Quantum State Smoothing

Ivonne Guevara and Howard Wiseman
Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia
(Received 13 March 2015; revised manuscript received 28 July 2015; published 30 October 2015)

Smoothing is an estimation method whereby a classical state (probability distribution for classical variables) at a given time is conditioned on all-time (both earlier and later) observations. Here we define a smoothed quantum state for a partially monitored open quantum system, conditioned on an all-time monitoring-derived record. We calculate the smoothed distribution for a hypothetical unobserved record which, when added to the real record, would complete the monitoring, yielding a pure-state "quantum trajectory." Averaging the pure state over this smoothed distribution yields the (mixed) smoothed quantum state. We study how the choice of actual unraveling affects the purity increase over that of the conventional (filtered) state conditioned only on the past record.

DOI: 10.1103/PhysRevLett.115.180407 PACS numbers: 03.65.Ta, 03.65.Yz, 42.50.Dv, 42.50.Lc

Estimation theory is used to assign values to parameters of interest, whose true values are unknown, using the available data. These parameters may evolve dynamically, and new data may arrive dynamically, through a continuous measurement process. Estimation in this instance is nontrivial because there can be noise associated with the measurement, noise affecting directly the dynamical system due to its environment, and initial uncertainty in the parameters. Optimal estimation theory can be formulated using the Bayesian approach to statistics, whereby the observer's knowledge of the parameters is described by a conditional probability distribution $\wp_C$, also known as a Bayesian state. If this state is conditional on measurements at earlier times, it is called a *filtered* state $\wp_F$, while if it is conditional on all-time (both earlier and later) measurements, it is called a *smoothed* state $\wp_S$. Smoothing uses more complete information than filtering, and so typically delivers a probability distribution that is purer (that is, having less entropy).

In the flourishing and important area of *quantum* estimation theory [1–12], much has been learnt from classical estimation theory. The analogy between quantum states and classical Bayesian states has been fruitful even in quantum foundations [13–15]. In particular, the stochastically evolving conditioned state ρ_C of an open quantum state, as introduced by physicists [16–21] and applied in quantum control [1,22–29], is now understood to be analogous to the classical filtered state [25,30–34] $\wp_F$ (and so for clarity we write it as ρ_F). However, the situation is very different regarding smoothing.

The term "quantum smoothing" was introduced by Tsang [35,36] in 2009 to mean smoothed estimation of classical parameters that affect the evolution of a quantum system, using the results of measurements on that system. It has been shown to be useful to the problem of estimating a stochastically varying optical phase using the all-time photocurrent record, both theoretically [37,38] and experimentally [39,40]. It has also been applied to the problem of estimating an unknown result from a measurement on a quantum system at one time, using records obtained both before and after that time, again both theoretically [41] and experimentally [42]. However, none of the above define a quantum smoothed state—that is, a positive operator ρ_S that is analogous to a Bayesian smoothed state $\wp_S$. There is a good reason for this lack, which is most easily stated in the Heisenberg picture [35]: quantum operators for a system at a given time commute with operators representing the results of earlier measurements on that system, but do not commute with operators representing the results of later measurements on that system [1].

In this Letter we show that there is a situation in which it *is* possible to define quantum state smoothing, producing a positive state ρ_S conditioned on both earlier and later results. The situation is that of open quantum systems under partial observation, which is the typical situation experimentally [26,28,29]. The system has couplings to several baths (all assumed Markovian). An experimenter, Alice, is able to monitor some of them, yielding the record $\mathbf{O}$ she observes. Other baths are not monitored by her, but hypothetically they could be monitored by another party, yielding results $\mathbf{U}$ unobserved by Alice. The "true" state ρ_T conditioned on both observed ($\mathbf{O}$) and unobserved ($\mathbf{U}$) records would be pure, while that conditioned only on $\mathbf{O}$ is mixed. The crucial point is that the record $\mathbf{U}$, comprising classical variables (c numbers), can be estimated by applying smoothing to the record $\mathbf{O}$, and in this way Alice can obtain a smoothed quantum state ρ_S. As in the classical case, this is typically purer than ρ_F, and a better approximation to ρ_T.

