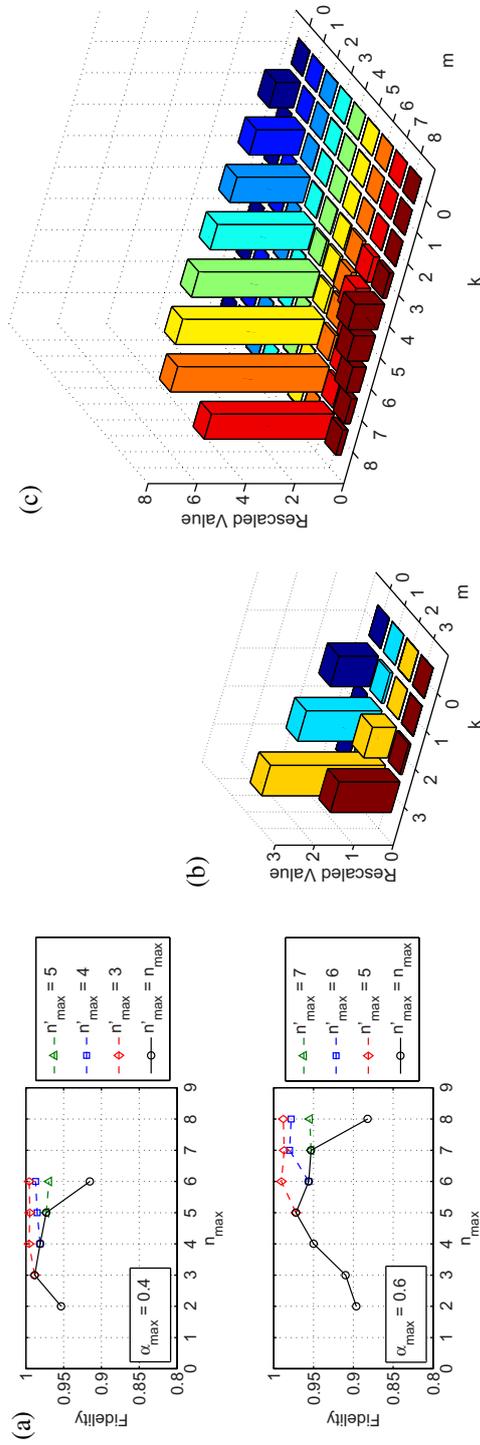


Figure 4. Effect of the photon number cut-off (n_{\max}) on the reconstruction of the photon creation process. (a) Worst-case process reconstruction fidelity as a function of n_{\max} for $\alpha_{\max} = 0.4$ (top) and $\alpha_{\max} = 0.6$ (bottom) and different n'_{\max} . The slight decreases of fidelity with increasing n_{\max} and constant n'_{\max} , observed in some cases, are numerical artefacts. (b) Diagonal values of the process tensor \mathcal{E}_{kk}^{mm} for $\alpha_{\max} = 0.6$ and $n_{\max} = 3$. The reconstructed process tensor has significant artefacts due to truncation errors. (c) Diagonal values of the process tensor \mathcal{E}_{kk}^{mm} for $\alpha_{\max} = 0.6$ and $n_{\max} = 8$. The reconstructed process tensor elements associated with input photon numbers 6–8 are invalid due to data insufficiency (for example, $|\langle \alpha_{\max} = 0.6 | n = 6 \rangle|^2 = 2.1 \times 10^{-6}$).

As evidenced by figure 4(a), the secondary cut-off points should be chosen close to $n'_{\max} = 3$ and 5, for $\alpha_{\max} = 0.4$ and 0.6, respectively. These values are much higher than those calculated in the supplementary information of [5]. Specifically, the method of [5] for the same values of $n_{\max} = 3, 5$ would require the maximum coherent-state amplitudes of 8 and 12, respectively. Further, the method of [6] requires $2N$ probe states for reconstruction in a Fock space of dimension $d = N + 1$, while our method poses no such constraints. In other words, for a given set of probe coherent states (defined by their number and maximum amplitude α_{\max}), the present reconstruction method provides much more information about the process tensor than previous methods.

