The Annals of Applied Probability
2004, Vol. 14, No. 1, 459–488

A NUMERICAL SCHEME FOR BSDEs

BY JIANFENG ZHANG

University of Southern California, Los Angeles

In this paper we propose a numerical scheme for a class of backward stochastic differential equations (BSDEs) with possible path-dependent terminal values. We prove that our scheme converges in the strong L^2 sense and derive its rate of convergence. As an intermediate step we prove an L^2-type regularity of the solution to such BSDEs. Such a notion of regularity, which can be thought of as the modulus of continuity of the paths in an L^2 sense, is new. Some other features of our scheme include the following: (i) both components of the solution are approximated by step processes (i.e., piecewise constant processes); (ii) the regularity requirements on the coefficients are practically "minimum"; (iii) the dimension of the integrals involved in the approximation is independent of the partition size.

1. Introduction. In this paper we are interested in the following *backward stochastic differential equation* (BSDE, for short):

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)\,dr - \int_t^T Z_r\,dW_r, \tag{1.1}$$

where W is a Brownian motion defined on some complete, filtered probability space $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{0\le t\le T})$ and $\xi \in \mathcal{F}_T$. The BSDEs of this kind, initiated by Bismut [3] and later developed by Pardoux–Peng [20], have been studied extensively in the past decade. We refer the readers to the books of El Karoui–Mazliak [9], Ma–Yong [16] and the survey paper of El Karoui–Peng–Quenez [10] for more information on both theory and application, especially in mathematical finance and stochastic control, for such equations.

A long-standing problem in the theory of BSDEs is to find an implementable numerical method. Many efforts have been made in this direction as well. For example, in the Markovian case, Douglas, Ma and Protter [8] established a numerical method for a class of forward–backward SDEs, a more general version of the BSDEs (1.1), based on a four step scheme developed by Ma–Protter–Yong [15]. In his Ph.D. thesis, Chevance [6] proposed a numerical method for BSDEs by using binomial approach to approximate the process Y. We should point out that, besides the Markovian requirement [i.e., the terminal value ξ has to be of the form $g(X_T)$, where X is some forward diffusion], both methods require rather high regularity of the coefficients.

Received April 2002; revised February 2003.
AMS 2000 subject classifications. Primary 60H10; secondary 65C30.
Key words and phrases. Backward SDEs, L^∞-Lipschitz functionals, step processes, L^2-regularity.

In the non-Markovian case where the terminal value ξ is allowed to depend on the history of a forward diffusion X, Bally [2] presented a discretization scheme and obtained its rate of convergence. The main idea therein was to use a random time partition to overcome some difficulty in approximating the *martingale integrand*—the process Z in (1.1). However, his scheme requires some extra approximations in order to give an actual implementation, and it involves computing multiple integrals whose dimension is proportional to the partition size, which is quite undesirable in implementation. Recently, Briand–Delyon–Mémin [5] and Ma–Protter–San Martin–Torres [14] proposed some numerical methods for BSDEs with path-dependent terminal values. In these works only weak convergence results are obtained. Finally, in [23] Zhang and Zheng suggested another method via PDE approach, which also requires high regularity of coefficients.

We note that the main difficulty in a numerical scheme for BSDEs usually lies in the approximation of the "martingale integrand" Z. In fact, in a sense the problem often comes down to the path regularity of Z. To our best knowledge, most existing methods either require high regularity conditions (e.g., [6, 15]) so as to guarantee the path regularity of the process Z or otherwise lack a good rate of convergence (e.g., [2, 14]).

In this paper we try to find some middle ground among the existing methods. We shall consider a class of BSDEs whose terminal value ξ takes the form $\Phi(X)$, where X is a diffusion process and $\Phi(\cdot)$ is a so-called L^∞-Lipschitz functional (see Section 2 for precise definition). With the help of some results in our previous work [18], we first prove that the martingale integrand Z satisfies a new type of path regularity, called the "*L^2-regularity*" in this paper, which in essence characterizes the modulus of continuity in a mean-square sense. With such a regularity result, and in light of an underlying discretization scheme, we then design a numerical scheme and study its rate of convergence. Our main results are the following: if Φ is an L^∞-Lipschitz functional, then the asymptotic rate of convergence is $\sqrt{(\log n)/n}$; if Φ is a so-called L^1-Lipschitz functional, or is of the form $g(X_T)$, then the rate of convergence will be $1/\sqrt{n}$. We should note that, while the latter rate is more or less standard (cf. [8]), the former rate of convergence, to our best knowledge, is new. In fact, it can be shown that such rate is indeed sharp (see Remark 4.3).

There are several other features of our numerical scheme. First, our approximating solutions are all step processes (i.e., piecewise constant processes), which is rather convenient in implementations for obvious reasons. Second, besides the L^∞ Lipschitz property of Φ, our method virtually requires only Lipschitz conditions, which is needed for the well posedness of the problem. Third, the high dimensionality caused by the non-Markovian nature of the BSDE is overcome significantly. We note that to implement a scheme which involves calculating a certain number of integrals, one needs to deal with two types of high dimensionality: One is the dimension of each integral and the other is the number of integrals involved.

In our scheme, the former dimension is equal to that of W and the latter number, which is exponential to the partition size in general, is linear to that size if the non-Markovian problem can be converted to a Markovian one by adding one (or some) state variable. We should note that this type of non-Markovian problem has lots of applications in finance theory (e.g., lookback options and Asian options) and the idea of converting it to a Markovian one has been exploited by many authors (see, e.g., [4, 7]).

The rest of the paper is organized as follows. In Section 2 we present some preliminaries. In Section 3 we establish the L^2-type regularity for the process Z, as well as that for X and Y. In Section 4 we review the Euler scheme used to discretize the forward diffusion X; then in Section 5 we discretize the BSDE and prove the rate of convergence. In Section 6 the numerical scheme is presented explicitly, with special attention to the BSDEs which can be converted to Markovian ones.

2. Preliminaries. Throughout this paper we let $T > 0$ be a *fixed* terminal time and $(\Omega, \mathcal{F}, \mathbf{F}, P)$ be a complete, filtered probability space on which is defined a standard Brownian motion W, such that $\mathbf{F} = \{\mathcal{F}_t\}_{0\le t\le T}$ is the natural filtration of W, augmented by all the P-null sets.

The following spaces will be frequently used in the sequel: Let $\mathbb{E}$ denote a generic Euclidean space with inner product $\langle \cdot, \cdot \rangle$ and norm $|\cdot|$.

- $\mathbb{D}$ is the space of all càdlàg functions defined on $[0, T]$.
- $C_{\mathrm{b}}^m([0, T] \times \mathbb{E})$ is the space of all continuous functions $\varphi : [0, T] \times \mathbb{E} \mapsto \mathbb{R}$, such that φ has uniformly bounded derivatives with respect to the spatial variables up to order m. We often denote $C_{\mathrm{b}}^m = C_{\mathrm{b}}^m([0, T] \times \mathbb{E})$ for simplicity, when the context is clear.
- $C^{1/2,1}([0, T] \times \mathbb{E})$ is the space of all continuous functions $\varphi : [0, T] \times \mathbb{E} \mapsto \mathbb{R}$, such that φ is uniformly $\frac{1}{2}$-Hölder continuous in t and uniformly Lipschitz continuous in the spatial variables. Again, we often denote $C^{1/2,1} = C^{1/2,1}([0, T] \times \mathbb{E})$ for simplicity, when the context is clear.
- For $1 \le p < \infty$, $L^p(\mathcal{F}_T)$ is the space of all $\mathcal{F}_T$-measurable and L^p-integrable random variables; $L^p(\mathbf{F})$ is the space of all $\mathbf{F}$-adapted processes ξ satisfying $\|\xi\|_{p,T}^p \triangleq E\{\int_0^T |\xi_t|^p \, dt\} < \infty$.

We consider the following (decoupled) forward–backward SDE:

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

where b, σ and f are deterministic functions and Φ is a deterministic *functional*. To simplify presentations, in what follows we assume that $X_t \in \mathbb{R}^d$, and W_t, Y_t and Z_t are all one-dimensional (noting that X and W may have different dimensions). But the results can be extended to cases with higher-dimensional W, Y and Z without significant difficulties. For simplicity we also denote the solution to (2.1) by $\Theta \stackrel{\triangle}{=} (X, Y, Z)$.

DEFINITION 2.1. A functional $\Phi : \mathbb{D}^d \mapsto \mathbb{R}$ is called L^∞-Lipschitz, if there exists a constant K such that

$$|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)| \le K \sup_{0\le t\le T} |\mathbf{x}_1(t) - \mathbf{x}_2(t)| \qquad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{D}^d; \tag{2.2}$$

and Φ is called L^1-Lipschitz, if it satisfies the following estimate:

$$|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)| \le K \int_0^T |\mathbf{x}_1(t) - \mathbf{x}_2(t)|\, dt \qquad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{D}^d. \tag{2.3}$$

Two typical examples of L^∞-Lipschitz and L^1-Lipschitz continuous functionals are $\Phi(\mathbf{x}) = \max_{0\le t\le T} |\mathbf{x}(t)|$ and $\Phi(\mathbf{x}) = \int_0^T \mathbf{x}(t)\, dt$, motivated by lookback options and Asian options, respectively. The following approximation result, due to Ma–Zhang [18], for L^∞-Lipschitz functional will be useful in the sequel.

LEMMA 2.2. *Suppose that Φ is an L^∞-Lipschitz functional satisfying the condition* (2.2). *Let $\Pi = \{\pi\}$ be a family of partitions of* $[0, T]$. *Then there exists a family of discrete functionals $\{g_\pi : \pi \in \Pi\}$ such that*

(i) *for each $\pi \in \Pi$, assuming $\pi : 0 = t_0 < \cdots < t_n = T$, we have that $g_\pi \in C_b^\infty(\mathbb{R}^{d(n+1)})$, and satisfies*

$$\sum_{i=0}^{n} |\partial_{x_i} g_\pi(x)| \le K \qquad \forall x = (x_0, \ldots, x_n) \in \mathbb{R}^{d(n+1)}, \tag{2.4}$$

where K is the same constant as that in (2.2).

(ii) *for any $\mathbf{x} \in \mathbb{D}^d$, it holds that*

$$\lim_{|\pi|\to 0} |g_\pi(\mathbf{x}(t_0), \ldots, \mathbf{x}(t_n)) - \Phi(\mathbf{x})| = 0. \tag{2.5}$$

We shall make use of the following *Standing assumptions*:

ASSUMPTION 2.3. The functions $b, \sigma, f \in C^{1/2,1}$; and Φ is an L^∞-Lipschitz functional. We use a common constant $K > 0$ to denote all the Lipschitz constants and assume that

$$\sup_{0\le t\le T} \{|b(t, 0)| + |\sigma(t, 0)| + |f(t, 0, 0, 0)|\} + |\Phi(\mathbf{0})| \le K,$$

where $\mathbf{0}$ is the constant function taking value 0 on $[0, T]$.

