The Annals of Applied Probability
1996, Vol. 6, No. 3, 940–968

NUMERICAL METHODS FOR FORWARD–BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

BY JIM DOUGLAS, JR.,[1] JIN MA[2] AND PHILIP PROTTER[3]

Purdue University

In this paper we study numerical methods to approximate the adapted solutions to a class of forward–backward stochastic differential equations (FBSDE's). The almost sure uniform convergence as well as the weak convergence of the scheme are proved, and the rate of convergence is proved to be as good as the approximation for the corresponding forward SDE. The idea of the approximation is based on the four step scheme for solving such an FBSDE, developed by Ma, Protter and Yong. For the PDE part, the combined characteristics and finite difference method is used, while for the forward SDE part, we use the first order Euler scheme.

1. Introduction. Let $(\Omega, \mathscr{F}, P; \{\mathscr{F}_t\}_{t\geq 0})$ be a filtered probability space satisfying the *usual conditions*. Assume that a standard d-dimensional Brownian motion $\{W_t\}_{t\geq 0}$ is defined on this space. We consider the following forward–backward stochastic differential equations (FBSDE's):

$$
\begin{aligned}
X_t &= x + \int_0^t b(s, X_s, Y_s, Z_s)\, ds + \int_0^t \sigma(s, X_s, Y_s)\, dW_s, \\
Y_t &= g(X_T) + \int_t^T \hat{b}(s, X_s, Y_s, Z_s)\, ds + \int_t^T \hat{\sigma}(s, X_s, Y_s, Z_s)\, dW_s,
\end{aligned}
\tag{1.1}
$$

where $t \in [0, T]$, (X, Y, Z) takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ and b, $\hat{b}$, σ, $\hat{\sigma}$ and g are smooth functions with appropriate dimensions; $T > 0$ is an arbitrarily prescribed number which stands for the *time duration*. By an "L^2-adapted solution" we mean a triple (X, Y, Z) which is $\{\mathscr{F}_t\}$-adapted and square integrable, such that the equations (1.1) are satisfied on $[0, T]$, P-almost surely. Such a stochastic differential equation has been found useful in applications, including stochastic control theory and mathematical finance (cf. [2], [7] and [8]). In previous work, Ma, Protter and Yong [12] studied the solvability of the adapted solution to the FBSDE; in particular, they designed a direct scheme, called the four step scheme (see Section 2 for a brief review), to solve the FBSDE explicitly.

Received May 1995; revised January 1996.
[1]Supported in part by NSF Grant DMS-92-07088.
[2]Supported in part by NSF Grant DMS-93-01516.
[3]Supported in part by NSF Grant INT-94-01109 and NSA Grant MDR 90494H2049.
AMS 1991 *subject classifications*. Primary 65U05; secondary 60H10, 65M06, 65M25.
Key words and phrases. Forward–backward stochastic differential equations, quasilinear parabolic equations, combined characteristics and finite difference method, Euler's scheme, weak convergence.

We note that in some applications the FBSDE (1.1) can be slightly simplified. That is, we may consider the FBSDE of the type

$$
\begin{aligned}
X_t &= x + \int_0^t b(s, X_s, Y_s)\, ds + \int_0^t \sigma(s, X_s, Y_s)\, dW_s, \\
Y_t &= E\left\{ g(X_T) + \int_t^T \hat{b}(s, X_s, Y_s)\, ds \,\middle|\, \mathscr{F}_t \right\},
\end{aligned}
\tag{1.2}
$$

where $t \in [0, T]$. Applying the usual technique using a martingale representation theorem, it is easily seen that (1.2) is equivalent to the FBSDE

$$
\begin{aligned}
X_t &= x + \int_0^t b(s, X_s, Y_s)\, ds + \int_0^t \sigma(s, X_s, Y_s)\, dW_s, \\
Y_t &= g(X_T) + \int_t^T \hat{b}(s, X_s, Y_s)\, ds + \int_t^T Z_s\, dW_s,
\end{aligned}
\tag{1.3}
$$

which is obviously a special form of (1.1). A first theoretical treatment of (1.2) can be found in Antonelli [1]. As a special case of (1.1), a more general treatment for (1.3) is contained in [12]. We note that the FBSDE's (1.2) and (1.3) have been found useful in the theory of mathematical finance. For instance, in the framework of stochastic recursive utility, the process X represents the "discounted weight process," or "wealth process," while the process Y defines a recursive differential utility (cf. [6] or [7] for more details). Also, in a model of term-structure of interest rates, Duffie, Ma, and Yong [8] considered a FBSDE of a form similar to (1.2), in which the process X is the short rate, while Y is the "consol rate" (or long term rate). Therefore, a satisfactory simulation result for the FBSDE (1.2) will have interest in its own right. In what follows we shall call (1.3) the "special case" and (1.1) the "general case."

For standard forward SDE's, there are two types of approximations typically considered: a pathwise convergence that typically converges at a rate $\mathscr{O}(1/\sqrt{n})$, and weak convergence to the terminal value $E\{f(X_T)\}$, where X is the true solution and f is an arbitrary smooth function. In the latter case, one approximates $E\{f(X_T)\}$ using a Monte Carlo technique once the law of X_T is known; thus it is the approximation of X_T that is needed, and since the Monte Carlo rate is slow, one is content to use a simple Euler scheme. We consider here both types of approximations for the forward–backward SDE's. Our technique allows the weak convergence to be a simple consequence of the pathwise convergence (which is not true in the usual forward case; note that its rate is faster). We obtain the same convergence rates as in the forward only case, an a priori surprising result.

It was shown in [12] that the solution, say θ, of a parabolic PDE plays a key role in solving FBSDE's; one uses θ to deduce a standard (forward) SDE which gives the component X. One then uses θ and X to obtain Y and Z. We have used this idea to construct a numerical scheme which first approximates θ using PDE numerical techniques, and then approximates X using SDE techniques. The two approximations have to mesh correctly, and the approximate solutions for θ have to have a certain regularity (e.g., Lipschitz property) so that the subsequent approximation for the forward SDE is feasible. It turns out that this can be done if the spatial mesh size h and the time mesh size Δt are essentially linearly related. In particular, we shall assume in this paper that *the condition* $h > C\,\Delta t$, *where* $C > 0$ *is a constant obtained from the coefficients, is fulfilled.*

For the PDE approximations we shall use a method combining the finite difference method and the method of characteristics; it was introduced earlier by Douglas and Russell [5] (see also Douglas [4]). This method allows us to treat the PDE in a more natural time-like variable and thus eliminate the first order term, which then facilitates an error analysis based on a maximum principle argument for the difference equations arising from the approximations. In practice, if the drift terms dominate the diffusion terms (i.e., in the so-called convection-dominated case), then this method will lead not to faster asymptotic rates but to smaller constants in the error estimate (cf., e.g., [5]), which can be just as (or more) important.

Since the special FBSDE's (1.2) and (1.3) are of independent interest, and the techniques of proofs and ideas are fundamental but more easily seen, we treat them separately in Section 4. We wish to point out that our techniques allow not only the approximation of (X, Y), but also that of the "extra" process Z that one needs to solve the FBSDE's in any sort of reasonable generality. This is significant because in some finance applications (for example), the process Z represents a hedging strategy, and thus we can give pathwise approximations of Z as well as weak (faster) approximations of $E\{f(X_T, Z_T)\}$. Again, these approximations are of the order $\mathscr{O}(1/\sqrt{n})$ and $\mathscr{O}(1/n)$, respectively, which are best possible for Euler schemes.

This paper is organized as follows. In Section 2 we formulate the problem and briefly review our four step scheme. In Section 3 we study the approximation for the quasilinear PDE arising in the special cases (1.2) and (1.3). In Section 4 we give our main result for the special case. In Section 5 we extend the results to the general case and give our final result.

2. Formulation of the problem. Let $(\Omega, \mathscr{F}, P)$ be a probability space carrying a standard d-dimensional Brownian motion $W = \{W_t: t \geq 0\}$ and let $\{\mathscr{F}_t\}$ be the σ-field generated by W (i.e., $\mathscr{F}_t = \sigma\{W_s: 0 \leq s \leq t\}$). We make the usual P-augmentation to each $\mathscr{F}_t$ so that $\mathscr{F}_t$ contains all the P-null sets of $\mathscr{F}$. Then $\{\mathscr{F}_t\}$ is right continuous and $\{\mathscr{F}_t\}$ satisfies the *usual hypotheses*. Let us consider the FBSDE (1.1). For the sake of simplicity, in what follows we will consider only the case in which $n = m = d = 1$, $\hat{\sigma}(t, x, y, z) \equiv z$ and σ is independent of Z. In other words, we content ourselves with an FBSDE that is slightly less general than (1.1) [but more general than (1.2)]:

$$
\begin{aligned}
X_t &= x + \int_0^t b(s, X_s, Y_s, Z_s)\,ds + \int_0^t \sigma(s, X_s, Y_s)\,dW_s,\\
Y_t &= g(X_T) + \int_t^T \hat{b}(s, X_s, Y_s, Z_s)\,ds + \int_t^T Z_s\,dW_s,
\end{aligned}
\tag{2.1}
$$

where $t \in [0,T]$. Here, X, Y and Z are now real-valued processes and b, $\hat{b}$, σ and g are real-valued functions. We note that the numerical study of the FBSDE of type (1.1) is also possible using our method, in view of the general theory for the four step scheme (see [12]), but some more complicated discussion involving the numerical scheme for inverse functions will be needed. To simplify presentation, we prefer not to include such a case in the present paper.

We first give the precise definition of an L^2-*adapted solution* to (2.1).

DEFINITION 2.1. A triple of processes (X,Y,Z): $[0,T] \times \Omega \to \mathbb{R}^3$ is called an L^2-*adapted solution* of the forward–backward SDE (2.1) if it is $\{\mathscr{F}_t\}$-adapted and square integrable and is such that it satisfies (2.1) almost surely.

Let us recall a standard Hölder space notation. For any bounded or unbounded region $G \subseteq \mathbb{R}$, $T > 0$ and $\alpha \in (0,1)$, we define $C^{1+\alpha/2,2+\alpha}([0,T] \times G)$ to be the space of all functions $\varphi(t,x)$ which are differentiable in t and twice differentiable in x with φ_t and φ_{xx} being $\alpha/2$- and α-Hölder continuous in $(t,x) \in [0,T] \times G$. The norm in $C^{1+\alpha/2,2+\alpha}([0,T] \times G)$ is defined by

$$\begin{aligned}\|\varphi\|_{1,2,\alpha;T,G} &= \|\varphi\|_{C_{T,G}} + \|\varphi_t\|_{C_{T,G}} + \|\varphi_x\|_{C_{T,G}} + \|\varphi_{xx}\|_{C_{T,G}} \\ &\quad + \sup_{(t,x)\neq(t',x')} \frac{|\varphi_t(t,x) - \varphi_t(t',x')| + |\varphi_{xx}(t,x) - \varphi_{xx}(t',x')|}{\left(|x-x'|^2 + |t-t'|\right)^{\alpha/2}},\end{aligned}$$

where $\|\cdot\|_{C_{T,G}}$ is the usual sup-norm on the closure of $[0,T] \times G$. When $G = \mathbb{R}$, we set $C^{1+\alpha/2,2+\alpha}([0,T] \times \mathbb{R}) = C^{1+\alpha/2,2+\alpha}$ and $\|\cdot\|_{1,2,\alpha;T,\mathbb{R}} = \|\cdot\|_{1,2,\alpha}$. For functions of the type $\varphi = \varphi(x)$, we define the space $C^{k+\alpha}(G)$ and $C^{k+\alpha} = C^{k+\alpha}(\mathbb{R})$ analogously, for $k = 1,2,\dots$.

We will make use of the following standing assumptions throughout the paper.

Standing assumptions.

(A1) The functions b, $\hat{b}$ and σ are continuously differentiable in t and twice continuously differentiable in x, y, z. Moreover, if we denote any one of these functions generically by ψ, then there exists a constant $\alpha \in (0,1)$, such that for fixed y and z, $\psi(\cdot,\cdot,y,z) \in C^{1+\alpha/2,2+\alpha}$. Furthermore, for some $L > 0$,

$$\|\psi(\cdot,\cdot,y,z)\|_{1,2,\alpha} \le L, \qquad \forall\,(y,z) \in \mathbb{R}^2.$$

(A2) The function σ satisfies

$$\mu \le \sigma(t,x,y) \le C, \qquad \forall\,(t,x,y) \in [0,T] \times \mathbb{R}^2, \tag{2.2}$$

where $0 < \mu \le C$ are two constants.

