SIAM J. NUMER. ANAL.
Vol. 46, No. 5, pp. 2636–2661
© 2008 Society for Industrial and Applied Mathematics

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

ON NUMERICAL APPROXIMATIONS OF FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS*

JIN MA†, JIE SHEN‡, AND YANHONG ZHAO‡

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* [J. Douglas, Jr., J. Ma, and P. Protter, *Ann. Appl. Probab.*, 6 (1996), pp. 940–968] but with a Hermite-spectral method to approximate the solution to the decoupling quasi-linear 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.

Key words. forward-backward stochastic differential equations, four step scheme, Hermite-spectral method, convergence rate

AMS subject classifications. Primary, 60H10; Secondary, 34F05, 93E03

DOI. 10.1137/06067393X

1. Introduction.

Since the seminal work of Pardoux and Peng [32] in early 1990s, the theory of backward stochastic differential equations (BSDEs) and forward-backward stochastic differential equations (FBSDEs) has grown into a ubiquitous tool in the fields of stochastic optimizations and mathematical finance (see, for example, the books of El Karoui and Mazliak [14] and Ma and Yong [28] and the survey of El Karoui, Peng, and Quenez [15] for a detailed account of both theory and applications of such SDEs). In the meantime, after the earlier works of Bally [1] and Douglas, Ma, and Protter [13], 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 the PDE method in the Markovian case (e.g., Chevance [8] and Zhang and Zheng [36]), random walk approximations (e.g., Briand, Delyon, and Mémin [6] and Ma et al. [26]), Malliavin calculus and Monte-Carlo method (e.g., Zhang [35], Ma and Zhang [29], and Bouchard and Touzi [5]), the quantization method (e.g., Bally and Pagès [2] and Bally, Pagès, and Printems [3]), and recently the regression-based Monte-Carlo method (Lemor, Gobet, and Warin [23] and [24]).

In the case of coupled FBSDEs, however, the results are very limited, largely due to the lack of a solution method itself in such cases. It has been well understood

*Received by the editors November 2, 2006; accepted for publication (in revised form) March 10, 2008; published electronically July 2, 2008. This work was supported in part by NSF grants DMS-0311915, DMS-0505427, and DMS-0610646 and a Purdue Research Foundation grant.
http://www.siam.org/journals/sinum/46-5/67393.html
†Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532 (jinma@usc.edu).
‡Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395 (shen@math.purdue.edu, zhao6@purdue.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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 and Yong [28]. Such an idea has led directly to most of the existing numerical results for FBSDEs: from the early work [13] to the recent improvements by Milstein and Tretyakov [31] and Delarue and Menozzi [12]. Extending such a decoupling idea and combining it with an optimal control method as well as Monte-Carlo simulation, Cvitanić and Zhang [9] and Bender and 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 FBSDE and its decoupling quasi-linear PDE, where the notorious "curse of dimensionality" is still a formidable difficulty for any numerical method. In fact, to the best of our 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 convergence rate $\frac{1}{2}$, except in [31], 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., [21]), 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 an effort towards this goal. We shall revisit the four step scheme again, but will replace 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 solution is approximated directly by Hermite functions as an orthonormal basis on the whole space, without using ad hoc artificial boundary conditions as in previous approaches using finite differences; (ii) the rate of convergence is related directly 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 FBSDEs seems to be able to achieve higher than first-order convergence rate.

Given the aforementioned spectral accuracy in space, it seems hopeful that one can produce arbitrarily higher-order global error by applying a higher-order scheme in SDEs in, e.g., [21], at least when the coefficients are smooth. We shall carry out some numerical simulations to validate this point. To compare with the existing results, we shall use an example proposed in Milstein and Tretyakov [31], 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 [16]. By the nature of these schemes, we expect the Euler scheme to be 1/2-order convergent, the Milstein first-order scheme to be first-order convergent, and the Platen–Wagner and SAB schemes to be 3/2-order convergent. Our simulation results give numerical evidence that this is exactly the case. Also, this

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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. A rigorous proof for higher order schemes would 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, albeit technical, can be extended to higher-dimensional cases with a tensor-product approach. However, such a method quickly becomes nonfeasible for spatial dimensions higher than three, unless some non–tensor-product method, such as those based on sparse grid (cf. [7] and [34]) or lattice rules (cf. [25], [19], and [22]), 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 $W = \{W(t) : t \geq 0\}$. We shall denote $\mathbf{F}^W = \{\mathcal{F}_t^W : t \geq 0\}$ to be the natural filtration generated by W, 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 $t \in [0, T]$,

$$\begin{cases} X(t) = x + \displaystyle\int_0^t b(s, \Theta(s))ds + \int_0^t \sigma(s, X(s), Y(s))dW(s), \\ \\ Y(t) = \varphi(X(T)) + \displaystyle\int_t^T g(s, \Theta(s))ds - \int_t^T Z(s)dW(s), \end{cases} \tag{2.1}$$

where $\Theta \stackrel{\triangle}{=} (X, Y, Z)$. 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 what follows.

(A1) The functions b, σ, g, and φ are continuously differentiable in all variables. Moreover, if we denote these functions by a generic one ψ, then there exists a constant $K > 0$ such that, for any $t \in [0, T]$, ψ 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 functions b, g, and φ satisfies the following growth conditions: for some $K > 0$,

$$|g(t, x, y, z)| \leq K(1 + |y| + |z|),$$

$$|\varphi(x)|, |b(t, x, y, z)| \leq K.$$

(A3) Function σ has bounded second derivatives, and there exist constants $0 < c < C$ such that

$$c \leq \sigma(t, x, y) \leq C \quad \forall\, (t, x, y) \in [0, T] \times \mathbb{R}^2. \tag{2.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Remark 2.1. We remark that assumptions (A1)–(A3) are stronger than is necessary for the well-posedness of FBSDE (2.1). In fact, in [10] it was shown that under (A1)–(A3) but without the differentiability assumption on the coefficients the FBSDE (2.1) already possesses a unique adapted solution over an arbitrarily prescribed time duration. The extra smoothness condition is needed only for our numerical scheme, and therefore is not essential, in principle. Note also, however, that even with the added differentiability conditions, our assumptions are still much weaker than that of [13].

Four step scheme. In [27] (see also [10]) it was shown that the *unique* adapted solution of FBSDE (2.1) can be obtained by the 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 quasi-linear parabolic PDE :

$$\begin{cases} u_t + \dfrac{1}{2}\sigma^2(t,x,u)u_{xx} + b(t,x,u,z(t,x,u,u_x))u_x \\ \quad + g(t,x,u,z(t,x,u,u_x)) = 0, \qquad (t,x)\in(0,T)\times\mathbb{R}, \\ u(T,x) = \varphi(x), \quad x\in\mathbb{R}. \end{cases} \tag{2.3}$$

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

$$X(t) = x + \int_0^t \tilde{b}(s,X(s))ds + \int_0^t \tilde{\sigma}(s,X(s))dW(s), \tag{2.4}$$

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

Step 4. Set

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

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 (2.3) and the other for the (forward) SDE (2.4). We shall adopt a Hermite–Galerkin method with numerical integration. To do this we find it convenient to rewrite PDE (2.3) in a divergence form, mainly for the sake of numerical stability:

$$\begin{cases} u_t + \partial_x\left(\dfrac{1}{2}\sigma^2(t,x,u)u_x\right) + \hat{b}(t,x,u,u_x)u_x \\ \quad + g(t,x,u,\sigma(t,x,u)u_x) = 0, \qquad (t,x)\in(0,T)\times\mathbb{R}, \\ u(T,x) = \varphi(x), \quad x\in\mathbb{R}, \end{cases} \tag{2.5}$$

where

$$\hat{b}(t,x,y,z) = b(t,x,y,\sigma(t,x,y)z) - \sigma(t,x,y)(\sigma_x(t,x,y) + \sigma_y(t,x,y)z).$$

Remark 2.2. We note here that the main difference between b and $\hat{b}$ is that the latter now has a linear growth in the variable z, and thus it no longer satisfies assumptions (A1) and (A2). We shall make the following adjustment in what follows

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

to facilitate the error analysis. Note that under (A1)–(A3) it is known (cf. [11]) that the solution to the PDE (2.3), whence (2.5), has a globally bounded gradient u_x. Thus in light of the four step scheme, if we introduce the "cut-off" version of the coefficient function $\hat{b}$,

$$\hat{b}_K(t,x,y,z) = b(t,x,y,\sigma(t,x,y)z) - \sigma(t,x,y)(\sigma_x(t,x,y) + \sigma_y(t,x,y)\gamma_K(z)),$$

where $\gamma_K(z) \in C^\infty(\mathbb{R})$, $\gamma_K(z) = z$ for $|z| \le K$, and $|\gamma_K(z)| \le K+1$, $|\gamma_K'(z)| \le 1$, then for K large enough, the solution of the PDE (2.5) with coefficient $\hat{b}_K$ will coincide with the original solution. In other words, we can simply consider the "truncated" coefficients $\hat{b}_K$ for the purpose of approximation, which again satisfies (A1) and (A2). For the sake of simplicity, in what follows we shall still use $\hat{b}$ to denote $\hat{b}_K$.

Hermite polynomials, Hermite functions, and their properties. Throughout this paper we shall denote by $L^2(\mathbb{R})$ the space of all square integrable functions $u : \mathbb{R} \mapsto \mathbb{R}$, with the inner product

$$\langle u, v\rangle \;=\; \int_{\mathbb{R}} u(x)v(x)dx \qquad \forall\, u, v \in L^2(\mathbb{R})$$

and the norm $\|u\|^2 \triangleq \int_{\mathbb{R}} |u(x)|^2 dx$. For simplicity, in what follows, without further specification we shall always denote by $\|\cdot\|$ the $L^2(\mathbb{R})$ norm.

