

NUMERICAL METHODS FOR FORWARD-BACKWARD
STOCHASTIC DIFFERENTIAL EQUATIONS

by

Jim Douglas, Jr., Jin Ma and Philip Protter
Purdue University

Technical Report #95-21

Department of Statistics
Purdue University

May 1995

NUMERICAL METHODS FOR FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

Jim Douglas, Jr.,¹ Jin Ma,² and Philip Protter³

Abstract. In this paper we study numerical methods to approximate the adapted solutions to a class of forward-backward stochastic differential equations (FBSDEs for short). 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 in [12]. 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.

Keywords. forward-backward stochastic differential equations, quasilinear parabolic equations, combined characteristics and finite difference method, Euler's scheme, weak convergence.

Mathematics subject classification (1991): Primary 65U05; Secondary 60H10, 65M06, 65M25.

§1. Introduction.

Let $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{t \geq 0})$ 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.

¹ Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395; supported in part by NSF grant #DMS 9207088.

² Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395; supported in part by NSF grant #DMS 9301516.

³ Departments of Mathematics and Statistics, Purdue University, West Lafayette, IN 47907-1395, supported in part by NSF grant #INT-9401109 and NSA grant #MDR 90494H2049.

We consider the following forward-backward stochastic differential equations:

$$(1.1) \quad \begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_s, Y_s, Z_s) dW_s; \\ Y_t = g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s, Z_s) ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s) dW_s, \end{cases} \quad t \in [0, T],$$

where (X, Y, Z) takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$ and $b, \widehat{b}, \sigma, \widehat{\sigma}$ 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 $\{\mathcal{F}_t\}$ -adapted, 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] [8]). In previous work Ma, Protter & Yong (1994) [12] studied the solvability of the adapted solution to the FBSDE; in particular, they designed a direct scheme, called the Four Step Scheme (see §2 for a brief review), to solve the FBSDE explicitly.

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

$$(1.2) \quad \begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s; \\ Y_t = E \left\{ g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s) ds \middle| \mathcal{F}_t \right\}. \end{cases} \quad t \in [0, T].$$

Applying the usual technique using a martingale representation theorem, it is easily seen that (1.2) is equivalent to the FBSDE

$$(1.3) \quad \begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s; \\ Y_t = g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s) ds + \int_t^T Z_s dW_s. \end{cases}$$

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 FBSDEs (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 & Yong (1994) [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 SDEs, there are two types of approximations typically considered: a pathwise convergence that typically converges at a rate $\mathcal{O}(\frac{1}{\sqrt{n}})$, 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 SDEs. 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 FBSDEs: 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 spacial mesh size h and the time mesh size Δt are essentially linearly related, with the condition $h > C\Delta t$, where $C > 0$ is a constant obtained from the coefficients, being 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 [4] (see also Douglas [5]). 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., [4]), which can be just as (or more) important.

Since the special FBSDE (1.2) and (1.3) are of independent interest, and the techniques of proofs and ideas are fundamental but more easily seen, we treat it separately in §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 FBSDEs 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 orders $\mathcal{O}(\frac{1}{\sqrt{n}})$ and $\mathcal{O}(\frac{1}{n})$, respectively, which are best possible for Euler schemes.

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

§2. Formulation of the Problem.

Let (Ω, \mathcal{F}, P) be a probability space carrying a standard d -dimensional Brownian motion $W = \{W_t : t \geq 0\}$, and let $\{\mathcal{F}_t\}$ be the σ -field generated by W (i.e., $\mathcal{F}_t = \sigma\{W_s : 0 \leq s \leq t\}$). We make the usual P -augmentation to each \mathcal{F}_t so that \mathcal{F}_t contains all the P -null sets of \mathcal{F} . Then $\{\mathcal{F}_t\}$ is right continuous and $\{\mathcal{F}_t\}$ satisfies the *usual hypotheses*. Let us consider the following forward-backward SDE:

$$(2.1) \quad \begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s, \\ Y_t = g(X_T) + \int_t^T \widehat{b}(s, X_s, Y_s, Z_s) ds + \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s) dW_s. \end{cases} \quad t \in [0, T].$$

Here, the processes X , Y and Z take values in \mathbb{R}^n , \mathbb{R}^m and $\mathbb{R}^{m \times d}$, respectively; and the functions b , \widehat{b} , σ , $\widehat{\sigma}$ and g take values in \mathbb{R}^n , \mathbb{R}^m , $\mathbb{R}^{n \times d}$, $\mathbb{R}^{m \times d}$ and \mathbb{R}^m , respectively. For the sake of simplicity, in what follows we will consider only the case in which $n = m = d = 1$, $\widehat{\sigma}(t, x, y, z) \equiv z$. The higher dimensional extension will be discussed in the last section of the paper.

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

Definition 2.1. A triple of processes $(X, Y, Z) : [0, T] \times \Omega \rightarrow \mathbb{R}^3$ is called an L^2 -adapted solution of the forward-backward SDE (2.1) if it is $\{\mathcal{F}_t\}$ -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 $C^{1+\frac{\alpha}{2}, 2+\alpha}([0, T] \times G)$ 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 $C^{1+\frac{\alpha}{2}, 2+\alpha}([0, T] \times G)$ is defined by

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

where $\|\cdot\|_{C_{T,G}}$ is the usual “sup”-norm on the closure of $[0, T] \times G$. When $G = \mathbb{R}$, we set $C^{1+\frac{\alpha}{2}, 2+\alpha}([0, T] \times \mathbb{R}) = C^{1+\frac{\alpha}{2}, 2+\alpha}$ 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 , \widehat{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 , $\psi(\cdot, \cdot, y, z) \in C^{1+\frac{\alpha}{2}, 2+\alpha}$. Furthermore, for some $L > 0$,