(A3) The function g belongs boundedly to $C^{4+\alpha}$ for some $\alpha \in (0,1)$ [one may assume that α is the same as that in (A1)].

REMARK 2.2. We should note here that the standing assumptions above are actually much stronger than those in [12], where the FBSDE was shown to be solvable, and thus they may not be optimal.

We now briefly review our four step scheme (see [12] for complete details).

The four step scheme.

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

$$z(t,x,y,p) = -p\sigma(t,x,y), \qquad \forall\,(t,x,y,p). \tag{2.3}$$

STEP 2. Using the function z above, solve the following quasilinear parabolic equation for $\theta(t,x)$ in $C^{1+\alpha/2,2+\alpha}$, for some $0<\alpha<1$:

$$\begin{aligned} &\theta_t + \tfrac{1}{2}\sigma(t,x,\theta)^2\theta_{xx} + b(t,x,\theta,z(t,x,\theta,\theta_x))\theta_x \\ &\quad + \hat{b}(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \qquad (t,x) \in (0,T)\times\mathbb{R}, \\ &\qquad\qquad \theta(T,x) = g(x), \qquad x \in \mathbb{R}. \end{aligned} \tag{2.4}$$

STEP 3. Using θ and z, solve the forward SDE

$$X_t = x + \int_0^t \tilde{b}(s,X_s)\,ds + \int_0^t \tilde{\sigma}(s,X_s)\,dW_s, \tag{2.5}$$

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

STEP 4. Set

$$\begin{aligned} Y_t &= \theta(t,X_t), \\ Z_t &= z(t,X_t,\theta(t,X_t),\theta_x(t,X_t)). \end{aligned} \tag{2.6}$$

Then, if this scheme is realizable, (X_t, Y_t, Z_t) will give an adapted solution of (2.1). In fact, in [12] it was proved that under reasonable conditions, the four step scheme is feasible. We summarize the results there in the following theorem, with modifications made to suit our future discussion. Since the arguments are standard, we give only a sketch of the proof.

THEOREM 2.3. *Suppose that the standing assumptions* (A1)–(A3) *hold. Then, the four step scheme defined above is applicable and any adapted solution to the FBSDE* (2.1) *must be the same as the one constructed from the four step scheme. Consequently, FBSDE* (2.1) *possesses a unique adapted solution.*

Furthermore, the unique classical solution θ to the quasilinear PDE (2.4) *belongs to the space $C^{2+\alpha/2, 4+\alpha}$, and all the partial derivatives of θ up to the second order in t and fourth order in x are bounded by a constant $K > 0$.*

SKETCH OF THE PROOF. The first assertion is a direct consequence of the results in [12]. To see the second assertion, note that by the result in [12] we know that the PDE (2.4) has a unique classical solution $\theta \in C^{1+\alpha/2, 2+\alpha}$ for some $\alpha \in (0,1)$. If we apply standard techniques (cf. [9] or [11]) using parabolic Schauder interior estimates to the difference quotients repeatedly, it is not hard to show that under our regularity and boundedness assumptions on the coefficients b, σ, $\hat{b}$ and g, one can improve the regularity of the solution to the desired order. □

3. Approximation of the PDE (2.4)—special case. In this section we study the numerical approximation scheme and its convergence analysis for the quasilinear parabolic PDE (2.4) corresponding to the *special* FBSDE (1.2), or equivalently, (1.3). We shall be interested in finding a *strong* approximation scheme which produces an approximate solution $(X^{(n)}, Y^{(n)})$ such that

$$E\left\{\sup_{0\le t\le T}\left|X_t^{(n)} - X_t\right|^2\right\} + E\left\{\sup_{0\le t\le T}\left|Y_t^{(n)} - Y_t\right|^2\right\} \to 0$$

and in determining its rate of convergence. Note that in this case the coefficients b, $\hat{b}$ and σ are independent of Z and only the (X, Y) part of the adapted solution need be considered; thus, the difficulty of the problem is reduced considerably. More precisely, in this case the corresponding PDE (2.4) now takes the simpler form

$$\begin{aligned} &\theta_t + \tfrac{1}{2}\sigma(t,x,\theta)^2\theta_{xx} + b(t,x,\theta)\theta_x + \hat{b}(t,x,\theta) = 0, \\ &\qquad\qquad\qquad\qquad (t,x) \in (0,T)\times\mathbb{R}, \\ &\theta(T,x) = g(x), \qquad x \in \mathbb{R}. \end{aligned} \tag{3.1}$$

We shall follow an idea of Douglas and Russell [5] and Douglas [4] that combines the method of characteristics with a finite difference procedure to design the approximation scheme. We discussed the advantages this brings to bear in the introduction.

Let us first standardize the PDE (3.1). Define $u(t,x) = \theta(T-t, x)$ and

$$\begin{aligned} \bar{\sigma}(t,x,y) &= \sigma(T-t,x,y), \\ \bar{b}(t,x,y) &= b(T-t,x,y), \\ \bar{\hat{b}}(t,x,y) &= \hat{b}(T-t,x,y). \end{aligned}$$

Then u satisfies the PDE

$$\begin{aligned} &u_t - \tfrac{1}{2}\bar{\sigma}^2(t,x,u)u_{xx} - \bar{b}(t,x,u)u_x - \bar{\hat{b}}(t,x,u) = 0, \\ &u(0,x) = g(x). \end{aligned} \tag{3.2}$$

To simplify notation we replace $\bar{\sigma}$, $\bar{b}$ and $\bar{\hat{b}}$ by σ, b and $\hat{b}$ themselves in the rest of this section. Following [5], we should first determine the characteristics of the first order nonlinear PDE

$$u_t - b(t, x, u)u_x = 0. \tag{3.3}$$

After transforming (3.3) into a first order system, it is not hard to show that the characteristic of (3.3) is given by the equation

$$\det\left| a_{ij}t'(s) - \delta_{ij}x'(s)\right| = 0, \tag{3.4}$$

where s is the parameter of the characteristic and (a_{ij}) is the matrix

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -b(t, x, u) & 0 \\ 0 & -1 & 0 \end{bmatrix}. \tag{3.5}$$

Therefore, (3.4) leads to

$$t'(s)b(t, x, u) - x'(s) = 0, \tag{3.6}$$

where (t, x, u) is evaluated along the characteristic curve $\mathscr{C}$: $(t(\cdot), x(\cdot))$. We replace the parameter of $\mathscr{C}$ by t and denote the arc length along $\mathscr{C}$ by τ. Then,

$$d\tau = \left[1 + b^2(t, x, u(t, x))\right]^{1/2} dt; \tag{3.7}$$

along $\mathscr{C}$,

$$\frac{\partial}{\partial \tau} = \frac{1}{\psi}\left\{\frac{\partial}{\partial t} - b\frac{\partial}{\partial x}\right\},$$

where

$$\psi(t, x) = \left[1 + b^2(t, x, u(t, x))\right]^{1/2}. \tag{3.8}$$

Thus, the equation (3.2) can be simplified to

$$\psi\frac{\partial u}{\partial \tau} = \frac{1}{2}\sigma^2(t, x, u)u_{xx} + \hat{b}(t, x, u), \qquad u(0, x) = g(x). \tag{3.9}$$

We shall design our numerical scheme based on (3.9).

Numerical scheme. Let $h > 0$ and $\Delta t > 0$ be fixed numbers. Let $x_i = ih$, $i = 0, \pm 1, \dots$ and $t^k = k\,\Delta t$, $k = 0, 1, \dots, N$, where $t^N = T$. For a function $f(t, x)$, let $f^k(\cdot) = f(t^k, \cdot)$ and let $f_i^k = f(t^k, x_i)$ denote the grid value of the function f. Define for each k the approximate solution w^k by the following recursive steps.

STEP 0. Set $w_i^0 = g(x_i)$, $i = \dots, -1, 0, 1, \dots$; use linear interpolation to obtain a function $w^0(x)$ defined on $x \in \mathbb{R}$.

Suppose that $w^{k-1}(x)$ is defined for $x \in \mathbb{R}$, let $w_i^{k-1} = w^{k-1}(x_i)$ and

$$\begin{aligned} &b_i^k = b(t^k, x_i, w_i^{k-1}), \qquad \sigma_i^k = \sigma(t^k, x_i, w_i^{k-1}), \\ &\hat{b}_i^k = \hat{b}(t^k, x_i, w_i^{k-1}), \qquad \bar{x}_i^k = x_i - b_i^k \Delta t, \qquad \overline{w}_k^{k-1} = w^{k-1}(\bar{x}_i^k), \\ &\delta_x^2(w)_i^k = h^{-1}[w_{i+1}^k - 2w_i^k + w_{i-1}^k]. \end{aligned} \tag{3.10}$$

STEP k. Obtain the grid values for the kth step approximate solution, denoted by $\{w_i^k\}$, via the difference equation

$$\frac{w_i^k - \overline{w}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(w)_i^k + (\hat{b})_i^k, \qquad -\infty < i < \infty. \tag{3.11}$$

Since by our assumption σ is bounded below positively and $\hat{b}$ and g are bounded, there exists a unique *bounded* solution of (3.11) as soon as an evaluation is specified for $w^{k-1}(x)$.

Finally, we use *linear interpolation* to extend the grid values of $\{w_i^k\}_{i=-\infty}^{\infty}$ to all $x \in \mathbb{R}$ to obtain the kth step approximate solution $w^k(\cdot)$.

REMARK 3.1. One must be careful with a numerical scheme to show that it does converge to the unique bounded solution (3.11) as desired and not to a (non-unique) unbounded solution. The following "localization" argument is essential, both theoretically and computationally, for this purpose, and it is also important to apply the maximum principle argument in Theorem 3.3. First, we note that the classical solution of the Cauchy problem (3.2) [therefore (3.9)] is actually viewed as the uniform limit of the solutions $\{u^R\}$ $(R \to \infty)$ to the *initial-boundary* problems (cf. [11] or [12]):

$$\begin{aligned} &u_t - \tfrac{1}{2}\bar{\sigma}(t,x,u)^2 u_{xx} - \bar{b}(t,x,u)u_x - \bar{\bar{b}}(t,x,u) = 0, \\ &u(0,x) = g(x), \qquad x \in \mathbb{R}; \\ &u(t,x) = g(x), \qquad |x| = R, \qquad 0 < t \le T. \end{aligned} \tag{3.2'}$$

Also, the (unique) bounded solution to (3.11) is derived in a similar way: by the boundedness assumption on the coefficients b and $\hat{b}$ and an assumption that σ is bounded from below by a positive constant, one can show that $\{w_i^k\}$ is the uniform limit of $w_i^{i_0,k}$, $i = -i_0, \dots, i_0$, $k = 0, 1, 2, \dots,$ where $\{w_i^{i_0,k}\}$ is the solution to the initial-boundary problem

$$\begin{aligned} &\frac{w_i^k - \overline{w}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(w)_i^k + (\hat{b})_i^k; \qquad -i_0 \le i \le i_0, \\ &w_i^0 = g(x_i), \qquad -i_0 \le i \le i_0, \\ &w_{\pm i_0}^k = g(x_{\pm i_0}), \qquad k = 0, 1, 2, \dots . \end{aligned} \tag{3.11'}$$

Therefore, if for fixed mesh size $h > 0$ we choose $R = i_0 h$, for some $i_0 = i_0(h)$, then $R \to \infty$ as $i_0 \to \infty$, and, uniformly as $i_0 \to \infty$ (possibly along a subsequence), we have

$$u(t,x) = \lim_{i_0\to\infty} u^R(t,x), \quad \text{uniformly in } (t,x),$$

$$w_i^k = \lim_{i_0\to\infty} w_i^{i_0,k}, \quad \text{uniformly in } i,k.$$

Consequently, we see that the quantities

$$\max_i \big|u(t^k,x_i) - w_i^k\big| \quad \text{and} \quad \max_{-i_0\le i\le i_0} \big|u^R(t^k,x_i) - w_i^{i_0,k}\big|$$

differ only by a (uniform) error $\varepsilon > 0$, which can be taken to be arbitrarily small as $i_0 h$ is sufficiently large. Because of this, in what follows we sometimes use the localized solutions when necessary without further specification. Note that if the localized solutions are used, then the error $|u^R(t^k, x_{\pm i_0}) - w_{\pm i_0}^{i_0,k}| \equiv 0$ for all $k = 0,1,2,\dots$. Therefore the maximum absolute value of the error $|u^R(t^k,x_i) - w_i^{i_0,k}|$, $i = -i_0,\dots,i_0$, will always occur in an "interior" point of $(-R,R)$, which will be essential in the maximum principle argument used in Theorem 3.3.

