On Numerical Approximations of Forward-Backward Stochastic Differential Equations *

Jin Ma  Jie Shen  and  Yanhong Zhao

Department of Mathematics, Purdue University
West Lafayette, IN 47907-1395, USA
majin@math.purdue.edu, shen@math.purdue.edu, zhao6@purdue.edu

Abstract. A numerical method for a class of forward-backward stochastic differential equations (FBSDEs) is proposed and analyzed. The method is designed around the Four Step Scheme (Douglas-Ma-Protter, 1996) but with a Hermite-spectral method to approximate the solution to the decoupled quasilinear PDE on the whole space. A rigorous synthetic error analysis is carried out for a fully discretized scheme, namely a first-order scheme in time and a Hermite-spectral scheme in space, of the FBSDEs. Equally important, a systematical numerical comparison is made between several schemes for the resulting decoupled forward SDE, including a stochastic version of the Adams-Bashforth scheme. It is shown that the stochastic version of the Adams-Bashforth scheme coupled with the Hermite-spectral method leads to a convergence rate of $\frac{3}{2}$ (in time) which is better than those in previously published work for the FBSDEs.

Keywords: Forward-backward SDEs, four step scheme, spectral method, Hermite functions, stochastic Adam-Bashforth scheme.

2000 Mathematics Subject Classification. Primary: 60H10, Secondary: 34F05, 93E03

1 Introduction

Since the seminal work of Pardoux-Peng [25] in early 1990’s, the theory of backward and forward-backward stochastic differential equations (BSDEs and FBSDEs, respectively, for short) has grown into an ubiquitous tool in the fields of stochastic optimizations and mathematical finance (see, for example, the books of El Karoui-Mazliak [12], Ma-Yong [21], and the survey of El Karoui-Peng-Quenez [13] for the detailed account for both theory and applications of such SDEs). In the meantime, after the earlier works of Bally [1] and Douglas-Ma-Protter [11], finding an efficient numerical scheme for both BSDEs and FBSDEs has also become an independent, but integral part of the theory. Tremendous efforts have been made during the past decade to circumvent the fundamental difficulties caused by the combination of the “backward” nature of the SDEs and the associated decoupling techniques for FBSDEs. In the “pure backward” (or “decoupled” forward-backward) case, various methods have been proposed. These include: PDE method in the Markovian case (cf. e.g., Chevance [7], Zhang-Zheng [28]), random walk approximations (cf. e.g., Briand-Delyon-Memin [6] and Ma-Protter-San Martin-Torres [19]), Malliavin calculus and Monte-Carlo method (cf. e.g., Zhang [27], Ma-Zhang [22], and Bouchard-Touzi [5]), and recently the quantization method (cf. e.g.,

---

*This work is supported in parts by NSF grants DMS-0311915, DMS-0505427, DMS-0610646 and a Purdue Research Foundation grant.
In the case of coupled FBSDEs, however, the results are very limited, due largely to the lack of solution method itself in such cases. It has been well-understood that in order to solve an FBSDE in an arbitrary duration, a (numerically) tractable method is to utilize the “decoupling PDE”, based on the so-called “Four Step Scheme” initiated in Ma-Protter-Yong [21]. Such an idea has led directly to most of the existing numerical results for FBSDEs: from the early work [11] to the recent improvements by Milstein-Tretyakov [24] and Delarue-Menozzi [10]. Extending such decoupling idea, and combining with an optimal control method as well as Monte-Carlo simulation, Cvitanic-Zhang [8] and Bender-Zhang [4] recently proposed a numerical scheme for FBSDEs without attacking the associated PDE directly. We should note that the Monte-Carlo method is most effective only when a single value of the solution is concerned at each computing cycle, which is quite different from the original problem where essentially the distribution of the solution at each time point was sought.

There are two main technical obstacles in developing numerical schemes for FBSDEs: the dimensionality and the rate of convergence. The former is a natural consequence of the close relation between the FBSDEs and its decoupling quasilinear PDE, where the notorious “curse of dimensionality” is still a formidable difficulty for any numerical method. In fact, to our best knowledge, there is still no efficient numerical method for high dimensional PDEs that is directly applicable to our case. We note that almost all the existing numerical schemes for FBSDEs are constructed with $\alpha = 1$ so their convergence rate is $\frac{1}{2}$, except in [24], where the Euler scheme for the forward SDE is replaced by Milstein’s first order scheme so the rate of convergence is improved to 1. We should mention that although higher order approximation is possible for forward SDEs (cf. e.g., [18]), it is by no means clear whether this can be extended to the coupled FBSDEs, even when all the coefficients are assumed to be smooth(!), given the intrinsic difficulties arising from the current numerical methods.

This then raises an interesting question: is it possible to design a numerical scheme that has better than first-order convergence rate, and at the same time is applicable (in the practical sense) to high dimensional cases? This paper is can be considered as the first step towards this goal. We shall revisit the Four Step Scheme again, but replacing the usual finite difference method for the PDE by a Hermite-spectral method. Several main features of the Hermite-spectral method are worth noting: (i) the PDE on the whole space is approximated directly by using Hermite functions which form an orthonormal basis on the whole space, while the whole space is replaced by a boundary domain with ad-hoc artificial boundary conditions in previous approaches using finite differences; (ii) the rate of convergence is related explicitly to the regularity of the coefficients, to wit, the higher regularity implies the higher convergence rate. We should remark that this last feature marks the main difference between a spectral method and traditional finite difference and finite element methods. In fact, as we pointed out before, even given smooth coefficients, none of the existing schemes for FBSDE seem to be able to achieve higher than first-order convergence rate.

Given the aforementioned spectral accuracy in space, it seems to be hopeful that one can produce arbitrarily higher order global error by applying higher order scheme in SDEs in, e.g., [18], at least when the coefficients are smooth. We shall carry out some numerical simulation to validate this point. To compare with the existing results, we shall use an example proposed in Milstein-Tretyakov [24], and we test several methods for the forward SDEs. These include: the standard Euler scheme, the Milstein scheme, the Platen-Wagner scheme, and the Stochastic Adam-Bashforth (SAB) scheme proposed in [14]. By the nature of these schemes, we expect that the Euler scheme should produce a
rate of convergence 1/2, the Milstein first-order scheme a rate of 1, and the Platen-Wagner and the SAB scheme with rate of 3/2. Our simulation results indicate that this is exactly the case. Also, this example shows that in the one-dimensional case our scheme has the best performance (in terms of computing time). Due to the apparent complexity of the error analysis and the length of the paper, in this paper we only give a rigorous proof of the rate of convergence with the Euler scheme for the forward SDE. It appears that a rigorous proof for higher order schemes will be quite similar, although conceivably much more tedious. We hope our numerical results are sufficiently convincing for this purpose.

As a final remark, we would like to point out that the numerical method in this paper can be, albeit technical, extended to higher-dimensional cases with a tensor-product approach. However, such a method quickly becomes non-feasible for spatial dimensions higher than three, unless some non tensor-product method, such as those based on sparse grid or lattice rules, is introduced. We plan to address the higher dimensional issue in a forthcoming work.

The rest of the paper is organized as follows. In Section 2 we give the necessary preliminaries; in Section 3 we introduce the Hermite-spectral method for the PDE and perform an error analysis. In Section 4 we study the synthetic error analysis of the full numerical scheme, and in Section 5 we carry out some numerical experiments.

2 Problem Formulation and Preliminaries

Throughout this paper we assume that \((\Omega, \mathcal{F}, P)\) is a complete probability space, on which is defined a d-dimensional Brownian motion \(B = \{B_t : t \geq 0\}\). We shall denote \(\mathcal{F}_B = \{\mathcal{F}_B^t : t \geq 0\}\) to be the natural filtration generated by \(B\), with usual \(P\)-augmentation so that it is right continuous and contains all the \(P\)-null sets in \(\mathcal{F}\).

We consider the following FBSDE: for \(t \in [0, T]\),

\[
\begin{aligned}
X(t) &= x + \int_0^t b(s, \Theta(s))ds + \int_0^t \sigma(s, X(s), Y(s))dW(s);
Y(t) &= \varphi(X(T)) + \int_0^T g(s, \Theta(s))ds - \int_0^T Z(s)dW(s),
\end{aligned}
\]

where \(\Theta \triangleq (X, Y, Z)\). To simplify presentation, in what follows we shall assume that all processes involved are one dimensional. Moreover, we shall make use of the following Standing Assumptions in the sequel.

(A1) The functions \(b, \sigma, g, \) and \(\varphi\) are continuously differentiable in all variables. Moreover, if we denote these functions by a generic one \(\psi\), then there exists a constant \(K > 0\), such that for any \(t \in [0, T]\), \(\psi\) satisfies the uniform Lipschitz condition:

\[
|\psi(t, x, y, z) - \psi(t, x', y', z')| \leq K(|x - x'| + |y - y'| + |z - z'|),
\]

(A2) The function \(b, \sigma, g, \) and \(\varphi\) satisfying the following growth conditions: for some \(K > 0\),

\[
\begin{aligned}
|b(t, x, y, z)| + |g(t, x, y, z)| &\leq K(|y| + |z|);
|\sigma(t, x, y)| &\leq K(1 + |y|); \\
|\varphi(x)| &\leq K.
\end{aligned}
\]
(A3) There exist constants $0 < c < C$, such that

$$c \leq \sigma(t, x, y) \leq C, \quad \forall (t, x, y) \in [0, T] \times \mathbb{R}^2.$$  \hspace{1cm} (2.3)

**Remark 2.1** We remark that the assumptions (A1)–(A3) are stronger than it is necessary for the well-posedness of FBSDE (2.1). In fact, in [9] it was shown that under (A1)–(A2) but without the differentiability assumption on the coefficients the FBSDE (2.1) already possesses a unique adapted solution over arbitrarily prescribed time duration. The extra smoothness condition is only needed for our numerical scheme, and therefore not essential, in principle. Note also, however, that even with the added differentiability conditions, our assumptions are still much weaker than that of [11].

**Four Step Scheme.** In [20] (see also [9]) it was shown that the unique adapted solution of FBSDE (2.1) can be obtained by following steps:

**Step 1.** Define a function $z : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ by

$$z(t, x, y, p) = -p \sigma(t, x, y), \quad \forall (t, x, y, p).$$

**Step 2.** Using the function $z$ above, solve the quasilinear parabolic PDE:

\[
\begin{align*}
  u_t + \frac{1}{2} \sigma^2(t, x, u)u_{xx} + b(t, x, u, z(t, x, u, u_x))u_x + g(t, x, u, \sigma(t, x, u)) &= 0, \quad (t, x) \in (0, T) \times \mathbb{R}, \\
  u(T, x) &= \varphi(x), \quad x \in \mathbb{R}.
\end{align*}
\]  \hspace{1cm} (2.4)

**Step 3.** Using the functions $u$ and $z$, solve the forward SDE:

$$X(t) = x + \int_0^t \tilde{b}(s, X(s))ds + \int_0^t \tilde{\sigma}(s, X(s))dW(s),$$  \hspace{1cm} (2.5)

where $\tilde{b} = b(t, x, u(t, x), z(t, x, u, u_x))$ and $\tilde{\sigma} = \sigma(t, x, u(t, x))$; and

**Step 4.** Setting

$$Y(t) = u(t, X(t)), \quad \text{and} \quad Z(t) = \sigma(t, X(t), u(t, X(t)))u_x(t, X(t)),$$

Then $(X, Y, Z)$ is the adapted solution to (2.1).

