

# Quantum state tomography with noninstantaneous measurements, imperfections, and decoherence

Pierre Six, Philippe Campagne-Ibarcq, Igor Dotsenko, Alain Sarlette, Benjamin Huard, Pierre Rouchon

► **To cite this version:**

Pierre Six, Philippe Campagne-Ibarcq, Igor Dotsenko, Alain Sarlette, Benjamin Huard, et al.. Quantum state tomography with noninstantaneous measurements, imperfections, and decoherence. Physical Review A, American Physical Society, 2016, 93, pp.12109. <10.1103/PhysRevA.93.012109>. <hal-01395584>

**HAL Id: hal-01395584**

**<https://hal.inria.fr/hal-01395584>**

Submitted on 11 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Quantum state tomography with noninstantaneous measurements, imperfections, and decoherence**P. Six,<sup>1</sup> Ph. Campagne-Ibarcq,<sup>2</sup> I. Dotsenko,<sup>3</sup> A. Sarlette,<sup>4</sup> B. Huard,<sup>2</sup> and P. Rouchon<sup>1,\*</sup><sup>1</sup>*Centre Automatique et Systèmes, Mines-ParisTech, PSL Research University, 60, Boulevard Saint-Michel, 75006 Paris, France*<sup>2</sup>*Laboratoire Pierre Aigrain, Ecole Normale Supérieure–PSL Research University, CNRS, Université Pierre et Marie Curie–Sorbonne Universités, Université Paris Diderot–Sorbonne Paris Cité, 24 rue Lhomond, 75231 Paris Cedex 05, France*<sup>3</sup>*Laboratoire Kastler-Brossel, Ecole Normale Supérieure–PSL Research University, Université Pierre et Marie Curie–Sorbonne Universités, CNRS, Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*<sup>4</sup>*INRIA Paris–Rocquencourt, Domaine de Voluceau, Boîte Postale 105, 78153 Le Chesnay Cedex, France*

(Received 8 October 2015; revised manuscript received 2 December 2015; published 12 January 2016)

Tomography of a quantum state is usually based on a positive-operator-valued measure (POVM) and on their experimental statistics. Among the available reconstructions, the maximum-likelihood (MaxLike) technique is an efficient one. We propose an extension of this technique when the measurement process cannot be simply described by an instantaneous POVM. Instead, the tomography relies on a set of quantum trajectories and their measurement records. This model includes the fact that, in practice, each measurement could be corrupted by imperfections and decoherence, and could also be associated with the record of continuous-time signals over a finite amount of time. The goal is then to retrieve the quantum state that was present at the start of this measurement process. The proposed extension relies on an explicit expression of the likelihood function via the effective matrices appearing in quantum smoothing and solutions of the adjoint quantum filter. It allows us to retrieve the initial quantum state as in standard MaxLike tomography, but where the traditional POVM operators are replaced by more general ones that depend on the measurement record of each trajectory. It also provides, aside from the MaxLike estimate of the quantum state, confidence intervals for any observable. Such confidence intervals are derived, as the MaxLike estimate, from an asymptotic expansion of multidimensional Laplace integrals appearing in Bayesian mean estimation. A validation is performed on two sets of experimental data: photon(s) trapped in a microwave cavity subject to quantum nondemolition measurements relying on Rydberg atoms, and heterodyne fluorescence measurements of a superconducting qubit.

DOI: [10.1103/PhysRevA.93.012109](https://doi.org/10.1103/PhysRevA.93.012109)**I. INTRODUCTION**

Determining efficiently the state of a system whose preparation is imperfectly known is instrumental to quantum physics experiments. Contrary to classical physics, the determination of a quantum state  $\bar{\rho}$ , its tomography, requires a large number  $N$  of independent measurements [1,2]. The state of a quantum system is indeed a statistical quantity in essence, as it encodes the statistics of outcomes for any upcoming measurement. These measurements are usually modeled using a positive-operator-valued measure (POVM) defined by non-negative self-adjoint operators  $\pi_n$  such that  $\sum_n \pi_n = I$ . The probability of measurement outcome  $n$  is then given by  $\text{Tr}[\bar{\rho}\pi_n]$ . For  $N$  large enough, the reconstruction of  $\bar{\rho}$  is based on the fact that  $\text{Tr}[\bar{\rho}\pi_n]$  should be close to  $N_n/N$ , with  $N_n$  the number of outcomes  $n$  among the  $N$  independent measurements:  $\sum_n N_n = N$ . Several reconstruction methods are available. Maximum entropy [3] and compressed sensing [4] methods are well adapted to informationally incomplete sets of measurements. For informationally complete sets of measurements, maximum likelihood (MaxLike) reconstruction [5] is usually used: it consists in taking as an estimate for  $\bar{\rho}$  the value  $\rho_{\text{ML}}$  that maximizes the likelihood function denoted by  $\mathcal{P}(Y|\rho)$ , the probability of the measurement data  $Y \equiv (N_n)_n$  knowing  $\rho$ . For the POVM  $(\pi_n)_n$ , the likelihood function is directly

given by

$$\mathcal{P}(Y|\rho) = \prod_n (\text{Tr}[\rho\pi_n])^{N_n}. \quad (1)$$

In such a usual setting, the measurement process is assumed to be instantaneous and free from imperfections and decoherence. In specific experimental situations, such as in [6], it is not very difficult to take into account some measurement imperfections. In general, the derivation of the likelihood function in the presence of imperfections and decoherence during the measurement has not been investigated. This is one of the precise goals of this paper. Since the seminal work of Belavkin, it is known how to take into account measurement imperfection and decoherence for quantum filtering [7]. We show here how to exploit the stochastic master equation governing  $\rho_t$ , the conditional state of the quantum filter, to derive a general expression of the likelihood function: this expression, given by (5), is a direct generalization of the above one where the  $\pi_n$  are replaced by the adjoint states at the initial times of the  $N$  measurement sequences. These adjoint states obey a backward master equation appearing in quantum smoothing [8,9], and they correspond to the effective operator  $E$  defining the past quantum state  $(\rho, E)$  introduced in [10].

When the support of the likelihood function is mainly concentrated around its maximum at  $\rho_{\text{ML}}$ , it is known that  $\rho_{\text{ML}}$  is a good approximation of the Bayesian mean estimate  $\rho_{\text{BM}}$  defined by (see, e.g., [11])

$$\rho_{\text{BM}} = \frac{\int_{\mathcal{D}} \rho \mathcal{P}(Y|\rho) \mathcal{P}_0(\rho) d\rho}{\int_{\mathcal{D}} \mathcal{P}(Y|\rho) \mathcal{P}_0(\rho) d\rho}, \quad (2)$$

\* pierre.rouchon@mines-paristech.fr

where  $\mathcal{D}$  is the convex set of density operators (here the underlying Hilbert space is of finite dimension) and  $\mathcal{P}_0(\rho)$  is some prior probability law of  $\rho$  (e.g., Gaussian unitary ensemble [12]). Such an approximation of  $\rho_{\text{BM}}$  by  $\rho_{\text{ML}}$  relies on the first terms of an asymptotic expansion of multidimensional Laplace integrals [13] under some regularity conditions.

We show here how such asymptotic expansion provides also a confidence interval for  $\text{Tr}(\rho_{\text{ML}}A)$ , where  $A$  is any Hermitian operator. Such a confidence interval is based on a similar approximation for the Bayesian variance of  $\text{Tr}(\rho_{\text{ML}}A)$ , denoted by  $\sigma_{\text{ML}}^2(A)$ . We provide in (10) an explicit expression that depends only on the first and second-order derivatives of the log-likelihood function at its maximum  $\rho_{\text{ML}}$ . When  $\rho_{\text{ML}}$  has full rank, we recover the usual asymptotics of MaxLike estimators involving the Fisher information matrix and the Cramér-Rao bound (see, e.g., [14]). When  $\rho_{\text{ML}}$  is rank-deficient, the likelihood reaches its maximum on the boundary of  $\mathcal{D}$ . In this case, such an explicit expression approximating the Bayesian variance is not usual and, as far as we know, seems to be new. From a practical viewpoint, our MaxLike estimator thus provides a statistically efficient reconstruction method—exploiting all measurements and information about the dynamics of the measurement process—along with a confidence bound about the estimate, provided the measurement model is correct. In principle, an uncertain parameter in the dynamics of the measurement model could also be MaxLike estimated by looking at the likelihood of the measurement data conditioned on the parameter value, although this is not covered in the present paper. Some robustness to model errors is illustrated in Sec. II D.

