Quantum Rauch-Tung-Striebel smoothed state

Kiarn T. Laverick ●*

Centre for Quantum Computation and Communication Technology (Australian Research Council),
Centre for Quantum Dynamics, Griffith University, Nathan, Queensland 4111, Australia

Smoothing is a technique that is used to estimate the state of a system using measurement information both prior and posterior to the estimation time. Two notable examples of this technique are the Rauch-Tung-Striebel and Mayne-Fraser-Potter smoothing techniques for linear Gaussian systems, both resulting in the optimal smoothed estimate of the state. However, when considering a quantum system, classical smoothing techniques can result in an estimate that is not a valid quantum state. Consequently, a different smoothing theory was developed explicitly for quantum systems. This theory has since been applied to the special case of linear Gaussian quantum (LGQ) systems, where, in deriving the LGQ state smoothing equations, the Mayne-Fraser-Potter technique was utilized. As a result, the final equations describing the smoothed state are closely related to the classical Mayne-Fraser-Potter smoothing equations. In this paper, I derive the equivalent Rauch-Tung-Striebel form of the quantum state smoothing equations, which further simplify the calculation for the smoothed quantum state in LGQ systems, and I provide insight into the dynamics of the smoothed quantum state. This form of the LGQ smoothing equations brings to light a property of the smoothed quantum state that was hidden in the Mayne-Fraser-Potter form, namely the nondifferentiability of the smoothed mean. By identifying the nondifferentiable part of the smoothed mean, I am then able to derive a necessary and sufficient condition for the quantum smoothed mean to be differentiable in the steady-state regime.

I. INTRODUCTION

Estimating an unknown state, i.e., a probability density function (PDF), of physical systems using indirect measurement results has been studied in great depth [1–13]. When restricting to the case of (classical) linear Gaussian (LG) systems, Kalman and Bucy [1] developed an optimal estimation technique, known as *filtering*, which conditions the estimate of the state on past measurement information, i.e., measurement information up until the time of estimation τ. This estimated state is referred to as the filtered state. While one of the appeals of the Kalman-Bucy filtering technique is its ability to actively update the estimate of the state in real time, there are other optimal estimation techniques that are more accurate [2–13].

One such technique was developed soon after the Kalman-Bucy filtering theory by Rauch, Tung, and Striebel [2,3]. This technique, referred to as *smoothing*, not only utilizes the past measurement information as the Kalman-Bucy filter does, but it also uses information gathered after the estimation time τ, i.e., the "future" measurement record, to provide a more accurate estimate of the state. The Rauch-Tung-Striebel (RTS)

smoothing technique, first, uses the Kalman-Bucy filtering technique to estimate the state until the final estimation time T. Once at the final estimation time T, the RTS smoothing equations run back over the estimated state, updating the results based on the future measurement record that has been gathered. This results in a smoothed estimate of the state. The only drawback of this smoothing technique, compared to filtering, is that the smoothed state cannot be obtained in real time since future information is required.

A similar technique, developed by Mayne [4], Fraser [5], and Potter [6] shortly after RTS smoothing, also utilized a past-future measurement record. Mayne-Fraser-Potter (MFP) smoothing, sometimes referred to as two-filter smoothing for reasons that will become apparent, also utilizes Kalman-Bucy filtering to condition on the past measurement record. However, to make use of the future information, they introduced a secondary filter, which I will refer to as a *retrofilter*, that ran backwards from a final uninformative state conditioning on the future measurement record. This state is the retrofiltered state. Combining the filtered state and the retrofiltered state together results in the MFP smoothed state. It has since been shown [6] that the RTS smoothed state and the MFP smoothed state are, in fact, identical.

Moving to quantum systems, an analogous problem to classical state estimation also exists, where instead of estimating a PDF one wishes to estimate the density matrix ρ. Similar to the classical filtering theory, one can estimate the quantum state based on the past measurement record. As a result, this technique is often referred to as quantum filtering [14–16]. Interestingly, if one restricts to linear Gaussian

---

quantum (LGQ) systems, the quantum filtering technique reduces to the Kalman-Bucy filtering theory [17–19]. Given this, one might assume that classical smoothing techniques, such as the RTS or MFP techniques, can also be applied to quantum state estimation. This is not the case. Applying the classical theory to quantum state estimation can result in an unphysical estimate of the quantum state [20–24]. This is due to the noncommutativity of the operators describing the system and the operators describing the future measurement outcomes in quantum theory, which is usually not present in the classical theory. Due to the failure of the classical theory, Guevara and Wiseman [22] devised a new smoothing theory specifically for quantum systems, namely the quantum state smoothing theory.

The quantum state smoothing theory [22,24,25] introduces a secondary measurement record that is unobserved by the observer, say Alice, but is observed by someone else, say Bob. The role of Bob's measurement is to gather information about the system that Alice's measurement may have missed. If Alice had access to Bob's record, hereby referred to as the "unobserved" record, she could condition the estimate of the state on both her past observed record $\overleftarrow{\mathbf{O}}$ and the unobserved record $\overleftarrow{\mathbf{U}}$ to obtain the true state of the system $\rho_{\mathrm{T}} = \rho_{\overleftarrow{\mathbf{O}}\,\overleftarrow{\mathbf{U}}}$, a state containing the maximum amount of information about the system given the measurements. Here, a leftward arrow indicates that it is the past record, and a bidirectional arrow indicates a past-future record. However, since Alice does not have access to the unobserved record, she cannot compute the true state of the system. The best [26] Alice can do is to estimate the true state based only on the observed record. Now, Alice can obtain a smoothed quantum state by averaging over all possible true states conditioned on her past-future observed record $\overleftrightarrow{\mathbf{O}}$ [22], i.e.,

$$\rho_{\mathrm{S}} = \mathbb{E}_{\overleftarrow{\mathbf{U}}\,|\,\overleftrightarrow{\mathbf{O}}}\{\rho_{\mathrm{T}}\}, \qquad (1)$$

where $\mathbb{E}_{A|B}\{C\}$ denotes the ensemble average of C over A conditioned on B, and it may appear without the A subscript when $A = C$.

Following the conception of the smoothed quantum state, the theory was applied to the special case of LGQ systems [24,27,28]. To derive the LGQ smoothed quantum state, the classical MFP smoothing technique was used. These quantum state smoothing equations, due to being closed-form equations, have been able to identify properties of the smoothed quantum state [24,27,28] that would have been difficult to find in the general case. Furthermore, the smoothed state is far simpler to compute for LGQ systems compared to even a simple system in the general case, allowing for easier verification of these properties. However, while the MFP forms of the quantum state smoothing equations have been very useful, they require the calculation of a retrofiltered state. This can be avoided by instead using the RTS form of the smoothing equations, making the smoothed state even simpler to compute for LGQ systems. Additionally, the RTS form of the quantum state smoothing equations are dynamical equations, and they can provide insight into the properties of the smoothed quantum state that would otherwise be hidden in the MFP form. As a case in point, I derive a necessary and

sufficient constraint for the mean of the smoothed quantum state to evolve smoothly in the steady-state regime.

This paper is structured as follows. First, in Sec. II, I will briefly review classical and quantum state smoothing for linear Gaussian systems. Next, in Sec. III, I will derive the RTS form of the smoothed quantum state. Finally, in Sec. IV, I will discuss this RTS form of the LGQ state smoothing equations, showing that, under the same condition that makes the classically smoothed mean a continuous function, the path of the smoothed mean is necessarily nondifferentiable. Furthermore, I identify a necessary and sufficient condition for the mean of the smoothed quantum state to be differentiable in the steady-state regime.

II. LINEAR GAUSSIAN STATE ESTIMATION

A. Classical

Consider a classical dynamical system. The state of the system is given by a probability density function $\wp(\check{\mathbf{x}})$, where $\check{\mathbf{x}} = (\check{x}_1, \check{x}_2, \ldots, \check{x}_M)^\top$ is a vector of M parameters that are required to completely characterize the system, and $\top$ denotes the transpose. Note, for clarity, the wedge accent will be used to denote a dummy variable to differentiate it from the corresponding random variable. For a LG system, the classical state will be a Gaussian distribution, i.e., $\wp(\check{\mathbf{x}}) = g(\check{\mathbf{x}}; \mathbb{E}\{\check{\mathbf{x}}\}, V)$, completely described by its mean $\mathbb{E}\{\check{\mathbf{x}}\}$ and covariance matrix $V = \mathbb{E}\{\check{\mathbf{x}}\check{\mathbf{x}}^\top\} - \mathbb{E}\{\check{\mathbf{x}}\}\mathbb{E}\{\check{\mathbf{x}}^\top\}$. To guarantee that the state remains Gaussian throughout the evolution of the system, provided that at the initial time t_0 the state is Gaussian with mean $\mathbb{E}\{\check{\mathbf{x}}(t_0)\} = \mathbf{x}_0$ and covariance $V(t_0) = V_0$, it is necessary that the following two constraints are satisfied [7–12,19]. First, the evolution of $\mathbf{x}$ must be described by a linear Langevin equation

$$d\mathbf{x} = A\mathbf{x}\,dt + E\,d\mathbf{v}_{\mathrm{p}}. \qquad (2)$$

Here A (the drift matrix) and E are constant matrices, and the process noise $d\mathbf{v}_{\mathrm{p}}$ is a vector of independent Wiener increments satisfying

$$\mathbb{E}\{d\mathbf{v}_{\mathrm{p}}\} = 0, \quad d\mathbf{v}_{\mathrm{p}}d\mathbf{v}_{\mathrm{p}}^\top = I_K dt, \qquad (3)$$

where I_k is the $k \times k$ identity matrix. Secondly, any measurement current $\mathbf{y}$ used to refine our estimate of the state must also be linear, i.e.,

$$\mathbf{y}\,dt = C\mathbf{x}\,dt + d\mathbf{v}_{\mathrm{m}}, \qquad (4)$$

where the measurement matrix C is a constant matrix, and the measurement noise $d\mathbf{v}_{\mathrm{m}}$ is a vector of independent Wiener increments satisfying similar conditions to Eq. (3). Usually, it is also assumed that the process noise and measurement noise are uncorrelated [1–12], i.e., $d\mathbf{v}_{\mathrm{p}}(d\mathbf{v}_{\mathrm{m}})^\top = 0$, however this assumption is not always true, i.e., there might be measurement backaction. For the sake of generality, I will consider the case in which the two noises may be correlated, described by the cross-correlation matrix [19,29–32]

$$\Gamma^\top dt = E\,d\mathbf{v}_{\mathrm{p}}(d\mathbf{v}_{\mathrm{m}})^\top. \qquad (5)$$

