Quantum state tomography with non-instantaneous measurements, imperfections and decoherence

P. Six,[1] Ph. Campagne-Ibarcq,[2] I. Dotsenko,[3] A. Sarlette,[4] B. Huard,[2] and P. Rouchon[1, *]

[1]*Centre Automatique et Systèmes, Mines-ParisTech,*
PSL Reseach University, 60, bd Saint-Michel, 75006 Paris, France.
[2]*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
[3]*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
[4]*INRIA Paris-Rocquencourt, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*
(Dated: November 27, 2015)

Tomography of a quantum state is usually based on 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 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 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 multi-dimensional 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 non-demolition measurements relying on Rydberg atoms; heterodyne fluorescence measurements of a superconducting qubit.

INTRODUCTION

Determining efficiently the state of a system whose preparation is imperfectly known is instrumental to quantum physics experiments. Contrarily to classical physics, the determination of a quantum state $\overline{\rho}$, its tomography, requires a large number N of independent measurements [1][2]. The state of a quantum system is indeed a statistical quantity by 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 π_n such that $\sum_n \pi_n = I$. The probability of measurement outcome n is then given by $\mathrm{Tr}\,(\overline{\rho}\pi_n)$. For N large enough, the reconstruction of $\overline{\rho}$ is based on the fact that $\mathrm{Tr}\,(\overline{\rho}\pi_n)$ 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 estimate for $\overline{\rho}$ the value ρ_{ML} that

maximizes the likelihood function denoted by $\mathbb{P}\,(Y \mid \rho)$, the probability of the measurement data $Y \equiv (N_n)_n$ knowing ρ. For the POVM $(\pi_n)_n$, the likelihood function is directly given by:

$$\mathbb{P}\,(Y \mid \rho) = \prod_n \left(\mathrm{Tr}\,(\rho\pi_n) \right)^{N_n}. \tag{1}$$

In such 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 presence of imperfections and decoherence during the measurement has not been investigated. This is precisely one of the goals of this paper. Since the seminal work of Belavkin, one knows how to take into account measurement imperfection and decoherence for quantum filtering [7]. We show here how to exploit the stochastic master equation governing ρ_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 π_n are replaced by the adjoint states at the initial times of the N measurement sequences. These adjoint states obey a backward mas-

ter equation appearing in quantum smoothing [8, 9] and correspond to the effective operator E defining the past quantum state (ρ, E) introduced in [10].

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

$$\rho_{BM} = \frac{\int_{\mathcal{D}} \rho \, \mathbb{P}\left(Y \mid \rho\right) \mathbb{P}_0(\rho) d\rho}{\int_{\mathcal{D}} \mathbb{P}\left(Y \mid \rho\right) \mathbb{P}_0(\rho) d\rho} \qquad (2)$$

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

We show here how such asymptotic expansion provides also a confidence interval for $\mathrm{Tr}\left(\rho_{ML}A\right)$, where A is any Hermitian operator. Such confidence interval is based on a similar approximation for the Bayesian variance of $\mathrm{Tr}\left(\rho_{ML}A\right)$, denoted by $\sigma_{ML}^2(A)$. 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 ρ_{ML}. When ρ_{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 ρ_{ML} is rank deficient, the likelihood reaches its maximum on the boundary of $\mathcal{D}$. In this case, such explicit expression approximating the Bayesian variance is not usual and, as far as we know, seems to be new. From a practical viewpoint, our Max-Like 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 the experimental validation section.

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 non-demolition 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 ρ_{ML} of $\overline{\rho}$, 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 physically interesting observable A, the usual 95% confidence interval via the approximation $\mathrm{Tr}\left(\rho_{ML}A\right) \pm 2\sigma_{ML}(A)$. This confidence interval just means that, if we perform another tomography with another similar data-set, the probability that the MaxLike estimation of $\mathrm{Tr}\left(\overline{\rho}A\right)$ remains between $\mathrm{Tr}\left(\rho_{ML}A\right) - 2\sigma_{ML}(A)$ and $\mathrm{Tr}\left(\rho_{ML}A\right) + 2\sigma_{ML}(A)$ is greater than 95%. This probability has nothing to do with the quantum stochastic character of $\overline{\rho}$. Here, $\overline{\rho}$ 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 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 towards its ground state and submitted to the heterodyne measurement of its fluorescence. In appendix, we give the main calculations yielding the asymptotic expression of the Bayesian variance for any observable A.

