Design of a Quantum Projection Filter

Qing Gao, Daoyi Dong, *Senior Member, IEEE,* Ian R. Petersen, *Fellow, IEEE,* and Steven X. Ding,

***Abstract*—This paper develops a quantum projection filtering approach to approximating the quantum filter equation based on *Itô* stochastic Taylor expansions and quantum information geometric techniques. The proposed approximation scheme is designed so that the truncated Taylor expansion of the difference between the true quantum trajectory and its approximation on a lower-dimensional differential submanifold is minimized through an orthogonal projection operation in the quantum fisher metric. In addition, a convenient design for a special class of open quantum systems is formulated. Simulation results from a numerical example demonstrate the approximation capability of the proposed quantum projection filter.**

***Index Terms*—Quantum filter; quantum projection filter; quantum information geometry; Itô stochastic Taylor expansions.**

I. Introduction

A quantum filter is a quantum stochastic differential equation that updates the posterior quantum density matrix based on the measurement records [1], [2], which is fundamentally useful in the implementation of many quantum technologies where online tracking of the state of a quantum system is essential [3]. Although the quantum filter has been successful experimentally for low-dimensional quantum plants, such as qubits and few-photon systems [4], as well as for near Gaussian systems, such as optical phase estimation and optomechanics [5], its implementation for general open quantum systems with high dimensionality has been beset with computation difficulties in practice. In general, for a quantum system that has a Hilbert space with dimension n, the number of elements of the quantum density matrix needed to be track online becomes $n^2 - 1$. A high computational burden arises when n is large. It is thus particularly useful to develop an approximation model for the quantum filter equation which is more computationally efficient.

In recent years, the problem of approximation or model reduction for the quantum filter has attracted great attention and there have been some successful approaches in the literature; to mention a few, see [6], [7], and [8]. A quantum version of the extended Kalman filter was developed in [6], where the system dynamics was linearized first and a Kalman filter was then designed for the linearized system. Like its classical counterpart, the approach in [6] works well for nearly linear quantum systems. An alternative approach was given in [7] based on the Volterra filter, which is applicable to complex nonlinear quantum systems, especially those with no simple Markovian models. In addition, a numerical approximation approach for online calculation of the quantum filter was developed in [8], in order to design a more efficient two-qubit feedback control scheme. It was demonstrated through simulation that using a small number of integration steps a high level of approximation accuracy can be achieved.

Among the existing results, one interesting approximation scheme is the quantum projection filtering approach which is motivated by the pioneering work on projection filtering for classical stochastic systems by Brigo, Hanzon and LeGland [9]. The basic idea of the projection filtering strategy is as follows. First, one designs a finite-dimensional differential submanifold embedded in the state space of the original filter and endows this submanifold with a metric structure (the Fisher metric is often used). Then, an orthogonal projection operation is defined in this metric and is used to construct a projected curve on the submanifold to approximate the original filter. Through a coordinate chart of the submanifold, this projected curve can be equivalently expressed as a set of dynamical equations which is called the *projection filter*. The first result on quantum projection filtering can be found in [10] where a highly nonlinear quantum model of a strongly coupled two-level atom in an optical cavity was considered. However, exact prior knowledge of an invariant set of the solution to the quantum filter equation was required in [10], which makes the approach in [10] less applicable in more complex cases. The work in [11] assumed no prior knowledge about the quantum density matrix and proposed an unsupervised learning identification algorithm to determine the structure of the submanifold instead. Unfortunately, the identification algorithm itself could be time consuming when dealing with a more general and complex open quantum system. An exponential quantum projection filtering approach for general open quantum systems can be found in recent work [12]. The quantum filter equation was converted into a Stratonovich form first, in order to make the stochastic differential equation compatible with the structure of differential manifolds. The coefficient functions of the quantum filter were then treated as tangent vectors along the quantum trajectory and were approximated using an orthogonal projection operation defined in the quantum Fisher metric.

In this paper, we propose a quantum projection filtering approach using *Itô* stochastic Taylor expansions and quantum information geometric techniques, motivated by the result in [13]. The quantum projection filter is designed by minimizing the truncated Taylor expansion of the difference between

This work was supported by the Alexander von Humboldt Foundation of Germany, the Australian Research Council's Discovery Projects funding scheme under the Projects DP190101566 and DP180101805, and the Air Force Office of Scientific Research under agreement number FA2386-16-1-4065.

Q. Gao and S. X. Ding are both with the Institute for Automatic Control and Complex Systems (AKS), University of Duisburg-Essen, 47057 Duisburg, Germany (e-mail: qing.gao.chance@gmail.com; steven.ding@uni-due.de).

D. Dong is with the School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2600, Australia (e-mail: daoyidong@gmail.com).

I. R. Petersen is with the Research School of Engineering, Australian National University, Canberra, ACT 2601, Australia (e-mail: i.r.petersen@gmail.com).

the original quantum trajectory and its approximation on a lower-dimensional submanifold in the mean square sense. In particular, a convenient design of the quantum projection filter is formulated for a special class of open quantum systems and is illustrated using simulation results from a numerical example.

Notation. i=$\sqrt{-1}$. Here we use the Roman type character i to distinguish the imaginary unit from the index i. For any two $n-$dimensional matrices A and B, $A\otimes B$ means the tensor product or Kronecker product of A and B, $[A,B]=AB-BA$ is the commutator of A and B. $\lambda_i(A), i=1,2,...,n$, are the eigenvalues of A, $\mathrm{Tr}(A)=\sum_{i=1}^n \lambda_i(A)$ is the trace of A, $A^\dagger$ represents the complex conjugate transpose of A, and $\|A\|_F=\sqrt{\mathrm{Tr}(A^\dagger A)}$ is the Hilbert-Schmidt norm of matrix A. I_p is the identity matrix with the dimensionality p. $\mathbb{R}^n$ represents the n-dimensional real vector space.

II. Problem Formulation

We first sketch the open quantum system model (see e.g., [1], [2] for more details). Let $\mathbb{H}_{\mathscr{Q}}$ be the Hilbert space of a finite-dimensional quantum system $\mathscr{Q}$ and suppose $\dim \mathbb{H}_{\mathscr{Q}}=n<\infty$. The quantum bath is modelled by a symmetric Fock space $\mathscr{E}$. Initially the composite quantum system consisting of the quantum system and the quantum bath is prepared in the state $\rho_{\mathrm{total}}=\rho_0\otimes\rho_{\mathrm{vacuum}}$, where ρ_0 is a state on $\mathbb{H}_{\mathscr{Q}}$ and ρ_{vacuum} is the vacuum state on the bath $\mathscr{E}$. The composite quantum system is supposed to be isolated and its temporal Heisenberg-picture evolution can be described by a unitary operator $U(t)$ on the tensor product Hilbert space $\mathbb{H}_{\mathscr{Q}}\otimes\mathscr{E}$, which satisfies the following Hudson-Parthasarathy equation[1]:

$$dU(t)=\left\{\left(-\mathrm{i}H-\frac{1}{2}L^\dagger L\right)dt+LdB^\dagger(t)-L^\dagger dB(t)\right\}U(t) \quad (1)$$

with the initial condition $U(0)=I$, where H is the system Hamiltonian, L is the coupling operator and $B(t)$ is the annihilation operator on $\mathscr{E}$.

A homodyne detector continuously monitors the field observable $Y_t=U^\dagger(t)(B(t)+B^\dagger(t))U(t)$ and generates a Wiener type classical signal. The observation process Y_t satisfies the so-called self-nondemolition property, i.e., $[Y_s,Y_t]=0$ for all $s\le t$. This enables us to interpret Y_t as a classical signal (photocurrent). Based on the *Itô* formula, Y_t satisfies

$$dY_t=U^\dagger(t)(L+L^\dagger)U(t)dt+d(B(t)+B^\dagger(t)). \quad (2)$$

With the system-observation pair of our model, i.e., Equations (1) and (2) , one can use quantum filtering theory [1] to optimally estimate the quantum state at time t using the following *Itô* quantum stochastic differential equation:

$$\begin{aligned} d\rho_t =&\mathscr{L}_{L,H}^\dagger(\rho_t)dt+(L\rho_t+\rho_t L^\dagger-\rho_t\,\mathrm{Tr}(\rho_t(L+L^\dagger)))\\ &\times\left(dY_t-\mathrm{Tr}(\rho_t(L+L^\dagger)dt\right), \quad (3)\end{aligned}$$

where $\mathscr{L}_{L,H}^\dagger(X)=-\mathrm{i}[H,X]+LXL^\dagger-\frac{1}{2}(L^\dagger LX+XL^\dagger L)$ is the adjoint Lindblad generator.