If all of the above criteria are satisfied, one can calculate the filtered estimate of the state $\wp_{\mathrm{F}}(\check{\mathbf{x}}) \equiv \wp(\check{\mathbf{x}}|\overleftarrow{\mathbf{O}}) = g(\check{\mathbf{x}}; \mathbf{x}_{\mathrm{F}}, V_{\mathrm{F}})$ by conditioning the state on the past measurement record

$\overleftarrow{\mathbf{O}}$. The filtered mean $\mathbf{x}_F := \mathbb{E}_{|\overleftarrow{\mathbf{O}}}\{\check{\mathbf{x}}\}$ and covariance $V_F = \mathbb{E}_{\check{\mathbf{x}}|\overleftarrow{\mathbf{O}}}\{\check{\mathbf{x}}\check{\mathbf{x}}^\top\} - \mathbf{x}_F\mathbf{x}_F^\top$ are given by the Kalman-Bucy filtering equations [1,19,29–32]

$$d\mathbf{x}_F = A\mathbf{x}_F dt + \mathcal{K}^+[V_F]d\mathbf{w}_F, \tag{6}$$

$$\frac{dV_F}{dt} = AV_F + V_F A^\top + D - \mathcal{K}^+[V_F]\mathcal{K}^+[V_F]^\top, \tag{7}$$

with initial conditions $\mathbf{x}_F(t_0) = \mathbf{x}_0$ and $V_F(t_0) = V_0$. Here $D = EE^\top$ is the diffusion matrix,

$$\mathcal{K}^\pm[V] = VC^\top \pm \Gamma^\top \tag{8}$$

is the optimal Kalman gain, where the minus version will appear shortly, and the vector of innovations is defined as $d\mathbf{w}_c = \mathbf{y}dt - C\mathbf{x}_c dt$ for a given conditioning c $\in$ {F, R, S}, representing filtering, retrofiltering, and smoothing, respectively, and it satisfies similar conditions to Eq. (3).

One can also calculate the smoothed estimate of the state $\wp_S(\check{\mathbf{x}}) \equiv \wp(\check{\mathbf{x}}|\overleftrightarrow{\mathbf{O}}) = g(\check{\mathbf{x}}; \mathbf{x}_S, V_S)$ by conditioning the estimate on the past-future measurement record $\overleftrightarrow{\mathbf{O}}$. The optimal smoothed mean $\mathbf{x}_S := \mathbb{E}_{|\overleftrightarrow{\mathbf{O}}}\{\check{\mathbf{x}}\}$ and covariance matrix $V_S = \mathbb{E}_{\check{\mathbf{x}}|\overleftrightarrow{\mathbf{O}}}\{\check{\mathbf{x}}\check{\mathbf{x}}^\top\} - \mathbf{x}_S\mathbf{x}_S^\top$ can be calculated in two ways. The first method is a maximum-likelihood argument and results in the RTS smoothing equations [2,3,32]

$$d\mathbf{x}_S = A\mathbf{x}_S dt + \tilde{D}V_F^{-1}(\langle\mathbf{x}\rangle_S - \langle\mathbf{x}\rangle_F)dt + \Gamma^\top d\mathbf{w}_S, \tag{9}$$

$$\frac{dV_S}{dt} = (\tilde{A} + \tilde{D}V_F^{-1})V_S + V_S(\tilde{A} + \tilde{D}V_F^{-1})^\top - \tilde{D}, \tag{10}$$

with the final conditions $\mathbf{x}_S(T) = \mathbf{x}_F(T)$ and $V_S(T) = V_F(T)$. Here, $\tilde{A} = A - \Gamma^\top C$, $\tilde{D} = D - \Gamma^\top\Gamma$.

The second method arises from a Bayesian argument, where one first introduces a retrofiltered state that runs backwards in time from the final estimation time. This retrofiltered state will also be Gaussian with mean $\mathbf{x}_R := \mathbb{E}_{|\overrightarrow{\mathbf{O}}}\{\check{\mathbf{x}}\}$ and covariance $V_R = \mathbb{E}_{\check{\mathbf{x}}|\overrightarrow{\mathbf{O}}}\{\check{\mathbf{x}}\check{\mathbf{x}}^\top\} - \mathbf{x}_R\mathbf{x}_R^\top$, given by

$$-d\mathbf{x}_R = -A\mathbf{x}_R dt + \mathcal{K}^-[V_R]d\mathbf{w}_R, \tag{11}$$

$$-\frac{dV_R}{dt} = -AV_R - V_R A^\top + D - \mathcal{K}^-[V_R]\mathcal{K}^-[V_R]^\top, \tag{12}$$

with $V_R^{-1}(T) = 0$ describing an uninformative state. Combining the filtered and retrofiltered state gives the smoothed state, with the mean and covariance described by the MFP smoothing equations [4–6]

$$\mathbf{x}_S = V_S[V_F^{-1}\mathbf{x}_F + V_R^{-1}\mathbf{x}_R], \tag{13}$$

$$V_S = (V_F^{-1} + V_R^{-1})^{-1}. \tag{14}$$

It has since been shown [6] that the RTS and MFP forms of the smoothed mean and covariance are identical, and this is easily verified by differentiating Eqs. (13) and (14) with respect to time.

B. Quantum

To begin, let us consider an open quantum system whose density matrix ρ, assuming a Markovian system, evolves according to the Lindblad master equation [19]

$$\hbar\frac{d\rho}{dt} = -[\hat{H}, \rho] + \mathcal{D}[\hat{\mathbf{c}}]\rho, \tag{15}$$

with initial condition $\rho(t_0) = \rho_0$. Here $\hat{H}$ is the system Hamiltonian describing unitary evolution, $\hat{\mathbf{c}} = (\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_K)$ is the vector of Lindblad operators describing the interaction between the system and environment, $[A, B] = AB - BA$ is the commutator, and the superoperator

$$\mathcal{D}[\hat{\mathbf{c}}]\rho = \hat{c}_k\rho\hat{c}_k^\dagger - \{\hat{c}_k^\dagger\hat{c}_k, \rho\}/2, \tag{16}$$

with $\{A, B\} = AB + BA$ being the anticommutator; I am using the Einstein summation convention over repeated indices.

For this quantum system to be analogous to a classical LG system, we require that the system observables have an unbounded spectrum. Thus, we will assume that the quantum system can be described by N bosonic modes with each mode described by a position $\hat{q}_k$ and momentum $\hat{p}_k$, which satisfy the commutation relation $[\hat{q}_k, \hat{p}_\ell] = i\hbar\delta_{k,\ell}$. We can then construct a $2N$ vector $\hat{\mathbf{x}} = (\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_N, \hat{p}_N)^\top$ describing all the modes of the bosonic system. For the system to be classified as a LGQ system [17–19], the Wigner function, a quasiprobability distribution defined as

$$W(\check{\mathbf{x}}) = (2\pi)^{-2N}\int d^{2N}\mathbf{b}\, \mathrm{Tr}[\rho e^{i\mathbf{b}^\top(\hat{\mathbf{x}}-\check{\mathbf{x}})}], \tag{17}$$

must initially be Gaussian and remain Gaussian, where the latter can be guaranteed when the Hamiltonian is quadratic and the vector of Lindblad operators is linear in $\hat{\mathbf{x}}$, i.e., $\hat{H} = \hat{\mathbf{x}}^\top G\hat{\mathbf{x}}/2$ and $\hat{\mathbf{c}} = B\hat{\mathbf{x}}$, respectively. Here, $d^{2N}\mathbf{b} = \prod_{k=1}^{2N}b_k$ is the usual vectorial integral measure. Note, since we are using the Wigner function to characterize the quantum state, the Hamiltonian must be symmetrically ordered in $\hat{q}_k$ and $\hat{p}_k$, meaning G needs to be a symmetric matrix. As the Wigner function is Gaussian, i.e., $W(\check{\mathbf{x}}) = g(\check{\mathbf{x}}; \langle\hat{\mathbf{x}}\rangle, V)$, all that is required to know the state is the mean $\langle\hat{\mathbf{x}}\rangle$ and covariance matrix, defined symmetrically, $V_{k\ell} = \langle\hat{x}_k\hat{x}_\ell + \hat{x}_\ell\hat{x}_k\rangle/2 - \langle\hat{x}_k\rangle\langle\hat{x}_\ell\rangle$, where the quantum expectation is $\langle\bullet\rangle = \mathrm{Tr}[\bullet\rho]$. Using the Lindblad master equation, we can compute the dynamic equations for the unconditioned mean and covariance,

$$d\langle\hat{\mathbf{x}}\rangle = A\langle\hat{\mathbf{x}}\rangle dt, \tag{18}$$

$$\frac{dV}{dt} = AV + VA^\top + D, \tag{19}$$

with initial conditions $\langle\hat{\mathbf{x}}\rangle(t_0) = \langle\hat{\mathbf{x}}\rangle_0$ and $V(t_0) = V_0$. Here

$$A = \Sigma(G + \mathrm{Im}[B^\dagger B]) \quad \text{and} \quad D = \hbar\Sigma\,\mathrm{Re}[B^\dagger B]\Sigma^\top, \tag{20}$$

with $\Sigma_{k\ell} = -i[\hat{x}_k, \hat{x}_\ell]/\hbar$ being a symplectic matrix. Additionally, since the position and momentum operators do not commute, it is necessary that the covariance matrix obeys the Schrödinger-Heisenberg uncertainty relation

$$V + \frac{i\hbar}{2}\Sigma \geqslant 0. \tag{21}$$

Without any measurement information, this would be the most accurate estimate of the state possible.