Next, for each integer $k \ge 1$, we denote the differential operator $\partial^k \triangleq \frac{d^k}{dx^k}$. Then the Sobolev space $H^m(\mathbb{R})$ is defined by

$$H^m(\mathbb{R}) = \{u \mid \partial^k u \in L^2(\mathbb{R}) \quad \forall\, 0 \le k \le m\},$$

with the seminorm $|u|_m \triangleq \|\partial^m u\|$, and the norm $\|u\|_m^2 \triangleq \sum_{k=0}^m \|\partial^k u\|^2$. We note that the operator ∂^k can be easily extended to the higher-dimensional case, namely, as a partial differential operator, in an obvious way. We define $H^{-1}(\mathbb{R})$ as the dual space of $H^1(\mathbb{R})$, and its norm will be denoted by $\|\cdot\|_{-1}$.

We recall that (cf., e.g., [17]) the *Hermite polynomials* $\{H_n(x)\}_{n\ge 0}$ are defined by

$$H_n(x) = (-1)^n \exp(x^2)\frac{d^n}{dx^n}\exp(-x^2), \qquad n = 0, 1, \ldots . \tag{2.6}$$

One can easily check that H_n is the solution to the recursive equations

$$\begin{cases} H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \\ H_n'(x) = 2nH_{n-1}(x), \\ H_0(x) = 1, \quad H_1(x) = 2x, \end{cases}$$

and the following orthogonality condition holds:

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2}dx = \gamma_n\delta_{mn}, \qquad m, n \ge 1, \tag{2.7}$$

where $\gamma_n = \sqrt{\pi}2^n n!$ and δ_{mn} is the Kronecker delta. For each $n \ge 1$, $H_n(x)$ is called the "Hermite polynomial of degree n," and the *Hermite function of degree* n is defined by

$$\widehat{H}_n(x) = \frac{1}{\sqrt{2^n n!}} e^{-x^2/2} H_n(x).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

It follows immediately that the Hermite functions $\widehat{H}_n(x)$ enjoy the following recurrence and orthogonal relations:

$$\widehat{H}_{n+1}(x) = x\sqrt{\frac{2}{n+1}}\widehat{H}_n(x) - \sqrt{\frac{n}{n+1}}\widehat{H}_{n-1}(x), \quad n \geq 1, \tag{2.8}$$

$$\frac{d}{dx}\widehat{H}_n(x) = \sqrt{\frac{n}{2}}\widehat{H}_{n-1}(x) - \sqrt{\frac{n+1}{2}}\widehat{H}_{n+1}(x), \tag{2.9}$$

$$\int_{-\infty}^{\infty} \widehat{H}_m(x)\widehat{H}_n(x)dx = \sqrt{\pi}\delta_{mn}. \tag{2.10}$$

Moreover, we derive from $H_n' = 2nH_{n-1}$ that $\widehat{H}_n' + x\widehat{H}_n = \sqrt{2n}\widehat{H}_{n-1}$. Hence, by introducing a new operator $Du \stackrel{\triangle}{=} \partial_x u + xu$, we obtain the following recursive relation:

$$D\widehat{H}_n = \sqrt{2n}\widehat{H}_{n-1}. \tag{2.11}$$

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 $\mathcal{P}_N$ be the space of polynomials of degree less than or equal to N. We then define

$$\mathcal{H}_N = e^{-x^2/2}\mathcal{P}_N \stackrel{\triangle}{=} \{e^{-x^2/2}p(x) :\ p(x) \in \mathcal{P}_N\}. \tag{2.12}$$

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

$$w_j = \frac{\sqrt{\pi}2^N N!}{(N+1)[H_N(x_j)]^2}, \quad 0 \leq j \leq N. \tag{2.13}$$

The pairs $\{(x_i, \omega_i)\}_{i=0}^N$ are known as the *Hermite–Gauss quadrature points and weights*, respectively (cf., e.g., [33]). Using the Hermite–Gauss quadrature points and weights, we can then define a discrete scalar product defined on $C(\mathbb{R})$ by

$$\langle f, g\rangle_N \stackrel{\triangle}{=} \sum_{j=0}^{N} f(x_j)g(x_j)\hat{w}_j \qquad \forall f, g \in C(\mathbb{R}), \tag{2.14}$$

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

$$\hat{w}_j \stackrel{\triangle}{=} w_j\ e^{x_j{}^2} = \frac{\sqrt{\pi}}{(N+1)\widehat{H}_N^2(x_j)}, \qquad 0 \leq j \leq N.$$

A direct consequence of the Hermite–Gauss quadrature is

$$\langle u, v\rangle\ =\ \langle u, v\rangle_N \qquad \forall\ u \in \mathcal{H}_m,\ v \in \mathcal{H}_n,\ n + m \leq 2N + 1. \tag{2.15}$$

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

$$H_D^m(\mathbb{R}) \stackrel{\triangle}{=} \{u \mid D^m u \in L^2(\mathbb{R})\},$$

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

In order to introduce our numerical approximation we need the following two operators:

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

- the "*orthogonal projection operator*" $P_N : L^2(\mathbb{R}) \to \mathcal{H}_N$, defined by
$$\langle u - P_N u, v\rangle = 0 \quad \forall u \in L^2(\mathbb{R}),\ v \in \mathcal{H}_N;$$
- the "*Hermite–Gauss interpolation operator*" $I_N : C(\mathbb{R}) \to \mathcal{H}_N$, defined by
$$I_N f \in \mathcal{H}_N \quad \text{such that} \quad I_N f(x_j) = f(x_j), \quad j = 0, 1, \ldots, N. \tag{2.16}$$

The following basic approximation results for P_N can be found in, e.g., [18].

LEMMA 2.3. *For any $u \in H^r_D(\mathbb{R})$, $N \in \mathbb{N}$, and $0 \le s \le r$, there exists a constant $C > 0$, independent of N and the function u, such that*

$$\|D^s(u - P_N u)\| \le C N^{\frac{s-r}{2}} \|D^r u\| \tag{2.17}$$

and

$$\|\partial_x(u - P_N u)\| \le C N^{\frac{1-r}{2}} \|D^r u\|. \tag{2.18}$$

3. Numerical schemes. In this section we present our numerical schemes, both for the PDE (2.5) and for the forward SDE (2.4). Although the scheme for the general quasi-linear PDE (2.5) 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 (2.5) becomes, with a slight abuse of notation in light of Remark 2.2,

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

where $\hat{b}(t,x,y,z) = b(t,x,y) - \sigma(t,x,y)(\sigma_x(t,x,y) + \sigma_y(t,x,y)\gamma_{\bar{K}}(z))$, for some $\bar{K} > 0$. We shall design a numerical scheme for (3.1) based on a semi-implicit discretization in time and the Hermite-collocation method in space, and discuss three types of schemes for the corresponding forward SDE (2.4). The error analysis for the combined scheme will be carried out in the next two sections.

Before we proceed further, we should note that the constant $\bar{K} > 0$ above is determined by the *true* solution, which is independent of the numerical scheme, and is useful only for error analysis. In what follows we shall denote all the constants depending on K in (A1)–(A3) and $\bar{K}$ above by a generic one, and still denote it by K. Then, from (A1) we see that $|\sigma_x(t,x,y)| \le K$ and $|\sigma_y(t,x,y)| \le K$ for all $(t,x,y) \in [0,T]\times\mathbb{R}^2$ follow easily. Furthermore, by (A3) it implies that σ_y satisfies the Lipschitz condition as

$$|\sigma_y(t,x,y) - \sigma_y(t,x,y')| \le K\,|y - y'|.$$

Hence,

$$\begin{aligned} &|\sigma(t,x,y)\sigma_y(t,x,y)\gamma_K(z) - \sigma(t,x,y')\sigma_y(t,x,y')\gamma_K(z')| \\ &\le |\sigma(t,x,y)\sigma_y(t,x,y)\gamma_K(z) - \sigma(t,x,y')\sigma_y(t,x,y)\gamma_K(z)| \\ &\quad + |\sigma(t,x,y')\sigma_y(t,x,y)\gamma_K(z) - \sigma(t,x,y')\sigma_y(t,x,y')\gamma_K(z)| \\ &\quad + |\sigma(t,x,y')\sigma_y(t,x,y')\gamma_K(z) - \sigma(t,x,y')\sigma_y(t,x,y')\gamma_K(z')| \\ &\le K^2(K+1)|y - y'| + CK(1+K)|y - y'| + CK|z - z'|. \end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Similarly, we have

$$|\sigma(t,x,y)\sigma_x(t,x,y) - \sigma(t,x,y')\sigma_x(t,x,y')| \le (K^2 + CK)|y - y'|.$$

For notational simplicity, we often allow the generic constant K to vary from line to line; then by the above arguments, the function $\hat{b}$ satisfies the Lipschitz condition

$$|\hat{b}(t,x,y,z) - \hat{b}(t,x,y',z')| \le K(|y - y'| + |z - z'|). \tag{3.2}$$

On the other hand, $\hat{b}$ satisfies the bounded property by boundedness of b and $\gamma_{_K}(z)$; i.e.,

$$|\hat{b}(t,x,y,z)| \le K. \tag{3.3}$$

3.1. Numerical scheme for the PDE (3.1). For any $M \in \mathbb{N}^*$, let $h = \frac{T}{M}$ be the length of the time step and set $t_k = kh$, $k = 0, 1, \ldots, M$.

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

$$\begin{aligned} &\frac{u^{k+1} - u^k}{h} + \frac{1}{2}\partial_x\Big(\sigma^2(t_k,x,u^{k+1})u_x^k\Big) \\ &+ \hat{b}(t_k,x,u^{k+1},u_x^{k+1})u_x^{k+1} + g(t_k,x,u^{k+1}) = 0. \end{aligned} \tag{3.4}$$

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

$$u^k - \frac{h}{2}\partial_x\Big(\sigma^2(t_k,x,u^{k+1})u_x^k\Big) = f_h^{k+1}(x), \quad \lim_{x\to\pm\infty} u^k(x) = 0, \tag{3.5}$$

where

$$f_h^{k+1}(x) = u^{k+1} + h\Big\{\hat{b}(t_k,x,u^{k+1},u_x^{k+1})u_x^{k+1} + g(t_k,x,u^{k+1})\Big\}.$$

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 let $\mathcal{H}_N$ be defined in (2.12). A Hermite-collocation method for (3.5) is as follows:

Let $u_N^M(x) = I_N\varphi$. For $k = M-1, \ldots, 0$, we compute $u_N^k \in \mathcal{H}_N$ such that

$$u_N^k(x_j) - \frac{h}{2}\partial_x\Big(I_N(\sigma^2(t_k,x,u_N^{k+1})\partial_x u_N^k)\Big)(x_j) = f_h^{k+1}(x_j), \quad j = 0,1,\ldots,N. \tag{3.6}$$

Using the identity (2.15) and integration by parts, we can rewrite (3.6) in the following variational formulation: Find $u_N^k \in \mathcal{H}_N$ such that

$$\langle u_N^k,\, v_N\rangle_N + \frac{h}{2}\langle\sigma^2(t_k,x,u_N^{k+1})\partial_x u_N^k,\, \partial_x v_N\rangle_N = \langle f_{h,N}^{k+1},\, v_N\rangle_N \quad \forall v_N \in \mathcal{H}_N, \tag{3.7}$$

where $f_{h,N}^{k+1} \in \mathcal{H}_N$ and $f_{h,N}^{k+1}(x_j) = f^{k+1}(x_j)$, $j = 0,1,\ldots,N$.

