EXPLICIT CONVERGENCE ANALYSES OF PDE-BASED FILTERING ALGORITHMS *

JIAYI KANG†, XIUQIONG CHEN ‡, AND STEPHEN S.-T. YAU§

Abstract. The direct method and the Yau-Yau algorithm are two crucial filtering methods in nonlinear filtering problems. Recently, a series of research works [22, 21, 32, 10, 19, 6, 28] in these two methods have gained great success in nonlinear numerical experiments. Therefore, a key issue is whether a general framework of convergence analyses can be developed for these numerical algorithms based on partial differential equations. The contributions of this work consist of two parts. The first one is that we set up a convergence analysis framework for the direct method and the Yau-Yau algorithm, which can show the deep connections between these two methods. They are different formulations of solving robust DMZ equation by reducing it to solve Yau-Yau PDE [2]. The second one is that we prove the explicit convergence rates of the Yau-Yau algorithm with the spectral method and the direct method with Gaussian approximation, in terms of the numbers of spectral basis and time steps.

Key words. Nonlinear filtering, DMZ equation, Yau-Yau algorithms.

MSC codes. 35K15, 60G35, 93E11

1. Introduction.

Nonlinear filtering has always played an important role in both commercial and military industries. Given the noisy observation data, the aim of the filtering problem is to find the best estimate of the unknown state. After many years of study, the discrete-time filtering problem can be unified inside the Bayesian filtering framework. With the development of filtering sensors, the time interval of data acquisition becomes shorter and shorter, and the continuous filtering model can better adapt to such changes than the discrete model. A continuous-time filtering system can be given as follows,

$$
\begin{aligned}
dx_t &= f(x_t)dt + g_V(x_t)dV_t,\\
dy_t &= h(x_t)dt + dW_t,
\end{aligned}
\tag{1.1}
$$

where x_t is some n-dimensional vector, y_t is some m-dimensional vector, the functions $f(\cdot):\mathbb{R}^n\to\mathbb{R}^n, g_V(\cdot):\mathbb{R}^n\to\mathbb{R}^{n\times n}$ are all assumed to be the Lipschitz and smooth, and $h(\cdot):\mathbb{R}^n\to\mathbb{R}^m$ is assumed to be smooth and the density function of the initial state x_0 is $\sigma_0(x)$. The $\{V_t\}_{t\geq 0}$ and $\{W_t\}_{t\geq 0}$ processes are independent Brownian motions and their covariance matrices are $E[dV_t dV_t^\top] = I_n dt$ and $E[dW_t dW_t^\top] = I_m dt$.

Given the sequential observations $\{y_s\}_{s=0}^t$, we define the σ-algebra $\mathcal{Y}_t$ generated by the $\{y_s|0\leq s\leq t\}$, i.e., $\mathcal{Y}_t := \sigma(\{y_s|0\leq s\leq t\})$. For any given function φ, we are interested in computing the conditional expectation $E[\varphi(x_t)|\mathcal{Y}_t]$, which is the optimal estimate of $\varphi(x_t)$ in the minimum mean square error sense. In the 1960s, Duncan,

---

*Submitted to the editors DATE. Xiuqiong Chen and Stephen S.-T. Yau are corresponding authors.

Funding: This work is supported by Tsinghua University Education Foundation fund and National Natural Science Foundation of China (NSFC) grant (12201631, 42450242).

†Beijing Key Laboratory of Topological Statistics and Applications for Complex Systems, Beijing Institute of Mathematical Sciences and Applications (BIMSA), Beijing, China (kangjiayi@bimsa.cn).

‡School of Mathematics, Renmin University of China, Beijing, China. (cxq0828@ruc.edu.cn).

§Beijing Key Laboratory of Topological Statistics and Applications for Complex Systems, Beijing Institute of Mathematical Sciences and Applications (BIMSA) and Department of Mathematical sciences, Tsinghua University, Beijing, China. (yau@uic.edu).

Mortensen, and Zakai independently derived Duncan-Mortensen-Zakai (DMZ) equation [11, 24, 40], which is satisfied by the unnormalized conditional density of the state in filtering system with independent noises. Let $Q(x) := g_V(x)g_V^\top(x)$ and we denote $\circ$ as Stratonovich integral in this paper. The unnormalized density function $\sigma(t,x)$ of x_t conditioned on the observation history $\mathcal{Y}_t$ satisfies the following DMZ equation:

$$\begin{cases} d\sigma(t,x) = L_0\sigma(t,x)dt + h^\top(x)\sigma(t,x)\circ dy_t, \\ \sigma(0,x) = \sigma_0(x), \end{cases} \tag{1.2}$$

where

$$L_0(*) := \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2}{\partial x_i \partial x_j}\left(Q_{i,j}(x)\cdot *\right) - \sum_{i=1}^{n}\frac{\partial}{\partial x_i}(f_i(x)\cdot *) - \frac{1}{2}h(x)^\top h(x)\cdot(*). \tag{1.3}$$

Generally speaking, the DMZ equation is hard to solve. In the 1980s, Davis considered this problem and proposed a type of measure transformation that converts the DMZ equation, a type of stochastic partial differential equation (SPDE), into a partial differential equation (PDE) with stochastic coefficients, which is also referred to as the Robust DMZ equation [9]. The stability of robust DMZ equation is proved in detail in 2000 by Yau and Yau [37]. For some special DMZ equations, the direct method can be used to solve them. The direct method was introduced in [35] and generalized in [36, 33, 15]. However, they all need to assume the observation function $h(x)$ is linear. Then, in 2003 , Yau and Lai solved the DMZ equation by transforming it into a series of ordinary differential equations (ODEs) when the initial distribution is Gaussian, which provided an important foundation for the following developments of direct methods [34]. Recently, Shi and Yau [29] designed an effective numerical method for time-invariant filtering problems by using the direct method and Gaussian approximation algorithm [29]. And soon, this work was extended to time- varying cases in [6]. As for solving the general DMZ equation, Yau and Yau, in 2008, developed a new algorithm called Yau-Yau algorithm to solve the "pathwise-robust" DMZ equation for the time-invariant system, and it has been proved theoretically that the Yau-Yau algorithm will converge to the true solution, as long as the growth rate of the observation function $h(x)$ is greater than that of the drift function $f(x)$ [39]. However, other advances in spectral methods [13], splitting-up method [41, 18] and kernel method [25] can only be established under the condition where h grows linearly or even bounded, whereas the Yau-Yau method has a broader range of applicability. Later, Luo and Yau generalized the "Yau-Yau" algorithm to the "time-varying" case [22, 21]. In 2020, Dong proposed to use the Legendre spectral method to solve the Yau-Yau algorithm [10]. Starting in 2019, a series of tensor algorithms are proposed to numerically implement higher dimensional Yau-Yau algorithms [32, 19].

In earlier publications [22, 21, 32, 10, 19, 5, 6, 28], the (PDE-based) filtering algorithms, such as the direct method and Galerkin Yau-Yau algorithm, demonstrated excellent numerical efficiency. Such a PDE-based filtering framework combines filtering algorithms and PDE solvers together. So, it is valuable to study the explicit convergence rate of this framework. The PDE-based numerical algorithms of solving the robust DMZ equations consists of two iterative steps: updating the initial value function and solving the Kolmogorov PDE. In the previous works [22, 10], the convergence analyses were focused on the forward Kolmogorov equation. However, there are no studies focused on the explicit convergence rate of the complete Yau-Yau algorithm and direct method.

So in this work, we start to formulate a suitable Galerkin spectral method for DMZ equation. Then, we calculate the explicit convergence rate of the Yau-Yau algorithm and the direct method. The technical approach of this paper has its roots in the stability and convergence of SPDE. The error analyses in this paper are motivated by several related papers such as [3], and [13].

The main contributions of this paper are listed as follows:

1. We proposed a truncated Gaussian approximation so that the direct method with N-th truncated Gaussian approximation can be viewed as a sort of Galerkin Yau-Yau algorithm according to Remark 3.8.
2. In Section 4, we proved the explicit convergence rate of global error expectation of the direct method with truncated Gaussian approximation and spectral Yau-Yau algorithm which is given in Theorem 4.8.
3. In Section 4, we develop a framework of the convergence analyses of the direct method with N-th truncated Gaussian approximation and spectral Yau-Yau algorithm, which is given in Corollary 4.9

In the rest of the paper, we assume that $\mathcal{H}$ with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is the Sobolev space $H^m(\Omega)$ for some $m \geq 0$, $\langle u, v \rangle_{H^m(\Omega)} = \int_\Omega \sum_{|\alpha| \leq m} D^\alpha u D^\alpha v \, dx$ with α as a multi-index denoting the order of derivatives and $\Omega \subset \mathbb{R}^n$ is a bounded domain with smooth boundary [12] And $(\mathcal{L}(\mathcal{H}), \|\cdot\|_{\mathcal{L}(\mathcal{H})})$ denotes the space of bounded linear operators mapping from $\mathcal{H}$ to $\mathcal{H}$, and $\|A\|_{\mathcal{L}(\mathcal{H})} := \sup_{x \in \mathcal{H}} \frac{\|Ax\|_{\mathcal{H}}}{\|x\|_{\mathcal{H}}}$. In section 2, we shall give some preliminary knowledge of the semi-group theory and different representations of SPDE. In section 3, we shall summarize the current PDE-based filtering algorithms. Finally, the main convergence results are given in section 4.

2. Preliminary knowledge.

2.1. Basics of Semi-group. At the beginning of this section, we shall introduce several important facts in semi-group theory.

DEFINITION 2.1. *[26] A C_0 semi-group $(S(t))_{t \geq 0}$ on $\mathcal{H}$ is a family of operators in $\mathcal{L}(\mathcal{H})$ satisfying $S(0) = I$, $S(t+s) = S(t)S(s)$ for all $s, t \geq 0$, and $t \to S(t)f$ is a continuous $\mathcal{H}$-valued function for all $f \in \mathcal{H}$. The subset $D(A) \subset \mathcal{H}$ is defined by*

$$D(A) := \left\{ f \in \mathcal{H} : \lim_{t \to 0} \frac{S(t)f - f}{t} \text{ exists in } \mathcal{H} \right\};$$

and we can define $Af := \lim_{t \to 0} \frac{S(t)f - f}{t}$, for $f \in D(A)$. $D(A)$ should be a dense set in $\mathcal{H}$. And $A : D(A) \subset \mathcal{H} \to \mathcal{H}$ is the infinitesimal generator of C_0 semi-group $S(t)$.

For any positive integer n, $(-A)^n$ can be determined trivially. Next, we shall introduce the fractional power of $(-A)^\delta$ with $\delta > 0$.