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

(A2) The function σ satisfies

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

where $0 < \mu \leq C$ 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 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 $z : [0, T] \times \mathbb{R}^3 \rightarrow \mathbb{R}$ by

$$(2.3) \quad z(t, x, y, p) = -p\sigma(t, x, y), \quad \forall (t, x, y, p).$$

Step 2. Using the function z above, solve the following quasilinear parabolic equation for $\theta(t, x)$:

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

Step 3. Using θ and z , solve the following forward SDE:

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

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

$$(2.6) \quad \begin{cases} Y_t = \theta(t, X_t), \\ Z_t = z(t, X_t, \theta(t, X_t), \theta_x(t, X_t)). \end{cases}$$

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 $C^{2+\frac{\alpha}{2}, 4+\alpha}$, 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 $\theta \in C^{1+\frac{\alpha}{2}, 2+\alpha}$ 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

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

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 following simpler form:

$$(3.1) \quad \begin{cases} \theta_t + \frac{1}{2}\sigma(t, x, \theta)^2\theta_{xx} + b(t, x, \theta)\theta_x + \widehat{b}(t, x, \theta) = 0, & (t, x) \in (0, T) \times \mathbb{R}, \\ \theta(T, x) = g(x), & x \in \mathbb{R}. \end{cases}$$

We shall follow an idea of Douglas and Russell [4] and Douglas [5] 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

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

Then u satisfies the PDE

$$(3.2) \quad \begin{cases} u_t - \frac{1}{2}\bar{\sigma}^2(t, x, u)u_{xx} - \bar{b}(t, x, u)u_x - \widetilde{\bar{b}}(t, x, u) = 0; \\ u(0, x) = g(x). \end{cases}$$

To simplify notation we replace $\bar{\sigma}$, \bar{b} and $\widetilde{\bar{b}}$ by σ , b and \widehat{b} themselves in the rest of this section. Following [4], we should first determine the characteristics of the first order nonlinear PDE

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

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 following equation:

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

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

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

Therefore, (3.4) leads to

$$(3.6) \quad t'(s)b(t, x, u) - x'(s) = 0,$$

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

$$(3.7) \quad d\tau = [1 + b^2(t, x, u(t, x))]^{\frac{1}{2}} dt;$$

along \mathcal{C} ,

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

where

$$(3.8) \quad \psi(t, x) = [1 + b^2(t, x, u(t, x))]^{\frac{1}{2}}.$$

Thus, the equation (3.3) can be simplified to

$$(3.9) \quad \begin{cases} \psi \frac{\partial u}{\partial \tau} = \frac{1}{2} \sigma^2(t, x, u) u_{xx} + \widehat{b}(t, x, u); \\ u(0, x) = g(x). \end{cases}$$

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 $f_i^k = f(t^k, x_i)$ 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 $w^{k-1}(t, x)$ is defined and let

$$(3.10) \quad \begin{cases} b_i^k = b(t^k, x_i, w_i^{k-1}); \sigma_i^k = \sigma(t^k, x_i, w_i^{k-1}); \widehat{b}_i^k = \widehat{b}(t^k, x_i, w_i^{k-1}); \\ \bar{x}_i^k = x_i - b_i^k \Delta t, \bar{w}_i^{k-1} = w^{k-1}(\bar{x}_i^k); \\ \delta_x^2(w)_i^k = h^{-2} [w_{i+1}^k - 2w_i^k + w_{i-1}^k]. \end{cases}$$

Step k: Obtain the grid values for the k -th step approximate solution, denoted by $\{w_i^k\}$, via the following difference equation:

$$(3.11) \quad \frac{w_i^k - \bar{w}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(w)_i^k + (\widehat{b})_i^k; \quad -\infty < i < \infty,$$

It is clear that, under the assumptions that σ be bounded below positively and that \widehat{b} be bounded, there exists a unique solution of (3.11) as soon as an evaluation is specified for $w^{k-1}(x)$.

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

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

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

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 following form:

$$(3.12) \quad \frac{u_i^k - \bar{u}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma(u)_i^k)^2 \delta_x^2(u)_i^k + \widehat{b}(u)_i^k + e_i^k, \quad -\infty < i < \infty, \quad k = 1, \dots, N,$$

where $\bar{u}_i^{k-1} = u^{k-1}(\bar{x}_i^k)$ and \bar{x}_i^k is the computed approximation of \bar{x} (see the definition following (3.17) below); $\widehat{b}(u)_i^k$ and $\sigma(u)_i^k$ correspond to \widehat{b}_i^k and σ_i^k defined in (3.10), except that the values $\{w_i^{k-1}\}$ are replaced by $\{u_i^{k-1}\}$; 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)

$$(3.13) \quad \psi(t^k, x_i) \frac{\partial u}{\partial \tau} \Big|_{(t^k, x_i)} = \frac{1}{2} \sigma^2(t^k, x_i, u_i^k) u_{xx} \Big|_{(t^k, x_i)} + \widehat{b}(t^k, x_i, u_i^k).$$

Therefore,

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

We have the following lemma.