It is readily seen that if a numerical scheme is designed along the lines of Four Step Scheme, then one essentially has to deal with two separate discretization schemes: one for the PDE (2.4), and one for the (forward) SDE (2.5). We should note that the PDE (2.4) takes the form of a Cauchy problem, an artificial “cut-off” procedure is usually needed (see, e.g., [11]). To avoid such an extra step, we shall adopt a Hermite-Galerkin method with numerical integration. To do this we found it convenient to rewrite PDE (2.4) in a divergence form, mainly for the sake of numerical stability:

\[
\begin{align*}
  u_t + \partial_x \left( \frac{1}{2} \sigma^2(t, x, u)u_x \right) + \tilde{b}(t, x, u, u_x)u_x + g(t, x, u, \sigma(t, x, u)) &= 0, \quad (t, x) \in (0, T) \times \mathbb{R}, \\
  u(T, x) &= \varphi(x), \quad x \in \mathbb{R}.
\end{align*}
\]  \hspace{1cm} (2.6)

where

$$\tilde{b}(t, x, y, z) = b(t, x, y, \sigma(t, x, y)) - \sigma(t, x, y)(\sigma_x(t, x, y) + \sigma_y(t, x, y)z).$$
Hermite polynomials, Hermite functions, and their properties

Throughout this paper we shall denote \( L^2(\mathbb{R}) \) to be the space of all square integrable functions \( u : \mathbb{R} \rightarrow \mathbb{R} \), with the inner product
\[
\langle u, v \rangle = \int_{\mathbb{R}} u(x)v(x)dx,
\]
and the norm \( \|u\|^2 \triangleq \int_{\mathbb{R}} |u(x)|^2dx \). For simplicity, in what follows, without further specification we shall always denote \( \| \cdot \| \) to be the \( L^2(\mathbb{R}) \) norm.

Next, for each integer \( k \geq 1 \), we denote the differential operator \( \partial^k \triangleq \frac{d^k}{dx^k} \). Then, the Sobolev spaces \( H^m(\mathbb{R}) \) is defined by
\[
H^m(\mathbb{R}) = \{ u \mid \partial^k u \in L^2(\mathbb{R}) \quad \forall 0 \leq k \leq m \},
\]
with the semi-norm \( |u|_m \triangleq \|\partial^m u\| \), and the norm \( \|u\|^2_m \triangleq \sum_{k=0}^{m} \|\partial^k u\|^2 \). We note that the operator \( \partial^k \) can be easily extended to the higher dimensional case, namely as a partial differential operator, in an obvious way.

We recall that (cf. e.g., [15]) the Hermite polynomials \( \{ H_n(x) \}_{n \geq 0} \) is defined by
\[
H_n(x) = (-1)^n \exp(x^2) \frac{d^n}{dx^n} \exp(-x^2), \quad n = 0, 1, \cdots \tag{2.7}
\]
One can easily check that \( H_n \) is the solution to the following recursive equations:
\[
\begin{aligned}
H_{n+1}(x) &= 2xH_n(x) - 2nH_{n-1}(x), \\
H_n'(x) &= 2nH_{n-1}(x), \\
H_0(x) &= 1, \quad H_1(x) = 2x;
\end{aligned}
\]
and the following orthogonality condition holds:
\[
\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2}dx = \gamma_n \delta_{mn}, \quad m, n \geq 1, \tag{2.8}
\]
where \( \gamma_n = \sqrt{2^n n!} \), and \( \delta_{mn} \) is the Kronecker delta. For each \( n \geq 1 \), \( H_n(x) \) is called the “Hermite polynomial of degree \( n \)”, and the Hermite function of degree \( n \) is defined by
\[
\tilde{H}_n(x) = \frac{1}{\sqrt{2^n n!}} e^{-x^2/2} H_n(x).
\]
It follows immediately that the Hermite functions \( \tilde{H}_n(x) \) enjoy the following recurrence and orthogonal relations:
\[
\begin{aligned}
\tilde{H}_{n+1}(x) &= x\sqrt{\frac{2}{n+1}} \tilde{H}_n(x) - \sqrt{\frac{n}{n+1}} \tilde{H}_{n-1}(x), \quad n \geq 1 \tag{2.9} \\
\frac{d}{dx} \tilde{H}_n(x) &= \sqrt{\frac{n}{2}} \tilde{H}_{n-1}(x) - \sqrt{\frac{n+1}{2}} \tilde{H}_{n+1}(x) \tag{2.10} \\
\int_{-\infty}^{\infty} \tilde{H}_m(x)\tilde{H}_n(x)dx &= \sqrt{\pi} \delta_{mn}. \tag{2.11}
\end{aligned}
\]
Moreover, we derive from \( H_n' = 2nH_{n-1} \) that \( \tilde{H}_n' + x\tilde{H}_n = \sqrt{2n} \tilde{H}_{n-1} \). Hence, by introducing a new operator \( Du \triangleq \partial_x u + xu \), we obtain the following recursive relation:
\[
D \tilde{H}_n = \sqrt{2n} \tilde{H}_{n-1}. \tag{2.12}
\]
Using the operator $D$ we can define a Hilbert space similar to the Sobolev space $H^m(\mathbb{R})$, which will be important for our error analysis. To begin with, for any positive integer $N$, let $P_N$ be the space of polynomials of degree less or equal than $N$. We then define

$$H_N = e^{-x^2/2}P_N \triangleq \{e^{-x^2/2}p(x) : p(x) \in P_N\}. \quad (2.13)$$

Next, for each $N \in \mathbb{N}$ let us denote $\{x_i\}_{i=0}^N$ to be the roots of the polynomial $H_{N+1}(x)$; and define

$$w_j = \frac{\sqrt{2^N N!}}{(N+1)|H_N(x_j)|^2}, \quad 0 \leq j \leq N. \quad (2.14)$$

The pairs $\{(x_i, \omega_i)\}_{i=0}^N$ are known as the Hermite-Gauss quadrature points and weights, respectively (cf. e.g., [26]). Using the Hermite-Gauss quadrature points and weights we can then define a discrete inner product in $L^2(\mathbb{R})$ by

$$\langle f, g \rangle_N \triangleq \sum_{j=0}^N f(x_j)g(x_j)\hat{w}_j, \quad \forall f, g \in L^2(\mathbb{R}), \quad (2.15)$$

where $\hat{w}_j$’s are the normalized Hermite-Gauss quadrature weights defined by

$$\hat{w}_j \triangleq w_j e^{x_j^2} = \frac{\sqrt{\pi}}{(N+1)H_N^2(x_j)}, \quad 0 \leq j \leq N.$$

A direct consequence of the Hermite-Gauss quadrature is:

$$\langle u, v \rangle = \langle u, v \rangle_N, \quad \forall u \cdot v \in H_{2N+1}. \quad (2.16)$$

For any integer $m \geq 0$, let us define

$$H^m_D(\mathbb{R}) \triangleq \{u \mid D^m u \in L^2(\mathbb{R})\},$$

equipped with norm $\|D^m u\|$. Moreover, for any real number $r \geq 0$, the space $H^r_D(\mathbb{R})$ and its norm can be defined by the usual space interpolation.

In order to introduce our numerical approximation we need the following two operators: 1) the “orthogonal projection operator” $P_N : L^2(\mathbb{R}) \rightarrow \mathcal{H}_N$, defined by

$$\langle u - P_Nu, v \rangle = 0, \quad \forall u \in L^2(\mathbb{R}), \ v \in \mathcal{H}_N;$$

2) the “Hermite-Gauss interpolation operator” $I_N : C(\mathbb{R}) \rightarrow \mathcal{H}_N$, defined by

$$I_N f \in \mathcal{H}_N, \text{ such that } I_N f(x_j) = f(x_j), \quad j = 0, 1, \ldots, N. \quad (2.17)$$

The following basic approximation results for $P_N$ can be found in, e.g., [16]:

**Lemma 2.2** For any $u \in H^r_D(\mathbb{R})$, $N \in \mathbb{N}$, and $0 \leq s \leq r$, there exits a constant $C > 0$, independent of $N$ and function $u$, such that

$$\|D^s(u - P_N u)\| \leq CN^{1-s} \|D^r u\|, \quad (2.18)$$

and

$$\|\partial_x(u - P_N u)\| \leq CN^{1} \|D^r u\|. \quad (2.19)$$
To end this section, let us give a backward discrete Gronwall inequality, which will be important in the proof of our main theorem. A similar, but simpler version of such an inequality can be found in [22].

**Lemma 2.3** Suppose that the sequence \( \{\xi_k, r_k, \eta_k\}_{k \geq 0} \) satisfies that \( \xi_k, r_k, \eta_k \geq 0 \), and that for some constants \( \alpha < 1 \) and \( \beta > -1 \),

\[
(1 - \alpha)\xi_k + \eta_k \leq (1 + \beta)\xi_{k+1} + r_k, \quad 0 \leq k \leq M - 1. \tag{2.20}
\]

Then for all \( 0 \leq k \leq M \), it holds that

\[
\xi_k + (1 + \beta)^{M-k-1} \sum_{j=k}^{M-1} \eta_j \leq \left(\frac{1 + \beta}{1 - \alpha}\right)^{M-k} \xi_M + (1 + \beta)^{M-k-1} \sum_{j=k}^{M-1} r_j. \tag{2.21}
\]

The proof of this lemma is based on a straightforward backward induction, and we omit it here.

## 3 Numerical Schemes

In this section we present our numerical schemes, both for the PDE (2.6) and for the forward SDE (2.5). Although the scheme for the general quasilinear PDE (2.6) can be treated the same way with more complicated notation, to simplify presentation we shall assume that the coefficients \( b \) and \( g \) are both independent of variable \( z \). We note that under such a simplification the PDE (2.6) becomes

\[
\begin{aligned}
\begin{cases}
  u_t + \partial_x \left( \frac{1}{2} \sigma^2(t, x, u)u_x \right) + \hat{b}(t, x, u, u_x)u_x + g(t, x, u) = 0, \\
  u(T, x) = \phi(x), \quad x \in \mathbb{R},
\end{cases}
\end{aligned}
\tag{3.1}
\]

where \( \hat{b}(t, x, y, z) = b(t, x, y) - \sigma(t, x, y)(\sigma_x(t, x, y) + \sigma_y(t, x, y)z) \), with a slight abuse of notations. We shall design a numerical scheme for (3.1) based on a semi-implicit discretization in time and Hermite-collocation method in space; and discuss three types of scheme for the corresponding forward SDE (2.5). The error analysis for the combined scheme will be carried out in the next two sections.

### 3.1 Numerical Scheme for the PDE (3.1)

For any \( M \in \mathbb{N} \), let \( h = \frac{T}{M} \) be the length of time step and set \( t_k = kh, \ k = 0, 1, \ldots, M \).