DISCRETE-TIME SETTING

Quantum filtering

We have at our disposal N realizations, starting from the same initial state $\overline{\rho}$ that we want to determine, and producing N measurement records $(y_t^{(n)})_{t=0,\ldots,T_n}$, indexed by $n \in \{1, \ldots, N\}$ 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 $(y_0^{(n)}, \ldots, y_{t-1}^{(n)})$. This filter, a discrete-time version of Belavkin's continuous-time filter, is defined by a family of completely positive maps $\boldsymbol{K}_{y,t}$ indexed by the

time t and the measurement outcome y (see, e.g., [18–20]). Moreover, for each t, the completely positive map $\boldsymbol{K}_t = \sum_y \boldsymbol{K}_{y,t}$ is trace preserving. For each y and density operator ξ, $\mathrm{Tr}\left(\boldsymbol{K}_{y,t}(\xi)\right)$ is the probability to measure y at time t knowing that the quantum state at t is ξ. The quantum filter reads:

$$\rho_{t+1}^{(n)} = \frac{\boldsymbol{K}_{y_t^{(n)},t}\left(\rho_t^{(n)}\right)}{\mathrm{Tr}\left(\boldsymbol{K}_{y_t^{(n)},t}\left(\rho_t^{(n)}\right)\right)}, \ t = 0, \ldots, T_n. \quad (3)$$

A discrete-time quantum filter has this Markovian structure, with $\boldsymbol{K}_{y,t}$ depending only on its physical settings. As the initial state $\overline{\rho}$ 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 $(y_t^{(n)})_{t=0,\ldots,T_n}$, we can associate a number $\mathbb{P}_n(\rho)$ being the probability of getting this record, assuming the initial state was $\rho_0^{(n)} = \rho$. Since $\mathrm{Tr}\left(\boldsymbol{K}_{y_t^{(n)},t}\left(\rho_t^{(n)}\right)\right)$ is the probability of having detected $y_t^{(n)}$ knowing $\rho_t^{(n)}$, a direct use of Bayes law yields $\mathbb{P}_n(\rho) = \prod_{t=0}^{T_n} \mathrm{Tr}\left(\boldsymbol{K}_{y_t^{(n)},t}\left(\rho_t^{(n)}\right)\right)$ with $\rho_t^{(n)}$ generated thanks to (3). Some elementary computations show that:

$$\mathbb{P}_n(\rho) = \mathrm{Tr}\left(\boldsymbol{K}_{y_{T_n}^{(n)},T_n} \circ \ldots \circ \boldsymbol{K}_{y_0^{(n)},0}(\rho)\right).$$

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

$$Y = \left\{y_t^{(n)} \mid n \in \{1,\ldots,N\}, \ t \in \{0,\ldots,T_n\}\right\},$$

knowing the initial state ρ, reads:

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

Adjoint-state derivation of the likelihood function

The adjoint map $\boldsymbol{K}_{y,t}^*$ of $\boldsymbol{K}_{y,t}$ is defined by $\mathrm{Tr}\left(A\boldsymbol{K}_{y,t}(B)\right) \equiv \mathrm{Tr}\left(\boldsymbol{K}_{y,t}^*(A)B\right)$ for all Hermitian operators A and B. Thus:

$$\mathbb{P}_n(\rho) = \mathrm{Tr}\left(\rho \ \boldsymbol{K}_{y_0^{(n)},0}^* \circ \ldots \circ \boldsymbol{K}_{y_{T_n}^{(n)},T_n}^*(I)\right),$$

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:

$$E_t^{(n)} = \frac{\boldsymbol{K}_{y_t^{(n)},t}^*\left(E_{t+1}^{(n)}\right)}{\mathrm{Tr}\left(\boldsymbol{K}_{y_t^{(n)},t}^*\left(E_{t+1}^{(n)}\right)\right)} \text{ for } t = T_n, \ldots, 0. \quad (4)$$

It defines a family of Hermitian non-negative operators $(E_t^{(n)})$ of trace one and depending only on the measurement data Y. We have $\boldsymbol{K}_{y_0^{(n)},0}^* \circ \ldots \circ \boldsymbol{K}_{y_{T_n}^{(n)},T_n}^*(I) = c_n E_0^{(n)}$ with c_n depending only on $(y_0^{(n)},\ldots,y_{T_n}^{(n)})$. Thus $\mathbb{P}_n(\rho) = c_n \mathrm{Tr}\left(\rho E_0^{(n)}\right)$ and we have:

$$\mathbb{P}(Y \mid \rho) = \prod_{n=1}^{N} c_n \mathrm{Tr}\left(\rho E^{(n)}\right), \quad (5)$$