We first review the necessary theoretical background. Then we explain how our smoothed quantum state ρ_S is quite different from the "past quantum state" (actually a pair of operators) introduced in Ref. [41]. We also show that our approach subsumes classical state estimation by smoothing, also known as the hidden Markov model (HMM) technique [43]. A HMM is applicable to quantum systems only when they are effectively classical (i.e., always diagonal in a particular basis) as in Refs. [44–46]. We apply our method to a genuinely quantum system (i.e., one which is not diagonal in a fixed basis)—a coherently driven two-level atom, the radiation from which is partly observed. We take the known record to be generated by homodyne detection, and the unknown record to be that corresponding to photon absorption, as this is the most intuitive picture of what happens to photons that are lost into the laboratory surroundings. These lost photons result in impurity in the standard (filtered) conditioned system state. We show that our smoothing technique can, on average, eliminate up to 26% of this impurity. Our investigations shed light on how well we can know the trajectory of a partially observed open quantum systems, and the relation between the quantum and classical versions of state smoothing.

Types of estimation.—Consider a classical dynamical system described by parameters $\mathbf{x}_t$ (bold font indicates a vector of parameters) which is monitored to yield a noisy output at each time $\mathbf{r}_t$. We denote a *measurement record* $\mathbf{R}_\Omega = \{\mathbf{r}_t : t \in \Omega\}$, where $\Omega \subseteq [t_0, T]$ is typically some finite time interval. Bayesian estimation involves data processing to infer the conditional classical state

$$\wp_{\mathbf{R}_\Omega}(\mathbf{x}_\tau) \equiv \Pr[\mathbf{x}_\tau^{\rm true} = \mathbf{x}_\tau | \mathbf{R}_\Omega; \wp_0], \tag{1}$$

where $\wp_0$ describes the *a priori* statistics of $\mathbf{x}$ at the initial time t_0. It is also useful to define the unnormalized state

$$\tilde{\wp}_{\mathbf{R}_\Omega}(\mathbf{x}_\tau) \equiv \wp_{\mathbf{R}_\Omega}(\mathbf{x}_\tau)\frac{\wp(\mathbf{R}_\Omega|\wp_0)}{\wp_{\rm ost}(\mathbf{R}_\Omega|\wp_0)} \propto \wp(\mathbf{R}_\Omega, \mathbf{x}_\tau|\wp_0). \tag{2}$$

Here, the $\wp(\mathbf{R}_\Omega|\wp_0)$ is the actual distribution for $\mathbf{R}_\Omega$, while $\wp_{\rm ost}(\mathbf{R}_\Omega|\wp_0)$ is an "ostensible" distribution for it—it is positive and normalized, but is otherwise arbitrary and does not depend on $\mathbf{x}_t$ [47].

There are three types of estimation worth distinguishing [48,49]: *filtering*, *retrofiltering* (as we call it), and *smoothing* (see Fig. 1). If—as in feedback control problems—for the time of interest τ there is only access to earlier results, $\overleftarrow{\mathbf{R}}_\tau \equiv \mathbf{R}_{[t_0,\tau)}$, the optimal protocol is filtering: $\wp_F(\mathbf{x}_\tau) \equiv \wp_{\overleftarrow{\mathbf{R}}_\tau}(\mathbf{x}_\tau)$. If there is access only to later results, $\vec{\mathbf{R}}_\tau \equiv \mathbf{R}_{[\tau,T)}$,

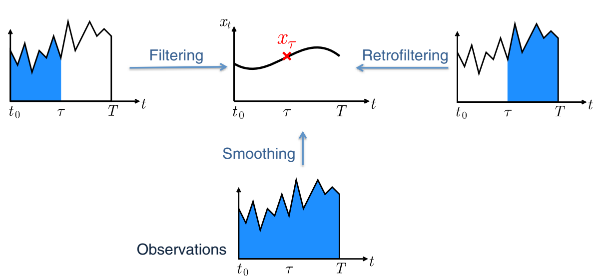

FIG. 1 (color online). Classical estimation classes depending on the measurement record considered relative to τ, time at which the signal is to be estimated. (Adapted from Ref. [35]).

the optimal protocol is retrofiltering: $\wp_R(\mathbf{x}_\tau) \equiv \wp_{\vec{\mathbf{R}}_\tau}(\mathbf{x}_\tau)$. As its name implies, this is simply the time reverse to filtering, but starting with an uninformative final state $\wp(\mathbf{x}_T) \propto 1$. Finally, if the all-time record $\overleftrightarrow{\mathbf{R}} \equiv \mathbf{R}_{[t_0,T)}$ is available, with $t_0 < \tau < T$, then all the information can be utilized by the technique of smoothing: $\wp_S(\mathbf{x}_\tau) = \wp_{\overleftrightarrow{\mathbf{R}}}(\mathbf{x}_\tau)$. This combines filtering and retrofiltering, using unnormalized states [36]:

$$\wp_S(\mathbf{x}_\tau) = \frac{\tilde{\wp}_R(\mathbf{x}_\tau)\tilde{\wp}_F(\mathbf{x}_\tau)}{\int d\mathbf{x}'_\tau \tilde{\wp}_R(\mathbf{x}'_\tau)\tilde{\wp}_F(\mathbf{x}'_\tau)}. \tag{3}$$

Here, one of the states (most conveniently the retrofiltered one) is defined using an uninformative prior, $\wp_0 \propto 1$, to prevent double counting of the *a priori* information.