We begin with a first-order semi-implicit time discretization: Let \( u^M(x) = \phi(x) \). For \( k = M - 1, \ldots, 0 \), we compute \( u^k \) by solving the the following discretized version of (3.1):

\[
\frac{u^{k+1} - u^k}{h} + \frac{1}{2} \partial_x \left( \sigma^2(t_k, x, u^{k+1})u_x^k \right) + \hat{b}(t_k, x, u^{k+1}, u_x^{k+1})u_x^{k+1} + g(t_k, x, u^{k+1}) = 0. \tag{3.2}
\]

In other words, at each time step, the problem is reduced to solving the following elliptic equation:

\[
u^k - \frac{h}{2} \partial_x \left( \sigma^2(t_k, x, u^{k+1})u_x^k \right) = f_h^{k+1}(x), \quad \lim_{x \to \pm\infty} u^k(x) = 0 \tag{3.3}\]

where

\[
f_h^{k+1}(x) = u^{k+1} + h \left\{ \hat{b}(t_k, x, u^{k+1}, u_x^{k+1})u_x^{k+1} + g(t_k, x, u^{k+1}) \right\}.
\]
Since the problem is set on the whole domain, it is natural to consider a *Hermite-spectral method*. To be more precise, let \( \{x_j\}_{j=0}^N \) be the Hermite-Gauss quadrature points and \( \mathcal{H}_N \) be defined in (2.13). A Hermite-collocation method for (3.3) is to find \( u_N^k \in \mathcal{H}_N \) such that

\[
u_N^k(x_j) - \frac{h}{2} \partial_x \left( \sigma^2(t_j, x, u_N^{k+1}) \partial_x u_N^k \right)(x_j) = f_{h}^{k+1}(x_j), \quad j = 0, 1, \cdots, N. \tag{3.4}
\]

Using the identity (2.16) and integration by parts, we can rewrite (3.4) in the following variational formulation: Find \( u_N^k \in \mathcal{H}_N \) such that

\[
\langle u_N^k, v_N \rangle_N + \frac{h}{2} \langle \sigma^2(t_j, x, u_N^{k+1}) \partial_x u_N^k, \partial_x v_N \rangle_N = \langle f_{h}^k, v_N \rangle_N, \quad \forall v_N \in \mathcal{H}_N, \tag{3.5}
\]

where \( f_{h}^k, f_{h}^j_N(x_j) = f^k(x_j), \quad j = 0, 1, \cdots, N, \)

We remark here that while the first order (time discretization) scheme is simple to use and analyze, in practice one often prefers higher-order schemes. These schemes could be easily constructed and often implemented in a straightforward manner, but their error analysis becomes rather tedious. In fact, as we will see in the next sections, the error analysis for the first-order semi-implicit scheme presented in (3.5) is already unpleasantly lengthy. As an example, we present the following second-order BDF Adam-Bashforth scheme which will be used in our numerical experiments without theoretical error analysis.

The second order BDF Adam-Bashforth Scheme.

Let \( u_N^M(x) = \varphi(x) \) and let \( u_N^{M-1}(x) \) be computed from the equation (3.2). For \( k = M - 2, \ldots, 1, 0 \), we compute \( u_N^k \) from:

\[
\begin{aligned}
-3u_N^{k+2} + 4u_N^{k+1} - u_N^k &= h^2 \frac{\sigma^2}{2} \partial_x^2 u_N^k + \partial_x (\frac{\sigma^2}{2} \partial_x u_N^k) + (2\partial_x u_N^{k+1} - \partial_x u_N^{k+2}) \hat{b} + g = 0,
\end{aligned}
\]

where

\[
\begin{aligned}
\sigma^2 &\triangleq \sigma^2(t_k, x, 2u_N^{k+1} - u_N^{k+2}), \quad g \triangleq g(t_k, x, 2u_N^{k+1} - u_N^{k+2}), \\
\hat{b} &\triangleq \hat{b}(t_k, x, 2u_N^{k+1} - u_N^{k+2}, \partial_x u_N^{k+1} - \partial_x u_N^{k+2}).
\end{aligned}
\]

We note that this scheme still leads to an elliptic equation of the form (3.3) for \( u_i^k \) at each time step. In particular, this second-order scheme should be used together with the \( \frac{1}{2} \)th-order SDE scheme presented below.

### 3.2 Numerical schemes for the SDE (2.5)

We now turn our attention to the discretization of the forward SDE (2.5). We begin with the simplest one, known as the "forward Euler-scheme". Assuming that for each \( N \in \mathbb{N} \) we have obtained the approximating solution to the PDE (3.1), denoted by \( u_N^k, k = 0, 1, \cdots, M, \) at each time \( t_k = kh \) we define recursively the approximate solution to the SDE (2.5) by:

\[
\begin{aligned}
X_0^N &= x, \\
X_{k+1}^N &= X_k^N + b(t_k, X_k^N, u_N^k(X_k^N))h + \sigma(t_k, X_k^N, u_N^k(X_k^N))\Delta_k W, \quad k = 0, 1, \cdots, M.
\end{aligned}
\]

where \( \Delta_k W \triangleq W(t_k + h) - W(t_k) \). For notational simplicity in what follows we shall simply denote \( X_k = X_k^N \), when context is clear.
It is well-understood that the Euler scheme is easy to implement and has the minimum requirements on the coefficients. Furthermore, if we denote the true solution to the PDE (3.1) by \( u \), denote \( u^k(x) = \hat{u}(t_k, x) \), and define an intermediate approximation \( \hat{X} \) by

\[
\begin{align*}
\hat{X}_0 &= x, \\
\hat{X}_{k+1} &= \hat{X}_k + b(t_k, \hat{X}_k, u^k(\hat{X}_k))h + (t_k, \hat{X}_k, u^k(\hat{X}_k))\Delta_k W, \\
& \quad + \frac{1}{2}\sigma(t_k, \hat{X}_k, u^k(\hat{X}_k))\{\sigma_x(t_k, \hat{X}_k, u^k(\hat{X}_k))\Delta_k W + \sigma_u(t_k, \hat{X}_k, u^k(\hat{X}_k))\partial_x u^k(\hat{X}_k)\}
\end{align*}
\]

(3.8)

then by the Fundamental Convergence Theorem (cf. e.g., [18]), the mean-square error of the Euler scheme is \( E|X(t_k) - \hat{X}_k|^2 \sim h \), where \( X \) is the true solution of (2.5) (namely, the rate of converges is \( \frac{1}{2} \)).

In order to obtain a higher order rate of convergence, one has to use a higher order scheme for the forward SDE. However, we should note that while in theory arbitrarily high-order schemes for (2.5) can be constructed with Taylor-Itô type expansions (see, e.g., [18]), the complexity of these schemes increases drastically. Consequently it often becomes too “expensive” in computational terms to implement. We will consider the following higher order schemes in our numerical experiments to test the numerical accuracy, and to compare our method with existing results.

A. Milstein Scheme (cf. [18], [24]).

This is a well-known scheme with first-order convergence rate. The recursive relation, adapted to our case, is given as follows.

\[
\begin{align*}
X_0 &= x, \\
X_{k+1} &= X_k + b(t_k, X_k, u^k_N(X_k))h + \frac{1}{2}\sigma(t_k, X_k, u^k_N(X_k))\{\sigma_x(t_k, X_k, u^k_N(X_k))\Delta_k W + \sigma_u(t_k, X_k, u^k_N(X_k))\partial_x u^k_N(X_k)\}
\end{align*}
\]

(3.9)

where \( \Delta_k W \overset{\Delta}{=} (W(t_{k+1}) - W(t_k))^2 \).

The next two schemes requires higher order regularity of the coefficients. We shall assume that all such requirements, whenever needed, are fulfilled without further specification.

B. Platen-Wagner Scheme (cf. [18]).

Using the idea of Taylor-Itô expansion up to order \( \frac{3}{2} \), and assuming that the coefficients are actually twice continuously differentiable, Platen and Wagner proposed the following scheme: for \( k = 0, 1, \ldots, M - 1 \),

\[
\begin{align*}
X_0 &= x, \\
X_{k+1} &= X_k + bh + \sigma \Delta_k W + \frac{1}{2} \sigma \sigma' \{\Delta_k^2 W - h\} + b' \sigma \Delta_k Z + \frac{1}{2} \big(b' h + \frac{1}{2} \sigma^2 \sigma'' \big) h^2 \\
& \quad + \big(b \sigma' + \frac{1}{2} \sigma^2 \sigma'' \big) \{ h \Delta_k W - \Delta_k Z \} + \frac{1}{2} \sigma \big( \sigma \sigma'' + (\sigma')^2 \big) \big( \frac{1}{2} \Delta_k^2 W - h \big) \Delta_k W
\end{align*}
\]

(3.10)

where \( \Delta_k^2 W \overset{\Delta}{=} (W(t_{k+1}) - W(t_k))^2 \), \( \Delta_k Z \overset{\Delta}{=} \int_{t_k}^{t_{k+1}} [W(t) - W(t_k)] dt \), \( b = b(t_k, X_k, u^k_N(X_k)) \), \( \sigma = \sigma(t_k, X_k, u^k_N(X_k)) \), and for \( \varphi = b, \sigma \),

\[
\varphi' = \varphi + x + \partial_x u^k_N, \quad \varphi'' = \varphi + 2 \partial_{xx} u^k_N + \varphi x + \partial_u (\partial_x u^k_N)^2 + \varphi u \partial_{x}^2 u^k_N.
\]

(3.11)

Here the partial derivatives of \( b \) and \( \sigma \) should be evaluated at \( (t_k, X_k, u^k_N(X_k)) \), and the approximate functions \( u^k_N, \frac{\partial u^k_N}{\partial x} \) and \( \frac{\partial^2 u^k_N}{\partial x^2} \) should be evaluated at point \( (X_k) \).
We should note that \( \Delta_k z = \int_{t_k}^{t_{k+1}} W(t) - W(t_k) dt \sim N(0, \frac{1}{2} h^2) \), and the covariance of \( \Delta_k W \) and \( \Delta_k W \) is \( \mathbb{E}(\Delta_k W \Delta_k Z) = \frac{1}{2} h^2 \). The presence of the \( \Delta_k Z \) obviously complicates the analysis of the scheme. This will be even more so when the order of expansion increases. Treating these terms effectively will become more important.