DEFINITION 2.2. *(The fractional power of $-A$ [23])*

Let $\delta > 0$ and $\delta \notin \mathbb{N}$. Then, there exists an $n_0 \in \mathbb{N}$ such that $n_0 > \delta > n_0 - 1$, if $\phi \in D(A^{n_0})$, then

$$(-A)^\delta \phi := \frac{1}{\Gamma(-\delta)} \int_0^\infty t^{-\delta-1} \left[S(t) - \sum_{k=0}^{n_0-1} \frac{t^k}{k!} A^k \right] \phi dt, \tag{2.1}$$

where $\Gamma(\cdot)$ is the gamma function.

Then, we shall introduce the important smoothing properties of the semi-group.

LEMMA 2.3. *(Smoothing properties of the semi-group [20])*

Let $\delta \geq 0$ and $0 \leq \gamma \leq 1$, then there exists a constant $C > 0$ such that

$$\|(-A)^{\delta} S(t)\|_{\mathcal{L}(\mathcal{H})} \leq Ct^{-\delta}, \ \|(-A)^{-\delta}(I - S(t))\|_{\mathcal{L}(\mathcal{H})} \leq Ct^{\delta}, \tag{2.2}$$

for $\forall\ t > 0$.

In the rest of the paper, we suppose the spectrum of A consists only of eigenvalues $(\lambda_n)_{n=1}^{\infty} \subset (-\infty, 0)$. In Remark 2.6 in Section 2.2, we shall explain why this assumption is reasonable. Assume $(\lambda_n)_{n=0}^{\infty}$ is ordered such that $\lambda_{n+1} \leq \lambda_n$ for all $n \in \mathbb{N}$ and let ϕ_n be the eigenvector corresponding to λ_n. We further assume that $\{\phi_1, \cdots, \phi_n, \dots\}$ is a basis of $\mathcal{H}$. According to [23], for any $\delta > 0$, if $(-A)^{\delta}\phi$ exists, we can rewrite the equation (2.1) as

$$(-A)^{\delta}\phi = \sum_{n=1}^{\infty} |\lambda_n|^{\delta} \langle \phi, \phi_n \rangle_{\mathcal{H}} \phi_n. \tag{2.3}$$

Definition 2.4 (The fractional domain space [23]). *Let $\delta > 0$, the fractional domain space $\mathcal{H}_{\delta}^{A}$ is defined as*

$$\mathcal{H}_{\delta}^{A} := \left\{ \phi \in \mathcal{H}, \sum_{n=1}^{\infty} |\lambda_n|^{2\delta} \langle \phi, \phi_n \rangle_{\mathcal{H}}^{2} < \infty \right\}. \tag{2.4}$$

Furthermore, there is a natural inner product structure for $\mathcal{H}_{\delta}^{A}$, which is

$$\langle \phi, \psi \rangle_{\mathcal{H}_{\delta}^{A}} := \sum_{n=1}^{\infty} |\lambda_n|^{2\delta} \langle \phi, \phi_n \rangle_{\mathcal{H}} \langle \psi, \phi_n \rangle_{\mathcal{H}}, \quad \forall \phi, \psi \in D((-A)^{\delta}). \tag{2.5}$$

So we define the norm $\|\phi\|_{\mathcal{H}_{\delta}^{A}}^{2} := \sum_{n=1}^{\infty} |\lambda_n|^{2\delta} \langle \phi, \phi_n \rangle_{\mathcal{H}}^{2}$. In view of equation (2.3), $\|\phi\|_{\mathcal{H}_{\delta}^{A}} = \|(-A)^{\delta}\phi\|_{\mathcal{H}}$ holds for $\forall\ \phi \in D((-A)^{\delta})$.

For more details on this topic, we offer some references such as [26, 20].

2.2. DMZ equation. We shall focus on the following linear SPDE:

$$d\sigma(t, x) = A(\sigma(t, x))dt + B(\sigma(t, x)) \circ dy_t. \tag{2.6}$$

For the system with independent noises,

$$A(\cdot) := L_0(\cdot) \text{ and } B(\cdot) := (\cdot)h^{\top}.$$

We shall denote $B = (B_i)_{i=1}^{m}$ and $y_t = (y_t^i)_{i=1}^{m}$. For the SPDE (2.6), we assume that A is the generator of some C_0 semi-group $(S(t))_{t \geq 0}$ [20]. For $0 \leq s \leq t$ with fixed t, we consider the differential $S(t - s)\sigma(s, x)$,

$$d(S(t - s)\sigma(s, x)) = -A(S(t - s)\sigma(s, x))ds + S(t - s)d\sigma(s, x), \tag{2.7}$$

and we use equation (2.6) for $d\sigma(s, x)$, which yields

$$\begin{aligned} & d(S(t - s)\sigma(s, x)) = -A(S(t - s)\sigma(s, x))ds \\ & + A(S(t - s)\sigma(s, x))ds + S(t - s)B(\sigma(s, x)) \circ dy_s \\ & = S(t - s)B(\sigma(s, x)) \circ dy_s. \end{aligned} \tag{2.8}$$

Integrate both sides of (2.8), we get

$$S(0)\sigma(t, x) = \sigma(t, x) = S(t)\sigma(0, x) + \int_0^t S(t - s)B(\sigma(s, x)) \circ dy_s, \tag{2.9}$$

where $S(0) = I$ is the identity operator.

Remark 2.5. For a simple case, if we assume B in (2.6) is a zero-operator in (2.6), then the SPDE (2.6) can be simplified as a PDE

$$d\sigma(t,x) = A(\sigma(0,x))dt. \tag{2.10}$$

We can find that (2.9) can be simplified as $\sigma(t,x) = S(t)\sigma(0,x)$. So, the semi-group can be used as the solution operator for a parabolic PDE.

Remark 2.6. In this paper, we assume that the eigenvalues of A are of the following order:

$$-\infty < \cdots \leq \lambda_n \cdots \leq \lambda_1 < 0. \tag{2.11}$$

Let ϕ_n denote the eigenvector corresponding to λ_n. It needs to be pointed out that if $\lambda_1 > 0$, we can use the transformation $\hat{\sigma}(t,x) := e^{-\eta t}\sigma(t,x)$ with $\eta > \lambda_1$. Observe that the normalized density functions of $\hat{\sigma}(t,x)$ and $\sigma(t,x)$ are the same. $\hat{\sigma}(t,x)$ obeys the following SPDE

$$d\hat{\sigma}(t,x) = A(\hat{\sigma}(t,x)) - \eta\hat{\sigma}(t,x) + B(\hat{\sigma}(t,x)) \circ dy_t,$$

which means that $\hat{A} := A - \eta$ and the eigenvalues of $\hat{A}$ associated with $\hat{\sigma}(t,x)$ are negative. Therefore, for DMZ equation (2.6), it is reasonable to assume that (2.11) holds.

2.3. The spectral Galerkin method (SGM) . In general, we cannot directly calculate the explicit solution of a SPDE, since solving it is an infinite-dimensional problem. Galerkin method was proposed to truncate the SPDE as a finite-dimensional differential equation. The SGM is one of the most standard methods for truncated SPDE and PDE problems.

Let us use $\mathcal{H}_N := \langle \phi_1, \cdots, \phi_N \rangle$ to represent the linear space spanned by the eigenvectors $\phi_1, \cdots, \phi_N$ with corresponding eigenvalue $\lambda_1, \cdots, \lambda_N$. For simplicity, we shall further assume $\{\phi_1, \cdots, \phi_N\}$ is orthonormal. And $P_N : \mathcal{H} \to \mathcal{H}_N$ is the orthogonal projection from $\mathcal{H}$ into $\mathcal{H}_N$. In the rest of this paper, we shall omit the x in $\sigma(t,x)$ for simplicity.

DEFINITION 2.7 (Finite N Galerkin approximation [3, 13]). *The finite N Galerkin approximation of* (2.6)*, which is defined as $\sigma^{(N)}(t,x) = \sum_{i=1}^N a_i(t)\phi_i(x)$ with $a_i(\cdot) \in L^2([0,T]) \cap C[0,T]$, is defined by the following SPDE,*

$$d\sigma^{(N)}(s) = A^{(N)}\sigma^{(N)}(s)ds + \sum_{j=1}^m P_N(B_j\sigma^{(N)}(s)) \circ dy_s^j, \tag{2.12}$$

where $\sigma^{(N)}(0) := P_N\sigma(0)$ and $A^{(N)} := P_N A P_N$.

In the rest of this subsection, we shall analyze the three forms of finite N Galerkin approximation.

2.3.1. The weak form of the SGM. The weak form is obtained by solving the following equation, for $k = 1, \cdots, N$, let $a_i(t) = \langle \sigma^{(N)}(t), \phi_i(x) \rangle_{\mathcal{H}_N}$.

We multiply both sides of (2.12) by ϕ_i and integrate them, which yields

$$\begin{aligned} \langle \sigma^{(N)}(t), \phi_i(x) \rangle_{\mathcal{H}_N} &= \langle \sigma(0), \phi_i \rangle_{\mathcal{H}_N} + \int_0^t \langle A^{(N)}\sigma^{(N)}(s), \phi_i(x) \rangle_{\mathcal{H}_N} ds \\ &+ \int_0^t \sum_{j=1}^m \langle P_N B_j(\sigma^{(N)}(s)), \phi_i \rangle_{\mathcal{H}_N} \circ dy_s^j. \end{aligned}$$

We observe that $\langle A^{(N)}\sigma^{(N)}(s), \phi_i(x)\rangle_{\mathcal{H}_N} = a_i(s)\langle A^{(N)}\phi_i(x), \phi_i(x)\rangle_{\mathcal{H}_N} = \lambda_i a(s)$,and $\langle P_N B_j(\sigma^{(N)}(s)), \phi_i\rangle_{\mathcal{H}_N} = \langle B_j(\sigma^{(N)}(s)), P_N\phi_i\rangle_{\mathcal{H}_N}$, since P_N is self-adjoint operator. So we have

$$a_i(t) = \langle \sigma(0), \phi_i\rangle_{\mathcal{H}_N} + \lambda_i \int_0^t a_i(s)ds + \int_0^t \sum_{j=1}^m \langle B_j(\sigma^{(N)}(s)), \phi_i\rangle_{\mathcal{H}_N} \circ dy_s^j. \tag{2.13}$$

We shall call (2.13) the weak form of the SGM of (2.6).