After developing the general theory, we report such quantum state reconstructions for two different sets of experimental data. The first set corresponds to the quantum nondemolition photon counting in a cavity using Rydberg atoms, and presented in [15]. The stochastic master equation is a discrete-time Markov chain whose state is the photon-number population. The second set corresponds to the fluorescence heterodyne measurements of a superconducting qubit presented in [16], a system described by a continuous-time stochastic master equation driven by two Wiener processes. In both cases, we compute the MaxLike estimates  $\rho_{\text{ML}}$  of  $\bar{\rho}$ , the initial state whose tomography we are supposed to make. For any time  $t$  between 0 and  $T$ , we can ignore the measurement outcomes between 0 and  $t$  and just retain the measurement outcomes between  $t$  and  $T$  for the tomography. Thus we can artificially investigate the result of such tomography on the quantum state at time  $t$ , namely the state that would result from decoherence between 0 and  $t$ . We also give, for a physically interesting observable  $A$ , the usual 95% confidence interval via the approximation  $\text{Tr}(\rho_{\text{ML}}A) \pm 2\sigma_{\text{ML}}(A)$ . This confidence interval just means that, if we perform another tomography with another similar dataset, the probability that the MaxLike estimation of  $\text{Tr}(\bar{\rho}A)$  remains between  $\text{Tr}(\rho_{\text{ML}}A) - 2\sigma_{\text{ML}}(A)$  and  $\text{Tr}(\rho_{\text{ML}}A) + 2\sigma_{\text{ML}}(A)$  is greater than 95%. This probability has nothing to do with the quantum stochastic character of  $\bar{\rho}$ . Here,  $\bar{\rho}$  is considered, from a classical statistical point of view, as an unknown constant parameter of a probability law. More generally, any constant parameter appearing in the quantum filter governing the conditional state can be estimated in a similar way; see, e.g., [17].

The following section is devoted to discrete-time systems. First, we show how to obtain, from the discrete-time formulation of the quantum filter, an explicit expression of the likelihood function; then, we give the expression of the confidence interval and use it on the first set of experimental data related to the detection of a photon creation quantum jump. In another section, we show how to apply the discrete-time formulation to the continuous-time stochastic master equation driven by Wiener processes in order to obtain a numerical algorithm for computing the adjoint states and the likelihood function. Then, we use this numerical algorithm to estimate the initial state of a qubit relaxing toward its ground state and submitted to the heterodyne measurement of its fluorescence. In the Appendix, we give the main calculations yielding the asymptotic expression of the Bayesian variance for any observable  $A$ .

## II. DISCRETE-TIME SETTING

### A. Quantum filtering

We have at our disposal  $N$  realizations, starting from the same initial state  $\bar{\rho}$  that we want to determine, and producing  $N$  measurement records  $(y_t^{(n)})_{t=0,\dots,T_n}$ , indexed by  $n \in \{1, \dots, N\}$  and where the time  $t$  corresponds to an integer between 0 and  $T_n$ , the duration of realization  $n$ . For each realization, the quantum filter provides the conditional state at  $t$ , denoted by  $\rho_t^{(n)}$ , conditioned on  $\rho_0^{(n)} = \rho$  and knowing the past measurements  $(y_0^{(n)}, \dots, y_{t-1}^{(n)})$ . This filter, a discrete-time version of Belavkin's continuous-time filter, is defined by a family of completely positive maps  $\mathbf{K}_{y,t}$  indexed by the time  $t$  and the measurement outcome  $y$  (see, e.g., [18–20]). Moreover, for each  $t$ , the completely positive map  $\mathbf{K}_t = \sum_y \mathbf{K}_{y,t}$  is trace-preserving. For each  $y$  and density operator  $\xi$ ,  $\text{Tr}[\mathbf{K}_{y,t}(\xi)]$  is the probability to measure  $y$  at time  $t$  knowing that the quantum state at  $t$  is  $\xi$ . The quantum filter reads

$$\rho_{t+1}^{(n)} = \frac{\mathbf{K}_{y_t^{(n)},t}(\rho_t^{(n)})}{\text{Tr}[\mathbf{K}_{y_t^{(n)},t}(\rho_t^{(n)})]}, \quad t = 0, \dots, T_n. \quad (3)$$

A discrete-time quantum filter has this Markovian structure, with  $\mathbf{K}_{y,t}$  depending only on its physical settings. As the initial state  $\bar{\rho}$  is unknown, one can only work with  $\rho_t^{(n)}$  generated by the above recurrence and starting from a guess  $\rho_0^{(n)} = \rho$ .

To each measurement record  $(y_t^{(n)})_{t=0,\dots,T_n}$  we can associate a number  $\mathcal{P}_n(\rho)$  being the probability of getting this record, assuming the initial state was  $\rho_0^{(n)} = \rho$ . Since  $\text{Tr}[\mathbf{K}_{y_t^{(n)},t}(\rho_t^{(n)})]$  is the probability of having detected  $y_t^{(n)}$  knowing  $\rho_t^{(n)}$ , a direct use of Bayes' law yields  $\mathcal{P}_n(\rho) = \prod_{t=0}^{T_n} \text{Tr}[\mathbf{K}_{y_t^{(n)},t}(\rho_t^{(n)})]$  with  $\rho_t^{(n)}$  generated thanks to (3). Some elementary computations show that

$$\mathcal{P}_n(\rho) = \text{Tr}[\mathbf{K}_{y_{T_n}^{(n)},T_n} \circ \dots \circ \mathbf{K}_{y_0^{(n)},0}(\rho)].$$

Since the  $N$  measurement records are independent realizations from the same initial state, the probability  $\mathcal{P}(Y|\rho)$  of the measurement data,

$$Y = \{y_t^{(n)} | n \in \{1, \dots, N\}, t \in \{0, \dots, T_n\}\},$$

knowing the initial state  $\rho$ , reads

$$\mathcal{P}(Y|\rho) = \prod_{n=1}^N \mathcal{P}_n(\rho).$$

### B. Adjoint-state derivation of the likelihood function

The adjoint map  $\mathbf{K}_{y,t}^*$  of  $\mathbf{K}_{y,t}$  is defined by  $\text{Tr}[A\mathbf{K}_{y,t}(B)] \equiv \text{Tr}[\mathbf{K}_{y,t}^*(A)B]$  for all Hermitian operators  $A$  and  $B$ . Thus

$$\mathcal{P}_n(\rho) = \text{Tr}[\rho \mathbf{K}_{y_0,0}^* \circ \cdots \circ \mathbf{K}_{y_{T_n},T_n}^*(I)],$$

where  $I$  is the identity operator. Consider the normalized adjoint quantum filter with the adjoint state  $E_t$  (see, e.g., [8,10]), with final condition  $E_{T_n+1}^{(n)} = I$  and governed by the following backward recurrence:

$$E_t^{(n)} = \frac{\mathbf{K}_{y_t,t}^*(E_{t+1}^{(n)})}{\text{Tr}[\mathbf{K}_{y_t,t}^*(E_{t+1}^{(n)})]} \text{ for } t = T_n, \dots, 0. \quad (4)$$

It defines a family of Hermitian non-negative operators ( $E_t^{(n)}$ ) of trace 1 and depending only on the measurement data  $Y$ . We have  $\mathbf{K}_{y_0,0}^* \circ \cdots \circ \mathbf{K}_{y_{T_n},T_n}^*(I) = c_n E_0^{(n)}$  with  $c_n$  depending only on  $(y_0^{(n)}, \dots, y_{T_n}^{(n)})$ . Thus  $\mathcal{P}_n(\rho) = c_n \text{Tr}[\rho E_0^{(n)}]$  and we have

$$\mathcal{P}(Y|\rho) = \prod_{n=1}^N c_n \text{Tr}[\rho E^{(n)}], \quad (5)$$

where  $E^{(n)}$  stands for  $E_0^{(n)}$ . This formula is a generalization of (1) where  $E^{(n)}$  replaces  $\pi_n$ .

### C. Quantum state tomography

The maximum likelihood (MaxLike) estimate  $\rho_{\text{ML}}$  of the hidden initial quantum state  $\bar{\rho}$  underlying the measurement data  $Y$  is given by maximizing the likelihood function  $\mathcal{P}(Y|\rho)$ . It is usual to consider the log-likelihood function  $f(\rho) = \log[\mathcal{P}(Y|\rho)]$ :

$$\begin{aligned} f(\rho) &= \sum_{n=1}^N \log c_n + \sum_{n=1}^N \log(\text{Tr}[\rho E^{(n)}]) \\ &= C + \sum_{n=1}^N \log(\text{Tr}[\rho E^{(n)}]), \end{aligned} \quad (6)$$

where  $C$  is a constant independent of  $\rho$ . Thus,

$$\rho_{\text{ML}} = \underset{\rho \in \mathcal{D}}{\text{argmax}} f(\rho), \quad (7)$$

where  $\mathcal{D}$  is the set of density operators. Assume that the underlying Hilbert space is finite-dimensional. Then  $\mathcal{D}$  is a closed convex set and  $f$  is a smooth concave function. Thus this optimization problem can be efficiently solved numerically (see, e.g., [21]). Moreover,  $\rho_{\text{ML}}$  is characterized by the following necessary and sufficient conditions: there exists a non-negative scalar  $\lambda_{\text{ML}}$  such that

$$[\rho_{\text{ML}}, \nabla f_{\text{ML}}] = 0 \text{ and } \lambda_{\text{ML}} P_{\text{ML}} \leq \nabla f_{\text{ML}} \leq \lambda_{\text{ML}} I, \quad (8)$$

where  $P_{\text{ML}}$  is the orthogonal projector on the range of  $\rho_{\text{ML}}$ , and  $\nabla f_{\text{ML}}$  is the gradient at  $\rho_{\text{ML}}$  of the log-likelihood:

$$\nabla f_{\text{ML}} = \sum_{n=1}^N \frac{E^{(n)}}{\text{Tr}[\rho_{\text{ML}} E^{(n)}]}. \quad (9)$$

The necessary and sufficient condition (8) is just the translation of the standard optimality criterion for a convex optimization problem (see, e.g., [21]):  $\rho_{\text{ML}}$  maximizes the log-likelihood function over the convex set of density operators, if and only if, for all density operators  $\rho$ ,  $\text{Tr}[(\rho - \rho_{\text{ML}})\nabla f_{\text{ML}}] \leq 0$ . When  $\rho_{\text{ML}}$  has full rank, it belongs to the interior of  $\mathcal{D}$ . Then  $P_{\text{ML}} = I$  and  $\nabla f_{\text{ML}}$  is collinear to  $I$ .