**C. The Stochastic Adam-Bashforth (SAB) Scheme (cf. [14])**

It is easily seen from (3.11) that Platen-Wagner’s \( \frac{1}{4} \)-order scheme requires \( b \) and \( \sigma \) are both twice differentiable. Recently, Brian D. Ewald and Roger Temam [14] constructed a stochastic version of the Adam-Bashforth scheme which does not involve the second derivatives of coefficient function \( b \) but still achieves the \( \frac{1}{4} \)-order convergence rate. This Stochastic Adam-Bashforth (SAB) scheme, adopted for a one-dimensional diffusion of the form:

\[
dX_t = b(t, X_t) dt + \sigma(t, X_t) dW(t),
\]

can be rewritten as follows:

\[
X_{k+2} = X_{k+1} + \frac{\Delta_t}{2} [3b(t_{k+1}, X_{k+1}) - b(t_k, X_k)] - \frac{3}{2} \Delta_t A_k(t_k, X_k) + B_k(t_k, X_k),
\]

(3.12)

where

\[
A_k(t, x) = [\sigma b_e](t, x) \Delta_k W,
\]

(3.13)

\[
B_k(t, x) = \sigma(t, x) \Delta_k W + [\sigma_t + b \sigma_x + \frac{1}{2} \sigma^2 \sigma_{xx}](t, x) I_{(0,1)} + [\sigma b_e](t, x) I_{(1,0)},
\]

\[
+ [\sigma \sigma_x](t, x) I_{(1,1)} + [\sigma ((\sigma_x)^2 + \sigma \sigma_{xx})](t, x) I_{(1,1,1)};
\]

(3.12)

with \( \Delta_k W = W_{t_{k+1}} - W_{t_k} \) and \( h = t_{k+1} - t_k \) as before, and the random coefficients \( I \)’s are defined by

\[
I_{(0,1)} = 2h \Delta_k W - \int_{t_k}^{t_{k+1}} [W(s) - W(t_{k+1})] ds
\]

\[
I_{(1,0)} = h \Delta_k W + \int_{t_k}^{t_{k+1}} [W(s) - W(t_{k+1})] ds
\]

(3.14)

and \( I_{(1,1)} \) are the iterated Itô integrals

\[
I_{(1,1)} = \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} dW(s) dW(t) - \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} dW(s) dW(t)
\]

(3.15)

To calculate the iterated Itô integrals \( I_{(1,1,1)} \), the useful formulae of Itô [17] are often been employed. More precisely, let us define a scaled Hermite polynomial \( h_n \) by

\[
H_n(x) = 2^{n/2} h_n(\sqrt{2} x), \quad n = 0, 1, 2, \ldots
\]

then we have (cf. [26])

\[
h_n(x) = (-1)^n e^{\frac{1}{2} x^2} \frac{d^n}{dx^n}(e^{-\frac{1}{2} x^2}); \quad n = 0, 1, 2, \ldots
\]
Thus the coefficients simplify the presentation, we shall only consider the following

\begin{equation}
\frac{n!}{a^n} \int_a^b f(t_1)g(t_2)\cdots g(t_n)dt_1\cdots dt_n = \|g\|^n h_n(\frac{\theta}{\|g\|}),
\end{equation}

where

\[\|g\| = \|g\|_{L^2([a,b])}\text{ and } \theta = \int_a^b g(t)dt.\]

Thus the coefficients \(I_{(1,1)}\) and \(I_{(1,1,1)}\) in (3.15) can be calculated explicitly as

\begin{align*}
I_{(1,1)} &= \frac{1}{2} \left( \Delta_{k+1}^2 W + 2\Delta_{k+1} W \Delta_k W - h \right) \\
I_{(1,1,1)} &= \frac{1}{6} \left( (\Delta_{k+1} W + \Delta_k W)^3 - (\Delta_k W)^3 - 6h\Delta_{k+1} W - 3h\Delta_k W \right). \tag{3.17}
\end{align*}

Using the above relations, we can rewrite the \(\frac{3}{2}\)-order SAB scheme for the one dimensional forward SDE (2.5) as

\begin{align*}
X_0 &= x, \\
X_{k+2} &= X_{k+1} + \frac{h}{2} \left[ 3b(t_{k+1}, X_{k+1}, u_N^{k+1}(X_{k+1})) - b \right], \\
&\quad -\frac{3}{2} \sigma b' \Delta_k W + \sigma \Delta_{k+1} W + \frac{1}{2} \sigma \sigma' \left( (\Delta_{k+1} W)^2 + 2(\Delta_{k+1} W)(\Delta_k W) - h \right) \\
&\quad + \left( b\sigma' + \frac{1}{2} \sigma^2 \sigma'' \right) \left( 2h\Delta_{k+1} W - \Delta_{k+1} Z \right) + b' \left( h\Delta_k W + \Delta_{k+1} Z \right) \\
&\quad + \frac{1}{6} \sigma \left( \sigma^2 + (\sigma')^2 \right) \left( (\Delta_{k+1} W + \Delta_k W)^3 - (\Delta_k W)^3 - 6h\Delta_{k+1} W - 3h\Delta_k W \right)
\end{align*}

where \(k = 0, 1, \ldots, M - 1\) and similarly,

\begin{align*}
b &= b(t_k, X_k, u_N^k(X_k)) \\
\sigma &= \sigma(t_k, X_k, u_N^k(X_k)) \\
b' &= \frac{\partial b}{\partial x} + \frac{\partial b}{\partial u} \frac{du_N^k}{dx} \\
\sigma' &= \frac{\partial \sigma}{\partial x} + \frac{\partial \sigma}{\partial u} \frac{du_N^k}{dx} \\
\sigma'' &= \frac{\partial^2 \sigma}{\partial x^2} + 2 \frac{\partial \sigma}{\partial x \partial u} \frac{du_N^k}{dx} + \frac{\partial^2 \sigma}{\partial u^2} \left( \frac{du_N^k}{dx} \right)^2 + \frac{\partial \sigma}{\partial u} \frac{d^2 u_N^k}{dx^2}
\end{align*}

where the partial derivatives of \(b\) and \(\sigma\) are evaluated at \((t_k, X_k, u_N^k(X_k))\), and the approximate functions \(u_N^k, \frac{du_N^k}{dx}\) and \(\frac{d^2 u_N^k}{dx^2}\) are evaluated at point \((X_k)\).

Finally, the components \(Y\) and \(Z\) of the solution to (2.1) are approximated as

\begin{equation}
Y_k = u_N^k(X_k), \quad Z_k = \sigma(t_k, X_k, Y_k) \frac{\partial u_N^k}{\partial x}(X_k), \tag{3.18}
\end{equation}

\section{Error analysis for the PDE approximation}

In this section we carry out an error analysis for the approximation of PDE (3.1). In order to simplify the presentation, we shall only consider the following Hermite-Galerkin scheme for (3.2):
Find \( u_N^k \in H_N \) such that
\[
\langle \frac{u_N^{k+1} - u_N^k}{h}, v_N \rangle + \frac{1}{2} \langle \partial_x \left( \sigma^2(t_k, x, u_N^{k+1})(u_N^k)_x \right), v_N \rangle \\
+ \langle \hat{b}(t_k, x, u_N^{k+1}(u_N^k)_x \rangle \partial_x (u_N^{k+1}) + g(t_k, x, u_N^{k+1}), v_N \rangle = 0, \quad \forall v_N \in H_N.
\]

We note that the only difference between this Hermite-Galerkin scheme and the Hermite-collocation scheme (3.5) is that in the latter the continuous inner product is replaced by the discrete inner product. It is well-known that for \( N \) sufficiently large, the difference between the two approaches are negligible.

Hereafter, we shall use “\( A \lesssim B \)” to mean that there exists a constant \( C > 0 \) independent of \( N \) or \( h \) such that \( A \leq CB \).

Our main result in this section is the following theorem.

**Theorem 4.1** Assume (A1)—(A3). Let \( u \) and \( u_N^k \) be the solutions of (3.1) and (4.1), respectively. Assume further that \( u, \phi \in L^\infty(0, T; H^2_D(\mathbb{R})) \) and \( \partial_t u \in L^2(0, T; H^3_D(\mathbb{R})) \) with some \( m > 1 \). Then, there exists \( h_0 > 0 \) such that for \( h \leq h_0 \), the scheme (3.2) is stable, and the following error estimates hold: for each \( k = 0, 1, \ldots, M \),
\[
\|u(t_k, \cdot) - u_N^k\| + \left( h \sum_{j=k}^M \|\partial_x (u(t_k, \cdot) - u_N^j)\|^2 \right)^{\frac{1}{2}} \lesssim N^{-\alpha}
\]
\[
\lesssim N^{-\alpha} \left( \|u\|_{L^\infty(0,T;H^m_D)} + \|u_t\|_{L^2(0,T;H^m_D)} + \|\tilde{\phi}\|_{L^\infty(0,T;H^m_D)} \right) + h.
\]

**Proof.** Let \( E^k \), \( k = M - 1, \ldots, 1, 0 \), be the truncation error defined by
\[
E^k(\cdot) \triangleq \frac{1}{h} \left( u(t_{k+1}, \cdot) - u(t_k, \cdot) \right) + \langle b(t_k, \cdot, u(t_{k+1}, \cdot)), \partial_x u(t_{k+1}, \cdot) \rangle \partial_x u(t_k, \cdot) + g(t_k, \cdot, u(t_k, \cdot)),
\]
where \( u(t_k, \cdot) \) is the exact solution to equation (3.1). In what follows we denote \( \hat{e}_N^k = P_N u(t_k, \cdot) - u_N^k \), \( \hat{e}_N = u(t, \cdot) - u_N^k = \hat{e}_N^k + \hat{e}_N^k \).

Next, multiplying \( e_N \in H_N \) on both sides of (4.2), using integration by parts, and then subtracting (4.1), we obtain that
\[
\langle E^k, v_N \rangle = \frac{1}{h} \left( e_N^{k+1} - e_N^k, v_N \right) - \langle \frac{1}{2} \sigma^2(t_k, \cdot, u(t_{k+1}, \cdot)), \partial_x u(t_k, \cdot)v_N \rangle \\
+ \frac{1}{2} \langle \sigma^2(t_k, \cdot, u_N^{k+1}) \partial_x u_N^k, \partial_x v_N \rangle \\
+ \langle b(t_k, \cdot, u(t_{k+1}, \cdot)), \partial_x u(t_{k+1}, \cdot) \rangle \partial_x u(t_k, \cdot) - \langle b(t_k, \cdot, u_N^{k+1}), \partial_x u_N^k \rangle \partial_x u_N^k, v_N \rangle \\
+ \langle g(t_k, \cdot, u(t_{k+1}, \cdot)), g(t_k, \cdot, u_N^{k+1}) \rangle, v_N \rangle.
\]

Or equivalently,
\[
\langle E^k, v_N \rangle = \frac{1}{h} \left( e_N^{k+1} - e_N^k, v_N \right) - \langle \frac{1}{2} \sigma^2(t_k, \cdot, u_N^{k+1}) \partial_x e_N^k, \partial_x v_N \rangle \\
- \langle \partial_x b(t_k, \cdot), \frac{1}{2} \sigma^2(t_k, \cdot, u(t_{k+1}, \cdot)) \partial_x v_N \rangle - \langle \frac{1}{2} \sigma^2(t_k, \cdot, u_N^{k+1}) \partial_x v_N \rangle \\
+ \langle b(t_k, \cdot, u(t_{k+1}, \cdot)), \partial_x u(t_{k+1}, \cdot) \rangle \partial_x u(t_k, \cdot) - \langle b(t_k, \cdot, u_N^{k+1}), \partial_x u_N^k \rangle \partial_x u_N^k, v_N \rangle \\
+ \langle g(t_k, \cdot, u(t_{k+1}, \cdot)), g(t_k, \cdot, u_N^{k+1}) \rangle, v_N \rangle.
\]
Since \( e_N^k = e_N^k + \tilde{e}_N^k \), we have

\[
\frac{1}{h} (e_{N+1}^k - \tilde{e}_N^k) \quad v_N = - \frac{1}{2} \sigma^2 (t_k, \cdot, u_{N+1}^{k+1}) \partial_x e_N^k \partial_x v_N
\]

\[
= - \frac{1}{h} (e_{N+1}^k - \tilde{e}_N^k, v_N) + \frac{1}{2} \sigma^2 (t_k, \cdot, u_{N+1}^{k+1}) \partial_x e_N^k \partial_x v_N
\]

\[
+ \langle \partial_x u(t_k, \cdot), \frac{1}{2} \sigma^2 (t_k, \cdot, u_{N+1}^{k+1}) \partial_x v_N - \frac{1}{2} \sigma^2 (t_k, \cdot, u_{N+1}^{k+1}) \partial_x v_N \rangle
\]

\[
- \langle b(t_k, \cdot, u_{N+1}^{k+1}), \partial_x u(t_{k+1}, \cdot) \rangle \partial_x v_N - b(t_k, \cdot, u_{N+1}^{k+1}, \partial_x u_{N+1}^{k+1}) \partial_x v_{N+1}
\]

\[
- \langle g(t_k, \cdot, u_{N+1}^{k+1}), - g(t_k, \cdot, u_{N+1}^{k+1}) \rangle + \langle E^k (\cdot), v_N \rangle.
\]