2.3.2. The mild form of the SGM. In the standard theory of SPDE, the weak form solution is equal to the mild form solution. We shall consider the semi-group $S^{(N)}(t)$ which can be viewed as the projection of $S(t)$ onto $\mathcal{H}_N$, i.e., $S^{(N)}(t) := P_N S(t) P_N$. For $0 \le s \le t$ with fixed t, we consider the differential $S^{(N)}(t-s)\sigma^{(N)}(s)$ by using the same method in (2.8), and obtain

$$\begin{aligned} &d((S^{(N)})(t-s)\sigma^{(N)}(s)) \\ &= -(A^{(N)})(S^{(N)}(t-s)\sigma^{(N)}(s))ds + S^{(N)}(t-s)d\sigma^{(N)}(s) \\ &= \sum_{j=1}^m S^{(N)}(t-s)(B_j\sigma^{(N)}(s)) \circ dy_s^j, \end{aligned} \tag{2.14}$$

where we use the fact $S^{(N)}(t-s)P_N = S^{(N)}(t-s)$ since $P_N^2 = P_N$. Integrate both sides of (2.14) on $[0, t]$, which yields

$$\sigma^{(N)}(t) = S^{(N)}(t)\sigma^{(N)}(0) + \int_0^t \sum_{j=1}^m S^{(N)}(t-s)(B_j\sigma^{(N)}(s)) \circ dy_s^j. \tag{2.15}$$

(2.15) is the mild form of the SGM for (2.6). Generally, the mild solution and weak solution are equivalent. For more general cases on this topic, please refer to [20] for details.

2.3.3. Fully discrete scheme of the mild SGM. We shall consider a suitable time discretization $0 = t_0 < t_1 < \cdots < t_l = T$ with $l \in \mathbb{Z}^+$, where $t_k = k\tau$, $0 \le k \le l$ and $\tau = \frac{T}{l}$. Considering (2.15) on $[t_k, t_{k+1}]$, we get

$$\sigma^{(N)}(t_{k+1}) = S^{(N)}(\tau)\sigma^{(N)}(t_k) + \int_{t_k}^{t_{k+1}} \sum_{j=1}^m S^{(N)}(t_{k+1}-s)(B_j\sigma^{(N)}(s)) \circ dy_s^j, \tag{2.16}$$

where $0 \le k \le l-1$.

However, there is no explicit form of the integration in (2.16). Motivated by the Euler scheme since τ is small, we shall define $\sigma_{t_k}^{(N)}$ as an approximation of $\sigma^{(N)}(t_k)$ at time t_k. So, the fully discrete scheme of mild SGM $\sigma_{t_k}^{(N)}$ is defined in the following iterative equation,

$$\begin{aligned} \sigma_{t_{k+1}}^{(N)} &= S^{(N)}(\tau)\sigma_{t_k}^{(N)} + S^{(N)}(\tau)\int_{t_k}^{t_{k+1}} \sum_{j=1}^m B_j(\sigma_{t_k}^{(N)}) \circ dy_k^j \\ &= S^{(N)}(\tau)\left[\sigma_{t_k}^{(N)} + \int_{t_k}^{t_{k+1}} \sum_{j=1}^m B_j(\sigma_s^{(N)}) \circ dy_s^j\right], \end{aligned} \tag{2.17}$$

where $\sigma_{t_0}^{(N)} = P_N\sigma_0$. Similar to (2.17), we can construct a recursive equation for σ_{t_k}.

3. The comparison for the PDE-based filtering algorithms. We shall introduce two numerical algorithms, the direct method and Yau-Yau algorithm, to solve the SPDE (2.6). We shall start with the numerical implementations of the fully discrete scheme of mild SGM defined in (2.17). The fully discrete scheme of the mild SGM can naturally divide the computation into two parts which is summarized into follows.

LEMMA 3.1. *Consider the numerical formulations of the fully discrete scheme in one step, i.e.,*

$$\sigma_{t_{k+1}} = S(\tau)\left[\sigma_{t_k} + \int_{t_k}^{t_{k+1}} B(\sigma_s)\circ dy_s\right], \tag{3.1}$$

where $\tau = t_{k+1} - t_k$. Then, (3.1) can be divided into two steps:

1. *(Update step): Solve the following SPDE with $\sigma_1(t_k, x) := \sigma_{t_k}$ as initial condition,*

$$d(\sigma_1(t,x)) = B(\sigma_1(t,x))\circ dy_t,\ t\in[t_k,t_{k+1}]. \tag{3.2}$$

2. *(Diffusion step): Solve the following PDE with $\sigma_2(t_k, x) := \sigma_1(t_{k+1}, x)$ as initial condition,*

$$d\sigma_2(t,x) = A(\sigma_2(t,x))dt,\ t\in[t_k,t_{k+1}]. \tag{3.3}$$

The solution of (3.3) can be viewed as the $\sigma_{t_{k+1}}$.

Proof. We take the differential of $\left[\sigma_{t_k} + \int_{t_k}^{t_{k+1}}\sum_{j=1}^m B_j(\sigma_s)\circ dy_s^j\right]$ $\sigma_1(t,x) := \sigma_1(t_k,x)e^{h^\top(y_t-y_{t_k})}$ which yields updating step (3.2). According to Remark 2.5, $S(\tau)$ is the solution operator of (3.3). So, we finish the proof. □

Remark 3.2. For DMZ equation (1.2), the update step in Lemma 3.1 can be solved explicitly, i.e.,

$$\begin{aligned} &d(\sigma_1(t_k,x)e^{h^\top(y_t-y_{t_k})}) = \sigma_1(t_k,x)d(e^{h^\top(y_t-y_{t_k})})(\text{Since } \sigma_1(t_k,x) \text{ is fixed})\\ &= \sigma_1(t_k,x)e^{h^\top(y_t-y_{t_k})}d(h^\top(y_t-y_{t_k}))\\ &= \sigma_1(t,x)h^\top\circ dy_t \ (\text{Since } h \text{ is independent of time})\end{aligned} \tag{3.4}$$

Then, it is easy to see that $\sigma_1(t_k,x)e^{h^\top(y_t-y_{t_k})}$ is the solution of (3.2) when $B(\cdot) = h^\top(\cdot)$. This is exactly the updating step of the direct method that appeared in [29, 6].

Now, we find that the key to the numerical method of DMZ equation is how to calculate diffusion step.

3.1. The closed-form solution of the diffusion step - The direct method. In this subsection, we shall introduce direct method which corresponds to closed-form solution of the diffusion step. At first, we shall list the important assumptions for direct method.

Assumption 3.3 (Assumptions for direct method).

1. Q is a positive-definite matrix with constant coefficients.
2. There exists a function ψ, such that $f - Q\cdot\nabla\psi$ is a linear function.
3. $\frac{1}{2}(\sum_{i,j=1}^n (Q_{i,j})\frac{\partial^2\psi}{\partial x_i\partial x_j} - 2f^\top\nabla(\psi) + \nabla(\psi)^\top(Q)\nabla(\psi) - h^\top h + \sum_{i=1}^n\frac{\partial f_i}{\partial x_i})$ is a quadratic function w.r.t. x.

We can summarize the direct method appearing in the works [29, 5] into the following proposition.

PROPOSITION 3.4 (The framework of direct method, Theorem 3.2 and 3.3 [29]). *Assume Assumption* 3.3 *holds. Using the transformation* $\tilde{\sigma}_{2,k}(t,x) := \sigma_{2,k}(t,x)e^{-\psi(x)}$, *the diffusion step of direct method is equivalent to the following parabolic equation:*

$$d\tilde{\sigma}_{2,k}(t,x) = (e^{-\psi(x)}L_0 e^{\psi(x)})\tilde{\sigma}_{2,k}(t,x)dt \ t \in [t_k, t_{k+1}]. \tag{3.5}$$

For (3.5), *the posterior density* $\tilde{\sigma}_{2,k}(t_{k+1},x)$ *is Gaussian if the initial value* $\tilde{\sigma}_{2,k}(t_k,x)$ *is Gaussian. So, we can approximate* $\tilde{\sigma}_{2,k}(t_k,x)$ *by a sum of Gaussian densities. We decompose the original PDE into several sub-PDEs with Gaussian initial conditions.*

So, there are two major directions to numerically solve (3.5) by the direct method.

- The first one is to numerically solve the kernel function of semi-group of (3.5) explicitly.
- The second one is to give explicit Gaussian approximation for any distribution.

And we shall summarize the current direct methods in the follows Table Table 1.

Methods	Systems	Reference
General framework	Time-invariant Yau filtering	[36, 33, 15]
Kernel function	Time-invariant Yau filtering	[38]
Kernel function	Time-varying Yau filtering	[5]
Gaussian approximation	Time-invariant Yau filtering	[28]
Gaussian approximation	Time-varying Yau filtering	[6]

TABLE 1
The development of direct method

3.2. The general approximation of the diffusion step - The Yau-Yau algorithms. In this subsection, we shall introduce the Galerkin Yau-Yau algorithms. We summarise the current Galerkin Yau-Yau algorithms [22, 21, 10, 19] into the following proposition.

PROPOSITION 3.5 (The framework of Galerkin Yau-Yau algorithms in independent noises cases). *The diffusion step of Yau-Yau algorithms is approximated by following three steps:*

1. *We shall choose a* N*-dimensional function space* $S_N \subset \mathcal{H}$ *where* S_N *is spanned by orthogonal basis* $\{\psi_j\}_{j=1}^N$.
2. *Then we approximate* $\sigma_{2,k}(t,x)$ *by* $\bar{\sigma}_{2,k}^{(N)}(t,x) := \sum_{j=1}^n a_{j|k}(t)\psi_j((x)$ *where the* $\{a_{j|k}(t)\}_{j=1}^N$ *can be computed via the following ODEs,*

$$\frac{d}{dt}a_{j|k}(t) = \frac{d}{dt}\langle \bar{\sigma}_2^{(N)}(t,x), \phi_j(x)\rangle = \mathcal{B}(\bar{\sigma}_{2,k}^{(N)}(t,x), \psi_j(x)), \tag{3.6}$$

   for $\forall 1 \le j \le N, t \in [t_k, t_{k+1}]$. *Here* $\mathcal{B}(u,v) := -\frac{1}{2}\sum_{i,j=1}^n Q_{i,l}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_l} + \sum_{i=1}^n \frac{\partial v}{\partial x_i} f_i(x)u(x) - \frac{1}{2}h(x)^\top h(x)uv$ *is* $\langle L_0 u, v\rangle$ *after integrating by part. The initial condition for* (3.6)

$$a_{j|k}(t_k) = \langle \sigma_{2,k}(t_k,x), \psi_j(x)\rangle, \forall 1 \le j \le N. \tag{3.7}$$

3. *Finally, we approximate the* $\sigma_1(t_{k+1},x)$ *by* $\bar{\sigma}_{2,k}^{(N)}(t_{k+1},x)$.

The central idea of the Yau-Yau Galerkin algorithms is to project the posterior density function into a finite-dimensional subspace S_N. However, there is a natural question that needs to be answered. How do we choose the basis functions ψ_j for S_N? The choice of the basis functions will affect the performance of the Yau-Yau Galerkin algorithm. In Table 2, we summarize current Galerkin Yau-Yau algorithms.