One must note that further research is needed for the inverse problem of determining the optimum α_{\max} for a chosen Fock space dimension $d = n_{\max}$. According to the numerical examples we studied, it is reasonable to choose α_{\max} such that $\langle \alpha_{\max} | n_{\max} \rangle^2 \approx 1/N$, where N is the total number of quadrature measurements.

4. Summary

We have presented a maximum-likelihood-based experimental data processing technique for the tomographic reconstruction of quantum optical processes. This technique relies on measuring the response of the process to various coherent probe states through optical homodyne tomography. The reconstruction applies directly to the obtained data, unlike the previous csQPT methods that involve intermediate reconstruction of density matrices of the output states. The range of probe states required for reconstruction has also been reduced. Complete positiveness and trace preservation/non-increase conditions are incorporated into the estimated process tensor by imposing *a priori* constraints, thus yielding physical results. The simplicity and robustness of this technique make it appealing for quantum process estimation, with applications extending to optical quantum computing and quantum communication.

Acknowledgments

This work was supported by the National Sciences and Engineering Research Council of Canada, the Canadian Institute for Advanced Research and the Mitacs Globalink program. We thank Sweta and R Kumar for fruitful pointers and discussions.

Appendix. Proof of a monotonic increase of log-likelihood for the diluted algorithm

In this section, we shall prove that the diluted algorithm of equation (29) ensures a monotonic increase of the log-likelihood value for $0 < \mu \ll 1$. We start by considering the $(k+1)$ th iteration

$$\hat{E}_{(k+1)} = \hat{\Lambda}_{(k)}^{-1} \hat{R}_{(k)} \hat{E}_{(k)} \hat{R}_{(k)} \hat{\Lambda}_{(k)}^{-1}. \quad (\text{A.1})$$

As per the diluted algorithm, $\hat{R}_{(k)}$ is modified as

$$\hat{R}'_{(k)} = \mu \hat{R}_{(k)} + (1 - \mu) \hat{I}_{\mathcal{H} \otimes \mathcal{K}}. \quad (\text{A.2})$$

To find an expression for the normalization operator $\hat{\Lambda}'_{(k)} = \hat{\lambda}'_{(k)} \otimes \hat{I}_{\mathcal{H}} = (\text{Tr}_{\mathcal{K}}[\hat{R}'_{(k)} \hat{E}_{(k)} \hat{R}'_{(k)}])^{1/2} \otimes \hat{I}_{\mathcal{H}}$, we first evaluate $\hat{R}'_{(k)} \hat{E}_{(k)} \hat{R}'_{(k)}$ to first order in μ :

$$\begin{aligned} \hat{R}'_{(k)} \hat{E}_{(k)} \hat{R}'_{(k)} &= [\mu \hat{R}_{(k)} + (1 - \mu) \hat{I}_{\mathcal{H} \otimes \mathcal{K}}] \hat{E}_{(k)} [\mu \hat{R}_{(k)} + (1 - \mu) \hat{I}_{\mathcal{H} \otimes \mathcal{K}}] \\ &= \hat{E}_{(k)} - 2\mu \hat{E}_{(k)} + \mu \hat{R}_{(k)} \hat{E}_{(k)} + \mu \hat{E}_{(k)} \hat{R}_{(k)} + O(\mu^2). \end{aligned} \quad (\text{A.3})$$

The matrix $\hat{\lambda}'_{(k)}$ can be obtained as

$$\begin{aligned} \hat{\lambda}'_{(k)} &= (\text{Tr}_{\mathcal{K}}[\hat{R}'_{(k)} \hat{E}_{(k)} \hat{R}'_{(k)}])^{1/2} \\ &= (\text{Tr}_{\mathcal{K}}[\hat{E}_{(k)} - 2\mu \hat{E}_{(k)} + \mu (\hat{R}_{(k)} \hat{E}_{(k)} + \hat{E}_{(k)} \hat{R}_{(k)})])^{1/2} \\ &= \left[(1 - 2\mu) \hat{I}_{\mathcal{H}} + 2\mu \text{Tr}_{\mathcal{K}} \left(\frac{\hat{R}_{(k)} \hat{E}_{(k)} + \hat{E}_{(k)} \hat{R}_{(k)}}{2} \right) \right]^{1/2} \\ &= \left[(1 - \mu) \hat{I}_{\mathcal{H}} + \mu \text{Tr}_{\mathcal{K}} \left(\frac{\hat{R}_{(k)} \hat{E}_{(k)} + \hat{E}_{(k)} \hat{R}_{(k)}}{2} \right) \right] \end{aligned} \quad (\text{A.4})$$