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

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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 (3.7) is already unpleasantly lengthy. As an example, we present the following second-order backward difference formula (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 (3.4).

For $k = M-2, \dots, 1, 0$, we compute u^k from

$$\frac{-3u_N^{k+2} + 4u_N^{k+1} - u_N^k}{2h} + \partial_x\Big(\frac{\sigma^2}{2}\partial_x u_N^k\Big) + (2\partial_x u_N^{k+1} - \partial_x u_N^{k+2})\hat{b} + g = 0, \tag{3.8}$$

where

$$\sigma^2 \triangleq \sigma^2(t_k, x, 2u_N^{k+1} - u_N^{k+2}), \quad g \triangleq g(t_k, x, 2u_N^{k+1} - u_N^{k+2}),$$

$$\hat{b} \triangleq \hat{b}(t_k, x, 2u_N^{k+1} - u_N^{k+2}, 2\partial_x u_N^{k+1} - \partial_x u_N^{k+2}).$$

We note that this scheme still leads to an elliptic equation of the form (3.5) for u^k at each time step. In particular, this second-order scheme should be used together with the $\frac{3}{2}$-order SDE scheme presented below.

3.2. Numerical schemes for the SDE (2.4). We now turn our attention to the discretization of the forward SDE (2.4). 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, $k = 0, 1, \dots, M$, at each time $t_k = kh$ we define recursively the approximate solution to the SDE (2.4) by

$$\begin{cases} X_0^N = x, \\ X_{k+1}^N = X_k^N + b(t_k, X_k^N, u_N^k(X_k^N))h + \sigma(t_k, X_k^N, u_N^k(X_k^N))\Delta_k W, \ k = 0, 1, \dots, M, \end{cases} \tag{3.9}$$

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 the 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 $u^k(x) \triangleq u(t_k, x)$, and define an intermediate approximation $\hat{X}$ by

$$\begin{cases} \hat{X}_0 = x, \\ \hat{X}_{k+1} = \hat{X}_k + b(t_k, \hat{X}_k, u^k(\hat{X}_k))h + \sigma(t_k, \hat{X}_k, u^k(\hat{X}_k))\Delta_k W, \quad k = 0, 1, \dots, M, \end{cases} \tag{3.10}$$

then by the *fundamental convergence theorem* (cf., e.g., [21]), 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 (2.4) (namely, the rate of convergence 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 (2.4) can be constructed with Taylor–Itô-type expansions (see, e.g., [21]), the complexity of these schemes increases drastically. Consequently it often

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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. [21], [31]).

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

(3.11)
$$\begin{cases} X_0 = x, \\ X_{k+1} = X_k + b(t_k, X_k, u_N^k(X_k))h + \sigma(t_k, X_k, u_N^k(X_k))\Delta_k W \\ \qquad +\dfrac{1}{2}\sigma(t_k, X_k, u_N^k(X_k))\{\sigma_x(t_k, X_k, u_N^k(X_k)) + \sigma_u(t_k, X_k, u_N^k(X_k))\partial_x u_N^k(X_k)\} \\ \qquad \times(\Delta_k^2 W - h), \qquad k = 0, 1, \ldots, M-1, \end{cases}$$

where $\Delta_k^2 W \stackrel{\triangle}{=} (W(t_{k+1}) - W(t_k))^2$.

The next two schemes require 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. [21]).

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 $k = 0, 1, \ldots, M-1$,

(3.12)
$$\begin{cases} X_0 = x, \\ X_{k+1} = X_k + bh + \sigma\Delta_k W + \dfrac{1}{2}\sigma\sigma'\{\Delta_k^2 W - h\} + b'\sigma\Delta_k\Psi + \dfrac{1}{2}\Big(bb' + \dfrac{1}{2}\sigma^2 b''\Big)h^2 \\ \qquad +\Big(b\sigma' + \dfrac{1}{2}\sigma^2\sigma''\Big)\{h\Delta_k W - \Delta_k W\} + \dfrac{1}{2}\sigma\Big(\sigma\sigma'' + (\sigma')^2\Big)\left\{\dfrac{1}{3}\Delta_k^2 W - h\right\}\Delta_k W, \end{cases}$$

where $\Delta_k^2 W \stackrel{\triangle}{=} (W(t_{k+1}) - W(t_k))^2$, $\Delta_k\Psi \stackrel{\triangle}{=} \int_{t_k}^{t_{k+1}}[W(t) - W(t_k)]dt$, $b = b(t_k, X_k, u_N^k(X_k))$, $\sigma = \sigma(t_k, X_k, u_N^k(X_k))$, and, for $\varphi = b$, σ,

$$\varphi' = \varphi_x + \varphi_u\partial_x u_N^k, \quad \varphi'' \stackrel{\triangle}{=} \varphi_{xx} + 2\varphi_{xu}\partial_x u_N^k + \varphi_{uu}(\partial_x u_N^k)^2 + \varphi_u\partial_x^2 u_N^k. \tag{3.13}$$

Here the partial derivatives of b and σ should be evaluated at $(t_k, X_k, u_N^k(X_k))$, and the approximate functions u_N^k, $\frac{du_N^k}{dx}$, and $\frac{d^2u_N^k}{dx^2}$ should be evaluated at point (X_k).

We should note that $\Delta_k\Psi = \int_{t_k}^{t_{k+1}} W(t) - W(t_k)dt \sim N(0, \frac{1}{3}h^3)$, and the covariance of $\Delta_k W$ and $\Delta_k\Psi$ is $\mathbb{E}\Big(\Delta_k W\Delta_k\Psi\Big) = \frac{1}{2}h^2$. The presence of the $\Delta_k\Psi$ 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. [16]).

It is easily seen from (3.13) that the Platen–Wagner $\frac{3}{2}$-order scheme requires that b and σ are both twice differentiable. Recently, Ewald and Temam [16] constructed a stochastic version of the Adam–Bashforth scheme which does not involve the second

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

derivatives of coefficient function b but still achieves the $\frac{3}{2}$-order convergence rate. This SAB scheme, adapted 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:

(3.14)
$$X_{k+2} = X_{k+1} + \frac{\Delta t}{2}[3b(t_{k+1}, X_{k+1}) - b(t_k, X_k)] - \frac{3}{2}\Delta t A_k(t_k, X_k) + B_k(t_k, X_k),$$

where

(3.15) $A_k(t, x) = [\sigma b_x](t, x)\Delta_k W,$

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

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

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

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

(3.16)
$$I_{(1,0)} = h\Delta_k W + \int_{t_{k+1}}^{t_{k+2}} [W(s) - W(t_{k+1})]ds,$$

and $I(1, 1)$, $I(1, 1, 1)$ are the iterated Itô integrals

(3.17)
$$I_{(1,1)} = \int_{t_k}^{t_{k+2}} \int_{t_k}^{t} dW(s)dW(t) - \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} dW(s)dW(t),$$

$$I_{(1,1,1)} = \int_{t_k}^{t_{k+2}} \int_{t_k}^{t} \int_{t_k}^{s} dW(r)dW(s)dW(t) - \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} \int_{t_k}^{s} dW(r)dW(s)dW(t).$$

To calculate the iterated Itô integrals $I_{(1,\ldots,1)}$, the useful formula of Itô [20] is often employed. More precisely, recall the scaled Hermite polynomials h_n defined by

$$h_n(x) = (-1)^n e^{\frac{1}{2}x^2} \frac{d^n}{dx^n}(e^{-\frac{1}{2}x^2}), \quad n = 0, 1, 2, \ldots .$$

Let $H_n^*(x, z) = z^{n/2}h_n(x/\sqrt{z})$. For any $g \in L^2([a, b])$, let $X_t = \int_a^t g(s)dW_s$. Applying Ito's formula repeatedly to $H_n^*(X_t, \langle X\rangle_t)$ from a to b and using the fact that $(\frac{\partial}{\partial_z} + 1/2\frac{\partial^2}{\partial_{x^2}})H_n^*(x, z) = 0$, we can derive

(3.18)
$$n! \int_a^b \int_a^{t_n} \cdots \int_a^{t_2} g(t_1)g(t_2)\cdots g(t_n)dW_{t_1}\cdots dW_{t_n} = \|g\|^n h_n\left(\frac{\theta}{\|g\|}\right),$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where

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

Thus the coefficients $I_{(1,1)}$ and $I_{(1,1,1)}$ in (3.17) can be calculated explicitly as

$$I_{(1,1)} = \frac{1}{2}\Big(\Delta_{k+1}^2 W + 2\Delta_{k+1}W\Delta_k W - h\Big), \tag{3.19}$$

$$I_{(1,1,1)} = \frac{1}{6}\Big((\Delta_{k+1}W + \Delta_k W)^3 - (\Delta_k W)^3 - 6h\Delta_{k+1}W - 3h\Delta_k W\Big).$$

Using the above relations, we can rewrite the $\frac{3}{2}$-order SAB scheme for the one-dimensional forward SDE (2.4) as

$$X_0 = x,$$

$$\begin{aligned} X_{k+2} = X_{k+1} &+ \frac{h}{2}\Big[3b(t_{k+1}, X_{k+1}, u_N^{k+1}(X_{k+1})) - b\Big] \\ &-\frac{3}{2}h\sigma b'\Delta_k W + \sigma\Delta_k W + \frac{1}{2}\sigma\sigma'\Big\{(\Delta_{k+1}W)^2 + 2(\Delta_{k+1}W)(\Delta_k W) - h\Big\} \\ &+\Big(\sigma_t + b\sigma' + \frac{1}{2}\sigma^2\sigma''\Big)\{2h\Delta_{k+1}W - \Delta_{k+1}\Psi\} + \sigma b'\{h\Delta_k W + \Delta_{k+1}\Psi\} \\ &+\frac{1}{6}\sigma\Big(\sigma\sigma'' + (\sigma')^2\Big)\Big\{(\Delta_{k+1}W + \Delta_k W)^3 - (\Delta_k W)^3 - 6h\Delta_{k+1}W - 3h\Delta_k W\Big\}, \end{aligned}$$

where $k = 0, 1, \ldots, M-1$. Similarly,

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

where the partial derivatives of b and σ 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

$$Y_k = u_N^k(X_k), \quad Z_k = \sigma(t_k, X_k, Y_k)\frac{\partial u_N^k}{\partial x}(X_k). \tag{3.20}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

4. Error analysis for the PDE approximation. In this section we carry out an error analysis for the approximation of the PDE (3.1). In order to simplify the presentation, we shall consider only the following *Hermite–Galerkin scheme* for (3.4).