Lemma 3.1. *Let $C = \|b\|_\infty$. Assume that $h > C\Delta t$. Then, there exists a constant $C_1 > 0$, depending only on $b, \widehat{b}, \sigma, u$ and T , such that for all $k = 0, \dots, N$ and $-\infty < i < \infty$,*

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

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 \widehat{b} that

$$(3.15) \quad \begin{aligned} |I_i^{3,k}| &= |\widehat{b}(t^k, x_i, u_i^k) - \widehat{b}(t^k, x_i, u_i^{k-1})| \\ &\leq \|\widehat{b}_u\|_\infty |u(t^k, x_i) - u(t^{k-1}, x_i)| \leq C^{1,1} \Delta t, \end{aligned}$$

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

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

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

$$(3.17) \quad \bar{x}_i^k = x_i - b(t^k, x_i, u_i^{k-1})\Delta t,$$

and $\bar{u}_i^{k-1} := u(t^{k-1}, \bar{x}_i^k)$. Let \bar{x} be the x -coordinate of the characteristic starting from (t^k, x_i) evaluated at $t = t^{k-1}$:

$$\bar{x} = x_i - \int_{t^{k-1}}^{t^k} b(t, x(t), u(t, x(t))) dt.$$

Since $\max\{|\bar{x}_i^k - x_i|, |\bar{x} - x_i|\} \leq \|b\|_\infty \Delta t < h$ by assumption,

$$\begin{aligned} |\bar{x}_i^k - \bar{x}| &\leq \int_{t^{k-1}}^{t^k} |b(t^k, x_i, u_i^{k-1}) - b(t, x(t), u(t, x(t)))| dt \\ &\leq \{ \|b_t\|_\infty \Delta t + \|b_x\|_\infty h + \|b_u\|_\infty (\|u_t\|_\infty \Delta t + \|u_x\|_\infty h) \} \Delta t \\ &\leq C^{1,3}(h + \Delta t)\Delta t. \end{aligned}$$

where $C^{1,3}$ depends only on $\|b_t\|_\infty, \|b_x\|_\infty, \|b_u\|_\infty, \|u_x\|_\infty$. Thus,

$$(3.18) \quad \left| \frac{u(t^{k-1}, \bar{x}) - u(t^{k-1}, \bar{x}_i^k)}{\Delta t} \right| \leq \frac{\|u\|_\infty |\bar{x} - \bar{x}_i^k|}{\Delta t} \leq C^{1,4}(h + \Delta t),$$

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

$$(3.19) \quad \begin{aligned} \frac{u(t^k, x_i) - u(t^{k-1}, \bar{x})}{\Delta t} &= \frac{1}{\Delta t} \int_{t^{k-1}}^{t^k} \frac{d}{dt} u(t, x(t)) dt \\ &= \frac{1}{\Delta t} \int_{t^{k-1}}^{t^k} (u_t - b(t, x, u)u_x)(t, x(t)) dt \\ &= \frac{1}{\Delta t} \int_{t^{k-1}}^{t^k} \psi \frac{\partial u}{\partial \tau}(t, x(t)) dt \\ &= \psi(t^k, x_i, u_i^k) \frac{\partial u}{\partial \tau} \Big|_{(t^k, x_i)} \\ &\quad + \frac{1}{\Delta t} \int_{t^{k-1}}^{t^k} \left\{ \psi \frac{\partial u}{\partial \tau}(t, x(t)) - \psi(t^k, x_i, u_i^k) \frac{\partial u}{\partial \tau} \Big|_{(t^k, x_i)} \right\} dt. \end{aligned}$$

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

$$(3.20) \quad \left| \frac{1}{\Delta t} \int_{t^{k-1}}^{t^k} \left\{ \psi \frac{\partial u}{\partial \tau}(t, x(t)) - \psi(t^k, x_i, u_i^k) \frac{\partial u}{\partial \tau} \Big|_{(t^k, x_i)} \right\} dt \right| \leq C^{1,5}(h + \Delta t),$$

where $C^{1,5}$ depends on uniform bounds of $\frac{\partial^2 u}{\partial \tau^2}$ along the characteristics (hence it depends on the bounds of u_{tt}, u_{tx} and u_{xx} and b). Combining (3.17)—(3.20), we have

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

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. \square

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

$$(3.21) \quad w^{h, \Delta t}(t, x) = \begin{cases} \sum_{k=1}^N w^k(x) 1_{(t^{k-1}, t^k]}(t), & t \in (0, T]; \\ w^0(x), & t = 0, \end{cases}$$

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) and that $h > C\Delta t$, where the constant C is defined in Lemma 3.1. Then,*

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

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

$$(3.22) \quad \begin{cases} \frac{\zeta_i^k - (\bar{u}_i^{k-1} - \bar{w}_i^{k-1})}{\Delta t} = \frac{1}{2} \left\{ (\sigma(u)_i^k)^2 \delta_x^2(u)_i^k - (\sigma_i^k)^2 \delta_x^2(w)_i^k \right\} + [\widehat{b}(u)_i^k - \widehat{b}_i^k] + e_i^k; \\ \zeta_i^0 = 0. \end{cases}$$

Note that

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

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

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

We can rewrite (3.22) as

$$(3.23) \quad \begin{cases} \frac{\zeta_i^k - \bar{\zeta}_i^{k-1}}{\Delta t} = \frac{1}{2} (\sigma_i^k)^2 \delta_x^2(\zeta)_i^k + I_i^k + e_i^k, \\ \zeta_i^0 = 0, \end{cases}$$

where

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

Now, using an argument similar to those used in Lemma 3.1 and Theorem 2.3, we find constants C_2 and $C_3 > 0$, independent of k and i , such that

$$(3.24) \quad |I_i^k| \leq C_2 |\zeta_i^{k-1}| + C_3(h + \Delta t).$$

Note that it follows from (3.23) that

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

so that, if we set $\|\zeta^k\| = \max_i |\zeta_i^k|$, a maximum principle argument, together with Lemma 3.1 and (3.24), shows that

$$(3.25) \quad \begin{aligned} \|\zeta^k\| &\leq \max_i |\bar{\zeta}_i^{k-1}| + \max_i \left\{ |I_i^k| + |e_i^k| \right\} \Delta t \\ &\leq \max_i |\bar{\zeta}_i^{k-1}| + C_2 \|\zeta^{k-1}\| \Delta t + (C_1 + C_3)(h + \Delta t) \Delta t, \end{aligned}$$

where C_1 is the constant in Lemma 3.1 and C_3 is that in (3.24). In order to estimate $\max_i |\bar{\zeta}_i^{k-1}|$, we adopt the argument in [4]. 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

$$(3.26) \quad \max_i |\bar{\zeta}_i^{k-1}| \leq \max_i |\zeta_i^{k-1}| + \max_i |u(t^{k-1}, \bar{x}_i^k) - I_1(u)(t^{k-1}, \bar{x}_i^k)|.$$