Basis Functions	Systems	Dimension	Reference
Generalized Hermite functions	Time-varying	framework	[22]
Generalized Hermite functions	Time-varying	$n = 1$	[21]
Generalized Hermite functions	Time-invariant	$n = 2$	[32]
Legendre polynomial	Time-varying	$n \leq 2$	[10]
Generalized Hermite functions	Time-invariant	$n \leq 6$	[19]

TABLE 2
Different basis functions for SGM of Yau-Yau algorithms

3.3. Summary of two PDE filtering algorithm. The framework for decomposing density by Gaussian was detailed by Genovese and Wasserman [14], who derived a convergence rate in Hellinger distance expressed as $\|\sqrt{p} - \sqrt{p_k}\|_{L^2(\Omega)} \leq C(\log m)^{\frac{1}{4}}/m^{\frac{1}{4}}$, where p is the density function of a bounded random variable, and p_k is its optimal estimation in mixed k Gaussian densities, with $k \sim m^{\frac{2}{3}}(\log m)^{\frac{2}{3}}$. Despite the utility of Gaussian models in distribution analysis, selecting the right parameters is challenging due to the limitations of conventional methods such as the EM algorithm [29, 27, 31]. To address these limitations, we propose the N-th truncated Gaussian approximation, which improves accuracy and parameter selection.

DEFINITION 3.6. *Let $p(x) \in \mathcal{H}$ be a density function. $(w_i, p_i)_{i=1}^N$, where $p_i \in \mathcal{H}$ is a Gaussian density and w_i is the real number, is called the $\mathcal{H}_N$ truncated Gaussian approximation of $p(x)$, if $\sum_{i=1}^N w_i P_N p_i(x)$ minimizes the following expression:*

$$(w_i^*, p_i^*)_{i=1}^N = \arg\min_{w_i \in \mathbb{R}, p_i} \|P_N p(x) - \sum_{i=1}^N w_i P_N p_i(x)\|_{\mathcal{H}}, \tag{3.8}$$

where p_i is Gaussian density. Here $P_N : \mathcal{H} \to \mathcal{H}_N$ is the projection operator and $\mathcal{H}_N = \langle \phi_1, \cdots, \phi_N \rangle$.

Next, we shall prove an important result on N order truncated Gaussian approximation.

LEMMA 3.7. *The $\mathcal{H}_N$ truncated Gaussian approximation for $p(x)$ is equivalent to Galerkin approximation in $\mathcal{H}_N$ for $p(x)$, i.e.,*

$$\sum_{i=1}^N w_i^* P_N p_i^*(x) = P_N p(x). \tag{3.9}$$

Proof. Firstly, the Galerkin approximation in $\mathcal{H}_N$ for $P_N p(x)$ is to find the optimal approximation for $p(x)$ in $\mathcal{H}_N$. That is to minimize the following expression

$$u^*(x) = \arg\min_{u(x) \in \mathcal{H}_N} \|p(x) - u(x)\|_{\mathcal{H}}. \tag{3.10}$$

Obviously, the solution is $u^*(x) = P_N p(x)$, where $P_N : \mathcal{H} \to \mathcal{H}_N$ is the projection operator.

The mixed Gaussian densities are dense in $\mathcal{H}$ (we provide a Theorem in the appendix for readers' convenience.), so that their projections to $\mathcal{H}_N$ are dense in $\mathcal{H}_N$ for any N. Now, we can select a set of Gaussian $(p_i(x))_{i=1}^N$ such that $(P_N p_i)_{i=1}^N$ are linearly independent. Recall $\mathcal{H}_N := \langle \phi_1, \cdots, \phi_N \rangle$ and we define the matrix $M_{i,j} := \langle P_N p_i, \phi_j \rangle_{\mathcal{H}}$ with $i, j = 1, \cdots, N$, so that we have $(P_N p_1, \cdots, P_N p_N)^\top = M(\phi_1, \cdots, \phi_N)^\top$. M is an invertible matrix by linear independence of $P_N p_i$. So, we can represent the $P_N p$ by using the linear combination of $(P_N p_1, \cdots, P_N p_N)^\top$. □

Remark 3.8. In real applications, one may choose the different Gaussian densities for Gaussian approximation in each steps. However, no matter what kinds of Gaussian densities $(p(x))_{i=1}^N$ we choose, the $\mathcal{H}_N$ truncated Gaussian approximation yields the unique optimal result $\sum_{i=1}^N w_i^* P_N p_i^*(x) = P_N p(x)$. So, according to Lemma 3.7, $\mathcal{H}_N$ truncated Gaussian approximation can be viewed as a sort of N-th order Galerkin spectral approximation.

From Remark 3.8, the direct method and Yau-Yau algorithms emerge as essential techniques for the DMZ equation. A unified approach to algorithm analysis for various PDEs can be achieved by introducing an N-th order Galerkin analysis framework, necessitating the development of suitable Galerkin basis functions. Motivated by signal processing and principal component analysis [30], basis function selection for SPDEs (2.6) is guided by eigenvalue maximization in $\mathcal{H}$. Initially, $v_1^* = \arg\max_{v \in \mathcal{H}} \frac{\langle Av, v \rangle_{\mathcal{H}}}{\langle v, v \rangle_{\mathcal{H}}}$ is chosen. Subsequently, functions are selected orthogonal to preceding ones, with $v_N^* = \arg\max_{v \perp V_{N-1}, v \in \mathcal{H}} \frac{\langle Av, v \rangle_{\mathcal{H}}}{\langle v, v \rangle_{\mathcal{H}}}$, where $V_{N-1} = \operatorname{span}\{v_1^*, \ldots, v_{N-1}^*\}$. These functions, $v_i^* = \phi_i$, based on the min-max principle, correspond to the first N eigenvalues. Galerkin convergence analysis utilizes $\mathcal{H}_N$, spanned by eigenvectors $\phi_1, \ldots, \phi_N$ with eigenvalues $\lambda_1, \ldots, \lambda_N$.

4. Convergence in spectral number and time discretization.

We recall $\mathcal{H}$ as the Sobolev space $H^d(\Omega)$ for $d \geq 0$, where $\Omega \subset \mathbb{R}^n$ is bounded with smooth boundary. The mapping $\sigma(t, x) : [0, T] \to \mathcal{H}$ signifies the posterior density. Due to the equivalence of mild and weak forms [20], our focus is on the mild form of SGM (2.15). In this analysis, we elucidate the convergence of the filtering algorithm via spectral methods. Initially, we present the explicit convergence rate for fully discrete scheme of SGM, akin to the direct method with truncated Gaussian approximation and Galerkin Yau-Yau algorithms. Following, we examine the time discretization convergence rate for this scheme.

4.1. Main Assumptions.

According to Remark 2.6, we shall add an assumption for the growth of $|\lambda_n|$. If the Ω is bounded with smooth boundary and A is the Laplace operator, then we can use Weyl law [1] which describes the asymptotic behavior (polynomial growth) of eigenvalues of the Laplace operator. And the result has been extended to the case that the Ω is a bounded manifold and A is the elliptic operator [4]. So, we naturally introduce the following assumption which is also considered by several related papers [7, 3].

Assumption 4.1. (Assumption of the diffusion step) Consider the filtering system (1.1) and the associated the DMZ equation (2.6). For the n-th eigenvalue λ_n of the unbounded operator A, we assume that

$$|\lambda_n| > Cn^\alpha, \ \forall n \geq 1, \tag{4.1}$$

where α and C are some positive numbers. And the eigenfunctions $\{\phi_1, \cdots, \phi_n, \cdots\}$ of A are the orthonormal basis for $\mathcal{H}$.

Luo and Yau proposed the moving-window trick [22] to overcome the weakness of bounded Ω. Recall that $\sum_{k=1}^{m} S^N(t-s)(B_k(\sigma(s)))$ appears in the mild form SGM (2.15). In what follows, the premise is that the operator B_k is Lipschitz and bounded in the norm $\|\cdot\|_{\mathcal{H}_\delta^A}$ defined in Definition 2.4. Such assumption is used in many related papers such as [3] [13], and [7].

Assumption 4.2. (Assumption of the updating step) Consider the filtering system (1.1) and the associated DMZ equation (2.6). We shall assume $\{B_k\}_{k=1}^m$ are Lipschitz operators on $\mathcal{H}$ which means that

$$\sum_{k=1}^{m} \|(B_k(X_1 - X_2))\|_{\mathcal{H}} \leq K_B \|X_1 - X_2\|_{\mathcal{H}} \text{ for } \forall X_1, X_2 \in \mathcal{H}, \tag{4.2}$$

where K_B is some positive constant. Furthermore, we shall assume that $\{B_k\}_{k=1}^m$ are bounded operators in the fractional domain space $\mathcal{H}_\delta^A$ with $\delta > 0$, i.e.,

$$\sum_{k=1}^{m} \|(B_k(X_1))\|_{\mathcal{H}_\delta^A} \leq K_{B_\delta}(1 + \|X_1\|_{\mathcal{H}_\delta^A}),\ \forall X_1 \in \mathcal{H}, \tag{4.3}$$

where K_{B^δ} is some positive constant.

PROPOSITION 4.3. *Given the linear SPDE* (2.6) *under Assumption* 4.2 *for* B, *for initial condition* $\sigma_0 \in \mathcal{H}_\delta^A$ *where* $1 > \delta > 0$ *and* $p \geq 2$, *there exists a unique mild solution* $\sigma(t)$ *satisfying follows,*

$$\sup_{t\in[0,T]} E[\|\sigma(t)\|_{\mathcal{H}_s^A}^p] < C(p)(E[\|\sigma(0)\|_{\mathcal{H}_s^A}^p] + 1), \tag{4.4}$$

for all $s \leq \delta$. *Furthermore, there exists a constant* C *depending on* p, *such that*

$$\left(E\left[\|\sigma(t_1) - \sigma(t_2)\|_{\mathcal{H}_s^A}^p\right]\right)^{\frac{1}{p}} \leq C(p)|t_1 - t_2|^{\min\left\{\frac{1}{2}, \delta - s\right\}}, \tag{4.5}$$

where $t_1, t_2 \in [0, T]$ *and for all* $s \leq \delta$.