Equation (3) is known as the quantum filter or the quantum stochastic master equation [1], and is a classical Itô stochastic differential equation driven by the Wiener type classical photocurrent signal $Y(t)$. Its solution ρ_t represents the best real time knowledge about the quantum state upon continuous homodyne detection and is essential in the design of state feedback quantum controller [3]. However, online calculation of (3) is equivalent to solving a system of n^2-1 recursive *Itô* stochastic differential equations and thus tends to be computationally expensive, especially for high-dimensional quantum systems. It is the goal of this paper to construct a lower-dimensional approximation for the filter equation (3).

[1]We have assumed $\hbar$=1 by using atomic units in this paper.

III. Quantum Projection Filter Design

This section will present the design procedure of a quantum projection filter that serves as an approximation for the quantum filter equation (3) and provide numerical simulation results to illustrate the approximation performance.

A. Basic idea

Instead of dealing with the nonlinear quantum filter equation (3), in practice it is often much easier to manipulate the following equivalent unnormalized linear form of (3):

$$d\bar{\rho}_t=\mathscr{L}_{L,H}^\dagger(\bar{\rho}_t)dt+\mathscr{D}(\bar{\rho}_t)dY_t, \quad (4)$$

where $\mathscr{D}(X)=LX+XL^\dagger$. The information state $\bar{\rho}_t$ propagated by (4) is the unnormalized quantum state corresponding to ρ_t, which satisfies $\rho_t=\bar{\rho}_t/\mathrm{Tr}(\bar{\rho}_t)$ and $\bar{\rho}_0=\rho_0$. The photocurrent Y_t is a Wiener process with bounded drift under some classical probability measure $\mathscr{P}$. Using Girsanov's theorem, however, an alternative measure $\mathscr{P}'$ equivalent to $\mathscr{P}$ can be always constructed such that Y_t becomes a Wiener process with zero drift on a finite time interval $[0,T_f]$, where $T_f>0$ is a fixed time called the final time (see page 458 in [3]). Let $(\Omega,\mathscr{F},\mathscr{P}')$ be a complete classical probability space on which we have a right continuous and complete filtration $\{\mathscr{F}_t\}_{t\ge0}$ of sub-σ fields of $\mathscr{F}$. In the sequel, we use $\mathbb{E}\{\cdot\}$ to denote the mathematical expectation operator with respect to the given probability measure $\mathscr{P}'$.

The state space of (4) is given by the following set of nonnegative self-adjoint operators:

$$\mathbb{Q}=\{\bar{\rho}\,|\,\bar{\rho}\ge0,\bar{\rho}=\bar{\rho}^\dagger\}. \quad (5)$$

Naturally, $\mathbb{Q}$ can be regarded as an n^2-dimensional real differential manifold and the solution $\bar{\rho}_t$ to Equation (4) is a curve on $\mathbb{Q}$ starting from time zero, see Fig. 1.

Now, we summarize the design procedure of the proposed dimension reduction approximation strategy using Fig 1:

Step 1. To choose an $m-$dimensional differential submanifold $\mathbb{S}=\{\bar{\rho}_\theta\}$ embedded in $\mathbb{Q}$ such that the entire submanifold $\mathbb{S}$ can be covered by a single coordinate chart $(\mathbb{S},\theta=(\theta_1,...,\theta_m)^T\in\Theta)$ and $\rho_0\in\mathbb{S}$, where $m<n^2$ is a positive integer and Θ is an open subset of $\mathbb{R}^m$ containing the origin.

Step 2. To endow $\mathbb{S}$ with a Riemannian metric structure $\langle\cdot,\cdot\rangle$.

Step 3. To construct a real curve $\theta_t=(\theta_1(t),\theta_2(t),...,\theta_m(t))$ on $\mathbb{R}^m$ such that its homeomorphic image on $\mathbb{S}$ through the coordinate chart $(\mathbb{S},\theta)$, i.e., $\bar{\rho}_{\theta_t}$, is as close to $\bar{\rho}_t$ as possible based on the Riemannian metric defined in Step 2.

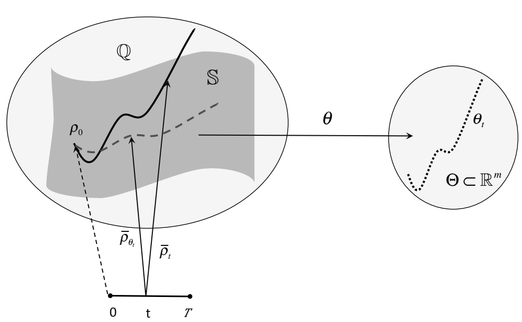

Fig. 1. Cartoon illustrating the basic setup of the optimal quantum projection filtering strategy. The dark line represents $\bar{\rho}_t$, the dashed line is $\bar{\rho}_{\theta_t}$ and the dotted line is θ_t.

For convenience of numerical implementation, θ_t is designed to satisfy the following It$\hat{o}$ stochastic differential equation:

$$d\theta_t = f(\theta_t)dt + g(\theta_t)dY_t, \tag{6}$$

where $f, g : \mathbb{R}^m \to \mathbb{R}^m$ are two real vector valued functions to be determined. Equation (6) is then the *quantum projection filter* to be designed in this paper.

We need to explain Step 3 further. To make the difference between two dynamical processes $\bar{\rho}_t$ and $\bar{\rho}_{\theta_t}$ as small as possible, the first requirement is that $\bar{\rho}_0 = \bar{\rho}_{\theta_0} = \rho_0$. Next, we need to minimize the following function of error named the subspace approximation error [14]:

$$e_t = \mathbb{E}\left\| d\bar{\rho}_t - d\bar{\rho}_{\theta_t}|_{\bar{\rho}_t = \bar{\rho}_{\theta_t}} \right\|_F^2, \tag{7}$$

with $e_0 = 0$. Then by integrating along time t, one can determine that the difference $\bar{\rho}_t - \bar{\rho}_{\theta_t}$ is minimized in the mean square sense. However, to minimize e_t in (7) is a mathematically tough task and we only provide an approximate solution in this paper.

Let t_0 and t be two stopping times satisfying

$$0 \leq t_0(\omega) \leq t(\omega) \leq T, \omega \in \Omega \tag{8}$$

with probability 1 and let k be a nonnegative integer. We will show later in Section IV-B that the quantum operator valued function $\bar{\rho}_\theta$ of θ has the following $k-$order It$\hat{o}$ stochastic Taylor expansion with respect to (6):

$$\bar{\rho}_{\theta_t} = \underbrace{\hat{\Upsilon}^k_{t_0,t}(\bar{\rho}_\theta)}_{k-\text{order truncation}} + O\left((2(t-t_0))^{\frac{k+1}{2}} \right) \tag{9}$$

with $\hat{\Upsilon}^0_{t_0,t}(\bar{\rho}_\theta) = \bar{\rho}_{\theta_{t_0}}$. Similarly, $\bar{\rho}_t$ admits the following It$\hat{o}$ stochastic Taylor expansion with respect to (4)

$$\bar{\rho}_t = \underbrace{\tilde{\Upsilon}^k_{t_0,t}(\bar{\rho})}_{k-\text{order truncation}} + O\left((2(t-t_0))^{\frac{k+1}{2}} \right) \tag{10}$$

with $\tilde{\Upsilon}^0_{t_0,t}(\bar{\rho}) = \bar{\rho}_{t_0}$.

It then follows from (9) and (10) that, for an positive integer $k \geq 1$ and a small time perturbation $0 < \sigma \ll 1$, the error function e_t in (7) satisfies

$$e_t = \mathbb{E}\left\| \bar{\rho}_{t+\sigma} - \bar{\rho}_{\theta_{t+\sigma}}|_{\bar{\rho}_t = \bar{\rho}_{\theta_t}} \right\|_F^2 \leq 2\mathbb{E}\left\| \tilde{\Upsilon}^k_{t,t+\sigma}(\bar{\rho}) - \hat{\Upsilon}^k_{t,t+\sigma}(\bar{\rho}_\theta) \right\|_F^2 \tag{11}$$

to a $k-$order approximation.