Apply the Peano Kernel Theorem (cf. [3] or [4]) to show that

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

where $h^* = \mathcal{O}(\Delta t)$ and $C_4 > 0$ is independent of k and i . This, together with (3.26), amounts to saying that (3.25) can be rewritten as

$$(3.27) \quad \begin{aligned} \|\zeta^k\| &\leq \|\zeta^{k-1}\| + C_2 \|\zeta^{k-1}\| \Delta t + C_5(h + \Delta t) \Delta t, \\ &= \|\zeta^{k-1}\| (1 + C_2 \Delta t) + C_5(h + \Delta t) \Delta t, \end{aligned}$$

where $C_5 \leq 2(C_2 + C_3 + C_4)$ is independent of k . It then follows from the Gronwall lemma and the bound on $\|\zeta^0\|$ that

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

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 is the constant given in Lemma 3.1. It is then obvious that the requirement $h > C\Delta t$ is fulfilled. Now let us choose

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

where $w^{h, \Delta t}$ is defined by (3.21). 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:*

(1) *for fixed $x \in \mathbb{R}$, $u^{(n)}(\cdot, x)$ is left continuous;*

(2) *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);*

(3) $\sup_{t, x} |u^{(n)}(t, x) - u(t, x)| = \mathcal{O}(\frac{1}{n})$.

Proof. The property (1) is obvious by definition (3.21). To see (3), we note that

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

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,

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

by virtue of Theorem 3.3 and the definitions of h and Δt . This proves (3).

To show (2), 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 $i < \ell < j - 1$, Theorem 3.3 implies that

$$\begin{aligned} |w^k(x_\ell) - w^k(x_{\ell+1})| &\leq |w^k(x_\ell) - u(t^k, x_\ell)| + |u(t^k, x_\ell) - u(t^k, x_{\ell+1})| \\ (3.28) \quad &\quad + |u(t^k, x_{\ell+1}) - w^k(x_{\ell+1})| \\ &\leq 2\|\zeta^k\| + \|u_x\|_\infty |x_\ell - x_{\ell+1}| \leq Kh = K(x_{\ell+1} - x_\ell), \end{aligned}$$

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

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

Hence,

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

where K is the same as that in (3.28). Similarly,

$$|w^k(x^2) - w^k(x_j)| \leq K|x^2 - x_j|.$$

Combining the above gives

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

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

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 following form:

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

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 SDEs, 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

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

Next, we set

$$(4.2) \quad \begin{cases} \theta^n(t, x) = u^{(n)}(T - t, x); \\ \tilde{b}^n(t, x) = b(t, x, \theta^n(t, x)); \\ \tilde{\sigma}^n(t, x) = \sigma(t, x, \theta^n(t, x)). \end{cases}$$

By Theorem 3.4, θ^n is left 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 ,

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

Also, from Theorem 3.4,

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

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

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

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

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

It is clear from the standing assumptions (A1)—(A3) and Theorem 3.4 that both \tilde{b} and $\tilde{\sigma}$ are left continuous in t and uniformly Lipschitz in x ; hence, both SDEs (4.5) and (4.6) above 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,*

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

Proof. To simplify notation, we shall suppress the sign “ \sim ” for the coefficients in the sequel. We first rewrite (4.5) as follows:

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

where

$$(4.8) \quad u_t^n = \int_0^t [b^n(\cdot, \bar{X}^n)_{\eta_n(s)} - b^n(s, \bar{X}_s^n)] ds + \int_0^t [\sigma^n(\cdot, \bar{X}^n)_{\eta_n(s)} - \sigma^n(s, \bar{X}_s^n)] dW_s.$$

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

$$(4.9) \quad \begin{aligned} E \left\{ \sup_{s \leq t} |X_s^n - \bar{X}_s^n|^2 \right\} &\leq 3E \left\{ \sup_{s \leq t} |u_s^n|^2 \right\} + 3K^2 t \int_0^t E \left\{ |X_s^n - \bar{X}_s^n|^2 \right\} ds \\ &\quad + 12K^2 \int_0^t E \left\{ |X_s^n - \bar{X}_s^n|^2 \right\} ds. \end{aligned}$$

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

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

and Gronwall’s inequality leads to

$$(4.10) \quad E \left\{ \sup_{s \leq t} |X_s^n - \bar{X}_s^n|^2 \right\} \leq 3e^{3K^2(T+4)} E \left\{ \sup_{s \leq t} |u_s^n|^2 \right\}.$$

Consequently, we turn our attention to $E \left\{ \sup_{s \leq t} |u_s^n|^2 \right\}$. Note that if $s \in [t^{n,k}, t^{n,k+1})$, for some $1 \leq k < n$, 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 $(n-k-1)\Delta t_n < T - s \leq (n-k)\Delta t_n$. Thus, by definitions (3.23) and (4.1), for every $x \in \mathbb{R}$