Let $u_N^M(x) = P_N\varphi$. For $k = M-1, \dots, 0$, we compute $u_N^k \in \mathcal{H}_N$ such that

$$\left\langle \frac{u_N^{k+1} - u_N^k}{h}, v_N \right\rangle + \frac{1}{2}\left\langle \partial_x\Big(\sigma^2(t_k,\cdot,u_N^{k+1})(u_N^k)_x\Big), v_N\right\rangle + \langle \hat{b}(t_k,\cdot,u_N^{k+1},(u_N^{k+1})_x)\partial_x(u_N^{k+1}) + g(t_k,\cdot,u_N^{k+1}), v_N\rangle = 0 \qquad \forall v_N \in \mathcal{H}_N. \tag{4.1}$$

We note that the only difference between this Hermite–Galerkin scheme and the Hermite-collocation scheme (3.7) 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 differences between the two approaches are negligible (cf. [18]).

Hereafter, we shall use "$A \lesssim B$" to mean that there exists a constant $C > 0$ independent of N or h such that $A \le CB$.

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 $u \in C([0,T]; H_D^m(\mathbb{R}))$, that $u_t \in L^2(0,T;H_D^m(\mathbb{R}))$ with some $m > 1$, and that $u_{tt} \in L^2(0,T;H^{-1}(\mathbb{R}))$. Then, there exists $h_0 > 0$ such that for $h \le h_0$ the scheme* (3.4) *is stable, and the following error estimates hold: for each $k = 0, 1, \dots, M$,*

$$\|u(t_k,\cdot) - u_N^k\| + \left(h\sum_{j=k}^{M}\|\partial_x(u(t_j,\cdot) - u_N^j)\|^2\right)^{\frac{1}{2}} \lesssim h\Big(\|u_{tt}\|_{L^2(0,T;H^{-1}(\mathbb{R}))} + \|u_t\|_{L^2(0,T;H^1(\mathbb{R}))}\Big) + N^{\frac{1-m}{2}}\Big(\|u\|_{C([0,T];H_D^m)} + \|u_t\|_{L^2(0,T;H_D^m)}\Big).$$

Proof. Let E^k, $k = M-1, \dots, 1, 0$, be the consistency error defined by

$$E^k(\cdot) \triangleq \frac{1}{h}\Big(u(t_{k+1},\cdot) - u(t_k,\cdot)\Big) + \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\,\partial_x u(t_{k+1},\cdot) + \partial_x\left(\frac{1}{2}\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))\partial_x u(t_k,\cdot)\right) + g(t_k,\cdot,u(t_{k+1},\cdot)), \tag{4.2}$$

where $u(t_k,\cdot)$ is the exact solution to (3.1). First, we shall obtain the estimate for the consistency error E^k. By using the integral residue form of the Taylor series, and the fact that $u(t_k,\cdot)$ is the exact solution to (3.1), we can easily rewrite $E^k(\cdot)$ as

$$\begin{aligned} E^k(\cdot) = {} & \frac{1}{h}\int_{t_k}^{t_{k+1}} u_{tt}(t,\cdot)(t-t_k)dt \\ & + g(t_k,\cdot,u(t_{k+1},\cdot)) - g(t_k,\cdot,u(t_k,\cdot)) \\ & + \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\,\partial_x u(t_{k+1},\cdot) \\ & - \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\,\partial_x u(t_k,\cdot) \\ & + \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\,\partial_x u(t_k,\cdot) - \hat{b}(t_k,\cdot,u(t_k,\cdot),\partial_x u(t_k,\cdot))\,\partial_x u(t_k,\cdot) \end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$+\partial_x\left(\frac{1}{2}(\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))-\sigma^2(t_k,\cdot,u(t_k,\cdot)))\partial_x u(t_k,\cdot)\right).$$

Under assumptions (A1)–(A3) (see also Remark 2.2), it is known that the solution to (3.1) must have bounded first-order derivative $\partial_x u$; i.e., $\|\partial_x u\|_{L^\infty((0,T)\times\mathbb{R})}\le C$ (see, e.g., [10]). We can then derive from assumptions (A1)–(A3) (see also Remark 2.2) that

$$\begin{aligned}\|E^k(\cdot)\|_{-1}\lesssim{}&\frac{1}{h}\int_{t_k}^{t_{k+1}}\|u_{tt}(t,\cdot)\|_{-1}(t-t_k)dt+h^{\frac{1}{2}}\left(\int_{t_k}^{t_{k+1}}\|\partial_t u(t,\cdot)\|^2dt\right)^{\frac{1}{2}}\\&+h^{\frac{1}{2}}\left(\int_{t_k}^{t_{k+1}}\|\partial_{tx}u(t,\cdot)\|^2dt\right)^{\frac{1}{2}}\\&+\|(\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))-\sigma^2(t_k,\cdot,u(t_k,\cdot)))\partial_x u(t_k,\cdot)\|.\end{aligned}$$

Applying the Schwarz inequality and using the boundedness of $\partial_x u$, we find

$$\|E^k(\cdot)\|_{-1}^2\lesssim h\int_{t_k}^{t_{k+1}}\|u_{tt}(t,\cdot)\|_{-1}^2dt+h\int_{t_k}^{t_{k+1}}\|u_t(t,\cdot)\|_1^2dt,$$

which implies that

$$h\sum_{k=0}^{M}\|E^k(\cdot)\|_{-1}^2\lesssim h^2\|u_{tt}\|^2_{L^2(0,T;H^{-1}(\mathbb{R}))}+h^2\|u_t\|^2_{L^2(0,T;H^1(\mathbb{R}))}.\tag{4.3}$$

In what follows we set $\hat{e}_N^k=P_Nu(t_k,\cdot)-u_N^k$, $\tilde{e}_N^k=u(t_k,\cdot)-P_Nu(t_k,\cdot)$, and thus we have $e_N^k=u(t_k,\cdot)-u_N^k=\hat{e}_N^k+\tilde{e}_N^k$.

Next, multiplying by $v_N\in\mathcal{H}_N$ on both sides of (4.2), using integration by parts, and then subtracting (4.1), we obtain that

$$\begin{aligned}\langle E^k,v_N\rangle=\frac{1}{h}\langle e_N^{k+1}-e_N^k\,,v_N\rangle&-\left\langle\frac{1}{2}\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))\partial_x u(t_k,\cdot)\,,\partial_x v_N\right\rangle\\&+\left\langle\frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x u_N^k\,,\partial_x v_N\right\rangle\\&+\langle\hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot)))\,\partial_x u(t_{k+1},\cdot)\\&\quad-\hat{b}(t_k,\cdot,u_N^{k+1},\partial_x u_N^{k+1})\partial_x u_N^{k+1}\,,v_N\rangle\\&+\langle g(t_k,\cdot,u(t_{k+1},\cdot))-g(t_k,\cdot,u_N^{k+1})\,,v_N\rangle.\end{aligned}$$

Or, equivalently,

$$\begin{aligned}\langle E^k,v_N\rangle=\frac{1}{h}\langle e_N^{k+1}-e_N^k\,,v_N\rangle&-\left\langle\frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x e_N^k\,,\partial_x v_N\right\rangle\\&-\left\langle\partial_x u(t_k,\cdot)\,,\frac{1}{2}\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))\partial_x v_N-\frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x v_N\right\rangle\\&+\langle\hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\,\partial_x u(t_{k+1},\cdot)\\&\quad-\hat{b}(t_k,\cdot,u_N^{k+1},\partial_x u_N^{k+1})\partial_x u_N^{k+1}\,,v_N\rangle\\&+\langle g(t_k,\cdot,u(t_{k+1},\cdot))-g(t_k,\cdot,u_N^{k+1})\,,v_N\rangle.\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Since $e_N^k = \tilde{e}_N^k + \hat{e}_N^k$, we have

(4.4)
$$\begin{aligned}
&\frac{1}{h}\langle \hat{e}_N^{k+1} - \hat{e}_N^k, v_N\rangle - \left\langle \frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x \hat{e}_N^k \,,\partial_x v_N \right\rangle \\
&\qquad = -\frac{1}{h}\langle \tilde{e}_N^{k+1} - \tilde{e}_N^k \,, v_N\rangle + \left\langle \frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x \tilde{e}_N^k \,,\partial_x v_N \right\rangle \\
&\qquad + \left\langle \partial_x u(t_k,\cdot)\,, \frac{1}{2}\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))\partial_x v_N - \frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x v_N \right\rangle \\
&\qquad\quad - \langle \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\ \partial_x u(t_{k+1},\cdot) \\
&\qquad - \hat{b}(t_k,\cdot,u_N^{k+1},\partial_x u_N^{k+1})\partial_x u_N^{k+1}\ , v_N\rangle \\
&\qquad - \langle g(t_k,\cdot,u(t_{k+1},\cdot)) - g(t_k,\cdot,u_N^{k+1})\ , v_N\rangle + \langle E^k(\cdot)\ , v_N\rangle.
\end{aligned}$$