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

Quantum-state tomography

The maximum likelihood (MaxLike) estimate ρ_{ML} of the hidden initial quantum state $\overline{\rho}$ underlying the measurement data Y is given by maximizing the likelihood function $\mathbb{P}(Y \mid \rho)$. It is usual to consider the Log-likelihood function $f(\rho) = \log\left(\mathbb{P}(Y \mid \rho)\right)$:

$$f(\rho) = \sum_{n=1}^{N} \log c_n + \sum_{n=1}^{N} \log\left(\mathrm{Tr}\left(\rho E^{(n)}\right)\right)$$
$$= C + \sum_{n=1}^{N} \log\left(\mathrm{Tr}\left(\rho E^{(n)}\right)\right) \quad (6)$$

where C is a constant independent from ρ. Thus,

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

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 solved numerically efficiently (see, e.g., [21]). Moreover, ρ_{ML} is characterized by the following necessary and sufficient conditions: there exists a nonnegative scalar λ_{ML} such that:

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

where P_{ML} is the orthogonal projector on the range of ρ_{ML} and ∇f_{ML} is the gradient at ρ_{ML} of the log-likelihood:

$$\nabla f_{ML} = \sum_{n=1}^{N} \frac{E^{(n)}}{\mathrm{Tr}\left(\rho_{ML} E^{(n)}\right)}. \quad (9)$$

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

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

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

$$\langle A \rangle_{BM} = \frac{\int_{\mathcal{D}} \mathrm{Tr}\,(\rho A)\ \exp(f(\rho))\mathbb{P}_0(\rho)\ d\rho}{\int_{\mathcal{D}} \exp(f(\rho))\mathbb{P}_0(\rho)\ d\rho}$$

can be approximated by $\langle A \rangle_{BM} \approx \mathrm{Tr}\,(\rho_{ML}A)$. Similarly, we prove in appendix that its Bayesian variance:

$$\sigma^2_{BM}(A) = \frac{\int_{\mathcal{D}} \mathrm{Tr}^2\,((\rho - \rho_{ML})A)\ \exp(f(\rho))\mathbb{P}_0(\rho)\ d\rho}{\int_{\mathcal{D}} \exp(f(\rho))\mathbb{P}_0(\rho)\ d\rho},$$

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

$$\sigma^2_{BM}(A) \approx \sigma^2_{ML}(A) \equiv \mathrm{Tr}\,\left(A_{\parallel}\ \boldsymbol{R}^{-1}(A_{\parallel})\right) \qquad (10)$$

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

$$\boldsymbol{R}(X) = \sum_{n=1}^{N} \frac{\mathrm{Tr}\left(X E^{(n)}_{\parallel}\right)}{\mathrm{Tr}^2\left(\rho_{ML}E^{(n)}\right)}E^{(n)}_{\parallel}$$
$$+ \tfrac{1}{2}(\lambda_{ML}I - \nabla f_{ML})X\rho^+_{ML} + \tfrac{1}{2}\rho^+_{ML}X(\lambda_{ML}I - \nabla f_{ML}) \quad (11)$$

with λ_{ML} and P_{ML} appearing in (8) and ρ^+_{ML} the Moore-Penrose pseudo-inverse of ρ_{ML}. Notice that the super-operator $\boldsymbol{R}$ is symmetric and non-negative for the Frobenius product. Thus, $\sigma^2_{ML}(A)$ is always non-negative as it should be.

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

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

We recover, up to this orthogonal projection, the standard Cramér-Rao bounds attached to MaxLike estimation: $-\nabla^2 f_{ML}$ stands for the Fisher information matrix.

The super-operator $\boldsymbol{R}$ defined by (11) is the prolongation of such Fisher-information matrix when ρ_{ML} lies on the boundary of $\mathcal{D}$. Notice its dependence on boundary curvature due to the fact that $(\lambda_{ML}I - \nabla f_{ML})X\rho^+_{ML}$ does not vanish in general. Notice that the expressions of $\langle A \rangle_{BM}$ and $\sigma^2_{BM}(A)$ do not depend on the prior distribution of probability $\mathbb{P}_0(\rho)$. 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 make the initial information $\mathbb{P}_0(\rho)$ outdated. The computations underlying (11) are given in appendix: they rely on a specific application of asymptotic expansions for multi-dimensional Laplace integrals given in [13, chapter 8].