Proof. By using Theorem 1 in [16], we can know that, if Assumption 4.2 is satisfied, then a unique mild solution $\sigma(t)$ exists. The equation (4.4) and (4.5) are proved in the regularity results of [17]. □

4.2. The convergence analyses. In this section, $C(a, b, c)$ denotes a constant depending only on a, b, c. We recall $S^{(N)}(t) = P_N S(t) P_N$ and $\mathcal{H}_N$ is the linear space spanned by the eigenvectors consisting of the first N eigenfunctions of the operator A. The convergence analysis unfolds in three stages: (i) error analysis for the finite N Galerkin approximation (2.12), presented in Theorem 4.6; (ii) convergence analysis for the mild SGM's fully discrete scheme (2.17), depicted in Theorem 4.8; and (iii) analysis of the Yau-Yau algorithm and direct method with truncated Gaussian approximation, pivotal to the fully discrete SGM schemes.

4.2.1. Error analysis of Finite N Galerkin approximation. The aim of this subsection is to prove the convergence of the finite N Galerkin approximation (2.12) of (2.6), i.e. to prove Theorem 4.6. To complete the proof, we first need to give three lemmas. The first two lemmas are designed to estimate the error term of $S(t)\sigma(0) - S^{(N)}(t)\sigma(0)$. The proof rely on a special version of the Burkholder-Davis-Gundy (BDG) inequality [8]. Finally, we prove Theorem 4.6 below by using Gronwall's inequality.

Lemma 4.4. *Let A be the operator of the diffusion term in* (2.6)*, and let $\mathcal{H}_N$ be the linear space spanned by the eigenvectors consisting of the first N eigenfunctions of the operator A. If Assumption* 4.1 *is satisfied and $\delta > 0$, then*

$$|\lambda_{N+1}|^\delta \|(P_N - I)X\|_{\mathcal{H}} \le \|X\|_{\mathcal{H}_\delta^A}, \quad \forall X \in \mathcal{H}_\delta^A, \delta > 0, \tag{4.6}$$

where P_N is the projection operator from $\mathcal{H}$ to $\mathcal{H}_N$ and I is the identity map.

Lemma 4.5. *Consider the DMZ equation* (2.6)*. Let the initial condition be $\sigma(0) \in \mathcal{H}_\delta^A$ with $0 < \delta < 1$. And both Assumption* 4.1 *and Assumption* 4.2 *are satisfied, then $\sigma^{(N)}$, the solution to* (2.12)*, follows*

$$\sup_{t\in[0,T]} E\Big[\|S^{(N)}(t)\sigma^{(N)}(0) - S(t)\sigma(0)\|_{\mathcal{H}}\Big] \le C(S(t),\delta)N^{-\alpha\delta}\|\sigma(0)\|_{H_\delta^A}. \tag{4.7}$$

The proofs of Lemma 4.4 and Lemma 4.5 can be found in Appendix B.

Theorem 4.6. *Consider the filtering system* (1.1) *and associated DMZ equation* (2.6)*. Assumption* 4.1 *and Assumption* 4.2 *are satisfied. Suppose that initial value $\sigma(0) \in \mathcal{H}_\delta^A$ with $0 < \delta < 1$, $p \ge 2$ and $\sigma^{(N)}$ is the solution of* (2.12)*, then there exists a constant $C := C(K_{B_\delta}, S(t), K_B, T, \sup_{t\in[0,T]} \|\sigma(t)\|_{\mathcal{H}_\delta^A})$ such that*

$$\sup_{t\in[0,T]} \left(E\left[\|\sigma^{(N)}(t) - \sigma(t)\|_{\mathcal{H}}^p\right]\right)^{\frac{1}{p}} \le CN^{-\alpha\delta}. \tag{4.8}$$

Proof. For any $t \in [0,T]$, we decompose the error between $\sigma^{(N)}$ and σ as follows

$$\begin{aligned}
&(E\|\sigma^{(N)}(t) - \sigma(t)\|_{\mathcal{H}}^p)^{\frac{1}{p}} \le \|S^{(N)}(t)\sigma^{(N)}(0) - S(t)\sigma(0)\|_{\mathcal{H}} \\
&+ \sum_{k=1}^m \left(E\left\|\int_0^t (S^{(N)}(t-s)B_k(\sigma^{(N)}(s) - \sigma(s)) - (S(t-s) - S^{(N)}(t-s))B_k(\sigma(s))) \circ dy_s\right\|^p\right)^{\frac{1}{p}} \\
&=: I_1 + I_2.
\end{aligned} \tag{4.9}$$

$\blacksquare$

Next, we only need to estimate the I_1, and I_2. For the first term I_1, it is bounded by using the same method that appears in (4.6),

$$I_1(t) \le C(S(t))N^{-\alpha\delta}\|\sigma(0)\|_{\mathcal{H}_\delta^A}. \tag{4.10}$$

Then, we shall transform I_2 into Itô's form which allows us to use BDG inequality in [8]. Now, we have

$$\begin{aligned}
I_2 &= \frac{1}{2}\sum_{k=1}^m \left(E\left\|\int_0^t (S^{(N)}(t-s)(B_k^2(\sigma^{(N)}(s) - \sigma(s)))ds\right\|_{\mathcal{H}}^p\right)^{\frac{1}{p}} \\
&+ \sum_{k=1}^m \left(E\left\|\int_0^t (S^{(N)}(t-s)(B_k(\sigma^{(N)}(s) - \sigma(s)))dy_s\right\|_{\mathcal{H}}^p\right)^{\frac{1}{p}} =: J_1 + J_2
\end{aligned} \tag{4.11}$$

The J_1 is dominated by three additional terms as follows:

$$\begin{aligned}
J_1(t) \le & \frac{1}{2}\sum_{k=1}^{m}\Big\{\Big(E\Big[\int_0^t \|(S(t-s)(B_k^2(\sigma^{(N)}(s)-\sigma(s)))\|_{\mathcal{H}}^p ds\Big]\Big)^{\frac{1}{p}} \\
& +\Big(E\Big[\int_0^t \|(S^{(N)}(t-s)-S(t-s))(B_k^2(P_N-I)(\sigma(s)-\sigma(t))\|_{\mathcal{H}}^p ds\Big]\Big)^{\frac{1}{p}} \\
& +\Big(E\Big[\int_0^t \|(S^{(N)}(t-s)-S(t-s))(B_k^2(P_N-I)(\sigma(t)))\|_{\mathcal{H}}^p ds\Big]\Big)^{\frac{1}{p}}\Big\} \\
=: & J_{11}+J_{12}+J_{13}.
\end{aligned} \tag{4.12}$$

Then, we shall estimate them separately. Trivally, we have P_N is a bounded operator. Together with Assumption 4.2, it yields

$$J_{11} \le C(S(t),K_B)\int_0^t E[\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p]^{\frac{1}{p}} ds. \tag{4.13}$$

The term J_{12} can be bounded by applying Lemma 4.5 and Assumption 4.2. Then, we get

$$\begin{aligned}
J_{12} \le & \frac{1}{2}K_B\int_0^t (E[\|(S^{(N)}(t)-S(t))(\sigma(s)-\sigma(t))\|_{\mathcal{H}}^p])^{\frac{1}{p}} ds \\
\le & \frac{1}{2}K_B C(S(t),\delta)N^{-\alpha\delta}(\int_0^t E[\|(\sigma(s)-\sigma(t))\|_{\mathcal{H}_\delta^A}^p]ds)^{\frac{1}{p}}.
\end{aligned} \tag{4.14}$$

According to Proposition 4.3, $\|\sigma(t)\|_{\mathcal{H}_\delta^A}$ is bounded, so J_{12} can be bounded as

$$J_{12} \le C(S(t),\delta,K_B,\|\sigma(0)\|_{\mathcal{H}_\delta^A},T)N^{-\alpha\delta}. \tag{4.15}$$

With the same procedure, we can bound J_{13} as follows,

$$J_{13} \le C(S(t),\delta,K_B,\|\sigma(0)\|_{\mathcal{H}_\delta^A},T)N^{-\alpha\delta}. \tag{4.16}$$

A combination of the estimates (4.13),(4.15) and (4.16) yields

$$J_1 \le C(S(t),\delta,K_B,\|\sigma(0)\|_{\mathcal{H}_\delta^A},T)N^{-\alpha\delta} + C(S(t),K_B)\int_0^t E[\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p]^{\frac{1}{p}} ds. \tag{4.17}$$

Next, we shall start to estimate the J_2. First, we apply the BDG inequality in [8] and get

$$\begin{aligned}
& \left(E\left\|\int_0^t \sum_{k=1}^{m}(S^{(N)}(t-s)(B_k\sigma^{(N)}(s))-S(t-s)(B_k\sigma(s)))dy_s\right\|_{\mathcal{H}}^p\right)^{\frac{1}{p}} \\
\le & \left(E[\int_0^t \sum_{k=1}^{m}\|(S^{(N)}(t-s)B_k(\sigma^{(N)}(s))-S(t-s)B_k(\sigma(s)))^2 ds\|_{\mathcal{H}}^{\frac{p}{2}}]\right)^{\frac{1}{p}}.
\end{aligned} \tag{4.18}$$

We can notice that the right-hand side is a kind of the norm. Therefore, by employing the embedding relations of the L^p spaces and the triangle inequality, we can decompose this term into three components,

$$
\begin{aligned}
J_2 \leq & (\int_0^t \sum_{k=1}^m E\|S(t-s)(B_k(\sigma^{(N)}(s)-\sigma(s)))\|_{\mathcal{H}}^p ds)^{\frac{1}{p}} \\
& +(\int_0^t \sum_{k=1}^m E\|(S^{(N)}(t-s)-S(t-s))(B_k P_N(\sigma(s)-\sigma(t)))\|_{\mathcal{H}}^p ds)^{\frac{1}{p}} \\
& +(\int_0^t \sum_{k=1}^m E\|(S^{(N)}(t-s)-S(t-s))(B_k P_N(\sigma(t)))\|_{\mathcal{H}}^p ds)^{\frac{1}{p}} \\
& =: J_{21}+J_{22}+J_{23}
\end{aligned} \tag{4.19}
$$

In a similar way as for J_{11}, we can estimate the J_{21} by the boundness of $S(t)$ and Assumption 4.2, i.e.,

$$
J_{21} \leq C(S(t), K_B)(\int_0^t E\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p ds)^{\frac{1}{p}}. \tag{4.20}
$$

For J_{21}, we can use the method in J_{12} and Lemma 4.5,

$$
J_{22} \leq C(S(t),\delta)N^{-\alpha\delta}(\int_0^t \sum_{k=1}^m E\|(B_k(\sigma(s)-\sigma(t)))\|_{\mathcal{H}_\delta^A}^p ds)^{\frac{1}{p}}. \tag{4.21}
$$

It is easy to see that

$$
\begin{aligned}
\sum_{k=1}^m \|B_k(\sigma(s)-\sigma(t))\|_{\mathcal{H}_\delta^A}^p & \leq (\sum_{k=1}^m \|B_k(\sigma(s)-\sigma(t))\|_{\mathcal{H}_\delta^A})^p \\
& \leq (2K_B \cdot \sup_{t\in[0,T]} \|\sigma(t)\|_{\mathcal{H}_\delta^A})^p.
\end{aligned} \tag{4.22}
$$