The following lemma collects some standard results in SDE and BSDE literature. We list them for ready references.

LEMMA 2.4. *Suppose that* $\widetilde{b}, \widetilde{\sigma} : \Omega \times [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ *and* $\widetilde{f} : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ *are* $\mathbf{F}$*-adapted random fields, such that*:

(a) *they are uniformly Lipschitz continuous with respect to* $x \in \mathbb{R}^d$, $y \in \mathbb{R}$ *and* $z \in \mathbb{R}$, *with the common Lipschitz constant* $K > 0$;

(b) $\widetilde{b}(t, 0), \widetilde{\sigma}(t, 0), \widetilde{f}(t, 0, 0) \in L^2(\mathbf{F})$.

For any $\xi \in L^2(\mathcal{F}_T)$, *denote* $\Theta = (X, Y, Z)$ *be the solution to following FBSDE*:

$$X_t = x + \int_0^t \widetilde{b}(s, X_s)\, ds + \int_0^t \widetilde{\sigma}(s, X_s)\, dW_s, \tag{2.6}$$

$$Y_t = \xi + \int_t^T \widetilde{f}(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s. \tag{2.7}$$

Then, we have the following estimates:

(i) *for any* $p \geq 2$, *there exists a constant* $C_p > 0$, *depending only on* T, K *and* p, *such that*

$$\begin{aligned} &E\Big\{\sup_{0\le t\le T} |X_t|^p\Big\} \\ &\quad \le C_p E\Big\{|x|^p + \int_0^T \big[|\widetilde{b}(t,0)|^p + |\widetilde{\sigma}(t,0)|^p\big]\, dt\Big\}, \end{aligned} \tag{2.8}$$

$$\begin{aligned} &E\Big\{\sup_{0\le t\le T} |Y_t|^p + \Big(\int_0^T |Z_t|^2\, dt\Big)^{p/2}\Big\} \\ &\quad \le C_p E\Big\{|\xi|^p + \int_0^T |\widetilde{f}(t,0,0)|^p\, dt\Big\}, \end{aligned} \tag{2.9}$$

$$\begin{aligned} &E\{|X_t - X_s|^p\} \\ &\quad \le C_p E\Big\{|x|^p + \sup_{0\le t\le T} |\widetilde{b}(t,0)|^p + \sup_{0\le t\le T} |\widetilde{\sigma}(t,0)|^p\Big\}|t-s|^{p/2}, \end{aligned} \tag{2.10}$$

$$\begin{aligned} &E\{|Y_t - Y_s|^p\} \\ &\quad \le C_p E\Big\{\Big[|\xi|^p + \sup_{0\le t\le T} |\widetilde{f}(t,0,0)|^p\Big]|t-s|^{p-1} \\ &\qquad\qquad + \Big\{\int_s^t |Z_r|^2\, dr\Big\}^{p/2}\Big\}. \end{aligned} \tag{2.11}$$

(ii) (*Stability*) *Let* $\Theta^\varepsilon = (X^\varepsilon, Y^\varepsilon, Z^\varepsilon)$ *be the solution to the perturbed FBSDE* (2.6) *and* (2.7) *in which the coefficients are replaced by* $\widetilde{b}^\varepsilon, \widetilde{\sigma}^\varepsilon, \widetilde{f}^\varepsilon$, *with initial* *state* x^ε *and terminal value* ξ^ε. *Assume that the assumptions* (a) *and* (b) *hold for all coefficients* b^ε, σ^ε *and* f^ε *and assume that* $\lim_{\varepsilon\to 0} x^\varepsilon = x$, *and for fixed* (x, y, z),

$$\lim_{\varepsilon\to 0} E\{|\widetilde{b}^\varepsilon(t,x) - \widetilde{b}(t,x)|^2 + |\widetilde{\sigma}^\varepsilon(t,x) - \widetilde{\sigma}(t,x)|^2\} = 0,$$

$$\lim_{\varepsilon\to 0} E\left\{|\xi^\varepsilon - \xi|^2 + \int_0^T |\widetilde{f}^\varepsilon(t,y,z) - \widetilde{f}(t,y,z)|^2\,dt\right\} = 0.$$

Then, *we have*

$$\lim_{\varepsilon\to 0} E\left\{\sup_{0\le t\le T}|X_t^\varepsilon - X_t|^2 + \sup_{0\le t\le T}|Y_t^\varepsilon - Y_t|^2 + \int_0^T |Z_t^\varepsilon - Z_t|^2\,dt\right\} = 0.$$

The next lemma contains some deeper results on the structure of the solution (Y, Z) to the BSDE (2.7). The proof of these results can be found in [17, 18].

LEMMA 2.5. *Assume that Assumption* 2.3 *holds*.

(i) *Suppose that* $\dim(X) = \dim(W)$, $\sigma\sigma^T \ge \delta I_d$ *for some constant* $\delta > 0$ *and that* Φ *satisfies the* L^∞*-Lipschitz condition* (2.2), *then the process* Z *admits a càdlàg version*.

(ii) *Suppose that* Φ *takes the special form*: $\Phi(X) = g(X_{t_0}, \dots, X_{t_n})$ *and that* $b, \sigma, f, g \in C_{\mathrm{b}}^1$. *Denote* $\nabla^i\Theta \stackrel{\triangle}{=} (\nabla X, \nabla^i Y, \nabla^i Z)$, $i = 0, \dots, n$, *to be the solution to the following* "*variational equations*":

$$\begin{aligned}\nabla X_t &= I_d + \int_0^t \partial_x b(r)\,\nabla X_r\,dr + \int_0^t \partial_x\sigma(r)\nabla X_r\,dW_r,\\ \nabla^i Y_t &= \sum_{j\ge i}\partial_j g\,\nabla X_{t_j} + \int_t^T \partial f(r)\,\nabla^i\Theta_r\,dr - \int_t^T \nabla^i Z_r\,dW_r,\end{aligned}\tag{2.12}$$

where I_d *is the* $d\times d$ *identity matrix*, $\partial_x b(r)$ *is a* $d\times d$ *matrix whose* j*th column is* $\partial_j b(r, X_r)$, $\partial\sigma$ *and* ∂f *are defined in a similar manner, and* $\partial f(r)\,\nabla^i\Theta_r \stackrel{\triangle}{=} \partial_x f(r)\,\nabla X_r + \partial_y f(r)\,\nabla^i Y_r + \partial_z f(r)\,\nabla^i Z_r$. *Then*, *for* $\forall\, t \in [0, T]$, *it holds that*

$$Z_t = \nabla Y_t\,[\nabla X_t]^{-1}\sigma(t, X_t),\tag{2.13}$$

where

$$\nabla Y_t = \sum_{i=1}^n \nabla^i Y_t \mathbb{1}_{[t_{i-1},t_i)}(t) + \nabla^n Y_{T-}\,\mathbb{1}_{\{T\}}(t).\tag{2.14}$$

REMARK 2.6. (i) Parts (i) and (ii) in Lemma 2.4 are standard for SDEs and BSDEs (see, e.g., [10, 11]). The part (ii) can be found in [16].

(ii) Lemma 2.5 gives sufficient conditions for the path regularity of Z. By Remark 3.2.3 of [22], Z is also càdlàg if, in addition to Assumption 2.3, σ is uniformly Lipschitz with respect to t.

To end this section, we give a useful representation formula for the martingale integrand Z in the simplest BSDE case.

LEMMA 2.7. *Assume that g is a Lipschitz continuous function and that $g(W_T) = E\{g(W_T)\} + \int_0^T \eta_t \, dW_t$ for some predictable process η. Then η_t is a martingale and $\eta_0 = \frac{1}{T} E\{g(W_T)W_T\}$.*

PROOF. Since g is Lipschitz, by [19] we know that there exists $\zeta \in L^2(\mathcal{F}_T)$ such that $\eta_t = E\{\zeta|\mathcal{F}_t\}$, thus η_t is a martingale. The formula for η_0 is a result of [17]. □

3. L^2-regularity. In this section we establish the first main result of this paper, which we shall call the *L^2-regularity* of the martingale integrand Z. Such a regularity, combined with the estimate for X and Y (3.2) below, plays a key role for deriving the rate of convergence of our numerical scheme in Section 5.

To begin with, let $\pi : 0 = t_0 < \cdots < t_n = T$ be a partition of $[0, T]$. We would like to estimate $\|Z_t - Z_{t_{i-1}}\|_2$ in terms of the partition size $|\pi| \overset{\triangle}{=} \max_i \Delta t_i$, where $\Delta t_i \overset{\triangle}{=} t_i - t_{i-1}$. We note that in general one cannot expect an estimate of $E\{|Z_t - Z_s|^2\}$ as strong as that for X (2.10) or that for Y (2.11). In light of the norm for the process Z in (2.9), we will try to derive an estimate in the space of $\bigoplus_{i=1}^n L^2([t_{i-1}, t_i] \times \Omega)$.

Our main result is the following theorem.

THEOREM 3.1. *Assume that Assumption* 2.3 *holds true and that Z is càdlàg. Let π be any partition of* $[0, T]$, *then the following estimate holds*:

$$\sum_{i=1}^n E\left\{\int_{t_{i-1}}^{t_i} [|Z_t - Z_{t_{i-1}}|^2 + |Z_t - Z_{t_i}|^2] \, dt\right\} \leq C(1 + |x|^2)|\pi|,$$

where $C > 0$ is a constant depending only on T and K, but independent of the partition π.

We note again that Lemma 2.5 and Remark 2.6 give some sufficient conditions for the path regularity of Z. The proof of Theorem 3.1 is quite lengthy, we split it into several lemmas. The first result is interesting in its own right.

LEMMA 3.2. *Assume Assumption* 2.3 *holds true. Then for $\forall\, p \geq 2$, there exists a constant $C_p > 0$, depending only on T, K and p, such that*,

$$\|Z_t\|_p \leq C_p(1 + |x|) \qquad a.e.\ t \in [0, T]. \tag{3.1}$$

Moreover, for any partition π, we have the following estimate:

$$\max_{1 \leq i \leq n} \sup_{t \in (t_{i-1}, t_i]} E\{|X_t - X_{t_{i-1}}|^2 + |Y_t - Y_{t_{i-1}}|^2\} \leq C(1 + |x|^2)|\pi|, \tag{3.2}$$

where $C > 0$ is a constant depending only on T and K, but independent of the partition π.

