Pure and Applied Mathematics Quarterly
Volume 18, Number 3, 1113–1146, 2022

Error estimate for the approximate solution to multivariate feedback particle filter

Wenhui Dong* and Xingbao Gao

Abstract: In this paper, based on the assumption that the gain function K has been optimally obtained in the multivariate feedback particle filter (FPF), we focus on the error estimate for the approximate solutions to the particle's density evolution equation, which is actually the forward Kolmogorov equation (FKE) satisfied by the "particle population". The approximation is essentially the unnormalized density of the states conditioning on the discrete observations with the given time discretization. Mainly owing to the representation of Brownian bridges for the Brownian motion, and the assumption on the coercivity condition, we prove that the mean square error of the approximate solution is of order equal to the square root of the time interval.

Keywords: Feedback particle filter, forward Kolmogorov equation, error estimate, Brownian bridges.

1. Introduction

This paper is concerned with the systems of diffusion processes which are modeled as the following stochastic differential equations (SDEs):

$$\begin{cases} dx_t = f(x_t, t)dt + g(x_t, t)dv_t, \\ dz_t = \hbar(x_t, t)dt + dw_t, \end{cases} \tag{1}$$

where $t \in [0, T]$, $x_t \in \mathbb{R}^d$ is the state at time t, the initial condition x_0 has a given probability density p_0^*, $z_t \in \mathbb{R}^m$ is the observation, and $\{v_t\}$, $\{w_t\}$ are two mutually independent standard Wiener processes taking values in $\mathbb{R}^d$ and $\mathbb{R}^m$, respectively. The mappings $f(\cdot, \cdot) : \mathbb{R}^d \times [0, T] \to \mathbb{R}^d$, $\hbar(\cdot, \cdot) : \mathbb{R}^d \times [0, T] \to \mathbb{R}^m$, $g(\cdot, \cdot) : \mathbb{R}^d \times [0, T] \to \mathbb{R}^{d\times d}$ have continuous bounded derivatives of all orders up to 4.

Received January 3, 2022.

*The author is supported by the Fundamental Research Funds for the Central Universities under grant no. GK202103002 and the start-up fund from Shaanxi Normal University.

The diffusion processes x_t and z_t are considered on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{Z}_t := \sigma\{z_s : 0 \leq s \leq t\}$ denote the $\mathbb{P}$-completed σ-field generated by the observations up to time t and $\mathcal{Z} = \{\mathcal{Z}_t\}$ for the associated filtration, $t \in [0, T]$. The goal of the nonlinear filtering (NLF) problem is to approximate the posterior distribution of the state given the history of observations $\mathcal{Z}_t$. It is a well-known fact in filtering theory [LS] that if $\phi \in C_b^2(\mathbb{R}^d)$, then the conditional probability measure $p_t^* := p^*(\cdot, t)$, $t \geq 0$, admits the following stochastic partial differential equation:

$$dp_t^*[\phi] = -p_t^*[\mathcal{L}^*\phi]dt + (p_t^*[\hbar^\top \phi] - p_t^*[\phi]p_t^*[\hbar^\top])(dz_t - p_t^*[\hbar]dt), \tag{2}$$

where

$$\mathcal{L}^*\phi = \sum_{i,j=1}^d \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial \phi}{\partial x_j}\right) + \sum_{i=1}^d a_i \frac{\partial \phi}{\partial x_i},$$
$$\text{with} \quad a_{ij} = \frac{1}{2}[gg^\top]_{ij}, \quad a_i = f_i - \sum_{j=1}^d \frac{\partial a_{ij}}{\partial x_j},$$

and $p_t^*[\hbar] = \int_{\mathbb{R}^d} \hbar(x)p_t^*(x)dx$. If the measure p_t^* is absolutely continuous with respect to the Lebesgue measure, i.e., $p_t^*(dx) = \mathbb{E}[1_{\{X_t \in dx\}}|\mathcal{Z}_t]$, then it is clear to know that the smooth density, also denoted as p_t^*, satisfies the Kushner equation [Kushner, LS, Rozovskii]:

$$dp_t^* + \mathcal{L}^* p_t^* dt = (h - p_t^*[\hbar])^\top p_t^*(dz_t - p_t^*[\hbar]dt), \tag{3}$$

with the initial density $p_0^* \in L^2(\mathbb{R}^d)$.

It is known that if f, $\hbar$, and g are linear functions, the solution is given by the finite-dimensional Kalman-Bucy filter. Generally speaking, we cannot get the analytic solution for the NLF problems and just obtain the approximate nonlinear filters by numerical methods [BCL] in most cases. Under suitable change of measures, the Kushner equation can be reduced into the following so-called Duncan-Mortensen-Zakai (DMZ) equation:

$$d\sigma_t = \mathcal{L}_0\sigma_t dt + \sigma_t \hbar^\top dy_t,\ \sigma_0(x) = p_0^*(x), \tag{4}$$

where

$$\mathcal{L}_0(\circ) := \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}[a_{ij}\circ] - \sum_{i=1}^d \frac{\partial(f_i\circ)}{\partial x_i}.$$

The equation (4) satisfied by the unnormalized conditional density function of the states, was derived by Duncan [Duncan], Mortensen [Mortensen] and Zakai [Zakai] independently in 1960s.

Once the Kushner or DMZ equation has been derived, several methods such as the splitting-up algorithm [BGR, GK, Ito, Nagase], the S^3 method [LMR] were proposed to solve them. We refer the interested readers to the survey paper [Gyongy] and references therein. Nevertheless, none of these algorithms are implementable for solving the Kushner or DMZ equation. The splitting-up method like the Trotter product formula from semigroup theory, was proposed to solve the DMZ equation [BGR]. This method is simply not implementable. Theoretically it is also hardly useful since it requires boundedness assumptions on f and $\hbar$. This means that the method cannot even be applied to the linear filtering problems. There is no implementable scheme for solving the DMZ equation until Yau and Yau in 2008 [YY] published a pioneering work on the first feasible algorithm, which is so-called Yau-Yau filtering algorithm to the "pathwise-robust" DMZ equation.

The Yau-Yau filtering algorithm [YY] is a major breakthrough in the NLF problems. The beauty of Yau-Yau filtering algorithm is that it is applicable to all practical engineering problems. The ingenious idea of Yau-Yau filtering algorithm is that the solution of the DMZ equation can be decomposed into the on-line and off-line parts. Historically, this is the first real time solution to NLF problems. Thus, the Yau-Yau filtering algorithm can solve the DMZ equation not only in the real-time manner but also in memoryless manner. Later, Luo and Yau in [LY] extended the Yau-Yau filtering algorithm to the most general settings of NLF problems, where f, $\hbar$ and g could depend on both time and states, and the variance of the noises $\{v_t\}$, $\{w_t\}$ are time dependent. Luo and Yau in [LY2] numerically verified the real time performance by the Yau-Yau filtering algorithm for the scalar NLF problems. In [YLY, DLY], the authors have investigated the numerical schemes based on the DMZ equation for the high-dimensional NLF problems, such as the finite-difference scheme [YLY], and Legendre Galerkin spectral method [DLY]. When solving the high-dimensional NLF problems, the highly computational demanding is required. Much attention has been paid to alleviating the so-called "curse of dimensionality" when the states of the NLF system encounter the high dimensions, for example by the sparse grid algorithm [LY1], the proper decomposition method [WLYZ], etc.

The particle filter (PF) [AMGC, GSS], as a simulation-based algorithm to approximate the filtering task [CHQZ], has been successfully applied in numerous fields, see [GSB, Gustaf] for some examples. The key step in the PF is the construction of N stochastic processes $\{X_t^i : 1 \le i \le N\}$. The value

$X_t^i \in \mathbb{R}^d$ is the i-th particle's state at time t. For each time t, the empirical distribution formed by the "particle population" is used to approximate the posterior distribution. One challenge of the PF is the particle degeneracy, namely only a few particles, or even one, have nonzero weight [DGA]. By virtue of the particle degeneracy which leads to decreased performance, or even filter divergence, a common approach to mitigate this in the PF is the sequential importance sampling, where particles are generated by their importance weights at every time step [CHQZ]. The resampling step makes PF practically useful, it however leads to other negative effects, such as sampling impoverishment and increased variance [GSS].

Recently, a feedback structure has been synchronized in the PF, namely the FPF [LL, YMM], which is an alternative feedback control-based approach to the construction of a PF for (1). Essentially, the particles evolve according to a controlled stochastic system, where the control is designed by the innovation method. The soundness of this filter has been demonstrated numerically in many other works besides [YMM, Yang]. Berntorp compared the efficiency of PF and FPF with different proposed gain-function approximate approaches in [Berntorp]. Radhakrishnan in [RM] proposed a novel gain function approximation in FPF for a nonlinear multidimensional stochastic system. Theoretically, the general error analysis of the FPF for general NLF systems has much less research literature as far as we know. In [TM1, TM2], they discussed the convergence analysis of FPF in the setting of the linear Gaussian systems and heavily relied on the assumption that the posterior density is Gaussian. Chen, Luo, Shi, and Yau in [CLSY] for the first time studied the error bound between the empirical distribution and real posterior distribution in the continuous-discrete FPF.

In this paper, however, based on the assumption that the K has been already optimally obtained in the FPF, we focus on the error estimate for the approximate solutions to the particle's density evolution equation. Actually, the density of the "particle population" evolves according to the FKE, which contains the observation process by virtue of the feedback structure of the FPF, thus we call it a generalized FKE in the sequel. Based on the technique of Brownian bridges for the Brownian motion, we propose an approximate solution to the generalized FKE for the corresponding multidimensional NLF systems. Essentially, approximations to the solution of the generalized FKE are considered which depends on the values of the observation process only at the times of a regular partition. The main contribution of this paper is that we show that, under the coercivity condition, the mean square error of the approximate solution to the generalized FKE is up to the order of the square root of the time increment.

The outline of the paper is as follows: we introduce some preliminary results of the multivariate FPF in section 2. In section 3, we give an error estimate of the approximate solutions to the generalized FKE in the FPF. The conclusion has been drawn in the end.

2. Multivariate feedback particle filter

2.1. Notations

With δ_{ij}, we mean the Dirac delta function, which is one when $i = j$ and zero otherwise. Assume all the vectors are column ones. The matrix $A^\top$ represents the transpose of the matrix A. $\nabla\phi$ is the gradient of the function ϕ with respect to x. $\nabla \cdot F$ is the divergence of the vector-valued function $F = (F_1, \cdots, F_d)^\top$ with $\nabla \cdot F = \sum_{i=1}^d \dfrac{\partial F_i}{\partial x_i}$. And further $\nabla \cdot A$ is the divergence of the matrix-valued function $A = [A]_{ij}$, $i = 1, \cdots, d$, $j = 1, \cdots, m$, and $\nabla^\top \cdot A = (\sum_{i=1}^d \dfrac{\partial A_{i1}}{\partial x_i}, \cdots, \sum_{i=1}^d \dfrac{\partial A_{im}}{\partial x_i})^\top$.