When  $\rho_{\text{ML}}$  is rank-deficient, it lies on the boundary of  $\mathcal{D}$ . Then  $P_{\text{ML}} < I$  and (8) means that the gradient of the log-likelihood is pointing outward  $\mathcal{D}$  and is orthogonal to the tangent space at  $\rho_{\text{ML}}$  to the submanifold of density operators with the same rank as  $\rho_{\text{ML}}$ .

When the likelihood function is concentrated around its maximum,  $\rho_{\text{ML}}$  appears to be an approximation of the Bayesian mean estimate  $\rho_{\text{BM}}$  whose definition has been recalled in (2). It is proven in the Appendix that  $\rho_{\text{BM}} \approx \rho_{\text{ML}}$  independent of the chosen prior distribution  $\mathcal{P}_0$ . Thus, for any Hermitian operator  $A$ , its Bayesian mean

$$\langle A \rangle_{\text{BM}} = \frac{\int_{\mathcal{D}} \text{Tr}[\rho A] \exp[f(\rho)] \mathcal{P}_0(\rho) d\rho}{\int_{\mathcal{D}} \exp[f(\rho)] \mathcal{P}_0(\rho) d\rho}$$

can be approximated by  $\langle A \rangle_{\text{BM}} \approx \text{Tr}[\rho_{\text{ML}} A]$ . Similarly, we prove in the Appendix that its Bayesian variance

$$\sigma_{\text{BM}}^2(A) = \frac{\int_{\mathcal{D}} \text{Tr}^2[(\rho - \rho_{\text{ML}})A] \exp[f(\rho)] \mathcal{P}_0(\rho) d\rho}{\int_{\mathcal{D}} \exp[f(\rho)] \mathcal{P}_0(\rho) d\rho},$$

which captures the mean uncertainty on the value of  $\langle A \rangle_{\text{BM}}$ , due to the fact that the hidden initial state  $\bar{\rho}$  is unknown, can be approximated by

$$\sigma_{\text{BM}}^2(A) \approx \sigma_{\text{ML}}^2(A) \equiv \text{Tr}[A_{\parallel} \mathbf{R}^{-1}(A_{\parallel})], \quad (10)$$

where  $B_{\parallel} = B - \frac{\text{Tr}[B P_{\text{ML}}]}{\text{Tr}[P_{\text{ML}}]} P_{\text{ML}} - (I - P_{\text{ML}})B(I - P_{\text{ML}})$  is the orthogonal projector of any Hermitian operator  $B$  on the tangent space at  $\rho_{\text{ML}}$  to the submanifold of Hermitian operators with zero trace and with ranks equal to the rank of  $\rho_{\text{ML}}$ . Here above, the linear super-operator  $\mathbf{R}$  reads for any Hermitian operator  $X$ ,

$$\begin{aligned} \mathbf{R}(X) &= \sum_{n=1}^N \frac{\text{Tr}(X E_{\parallel}^{(n)})}{\text{Tr}^2(\rho_{\text{ML}} E^{(n)})} E_{\parallel}^{(n)} + \frac{1}{2}(\lambda_{\text{ML}} I - \nabla f_{\text{ML}}) X \rho_{\text{ML}}^+ \\ &\quad + \frac{1}{2} \rho_{\text{ML}}^+ X (\lambda_{\text{ML}} I - \nabla f_{\text{ML}}) \end{aligned} \quad (11)$$

with  $\lambda_{\text{ML}}$  and  $P_{\text{ML}}$  appearing in (8) and  $\rho_{\text{ML}}^+$  the Moore-Penrose pseudoinverse of  $\rho_{\text{ML}}$ . Notice that the superoperator  $\mathbf{R}$  is symmetric and non-negative for the Frobenius product. Thus,  $\sigma_{\text{ML}}^2(A)$  is always non-negative, as it should be.

When  $\rho_{\text{ML}}$  has full rank,  $P_{\text{ML}} = I$  and  $\nabla f_{\text{ML}} = \lambda_{\text{ML}} I$ . Then  $\mathbf{R}(X)$  corresponds to the orthogonal projection (for the Frobenius product) of  $-\nabla^2 f_{\text{ML}}(X)$  onto the subspace of Hermitian operators of zero trace. Here,  $\nabla^2 f_{\text{ML}}$  corresponds

to the Hessian of  $f$  at  $\rho_{\text{ML}}$ :

$$\nabla^2 f_{\text{ML}}(X) = - \sum_{n=1}^N \frac{\text{Tr}(X E^{(n)})}{\text{Tr}^2(\rho_{\text{ML}} E^{(n)})} E^{(n)}.$$

We recover, up to this orthogonal projection, the standard Cramér-Rao bounds attached to MaxLike estimation:  $-\nabla^2 f_{\text{ML}}$  stands for the Fisher information matrix. The superoperator  $\mathbf{R}$  defined by (11) is the prolongation of such a Fisher-information matrix when  $\rho_{\text{ML}}$  lies on the boundary of  $\mathcal{D}$ . Notice its dependence on boundary curvature due to the fact that  $(\lambda_{\text{ML}} I - \nabla f_{\text{ML}}) X \rho_{\text{ML}}^+$  does not vanish in general. Notice that the expressions of  $\langle A \rangle_{\text{BM}}$  and  $\sigma_{\text{BM}}^2(A)$  do not depend on the prior distribution of probability  $\mathcal{P}_0(\rho)$ . This can be easily understood, as these expressions are the first terms of asymptotic expansions when  $f(\rho)$  grows large, i.e., when the amount of information brought by the measurement records makes the initial information  $\mathcal{P}_0(\rho)$  outdated. The computations underlying (11) are given in the Appendix: they rely on a specific application of asymptotic expansions for multidimensional Laplace integrals given in Chap. 8 of [13].

#### D. Experimental validation for QND photon counting

We apply in the following the MaxLike reconstruction to the state  $\bar{\rho}$  of the light field stored in a cavity based on the experiment reported in [15] and, more precisely, experimental data associated with Fig. 4(b) therein. The field of a very high-quality superconducting cavity (with a frequency of 51 GHz and a photon lifetime of  $T_c = 65$  ms) is initially in a thermal state (with a temperature of 0.8 K and a mean number of thermal photons of  $n_b = 0.06$ ). At time  $t = 0^-$ , a single photon is injected into the field. Due to experimental imperfections, in reality 1.26 photons are injected on average. The QND photon number measurement consists of a long sequence of atomic probes (samples of individual atoms prepared in a highly excited circular Rydberg state) crossing the cavity mode one by one and separated by  $86 \mu\text{s}$ . The measurement starts at  $t = -172$  ms and has a duration of 344 ms, which is large compared to the photon lifetime. The main sources of measurement imperfections are random atomic occupation of samples, nonconstant atom-photon interaction from sample to sample, nonideal atom state detection, etc. The decoherence is mainly due to the limited cavity lifetime  $T_c$  leading to photon losses. For more details on the considered experiment, please refer to [15] and references therein.