Denote

$$I_1 = -\langle g(t_k,\cdot,u(t_{k+1},\cdot)) - g(t_k,\cdot,u_N^{k+1})\ , v_N\rangle,$$

$$I_2 = -\langle \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\ \partial_x u(t_{k+1},\cdot) - \hat{b}(t_k,\cdot,u_N^{k+1},\partial_x u_N^{k+1})\partial_x u_N^{k+1}\ , v_N\rangle,$$

$$I_3 = \left\langle \partial_x u(t_k,\cdot)\,, \frac{1}{2}\sigma^2(t_k,\cdot,u(t_{k+1},\cdot))\partial_x v_N - \frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x v_N \right\rangle,$$

$$I_4 = \left\langle \frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x \tilde{e}_N^k \,,\partial_x v_N \right\rangle,$$

$$I_5 = -\frac{1}{h}\langle \tilde{e}_N^{k+1} - \tilde{e}_N^k \,, v_N\rangle,$$

$$I_6 = \langle E^k(\cdot)\ , v_N\rangle.$$

We shall estimate these terms separately under assumptions (A1)–(A3) (see also Remark 2.2). To begin with, noting the Lipschitz property of g and applying the Cauchy–Schwarz inequality, we have, for some constant $C_1 > 0$ and any $\varepsilon > 0$,

$$\begin{aligned}
|I_1| &\le C_1\langle |u(t_{k+1},\cdot) - u_N^{k+1}|\ , |v_N|\rangle = C_1\langle |e_N^{k+1}|\ , |v_N|\rangle \\
&\le \frac{C_1 h}{2\varepsilon}\|e_N^{k+1}\|^2 + \frac{C_1\varepsilon}{2h}\|v_N\|^2 \le \frac{C_1 h}{\varepsilon}\|\hat{e}_N^{k+1}\|^2 + \frac{C_1 h}{\varepsilon}\|\tilde{e}_N^{k+1}\|^2 + \frac{C_1\varepsilon}{2h}\|v_N\|^2.
\end{aligned} \tag{4.5}$$

Similarly, we have

$$|I_4| \le C_4\langle |\partial_x \tilde{e}_N^k|\ , |\partial_x v_N|\rangle \le \frac{C_4 h}{2\varepsilon}\|\partial_x \tilde{e}_N^k\|^2 + \frac{C_4\varepsilon}{2h}\|\partial_x v_N\|^2; \tag{4.6}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$$
\begin{aligned}
|I_2| &\le |\langle \hat{b}(t_k,\cdot,u(t_{k+1},\cdot),\partial_x u(t_{k+1},\cdot))\ \partial_x u(t_{k+1},\cdot) \\
&\qquad - \hat{b}(t_k,\cdot,u_N^{k+1},\partial_x u_N^{k+1})\partial_x u(t_{k+1},\cdot)\ , v_N\rangle| \\
&\quad + |\langle \hat{b}(t_k,\cdot,u_N^{k+1},\partial_x u_N^{k+1}))(\partial_x u(t_{k+1},\cdot) - \partial_x u_N^{k+1})\ , v_N\rangle| \\
&\le |\langle \|\partial_x u\|_{L^\infty((0,T)\times\mathbb{R})} K(|e_N^{k+1}| + |\partial_x e_N^{k+1}|),\ |v_N|\rangle| + K|\langle |\partial_x e_N^{k+1}|, |v_N|\rangle| \\
&\le C_2\langle |e_N^{k+1}| + |\partial_x e_N^{k+1}|, |v_N|\rangle + C_2\langle |\partial_x e_N^{k+1}|, |v_N|\rangle \qquad (4.7)\\
&\le C_2\langle |\hat{e}_N^{k+1}| + |\partial_x \hat{e}_N^{k+1}| + |\tilde{e}_N^{k+1}| + |\partial_x \tilde{e}_N^{k+1}|\ , |v_N|\rangle \\
&\quad + C_2\langle |\partial_x \hat{e}_N^{k+1}| + |\partial_x \tilde{e}_N^{k+1}|\ , |v_N|\rangle \\
&\le \frac{C_2 h}{\varepsilon}(\|\hat{e}_N^{k+1}\|^2 + \|\tilde{e}_N^{k+1}\|^2 + \|\partial_x \tilde{e}_N^{k+1}\|^2) \\
&\quad + \frac{C_2\varepsilon}{h}\|v_N\|^2 + C_2\varepsilon_1 h\|\partial_x \hat{e}_N^{k+1}\|^2 + \frac{C_2}{\varepsilon_1 h}\|v_N\|^2,
\end{aligned}
$$

where C_2 depends only on $K, \|\partial_x u\|_{L^\infty((0,T)\times\mathbb{R})}$ and ε_1 is to be determined later;

$$
\begin{aligned}
|I_3| &\le CK\ \|\partial_x u\|_{L^\infty((0,T)\times\mathbb{R})}\ \langle |e_N^{k+1}|\ , |\partial_x v_N|\rangle \le \frac{C_3 h}{2\varepsilon}\|e_N^{k+1}\|^2 + \frac{C_3\varepsilon}{2h}\|\partial_x v_N\|^2 \qquad (4.8)\\
&\le \frac{C_3 h}{\varepsilon}(\|\hat{e}_N^{k+1}\|^2 + \|\tilde{e}_N^{k+1}\|^2) + \frac{C_3\varepsilon}{2h}\|\partial_x v_N\|^2,
\end{aligned}
$$

where $C_3 = CK\ \|\partial_x u\|_{L^\infty((0,T)\times\mathbb{R})}$. Finally, let $\tilde{E}_N^k = -\frac{1}{h}(\tilde{e}_N^{k+1} - \tilde{e}_N^k)$; then we have

$$
|I_5| \le \frac{h}{2\varepsilon}\|\tilde{E}_N^k\|^2 + \frac{\varepsilon}{2h}\|v_N\|^2 \quad \text{and} \quad |I_6| \le \frac{h}{2\varepsilon}\|E^k(\cdot)\|_{-1}^2 + \frac{\varepsilon}{2h}\|v_N\|_1^2. \qquad (4.9)
$$

To obtain the desired estimate, let us now look at the left-hand side of (4.4) with $v_N = -h\ \hat{e}_N^k$. Note that

$$
\begin{aligned}
&\frac{1}{h}\langle \hat{e}_N^{k+1} - \hat{e}_N^k\ , -h\ \hat{e}_N^k\rangle = \frac{1}{2}(\|\hat{e}_N^k\|^2 - \|\hat{e}_N^{k+1}\|^2 + \|\hat{e}_N^k - \hat{e}_N^{k+1}\|^2), \\
&-\left\langle \partial_x \hat{e}_N^k\ , \frac{1}{2}\sigma^2(t_k,\cdot,u_N^{k+1})\partial_x v_N \right\rangle \ge ch\|\partial_x \hat{e}_N^k\|^2, \qquad (4.10)
\end{aligned}
$$

where $c > 0$ is a lower bound of σ^2, with a slight abuse of notation (compare to the constant $c > 0$ in (A3)). Thanks to the Schwarz inequality, and denoting $C = 5\max_{1\le i\le 4}\{C_i\} + 1$, we derive from (4.5)–(4.10) that

$$
\begin{aligned}
\|\hat{e}_N^k\|^2 + ch\|\partial_x \hat{e}_N^k\|^2 &\le \left(\frac{Ch}{\varepsilon} + \frac{1}{2}\right)\|\hat{e}_N^{k+1}\|^2 + C\varepsilon_1 h\|\partial_x \hat{e}_N^{k+1}\|^2 \qquad (4.11)\\
&\quad + \left(\frac{Ch}{\varepsilon_1} + C\varepsilon h + \varepsilon h + \frac{1}{2}\right)\|\hat{e}_N^k\|^2 \\
&\quad + C\varepsilon h\|\partial_x \hat{e}_N^k\|^2 + \frac{Ch}{\varepsilon}(\|\tilde{e}_N^{k+1}\|^2 + \|\partial_x \tilde{e}_N^{k+1}\|^2) \\
&\quad + \frac{Ch}{\varepsilon}\|\partial_x \tilde{e}_N^k\|^2 + \frac{h}{\varepsilon}\|E^k(\cdot)\|_{-1}^2 + \frac{h}{\varepsilon}\|\tilde{E}_N^k\|^2.
\end{aligned}
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Therefore, (4.11) can now be rewritten as

$$\begin{aligned}\left(\frac{1}{2}-h\left(C\varepsilon+\varepsilon+\frac{C}{\varepsilon_1}\right)\right)\|\hat{e}_N^k\|^2+(c-C\varepsilon)h\|\partial_x\hat{e}_N^k\|^2\\ \le\left(\frac{Ch}{\varepsilon}+\frac{1}{2}\right)\|\hat{e}_N^{k+1}\|^2+C\varepsilon_1 h\|\partial_x\hat{e}_N^{k+1}\|^2\\ +\frac{h}{\varepsilon}\Big(C\|\tilde{e}_N^{k+1}\|^2+C\|\partial_x\tilde{e}_N^k\|^2+C\|\partial_x\tilde{e}_N^{k+1}\|^2+\|\tilde{E}_N^k\|^2+\|E^k(\cdot)\|_{-1}^2\Big).\end{aligned}\tag{4.12}$$