Quantum analogues of (retro)filtering.—An extension of these results to quantum mechanics has been done partially. Quantum trajectory theory [16] is the analogue of classical state filtering. A quantum trajectory describes the path (in "density operator space") of the state of the quantum system through time, conditioned on the measurement result $\mathbf{r}_t$ in each infinitesimal interval $[t, t+dt)$. Note that the path may be continuous (quantum diffusion) or discontinuous (quantum jumps) [1,16]. This process is described by a set (indexed by $\mathbf{r}_t$) of measurement operations (completely positive maps) $\mathcal{M}_{\mathbf{r}_t}$, that evolve the state forward in time [47]:

$$\tilde{\rho}_F(t+dt) = \mathcal{M}_{\mathbf{r}_t}\tilde{\rho}_F(t). \tag{4}$$

Starting with $\rho(t_0) = \rho_0$, this procedure generates the state conditioned on the whole past record: $\tilde{\rho}_F(\tau) = \tilde{\rho}_{\overleftarrow{\mathbf{R}}_\tau}(\tau)$. This is an unnormalized state (as indicated by the tilde), analogous to Eq. (2). That is, the normalized version $\rho_F(\tau)$ generates the correct filtered probability distribution $\wp_F(x_\tau)$ for any system observable $\hat{X}_\tau$, while

$$\mathrm{Tr}[\tilde{\rho}_F(\tau)]\wp_{\rm ost}(\overleftarrow{\mathbf{R}}_\tau|\rho_0) = \wp(\overleftarrow{\mathbf{R}}_\tau|\rho_0). \tag{5}$$

The corresponding analogue for Bayesian state retrofiltering has been set out in Ref. [35]; it is the solution of the adjoint of Eq. (4),

$$\hat{E}_R(t) = \mathcal{M}^\dagger_{\mathbf{r}_t}\hat{E}_R(t+dt). \tag{6}$$

In this case the *effect operator* evolves backwards from the final uninformative effect $\hat{E}(T) = I$ towards $\hat{E}_R(\tau) \equiv \hat{E}_{\vec{\mathbf{R}}_\tau}(\tau)$, conditioned on the record $\vec{\mathbf{R}}_\tau$ in the future of τ. This solution $\hat{E}_{\vec{\mathbf{R}}_\tau}(\tau)$ determines the statistics of $\vec{\mathbf{R}}_\tau$:

$$\mathrm{Tr}[\hat{E}_R(\tau)\rho_\tau]\wp_{\rm ost}(\vec{\mathbf{R}}_\tau|\rho_\tau) = \wp(\vec{\mathbf{R}}_\tau|\rho_\tau). \tag{7}$$

Quantum smoothing?—A naive approach to construct a quantum smoothed state, given the quantum analogues of filtering ρ_F and retrofiltering $\hat{E}_R$, would be to combine them directly as in Eq. (3) so that $\rho_S(\tau) \propto \rho_F(\tau)\hat{E}_R(\tau)$. However, as pointed out in Ref. [35], the result is not in general Hermitian or (even if symmetrized) positive semidefinite.

Therefore, it cannot correspond to a physical state. As discussed in the introduction, there is a deep reason for this, which is why Tsang gave quantum smoothing the restricted meaning of estimating an external *classical* parameter $\mathbf{x}$.

The filtered state $\rho_F(\tau)$ and the retrofiltered effect $\hat{E}_R(\tau)$ are sufficient to best estimate, from the all-time record $\overleftrightarrow{\mathbf{R}}$, the result of any measurement performed on the system at time τ. For this reason, Gammelmark *et al.* declared the pair $\Xi = (\rho_F, \hat{E}_R)$ to be "the past quantum state" [41]. They did not define a quantum state in the usual sense (i.e., a density operator) combining the future and past information. While our notion of quantum state smoothing also makes use of filtered states and retrofiltered effects, it is quite different in that: it applies only to partially observed open quantum systems; it defines a quantum state in the usual sense; and it can be compared directly to quantum state filtering by measures such as purity, similarly to the classical case.