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

More generally,

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

Using this fact, it is easily seen that

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

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

$$I_1 \leq \int_0^t \left\{ \|b_t\|_\infty |\eta^n(s) - s| + \|b_x\|_\infty |\bar{X}_{\eta^n(s)}^n - X_s^n| \right\} ds,$$

and

$$I_2 \leq K \|b_y\|_\infty \cdot \int_0^t |\bar{X}_{\eta^n(s)}^n - X_s^n| ds.$$

Thus,

$$(4.13) \quad \left| \int_0^t b^n(\cdot, \bar{X}^n)_{\eta^n(s)} - b^n(s, X_s^n) ds \right| \leq \tilde{K} \int_0^t \left\{ |\eta^n(s) - s| + |\bar{X}_{\eta^n(s)}^n - X_s^n| \right\} ds,$$

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

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

$$\begin{aligned}
(4.14) \quad & E \left\{ \sup_{u \leq t} \left| \int_0^t b^n(\cdot, \bar{X}^n)_{\eta^n(s)} - b^n(s, X_s^n) ds \right|^2 \right\} \\
& \leq 2\tilde{K}^2 \left\{ T \int_0^t E |\bar{X}_{\eta^n(s)}^n - X_s^n|^2 ds + \frac{T^4}{4n^2} \right\}.
\end{aligned}$$

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

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

Combining (4.14) and (4.15), we get

$$(4.16) \quad E\left\{\sup_{s \leq t} |u_s^k|^2\right\} \leq \tilde{K}^2(4T + 16) \int_0^t E|\bar{X}_{\eta^n(s)}^n - X_s^n|^2 ds + \tilde{K}^2 T(T + \frac{16}{3}) \frac{1}{n^2}.$$

Thus, by (4.10),

$$(4.17) \quad E\left\{\sup_{s \leq t} |X_s^n - \bar{X}_s^n|^2\right\} \leq 3e^{3K^2(T+4)} \left\{ \tilde{K}^2(4T + 16) \int_0^t E|\bar{X}_{\eta^n(s)}^n - X_s^n|^2 ds + \tilde{K}^2 T(T + \frac{16}{3}) \frac{1}{n^2} \right\}.$$

Finally, noting that $|\bar{X}_{\eta^n(s)}^n - X_s^n| \leq |\bar{X}_{\eta^n(s)}^n - \bar{X}_s^n| + |\bar{X}_s^n - X_s^n|$ and that

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

we see as before that

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

Therefore, (4.17) becomes

$$(4.18) \quad E\left\{\sup_{s \leq t} |X_s^n - \bar{X}_s^n|^2\right\} \leq C_1 \frac{1}{n} + C_2 \frac{1}{n^2} + C_3 \int_0^t E\left\{\sup_{r \leq s} |\bar{X}_r^n - X_r^n|^2\right\} ds,$$

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

$$\alpha_n(t) \leq \beta_n e^{C_T t}, \quad \forall t \in [0, T],$$

where $\beta_n = C_1 n^{-1} + C_2 n^{-2}$ and $C_T = C_3 T$. In particular, by slightly changing the constants, we have

$$(4.19) \quad \alpha_n(T) = E\left\{\sup_{0 \leq t \leq T} |\bar{X}_t^n - X_t^n|^2\right\} \leq \frac{C_1}{n} + \frac{C_2}{n^2} = \mathcal{O}\left(\frac{1}{n}\right),$$

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,

$$(4.20) \quad E \left\{ \sup_{0 \leq t \leq T} |\bar{X}_t^n - X_t| \right\} + E \left\{ \sup_{0 \leq t \leq T} |\bar{Y}_t^n - Y_t| \right\} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right),$$

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,

$$(4.21) \quad \left| E\{f(\bar{X}_T^n)\} - E\{f(X_T)\} \right| \leq \frac{K}{n},$$

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 sign “ \sim ” for \tilde{b} and $\tilde{\sigma}$ to simplify notation. Set

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

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

$$(4.21) \quad \begin{aligned} E \left\{ \sup_{s \leq t} |\bar{X}_s^n - X_s|^2 \right\} &\leq 2E \left\{ \sup_{s \leq t} |\bar{X}_s^n - X_s^n|^2 \right\} + 2E \left\{ \sup_{s \leq t} |X_s^n - X_s|^2 \right\} \\ &= \mathcal{O}\left(\frac{1}{n}\right) + 2E \left\{ \sup_{s \leq t} |X_s^n - X_s|^2 \right\}. \end{aligned}$$

Further, observe that

$$\begin{aligned} &E \left\{ \sup_{s \leq t} |X_s^n - X_s|^2 \right\} \\ &\leq 2T \int_0^t E |b^n(s, X_s^n) - b(s, X_s)|^2 ds + 8 \int_0^t E |\sigma^n(s, X_s^n) - \sigma(s, X_s)|^2 ds \\ &\leq 4T \int_0^t E |b^n(s, X_s^n) - b^n(s, X_s)|^2 ds + 16 \int_0^t E |\sigma^n(s, X_s^n) - \sigma^n(s, X_s)|^2 ds \\ &\quad + 4(T+4) \int_0^t \varepsilon_n(s) ds \\ &\leq 4(T+4)K^2 \int_0^t E \left\{ \sup_{r \leq s} |X_r^n - X_r|^2 \right\} ds + 4(T+4) \int_0^t \varepsilon_n(s) ds. \end{aligned}$$

Applying Gronwall’s inequality, we get

$$(4.22) \quad E \left\{ \sup_{s \leq t} |X_s^n - X_s|^2 \right\} \leq 4(T+4) \int_0^t \varepsilon_n(s) ds \cdot e^{4(T+4)K^2} \leq \frac{\tilde{C}}{n^2},$$