Therefore, the quantum projection filter equation in (6), or equivalently its two coefficient functions f and g at time t, can be then determined by solving the following optimization problem for $k = 1$ and $k = 2$:

Problem 3.1: $\min_{f,g} \mathbb{E}\left\| \hat{\Upsilon}^k_{t,t+\delta}(\bar{\rho}_\theta) - \tilde{\Upsilon}^k_{t,t+\delta}(\bar{\rho}) \right\|_F^2$.

Remark 3.1. By the quantum projection filtering strategy, one can minimize the truncated Taylor expansion of the difference between $\bar{\rho}_t$ and $\bar{\rho}_{\theta_t}$ up to order 2 in the mean square sense. This is because only two independent functions, i.e., f and g are used to solve Problem 3.1. Using a quantum projection filter with more complex formulation than that of (6) might lead to higher approximation accuracy.

B. Quantum projection filter

In this paper, the submanifold $\mathbb{S}$, which can be covered by a single coordinate chart $(\mathbb{S}, \theta = (\theta_1, ..., \theta_m)^T \in \Theta)$, is designed to be an $m-$dimensional smooth differential manifold consisting of an exponential family of unnormalized quantum density operators:

$$\mathbb{S} = \{\bar{\rho}_\theta\} = \left\{ e^{\frac{1}{2}\sum_{i=1}^m \theta_i A_i} \rho_0 e^{\frac{1}{2}\sum_{i=1}^m \theta_i A_i} \right\}, \tag{12}$$

where the submanifold operators $A_i \in \mathbb{A}, i \in \{1, 2, ..., m\}$ are pre-designed and mutually commutating.

By using this submanifold, an explicit design of the quantum projection filter (6) can be obtained.

Theorem 3.1. The solution to Problem 3.1 for $k = 1$ is

$$g(\theta_t) = R(\theta_t)^{-1}\Gamma(\theta_t), \tag{13}$$

where $\Gamma(\theta_t)$ is an $m-$dimensional column vector of real functions on θ_t and the jth element is given by

$$\Gamma_j(\theta_t) = \text{Tr}(\bar{\rho}_{\theta_t}(A_j L + L^\dagger A_j)), j \in \{1, 2, ..., m\}. \tag{14}$$

By choosing the coefficient function g as in (13), then Problem 3.1 for $k = 2$ is solved if

$$f(\theta_t) = R(\theta_t)^{-1}\Xi(\theta_t), \tag{15}$$

where $\Xi(\theta_t)$ is an $m-$dimensional column vector of real functions on θ_t and the jth element is given by

$$\Xi_j(\theta_t) = \text{Tr}(\bar{\rho}_{\theta_t}(\mathscr{L}_{L,H}(A_j)) - \frac{1}{2}g(\theta_t)^T \Delta_j g(\theta_t), \tag{16}$$

for $j \in \{1, 2, ..., m\}$. Here Δ_j is an $m \times m$ matrix of real function on θ_t and $\Delta_j(p, q) = \text{Tr}(\bar{\rho}_{\theta_t} A_p A_q A_j), p, q \in \{1, 2, ..., m\}$.

In fact, letting the submanifold operators A_i in (12) be mutually commutating leads to further simplification of the design scheme. On the finite dimensional Hilbert space $\mathscr{H}_{\mathscr{Q}}$, one can always find a unitary operator $\bar{U}$ by which the set of self-adjoint operators $\{A_i\}$ can be simultaneously diagonalized. That is, $A_i = \bar{U}^\dagger \text{Diag}(\lambda_1(A_i), \lambda_2(A_i), ..., \lambda_m(A_i))\bar{U}$. Let $\bar{\theta}$ be a linear transformation of θ:

$$\bar{\theta} = \begin{bmatrix} \bar{\theta}_1 \\ \bar{\theta}_2 \\ \vdots \\ \bar{\theta}_m \end{bmatrix} = \begin{bmatrix} \lambda_1(A_1) & \lambda_1(A_2) & \dots & \lambda_1(A_m) \\ \lambda_2(A_1) & \lambda_2(A_2) & \dots & \lambda_2(A_m) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_m(A_1) & \lambda_m(A_2) & \dots & \lambda_m(A_m) \end{bmatrix} \theta, \tag{17}$$

and let $\bar{A}_i = \bar{U}^\dagger \mathrm{Diag}\{\delta_{1i},..,\delta_{ni}\}\bar{U}, i = 1,2,...,m$ be a set of projection operators.

Remark 3.2. It is noted that since (17) does not necessarily suggest a homeomorphic mapping, thus the new vector $\bar{\theta}$ is not necessarily a coordinate system of the submanifold $\mathbb{S}$ either.

Suppose that $\mathrm{Tr}(\rho_0 \bar{A}_j) \neq 0, i = 1,2,...,m$. With the linear transformation (17), the submanifold (12) can be rewritten as

$$\mathbb{S} = \{\bar{\rho}_\theta\} := \left\{ e^{\frac{1}{2}\sum_{i=1}^m \bar{\theta}_i \bar{A}_i} \rho_0 e^{\frac{1}{2}\sum_{i=1}^m \bar{\theta}_i \bar{A}_i} \right\}, \tag{18}$$

and the quantum projection filter constructed in Theorem 3.1 yields the following simplified decoupled system of stochastic differential equations:

$$\begin{aligned} d\bar{\theta}_j(t) = &\left\{ \frac{\mathrm{Tr}(\bar{\rho}_{\theta_t}\mathscr{L}_{H,L}(\bar{A}_j))}{e^{\bar{\theta}_j(t)}\mathrm{Tr}(\rho_0 \bar{A}_j)} - \frac{1}{2}\left(\frac{\mathrm{Tr}(\bar{\rho}_{\theta_t}(\bar{A}_j L + L^\dagger \bar{A}_j)}{e^{\bar{\theta}_j(t)}\mathrm{Tr}(\rho_0 \bar{A}_j)}\right)^2 \right\} dt \\ &+ \frac{\mathrm{Tr}(\bar{\rho}_{\theta_t}(\bar{A}_j L + L^\dagger \bar{A}_j)}{e^{\bar{\theta}_j(t)}\mathrm{Tr}(\rho_0 \bar{A}_j)} dY(t), j = 1,2,...,m. \end{aligned} \tag{19}$$

This motivates a convenient design guideline for the submanifold. That is, the submanifold operators A_i could be chosen as a set of projectors satisfying $A_i A_j = \delta_{ij} P_j$. In particular, this design leads to a significantly simplified optimal quantum projection filter, for a special class of open quantum systems where the following assumption holds:

Assumption 3.1. The coupling operator L is self-adjoint, that is, $L = L^\dagger$.

This assumption is practically reasonable in many experimental settings; e.g., trapping a cold atomic ensemble in an optical cavity [2]. Since L is self-adjoint, it admits a spectral decomposition $L = \sum_{i=1}^{n_0} \lambda_i P_{L_i}$, where $n_0 \leq n$ is the number of nonzero eigenvalues of L, the set $\{\lambda_i\}$ contains all of the nonzero real eigenvalues of L, and $\{P_{L_i}\}$ is a set of projection operators that satisfies $P_{L_i} P_{L_j} = \delta_{ij} P_{L_j}$. Then one has the following result:

Theorem 3.2. Suppose that $\mathrm{Tr}(\rho_0 P_{L_i}) \neq 0, i = 1,2,...,m$. By designing the submanifold (12) according to

$$\begin{cases} m = n_0, \\ A_i = P_{L_i}, \end{cases} \tag{20}$$

the optimal quantum projection filter designed in Theorem 3.1 becomes

$$d\theta_j(t) = \left(\mathrm{i}\frac{\mathrm{Tr}(\bar{\rho}_{\theta_t}[H, A_j])}{e^{\bar{\theta}_j(t)}\mathrm{Tr}(\rho_0 A_j)} - 2\lambda_j^2 \right) dt + 2\lambda_j dY(t), \tag{21}$$

for $j \in \{1,2,...,m\}$. Moreover, the order 1 Itô stochastic Taylor expansion for $\bar{\rho}_t - \bar{\rho}_{\theta_t}$ is identically zero for all $t \geq 0$.