If one wishes to obtain a better estimate on the quantum state, it is necessary to gather more information about the

system by measuring the environment. In the event of a continuous monitoring, we can condition the evolution of the state on the past measurement results and obtain the so called quantum filtering equation [19,33],

$$\hbar d\rho_F = -i[H, \rho_F]dt + \mathcal{D}[\hat{\mathbf{c}}]\rho_F dt + \sqrt{\hbar}d\mathbf{w}_F^\top \mathcal{H}[M^\dagger \hat{\mathbf{c}}]\rho_F, \tag{22}$$

with $\rho_F(t_0) = \rho_0$. For reasons that will become apparent, I have restricted the measurements to diffusive-type measurements, such as homodyne or heterodyne schemes, with the matrix M (assumed time-independent for simplicity) characterizing the particular unraveling. Here the filtered innovation is a vector of independent Wiener increments defined by $d\mathbf{w}_F = \mathbf{y}dt - \langle M^\dagger \hat{\mathbf{c}} + \hat{\mathbf{c}}^\dagger M^\top \rangle_F dt$, satisfying conditions similar to Eq. (3), with the conditioned expectation value defined as $\langle \bullet \rangle_c = \text{Tr}[\bullet \rho_c]$ with $c \in \{F, T, S\}$, representing the filtered, true, and smoothed states, respectively, and $\mathbf{y}dt$ being the measurement current, and the superoperator

$$\mathcal{H}[M^\dagger \hat{\mathbf{c}}]\rho = M^\dagger \hat{\mathbf{c}}\rho + \rho\hat{\mathbf{c}}^\dagger M - \text{Tr}[M^\dagger \hat{\mathbf{c}}\rho + \rho\hat{\mathbf{c}}^\dagger M]\rho. \tag{23}$$

To ensure that the resulting filtered state is a valid quantum state, it is necessary and sufficient [33] that the matrix M satisfies $MM^\dagger = \text{diag}(\eta_1, \eta_2, \ldots, \eta_K)$, where $0 \leqslant \eta_k \leqslant 1 \, \forall k$ and it can be interpreted as the fraction of the channel $\hat{c}_k$ that has been observed. Note, it must be the case that at least one $\eta_k > 0$, otherwise no measurement has been performed and both filtering and smoothing are redundant.

When considering an LGQ system, in order to keep the filtered Wigner function Gaussian throughout the entire evolution, we require the measurement current to be linear in $\hat{\mathbf{x}}$ (which is the case for diffusive type measurements), i.e.,

$$\mathbf{y}dt = C\hat{\mathbf{x}}\,dt + d\mathbf{v}_m, \tag{24}$$

where the measurement matrix $C = 2\sqrt{\hbar^{-1}}T^\top \tilde{B}$, $T^\top = [\text{Re}(M^\top), \text{Im}(M^\top)]$, $\tilde{B}^\top = [\text{Re}(B^\top), \text{Im}(B^\top)]$, and the measurement noise, for simplicity, is assumed to be white, i.e., it satisfies the properties in Eq. (3). As was the case for the unconditioned state, since the filtered state has been restricted to have a Gaussian Wigner function, $W_F(\check{\mathbf{x}}) = g(\check{\mathbf{x}}; \langle\hat{\mathbf{x}}\rangle_F, V_F)$, we only need information about the mean and covariance matrix to specify the state. Using Eq. (22), one can obtain the equations for the filtered mean and covariance matrix [17–19],

$$d\langle\hat{\mathbf{x}}\rangle_F = A\langle\hat{\mathbf{x}}\rangle_F dt + \mathcal{K}^+[V_F]d\mathbf{w}_F, \tag{25}$$

$$\frac{dV_F}{dt} = AV_F + V_F A^\top + D - \mathcal{K}^+[V_F]\mathcal{K}^+[V_F]^\top, \tag{26}$$

with initial conditions $\langle\hat{\mathbf{x}}\rangle_F(t_0) = \langle\hat{\mathbf{x}}\rangle_0$ and $V_F(t_0) = V_0$. Here $\mathcal{K}^+[V]$ is defined in Eq. (8), with $\Gamma = -\sqrt{\hbar}T^\top S\tilde{B}\Sigma^\top$ and $S = \begin{bmatrix} 0 & I_K \\ -I_K & 0 \end{bmatrix}$. Note, since the filtered state ρ_F is a valid quantum state, it is the case that V_F satisfies the Schrödinger-Heisenberg uncertainty relation.

If one wishes to obtain an even more accurate estimate of the state, it is possible that we can condition the estimate of the state on not only the past measurement information but also the future measurement record. This leads us to the quantum state smoothing theory of Guevara and Wiseman [22].

III. DERIVING THE QUANTUM RAUCH-TUNG-STRIEBEL SMOOTHED STATE

To calculate a smoothed quantum state, one needs to first introduce the true state ρ_T. As mentioned in the Introduction, the true state is an estimate conditioned on two independent measurement records, the past record observed by Alice and the past record observed by Bob. This situation is merely an extension of the filtering scenario to multiple independent measurement records, and thus it is simple to see that the resulting stochastic master equation will be

$$\hbar d\rho_T = -i[H, \rho_T]dt + \mathcal{D}[\hat{\mathbf{c}}]\rho_T dt$$

$$+ \sum_{r \in \{o,u\}} \sqrt{\hbar}d\,\mathbf{w}_r^\top \mathcal{H}[M_r^\dagger \hat{\mathbf{c}}]\rho_T, \tag{27}$$

where, assuming Alice and Bob have the same initial information about the quantum state, the initial condition is $\rho_T(t_0) = \rho_0$. Note, this may not always be the case and it is trivial to adapt to the general case, however this assumption seems like a scenario that would occur frequently and is worth considering. As before, we are restricting to the case in which both Alice and Bob measure there respective fractions of the measurement channels using a diffusive-type measurements. We have introduced the (r)ecord subscript to distinguish between the record (o)bserved by Alice and the record (u)nobserved by Alice (Bob's record), where the matrices M_o and M_u describe how Alice and Bob have unraveled the system, respectively. The observed and unobserved innovations are defined as $d\mathbf{w}_r = \mathbf{y}_r dt - \langle M_r^\dagger \hat{\mathbf{c}} + \hat{\mathbf{c}}^\dagger M_r^\top \rangle_T dt$, where we have assumed both records are uncorrelated, i.e., $d\mathbf{w}_o d\mathbf{w}_u^\top = 0$, and $\mathbf{y}_r dt$ is the corresponding measurement current for Alice and Bob.

As was the case for the filtered state, to ensure that the true state is a valid quantum case, it is necessary and sufficient that both M_o and M_u must satisfy $M_r M_r^\dagger = \text{diag}(\eta_{r,1}, \eta_{r,2}, \ldots, \eta_{r,K})$ with $0 \leqslant \eta_{r,k} \leqslant 1$ and $\eta_{o,k} + \eta_{u,k} \leqslant 1 \, \forall k$. Note, although it is often convenient to assume that $M_o M_o^\dagger + M_u M_u^\dagger = I_K$ [24,27,28], that is, together Alice's and Bob's measurements constitute a perfect measurement of the system and the resulting true state will be pure, it not a necessary assumption and what follows will hold generally. As was the case with filtering, it must be that at least one $\eta_{o,k} > 0$ and $\eta_{u,k'} > 0$.

For an LGQ system, to ensure that the true state remains Gaussian, as was the case for filtering, we require that both Alice's and Bob's measurements are linear,

$$\mathbf{y}_r dt = C_r \langle\hat{\mathbf{x}}\rangle_T dt + d\mathbf{w}_r, \tag{28}$$

where, for simplicity, I have defined the currents in terms of the true mean $\langle\hat{\mathbf{x}}\rangle_T$ and the innovation, where, in this form, it should be clear that the observed innovation $d\mathbf{w}_o = \mathbf{y}_o dt - C_o \langle\hat{\mathbf{x}}\rangle_T dt$ is different from Alice's filtered innovation $d\mathbf{w}_F = \mathbf{y}_o dt - C_o \langle\hat{\mathbf{x}}\rangle_F dt$. Given this restriction, the Wigner function for the true state will be Gaussian, $W_T(\check{\mathbf{x}}) = g(\check{\mathbf{x}}, \langle\hat{\mathbf{x}}\rangle_T, V_T)$, with the true mean $\langle\hat{\mathbf{x}}\rangle_T$ and covariance V_T [27,28],

$$d\langle\hat{\mathbf{x}}\rangle_T = A\langle\hat{\mathbf{x}}\rangle_T dt + \sum_{r \in \{o,u\}} \mathcal{K}_r^+[V_T]d\mathbf{w}_r, \tag{29}$$

$$\frac{dV_\mathrm{T}}{dt} = AV_\mathrm{T} + V_\mathrm{T}A^\top + D - \sum_{\mathrm{r}\in\{\mathrm{o,u}\}} \mathcal{K}_\mathrm{r}^+[V_\mathrm{T}]\mathcal{K}_\mathrm{r}^+[V_\mathrm{T}]^\top, \quad (30)$$

with initial conditions $\langle\hat{\mathbf{x}}\rangle_\mathrm{T}(t_0) = \langle\hat{\mathbf{x}}\rangle_0$ and $V_\mathrm{T}(t_0) = V_0$. Here, $\mathcal{K}_\mathrm{r}^\pm[V] = VC_\mathrm{r}^\top \pm \Gamma_\mathrm{r}^\top$, where C_r and Γ_r are defined in the same way as before, with the appropriate M_r used in both cases.