Quantum state smoothing.—The key idea for quantum state smoothing is illustrated in Fig. 2. Consider an open quantum system with two groups of output channels (b, c). An observer Alice monitors the first group, b, yielding the all-time measurement record $\overleftrightarrow{\mathbf{O}}$. A hypothetical observer monitors the second group c, yielding a record $\overleftrightarrow{\mathbf{U}}$ unobserved by Alice. The true state of the system $\rho_{\mathrm{T}}(t) \equiv \rho_{\overleftarrow{\mathbf{O}}_t, \overleftarrow{\mathbf{U}}_t}(t)$ is conditioned on both measurement records. If ρ_0 is pure then $\rho_T(t)$ will be pure for all times; no extra conditioning could possibly give a better (more pure) state. However Alice's conditional state, calculated in the conventional way (filtering),

$$\rho_F(t) \equiv \rho_{\overleftarrow{\mathbf{O}}_t}(t) = \mathrm{E}_{\overleftarrow{\mathbf{U}}_t|\overleftarrow{\mathbf{O}}_t}[\rho_{\overleftarrow{\mathbf{O}}_t, \overleftarrow{\mathbf{U}}_t}(t)], \tag{8}$$

will be mixed, because of the averaging over the unobserved record—here $\mathrm{E}_{A|B}[X]$ means the expected value of X, averaged over A, for a given B. Note that this averaging does not have to be done explicitly—it is implicit in the quantum trajectory theory as in Eq. (4) and is independent of how the channels c are monitored.

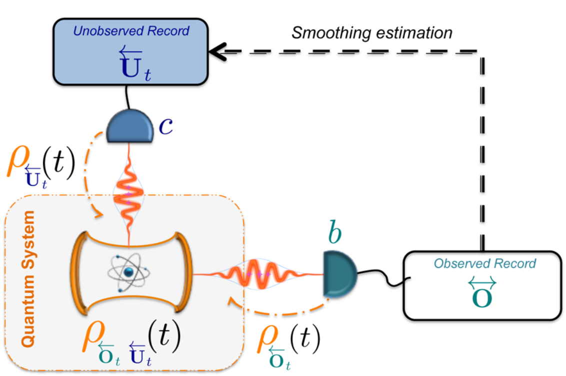

FIG. 2 (color online). The quantum state smoothing problem: to best approximate the unknown true state of a quantum system, conditioned on both observed ($\mathbf{O}$) and unobserved ($\mathbf{U}$) records, given access only to $\mathbf{O}$. This requires one to estimate $\overleftarrow{\mathbf{U}}_t$ up to time t using the full record for $\overleftrightarrow{\mathbf{O}}$ (before and after t).

The crucial insight is that Alice can do better, when averaging over $\overleftarrow{\mathbf{U}}_t$, by using information in the future of t, to define a positive-definite smoothed quantum state

$$\rho_S(t) = E_{\overleftarrow{\mathbf{U}}_t|\overleftrightarrow{\mathbf{O}}}[\rho_{\overleftarrow{\mathbf{O}}_t, \overleftarrow{\mathbf{U}}_t}(t)] \equiv \sum_{\overleftarrow{\mathbf{U}}_t} \wp_S(\overleftarrow{\mathbf{U}}_t)\rho_{\overleftarrow{\mathbf{O}}_t, \overleftarrow{\mathbf{U}}_t}(t), \tag{9}$$

Here, $\wp_S(\overleftarrow{\mathbf{U}}_t) = \wp_{\overleftrightarrow{\mathbf{O}}}(\overleftarrow{\mathbf{U}}_t) = \Pr[\overleftarrow{\mathbf{U}}_t^{\mathrm{true}} = \overleftarrow{\mathbf{U}}_t|\overleftrightarrow{\mathbf{O}}, \rho_0]$ is the probability distribution for the unobserved record prior to t, obtained by smoothing from the all-time observed record $\overleftrightarrow{\mathbf{O}}$. Note that $\wp_S(\overleftarrow{\mathbf{U}}_t)$ yields exactly the same type of information as the past quantum state of Ref. [41], except that it is more general—it specifies the probability of a continuous monitoring record $\overleftarrow{\mathbf{U}}$, not just a result of a measurement at one point in time, and is conditioned on another record, $\overleftrightarrow{\mathbf{O}}$, covering the same time interval, not just records strictly earlier and strictly later. Unlike the classical stochastic process $\mathbf{x}_t$ considered previously by Tsang [35], the record $\overleftarrow{\mathbf{U}}_t$ is of quantum origin—its statistics are undefined without a quantum system. Nevertheless, it is still a time series of c numbers with well-defined statistics and so there is no conceptual problem in applying his theory of quantum smoothing to obtain $\wp_S(\overleftarrow{\mathbf{U}}_t)$, and, thereby, $\rho_S(t)$.