We now choose \( v_N = - h \tilde{e}_N^k \), and denote

\[
I_1 = - \langle g(t_k, \cdot, u_{N+1}^{k+1}), - g(t_k, \cdot, u_{N+1}^{k+1}) \rangle = g(t_k, \cdot, u_{N+1}^{k+1}), v_N \rangle
\]

Similarly, we have

\[
I_4 \leq C_4 (\partial_x e_N^k | \partial_x v_N) < \frac{C_4 h}{2 \varepsilon} \| e_N^k \|^2 + \frac{C_4 \varepsilon}{2h} \| \partial_x v_N \|^2.
\]

Second, note that under assumptions (A1) and (A2), the solution to (3.1) must have bounded first order derivative \( \partial_x u \) (see, e.g., [9]). Thus, one has

\[
I_2 = - \langle b(t_k, \cdot, u_{N+1}^{k+1}), \partial_x u(t_{k+1}, \cdot) \rangle \partial_x u(t_{k+1}, \cdot) - b(t_k, \cdot, u_{N+1}^{k+1}, \partial_x u_{N+1}^{k+1}) \partial_x u(t_{k+1}, \cdot) \rangle
\]

\[
- \langle b(t_k, \cdot, u_{N+1}^{k+1}, \partial_x u_{N+1}^{k+1}) \rangle \partial_x v_N \rangle
\]

\[
\leq K \| \partial_x u \|_{\infty} (|e_{N+1}^{k+1}| + |\partial_x e_{N+1}^{k+1}|, |v_N|) + |b|_{\infty} (|\partial_x e_{N+1}^{k+1}|, |v_N|)
\]

\[
\leq K \| \partial_x u \|_{\infty} (|e_{N+1}^{k+1}| + |\partial_x e_{N+1}^{k+1}| + |e_{N+1}^{k+1}|, |v_N|) + |b|_{\infty} (|\partial_x e_{N+1}^{k+1}|, |v_N|)
\]

\[
\leq \frac{C_2 h}{\varepsilon} (|e_{N+1}^{k+1}|^2 + |\partial_x e_{N+1}^{k+1}|^2) + \frac{C_2 \varepsilon}{h} \| v_N \|^2 + C_2 \varepsilon h \| e_{N+1}^{k} \|^2 + \frac{C_2}{\varepsilon h} \| v_N \|.\]

where \( C_2 = 4 \max \{ K \| \partial_x u \|_{\infty}, 1 \| \| \} \) and \( \varepsilon_1 \) is to be determined later. Similarly, we have

\[
I_3 \leq L \| \partial_x u \|_{\infty} (|e_{N+1}^{k+1}|, \partial_x v_N) \leq \frac{C_3 h}{2 \varepsilon} \| e_{N+1}^{k+1} \|^2 + \frac{C_3 \varepsilon}{2h} \| \partial_x v_N \|^2
\]

\[
\leq \frac{C_3 h}{\varepsilon} \| e_{N+1}^{k+1} \|^2 + \frac{C_3 \varepsilon}{2h} \| \partial_x v_N \|^2.
\]
where \( C_4 = L \| \partial_x u \|_\infty \). Finally, if we denote \( \tilde{E}^k_N = -\frac{1}{2} (\tilde{e}^k_N - \tilde{e}^k_N) \), then we have

\[
I_5 \leq \frac{h}{2\epsilon} \| \tilde{E}^k_N \|^2 + \frac{\epsilon}{2h} \| v_N \|^2, \quad \text{and} \quad I_6 \leq \frac{h}{2\epsilon} \| E^k(\cdot) \|^2 + \frac{\epsilon}{2h} \| v_N \|^2.
\]  

(4.8)

To obtain the desired estimate, let us now look at the left hand side of (4.3) with \( v_N = -h \tilde{e}^k_N \). Note that

\[
\frac{1}{h} (\tilde{e}^k_N - \tilde{e}^k_N, -h \tilde{e}^k_N) = \frac{1}{2} \| \tilde{e}^k_N \|^2 - \| \tilde{e}^k_N \|^2 + \| \tilde{e}^k_N - \tilde{e}^k_N \|^2),
\]

(4.9)

\[-\langle \partial_x \tilde{e}^k_N, \frac{1}{2} \sigma^2(t, \cdot, v^{N+1}_N) \partial_x v_N \rangle \geq ch \| \partial_x \tilde{e}^k_N \|^2,
\]

(4.10)

where \( c > 0 \) is a lower bound of \( \sigma^2 \), with a slight abuse of notation (compare to the constant \( c > 0 \) in (A3)). Thanks again the Schwarz inequality, and denoting \( C = 5 \max_{1 \leq i \leq 6} \{ C_i \} + 1 \), we derive from (4.4)—(4.9) that

\[
\| \tilde{e}^k_N \|^2 + ch \| \partial_x \tilde{e}^k_N \|^2 \leq \left( \frac{Ch}{\epsilon} + \frac{1}{2} \right) \| \tilde{e}^k_N \|^2 + C \epsilon h \| \partial_x \tilde{e}^k_N \|^2 + \left( \frac{Ch}{\epsilon} + 1 \right) \| \tilde{e}^k_N \|^2 + C \epsilon + ch + \frac{h}{\epsilon} \| E^k(\cdot) \|^2 + \frac{h}{\epsilon} \| \tilde{E}^k_N \|^2.
\]

(4.11)

Note that the truncation error \( E^k(\cdot) \) is of first order, i.e.,

\[
\| E^k(\cdot) \|^2 \leq C h^2.
\]

(4.12)

We see that (4.11) can now be rewritten as

\[
\left( \frac{1}{2} - h \left( \frac{C \epsilon + \epsilon + \frac{C}{\epsilon_1} \right) \right) \| \tilde{e}^k_N \|^2 + \left( c - C \epsilon \right) h \| \partial_x \tilde{e}^k_N \|^2
\]

\[
\leq \left( \frac{Ch}{\epsilon} + \frac{1}{2} \right) \| \tilde{e}^k_N \|^2 + C \epsilon h \| \partial_x \tilde{e}^k_N \|^2
\]

(4.13)

\[+ \frac{h}{\epsilon} \left( C \| \tilde{e}^k_N \|^2 + \| \tilde{e}^k_N \|^2 + C \| \partial_x \tilde{e}^k_N \|^2 + \| \tilde{E}^k_N \|^2 + C h^2 \right).
\]

Now, let us fix \( \epsilon \) small enough so that \( c - C \epsilon > 0 \) (for example, \( \epsilon = \frac{2\epsilon_1}{2\epsilon_1 + 1} \) so that \( c - C \epsilon = 1/2 \)). Multiplying both sides of (4.13) by 2, choosing \( \epsilon_1 = \frac{c - C \epsilon}{2c} \), setting \( C_9 \triangleq \max \{ C + 1, C_8 \}, \ C_\epsilon \triangleq 2(C \epsilon + \frac{C}{\epsilon_1}) \), and finally assuming that the time step \( h \) satisfies the conditions

\[
hC_\epsilon < \frac{1}{2}, \quad \frac{2Ch}{\epsilon} \leq 1,
\]

(4.14)

the inequality (4.13) then becomes

\[
(1 - hC_\epsilon) \| \tilde{e}^k_N \|^2 + 2(c - C \epsilon) h \| \partial_x \tilde{e}^k_N \|^2
\]

(4.15)

\[\leq \left( \frac{2Ch}{\epsilon} + 1 \right) \| \tilde{e}^k_N \|^2 + (c - C \epsilon) h \| \partial_x \tilde{e}^k_N \|^2 + \frac{hC_9}{\epsilon} \left( \| \tilde{e}^k_N \|^2 + \| \tilde{e}^k_N \|^2 + \| \partial_x \tilde{e}^k_N \|^2 + \| \tilde{E}^k_N \|^2 + h^2 \right).
\]

We now apply the backward Gronwall inequality (Lemma 2.3) with \( \alpha = hC_\epsilon, \ \beta = 2Ch/\epsilon, \ \xi_k = \| \tilde{e}^k_N \|^2, \ \eta_k = 2(c - C \epsilon) \| \partial_x \tilde{e}^k_N \|^2, \) and

\[
r_k = (c - C \epsilon) h \| \partial_x \tilde{e}^k_N \|^2 + \frac{hC_9}{\epsilon} \left( \| \tilde{e}^k_N \|^2 + \| \tilde{e}^k_N \|^2 + \| \partial_x \tilde{e}^k_N \|^2 + \| \tilde{E}^k_N \|^2 + h^2 \right), \quad k = 1, 2, \cdots.
\]

14
Comparing (4.15) with (2.20), and noting the time step conditions (4.14), we derive from (2.21) that

\[
\|e_N^k\|^2 + 2(c - C\varepsilon)h\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)}\sum_{j=k}^{M-1}\|\partial_x e_N^j\|^2 \\
\leq \left(\frac{2Ch + 1}{1 - hC\varepsilon}\right)^{(M-k)}\|e_N^k\|^2 + (c - C\varepsilon)h\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)}\sum_{j=k}^{M-1}\|\partial_x e_N^{j+1}\|^2 \\
+ \frac{C_0h}{\varepsilon}\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)}\sum_{j=k}^{M-1}\left(\|e_N^{j+1}\|^2 + \|e_N^j\|^2 + \|\partial_x e_N^{j+1}\|^2 + \|\tilde{E}_N^k\|^2 + h^2\right).
\]

By a simple cancellation we see that the above inequality can be easily reduced to

\[
\|e_N^k\|^2 + (c - C\varepsilon)h\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)}\sum_{j=k}^{M-1}\|\partial_x e_N^j\|^2 \\
\leq \left(\frac{2Ch + 1}{1 - hC\varepsilon}\right)^{(M-k)}\|e_N^k\|^2 + (c - C\varepsilon)h\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)}\|\partial_x e_N^M\|^2 \\
+ \frac{C_0h}{\varepsilon}\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)}\sum_{j=k}^{M-1}\left(\|e_N^{j+1}\|^2 + \|e_N^j\|^2 + \|\partial_x e_N^{j+1}\|^2 + \|\tilde{E}_N^k\|^2 + h^2\right).
\]