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

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

by Theorem 3.4 and (4.22). The estimate (4.20) then follows from an easy application of 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 $\sup_{(t,x)} |\theta^n(t, x) - \theta(t, x)| \leq Cn^{-1}$. We modify \bar{X}_t^n 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 :

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

It is then standard [10, p.460] that

$$(4.23) \quad \left| E\{f(X_T^n)\} - E\{f(\bar{X}_T^{n,k})\} \right| \leq \frac{C_1}{k}.$$

On the other hand, we have

$$(4.24) \quad |E\{f(X_T)\} - E\{f(X_T^n)\}| \leq KE\{|X_T - X_T^n|\} \leq E\left\{\sup_{0 \leq t \leq T} |X_t - X_t^n|\right\} \leq \frac{C_2}{n}$$

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

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

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

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 §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

$$(5.2) \quad \begin{cases} 0 = \theta_t + \frac{1}{2}\sigma^2(t, x, \theta)\theta_{xx} + b(t, x, \theta, -\sigma(t, x, \theta)\theta_x)\theta_x + \widehat{b}(t, x, \theta, -\sigma(t, x, \theta)\theta_x); \\ \theta(T, x) = g(x). \end{cases}$$

Define b_0 and \widehat{b}_0 by

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

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

$$(5.4) \quad \begin{cases} u_t = \frac{1}{2}\bar{\sigma}^2(t, x, u)u_{xx} + \bar{b}_0(t, x, u, u_x)u_x + \widehat{\bar{b}}_0(t, x, u, u_x); \\ u(0, x) = g(x). \end{cases}$$

We will again drop the sign “ $-$ ” in the sequel. Note that Theorem 2.3 holds for the general case; hence, the solution to (5.4) will be bounded in $C^{2+\frac{\alpha}{2}, 4+\alpha}$ 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:

$$(5.5) \quad \begin{cases} u_t = \frac{1}{2}\bar{\sigma}^2(t, x, u)u_{xx} + b_0(t, x, u, v)u_x + \widehat{b}_0(t, x, u, v); \\ v_t = \frac{1}{2}\bar{\sigma}^2(t, x, u)v_{xx} + B_0(t, x, u, v)v_x + \widehat{B}_0(t, x, u, v); \\ u(0, x) = g(x), \quad v(0, x) = g'(x), \end{cases}$$

where

$$(5.6) \quad \begin{cases} B_0(t, x, y, z) = \sigma(t, x, y)[\sigma_x(t, x, y) + \sigma_y(t, x, y)z] + b(t, x, y, z) + b_z(t, x, y, z) \\ \quad + \widehat{b}_z(t, x, y, z); \\ \widehat{B}_0(t, x, y, z) = [b_x(t, x, y, z) + b_y(t, x, y, z)z]z + \widehat{b}(t, x, y, z)z + \widehat{b}_x(t, x, y, z). \end{cases}$$

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 $\widehat{B}_0(t, x, u(t, x), v(t, x))$ are uniformly bounded for all (t, x) . Also, B_0 and \widehat{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

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

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

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

It is easy to check that the two characteristic curves $\mathcal{C}_i : (t, x_i(t))$, $i = 1, 2$, are determined by the ODEs

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

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

$$(5.7) \quad d\tau_1 = \psi_1(t, x_1(t))dt; \quad d\tau_2 = \psi_2(t, x_2(t))dt,$$

where

$$(5.8) \quad \begin{cases} \psi_1(t, x) = [1 + b_0^2(t, x, u(t, x), v(t, x))]^{1/2}; \\ \psi_2(t, x) = [1 + B_0^2(t, x, u(t, x), v(t, x))]^{1/2}. \end{cases}$$

Thus, along \mathcal{C}_1 and \mathcal{C}_2 , respectively,

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

and (5.5) can be simplified to

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

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

$$(5.10) \quad \begin{cases} (\widehat{b}_0)_i^k = \widehat{b}_0(t^k, x_i, U_i^{k-1}, V_i^{k-1}); & (\widehat{B}_0)_i^k = \widehat{B}_0(t^k, x_i, U_i^{k-1}, V_i^{k-1}); \\ \sigma_i^k = \sigma(t^k, x_i, U_i^{k-1}); \\ \bar{x}_i^k = x_i + b_0(t^k, x_i, U_i^{k-1}, V_i^{k-1})\Delta t; & \bar{\mathbf{x}}_i^k = x_i + B_0(t^k, x_i, U_i^{k-1}, V_i^{k-1})\Delta t, \end{cases}$$

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

Step k: Determine the k -th step grid values (U^k, V^k) by the system of difference equations

$$(5.11) \quad \begin{cases} \frac{U_i^k - \bar{U}_i^{k-1}}{\Delta t} = \frac{1}{2} (\sigma_i^k)^2 \delta_x^2(U)_i^k + (\widehat{b}_0)_i^k; \\ \frac{V_i^k - \bar{V}_i^{k-1}}{\Delta t} = \frac{1}{2} (\sigma_i^k)^2 \delta_x^2(V)_i^k + (\widehat{B}_0)_i^k. \end{cases}$$

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

We shall follow the argument in §3 to prove convergence. We point out that, unlike in the previous case, the functions B_0 and \widehat{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 $Q_M := \{(t, x, y, z) \in [0, T] \times \mathbb{R}^3 : |y| \leq M, |z| \leq M\}$, 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 \widehat{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 $\{a_k : k = 1, \dots, n\}$ is a finite sequence such that $a_k \geq 0, \forall k$, and $a_0 \leq \alpha$. Also, assume that the following recursive relation holds:*

$$(5.13) \quad a_k \leq a_{k-1} + \frac{C}{n}(1 + a_{k-1}^2), \quad k = 1, \dots, n,$$

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 $\sup_n \sup_{0 \leq k \leq n} a_k \leq M$.