We now show how to calculate Eq. (9). For definiteness and simplicity, we consider a single channel (b) yielding homodyne photocurrent y_t and a single channel (c) yielding an unobserved photon count n_t. These processes are related to the dynamics of the quantum system via the joint measurement operation $\mathcal{M}_{n_t, y_t}$ defined such that $\mathcal{M}_{y_t} = \sum_{n_t=0}^{1} \wp_{\mathrm{ost}}(n_t|\overleftarrow{Y}_t)\mathcal{M}_{n_t, y_t}$, for a convenient choice of $\wp_{\mathrm{ost}}(n_t|\overleftarrow{Y}_t)$ [47]. By standard techniques [1], $\mathcal{M}_{n_t, y_t}$ lets us generate a typical sample of the all-time records $\overleftrightarrow{O} = \overleftrightarrow{Y}$ and $\overleftrightarrow{\mathbf{U}} = \overleftrightarrow{N}^{\mathrm{true}}$. For all but one purpose (see below), the latter is irrelevant, but using the former we calculate the filtered state $\rho_{\overleftarrow{Y}_t}(t)$ from Eq. (4) and the retrofiltered effect $\hat{E}_{\overrightarrow{Y}_t}(t)$ from Eq. (6), with $\mathcal{M}_{y_t}$ in place of $\mathcal{M}_{\mathbf{r}_t}$ [50]. Next, we generate a large ensemble $\mathfrak{E}_{\mathrm{ost}}$ of random samples of $\overleftrightarrow{U} = \overleftrightarrow{N}$, according to the ostensible distribution $\wp_{\mathrm{ost}}(n_t|\overleftarrow{Y}_t)$. For each sample we calculate an associated pure state $\tilde{\rho}_{\overleftarrow{Y}_t, \overleftarrow{N}_t}(t)$, conditioned on both records, from Eq. (4) with $\mathcal{M}_{n_t, y_t}$ in place of $\mathcal{M}_{\mathbf{r}_t}$ [51].

Elementary manipulation of probabilities [47] gives $\wp_S(\overleftarrow{N}_t) \equiv \wp(\overleftarrow{N}_t|\overleftrightarrow{Y}) \propto \wp(\overrightarrow{Y}_t|\overleftarrow{N}_t, \overleftarrow{Y}_t)\wp(\overleftarrow{N}_t|\overleftarrow{Y}_t)$. Using the equations for multiple channels corresponding to Eq. (7),

$$\wp(\overrightarrow{Y}_t|\overleftarrow{N}_t, \overleftarrow{Y}_t) \quad = \mathrm{Tr}[\hat{E}_{\overrightarrow{Y}_t}\rho_{\overleftarrow{N}_t, \overleftarrow{Y}_t}]\wp_{\mathrm{ost}}(\overrightarrow{Y}_t), \tag{10}$$

and to Eq. (5),

$$\mathrm{Tr}[\hat{E}_{\vec{Y}_t}\tilde{\rho}_{\overleftarrow{N}_t,\overleftarrow{Y}_t}]\wp_{\mathrm{ost}}(\overleftarrow{N}_t|\overleftarrow{Y}_t) = \mathrm{Tr}[\hat{E}_{\vec{Y}_t}\rho_{\overleftarrow{N}_t,\overleftarrow{Y}_t}]\wp(\overleftarrow{N}_t|\overleftarrow{Y}_t), \quad (11)$$

we finally obtain, from Eq. (9),

$$\rho_S(t) \propto \sum_{\overleftarrow{N}_t}\wp_{\mathrm{ost}}(\overleftarrow{N}_t|\overleftarrow{Y}_t) \times \rho_{\overleftarrow{Y}_t,\overleftarrow{N}_t}(t)\mathrm{Tr}[\hat{E}_{\vec{Y}_t}(t)\tilde{\rho}_{\overleftarrow{Y}_t,\overleftarrow{N}_t}(t)]. \quad (12)$$

We can approximate this weighted average over all possible unobserved records using the ensemble $\mathfrak{E}_{\mathrm{ost}}$ drawn from the appropriate ostensible distribution, as discussed in the preceding paragraph. This is the method we use below to find the smoothed quantum state.