To analyze the convergence of the approximation, we need to derive an error equation for the procedure. First, note that along the characteristic curve $\mathscr{C}$,

$$\begin{aligned}
\psi\frac{\partial u}{\partial\tau} &\approx \psi\frac{u(t^k,x) - u(t^{k-1},\bar{x})}{\Delta\tau} \\
&\approx \psi(x)\frac{u(t^k,x) - u(t^{k-1},\bar{x})}{\left[(x-\bar{x})^2 + (\Delta t)^2\right]^{1/2}} \\
&= \frac{u(t^k,x) - u(t^{k-1},\bar{x})}{\Delta t},
\end{aligned}$$

where $\bar{x}$ is the location of the characteristic starting from (t^k,x) at $t = t^{k-1}$. Therefore, the solution of (3.9) satisfies a difference equation of the form

$$\frac{u_i^k - \bar{u}_i^{k-1}}{\Delta t} = \frac{1}{2}\big(\sigma(u)_i^k\big)^2 \delta_x^2(u)_i^k + \hat{b}(u)_i^k + e_i^k, \tag{3.12}$$

$$-\infty < i < \infty,\ k = 1,\dots,N,$$

where $\bar{u}_i^{k-1} = u^{k-1}(\bar{\mathbf{x}}_i^k)$ and $\bar{\mathbf{x}}_i^k$ is an approximation of $\bar{x}$ [see the definition following (3.17) below]; $\hat{b}(u)_i^k$ and $\sigma(u)_i^k$ correspond to $\hat{b}_i^k$ and σ_i^k defined in (3.10), except that the values $\{w_i^{k-1}\}$ are replaced by $\{w_i^{k-1}\}$; e_i^k is the error term to be estimated.

In order to estimate the error terms $\{e_i^k\}$, we first observe that at each grid point (t^k,x_i),

$$\psi(t^k,x_i)\frac{\partial u}{\partial\tau}\bigg|_{(t^k,x_i)} = \frac{1}{2}\sigma^2(t^k,x_i,u_i^k)u_{xx}\big|_{(t_k,x_i)} + \hat{b}(t^k,x_i,u_i^k). \tag{3.13}$$

Therefore,

$$
\begin{aligned}
e_i^k &= \left\{ \frac{u_i^k - \bar{u}_i^{k-1}}{\Delta t} - \psi(t^k, x_i) \frac{\partial u}{\partial \tau}\bigg|_{(t^k,\, x_i)} \right\} \\
&\quad + \left\{ \frac{1}{2}\sigma^2(t^k, x_i, u_i^k) u_{xx}|_{(t^k,\, x_i)} - \frac{1}{2}\big(\sigma(u)_i^k\big)^2 \delta^2(u)_i^k \right\} \\
&\quad + \left\{ \hat{b}(t^k, x_i, u_i^k) - \hat{b}(u)_i^k \right\} \\
&= I_i^{1,k} + I_i^{2,k} + I_i^{3,k}, \qquad -\infty < i < \infty,\ k = 1, \dots, N.
\end{aligned} \tag{3.14}
$$

We have the following lemma.

LEMMA 3.2. *There exists a constant $C_1 > 0$, depending only on b, $\hat{b}$, σ, u and T, such that for all $k = 0, \dots, N$ and $-\infty < i < \infty$,*

$$|e_i^k| \le C_1(h + \Delta t).$$

PROOF. We shall estimate $I_i^{1,k}$, $I_i^{2,k}$ and $I_i^{3,k}$ separately. Note that, by Theorem 2.3, the partial derivatives u_t, u_x, u_{xx} and u_{xxx} are uniformly bounded. Hence, it is easy to see from the uniform Lipschitz conditions on $\hat{b}$ that

$$
\begin{aligned}
|I_i^{3,k}| &= \left| \hat{b}(t^k, x_i, u_i^k) - \hat{b}(t^k, x_i, u_i^{k-1}) \right| \\
&\le \|\hat{b}_u\|_\infty \left| u(t^k, x_i) - u(t^{k-1}, x_i) \right| \le C^{1,1} \Delta t,
\end{aligned} \tag{3.15}
$$

where $C^{1,1} := \|\hat{b}_u\|_\infty \|u_t\|_\infty < \infty$. Similarly,

$$
\begin{aligned}
|I_i^{2,k}| &\le \frac{1}{2}\Bigg\{ \left| \sigma^2(t^k, x_i, u_i^k) - \sigma^2(t^k, x_i, u_i^{k-1}) \right| \left| u_{xx}(t^i, x_i) \right| \\
&\quad + \left| \sigma^2(t^k, x_i, u_i^{k-1}) \right| \left| u_{xx}(t^k, x_i) - \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} \right| \Bigg\} \\
&\le \frac{1}{2}\left\{ 2\|\sigma\|_\infty \|\sigma_u\|_\infty \|u_t\|_\infty \|u_{xx}\|_\infty \Delta t + \|\sigma\|_\infty^2 \|u_{xxx}\|_\infty h \right\} \\
&\le C^{1,2}(h + \Delta t),
\end{aligned} \tag{3.16}
$$

where $C^{1,2} := \max\{2\|\sigma\|_\infty \|\sigma_u\|_\infty \|u_t\|_\infty \|u_{xx}\|_\infty, \|\sigma\|_\infty^2 \|u_{xxx}\|_\infty\} < \infty$. Thus, it remains to estimate $I_i^{1,k}$. For each k and i, set

$$\bar{\mathbf{x}}_i^k := x_i - b(t^k, x_i, u_i^{k-1}) \Delta t, \qquad \bar{u}_i^{k-1} := u(t^{k-1}, \bar{\mathbf{x}}_i^k). \tag{3.17}$$

Let $\{x(t)\colon t^{k-1} \le t \le t^k\}$ be the characteristic such that $x(t^k) = x_i$. Since

$$x(t) = x_i - \int_t^{t^k} b(s, x(s), u(s, x(s)))\, ds, \qquad t^{k-1} \le t \le t^k,$$

we have $\sup_{t^{k-1}\le t\le t^k}|x(t)-x_i \le \|b\|_\infty \Delta t$. Therefore, denoting $\bar{x}=x(t^{k-1})$, we obtain easily that

$$\begin{aligned} |\bar{\mathbf{x}}_i^k-\bar{x}| &\le \int_{t^{k-1}}^{t^k}\left|b\left(t^k,x_i,u_i^{k-1}\right)-b(t,x(t),u(t,x(t)))\right|dt \\ &\le \left\{\|b_t\|_\infty+\|b_x\|_\infty\|b\|_\infty+\|b_u\|_\infty\left(\|u_t\|_\infty+\|u_x\|_\infty\|b\|_\infty\right)\right\}\Delta t^2 \\ &\le C^{1,3}\Delta t^2, \end{aligned} \tag{3.18}$$

where $C^{1,3}=\|b_t\|_\infty+\|b_x\|_\infty\|b\|_\infty+\|b_u\|_\infty(\|u_t\|_\infty+\|u_x\|_\infty\|b\|_\infty)$. Thus,

$$\left|\frac{u(t^{k-1},\bar{x})-u\left(t^{k-1},\bar{\mathbf{x}}_i^k\right)}{\Delta t}\right|\le\frac{\|u_x\|_\infty|\bar{x}-\bar{\mathbf{x}}_i^k|}{\Delta t}\le C^{1,4}\Delta t, \tag{3.19}$$

where $C^{1,4}=\|u_x\|_\infty C^{1,3}$. Now by integrating along the characteristic from $(t^{k-1},\bar{x})$ to (t^k,x_i), we see that

$$\begin{aligned} &\frac{u\left(t^k,x_i\right)-u(t^{k-1},\bar{x})}{\Delta t} \\ &\qquad=\frac{1}{\Delta t}\int_{t^{k-1}}^{t^k}\frac{d}{dt}u(t,x(t))\,dt \\ &\qquad=\frac{1}{\Delta t}\int_{t^{k-1}}^{t^k}(u_t-b(t,x,u)u_x)(t,x(t))\,dt \\ &\qquad=\frac{1}{\Delta t}\int_{t^{k-1}}^{t^k}\psi\frac{\partial u}{\partial\tau}(t,x(t))\,dt \\ &\qquad=\psi\left(t^k,x_i\right)\frac{\partial u}{\partial\tau}\bigg|_{(t^k,x_i)} \\ &\qquad\quad+\frac{1}{\Delta t}\int_{t^{k-1}}^{t^k}\left\{\psi\frac{\partial u}{\partial\tau}(t,x(t))-\psi\left(t^k,x_i\right)\frac{\partial u}{\partial\tau}\bigg|_{(t^k,x_i)}\right\}dt. \end{aligned} \tag{3.20}$$

Applying Theorem 2.3 and using the boundedness of the function b, one can easily deduce that

$$\left|\frac{1}{\Delta t}\int_{t^{k-1}}^{t^k}\left\{\psi\frac{\partial u}{\partial\tau}(t,x(t))-\psi\left(t^k,x_i\right)\frac{\partial u}{\partial\tau}\bigg|_{(t^k,x_i)}\right\}dt\right|\le C^{1,5}(h+\Delta t), \tag{3.21}$$

where $C^{1,5}$ depends on uniform bounds of $\partial^2u/\partial\tau^2$ along the characteristics (hence it depends on the bounds of u_{tt}, u_{tx} and b). Combining (3.18)–(3.21), we have

$$|I_i^{1,k}|=\left|\psi\frac{\partial u}{\partial\tau}\bigg|_{(t^k,x_i)}-\frac{u\left(t^k,x_i\right)-\bar{u}_i^{k-1}}{\Delta t}\right|\le\tilde{C}^{1,3}(h+\Delta t),$$

where $\tilde{C}^{1,3}=C^{1,3}+C^{1,4}+C^{1,5}$. If we set $C_1=C^{1,1}+C^{1,2}+\tilde{C}^{1,3}$, we have proven the lemma. □

We now study the consistency of our numerical scheme; namely, we shall prove that the approximate solution obtained from the difference equation

(3.11) converges to the true solution, in a certain sense. To do this, let us first construct the *approximate solution* defined on $[0,T] \times \mathbb{R}$ as follows. For given $h > 0$ and $\Delta t > 0$, set

$$w^{h,\Delta t}(t,x) = \begin{cases} \sum_{k=1}^{N} w^k(x)\mathbf{1}_{(t^{k-1},t^k]}(t), & t \in (0,T], \\ w^0(x), & t = 0, \end{cases} \tag{3.22}$$

where $w^k(\cdot)$, $k = 1,\dots,N$, are the functions extended from the solutions of the difference equations (3.11) by linear interpolations. In other words, for each k and i, $w^{h,\Delta t}(t^k, x_i) = w_i^k$, where $\{w_i^k\}$ is the solution to (3.11).

Let us now define a function $\zeta(t,x) = u(t,x) - w^{h,\Delta t}(t,x)$ for $(t,x) \in [0,T] \times \mathbb{R}$; as before, let $\zeta_i^k = \zeta(t^k, x_i) = u_i^k - w_i^k$. We first prove a theorem analogous to one in [3].