Proof. Theorem 3.2 can be derived from Theorem 3.1 and the proof is omitted. □

The following model reduction result is a corollary of Theorem 3.2.

Corollary 3.1. For open quantum systems with a self-adjoint coupling operator, if the system Hamiltonian $H = 0$, then $\bar{\rho}_{\theta_t} \equiv \bar{\rho}_t$ by designing the submanifold according to (20).

C. A numerical example

A numerical example is given in this subsection to demonstrate the approximation performance of the proposed quantum projection filtering approach. The evolution of the open quantum system is given as in (1) with

$$H = H_u u_t = \begin{bmatrix} 0 & -j & 0 & 0 \\ j & 0 & 0 & 0 \\ 0 & 0 & 0 & -j \\ 0 & 0 & j & 0 \end{bmatrix} u_t$$

and

$$L = \mathrm{Diag}\{1,-1,1,-1\},$$

where u_t is an external signal applied to the quantum system through the channel H_u. One example in practice is that u_t is the control input and H_u is the control channel. Initially the quantum state is given by $\rho_0 = \mathrm{Diag}\{1,2,3,4\}/10$.

Therefore the system is defined in the Hilbert space $\mathscr{H}_\mathscr{Q} = \mathbb{C}^4$ and generally **15** stochastic differential equations need to be solved online in order to solve the quantum filter equation (6) in time. However, since the coupling operator L is self-adjoint, it follows from Theorem 3.2 that one can design a differential submanifold according to (20) such that the quantum state can be approximately obtained by solving **4** stochastic differential equations. To be specific, the four submanifold operators can be given by

$$\begin{aligned} A_1 = P_{L_1} = \mathrm{Diag}\{1,0,0,0\}, A_2 = P_{L_2} = \mathrm{Diag}\{0,1,0,0\} \\ A_3 = P_{L_3} = \mathrm{Diag}\{0,0,1,0\}, A_4 = P_{L_4} = \mathrm{Diag}\{0,0,0,1\} \end{aligned}$$

respectively, where $P_{L_i}, i = 1,2,3,4$ are four projection operators of the coupling operator L corresponding to four eigenvalues $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 1$ and $\lambda_4 = -1$, respectively, through spectral decomposition.

In real experiments, the photocurrent Y_t is a classical signal generated by a homodyne detector and can be used to drive both the quantum filter equation (3) and the quantum projection filter equation (21). In order to simulate Y_t in our case, we replace the term $dY_t - \mathrm{Tr}(\rho_t(L + L^\dagger)dt$ by the instantaneous increment of a Wiener process $dW(t)$, based on the result on page 36 in [1]. Then the quantum filter (3) is calculated first and the photocurrent Y_t is simulated from $dY(t) = \mathrm{Tr}(\rho_t(L + L^\dagger))dt + dW(t)$. The external signal u_t is chosen to be $u_t = 5e^{-t}sin(2t)$ as shown in Fig. 2. Monte Carlo simulations have been conducted by using the discretization approach as in [17], in order to sovle the stochastic differential equations involved. The simulation parameters used are as follows: the simulation interval $t \in [0,T]$ with $T = 3$, the normally distributed variance is $\delta t = T/2^{12}$ with $N_0 = 2^{12}$, and the step size is chosen to be $\Delta t = \delta t$.

The performance of the proposed approximation filtering scheme is demonstrated by comparing the photocurrent simulated from $dY(t) = \mathrm{Tr}(\rho_t(L + L^\dagger))dt + dW(t)$ with the one simulated from $d\tilde{Y}(t) = \frac{\mathrm{Tr}(\bar{\rho}_{\theta_t}(L+L^\dagger))}{\mathrm{Tr}(\bar{\rho}_{\theta_t})}dt + dW(t)$. Simulation results from one particular experiment is presented in Fig. 3. It is also verified by a number of simulations that $Y(t) \equiv \tilde{Y}(t)$ when $u_t \equiv 0$, which coincides with Corollary 3.1.

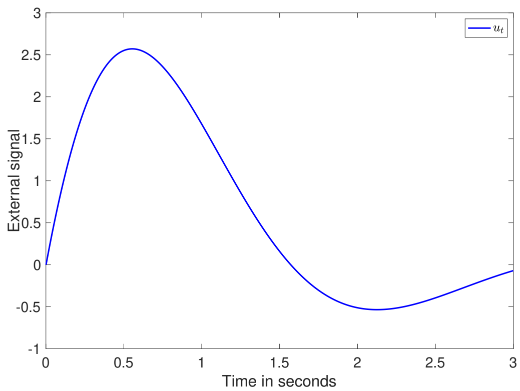

Fig. 2. The input signal u_t.

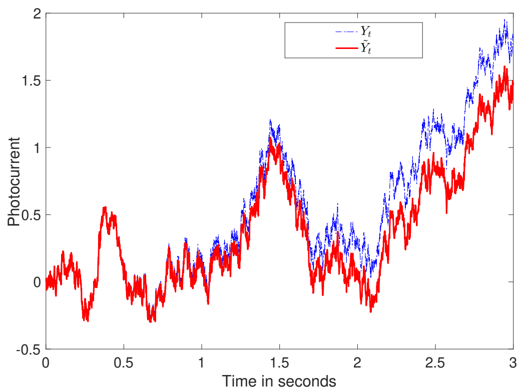

Fig. 3. Approximation performance of the quantum projection filter.

IV. Derivation of Theorem 3.1

This section will show the detailed derivation procedure of Theorem 3.1 following the basic idea sketched in Section III-A. After introducing two key technical tools in Sections IV-A and IV-B, the proof of Theorem 3.1 will be given in Section IV-C.

A. An orthogonal projection

We start from the geometric structure of the differential manifold $\mathbb{S}$ formulated in (12) using quantum information geometry theory [16]. Let $\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$ denote the tangent vector space of the manifold $\mathbb{S}$ at the point $\bar{\rho}_\theta$. The *mixture representation* ($m-$representation) of the natural basis vector of $\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$ can be obtained through a direct calculation as

$$\bar{\partial}_i := \frac{\partial \bar{\rho}_\theta}{\partial \theta_i} = \frac{1}{2}(A_i \bar{\rho}_\theta + \bar{\rho}_\theta A_i). \tag{22}$$

Assume that $\{\bar{\partial}_i\}$ are linearly independent. Then

$$\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S}) = \mathrm{Span}\{\bar{\partial}_i\}. \tag{23}$$

Now we endow $\mathbb{S}$ with a Riemannian metric. The symmetrized inner product is employed to define the inner product $\{\ll,\gg_{\bar{\rho}_\theta}, \bar{\rho}_\theta \in \mathbb{S}\}$ on $\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$:

$$\ll A, B \gg_{\bar{\rho}_\theta} = \frac{1}{2}\mathrm{Tr}(\bar{\rho}_\theta AB + \bar{\rho}_\theta BA), \forall A, B \in \mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S}). \tag{24}$$

Based on this inner product, we define another useful representation called the $e-representation$ of a tangent vector $X \in \mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$ as the self-adjoint operator $X^{(e)}$ satisfying

$$\ll X^{(e)}, A \gg_{\bar{\rho}_\theta} = \mathrm{Tr}\left(X^{(m)} A\right), \forall A \in \mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S}), \tag{25}$$

which implies that $\bar{\partial}_i^{(e)} = A_i$. Using this $e-$representation, we define an inner product $\langle,\rangle$ on $\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$ by

$$\begin{aligned}\langle X, Y\rangle_{\bar{\rho}_\theta} &= \ll X^{(e)}, Y^{(e)} \gg_{\bar{\rho}_\theta} \\ &= \mathrm{Tr}\left(X^{(m)} Y^{(e)}\right), \forall X, Y \in \mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S}).\end{aligned} \tag{26}$$

Then $r = \langle,\rangle$ forms a Riemmanian metric on $\mathbb{S}$ which may be regarded as a quantum version of the Fisher metric [18]. Each component of the quantum Fisher metric is given by a real-valued function of θ:

$$r_{ij}(\theta) = \langle \bar{\partial}_i, \bar{\partial}_j \rangle_{\bar{\rho}_\theta} = \ll \bar{\partial}_i^{(e)}, \bar{\partial}_j^{(e)} \gg_{\bar{\rho}_\theta} = \mathrm{Tr}(\bar{\rho}_\theta A_i A_j). \tag{27}$$

The quantum Fisher information matrix is an $m \times m$ dimensional real matrix given by $R(\theta) = (r_{ij}(\theta))$.