Here, the quantum and adjoint states  $\rho_t$  and  $E_t$  are diagonal in the Fock basis and truncated to a maximum of seven photons: they are described by vectors of dimension 8. The partial Kraus maps  $\mathbf{K}_{y,t}$  reduced to  $8 \times 8$  real matrices, and the computations of the  $E^{(n)}$  rely on the transpose of these real matrices. We do not detail here the precise expressions of the different  $\mathbf{K}_{y,t}$ : they can be deduced from [15]. For any  $t$  between  $-100$  ms and  $+150$  ms, we compute the MaxLike estimation  $\rho_{\text{ML}}(t)$  of  $\bar{\rho}(t)$ , the state at time  $t$ , based on the measurement outcomes between  $t$  and  $+172$  ms. In a tomographic spirit, this MaxLike reconstruction takes into account the precise model of the “measurement” of state  $\bar{\rho}(t)$ : QND interaction with the atomic probes while the cavity is decohering, between time  $t$  and our last information at  $+172$  ms. However, it does not assume anything about the target state prior to measurement, i.e.,

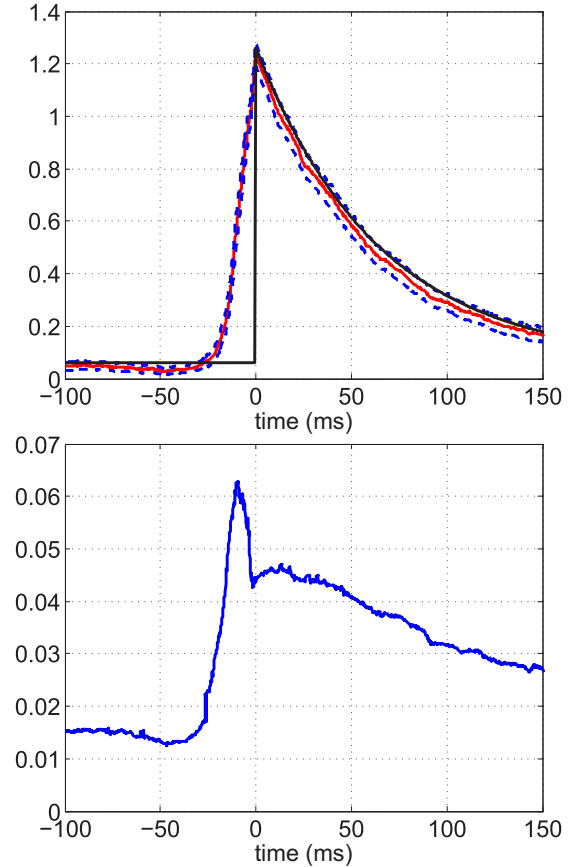


FIG. 1. MaxLike quantum state tomography based on the QND photon measurements reported in Fig. 4(b) of [15] for a photon creation quantum jump induced at  $t = 0$ . The latter is not taken into account in the measurement model. Top: solid red line corresponds to the MaxLike mean photon number  $\text{Tr}[\rho_{\text{ML}}(t) \mathbf{a}^\dagger \mathbf{a}]$  of the tomography target  $\bar{\rho}(t)$ , and based on measurements between  $t$  and  $+172$  ms. The two dashed blue lines correspond to  $\text{Tr}[\rho_{\text{ML}} \mathbf{a}^\dagger \mathbf{a}] \pm 2\sigma_{\text{ML}}(\mathbf{a}^\dagger \mathbf{a})$ , a 95% confidence interval. The black line corresponds to the mean photon number  $\langle \mathbf{a}^\dagger \mathbf{a} \rangle_t = \text{Tr}[\bar{\rho}(t) \mathbf{a}^\dagger \mathbf{a}]$  given by (12). Bottom:  $2\sigma_{\text{ML}}(\mathbf{a}^\dagger \mathbf{a})$  based on (10). See the text for more explanations.

it neglects anything that happened before  $t$ . Moreover, the addition of a photon at  $t = 0^-$  is not taken into account in the model, which means that for all  $t < 0$  there is in fact a mismatch in the model of the measurement process lasting from  $t$  to  $+172$  ms; this allows us to get an idea about the robustness of  $\rho_{\text{ML}}$  to model errors. For any time  $t$ , the rank of  $\rho_{\text{ML}}(t)$  is strictly less than 8 (between 2 and 5), and we have checked that it satisfies the characterization (8).

The results for  $N = 1390$  measurement trajectories are illustrated in Fig. 1. To evaluate the precision of such a MaxLike estimation, we have computed the variance given in (10) for the photon-number operator  $\mathbf{a}^\dagger \mathbf{a}$ . The resulting 95% confidence intervals show that, for all  $t$  between  $-100$  and  $+150$  ms, the mean photon numbers are evaluated with estimated uncertainties between  $\pm 0.015$  and  $\pm 0.065$ . We also compare the MaxLike reconstruction for  $t$  between  $-100$  and  $+150$  ms with the mean photon number  $\langle \mathbf{a}^\dagger \mathbf{a} \rangle_t = \text{Tr}[\bar{\rho}(t) \mathbf{a}^\dagger \mathbf{a}]$  given by

$$\langle \mathbf{a}^\dagger \mathbf{a} \rangle_t = \begin{cases} n_b & \text{for } t < 0, \\ n_b + (\langle \mathbf{a}^\dagger \mathbf{a} \rangle_0 - n_b) e^{-t/T_c} & \text{for } t \geq 0, \end{cases} \quad (12)$$



where  $\langle \mathbf{a}^\dagger \mathbf{a} \rangle_0 = 1.26$ ;  $T_c = 65$  ms and  $n_b = 0.06$  are derived from [15]. We observe that  $\langle \mathbf{a}^\dagger \mathbf{a} \rangle_t$  remains almost inside the 95% confidence tube except for  $t$  between  $-50$  ms and 0. This is a very positive result, and incidentally it is much better than the “backward reconstruction” based on future measurements performed in [15] [Fig. 4(b), blue curve], which features a significant offset. Let us briefly comment on the MaxLike estimation at each  $t$ .

At time 0 ms, the MaxLike photon number reaches its maximum,  $1.237 \pm 0.045$ . It corresponds to the programed injection of [15]. Our MaxLike estimation is a tomography of that state, i.e., it never includes information about what we have injected into the system before  $t = 0$ , and hence this is an independent confirmation of the programed injection. For  $t > 0$ , we expect the mean photon number to decay due to the finite photon lifetime. The MaxLike estimate confirms such behavior—again, just by making the tomography of  $\bar{\rho}(t)$  using measurements obtained after  $t$ , i.e., without any information on how  $\bar{\rho}(t)$  might have been constructed. For  $t$  between  $-50$  ms and 0, the observed mismatch is expected, since the photon jump at  $t = 0$  is not taken into account in the tomography model for the MaxLike reconstruction. Yet for  $t < -50$  ms we get a MaxLike mean photon number of  $0.042 \pm 0.015$ , close to the thermal photon number 0.06 observed in [15]. This is a notable result, illustrating that the MaxLike estimate is robust to the erroneous tomography model, which neglects the photon injection at  $t = 0$ . In other words, our MaxLike estimate appropriately gives more credit to measurements obtained right after  $\bar{\rho}(t)$  was prepared than to measurements further in the future, although it takes all these measurements into account with a (supposedly) appropriate weight. Hence, when  $t$  becomes negative, one simply has to collect enough measurements between  $t$  and 0 to recover a correct MaxLike estimation of  $\bar{\rho}(t)$ . Due to the moderate amount  $N$  of measurement trajectories here, having sufficiently many measurements means requiring that  $t < -50$  ms.

### III. ADAPTATION TO CONTINUOUS TIME

Most open quantum systems are modeled in continuous time. This section shows how to recover and exploit the discrete-time setting when  $dt$ , the sampling time, is much smaller than the characteristic time scales for systems modeled with stochastic master differential equations driven by Wiener processes.

#### A. Diffusive quantum filtering and adjoint state

The stochastic master equations admit here the following form (see, e.g., [7,22]):

$$d\rho_t = \left( -\frac{i}{\hbar} [H, \rho_t] + \sum_{\nu} L_{\nu} \rho_t L_{\nu}^{\dagger} - \frac{1}{2} (L_{\nu}^{\dagger} L_{\nu} \rho_t + \rho_t L_{\nu}^{\dagger} L_{\nu}) \right) dt + \sum_{\nu} \sqrt{\eta_{\nu}} (L_{\nu} \rho_t + \rho_t L_{\nu}^{\dagger} - \text{Tr}[(L_{\nu} + L_{\nu}^{\dagger}) \rho_t] \rho_t) dW_{\nu,t} \quad (13)$$

with  $dW_{\nu,t} = dy_{\nu,t} - \sqrt{\eta_{\nu}} \text{Tr}[(L_{\nu} + L_{\nu}^{\dagger}) \rho_t] dt$  given by the measurement  $y_{\nu,t}$ , attached to the operator  $L_{\nu}$ , with efficiency  $\eta_{\nu}$  between 0 and 1. Here,  $dW_{\nu,t}$  are independent scalar Wiener processes and  $H$  is the Hamiltonian that could depend on time  $t$  via, e.g., some time-varying coherent drives.