Experimental validation for QND photon counting

We apply in the following the MaxLike reconstruction to the state $\overline{\rho}$ of the light field stored in a cavity based on the experiment reported in [15] and, more precisely, experimental data associated to Fig. 4b therein. The field of a very high-quality superconducting cavity (frequency of 51 GHz, photon lifetime of $T_c = 65$ ms) is initially in a thermal state (temperature of 0.8 K, 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 μs. The measurement starts at $t = -172$ ms and has a duration of 344 ms, large compared to the photon lifetime. The main sources of measurement imperfections are random atomic occupation of samples, non-constant atom-photon interaction from sample to sample, non-ideal 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 ρ_t and E_t are diagonal in the Fock basis and truncated to a maximum of 7 photons: they are described by vectors of dimension 8. The partial Kraus maps $\boldsymbol{K}_{y,t}$ reduced to 8×8 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 $\boldsymbol{K}_{y,t}$: they can be deduced from [15]. For any t between -100 ms and $+150$ ms, we compute the MaxLike estimation $\rho_{ML}(t)$ of $\overline{\rho}(t)$, 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 $\overline{\rho}(t)$: 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. 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 ρ_{ML} to model errors. For any time t, the rank of $\rho_{ML}(t)$ 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 on figure 1. In order to evaluate the precision of such MaxLike estimation, we have computed the variance given in (10) for the photon-number operator $\boldsymbol{a}^\dagger \boldsymbol{a}$. 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 ± 0.015 and ± 0.065. We also compare the MaxLike reconstruction for t between -100 and $+150$ ms with the mean photon-number $\langle \boldsymbol{a}^\dagger \boldsymbol{a} \rangle_t = \mathrm{Tr}\left(\bar{\rho}(t)\boldsymbol{a}^\dagger \boldsymbol{a}\right)$ given by:

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

where $\langle \boldsymbol{a}^\dagger \boldsymbol{a} \rangle_0 = 1.26$, $T_c = 65$ ms and $n_b = 0.06$ are derived from [15]. We observe that $\langle \boldsymbol{a}^\dagger \boldsymbol{a} \rangle_t$ remains almost inside the 95%-confidence tube except for t between -50 ms and 0. This is a very positive result and incidentally, 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 ± 0.045. It corresponds to the programmed 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 programmed 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 ± 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.
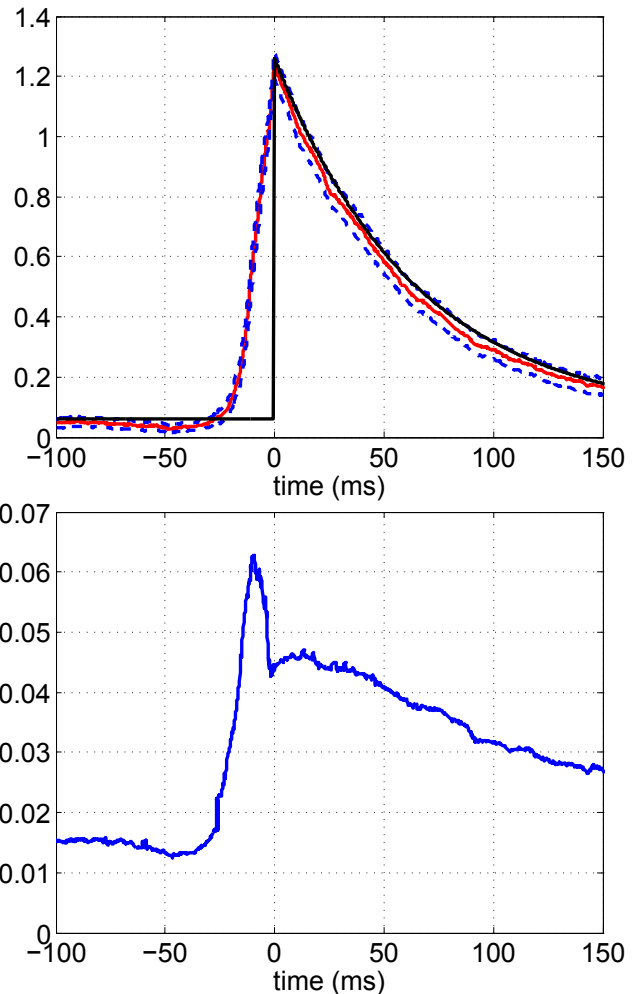

FIG. 1. (Color online) MaxLike quantum-state tomography based on the QND photon measurements reported in [15, figure 4b] for a photon creation quantum jump induced at $t = 0$. The latter is not taken into account in the measurement model. Top: solid red line corresponds to the MaxLike mean photon-number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

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.

Diffusive quantum filtering and adjoint-state

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

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

$$(13)$$

with $dW_{\nu,t} = dy_{\nu,t} - \sqrt{\eta_\nu}\mathrm{Tr}\left((L_\nu + L_\nu^\dagger)\rho_t \right) dt$ given by the measurement $y_{\nu,t}$, attached to the operator L_ν, with efficiency η_ν 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 Itō rules:

$$\rho_{t+dt} = \frac{\boldsymbol{K}_{dy_t}(\rho_t)}{\mathrm{Tr}\left(\boldsymbol{K}_{dy_t}(\rho_t) \right)},$$

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

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

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:

$$\mathbb{P}\left(dy_t \in \prod_\nu [\xi_\nu, \xi_\nu + d\xi_\nu] \mid \rho_t \right)$$
$$= \mathrm{Tr}\left(\boldsymbol{K}_\xi(\rho_t) \right) \prod_\nu e^{-\xi_\nu^2/2dt} \frac{d\xi_\nu}{\sqrt{2\pi dt}}.$$