Let $\mathscr{T}_{\bar{\rho}_t}(\mathbb{Q})$ denote the tangent space of the higher dimensional differential manifold $\mathbb{Q}$ at each point $\bar{\rho}_t$. Apparently, $\mathscr{T}_{\bar{\rho}_t}(\mathbb{Q})$ is identified with the following set of all self-adjoint operators on the Hilbert space $\mathscr{H}_{\mathscr{Q}}$:

$$\mathbb{A} = \{A | A = A^\dagger\}. \tag{28}$$

Then an orthogonal projection operation $\Pi_{\bar{\rho}_\theta}$ from vectors in $\mathbb{A}$ to vectors in $\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$ can be defined as follows:

$$\begin{aligned}\Pi_{\bar{\rho}_\theta} : \mathbb{A} &\longrightarrow \mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S}) \\ \nu &\longmapsto \sum_{i=1}^{m}\sum_{j=1}^{m} r^{ij}(\theta) \langle \nu, \bar{\partial}_j \rangle_{\bar{\rho}_\theta} \bar{\partial}_i,\end{aligned} \tag{29}$$

where the matrix $\left(r^{ij}(\theta)\right)$ is the inverse of the quantum information matrix $R(\theta)$.

This projection operation $\Pi_{\bar{\rho}_\theta}$ in (29) maps any vector in $\mathbb{A}$ to their corresponding best approximation in $\mathscr{T}_{\bar{\rho}_\theta}(\mathbb{S})$ and is one of two key technical tools in our approach.

B. Itô stochastic Taylor expansions for quantum operator valued functions

In this subsection, we formulate It$\hat{o}$ stochastic Taylor expansions for two quantum operator valued functions $\bar{\rho}_{\theta_t}$ and $\bar{\rho}_t$, which serve as the other key technical tool in our approach. The basic idea and notation used follow from those of [15] where classical stochastic differential equations are considered.

The following finite list

$$\alpha = (j_1, j_2, ..., j_l), \tag{30}$$

is named a *multi-index* of length l, where l is a positive integer and $j_i \in \{0,1\}$ for $i \in \{1,2,...,l\}$. For completeness, we denote the empty list, or equivalently, the multi-index of length zero, by $()$. In addition, we write $l(\alpha)$, the length of α and $z(\alpha)$ the number of zeros in α, respectively. The set of all multi-indices is denoted by $\mathscr{M}$.

Given a multi-index $\alpha \in \mathscr{M}$ with $l(\alpha) \geq 1$, we define: $-\alpha$ to be the multi-index obtained by removing the first element of

α; $\alpha-$ the multi-index obtained by removing the last element of α; α_1 the first element of α; and α_{-1} the last element of α. Finally, for any two multi-indices $\alpha=(j_1,j_2,...,j_l)$ and $\beta=(\bar{j}_1,\bar{j}_2,...,\bar{j}_k)$, we define the concatenation operation $*$ on $\mathscr{M}$ by

$$\alpha*\beta=(j_1,j_2,...,j_l,\bar{j}_1,\bar{j}_2,...,\bar{j}_k). \tag{31}$$

The multi-indices introduced above can be used to enumerate stochastic integrals with respect to the Wiener process Y_t in (2) and time t. For any time t, we define $Y_t^1:=Y_t$ and $Y_t^0:=t$. The multi-integral for any function a of time associated with a multi-index α is defined by

$$\mathrm{I}^{\alpha}_{t_1,t_2}(a)=\begin{cases}a(t_2) & \alpha=()\\ \int_{t_1}^{t_2}\mathrm{I}^{\alpha-}_{t_1,s}(a)dY_s^{\alpha_{-1}} & \text{otherwise}\end{cases} \tag{32}$$

where $t_1\le t_2$ are two time points.

Now we are ready to present the Itô stochastic Taylor expansions for $\bar{\rho}_{\theta_t}$ and $\bar{\rho}_t$, respectively.

First, consider the self-adjoint operator valued function $\bar{\rho}_{\theta_t}:\mathbb{R}^m\to\mathbb{Q}$, where $\mathbb{Q}$ is the quantum state space defined in (5). Let $\mathscr{D}_\theta\bar{\rho}_\theta$ be the derivative of $\bar{\rho}_\theta$ with respect to the vector θ, which is given by

$$\mathscr{D}_\theta\bar{\rho}_\theta=\left(\frac{\partial\bar{\rho}_\theta}{\partial\theta_1},\frac{\partial\bar{\rho}_\theta}{\partial\theta_2},...,\frac{\partial\bar{\rho}_\theta}{\partial\theta_m}\right)^T. \tag{33}$$

Structure of higher order derivatives is thereby uniquely defined as

$$\mathscr{D}^i_{\theta^i}\bar{\rho}_\theta=\underbrace{\mathscr{D}_\theta(\mathscr{D}_\theta(...(\mathscr{D}_\theta\bar{\rho}_\theta)...))}_{i\text{ successive derivative operations}}. \tag{34}$$

A differential operator for $\bar{\rho}_{\theta_t}$ associated with a multi-index α is defined as

$$\mathrm{L}_\alpha(\bar{\rho}_\theta)=\begin{cases}\bar{\rho}_\theta & \alpha=()\\ \mathrm{L}^{\alpha_1}(\mathrm{L}_{-\alpha}(\bar{\rho}_\theta)) & \text{otherwise}\end{cases} \tag{35}$$

where

$$\begin{aligned}\mathrm{L}^0(\bar{\rho}_\theta) &= (f^T\otimes I_n)\mathscr{D}_\theta\bar{\rho}_\theta+\frac{1}{2}\left((g\otimes g)^T\otimes I_n\mathscr{D}^2_{\theta^2}\bar{\rho}_\theta\right),\\ \mathrm{L}^1(\bar{\rho}_\theta) &= (g^T\otimes I_n)\mathscr{D}_\theta\bar{\rho}_\theta.\end{aligned} \tag{36}$$

Applying Itô formula to (6), one has

$$d\bar{\rho}_{\theta_t}=\mathrm{L}^0(\bar{\rho}_{\theta_t})dt+\mathrm{L}^1(\bar{\rho}_{\theta_t})dY_t. \tag{37}$$

Next, consider the self-adjoint operator valued function $\bar{\rho}_t:\mathbb{R}\to\mathbb{Q}$. Like in (35), we define a differential operator for $\bar{\rho}_t$ associated with a multi index α as

$$\mathrm{D}_\alpha(\bar{\rho})=\begin{cases}\bar{\rho} & \alpha=()\\ \mathrm{D}^{\alpha_1}(\mathrm{D}_{-\alpha}(\bar{\rho})) & \text{otherwise}\end{cases} \tag{38}$$

with $\mathrm{D}^0=\mathscr{L}^\dagger_{L,H}$ and $\mathrm{D}^1=\mathscr{D}$ as defined below (3) and (4), respectively.

Remark 4.1. One observes that both the differential operators L_α and D_α contain a total of $l(\alpha)+z(\alpha)$ derivatives.

For any nonnegative integer k, we define a subset $\Lambda_k\subset\mathscr{M}$ to be

$$\Lambda_k=\{\alpha\in\mathscr{M}:l(\alpha)+z(\alpha)\le k\}. \tag{39}$$

Then we have the following result.

Theorem 4.1. Let t_0 and t be two stopping times with

$$0\le t_0(\omega)\le t(\omega)\le T,\omega\in\Omega \tag{40}$$

with probability 1. For any multi-index $\alpha\in\mathscr{M}$, suppose that $\mathbb{E}\|\mathrm{L}_\alpha(\bar{\rho}_{\theta_t})\|_F^2$ and $\mathbb{E}\|\mathrm{D}_\alpha(\bar{\rho}_t)\|_F^2$ are both bounded for all time t. Then for any nonnegative integer k, the Itô stochastic Taylor expansion of the function $\bar{\rho}_{\theta_t}$ at time t_0 is given by (9) with

$$\hat{\Upsilon}^k_{t_0,t}(\bar{\rho}_\theta)=\sum_{\alpha\in\Lambda_k}\mathrm{I}^\alpha_{t_0,t}(1)(\mathrm{L}_\alpha(\bar{\rho}_\theta)|_{t=t_0}); \tag{41}$$

the Itô stochastic Taylor expansion of the function $\bar{\rho}_t$ at time t_0 is given by (10) with

$$\tilde{\Upsilon}^k_{t_0,t}(\bar{\rho})=\sum_{\alpha\in\Lambda_k}\mathrm{I}^\alpha_{t_0,t}(1)(\mathrm{D}_\alpha(\bar{\rho})|_{t=t_0}). \tag{42}$$

Proof. See Appendix.