Following [17,23,24], such quantum filters admit the following infinitesimal discrete-time formulations based on Itô rules:

$$\rho_{t+dt} = \frac{\mathbf{K}_{dy_t}(\rho_t)}{\text{Tr}[\mathbf{K}_{dy_t}(\rho_t)]},$$

where the complete positive maps  $\mathbf{K}_{dy_t}$  depend on  $dy_t = (dy_{\nu,t})$  according to

$$\mathbf{K}_{dy_t}(\rho_t) = M_{dy_t} \rho_t M_{dy_t}^{\dagger} + \sum_{\nu} (1 - \eta_{\nu}) L_{\nu} \rho_t L_{\nu}^{\dagger} dt$$

with

$$M_{dy_t} = I + \left[ -\frac{i}{\hbar} H - \frac{1}{2} \left( \sum_{\nu} L_{\nu}^{\dagger} L_{\nu} \right) \right] dt + \sum_{\nu} \sqrt{\eta_{\nu}} dy_{\nu,t} L_{\nu}.$$

The probability of outcome  $dy_t = (dy_{\nu,t})$  is given then by the following distribution based on Gaussian laws of variance  $dt$ :

$$\begin{aligned} \mathcal{P} \left( dy_t \in \prod_{\nu} [\xi_{\nu}, \xi_{\nu} + d\xi_{\nu}] \middle| \rho_t \right) \\ = \text{Tr}[\mathbf{K}_{\xi}(\rho_t)] \prod_{\nu} e^{-\xi_{\nu}^2/2dt} \frac{d\xi_{\nu}}{\sqrt{2\pi dt}}. \end{aligned} \quad (14)$$

Take a sampling time  $dt$  much smaller than the time constant involved in the Hamiltonian  $H$  and in the decoherence operator  $L_{\nu}$ . Then we can exploit the above formulations similar to discrete-time quantum filtering and compute the normalized adjoint states  $E_t^{(n)}$ . They are associated with the measurement data  $(y_t^{(n)})_{0 \leq t \leq T_n}$  corresponding to the quantum trajectory number  $n$ , via the following discrete-time formulation:

$$E_t = \frac{\mathbf{K}_{dy_t^{(n)}}^*(E_{t+dt})}{\text{Tr}[\mathbf{K}_{dy_t^{(n)}}^*(E_{t+dt})]} \text{ with } E_{T_n+dt} = I/\text{Tr}(I)$$

with  $T_n = \mathcal{N}_n dt$  for  $\mathcal{N}_n$  a large integer and where the adjoint  $\mathbf{K}_{dy_t}^*$  of  $\mathbf{K}_{dy_t}$  reads

$$\mathbf{K}_{dy_t}^*(\rho_t) = M_{dy_t}^{\dagger} \rho_t M_{dy_t} + \sum_{\nu} (1 - \eta_{\nu}) L_{\nu}^{\dagger} \rho_t L_{\nu} dt.$$

After having obtained, for the  $N$  quantum trajectories, the value at  $t = 0$  of the adjoint states  $(E^{(n)})_{n=1, \dots, N}$ , we can directly use the MaxLike quantum-state tomography developed in the preceding section.

#### B. Quantum state tomography for a qubit

For a two-level system, the MaxLike estimation developed in the previous sections admits a simpler formulation with the

Bloch sphere variables that can be used for both  $\rho$  and  $E$ :

$$\rho = \frac{I + x\sigma_x + y\sigma_y + z\sigma_z}{2},$$

$$E = \frac{I + e_x\sigma_x + e_y\sigma_y + e_z\sigma_z}{2},$$

where  $(x, y, z)$  and  $(e_x, e_y, e_z)$  correspond to the coordinate of vectors, with  $\sigma_x, \sigma_y,$  and  $\sigma_z$  the three Pauli matrices. Here the convex set  $\mathcal{D}$  corresponds to the unit ball  $x^2 + y^2 + z^2 \leq 1$  and the Frobenius product between operators to the Euclidian product between vectors in three-dimensional Euclidian space. Then the gradient of the log-likelihood function (9) becomes the vector

$$\nabla f_{\text{ML}} = \sum_n \frac{1}{1 + x_{\text{ML}}e_x^{(n)} + y_{\text{ML}}e_y^{(n)} + z_{\text{ML}}e_z^{(n)}} \begin{pmatrix} e_x^{(n)} \\ e_y^{(n)} \\ e_z^{(n)} \end{pmatrix}.$$

The characterization of  $\rho_{\text{ML}}$  given in (8) becomes as follows: if  $x_{\text{ML}}^2 + y_{\text{ML}}^2 + z_{\text{ML}}^2 < 1$ , then  $\nabla f_{\text{ML}} = 0$ ; if  $x_{\text{ML}}^2 + y_{\text{ML}}^2 + z_{\text{ML}}^2 = 1$ , then  $\nabla f_{\text{ML}} = \lambda_{\text{ML}} \begin{pmatrix} x_{\text{ML}} \\ y_{\text{ML}} \\ z_{\text{ML}} \end{pmatrix}$  with  $\lambda_{\text{ML}} \geq 0$ . The superoperator  $\mathbf{R}$  defined in (11) becomes a  $3 \times 3$  symmetric non-negative matrix. It reads as follows:

(i) When  $x_{\text{ML}}^2 + y_{\text{ML}}^2 + z_{\text{ML}}^2 < 1$ , we have

$$\mathbf{R} = \sum_n \frac{\begin{pmatrix} e_x^{(n)} e_x^{(n)} & e_x^{(n)} e_y^{(n)} & e_x^{(n)} e_z^{(n)} \\ e_y^{(n)} e_x^{(n)} & e_y^{(n)} e_y^{(n)} & e_y^{(n)} e_z^{(n)} \\ e_z^{(n)} e_x^{(n)} & e_z^{(n)} e_y^{(n)} & e_z^{(n)} e_z^{(n)} \end{pmatrix}}{(1 + x_{\text{ML}}e_x^{(n)} + y_{\text{ML}}e_y^{(n)} + z_{\text{ML}}e_z^{(n)})^2}.$$

It is usually of rank 3 and can be inverted on any operator of the form  $A = \frac{a\sigma_x + b\sigma_y + c\sigma_z}{2}$  associated with the vector  $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$  to get a variance estimation via (10) where the trace is replaced by the Euclidean scalar product.

(ii) When  $x_{\text{ML}}^2 + y_{\text{ML}}^2 + z_{\text{ML}}^2 = 1$  and  $\lambda_{\text{ML}} > 0$  large enough, we have

$$\mathbf{R} = \sum_n \frac{\begin{pmatrix} e_{\parallel x}^{(n)} e_{\parallel x}^{(n)} & e_{\parallel x}^{(n)} e_{\parallel y}^{(n)} & e_{\parallel x}^{(n)} e_{\parallel z}^{(n)} \\ e_{\parallel y}^{(n)} e_{\parallel x}^{(n)} & e_{\parallel y}^{(n)} e_{\parallel y}^{(n)} & e_{\parallel y}^{(n)} e_{\parallel z}^{(n)} \\ e_{\parallel z}^{(n)} e_{\parallel x}^{(n)} & e_{\parallel z}^{(n)} e_{\parallel y}^{(n)} & e_{\parallel z}^{(n)} e_{\parallel z}^{(n)} \end{pmatrix}}{(1 + x_{\text{ML}}e_x^{(n)} + y_{\text{ML}}e_y^{(n)} + z_{\text{ML}}e_z^{(n)})^2} + \lambda_{\text{ML}} \begin{pmatrix} 1 - x_{\text{ML}}x_{\text{ML}} & -x_{\text{ML}}y_{\text{ML}} & -x_{\text{ML}}z_{\text{ML}} \\ -y_{\text{ML}}x_{\text{ML}} & 1 - y_{\text{ML}}y_{\text{ML}} & -y_{\text{ML}}z_{\text{ML}} \\ -z_{\text{ML}}x_{\text{ML}} & -z_{\text{ML}}y_{\text{ML}} & 1 - z_{\text{ML}}z_{\text{ML}} \end{pmatrix}$$

where  $e_{\parallel \xi}^{(n)} = e_{\xi}^{(n)} - s^{(n)}\xi_{\text{ML}}$  for  $\xi = x, y, z$  and  $s^{(n)} = e_x^{(n)}x_{\text{ML}} + e_y^{(n)}y_{\text{ML}} + e_z^{(n)}z_{\text{ML}}$ . Its rank is less than or equal to 2. The inverse of  $\mathbf{R}$  appearing in (10) corresponds here to the Moore-Penrose pseudoinverse. It is evaluated on the vector associated with  $A_{\parallel}$ :

$$A_{\parallel} = \frac{(a - sx_{\text{ML}})\sigma_x + (b - sy_{\text{ML}})\sigma_y + (c - sz_{\text{ML}})\sigma_z}{2}$$

with  $s = ax_{\text{ML}} + by_{\text{ML}} + cz_{\text{ML}}$ .

### C. Experimental validation for a qubit with fluorescence heterodyne measurements

MaxLike quantum state tomography is conducted on a superconducting qubit whose fluorescence field is measured using a heterodyne detector [25]. For the detailed physics of this experiment, see [16,24]. The measurement model is described by a stochastic master equation of the form (13) with  $H = 0$  and  $\nu = 1, 2, 3$ :

$$L_1 = \sqrt{\frac{1}{2T_1}} \frac{\sigma_x - i\sigma_y}{2}, \quad L_2 = iL_1, \quad L_3 = \sqrt{\frac{1}{2T_\phi}} \sigma_z$$

with  $\eta_1 = \eta_2 = 0.24$  the efficiency of the heterodyne measurement and with  $\eta_3 = 0$  corresponding to an unmonitored dephasing channel. The measurement and dephasing time constants are  $T_1 = 4.15 \mu\text{s}$  and  $T_\phi = 35 \mu\text{s}$ .