Let $L^2(\mathbb{R}^d)$ mean the Hilbert space of square integrable functions at a given time. The inner product between $u = u(x)$ and $v = v(x)$ is $(u, v) := \int_{\mathbb{R}^d} u^\top v dx$. And $u \in L^\infty(\mathbb{R}^d)$ means that

$$|u|_\infty := ess \sup_{x\in\mathbb{R}^d} |u(x)| := \inf_{|E|=0, E\subset\mathbb{R}^d} \left(\sup_{\mathbb{R}^d\setminus E} |u(x)| \right).$$

For the vector-valued function $F = (F_1, \cdots, F_m)^\top$ and matrix-valued function $A = [A_{ij}]$, $i = 1, \cdots, d$ and $j = 1, \cdots, m$, if $F, A \in L^\infty(\mathbb{R}^d)$, we denote $|F|_\infty := \max_{1\le j\le m} |F_j|_\infty$ and $|A|_\infty := \max_{1\le i\le d, 1\le j\le m} |A_{ij}|_\infty$, respectively. The notation $H^1(\mathbb{R}^d)$ means the function space where the function and its first derivative are in $L^2(\mathbb{R}^d)$. If $u \in H^1(\mathbb{R}^d)$, the norm is defined by

$$\|u\| := \left(|u|^2 + \sum_{i=1}^d |\frac{\partial u}{\partial x_i}|^2 \right)^{1/2},$$

where $|\circ|$ is the $L^2(\mathbb{R}^d)$ norm which is induced by the inner product in $L^2(\mathbb{R}^d)$ with $|\circ| = \left(\int_{\mathbb{R}^d} \circ^2 dx\right)^{1/2}$. Recall $H^1(\mathbb{R}^d)$ has dual $H^{-1}(\mathbb{R}^d)$ and

$$H^1(\mathbb{R}^d) \subset L^2(\mathbb{R}^d) \subset H^{-1}(\mathbb{R}^d).$$

Further, denote $\langle\cdot,\cdot\rangle$ for the duality between $H^1(\mathbb{R}^d)$ and $H^{-1}(\mathbb{R}^d)$. The derivatives are interpreted in the weak sense in the sequel.

2.2. Generalized forward Kolmogorov equation

The particles propagate according to a optimally controlled SDE with a gain feedback form defined as follows:

$$dx_t^i = f(x_t^i, t)dt + g(x_t^i, t)dv_t^i + dU_t^i, \tag{5}$$

the control input $dU_t^i = u(x_t^i, t)dt + K(x_t^i, t)dz_t$, for $i = 1, \cdots, N$, where $x_t^i \in \mathbb{R}^d$ is the state for the ith particle at time t, the initial condition $x_0^i \sim p_0^*$, and $\{v_t^i\}$ are mutually independent standard Wiener Processes. Both v_t^i and x_0^i are mutually independent and also independent of $\{x_t, z_t\}$.

Throughout the paper, we denote conditional distribution of a particle x_t^i given $\mathcal{Z}_t$ by $p(x,t)$, or $p_t(x)$. Additionally, certain admissibility requirements are imposed on the control input U_t^i in (5).

Definition 2.1 (Admissible input)**.** *The control input U_t^i is admissible if the random variables $u(x,t)$ and $K(x,t)$ are $\mathcal{Z}_t$ measurable for each t. Moreover, at each fixed time t,*

$$\mathbb{E}\left[|u(x_t^i, t)|\right] := \mathbb{E}\left[\sum_{l=1}^{d} |u_l(x_t^i, t)|\right] < \infty, \tag{6}$$

$$\mathbb{E}\left[|K(x_t^i, t)|^2\right] := \mathbb{E}\left[\sum_{l=1}^{d}\sum_{j=1}^{m} |K_{lj}(x_t^i, t)|^2\right] < \infty,$$

and

$$\lim_{|x|\to\infty} u(x,t)p(x,t) = 0, \tag{7}$$

$$\lim_{|x|\to\infty} K(x,t)p(x,t) = 0, \tag{8}$$

with probability one.

The optimal control (u, K) is the minimizer of an optimization problem. Its Euler-Lagrange boundary value problem (E-L BVP) is obtained via the analysis of first variation. Specifically, the gain function $K : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^{d\times m}$ is the solution to

$$\nabla \cdot (p_t K) = -(\hbar - p_t[\hbar])^\top p_t, \tag{9}$$

with the boundary conditions $\lim_{|x|\to\infty} p_t(x)K(x,t) = 0$, and the function $u : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^d$ is obtained as

$$u = -\frac{1}{2}K(\hbar + p_t[\hbar]) + \Omega(x,t), \tag{10}$$

and $\Omega = (\Omega_1, \cdots, \Omega_d)^\top$ is the Wong-Zakai correction term

$$\Omega_l := \frac{1}{2}\sum_{k=1}^{d}\sum_{s=1}^{m} K_{ks}(x,t)\frac{\partial K_{ls}}{\partial x_k}(x,t), \tag{11}$$

for $l = 1, \cdots, d$.

Moreover, if we consider the process x_t^i that evolves according to the particle filter model (5), the evolution of the density p_t, i.e., the conditional probability density of x_t^i given the filtration $\mathcal{Z}_t$. p_t satisfies the generalized FKE, which has been derived in Proposition 3.2.1 [Yang], as follows

$$dp_t = \mathcal{L}^* p_t dt - \nabla\cdot(p_t K)dz_t - \nabla\cdot(p_t u)dt + \frac{1}{2}\sum_{l,k=1}^{d}\frac{\partial^2}{\partial x_l \partial x_k}(p_t[KK^\top]_{lk})dt. \tag{12}$$

Denote $K_j = (K_{1j}, \cdots, K_{dj})^\top$ for the j-th column of K, $j = 1, \cdots, m$, $\nabla\cdot(p_t K) = (\nabla\cdot(p_t K_1), \cdots, \nabla\cdot(p_t K_m))$. And B_j defines the operator mapping $H^1(\mathbb{R})$ into $L^2(\mathbb{R}^d)$ given by $B_j\phi := \nabla\cdot(\phi K_j) = \sum_{i=1}^{d}\dfrac{\partial(\phi K_{ij})}{\partial x_i} = K_j^\top \nabla\phi + \nabla\cdot K_j \phi$, for $j = 1, \cdots, m$. Thus, we can rewrite (12) as

$$dp_t = \mathcal{L}^+ p_t dt - \nabla\cdot(p_t K)dz_t, \tag{13}$$

where

$$\mathcal{L}^+\phi := \sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(\gamma_{ij}\frac{\partial\phi}{\partial x_j}\right) - \sum_{i=1}^{d}\frac{\partial(\bar{\gamma}_i\phi)}{\partial x_i},$$

with $\gamma_{ij} := a_{ij} + \dfrac{1}{2}[KK^\top]_{ij}$, and $\bar{\gamma}_i := f_i + u_i - \sum_{j=1}^{d}\dfrac{\partial\gamma_{ij}}{\partial x_j}$. Actually, the (13) is equivalent to

$$p_t(x) = p_0(x) + \int_0^t \mathcal{L}^+ p_s(x)ds - \sum_{j=1}^{m}\int_0^t B_j(s)p_s(x)dz_s^j, \tag{14}$$

with $z_t = (z_t^1, \cdots, z_t^m)^\top$, $t \in [0,T]$.

3. Error estimate for the approximation

3.1. Premilinaries

For $p > 1$ and $r \in \mathbb{R}$, denote $L^{p,r}(\mathbb{R}^d)$ the space of real valued Lebesgue measurable functions on $\mathbb{R}^d$ with finite norm

$$\|\phi\|_{p,r} = \left(\int_{\mathbb{R}^d} (1 + \|x\|^2)^{r/2} |f(x)|^p dx \right)^{1/p}. \tag{15}$$

$W^{m,p,r}(\mathbb{R}^d)$ is then the subset of $L^{p,r}(\mathbb{R}^d)$ consisting of functions whose generalized derivatives up to order m belong to $L^{p,r}(\mathbb{R}^d)$. The norm in $W^{m,p,r}(\mathbb{R}^d)$ is defined by

$$\|\phi\|_{m,p,r} = \left(\sum_{|\gamma| \leq m} \frac{|\gamma|!}{\gamma^1! \cdots \gamma^d!} \cdot \int_{\mathbb{R}^d} (1 + \|x\|^2)^{r/2} |D^\gamma f(x)|^p dx \right)^{1/p}, \tag{16}$$

where $\gamma = (\gamma^1, \cdots, \gamma^d)$ is a multi-index of non-negative integers, $D^\gamma = \dfrac{\partial \gamma^1}{\partial x_1^{\gamma^1}} \cdots \dfrac{\partial \gamma^d}{\partial x_d^{\gamma^d}}$.

Assumption A1: We shall suppose $p_0 \in W^{4,4,1}$, so that

$$\mathbb{E}[\sup_{t \in [0,T]} \|p_t\|_{4,4,1}] < \infty. \tag{17}$$

It is a consequence of the Theorem 2.2 in [KR] that if p_t is the solution of the DMZ equation (4), then $\mathbb{E}[\sup_{t \in [0,T]} \|p_t\|_{m,p,r}] \leq C\mathbb{E}[\|p_0\|^p_{m,p,r}]$, for some constant C.

On the condition that the gain function K has been already optimally obtained in (9), where K and u have certain smooth properties fairly well as we need, then we shall suppose that the following coercivity condition is satisfied:

Assumption A2: There exist $\alpha > 0$ and $\lambda \in \mathbb{R}$ such that for all $v \in H^1(\mathbb{R}^d)$ with compact support and for almost all $t \in [0, T]$,

$$\alpha \|v\|^2 + \sum_{j=1}^{m} |B_j v|^2 + 2\langle \mathcal{L}^+ v, v \rangle \leq \lambda |v|^2. \tag{18}$$

Remark 3.1. *The coercivity condition (18) can be reached under some additional restrictions on the control input K and u.*

(C.1) *With* $\tilde{\gamma} := gg^\top + KK^\top$*, there exists* $\alpha_0 > 0$*, such that* $\xi^\top \tilde{\gamma} \xi \geq \alpha_0 \xi^\top \xi$*, for all* $\xi \in \mathbb{R}^d$*.*

(C.2) *With* $\bar{\gamma} := (\bar{\gamma}_1, \cdots, \bar{\gamma}_d)^\top$*, we assume that* $\bar{\gamma}$ *and its derivatives are in* $L^\infty(\mathbb{R}^d)$*.*

(C.3) *With* K_j*,* $j = 1, \cdots, m$*, we need to assume that* $K_j \in L^\infty(\mathbb{R}^d)$ *and* $\nabla K_j \in L^\infty(\mathbb{R}^d)$*, respectively. In addition, we assume that*

$$\alpha_0 - \sum_{j=1}^{m} d^3 |K_j|_\infty (1 + |\nabla K_j|_\infty + |K_j|_\infty) > 0.$$

From the definitions of $\mathcal{L}^+$ *and* B_j*,* $j = 1, \cdots, m$*, and the additional assumptions (C.1)–(C.3) shown above, they implies that*