We choose ε small enough so that $c-C\varepsilon>0$. Multiplying both sides of (4.12) by 2, choosing $\varepsilon_1=\frac{c-C\varepsilon}{2C}$, setting $C_5\triangleq C+1$, $C_\varepsilon\triangleq 2(C\varepsilon+\varepsilon+\frac{C}{\varepsilon_1})$, and assuming that the time step h satisfies the conditions

$$h<h_0=\min\left(\frac{1}{2C_\varepsilon},\frac{\varepsilon}{2C}\right),\tag{4.13}$$

inequality (4.12) becomes

$$\begin{aligned}(1-hC_\varepsilon)\|\hat{e}_N^k\|^2+2(c-C\varepsilon)h\|\partial_x\hat{e}_N^k\|^2\le\left(\frac{2Ch}{\varepsilon}+1\right)\|\hat{e}_N^{k+1}\|^2+(c-C\varepsilon)h\|\partial_x\hat{e}_N^{k+1}\|^2\\ +\frac{hC_5}{\varepsilon}\Big(\|\tilde{e}_N^{k+1}\|^2+\|\partial_x\tilde{e}_N^k\|^2+\|\partial_x\tilde{e}_N^{k+1}\|^2+\|\tilde{E}_N^k\|^2+\|E^k(\cdot)\|_{-1}^2\Big).\end{aligned}\tag{4.14}$$

We now apply a standard discrete backward Gronwall inequality and, noticing that $\hat{e}_N^M=0$, we obtain

$$\|\hat{e}_N^k\|^2+h\sum_{j=k}^M\|\partial_x\hat{e}_N^j\|^2\lesssim h\sum_{j=k}^M\Big(\|\tilde{e}_N^j\|^2+\|\partial_x\tilde{e}_N^j\|^2+\|\tilde{E}_N^j\|^2+\|E^j(\cdot)\|_{-1}^2\Big).\tag{4.15}$$

We now estimate the terms on the right-hand side. First, by virtue of Lemma 2.3 with $s=0,1$, respectively, and $r=m$, we have

$$\begin{aligned}\|\tilde{e}_N^{k+1}\|^2+\|\tilde{e}_N^k\|^2\le C_6N^{-m}(\|D^mu(t_{k+1},.)\|^2+\|D^mu(t_k,.)\|^2),\\ \|\partial_x\tilde{e}_N^{k+1}\|^2\le C_6N^{1-m}\|D^mu(t_{k+1},.)\|^2.\end{aligned}\tag{4.16}$$

It remains to estimate $\|\tilde{E}_N^k\|$. Let $\tilde{e}_N(t,\cdot)=u(t,\cdot)-P_Nu(t,\cdot)$. 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 (2.17) again, we obtain that

$$\|\tilde{E}_N^k\|^2\le\frac{1}{h}\int_{t_k}^{t_{k+1}}\|\partial_t\tilde{e}_N(t,\cdot)\|^2\,dt.\tag{4.17}$$

We then derive from the above and (4.15) that

$$\begin{aligned}\|\hat{e}_N^k\|^2+h\sum_{j=k}^M\|\partial_x\hat{e}_N^j\|^2\lesssim h^2\Big(\|u_{tt}\|^2_{L^2(0,T;H^{-1}(\mathbb{R}))}+\|u_t\|^2_{L^2(0,T;H^1(\mathbb{R}))}\Big)\\ +N^{1-m}\Big(\|u\|^2_{C([0,T];H_D^m)}+\|u_t\|^2_{L^2(0,T;H_D^m)}\Big),\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

proving the theorem. □

Remark 4.2. We would like to point out that one of the main purposes of Theorem 4.1 is to display a significant feature of the spectral method, that is, that the rate of convergence of the scheme increases as the smoothness of the solutions (i.e., m in the assumption of the above theorem) increases. Assumptions (A1)–(A3) are used only to guarantee the convergence of the numerical schemes proposed in this paper (in fact, all but the Platen–Wagner scheme where the second derivative of the drift b is required). The assumption on smoothness of the solution is used to prove the rate of convergence, which should always be true if the coefficients are sufficiently regular. We should note, however, that the explicit relation between the regularities of the coefficients and that of the solution seems to be a more subtle issue due to the quasi linearity of the PDE. We thus prefer making the exact smoothness requirement needed on the solution rather than on the coefficients.

We note also that by using a tensor product approach, the results in the above theorem can be directly extended to the high-dimensional case, although such a tensor product approach will become prohibitively expensive for space dimension ≥ 4.

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* (3.9)*. Define*

$$Y_k \triangleq u_N^k(X_k), \qquad Z_k \triangleq \sigma(t_k, X_k, Y_k)\partial_x(u_N^k)(X_k), \qquad k = 0, 1, \ldots, M, \tag{5.1}$$

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

$$\begin{aligned}\sup_{0\leq k\leq M} \mathbb{E}\Big[(X(t_k) - X_k)^2 + (Y(t_k) - Y_k)^2 + (Z(t_k) - Z_k)^2\Big]^{1/2} \\ \lesssim \sqrt{h} + h\Big(\|u_{tt}\|_{L^2(0,T;H^{-1}(\mathbb{R}))} + \|u_t\|_{L^2(0,T;L^2(\mathbb{R}))}\Big) \\ + N^{\frac{1-m}{2}}\Big(\|u\|_{C([0,T];H_D^m)} + \|u_t\|_{L^2(0,T;H_D^m)}\Big).\end{aligned} \tag{5.2}$$

Proof. Recall the intermediate Euler scheme (3.10):

$$\begin{cases} \hat{X}_0 = x, \\ \hat{X}_{k+1} = \hat{X}_k + b(t_k, \hat{X}_k, u^k(\hat{X}_k))h + \sigma(t_k, \hat{X}_k, u^k(\hat{X}_k))\Delta_k W, \end{cases} \tag{5.3}$$

where $u^k(x) = u(t_k, x)$ and $u(\cdot, \cdot)$ is the solution to the original PDE (3.1). Since the true solution u is at least uniformly Lipschitz under (A1)–(A3), we can apply the *fundamental convergence theorem* for SDEs (cf. [21] or [30]) to conclude that

$$\mathbb{E}|X(t_k) - \hat{X}_k|^2 \leq Ch \quad \forall\, k = 0, 1, \ldots, M. \tag{5.4}$$

Thus it remains to evaluate $\mathbb{E}|\hat{X}_k - X_k|^2$. To this end, note that

$$\begin{aligned}\hat{X}_{k+1} - X_{k+1} = \hat{X}_k - X_k + [b(t_k, \hat{X}_k, u(t_k, \hat{X}_k)) - b(t_k, X_k, u_N^k(X_k))]h \\ + [\sigma(t_k, \hat{X}_k, u(t_k, \hat{X}_k)) - \sigma(t_k, X_k, u_N^k(X_k))]\Delta_k W.\end{aligned} \tag{5.5}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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 $\mathbb{E}[(\Delta_k W)F] = 0$ and $\mathbb{E}[(\Delta_k W)F]^2 = h\mathbb{E}[F]^2$ for all $F \in L^2(\mathcal{F}_{t_k};\mathbb{R})$. Therefore, squaring both sides of (5.5), applying the Cauchy–Schwartz inequality, and then using the Lipschitz assumption on the coefficients b and σ, we obtain that

$$\begin{aligned}
&\mathbb{E}|\hat{X}_{k+1} - X_{k+1}|^2 \\
&\qquad \le (1+h)\mathbb{E}|\hat{X}_k - X_k|^2 + (h^2+h)\mathbb{E}|b(t_k,\hat{X}_k,u^k(\hat{X}_k)) - b(t_k,X_k,u_N^k(X_k))|^2 \\
&\qquad\quad + h\,\mathbb{E}|\sigma(t_k,\hat{X}_k,u^k(\hat{X}_k)) - \sigma(t_k,X_k,u_N^k(X_k))|^2 \qquad (5.6)\\
&\qquad \le (1+h+2K^2(h^2+h))\mathbb{E}|\hat{X}_k - X_k|^2 + 2K^2(h^2+h)\mathbb{E}|u^k(\hat{X}_k) - u_N^k(X_k)|^2.
\end{aligned}$$

Now, we write

$$u^k(\hat{X}_k) - u_N^k(X_k) = [u^k(\hat{X}_k) - u^k(X_k)] + [u^k(X_k) - u_N^k(X_k)] \qquad (5.7)$$

and note that

$$\mathbb{E}|u^k(\hat{X}_k) - u^k(X_k)|^2 \le K^2\mathbb{E}|\hat{X}_k - X_k|^2. \qquad (5.8)$$

We need only to estimate $\mathbb{E}|u^k(X_k) - u_N^k(X_k)|^2$. To this end, let us first establish the following simple inequality:

$$\|f\|_{L^\infty(\mathbb{R})} \le \|f\|_{L^2(\mathbb{R})} + \|f'\|_{L^2(\mathbb{R})} \qquad \forall\, f \in\ H^1(\mathbb{R}). \qquad (5.9)$$

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 $|f(x_n)| \le \|f\|_{L^2(n,n+1)} \le \|f\|_{L^2(\mathbb{R})}$. Hence, for any given $x \in [n, n+1)$, we have

$$|f(x)| \le |f(x) - f(x_n)| + |f(x_n)| = \left|\int_{x_n}^x f'(t)dt\right| + |f(x_n)| \le \|f\|_{L^2(\mathbb{R})} + \|f'\|_{L^2(\mathbb{R})}.$$

Thanks to (5.9), we obtain that

$$\begin{aligned}
\mathbb{E}|u^k(X_k) - u_N^k(X_k)|^2 &= \int_{-\infty}^{\infty} |u(t_k,x) - u_N^k(x)|^2 f_{X_k}(x)dx \\
&\le \|u(t_k,\cdot) - u_N^k(\cdot)\|^2_{L^\infty(\mathbb{R})} \int_{-\infty}^{\infty} f_{X_k}(x)dx \qquad (5.10)\\
&\le 2\|u(t_k,\cdot) - u_N^k\|^2 + 2\|\partial_x(u(t_k,\cdot) - u_N^k)\|^2,
\end{aligned}$$

where f_{X_k} is the probability density function of random variable X_k.