We have at our disposal  $N = 4 \times 10^4$  quantum trajectories with the same length  $T_n = T = 9.2 \mu\text{s}$  with a sampling time  $dt = 200 \text{ ns} \approx \frac{1}{20}T_1$ . Each trajectory is supposed to start at time  $t = 0$  from the same initial state  $\bar{\rho}_0$  close to  $(|g\rangle + |e\rangle)/\sqrt{2}$ . In a first test, probably closest to experimental needs, the goal is to check this fact by performing a MaxLike tomographic estimation of  $\bar{\rho}_0$  based on records of the continuous fluorescence signals between 0 and final time  $T$ . For the trajectory number  $n$ , the measurement record corresponds to  $2 \times 47$  real values corresponding to  $\int_{t-dt}^t dy_1^{(n)}$  and  $\int_{t-dt}^t dy_2^{(n)}$  for  $t = dt, 2dt, \dots, 47dt$ . From the measurements between 0 and  $T$ , we get an estimation of the quantum state  $\bar{\rho}_0$  with a 95% confidence interval using the above formula for  $\sigma_{\text{ML}}^2(A)$  and  $A = \sigma_x, \sigma_y,$  and  $\sigma_z$ :

$$\text{Tr}(\bar{\rho}_0\sigma_x) = 0.99 \pm 0.06,$$

$$\text{Tr}(\bar{\rho}_0\sigma_y) = -0.03 \pm 0.07,$$

$$\text{Tr}(\bar{\rho}_0\sigma_z) = -0.10 \pm 0.19.$$

This estimated value of the initial state is consistent with a gate error of a few percent in the preparation of  $(|g\rangle + |e\rangle)/\sqrt{2}$  starting from the thermal state with less than 1% excitation.

To further validate how the MaxLike tomography can reconstruct different states of the qubit, we next perform, as for the previous experiment, the tomography of the state  $\bar{\rho}_t$  obtained at various times  $t$ , using  $N$  records of the continuous fluorescence signals between time  $t$  and final time  $T$ . Figure 2 shows, for  $t$  between 0 and  $5 \mu\text{s}$ , the resulting estimates  $x_{\text{ML}}(t)$  and  $y_{\text{ML}}(t)$  (red solid lines) with their 95% confidence intervals (blue dashed lines). The black circle marks correspond to the average over a much larger set of  $N^* = 3 \times 10^6$  trajectories of the normalized signals,  $\sqrt{\frac{2T_1}{\eta}} \int_{t-dt}^t dy_1$  and  $\sqrt{\frac{2T_1}{\eta}} \int_{t-dt}^t dy_2$ . Since  $\langle dy_{\nu,t} \rangle = \sqrt{\eta_\nu} \text{Tr}[(L_\nu + L_\nu^\dagger)\rho_t] dt$ , these black circles are meant to provide reliable estimations of  $\text{Tr}(\bar{\rho}_t\sigma_x)$  and  $\text{Tr}(\bar{\rho}_t\sigma_y)$  at times  $t = 0, dt, 2dt, \dots, 25dt$ , where  $\bar{\rho}_t$  starts at  $\bar{\rho}_0$  and follows the Lindblad master equation corresponding to unread measurements. Thus the message is that our MaxLike method appears to obtain consistent reconstructions, as its confidence interval covers the very precise statistics (black circles) of a more standard model.

Figure 3 corresponds, for the same set of  $N$  measurement trajectories, to the MaxLike reconstruction of  $z$ . Contrary to  $x$  and  $y$ , this Bloch coordinate cannot be recovered directly by

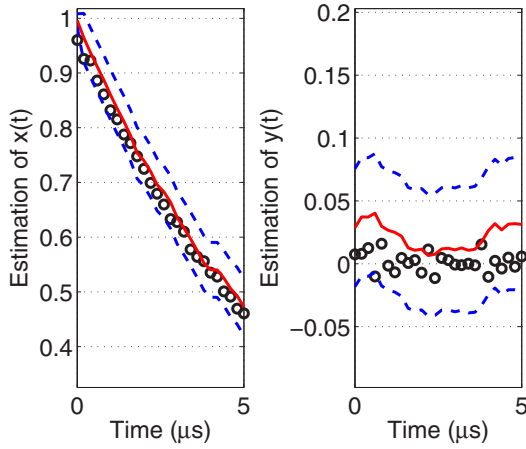


FIG. 2. Comparison of MaxLike reconstruction for  $x$  (left plot: solid and dashed curves) and  $y$  (right plot: solid and dashed curves) exploiting the measurement data set associated with  $N = 4 \times 10^4$  experimental trajectories, with an ensemble average of the normalized fluorescence signals (black circle) exploiting a much larger data set associated with  $N^* = 3 \times 10^6$  experimental trajectories. The red solid lines correspond to  $x_{ML}$  and  $y_{ML}$ , while the blue dashed lines correspond to the confidence intervals  $x_{ML} \pm 2\sigma_{ML}(\sigma_x)$  and  $y_{ML} \pm 2\sigma_{ML}(\sigma_y)$  obtained from (10). The black circles are generally admitted to represent reliable estimates of  $\text{Tr}(\bar{\rho}_t \sigma_x)$  and  $\text{Tr}(\bar{\rho}_t \sigma_y)$ ; see the text for more explanations.

an ensemble average of the measurement signal, so there are no black circles for validation.

In Fig. 4 we use the same set of  $N = 4 \times 10^4$  measurement records to compute the ensemble average of the normalized heterodyne signals (black circles). With this smaller dataset, the black circles do not provide a statistically accurate estimation of  $\text{Tr}(\bar{\rho}_t \sigma_x)$  and  $\text{Tr}(\bar{\rho}_t \sigma_y)$  anymore. On the contrary, the noise attached to such an ensemble average is large compared to MaxLike estimation of  $x$  and  $y$ . This shows that, as expected, our MaxLike estimation by exploiting all the data

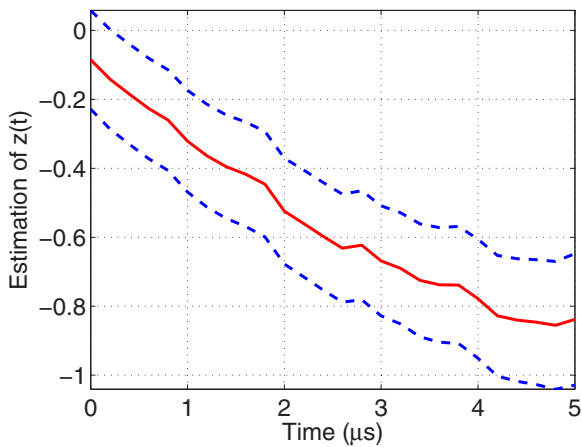


FIG. 3. MaxLike reconstruction for  $z$  exploiting a measurement data set with  $N = 4 \times 10^4$  experimental trajectories; the red solid line corresponds to  $z_{ML}$  and the blue dashed line to the confidence interval  $z_{ML} \pm 2\sigma_{ML}(\sigma_z)$  obtained from (10); see the text for more explanations.

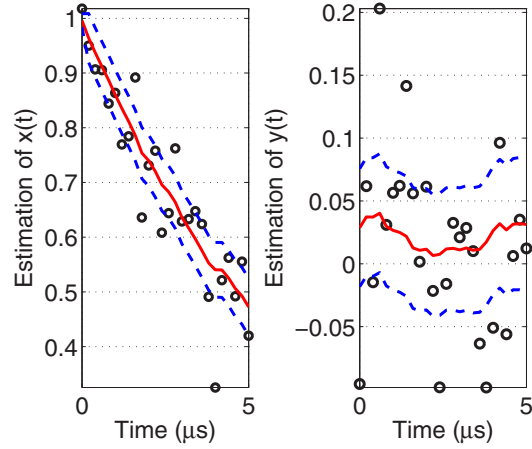


FIG. 4. Similar to Fig. 2, but the ensemble average of the normalized fluorescence signals (black circles) uses the same dataset as the MaxLike reconstruction. The MaxLike reconstruction appears to be significantly less noisy, illustrating the efficiency of our method; see the text for more explanations.

between  $t$  and  $T$  has superior statistical power compared to an estimation of  $\bar{\rho}_t$  from measurement outputs obtained at time  $t$  only.

#### IV. CONCLUDING REMARKS

We have proposed a method that takes into account imperfection and decoherence for quantum state tomography based on noninstantaneous measurements. This method is well adapted to MaxLike estimation since it provides directly the likelihood function. The latter is written in terms of sort of backpropagated POVM operators that depend on the measurement trajectory, and which characterize the information the trajectory of outputs carries on about the initial state to be estimated. We have given an approximation of the confidence interval to complete the MaxLike state estimate. In essence, given many copies of an initial quantum state, we can run it through any well-characterized device giving a sequence of outputs for each copy, and from this we can estimate efficiently what the initial state was, with an associated confidence interval. The estimator appears to rely mostly on measurement outputs obtained just after the state was prepared, much like a standard POVM; however, it can compensate for “missing statistical power” by relying on further evolution of the trajectory. This seems to guarantee robustness to uncertainties in the model of the dynamic measurement apparatus, as illustrated in Sec. II D.