Now that the true state has been calculated, we can begin to derive the smoothed quantum state for LGQ systems. To begin, as we are interested in the Wigner function of the smoothed state $W_\mathrm{S}(\check{\mathbf{x}})$ in the LGQ setting, we can apply Eq. (17) to both sides of Eq. (1), where, by the linearity of the trace, we obtain

$$W_\mathrm{S}(\check{\mathbf{x}}) = \mathbb{E}_{\overleftarrow{\mathbf{U}}|\overleftrightarrow{\mathbf{O}}}\{W_\mathrm{T}(\check{\mathbf{x}})\}. \quad (31)$$

As the Wigner function of the true state is restricted to a Gaussian, we know that it only depends on the mean $\langle\hat{\mathbf{x}}\rangle_\mathrm{T}$ and covariance matrix V_T, the latter of which is deterministic with the former depending explicitly on both $\overleftarrow{\mathbf{O}}$ and $\overleftarrow{\mathbf{U}}$. Hence, averaging over $\overleftarrow{\mathbf{U}}$ with a fixed $\overleftrightarrow{\mathbf{O}}$ will be equivalent to averaging over the true mean $\langle\hat{\mathbf{x}}\rangle_\mathrm{T}$ for a fixed observed record. To make the notation simpler for the derivation to come, we take $\overset{\circ}{\check{\mathbf{x}}} = \langle\hat{\mathbf{x}}\rangle_\mathrm{T}$, where the ellipse accent will be referred to as a "halo" and will denote intermediary variables between the true state and the filtered/smoothed state. Making these changes to Eq. (31), we get

$$W_\mathrm{S}(\check{\mathbf{x}}) = \mathbb{E}_{\overset{\circ}{\check{\mathbf{x}}}|\overleftrightarrow{\mathbf{O}}}\{W_\mathrm{T}(\check{\mathbf{x}})\} = \int d^{2N}\overset{\circ}{\check{\mathbf{x}}}\, \wp(\overset{\circ}{\check{\mathbf{x}}}|\overleftrightarrow{\mathbf{O}})g(\check{\mathbf{x}};\overset{\circ}{\check{\mathbf{x}}}, V_\mathrm{T}). \quad (32)$$

We also know that $\wp(\overset{\circ}{\check{\mathbf{x}}}|\overleftarrow{\mathbf{O}})$ will be a Gaussian distribution. To see why this is the case, we rewrite Eq. (29) as

$$d\overset{\circ}{\check{\mathbf{x}}} = A\overset{\circ}{\check{\mathbf{x}}}dt + \bar{E}d\bar{\mathbf{v}}_\mathrm{p}, \quad (33)$$

where $\bar{E}d\bar{\mathbf{v}}_\mathrm{p} = \sum_{\mathrm{r}\in\{\mathrm{o,u}\}}\mathcal{K}_\mathrm{r}^+[V_\mathrm{T}]d\mathbf{w}_\mathrm{r}$, with the observed measurement current

$$\mathbf{y}_\mathrm{o}dt = C_\mathrm{o}\overset{\circ}{\check{\mathbf{x}}}dt + d\mathbf{w}_\mathrm{o}, \quad (34)$$

where it is clear that $\overset{\circ}{\check{\mathbf{x}}}$ is described by a linear Langevin equation, as in Eq. (2), and by conditioning on a linear measurement current the resulting PDF will be Gaussian.

As $\wp(\overset{\circ}{\check{\mathbf{x}}}|\overleftrightarrow{\mathbf{O}})$ is a classical object, we can simply apply classical smoothing theory in order to compute this PDF, and since this PDF is Gaussian, $\wp(\overset{\circ}{\check{\mathbf{x}}}|\overleftrightarrow{\mathbf{O}}) = g(\overset{\circ}{\check{\mathbf{x}}};\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}, \overset{\circ}{V}_\mathrm{S})$, we only need to determine the mean $\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}$ and covariance $\overset{\circ}{V}_\mathrm{S}$. At this point there are two paths we can take, i.e., we can use either the MFP or RTS smoothing technique on Eq. (33) to obtain equations for the haloed smoothed mean and covariance. As stated earlier, I will choose the latter. Applying Eqs. (9) and (10) to Eq. (33) results in

$$d\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S} = A\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}dt + \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1}(\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S} - \overset{\circ}{\check{\mathbf{x}}}_\mathrm{F})dt + \bar{\Gamma}^\top d\mathbf{w}_\mathrm{S}, \quad (35)$$

$$\frac{d\overset{\circ}{V}_\mathrm{S}}{dt} = (\bar{A} + \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1})\overset{\circ}{V}_\mathrm{S} + \overset{\circ}{V}_\mathrm{S}(\bar{A} + \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1})^\top - \bar{D}, \quad (36)$$

with final conditions $\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}(T) = \overset{\circ}{\check{\mathbf{x}}}_\mathrm{F}(T)$ and $\overset{\circ}{V}_\mathrm{S}(T) = \overset{\circ}{V}_\mathrm{F}(T)$. Here $\bar{A} = A - \bar{\Gamma}^\top C_\mathrm{o}$, $\bar{D} = \bar{E}\bar{E}^\top - \bar{\Gamma}^\top\bar{\Gamma}$, $\bar{\Gamma}^\top dt = \bar{E}d\bar{\mathbf{v}}_\mathrm{p}d\mathbf{w}_\mathrm{o}^\top = \mathcal{K}_\mathrm{o}^+[V_\mathrm{T}]dt$, and $d\mathbf{w}_\mathrm{S} = \mathbf{y}_\mathrm{o}dt - C_\mathrm{o}\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}dt$. Additionally, I have introduced both the haloed filtered mean $\overset{\circ}{\check{\mathbf{x}}}_\mathrm{F} = \mathbb{E}_{|\overleftarrow{\mathbf{O}}}\{\overset{\circ}{\check{\mathbf{x}}}\}$ and covariance $\overset{\circ}{V}_\mathrm{F} = \mathbb{E}_{\overset{\circ}{\check{\mathbf{x}}}|\overleftarrow{\mathbf{O}}}\{\overset{\circ}{\check{\mathbf{x}}}\overset{\circ}{\check{\mathbf{x}}}^\top\} - \overset{\circ}{\check{\mathbf{x}}}_\mathrm{F}\overset{\circ}{\check{\mathbf{x}}}_\mathrm{F}^\top$,

which are obtained from the filtered PDF $\wp(\overset{\circ}{\check{\mathbf{x}}}|\overleftarrow{\mathbf{O}})$. Note, in Eqs. (35) and (36) I have assumed that $\overset{\circ}{V}_\mathrm{F}$ is invertible, which may not always be the case. In the event that $\overset{\circ}{V}_\mathrm{F}$ is not invertible, the smoothed quantum state can still be computed with slight modifications; see Appendix for details.

While the haloed filtered mean and covariance do satisfy their own differential equations, it has been shown [27] that $\overset{\circ}{\check{\mathbf{x}}}_\mathrm{F} = \langle\hat{\mathbf{x}}\rangle_\mathrm{F}$ and $\overset{\circ}{V}_\mathrm{F} = V_\mathrm{F} - V_\mathrm{T}$, thus the specific equations are not important for computing the haloed smoothed mean and covariance matrices. However, for simplicity, the haloed covariance matrix will be used more often for simplicity. It should also be emphasized that $\overset{\circ}{V}_\mathrm{F}$ is the covariance of a classical state and thus is not required to satisfy the Schrödinger-Heisenberg uncertainty relation, unlike V_F and V_T.

Finally, we have all the necessary information to compute the smoothed quantum state for LGQ systems. Returning to Eq. (32), we can substitute in the Gaussian PDF, $\wp(\overset{\circ}{\check{\mathbf{x}}}|\overleftrightarrow{\mathbf{O}}) = g(\overset{\circ}{\check{\mathbf{x}}};\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}, \overset{\circ}{V}_\mathrm{S})$, obtaining

$$g(\check{\mathbf{x}}; \langle\hat{\mathbf{x}}\rangle_\mathrm{S}, V_\mathrm{S}) = \int d^{2N}\overset{\circ}{\check{\mathbf{x}}}\, g(\overset{\circ}{\check{\mathbf{x}}};\overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}, \overset{\circ}{V}_\mathrm{S})g(\check{\mathbf{x}};\overset{\circ}{\check{\mathbf{x}}}, V_\mathrm{T}), \quad (37)$$

where, using the fact that convolving two Gaussian functions will result in another Gaussian, I have replaced the Wigner function of the smoothed state with its Gaussian with mean $\langle\hat{\mathbf{x}}\rangle_\mathrm{S}$ and covariance matrix V_S. Using the properties of such a convolution, we find that $\langle\hat{\mathbf{x}}\rangle_\mathrm{S} = \overset{\circ}{\check{\mathbf{x}}}_\mathrm{S}$ and $V_\mathrm{S} = \overset{\circ}{V}_\mathrm{S} + V_\mathrm{T}$. Thus, using $\frac{dV_\mathrm{S}}{dt} = \frac{d\overset{\circ}{V}_\mathrm{S}}{dt} + \frac{dV_\mathrm{T}}{dt}$, we obtain the RTS form of the quantum state smoothing equations,

$$d\langle\hat{\mathbf{x}}\rangle_\mathrm{S} = A\langle\hat{\mathbf{x}}\rangle_\mathrm{S}dt + \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1}(\langle\hat{\mathbf{x}}\rangle_\mathrm{S} - \langle\hat{\mathbf{x}}\rangle_\mathrm{F})dt + \mathcal{K}_\mathrm{o}^+[V_\mathrm{T}]d\mathbf{w}_\mathrm{S}, \quad (38)$$

$$\frac{dV_\mathrm{S}}{dt} = (\bar{A} + \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1})V_\mathrm{S} + V_\mathrm{S}(\bar{A} + \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1})^\top + Q, \quad (39)$$

with the final condition $V_\mathrm{S}(T) = \overset{\circ}{V}_\mathrm{F}(T) + V_\mathrm{T}(T) = V_\mathrm{F}(T)$ and

$$Q = D - \Gamma_\mathrm{o}^\top\Gamma_\mathrm{o} + V_\mathrm{T}C_\mathrm{o}^\top C_\mathrm{o}V_\mathrm{T} - \bar{D}\overset{\circ}{V}_\mathrm{F}^{-1}V_\mathrm{T} - V_\mathrm{T}\overset{\circ}{V}_\mathrm{F}^{-1}\bar{D} - 2\bar{D}. \quad (40)$$

One advantage of the RTS form of the LGQ state smoothing equations is that there is no need to compute the retrofiltered effect, and we only have to focus on computing the necessary quantities to obtain the smoothed state. While this does not reduce the overall number of differential equations (ordinary or stochastic) that need to be solved to compute the smoothed quantum state, it does remove the extra step of computing the smoothed mean and covariance using the MFP equations. However, the greatest advantage lies in the fact that the RTS equations are dynamical equations that can provide insight into how the smoothed state evolves. This is hidden in the MFP forms and also in the general expression in Eq. (1).