Since as \( M \to \infty \) (recall that \( M \) is the number of time steps), one has

\[
\left(\frac{2Ch + 1}{\varepsilon}\right)^{M-k} \leq \left(\frac{2Ch}{\varepsilon} + 1\right)^{M} = \left(\frac{2CT}{\varepsilon M} + 1\right)^{M} \to e^{\frac{2CT}{M}},
\]

\[
\left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)} \geq \left(\frac{1}{1 + hC\varepsilon}\right)^{M} = \left(\frac{1}{1 + \frac{2Ch}{M}}\right)^{M} \to e^{-TC\varepsilon},
\]

we see that the sequence \( \left(\frac{2Ch + 1}{\varepsilon}\right)^{M-k} \) is bounded from above by some constant \( \alpha_1 > 0 \), and the sequence \( \left(\frac{2Ch + 1}{1 + hC\varepsilon}\right)^{(M-k-1)} \) is bounded from below by \( \alpha_2 > 0 \). We note that both constants \( \alpha_1 \) and \( \alpha_2 \) depend only on the constants \( c \), \( C \) and \( T \) for fixed \( \varepsilon \), and the bounds are all uniformly in \( k \). We now take a closer look at all the errors on the right hand side above. First, by virtue of Lemma 2.2

\[
\|\tilde{E}_N^k\|^2 \leq C_T N^{-m} \|D^m u\|^2, \quad \|\partial_x \tilde{E}_N^k\|^2 \leq C_T N^{-m} \|D^m u\|^2.
\]

It remains to estimate \( \|\tilde{E}_N^k\| \). Let \( \tilde{e}_N(t, \cdot) = u(t, \cdot) - P_N u(t, \cdot) \). Then,

\[
\tilde{E}_N^k = -\frac{1}{h} (\tilde{e}_N^{k+1} - \tilde{e}_N^k) = -\frac{1}{h} \int_{t_k}^{t_{k+1}} \partial_t \tilde{e}_N(t, \cdot) dt.
\]

Applying the Schwarz inequality and (2.18) again we obtain that

\[
\|\tilde{E}_N^k\|^2 \leq \frac{1}{h} \int_{t_k}^{t_{k+1}} \|\partial_t \tilde{e}_N(t, \cdot)\|^2 dt.
\]
for $0 \leq x \leq \frac{1}{2}$,

$$
\|\hat{e}_N^k\|^2 + (c - C\varepsilon)\alpha_2 h \sum_{j=k}^{M-1} \|\hat{e}_N^j\|^2 \leq 4C\varepsilon T \alpha_1 \|\hat{e}_N^M\|^2 + (c - C\varepsilon)\alpha_1 h \|\hat{e}_N^M\|^2 + \frac{C_T}{\varepsilon} (N^{1-m} \|u\|_{L^\infty(0,T,H^m_{TB})}^2 + h^2) + \frac{C_0}{\varepsilon} N^{-m} \|\hat{u}\|_{L^2(0,T,H^m_{TB})}^2
$$

and combining (4.19) with (4.17), and noting that $u_N^M = \varphi$ and $e_N^k = \hat{e}_N^k - \tilde{e}_N^k$, we obtain that

$$
\|\hat{e}_N^k\|^2 + (c - C\varepsilon)\alpha_2 h \sum_{j=k}^{M-1} \|\hat{e}_N^j\|^2 \leq 2\|\hat{e}_N^k\|^2 + 2(c - C\varepsilon)\alpha_2 h \sum_{j=k}^{M-1} \|\hat{e}_N^j\|^2 + 2\|\tilde{e}_N^k\|^2 + 2(c - C\varepsilon)\alpha_2 h \sum_{j=k}^{M-1} \|\hat{e}_N^j\|^2
$$

$$
\leq 4C\varepsilon T \alpha_1 \|\hat{e}_N^M\|^2 + (c - C\varepsilon)\alpha_1 h \|\hat{e}_N^M\|^2 + 2\|\tilde{e}_N^k\|^2 + 4(c - C\varepsilon)h \|\hat{e}_N^M\|^2 + \frac{C_T}{\varepsilon} (N^{1-m} \|u\|_{L^\infty(0,T,H^m_{TB})}^2 + h^2) + \frac{C_0}{\varepsilon} N^{-m} \|\hat{u}\|_{L^2(0,T,H^m_{TB})}^2
$$

$$
\lesssim N^{1-m} \left( \|u\|_{L^\infty(0,T,H^m_{TB})}^2 + \|\hat{u}\|_{L^2(0,T,H^m_{TB})}^2 + \|\varphi\|_{L^\infty(0,T,H^m_{TB})}^2 \right) + h^2.
$$

The conclusion of the theorem now follows easily.

## 5 The synthetic error analysis

In this section we present an error analysis for the synthesized numerical scheme. Our main result is the following.

**Theorem 5.1** Assume that (A1)-(A3) and the assumptions in theorem 4.1 hold. Let $(X, Y, Z)$ be the adapted solution to the FBSDE (2.1), and let $(X_k)_{k=0}^M$ be the solution to the Euler scheme (3.7). Define

$$
Y_k \triangleq u_N^k(X_k), \quad Z_k \triangleq \sigma(t_k, X_k, Y_k)\partial_x(u_N^k)(X_k), \quad k = 0, 1, \cdots, M,
$$

where $(u_N^k)_{k=0}^M$ is the solution to (4.1). Then, the following error estimate holds:

$$
\sup_{0 \leq k \leq M} \mathbb{E} \left[ (X(t_k) - X_k)^2 + (Y(t_k) - Y_k)^2 + (Z(t_k) - Z_k)^2 \right]^{1/2} \lesssim \sqrt{h} + N^{1-m} \sqrt{\Upsilon(u, \phi)}.
$$

where $\Upsilon(u, \phi) \triangleq \|u\|_{L^\infty(0,T,H^m_{TB})}^2 + \|\partial_t u\|_{L^2(0,T,H^m_{TB})}^2 + \|\varphi\|_{L^\infty(0,T,H^{m+1}_{TB})}^2$.

**Proof.** Recall the intermediate Euler scheme (3.8):

$$
\begin{cases}
\hat{X}_0 = x, \\
\hat{X}_{k+1} = \hat{X}_k + b(t_k, \hat{X}_k, w^k(\hat{X}_k))h + \sigma(t_k, \hat{X}_k, u^k(\hat{X}_k))\Delta_k W_k
\end{cases}
$$

where $w^k(x) = u(t_k, x)$, and $u(\cdot, \cdot)$ is the solution to the original PDE (3.1). Since the true solution is at least uniformly Lipschitz under (A1)–(A3), we can apply the fundamental convergence theorem for SDEs (cf. [18] or [23]) to conclude that

$$
\mathbb{E} |X(t_k) - \hat{X}_k|^2 \leq Ch \quad \forall \ k = 0, 1, \cdots, M.
$$
Thus it remains to evaluate $\mathbb{E}[\hat{X}_k - X_k]^2$. To this end, note that
\[
\hat{X}_{k+1} - X_{k+1} = \hat{X}_k - X_k + [b(t_k, \hat{X}_k, u(t_k, \hat{X}_k)) - b(t_k, X_k, u^k_N(X_k))]h + [\sigma(t_k, \hat{X}_k, u(t_k, \hat{X}_k)) - \sigma(t_k, X_k, u^k_N(X_k))]\Delta_k W. \tag{5.5}
\]
Since $\Delta_k W = W(t_{k+1}) - W(t_k)$ is independent of $\mathcal{F}_{t_k}$ with $\mathbb{E}[\Delta_k W] = 0$ and $\mathbb{E}[\Delta_k W]^2 = h$, one has $\mathbb{E}[(\Delta_k W)^2] = 0$ and $\mathbb{E}[(\Delta_k W)^2] = h\mathbb{E}[F]^2$ for all $F \in L^2(\mathcal{F}_{t_k}; \mathbb{R})$. Therefore, squaring both sides of (5.5), applying Cauchy-Schwarz inequality, and then using the Lipschitz assumption on the coefficients $b$ and $\sigma$, we obtain that
\[
\mathbb{E}[\hat{X}_{k+1} - X_{k+1}]^2 \leq (1 + h)\mathbb{E}[\hat{X}_k - X_k]^2 + (h^2 + h)\mathbb{E}[b(t_k, \hat{X}_k, u^k(\hat{X}_k)) - b(t_k, X_k, u^k_N(X_k))]^2
\]
\[+ h \mathbb{E}[\sigma(t_k, \hat{X}_k, u^k(\hat{X}_k)) - \sigma(t_k, X_k, u^k_N(X_k))]^2 \leq (1 + h + 2K^2(h^2 + h))\mathbb{E}[\hat{X}_k - X_k]^2 + 2K^2(h^2 + h)\mathbb{E}[u^k(\hat{X}_k) - u^k_N(X_k)]^2. \tag{5.6}
\]
Now, we write
\[
u^k(\hat{X}_k) - u^k_N(X_k) = [u^k(\hat{X}_k) - u^k(\hat{X}(t_k))] + [u^k(\hat{X}(t_k)) - u^k_N(X_k)] + [u^k_N(X(t_k)) - u^k_N(X_k)], \tag{5.7}
\]
and note that
\[
\mathbb{E}[u^k(\hat{X}_k) - u^k(\hat{X}(t_k))]^2 \leq K^2\mathbb{E}[\hat{X}_k - X(t_k)]^2 \leq CK^2 h, \tag{5.8}
\]
\[
\mathbb{E}[u^k_N(X(t_k)) - u^k_N(X_k)]^2 \leq K^2\mathbb{E}[X(t_k) - X_k]^2 \leq 2K^2\mathbb{E}[X(t_k) - \hat{X}_k]^2 + 2K^2\mathbb{E}[\hat{X}_k - X_k]^2 \leq 2CK^2 h + 2K^2\mathbb{E}[\hat{X}_k - X_k]^2, \tag{5.9}
\]
we need only to estimate $\mathbb{E}[u^k(\hat{X}(t_k)) - u^k_N(X(t_k))]^2$. To this end, let us first establish the following simple inequality:
\[
\|f\|_{L^\infty(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})} + \|f’\|_{L^2(\mathbb{R})}. \tag{5.10}
\]
Indeed, we first observe that $f \in H^1(\mathbb{R})$ implies that $f \in C(\mathbb{R})$. Then, for any unit interval $[n, n+1)$, $n \in \mathbb{Z}$, there must exist $x_n \in [n, n+1)$ such that $|f(x_n)| \leq \|f\|_{L^2(n,n+1)} \leq \|f\|_{L^2(\mathbb{R})}$. Hence, for any given $x \in [n, n+1)$, we have
\[
|f(x)| \leq |f(x) - f(x_n)| + |f(x_n)| = \int_{x_n}^x |f’(t)| dt + |f(x_n)| \leq \|f\|_{L^2(\mathbb{R})} + \|f’\|_{L^2(\mathbb{R})}.
\]
Thanks to (5.10), we obtain that
\[
\mathbb{E}[u^k(X(t_k)) - u^k_N(X(t_k))]^2 = \int_{-\infty}^\infty |u(t_k, x) - u^k_N(x)|^2 f_X(t_k)(x) dx \tag{5.11}
\]
\[\leq \|u(t_k, \cdot) - u^k_N(\cdot)\|_{L^\infty(\mathbb{R})}^2 \int_{-\infty}^\infty f_X(t_k)(x) dx \tag{5.12}
\]
\[\leq 2\|u(t_k, \cdot) - u^k_N(\cdot)\|^2 + 2\|\partial_x(u(t_k, \cdot) - u^k_N(\cdot))\|^2,
\]
where $f$ is the probability density function.
In the above we used the fact that

\[ E[\hat{X}_{k+1} - X_{k+1}]^2 \leq (1 + h)E[\hat{X}_k - X_k]^2 + C(h^2 + h)\left( E[\hat{X}_k - X_k]^2 + \hat{h} \right) + Ch\left( E[\hat{X}_k - X_k]^2 + \hat{h} \right) + C(h^2 + h)\left( \|u(t_k, \cdot) - u_k\|^2 + \|\partial_x(u(t_k, \cdot) - u_k)\|^2 \right) \]

\[ \leq (1 + Ch)E[\hat{X}_k - X_k]^2 + Ch^2 + Ch\left( \|u(t_k, \cdot) - u_k\|^2 + \|\partial_x(u(t_k, \cdot) - u_k)\|^2 \right). \]