The proposed method is directly applicable to reconstruct the state  $\rho_0$  from a standard projective measurement performed on  $\rho_t$  with  $t > 0$ . If on the measurement run  $n$  the final projective measurement at  $T_n$  yields the known state  $\rho_n$ , then the backward computations (4) just start with  $E_{T_n}^{(n)} = \rho_n$  instead of  $E_{T_n}^{(n)} = I/\text{Tr}(I)$ . If measurements are obtained as the state evolves from  $\rho_0$  to  $\rho_t$ , one just applies our standard filter with this different initial condition. If no measurements are obtained between 0 and  $t$ , e.g., the quantum state just undergoes decoherence, then one can obtain the corresponding filter by modeling the decoherence as hypothetical unread



measurements in the environment, and replacing the corresponding  $\mathbf{K}_{y,t}^*$  associated to read measurements by the unital linear map  $\mathbf{K}_t^* = \sum_y \mathbf{K}_{y,t}^*$ . In this case, the reconstructed MaxLike estimation will appropriately take into account that the POVM was performed not on  $\rho_0$ , but on a state  $\rho_t$  that has evolved from  $\rho_0$  according to some known dynamics.

It appears possible to extend the proposed method to measurement protocols in which a meter is coupled to the system of interest, and, after a small time, a (projective) measurement on the meter only is done. For example, it may be useful for the Wigner tomography of a quantum oscillator, such as those realized, e.g., in [26], to take into account higher-order Hamiltonian distortions and some decoherence effects during the joint evolution of system and meter.

### ACKNOWLEDGMENTS

This work was supported partially by the EMERGENCES grant QUMOTEL of Ville de Paris, by the IDEX program ANR-10-IDEX-0001-02 PSL\*, and by the ANR, Projet Blanc EMAQS ANR-2011-BS01-017-01.

### APPENDIX: ASYMPTOTIC EXPANSIONS UNDERLYING (10)

To formalize the fact that the likelihood function  $\rho \mapsto e^{f(\rho)}$  is concentrated around its maximum, we set  $f(\rho) = \Omega \bar{f}(\rho)$ ,  $\Omega > 0$  and large with  $\bar{f}(\rho)$  a normalized log-likelihood with bounded variations: for any density operators  $\rho_1$  and  $\rho_2$ ,  $|\bar{f}(\rho_1) - \bar{f}(\rho_2)| \leq 1$ . Throughout this appendix, we assume that  $\bar{f}(\rho)$  is maximal for a unique density matrix  $\rho_{\text{ML}}$  and that its Hessian matrix  $\nabla^2 \bar{f}_{\text{ML}}$  is negative definite. The dimension of the underlying finite-dimensional Hilbert space is denoted here by the integer  $n$ .

For any scalar function  $g$  of  $\rho$ , we consider thus the asymptotic expansion versus  $\Omega$  of the following multidimensional integral of Laplace type:

$$I_g(\Omega) = \int_{\mathcal{D}} g(\rho) e^{\Omega \bar{f}(\rho)} d\rho.$$

The domain of integration  $\mathcal{D}$  is a compact convex subset of the set of  $n \times n$  Hermitian matrices of trace 1. Here,  $d\rho$  stands for the standard Euclidian volume element on  $\mathcal{D}$ , derived from the Frobenius product between  $n \times n$  Hermitian matrices. We denote by  $\mathcal{D}_{n-1}$  the set of density matrices of rank less than or equal to  $n - 1$ . It corresponds to the boundary of  $\mathcal{D}$ . Here, we denote by  $\nabla \bar{f}$  and  $\nabla^2 \bar{f}$  the gradient and Hessian of  $\bar{f}$  that is considered to be a scalar function on the set of Hermitian matrices of unit-trace.

Assume that  $\rho_{\text{ML}}$  has full rank, i.e.,  $\bar{f}$  reaches its maximum in the interior of  $\mathcal{D}$  and  $\nabla \bar{f}_{\text{ML}} = 0$ . Then we can use Eq. 8.3.52 of Ref. [13] to get the following equivalent for  $I_g(\Omega)$  using  $\dim \mathcal{D} = n^2 - 1$ :

$$I_g(\Omega) = \frac{\left(\frac{2\pi}{\Omega}\right)^{\frac{n^2-1}{2}} e^{\Omega \bar{f}(\rho_{\text{ML}})}}{\sqrt{|\det(\nabla^2 \bar{f}_{\text{ML}})|}} \left[ g(\rho_{\text{ML}}) + O\left(\frac{1}{\Omega}\right) \right]. \quad (\text{A1})$$

Thus,  $I_g(\Omega) = I_1(\Omega)[g(\rho_{\text{ML}}) + O(\frac{1}{\Omega})]$ . With  $g(\rho) = \text{Tr}(\rho A) \mathcal{P}_0(\rho)$  for any operator  $A$ , we get, assuming

$$\mathcal{P}_0(\rho_{\text{ML}}) > 0,$$

$$\frac{\int_{\mathcal{D}} \text{Tr}(A\rho) e^{\Omega \bar{f}(\rho)} \mathcal{P}_0(\rho) d\rho}{\int_{\mathcal{D}} e^{\Omega \bar{f}(\rho)} \mathcal{P}_0(\rho) d\rho} = \text{Tr}(A\rho_{\text{ML}}) + O\left(\frac{1}{\Omega}\right). \quad (\text{A2})$$

If  $g(\rho_{\text{ML}}) = 0$ , then  $I_g(\Omega) = O(\frac{1}{\Omega})$  and we have to use the next term in the asymptotic expansion. It appears that, when  $g$  and its gradient vanish at  $\rho_{\text{ML}}$ , Eqs. 8.3.50 and 8.3.53 of Ref. [13] yield an explicit expression of this term with respect to  $\nabla^2 g_{\text{ML}}$ , the Hessian of  $g$  at  $\rho_{\text{ML}}$ . For  $g(\rho) = h(\rho) \text{Tr}[(\rho - \rho_{\text{ML}})A]$ , where  $h$  is any scalar function of  $\rho$  and  $A$  is any Hermitian matrix of zero-trace, we have  $\nabla^2 g_{\text{ML}}(B) = 2h(\rho_{\text{ML}}) \text{Tr}(AB)A$  for any zero-trace Hermitian matrix  $B$ . For such a special form of  $g$ , Eqs. 8.3.50 and 8.3.53 of [13] give

$$I_g(\Omega) = - \left[ \frac{h(\rho_{\text{ML}})}{\Omega} + O\left(\frac{1}{\Omega^2}\right) \right] \dots \left( \frac{\left(\frac{2\pi}{\Omega}\right)^{\frac{n^2-1}{2}} e^{\Omega \bar{f}(\rho_{\text{ML}})} \text{Tr}[A \nabla^2 \bar{f}_{\text{ML}}^{-1}(A)]}{\sqrt{|\det(\nabla^2 \bar{f}_{\text{ML}})|}} \right). \quad (\text{A3})$$

Thus with  $h = \mathcal{P}_0$ , we get

$$\frac{\int_{\mathcal{D}} \text{Tr}^2[A(\rho - \rho_{\text{ML}})] e^{\Omega \bar{f}(\rho)} \mathcal{P}_0(\rho) d\rho}{\int_{\mathcal{D}} e^{\Omega \bar{f}(\rho)} \mathcal{P}_0(\rho) d\rho} = - \frac{\text{Tr}[A \nabla^2 \bar{f}_{\text{ML}}^{-1}(A)]}{\Omega} + O\left(\frac{1}{\Omega^2}\right). \quad (\text{A4})$$

Notice that  $\Omega \nabla^2 \bar{f}$  corresponds to the Hessian of the restriction of  $f$  to  $\mathcal{D}$ . Some usual calculations show that this Hessian coincides with  $-\mathbf{R}$ , where  $\mathbf{R}$  is defined in (11) with  $\nabla f_{\text{ML}} - \lambda_{\text{ML}} I = 0$  and  $P_{\text{ML}} = I$ . This explains approximation (10) since  $A$  is of zero trace here and thus coincides with  $A_{\parallel}$  because  $\rho_{\text{ML}}$  has full rank.