IV. DIFFERENTIABILITY OF THE QUANTUM SMOOTHED MEAN

To show just one of the perks of having the dynamical equations, I will investigate nondifferentiability of the smoothed mean. It may not be terribly surprising that the mean of the smoothed quantum state is nondifferentiable as this is

also the case in classical smoothing. However, there is a slight, perhaps predictable difference between the innovation terms in the classical and quantum cases. Specifically, we see that, in the quantum case, there is a dependence on the true covariance V_T, which is reasonable as this is the minimum uncertainty in the mean for the quantum system given Alice's and Bob's measurements. Thus it is reasonable to expect that Alice's measurement would change her estimate of the mean by at least an amount proportional to V_T. In contrast, the innovation in the classical case only depends on the cross-correlation matrix Γ since the true covariance is zero, corresponding to a δ function PDF, where the Schrödinger-Heisenberg uncertainty relation prevents $V_T \to 0$ in quantum systems. Furthermore, because of the uncertainty relation, one might expect that the quantum smoothed mean will always be nondifferentiable. This is not the case.

In the classical case, all that is required for the smoothed mean $\mathbf{x}_S$ to be differentiable is that the cross-correlation matrix $\Gamma = 0$, meaning that Alice's measurement is uncorrelated with the noise affecting the system. However, this condition is not sufficient in general for the quantum case. To find the condition for differentiability in the quantum case, we will assume that the mean is differentiable, that is, $d\langle\hat{\mathbf{x}}\rangle_S \propto dt$, over the interval $[\tau_1, \tau_2]$. It is easy to see by looking at Eq. (38) that the only term that is preventing the mean from being differentiable is the innovation term $\mathcal{K}_o^+[V_T]d\mathbf{w}_S$. Thus, in order for the mean to be differentiable, this innovation term must either be proportional to dt or vanish over the interval. It is impossible for the former to be true as for the innovation term to be proportional to dt, since $d\mathbf{w}_S^\top d\mathbf{w}_S \propto dt$, it must be the case that $\mathcal{K}_o^+[V_T] = Rd\mathbf{w}_S^\top$, where R is an arbitrary matrix. This cannot be the case as V_T, C_o, and Γ_o are all independent of the observed measurement at the estimation time. The latter, on the other hand, is possible and occurs when $\mathcal{K}_o^+[V_T] = 0$. Note, one could also make this term vanish by considering $K_o^+[V_T] \propto dt$. However, this means that C_o and Γ_o are of order dt, as V_T cannot be by the Schrödinger-Heisenberg uncertainty relation, which causes the filtered mean, and consequently the smoothed mean, to reduce the unconditioned estimate. Thus, this case can be ignored. Since we are considering a fixed measurement scheme, i.e., C_o and Γ_o are time-independent, it is impossible for $\mathcal{K}_o^+[V_T] = 0$ to be satisfied at all times. Thus, to have a differentiable smoothed mean over a noninfinitesimal time interval, I will only consider time intervals in the steady-state regime, i.e., $\tau_1 \geqslant \tau^{ss}$, where τ^{ss} is the time taken for the true covariance to reach steady state.

Under this condition, the steady state of the true covariance V_T^{ss} satisfies

$$0 = AV_T^{ss} + V_T^{ss}A^\top + D - \mathcal{K}_u^+[V_T^{ss}]\mathcal{K}_u^+[V_T^{ss}]^\top. \quad (41)$$

At this point, we can see that the steady-state covariance of the true state is identical to the steady state of a single record filtered state, like Eq. (26). However, rather than using Alice's past record, it is Bob's past record that is used. For comparison, Bob's filtered state is $W_{\overleftarrow{U}}(\check{\mathbf{x}}) = g(\check{\mathbf{x}}; \langle\hat{\mathbf{x}}\rangle_{\overleftarrow{U}}, V_{\overleftarrow{U}})$, with

$$d\langle\hat{\mathbf{x}}\rangle_{\overleftarrow{U}} = A\langle\hat{\mathbf{x}}\rangle_{\overleftarrow{U}}dt + \mathcal{K}_u^+[V_{\overleftarrow{U}}]d\mathbf{w}_{\overleftarrow{U}}, \quad (42)$$

$$\frac{dV_{\overleftarrow{U}}}{dt} = AV_{\overleftarrow{U}} + V_{\overleftarrow{U}}A^\top + D - \mathcal{K}_u^+[V_{\overleftarrow{U}}]\mathcal{K}_u^+[V_{\overleftarrow{U}}], \quad (43)$$

where $d\mathbf{w}_{\overleftarrow{U}} = \mathbf{y}_u dt - C_u\langle\hat{\mathbf{x}}\rangle_{\overleftarrow{U}}dt$. Thus, we have that if the mean of the smoothed quantum state is differentiable in the steady-state regime, then $V_T^{ss} = V_{\overleftarrow{U}}^{ss}$.

Importantly, the converse of this is also true, that is, if $V_T^{ss} = V_{\overleftarrow{U}}^{ss}$, then the mean of the smoothed quantum state is differentiable. This is simple to see since the steady state of the true covariance in general satisfies

$$0 = AV_T^{ss} + V_T^{ss}A^\top + D - \mathcal{K}_o^+[V_T^{ss}]\mathcal{K}_o^+[V_T^{ss}]^\top$$
$$- \mathcal{K}_u^+[V_T^{ss}]\mathcal{K}_u^+[V_T^{ss}]^\top, \quad (44)$$

and the steady-state of Bob's filtered covariance satisfies

$$0 = AV_{\overleftarrow{U}}^{ss} + V_{\overleftarrow{U}}^{ss}A^\top + D - \mathcal{K}_u^+[V_{\overleftarrow{U}}^{ss}]\mathcal{K}_u^+[V_{\overleftarrow{U}}^{ss}]^\top. \quad (45)$$

Since $V_T^{ss} = V_{\overleftarrow{U}}^{ss}$, Eq. (44) reduces to $\mathcal{K}_o^+[V_T^{ss}]\mathcal{K}_o^+[V_T^{ss}]^\top = 0$ giving $\mathcal{K}_o^+[V_T^{ss}] = 0$. As we have already shown, under this condition $d\langle\hat{\mathbf{x}}\rangle_S \propto dt$ and it is differentiable. As a result, we have the necessary and sufficient condition that $V_T^{ss} = V_{\overleftarrow{U}}^{ss}$ for a differentiable mean of the smoothed quantum state in the steady-state regime.

Note, this is true more generally outside the steady-state regime, when $V_T(t) = V_{\overleftarrow{U}}(t)$ over a time interval $[\tau_1, \tau_2]$. To see this, we note that since $V_T(t) = V_{\overleftarrow{U}}(t)$, their derivatives at all times in the interval are also equal. Subtracting Eq. (30) from Eq. (43) gives $\mathcal{K}_o^+[V_T]\mathcal{K}_o^+[V_T]^\top = 0$ and thus $\mathcal{K}_o^+[V_T] = 0$ for all times in the interval. Hence, in the general case, it is sufficient for a differentiable smoothed mean over the interval $[\tau_1, \tau_2]$ if $V_T(t) = V_{\overleftarrow{U}}(t)$. However, it is not a necessary condition.

This differentiability condition contains useful information about the types of systems and measurements that could exhibit a differentiable smoothed mean. Specifically, we see from $V_T^{ss} = V_{\overleftarrow{U}}^{ss}$ that Alice's measurement cannot reduce the uncertainty of the true state beyond what Bob's measurement already has. This immediately means that the quantum system must have more than one Lindblad operator, otherwise the portion of the channel that Alice monitors will always reduce the uncertainty by a fraction more when combine with Bob's measurement.

As an example, I will consider a single-mode ($N = 1$) open quantum system described by the master equation,

$$\hbar\frac{d\rho}{dt} = -i\chi[(\hat{q}\hat{p} + \hat{p}\hat{q})/2, \rho] + \gamma\mathcal{D}[\hat{q} + i\hat{p}]\rho + \kappa\mathcal{D}[\hat{q}]\rho. \quad (46)$$