(19)

$$\begin{aligned}
-\langle \mathcal{L}^+ v, v\rangle =& -\sum_{i,j=1}^{d} \int_{\mathbb{R}^d} \frac{\partial}{\partial x_i}\left(\gamma_{ij}\frac{\partial v}{\partial x_j}\right) v(x)dx + \sum_{i=1}^{d}\int_{\mathbb{R}^d} \frac{\partial(\bar{\gamma}_i v)}{\partial x_i} v(x)dx \\
=& \sum_{i,j=1}^{d} \int_{\mathbb{R}^d} \gamma_{ij}\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}dx + \sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial \bar{\gamma}_i}{\partial x_i} v^2(x)dx + \sum_{i=1}^{d}\int_{\mathbb{R}^d}\bar{\gamma}_i\frac{\partial v}{\partial x_i}v(x)dx \\
=& \sum_{i,j=1}^{d} \int_{\mathbb{R}^d} \gamma_{ij}\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}dx + \sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial \bar{\gamma}_i}{\partial x_i} v^2(x)dx + \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\bar{\gamma}_i\frac{\partial (v^2)}{\partial x_i}dx \\
=& \sum_{i,j=1}^{d} \int_{\mathbb{R}^d} \gamma_{ij}\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}dx + \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial \bar{\gamma}_i}{\partial x_i} v^2(x)dx \\
\geq& \frac{1}{2}\alpha_0 \int_{\mathbb{R}^d}\left[\sum_{i=1}^{d}\left(\frac{\partial v}{\partial x_i}\right)^2\right]dx + \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial \bar{\gamma}_i}{\partial x_i} v^2(x)dx \\
\geq& \frac{1}{2}\alpha_0 \int_{\mathbb{R}^d}\left[\sum_{i=1}^{d}\left(\frac{\partial v}{\partial x_i}\right)^2 + v^2(x)\right]dx - \frac{1}{2}\alpha_0\int_{\mathbb{R}^d} v^2(x)dx \\
&- \frac{d}{2}|\nabla\bar{\gamma}|_\infty \int_{\mathbb{R}^d} v^2(x)dx \\
=& \frac{1}{2}\alpha_0\|v\| - \frac{1}{2}(\alpha_0 + d|\nabla\bar{\gamma}|_\infty)|v|,
\end{aligned}$$

(20)

$$\begin{aligned}
(B_j v, v) =& \int_{\mathbb{R}^d}\sum_{i=1}^{d}\frac{\partial(vK_{ij})}{\partial x_i}v(x)dx \\
=& \sum_{i=1}^{d}\int_{\mathbb{R}^d} K_{ij}\frac{\partial v}{\partial x_i}v(x)dx + \sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial K_{ij}}{\partial x_i}v^2(x)dx
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d} K_{ij}\frac{\partial(v^2)}{\partial x_i}dx + \sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial K_{ij}}{\partial x_i}v^2(x)dx \\
&= \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{\partial K_{ij}}{\partial x_i}v^2(x)dx \\
&\leq \frac{d}{2}|\nabla K_j|_\infty |v|.
\end{aligned}
$$

and

(21)

$$
\begin{aligned}
|B_j v|^2 = &\int_{\mathbb{R}^d}\left[\sum_{i=1}^{d}\frac{\partial(vK_{ij})}{\partial x_i}\right]\left[\sum_{l=1}^{d}\frac{\partial(vK_{lj})}{\partial x_l}\right]dx \\
= &\sum_{i,l=1}^{d}\int_{\mathbb{R}^d} K_{ij}K_{lj}\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_l}dx + 2\sum_{i,l=1}^{d}\int_{\mathbb{R}^d} K_{ij}\frac{\partial K_{lj}}{\partial x_l}\frac{\partial v}{\partial x_i}v(x)dx \\
&+ \sum_{i,l=1}^{d}\int_{\mathbb{R}^d}\frac{\partial K_{ij}}{\partial x_i}\frac{\partial K_{lj}}{\partial x_l}v^2dx \\
\leq &d^2|K_j|_\infty \sum_{i,l=1}^{d}\int_{\mathbb{R}^d}\left|\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_l}\right|dx + 2d^2|K_j|_\infty|\nabla K_j|_\infty\sum_{i,l=1}^{d}\int_{\mathbb{R}^d}\left|\frac{\partial v}{\partial x_i}v\right|dx \\
&+ d^2|K_j|_\infty^2\int_{\mathbb{R}^d}v^2dx \\
\leq &d^3|K_j|_\infty\int_{\mathbb{R}^d}\sum_{i=1}^{d}\left|\frac{\partial v}{\partial x_i}\right|^2 dx + d^4|K_j|_\infty|\nabla K_j|_\infty\|v\|^2 + d^3|K_j|_\infty^2\|v\|^2 \\
\leq &d^3|K_j|_\infty\left(1 + d|\nabla K_j|_\infty + |K_j|_\infty\right)\|v\|^2,
\end{aligned}
$$

thus by taking

$$
\alpha := \alpha_0 - \sum_{j=1}^{m} d^3|K_j|_\infty\left(1 + |\nabla K_j|_\infty + |K_j|_\infty\right),
$$

and

$$
\lambda := \alpha_0 + d|\nabla\bar{\gamma}|_\infty,
$$

in (18)*, the coercivity condition can be obtained.*

Assumption A3: Assume that the observation process z_t in (14) is an Brownian motion.

We wish to consider approximation to the function p_t which depends on the values of the observation process as only a finite number of times under the assumptions A1–A3. For $N_t \in Z^+$, denote

$$h = T/N_t, \tag{22}$$

and $\mathcal{Z}^k := \sigma\{z_h, z_{2h}, \cdots, z_{kh}\}$, $k = 0, \cdots, N_t$. The following related results hold under the assumption A3 that the observation z_t is a Brownian motion.

Lemma 3.1. *If $(k-1)h < t \leq kh$, then $\mathbb{E}[p_t \mid \mathcal{Z}^{N_t}] = \mathbb{E}[p_t \mid \mathcal{Z}^k]$.*

We give a brief proof of Lemma 3.1 in Appendix A.1.

Remark 3.2. *Denote $\Delta z_k^j := z_{kh}^j - z_{(k-1)h}^j$, it is well known that for $(k-1)h \leq t \leq kh$,*

$$\mathbb{E}[z_t^j - z_{(k-1)h}^j \mid \mathcal{Z}^k] = (t-(k-1)h)\frac{\Delta z_k^j}{h}, \tag{23}$$

and thus

$$\mathbb{E}[z_{kh}^j - z_t^j \mid \mathcal{Z}^k] = (kh-t)\frac{\Delta z_k^j}{h}, \tag{24}$$

which are all by virtue of the assumption that z_t are Gaussian.

Lemma 3.2. *For $(k-1)h \leq t \leq kh$, we have*

$$\mathbb{E}[(z_t^j - z_{(k-1)h}^j)^2 \mid \mathcal{Z}^k] = (t-(k-1)h)^2\left(\frac{\Delta z_k^j}{h}\right)^2 + (t-(k-1)h)\frac{kh-t}{h}. \tag{25}$$

The proof of Lemma 3.2 is extremely standard and it is based on the semimartingale decomposition of z_t^j [JY], thus we omit the proof here and include it in Appendix A.2 for the readers' convenience.

Remark 3.3. *The conditional variance of $(z_t^j - z_{(k-1)h}^j)$ given $\mathcal{Z}^k$ is deterministic and equal to $\dfrac{(t-(k-1)h)(kh-t)}{h}$, i.e., $var[z_t^j - z_{(k-1)h}^j] := \mathbb{E}[((z_t^j - z_{(k-1)h}^j) - \mathbb{E}[(z_t^j - z_{(k-1)h}^j) \mid \mathcal{Z}^k])^2 \mid \mathcal{Z}^k] = (t-(k-1)h)\dfrac{(kh-t)}{h}$.*

Definition 3.1 ([Oksendal] Brownian bridge)**.** *For fixed $a, b \in \mathbb{R}$, the process Y_t is called the Brownian bridge from a to b, if there exists an Brownian motion B_t such that*

$$dY_t = \frac{b-Y_t}{1-t}dt + dB_t;\ 0 \leq t < 1,\ Y_0 = a,$$

or equivalently,

$$Y_t = a(1-t) + bt + (1-t)\int_0^t \frac{dB_s}{1-s};\ 0 \le t < 1,$$

and $\lim_{t\to 1} Y_t = b$.

The representation of Brownian bridge for Brownian motion can also be referred as to the book [JY]. We clarify it in the following lemma for our use.

Lemma 3.3 ([JY])**.** *For* $(k-1)h \le t \le kh$*, we consider the enlarged filtration* $\{\hat{\mathcal{Z}}_t\} := \mathcal{Z}_t \vee \sigma\{z_{kh}\}$*. Then, there exists a* $\{\hat{\mathcal{Z}}_t\}$*-Brownian motion* $\eta_t^{j,k}$*, which is independent of* $\mathcal{Z}^{N_t}$*, such that the* $\{\hat{\mathcal{Z}}_t\}$*-semimartingale decomposition of* z_t^j

$$z_t^j = z_{(k-1)h}^j + \eta_t^{j,k} + \int_{(k-1)h}^t \frac{z_{kh}^j - z_u^j}{kh-u}du \tag{26}$$

holds. As pointed out in [JY], the final integral in (26) *is absolutely convergent.*

From (14), for $(k-1)h \le t \le kh$,

$$p_t = p_{(k-1)h} + \int_{(k-1)h}^t \mathcal{L}^+ p_s ds - \sum_{j=1}^m \int_{(k-1)h}^t B_j(s)p_s dz_s^j. \tag{27}$$

By the Itô formula, we have

$$\begin{aligned} |p_t|^2 =& |p_{(k-1)h}|^2 + 2\int_{(k-1)h}^t \langle \mathcal{L}_s^+ p_s, p_s\rangle ds - 2\sum_{j=1}^m \int_{(k-1)h}^t (B_j(s)p_s, p_s) dz_s^j \\ &+ \sum_{j=1}^m \int_{(k-1)h}^t |B_j(s)p_s|^2 ds. \end{aligned} \tag{28}$$

Taking expectation on both sides of (28), we obtain that

$$\mathbb{E}[|p_t|^2] = \mathbb{E}[|p_{(k-1)h}|^2] + 2\int_{(k-1)h}^t \mathbb{E}[\langle \mathcal{L}_s^+ p_s, p_s\rangle] ds + \sum_{j=1}^m \int_{(k-1)h}^t \mathbb{E}[|B_j(s)p_s|^2] ds. \tag{29}$$

Using the $\hat{\mathcal{Z}}_t$-semimartingale decomposition of the components of z_t, we have

$$p_t = p_{(k-1)h} + \int_{(k-1)h}^t \mathcal{L}^+ p_s ds - \sum_{j=1}^m \int_{(k-1)h}^t B_j(s)p_s d\eta_s^{j,k} \tag{30}$$

$$-\sum_{j=1}^{m}\int_{(k-1)h}^{t} B_j(s)p_s\frac{z^j_{kh}-z^j_s}{kh-s}ds$$

For a random variable, denote $\hat{\psi} := \mathbb{E}[\psi \mid \mathcal{Z}^{N_t}]$. Using Fubini's theorem, Lemma 3.1, and noting that the $\eta^{j,k}$ are independent of $\mathcal{Z}^{N_t}$, we have

$$\hat{p}_t = \hat{p}_{(k-1)h} + \int_{(k-1)h}^{t}\mathcal{L}^+\hat{p}_s ds - \sum_{j=1}^{m}\int_{(k-1)h}^{t} B_j(s)\frac{\widehat{p_s(z^j_{kh}-z^j_s)}}{kh-s}ds. \tag{31}$$

Thus,

$$\begin{aligned}(\hat{p}_t)^2 =& (\hat{p}_{(k-1)h})^2 + 2\int_{(k-1)h}^{t}\langle\mathcal{L}^+\hat{p}_s,\hat{p}_s\rangle ds\\ &-2\sum_{j=1}^{m}\int_{(k-1)h}^{t}\left(B_j(s)\frac{\widehat{p_s(z^j_{kh}-z^j_s)}}{kh-s},\hat{p}_s\right)ds.\end{aligned} \tag{32}$$

3.2. The error estimate

Before we give out the error estimate between p_t and $\hat{p}_t$, we firstly obtain the following result in Theorem 3.1.