PROOF. First we assume that all the conditions in Lemma 2.5(ii) hold true (recalling that a function in $C^{1/2,1}$ is only Lipschitz continuous, but not necessarily differentiable on x, y, z!), and that g satisfies (2.4). For $\forall p \geq 2$, by (2.8) we have $E\{\sup_{0\leq t\leq T} |\nabla X_t|^p\} \leq C_p$ and by (2.9) we get, for $i = 0, \dots, n$,

$$\begin{aligned} &E\left\{\sup_{0\leq t\leq T} |\nabla^i Y_t|^p\right\} \\ &\quad \leq C_p E\left\{\left|\sum_{j\geq i} \partial_j g\, \nabla X_{t_j}\right|^p + \int_0^T |\partial_x f(r)\, \nabla X_r|^p\, dr\right\} \\ &\quad \leq C_p E\left\{\left[\left(\sum_{j=0}^n |\partial_j g|\right)^p + 1\right] \sup_{0\leq t\leq T} |\nabla X_t|^p\right\} \\ &\quad \leq C_p. \end{aligned}$$

Since ∇X is the solution to the linear SDE (2.12), one can easily check that $[\nabla X]^{-1}$ also satisfies a linear SDE and it holds that

$$E\left\{\sup_{0\leq t\leq T} \left|[\nabla X_t]^{-1}\right|^p\right\} \leq C_p.$$

Now recalling (2.13) and (2.14) and applying Hölder's inequality, we have

$$\begin{aligned} \|Z_t\|_p &\leq \|\nabla Y_t\|_{3p} \left\|[\nabla X_t]^{-1}\right\|_{3p} \|\sigma(t, X_t)\|_{3p} \\ &\leq C_p(1 + \|X_t\|_{3p}) \leq C_p(1 + |x|). \end{aligned}$$

For the general case, let $\{\pi\}$ be a sequence of partitions of $[0, T]$, such that $|\pi| \to 0$. By Lemma 2.2, we may choose smooth functions $\{g_\pi\}$ satisfying (2.4) and (2.5). Now let b^π, σ^π, f^π be smooth molifiers of b, σ, f, respectively. For each π, by the above arguments we know $\|Z_t^\pi\|_p \leq C_p(1 + |x|)$, where $C_p > 0$ is independent of π. Moreover, applying Lemma 2.4(ii) we get that

$$\lim_{|\pi|\to 0} E\left\{\int_0^T |Z_t^\pi - Z_t|^2\, dt\right\} = 0. \tag{3.3}$$

Therefore, for a.e. $t \in [0, T]$, there exists a subsequence of $\{\pi\}$ such that $\lim_{|\pi|\to 0} Z_t^\pi = Z_t$, in probability. Applying Fatou's lemma, we then obtain that $\|Z_t\|_p \leq C_p(1 + |x|)$.

To finish the proof we note that in (3.2) the estimate for X is a direct consequence of (2.10), while by applying (3.1) and (2.11) we can easily get $E\{|Y_t - Y_{t_{i-1}}|^2\} \leq C(1 + |x|^2)\, \Delta t_i$. The proof is now complete. □

The following technical lemma is the building block of the proof of Theorem 3.1.

LEMMA 3.3. *Assume* $\Lambda \in L^2(\mathbf{F})$ *and* $\eta^j \in L^2(\mathbf{F})$, $j = 1, \dots, n$. *Denote* $\xi_t^j \triangleq \alpha_j + \int_0^t \eta_r^j \, dW_r$, *where* α_j *are some constants. Then*

$$E\left\{\sum_{i=1}^{n} \int_{t_{i-1}}^{t_{i+1}} \left|\sum_{j \geq i} \eta_r^j\right|^2 dr \, \Lambda_{t_{i-1}}\right\} \leq C E\left\{\left(\sup_{0 \leq t \leq T} \sum_{j=1}^{n} |\xi_t^j|\right)^2 \Lambda_T^*\right\},$$

where $t_{n+1} \triangleq t_n$ *and* $\Lambda_t^* \triangleq \sup_{0 \leq s \leq t} |\Lambda_s|$.

PROOF. By Itô's formula we have

$$E\left\{\int_{t_{i-1}}^{t_{i+1}} \eta_r^{j_1} \eta_r^{j_2} \, dr \middle| \mathcal{F}_{t_{i-1}}\right\} = E\{\xi_{t_{i+1}}^{j_1} \xi_{t_{i+1}}^{j_2} - \xi_{t_{i-1}}^{j_1} \xi_{t_{i-1}}^{j_2} | \mathcal{F}_{t_{i-1}}\}.$$

Note that since $\Lambda_{t_{i-1}}^* \in \mathcal{F}_{t_{i-1}}$, it holds obviously that

$$E\left\{\int_{t_{i-1}}^{t_{i+1}} \eta_r^{j_1} \eta_r^{j_2} \, dr \, \Lambda_{t_{i-1}}^*\right\} = E\{(\xi_{t_{i+1}}^{j_1} \xi_{t_{i+1}}^{j_2} - \xi_{t_{i-1}}^{j_1} \xi_{t_{i-1}}^{j_2}) \Lambda_{t_{i-1}}^*\}.$$

Denote $\Lambda_{t_{n+1}}^* \triangleq \Lambda_T^*$ and $\Lambda_{t_{-1}}^* \triangleq 0$. Since Λ_t^* is increasing, by some simple calculation and applying the Abel transformation one can show that

$$\begin{aligned}
&E\left\{\sum_{i=1}^{n} \int_{t_{i-1}}^{t_{i+1}} \left|\sum_{j \geq i} \eta_r^j\right|^2 dr \, \Lambda_{t_{i-1}}\right\} \\
&\quad \leq E\left\{\sum_{i=1}^{n} \int_{t_{i-1}}^{t_{i+1}} \left|\sum_{j \geq i} \eta_r^j\right|^2 dr \, \Lambda_{t_{i-1}}^*\right\} \\
&\quad = E\left\{\sum_{i=1}^{n} \sum_{j_1, j_2 \geq i} \int_{t_{i-1}}^{t_{i+1}} \eta_r^{j_1} \eta_r^{j_2} \, dr \, \Lambda_{t_{i-1}}^*\right\} \\
&\quad = E\left\{\sum_{i=1}^{n} \sum_{j_1, j_2 \geq i} (\xi_{t_{i+1}}^{j_1} \xi_{t_{i+1}}^{j_2} - \xi_{t_{i-1}}^{j_1} \xi_{t_{i-1}}^{j_2}) \Lambda_{t_{i-1}}^*\right\} \\
&\quad = E\left\{\sum_{j_1, j_2 = 1}^{n} \left[\sum_{1 \leq i \leq j_1 \wedge j_2} \xi_{t_{i+1}}^{j_1} \xi_{t_{i+1}}^{j_2} (\Lambda_{t_{i-1}}^* - \Lambda_{t_{i+1}}^*) \right.\right. \\
&\qquad\qquad + \xi_{t_{j_1 \wedge j_2}}^{j_1} \xi_{t_{j_1 \wedge j_2}}^{j_2} \Lambda_{t_{j_1 \wedge j_2}}^* \\
&\qquad\qquad + \xi_{t_{j_1 \wedge j_2 + 1}}^{j_1} \xi_{t_{j_1 \wedge j_2 + 1}}^{j_2} \Lambda_{t_{j_1 \wedge j_2 + 1}}^* \\
&\qquad\qquad \left.\left. - \xi_{t_0}^{j_1} \xi_{t_0}^{j_2} \Lambda_{t_0}^* - \xi_{t_1}^{j_1} \xi_{t_1}^{j_2} \Lambda_{t_1}^* \right]\right\}.
\end{aligned} \tag{3.4}$$

Note that

$$
\begin{aligned}
E\Bigg\{ \sum_{j_1,j_2=1}^{n} \xi^{j_1}_{t_{j_1\wedge j_2}} \xi^{j_2}_{t_{j_1\wedge j_2}} \Lambda^*_{t_{j_1\wedge j_2}} \Bigg\}
&\le 2E\Bigg\{ \sum_{j_1=1}^{n} \sum_{j_2=j_1}^{n} \xi^{j_1}_{t_{j_1}} \xi^{j_2}_{t_{j_1}} \Lambda^*_{t_{j_1}} \Bigg\} \\
&= 2E\Bigg\{ \sum_{j_1=1}^{n} \sum_{j_2=j_1}^{n} \xi^{j_1}_{t_{j_1}} \xi^{j_2}_{T} \Lambda^*_{t_{j_1}} \Bigg\} \\
&= 2E\Bigg\{ \sum_{j_1=1}^{n} \sum_{j_2=j_1}^{n} E\Big\{ \xi^{j_1}_{T} \sqrt{\Lambda^*_{t_{j_1}}} \Big| \mathcal{F}_{t_{j_1}} \Big\} \xi^{j_2}_{T} \sqrt{\Lambda^*_{t_{j_1}}} \Bigg\} \\
&\le 2E\Bigg\{ \sum_{j_1=1}^{n} E\Big\{ |\xi^{j_1}_{T}| \sqrt{\Lambda^*_{T}} \Big| \mathcal{F}_{t_{j_1}} \Big\} \sum_{j_2=j_1}^{n} |\xi^{j_2}_{T}| \sqrt{\Lambda^*_{T}} \Bigg\} \\
&\le 2E\Bigg\{ \sum_{j_1=1}^{n} E\Big\{ |\xi^{j_1}_{T}| \sqrt{\Lambda^*_{T}} \Big| \mathcal{F}_{t_{j_1}} \Big\} \sum_{j=1}^{n} |\xi^{j}_{T}| \sqrt{\Lambda^*_{T}} \Bigg\} \\
&= 2E\Bigg\{ \sum_{j_1=1}^{n} |\xi^{j_1}_{T}| \sqrt{\Lambda^*_{T}}\, E\Bigg\{ \sum_{j=1}^{n} |\xi^{j}_{T}| \sqrt{\Lambda^*_{T}} \Big| \mathcal{F}_{t_{j_1}} \Bigg\} \Bigg\} \\
&\le 2E\Bigg\{ \sum_{j_1=1}^{n} |\xi^{j_1}_{T}| \sqrt{\Lambda^*_{T}} \sup_{0\le t\le T} E\Bigg\{ \sum_{j=1}^{n} |\xi^{j}_{T}| \sqrt{\Lambda^*_{T}} \Big| \mathcal{F}_{t} \Bigg\} \Bigg\} \\
&\le 2E\Bigg\{ \Bigg(\sup_{0\le t\le T} E\Bigg\{ \sum_{j=1}^{n} |\xi^{j}_{T}| \sqrt{\Lambda^*_{T}} \Big| \mathcal{F}_{t} \Bigg\} \Bigg)^2 \Bigg\} \\
&\le CE\Bigg\{ \Bigg(\sum_{j=1}^{n} |\xi^{j}_{T}| \sqrt{\Lambda^*_{T}} \Bigg)^2 \Bigg\} \\
&= CE\Bigg\{ \Bigg(\sum_{j=1}^{n} |\xi^{j}_{T}| \Bigg)^2 \Lambda^*_{T} \Bigg\},
\end{aligned}
$$

where the last inequality is thanks to Doob's inequality. Analogously we have

$$
E\Bigg\{ \sum_{j_1,j_2=1}^{n} \xi^{j_1}_{t_{j_1\wedge j_2+1}} \xi^{j_2}_{t_{j_1\wedge j_2+1}} \Lambda^*_{t_{j_1\wedge j_2+1}} \Bigg\} \le CE\Bigg\{ \Bigg(\sum_{j=1}^{n} |\xi^{j}_{T}| \Bigg)^2 \Lambda^*_{T} \Bigg\}.
$$