C. Proof of Theorem 3.1

This subsection proves Theorem 3.1 based on the orthogonal projection operation defined in (29) and Theorem 4.1.

Since $\bar{\rho}_{\theta_t}$ is independent of the increment $Y_{t+\delta}-Y_t$, it follows from Theorem 4.1 that for $k=1$

$$\begin{aligned}&\mathbb{E}\left\|\hat{\Upsilon}^k_{t,t+\delta}(\bar{\rho}_\theta)-\tilde{\Upsilon}^k_{t,t+\delta}(\bar{\rho})\right\|_F^2\\ =&\mathbb{E}\left\|\sum_{\alpha\in\Lambda_1\setminus\{()\}}\mathrm{I}^\alpha_{t,t+\delta}(1)(\mathrm{L}_\alpha(\bar{\rho}_{\theta_t})-\mathrm{D}_\alpha(\bar{\rho}_{\theta_t}))\right\|_F^2\\ =&\mathbb{E}\left\|\mathrm{L}_{(1)}(\bar{\rho}_{\theta_t})-\mathrm{D}_{(1)}(\bar{\rho}_{\theta_t})\right\|_F^2\mathbb{E}\left(\mathrm{I}^{(1)}_{t,t+\delta}(1)\right)^2\\ =&\mathbb{E}\left\|\mathrm{L}^1(\bar{\rho}_{\theta_t})-\mathrm{D}^1(\bar{\rho}_{\theta_t})\right\|_F^2\delta.\end{aligned} \tag{43}$$

Based on the orthogonal projection operation $\Pi_{\bar{\rho}_\theta}$ in (29), Problem 3.1 with $k=1$ is solved if

$$\begin{aligned}&\mathrm{L}^1(\bar{\rho}_{\theta_t})=(g^T\otimes I_n)\mathscr{D}_\theta\bar{\rho}_{\theta_t}\\ =&\Pi_{\bar{\rho}_\theta}(\mathrm{D}^1(\bar{\rho}_{\theta_t}))\\ =&\sum_{i=1}^m\sum_{j=1}^m r^{ij}(\theta)\left\langle\mathrm{D}^1(\bar{\rho}_{\theta_t}),\bar{\partial}_j\right\rangle_{\bar{\rho}_\theta}\bar{\partial}_i\\ =&\sum_{i=1}^m\sum_{j=1}^m r^{ij}(\theta)\mathrm{Tr}(\bar{\rho}_{\theta_t}(A_jL+L^\dagger A_j))\bar{\partial}_i\\ =&\left(\begin{bmatrix}\sum_{j=1}^m r^{1j}(\theta)\mathrm{Tr}(\bar{\rho}_{\theta_t}(A_jL+L^\dagger A_j))\\ \sum_{j=1}^m r^{2j}(\theta)\mathrm{Tr}(\bar{\rho}_{\theta_t}(A_jL+L^\dagger A_j))\\ \vdots\\ \sum_{j=1}^m r^{mj}(\theta)\mathrm{Tr}(\bar{\rho}_{\theta_t}(A_jL+L^\dagger A_j))\end{bmatrix}\otimes I_n\right)^T\mathscr{D}_\theta\bar{\rho}_{\theta_t},\end{aligned} \tag{44}$$

which yields (13).

Now let $k=2$ and suppose that g is already designed as in (13). A simple calculation shows that $\delta, Y_{t+\delta}-Y_t$ and

$\frac{(Y_{t+\delta}-Y_t)^2-\delta}{2}$ are orthogonal with respect to the expectation operation $\mathbb{E}$. We have

$$\begin{aligned}
&\mathbb{E}\left\|\sum_{\alpha\in\Lambda_2\backslash\{()\}} \mathrm{I}^{\alpha}_{t,t+\delta}(1)(\mathrm{L}_\alpha(\bar\rho_{\theta_t})-\mathrm{D}_\alpha(\bar\rho_{\theta_t}))\right\|_F^2\\
=&\mathbb{E}\left\|\sum_{\alpha\in\{(0),(1),(1,1)\}} \mathrm{I}^{\alpha}_{t,t+\delta}(1)(\mathrm{L}_\alpha(\bar\rho_{\theta_t})-\mathrm{D}_\alpha(\bar\rho_{\theta_t}))\right\|_F^2 \qquad (45)\\
=&\mathbb{E}\left\|(\mathrm{L}^0(\bar\rho_{\theta_t})-\mathrm{D}^0(\bar\rho_{\theta_t}))\delta+(\mathrm{L}^1(\bar\rho_{\theta_t})-\mathrm{D}^1(\bar\rho_{\theta_t}))(Y_{t+\delta}-Y_t)\right.\\
&\left.+(\mathrm{L}^1(\mathrm{L}^1(\bar\rho_{\theta_t}))-\mathrm{D}^1(\mathrm{D}^1(\bar\rho_{\theta_t})))\frac{(Y_{t+\delta}-Y_t)^2-\delta}{2}\right\|_F^2\\
=&\mathbb{E}\left\|(f^T\otimes I_n)\mathscr{D}_\theta\bar\rho_{\theta_t}+\frac{1}{2}((g\otimes g)^T\otimes I_n\mathscr{D}^2_{\theta^2}\bar\rho_{\theta_t})-\mathrm{D}^0(\bar\rho_{\theta_t}))\right\|_F^2\delta\\
&+\mathscr{R}(\theta_t), \qquad (46)
\end{aligned}$$

where $\mathscr{R}(\theta_t) = \mathbb{E}\left\|(\mathrm{L}^1\mathrm{L}^1(\bar\rho_{\theta_t})-\mathrm{D}^1\mathrm{D}^1(\bar\rho_{\theta_t}))\right\|_F^2\frac{\delta^2}{2} + \mathbb{E}\left\|(\mathrm{L}^1(\bar\rho_{\theta_t})-\mathrm{D}^1(\bar\rho_{\theta_t}))\right\|_F^2\delta$. Since $\mathscr{R}(\theta_t)$ is independent of the coefficient function f, the solution to Problem 3.1 with $k=2$ is determined by

$$\begin{aligned}
&(f^T\otimes I_n)\mathscr{D}_\theta\bar\rho_\theta=\Pi_{\bar\rho_\theta}(\Upsilon(\bar\rho_{\theta_t}))\\
=&\left(\begin{bmatrix}\sum_{j=1}^m r^{1j}(\theta)\left\langle\Upsilon(\bar\rho_{\theta_t}),\bar\partial_j\right\rangle_{\bar\rho_\theta}\\ \sum_{j=1}^m r^{2j}(\theta)\left\langle\Upsilon(\bar\rho_{\theta_t}),\bar\partial_j\right\rangle_{\bar\rho_\theta}\\ \vdots\\ \sum_{j=1}^m r^{mj}(\theta)\left\langle\Upsilon(\bar\rho_{\theta_t}),\bar\partial_j\right\rangle_{\bar\rho_\theta}\end{bmatrix}\otimes I_n\right)^T\mathscr{D}_\theta\bar\rho_{\theta_t} \qquad (47)
\end{aligned}$$

where $\Upsilon(\bar\rho_{\theta_t})=\mathrm{D}^0(\bar\rho_{\theta_t}))-\frac{1}{2}((g\otimes g)^T\otimes I_n)\mathscr{D}^2_{\theta^2}\bar\rho_{\theta_t})$, which yields (15).

The proof of Theorem 3.1 is thus completed. □

V. Conclusion

A quantum projection filter is designed by minimizing the order 1 and order 2 It$\hat{o}$ stochastic Taylor expansion of the difference between the true quantum trajectory and the approximation quantum density matrix on a lower-dimensional differential submanifold, using an orthogonal projection operation in the quantum Fisher metric. Simulations from a numerical example demonstrate the approximation performance of the designed quantum projection filter.