Note, while this is a toy system to illustrate the differentiability condition for quantum state smoothing, in principle it could be constructed using linear optics. In fact, this system is similar to an optical parametric oscillator [18,19], where the only difference is that the additional Lindblad operator $\hat{c}_2 = \hat{q}$, as mentioned before, is required. In this system, since the variance in the $\hat{p}$ quadrature is bounded by the squeezing Hamiltonian, we can expect that, when κ is large enough, the amount of information about the position quadrature gained by monitoring $\hat{c}_2$ would reduce the uncertainty in $\hat{q}$ by enough so that the state is pure. Thus, for this system, if Bob were
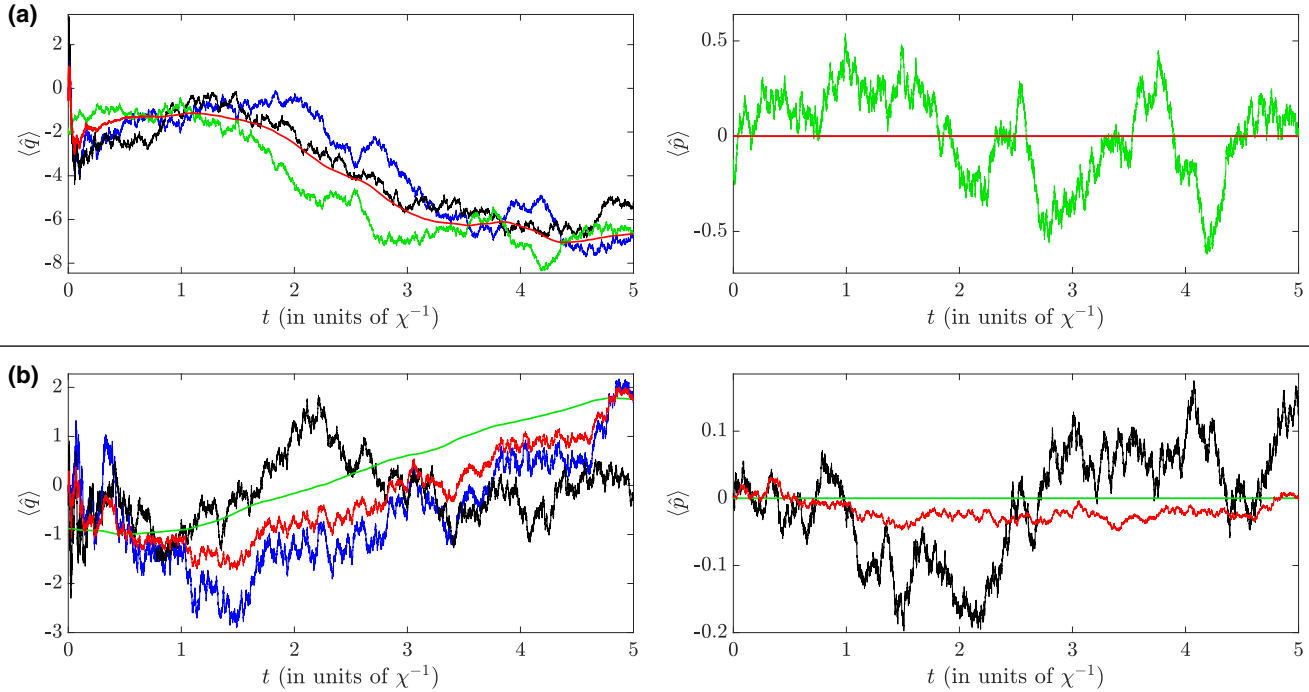
FIG. 1. (a) A single realization of the q- and p-quadratures (left and right, respectively) for the quantum system Eq. (46). The top two graphs are when Alice is measuring the γ-channel with homodyne phase $\theta_{\gamma,\mathrm{o}} = \pi/8$ and Bob is measuring the κ-channel with $\theta_{\kappa,\mathrm{u}} = 0$, where $g = 1$. The quantum smoothed mean (red line) initially evolves stochastically. However, at $t \approx 0.8$, a sufficient time has passed for the system to reach steady state and the quantum smoothed mean begins to evolve continuously since $V_{\mathrm{T}}^{\mathrm{ss}} = V_{\underleftarrow{\mathbf{U}}}^{\mathrm{ss}}$. Whereas the other three estimates—the true mean (black line), the filtered mean (blue line), and the classically smoothed mean (green line)—naively applied to the quantum system, all evolve stochastically. For this case, $\langle \hat{p} \rangle_{\mathrm{T}}(t) = \langle \hat{p} \rangle_{\mathrm{F}}(t) = 0$. (b) A single realization of the q- and p-quadratures (left and right, respectively) for the reverse scenario is considered for Alice and Bob. That is, Alice measures the κ-channel with $\theta_{\kappa,\mathrm{o}} = 0$ and Bob measures the γ-channel with homodyne phase $\theta_{\gamma,\mathrm{u}} = \pi/8$, in this case with $g = 0.1$. While Alice's measurement has a zero cross-correlation matrix, i.e., $\Gamma_{\mathrm{o}} = 0$, the quantum smoothed mean (red line) evolves stochastically, since $V_{\mathrm{T}}^{\mathrm{ss}} \neq V_{\underleftarrow{\mathbf{U}}}^{\mathrm{ss}}$. I have included the true mean (black line) and filtered mean (blue line) for completeness, where $\langle \hat{p} \rangle_{\mathrm{F}}(t) = 0$. Also, in this case, since Alice's cross-correlation matrix is zero, the classically smoothed mean evolves continuously. In both cases, $\hbar = 2$ and the initial conditions are taken to be $\langle \hat{\mathbf{x}} \rangle_0 = (0, 0)^\top$ and $V_0 = \mathrm{diag}(10, (1 + g)/2)$, where V_0 was chosen to be a finite version of the unconditioned steady state.

to perfectly monitor $\hat{c}_2$ (the κ-channel), with Alice perfectly monitoring $\hat{c}_1$ (the γ-channel), then the steady state of the true state would be pure, as both Alice and Bob have performed a perfect monitoring, and the steady state of Bob's filtered state will also be pure with the same covariance, satisfying the differentiability condition.

To make this more formal, time will be measured in units of χ^{-1} and I will consider the case $\gamma = \chi$. For this system, the drift and diffusion matrices are $A = \mathrm{diag}(0, -2)$ and $D = \hbar \cdot \mathrm{diag}(1, 1 + g)$, where $g = \kappa/\chi$. I will assume that Alice and Bob both perform homodyne measurements. For homodyne measurements, the matrix $M_{\mathrm{r}} = \mathrm{diag}(\sqrt{\eta_{\gamma,\mathrm{r}}} \exp[i\theta_{\gamma,\mathrm{r}}], \sqrt{\eta_{\kappa,\mathrm{r}}} \exp[i\theta_{\kappa,\mathrm{r}}])$, with the resulting measurement matrix

$$C_{\mathrm{r}} = 2\sqrt{\hbar^{-1}} \begin{bmatrix} \sqrt{\eta_{\gamma,\mathrm{r}}} \cos \theta_{\gamma,\mathrm{r}} & \sqrt{\eta_{\gamma,\mathrm{r}}} \sin \theta_{\gamma,\mathrm{r}} \\ \sqrt{g\eta_{\kappa,\mathrm{r}}} \cos \theta_{\kappa,\mathrm{r}} & 0 \end{bmatrix}, \qquad (47)$$

and cross-correlation matrix

$$\Gamma_{\mathrm{r}} = -\sqrt{\hbar} \begin{bmatrix} \sqrt{\eta_{\gamma,\mathrm{r}}} \cos \theta_{\gamma,\mathrm{r}} & \sqrt{\eta_{\gamma,\mathrm{r}}} \sin \theta_{\gamma,\mathrm{r}} \\ 0 & \sqrt{g\eta_{\kappa,\mathrm{r}}} \sin \theta_{\kappa,\mathrm{r}} \end{bmatrix}. \qquad (48)$$

Here, $\eta_{n,\mathrm{r}}$ and $\theta_{n,\mathrm{r}}$ are the measurement efficiency and homodyne phase, respectively, for each channel $n \in \{\gamma, \kappa\}$ and observer $\mathrm{r} \in \{\mathrm{o}, \mathrm{u}\}$.

To demonstrate the smooth evolution of the quantum smoothed mean, Bob must measure the position quadrature of the κ-channel perfectly, i.e., $\eta_{\gamma,\mathrm{u}} = 0$, $\eta_{\kappa,\mathrm{u}} = 1$, and $\theta_{\kappa,\mathrm{u}} = 0$. Crucially, g is chosen to be unity as this is the point where the steady-state solution for Bob's filtered covariance is pure and is equal to the true state. Alice, on the other hand, perfectly monitors the γ channel, i.e., $\eta_{\gamma,\mathrm{o}} = 1$, with some homodyne phase $\theta_{\gamma,\mathrm{o}} = \pi/8$. Note, the particular homodyne phase for Alice does not matter in the slightest for the theory and was just chosen for the simulations. As Fig. 1(a) shows, after a sufficient time ($t \approx 0.8$) for the true covariance to reach its steady state, the quantum smoothed mean (red line) begins to evolve smoothly, as one would expect from a differentiable function, whereas the other three estimates, i.e., the true, the filtered, and the classically smoothed means, all evolve stochastically over the entire time interval. The classically smoothed mean was computed using Eqs. (9) and (10), where the measurement matrix C and the cross-correlation matrix Γ

have been replaced with Alice's measurement matrix C_o and cross-correlation matrix Γ_o.

Now, if we reverse the channels that Alice and Bob measure, i.e., $\eta_{\gamma,u} = 1$, $\theta_{\gamma,u} = \pi/8$ and $\eta_{\kappa,o} = 1$, $\theta_{\kappa,o} = 0$, we would expect the smoothed mean to be nondifferentiable throughout the entire evolution, like the filtered and true state. This is clearly the case, as seen in Fig. 1(b). Note, for this case $g \neq 1$ so that Alice's filtered and smoothed states do not reduce to the true state when the system reaches steady state.

The fact that the quantum smoothed mean evolves stochastically might bring into question whether this technique should be referred to as smoothing. While it is idiosyncratic that a "smoothing" technique does not provide a smooth estimate, I reiterate that the classically smoothed mean suffers from the same issue when $\Gamma_o \neq 0$, as seen in Fig. 1(b). I also remind the reader that the smoothing technique refers to using future measurement information as well as the past information to obtain an estimate [7–13] and not necessarily obtaining a smooth estimate. Thus, I believe there is no issue in referring to this quantum state estimation technique as smoothing.

V. CONCLUSION

In this paper, I have derived the Rauch-Tung-Striebel forms of the quantum state smoothing equations for LGQ systems. These forms make it easier to compute the smoothed quantum state, as there is no need to compute or even mention the retro-filtered effect, while also providing insight into the dynamics of the smoothed quantum state. From these equations, I have derived a necessary and sufficient condition for the quantum smoothed mean to be differentiable in the steady-state limit. These equations could prove very useful in the future in identifying more properties of the smoothed state. A good example in the existing literature could be to provide an explanation or even an analytic solution for the optimal measurement strategies for Alice and Bob presented in Ref. [28]. Additionally, there may even be uses for this smoothing technique outside of quantum mechanics. This result would hold for any classical linear Gaussian system with a minimum bound on the covariance, i.e., a system with an uncertainty relation. Lastly, it would be interesting to see if the sufficient condition $V_T(t) = V_{\overleftarrow{U}}(t)$ for a differentiable quantum smoothed mean is sufficient outside the LGQ regime.