Theorem 3.1. *Define $\Delta p_t = p_t - \hat{p}_t$, for $(k-1)h \le t \le kh$, we have*

(33)

$$\begin{aligned}&\mathbb{E}[|\Delta p_t|^2] =\\ &\mathbb{E}[|\hat{p}_{(k-1)h}|^2] + 2\int_{(k-1)h}^{t}\mathbb{E}[\langle\mathcal{L}^+_s\Delta p_s,\Delta p_s\rangle]ds + \sum_{j=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}[|B_j(s)\Delta p_s|^2]ds\\ &+\sum_{j=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}[|B_j(s)\hat{p}_s|^2]ds - 2\sum_{j=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}\left[\left(B_j(s)\frac{\widehat{p_s(z^j_{kh}-z^j_s)}}{kh-s},\hat{p}_s\right)\right]ds.\end{aligned}$$

Proof. By applying the Fubini's theorem, we have

$$\begin{aligned}&\mathbb{E}\left[\int_{(k-1)h}^{t}\langle\mathcal{L}^+_s(p_s-\hat{p}_s),p_s-\hat{p}_s\rangle ds \mid \mathcal{Z}^k\right]\\ =&\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\langle\mathcal{L}^+_s p_s,p_s\rangle - \langle\mathcal{L}^+_s p_s,\hat{p}_s\rangle - \langle\mathcal{L}^+_s\hat{p}_s,p_s\rangle\right)ds \mid \mathcal{Z}^k\right]\end{aligned} \tag{34}$$

$$
\begin{aligned}
&+\mathbb{E}\left[\int_{(k-1)h}^{t}\langle\mathcal{L}_s^{+}\hat{p}_s,\hat{p}_s\rangle ds \mid \mathcal{Z}^k\right]\\
=&\mathbb{E}\left[\int_{(k-1)h}^{t}\langle\mathcal{L}_s^{+}p_s,p_s\rangle ds \mid \mathcal{Z}^k\right]-\int_{(k-1)h}^{t}\langle\mathcal{L}_s^{+}\hat{p}_s,\hat{p}_s\rangle ds,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\int_{(k-1)h}^{t}|B_j(s)(p_s-\hat{p}_s)|^2 ds \mid \mathcal{Z}^k\right]\\
=&\mathbb{E}\left[\int_{(k-1)h}^{t}|B_j(s)p_s|^2 ds \mid \mathcal{Z}^k\right]-\int_{(k-1)h}^{t}|B_j(s)\hat{p}_s|^2 ds
\end{aligned}
\tag{35}
$$

Since

$$
\begin{aligned}
\mathbb{E}[|\Delta p_t|^2]&=\mathbb{E}[(p_t-\hat{p}_t,p_t-\hat{p}_t)]=\mathbb{E}[(p_t,p_t)-2(p_t,\hat{p}_t)+(\hat{p}_t,\hat{p}_t)]\\
&=\mathbb{E}[|p_t|^2]-|\hat{p}_t|^2,
\end{aligned}
\tag{36}
$$

and substituting (29) and (32) in (36), we obtain that

(37)

$$
\begin{aligned}
&\mathbb{E}[|\Delta p_t|^2]=\\
&\quad\mathbb{E}[|\Delta p_{(k-1)h}|^2]+2\int_{(k-1)h}^{t}\mathbb{E}[\langle\mathcal{L}_s^{+}p_s,p_s\rangle-\langle\mathcal{L}^{+}\hat{p}_s,\hat{p}_s\rangle]ds\\
&\quad+\sum_{j=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}[|B_j(s)p_s|^2]ds-2\sum_{j=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}\left[\left(B_j(s)\frac{p_s(\widehat{z_{kh}^{j}-z_s^{j}})}{kh-s},\hat{p}_s\right)\right]ds.
\end{aligned}
$$

The result can be reached by using $\mathbb{E}[\langle\mathcal{L}_s^{+}p_s,p_s\rangle-\langle\mathcal{L}^{+}\hat{p}_s,\hat{p}_s\rangle]=\mathbb{E}[\langle\mathcal{L}_s^{+}\Delta p_s,\Delta p_s\rangle]$ in (37) and substituting the term $\sum_{j=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}[|B_j(s)p_s|^2]ds$ in (37) by the (35). □

Remark 3.4. *Derivatives of p_t (or $\hat{p}_t$) in the x variables satisfy the equation obtained from that of p_t (or $\hat{p}_t$) obtained by differentiating, i.e.,*

$$
d\left(\frac{\partial p}{\partial x_i}\right)=\left(\frac{\partial}{\partial x_i}\mathcal{L}^{+}p\right)dt-\sum_{j=1}^{m}\left(\frac{\partial}{\partial x_i}B_jp\right)dz_t^j,
\tag{38}
$$

and similarly for higher derivatives. Furthermore, if for example $a(t,x)$ is *certain coefficient function, for $(k-1)h \le t \le kh$, we have*

(39)
$$a(t,x)\frac{\partial p}{\partial x_i}(t) = a((k-1)h,x)\frac{\partial p}{\partial x_i}((k-1)h) + \int_{(k-1)h}^{t} a(s,x)\frac{\partial \mathcal{L}^+}{\partial x_i}p(s)ds + \int_{(k-1)h}^{t}\frac{\partial a(s,x)}{\partial s}\cdot\frac{\partial p(s)}{\partial x_i}ds - \sum_{j=1}^{m}\int_{(k-1)h}^{t} a(s,x)\frac{\partial B_j}{\partial x_i}p(s)dz_s^j.$$

Theorem 3.2. *There is a constant C_0, which is independent of t, j, and k, such that for $(k-1)h \le t \le kh$, $1 \le j \le m$,*

(40)
$$\left|\int_{(k-1)h}^{t}\mathbb{E}|B_j(s)\hat{p}_s|^2 ds - 2\int_{(k-1)h}^{t}\mathbb{E}\left(\hat{p}_s, B_j(s)\frac{\widehat{p_s(z_{kh}^j - z_s^j)}}{kh-s}\right)ds\right| \le C_0 h^2.$$

Proof. Since $\hat{p}_s$ is $\hat{\mathcal{Z}}_s$-adapted, and recalling the $\hat{\mathcal{Z}}_s$ semimartingale decomposition of z_s as introduced in Lemma 3.3, i.e., $dz_s^j = d\eta_s^{j,k} + \frac{z_{kh}^j - z_s^j}{kh-s}ds$, with $\eta_s^{j,k}$ is a Brownian motion independent of $\mathcal{Z}^k$, we have

(41)
$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{t}(\hat{p}_s, B_j(s)p_s)dz_s^j \mid \mathcal{Z}^k\right]\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{t}(\hat{p}_s, B_j(s)p_s)d\eta_s^{j,k} \mid \mathcal{Z}^k\right]\right]\\
&+\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{t}(\hat{p}_s, B_j(s)p_s)\frac{z_{kh}^j - z_s^j}{kh-s}ds \mid \mathcal{Z}^k\right]\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{t}(\hat{p}_s, B_j(s)p_s)\frac{z_{kh}^j - z_s^j}{kh-s}ds \mid \mathcal{Z}^k\right]\right]\\
=&\int_{(k-1)h}^{t}\mathbb{E}\left(\hat{p}_s, B_j(s)\frac{\widehat{p_s(z_{kh}^j - z_s^j)}}{kh-s}\right)ds,
\end{aligned}$$

with the fact that η is a Brownian motion independent of $\mathcal{Z}^k$ and the integral with respect to η is 0.

Thus, we will consider the term $\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{t}(\hat{p}_s, B_j(s)p_s)dz_s^j \mid \mathcal{Z}^k\right]\right]$ instead of the $\int_{(k-1)h}^{t}\mathbb{E}(\hat{p}_s, B_j(s)\frac{\widehat{p_s(z_{kh}^j - z_s^j)}}{kh-s})ds$ in the statement of the theorem (40). Based on the introduce in Remark 3.4, for $B_j(t) = K_j^\top(t)\nabla + \nabla \cdot K_j(t)$, we have

$$B_j^{'}(t) = \frac{\partial B_j}{\partial t} = \frac{\partial K_j^\top}{\partial t}(t)\nabla + \nabla \cdot \frac{\partial K_j}{\partial t}(t), \tag{42}$$

and for

$$\mathcal{L}_t^+ := \mathcal{L}^+(t) = \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\gamma_{ij}(t) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(f_i(t) + u_i(t)), \tag{43}$$

we have

$$\mathcal{L}^{+'}(t) = \frac{\partial \mathcal{L}^+}{\partial t} = \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\gamma_{ij}(t) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left(\frac{\partial f_i}{\partial t}(t) + \frac{\partial u_i}{\partial t}(t)\right). \tag{44}$$

Thus, we obtain that

(45)

$$\begin{aligned} B_j(s)\hat{p}_s =& B_j((k-1)h)\hat{p}_{(k-1)h} + \int_{(k-1)h}^{s} B_j(u)\mathcal{L}_u^+ \hat{p}_u du \\ &- \sum_{l=1}^{m}\int_{(k-1)h}^{s} B_j(u)B_l(u)\frac{p_u(\widehat{z_{kh}^l - z_u^l})}{kh-u}du + \int_{(k-1)h}^{s} B_j^{'}(u)\hat{p}_u du, \end{aligned}$$

and by Itô formula, we have

$$\begin{aligned} &|B_j(s)\hat{p}_s|^2 \\ =& |B_j((k-1)h)\hat{p}_{(k-1)h}|^2 + 2\int_{(k-1)h}^{s} \langle B_j(u)\mathcal{L}_u^+\hat{p}_u + B_j^{'}(u)\hat{p}_u, B_j(u)\hat{p}_u\rangle du \\ &- 2\sum_{l=1}^{m}\int_{(k-1)h}^{s}\left(B_j(u)\hat{p}_u, B_j(u)B_l(u)\frac{p_u(\widehat{z_{kh}^l - z_u^l})}{kh-u}\right)du. \end{aligned} \tag{46}$$

Denote $\phi_u := B_j(u)\mathcal{L}^+(u) + B_j^{'}(u)p_u$, and similarly as (45), we have

(47)

$$B_j(s)p_s = B_j((k-1)h)p_{(k-1)h} + \int_{(k-1)h}^{s} \phi_u du - \sum_{l=1}^{m}\int_{(k-1)h}^{s} B_j(u)B_l(u)p_u dz_u^l.$$

Recalling that

$$\hat{p}_s = \hat{p}_{(k-1)h} + \int_{(k-1)h}^{s} \mathcal{L}_u^+ \hat{p}_u du - \sum_{i=1}^{m} \mathbb{E}\left[\int_{(k-1)h}^{s} B_i(u) p_u dz_u^i \mid \mathcal{Z}^k\right], \tag{48}$$

we next consider the $(\hat{p}_s, B_j(s)p_s)$ in turn, which is the sum of the inner products of the terms in (48) and (47). Define $\tilde{A}, \tilde{B}, \tilde{C}_i$, $i = 1, \cdots, m$, for the terms on the right hand side of (48), and $\tilde{X}, \tilde{Y}, \tilde{Z}_l$, $l = 1, \cdots, m$ for the terms on the right of (47). We can see the discussion in turn shown in the following step 1-3.