Example.—Consider a two-level atom, with driving Hamiltonian $\hat{H} = (\Omega/2)\hat{\sigma}_x$ in the interaction frame, and radiative damping described by a Lindblad operator $\sqrt{\gamma}\hat{\sigma}_-$ [1]. We take a fraction η of the fluorescence to be observed by homodyne detection, so $\hat{b} = \sqrt{\gamma\eta}\hat{\sigma}_-$. The remainder is absorbed by the environment, which we model as an unobserved record of photon counts, as discussed above, with $\hat{c} = \sqrt{\gamma(1-\eta)}\hat{\sigma}_-$. For a fixed $\overleftrightarrow{Y}$ we can compare ρ_S with ρ_F on the interval $[0, T]$. At the final time $\rho_S(T) = \rho_F(T)$ because there is no more future record $\vec{Y}_T$ to give extra information to $\rho_S(T)$. Also, we take the initial state to be pure, $\rho_0 = |1\rangle\langle 1|$, which guarantees that ρ_T is pure and that $\rho_S(0) = \rho_F(0)$.

To evaluate the advantage gained by smoothing over filtering, we use the purity,

$$P[\rho_C(t)] = \mathrm{Tr}[\rho_C^2(t)], \quad (13)$$

where ρ_C could be either ρ_F or ρ_S. If (as is the case in simulations) we know the true unobserved record $\overleftrightarrow{N}^{\mathrm{true}}$ we can also calculate the fidelity of the conditioned state to the true state $\rho_T(t) = \rho_{\overleftarrow{Y}_t,\overleftarrow{N}_t^{\mathrm{true}}}(t)$,

$$F[\rho_T(t), \rho_C(t)] = \mathrm{Tr}[\rho_T(t)\rho_C(t)]. \quad (14)$$

It is easy to show [47] that these measures are related by

$$\mathrm{E}\{P[\rho_C(t)]\} = \mathrm{E}\{F[\rho_T(t), \rho_C(t)]\} \quad (15)$$

where the ensemble averages here are over the actual distributions for $\overleftrightarrow{N}^{\mathrm{true}}$ and $\overleftrightarrow{Y}$.

In Fig. 3(a) we show typical trajectories, for Y homodyne $[\Phi = \pi/2]$ for a randomly generated true state $\rho_T(t) = \rho_{\overleftarrow{Y}_t,\overleftarrow{N}_t}(t)$ featuring one jump at $t \approx 1.8$. In this, case, ρ_T, ρ_S, and ρ_F are all confined to the Y-Z plane of the Bloch sphere, as shown. We plot Eqs. (13)–(14) in Figs. 3(b) and 3(c), respectively. It is notable from (b) that ρ_S *anticipates* the jump in ρ_T and its uncertainty about the timing of the jump leads to a lower purity in the region of the jump than the nonanticipating ρ_F. Similarly, 3(c) shows that the fidelity of ρ_S to ρ_T decreases below that of ρ_F prior to the jump, but is higher after the jump. In Fig. 3(d) we see that if there is no jump, the fidelity with ρ_T is always greater for ρ_S.

We confirm that smoothing enables better state estimation on average by calculating the average purity, for 10^3

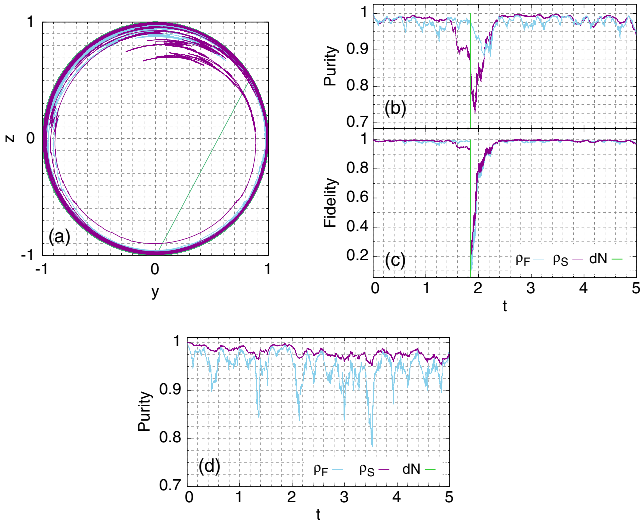

FIG. 3 (color online). (a) Trajectories in the Bloch sphere for our model system with $\Omega = 20$, $\gamma = 1$, and $\eta = 10/11$. The states shown are ρ_F (filtered, blue), ρ_S (smoothed, purple) and ρ_T (true, green) for a case where the true record includes a jump. We also plot the purities (b) and fidelities with ρ_T (c) of these ρ_F and ρ_S. The purities for a record with no jump are shown in (d). To compute ρ_S we average over an ensemble of 10^4 hypothetical unobserved records $\overleftrightarrow{N}$.