THEOREM 3.3. *Assume* (A1)–(A3). *Then*

$$\sup_{k,i} |\zeta_i^k| = \mathscr{O}(h + \Delta t).$$

PROOF. First, by subtracting (3.11) from (3.12), we see that $\{\zeta_i^k\}$ satisfies the difference equation

$$\begin{aligned} \frac{\zeta_i^k - \left(\bar{u}_i^{k-1} - \bar{w}_i^{k-1}\right)}{\Delta t} &= \frac{1}{2}\left\{\left(\sigma(u)_i^k\right)^2 \delta_x^2(u)_i^k - \left(\sigma_i^k\right)^2 \delta^2(w)_i^k\right\} \\ &\quad + \left[\hat{b}(u)_i^k - \hat{b}_i^k\right] + e_i^k, \qquad \zeta_i^0 = 0. \end{aligned} \tag{3.23}$$

Note that

$$\begin{aligned} \bar{u}_i^{k-1} - \bar{w}_i^{k-1} &= \left[u\left(t^{k-1}, \bar{\mathbf{x}}_i^k\right) - u\left(t^{k-1}, \bar{x}_i^k\right)\right] + \left[u\left(t^{k-1}, \bar{x}_i^k\right) - w^{k-1}\left(\bar{x}_i^k\right)\right] \\ &= \bar{\zeta}_i^{k-1} + \left[u\left(t^{k-1}, \bar{\mathbf{x}}_i^k\right) - u\left(t^{k-1}, \bar{x}_i^k\right)\right], \end{aligned}$$

where $\bar{\zeta}_i^{k-1} = u(t^{k-1}, \bar{x}_i^k) - w^{k-1}(\bar{x}_i^k)$. Also,

$$\begin{aligned} &\left(\sigma(u)_i^k\right)^2 \delta_x^2(u)_i^k - \left(\sigma_i^k\right)^2 \delta_x^2(w)_i^k \\ &\qquad = \left(\sigma_i^k\right)^2 \delta_x^2(\zeta)_i^k + \left[\sigma^2\left(t^k, x_i, u_i^{k-1}\right) - \sigma^2\left(t^k, x_i, w_i^{k-1}\right)\right] \delta_x^2(u)_i^k. \end{aligned}$$

We can rewrite (3.23) as

$$\frac{\zeta_i^k - \bar{\zeta}_i^{k-1}}{\Delta t} = \frac{1}{2}\left(\sigma_i^k\right)^2 \delta_x^2(\zeta)_i^k + I_i^k + e_i^k, \qquad \zeta_i^0 = 0, \tag{3.24}$$

where

$$\begin{aligned} I_i^k = &-\frac{u\left(t^{k-1}, \bar{\mathbf{x}}_i^k\right) - u\left(t^{k-1}, \bar{x}_i^k\right)}{\Delta t} \\ &+ \frac{1}{2}\left[\sigma^2\left(t^k, x_i, u_i^{k-1}\right) - \sigma^2\left(t^k, x_i, w_i^{k-1}\right)\right] \delta_x^2(u)_i^k + \left[\hat{b}(u)_i^k - \hat{b}_i^k\right]. \end{aligned}$$

It is clear that, by Theorem 2.3 and the estimate (3.18), we can find constants C_2 and $C_3 > 0$, independent of k and i, such that

$$\left|\tfrac{1}{2}\left[\sigma^2\left(t^k, x_i, u_i^{k-1}\right) - \sigma^2\left(t^k, x_i, w_i^{k-1}\right)\right]\delta_x^2(u)_i^k + \left[\hat{b}(u)_i^k - \hat{b}_i^k\right]\right| \le C_2|\zeta_i^{k-1}|$$

and

$$\left|\frac{u\left(t^{k-1}, \overline{\mathbf{x}}_i^k\right) - u\left(t^{k-1}, \bar{x}_i^k\right)}{\Delta t}\right| \le C_3\,\Delta t.$$

Consequently we have

$$|I_i^k| \le C_2|\zeta_i^{k-1}| + C_3\,\Delta t. \tag{3.25}$$

Note that it follows from (3.24) that

$$\zeta_i^k = \bar{\zeta}_i^{k-1} + \left\{\tfrac{1}{2}\left(\sigma_i^k\right)^2\delta_x^2(\zeta)_i^k + I_i^k + e_i^k\right\}\Delta t.$$

We wish to apply a maximum principle argument, as was done in [5], to bound ζ_i^k. In order to do so, we consider the localized solutions of u and w as described in Remark 3.1. Note that for such solutions, the maximum absolute value of ζ_i^k (by a slight abuse of notation) occurs at an "interior" mesh point $x_{i(k)}^k$, where $-R < i(k)h < R$ for some large $R > 0$. Now, if we set $\|\zeta^k\| = \max_i|\zeta_i^k|$, then a maximum principle argument similar to that in [5], together with Lemma 3.1 and (3.25), shows that

$$\begin{aligned}\|\zeta^k\| &\le \max_i|\bar{\zeta}_i^{k-1}| + \max_i\left\{|I_i^k| + |e_i^k|\right\}\Delta t\\ &\le \max_i|\bar{\zeta}_i^{k-1}| + C_2\|\zeta^{k-1}\|\,\Delta t + (C_1 + C_3)(h + \Delta t)\,\Delta t,\end{aligned} \tag{3.26}$$

where C_1 is the constant in Lemma 3.1 and C_2 and C_3 are those in (3.25). Note that the constants C_1, C_2, C_3 are independent of the localization; therefore by taking the limit we see that (3.26) should hold for the "global solution" as well.

In order to estimate $\max_i|\bar{\zeta}_i^{k-1}|$, we adopt the argument in [5]. Namely, if $I_1(u)(t^k, \cdot)$ denotes the linear interpolate of the grid values $\{u_i^k\}_{i=-\infty}^{\infty}$ and $w^k(\cdot)$ the linear interpolate of $\{w_i^k\}_{i=-\infty}^{\infty}$, then

$$\max_i|\bar{\zeta}_i^{k-1}\| \le \max_i|\zeta_i^{k-1}| + \max_i\left|u\left(t^{k-1}, \bar{x}_i^k\right) - I_1(u)\left(t^{k-1}, \bar{x}_i^k\right)\right|. \tag{3.27}$$

Apply the Peano kernel theorem (cf. [3] or 5]) to show that

$$\max_i\left|u\left(t^{k-1}, \bar{x}_i^k\right) - I_1(u)\left(t^{k-1}, \bar{x}_i^k\right)\right| \le C_4 h^* h,$$

where $h^* = \mathscr{O}(\Delta t)$ and $C_4 > 0$ is independent of k and i. This, together with (3.27), amounts to saying that (3.26) can be rewritten as

$$\begin{aligned}\|\zeta^k\| &\le \|\zeta^{k-1}\| + C_2\|\zeta^{k-1}\|\,\Delta t + C_5(h + \Delta t)\,\Delta t,\\ &= \|\zeta^{k-1}\|(1 + C_2\,\Delta t) + C_5(h + \Delta t)\,\Delta t,\end{aligned} \tag{3.28}$$

where $C_5 \le 2(C_1 + C_3 + C_3)$ is independent of k. It then follows from the Gronwall lemma and the bound on $\|\zeta^0\|$ that

$$\|\zeta^k\| \le \mathscr{O}(h + \Delta t),$$

which proves the theorem. □

Our next goal is to construct for each n an approximate solution $u^{(n)}$ that converges to the true solution u in the FSDE (2.5) in a satisfactory way as n tends to infinity. To this end, let $n \in \mathbb{N}$ be given. Let $\Delta t = T/n$ and $h = 2C\,\Delta t$, where $C = \|b\|_\infty$. We note here that such a choice of h is only for the convenience of actual computation, since $h > C\,\Delta t$ implies that $|\bar{x}_i^k - x_i| \le \|b\|_\infty\,\Delta t < h$. Hence $\bar{x}_i^k$ do not go beyond the interval (x_{i-1}^k, x_{i+1}^k). Now let us choose

$$u^{(n)}(t,x) = w^{2CT/n,T/n}(t,x), \qquad (t,x) \in [0,T] \times \mathbb{R},$$

where $w^{h,\Delta t}$ is defined by (3.22). Our main theorem of this section is the following.

THEOREM 3.4. *Suppose that* (A1)–(A3) *hold. Then the sequence* $\{u^{(n)}(\cdot,\cdot)\}$ *enjoys the following properties*:

(i) *For fixed* $x \in \mathbb{R}$, $u^{(n)}(\cdot, x)$ *is left continuous.*
(ii) *For fixed* $t \in [0,T]$, $u^{(n)}(t,\cdot)$ *is Lipschitz, uniformly in* t *and* n (*i.e., the Lipschitz constant is independent of* t *and* n).
(iii) $\sup_{t,x}|u^{(n)}(t,x) - u(t,x)| = \mathcal{O}(1/n)$.

PROOF. The property (i) is obvious by definition (3.22). To see (iii), we note that

$$u^{(n)}(t,x) - u(t,x) = \big[w^0(x) - u(0,x)\big]\mathbf{1}_{\{0\}}(t) + \sum_{k=1}^{N}\big[w^k(x) - u(t,x)\big]\mathbf{1}_{(t^{k-1},t^k]}(t).$$

Since for each fixed $t \in (t^{k-1}, t^k]$, $k > 0$ or $t = 0$, we have $u^{(n)}(t,x) = w^k(x)$ for $k > 0$ or $k = 0$ if $t = 0$. Thus,

$$\begin{aligned}\sup_x\big|w^k(x) - u(t,x)\big| &\le \|\zeta^k\| + \sup_x\big|I_1(u)(t^k,x) - u(t^k,x)\big| + \sup_x\big|u(t^k,x) - u(t,x)\big| \\ &\le \|\zeta^k\| + o(h+\Delta t) + \|u_t\|_\infty\,\Delta t = \mathcal{O}(h+\Delta t) = \mathcal{O}\left(\frac{1}{n}\right),\end{aligned}$$

by virtue of Theorem 3.3 and the definitions of h and Δt. This proves (iii).

To show (ii), let n and t be fixed, and assume that $t \in (t^k, t^{k+1}]$. Then $u^{(n)}(t,x) = w^k(x)$ is obviously Lipschitz in x. So it remains to determine the Lipschitz constant of every w^k. Let x^1 and x^2 be given. We may assume that $x^1 \in [x_i, x_{i+1})$ and $x^2 \in [x_j, x_{j+1})$, with $i < j$. For $i < l < j-1$, Theorem 3.3 implies that

$$\begin{aligned}\big|w^k(x_l) - w^k(x_{l+1})\big| &\le \big|w^k(x_l) - u\big(t^k,x_l\big)\big| + \big|u\big(t^k,x_l\big) - u\big(t^k,x_{l+1}\big)\big| \\ &\quad + \big|u\big(t^k,x_{l+1}\big) - w^k(x_{l+1})\big| \\ &\le 2\|\zeta^k\| + \|u_x\|_\infty|x_l - x_{l+1}| \le Kh = K(x_{l+1} - x_l),\end{aligned} \tag{3.29}$$

where K is a constant independent of k, l and n. Further, for $x^1 \in [x_i, x_{i+1})$,

$$w^k(x^1) = w^k(x_{i+1}) + \frac{w^k(x_{i+1}) - w^k(x_i)}{x_{i+1} - x_i}\left(x^1 - x_{i+1}\right).$$

Hence,

$$\left|w^k(x^1) - w^k(x_{i+1})\right| = \left|\frac{w^k(x_{i+1}) - w^k(x_i)}{x_{i+1} - x_i}\right| |x^1 - x_{i+1}| \le K|x^1 - x_{i+1}|,$$

where K is the same as that in (3.29). Similarly,

$$\left|w^k(x^2) - w^k(x_j)\right| \le K|x^2 - x_j|.$$

Combining the above gives

$$\begin{aligned}\left|w^k(x^1) - w^k(x^2)\right| &\le \left|w^k(x^1) - w^k(x_{i+1})\right| + \sum_{l=1}^{j-1}\left|w^k(x_l) - w^k(x_{l+1})\right| \\ &\quad + \left|w^k(x_j) - w^k(x^2)\right| \\ &\le K\left\{\left(x_{i+1} - x^1\right) + \sum_{l=1}^{j-1}(x_{l+1} - x_l) + \left(x^2 - x_{j+1}\right)\right\} \\ &= K|x^2 - x^1|.\end{aligned}$$