Then (3.4) leads to that

$$
\begin{aligned}
&E\Biggl\{\sum_{i=1}^{n}\int_{t_{i-1}}^{t_{i+1}}\Biggl|\sum_{j\ge i}\eta_r^j\Biggr|^2 dr\,\Lambda_{t_{i-1}}\Biggr\}\\
&\qquad\le E\Biggl\{\sum_{i=1}^{n}\Biggl(\sum_{j=1}^{n}|\xi_{t_{i+1}}^j|\Biggr)^2(\Lambda_{t_{i+1}}^*-\Lambda_{t_{i-1}}^*)\\
&\qquad\qquad+C\Lambda_T^*\Biggl(\sum_{j=1}^{n}|\xi_T^j|\Biggr)^2-\Biggl(\sum_{j=1}^{n}|\xi_{t_0}^j|\Biggr)^2\Lambda_0^*-\Biggl(\sum_{j=1}^{n}|\xi_{t_1}^j|\Biggr)^2\Lambda_{t_1}^*\Biggr\}\\
&\qquad\le E\Biggl\{\sup_{0\le t\le T}\Biggl(\sum_{j=1}^{n}|\xi_t^j|\Biggr)^2\Biggl[\sum_{i=0}^{n}(\Lambda_{t_{i+1}}^*-\Lambda_{t_{i-1}}^*)+C\Lambda_T^*\Biggr]\Biggr\}\\
&\qquad\le CE\Biggl\{\sup_{0\le t\le T}\Biggl(\sum_{j=1}^{n}|\xi_t^j|\Biggr)^2\Lambda_T^*\Biggr\}.
\end{aligned}
$$

This proves the lemma. □

Let $\xi^n=\xi$ and $\xi^j=0$ for $j=1,\dots,n-1$. Then the following result is a direct consequence of Lemma 3.3.

COROLLARY 3.4. *If* $\xi=\alpha+\int_0^t\eta_r\,dW_r$ *and* $\eta,\Lambda\in L^2(\mathbf{F})$, *then*

$$
E\Biggl\{\sum_{i=1}^{n}\int_{t_{i-1}}^{t_{i+1}}|\eta_r|^2\,dr\,\Lambda_{t_{i-1}}\Biggr\}\le CE\Bigl\{\sup_{0\le t\le T}|\xi_t|^2\Lambda_T^*\Bigr\}.
$$

We now turn to the proof of Theorem 3.1. We would like to use (3.3) and take advantage of the representation formula (2.13) of Z^π. But one should be careful that (3.3) does *not* imply $\lim_{|\pi|\to 0}E\{|Z_{t_{i-1}}^\pi-Z_{t_{i-1}}|^2\}=0$.

PROOF OF THEOREM 3.1. We fix a partition $\pi_0: 0=t_0<\cdots<t_n=T$ and will prove the theorem for π_0.

Let $\pi: 0=s_0<\cdots<s_m=T$ be any partition of $[0,T]$ finer than π_0 and without loss of generality, we assume $t_i=s_{l_i}$ for $i=1,\dots,n$. For $\varphi=b,\sigma,f$, let $\varphi^\pi\in C_{\mathrm{b}}^1$ be smooth molifiers of φ such that the derivatives of φ are bounded by K and $\lim_{|\pi|\to 0}\varphi^\pi=\varphi$. Moreover, since Φ satisfies the L^∞-Lipschitz condition (2.2), by virtue of Lemma 2.2 one can find $g^\pi\in C^1(\mathbb{R}^{d(m+1)})$ satisfying (2.4) and (2.5). Let $\Theta^\pi\triangleq(X^\pi,Y^\pi,Z^\pi)$ denote the adapted solution to the

following FBSDE:

$$
\begin{aligned}
X_t^\pi &= x + \int_0^t b^\pi(r, X_r^\pi)\,dr + \int_0^t \sigma^\pi(r, X_r^\pi)\,dW_r,\\
Y_t^\pi &= g^\pi\big(X_{s_0}^\pi, \dots, X_{s_m}^\pi\big) + \int_t^T f^\pi(r, \Theta_r^\pi)\,dr - \int_t^T Z_r^\pi\,dW_r.
\end{aligned} \tag{3.5}
$$

Now by (2.5), applying Lemma 2.4(ii) we know that

$$
\lim_{|\pi|\to 0} E\left\{\sup_{0\le t\le T}\big[|X_t^\pi - X_t|^2 + |Y_t^\pi - Y_t|^2\big] + \int_0^T |Z_t^\pi - Z_t|^2\,dt\right\} = 0. \tag{3.6}
$$

By (3.6) there exists a subsequence of $\{\pi\}$ such that $\lim_{|\pi|\to 0} E\{|Z_t^\pi - Z_t|^2\} = 0$, for dt-a.s. t. From now on we always assume π is in that subsequence. Obviously we may find a sequence $r_k \downarrow 0$ such that

$$
\lim_{|\pi|\to 0} E\{|Z_{t_i+r_k}^\pi - Z_{t_i+r_k}|^2\} = 0 \qquad \forall i, \forall k. \tag{3.7}
$$

Without loss of generality, we assume $t_i + r_k \in (t_i, t_{i+1})$, for all i and k. Note that, for $t \in [t_{i-1}, t_i)$,

$$
\begin{aligned}
&E\{|Z_t - Z_{t_{i-1}}|^2 + |Z_t - Z_{t_i}|^2\}\\
&\quad \le CE\Big\{|Z_t - Z_t^\pi|^2 + |Z_t^\pi - Z_{t_{i-1}}^\pi|^2 + |Z_{t_{i-1}}^\pi - Z_{t_{i-1}+r_k}^\pi|^2\\
&\qquad + |Z_{t_{i-1}+r_k}^\pi - Z_{t_{i-1}+r_k}|^2 + |Z_{t_{i-1}+r_k} - Z_{t_{i-1}}|^2\\
&\qquad + |Z_{t_{i-1}}^\pi - Z_{t_i+r_k}^\pi|^2 + |Z_{t_i+r_k}^\pi - Z_{t_i+r_k}|^2 + |Z_{t_i+r_k} - Z_{t_i}|^2\Big\}.
\end{aligned} \tag{3.8}
$$

By (3.6), (3.7) and the right continuity of Z, to prove the theorem it suffices to estimate $E\{|Z_t^\pi - Z_{t_{i-1}}^\pi|^2\}$ for $t \in (t_{i-1}, t_{i+1})$.

To this end we recall Lemma 2.5 with the coefficients φ replaced by φ^π and denote $\nabla X^\pi, \nabla Y^\pi$ as the terms corresponding to ∇X and ∇Y, respectively. For the convenience of application, we shall rewrite ∇Y^π in some other form. Note that for each i (2.12) is linear. Let (γ^0, ζ^0) and (γ^j, ζ^j), $j = 1, \dots, m$, be the adapted solutions to the BSDEs

$$
\begin{aligned}
\gamma_t^0 &= \int_t^T [f_x^\pi(r)\,\nabla X_r^\pi + f_y^\pi(r)\gamma_r^0 + f_z^\pi(r)\zeta_r^0]\,dr - \int_t^T \zeta_r^0\,dW_r,\\
\gamma_t^j &= \partial_j g^\pi\,\nabla X_{t_j}^\pi + \int_t^T [f_y^\pi(r)\gamma_r^j + f_z^\pi(r)\zeta_r^j]\,dr - \int_t^T \zeta_r^j\,dW_r,
\end{aligned} \tag{3.9}
$$

respectively, then we have the following decomposition:

$$
\nabla^i Y_t^\pi = \gamma_t^0 + \sum_{j=i}^m \gamma_t^j, \qquad t \in [s_{i-1}, s_i). \tag{3.10}
$$

We may simplify (3.10) further. Let us define

$$
\begin{aligned}
\Lambda_t &\triangleq \exp\left\{-\int_0^t f_y^\pi(r)\,dr\right\},\\
M_t &\triangleq \exp\left\{\int_0^t f_z^\pi(r)\,dW_r - \frac{1}{2}\int_0^t |f_z^\pi(r)|^2\,dr\right\}.
\end{aligned}
\tag{3.11}
$$

Since f_z^π is uniformly bounded, by Girsanov's theorem (see, e.g., [12]) we know that M is a P-martingale on $[0,T]$, and $\widetilde{W}_t \triangleq W_t - \int_0^t f_z^\pi(r)\,dr$, $t\in[0,T]$, is an $\mathbf{F}$-Brownian motion on the new probability space $(\Omega,\mathcal{F},\widetilde{P})$, where $\widetilde{P}$ is defined by $\frac{d\widetilde{P}}{dP} = M_T$. Moreover, noting that f_y^π and f_z^π are uniformly bounded, one can deduce easily from (3.11) that, for $\forall\, p\ge 1$, there exists a constant C_p depending only on T, K and p, such that

$$
\begin{aligned}
&\sup_{0\le t\le T}\big[|\Lambda_t|^p + |\Lambda_t^{-1}|^p\big] \le C_p,\\
&E\Big\{\sup\nolimits_{0\le t\le T}\big[|M_t|^p + |M_t^{-1}|^p\big]\Big\} \le C_p,\\
&|\Lambda_t-\Lambda_s|^p + |\Lambda_t^{-1}-\Lambda_s^{-1}|^p \le C_p|t-s|^p,\\
&E\big\{|M_t-M_s|^p + |M_t^{-1}-M_s^{-1}|^p\big\} \le C_p|t-s|^{p/2}.
\end{aligned}
\tag{3.12}
$$

Now we define

$$
\begin{aligned}
&\widetilde{\xi}^0 \triangleq \int_0^T f_x^\pi(r)\,\nabla X_r^\pi\,\Lambda_r^{-1}\,dr, \qquad \widetilde{\zeta}_t^0 \triangleq \zeta_t^0\Lambda_t^{-1},\\
&\widetilde{\gamma}_t^0 \triangleq \gamma_t^0\Lambda_t^{-1} + \int_0^t f_x^\pi(r)\,\nabla X_r^\pi\,\Lambda_r^{-1}\,dr,\\
&\widetilde{\xi}^i \triangleq \partial_i g^\pi\,\nabla X_{s_i}^\pi\Lambda_T^{-1}, \qquad \widetilde{\zeta}_t^i \triangleq \zeta_t^i\Lambda_t^{-1}, \qquad \widetilde{\gamma}_t^i \triangleq \gamma_t^i\Lambda_t^{-1}.
\end{aligned}
\tag{3.13}
$$