Appendix

Proof of Theorem 4.1. We use three steps to prove that $\bar\rho_{\theta_t}$ has It$\hat{o}$ stochastic Taylor expansion satisfying (9) and (41). The part about the function $\bar\rho_t$ in Theorem 4.1 can be derived following a similar procedure and the corresponding proof is omitted.

Step 1. We show that

$$\bar\rho_{\theta_t} = \hat\Upsilon^k_{t_0,t}(\bar\rho_\theta)+\underbrace{\sum_{\alpha\in\mathscr{R}(\Lambda_k)}\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta))}_{\text{Truncation error}}, \qquad (48)$$

where $\hat\Upsilon^k_{t_0,t}(\bar\rho_\theta)$ is the $k-$order truncation term given by (41), and the *remainder set* $\mathscr{R}(\Lambda_k)$ of Λ_k is defined by

$$\begin{aligned}
&\mathscr{R}(\Lambda_k)=\{\beta\in\mathscr{M}\backslash\Lambda_k:-\beta\in\Lambda_k\}\\
=&\{\beta\in\mathscr{M}\backslash\Lambda_k:\beta=(0)*\alpha \text{ or } (1)*\alpha,\alpha\in\Lambda_k\}. \qquad (49)
\end{aligned}$$

The proof is given by induction on the integer k. For the case that $k=0$, or equivalently, $\Lambda_k=\{()\}$ with the remainder set $\mathscr{R}(\Lambda_k)=\{(0),(1)\}$, (41) becomes

$$\bar\rho_{\theta_t}=\bar\rho_{\theta_0}+\int_{t_0}^t \mathrm{L}^0(\bar\rho_{\theta_s})ds+\int_{t_0}^t \mathrm{L}^1(\bar\rho_{\theta_s})dY_s, \qquad (50)$$

which is exactly the It$\hat{o}$ formula in (37).

Based on the definitions in (39) and (49), one has that for any integer $j\geq 1$,

$$\begin{aligned}
\Lambda_{j+1}\backslash\Lambda_j&=\{\beta\in\mathscr{M}:l(\beta)+n(\beta)=j+1\}\\
&=\{\beta\in\mathscr{M}\backslash\Lambda_j:l(-\beta)+n(-\beta)=j,\text{ or } j-1\}\\
&\subset\{\beta\in\mathscr{M}\backslash\Lambda_j:l(-\beta)+n(-\beta)\leq j\}\\
&=\{\beta\in\mathscr{M}\backslash\Lambda_j:-\beta\in\Lambda_j\}\mathscr{R}(\Lambda_j). \qquad (51)
\end{aligned}$$

Using the induction hypothesis that (41) holds for the case that $k=j\geq 1$, and based on the It$\hat{o}$ formula, we have

$$\begin{aligned}
\bar\rho_{\theta_t}&=\sum_{\alpha\in\Lambda_j}\mathrm{I}^{\alpha}_{t_0,t}(1)(\mathrm{L}_\alpha(\bar\rho_\theta)|_{t=t_0})+\sum_{\alpha\in\mathscr{R}(\Lambda_j)}\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta))\\
&=\sum_{\alpha\in\Lambda_j}\mathrm{I}^{\alpha}_{t_0,t}(1)(\mathrm{L}_\alpha(\bar\rho_\theta)|_{t=t_0})+\sum_{\alpha\in\mathscr{R}(\Lambda_j)\backslash(\Lambda_{j+1}\backslash\Lambda_j)}\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta))\\
&\quad+\sum_{\alpha\in\Lambda_{j+1}\backslash\Lambda_j}\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta))\\
&=\sum_{\alpha\in\Lambda_j}\mathrm{I}^{\alpha}_{t_0,t}(1)(\mathrm{L}_\alpha(\bar\rho_\theta)|_{t=t_0})+\sum_{\alpha\in\mathscr{R}(\Lambda_j)\backslash(\Lambda_{j+1}\backslash\Lambda_j)}\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta))\\
&\quad+\sum_{\alpha\in\Lambda_{j+1}\backslash\Lambda_j}\left(\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta)|_{t=t_0})+\sum_{z=0}^{1}\mathrm{I}^{(z)*\alpha}_{t_0,t}(\mathrm{L}_{(z)*\alpha}(\bar\rho_\theta))\right)\\
&=\sum_{\alpha\in\Lambda_{j+1}}\mathrm{I}^{\alpha}_{t_0,t}(1)(\mathrm{L}_\alpha(\bar\rho_\theta)|_{t=t_0})+\sum_{\alpha\in\bar{\mathscr{R}}}\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\alpha(\bar\rho_\theta)) \qquad (52)
\end{aligned}$$

where

$$\begin{aligned}
\bar{\mathscr{R}}&=[\mathscr{R}(\Lambda_j)\backslash(\Lambda_{j+1}\backslash\Lambda_j)]\bigcup\left[\bigcup_{z=0}^{1}\{(z)*\alpha:\alpha\in\Lambda_{j+1}\backslash\Lambda_j\}\right]\\
&=\left[\{\beta\in\mathscr{M}\backslash\Lambda_j:-\beta\in\Lambda_j\}\backslash\{\beta\in\mathscr{M}\backslash\Lambda_j:\beta\in\Lambda_{j+1}\}\right]\\
&\qquad\bigcup\{\beta\in\mathscr{M}:-\beta\in\Lambda_{j+1}\backslash\Lambda_j\}\\
&=\{\beta\in\mathscr{M}\backslash\Lambda_{j+1}:-\beta\in\Lambda_j\}\\
&\qquad\bigcup\{\beta\in\mathscr{M}\backslash\Lambda_{j+1}:-\beta\in\Lambda_{j+1}\backslash\Lambda_j\}\\
&=\{\beta\in\mathscr{M}\backslash\Lambda_{j+1}:-\beta\in\Lambda_{j+1}\}\\
&=\mathscr{R}(\Lambda_{j+1}). \qquad (53)
\end{aligned}$$

Then (41) holds for the case that $k=j+1$ by combining (52) and (53). Eq. (48) can be obtained by mathematical induction.

Step 2. For any two multi-indices $\alpha,\beta\in\mathscr{M}$, let $\mathbb{E}\|\mathrm{L}_\beta(\bar\rho_{\theta_t})\|_F^2\leq R$ for a constant R and we have

$$\mathbb{E}\left\|\mathrm{I}^{\alpha}_{t_0,t}(\mathrm{L}_\beta(\bar\rho_\theta))\right\|_F^2\leq R(t-t_0)^{l(\alpha)+z(\alpha)}. \qquad (54)$$

The proof is given by induction on $l(\alpha)\geq 1$. Let $t_0<t_1<t_2...<t_p<t$ be a partition of the time interval $[t_0,t]$ and let the positive integer p be large enough.

First, consider the case that $l(\alpha)=1$ with $\alpha=(0)$. Then based on the *Hölder* inequality, one has

$$\begin{aligned}
&\mathbb{E}\left\|\mathrm{I}_{t_0,t}^{\alpha}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta}))\right\|_F^2\\
=&\mathbb{E}\left\|\int_{t_0}^{t}\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s})ds\right\|_F^2=\mathbb{E}\left\|\lim_{p\to\infty}\sum_{i=0}^{p}\mathrm{L}_{\beta}(\bar{\rho}_{\theta_{t_i}})(t_{i+1}-t_i)\right\|_F^2\\
\leq&\mathbb{E}\left\{\lim_{p\to\infty}\sum_{i=0}^{p}\|\mathrm{L}_{\beta}(\bar{\rho}_{\theta_{t_i}})\|_F(t_{i+1}-t_i)\right\}^2\\
=&\mathbb{E}\left\{\int_{t_0}^{t}\|\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s})\|_F ds\right\}^2\leq(t-t_0)\mathbb{E}\left\{\int_{t_0}^{t}\|\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s})\|_F^2 ds\right\}\\
=&R(t-t_0)^2=R(t-t_0)^{l(\alpha)+z(\alpha)}. \qquad (55)
\end{aligned}$$