realisations of $\overleftrightarrow{Y}$. Recall from Eq. (15) that higher purity means higher fidelity with the true state. We plot this in Fig. 4 for two different local oscillator phases: $\Phi = \pi/2$ (Y homodyne) in 4(a) and $\Phi = 0$ (X homodyne) in 4(b). Because the driving of the atom causes $\hat{\sigma}_y$ to oscillate at a frequencies $\Omega \gg \gamma$, it is harder to track the state of the atom using Y-homodyne detection, and the purity of the filtered

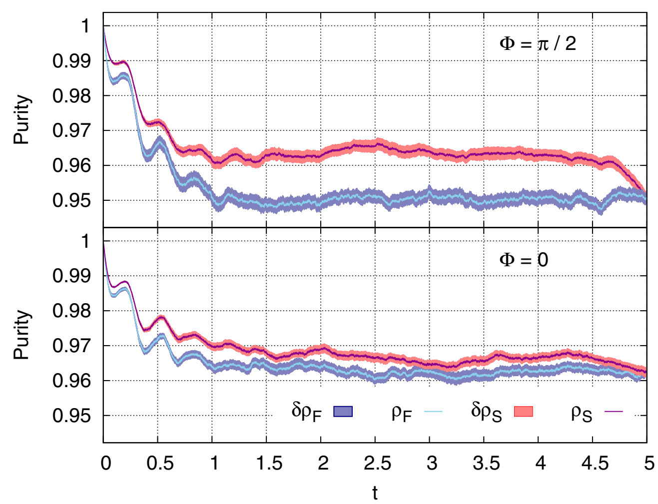

FIG. 4 (color online). Average purity for the case considered for filtered and smoothed states. The average has been calculated with 10^3 $\overleftrightarrow{Y}$ records, each one of them calculated with 10^4 estimated $\overleftrightarrow{N}$ records. The figure shows the results obtained form a Y homodyne measure (top) or X homodyne (bottom).

state is lower than for X-homodyne detection [52]. It is the former case which shows the greatest improvement in purity under smoothing: about 26% of the purity lost, because of the unobserved radiation, is recovered in 4(a) compared to about 12% in 4(b).

One can easily show [47] that our theory of quantum state smoothing includes as a special case the HMM that applies [41,44,45] to quantum systems that (unlike our atomic example) have no coherences and so are effectively classical. For genuinely quantum systems there are many questions about state smoothing to explore, including the following: what happens if one assumes the unobserved unraveling to be different from photon counting; is there a relation between quantum state smoothing and the "most probable path" formalism of Refs. [53,54]; does the HMM inevitably emerge in the classical limit, and does quantum smoothing necessarily work best in that limit; and what experiments would show uniquely quantum aspects of quantum state smoothing?

This research was supported by the Australian Research Council Program CE110001027.