When  $\rho_{\text{ML}}$  is rank-deficient, then  $\bar{f}$  reaches its maximum on the boundary of  $\mathcal{D}$  and the computations are more complicated. We only provide here the main steps for  $\rho_{\text{ML}}$  of rank  $n - 1$ . Moreover, we assume that  $\nabla \bar{f}_{\text{ML}} \neq 0$ . The other cases of rank between 1 and  $n - 2$  can be conducted in a similar way and will be detailed in a forthcoming publication. According to Eqs. 8.3.10 and 8.3.11 of Ref. [13], the first term of the asymptotic expansion of  $I_g(\Omega)$  coincides with the first term of the asymptotic expansion of a boundary integral localized on  $\mathcal{V}$ , an open small neighborhood of  $\rho_{\text{ML}}$  on  $\mathcal{D}_{n-1}$ :

$$J_g(\Omega) = \frac{1}{\Omega} \int_{\mathcal{V}} g(\rho) \frac{\text{Tr}[\nabla \bar{f}(\rho) N(\rho)]}{\text{Tr}[(\nabla \bar{f}(\rho))^2]} e^{\Omega \bar{f}(\rho)} d\Sigma,$$

where the Hermitian operator  $\nabla \bar{f}$  is the gradient of  $\bar{f}$ , the Hermitian operator  $N(\rho)$  corresponds to the unitary normal to  $\mathcal{D}_{n-1}$  at  $\rho$ , and where  $d\Sigma$  is the volume element on  $\mathcal{V}$  considered to be a Riemannian submanifold of the Euclidian space of Hermitian matrices equipped with the Frobenius product. With  $g_1(\rho) = g(\rho) \frac{\text{Tr}[\nabla \bar{f}(\rho) N(\rho)]}{\text{Tr}[(\nabla \bar{f}(\rho))^2]}$ , we have to evaluate the integral  $\int_{\mathcal{V}} g_1(\rho) e^{\Omega \bar{f}(\rho)} d\Sigma$ . Once again, we exploit Eq. 8.3.52 of Ref. [13] for this integral since  $\bar{f}$  restricted to  $\mathcal{V}$  reaches its maximum in the interior of  $\mathcal{V}$  with an invertible

Hessian. We get

$$J_g(\Omega) = \frac{1}{\Omega} \frac{K_{\text{ML}} e^{\Omega \bar{f}(\rho_{\text{ML}})}}{\Omega^{\frac{n^2-2}{2}}} \left[ g_1(\rho_{\text{ML}}) + O\left(\frac{1}{\Omega}\right) \right],$$

where  $K_{\text{ML}}$  is a positive constant, independent of  $g_1$ , and it can be expressed via the Hessian of  $\bar{f}$  restricted to the Riemannian submanifold  $\mathcal{V}$ . After some simple computations, we recover the asymptotic expansion (A2).

For  $g(\rho) = h(\rho) \text{Tr}^2[(\rho - \rho_{\text{ML}})A]$  and the corresponding  $g_1(\rho)$ , we exploit Eqs. 8.3.50 and 8.3.53 of Ref. [13]. We get the following analog to (A3):

$$J_g(\Omega) = - \left[ \frac{h_1(\rho_{\text{ML}})}{\Omega} + O\left(\frac{1}{\Omega^2}\right) \right] \dots \left( \frac{K_{\text{ML}} e^{\Omega \bar{f}(\rho_{\text{ML}})} \text{Tr}[\tilde{A} \widetilde{\nabla^2 \bar{f}}_{\text{ML}}^{-1}(\tilde{A})]}{\Omega^{\frac{n^2}{2}}} \right)$$

with  $h_1(\rho) = h(\rho) \frac{\text{Tr}[\widetilde{\nabla^2 \bar{f}}(\rho)N(\rho)]}{\text{Tr}[\widetilde{\nabla^2 \bar{f}}(\rho)^2]}$ . Here, the Hermitian matrix  $\tilde{A}$  is given by the Hessian at  $\rho_{\text{ML}}$  of the restriction of  $\rho \mapsto \text{Tr}^2[(\rho - \rho_{\text{ML}})A]$  to the Riemannian submanifold  $\mathcal{V}$ . Thus  $\tilde{A}$  is equal to the orthogonal projection onto the tangent space at  $\rho_{\text{ML}}$  to the submanifold  $\mathcal{V}$ . Similarly,  $\widetilde{\nabla^2 \bar{f}}_{\text{ML}}$  corresponds to the Hessian at  $\rho_{\text{ML}}$  of  $\bar{f}$  restricted to the submanifold  $\mathcal{V}$ . With the two previous asymptotic expansions for the boundary integral  $J_g(\Omega)$ , we recover (A4) with  $A_{\parallel}$  instead of  $A$  and  $\widetilde{\nabla^2 \bar{f}}_{\text{ML}}$  instead of  $\nabla^2 \bar{f}_{\text{ML}}$ . Some additional calculations show that  $\tilde{A}$  corresponds to  $A_{\parallel}$ , and  $\Omega \widetilde{\nabla^2 \bar{f}}_{\text{ML}}$  corresponds to the opposite of  $\mathbf{R}$  defined in (11).

---

[1] M. G. A. Paris and J. Rehacek, *Quantum State Estimation*, Lecture Notes in Physics (Springer, 2004).

[2] See also the lectures at Collège de France of Serge Haroche Cours 2009–2010: Synthèse et reconstructions d'états quantiques (in French), [http://www.college-de-france.fr/site/serge-haroche/p1346267508401\\_content.htm](http://www.college-de-france.fr/site/serge-haroche/p1346267508401_content.htm).

[3] V. Buzek, *Quantum Tomography from Incomplete Data via MaxEnt Principle*, Lecture Notes in Physics Vol. 649 (Springer-Verlag, Berlin, 2004), pp. 180–234.

[4] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, *Phys. Rev. Lett.* **105**, 150401 (2010).

[5] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, *Rev. Mod. Phys.* **81**, 299 (2009).

[6] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouiri, and P. Grangier, Generation of optical ‘Schrodinger cats’ from photon number states, *Nature (London)* **448**, 784 (2007).

[7] V. P. Belavkin, Quantum stochastic calculus and quantum nonlinear filtering, *J. Multivariate Anal.* **42**, 171 (1992).

[8] M. Tsang, Time-Symmetric Quantum Theory of Smoothing, *Phys. Rev. Lett.* **102**, 250403 (2009).

[9] I. Guevara and H. Wiseman, Quantum State Smoothing, *Phys. Rev. Lett.* **115**, 180407 (2015).

[10] S. Gammelmark, B. Julsgaard, and K. Mølmer, Past Quantum States of a Monitored System, *Phys. Rev. Lett.* **111**, 160401 (2013).

[11] R. Blume-Kohout, Optimal, reliable estimation of quantum states, *New J. Phys.* **12**, 043034 (2010).

[12] M. L. Mehta, *Random Matrices*, 3rd ed. (Elsevier, Amsterdam, 2004).

[13] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals* (Dover, New York, 1986).

[14] O. Cappé, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models*, Springer Series in Statistics (Springer, Berlin, 2005).

[15] T. Rybarczyk, B. Peaudecerf, M. Penasa, S. Gerlich, B. Julsgaard, K. Mølmer, S. Gleyzes, M. Brune, J. M. Raimond, S. Haroche, and I. Dotsenko, Forward-backward analysis of the photon-number evolution in a cavity, *Phys. Rev. A* **91**, 062116 (2015).

[16] P. Campagne-Ibarcq, L. Bretheau, E. Flurin, A. Auffèves, F. Mallet, and B. Huard, Observing Interferences between Past and Future Quantum States in Resonance Fluorescence, *Phys. Rev. Lett.* **112**, 180402 (2014).

[17] P. Six, Ph. Campagne-Ibarcq, L. Bretheau, B. Huard, and P. Rouchon, Parameter estimation from measurements along quantum trajectories, *Proceedings of the 54th Annual Conference on Decision and Control (CDC)* (IEEE, 2015), pp 7742–7748.

[18] I. Dotsenko, M. Mirrahimi, M. Brune, S. Haroche, J.-M. Raimond, and P. Rouchon, Quantum feedback by discrete quantum non-demolition measurements: Towards on-demand generation of photon-number states, *Phys. Rev. A* **80**, 013805 (2009).

[19] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).

[20] A. Somaraju, I. Dotsenko, C. Sayrin, and P. Rouchon, Design and stability of discrete-time quantum filters with measurement imperfections, in *American Control Conference (ACC)* (IEEE, Montreal, QC, 2012), pp. 5084–5089.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2009).

[22] A. Barchielli and M. Gregoratti, *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case* (Springer-Verlag, Berlin, 2009).

[23] P. Rouchon and J. F. Ralph, Efficient quantum filtering for quantum feedback control, *Phys. Rev. A* **91**, 012118 (2015).

[24] Ph. Campagne-Ibarcq, Quantum backaction and feedback in superconducting circuits, Ph.D. thesis, Ecole Normale Supérieure de Paris, 2015.

[25] P. Campagne-Ibarcq, P. Six, L. Bretheau, A. Sarlette, M. Mirrahimi, P. Rouchon, and B. Huard, Observing quantum state diffusion by heterodyne detection of fluorescence, *Phys. Rev. X* **6**, 011002 (2016).

[26] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, Confining the state of light to a quantum manifold by engineered two-photon loss, *Science* **347**, 853 (2015).