Using (4.22), J_{22} can be bounded by

$$
J_{22} \leq C(S(t),\delta,K_B,T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})N^{-\alpha\delta}. \tag{4.23}
$$

The estimate of J_{23} can be calculated as the same reason in J_{22}, and we can obtain

$$
J_{23} \leq C(S(t),\delta,K_B,T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})N^{-\alpha\delta}. \tag{4.24}
$$

Coming back to the (4.19), by using (4.20),(4.23) and (4.24) we conclude that

$$
\begin{aligned}
J_2 \leq & \Big(C(S(t),K_B)(\int_0^t E\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p ds)^{\frac{1}{p}} \\
& +C(S(t),\delta,K_B,T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})\Big) N^{-\alpha\delta},
\end{aligned} \tag{4.25}
$$

where the last inequality holds, since $(\int_0^t E\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^2 ds)^{\frac{1}{2}} \leq (\int_0^t E\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p ds)^{\frac{1}{p}}$. Now, we have finished estimating the I_2. Combining the (4.17), (4.25), (4.10), (4.11) and back to (4.9), we have

$$
\begin{aligned}
& (E\|\sigma^{(N)}(t)-\sigma(t)\|_{\mathcal{H}}^p)^{\frac{1}{p}} \leq \tilde{C}_1(S(t),\delta,K_B,T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})N^{-\alpha\delta} \\
& + \tilde{C}_2(S(t),K_B)(\int_0^t E\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p ds)^{\frac{1}{p}}.
\end{aligned} \tag{4.26}
$$

Taking the p power for both sides and using mean-value inequality, we have

$$
\begin{aligned}
E\|\sigma^{(N)}(t)-\sigma(t)\|_{\mathcal{H}}^p &\le \hat{C}_1(S(t),\delta,K_B,T,\|\sigma(0)\|_{\mathcal{H}_\delta^A},p)N^{-p\alpha\delta} \\
&+\hat{C}_2((S(t),K_B),p)(\int_0^t E\|\sigma^{(N)}(s)-\sigma(s)\|_{\mathcal{H}}^p ds).
\end{aligned} \tag{4.27}
$$

Donating the $e(t) := E\|\sigma^{(N)}(t)-\sigma(t)\|_{\mathcal{H}}^p$ and taking the differential for both side of (4.27), it yields $\frac{d}{dt}e(t) \le \hat{C}_2 e(t)$. By standard Gronwall's inequality, we have

$$
e(t) \le C(S(t),\delta,K_B,T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})N^{-p\alpha\delta}. \tag{4.28}
$$

So, we finally finish the proof. □

4.2.2. Error analysis of the fully discrete scheme of mild SGM . This subsection is devoted to analyzing the error between the conditional density function $\sigma(t_k)$ and its approximation $\sigma_{t_k}^{(N)}$ by the fully discrete scheme of the mild SGM. We first give a lemma which will be used later.

LEMMA 4.7. *Consider the DMZ equation* (2.6)*. Assumptions Assumption* 4.1 *and Assumption* 4.2 *are satisfied and with the initial condition $\sigma(0) \in \mathcal{H}_\delta^A$ where $\delta > 0$, for any $t \in [0,T]$, there is*

$$
\sum_{k=1}^m \|(S^{(N)}(t)-S^{(N)}([t]_\tau))P_N B_k(\sigma(s))\|_{\mathcal{H}} \le C(S(t),\tau,K_{B_\delta})(1+\|\sigma\|_{\mathcal{H}_\delta^A})\tau^\delta, \tag{4.29}
$$

where $[t]_\tau := \min\{t_i, t_i \ge t, i=0,1,\cdots,K\}$ and $t_i = i\tau, \tau = \frac{T}{K}$.

Proof. Firstly,

$$
\begin{aligned}
\sum_{k=1}^m \|S^{(N)}(t)-S^{(N)}([t]_\tau)B_k(\sigma)\|_{\mathcal{H}} &= \sum_{k=1}^m \|P_N(I-S(t-[t]_\tau))S([t]_\tau))B_k(\sigma)\|_{\mathcal{H}} \\
&\le C(S(t),\tau)(t-[t]_\tau)^\delta \sum_{k=1}^m \|B_k(\sigma)\|_{\mathcal{H}_\delta^A}^2 \\
&\le C(S(t),\tau,K_{B^\delta})\tau^\delta(1+\|\sigma\|_{\mathcal{H}_\delta^A}).
\end{aligned}
$$

where the first inequality comes from the Lemma 2.3 and the second one is due to Assumption 4.2. □

The main goal in this paper is to analyze the error between $\sigma(t_k)$ and $\sigma_{t_k}^{(N)}$. Now, we present the main results of this paper.

THEOREM 4.8 (Main Theorem). *Let $0 < \delta < 1$, assume $\sigma(0) \in \mathcal{H}_\delta^A$ and that Assumptions Assumption* 4.1 *and Assumption* 4.2 *are satisfied. Given $\sigma_{t_k}^{(N)}$ as the fully discrete schemes of mild SGM for* (2.6)*, as defined in* (2.17)*, then there exists a constant $C := C(S(t),K_B,K_{B^\delta},T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})$ such that*

$$
\sup_{0\le k\le K}(E\|\sigma(t_k)-\sigma_{t_k}^{(N)}\|_{\mathcal{H}}^p)^{\frac{1}{p}} \le C(N^{-\alpha\delta}+\tau^{\min\{\delta,\frac{1}{2}\}}), \tag{4.30}
$$

where α is defined in the Assumption 4.1*, and $\sigma(t)$ is the solution of* (2.6)*.*

Proof. Using Theorem 4.6, and the triangle inequality, it suffices to prove that

$$\sup_{0\le k\le K}(E\|\sigma^{(N)}(t_k)-\sigma^{(N)}_{t_k}\|^p_{\mathcal{H}})^{\frac{1}{p}}\le C\tau^{\delta}. \tag{4.31}$$

According to the scheme (2.17), we have

$$\begin{aligned}\sigma^{(N)}_{t_k}&=(S^{(N)}(\tau))^k\sigma^{(N)}(0)+\sum_{i=1}^{k-1}S^{(N)}(i\tau)\sum_{j=1}^{m}B_j(\sigma^{(N)}_{t_{k-i}})\circ\Delta y^j_k,\\&=(S^{(N)}(t_k)\sigma^{(N)}(0)+\sum_{i=1}^{k-1}S^{(N)}(i\tau)\sum_{j=1}^{m}B_j(\sigma^{(N)}_{t_{k-i}})\circ\Delta y^j_k,\end{aligned} \tag{4.32}$$

where $(S^{(N)}(\tau))^k=S^{(N)}(k\tau)$ holds since it is a semi-group. And we decompose the error between $\sigma^{(N)}(t_k)$ and $\sigma^{(N)}_{t_k}$ as

$$\begin{aligned}(E\|\sigma^{(N)}(t_k)-\sigma^{(N)}_{t_k}\|^p_{\mathcal{H}})^{\frac{1}{p}}=\Big(E\Big\|\sum_{i=0}^{k-1}\sum_{j=1}^{m}\Big\|\int_{t_i}^{t_{i+1}}\Big(S^{(N)}(t_k-s)P_NB_j(\sigma^{(N)}(s))\\-S^{(N)}(t_k-t_i)P_NB_j(\sigma^N_i)\Big)\circ dy^j_s\Big\|^p_{\mathcal{H}}\Big)^{\frac{1}{p}}\end{aligned} \tag{4.33}$$

The stratonovich stochastic integral in (4.33) can be transformed into Itô's form. Then according to $S^{(N)}=S^{(N)}P_N$ and for any $k=1,\cdots,K$, the error can be bounded by the triangle inequality,

$$\begin{aligned}&(E\|\sigma^{(N)}(t_k)-\sigma^N_k\|^p_{\mathcal{H}})^{\frac{1}{p}}\le\\&\frac{1}{2}\Big(E\Big\|\sum_{i=0}^{k-1}\sum_{j=1}^{m}\int_{t_i}^{t_{i+1}}S^{(N)}(t_k-s)B_j^2(\sigma^{(N)}(s)-\sigma^N_{t_i})ds\Big\|^p_{\mathcal{H}}\Big)^{\frac{1}{p}}\\&+\Big(E\Big\|\sum_{i=0}^{K-1}\sum_{j=1}^{m}\int_{t_i}^{t_{i+1}}[(S^{(N)}(t_k-s)B_j(\sigma^{(N)}(s)-\sigma^N_{t_i})]dy^j_s\Big\|^p_{\mathcal{H}}\Big)^{\frac{1}{p}}\\&=:I_1+I_2.\end{aligned} \tag{4.34}$$

Next, we only need to estimate the I_1, and I_2. For the first term I_1 we can notice that the right-hand side is a kind of norm. By choose $k=K$, we get

$$\begin{aligned}I_1\le&\frac{1}{2}\Big(\sum_{i=0}^{K-1}\int_{t_i}^{t_{i+1}}\sum_{j=1}^{m}E\|(S^{(N)}(t_{i+1}-s)-S^{(N)}(t_{i+1}-t_i))B_j^2(\sigma^{(N)}(s))\|^p_{\mathcal{H}}ds\Big)^{\frac{1}{p}}\\&+\frac{1}{2}\Big(\sum_{i=0}^{K-1}\int_{t_i}^{t_{i+1}}\sum_{j=1}^{m}(E\|S^{(N)}(t_{i+1}-t_i)B_j^2(\sigma^{(N)}(s)-\sigma^{(N)}(t_i))\|^p_{\mathcal{H}}ds\Big)^{\frac{1}{p}}\\&+\frac{1}{2}\Big(\sum_{i=0}^{K-1}\int_{t_i}^{t_{i+1}}E\|S^{(N)}(t_{i+1}-t_i)B_j^2(\sigma^{(N)}(t_i)-\sigma^N_{t_i}))\|^p_{\mathcal{H}}ds\Big)^{\frac{1}{p}}\\&=:I_{11}+I_{12}+I_{13}.\end{aligned}$$