$$(\hat{\lambda}'_{(k)})^{-1} = \left[(1 + \mu) \hat{I}_{\mathcal{H}} - \mu \text{Tr}_{\mathcal{K}} \left(\frac{\hat{R}_{(k)} \hat{E}_{(k)} + \hat{E}_{(k)} \hat{R}_{(k)}}{2} \right) \right], \quad (\text{A.5})$$

where in equation (A.4), we have used the trace preserving condition from equation (9). Thus, one has

$$\begin{aligned} (\hat{\Lambda}'_{(k)})^{-1} &= (\hat{\lambda}'_{(k)})^{-1} \otimes \hat{I}_{\mathcal{K}}, \\ &= (1 + \mu) \hat{I}_{\mathcal{H} \otimes \mathcal{K}} - \mu \hat{Y}_{(k)}, \end{aligned} \quad (\text{A.6})$$

where

$$\hat{Y}_{(k)} = \text{Tr}_{\mathcal{K}} \left(\frac{\hat{R}_{(k)} \hat{E}_{(k)} + \hat{E}_{(k)} \hat{R}_{(k)}}{2} \right) \otimes \hat{I}_{\mathcal{K}}.$$

Equation (A.1) can now be written as

$$\begin{aligned} \hat{E}_{(k+1)} &= (\hat{\Lambda}'_{(k)})^{-1} \hat{R}'_{(k)} \hat{E}_{(k)} \hat{R}'_{(k)} (\hat{\Lambda}'_{(k)})^{-1} \\ &= [(1 + \mu) \hat{I}_{\mathcal{H} \otimes \mathcal{K}} - \mu \hat{Y}_{(k)}] [\hat{E}_{(k)} - 2\mu \hat{E}_{(k)} + \mu \hat{R}_{(k)} \hat{E}_{(k)} + \mu \hat{E}_{(k)} \hat{R}_{(k)}] \\ &\quad \times [(1 + \mu) \hat{I}_{\mathcal{H} \otimes \mathcal{K}} - \mu \hat{Y}_{(k)}] \\ &= \hat{E}_{(k)} + \Delta \hat{E}_{(k)}, \end{aligned} \quad (\text{A.7})$$

where $\Delta \hat{E}_{(k)} = \mu (\hat{R}_{(k)} \hat{E}_{(k)} + \hat{E}_{(k)} \hat{R}_{(k)} - \hat{Y}_{(k)} \hat{E}_{(k)} - \hat{E}_{(k)} \hat{Y}_{(k)})$. The log-likelihood at the $(k + 1)$ th iteration is given by

$$\begin{aligned} \mathcal{L}(\hat{E}_{(k+1)}) &= \sum_{m,i} \ln(\text{Tr}[(\hat{E}_{(k)} + \Delta \hat{E}_{(k)}) \rho_m^{\text{T}} \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]) \\ &= \sum_{m,i} \ln(\text{Tr}[\hat{E}_{(k)} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]) + \sum_{m,i} \ln \left(1 + \frac{\text{Tr}[\Delta \hat{E}_{(k)} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]}{\text{Tr}[\hat{E}_{(k)} \hat{\rho}_m^{\text{T}} \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]} \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{m,i} \ln(\text{Tr}[\hat{E}_{(k)} \hat{\rho}_m^T \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]) + \sum_{m,i} \left(\frac{\text{Tr}[\Delta \hat{E}_{(k)} \hat{\rho}_m^T \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]}{\text{Tr}[\hat{E}_{(k)} \hat{\rho}_m^T \otimes \hat{\Pi}(x_{i,m}, \theta_{i,m})]} \right) \\
&= \mathcal{L}(\hat{E}_{(k)}) + \text{Tr}[\Delta \hat{E}_{(k)} \hat{R}_{(k)}]. \tag{A.8}
\end{aligned}$$