Take a sampling-time dt much smaller than the time-constant involved in the Hamiltonian H and in the decoherence operator L_ν. 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 to the measurement data $(y_t^{(n)})_{0 \le t \le T_n}$ corresponding to the quantum trajectory number n, via the following discrete-time formulation:

$$E_t = \frac{\boldsymbol{K}_{dy_t^{(n)}}^*(E_{t+dt})}{\mathrm{Tr}\left(\boldsymbol{K}_{dy_t^{(n)}}^*(E_{t+dt}) \right)} \text{ with } E_{T_n+dt} = I/\mathrm{Tr}(I)$$

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

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

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 previous section.

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 ρ 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 σ_x, σ_y and σ_z the three Pauli matrices. Here the convex set $\mathcal{D}$ corresponds to the unit ball $x^2 + y^2 + z^2 \le 1$ and the Frobenius product between operators to the Euclidian product between vectors in the 3-dimensional Euclidian space. Then, the gradient of the log-likelihood function (9) becomes the vector:

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

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

- when $x_{ML}^2 + y_{ML}^2 + z_{ML}^2 < 1$, we have:

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

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 to 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.

- When $x_{ML}^2 + y_{ML}^2 + z_{ML}^2 = 1$ and $\lambda_{ML} > 0$ large enough, we have:

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

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

$$A_\| = \frac{(a - s x_{ML})\sigma_x + (b - s y_{ML})\sigma_y + (c - s z_{ML})\sigma_z}{2}$$

with $s = a x_{ML} + b y_{ML} + c z_{ML}$.

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 $T_1 = 4.15$ μs and $T_\phi = 35$ μs.

We have at our disposal $N = 4.10^4$ quantum trajectories with the same length $T_n = T = 9.2$ μs with a sampling-time $dt = 200$ ns $\approx \frac{1}{20}T_1$. Each trajectory is supposed to start at time $t = 0$ from the same initial state $\overline{\rho}_0$ 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 $\overline{\rho}_0$ based on records of the continuous fluorescence signals between 0 and final time T. For the trajectory number n, the measurement record corresponds to 2×47 real

values corresponding to $\int_{t-dt}^{t} dy_1^{(n)}$ and $\int_{t-dt}^{t} dy_2^{(n)}$ for $t = dt, 2dt, \ldots, 47dt$. From the measurements between 0 and T, we get an estimation of the quantum state $\overline{\rho}_0$ with a 95% confidence interval using the above formula for $\sigma_{ML}^2(A)$ and $A = \sigma_x, \sigma_y$ and σ_z:

$$\mathrm{Tr}\left(\overline{\rho}_0 \sigma_x\right) = 0.99 \pm 0.06$$
$$\mathrm{Tr}\left(\overline{\rho}_0 \sigma_y\right) = -0.03 \pm 0.07$$
$$\mathrm{Tr}\left(\overline{\rho}_0 \sigma_z\right) = -0.10 \pm 0.19$$

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 $\overline{\rho}_t$ 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 μs, the resulting estimates $x_{ML}(t)$ and $y_{ML}(t)$ (red solid lines) with their 95% confidence intervals (blue dashed lines). The black circle marks correspond to average over a much larger set of $N^* = 3.10^6$ 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 $\langle dy_{\nu,t} \rangle = \sqrt{\eta_\nu} \mathrm{Tr}\left((L_\nu + L_\nu^\dagger)\rho_t\right) dt$, these black circles are meant to provide reliable estimations of $\mathrm{Tr}\left(\overline{\rho}_t \sigma_x\right)$ and $\mathrm{Tr}\left(\overline{\rho}_t \sigma_y\right)$ at times $t = 0, dt, 2dt, \ldots, 25dt$, where $\overline{\rho}_t$ starts at $\overline{\rho}_0$ 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. Contrarily to x and y, this Bloch coordinate cannot be recovered directly by an ensemble average of the measurement signal, so there are no black circles for validation.

On Figure 4 we use the same set of $N = 4.10^4$ 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 $\mathrm{Tr}\left(\overline{\rho}_t \sigma_x\right)$ and $\mathrm{Tr}\left(\overline{\rho}_t \sigma_y\right)$ anymore. To the contrary, the noise attached to such 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 between t and T has superior statistical power, compared to an estimation of $\overline{\rho}_t$ from measurement outputs obtained at time t only.