By using Lemma 4.7, I_{11} is bounded as follows

$$\begin{aligned} I_{11} &\le C(S(t),K_{B^\delta})\left(\sum_{i=0}^{K-1}\int_{t_i}^{t_{i+1}}\tau^{2\delta}(1+\|\sigma^{(N)}\|_{\mathcal{H}_\delta^A})^p ds\right)^{\frac{1}{p}} \\ &= C(S(t),K_{B^\delta})\tau^\delta\left(\int_0^T(1+\|\sigma^{(N)}(s)\|_{\mathcal{H}_\delta^A})^p ds\right)^{\frac{1}{p}} \\ &\le C(S(t),K_{B^\delta},\|\sigma(0)\|_{\mathcal{H}_\delta^A},T)\tau^\delta, \end{aligned} \tag{4.35}$$

where the last inequality holds due to (4.4). Then, we estimate I_{12}, and we have

$$I_{12}\le C(S(t),K_B)\left(\sum_{i=0}^{K-1}\int_{t_i}^{t_{i+1}}E\|\sigma^{(N)}(s)-\sigma^{(N)}(t_i)\|_{\mathcal{H}}^p ds\right)^{\frac{1}{p}} \tag{4.36}$$

According to (4.5) in Proposition 4.3, we get

$$\begin{aligned} I_{12} &\le C(S(t),K_{B^\delta},T,\|\sigma(0)\|_{\mathcal{H}_\delta^A})\tau^{\min\left\{\delta,\frac{1}{2}\right\}} \\ &+ C(S(t),K_B,T)\left(\sum_{i=0}^{K-1}\sup_{s\in[t_i,t_{i+1}]}E\|\sigma^{(N)}(t_i)-\sigma_{t_i}^{(N)}\|_{\mathcal{H}}^p\tau\right)^{\frac{1}{p}}. \end{aligned} \tag{4.37}$$

Next, we shall estimate I_{13}. By using Assumption 4.2, we have

$$I_{13}\le C(S(t),K_B,T)\left(\sum_{i=0}^{K-1}\sup_{s\in[t_i,t_{i+1}]}E\|\sigma^{(N)}(t_i)-\sigma_{t_i}^{(N)}\|_{\mathcal{H}}^p\tau\right)^{\frac{1}{p}}. \tag{4.38}$$

Collecting the above estimations for I_{11}, I_{12}, I_{13} together, we have that

$$\begin{aligned} I_1^p &\le C_1(S(t),K_B,K_{B^\delta},T,p)\tau^{p\min\left\{\delta,\frac{1}{2}\right\}} \\ &+ C_2(S(t),K_B,T,p)(\sum_{i=0}^{k-1}E\|\sigma^{(N)}(t_i)-\sigma_i^{(N)}\|_{\mathcal{H}}^p\tau). \end{aligned} \tag{4.39}$$

To estimate I_2, we utilize the BDG inequality in [8] and get

$$\begin{aligned} I_2 &\le C(S(t),K_B,T)\left(\sum_{i=0}^{K-1}\sup_{s\in[t_i,t_{i+1}]}E\|\sigma^{(N)}(s)-\sigma_{t_i}^{(N)}\|_{\mathcal{H}}^p\tau\right)^{\frac{1}{p}} \\ &\le C(S(t),K_B,T)\left(\sum_{i=0}^{K-1}\sup_{s\in[t_i,t_{i+1}]}E\|\sigma^{(N)}(s)-\sigma_{t_i}^{(N)}\|_{\mathcal{H}}^p\tau\right)^{\frac{1}{p}} \end{aligned} \tag{4.40}$$

Let $e(k):=E\|\sigma^{(N)}(t_k)-\sigma_{t_k}^{(N)}\|_{\mathcal{H}}^p$. Using (4.34), (4.39) and (4.40), we have

$$e(k)\le\tilde{C}_1(S(t),K_B,K_{B^\delta},T,p))\tau^{p\min\left\{\delta,\frac{1}{2}\right\}}+\tilde{C}_2(S(t),K_B,T,p)(\sum_{i=0}^{k-1}e(i)\tau). \tag{4.41}$$

And consider $e(k+1)-e(k)$, we get

$$|e(k+1)-e(k)|\le\tau\tilde{C}_2e(k). \tag{4.42}$$

Trivially, we can choose a constant C_0 such that,

$$e(1) \le C_0 \tilde{C}_1(S(t), K_B, K_{B^\delta}, T, p))\tau^{p \min\left\{\delta, \frac{1}{2}\right\}} e^{\tilde{C}_2(S(t), K_B, T, p)\tau}. \tag{4.43}$$

We construct a function

$$\hat{e}(k) := C_0 \tilde{C}_1(S(t), K_B, K_{B^\delta}, T)) e^{\tilde{C}_2(S(t), K_B, T)\tau k}.$$

If we assume that for any $k < n$ the $e(k) < \hat{e}(k)$ holds, then by (4.42), we have

$$\begin{aligned} e(n+1) &< |e(n+1) - e(n)| + e(n) \le (1 + \tau \tilde{C}_2)\hat{e}(n) \\ &\le e^{\tilde{C}_2 \tau} \hat{e}(n) = \hat{e}(n+1). \end{aligned} \tag{4.44}$$

So we prove $e(k) \le \hat{e}(k)$ for any k by induction. Then,

$$E\|\sigma^{(N)}(t_k) - \sigma^{(N)}_{t_k}\|^p_{\mathcal{H}} \le \hat{e}(K) \le C(S(t), K_B, K_{B^\delta}, T)\tau^{p \min\left\{\delta, \frac{1}{2}\right\}} \tag{4.45}$$

So, we finish the proof. □

Corollary 4.9 (Convergence results of the direct method and Galerkin Yau-Yau algorithm). *Let $0 < \delta < 1$, assume that $\sigma(0) \in \mathcal{H}^A_\delta$ and the Assumption* 4.1 *and Assumption* 4.2 *are satisfied. $\sigma_1(t_k, x)$ is one of the following:*

1. *$\sigma_1(t_k, x)$ is the numerical approximation of the direct method with N-th order truncated Gaussian approximation, which is defined in Proposition* 3.4.
2. *$\sigma_1(t_k, x)$ is the numerical approximation of Galerkin Yau-Yau algorithm with basis $\mathcal{H}_N = \langle \phi_1, \cdots, \phi_N \rangle$, which is defined in Proposition* 3.5.

Then, there exists a constant $C := C(S(t), K_B, K_{B^\delta}, T, \|\sigma(0)\|_{\mathcal{H}^A_\delta})$ such that

$$\sup_{0 \le k \le K} (E\|\sigma(t_k) - \sigma_1(t_k, x)\|^p_{\mathcal{H}})^{\frac{1}{p}} \le C(N^{-\alpha\delta} + \tau^{\min\left\{\delta, \frac{1}{2}\right\}}), \tag{4.46}$$

where α is defined in Assumption 4.1*, and $\sigma(t)$ is the solution of* (2.6).

Proof of Corollary 4.9 The two types of $\sigma_1(t_k, x)$ correspond to the fully discrete schemes of mild SGM. The proof is direct consequence of Theorem 4.8. □

4.3. Numerical Experiments. In this subsection, we chose this model specifically as other comparative algorithms [13, 41] cannot handle cubic sensor problems, highlighting our method's unique advantage. We implemented the Hermite Spectral Yau-Yau algorithm [21, 32] with scaling factor 2.4637. All numerical experiments were conducted in Python on a Mac Pro 2024 laptop. The system dynamics are defined as:

$$\begin{cases} dx_t = dV_t, \quad x_0 \sim \sigma_0 = e^{-x^4} \\ dy_t = x_t^3 dt + dW_t, \quad y_0 = 0 \end{cases}$$

where dW_t and dV_t are scalar independent Brownian motion processes. We investigate the relationship between the convergence rate, the number of spectral functions N, and the time discretization step size τ. For evaluation metrics, we employ the commonly used Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}, \quad \text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \tag{4.47}$$

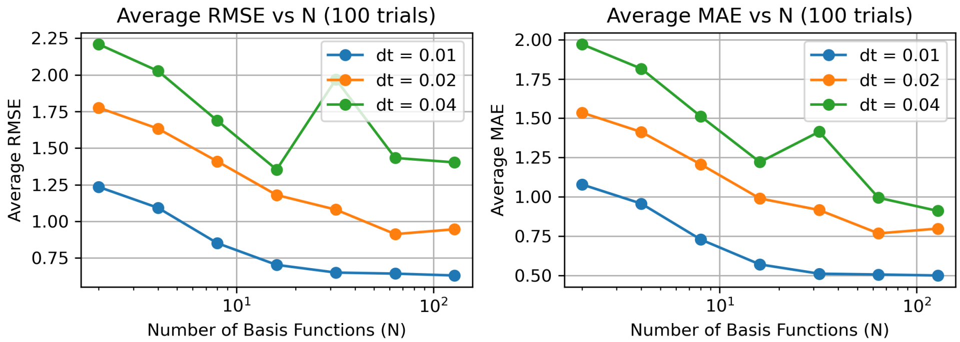

FIG. 1. *RMSE and MAE via N*

where x_i represents the true state and $\hat{x}_i$ represents the estimated state. All reported results are averaged over 100 Monte Carlo simulation runs. We conducted a parameter sweep over time steps $\tau = \{0.01, 0.02, 0.04\}$ and number of basis functions $N = \{2, 4, 8, 16, 32, 64, 128\}$, with total steps fixed at 1000. The theorem suggests an error bound of $C(N^{-\alpha\delta} + \tau^{\min\{\delta, \frac{1}{2}\}})$, where error decreases with increasing N and decreasing τ. The results are presented in Figure 1.

As shown in Figure 1, our results confirm the theoretical predictions. We observe that for a fixed τ, increasing the number of basis functions N consistently reduces RMSE and MAE across all tested time steps, supporting the theoretical prediction of improved approximation with larger N. Conversely, when N is fixed, smaller values of τ (i.e., finer time discretization) lead to lower RMSE and MAE, aligning with expectations from our main theorem regarding the roles of N and τ.

From the MAE perspective, performance improves monotonically with increasing N. However, RMSE exhibits a non-monotonic trend: while initially decreasing with larger N, it eventually plateaus or slightly increases, particularly noticeable at $\tau = 0.04$. This suggests the existence of an optimal N for each τ value. We observe that the optimal N values ($N = 16$ for $\tau = 0.04$, $N = 64$ for $\tau = 0.02$, $N = 128$ for $\tau = 0.01$) approximately follow the relationship $N \sim \tau^{-\frac{\alpha\delta}{2}} \approx \tau^{-1.05}$, which is consistent with our theoretical analysis.

5. Conclusion.

In this paper, we develop a convergence analysis framework specifically for the direct and Yau-Yau algorithms, further introducing convergence analyses concerning spectral number and time steps. Our findings reveal that, for smooth enough $\sigma(t) \in \mathcal{H}_\delta$, i.e. $\delta \geq 0.5$, the error upper bound for time discretization is of order 0.5, and the convergence speed for the spectral number is $N^{-\alpha\delta}$. This implies a relationship between the time discretization step and the spectral number i.e. $\tau^{-\frac{\alpha\delta}{2}} \approx N$, as corroborated by numerical experiments across a series of works, such as the direct method appeared in [29, 6] and Yau-Yau algorithms [22, 21, 32, 10, 19].