To prove the monotonicity of the log-likelihood functional, we require that $\text{Tr}[\Delta \hat{E}_{(k)} \hat{R}_{(k)}] = \mu \text{Tr}[\hat{R}_{(k)} \hat{E}_{(k)} (\hat{R}_{(k)} - \hat{Y}_{(k)}) + (\hat{R}_{(k)} - \hat{Y}_{(k)}) \hat{E}_{(k)} \hat{R}_{(k)}] \geq 0$ holds at each iteration, i.e.

$$\text{Tr}[\hat{R}_{(k)} \hat{E}_{(k)} (\hat{R}_{(k)} - \hat{Y}_{(k)}) + (\hat{R}_{(k)} - \hat{Y}_{(k)}) \hat{E}_{(k)} \hat{R}_{(k)}] \geq 0 \quad \forall k. \tag{A.9}$$

Starting with the left-hand side (dropping subscripts):

$$\begin{aligned}
\text{Tr}[\hat{R} \hat{E} (\hat{R} - \hat{Y}) + (\hat{R} - \hat{Y}) \hat{E} \hat{R}] &= \text{Tr}[2\hat{R} \hat{E} \hat{R} - \hat{R} \hat{E} \hat{Y} - \hat{Y} \hat{E} \hat{R}] \\
&= \text{Tr}[2\hat{R} \hat{E} \hat{R} - 2\hat{R} \hat{E} \hat{Y} - 2\hat{Y} \hat{E} \hat{R} + (\hat{R} \hat{E} \hat{Y} + \hat{Y} \hat{E} \hat{R})]. \tag{A.10}
\end{aligned}$$

Considering the expression $(\hat{R} \hat{E} \hat{Y} + \hat{Y} \hat{E} \hat{R})$:

$$\begin{aligned}
&\text{Tr}[\hat{R} \hat{E} \hat{Y} + \hat{Y} \hat{E} \hat{R}] \\
&= \text{Tr}[(\hat{R} \hat{E} + \hat{E} \hat{R}) \hat{Y}] \\
&= \text{Tr} \left[2 \frac{\hat{R} \hat{E} + \hat{E} \hat{R}}{2} \left(\text{Tr}_{\mathcal{K}} \left(\frac{\hat{R} \hat{E} + \hat{E} \hat{R}}{2} \right) \otimes \hat{I}_{\mathcal{K}} \right) \right] \\
&= \text{Tr} \left[2 \left(\text{Tr}_{\mathcal{K}} \left(\frac{\hat{R} \hat{E} + \hat{E} \hat{R}}{2} \right) \right)^2 \right] \tag{A.11}
\end{aligned}$$

$$= \text{Tr} \left[2 \left(\text{Tr}_{\mathcal{K}} \left(\frac{\hat{R} \hat{E} + \hat{E} \hat{R}}{2} \right) \right)^2 \text{Tr}_{\mathcal{K}}(\hat{E}) \right] \tag{A.12}$$

$$= \text{Tr} \left[2\hat{E} \left(\left(\text{Tr}_{\mathcal{K}} \left(\frac{\hat{R} \hat{E} + \hat{E} \hat{R}}{2} \right) \right)^2 \otimes \hat{I}_{\mathcal{K}} \right) \right]$$

$$= \text{Tr}[2\hat{E} \hat{Y}^2]$$

$$= \text{Tr}[2\hat{Y} \hat{E} \hat{Y}]. \tag{A.13}$$

We have used a property of the partial trace $\text{Tr}_{\mathcal{K}}$: $\text{Tr}[\hat{A}(\hat{B} \otimes \hat{I}_{\mathcal{K}})] = \text{Tr}[\hat{B} \text{Tr}_{\mathcal{K}}(\hat{A})]$ to arrive at (A.11) and used equation (9) to arrive at (A.12). Substituting (A.13) into

equation (A.10), we have

$$\begin{aligned} & \text{Tr}[\hat{R}\hat{E}(\hat{R} - \hat{Y}) + (\hat{R} - \hat{Y})\hat{R}\hat{E}] \\ &= 2 \text{Tr}[\hat{R}\hat{E}\hat{R} - \hat{R}\hat{E}\hat{Y} - \hat{Y}\hat{E}\hat{R} + \hat{Y}\hat{E}\hat{Y}] \\ &= 2 \text{Tr}[(\hat{R}\hat{E}^{1/2} - \hat{Y}\hat{E}^{1/2})(\hat{E}^{1/2}\hat{R} - \hat{E}^{1/2}\hat{Y})] \end{aligned} \quad (\text{A.14})$$