Step 1: We shall consider $(\tilde{A}, \tilde{X} + \tilde{Y} - \tilde{Z})$ with $\tilde{Z} = \sum_{l=1}^{m} \tilde{Z}_l$.

$$\begin{aligned} &\mathbb{E}\left[\int_{(k-1)h}^{t} (\hat{p}_{(k-1)h}, B_j(s)p_s) dz_s^j\right] \\ =&\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{t} (\hat{p}_{(k-1)h}, B_j(s)p_s) dz_s^j \mid \mathcal{Z}^{k-1}\right]\right] \\ =&\mathbb{E}\left[\left(\hat{p}_{(k-1)h}, \mathbb{E}\left[\int_{(k-1)h}^{t} B_j(s)p_s dz_s^j \mid \mathcal{Z}^{k-1}\right)\right]\right] \\ =&0, \end{aligned} \tag{49}$$

with $\hat{p}_{(k-1)h}$ is $\mathcal{Z}^{k-1} \subset \mathcal{Z}_{(k-1)h}$ measurable, z_t is in fact a Brownian motion, and by using the Fubini's theorem.

Step 2: Consider $(\tilde{B}, \tilde{X} + \tilde{Y} - \tilde{Z})$.

In $\tilde{B} = \int_{(k-1)h}^{s} \mathcal{L}_u^+ \hat{p}_u du$, by Itô formula, we can expand $\mathcal{L}_u^+ \hat{p}_u$ as follows:

$$\begin{aligned} \mathcal{L}_u^+ \hat{p}_u =& \mathcal{L}_{(k-1)h}^+ \hat{p}_{(k-1)h} + \int_{(k-1)h}^{u} \left((\mathcal{L}_r^+)^2 \hat{p}_r + \mathcal{L}_r^{+'} \hat{p}_r\right) dr \\ &- \sum_{l=1}^{m} \mathbb{E}\left[\int_{(k-1)h}^{u} \mathcal{L}_r^+ B_l(r) p_r dz_r^l \mid \mathcal{Z}^k\right], \end{aligned} \tag{50}$$

thus, we rewrite $\tilde{B} := \tilde{B}_1 + \tilde{B}_2 - \sum_{l=1}^{m} \tilde{B}_{3l}$.

Taking care of the $(\tilde{B}_1, \tilde{X} + \tilde{Y} - \tilde{Z})$ term firstly. In $\int_{(k-1)h}^{s} \mathcal{L}_u^+ \hat{p}_u du$, substituting the constant term gives $\mathcal{L}_{(k-1)h}^+ \hat{p}_{(k-1)h}(s - (k-1)h)$. And again taking expectation, we have

$$\mathbb{E}\left[\int_{(k-1)h}^{t} \left(\mathcal{L}_{(k-1)h}^+ \hat{p}_{(k-1)h}, B_j(s)p_s\right)(s - (k-1)h) dz_s^j\right] = 0. \tag{51}$$

For $l = 1, \cdots, m$, denote the next lower order terms in (50) as

$$I_l(u) := \mathbb{E}\left[\int_{(k-1)h}^{u} \mathcal{L}_r^+ B_l(r) p_r dz_r^l \mid \mathcal{Z}^k\right] := \psi(x, u, \omega). \tag{52}$$

Recalling the norm used is that in $L^2(\mathbb{R}^d)$ and by using Jensen's inequality, we see that

$$|I_l(u)|^2 = \int_{\mathbb{R}^d} \psi(x, s, \omega)^2 dx \le \mathbb{E}\left[\left|\int_{(k-1)h}^{u} \mathcal{L}_r^+ B_l(r) p_r dz_r^l\right|^2 \mid \mathcal{Z}^k\right], \tag{53}$$

thus,

$$\mathbb{E}[|I_l(u)|^2] \le \int_{(k-1)h}^{u} \mathbb{E}|\mathcal{L}_r^+ B_l(r) p_r|^2 dr \le C_1(u - (k-1)h), \tag{54}$$

where the last inequality is due to (1.6) of Pardoux [Pardoux].

Now

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_{(k-1)h}^{t} \left(\int_{(k-1)h}^{s} I_l(u) du, B_j(s) p_s\right) dz_s^j\right]\right| \\
\le &\int_{(k-1)h}^{t} \mathbb{E}\left|\left(\int_{(k-1)h}^{s} I_l(u) du, B_j(s) p_s \frac{(z_{kh}^j - z_s^j)}{kh - s}\right)\right| ds \\
\le &\int_{(k-1)h}^{t} \left(\mathbb{E}\left|\int_{(k-1)h}^{s} I_l(u) du\right|^2\right)^{1/2} \cdot \left(\mathbb{E}\left|B_j(s) p_s (z_{kh}^j - z_s^j)\right|^2\right)^{1/2} \frac{1}{kh - s} ds \\
\overset{(54)}{\le} &C_2 \int_{(k-1)h}^{t} \left((s - (k-1)h) \int_{(k-1)h}^{s} (u - (k-1)h) du\right)^{1/2} (kh - s)^{-1/2} ds \\
\le &C_2 \int_{(k-1)h}^{t} (s - (k-1)h)^{3/2} (kh - s)^{-1/2} ds \\
\le &C_2 h^{3/2} \int_{(k-1)h}^{t} (kh - s)^{-1/2} ds \le 2C_2 h^2,
\end{aligned} \tag{55}$$

thus, by substituting the $I_l(u)$ in (55),

$$\left|\mathbb{E}\left[\int_{(k-1)h}^{t} \left(\int_{(k-1)h}^{s} \mathbb{E}\left[\int_{(k-1)h}^{u} \mathcal{L}_r^+ B_l(r) p_r dz_r^l \mid \mathcal{Z}^k\right] du, B_j(s) p_s\right) dz_s^j\right]\right| \le C_3 h^2, \tag{56}$$

for some constant C_3. Actually, this is a bound for the $(\tilde{B}_{3l}, \tilde{X}+\tilde{Y}-\tilde{Z})$ terms. Finally, let us consider the $(\tilde{B}_2, \tilde{X}+\tilde{Y}-\tilde{Z})$ term as

(57)
$$\left|\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\int_{(k-1)h}^{s}\int_{(k-1)h}^{u}((\mathcal{L}_r^{+})^2\hat{p}_r+\mathcal{L}_r^{+'}\hat{p}_r)drdu, B_j(s)p_s\right)dz_s^j\right]\right|\leq C_4h^{5/2},$$

and thus,

$$\left|\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\int_{(k-1)h}^{s}\mathcal{L}_u^{+}\hat{p}_u du, B_j(s)p_s\right)dz_s^j\right]\right| \leq C_5h^2. \tag{58}$$

Step 3: To discuss the $(\tilde{C}_i, \tilde{X}+\tilde{Y}-\tilde{Z})$ term, we consider now the i-th term of the sum in (48), i.e.,

$$\mathbb{E}\left[\int_{(k-1)h}^{s} B_i(u)p_u dz_u^i \mid \mathcal{Z}^k\right]. \tag{59}$$

As shown in the (47), we substitute the

$$B_i(u)p_u = B_i((k-1)h)p_{(k-1)h} + \int_{(k-1)h}^{u}\phi_r dr - \sum_{l=1}^{m}\int_{(k-1)h}^{u} B_i(r)B_l(r)p_r dz_r^l$$

in the (59), and denote

$$\begin{aligned}
\tilde{C}_{i1} :=& \mathbb{E}\left[\int_{(k-1)h}^{s} B_i((k-1)h)p_{(k-1)h}dz_u^i \mid \mathcal{Z}^k\right] \\
=&\mathbb{E}\left[B_i((k-1)h)p_{(k-1)h}\int_{(k-1)h}^{s} dz_u^i \mid \mathcal{Z}^k\right] \\
=&\mathbb{E}\left[B_i((k-1)h)p_{(k-1)h}(z_s^i - z_{(k-1)h}^i) \mid \mathcal{Z}^k\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[B_i((k-1)h)p_{(k-1)h}(z_s^i - z_{(k-1)h}^i) \mid \hat{\mathcal{Z}}_{(k-1)h}\right] \mid \mathcal{Z}^k\right] \\
=&\mathbb{E}\left[B_i((k-1)h)p_{(k-1)h}\mathbb{E}\left[(z_s^i - z_{(k-1)h}^i) \mid \hat{\mathcal{Z}}_{(k-1)h}\right] \mid \mathcal{Z}^k\right] \\
=&\mathbb{E}\left[B_i((k-1)h)p_{(k-1)h}\Delta z_k^i(s-(k-1)h)/h \mid \mathcal{Z}^k\right] \\
=&B_i((k-1)h)\hat{p}_{(k-1)h}\Delta z_k^i(s-(k-1)h)/h,
\end{aligned} \tag{60}$$

$$\tilde{C}_{i2} := \mathbb{E}\left[\int_{(k-1)h}^{s}\left(\int_{(k-1)h}^{u}\phi_r dr\right)dz_u^i \mid \mathcal{Z}^k\right], \tag{61}$$

$$\tilde{C}_{i3l} := \mathbb{E}\left[\int_{(k-1)h}^{s}\left(\int_{(k-1)h}^{u} B_i(r)B_l(r)p_r dz_r^l\right) dz_u^i \mid \mathcal{Z}^k\right], \tag{62}$$

with $l = 1, \cdots, m$. Then, we need to consider

$$\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\tilde{C}_{i1} + \tilde{C}_{i2} - \sum_{l=1}^{m}\tilde{C}_{i3l}, B_j(s)p_s\right) dy_s^j\right], \tag{63}$$

with the expansion (47) is substituted for $B_j(s)p_s$.

Thus, we must obtain the bounds for the expected values of all inner product terms arising from these expansions shown in (63). Consider first $(\tilde{C}_{i1}, \tilde{X})$.

$$\begin{aligned} &2\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\tilde{C}_{i1}, \tilde{X}\right) dz_s^i\right] \\ =&\mathbb{E}\left[2\int_{(k-1)h}^{t}\left(B_i((k-1)h)\hat{p}_{(k-1)h}, B_i((k-1)h)p_{(k-1)h}\right)\right. \\ &\left.\qquad\times(s-(k-1)h)\frac{\Delta z_k^i}{h}\frac{(z_{kh}^j - z_s^j)}{kh-s}ds\right], \end{aligned} \tag{64}$$

where this is 0 if $i \neq j$ because the components of z are independent. If $i = j$, we take expectation conditioned on $\hat{\mathcal{Z}}_{(k-1)h}$ first and obtain

$$\begin{aligned} &\mathbb{E}\left[2\int_{(k-1)h}^{t}\left|B_i((k-1)h)\hat{p}_{(k-1)h}\right|^2(s-(k-1)h)ds\frac{(\Delta z_k^i)^2}{h^2}\right] \\ =&\frac{(t-(k-1)h)^2}{h}\mathbb{E}\left|B_i((k-1)h)\hat{p}_{(k-1)h}\right|^2 \\ \leq&(t-(k-1)h)\mathbb{E}\left|B_i((k-1)h)\hat{p}_{(k-1)h}\right|^2. \end{aligned} \tag{65}$$

This term is $O(h)$, but we shall see below that it cancels, with the observation that this term $\dfrac{(t-(k-1)h)^2}{h}\mathbb{E}\left|B_i((k-1)h)\hat{p}_{(k-1)h}\right|^2 \geq 0$.