CONCLUDING REMARKS

We have proposed a new method to take into account imperfection and decoherence for quantum-state tomography based on non-instantaneous measurements. This
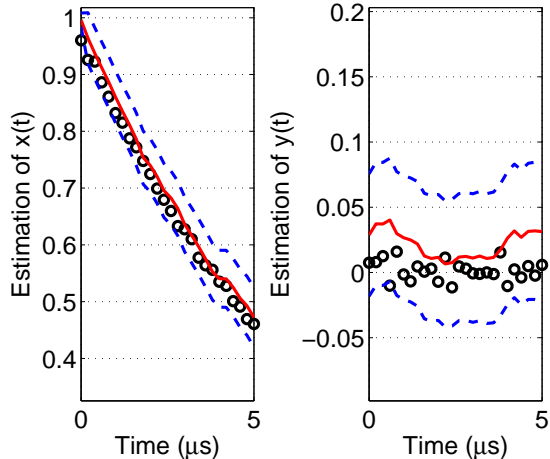
FIG. 2. (Color online) Comparison of MaxLike reconstruction for x (left plot: solid and dashed curves) and y (right plot: solid and dashed curves) exploiting the measurement data set associated to $N = 4.10^4$ experimental trajectories; with an ensemble average of the normalized fluorescence signals (black circle) exploiting a much larger data set associated to $N^* = 3.10^6$ experimental trajectories; the red solid lines correspond to x_{ML} and y_{ML}; the blue dashed lines 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 $\mathrm{Tr}\,(\overline{\rho}_t \sigma_x)$ and $\mathrm{Tr}\,(\overline{\rho}_t \sigma_y)$; see text for more explanations.

method is well adapted to MaxLike estimation since it provides directly the likelihood function. The latter is written in terms of a sort of back-propagated POVM operators which 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 mostly rely 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 our experimental validation.

The proposed method is directly applicable to reconstruct the state ρ_0 from a standard projective measurement performed on ρ_t with $t > 0$. If on the measurement run n the final projective measurement at T_n yields the known state ρ_n, then the backward computations (4) just start with $E_{T_n}^{(n)} = \rho_n$ instead of $E_{T_n}^{(n)} = I/\mathrm{Tr}\,(I)$. If measurements are obtained between as the state evolves from
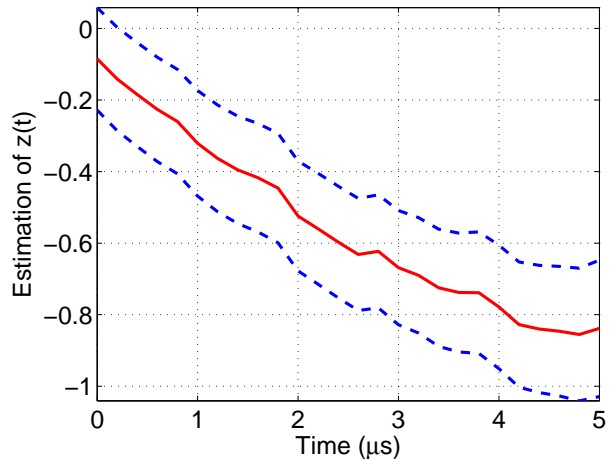

FIG. 3. (Color online) MaxLike reconstruction for z exploiting a measurement data set with $N = 4.10^4$ experimental trajectories; the red solid line corresponds to z_{ML} and the blue dashed one to the confidence interval $z_{ML} \pm 2\sigma_{ML}(\sigma_z)$ obtained from (10); see text for more explanations.
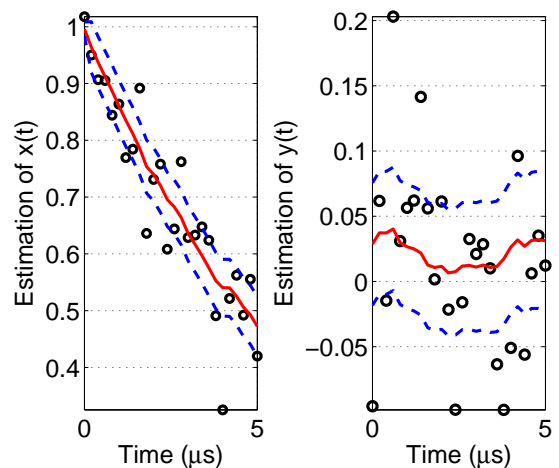

FIG. 4. (Color online) Similar to figure 2 but the ensemble average of the normalized fluorescence signals (black circles) use the same dataset as the MaxLike reconstruction. The MaxLike reconstruction appears to be significantly less noisy, illustrating the efficiency of our method; see text for more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

ρ_0 to ρ_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 $\boldsymbol{K}_{y_t,t}^*$ associated to read measurements by the unital linear map $\boldsymbol{K}_t^* = \sum_y \boldsymbol{K}_{y,t}^*$. In this case, the reconstructed MaxLike estimation will appropriately take into account that the POVM was performed not on ρ_0,

but on a state ρ_t that has evolved from ρ_0 according to some known dynamics.