Since the constant K is independent of t and n, the theorem is proved. □

4. Approximation of the FSDE (2.5)—special case. We now use the approximate solution derived in the previous section to construct an approximation of the FSDE (2.5). First, we recall that the FSDE to be approximated has the form

$$X_t = x + \int_0^t \tilde{b}(s, X_s)\, ds + \int_0^t \tilde{\sigma}(s, X_s)\, dW_s, \tag{4.0}$$

where $\tilde{b}(t, x) = b(t, x, \theta(t, x))$ and $\tilde{\sigma}(t, x) = \sigma(t, x, \theta(t, x))$, $(t, x) \in [0, T] \times \mathbb{R}$. In order to define the approximate SDE's, we first define some quantities. For each $n \in \mathbb{N}$, set $\Delta t_n = T/n$ and $t^{n,k} = k\,\Delta t_n$, $k = 0, 1, 2, \dots, n$. Also, let

$$\begin{aligned}\eta^n(t) &= \sum_{k=0}^{n-1} t^{n,k} 1_{[t^{n,k}, t^{n,k+1})}(t), \qquad t \in [0, T), \\ \eta^n(T) &= T.\end{aligned} \tag{4.1}$$

Next, we set

$$\begin{aligned}\theta^n(t, x) &= u^{(n)}(T - t, x), \\ \tilde{b}^n(t, x) &= b(t, x, \theta^n(t, x)), \\ \tilde{\sigma}^n(t, x) &= \sigma(t, x, \theta^n(t, x)).\end{aligned} \tag{4.2}$$

By Theorem 3.4, θ^n is right continuous in t and uniformly Lipschitz in x, with the Lipschitz constant being independent of t and n; thus, so also are the functions $\tilde{b}^n$ and $\tilde{\sigma}^n$. We henceforth assume that there exists a constant K such that, for all t and n,

$$|\tilde{b}^n(t,x) - \tilde{b}^n(t,x')| + |\tilde{\sigma}^n(t,x) - \tilde{\sigma}^n(t,x')| \le K|x - x'|, \quad x, x' \in \mathbb{R}. \tag{4.3}$$

Also, from Theorem 3.4,

$$\sup_{t,x} |\tilde{b}^n(t,x) - \tilde{b}(t,x)| + \sup_{t,x} |\tilde{\sigma}^n(t,x) - \tilde{\sigma}(t,x)| = \mathscr{O}\left(\frac{1}{n}\right). \tag{4.4}$$

We now introduce two SDE's: the first one is a *discretized SDE* given by

$$\overline{X}_t^n = x + \int_0^t \tilde{b}^n(\cdot, \overline{X}_\cdot^n)_{\eta^n(s)}\, ds + \int_0^t \tilde{\sigma}^n(\cdot, \overline{X}_\cdot^n)_{\eta^n(s)}\, dW_s, \tag{4.5}$$

where η^n is defined by (4.1). The other is an *intermediate approximate SDE* given by

$$X_t^n = x + \int_0^t \tilde{b}^n(s, X_s^n)\, ds + \int_0^t \tilde{\sigma}^n(s, X_s^n)\, dW_s. \tag{4.6}$$

It is clear from the properties of $\tilde{b}^n$ and $\tilde{\sigma}^n$ mentioned above that both SDE's (4.5) and (4.6) possess unique strong solutions.

Our first result of this section is the following lemma. The proof of the lemma is more or less standard in the context of first order Euler approximations, but contains some special considerations due to the structure of the approximate solution to the PDE (2.4). We provide details for completeness.

LEMMA 4.1. *Assume* (A1)–(A3). *Then*

$$E\left\{\sup_{0\le t\le T} |\overline{X}_t^n - X_t^n|^2\right\} = \mathscr{O}\left(\frac{1}{n}\right).$$

PROOF. To simplify notation, we shall suppress the tilde (˜) for the coefficients in the sequel. We first rewrite (4.5) as

$$\overline{X}_t^n = X_0 + u_t^n + \int_0^t b^n(s, \overline{X}_s^n)\, ds + \int_0^t \sigma^n(s, \overline{X}_s^n)\, dW_s, \tag{4.7}$$

where

$$\begin{aligned} u_t^n = &\int_0^t \left[b^n(\cdot, \overline{X}_\cdot^n)_{\eta_n(s)} - b^n(s, \overline{X}_s^n)\right] ds \\ &+ \int_0^t \left[\sigma^n(\cdot, \overline{X}_\cdot^n)_{\eta_n(s)} - \sigma^n(s, \overline{X}_s^n)\right] dW_s. \end{aligned} \tag{4.8}$$

Applying Doob's maximal quadratic inequality, Jensen's inequality and the Lipschitz property (4.3) of the coefficients, we have

$$\begin{aligned} E\left\{\sup_{s\le t} |X_s^n - \overline{X}_s^n|^2\right\} \le\, &3E\left\{\sup_{s\le t} |u_s^n|^2\right\} + 3K^2 t \int_0^t E\{|X_s^n - \overline{X}_s^n|^2\}\, ds \\ &+ 12K^2 \int_0^t E\{|X_s^n - \overline{X}_s^n|^2\}\, ds. \end{aligned} \tag{4.9}$$

Now, set $\alpha_n(t) = E\{\sup_{s\le t}|X_s^n - \overline{X}_s^n|^2\}$. Then, from (4.9),

$$\alpha_n(t) \le 3E\Big\{\sup_{s\le t}|u_s^n|^2\Big\} + 3K^2(T+4)\int_0^t \alpha_n(s)\,ds$$

and Gronwall's inequality leads to

$$E\Big\{\sup_{s\le t}|X_s^n - \overline{X}_s^n|^2\Big\} \le 3e^{3K^2(T+4)}E\Big\{\sup_{s\le t}|u_s^n|^2\Big\}. \tag{4.10}$$

Consequently, we turn our attention to $E\{\sup_{s\le t}|u_s^n|^2\}$. Note that if $s \in [t^{n,k}, t^{n,k+1})$, for some $1 \le k < n$, then $\eta^n(s) = k\,\Delta t_n$ [whence $T - \eta^n(s) = (n-k)\,\Delta t_n$ as $T = n\,\Delta t_n$] and $T - s \in ((n-k-1)\,\Delta t_n, (n-k)\,\Delta t_n)$. Thus, by definitions (3.22) and (4.1), for every $x \in \mathbb{R}$,

$$\begin{aligned}\theta^n(\eta^n(s), x) &= u^{(n)}(T - \eta^n(s), x) = u^{(n)}((n-k)\,\Delta t_n, x)\\ &= u^{(n)}(T - s, x) = \theta^n(s, x).\end{aligned}$$

More generally,

$$b^n(s,x) = b(s,x,\theta^n(s,x)) = b(s,x,\theta^n(\eta^n(s),x)) \quad \forall\,(s,x) \in [0,T]\times\mathbb{R}. \tag{4.11}$$

Using this fact, it is easily seen that

$$\begin{aligned}&\left|\int_0^t b^n(\cdot,\overline{X}_\cdot^n)_{\eta_n(s)} - b^n(s,X_s^n)\,ds\right|\\ &\quad\le \int_0^t \Big|b\Big(\eta^n(s), \overline{X}_{\eta^n(s)}^n, \theta^n\Big(\eta^n(s), \overline{X}_{\eta^n(s)}^n\Big)\Big) - b(s, X_s^n, \theta^n(s, X_s^n))\Big|\,ds\\ &\quad\le \int_0^t \Big\{\Big|b\Big(\eta^n(s), \overline{X}_{\eta^n(s)}^n, \theta^n\Big(s, \overline{X}_{\eta^n(s)}^n\Big)\Big) - b\Big(s, X_s^n, \theta^n\Big(s, \overline{X}_{\eta^n(s)}^n\Big)\Big)\Big|\\ &\qquad\qquad + \Big|b\Big(s, X_s^n, \theta^n\Big(s, \overline{X}_{\eta^n(s)}^n\Big)\Big) - b(s, X_s^n, \theta^n(s, X_s^n))\Big|\Big\}\,ds\\ &\quad= I_1 + I_2.\end{aligned} \tag{4.12}$$

Using the boundedness of the functions b_t, b_x and b_y, we see that

$$I_1 \le \int_0^t \Big\{\|b_t\|_\infty|\eta^n(s) - s| + \|b_x\|_\infty|\overline{X}_{\eta^n(s)}^n - X_s^n|\Big\}\,ds$$

and

$$I_2 \le K\|b_y\|_\infty\int_0^t |\overline{X}_{\eta^n(s)}^n - X_s^n|\,ds.$$

Thus,

$$\begin{aligned}&\left|\int_0^t b^n(\cdot,\overline{X}_\cdot^n)_{\eta^n(s)} - b^n(s,X_s^n)\,ds\right|\\ &\qquad\le \tilde{K}\int_0^t \Big\{|\eta^n(s) - s| + |\overline{X}_{\eta^n(s)}^n - X_s^n|\Big\}\,ds,\end{aligned} \tag{4.13}$$

where $\tilde{K}$ depends only on K, $\|b_t\|_\infty$, $\|b_x\|_\infty$ and $\|b_y\|_\infty$. Since

$$\int_0^t |\eta^n(s) - s|\, ds = \sum_{k=0}^{n-1} \int_{t^k \wedge t}^{t^{k+1} \wedge t} (s - t^k)\, ds \le \frac{1}{2} \sum_{k=0}^{n-1} (\Delta t_n)^2 = \frac{T^2}{2n},$$

$$\tag{4.14} E\left\{ \sup_{u \le t} \left| \int_0^t b^n(\cdot, \bar{X}^n_\cdot)_{\eta^n(s)} - b^n(s, X^n_s)\, ds \right|^2 \right\} \le 2\tilde{K}^2 \left\{ T \int_0^t E|\bar{X}^n_{\eta^n(s)} - X^n_s|^2\, ds + \frac{T^4}{4n^2} \right\}.$$

Using the same reasoning for σ with Doob's inequality, we can see that

$$\tag{4.15} \begin{aligned} E\left\{ \sup_{u \le t} \left| \int_0^t \sigma^n(\cdot, \bar{X}^n_\cdot)_{\eta^n(s)} - \sigma^n(s, X^n_s)\, dW_s \right|^2 \right\} &\le 8\tilde{K}^2 \left\{ \int_0^t E|\bar{X}^n_{\eta^n(s)} - X^n_s|^2\, ds + \int_0^t (s - \eta^n(s))^2\, ds \right\} \\ &\le 8\tilde{K}^2 \left\{ \int_0^t E|\bar{X}^n_{\eta^n(s)} - X^n_s|^2\, ds + \frac{T}{3n^2} \right\}. \end{aligned}$$

Combining (4.14) and (4.15), we get

$$\tag{4.16} \begin{aligned} E\left\{ \sup_{s \le t} |u^k_s|^2 \right\} &\le \tilde{K}^2 (4T + 16) \int_0^t E|\bar{X}^n_{\eta^n(s)} - X^n_s|^2\, ds \\ &\quad + \tilde{K}^2 T \left(T + \frac{16}{3} \right) \frac{1}{n^2}. \end{aligned}$$

Thus, by (4.10),

$$\tag{4.17} \begin{aligned} E\left\{ \sup_{s \le t} |X^n_s - \bar{X}^n_s|^2 \right\} &\le 3 \exp\left(3K^2 (T + 4)\right) \left\{ \tilde{K}^2 (4T + 16) \int_0^T E|\bar{X}^n_{\eta^n(s)} - X^n_s|^2\, ds \right. \\ &\quad \left. + \tilde{K}^2 T \left(T + \frac{16}{3} \right) \frac{1}{n^2} \right\}. \end{aligned}$$

Finally, noting that $|\bar{X}^n_{\eta^n(s)} - X^n_s| \le |\bar{X}^n_{\eta^n(s)} - \bar{X}^n_s| + |\bar{X}^n_s - X^n_s|$ and that

$$\bar{X}^n_{\eta_k(s)} - \bar{X}^n_s = b^n(\cdot, \bar{X}^n_\cdot)_{\eta^n(s)}(s - \eta^n(s)) + \sigma(\cdot, \bar{X}^n_\cdot)_{\eta^n(s)}(W_s - W_{\eta^n(s)}),$$

we see as before that

$$\begin{aligned} \int_0^t E|\bar{X}^n_{\eta^n(s)} - \bar{X}^n_s|^2\, ds &\le 2 \int_0^t \left\{ \|b\|^2_\infty (s - \eta^n(s))^2 + \|\sigma\|^2_\infty |s - \eta^n(s)| \right\} ds \\ &\le \frac{2\|b\|^2_\infty T}{3} \frac{1}{n^2} + \|\sigma\|^2_\infty T \frac{1}{n}. \end{aligned}$$

Therefore, (4.17) becomes

$$(4.18)\quad E\Big\{\sup_{s\le t}|X_s^n-\bar X_s^n|^2\Big\}\le C_1\frac{1}{n}+C_2\frac{1}{n^2}+C_3\int_0^t E\Big\{\sup_{r\le s}|\bar X_r^n-X_r^n|^2\Big\}\,ds,$$

where C_1, C_2 and C_3 are constants depending only on the coefficients b, σ and K and can be calculated explicitly from (4.17). Now, we conclude from (4.18) and Gronwall's inequality that

$$\alpha_n(t)\le\beta_n\exp(C_T),\qquad\forall\, t\in[0,T],$$

where $\beta_n=C_1n^{-1}+C_2n^{-2}$ and $C_T=C_3T$. In particular, by slightly changing the constants, we have

$$(4.19)\qquad \alpha_n(T)=E\Big\{\sup_{0\le t\le T}|\bar X_t^n-X_s^n|^2\Big\}\le\frac{C_1}{n}+\frac{C_2}{n^2}=\mathscr{O}\Big(\frac{1}{n}\Big),$$

proving the lemma. □

Our main result of the section is the following theorem.