On the other hand, considering the case that $l(\alpha)=1$ with $\alpha=(1)$, we have

$$\begin{aligned}
&\mathbb{E}\|\mathrm{I}_{t_0,t}^{\alpha}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta}))\|_F^2\\
=&\mathbb{E}\left\|\int_{t_0}^{t}\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s})dY_s\right\|_F^2=\mathbb{E}\left\|\lim_{p\to\infty}\sum_{i=0}^{p}\mathrm{L}_{\beta}(\bar{\rho}_{\theta_{t_i}})(Y_{t_{i+1}}-Y_{t_i})\right\|_F^2\\
=&\mathbb{E}\left\{\lim_{p\to\infty}\sum_{i=0}^{p}\|\mathrm{L}_{\beta}(\bar{\rho}_{\theta_{t_i}})\|_F^2\right\}\mathbb{E}(Y_{t_{i+1}}-Y_{t_i})^2\\
\leq&\lim_{p\to\infty}\sum_{i=0}^{p}\mathbb{E}\left\{\|\mathrm{L}_{\beta}(\bar{\rho}_{\theta_{t_i}})\|_F^2\right\}(t_{i+1}-t_i)\\
=&\mathbb{E}\left\{\int_{t_0}^{t}\|\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s})\|_F^2 ds\right\}\\
\leq&R(t-t_0)=R(t-t_0)^{l(\alpha)+z(\alpha)}. \qquad (56)
\end{aligned}$$

Suppose that (54) holds for the case that $l(\alpha)=j\geq 1$. Now let $l(\alpha)=j+1$ with $\alpha_{-1}=0$. Then $l(\alpha)+z(\alpha)\geq 2z(\alpha)\geq 2$, and $l(\alpha-)+n(\alpha-)=l(\alpha)+z(\alpha)-2$. By using the inductive hypothesis, one has

$$\begin{aligned}
&\mathbb{E}\|\mathrm{I}_{t_0,t}^{\alpha}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta}))\|_F^2=\mathbb{E}\left\|\int_{t_0}^{t}\mathrm{I}_{t_0,s}^{\alpha-}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s}))ds\right\|_F^2\\
\leq&(t-t_0)\int_{t_0}^{t}\mathbb{E}\left\|\mathrm{I}_{t_0,s}^{\alpha-}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s}))\right\|_F^2 ds\\
\leq&(t-t_0)\int_{t_0}^{t}R(s-t_0)^{l(\alpha)+z(\alpha)-2}ds\\
\leq&\frac{R(t-t_0)^{l(\alpha)+z(\alpha)}}{l(\alpha)+z(\alpha)-1}\leq R(t-t_0)^{l(\alpha)+z(\alpha)}. \qquad (57)
\end{aligned}$$

On the other hand, let $\alpha_{-1}=1$. Then $l(\alpha)+z(\alpha)\geq l(\alpha)\geq 1$, and $l(\alpha-)+n(\alpha-)=l(\alpha)+z(\alpha)-1$. By using the inductive assumption, one has

$$\begin{aligned}
&\mathbb{E}\|\mathrm{I}_{t_0,t}^{\alpha}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta_t}))\|_F^2\\
=&\mathbb{E}\left\|\int_{t_0}^{t}\mathrm{I}_{t_0,s}^{\alpha-}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s}))dY_s\right\|_F^2\leq\mathbb{E}\left\{\int_{t_0}^{t}\|\mathrm{I}_{t_0,s}^{\alpha-}(\mathrm{L}_{\beta}(\bar{\rho}_{\theta_s}))\|_F^2 ds\right\}\\
\leq&\int_{t_0}^{t}R(s-t_0)^{l(\alpha)+z(\alpha)-1}ds\\
\leq&\frac{R(t-t_0)^{l(\alpha)+z(\alpha)}}{l(\alpha)+z(\alpha)}\leq R(t-t_0)^{l(\alpha)+z(\alpha)}. \qquad (58)
\end{aligned}$$

Then (54) holds for the case that $l(\alpha)=j+1$ by combining (57) and (58), and can be verified by mathematical induction.

Step 3. In the final step, we derive an upper bound of the truncation error term in (48).

For any multi-index $\alpha\in\mathscr{R}(\Lambda_k)$, it can be shown that

$$l(\alpha)+z(\alpha)\in\{k+1,k+2\}. \qquad (59)$$

In addition, the number of multi-indices belonging to $\mathscr{R}(\Lambda_k)$ is 2^{k+1}. When the time t is within a small neighbourhood of t_0, it follows from Step 2 in this proof that an upper bound on the truncation error term in (48) can be determined as

$$\begin{aligned}
&\mathbb{E}\left\|\sum_{\alpha\in\mathscr{R}(\Lambda_k)}\mathrm{I}_{t_0,t}^{\alpha}(\mathrm{L}_{\alpha}(\bar{\rho}_{\theta}))\right\|_F^2\\
\leq&\sum_{\alpha\in\mathscr{R}(\Lambda_k)}\mathbb{E}\left\|\mathrm{I}_{t_0,t}^{\alpha}(\mathrm{L}_{\alpha}(\bar{\rho}_{\theta}))\right\|_F^2\\
\leq&\sum_{\alpha\in\mathscr{R}(\Lambda_k)}R(t-t_0)^{l(\alpha)+z(\alpha)}\leq R(2(t-t_0))^{k+1}. \qquad (60)
\end{aligned}$$

The proof is then completed. □

REFERENCES

[1] L. Bouten, R. van Handel and M. R. James, "An introduction to quantum filtering," *SIAM Journal on Control and Optimization*, vol. 46, no. 6, pp. 2199-2241, 2007.

[2] H. -P. Breuer and Petruccione F. *The Theory of Open Quantum Systems.* Oxford, U.K.: Oxford University Press, 2002.

[3] M. Mirrahimi and R. van Handel, "Stabilizing feedback control for quantum systems," *SIAM Journal on Control and Optimization*, vol. 46, pp. 445-467, 2007.

[4] Q. Gao, D. Dong, I. R. Petersen and H. Rabitz, "Fault tolerant filtering and fault detection for quantum systems driven by fields in single photon states," *Journal of Mathematical Physics*, vol. 57, no. 6, p.062201,2016.

[5] H. Yonezawa, D. Nakane, T. A. Wheatley *et al.* "Quantum-enhanced optical-phase tracking," *Science*, vol. 337, pp. 1514-1517, 2012.

[6] M. F. Emzir, M. J. Woolley and I. R. Petersen, "A quantum extended Kalman filter," *Journal of Physics A: Mathematical and Theoretical,* vol. 50, no. 22, p. 225301, 2017.

[7] M. Tsang, "Volterra filters for quantum estimation and detection," *Physical Review A,* vol. 92, p. 062119, 2015.

[8] P. Rouchon and J. F. Ralph, "Efficient quantum filtering for quantum feedback control," *Physical Review A*, vol. 91, p. 012118, 2015.

[9] D. Brigo, B. Hanzon, and F. LeGland, "A differential geometric approach to nonlinear filtering: the projection filter," *IEEE Transactions on Automatic Control,* vol. 43, no. 2, pp. 247-252, 1998.

[10] R. van Handel and H. Mabuchi, "Quantum projection filter for a highly nonlinear model in cavity QED," *Journal of Optics B: Quantum and Semiclassical Optics,* vol. 7, S226-S236, 2005.

[11] A. Nielsen, A. Hopkins and H. Mabuchi, "Quantum filter reduction for measurement-feedback control via unsupervised manifold learning," *New Journal of Physics*, vol. 11, p. 105043, 2009.

[12] Q. Gao, G. Zhang, and I. R. Petersen, "An exponential quantum projection filter for open quantum systems," *Automatica*, vol. 99, pp. 59-68, 2019.

[13] J. Armstrong, D. Brigo and E. R. Ferrucci, "Optimal approximation of SDEs on submanifolds: the Itô-vector and Itô-jet projections," *Proceedings of London Mathemtical Society,* vol.119, no. 3, pp. 176-213, 2019.

[14] M. Rathinam and L. R. Petzold, "A new look at proper orthogonal decomposition," *SIAM Journal of Numerical Analysis,* vol. 41, no. 5, pp. 1893-1925, 2003.

[15] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations. 3rd Edition,* Berlin, New York, Springer, 1999.

[16] S. Amari and H. Nagaoka, *Methods of Information Geometry.* Oxford: Oxford University Press, 2000.

[17] D. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations. " *SIAM Review*, vol. 43, pp. 525-546, 2001.

[18] P. Gibilisco, D. Imparato and T. Isola, "Quantum covariance, quantum fisher information and the uncertainty principle," *IEEE Transactions on Information Theory,* vol. 55, no. 1, pp. 439-443, 2009.