APPENDIX: NONINVERTIBLE HALOED FILTERED COVARIANCE

To show how to treat the smoothed quantum state when $\overleftrightarrow{V}_F^{-1}$ does not exist, I will go back to the more general definitions and consider the definition of the filtered state in terms

of the true state, that is, $\rho_F = \mathbb{E}_{\overleftarrow{U}|\overleftarrow{O}}\{\rho_T\}$ [22]. Following similar steps to the derivation of the smoothed quantum state in Sec. III, I arrive at

$$W_F(\check{x}) = \int d^{2N}\overleftrightarrow{x}\, \wp(\check{x}|\overleftarrow{O})W_T(\check{x}). \tag{A1}$$

In particular, I will look at the probability distribution $\wp(\check{x}|\overleftarrow{O}) = g(\check{x}; \check{x}_F, \overleftrightarrow{V}_F)$, where the Gaussianity follows from a similar argument to $\wp(\check{x}|\overrightarrow{O})$. As $\overleftrightarrow{V}_F$ is a real symmetric matrix, we can make use of the eigendecomposition $\overleftrightarrow{V}_F = P^\top \Lambda P$, where P is a matrix containing the eigenvectors of $\overleftrightarrow{V}_F$, and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{2N})$, with λ_i being an eigenvalue of $\overleftrightarrow{V}_F$. Performing a change of basis to the eigenbasis, the Wigner function of the filtered state becomes

$$g(\check{z}; \langle\hat{z}\rangle_F, PV_F P^\top) = \int d^{2N}\overleftrightarrow{z}\, g(\overleftrightarrow{z}; \overleftrightarrow{z}_F, \Lambda)g(\check{z}; \overleftrightarrow{z}, PV_T P^\top), \tag{A2}$$

where $\check{z} = P\check{x}$, $\langle\hat{z}\rangle = P\langle\hat{x}\rangle_F$, $\overleftrightarrow{z} = P\overleftrightarrow{x}$, and $\overleftrightarrow{z}_F = P\overleftrightarrow{x}_F$. Now, since Λ is a diagonal matrix, the Gaussian PDF can be factorized as $g(\overleftrightarrow{z}; \overleftrightarrow{z}_F, \Lambda) = \prod_{k=1}^{2N} g(\overleftrightarrow{z}_k; \overleftrightarrow{z}_{F,k}, \lambda_k)$, and we can consider the scenario in which $\overleftrightarrow{V}_F$ has at least one zero eigenvalue.

Let us assume, without loss of generality, that $\lambda_k > 0\ \forall k \leqslant s$ and the remaining $2N - s$ eigenvalues are zero. Using the fact that $\lim_{\sigma\to 0} g(\check{x}; a, \sigma^2) = \delta(\check{x} - a)$, the Gaussian PDF becomes

$$g(\overleftrightarrow{z}; \overleftrightarrow{z}_F, \Lambda) = \prod_{k=1}^{s} g(\overleftrightarrow{z}_k; \overleftrightarrow{z}_{F,k}, \lambda_k) \prod_{j=s+1}^{2N} \delta(\overleftrightarrow{z}_j - \overleftrightarrow{z}_{F,j}). \tag{A3}$$

Computing the integral in Eq. (A2), we find that the transformed filtered mean and covariance are

$$\langle\hat{z}\rangle_F = [\overleftrightarrow{z}_{F,1}, \ldots, \overleftrightarrow{z}_{F,s}, \langle\hat{z}\rangle_{T,s+1}, \ldots, \langle\hat{z}\rangle_{T,2N}]^\top, \tag{A4}$$

$$PV_F P^\top = \Lambda + PV_T P^\top. \tag{A5}$$

Thus, when $\overleftrightarrow{V}_F$ has at least one zero eigenvalue, the corresponding components of the mean and covariance matrix are equal to the same components of the transformed true mean and covariance.

Moving on to quantum state smoothing, beginning with Eq. (37), if we perform the same change of basis as we did for the filtered state, we obtain

$$g(\check{z}; \langle\hat{z}\rangle_S, PV_S P^\top)$$
$$= \int d^{2N}\overleftrightarrow{z}\, g(\overleftrightarrow{z}; \overleftrightarrow{z}_S, P\overleftrightarrow{V}_S P^\top)g(\check{z}; \overleftrightarrow{z}, PV_T P^\top), \tag{A6}$$

where $\langle\hat{z}\rangle_S = P\langle\hat{x}\rangle_S$ and $\overleftrightarrow{z}_S = P\overleftrightarrow{x}_S$. However, unlike in the filtering case, $P\overleftrightarrow{V}_S P^\top$ is not necessarily a diagonal matrix and thus the Gaussian cannot be completely factorized. However, we can still expect the Gaussian to factorize as

$$g(\overleftrightarrow{z}; \overleftrightarrow{z}_S, P\overleftrightarrow{V}_S P^\top) = g(\overleftrightarrow{z}'; \overleftrightarrow{z}_S', \overleftrightarrow{V}_S') \prod_{j=s+1}^{2N} \delta(\overleftrightarrow{z}_j - \overleftrightarrow{z}_{S,j}), \tag{A7}$$

because conditioning the estimate of $\langle\hat{z}\rangle_T$ on more information cannot make those components of the probability distribution any more certain than a δ function. Here $\overleftrightarrow{z}'$, $\overleftrightarrow{z}_S'$ are the first s

components of $\breve{\mathbf{z}}$ and $\breve{\mathbf{z}}_S$, respectively, and $\breve{V}_S'$ is the first $s \times s$ block of $P\breve{V}_S P^\top$. Computing the integral in Eq. (A7) gives

$$\langle \hat{\mathbf{z}} \rangle_S = [\breve{\mathbf{z}}_S', \langle \hat{z} \rangle_{T,s+1}, \ldots, \langle \hat{z} \rangle_{T,2N}]^\top, \qquad (A8)$$

$$P V_S P^\top = \begin{bmatrix} \breve{V}_S' & 0 \\ 0 & 0 \end{bmatrix} + P V_T P^\top. \qquad (A9)$$

We see that, like the filtered state, the components of the smoothed mean that have an eigenvalue of zero for the $\breve{V}_F(t)$ are equal to the corresponding components of the true mean and similarly for the smoothed covariance matrix. The remaining components of the mean and covariance are computed using the remaining elements of the filtered mean and covariance. This now gives us a method to compute the smoothed quantum state when $\breve{V}_F^{-1}$ does not exist at a particular time t. In the event that this occurs, the MFP form in Ref. [27] will be more useful because when the inverse does not exist at time t, it will only affect the smoothed mean and covariance at t, which can easily be correct.

As an aside, this analysis highlights an interesting property of the smoothed quantum state. In the event that all the eigenvalues of the haloed filtered covariance are zero, using Eqs. (A8) and (A9) the smoothed mean and covariance will be equal to the true mean and true covariance, respectively. While this in itself is not particularly interesting, it becomes interesting when we consider the initial conditions for the smoothed state. As we have assumed throughout this paper, we have $V_F(t_0) = V_T(t_0) = V_0$, i.e., $\breve{V}_F(t_0) = 0$. Thus, we have that initially the smoothed quantum state must have mean $\langle \hat{\mathbf{x}} \rangle_S(t_0) = \langle \hat{\mathbf{x}} \rangle_T(t_0) = \langle \hat{\mathbf{x}} \rangle_0$ and covariance $V_S(t_0) = V_T(t_0) = V_0$, meaning that, in this case, the smoothed quantum state will always start in the same state as both the filtered and true quantum state. Moreover, this fact holds irrespective of whether the true state is pure or not. As pointed out earlier, we can see this occur in Fig. 1, where we also see that the classically smoothed state does not coincide with the filtered and true means initially. This is because, in the classical case, it is always assumed that the classical true state (the state of maximal knowledge) is a δ function causing the condition to be violated. Note, this does not mean that the quantum smoothed mean can be computed forward in time, just that it is constrained at both t_0 and T.

While the MFP form of the quantum state smoothing equations might be more useful in general, there is a special case when P is time-independent over an interval $[\tau_1, \tau_2]$ with $\lambda_k = 0$ for $k > s$. In this case, the Moore-Penrose pseudoinverse can be used in Eqs. (38) and (39) instead of the usual matrix inverse. Over this interval, following the reasoning prior to this, $\langle \hat{\mathbf{z}} \rangle_S$ and $\breve{V}_S$ will be of the forms Eq. (A8) and (A9), respectively, over the interval. I will show that taking the pseudoinverse of the haloed filtered covariance causes the smoothed mean and covariance to evolve as the corresponding components of the true state would, while the remaining components evolve in a similar manner to a system in which the inverse exists.