THEOREM 4.2. *Assume* (A1)–(A3) *hold. Let* $\bar X^n$ *be the solution of the discretized SDE* (4.5) *and define* $\bar Y^n$ *by* $\bar Y_t^n=\theta^n(t,\bar X_t^n)$, $t\in[0,T]$, *where* θ^n *is given by* (4.2). *Then*

$$(4.20a)\qquad E\Big\{\sup_{0\le t\le T}|\bar X_t^n-X_t|\Big\}+E\Big\{\sup_{0\le t\le T}|\bar Y_t^n-Y_t|\Big\}=\mathscr{O}\Big(\frac{1}{\sqrt n}\Big),$$

where (X,Y) *is the adapted solution to the FBSDE* (2.1).

Moreover, if f *is any uniformly Lipschitz* C^2 *function, then for* n *large enough,*

$$(4.20b)\qquad \big|E\{f(\bar X_T^n)\}-E\{f(X_T)\}\big|\le\frac{K}{n},$$

where K *is a constant depending only on* f, σ, b, $\hat b$ *and* g.

PROOF. Recall that at the beginning of the proof of Lemma 4.1, we have suppressed the tilde (˜) for $\tilde b$ and $\tilde\sigma$ to simplify notation. Set

$$\varepsilon^n(t)=\Big\{\sup_x|b^n(t,x)-b(t,x)|^2+\sup_x|\sigma^n(t,x)-\sigma(t,x)|^2\Big\},$$

where b, b^n, σ and σ^n are defined by (4.0) and (4.2). Then, from (4.4) we know that $\sup_t|\varepsilon^n(t)|=\mathscr{O}(1/n^2)$. Now, applying Lemma 4.1, we have

$$(4.21)\qquad \begin{aligned} E\Big\{\sup_{s\le t}|\bar X_s^n-X_s|^2\Big\}&\le 2E\Big\{\sup_{s\le t}|\bar X_s^n-X_s^n|^2\Big\}+2E\Big\{\sup_{s\le t}|X_s^n-X_s|^2\Big\}\\ &=\mathscr{O}\Big(\frac{1}{n}\Big)+2E\Big\{\sup_{s\le t}|X_s^n-X_s|^2\Big\}.\end{aligned}$$

Further, observe that

$$
\begin{aligned}
&E\Big\{\sup_{s\le t}|X^n_s - X_s|^2\Big\}\\
&\quad\le 2T\int_0^t E\big|b^n(s,X^n_s) - b(s,X_s)\big|^2\,ds + 8\int_0^t E\big|\sigma^n(s,X^n_s) - \sigma(s,X_s)\big|^2\,ds\\
&\quad\le 4T\int_0^t E\big|b^n(s,X^n_s) - b^n(s,X_s)\big|^2\,ds\\
&\qquad + 16\int_0^t E\big|\sigma^n(s,X^n_s) - \sigma^n(s,X_s)\big|^2\,ds + 4(T+4)\int_0^t \varepsilon_n(s)\,ds\\
&\quad\le 4(T+4)K^2\int_0^t E\Big\{\sup_{r\le s}|X^n_r - X_r|^2\Big\}\,ds + 4(T+4)\int_0^t \varepsilon_n(s)\,ds.
\end{aligned}
$$

Applying Gronwall's inequality, we get

$$
(4.22)\quad E\Big\{\sup_{s\le t}|X^n_s - X_s|^2\Big\} \le 4(T+4)\left(\int_0^t \varepsilon_n(s)\,ds\right)\exp(4(T+4)K^2) \le \frac{\tilde C}{n^2},
$$

where $\tilde C$ is a constant depending only on K and T. Now, note that the functions θ and θ^n are both uniformly Lipschitz in x. So, if we denote their Lipschitz constants by the same L, then

$$
\begin{aligned}
E\Big\{\sup_{0\le t\le T}|Y_t - \bar Y^n_t|^2\Big\} &\le 2E\Big\{\sup_{0\le t\le T}\big|\theta(t,X_t) - \theta^n\big(t,\bar X^n_t\big)\big|^2\Big\}\\
&\quad + 2E\Big\{\sup_{0\le t\le T}\big|\theta^n\big(t,\bar X^n_t\big) - \theta\big(t,\bar Y^n_t\big)\big|^2\Big\}\\
&\le 2L^2E\Big\{\sup_{0\le t\le T}|X_t - \bar X^n_t|^2\Big\} + 2\sup_{(t,x)}|\theta(t,x) - \theta^n(t,x)|^2\\
&= \mathscr{O}\left(\frac{1}{n}\right),
\end{aligned}
$$

by Theorem 3.4 and (4.22). The estimate (4.20a) then follows from an easy application of the Cauchy–Schwarz inequality. To prove (4.21), let us begin by assuming from Theorem 3.4, without loss of generality (e.g., by taking n large enough), that $\sup_{(t,x)}|\theta^n(t,x) - \theta(t,x)| \le Cn^{-1}$. We modify $\bar X^n_t$ as defined by (4.5) by fixing n and approximating the solution X^n of (4.6) by a standard Euler scheme indexed by k:

$$
\bar X^{n,k}_t = x + \int_0^t b\big(\cdot,\bar X^{n,k}_\cdot\big)_{\eta^k(s)}\,ds + \int_0^t \sigma\big(\cdot,\bar X^{n,k}_\cdot\big)_{\eta^k(s)}\,dW_s.
$$

It is then standard ([10], page 460) that

$$
(4.23)\qquad \big|E\{f(X^n_T)\} - E\big\{f\big(\bar X^{n,k}_T\big)\big\}\big| \le \frac{C_1}{k}.
$$

On the other hand, we have

$$
\begin{aligned}
\big|E\{f(X_T)\} - E\{f(X_T^n)\}\big| &\le KE\{|X_T - X_T^n|\} \\
&\le E\Big\{\sup_{0\le t\le T}|X_t - X_t^n|\Big\} \le \frac{C_2}{n}
\end{aligned} \tag{4.24}
$$

for Lipschitzian f, by (4.22). Therefore, noting that $\overline{X}_t^n$ as defined by (4.5) is just $\overline{X}_t^{n,n}$, the triangle inequality, (4.23), and (4.24) lead to (4.21). □

5. The general case. In this section we generalize the results in the previous sections to the general case. Namely, we shall consider the FBSDE

$$
\begin{aligned}
X_t &= x + \int_0^t b(t, X_t, Y_t, Z_t)\,dt + \int_0^t \sigma(t, X_t, Y_t)\,dW_t, \\
Y_t &= g(X_T) + \int_t^T \hat{b}(t, X_t, Y_t, Z_t)\,dt + \int_t^T Z_t\,dW_t,
\end{aligned} \tag{5.1}
$$

and we shall design a numerical scheme that approximates not only the processes (X, Y), but also the process Z, which in some applications is the most interesting part. For example, in an option pricing model (see, e.g., [2]), the process Z represents a hedging strategy and therefore schemes approximating Z are of intrinsic interest.

Using the four step scheme described in Section 2, one can easily deduce that in this case the function $z(t, x, y, p)$ in Step 1 is given by

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

Therefore, the PDE (2.4) becomes

$$
\begin{aligned}
0 = \theta_t &+ \tfrac{1}{2}\sigma^2(t, x, \theta)\theta_{xx} + b(t, x, \theta, -\sigma(t, x, \theta)\theta_x)\theta_x \\
&+ \hat{b}(t, x, \theta, -\sigma(t, x, \theta)\theta_x), \qquad \theta(T, x) = g(x).
\end{aligned} \tag{5.2}
$$

Define b_0 and $\hat{b}_0$ by

$$
\begin{aligned}
b_0(t, x, y, z) &= b(t, x, y, -\sigma(t, x, y)z), \\
\hat{b}_0(t, x, y, z) &= \hat{b}(t, x, y, -\sigma(t, x, y)z).
\end{aligned} \tag{5.3}
$$

One can check that, if σ, b and $\hat{b}$ satisfy (A1)–(A3), then so do the functions σ, b_0 and $\hat{b}_0$. Further, if we again set $u(t, x) = \theta(T - t, x)$, $\forall\ (t, x)$, then (5.2) becomes

$$
\begin{aligned}
u_t = \tfrac{1}{2}\overline{\sigma}^2(t, x, u)u_{xx} + \overline{b}_0(t, x, u, u_x)u_x &+ \overline{\hat{b}}_0(t, x, u, u_x), \\
&u(0, x) = g(x).
\end{aligned} \tag{5.4}
$$

We will again drop the overbar $(\bar{\ })$ in the sequel. Note that Theorem 2.3 holds for the general case; hence, the solution to (5.4) will be bounded in $C^{2+\alpha/2, 4+\alpha}$ for some $\alpha \in (0, 1)$.

A way to approximate the process Z is to have a numerical scheme that approximates θ_x, or equivalently u_x. To do this, let us define the function

$v(t, x) = u_x(t, x)$; then using the technique mentioned in the proof of Theorem 2.3, one can first "differentiate" equation (5.4) and then show that (u, v) satisfies a parabolic system

$$\begin{aligned} u_t &= \tfrac{1}{2}\bar{\sigma}^2(t, x, u)u_{xx} + b_0(t, x, u, v)u_x + \hat{b}_0(t, x, u, v), \\ v_t &= \tfrac{1}{2}\bar{\sigma}^2(t, x, u)v_{xx} + B_0(t, x, u, v)v_x + \hat{B}_0(t, x, u, v), \\ u(0, x) &= g(x), \qquad v(0, x) = g'(x), \end{aligned} \tag{5.5}$$

where

$$\begin{aligned} B_0(t, x, y, z) &= \sigma(t, x, y)\left[\sigma_x(t, x, y) + \sigma_y(t, x, y)z\right] \\ &\quad + b(t, x, y, z) + b_z(t, x, y, z) + \hat{b}_z(t, x, y, z), \\ \hat{B}_0(t, x, y, z) &= \left[b_x(t, x, y, z) + b_y(t, x, y, z)z\right]z \\ &\quad + \hat{b}(t, x, y, z)z + \hat{b}_x s(t, x, y, z). \end{aligned} \tag{5.6}$$

REMARK 5.1. Since u and v are uniformly bounded by Theorem 2.3, the functions $B_0(t, x, u(t, x), v(t, x))$ and $\hat{B}_0(t, x, u(t, x), v(t, x))$ are uniformly bounded for all (t, x). Also, B_0 and $\hat{B}_0$ are Lipschitz in x, y and z, uniformly in t and x and locally uniformly in y and z.