Then by (3.9) we have, for $i=0,\dots,m$,

$$
\widetilde{\gamma}_t^i = \widetilde{\xi}^i - \int_t^T \widetilde{\zeta}_r^i\,d\widetilde{W}_r, \qquad t\in[0,T].
$$

Therefore, by the Bayes rule (see, e.g., [12], Lemma 3.5.3) we have, for $t\in[0,T]$ and $i=1,\dots,m$,

$$
\begin{aligned}
\gamma_t^i &= \widetilde{\gamma}_t^i\Lambda_t = \widetilde{E}\{\widetilde{\xi}^i|\mathcal{F}_t\}\Lambda_t\\
&= E\{M_T\widetilde{\xi}^i|\mathcal{F}_t\}M_t^{-1}\Lambda_t = \xi_t^i M_t^{-1}\Lambda_t,\\
\gamma_t^0 &= \widetilde{\gamma}_t^0\Lambda_t - \int_0^t f_x^\pi(r)\,\nabla X_r^\pi\,\Lambda_r^{-1}\,dr\,\Lambda_t\\
&= \xi_t^0 M_t^{-1}\Lambda_t - \int_0^t f_x^\pi(r)\,\nabla X_r^\pi\,\Lambda_r^{-1}\,dr\,\Lambda_t,
\end{aligned}
\tag{3.14}
$$

where, for $i = 0, 1, \dots, m$,

$$\xi_t^i \triangleq E\{M_T \widetilde{\xi}^i | \mathcal{F}_t\} = E\{M_T \widetilde{\xi}^i\} + \int_0^t \eta_r^i \, dW_r. \tag{3.15}$$

Note that f_z^π is bounded, thus $M_T \in L^p(\mathcal{F}_T)$ and $\nabla X^\pi \in L^p(\mathbf{F})$ for all $p \geq 2$. Therefore for each $p \geq 1$, (2.4) leads to

$$E\left\{\sum_{j=1}^n |M_T \widetilde{\xi}^j|\right\}^p \leq C_p E\left\{|M_T|^p \sup_{0 \leq t \leq T} |\nabla X_t^\pi|^p\right\} \leq C_p. \tag{3.16}$$

In particular, for each j, $M_T \widetilde{\xi}^j \in L^2(\mathcal{F}_T)$, thus (3.15) makes sense.

Now we fix i_0. For $t \in [s_{i-1}, s_i)$, by applying Lemma 2.5, (3.10) and (3.14) imply that

$$Z_t^\pi = \left[\left(\xi_t^0 + \sum_{j \geq i} \xi_t^j\right) M_t^{-1} - \int_0^t f_x(r) \, \nabla X_r^\pi \, \Lambda_r^{-1} \, dr\right] \Lambda_t [\nabla X_t^\pi]^{-1} \sigma^\pi(t, X_t^\pi).$$

Therefore,

$$\left|Z_t^\pi - Z_{t_{i_0-1}}^\pi\right| \leq I_t^1 + I_t^2 + I_t^3, \tag{3.17}$$

where (recalling that $t_{i_0-1} = s_{l_{i_0-1}}$!)

$$\begin{aligned}
I_t^1 &\triangleq \left|\left[\xi_t^0 + \sum_{j \geq i} \xi_t^j\right] - \left[\xi_{t_{i_0-1}}^0 + \sum_{j \geq l_{i_0-1}+1} \xi_{t_{i_0-1}}^j\right]\right| \\
&\qquad \times \left|M_{t_{i_0-1}}^{-1} \Lambda_{t_{i_0-1}} \left[\nabla X_{t_{i_0-1}}^\pi\right]^{-1} \sigma^\pi\left(t_{i_0-1}, X_{t_{i_0-1}}^\pi\right)\right|, \\
I_t^2 &\triangleq \left|\xi_t^0 + \sum_{j \geq i} \xi_t^j\right| \Big| M_t^{-1} \Lambda_t [\nabla X_t^\pi]^{-1} \sigma^\pi(t, X_t^\pi) \\
&\qquad\qquad - M_{t_{i_0-1}}^{-1} \Lambda_{t_{i_0-1}} \left[\nabla X_{t_{i_0-1}}^\pi\right]^{-1} \sigma^\pi\left(t_{i_0-1}, X_{t_{i_0-1}}^\pi\right)\Big|, \\
I_t^3 &\triangleq \Big| \int_0^t f_x^\pi(r) \, \nabla X_r^\pi \, \Lambda_r^{-1} \, dr \, \Lambda_t [\nabla X_t^\pi]^{-1} \sigma^\pi(t, X_t^\pi) \\
&\qquad - \int_0^{t_{i_0-1}} f_x^\pi(r) \, \nabla X_r^\pi \, \Lambda_r^{-1} \, dr \, \Lambda_{t_{i_0-1}} \left[\nabla X_{t_{i_0-1}}^\pi\right]^{-1} \sigma^\pi\left(t_{i_0-1}, X_{t_{i_0-1}}^\pi\right)\Big|.
\end{aligned}$$

We assume $t \in (t_{i_0-1}, t_{i_0+1})$. Recalling (3.12) and applying Lemma 2.4 one can easily show that

$$E\{|I_t^3|^2\} \le C(1+|x|^2)|\pi_0|. \tag{3.18}$$

Recalling (3.15), (3.13) and (3.12) we have

$$\begin{aligned}\sum_{j=0}^{m}|\xi_t^j| &\le \sum_{j=0}^{m} E\{|M_T\widetilde{\xi}^j||\mathcal{F}_t\} \\ &\le E\biggl\{\biggl|M_T\int_0^T f_x^\pi(r)\,\nabla X_r^\pi\,\Lambda_r^{-1}\,dr\biggr| \\ &\qquad + |M_T\Lambda_T^{-1}|\sum_{j=1}^{m}|g_j^\pi\,\nabla X_{s_j}^\pi|\Big|\mathcal{F}_t\biggr\} \\ &\le CE\Bigl\{M_T\sup_{0\le s\le T}|\nabla X_s^\pi|\Big|\mathcal{F}_t\Bigr\},\end{aligned}\tag{3.19}$$

where the last inequality is due to (2.4) and the fact that Λ^{-1} and f_x^π are uniformly bounded. Then by applying (3.16), (3.12) and Lemma 2.4, one can show that

$$E\{|I_t^2|^2\} \le C(1+|x|^2)|\pi_0|. \tag{3.20}$$

It remains to estimate I_t^1. To this end we denote

$$\begin{aligned}\Gamma_t &\triangleq \sup_{0\le s\le t}\{1+|\nabla X_s^\pi|+|[\nabla X_s^\pi]^{-1}|+|M_s^{-1}|\},\\ \bar{\Gamma}_t &\triangleq \sup_{0\le s\le t}\{1+|X_s^\pi|\}.\end{aligned}$$

Noting that Λ is bounded and that $\Gamma_{t_{i_0-1}}, \bar{\Gamma}_{t_{i_0-1}} \in \mathcal{F}_{t_{i_0-1}}$, and recalling that $t_{i_0-1} = s_{l_{i_0-1}}$, by (3.15) we have

$$\begin{aligned}E\{|I_t^1|^2\} &\le CE\biggl\{\Gamma_{t_{i_0-1}}^4\bar{\Gamma}_{t_{i_0-1}}^2\biggl|\biggl[\xi_t^0+\sum_{j\ge i}\xi_t^j\biggr] \\ &\qquad - \biggl[\xi_{t_{i_0-1}}^0+\sum_{j\ge l_{i_0-1}+1}\xi_{t_{i_0-1}}^j\biggr]\biggr|^2\biggr\} \\ &\le CE\biggl\{\Gamma_{t_{i_0-1}}^4\bar{\Gamma}_{t_{i_0-1}}^2\biggl[\Bigl|\xi_t^0-\xi_{t_{i_0-1}}^0\Bigr|^2 \\ &\qquad + \biggl|\sum_{l_{i_0-1}<j<i}\xi_t^j\biggr|^2+\biggl|\sum_{j>l_{i_0-1}}(\xi_t^j-\xi_{t_{i_0-1}}^j)\biggr|^2\biggr]\biggr\}\end{aligned}\tag{3.21}$$

$$
\begin{aligned}
&= CE\Big\{\Gamma^4_{t_{i_0-1}}\bar\Gamma^2_{t_{i_0-1}}\Big|E\Big\{\sum_{l_{i_0-1}<j<i}\xi^j_T\Big|\mathcal{F}_t\Big\}\Big|^2 \\
&\qquad + \Gamma^4_{t_{i_0-1}}\bar\Gamma^2_{t_{i_0-1}}E\Big\{\Big|\xi^0_t-\xi^0_{t_{i_0-1}}\Big|^2 \\
&\qquad\qquad + \Big|\sum_{j>l_{i_0-1}}(\xi^j_t-\xi^j_{t_{i_0-1}})\Big|^2\Big|\mathcal{F}_{t_{i_0-1}}\Big\}\Big\} \\
&\le CE\Big\{\Gamma^4_{t_{i_0-1}}\bar\Gamma^2_{t_{i_0-1}}\Big[\Big|\sum_{l_{i_0-1}<j<i}\xi^j_T\Big|^2+\int_{t_{i_0-1}}^t|\eta^0_r|^2\,dr \\
&\qquad\qquad + \int_{t_{i_0-1}}^t\Big|\sum_{j>l_{i_0-1}}\eta^j_r\Big|^2\,dr\Big]\Big\} \\
&\le CE\Big\{\Gamma^4_{t_{i_0-1}}\bar\Gamma^2_{t_{i_0-1}}\Big[\Big(\sum_{j=l_{i_0-1}}^{l_{i_0+1}}|\xi^j_T|\Big)^2+\int_{t_{i_0-1}}^{t_{i_0+1}}|\eta^0_r|^2\,dr \\
&\qquad\qquad + \int_{t_{i_0-1}}^{t_{i_0+1}}\Big|\sum_{j>l_{i_0-1}}\eta^j_r\Big|^2\,dr\Big]\Big\}.
\end{aligned}
$$