Next consider a term of the form $(\tilde{C}_{i1}, \tilde{Z}_l)$.

$$\begin{aligned} &\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\tilde{C}_{i1}, \int_{(k-1)h}^{s} B_j(u)B_l(u)p_u dz_u^l\right) dz_s^l\right] \\ =&\mathbb{E}\left[\int_{(k-1)h}^{t}\left(B_i((k-1)h)\hat{p}_{(k-1)h}\frac{\Delta z_k^i}{h}(s-(k-1)h),\right.\right. \end{aligned} \tag{66}$$

$$
\begin{aligned}
&\qquad\qquad \int_{(k-1)h}^{s} B_j(u)B_l(u)p_u dz_u^l\Big) dz_s^l\Big]\\
=&\mathbb{E}\left[\int_{(k-1)h}^{t}\Big(B_i((k-1)h)\hat{p}_{(k-1)h}\frac{\Delta z_k^i}{h}(s-(k-1)h),\right.\\
&\qquad\qquad \left.\int_{(k-1)h}^{s} B_j(u)B_l(u)p_u dz_u^l\Big)\frac{z_{kh}^j - z_s^j}{kh-s}ds\right],\\
=&\frac{\delta_{ij}}{h}\mathbb{E}\left[\int_{(k-1)h}^{t}(s-(k-1)h)\Big(B_i((k-1)h)\hat{p}_{(k-1)h},\right.\\
&\qquad\qquad\qquad \left.\int_{(k-1)h}^{s} B_j(u)B_l(u)p_u dz_u^l\Big) ds\right]\\
=&0,
\end{aligned}
$$

where the last second equality holds by taking expectation conditioned on $\hat{\mathcal{Z}}_{(k-1)h}$ first, and the last equality follows that the components of z are independent in all cases.

For the term $(\tilde{C}_{i1}, \tilde{Y})$, we have

$$
\begin{aligned}
&\left|\mathbb{E}\left[\int_{(k-1)h}^{t}(\tilde{C}_{i1},\tilde{Y})dz_s^j\right]\right|\\
=&\left|\mathbb{E}\left[\int_{(k-1)h}^{t}\left(B_i((k-1)h)\hat{p}_{(k-1)h}\frac{\Delta z_k^i}{h}(s-(k-1)h),\int_{(k-1)h}^{s}\phi_u du\right)dz_s^j\right]\right|\\
=&\delta_{ij}\left|\mathbb{E}\left[\left(\frac{1}{h}B_i((k-1)h)\hat{p}_{(k-1)h},\int_{(k-1)h}^{t}(s-(k-1)h)\int_{(k-1)h}^{s}\phi_u duds\right)\right]\right|\\
\le& C_6 h^2.
\end{aligned}
\tag{67}
$$

Before taking account of the term $(\tilde{C}_{i2}, \tilde{X}+\tilde{Y}-\tilde{Z})$, we know the term $\int_{(k-1)h}^{u}\phi_r dr$ gives rise to integral as

$$
\begin{aligned}
\mathbb{E}|\hat{C}_{i2}|^2 =&\mathbb{E}\left|\mathbb{E}\left[\int_{(k-1)h}^{s}\left(\int_{(k-1)h}^{u}\phi_r dr\right)dz_u^i \mid \mathcal{Z}^k\right]\right|^2\\
\le&\mathbb{E}\left[\mathbb{E}\left[\left|\int_{(k-1)h}^{s}\left(\int_{(k-1)h}^{u}\phi_r dr\right)dz_u^i\right|^2 \mid \mathcal{Z}^k\right]\right]
\end{aligned}
\tag{68}
$$

$$\leq \int_{(k-1)h}^{s} \mathbb{E}\left|\int_{(k-1)h}^{u} \phi_r dr\right|^2 du \leq C_7(s-(k-1)h)^3.$$

As shown before we obtain that

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_{(k-1)h}^{t} (\tilde{C}_{i2}, B_j(s)p_s) dz_s^j\right]\right| \\
=&\left|\mathbb{E}\left[\int_{(k-1)h}^{t} (\tilde{C}_{i2}, B_j(s)p_s) \frac{(z_{kh}^j - z_s^j)}{kh-s} ds\right]\right| \\
\leq& \int_{(k-1)h}^{t} \mathbb{E}\left|\left(\tilde{C}_{i2}, B_j(s)p_s \frac{(z_{kh}^j - z_s^j)}{kh-s}\right)\right| ds \\
\leq& \int_{(k-1)h}^{t} \left(\mathbb{E}|\tilde{C}_{i2}|^2\right)^{1/2} \left(\mathbb{E}|B_j(s)p_s(z_{kh}^j - z_s^j)|^2\right)^{1/2} \frac{ds}{kh-s} \\
\overset{(68)}{\leq}& C_7 \int_{(k-1)h}^{t} (s-(k-1)h)^{3/2}(kh-s)^{-1/2} ds \\
\leq& C_7 h^{3/2} \int_{(k-1)h}^{t} (kh-s)^{-1/2} ds \leq C_7 h^2.
\end{aligned} \tag{69}$$

For the term $(\tilde{C}_{i3l}, \tilde{X})$, we must consider the following expression

$$\mathbb{E}\left[\int_{(k-1)h}^{t} (\tilde{C}_{i3l}, B_{(k-1)h}^j p_{(k-1)h}) dz_s^j\right]. \tag{70}$$

This appears to be only $O(h^{3/2})$; however, we can take expansion on $B_i(r)B_l(r)$ which is in $\tilde{C}_{i3l}$ as follows:

(71)

$$B_i(r)B_l(r)p_r = B_i((k-1)h)B_l((k-1)h)p_{(k-1)h} + \int_{(k-1)h}^{r} \frac{d(B_i(v)B_l(v)p_v)}{dv} dv.$$

The constant term of the $\tilde{C}_{i3l}$ then gives

(72)

$$\mathbb{E}[B_i((k-1)h)B_l((k-1)h)p_{(k-1)h} \int_{(k-1)h}^{s} (z_u^l - z_{(k-1)h}^l) dz_u^i \mid \mathcal{Z}^k] := \overline{\tilde{C}_{i3l}}.$$

If $l \neq i$ in (72), we have

(73)
$$\begin{aligned}\overline{\tilde{C}_{i3l}} =& \frac{\Delta z_k^l \Delta z_k^i}{h^2} B_i((k-1)h) B_l((k-1)h)\hat{p}_{(k-1)h} \int_{(k-1)h}^{s} (u-(k-1)h)du \\ =& \frac{\Delta z_k^l \Delta z_k^i}{h^2} B_i((k-1)h) B_l((k-1)h)\hat{p}_{(k-1)h} \frac{(s-(k-1)h)^2}{2},\end{aligned}$$

substituting (73) in (70), the expectation is zero.

If $l = i$ in (72), we have

(74)
$$\begin{aligned}\overline{\tilde{C}_{i3l}} =& B_i^2((k-1)h)\hat{p}_{(k-1)h}\mathbb{E}\left[\int_{(k-1)h}^{s} (z_u^i - z_{(k-1)h}^i)dz_u^i \mid \mathcal{Z}^k\right] \\ =& B_i^2((k-1)h)\hat{p}_{(k-1)h}\mathbb{E}\left[(z_s^i - z_{(k-1)h}^i)^2 - (s-(k-1)h) \mid \mathcal{Z}^k\right] \\ =& B_i^2((k-1)h)\hat{p}_{(k-1)h} \cdot \left[\frac{1}{2}\frac{(\Delta z_k^i)^2}{h^2}(s-(k-1)h)^2\right. \\ & \left. + \frac{(s-(k-1)h)(kh-s)}{2h} - \frac{(s-(k-1)h)}{2}\right] \\ =& B_i^2((k-1)h)\hat{p}_{(k-1)h} \cdot \left[\frac{1}{2}\frac{(\Delta z_k^i)^2}{h^2}(s-(k-1)h)^2 - \frac{(s-(k-1)h)^2}{2h}\right],\end{aligned}$$

the last second equality follows Lemma 3.2. Substituting (74) which is the constant term of $\tilde{C}_{i3l}$ in (70), we obtain that the term involving $\dfrac{(s-(k-1)h)^2}{2h}$ has expectation 0, and the other term gives

$$\begin{aligned}\frac{1}{2h^3}\mathbb{E}\Bigg[(\Delta z_k^i)^2 \int_{(k-1)h}^{t} \Big(& B_i^2((h-1)h)\hat{p}_{(k-1)h}, \\ & B_j((h-1)h)p_{(k-1)h}\Big)(s-(k-1)h)^2(\Delta z_k^j)\Bigg],\end{aligned} \tag{75}$$

where both when $i = j$ and $i \neq j$ the expectation is 0.

Arguments similar to those above can show the remaining terms in (70) are less than $C_8 h^2$, for some constant C_8.

Take a turning to consider $(\tilde{C}_{i3l}, \tilde{Y})$, we have

$$|\mathbb{E}\int_{(k-1)h}^{t} (\tilde{C}_{i3l}, \tilde{Y}) dz_s^j| \tag{76}$$

$$
\begin{aligned}
=&\left|\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\mathbb{E}\left[\int_{(k-1)h}^{s}\int_{(k-1)h}^{u}B_i(r)B_l(r)p_r dz_r^l dz_u^i \mid \mathcal{Z}^k\right],\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\int_{(k-1)h}^{s}\phi_u du\right)dz_s^j\right]\right|\\
\leq& C_9\int_{(k-1)h}^{t}\left(\int_{(k-1)h}^{s}\int_{(k-1)h}^{u}drdu\right)^{1/2}(s-(k-1)h)^{1/2}\frac{(kh-s)^{1/2}}{kh-s}ds\\
\leq& C_9 h^{5/2}.
\end{aligned}
$$

Finally, we are left with terms in $(\tilde{C}_{i3l},\tilde{Z}_\alpha)$, for $\alpha=1,\cdots,m$. Let us consider

$$
\begin{aligned}
&\left|\mathbb{E}\left[\int_{(k-1)h}^{t}(\tilde{C}_{i3l},\tilde{Z}_\alpha)dz_s^j\right]\right|\\
=&\left|\mathbb{E}\left[\int_{(k-1)h}^{t}\left(\mathbb{E}\left[\int_{(k-1)h}^{s}\int_{(k-1)h}^{u}B_i(r)B_l(r)p_r dz_r^l dz_u^i \mid \mathcal{Z}^k\right],\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\int_{(k-1)h}^{s}B_j(u)B_\alpha(u)p_u dz_u^\alpha\right)\frac{z_{kh}^j-z_s^j}{kh-s}ds\right]\right|\\
\leq&\int_{(k-1)h}^{t}\left(\mathbb{E}\left|\int_{(k-1)h}^{s}\int_{(k-1)h}^{u}B_i(r)B_l(r)p_r dz_r^l dz_u^i\right|^2\right)^{1/2}\\
&\quad\cdot\left(\mathbb{E}\left|\int_{(k-1)h}^{s}B_j(u)B_\alpha(u)p_u dz_u^\alpha(z_{kh}^j-z_s^j)\right|^2\right)^{1/2}\frac{ds}{kh-s}\\
\leq& h^{3/2}\int_{(k-1)h}^{t}\frac{ds}{(kh-s)^{1/2}}\leq C_{10}h^2.
\end{aligned}
\tag{77}
$$

From the above estimates on Step 1-3, therefore, we have

$$
\begin{aligned}
&\left|2\int_{(k-1)h}^{t}\mathbb{E}\left(\hat{p}_s, B_j(s)\frac{\widehat{p_s(z_{kh}^j-z_s^j)}}{kh-s}\right)ds\right.\\
&\qquad\left.-(t-(k-1)h)\mathbb{E}|B_j((k-1)h)\hat{p}_{(k-1)h}|^2\right|\\
\leq& M_1h^2,
\end{aligned}
\tag{78}
$$

where $M_1=C_2+\cdots C_{10}$.

Consider now the term

$$\int_{(k-1)h}^{t} \mathbb{E}\,|B_j(s)\hat{p}_s|^2\,ds \tag{79}$$

in the statement of the theorem (40).