Proof. Let $A(\cdot)$ be the solution to the following ODE:

$$(5.14) \quad \begin{cases} \frac{dA(t)}{dt} = \frac{C}{n}(1 + A^2(t)), & 0 \leq t \leq n. \\ A(0) = \alpha, \end{cases}$$

Since $\frac{dA}{dt} > 0$, A is increasing. Thus, for each $k = 1, \dots, n$, it holds that

$$(5.15) \quad A(k) = A(k-1) + \frac{C}{n} \int_{k-1}^k (1 + A^2(r)) dr \geq A(k-1) + \frac{C}{n}(1 + A^2(k-1)).$$

Noting that $A(0) = \alpha \geq a_0$, a simple induction using (5.13) and (5.15) then shows that $A(k) \geq a_k, k = 1, \dots, n$. It suffices to determine the bound for $A(t), 0 \leq t \leq n$. 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

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

Consequently, we obtain that

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

proving the lemma. □

We now give a crucial result of this section.

Theorem 5.4. *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 , \widehat{b} and T , such that*

$$(5.16) \quad \sup_n \sup_{k,i} \{|U_i^k| + |V_i^k|\} \leq M_1.$$

Proof. Let us rewrite (5.11) as

$$\begin{cases} U_i^k = \bar{U}_i^{k-1} + \left\{ \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(U)_i^k + (\widehat{b}_0)_i^k \right\} \Delta t; \\ V_i^k = \bar{V}_i^{k-1} + \left\{ \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(V)_i^k + (\widehat{B}_0)_i^k \right\} \Delta t. \end{cases}$$

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

$$(5.17) \quad \begin{cases} \|U^k\| \leq \|\bar{U}^{k-1}\| + \|(\widehat{b}_0)^k\| \Delta t \leq \|U^{k-1}\| + \|(\widehat{b}_0)^k\| \Delta t; \\ \|V^k\| \leq \|\bar{V}^{k-1}\| + \|(\widehat{B}_0)^k\| \Delta t \leq \|V^{k-1}\| + \|(\widehat{B}_0)^k\| \Delta t. \end{cases}$$

Since \widehat{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 \widehat{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 , \widehat{b} and their first order partial derivatives, such that $|\widehat{B}_0(t, x, y, z)| \leq K(1 + z^2)$. Therefore,

$$|(\widehat{B}_0)_i^k| = |\widehat{B}_0(t^k, x_i, U_i^{k-1}, V_i^{k-1})| \leq K(1 + |V_i^{k-1}|^2),$$

and the second inequality in (5.17) leads to

$$(5.18) \quad \|V^k\| \leq \|V^{k-1}\| + K(1 + \|V^{k-1}\|^2) \Delta t = \|V^{k-1}\| + \frac{TK}{n}(1 + \|V^{k-1}\|^2),$$

since $\Delta t = T/n$. Hence, the result follows from Lemma 5.3. □

We can now follow the arguments in §3 line by line. What follows is essentially a somewhat detailed sketch of the proof of Theorem 5.6 below. First, we evaluate the first equation in (5.9) along \mathcal{C}_1 and the second one along \mathcal{C}_2 to get an analogue of (3.12):

$$\begin{cases} \frac{u_i^k - \widehat{u}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma(u_i^k)^2 \delta_x^2(u_i^k) + \widehat{b}_0(u, v)_i^k + (e_1)_i^k); \\ \frac{v_i^k - \widehat{v}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma(u_i^k)^2 \delta_x^2(u_i^k) + \widehat{B}_0(u, v)_i^k + (e_2)_i^k, \end{cases}$$

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 $\widehat{u}_i^{k-1} = u(t^{k-1}, \widehat{x}_i)$, $\widehat{v}_i^{k-1} = v(t^{k-1}, \widehat{x}_i)$, with

$$\widehat{x}_i^k = x_i + b_0(t^k, x_i, u_i^{k-1}, v_i^{k-1})\Delta t; \quad \widehat{x}_i^k = x_i + B_0(t^k, x_i, u_i^{k-1}, v_i^{k-1})\Delta t.$$

Also, $\sigma(u)_i^k$, $\widehat{b}_0(u, v)_i^k$ and $\widehat{B}_0(u, v)_i^k$ are analogous to σ_i^k , $(\widehat{b}_0)_i^k$ and $(\widehat{B}_0)_i^k$, 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 3.1 to see that

$$(5.19) \quad \sup_{k,i} \{|(e_1)_i^k| + |(e_2)_i^k|\} \leq \mathcal{O}(h + \Delta t),$$

provided $h > C\Delta t$, where C is a constant depending on the bounds of the coefficients and the true solution v .

Remark 5.5. In Lemma 3.1 we required the relation $h > C\Delta t$, where $C = \|b\|_\infty$. 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 , \widehat{b} , σ), their partial derivatives, and the bound on $v = u_x$. But 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 5.4.