$$= 2 \text{Tr}[(\hat{E}^{1/2}\hat{R} - \hat{E}^{1/2}\hat{Y})^\dagger(\hat{E}^{1/2}\hat{R} - \hat{E}^{1/2}\hat{Y})] \quad (\text{A.15})$$

$$= 2 \text{Tr}[\hat{X}^\dagger\hat{X}] \geq 0 \quad \text{where } \hat{X} = \hat{E}^{1/2}\hat{R} - \hat{E}^{1/2}\hat{Y}, \quad (\text{A.16})$$

where in (A.14), the positive semidefiniteness of \hat{E} allows us to factorize it as $\hat{E} = \hat{E}^{1/2}\hat{E}^{1/2}$, with $\hat{E}^{1/2}$ being a Hermitian operator. Equation (A.15) follows as \hat{R} and \hat{Y} are also Hermitian operators. We arrive at inequality (A.16) as the operator $X^\dagger X$ is positive semidefinite and thus has non-negative trace. This completes the proof that the log-likelihood monotonically increases with k for $0 < \mu \ll 1$.

References

- [1] Poyatos J F, Cirac J I and Zoller P 1997 *Phys. Rev. Lett.* **78** 390
- [2] D'Ariano G M and Lo Presti P 2001 *Phys. Rev. Lett.* **86** 4195
- [3] Mohseni M, Rezakhani A T and Lidar D A 2008 *Phys. Rev. A* **77** 032322
- [4] Fiurášek J and Hradil Z 2001 *Phys. Rev. A* **63** 020101
- [5] Lobino M, Korystov D, Kupchak C, Figueroa E, Sanders B C and Lvovsky A I 2008 *Science* **322** 563
- [6] Rahimi-Keshari S, Scherer A, Mann A, Rezakhani A T, Lvovsky A I and Sanders B C 2011 *New J. Phys.* **13** 013006
- [7] Lobino M, Kupchak C, Figueroa E and Lvovsky A I 2009 *Phys. Rev. Lett.* **102** 203601
- [8] Lundeen J S *et al* 2009 *Nature Phys.* **5** 27
- [9] Zhang L *et al* 2012 *Nature Photon.* **6** 364
- [10] Lvovsky A I 2004 *J. Opt. B: Quantum Semiclass. Opt.* **6** S556
- [11] Ježek M, Fiurášek J and Hradil Z 2003 *Phys. Rev. A* **68** 012305
- [12] Lvovsky A I and Raymer M G 2009 *Rev. Mod. Phys.* **81** 299
- [13] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
- [14] Hradil Z, Řeháček J, Fiurášek J and Ježek M 2004 Maximum-likelihood methods in quantum mechanics *Quantum State Estimation (Springer Lecture Notes in Physics vol 649)* ed M G A Paris and J Řeháček (Berlin: Springer)
- [15] Banaszek K, D'Ariano G M, Paris M G A and Sacchi M F 1999 *Phys. Rev. A* **61** 010304
- [16] Leonhardt U 1997 *Measuring the Quantum State of Light (Cambridge Studies in Modern Optics)* (Cambridge: Cambridge University Press)
- [17] Killoran N and Lütkenhaus N 2011 *Phys. Rev. A* **83** 052320
- [18] Řeháček J, Mogilevtsev D and Hradil Z 2008 *New J. Phys.* **10** 043022
- [19] Blume-Kohout R 2012 arXiv:1202.5270v [quant-ph]
- [20] Christandl M and Renner R 2011 arXiv:1108.5329v1 [quant-ph]
- [21] Řeháček J, Hradil Z, Knill E and Lvovsky A I 2007 *Phys. Rev. A* **75** 042108
- [22] Glancy S, Knill E and Girard M 2012 arXiv:1205.4043v1 [quant-ph]
- [23] Zavatta A, Viciani S and Bellini M 2005 *Phys. Rev. A* **72** 023820
- [24] Lvovsky A I, Sanders B C and Tittel W 2009 *Nature Photon.* **3** 706–14