We shall introduce a numerical scheme based on (5.6) which produces a sequence of approximate solutions $\{U^{(n)}, V^{(n)}\}_{n=1}^{\infty}$ such that

$$\sup_{t,x}\left\{\left|U^{(n)}(t, x) - u(t, x)\right| + \left|V^{(n)}(t, x) - v(t, x)\right|\right\} = \mathscr{O}\left(\frac{1}{n}\right).$$

Following the idea presented in Section 3, we first determine the characteristics of the first order system

$$\begin{aligned} u_t - b_0(t, x, u, v)u_x &= 0, \\ v_t - B_0(t, x, u, v)v_x &= 0. \end{aligned}$$

It is easy to check that the two characteristics curves $\mathscr{C}_i$: $(t, x_i(t))$, $i = 1, 2$, are determined by the ODE's

$$\begin{aligned} dx_1(t) &= b_0\big(t, x_1(t), u(t, x_1(t)), v(t, x_1(t))\big)\,dt, \\ dx_2(t) &= B_0\big(t, x_2(t), u(t, x_2(t)), v(t, x_2(t))\big)\,dt. \end{aligned}$$

Let τ_1 and τ_2 be the arc lengths along $\mathscr{C}_1$ and $\mathscr{C}_2$, respectively. Then,

$$d\tau_1 = \psi_1(t, x_1(t))\,dt, \qquad d\tau_2 = \psi_2(t, x_2(t))\,dt, \tag{5.7}$$

where

$$\begin{aligned} \psi_1(t, x) &= \left[1 + b_0^2(t, x, u(t, x), v(t, x))\right]^{1/2}, \\ \psi_2(t, x) &= \left[1 + B_0^2(t, x, u(t, x), v(t, x))\right]^{1/2}. \end{aligned} \tag{5.8}$$

Thus, along $\mathscr{C}_1$ and $\mathscr{C}_2$, respectively,

$$\psi_1 \frac{\partial}{\partial \tau_1} = \left\{ \frac{\partial}{\partial t} - b_0 \frac{\partial}{\partial x} \right\}, \qquad \psi_2 \frac{\partial}{\partial \tau_2} = \left\{ \frac{\partial}{\partial t} - B_0 \frac{\partial}{\partial x} \right\}$$

and (5.5) can be simplified to

$$\begin{aligned} \psi_1 \frac{\partial u}{\partial \tau_1} &= \frac{1}{2}\sigma^2(t,x,u)u_{xx} + \hat{b}_0(t,x,u,v), \\ \psi_2 \frac{\partial v}{\partial \tau_2} &= \frac{1}{2}\sigma^2(t,x,u)v_{xx} + \hat{B}_0(t,x,u,v). \end{aligned} \tag{5.9}$$

Numerical scheme. For any $n \in \mathbb{N}$, let $\Delta t = T/n$. Let $h > 0$ be given. Let $t^k = k\,\Delta t$, $k = 0, 1, 2, \dots,$ and $x_i = ih$, $i = \dots, -1, 0, 1, \dots,$ as before.

STEP 0. Set $U_i^0 = g(x_i)$, $V_i^0 = g'(x_i)$ $\forall$ i, and extend U^0 and V^0 to all $x \in \mathbb{R}$ by linear interpolation.

Next, suppose that U^{k-1} and V^{k-1} are defined such that $U^{k-1}(x_i) = U_i^{k-1}$ and $V^{k-1}(x_i) = V_i^{k-1}$, and let

$$\begin{aligned} \left(\hat{b}_0\right)_i^k &= \hat{b}_0\left(t^k, x_i, U_i^{k-1}, V_i^{k-1}\right), \\ \left(\hat{B}_0\right)_i^k &= \hat{B}_0\left(t^k, x_i, U_i^{k-1}, V_i^{k-1}\right), \\ \sigma_i^k &= \sigma\left(t^k, x_i, U_i^{k-1}\right), \\ \bar{x}_i^k &= x_i + b_0\left(t^k, x_i, U_i^{k-1}, V_i^{k-1}\right)\Delta t, \\ \bar{\mathbf{x}}_i^k &= x_i + B_0\left(t^k, x_i, U_i^{k-1}, V_i^{k-1}\right)\Delta t \end{aligned} \tag{5.10}$$

and $\overline{U}_i^{k-1} = U^{k-1}(\bar{x}_i^k)$ and $\overline{V}_i^{k-1} = V^{k-1}(\bar{\mathbf{x}}_i^k)$.

STEP k. Determine the kth step grid values (U^k, V^k) by the system of difference equations

$$\begin{aligned} \frac{U_i^k - \overline{U}_i^{k-1}}{\Delta t} &= \frac{1}{2}\left(\sigma_i^k\right)^2 \delta_x^2(U)_i^k + \left(\hat{b}_0\right)_i^k, \\ \frac{V_i^k - \overline{V}_i^{k-1}}{\Delta t} &= \frac{1}{2}\left(\sigma_i^k\right)^2 \delta_x^2(V)_i^k + \left(\hat{B}_0\right)_i^k. \end{aligned} \tag{5.11}$$

We then extend the grid values $\{U_i^k\}$ and $\{V_i^k\}$ to the function $U^k(x)$ and $V^k(x)$, $x \in \mathbb{R}$, by linear interpolation.

We shall follow the argument in Section 3 to prove convergence. We point out that, unlike in the previous case, the functions B_0 and $\hat{B}_0$ [see definition (5.6)] are neither uniformly bounded nor uniformly Lipschitz. The arguments are thus more delicate. It turns out that this difficulty can be overcome if one can show that the solutions $\{U_i^k\}$ and $\{V_i^k\}$ to the difference equation (5.11) are uniformly bounded for all k and i and the bound is independent of n.

Indeed, if this is the case, then because the true solutions u and $v = u_x$ are uniformly bounded, we can restrict ourselves to the set $Q_M := \{(t, x, y, z) \in [0, T] \times \mathbb{R}^3 \colon |y| \le M, |z| \le M\}$, where M depends on the bounds of (u, v) and $\{U^k, V^k\}$, and all previous estimates will go through, with constants now depending possibly on uniform bounds of B_0 and $\hat{B}_0$ and their partial derivatives over Q_M. To justify this argument, let us first prove a lemma that has intrinsic interest.

LEMMA 5.2. *Suppose that* $\{a_k \colon k = 1, \dots, n\}$ *is a finite sequence such that* $a_k \ge 0$, $\forall$ k *and* $a_0 \le \alpha$. *Also assume that the following recursive relation holds*:

$$a_k \le a_{k-1} + \frac{C}{n}\left(1 + a_{k-1}^2\right), \qquad k = 1, \dots, n, \tag{5.13}$$

where $C > 0$ *is a constant independent of* k *and* n. *Then there exists a constant* $M > 0$ *depending only on* C *and* α, *such that* $\sup_n \sup_{0 \le k \le n} a_k \le M$.

PROOF. Let $A(\cdot)$ be the solution to the ODE

$$\frac{dA(t)}{dt} = \frac{C}{n}\left(1 + A^2(t)\right), \qquad A(0) = \alpha, \tag{5.14}$$

where $0 \le t \le n$. Since $dA/dt > 0$, A is increasing. Thus, for each $k = 1, \dots, n$, it holds that

$$\begin{aligned} A(k) &= A(k-1) + \frac{C}{n}\int_{k-1}^{k}\left(1 + A^2(r)\right) dr \\ &\ge A(k-1) + \frac{C}{n}\left(1 + A^2(k-1)\right). \end{aligned} \tag{5.15}$$

Noting that $A(0) = \alpha \ge a_0$, a simple induction using (5.13) and (5.15) then shows that $A(k) \ge a_k$, $k = 1, \dots, n$. It suffices to determine the bound for $A(t)$, $0 \le t \le n$, but by solving (5.14), we have

$$C = \int_0^n \frac{dA(t)}{1 + A^2(t)} = \arctan A(n) - \arctan A(0) = \arctan A(n) - \arctan \alpha,$$

hence

$$\sup_{0 \le k \le n} A(k) = A(n) = \tan(C + \arctan \alpha) := M < \infty.$$

Consequently, we obtain that

$$\sup_n \sup_{0 \le k \le n} a_k \le \sup_n \sup_{0 \le k \le n} A(k) = M,$$

proving the lemma. □

We now give a crucial result of this section.

THEOREM 5.3. *Suppose that* (A1)–(A3) *hold. Then the solutions* $\{U_i^k, V_i^k\}$ *to the system of difference equations* (5.11) *are uniformly bounded in* k *and* i, *and the bound is independent of* n. *More precisely, there exists a constant* $M_1 > 0$, *depending only on* b, σ, $\hat{b}$ *and* T, *such that*

$$\sup_n \sup_{k,i} \left\{ |U_i^k| + |V_i^k| \right\} \le M_1. \tag{5.16}$$

PROOF. Let us rewrite (5.11) as

$$U_i^k = \overline{U}_i^{k-1} + \left\{ \tfrac{1}{2} \left(\sigma_i^k \right)^2 \delta_x^2 (U)_i^k + \left(\hat{b}_0 \right)_i^k \right\} \Delta t,$$

$$V_i^k = \overline{V}_i^{k-1} + \left\{ \tfrac{1}{2} \left(\sigma_i^k \right)^2 \delta_x^2 (V)_i^k + \left(\hat{B}_0 \right)_i^k \right\} \Delta t.$$

Since $|\overline{U}_i^{k-1}| \le \max_i |U_i^{k-1}| = \|U^{k-1}\|$ and $|\overline{V}_i^{k-1}| \le \|V^{k-1}\|$, a maximum principle argument shows that

$$\begin{aligned} \|U^k\| &\le \|\overline{U}^{k-1}\| + \left\| \left(\hat{b}_0 \right) \right\| \Delta t \le \|U^{k-1}\| + \left\| \left(\hat{b}_0 \right)^k \right\| \Delta t, \\ \|V^k\| &\le \|\overline{V}^{k-1}\| + \left\| \left(\hat{B}_0 \right)^k \right\| \Delta t \le \|V^{k-1}\| + \left\| \left(\hat{B}_0 \right)^k \right\| \Delta t. \end{aligned} \tag{5.17}$$

Since $\hat{b}_0$ is uniformly bounded, it is easy to check by iteration that $\sup_n \sup_k \|U^k\| < \infty$. It remains to show that the same is true for V_i^k as well. To this end, we first observe from the definition (5.6) that $\hat{B}_0$ is of quadratic growth in z, uniformly in (t, x, y). Namely, there exists a constant $K > 0$, depending only on the bounds of b, σ, $\hat{b}$ and their first order partial derivatives, such that $|\hat{B}_0(t, x, y, z)| \le K(1 + z^2)$. Therefore,

$$\left| \left(\hat{B}_0 \right)_i^k \right| = \left| \hat{B}_0 \left(t^k, x_i, U_i^{k-1}, V_i^{k-1} \right) \right| \le K \left(1 + |V_i^{k-1}|^2 \right)$$

and the second inequality in (5.17) leads to

$$\begin{aligned} \|V^k\| &\le \|V^{k-1}\| + K \left(1 + \|V^{k-1}\|^2 \right) \Delta t \\ &= \|V^{k-1}\| + \frac{TK}{n} \left(1 + \|V^{k-1}\|^2 \right), \end{aligned} \tag{5.18}$$

since $\Delta t = T/n$. Hence, the result follows from Lemma 5.2. □

We can now follow the arguments in Section 3 line by line. What follows is essentially a somewhat detailed sketch of the proof of Theorem 5.5. First, we evaluate the first equation in (5.9) along $\mathscr{C}_1$ and the second one along $\mathscr{C}_2$ to get an analogue of (3.12):

$$\frac{u_i^k - \hat{u}_i^{k-1}}{\Delta t} = \frac{1}{2} \left(\sigma(u)_i^k \right)^2 \delta_x^2 (u)_i^k + \hat{b}_0(u, v)_i^k + (e_1)_i^k,$$

$$\frac{v_i^k - \hat{v}_i^{k-1}}{\Delta t} = \frac{1}{2} \left(\sigma(u)_i^k \right)^2 \delta_x^2 (u)_i^k + \hat{B}_0(u, v)_i^k + (e_2)_i^k,$$

where $u_i^k = u(t^k, x_i)$, $v_i^k = v(t^k, x_i)$ [recall that $(u, v) = (u, u_x)$ is the true solution of (5.6)] and $\hat{u}_i^{k-1} = u(t^{k-1}, \hat{x}_i)$, $\hat{v}_i^k = v(t^{k-1}, \hat{\mathbf{x}}_i)$, with

$$\hat{x}_i^k = x_i + b_0\big(t^k, x_i, u_i^{k-1}, v_i^{k-1}\big)\,\Delta t, \qquad \hat{\mathbf{x}}_i^k = x_i + B_0\big(t^k, x_i, u_i^{k-1}, v_i^{k-1}\big)\,\Delta t.$$

Also, $\sigma(u)_i^k$, $\hat{b}_0(u,v)_i^k$ and $\hat{B}_0(u,v)_i^k$ are analogous to σ_i^k, $(\hat{b}_0)_i^k$ and $(\hat{B}_0)_i^k$ except that U_i^{k-1} and V_i^{k-1} are replaced by u_i^{k-1} and v_i^{k-1}.