Apply the forward discrete Gronwall inequality, and notice \( \hat{X}_0 = X_0 = x \), we obtain, for \( k = 1, 2, \cdots, M \),

\[ E[\hat{X}_k - X_k]^2 \leq (1 + Ch)^kE[\hat{X}_0 - X_0]^2 + C(1 + Ch)^{k-1} \sum_{j=1}^k \hat{h}^2 \]

\[ +C(1 + Ch)^{k-1}h \sum_{j=1}^k \left( \|u(t_j, \cdot) - u_j\|^2 + \|\partial_x(u(t_j, \cdot) - u_j)\|^2 \right) \]

\[ \leq C(1 + Ch)^MT\left[ h + N^{1-m}Y(u, \phi) \right]. \]

In the above we used the fact that \( k \leq M = T/h \). Further, noting \( T = hM \) we also have

\[ (1 + Ch)^M = (1 + \frac{CT}{M})M \to e^{CT}, \quad \text{as} \ M \to \infty, \]

consequently we must have

\[ E[\hat{X}_k - X_k]^2 \lesssim h + N^{1-m}Y(u, \phi), \]

where \( C \) is a generic constant that is independent of the choice of the number of time steps \( M \) and the number of Hermite functions \( N \). This, together with (5.4), yields that final mean square error estimate for \( X(t_k) - X_k \):

\[ E[X(t_k) - X_k]^2 \lesssim h + N^{1-m}Y(u, \phi). \tag{5.13} \]

The error estimates for \( E[Y(t_k) - Y_k]^2 \) and \( E[Z(t_k) - Z_k]^2 \) then follows easily from the relation (3.18), (5.1) and the estimate (5.13). The proof is now complete. \( \blacksquare \)

**Remark 5.2** From the proof of Theorem 5.1 we see that if we combine the higher order schemes for both PDE (2.6) and the resulting decoupled SDE (2.5), then we will obtain a higher order rate of convergence. The proof will be completely similar, but conceivably much more lengthy and tedious. We shall discuss this issue with extensive numerical examples in the next section.

## 6 Numerical Simulations and Conclusions

### 6.1 Implementation details

Let \( \{h_j(x)\}_{j=0,1,\ldots,N} \) be the Lagrange interpolation polynomials in \( P_N \) associated with the Gauss points \( \{x_j\}_{j=0,1,\ldots,N} \), which are zeros of \( H_{N+1}(x) \). We define \( \hat{h}_j(x) = h_j(x) \frac{x - x_j}{e^{-x^2/2}}, \) then

\[ \hat{h}_j(x) = \delta_{ij}, \quad \mathcal{H}_N = \text{span}\{\hat{h}_j : j = 0, 1, \ldots, N\} \]
so any function \( g \in \mathcal{H}_N \) can be written as

\[
g(x) = \sum_{j=0}^{N} g(x_j) \hat{h}_j(x).
\]

We derive easily from \( \hat{h}_j'(x_i) \)

\[
\hat{h}_j'(x_i) = \frac{e^{-x_i^2/2}}{e^{-x_j^2/2}} (h_j'(x_i) - x_i \delta_{ij}) = \left\{ \begin{array}{ll} \frac{\hat{H}_N(x_i)}{(x_i-x_j)H_N(x_j)}, & i \neq j \\ 0, & i = j \end{array} \right.
\]

Hence, \( \hat{h}_j'(x_i) \) can be computed in a stable way for \( N \) large.

Let us denote

\[
\bar{u} = (u_k^N(x_0), u_k^N(x_1), \ldots, u_k^N(x_N))^T, \quad \bar{f} = (f^{k+1}(x_0), f^{k+1}(x_1), \ldots, f^{k+1}(x_N))^T
\]

\[
s_{ij} = (\alpha(x)\hat{h}_j, \hat{h}_i)_N = \sum_{l=0}^{N} \alpha(x_l)\hat{h}_j'(x_l)\hat{h}_i'(x_l)\hat{w}_l + \alpha(x_i)\hat{w}_i \delta_{ij}
\]

with \( \alpha(x_i) = \frac{b}{2} \sigma^2(t_k, x_i, u_k^{k+1}(x_i)) \). Then, the equation (3.5) is reduced to the following linear system:

\[
(S + W)\bar{u} = W\bar{f},
\]

where

\[
W = \text{diag}(\hat{w}_0, \hat{w}_1, \ldots, \hat{w}_N), S = (s_{ij})_{i,j=0,1,\ldots,N}
\]

### 6.2 Numerical results and discussions

In order to make sensible comparisons to the existing results, we shall consider an example proposed by Milstein and Tretyakov [24], and use our spectral method to carry out the numerical approximation for \( u(t, x) \), and then use different stochastic numerical schemes for the resulting decoupled SDE. In particular, we shall use the standard Euler scheme, first-order Milstein scheme, \( 3/2 \)-order Platen-Wagner strong scheme, as well as \( 3/2 \)-order stochastic Adam-Bashforth scheme (SAB) scheme, respectively.

**Example.** (Milstein-Tretyakov [24]). Consider the following FBSDE:

\[
\begin{cases}
dX = \frac{X(1 + X^2)}{(2 + X^2)^3} dt + \frac{1 + X^2}{2 + X^2} \sqrt{1 + 2Y^2 + \exp(-\frac{2X^2}{t+1})} dW(t), & X(0) = x, \\
dY = -g(t, X, Y) dt - f(t, X, Y) Z dt + Z dW(t), & Y(T) = \exp(-\frac{X^2(T)}{T+1})
\end{cases}
\]

where

\[
g(t, x, u) = \frac{1}{t + 1} \exp\left( -\frac{x^2}{t + 1} \right) \left[ 4x^2(1 + x^2) \left( \frac{1 + x^2}{2 + x^2} \right)^3 + \left( \frac{1 + x^2}{2 + x^2} \right)^2 \left( 1 - \frac{2x^2}{t + 1} \right) - \frac{x^2}{t + 1} \right],
\]

\[
f(t, x, u) = \frac{x}{(2 + x^2)^2} \sqrt{\frac{1 + u^2 + \exp(-\frac{2x^2}{t+1})}{1 + 2u^2}}.
\]
Then the corresponding Cauchy problem has the form, for \( t < T, x \in \mathbb{R} \)

\[
\begin{aligned}
0 &= \frac{\partial u}{\partial t} + \frac{1}{2} \left( 1 + x^2 \right)^2 \frac{1 + 2u^2}{1 + u^2} \frac{\partial^2 u}{\partial x^2} + \frac{2x(1 + x^2)}{(2 + x^2)^3} \frac{\partial u}{\partial x} \\
+ &\frac{1}{t+1} \exp\left( -\frac{x^2}{t+1} \right) \left[ 4x^2(1 + x^2) + \left( \frac{1 + x^2}{2 + x^2} \right)^2 \left( 1 - \frac{2x^2}{t+1} \right) - \frac{x^2}{t+1} \right], \\
\end{aligned}
\]

(6.2)

It is easy to verify that the solution to problem (6.2) is

\[
u(t, x) = \exp\left( -\frac{x^2}{t+1} \right).
\]

(6.3)

Since \( Y(t) = u(t, X(t)) \), plugging \( Y(t) \) into the forward equation in (6.1) we get

\[
dX = \frac{X(1 + X^2)}{(2 + X^2)^3} dt + \frac{1 + X^2}{2 + X^2} dW(t), \quad X(0) = x.
\]

(6.4)

One can then easily check that the solution to this equation can be expressed as

\[
X(t) = \Lambda(x + \arctan x + W(t)),
\]

where the function \( \Lambda(z) \) is defined by the equation

\[
\Lambda(z) + \arctan \Lambda(z) = z.
\]

(6.5)

Differentiating (6.5) with respect to \( z \), we get

\[
\Lambda'(z) = \frac{1 + \Lambda^2(z)}{2 + \Lambda^2(z)} > 0 \quad \text{for all} \quad z \in \mathbb{R}, \quad \text{and} \quad \Lambda''(z) = \frac{2\Lambda(1 + \Lambda^2)}{(2 + \Lambda^2)^3}.
\]

(6.6)

which implies that \( \Lambda(z) \) is an one-to-one function. Hence \( X(0) = \Lambda(x + \arctan x) = x \). Furthermore, due to the Itô formula, we have

\[
dX = \Lambda'(x + \arctan x + W(t)) dW(t) + \frac{1}{2} \Lambda''(x + \arctan x + W(t)) dW(t)
\]

\[
= \frac{X(1 + X^2)}{(2 + X^2)^3} dt + \frac{1 + X^2}{2 + X^2} dW(t),
\]

Hence, the solution to (6.1) is

\[
X(t) = \Lambda(x + \arctan x + W(t)),
\]

\[
Y(t) = \exp\left( -\frac{X^2}{t+1} \right),
\]

\[
Z(t) = -\frac{2X(1 + X^2)}{(t+1)(2 + X^2)} \exp\left( -\frac{X^2}{t+1} \right).
\]

We first carry out numerical tests on the parabolic equation (6.2) using the second-order (in time) scheme (3.6). In Tables 1 to 3, we tabulate the errors \( \max_k \| u(t_k, x) - u_N^k(x) \|_{L^2(\mathbb{R})} \), \( \max_k \| \frac{\partial u}{\partial x}(t_k, x) - \frac{\partial u}{\partial x} u_N^k(x) \|_{L^2(\mathbb{R})} \), \( \max_k \| \frac{\partial^2 u}{\partial x^2}(t_k, x) - \frac{\partial^2 u}{\partial x^2} u_N^k(x) \|_{L^2(\mathbb{R})} \), respectively. We note that it is important to measure the errors of the approximation \( u_N^k(x) \) to the first and second derivatives of \( u(t_k, x) \) since
Table 1: $\max_k \|u(t_k, x) - u_N^k(x)\|_{L^2(\mathbb{R})}$

<table>
<thead>
<tr>
<th>(h) (\times) (N)</th>
<th>64</th>
<th>70</th>
<th>100</th>
<th>128</th>
<th>150</th>
<th>RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0049</td>
<td>0.0049</td>
<td>0.0049</td>
<td>0.0049</td>
<td>0.0049</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>1.5391e-004</td>
<td>9.6229e-005</td>
<td>7.8519e-005</td>
<td>7.8554e-005</td>
<td>7.8555e-005</td>
<td>1.7344</td>
</tr>
<tr>
<td>0.05</td>
<td>1.5389e-004</td>
<td>8.3279e-005</td>
<td>1.3177e-005</td>
<td>1.3167e-005</td>
<td>1.3168e-005</td>
<td>1.9492</td>
</tr>
<tr>
<td>0.005</td>
<td>1.5388e-004</td>
<td>8.3266e-005</td>
<td>3.9669e-006</td>
<td>8.4446e-007</td>
<td>8.3360e-007</td>
<td>1.9822</td>
</tr>
</tbody>
</table>

the first-order Milstein scheme (3.9) and 2-order Platen-Wagner scheme (3.10) need to use these derivatives. In the last column of these tables, the rates of convergence in time with \(N = 150\) are reported. We observe that essentially second-order accuracy in time and exponential convergence in space are achieved for all three quantities.