Beginning with the mean, the stochastic differential equation for the transformed mean is

$$d\langle \hat{\mathbf{z}} \rangle_S = \bar{A}^* \langle \hat{\mathbf{z}} \rangle_S dt + \bar{D}^* \Lambda^+ (\langle \hat{\mathbf{z}} \rangle_S - \langle \hat{\mathbf{z}} \rangle_F) dt + P\mathcal{K}_o^+[V_T]\mathbf{y}_o dt, \qquad (A10)$$

where I have used $P^\top P = I_{2N}$ and an asterisked matrix denotes a transformation by P, i.e., $F^* = PFP^\top$, and Λ^+ is the pseudoinverse of Λ. For a diagonal matrix, the pseudoinverse is simple to compute by inverting the nonzero elements and leaving the remaining elements unchanged. Looking at the kth component of the mean, we have

$$d\langle \hat{z}_k \rangle_S = \bar{A}_{k,\ell}^* \langle \hat{z}_\ell \rangle_S dt + (\bar{D}\Lambda)_{k,\ell} (\langle \hat{z}_\ell \rangle_S - \langle \hat{z}_\ell \rangle_F) dt$$
$$+ (P\mathcal{K}_o^+[V_T])_{k,\ell} y_{o,\ell} dt, \qquad (A11)$$

where for comparison, the evolution of the kth component of the transformed true mean is

$$d\langle \hat{z}_k \rangle_T = \bar{A}_{k,\ell}^* \langle \hat{z}_\ell \rangle_T dt + (P\mathcal{K}_o^+[V_T])_{k,\ell} y_{o,\ell} dt$$
$$+ (P\mathcal{K}_u^+[V_T])_{k,\ell} dw_{u,\ell}. \qquad (A12)$$

When $k > s$, since the evolution of the smoothed mean must be the same as the evolution of the true mean over the time interval, as $\langle \hat{z}_k \rangle_S(\tau_2) = \langle \hat{z}_k \rangle_T(\tau_2)$, it must be the case that the matrices must have the following block forms:

$$\bar{A}^* = \begin{bmatrix} \bar{A}_{00}^* & \bar{A}_{01}^* \\ 0 & \bar{A}_{11}^* \end{bmatrix}, \quad \bar{D}^* = \begin{bmatrix} \bar{D}_{00}^* & 0 \\ 0 & 0 \end{bmatrix}, \qquad (A13)$$

where the blocks are divided so that the diagonal matrices have dimensions $s \times s$ and $(2N - s) \times (2N - s)$.

We can understand why this must be the case because if $\bar{A}^*$ was of another form, the first term in Eq. (A11) would cause the evolution of the kth component to be influenced by components other than the true mean and hence would cause the kth components to deviate from the true mean. A similar reasoning follows for the form of $\bar{D}^*$ to eliminate the second term. Note, since the lower half of $\bar{D}^*$ is zero, we see, using $\bar{D} = \mathcal{K}_u^+[V_T]\mathcal{K}_u^+[V_T]^\top$, that $(\mathcal{K}_u^+[V_T])_{k,\ell} = 0$ for all ℓ when $k > s$. This will eliminate the final term in Eq. (A12) and the kth components will evolve identically.

All that remains is to show that $\bar{A}^*$ and $\bar{D}^*$ are of the forms in Eq. (A13). Consider the differential equation for the haloed filtered covariance [27,28],

$$\frac{d\breve{V}_F}{dt} = \bar{A}\breve{V}_F + \breve{V}_F\bar{A}^\top + \bar{D} - \breve{V}_F C_o^\top C_o \breve{V}_F. \qquad (A14)$$

Using $\breve{V}_F = P^\top \Lambda P$, we obtain

$$\frac{d\Lambda}{dt} = \bar{A}^*\Lambda + \Lambda\bar{A}^{*\top} + \bar{D}^* - \Lambda P^\top C_o^\top C_o P \Lambda, \qquad (A15)$$

and for the k, ℓth element we have

$$\frac{d\lambda_k}{dt}\delta_{k,\ell} = \lambda_\ell \bar{A}_{k,m}^* \delta_{m,\ell} + \lambda_k \delta_{k,m} \bar{A}_{m,\ell}^{*\top} + \bar{D}_{k,\ell}^*$$
$$- \lambda_k \lambda_\ell \delta_{k,m} (P C_o^\top C_o P^\top)_{m,n} \delta_{n,\ell}. \qquad (A16)$$

Looking at $k, \ell > s$ and remembering that the zero eigenvalues do not change over the interval, we have $\bar{D}_{k,\ell}^* = 0$. By looking at the k, kth element of $\bar{D}^*$ with $k > s$, we have $\bar{D}^* = \sum_\ell (P\mathcal{K}_u^+[V_T])_{k,\ell}^2 = 0$ and thus $(P\mathcal{K}_u^+[V_T])_{k,\ell} = 0$ for all ℓ, showing that $\bar{D}^*$ is of the form in Eq. (A13). Now, returning to Eq. (A16), if we consider the case in which $k > s$ and $\ell < s + 1$, we have $\lambda_\ell \bar{A}_{k,\ell}^* + \bar{D}_{k,\ell}^* = 0$. As we have just shown, in this regime $\bar{D}_{k,\ell}^* = 0$, and we know $\lambda_\ell \neq 0$, resulting in $\bar{A}_{k,\ell}^* = 0$. Thus $\bar{A}^*$ is also of the form in Eq. (A13) and

completes the proof that the components of the transformed smoothed mean with $k > s$ are equal to the true mean during the interval. Moving onto the haloed covariance matrix, consider that the transformed differential equation for the haloed smoothed mean is

$$\frac{d\bar{\bar{V}}_S^*}{dt} = (\bar{A}^* + \bar{D}^*\Lambda^+)\bar{\bar{V}}_S^* + \bar{\bar{V}}_S^*(\bar{A}^* + \bar{D}^*\Lambda^+)^\top - \bar{D}^*. \quad (A17)$$

It is easy to show using Eq. (A13)) that, given the forms of $\bar{A}^*$ and $\bar{D}^*$, the haloed covariance is of the form

$$\bar{\bar{V}}_S^* = \begin{bmatrix} (\bar{\bar{V}}_S^*)_{00} & 0 \\ 0 & 0 \end{bmatrix}, \quad (A18)$$

and will remain in this form over the inerval $[\tau_1, \tau_2]$.

[1] R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory, J. Basic Eng. **83**, 95 (1961).

[2] H. E. Rauch, Solutions to the linear smoothing problem, IEEE Trans. Autom. Contr. **8**, 371 (1963).

[3] H. E. Rauch, F. Tung, and C. T. Striebel, Maximum likelihood estimates of linear dynamic systems, AIAA J. **3**, 1445 (1965).

[4] D. Q. Mayne, A solution of the smoothing problem for linear dynamic systems, Automatica **4**, 73 (1966).

[5] D. C. Fraser, A new technique for the optimal smoothing of data, Ph.D. thesis, Massachusetts Institute of Technology, 1967.

[6] D. Fraser and J. Potter, The optimum linear smoother as a combination of two optimum linear filters, IEEE Trans. Autom. Contr. **14**, 387 (1969).

[7] H. L. Weinert, *Fixed Interval Smoothing for State Space Models* (Kluwer Academic, New York, 2001).

[8] S. Haykin, *Kalman Filtering and Neural Networks* (Wiley, New York, 2001).

[9] H. L. Van Trees and K. L. Bell, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*, 2nd ed. (Wiley, New York, 2013).

[10] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 4th ed. (Wiley, New York, 2012).

[11] G. A. Einicke, *Smoothing, Filtering and Prediction: Estimating the Past, Present and Future* (InTech Rijeka, Rijeka, 2012).

[12] B. Friedland, *Control System Design: An Introduction to State-space Methods* (Courier, New York, 2012).

[13] S. Särkkä, *Bayesian Filtering and Smoothing* (Cambridge University Press, Cambridge, 2013), Vol. 3.

[14] V. P. Belavkin, *Information, Complexity and Control in Quantum Physics*, edited by A. Blaquíere, S. Dinar, and G. Lochak (Springer, New York, 1987).

[15] V. P. Belavkin, Quantum continual measurements and a posteriori collapse on ccr, Commun. Math. Phys. **146**, 611 (1992).

[16] V. P. Belavkin, Measurement, filtering and control in quantum open dynamical systems, Rep. Math. Phys. **43**, A405 (1999).

[17] A. C. Doherty and K. Jacobs, Feedback control of quantum systems using continuous state estimation, Phys. Rev. A **60**, 2700 (1999).

[18] H. M. Wiseman and A. C. Doherty, Optimal Unravellings for Feedback Control in Linear Quantum Systems, Phys. Rev. Lett. **94**, 070405 (2005).

[19] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2010).

[20] M. Tsang, Optimal waveform estimation for classical and quantum systems via time-symmetric smoothing, Phys. Rev. A **80**, 033840 (2009).

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

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

[23] K. Ohki, A smoothing theory for open quantum systems: The least mean square approach, in *54th IEEE Conference on Decision and Control (CDC)* (IEEE, Piscataway, NJ, 2015), pp. 4350–4355.

[24] K. T. Laverick, A. Chantasri, and H. M. Wiseman, General criteria for quantum state smoothing with necessary and sufficient criteria for linear gaussian quantum systems, Quantum Stud.: Math. Found. **8**, 37 (2021).

[25] A. Chantasri, I. Guevara, and H. M. Wiseman, Quantum state smoothing: why the types of observed and unobserved measurements matter, New J. Phys. **21**, 083039 (2019).

[26] A. Chantasri, I. Guevara, K. T. Laverick, and H. M. Wiseman, Unifying theory of quantum state estimation using past and future information, Phys. Rep. **930**, 1 (2021).

[27] K. T. Laverick, A. Chantasri, and H. M. Wiseman, Quantum State Smoothing for Linear Gaussian Systems, Phys. Rev. Lett. **122**, 190402 (2019).

[28] K. T. Laverick, A. Chantasri, and H. M. Wiseman, Linear Gaussian quantum state smoothing: Understanding the optimal unravelings for Alice to estimate Bob's state, Phys. Rev. A **103**, 012213 (2021).

[29] T. Kailath and P. Frost, An innovations approach to least-squares estimation–part ii: Linear smoothing in additive white noise, IEEE Trans. Autom. Contr. **13**, 655 (1968).

[30] T. Kailath, The innovations approach to detection and estimation theory, Proc. IEEE **58**, 680 (1970).

[31] T. Kailath, Some new algorithms for recursive estimation in constant linear systems, IEEE Trans. Inf. Theor. **19**, 750 (1973).

[32] F. Badawi, A. Lindquist, and M. Pavon, A stochastic realization approach to the smoothing problem, IEEE Trans. Autom. Contr. **24**, 878 (1979).

[33] A. Chia and H. M. Wiseman, Complete parametrizations of diffusive quantum monitorings, Phys. Rev. A **84**, 012119 (2011).