Next, we estimate the error $\{(e_1)_i^k\}$ and $\{(e_2)_i^k\}$ in the same fashion as in Lemma 3.2 to see that

$$\sup_{k,i}\Big\{\big|(e_1)_i^k\big| + \big|(e_2)_i^k\big|\Big\} \le \mathscr{O}(h + \Delta t). \tag{5.19}$$

REMARK 5.4. In Lemma 3.2 we used the bound $\|b\|_\infty$ [see (3.18)]. The analogue for (5.19) is that $C = \max\{\|B_0\|_\infty, \|b\|_\infty\}$. In theory, the definition (5.6) implies that $\|B_0\|_\infty$ is always computable using the bounds of the coefficients (i.e., $b, \hat{b}, \sigma$) their partial derivatives and the bound on $v = u_x$. However, in practice, the a priori estimate of $\|v\|_\infty$ is not easy to obtain. However, in a computational process one could always replace $\|v\|_\infty$ by the a priori bound of the approximate solution $\{V_i^k\}$ derived in Theorem 5.3.

We now define as we did in (3.22) the approximate solutions $U^{(n)}$ and $V^{(n)}$ by

$$\begin{aligned} U^{(n)}(t,x) &= \begin{cases} \sum_{k=1}^{n} U^k(x) 1_{(((k-1)T)/n,\, kT/n]}(t), & t \in (0,T],\\ U^0(x), & t = 0,\end{cases}\\ V^{(n)}(t,x) &= \begin{cases} \sum_{k=1}^{n} V^k(x) 1_{(((k-1)T)/n,\, kT/n]}(t), & t \in (0,T],\\ V^0(x), & t = 0.\end{cases}\end{aligned} \tag{5.20}$$

Let $\xi(t,x) = u(t,x) - U^n(t,x)$ and $\zeta(t,x) = v(t,x) - V^n(t,x)$. We can derive the analogue of (3.24):

$$\frac{\xi_i^k - \hat{\xi}_i^{k-1}}{\Delta t} = \frac{1}{2}\big(\sigma_i^k\big)^2 \delta_x^2(\xi)_i^k + (I_1)_i^k + (e_1)_i^k,$$

$$\frac{\zeta_i^k - \hat{\zeta}_i^{k-1}}{\Delta t} = \frac{1}{2}\big(\sigma_i^k\big)^2 \delta_x^2(\zeta)_i^k + (I_2)_i^k + (e_2)_i^k,$$

where

$$\begin{aligned}(I_1)_i^k = &-\frac{u\big(t^{k-1}, \hat{x}_i^k\big) - u\big(t^{k-1}, \bar{x}_i^k\big)}{\Delta t}\\ &+ \frac{1}{2}\big[\sigma^2\big(t^k, x_i, u_i^{k-1}\big) - \sigma^2\big(t^k, x_i, U_i^{k-1}\big)\big]\delta_x^2(u)_i^k\\ &+ \Big[\hat{b}_0(u,v)_i^k - \big(\hat{b}_0\big)_i^k\Big]\end{aligned}$$

$$(I_2)_i^k = -\frac{v\left(t^{k-1}, \hat{\mathbf{x}}_i^k\right) - v\left(t^{k-1}, \overline{\mathbf{x}}_i^k\right)}{\Delta t}$$
$$+\frac{1}{2}\left[\sigma^2\left(t^k k, x_i, v_i^{k-1}\right) - \sigma^2\left(t^k, x_i, V_i^{k-1}\right)\right]\delta_x^2(v)_i^k$$
$$+\left[\hat{B}_0(u,v)_i^k - \left(\hat{B}_0\right)_i^k\right].$$

Using the uniform Lipschitz property of $\hat{b}_0$ in y and z, one shows that

$$\left|(I_1)_i^k\right| \le C_2\{|\xi_i^{k-1}| + |\zeta_i^{k-1}|\} + C_3(h + \Delta t) \qquad \forall\, k, i. \tag{5.21}$$

On the other hand, note that the true solution (u, v) is uniformly bounded and that $\{U_i^k\}$ and $\{V_i^k\}$ are also uniformly bounded by Lemma 5.1. We can use the locally uniform boundedness and local Lipschitz property of the function $\hat{B}_0$ (in y and z) to get

$$\left|\hat{B}_0(u,v)_i^k - \left(\hat{B}_0\right)_i^k\right| \le C_4\left(|\xi_i^{k-1}| + |\zeta_i^{k-1}|\right) \qquad \forall\, k, i,$$

where C_4 depends only on the bounds of u, v, $\{U_i^k\}$, $\{V_i^k\}$ and those of σ, b, $\hat{b}$ and their partial derivatives. Consequently,

$$\left|(I_2)_i^k\right| \le C_2'\{|\xi_i^{k-1}| + |\zeta_i^{k-1}|\} + C_3'(h + \Delta t) \qquad \forall\, k, i. \tag{5.22}$$

Use of the maximum principle and the estimates (5.19), (5.21) and (5.22) leads to

$$\|\xi^k\| \le \|\xi^{k-1}\| + C_2\left(\|\xi^{k-1}\| + \|\zeta^{k-1}\|\right)\Delta t + C_5(h + \Delta t)\,\Delta t,$$
$$\|\zeta^k\| \le \|\zeta^{k-1}\| + C_2'\left(\|\xi^{k-1}\| + \|\zeta^{k-1}\|\right)\Delta t + C_5'(h + \Delta t)\,\Delta t.$$

Add the two inequalities above and apply Gronwall's lemma; we see that

$$\sup_k\left(\|\xi^k\| + \|\zeta^k\|\right) = \mathscr{O}(h + \Delta t).$$

Arguments similar to those in Theorem 3.4 complete the proof of the following theorem.

THEOREM 5.5. *Suppose that* (A1)–(A3) *hold. Then*

$$\sup_{(t,x)}\left\{\left|U^{(n)}(t,x) - u(t,x)\right| + \left|V^{(n)}(t,x) - u_x(t,x)\right|\right\} = \mathscr{O}\left(\frac{1}{n}\right).$$

Moreover, for each fixed $x \in \mathbb{R}$, $U^{(n)}(\cdot, x)$ and $V^{(n)}(\cdot, x)$ are left-continuous; for fixed $t \in [0, T]$, $U^{(n)}(t, \cdot)$ and $V^{(n)}(t, \cdot)$ are uniformly Lipschitz, with the same Lipschitz constant that is independent of n.

Using Theorem 5.5, we can now approximate the SDE (2.5) as before without any extra work. In fact, if we set

$$\theta^n(t,x) = U^{(n)}(T - t, x), \qquad \theta_x^n(T - t, x) = V^{(n)}(t, x) \tag{5.24}$$

and

$$\tilde{b}^n(t,x) = b_0(t,x,\theta^n(t,x),\theta^n_x(t,x)), \qquad \tilde{\sigma}^n(t,x) = \sigma(t,x,\theta^n(t,x)),$$

then it follows from Theorem 5.6 that

$$\sup_{(t,x)} \left\{ \left| \tilde{b}(t,x) - \tilde{b}^n(t,x) \right| + \left| \tilde{\sigma}(t,x) - \tilde{\sigma}^n(t,x) \right| \right\} = \mathscr{O}\left(\frac{1}{n}\right)$$

and that, for fixed $t \in [0,T]$, $\tilde{b}^n(t,\cdot)$ and $\tilde{\sigma}^n(t,\cdot)$ are uniform Lipschitz, with the Lipschitz constant independent of n. Thus, if we again let $\overline{X}^n$ denote the solution to the discretized SDE,

$$\overline{X}^n_t = x + \int_0^t \tilde{b}^n\left(\cdot, \overline{X}^n_\cdot\right)_{\eta^n(s)} ds + \int_0^t \tilde{\sigma}^n\left(\cdot, \overline{X}^n_\cdot\right)_{\eta^n(s)} dW_s, \tag{5.25}$$

where η^n is defined by (4.1), then one can easily show the following final result of the paper, which is the analogous to Theorem 4.2.

THEOREM 5.6. *Suppose that the standing assumptions* (A1)–(A3) *hold. Then the adapted solution* (X,Y,Z) *to the FBSDE* (2.1) *can be approximated by a sequence of adapted processes* $(\overline{X}^n, \overline{Y}^n, \overline{Z}^n)$, *where* $\overline{X}^n$ *is the solution to the discretized SDE* (5.25) *and, for* $t \in [0,T]$,

$$\overline{Y}^n_t := \theta^n\left(t, \overline{X}^n_t\right); \qquad \overline{Z}^n_t := -\sigma\left(t, \overline{X}^n_t, \theta^n\left(t, \overline{X}^n_t\right)\right)\theta^n_x\left(t, \overline{X}^n_t\right),$$

with θ^n *and* θ^n_x *being defined by* (5.24) *and* $U^{(n)}$ *and* $V^{(n)}$ *by* (5.20). *Furthermore,*

$$E\left\{ \sup_{0 \le t \le T} |\overline{X}^n_t - X_t| + \sup_{0 \le t \le T} |\overline{Y}^n_t - Y_t| + \sup_{0 \le t \le T} |\overline{Z}^n_t - Z_t| \right\} = \mathscr{O}\left(\frac{1}{\sqrt{n}}\right).$$

If f is C^2 and uniformly Lipschitz, then, for n large enough,

$$\left| E\left\{ f\left(\overline{X}^n_T, \overline{Z}^n_T\right)\right\} - E\{f(X_T, Z_T)\} \right| \le \frac{K}{n}$$

for a constant K.

Acknowledgment. We wish to thank an anonymous referee for a careful reading of the manuscript and resultant constructive suggestions.

REFERENCES

[1] ANTONELLI, F. (1993). Backward–forward stochastic differential equations. *Ann. Appl. Probab.* **3** 777–793.
[2] CVITANIĆ, J. and MA, J. (1996). Hedging options for a large investor and forward–backward SDE's. *Ann. Appl. Probab.* **6** 370–398.
[3] DAVIS, P. (1975). *Interpolation and Approximation.* Dover, New York.
[4] DOUGLAS, J., JR. (1983). Finite difference methods for two-phase incompressible flow in porous media. *SIAM J. Numer. Anal.* **20** 681–696.
[5] DOUGLAS, J., JR. and RUSSELL, T. F. (1982). Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM J. Numer. Anal.* **19** 871–885.

[6] DUFFIE, D. and EPSTEIN, L. G. (1992). Stochastic differential utility. *Econometrica* **61** 353–394.

[7] DUFFIE, D., GEOFFARD, P-Y. and SKIADAS, C. (1992). Efficient and equilibrium allocations with stochastic differential utility. Preprint.

[8] DUFFIE, D., MA, J. and YONG, J. (1994). Black's console rate conjecture. *Ann. Appl. Probab.* **5** 356–382.

[9] GILBARG, D. and TRUDINGER, N. S. (1983). *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin.

[10] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin.

[11] LADYZENSKAJA, O. A., SOLONNIKOV, V. A. and URAL'CEVA, N. N. (1968). *Linear and Quasilinear Equations of Parabolic Type*. Amer. Math. Soc., Providence, RI.

[12] MA, J., PROTTER, P. and YONG, J. (1994). Solving forward–backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields* **98** 339–359.

DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907-1395
E-MAIL: douglas@math.purdue.edu
majin@math.purdue.edu
protter@math.purdue.edu