Applying Lemma 3.3 and Corollary 3.4, (3.21) leads to that

$$
\begin{aligned}
&\sum_{i=1}^n E\Big\{\int_{t_{i-1}}^{t_i}|I^1_t|^2\,dt+\Delta t_i\big[|I^1_{t_{i-1}+r_k}|^2+|I^1_{t_i+r_k}|^2\big]\Big\} \\
&\quad\le \sum_{i=1}^n C\,\Delta t_i\;E\Big\{\Gamma^4_T\bar\Gamma^2_T\Big(\sum_{l_{i-1}+1\le j\le l_{i+1}}|\xi^j_T|\Big)^2 \\
&\qquad\qquad + \Gamma^4_{t_{i-1}}\bar\Gamma^2_{t_{i-1}}\Big[\int_{t_{i-1}}^{t_{i+1}}|\eta^0_r|^2\,dr \\
&\qquad\qquad\qquad + \int_{t_{i-1}}^{t_{i+1}}\Big|\sum_{k\ge i}\Big(\sum_{l_{k-1}+1\le j\le l_k}\eta^j_r\Big)\Big|^2\,dr\Big]\Big\} \\
&\quad\le C|\pi_0|E\Big\{\Gamma^4_T\bar\Gamma^2_T\Big[\sum_{i=1}^n\Big(\sum_{l_{i-1}+1\le j\le l_{i+1}}|\xi^j_T|\Big)^2 \\
&\qquad\qquad + \sup_{0\le t\le T}|\xi^0_t|^2+\Big(\sup_{0\le t\le T}\sum_{k=1}^n\Big|\sum_{l_{k-1}+1\le j\le l_k}\xi^j_t\Big|\Big)^2\Big]\Big\}
\end{aligned}
\tag{3.22}
$$

$$\leq C|\pi_0| E\left\{\Gamma_T^4 \bar{\Gamma}_T^2 \left(\sup_{0\leq t\leq T} \sum_{j=0}^m |\xi_t^j|\right)^2\right\}$$

$$\leq C|\pi_0| \|\Gamma_T^4\|_3 \|\bar{\Gamma}_T^2\|_3 \left\|\left(\sup_{0\leq t\leq T} \sum_{j=0}^m |\xi_t^j|\right)^2\right\|_3$$

$$\leq C(1+|x|^2)|\pi_0| \left\|\sup_{0\leq t\leq T} \sum_{j=0}^m |\xi_t^j|\right\|_6^2.$$

Now by (3.19) and applying Doob's inequality one has

$$\begin{aligned} E\left\{\left(\sup_{0\leq t\leq T} \sum_{j=0}^m |\xi_t^j|\right)^6\right\} &\leq CE\left\{\sup_{0\leq t\leq T} \left|E\left\{M_T \sup_{0\leq s\leq T} |\nabla X_s^\pi| \Big| \mathcal{F}_t\right\}\right|^6\right\} \\ &\leq CE\left\{M_T^6 \sup_{0\leq t\leq T} |\nabla X_t^\pi|^6\right\} \\ &\leq C, \end{aligned}$$

which, together with (3.22), implies that

$$\sum_{i=1}^n E\left\{\int_{t_{i-1}}^{t_i} |I_t^1|^2\,dt + \Delta t_i\big[|I_{t_{i-1}+r_k}^1|^2 + |I_{t_i+r_k}^1|^2\big]\right\} \leq C(1+|x|^2)|\pi_0|. \tag{3.23}$$

Combining (3.23), (3.20) and (3.18), we deduce from (3.17) that

$$\begin{aligned} &\sum_{i=1}^n E\left\{\int_{t_{i-1}}^{t_i} |Z_t^\pi - Z_{t_{i-1}}^\pi|^2\,dt + \Delta t_i\,\big[|Z_{t_{i-1}+r_k}^\pi - Z_{t_{i-1}}^\pi|^2 + |Z_{t_i+r_k}^\pi - Z_{t_{i-1}}^\pi|^2\big]\right\} \\ &\qquad \leq C(1+|x|^2)|\pi_0|, \end{aligned}$$

which, combined with (3.8), leads to

$$\begin{aligned} &\sum_{i=1}^n E\left\{\int_{t_{i-1}}^{t_i} \big[|Z_t - Z_{t_{i-1}}|^2 + |Z_t - Z_{t_i}|^2\big]\,dt\right\} \\ &\qquad \leq CE\left\{\int_0^T |Z_t^\pi - Z_t|^2\,dt + \sum_{i=1}^n \Delta t_i\,|Z_{t_{i-1}+r_k}^\pi - Z_{t_{i-1}+r_k}|^2\right\} \\ &\qquad\quad + C\sum_{i=1}^n \Delta t_i\, E\big\{|Z_{t_{i-1}+r_k} - Z_{t_{i-1}}|^2 + |Z_{t_i+r_k} - Z_{t_i}|^2\big\} \\ &\qquad\quad + C(1+|x|^2)|\pi_0|. \end{aligned} \tag{3.24}$$

Recalling (3.6) and (3.7) and letting $|\pi| \to 0$, (3.24) implies that

$$\sum_{i=1}^{n} E\left\{\int_{t_{i-1}}^{t_i} [|Z_t - Z_{t_{i-1}}|^2 + |Z_t - Z_{t_i}|^2]\, dt\right\}$$

$$(3.25) \qquad \leq C \sum_{i=1}^{n} \Delta t_i \, E\{|Z_{t_{i-1}+r_k} - Z_{t_{i-1}}|^2 + |Z_{t_i+r_k} - Z_{t_i}|^2\} + C(1+|x|^2)|\pi_0|.$$

Finally, since Z is càdlàg, applying Lemma 3.2 and Fatou's lemma we get that $E\{|Z_t|^p\} \leq C_p(1+|x|)$ for all $t \in [0,T]$, which implies that $\{Z_t\}_{0\leq t\leq T}$ is uniformly integrable. Now let $r_k \to 0$, again by the assumption that Z is càdlàg, we have

$$\lim_{k\to\infty} E\{|Z_{t_{i-1}+r_k} - Z_{t_{i-1}}|^2 + |Z_{t_i+r_k} - Z_{t_i}|^2\} = 0,$$

which, combined with (3.25), proves the theorem. □

4. Discretization of the FSDE. In this section, we briefly review the Euler scheme for the forward diffusion X. Let $\pi : 0 = t_0 < t_1 < \cdots < t_n = T$ be a partition of $[0,T]$. Define $\pi(t) \triangleq t_{i-1}$, for $t \in [t_{i-1}, t_i)$. Let X^π be the solution of the following SDE:

$$(4.1) \qquad X_t^\pi = x + \int_0^t b(\pi(r), X_{\pi(r)}^\pi)\, dr + \int_0^t \sigma(\pi(r), X_{\pi(r)}^\pi)\, dW_r,$$

and we define a "step process" $\hat{X}^\pi$ as follows.

$$(4.2) \qquad \hat{X}_t^\pi \triangleq X_{\pi(t)}^\pi, \qquad t \in [0,T].$$

The following estimates are well known (see, e.g., [13]).

LEMMA 4.1. *Assume that b and σ satisfy the conditions in Assumption* 2.3. *Then for X and X^π defined as in* (1.1) *and* (4.1), *respectively*, *there exists a constant C*, *depending only on T and K*, *such that*

$$E\left\{\sup_{0\leq t\leq T} |X_t^\pi|^4\right\} \leq C(1+|x|^4),$$

$$E\left\{\sup_{0\leq t\leq T} |X_t - X_t^\pi|^2\right\} \leq C(1+|x|^2)|\pi|.$$

We now turn our attention to the estimates involving $\hat{X}^\pi$. The following theorem is more or less standard. For completeness we shall sketch a proof.

THEOREM 4.2. *Assume that b and σ satisfy the conditions in Assumption* 2.3. *Then there exists a constant $C > 0$, depending only on T and K, such that the following estimates hold*:

$$\sup_{0\le t\le T} E\{|X_t - \hat{X}_t^\pi|^2\} \le C(1+|x|^2)|\pi|;$$

$$E\left\{\sup_{0\le t\le T} |X_t - \hat{X}_t^\pi|^2\right\} \le C(1+|x|^2)|\pi| \log\frac{1}{|\pi|}.$$

PROOF. First, for $t \in [t_{i-1}, t_i)$, applying Lemmas 3.2 and 4.1 we have

$$\begin{aligned} E\{|X_t - \hat{X}_t^\pi|^2\} &= E\{|X_t - X_{t_{i-1}}^\pi|^2\} \\ &\le 2E\{|X_t - X_{t_{i-1}}|^2 + |X_{t_{i-1}} - X_{t_{i-1}}^\pi|^2\} \\ &\le C(1+|x|^2)|\pi|. \end{aligned}$$

Next, recalling (4.1) one can easily get

$$\begin{aligned} \sup_{0\le t\le T} |X_t - \hat{X}_t^\pi| &\le \sup_{0\le t\le T} |X_t - X_t^\pi| + \max_{1\le i\le n} \sup_{t_{i-1}\le t<t_i} |X_t^\pi - X_{t_{i-1}}^\pi| \\ &\le \sup_{0\le t\le T} |X_t - X_t^\pi| + \max_{1\le i\le n} [|b(t_{i-1}, X_{t_{i-1}}^\pi)|\,\Delta t_i] \\ &\quad + \max_{1\le i\le n}\Big[|\sigma(t_{i-1}, X_{t_{i-1}}^\pi)| \sup_{t_{i-1}\le t<t_i} |W_t - W_{t_{i-1}}|\Big]. \end{aligned} \tag{4.3}$$

Now using Lemmas 2.8 and 4.1, we infer from (4.3) that

$$\begin{aligned} &E\left\{\sup_{0\le t\le T} |X_t - \hat{X}_t^\pi|^2\right\} \\ &\quad\le CE\left\{\sup_{0\le t\le T} |X_t - X_t^\pi|^2 + |\pi|^2 \max_{1\le i\le n} |b(t_{i-1}, X_{t_{i-1}}^\pi)|^2\right. \\ &\qquad\left. + \max_{1\le i\le n} |\sigma(t_{i-1}, X_{t_{i-1}}^\pi)|^2 \max_{1\le i\le n} \sup_{t_{i-1}\le t<t_i} |W_t - W_{t_{i-1}}|^2\right\} \\ &\quad\le C(1+|x|^2)\left[|\pi| + \sqrt{E\left\{\max_{1\le i\le n} \sup_{t_{i-1}\le t<t_i} |W_t - W_{t_{i-1}}|^4\right\}}\right] \\ &\quad\le C(1+|x|^2)\left[|\pi| + \sqrt{E\left\{\max_{1\le i\le n} (\Delta t_i)^2 N_i^4\right\}}\right], \end{aligned} \tag{4.4}$$

where N_i, $i = 1, 2, \dots, n$, are i.i.d. random variables with standard Normal distribution. Denote $C_\varepsilon \overset{\triangle}{=} 2\varepsilon \log(1/\varepsilon)$, then we have $C_{|\pi|}/\Delta t_i \ge 2\log(1/|\pi|)$ and

$$
\begin{aligned}
&E\Big\{\max_{1\le i\le n}(\Delta t_i)^2 N_i^4\Big\}\\
&\quad = E\Big\{\max_{1\le i\le n}(\Delta t_i)^2 N_i^4 \mathbb{1}_{\{\max_i \Delta t_i N_i^2 \le C_{|\pi|}\}}\Big\}\\
&\qquad + E\Big\{\max_{1\le i\le n}(\Delta t_i)^2 N_i^4 \mathbb{1}_{\{\max_i \Delta t_i N_i^2 \ge C_{|\pi|}\}}\Big\} \qquad (4.5)\\
&\quad \le C_{|\pi|}^2 + \sum_{i=1}^n (\Delta t_i)^2 E\big\{N_i^4 \mathbb{1}_{\{N_i^2 \ge C_{|\pi|}/\Delta t_i\}}\big\}\\
&\quad \le C_{|\pi|}^2 + T|\pi| E\big\{N_1^4 \mathbb{1}_{\{N_1^2 \ge 2\log(1/|\pi|)\}}\big\}.
\end{aligned}
$$