Recalling the product formula of $|B_j(s)\hat{p}_s|^2$ in (46), i.e.,

$$\begin{aligned}|B_j(s)\hat{p}_s|^2 =&|B_j((k-1)h)\hat{p}_{(k-1)h}|^2\\ &+2\int_{(k-1)h}^{s}\langle B_j(u)\mathcal{L}_u^{+}\hat{p}_u + B_j^{'}(u)\hat{p}_u, B_j(u)\hat{p}_u\rangle du\\ &-2\sum_{l=1}^{m}\int_{(k-1)h}^{s}\left(B_j(u)\hat{p}_u, B_j(u)B_l(u)\frac{p_u(\widehat{z_{kh}^l - z_u^l})}{kh-u}\right)du,\end{aligned} \tag{80}$$

and substituting it in (79), we need to consider the following expression

$$\begin{aligned}&\int_{(k-1)h}^{t}\mathbb{E}|B_j((k-1)h)\hat{p}_{(k-1)h}|^2 ds\\ &+2\int_{(k-1)h}^{t}\mathbb{E}\left[\int_{(k-1)h}^{s}\langle B_j(u)\mathcal{L}_u^{+}\hat{p}_u + B_j^{'}(u)\hat{p}_u, B_j(u)\hat{p}_u\rangle du\right]ds\\ &-2\sum_{l=1}^{m}\int_{(k-1)h}^{t}\mathbb{E}\left[\int_{(k-1)h}^{s}\left(B_j(u)\hat{p}_u, B_j(u)B_l(u)\frac{p_u(\widehat{z_{kh}^l - z_u^l})}{kh-u}\right)du\right]ds.\end{aligned} \tag{81}$$

We have first a term in the expression (81)

$$(t-(k-1)h)\mathbb{E}|B_j((k-1)h)\hat{p}_{(k-1)h}|^2, \tag{82}$$

which is going to be canceled with (65).

The second integral in (81) is bounded by $C_{11}h^2$, shown as

$$\begin{aligned}&2\int_{(k-1)h}^{t}\mathbb{E}\left[\int_{(k-1)h}^{s}\langle B_j(u)\mathcal{L}_u^{+}\hat{p}_u + B_j^{'}(u)\hat{p}_u, B_j(u)\hat{p}_u\rangle du\right]ds\\ \le& C_{11}\int_{(k-1)h}^{t}\int_{(k-1)h}^{s} duds \le C_{11}h^2\end{aligned} \tag{83}$$

for some C_{11}.

Consider now a term of the form

$$\int_{(k-1)h}^{s}\left(B_j(u)B_l(u)\frac{p_u(\widehat{z_{kh}^l - z_u^l})}{kh-u}, B_j(u)\hat{p}_u\right)du, \tag{84}$$

which is equivalent to

$$\mathbb{E}\left[\int_{(k-1)h}^{s} \left(B_j(u)B_l(u)p_u, B_j(u)\hat{p}_u\right) dz_u^l \mid \mathcal{Z}^k\right]. \tag{85}$$

Using the Itô formula which is the similar argument to $B_j(u)\hat{p}_u$ and $B_j(u)B_l(u)p_u$, respectively, the term in the norm of $L^2(\mathbb{R}^d)$ can be written as

$$\begin{aligned}(86)\qquad &(B_j(u)B_l(u)p_u, B_j(u)\hat{p}_u)\\ &=(B_j((k-1)h)B_l((k-1)h)p_{(k-1)h}, B_j((k-1)h)\hat{p}_{(k-1)h}) + O(\sqrt{h}).\end{aligned}$$

By the fact that $\hat{p}_{(k-1)h}$ is $\mathcal{Z}^{k-1} \subset \mathcal{Z}_{(k-1)h}$ measurable, we obtain that

$$\begin{aligned}\mathbb{E}\Bigg[\mathbb{E}\Bigg[\int_{(k-1)h}^{s} \Big(&B_j((k-1)h)B_l((k-1)h)p_{(k-1)h},\\ &B_j((k-1)h)\hat{p}_{(k-1)h}\Big)\, dz_u^l \mid \mathcal{Z}^k\Bigg]\Bigg] = 0,\end{aligned} \tag{87}$$

and thus,

$$\begin{aligned}(88)\qquad &\mathbb{E}\left[\int_{(k-1)h}^{s} \left(B_j(u)B_l(u)\frac{p_u(\widehat{z_{kh}^l - z_u^l})}{kh-u}, B_j(u)\hat{p}_u\right) du\right]\\ &=\mathbb{E}\left[\mathbb{E}\left[\int_{(k-1)h}^{s}(B_j(u)B_l(u)p_u, B_j(u)\hat{p}_u)dz_u^l \mid \mathcal{Z}^k\right]\right]\\ &=\mathbb{E}\left[\int_{(k-1)h}^{s} \mathbb{E}\left[O(\sqrt{h})dz_u^l \mid \mathcal{Z}^k\right]\right] = 0.\end{aligned}$$

Therefore, we obtain that

(89)
$$\left|\int_{(k-1)h}^{t} \mathbb{E}\left|B_j(s)\hat{p}_s\right|^2 ds - (t-(k-1)h)\mathbb{E}|B_j((k-1)h)\hat{p}_{(k-1)h}|^2\right| \leq M_2h^2.$$

From (78) and (89), we have

(90)
$$\left|\int_{(k-1)h}^{t} \mathbb{E}\left[|B_j(s)\hat{p}_s|^2\right] ds - 2\int_{(k-1)h}^{t} \mathbb{E}\left[\left(\hat{p}_s, B_j(s)\frac{p_s(\widehat{z_{kh}^j - z_s^j})}{kh-s}\right)\right] ds\right| \leq C_0h^2,$$

where $C_0 = M_1 + M_2$. $\square$

We can now get the main result of the error for the approximate solution to the generalized FKE.

Theorem 3.3. *Suppose $C = mC_0$, then for $(k-1)h < t \le kh$, we have*

$$\mathbb{E}|\Delta p_t|^2 \le C(e^{\lambda h} + \cdots e^{k\lambda h}) \le kCe^{\mu T}h^2, \tag{91}$$

where $\mu = \lambda \vee 0$. Furthermore, in particular we obtain $\mathbb{E}|\Delta p_T|^2 \le CTe^{\mu T}h$.

Proof. We will get the result by induction. Firstly, we have the following bounds by virtue of the (90).

$$\begin{aligned} &\left| \sum_{j=1}^{m} \int_{(k-1)h}^{t} \mathbb{E}|B_j(s)\hat{p}_s|^2 ds - 2\sum_{j=1}^{m} \int_{(k-1)h}^{t} \mathbb{E}\left(\hat{p}_s, B_j(s) \frac{p_s\widehat{(z^j_{kh} - z^j_s)}}{kh - s} \right) ds \right| \\ \le & mC_0h^2 := Ch^2. \end{aligned} \tag{92}$$

By using the coercivity condition (18) in Assumption A2, (92), and the statement in Theorem 3.2, we have for $0 \le t \le h$,

$$\mathbb{E}|\Delta p_t|^2 \le Ch^2 + \int_0^t \lambda \mathbb{E}|\Delta p_s|^2 ds, \tag{93}$$

with $\hat{p}_0 = p_0$. Thus, by Gronwall's inequality, we obtain that

$$\mathbb{E}|\Delta p_t|^2 \le Ch^2 e^{\lambda t}. \tag{94}$$

In particular,

$$\mathbb{E}|\Delta p_t|^2 \le Ch^2 e^{\lambda h}. \tag{95}$$

Suppose the result is proved for $t \le (k-1)h$, i.e.,

$$\mathbb{E}|\Delta p_t|^2 \le C(e^{\lambda h} + \cdots + e^{(k-1)\lambda h})h^2, \tag{96}$$

so that particularly,

$$\mathbb{E}|\Delta p_{(k-1)h}|^2 \le C(e^{\lambda h} + \cdots + e^{(k-1)\lambda h})h^2. \tag{97}$$

By using the coercivity condition in Theorem 3.1 again, and the bound in (92), for $(k-1)h \le t \le kh$, we have

$$\mathbb{E}|\Delta p_t|^2 \le C(e^{\lambda h} + \cdots + e^{(k-1)\lambda h})h^2 + Ch^2 + \int_{(k-1)h}^{t} \lambda \mathbb{E}|\Delta p_s|^2 ds. \tag{98}$$

Therefore, again by using the Gronwall's inequality, we have

$$\begin{aligned} \mathbb{E}|\Delta p_t|^2 &\le C(1 + e^{\lambda h} + \cdots + e^{(k-1)\lambda h})h^2 e^{\lambda(t-(k-1)h)} \\ &\le C(1 + e^{\lambda h} + \cdots + e^{(k-1)\lambda h})h^2 e^{\lambda h} \\ &= C(e^{\lambda h} + \cdots + e^{k\lambda h})h^2 \\ &\le kCe^{\mu T}h^2, \end{aligned} \tag{99}$$

with $\mu = \lambda \vee 0$, and further, $\mathbb{E}|\Delta p_T|^2 \le N_t C e^{\mu T} h^2 = CTe^{\mu T} h$. □

4. Conclusion

In this paper, we firstly validated that the error estimate of the approximate solution for the generalized FKE in the FPF has the order $O(\sqrt{h})$, where the h is the time interval, in the sense of the mean square error. Moreover, we will validate that the approximate solution can lead the multivariate FPF to attaining a more efficient filtering manner by the numerical simulations, which is one of our future research focuses.

Appendix

A.1. The proof of Lemma 3.1

Proof.

$$\begin{aligned} &\mathbb{E}[p_t \mid \mathcal{Z}^{N_t}] = \mathbb{E}[p_t \mid z_h, z_{2h}, \cdots, z_{N_t h}] \\ =&\mathbb{E}[p_t \mid z_h, z_{2h}, \cdots, z_{(k-1)h}, z_{kh}, z_{(k+1)h} - z_{kh}, \cdots, z_{N_t h} - z_{(N_t-1)h}] \\ =&\mathbb{E}[p_t \mid z_h, z_{2h}, \cdots, z_{kh}], \end{aligned} \tag{100}$$

where the last equality holds following the fact that $\sigma\{z_t\} \vee \sigma\{z_h, z_{2h}, \cdots, z_{kh}\}$ is independent of $\sigma\{z_{(k+1)h} - z_{kh}, \cdots, z_{N_t h} - z_{(N_t-1)h}\}$. □

A.2. The proof of Lemma 3.2

Proof. For $(k-1)h \le t \le kh$, we consider the enlarged filtration $\{\hat{\mathcal{Z}}_t\}$, the $\{\hat{\mathcal{Z}}_t\}$-semimartingale decomposition of z_t^j is

$$z_t^j = z_{(k-1)h}^j + \eta_t^{j,k} + \int_{(k-1)h}^t \frac{z_{kh}^j - z_u^j}{kh-u} du, \tag{101}$$

where η in this decomposition is a $\{\hat{\mathcal{Z}}_t\}$-Brownian motion which is independent of $\mathcal{Z}^{N_t}$. As pointed out in [JY], the final integral in (101) is absolutely convergent. Now,

$$(z_t^j - z_{(k-1)h}^j)^2 = 2\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j) dz_u^j + (t-(k-1)h), \tag{102}$$

thus the $\{\hat{\mathcal{Z}}_t\}$ decomposition is

$$\begin{aligned}(z_t^j - z_{(k-1)h}^j)^2 =& 2\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j) d\eta_u^{j,k} \\ &+ 2\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j)\frac{z_{kh}^j - z_u^j}{kh-u} du \\ &+ (t-(k-1)h).\end{aligned} \tag{103}$$