Combining (5.8)–(5.10), noting (5.7), and assuming without loss of generality that $h < 1$, it follows easily from (5.6) that, for $k = 1, 2, \ldots, M$,

$$\begin{aligned}
\mathbb{E}|\hat{X}_{k+1} - X_{k+1}|^2 &\le (1+Ch)\mathbb{E}|\hat{X}_k - X_k|^2 \\
&\quad + Ch\Big(\|u(t_k,\cdot) - u_N^k\|^2 + \|\partial_x(u(t_k,\cdot) - u_N^k)\|^2\Big).
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We now apply a standard forward discrete Gronwall inequality and, noticing that $\hat{X}_0 = X_0 = x$ and using Theorem 4.1, we obtain, for $k = 1, 2, \ldots, M$,

$$\mathbb{E}|\hat{X}_k - X_k|^2 \lesssim \mathbb{E}|\hat{X}_0 - X_0|^2 + h\sum_{j=1}^{k}\Big(\|u(t_j,\cdot) - u_N^j\|^2 + \|\partial_x(u(t_j,\cdot) - u_N^j)\|^2\Big)$$

$$\lesssim h^2\Big(\|u_{tt}\|^2_{L^2(0,T;H^{-1}(\mathbb{R}))} + \|u_t\|^2_{L^2(0,T;L^2(\mathbb{R}))}\Big)$$

$$+ N^{1-m}\Big(\|u\|^2_{C([0,T];H_D^m)} + \|u_t\|^2_{L^2(0,T;H_D^m)}\Big).$$

This, together with (5.4), yields the final mean-square error estimate:

$$\mathbb{E}|X(t_k) - X_k|^2 \lesssim h + h^2\Big(\|u_{tt}\|^2_{L^2(0,T;H^{-1}(\mathbb{R}))} + \|u_t\|^2_{L^2(0,T;L^2(\mathbb{R}))}\Big)$$

$$+ N^{1-m}\Big(\|u\|^2_{C([0,T];H_D^m)} + \|u_t\|^2_{L^2(0,T;H_D^m)}\Big).$$

The error estimates for $\mathbb{E}|Y(t_k) - Y_k|^2$ and $\mathbb{E}|Z(t_k) - Z_k|^2$ then follow easily from the relation (3.20), (5.1), and the estimate (5.11). The proof is now complete. ∎

Remark 5.2. From the proof of Theorem 5.1 we see that if we combine the higher-order schemes for both PDE (2.5) and the resulting decoupled SDE (2.4), then we will obtain a higher-order rate of convergence. The proof would be completely similar, but conceivably much more lengthy and tedious. We shall discuss this issue with extensive numerical examples in the next section.

Unlike the results in Theorem 4.1, the results in the above theorem cannot be directly extended to the high-dimensional case due (only) to the fact that the embedding (5.9) is no longer valid in a high space dimension.

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 Hermite–Gauss points $\{x_j\}_{j=0,1,\ldots,N}$, which are zeros of $H_{N+1}(x)$. We define $\hat{h}_j(x) = h_j(x)\frac{e^{-x^2/2}}{e^{-x_j^2/2}}$; then

$$\hat{h}_j(x_i) = \delta_{ij}, \quad \mathcal{H}_N = \text{span}\{\hat{h}_j : j = 0, 1, \ldots, N\},$$

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 $h_j^{'}(x_i)$ that

$$\hat{h}_j^{'}(x_i) = \frac{e^{-x_i^2/2}}{e^{-x_j^2/2}}(h_j^{'}(x_i) - x_i\delta_{ij}) = \begin{cases} \frac{\widehat{H}_N(x_i)}{(x_i - x_j)\widehat{H}_N(x_j)}, & i \neq j, \\ 0, & i = j. \end{cases}$$

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

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let us denote

$$\bar{u} = (u_N^k(x_0), u_N^k(x_1), \dots, u_N^k(x_N))^T, \quad \bar{f} = (f^{k+1}(x_0), f^{k+1}(x_1), \dots, f^{k+1}(x_N))^T$$

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

with $\alpha(x_i) = \frac{h}{2}\sigma^2(t_k, x_i, u_N^{k+1}(x_i))$. Then, (3.7) is reduced to the following linear system:

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

where

$$W = \mathrm{diag}(\hat{w}_0, \hat{w}_1, \dots, \hat{w}_N), \quad S = (s_{ij})_{i,j=0,1,\dots,N}.$$

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 [31], 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, $\frac{3}{2}$-order Platen–Wagner strong scheme, as well as the $\frac{3}{2}$-order stochastic SAB scheme, respectively.

Example (Milstein–Tretyakov [31]). Consider the following FBSDE:

(6.1)
$$\begin{cases} dX = \dfrac{X(1+X^2)}{(2+X^2)^3}dt + \dfrac{1+X^2}{2+X^2}\sqrt{\dfrac{1+2Y^2}{1+Y^2+\exp(-\frac{2X^2}{t+1})}}dW(t), \quad X(0) = x, \\ \\ dY = -g(t,X,Y)dt - f(t,X,Y)Zdt + ZdW(t), \quad Y(T) = \exp\left(-\dfrac{X^2(T)}{T+1}\right), \end{cases}$$

where

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

$$f(t,x,u) = \frac{x}{(2+x^2)^2}\sqrt{\frac{1+u^2+\exp(-\frac{2x^2}{t+1})}{1+2u^2}}\,.$$

Then the corresponding Cauchy problem has the following form, for $t < T, x \in \mathbb{R}$:

(6.2)
$$\begin{cases} 0 = \dfrac{\partial u}{\partial t} + \dfrac{1}{2}\left(\dfrac{1+x^2}{2+x^2}\right)^2 \dfrac{1+2u^2}{1+u^2+\exp(-\frac{2x^2}{t+1})}\dfrac{\partial^2 u}{\partial x^2} + \dfrac{2x(1+x^2)}{(2+x^2)^3}\dfrac{\partial u}{\partial x} \\ \\ \qquad + \dfrac{1}{t+1}\exp\left(-\dfrac{x^2}{t+1}\right)\left[\dfrac{4x^2(1+x^2)}{(2+x^2)^3} + \left(\dfrac{1+x^2}{2+x^2}\right)^2\left(1 - \dfrac{2x^2}{t+1}\right) - \dfrac{x^2}{t+1}\right], \\ \\ u(T,x) = \exp(-\frac{x^2}{T+1}). \end{cases}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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

$$u(t,x) = \exp\left(-\frac{x^2}{t+1}\right). \tag{6.3}$$

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

$$dX = \frac{X(1+X^2)}{(2+X^2)^3}dt + \frac{1+X^2}{2+X^2}dW(t), \quad X(0) = x. \tag{6.4}$$

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

$$\Lambda(z) + \arctan \Lambda(z) = z. \tag{6.5}$$

Differentiating (6.5) with respect to z, we get

$$\Lambda'(z) = \frac{1+\Lambda^2(z)}{2+\Lambda^2(z)} > 0, \quad \forall\ z \in \mathbb{R}, \quad \text{and} \quad \Lambda''(z) = \frac{2\Lambda(1+\Lambda^2)}{(2+\Lambda^2)^3}, \tag{6.6}$$

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

$$\begin{aligned} dX &= \Lambda'(x + \arctan x + W(t))dW(t) + \frac{1}{2}\Lambda''(x + \arctan x + W(t))dW(t) \\ &= \frac{X(1+X^2)}{(2+X^2)^3}dt + \frac{1+X^2}{2+X^2}dW(t). \end{aligned}$$

Hence, the solution to (6.1) is

$$\begin{aligned} X(t) &= \Lambda(x + \arctan x + W(t)), \\ Y(t) &= \exp\left(-\frac{X^2}{t+1}\right), \\ Z(t) &= -\frac{2X(1+X^2)}{(t+1)(2+X^2)} \exp\left(-\frac{X^2}{t+1}\right). \end{aligned}$$

We first carry out numerical tests on the parabolic equation (6.2) using the second-order (in time) scheme (3.8). In Tables 1–3, we tabulate the errors $\max_k \|u(t_k,x) - u_N^k(x)\|_{L^2(\mathbb{R})}$, $\max_k \|\frac{\partial u}{\partial x}(t_k,x) - \frac{d}{dx}u_N^k(x)\|_{L^2(\mathbb{R})}$, and $\max_k \|\frac{\partial^2 u}{\partial x^2}(t_k,x) - \frac{d^2}{dx^2}u_N^k(x)\|_{L^2(\mathbb{R})}$, 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 the first-order Milstein scheme (3.11) and $\frac{3}{2}$-order Platen–Wagner scheme (3.12) 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.

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 (3.7). For the sake of comparison with the results in [31], we

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

TABLE 1
$\max_k \|u(t_k,x) - u_N^k(x)\|_{L^2(\mathbb{R})}$.

$h\backslash N$	64	70	100	128	150	Rate
0.5	0.0049	0.0049	0.0049	0.0049	0.0049	
0.2	0.0010	0.0010	0.0010	0.0010	0.0010	1.7344
0.05	1.5391e-004	9.0229e-005	7.8519e-005	7.8554e-005	7.8555e-005	1.8351
0.02	1.5389e-004	8.3270e-005	1.3177e-005	1.3167e-005	1.3168e-005	1.9492
0.005	1.5388e-004	8.3266e-005	3.9669e-006	8.4446e-007	8.4360e-007	1.9822

TABLE 2
$\max_k \|\frac{\partial u}{\partial x}(t_k,x) - \frac{d}{dx}u_N^k(x)\|_{L^2(\mathbb{R})}$.