Extension of the proposed method to measurement protocols where a meter is coupled to the system of interest and, after a small time a (projective) measurement on the meter only is done, appears to be possible. For example, it could 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.

———————

* pierre.rouchon@mines-paristech.fr

[1] M.G.A Paris and J. Rehacek. Quantum State Estimation. 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).

[3] V. Buzek. Lecture Notes in Physics, volume 649, chapter Quantum tomography from incomplete data via MaxEnt principle, pages 180–234. Springer-Verlag Berlin, 2004.

[4] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. Phys. Rev. Lett., 105(15):150401–, October 2010.

[5] A. I. Lvovsky and M. G. Raymer. Continuous-variable optical quantum-state tomography. Rev. Mod. Phys., 81(1):299–332, March 2009.

[6] Alexei Ourjoumtsev, Hyunseok Jeong, Rosa Tualle-Brouri, and Philippe Grangier. Generation of optical /'schrodinger cats/' from photon number states. Nature, 448(7155):784–786, August 2007.

[7] V.P. Belavkin. Quantum stochastic calculus and quantum nonlinear filtering. Journal of Multivariate Analysis, 42(2):171–201, 1992.

[8] Mankei Tsang. Time-symmetric quantum theory of smoothing. Phys. Rev. Lett., 102(25):250403–, June 2009.

[9] I. Guevara and H. Wiseman. Quantum state smoothing. arXiv:1503.02799 [quant-ph], 2015.

[10] S. Gammelmark, B. Julsgaard, and K. Mølmer. Past quantum states of a monitored system. Phys. Rev. Lett., 111(16):160401–, October 2013.

[11] Robin Blume-Kohout. Optimal, reliable estimation of quantum states. New Journal of Physics, 12(4):043034–, 2010.

[12] M. L. Mehta. Random matrices (3rd ed). Elsevier, Academic Press, 2004.

[13] N. Bleistein and R.A. Handelsman. Asymptotic Expansions of Integrals. Dover, 1986.

[14] O. Cappé, E. Moulines, and T. Ryden. Inference in Hidden Markov Models. Springer series in statistics, 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, Jun 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, May 2014.

[17] Ph. Campagne-Ibarcq Six, L. Bretheau, B. Huard, and P. Rouchon. Parameter estimation from measurements along quantum trajectories. CDC 2015, arXiv:1503.06149 [math.OC], 2015.

[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. Physical Review A, 80: 013805-013813, 2009.

[19] H.M. Wiseman and G.J. Milburn. Quantum Measurement and Control. Cambridge University Press, 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, pages 5084–5089, 2012.

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

[22] A. Barchielli and M. Gregoratti. Quantum Trajectories and Measurements in Continuous Time: the Diffusive Case. Springer Verlag, 2009.

[23] P. Rouchon and J. F. Ralph. Efficient quantum filtering for quantum feedback control. Phys. Rev. A, 91(1):012118–, January 2015.

[24] Ph. Campagne-Ibarcq. Quantum backaction and feedback in superconducting circuits. PhD 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. submitted. http://arxiv.org/abs/1511.01415, 2015.

[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(6224):853–857, February 2015.

Asymptotic expansions underlying (10)

To formalize the fact that the likelihood function $\rho \mapsto e^{f(\rho)}$ is concentrated around its maximum, we set $f(\rho) = \Omega \overline{f}(\rho)$, $\Omega > 0$ and large with $\overline{f}(\rho)$ a normalized log-likelihood with bounded variations: for any density operators ρ_1 and ρ_2, $|\overline{f}(\rho_1) - \overline{f}(\rho_2)| \leq 1$. Throughout this appendix, we assume that $\overline{f}(\rho)$ is maximal for a unique density matrix ρ_{ML} and that its Hessian matrix $\nabla^2 \overline{f}_{ML}$ 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 ρ, we consider thus the asymptotic expansion versus Ω of the following multi-dimensional integral of Laplace type:

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

The domain of integration $\mathcal{D}$ is a compact convex subset of the set of $n \times n$ Hermitian matrices of trace one. 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 or equal to $n-1$. It corresponds to the boundary of $\mathcal{D}$. Here, we denote by $\nabla\overline{f}$ and $\nabla^2\overline{f}$ the gradient and Hessian of $\overline{f}$ that is considered as a scalar function on the set of Hermitian matrices of unit-trace.