[1] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2010).
[2] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering (Academic Press, New York, 1976), Vol. 123.
[3] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).
[5] A. S. Holevo, *Quantum Systems, Channels, Information* (de Gruyter, Berlin, 2012).
[6] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, England, 1992).
[7] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106**, 090401 (2011).
[8] M. Tsang and R. Nair, Phys. Rev. A **86**, 042115 (2012).
[9] M. Tsang, New J. Phys. **15**, 073005 (2013).
[10] M. Tsang, Phys. Rev. Lett. **102**, 253601 (2009).
[11] M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H. A. Bachor, and W. P. Bowen, Nat. Photonics **7**, 229 (2013).
[12] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A **54**, R4649 (1996).
[13] C. W. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics* (Springer, Berlin, 2004).
[14] C. M. Caves, C. A. Fuchs, and R. Schack, Phys. Rev. A **65**, 022305 (2002).
[15] C. A. Fuchs, N. D. Mermin, and R. Schack, Am. J. Phys. **82**, 749 (2014).
[16] H. J. Carmichael, *An Open Systems Approach to Quantum Optics* (Springer, Berlin, 1993).
[17] J. Dalibard, Y. Castin, and K. Mølmer, Phys. Rev. Lett. **68**, 580 (1992).
[18] C. W. Gardiner, A. S. Parkins, and P. Zoller, Phys. Rev. A **46**, 4363 (1992).
[19] H. M. Wiseman and G. J. Milburn, Phys. Rev. A **47**, 1652 (1993).
[20] A. Barchielli, Int. J. Theor. Phys. **32**, 2221 (1993).
[21] A. N. Korotkov, Phys. Rev. B **60**, 5737 (1999).
[22] H. M. Wiseman and G. J. Milburn, Phys. Rev. Lett. **70**, 548 (1993).
[23] H. M. Wiseman, Phys. Rev. Lett. **75**, 4587 (1995).
[24] D. B. Horoshko and S. Y. Kilin, Phys. Rev. Lett. **78**, 840 (1997).
[25] A. C. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. M. Tan, Phys. Rev. A **62**, 012105 (2000).
[26] W. P. Smith, J. E. Reiner, L. A. Orozco, S. Kuhr, and H. M. Wiseman, Phys. Rev. Lett. **89**, 133601 (2002).
[27] H. M. Wiseman and A. C. Doherty, Phys. Rev. Lett. **94**, 070405 (2005).
[28] P. Bushev, D. Rotter, A. Wilson, F. Dubin, C. Becher, J. Eschner, R. Blatt, V. Steixner, P. Rabl, and P. Zoller, Phys. Rev. Lett. **96**, 043003 (2006).
[29] R. Vijay, C. Macklin, D. H. Slichter, S. J. Weber, K. W. Murch, R. Naik, A. N. Korotkov, and I. Siddiqi, Nature (London) **490**, 77 (2012).
[30] V. P. Belavkin, *Lecture Notes in Control and Information Sciences* (Springer, Berlin, 1988), p. 245.
[31] V. P. Belavkin, Rep. Math. Phys. **43**, A405 (1999).
[32] M. R. James, Phys. Rev. A **69**, 032108 (2004).
[33] H. F. Hofmann, G. Mahler, and O. Hess, Phys. Rev. A **57**, 4877 (1998).
[34] A. N. Korotkov, Phys. Rev. B **63**, 115403 (2001).
[35] M. Tsang, Phys. Rev. Lett. **102**, 250403 (2009).
[36] M. Tsang, Phys. Rev. A **80**, 033840 (2009).
[37] M. Tsang, J. H. Shapiro, and S. Lloyd, Phys. Rev. A **79**, 053843 (2009).
[38] D. W. Berry, M. J. W. Hall, and H. M. Wiseman, Phys. Rev. Lett. **111**, 113601 (2013).
[39] T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arao, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington, Phys. Rev. Lett. **104**, 093601 (2010).
[40] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph *et al.*, Science **337**, 1514 (2012).
[41] S. Gammelmark, B. Julsgaard, and K. Mølmer, Phys. Rev. Lett. **111**, 160401 (2013).
[42] D. Tan, S. J. Weber, I. Siddiqi, K. Mølmer, and K. W. Murch, Phys. Rev. Lett. **114**, 090403 (2015).
[43] W. H. Press, A. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, New York, 2010).
[44] M. A. Armen, A. E. Miller, and H. Mabuchi, Phys. Rev. Lett. **103**, 173601 (2009).
[45] S. Gammelmark, K. Mølmer, W. Alt, T. Kampschulte, and D. Meschede, Phys. Rev. A **89**, 043839 (2014).
[46] T. Rybarczyk, B. Peaudecerf, M. Penasa, S. Gerlich, B. Julsgaard, K. Mølmer, S. Gleyzes, M. Brune, J. M. Raimond, S. Haroche *et al.*, Phys. Rev. A **91**, 062116 (2015).
[47] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.115.180407 for more detailed analytical calculations.

[48] E. Pardoux, Stochastics **6**, 38 (1981).
[49] J. L. Junkins, *Optical Estimation of Dynamic Systems* (Chapman and Hall, New York, 2011).
[50] We tested the correctness of our numerical calculations for the two-level atom example below by verifying the t independence of $\wp(\overleftrightarrow{Y}) = \mathrm{Tr}[\hat{E}_{\bar{Y}_t}(t)\rho_{\bar{Y}_t}(t)]$.
[51] We tested this numerically for the example below by verifying that the ensemble average of the calculated "doubly" filtered states over the random sample of $\overleftrightarrow{N}$ coincides with the filtered states only conditioned on $\bar{Y}_t$, i.e., $E_{\bar{N}_t}[\rho_{\bar{N}_t,\bar{Y}_t}(t)] = \rho_{\bar{Y}_t}(t),\ \forall\ t \in [t_0, T]$.
[52] D. J. Atkins, Z. Brady, K. Jacobs, and H. M. Wiseman, Europhys. Lett. **69**, 163 (2005).
[53] A. Chantasri, J. Dressel, and A. N. Jordan, Phys. Rev. A **88**, 042110 (2013).
[54] S. J. Weber, A. Chantasri, J. Dressel, A. N. Jordan, K. W. Murch, and I. Siddiqi, Nature (London) **511**, 570 (2014).