Denoting $\varrho_t := \mathbb{E}[(z_t^j - z_{(k-1)h}^j)^2 \mid \mathcal{Z}^k]$ and taking expectation at both sides of (103) conditioned on $\mathcal{Z}^k$, we have

$$\begin{aligned}\varrho_t =& 2\mathbb{E}[\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j) d\eta_u^{j,k} \mid \mathcal{Z}^k] \\ &+ 2\mathbb{E}[\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j)\frac{z_{kh}^j - z_u^j}{kh-u} du \mid \mathcal{Z}^k] + (t-(k-1)h) \\ =& 2\mathbb{E}[\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j)\frac{z_{kh}^j - z_u^j}{kh-u} du \mid \mathcal{Z}^k] + (t-(k-1)h),\end{aligned} \tag{104}$$

where the last equality holds following that η is a $\{\hat{\mathcal{Z}}_t\}$-Brownian motion and independent of $\mathcal{Z}^{N_t}$ such that

$$\mathbb{E}[\int_{(k-1)h}^t (z_u^j - z_{(k-1)h}^j) d\eta_u^{j,k} \mid \mathcal{Z}^k] \tag{105}$$

$$
\begin{aligned}
&= \mathbb{E}[\mathbb{E}[\int_{(k-1)h}^{t} (z_u^j - z_{(k-1)h}^j) d\eta_u^{j,k} \mid \mathcal{Z}^{N_t}] \mid \mathcal{Z}^k] = 0.
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[\int_{(k-1)h}^{t} (z_u^j - z_{(k-1)h}^j) \frac{z_{kh}^j - z_u^j}{kh-u} du \mid \mathcal{Z}^k] \\
&= \mathbb{E}[\int_{(k-1)h}^{t} \frac{(z_u^j - z_{(k-1)h}^j)(z_{kh}^j - z_{(k-1)h}^j)}{kh-u} du \mid \mathcal{Z}^k] \\
&\quad - \mathbb{E}[\int_{(k-1)h}^{t} \frac{(z_u^j - z_{(k-1)h}^j)^2}{kh-u} du \mid \mathcal{Z}^k] \\
&= \int_{(k-1)h}^{t} \frac{\mathbb{E}[(z_u^j - z_{(k-1)h}^j) \mid \mathcal{Z}^k]\mathbb{E}[(z_{kh}^j - z_{(k-1)h}^j) \mid \mathcal{Z}^k]}{kh-u} du \\
&\quad - \int_{(k-1)h}^{t} \frac{\varrho_u}{kh-u} du \\
&\overset{(23),(24)}{=} \int_{(k-1)h}^{t} \frac{(\Delta z_k^j)^2}{h} \frac{u-(k-1)h}{kh-u} du - \int_{(k-1)h}^{t} \frac{\varrho_u}{kh-u} du.
\end{aligned} \tag{106}
$$

Substituting (106) back in to (104), we obtain that

$$
\varrho_t = 2\int_{(k-1)h}^{t} \frac{(\Delta z_k^j)^2}{h} \frac{u-(k-1)h}{kh-u} du - \int_{(k-1)h}^{t} \frac{\varrho_u}{kh-u} du + (t-(k-1)h). \tag{107}
$$

If we write $C = \frac{(\Delta z_k^j)^2}{h}$, ϱ_t will satisfy the equation as following

$$
(kh-u)d\varrho_u + 2\varrho_u du = 2C(u-(k-1)h)du + (kh-u)du. \tag{108}
$$

Taking $(kh-u)^{-3}$ as an integrating factor in (108), we have

$$
d\left(\frac{\varrho_u}{(kh-u)^2}\right) = \frac{(1-2C)}{(kh-u)^2} du + \frac{2Ch}{(kh-u)^3} du. \tag{109}
$$

Integrating both sides of the (109) from $(k-1)h$ to t, with $t < kh$, we have

$$
\varrho_t = (\Delta z_k^j)^2 \left(\frac{s-(k-1)h}{h}\right)^2 + (kh-t)\frac{t-(k-1)h}{h} \tag{110}
$$

Clearly, the (110) also holds when $t = kh$. $\square$

Acknowledgements

This work is supported in part by the Fundamental Research Funds for the Central Universities (grant no. GK202103002), and in part by the start-up fund from Shaanxi Normal University. The authors would like to thank Xue Luo for giving us lots of useful advice.

References

[AMGC] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, *A tutorial on particle filters for nonlinear/non-Gaussian Bayesian tracking*, IEEE Trans. Signal Process. **50**(2) (2002), 174–188.

[BCL] A. Budhiraja, L. Chen, and C. Lee, *A survey of numerical methods for nonlinear filtering problems*, Physica D: Nonlinear Phenomena **230**(1-2) (2007), 27–36. MR2345200

[Berntorp] K. Berntorp, *Comparison of gain function approximation mehtods in the feedback particle filter*, In 2018 21st International Conference on Information Fusion, (2018): 123–130.

[BGR] A. Bensoussan, R. Glowinski, and A. Rascanu, *Approximation of the Zakai equation by the splitting up method*, SIAM J. Control Optim. **28** (1990), 1420–1431. MR1075210

[CHQZ] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, Spectral Methods in Fluid Dynamics, Springer, New York, NY, USA, 1988. MR2340254

[CLSY] X. Q. Chen, X. Luo, J. Shi, and S. S.-T. Yau, *General convergence result for continuous-discrete feedback particle filter, Int. J. Control*, DOI: 10.1080/00207179.2021.1948105, 2022.

[Duncan] T. Duncan, *Probability density for diffusion processes with applications to nonlinear filtering theory*, Ph. D. dissertation, Stanford Univ., Stanford, CA, USA, 1967.

[DGA] A. Doucet, S. Godsill, and C. Andrieu, *On sequential Monte Carlo sampling methods for Bayesian filtering*, Statist. Computi., **10**(3) (2000), 197–208.

[DLY] W. H. Dong, X. Luo, and S. S.-T. Yau, *Solving nonlinear filtering problems in real time by Legendre Galerkin spectral method*, IEEE Trans. Automat. Control **66**(4) (2021), 1559–1572. MR4240187

[Gyongy] I. Gyongy, *Approximations of stochastic partial differential equations*, In Lecture notes in Pure and Applied Mathematics 277 (2002), 287–307, New York, Dekker. MR1919514

[Gustaf] F. Gustafsson, *Particle filter theory and practice with positioning applications*, IEEE Aerosp. Electron. Syst. Mag. **25**(7) (2010), 53–82.

[GK] I. Gyongy and N. Krylov, *On the splitting-up method and stochastic partial differential equation*, Ann. Probab. **31** (2003), 564–591. MR1964941

[GSB] G. Grisetti, C. Stachniss, and W. Burgard, *Improved techniques for grid mapping with Rao-Blackwellized particle filters*, IEEE Trans. Robot. **23**(1) (2007), 34–46.

[GSS] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*, Radar and Signal Process., IEE Proceedi. F **140**(2) (1993), 107–113.

[Ito] K. Ito, *Approximation of the Zakai equation for nonlinear filtering*, SIAM J. Control. Optim. **34** (1996), 620–634. MR1377715

[JY] T. Jeulin and M. Yor, Inégalités de Hardy, semimartingale et faux amis. Seminare de Probabilités XIII, Lecture Notes in Math. 721, Springer-Verlag, Berlin, Heidelberg, New York, 1979. MR0544805

[Kushner] H. J. Kushner, *Dynamical equations for nonlinear filtering*, J. Differential Equations **3** (1967), 179–190. MR0213182

[KR] N. V. Krylov and B. L. Rozovskii, *Stochastic partial differential equations and diffusion processes*, Russian Math. Surveys **37** (1982), 81–105. MR0683274

[LL] J. M. Lasry and P. L. Lions, *Mean filed games*, Jpn. J. Math. **2**(1) (2007), 229–260. MR2295621

[LMR] S. Lototsky, R. Mikulevicius, and B, Rozovskii, *Nonlinear filtering revisited: A spectral approach*, SIAM J. Control Optim. **35**(2) (1997), 435–461. MR1436632

[LS] R. S. Lipster and A. N. Shiryayev, The Statistics of Random Process, I, II, Springer-Verlag, Berlin, 1977. MR0474486

[LY] X. Luo and S. S.-T. Yau, *Complete real time solution of the general nonlinear filtering problem without memory*, IEEE Trans. Automat. Control **58**(10) (2013), 2563–2578. MR3106062

[LY1] X. Luo and S. S.-T. Yau, *Hermite spectral method with hyperbolic cross approximation to high-dimensional parabolic pdes*, SIAM J. Numer. Anal. **51**(6) (2013), 3186–3212. MR3131484

[LY2] X. Luo and S. S.-T. Yau, *Hermite spectral method to 1D forward kolmogorov equation and its application to nonlinear filtering problems*, IEEE Trans. Automat. Control **58**(10) (2013), 2495–2507. MR3106057

[Mortensen] R. Mortensen, *Optimal control of continuous time stochastic systems*, Ph. D. dissertation, Univ. California, Berkeley, CA, USA, 1996.

[Nagase] N. Nagase, *Remarks on nonlinear stochastic partial differential equations: An application of the splitting-up method*, SIAM J. Control Optim. **33** (1995), 1716–1730. MR1358091

[Oksendal] B. Oksendal, Stochastic differential equations: an introduction with applications, Berlin: Springer-Verlag, 2005. MR2001996

[Pardoux] E. Pardoux, *Stochastic partial differential equations and filtering of diffusion processes*, Stochastics **3** (1979), 127–167. MR0553909

[Rozovskii] B. L. Rozovskii, Stochastic Evolution Systems, Linear Theory and Application to Nonlinear Filtering, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991. MR1135324

[RM] A. Radhakrishnan and S. Meyn, *Gain function tracking in the feedback particle filter*, In 2019 American Control Conference, (2019): 5352–5359.

[TM1] A. Taghvaei and P. Mehta, *Error analysis of the stochastic linear feedback particle filter*, Proceedings of 2018 IEEE Conference on Decision and Control, (2018): 7194–7199.

[TM2] A. Taghvaei and P. Mehta, *Error analysis of the linear feedback particle filter, 2018 Annual* American Control Conference, (2018): 4261–4266.

[WLYZ] Z. Wang, X. Luo, S. S.-T. Yau, and Z. Zhang, *Proper orthogonal decomposition method to nonlinear filtering problems in* *medium-high dimension*, IEEE Trans. Automat. Control **65**(4) (2020), 1613–1624. MR4085532

[Yang] T. Yang, Feedback particle filter and its applications, Ph. D. thesis, University of Illinois at Urbana-Champaign, 2014. MR3322048

[YLY] M. H. Yueh, W. W. Lin, and S. T. Yau, *An efficient numerical method for solving high-dimensional nonlinear filtering problems*, Commun. Inf. Syst. **14** (2014), 243–262. MR3340942

[YMM] T. Yang, P. G. Metha and S. P. Meyn, *Feedback particle filter*, IEEE Trans. Automat. Control **58**(10) (2013), 2465–2480. MR3106055

[YY] S. T. Yau and S. S.-T. Yau, *Real time solution of nonlinear filtering problem without memory II*, SIAM J. Control Optim. **47**(1) (2008), 163–195. MR2373467

[Zakai] M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrsch. Verw. Gebiete **11** (1969), 230–243. MR0242552

Wenhui Dong
School of Mathematics and Statistics
Shaanxi Normal University
710119 Xi'an
P. R. China
E-mail: dongwh@snnu.edu.cn

Xingbao Gao
School of Mathematics and Statistics
Shaanxi Normal University
710119 Xi'an
P. R. China
E-mail: xinbaog@snnu.edu.cn