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

$$I_g(\Omega) = \frac{\left(\frac{2\pi}{\Omega}\right)^{\frac{n^2-1}{2}} e^{\Omega\overline{f}(\rho_{ML})}}{\sqrt{\left|\det\left(\nabla^2\overline{f}_{ML}\right)\right|}}\left(g(\rho_{ML}) + O\left(\frac{1}{\Omega}\right)\right) \quad (14)$$

Thus, $I_g(\Omega) = I_1(\Omega)\left(g(\rho_{ML}) + O\left(\frac{1}{\Omega}\right)\right)$. With $g(\rho) = \text{Tr}\left(\rho A\right)\mathbb{P}_0(\rho)$ for any operator A, we get, assuming $\mathbb{P}_0(\rho_{ML}) > 0$:

$$\frac{\int_{\mathcal{D}} \text{Tr}(A\rho)e^{\Omega\overline{f}(\rho)}\mathbb{P}_0(\rho)d\rho}{\int_{\mathcal{D}} e^{\Omega\overline{f}(\rho)}\mathbb{P}_0(\rho)d\rho} = \text{Tr}\left(A\rho_{ML}\right) + O\left(\frac{1}{\Omega}\right). \quad (15)$$

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

$$I_g(\Omega) = -\left(\frac{h(\rho_{ML})}{\Omega} + O\left(\frac{1}{\Omega^2}\right)\right)\cdots$$
$$\cdots\left(\frac{\left(\frac{2\pi}{\Omega}\right)^{\frac{n^2-1}{2}} e^{\Omega\overline{f}(\rho_{ML})}\text{Tr}\left(A\ \nabla^2\overline{f}_{ML}^{-1}(A)\right)}{\sqrt{\left|\det\left(\nabla^2\overline{f}_{ML}\right)\right|}}\right). \quad (16)$$

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

$$\frac{\int_{\mathcal{D}} \text{Tr}^2(A(\rho - \rho_{ML}))e^{\Omega\overline{f}(\rho)}\mathbb{P}_0(\rho)d\rho}{\int_{\mathcal{D}} e^{\Omega\overline{f}(\rho)}\mathbb{P}_0(\rho)d\rho}$$
$$= -\frac{\text{Tr}\left(A\ \nabla^2\overline{f}_{ML}^{-1}(A)\right)}{\Omega} + O\left(\frac{1}{\Omega^2}\right). \quad (17)$$

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

When ρ_{ML} is rank deficient, then $\overline{f}$ reaches its maximum on the boundary of $\mathcal{D}$ and the computations are more complicated. We only provide here the main steps for ρ_{ML} of rank $n-1$. Moreover, we assume that $\nabla\overline{f}_{ML} \neq 0$. 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 [13, eq. 8.3.10 and 8.3.11], 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 neighbourhood of ρ_{ML} on $\mathcal{D}_{n-1}$:

$$J_g(\Omega) = \frac{1}{\Omega}\int_{\mathcal{V}} g(\rho)\frac{\text{Tr}\left(\nabla\overline{f}(\rho)N(\rho)\right)}{\text{Tr}\left(\left(\nabla\overline{f}(\rho)\right)^2\right)}e^{\Omega\overline{f}(\rho)}d\Sigma$$

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

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

where K_{ML} is a positive constant, independent of g_1 and that can be expressed via the Hessian of $\overline{f}$ restricted to the Riemannian submanifold $\mathcal{V}$. After some simple computations, we recover the asymptotic expansion (15).

For $g(\rho) = h(\rho)\text{Tr}^2\left((\rho - \rho_{ML})A\right)$ and the corresponding $g_1(\rho)$, we exploit [13, eq. 8.3.50 and 8.3.53]. We get the following analogue to (16):

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

with $h_1(\rho) = h(\rho)\frac{\text{Tr}\left(\nabla\overline{f}(\rho)N(\rho)\right)}{\text{Tr}\left(\left(\nabla\overline{f}(\rho)\right)^2\right)}$. Here, the Hermitian matrix $\widetilde{A}$ is given by the Hessian at ρ_{ML} of the restriction of $\rho \mapsto \text{Tr}^2\left((\rho - \rho_{ML})A\right)$ to the Riemannian submanifold $\mathcal{V}$. Thus $\widetilde{A}$ is equal to the orthogonal projection onto the

tangent space at ρ_{ML} to the submanifold $\mathcal{V}$. Similarly, $\widetilde{\nabla^2 \overline{f}}_{ML}$ corresponds to the Hessian at ρ_{ML} of $\overline{f}$ restricted to the submanifold $\mathcal{V}$. With the two previous asymptotic expansions for the boundary integral $J_g(\Omega)$, we recover (17) with $A_\parallel$ instead of A and $\widetilde{\nabla^2 \overline{f}}_{ML}$ instead of $\nabla^2 \overline{f}_{ML}$. Some additional calculations show that $\widetilde{A}$ corresponds to $A_\parallel$ and $\Omega \widetilde{\nabla^2 \overline{f}}_{ML}$ to the opposite of $\boldsymbol{R}$ defined in (11).