Table 2: $\max_k \|\frac{\partial}{\partial x} u(t_k, x) - \frac{\partial}{\partial x} u_N^k(x)\|_{L^2(\mathbb{R})}$

<table>
<thead>
<tr>
<th>(h) (\times) (N)</th>
<th>64</th>
<th>70</th>
<th>100</th>
<th>128</th>
<th>150</th>
<th>RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0025</td>
<td>0.0074</td>
<td>0.0074</td>
<td>0.0074</td>
<td>0.0074</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0018</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0017</td>
<td>1.6052</td>
</tr>
<tr>
<td>0.05</td>
<td>9.6178e-004</td>
<td>1.3741e-004</td>
<td>1.3685e-004</td>
<td>1.3685e-004</td>
<td>1.3685e-004</td>
<td>1.8174</td>
</tr>
<tr>
<td>0.02</td>
<td>9.6179e-004</td>
<td>5.4862e-005</td>
<td>2.3165e-005</td>
<td>2.3152e-005</td>
<td>2.3155e-005</td>
<td>1.9391</td>
</tr>
<tr>
<td>0.005</td>
<td>9.6178e-004</td>
<td>5.4859e-005</td>
<td>3.7980e-006</td>
<td>1.4925e-006</td>
<td>1.9777</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: $\max_k \|\frac{\partial^2}{\partial x^2} u(t_k, x) - \frac{\partial^2}{\partial x^2} u_N^k(x)\|_{L^2(\mathbb{R})}$

<table>
<thead>
<tr>
<th>(h) (\times) (N)</th>
<th>64</th>
<th>70</th>
<th>100</th>
<th>128</th>
<th>150</th>
<th>RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0230</td>
<td>0.0226</td>
<td>0.0224</td>
<td>0.0224</td>
<td>0.0224</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0188</td>
<td>0.0112</td>
<td>0.0055</td>
<td>0.0055</td>
<td>0.0055</td>
<td>1.5326</td>
</tr>
<tr>
<td>0.05</td>
<td>7.7079e-004</td>
<td>4.5192e-004</td>
<td>4.5174e-004</td>
<td>4.5174e-004</td>
<td>4.5174e-004</td>
<td>1.8029</td>
</tr>
<tr>
<td>0.02</td>
<td>7.7061e-004</td>
<td>7.8042e-005</td>
<td>7.7166e-005</td>
<td>7.7166e-005</td>
<td>7.7166e-005</td>
<td>1.9293</td>
</tr>
<tr>
<td>0.005</td>
<td>7.7058e-004</td>
<td>6.0486e-005</td>
<td>8.0385e-006</td>
<td>8.0385e-006</td>
<td>1.6310</td>
<td></td>
</tr>
</tbody>
</table>

Next, we examine the errors of the full simulation for the FBSDE with four different schemes presented in Section 3 for the forward SDE and with the Hermite-collocation scheme (3.5). For the sake of comparison with the results in [24], we also computed the same problem with a Monte Carlo simulation (with \(S = 1000\) independent realizations of \(X(T)\) and \(X_N\)) for the forward SDE. The results of the Euler scheme with Monte Carlo simulation are reported in Table 6.2. The averages presented in the table are computed as follows:

\[
E(X(T) - X_N)^2 = \frac{1}{S} \sum_{k=1}^{S} \left( X^{(k)}(T) - X_N^{(k)} \right)^2 \pm 2 \frac{D_S}{S}
\]
where

\[ D_S = \frac{1}{S} \sum_{k=1}^{S} (X^{(k)}(T) - X^{(k)}_N)^4 - \left[ \frac{1}{S} \sum_{k=1}^{S} (X^{(k)}(T) - X^{(k)}_N)^2 \right]^2 \]

Hence,

\[ \left[ E(X(T) - X_N)^2 \right]^{1/2} = \sqrt{\frac{1}{S} \sum_{k=1}^{S} (X^{(k)}(T) - X^{(k)}_N)^2} \]

\[ + \sqrt{\frac{1}{S} \sum_{k=1}^{S} (X^{(k)}(T) - X^{(k)}_N)^2} \pm 2\sqrt{\frac{D_S}{S}} - \sqrt{\frac{1}{S} \sum_{k=1}^{S} (X^{(k)}(T) - X^{(k)}_N)^2} \]

\(X^{(k)}(T)\) and \(X^{(k)}_N\) are independent realizations of \(X(T)\) and \(X_N\), respectively, \(k = 1, 2, \ldots, S = 1000\).

Table 4: Euler Scheme with 1000 Monte Carlo simulation realizations with \(N = 150\)

<table>
<thead>
<tr>
<th>h</th>
<th>(E(X(T) - X_N)^2)^{1/2}</th>
<th>Rate</th>
<th>(E(Y(T) - Y_N)^2)^{1/2}</th>
<th>Rate</th>
<th>(E(Z(T) - Z_N)^2)^{1/2}</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.2453 ± 0.0121</td>
<td></td>
<td>0.0325 ± 0.0022</td>
<td></td>
<td>0.0121 ± 7.9404e-004</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.1605 ± 0.0083</td>
<td>0.4643</td>
<td>0.0213 ± 0.0015</td>
<td></td>
<td>0.0079 ± 5.0520e-004</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>0.0745 ± 0.0037</td>
<td>0.5527</td>
<td>0.0101 ± 0.0007</td>
<td></td>
<td>0.0037 ± 2.2556e-004</td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>0.0479 ± 0.0024</td>
<td>0.482</td>
<td>0.0064 ± 0.0004</td>
<td></td>
<td>0.0023 ± 1.4899e-004</td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>0.0259 ± 0.0013</td>
<td>0.4435</td>
<td>0.0036 ± 0.0002</td>
<td></td>
<td>0.0013 ± 7.6865e-005</td>
<td></td>
</tr>
</tbody>
</table>

For conciseness, we now omit the Monte Carlo errors, which, as can be seen from Table 6.2, are significantly small than the approximation errors. In what follows, we denote \([E(X(T) - X_N)^2]\)^{1/2} by \(X\), \([E(Y(T) - Y_N)^2]\)^{1/2} by \(Y\) and \([E(Z(T) - Z_N)^2]\)^{1/2} by \(Z\).

In the following three tables, we report these errors by using the three schemes with 1000 independent realizations of \(X(T)\) and \(X_N\).

Table 5: First Order SDE Scheme with \(N = 150\)

<table>
<thead>
<tr>
<th>h</th>
<th>X</th>
<th>Rate</th>
<th>Y</th>
<th>Rate</th>
<th>Z</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5</td>
<td>0.0937</td>
<td></td>
<td>0.0131</td>
<td></td>
<td>0.0048</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>0.0394</td>
<td>0.9455</td>
<td>0.0052</td>
<td>1.0084</td>
<td>0.0021</td>
<td>0.9022</td>
</tr>
<tr>
<td>.05</td>
<td>0.0095</td>
<td>1.0261</td>
<td>0.0013</td>
<td>1.0000</td>
<td>4.7994e-004</td>
<td>1.0647</td>
</tr>
<tr>
<td>.02</td>
<td>0.0039</td>
<td>0.9717</td>
<td>5.3451e-004</td>
<td>0.9700</td>
<td>1.9852e-004</td>
<td>0.9634</td>
</tr>
<tr>
<td>.005</td>
<td>9.5023e-004</td>
<td>1.0186</td>
<td>1.2996e-004</td>
<td>1.0201</td>
<td>4.8413e-005</td>
<td>1.0179</td>
</tr>
</tbody>
</table>

Observe that all three schemes produce expected convergence rates. The \(\frac{3}{2}\)-order SAB scheme is slightly less accurate than the \(\frac{3}{2}\)-order strong scheme but the convergence rates of these two schemes are essentially the same.
Table 6: $\frac{3}{2}$ order strong scheme with $N = 200$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$X$</th>
<th>Rate</th>
<th>$Y$</th>
<th>Rate</th>
<th>$Z$</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5</td>
<td>0.0425</td>
<td>0.0041</td>
<td>0.0025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>0.0115</td>
<td>1.4266</td>
<td>0.0015</td>
<td>1.0974</td>
<td>5.4051e-004</td>
<td>1.6714</td>
</tr>
<tr>
<td>.05</td>
<td>0.0015</td>
<td>1.4693</td>
<td>1.7410e-004</td>
<td>1.5535</td>
<td>7.5482e-005</td>
<td>1.4201</td>
</tr>
<tr>
<td>.02</td>
<td>3.3583e-004</td>
<td>1.6333</td>
<td>3.6401e-005</td>
<td>1.7080</td>
<td>1.7780e-005</td>
<td>1.5779</td>
</tr>
<tr>
<td>.005</td>
<td>3.9962e-005</td>
<td>1.5355</td>
<td>7.6776e-006</td>
<td>1.1226</td>
<td>2.2442e-006</td>
<td>1.4930</td>
</tr>
</tbody>
</table>

Table 7: $\frac{3}{2}$ order SAB scheme with $N = 200$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$X$</th>
<th>Rate</th>
<th>$Y$</th>
<th>Rate</th>
<th>$Z$</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5</td>
<td>0.1301</td>
<td>0.0177</td>
<td>0.0063</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>0.0400</td>
<td>1.2872</td>
<td>0.0053</td>
<td>1.3160</td>
<td>0.0020</td>
<td>1.2522</td>
</tr>
<tr>
<td>.05</td>
<td>0.0048</td>
<td>1.5294</td>
<td>6.6087e-004</td>
<td>1.5018</td>
<td>2.3552e-004</td>
<td>1.5430</td>
</tr>
<tr>
<td>.02</td>
<td>0.0013</td>
<td>1.4256</td>
<td>1.6757e-004</td>
<td>1.4975</td>
<td>6.5992e-005</td>
<td>1.3885</td>
</tr>
<tr>
<td>.005</td>
<td>1.5749e-004</td>
<td>1.5226</td>
<td>2.0907e-005</td>
<td>1.5014</td>
<td>7.9073e-006</td>
<td>1.5305</td>
</tr>
</tbody>
</table>

7 Concluding remarks

We presented a numerical method for a class of forward-backward stochastic differential equations (FBSDEs). The method is based on the Four Step Scheme with a Hermite-spectral method to approximate the solution to the decoupled quasilinear PDE on the whole space. The use of Hermite-spectral method not only avoids the use of artificial far field boundary conditions but also leads to spectrally accurate results in space. We carried out a rigorous error analysis for a fully discretized scheme for the FBSDEs with a first-order scheme in time and a Hermite-spectral scheme in space, and indicated that similar analysis can be extended to higher-order schemes in time. We presented detailed numerical comparisons between several schemes for the resulting decoupled forward SDE and showed that the stochastic version of the Adams-Bashforth scheme coupled with the Hermite-spectral method leads to a convergence rate of $\frac{3}{2}$ (in time).

Although the analysis and computation is performed for the one-dimensional case, it is clear that similar analysis can be carried out for higher-dimensional cases, while computationally a direct tensor-product extension of the algorithm quickly becomes prohibitive as the dimension increases since . In a forthcoming work, we plan to introduce new elliptic solvers based on lattice rules and sparse grids which would allow us to handle FBSDEs with relatively large dimensional problems.

References