Moreover, by direct calculation we have, for $\forall a \ge 1$,

$$
E\big\{N_1^4 \mathbb{1}_{\{N_1^2 > a\}}\big\} \le C\sqrt{a^3}\exp\Big(-\frac{a}{2}\Big). \tag{4.6}
$$

Setting $a = 2\log(1/|\pi|)$ and assuming that the partition π is fine enough so that $a \ge 1$, we obtain from (4.6) and (4.5) that

$$
\begin{aligned}
E\Big\{\max_{0\le i\le n-1}(\Delta t_i)^2 N_i^4\Big\} &\le C_{|\pi|}^2 + C|\pi|\Big(\log\frac{1}{|\pi|}\Big)^{3/2}\exp\Big(-\log\frac{1}{|\pi|}\Big)\\
&\le C|\pi|^2\Big(\log\frac{1}{|\pi|}\Big)^2.
\end{aligned}
$$

This, together with (4.4), proves the theorem. $\square$

REMARK 4.3. We note that $|\pi|\log(1/|\pi|)$ is indeed sharp. In fact, let $X_t = W_t$ and π be a partition with equal size, one can show that $[|\pi| \times \log(1/|\pi|)]^{-1}\sup_{0\le t\le T}|X_t - \hat{X}_t^\pi|^2 \to \sqrt{2}$ in distribution (see, e.g., [1], Proposition 1).

The following result is a direct consequence of Theorem 4.2.

COROLLARY 4.4. *Assume all the conditions in Theorem* 4.2 *hold. If* $\Phi : \mathbb{D}^d \mapsto \mathbb{R}$ *satisfies the* L^∞*-Lipschitz condition* (2.2), *then*

$$
E\big\{|\Phi(X) - \Phi(\hat{X}^\pi)|^2\big\} \le C(1+|x|^2)|\pi|\log\frac{1}{|\pi|}.
$$

Moreover, if Φ *satisfies the* L^1*-Lipschitz condition* (2.3) *or it takes the form* $\Phi(X) = g(X_T)$, *then*

$$
E\big\{|\Phi(X) - \Phi(\hat{X}^\pi)|^2\big\} \le C(1+|x|^2)|\pi|.
$$

5. Discretization of the BSDE. In this section, we construct an approximating solution (Y, Z) by using the "step processes." Let $\pi : 0 = t_0 < \cdots < t_n = T$ be any partition on $[0, T]$ and define the approximating pairs (Y^π, Z^π) recursively (in a backward manner), such that

$$Y^\pi_{t_n} = \xi^\pi, \qquad Z^{\pi,1}_{t_n} = 0,$$

$$Y^\pi_t = Y^\pi_{t_i} + f\big(t_i, \Theta^{\pi,1}_{t_i}\big)\,\Delta t_i - \int_t^{t_i} Z^\pi_r\,dW_r, \tag{5.1}$$

$$t \in [t_{i-1}, t_i),\ i = n, n-1, \dots, 1,$$

where $\xi^\pi \in L^2(\mathcal{F}_T)$, $\Theta^{\pi,1}_{t_i} \triangleq (X^\pi_{t_i}, Y^\pi_{t_i}, Z^{\pi,1}_{t_i})$ and

$$Z^{\pi,1}_{t_i} \triangleq \frac{1}{\Delta t_{i+1}} E\bigg\{\int_{t_i}^{t_{i+1}} Z^\pi_r\,dr \bigg| \mathcal{F}_{t_i}\bigg\}, \qquad i = 0, \dots, n-1. \tag{5.2}$$

We should point out here that the family $\{(Y^\pi, Z^\pi)\}$ is *different* from that in (3.5). Also, we note that in (5.1) the terminal value of the BSDE is $Y^\pi_{t_i} + f(t_i, \Theta^{\pi,1}_{t_i})\,\Delta t_i$.

REMARK 5.1. As we shall see in (6.4) below, if one chooses ξ^π appropriately, then $Z^{\pi,1}_{t_i} = Z^\pi_{t_i}$. Therefore the computation of the integral and the conditional expectation in (5.2) becomes redundant and can be omitted. This will be significant in implementation.

To prove the convergence of (5.1), let us introduce a new notion regarding the partition π.

DEFINITION 5.2. Let $K > 0$ be a constant. A partition π is called K-uniform if $\Delta t_i \ge |\pi|/K$ for $i = 1, \dots, n$.

THEOREM 5.3. *Assume that all the conditions in Theorem* 3.1 *hold true and suppose that the partition π is K-uniform. Then the following estimate holds*:

$$\max_{0\le i\le n} E\big\{|Y_{t_i} - Y^\pi_{t_i}|^2\big\} + E\bigg\{\int_0^T |Z_r - Z^\pi_r|^2\,dr\bigg\} \le C\big[(1+|x|^2)|\pi| + E\{|\xi - \xi^\pi|^2\}\big],$$

where $\xi = \Phi(X)$ and C depends only on T and K.

Before we prove the theorem, let us first give a "discrete version" of the Gronwall inequality for convenience.

LEMMA 5.4. *Assume $\pi : 0 = t_0 < \cdots < t_n = T$. If $a_i, b_i, i = 0, 1, \dots, n$, satisfy that $a_n \ge 0$, $b_i \ge 0$ and that $a_{i-1} \le (1 + C\,\Delta t_i)a_i + b_i$ for $i = 1, \dots, n$.*

Then we have $\max_{0\le i\le n} a_i \le e^{CT}[a_n + \sum_{i=1}^n b_i]$.

PROOF. Note that $1 + C\,\Delta t_i \le e^{C\,\Delta t_i}$. By induction one can easily show that

$$a_i \le e^{C(T-t_i)}\Biggl[a_T + \sum_{j=i+1}^n b_j\Biggr],$$

which clearly proves the lemma. □

PROOF OF THEOREM 5.3. We first denote, for $i = 1, \dots, n$,

$$I_{i-1} \triangleq E\biggl\{|Y_{t_{i-1}} - Y^\pi_{t_{i-1}}|^2 + \int_{t_{i-1}}^{t_i} |Z_r - Z^\pi_r|^2\,dr\biggr\}.$$

By (2.1) and (5.1) we have

$$\begin{aligned} &Y_{t_{i-1}} - Y^\pi_{t_{i-1}} + \int_{t_{i-1}}^{t_i} (Z_r - Z^\pi_r)\,dW_r \\ &\qquad = Y_{t_i} - Y^\pi_{t_i} + \int_{t_{i-1}}^{t_i} f(r,\Theta_r)\,dr - f(t_i,\Theta^{\pi,1}_{t_i})\,\Delta t_i. \end{aligned}$$

Squaring both sides and then taking expectations, also noting that $Y_{t_{i-1}} - Y^\pi_{t_{i-1}}$ is uncorrelated with $\int_{t_{i-1}}^{t_i}(Z_r - Z^\pi_r)\,dW_r$, we get

$$\begin{aligned} I_{i-1} &= E\biggl\{\biggl[(Y_{t_i} - Y^\pi_{t_i}) + \int_{t_{i-1}}^{t_i} \bigl(f(r,\Theta_r) - f\bigl(t_i,\Theta^{\pi,1}_{t_i}\bigr)\bigr)\,dr\biggr]^2\biggr\} \\ &\le E\biggl\{|Y_{t_i} - Y^\pi_{t_i}| \\ &\qquad + C\int_{t_{i-1}}^{t_i} [\sqrt{t_i - r} + |X_r - X^\pi_{t_i}| + |Y_r - Y^\pi_{t_i}| + |Z_r - Z^{\pi,1}_{t_i}|]^2\,dr\biggr\}. \end{aligned}$$

Note that for any $\varepsilon > 0$, one has

$$(a+b)^2 = a^2 + b^2 + 2ab \le \biggl(1 + \frac{\Delta t_i}{\varepsilon}\biggr)a^2 + \biggl(1 + \frac{\varepsilon}{\Delta t_i}\biggr)b^2.$$

Thus we have

$$\begin{aligned} I_{i-1} &\le E\biggl\{\biggl(1 + \frac{C\Delta t_i}{\varepsilon}\biggr)|Y_{t_i} - Y^\pi_{t_i}|^2 \\ &\qquad + C\biggl(1 + \frac{\varepsilon}{\Delta t_i}\biggr) \\ &\qquad \times \biggl(\int_{t_{i-1}}^{t_i} [\sqrt{|\pi|} + |X_r - X^\pi_{t_i}| + |Y_r - Y_{t_i}| + |Z_r - Z^{\pi,1}_{t_i}|]\,dr\biggr)^2\biggr\} \end{aligned} \tag{5.3}$$

$$\le E\bigg\{\bigg(1+\frac{C\Delta t_i}{\varepsilon}\bigg)|Y_{t_i}-Y^{\pi}_{t_i}|^2 + C(\varepsilon+\Delta t_i)\int_{t_{i-1}}^{t_i}\big[|\pi|+|X_r-X_{t_i}|^2+|X_{t_i}-X^{\pi}_{t_i}|^2 + |Y_r-Y_{t_i}|^2+|Z_r-Z_{t_i}|^2+|Z_{t_i}-Z^{\pi,1}_{t_i}|^2\big]\,dr\bigg\}$$

$$\le E\bigg\{\bigg(1+\frac{C\Delta t_i}{\varepsilon}\bigg)|Y_{t_i}-Y^{\pi}_{t_i}|^2 + C\int_{t_{i-1}}^{t_i}|Z_r-Z_{t_i}|^2\,dr + C(\varepsilon+\Delta t_i)\,\Delta t_i|Z_{t_i}-Z^{\pi,1}_{t_i}|^2 + C|\pi|^2\bigg\},$$

thanks to Lemmas 3.2 and 4.1.