We now define as we did in (3.21) the approximate solutions $U^{(n)}$ and $V^{(n)}$ by

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

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 (3.23):

$$\begin{cases} \frac{\xi_i^k - \widehat{\xi}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(\xi)_i^k + (I_1)_i^k + (e_1)_i^k; \\ \frac{\zeta_i^k - \widehat{\zeta}_i^{k-1}}{\Delta t} = \frac{1}{2}(\sigma_i^k)^2 \delta_x^2(\zeta)_i^k + (I_2)_i^k + (e_2)_i^k, \end{cases}$$

where

$$\begin{aligned} (I_1)_i^k &= -\frac{u(t^{k-1}, \widehat{x}_i^k) - u(t^{k-1}, \bar{x}_i^k)}{\Delta t} \\ &\quad + \frac{1}{2}[\sigma^2(t^k, x_i, u_i^{k-1}) - \sigma^2(t^k, x_i, U_i^{k-1})]\delta_x^2(u)_i^k + [\widehat{b}_0(u, v)_i^k - (\widehat{b}_0)_i^k]; \\ (I_2)_i^k &= -\frac{v(t^{k-1}, \widehat{x}_i^k) - v(t^{k-1}, \bar{x}_i^k)}{\Delta t} \\ &\quad + \frac{1}{2}[\sigma^2(t^k, x_i, v_i^{k-1}) - \sigma^2(t^k, x_i, V_i^{k-1})]\delta_x^2(v)_i^k + [\widehat{B}_0(u, v)_i^k - (\widehat{B}_0)_i^k]; \end{aligned}$$

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

$$(5.21) \quad |(I_1)_i^k| \leq C_2\{|\xi_i^{k-1}| + |\zeta_i^{k-1}|\} + C_3(h + \Delta t), \quad \forall k, i.$$

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 \widehat{B}_0 (in y and z) to get

$$|\widehat{B}_0(u, v)_i^k - (\widehat{B}_0)_i^k| \leq C_4(|\xi_i^{k-1}| + |\zeta_i^{k-1}|), \quad \forall k, i,$$

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

$$(5.22) \quad |(I_2)_i^k| \leq C'_2\{|\xi_i^{k-1}| + |\zeta_i^{k-1}|\} + C'_3(h + \Delta t), \quad \forall k, i.$$

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

$$\begin{aligned} \|\xi^k\| &\leq \|\xi^{k-1}\| + C_2(\|\xi^{k-1}\| + \|\zeta^{k-1}\|)\Delta t + C_5(h + \Delta t)\Delta t; \\ \|\zeta^k\| &\leq \|\zeta^{k-1}\| + C'_2(\|\xi^{k-1}\| + \|\zeta^{k-1}\|)\Delta t + C'_5(h + \Delta t)\Delta t; \end{aligned}$$

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

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

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

Theorem 5.6. *Suppose that (A1)–(A3) hold. Then,*

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

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

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

and

$$\tilde{b}^n(t, x) = b_0(t, x, \theta^n(t, x), \theta_x^n(t, x)); \quad \tilde{\sigma}^n(t, x) = \sigma(t, x, \theta^n(t, x)),$$

then, it follows from Theorem 5.6 that

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

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 \bar{X}^n denote the solution to the discretized SDE,

$$(5.25) \quad \bar{X}_t^n = x + \int_0^t \tilde{b}^n(\cdot, \bar{X}^n)_{\eta^n(s)} ds + \int_0^t \tilde{\sigma}^n(\cdot, \bar{X}^n)_{\eta^n(s)} dW_s,$$

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 5.7. *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 $(\bar{X}^n, \bar{Y}^n, \bar{Z}^n)$, where \bar{X}^n is the solution to the discretized SDE (5.25) and, for $t \in [0, T]$,*

$$\bar{Y}_t^n := \theta^n(t, \bar{X}_t^n); \quad \bar{Z}_t^n := -\sigma(t, \bar{X}_t^n, \theta^n(t, \bar{X}_t^n))\theta_x^n(t, \bar{X}_t^n),$$

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

$$E\left\{\sup_{0 \leq t \leq T} |\bar{X}_t^n - X_t| + \sup_{0 \leq t \leq T} |\bar{Y}_t^n - Y_t| + \sup_{0 \leq t \leq T} |\bar{Z}_t^n - Z_t|\right\} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

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

$$\left|E\{f(\bar{X}_T^n, \bar{Z}_T^n)\} - E\{f(X_T, Z_T)\}\right| \leq \frac{K}{n},$$

for a constant K . □

References.

- [1] F. Antonelli, *Backward-Forward Stochastic Differential Equations*, *The Annals of Applied Probability*, **3** (1993), pp. 777-793.
- [2] J. Cvitanic and J. Ma, *Hedging Options for a Large Investor and Forward-Backward SDEs*, Preprint (1994).
- [3] P. Davis, *Interpolation and Approximation*, Dover, New York (1975).
- [4] J. Douglas, Jr. and T.F. Russell, *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.*, Vol. 19, (1982), pp. 871-885.
- [5] J. Douglas, Jr., *Finite Difference Methods for Two-Phase Incompressible Flow in Porous Media*, *SIAM J. Numer. Anal.*, Vol. 20, 4 (1983), pp. 681-696.
- [6] D. Duffie and L.G. Epstein, *Stochastic Differential Utility*, *Econometrica*, Vol. 61, **2** (1992), pp. 353-394.
- [7] D. Duffie, P-Y. Geoffard, and C. Skiadas, *Efficient and Equilibrium Allocations with Stochastic Differential Utility*, Preprint (1992).
- [8] D. Duffie, J. Ma and J. Yong, *Black's Console Rate Conjecture*, to appear in *The Annals of Applied Probability* (1994).

- [9] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag (1983).
- [10] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag (1992).
- [11] O.A. Ladyzenskaja, V.A. Solonnikov and N.N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, Providence, R.I: Amer. Math. Soc. (1968).
- [12] J. Ma, P. Protter and J. Yong, *Solving Forward-Backward Stochastic Differential Equations Explicitly—A Four Step Scheme*, *Probability Theory and Related Fields*, 98 (1994), pp. 339–359.