Appendix A. Appendix for Section IV. Proof of Lemma 4.4

By Assumption 4.1, we know that $(\phi_i)_{i=1}^{\infty}$ is an orthogonal basis of $\mathcal{H}$. So, we have

$$|\lambda_{N+1}|^{2\delta} \|(P_N - I)X\|_{\mathcal{H}}^2 = |\lambda_{N+1}|^{2\delta} \sum_{i=N+1}^{\infty} \langle X, \phi_i \rangle^2. \tag{A.1}$$

We know that $|\lambda_i| \geq |\lambda_{N+1}|$ for $i \geq N+1$ from (2.11) and using the Definition Definition 2.4, we have

$$\begin{aligned}|\lambda_{N+1}|^{2\delta} \sum_{i=N+1}^{\infty} \langle X, \phi_i \rangle^2 &\leq \sum_{i=N+1}^{\infty} |\lambda_i|^{2\delta} \langle X, \phi_i \rangle^2 \\ &\leq \|(-A)^{\delta} X\|_{\mathcal{H}}^2 = \|X\|_{\mathcal{H}_\delta^A}^2 .\end{aligned} \tag{A.2}$$

So, we finish the proof. □

Proof of Lemma 4.5 According to Lemma 4.4, we have

$$|\lambda_{N+1}|^{\delta} \|(P_N - I)X\|_{\mathcal{H}} \leq \|X\|_{\mathcal{H}_\delta^A}, \quad \forall X \in \mathcal{H}_\delta^A, \delta > 0. \tag{A.3}$$

Firstly,

$$\begin{aligned}&\|S^{(N)}(t)\sigma^{(N)}(0) - S(t)\sigma(0)\|_{\mathcal{H}} \leq \|S(t)\sigma^{(N)}(0) - S(t)\sigma(0)\|_{\mathcal{H}} \\ &+ \|S^{(N)}(t)\sigma^{(N)}(0) - S(t)\sigma^{(N)}(0)\|_{\mathcal{H}} \\ &= \|(P_N - I)S(t)\sigma(0)\|_{\mathcal{H}} + \|(P_N - I)(S(t)P_N)\sigma(0)\|_{\mathcal{H}} \\ &\leq 2\|(P_N - I)S(t)\sigma(0)\|_{\mathcal{H}}.\end{aligned} \tag{A.4}$$

According to Lemma 2.3, $S(t)$ is the bounded operator. Combining (A.3), we get

$$\begin{aligned}&\|S^{(N)}(t)\sigma^{(N)}(0) - S(t)\sigma(0)\|_{\mathcal{H}} \leq 2\|(P_N - I)S(t)\sigma(0)\|_{\mathcal{H}} \\ &\leq C(S(t))|\lambda_{N+1}|^{-\delta} \|\sigma(0)\|_{\mathcal{H}_\delta^A}\end{aligned} \tag{A.5}$$

Finally, the proof is completed by the estimation of $|\lambda_{N+1}|$ in Assumption 4.1. □

REFERENCES

[1] B. Aupetit, *A primer on spectral theory*, Springer Science & Business Media, 2012.

[2] B. Balaji, *Universal nonlinear filtering using Feynman path integrals II: the continuous-continuous model with additive noise*, PMC Physics A, 3 (2009), pp. 1–28.

[3] A. Barth and A. Lang, *Simulation of stochastic partial differential equations using finite element methods*, International Journal of Probability and Stochastic Processes, 84 (2012), pp. 217–231.

[4] W. Bordeaux Montrieux and J. Sjöstrand, *Almost sure Weyl asymptotics for non-self-adjoint elliptic operators on compact manifolds*, in Annales de la Faculté des sciences de Toulouse: Mathématiques, vol. 19, 2010, pp. 567–587.

[5] X. Chen, X. Luo, and S. S.-T. Yau, *Direct method for time-varying nonlinear filtering problems*, IEEE Transactions on Automatic Control, 53 (2017), pp. 630–639.

[6] X. Chen, J. Shi, and S. S.-T. Yau, *Real-time solution of time-varying Yau filtering problems via direct method and Gaussian approximation*, IEEE Transactions on Automatic Control, 64 (2019), pp. 1648–1654, https://doi.org/10.1109/TAC.2018.2853576.

[7] S. Cox and E. Hausenblas, *Pathwise space approximations of semi-linear parabolic SPDEs with multiplicative noise*, International Journal of Computer Mathematics, 89 (2012), pp. 2460–2478.

[8] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge university press, 2014.

[9] M. H. A. Davis, *On a multiplicative functional transformation arising in nonlinear filtering theory*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 54 (1980), pp. 125–139.

[10] W. Dong, X. Luo, and S. S.-T. Yau, *Solving nonlinear filtering problems in real time by Legendre Galerkin spectral method*, IEEE Transactions on Automatic Control, 66 (2021), pp. 1559–1572, https://doi.org/10.1109/TAC.2020.3002979.

[11] T. E. Duncan, *Probability densities for diffusion processes with applications to nonlinear filtering system (Ph.D. dissertation)*, PhD dissertation for Stanford University., (1967).

[12] L. C. Evans, *Partial differential equations (graduate studies in mathematics, vol. 19)*, Instructor, 67 (2009).
[13] R. Frey, T. Schmidt, and L. Xu, *On Galerkin approximations for the Zakai equation with diffusive and point process observations*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 2036–2062.
[14] C. R. Genovese and L. Wasserman, *Rates of convergence for the Gaussian mixture sieve*, The Annals of Statistics, 28 (2000), pp. 1105–1127.
[15] G.-Q. Hu and S. S.-T. Yau, *Finite-dimensional filters with nonlinear drift. XV. new direct method for construction of universal finite-dimensional filter*, IEEE Transactions on Aerospace and Electronic Systems, 38 (2002), pp. 50–57.
[16] A. Jentzen and M. Röckner, *Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise*, Journal of Differential Equations, 252 (2012), pp. 114–136.
[17] R. Kruse and S. Larsson, *Optimal regularity for semilinear stochastic partial differential equations with multiplicative noise*, Electronic Journal of Probability, 17 (2012), pp. 1–19.
[18] A. Lanconelli and R. Scorolli, *A small time approximation for the solution to the Zakai equation*, Potential Analysis, (2023), pp. 1–11.
[19] S. Li, Z. Wang, S. S.-T. Yau, and Z. Zhang, *Solving nonlinear filtering problems using a tensor train decomposition method*, IEEE Transactions on Automatic Control, (2022).
[20] L. Lorenzi and A. Rhandi, *Semi-groups of bounded operators and second-order elliptic and parabolic partial differential equations*, CRC Press, 2021.
[21] X. Luo and S. S.-T. Yau, *Hermite spectral method to 1-d forward Kolmogorov equation and its application to nonlinear filtering problems*, IEEE Transactions on Automatic Control, 58 (2013), pp. 2495–2507.
[22] X. Luo and S.-T. Yau, *Complete real time solution of the general nonlinear filtering problem without memory*, IEEE Transactions on Automatic Control, 58 (2013), pp. 2563–2578.
[23] C. Martinez and M. Sanz, *The theory of fractional powers of operators*, Elsevier, 2001.
[24] R. E. Mortensen, *Optimal control of continuous time stochastic systems (Ph.D. dissertation)*, (PhD dissertation). University of California, Berkeley., (1966).
[25] Y. Nakano, *Kernel-based collocation methods for Zakai equations*, Stochastics and Partial Differential Equations: Analysis and Computations, 7 (2019), pp. 476–494.
[26] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer Science & Business Medi, 2012.
[27] R. A. Redner and H. F. Walker, *Mixture densities, maximum likelihood and the em algorithm*, SIAM review, 26 (1984), pp. 195–239.
[28] J. Shi, X. Chen, W. Dong, and S. S. T. Yau, *New classes of finite dimensional filters with non-maximal rank*, IEEE Control Systems Letters, 1 (2017), pp. 233–237.
[29] J. Shi, Z. Yang, and S. S.-T. Yau, *Direct method for Yau filtering system with nonlinear observations*, International Journal of Control, (2017), pp. 1–10.
[30] L. I. Smith, *A tutorial on principal components analysis*, (2002).
[31] F. Tuerlinckx, F. Rijmen, G. Verbeke, and P. De Boeck, *Statistical inference in generalized linear mixed models: A review*, British Journal of Mathematical and Statistical Psychology, 59 (2006), pp. 225–255.
[32] Z. Wang, X. Luo, S. S.-T. Yau, and Z. Zhang, *Proper orthogonal decomposition method to nonlinear filtering problems in medium-high dimension*, IEEE Transactions on Automatic Control, 65 (2019), pp. 1613–1624.
[33] S. S.-T. Yan and G.-Q. Hu, *Finite-dimensional filters with nonlinear drift X: Explicit solution of DMZ equation*, IEEE Transactions on Automatic Control, 46 (2001), pp. 142–148.
[34] S.-T. Yau and Y. T. Lai, *Explicit solution of DMZ equation in nonlinear filtering via solution of ODEs*, IEEE Transactions on Automatic Control, 48 (2003), pp. 505–508.
[35] S.-T. Yau and S. Yau, *New direct method for Kalman-Bucy filtering system with arbitrary initial condition*, in Proceedings of 1994 33rd IEEE Conference on Decision and Control, vol. 2, IEEE, 1994, pp. 1221–1225.
[36] S.-T. Yau and S. S.-T. Yau, *Finite-dimensional filters with nonlinear drift. III: Duncan-Mortensen-Zakai equation with arbitrary initial condition for the linear filtering system and the benes filtering system*, IEEE Transactions on Automatic Control, 33 (1997), pp. 1277–1294.
[37] S.-T. Yau and S. S. T. Yau, *Real time solution of nonlinear filtering problem without memory I*, Mathematical Research Letters, (2000).
[38] S.-T. Yau and S. S.-T. Yau, *Nonlinear filtering and time varying Schrodinger equation*, IEEE Transactions on Automatic Control, 40 (2004), pp. 284–292.
[39] S.-T. Yau and S.-T. Yau, *Real time solution of the nonlinear filtering problem without memory*

II, SIAM Journal on Control and Optimization, (2008).

[40] M. Zakai, *On the optimal filtering of diffusion processes*, Probability Theory and Related Fields, 11 (1969), pp. 230–243.

[41] F. Zhang, Y. Zou, S. Chai, R. Zhang, and Y. Cao, *Splitting-up spectral method for nonlinear filtering problems with correlation noises*, Journal of Scientific Computing, 93 (2022), p. 25.