$h\backslash N$	64	70	100	128	150	Rate
0.5	0.0075	0.0074	0.0074	0.0074	0.0074	
0.2	0.0018	0.0017	0.0017	0.0017	0.0017	1.6052
0.05	0.0017	9.6178e-004	1.3741e-004	1.3685e-004	1.3685e-004	1.8174
0.02	0.0017	9.6179e-004	5.4862e-005	2.3165e-005	2.3152e-005	1.9391
0.005	0.0017	9.6178e-004	5.4859e-005	3.7980e-006	1.4925e-006	1.9777

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

$h\backslash N$	64	70	100	128	150	Rate
0.5	0.0230	0.0226	0.0224	0.0224	0.0224	
0.2	0.0188	0.0112	0.0055	0.0055	0.0055	1.5326
0.05	0.0188	0.0112	7.7070e-004	4.5192e-004	4.5174e-004	1.8029
0.02	0.0188	0.0112	7.7061e-004	7.8042e-005	7.7116e-005	1.9293
0.005	0.0188	0.0112	7.7058e-004	6.0488e-005	8.0385e-006	1.6310

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 4. The averages presented in the table are computed as follows:

$$E(X(T) - X_N)^2 = \frac{1}{S}\sum_{k=1}^{S}\left(X^{(k)}(T) - X_N^{(k)}\right)^2 \pm 2\sqrt{\frac{D_S}{S}},$$

where

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

Hence,

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

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $X^{(k)}(T)$ and $X_N^{(k)}$ 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* $N = 150$.

h	$\left[E(X(T) - X_N)^2\right]^{1/2}$	Rate	$\left[E(Y(T) - Y_N)^2\right]^{1/2}$	$\left[E(Z(T) - Z_N)^2\right]^{1/2}$
0.5	0.2453 ± 0.0121		0.0325 ± 0.0022	$0.0121 \pm 7.9404e-004$
0.2	0.1603 ± 0.0083	0.4643	0.0213 ± 0.0015	$0.0079 \pm 5.0520e-004$
0.05	0.0745 ± 0.0037	0.5527	0.0101 ± 0.0007	$0.0037 \pm 2.2556e-004$
0.02	0.0479 ± 0.0024	0.482	0.0064 ± 0.0004	$0.0023 \pm 1.4899e-004$
0.005	0.0259 ± 0.0013	0.4435	0.0036 ± 0.0002	$0.0013 \pm 7.6865e-005$

For conciseness, we now omit the Monte Carlo errors, which, as can be seen from Table 4, are significantly smaller 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 Tables 5–7, 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 $N = 150$.

h	X	Rate	Y	Rate	Z	Rate
0.5	0.0937		0.0131		0.0048	
0.2	0.0394	0.9455	0.0052	1.0084	0.0021	0.9022
0.05	0.0095	1.0261	0.0013	1.000	4.7994e-004	1.0647
0.02	0.0039	0.9717	5.3451e-004	0.9700	1.9852e-004	0.9634
0.005	9.5023e-004	1.0186	1.2996e-004	1.0201	4.8413e-005	1.0179

TABLE 6
$\frac{3}{2}$*-order strong scheme with* $N = 200$.

h	X	Rate	Y	Rate	Z	Rate
0.5	0.0425		0.0041		0.0025	
0.2	0.0115	1.4266	0.0015	1.0974	5.4051e-004	1.6714
0.05	0.0015	1.4693	1.7410e-004	1.5535	7.5482e-005	1.4201
0.02	3.3583e-004	1.6333	3.6401e-005	1.7080	1.7780e-005	1.5779
0.005	3.9962e-005	1.5355	7.6776e-006	1.1226	2.2442e-006	1.4930

TABLE 7
$\frac{3}{2}$*-order SAB scheme with* $N = 200$.

h	X	Rate	Y	Rate	Z	Rate
0.5	0.1301		0.0177		0.0063	
0.2	0.0400	1.2872	0.0053	1.3160	0.0020	1.2522
0.05	0.0048	1.5294	6.6087e-004	1.5018	2.3552e-004	1.5430
0.02	0.0013	1.4256	1.6757e-004	1.4975	6.5992e-005	1.3885
0.005	1.5749e-004	1.5226	2.0907e-005	1.5014	7.9073e-006	1.5305

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.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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 decoupling quasi-linear PDE on the whole space. The use of the 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, the numerical scheme, and most of the analysis, can be extended to higher-dimensional cases. More precisely, the Hermite-spectral method for the PDE can be extended to high-dimensional cases in a straightforward matter using a tensor-product approach. Furthermore, the results in Theorem 4.1 still hold for high-dimensional cases; unfortunately, it appears to be very difficult to establish the results in Theorem 5.1 for high-dimensional cases due mainly to the fact that the Sobolev-type inequality (5.9) is no longer valid for higher-dimensional cases. On the other hand, the aforementioned tensor-product approach is feasible only for two- or three-dimensional problems, as the number of unknowns grows exponentially fast as the dimension increases, and the computational cost quickly becomes prohibitive for problems of four or more dimensions. 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 in relatively large space dimensions.

Acknowledgment. The authors would like to thank an anoymous referee whose comments and suggestions have led to numerous improvements to the paper.

REFERENCES

[1] V. Bally, *Approximation scheme for solutions of BSDE*, in Backward Stochastic Differential Equations (Paris, 1995–1996), Pitman Res. Notes Math. Ser. 364, Longman, Harlow, UK, 1997, pp. 177–191.

[2] V. Bally and G. Pagès, *A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems*, Bernoulli, 9 (2003), pp. 1003–1049.

[3] V. Bally, G. Pagès, and J. Printems, *A quantization tree method for pricing and hedging multidimensional American options*, Math. Finance, 15 (2005), pp. 119–168.

[4] C. Bender and J. Zhang, *Time discretization and Markovian iteration for coupled FBSDEs*, Ann. Appl. Probab., 18 (2008), pp. 143–177.

[5] B. Bouchard and N. Touzi, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl., 111 (2004), pp. 175–206.

[6] P. Briand, B. Delyon, and J. Mémin, *Donsker-type theorem for BSDEs*, Electron. Comm. Probab., 6 (2001), pp. 1–14.

[7] H.-J. Bungartz and T. Dornseifer, *Sparse grids: Recent developments for elliptic partial differential equations*, in Multigrid Methods V (Stuttgart, 1996), Lect. Notes Comput. Sci. Eng. 3, Springer, Berlin, 1998, pp. 45–70.

[8] D. Chevance, *Numerical methods for backward stochastic differential equations*, in Numerical Methods in Finance, Publ. Newton Inst., Cambridge University Press, Cambridge, UK, 1997, pp. 232–244.

[9] J. Cvitanić and J. Zhang, *The steepest descent method for forward-backward SDEs*, Electron. J. Probab., 10 (2005), pp. 1468–1495.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[10] F. Delarue, *On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case*, Stochastic Process. Appl., 99 (2002), pp. 209–286.
[11] F. Delarue, *Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme*, in Séminaire de Probabilités XXXVII, Lecture Notes in Math. 1832, Springer, Berlin, 2003, pp. 290–332.
[12] F. Delarue and S. Menozzi, *A forward-backward stochastic algorithm for quasi-linear PDEs*, Ann. Appl. Probab., 16 (2006), pp. 140–184.
[13] J. Douglas, Jr., J. Ma, and P. Protter, *Numerical methods for forward-backward stochastic differential equations*, Ann. Appl. Probab., 6 (1996), pp. 940–968.
[14] N. El Karoui and L. Mazliak, eds., *Backward Stochastic Differential Equations* (Paris, 1995–1996), Pitman Res. Notes Math. Ser. 364, Longman, Harlow, UK, 1997.
[15] N. El Karoui, S. Peng, and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance, 7 (1997), pp. 1–71.
[16] B. D. Ewald and R. Témam, *Numerical analysis of stochastic schemes in geophysics*, SIAM J. Numer. Anal., 42 (2005), pp. 2257–2276.
[17] I. M. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed., Academic Press, San Diego, CA, 2000.
[18] B.-Y. Guo, J. Shen, and C.-L. Xu, *Spectral and pseudospectral approximations using Hermite functions: Application to the Dirac equation*, Adv. Comput. Math., 19 (2003), pp. 35–55.
[19] F. J. Hickernell, *Lattice rules: How well do they measure up?*, in Random and Quasi-Random Point Sets, Lecture Notes in Statist. 138, Springer, New York, 1998, pp. 109–166.
[20] K. Itô, *Multiple Wiener integral*, J. Math. Soc. Japan, 3 (1951), pp. 157–169.
[21] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.
[22] F. Y. Kuo, I. H. Sloan, and H. Woźniakowski, *Lattice rules for multivariate approximation in the worst case setting*, in Monte Carlo and Quasi-Monte Carlo Methods 2004, Springer, Berlin, 2006, pp. 289–330.
[23] J.-P. Lemor, E. Gobet, and X. Warin, *A regression-based Monte Carlo method for backward stochastic differential equations*, Ann. Appl. Probab., 15 (2005), pp. 2172–2202.
[24] J.-P. Lemor, E. Gobet, and X. Warin, *Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations*, Bernoulli, 12 (2006), pp. 889–916.
[25] D. Li and F. J. Hickernell, *Trigonometric spectral collocation methods on lattices*, in Recent Advances in Scientific Computing and Partial Differential Equations (Hong Kong, 2002), Contemp. Math. 330, AMS, Providence, RI, 2003, pp. 121–132.
[26] J. Ma, P. Protter, J. San Martin, and S. Torres, *Numerical method for backward stochastic differential equations*, Ann. Appl. Probab., 12 (2002), pp. 302–316.
[27] J. Ma, P. Protter, and J. Yong, *Solving forward-backward stochastic differential equations explicitly—a four step scheme*, Probab. Theory Related Fields, 98 (1994), pp. 339–359.
[28] J. Ma and J. Yong, *Forward-backward stochastic differential equations and their applications*, Lecture Notes in Math. 1702, Springer, Berlin, 1999.
[29] J. Ma and J. Zhang, *Representations and regularities for solutions to BSDEs with reflections*, Stochastic Process. Appl., 115 (2005), pp. 539–569.
[30] G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer, Berlin, 2004.
[31] G. N. Milstein and M. V. Tretyakov, *Numerical algorithms for forward-backward stochastic differential equations*, SIAM J. Sci. Comput., 28 (2006), pp. 561–582.
[32] É. Pardoux and S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, in Stochastic Partial Differential Equations and Their Applications (Charlotte, NC, 1991), Lecture Notes in Control and Inform. Sci. 176, Springer, Berlin, 1992, pp. 200–217.
[33] G. Szegő, *Orthogonal Polynomials*, 4th ed., Amer. Math. Soc. Colloq. Publ. XXIII, AMS, Providence, RI, 1975.
[34] B. Volker, N. Erich, and R. Klaus, *High dimensional polynomial interpolation on sparse grids*, Adv. Comput. Math., 12 (2000), pp. 273–288.
[35] J. Zhang, *A numerical scheme for BSDEs*, Ann. Appl. Probab., 14 (2004), pp. 459–488.
[36] Y. Zhang and W. Zheng, *Discretizing a backward stochastic differential equation*, Int. J. Math. Math. Sci., 32 (2002), pp. 103–116.

Downloaded 12/01/16 to 128.125.208.57. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.