Since π is K-uniform, we have $\Delta t_i \le K\Delta t_{i+1}$. Therefore,

$$\begin{aligned} E\{\Delta t_i\,|Z_{t_i}-Z^{\pi,1}_{t_i}|^2\} &\le C E\{\Delta t_{i+1}\,|Z_{t_i}-Z^{\pi,1}_{t_i}|^2\} \\ &= \frac{C}{\Delta t_{i+1}} E\bigg\{\bigg[E\bigg\{\int_{t_i}^{t_{i+1}}(Z^{\pi}_r-Z_{t_i})\,dr\,\bigg|\mathcal{F}_{t_i}\bigg\}\bigg]^2\bigg\} \\ &\le C E\bigg\{\int_{t_i}^{t_{i+1}}|Z^{\pi}_r-Z_{t_i}|^2\,dr\bigg\} \\ &\le C\bigg\{\int_{t_i}^{t_{i+1}}\big[|Z^{\pi}_r-Z_r|^2+|Z_r-Z_{t_i}|^2\big]\,dr\bigg\}. \end{aligned}$$

Thus (5.3) leads to

$$I_{i-1}\le E\bigg\{\bigg(1+\frac{C\Delta t_i}{\varepsilon}\bigg)|Y_{t_i}-Y^{\pi}_{t_i}|^2 + C_1(\varepsilon+\Delta t_i)\int_{t_i}^{t_{i+1}}|Z^{\pi}_r-Z_r|^2\,dr + C\int_{t_{i-1}}^{t_{i+1}}|Z_r-Z_{t_i}|^2\,dr + C|\pi|^2\bigg\}.$$

Now choosing $\varepsilon = 1/(4C_1)$, then for $|\pi|\le 1/(4C_1)$, we have

$$I_{i-1}\le E\bigg\{(1+C\Delta t_i)|Y^{\pi}_{t_i}-Y_{t_i}|^2+\tfrac{1}{2}\int_{t_i}^{t_{i+1}}|Z^{\pi}_r-Z_r|^2\,dr + C\int_{t_{i-1}}^{t_{i+1}}|Z_r-Z_{t_i}|^2\,dr + C|\Delta t_i|^2\bigg\}.$$

Therefore,

$$I_{i-1}+\tfrac{1}{2}E\bigg\{\int_{t_i}^{t_{i+1}}|Z^{\pi}_r-Z_r|^2\,dr\bigg\} \le (1+C\,\Delta t_i)I_i + C E\bigg\{\int_{t_{i-1}}^{t_{i+1}}|Z_r-Z_{t_i}|^2\,dr+|\pi|^2\bigg\}. \tag{5.4}$$

Applying Lemma 5.4 and recalling Theorem 3.1, we have

$$\max_{0\le i\le n} I_i \le CE\left\{|\xi-\xi^\pi|^2+\sum_{i=1}^n\int_{t_{i-1}}^{t_i}\left[\left|Z_r-Z_{t_{i-1}}\right|^2+\left|Z_r-Z_{t_i}\right|^2\right]dr + (1+|x|^2)|\pi|\right\} \tag{5.5}$$
$$\le C\left[(1+|x|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\right].$$

Moreover, summing both sides of (5.4) over i from 0 to $n-1$, we have

$$\sum_{i=0}^{n-2} I_i+\tfrac{1}{2}E\left\{\int_0^T|Z_r^\pi-Z_r|^2\,dr\right\}$$
$$\le\sum_{i=0}^{n-1}(1+C\,\Delta t_i)I_i+C\left[(1+|x|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\right].$$

Therefore,

$$E\left\{\int_0^T|Z_r^\pi-Z_r|^2\,dr\right\} \le 2(1+C\,\Delta t_{n-1})I_{n-1} + C\left[\sum_{i=0}^{n-1}\Delta t_i\,I_i+(1+|x|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\right]. \tag{5.6}$$

Plugging (5.5) into (5.6), one can easily prove that

$$E\left\{\int_0^T|Z_r-Z_r^\pi|^2\,dr\right\}\le CE\{|\xi-\xi^\pi|^2+C(1+|x|^2)|\pi|\},$$

which, combined with (5.5), proves the theorem. □

REMARK 5.5. If f is independent of z, then the "K-uniform" assumption of π is *not* necessary. In fact, in such a case (5.3) does not involve the process Z. Thus, one may apply Lemma 5.4 directly on (5.3) and obtain a simplified proof.

Now let us define the following two step processes:

$$\widehat{Y}_t^\pi\triangleq Y_{t_{i-1}}^\pi,\qquad \widehat{Z}_t^\pi\triangleq Z_{t_{i-1}}^{\pi,1}\qquad\text{for } t\in[t_{i-1},t_i).$$

Then we have the following theorem.

THEOREM 5.6. *Assume that all the conditions in Theorem* 3.1 *hold true and let* $K>0$ *be given. Then for any* K*-uniform partition* π, *the following estimate* *holds*:

$$\sup_{0\le t\le T} E\{|Y_t-\widehat{Y}_t^\pi|^2\}+E\left\{\int_0^T|Z_r-\widehat{Z}_r^\pi|^2\,dr\right\}$$
$$\le C\big[(1+|x|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\big].$$

PROOF. First, combining Lemma 3.2, Theorem 5.3 and Lemma 2.4 we have, for $\forall\, t\in[t_{i-1},t_i)$,

$$\begin{aligned}E\{|Y_t-\widehat{Y}_t^\pi|^2\}&\le 2E\{|Y_t-Y_{t_{i-1}}|^2+|Y_{t_{i-1}}-Y_{t_{i-1}}^\pi|^2\}\\&\le C\big[(1+|x|^2+E|\xi|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\big]\\&\le C\big[(1+|x|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\big].\end{aligned}\tag{5.7}$$

To estimate $Z-\hat{Z}^\pi$, we recall that a conditional expectation minimizes the conditional mean square error. By (5.2) one can easily show that

$$E\left\{\int_{t_{i-1}}^{t_i}|Z_r^\pi-Z_{t_{i-1}}^{\pi,1}|^2\,dr\right\}\le E\left\{\int_{t_{i-1}}^{t_i}|Z_r^\pi-Z_{t_{i-1}}|^2\,dr\right\}.$$

Therefore, by Theorems 3.1 and 5.3, we have

$$\begin{aligned}E\left\{\int_0^T|Z_t-\widehat{Z}_t^\pi|^2\,dt\right\}&\le 2E\left\{\int_0^T|Z_t-Z_t^\pi|^2\,dt+\sum_{i=1}^n\int_{t_{i-1}}^{t_i}|Z_t^\pi-Z_{t_{i-1}}^{\pi,1}|^2\,dt\right\}\\&\le 2E\left\{\int_0^T|Z_t-Z_t^\pi|^2\,dt+\sum_{i=1}^n\int_{t_{i-1}}^{t_i}|Z_t^\pi-Z_{t_{i-1}}|^2\,dt\right\}\\&\le CE\left\{\int_0^T|Z_t-Z_t^\pi|^2\,dt+\sum_{i=1}^n\int_{t_{i-1}}^{t_i}|Z_t-Z_{t_{i-1}}|^2\,dt\right\}\\&\le C\big[(1+|x|^2)|\pi|+E\{|\xi-\xi^\pi|^2\}\big].\end{aligned}$$

This, combined with (5.7), proves the theorem. □

6. Numerical schemes. In this section we propose a numerical scheme based on the results of previous sections. To present our scheme explicitly, we first recall the process $\hat{X}^\pi$ defined by (4.2) and define the "discrete" functional $\xi^\pi \stackrel{\triangle}{=} \Phi(\hat{X}^\pi)$ in (5.1). Next, note that we can rewrite (4.1) as

$$X_{t_i}^\pi = X_{t_i}(t_{i-1},X_{t_{i-1}}^\pi),\tag{6.1}$$

where, for $0\le s<t\le T$ and $x\in\mathbb{R}$,

$$X_t(s,x)\stackrel{\triangle}{=}x+b(s,x)(t-s)+\sigma(s,x)(W_t-W_s).\tag{6.2}$$

Further, for each $0\le i\le n$, we denote $\mathbf{x}^{(i)}=(x_0,\dots,x_i)$. Our scheme is as follows.

THEOREM 6.1. *Let $\xi = \Phi(X)$ and $\xi^\pi = \Phi(\hat{X}^\pi)$, (Y, Z) and (Y^π, Z^π) are solutions to BSDEs* (2.1) *and* (5.1), *respectively. Assume that all the conditions in Theorem* 3.1 *hold true. Define $u_i^\pi, v_i^\pi : \mathbb{R}^{d(i+1)} \to \mathbb{R}$, $i = 0, \dots, n$, as follows*:

$$u_n^\pi(\mathbf{x}^{(n)}) \triangleq \Phi\left(\sum_{i=1}^n x_{i-1}\mathbb{1}_{[t_{i-1},t_i)}(\cdot) + x_n\mathbb{1}_{\{T\}}(\cdot)\right), \qquad v_n^\pi(\mathbf{x}^{(n)}) \triangleq 0,$$

$$\begin{aligned} U_i^\pi(\mathbf{x}^{(i-1)}, \omega) &\triangleq u_i^\pi(\mathbf{x}^{(i-1)}, X_{t_i}(t_{i-1}, x_{i-1})) \\ &\quad + f\Big(t_i, X_{t_i}(t_{i-1}, x_{i-1}), u_i^\pi(\mathbf{x}^{(i-1)}, X_{t_i}(t_{i-1}, x_{i-1})), \\ &\qquad\qquad v_i^\pi(\mathbf{x}^{(i-1)}, X_{t_i}(t_{i-1}, x_{i-1}))\Big)\Delta t_i, \end{aligned} \tag{6.3}$$

$$u_{i-1}^\pi(\mathbf{x}^{(i-1)}) \triangleq E\{U_i^\pi(\mathbf{x}^{(i-1)}, \omega)\},$$

$$v_{i-1}^\pi(\mathbf{x}^{(i-1)}) \triangleq \frac{1}{\Delta t_i} E\{U_i^\pi(\mathbf{x}^{(i-1)}, \omega)[W_{t_i} - W_{t_{i-1}}]\}.$$

Then we have

$$\widehat{Y}_{t_i}^\pi = Y_{t_i}^\pi = u_i^\pi(X_{t_0}^\pi, \dots, X_{t_i}^\pi), \qquad \widehat{Z}_{t_i}^\pi = Z_{t_i}^{\pi,1} = Z_{t_i}^\pi = v_i^\pi(X_{t_0}^\pi, \dots, X_{t_i}^\pi). \tag{6.4}$$

If f is independent of z, or if π is K-uniform, then it holds that

$$\sup_{0 \le t \le T} E\{|Y_t - \widehat{Y}_t^\pi|^2\} + E\left\{\int_0^T |Z_t - \widehat{Z}_t^\pi|^2\,dt\right\} \le C(1 + |x|^2)|\pi| \log \frac{1}{|\pi|}. \tag{6.5}$$

Moreover, if Φ satisfies the L^1-Lipschitz condition (2.3) *or it takes the form $\Phi(X) = g(X_T)$, then we have*

$$\sup_{0 \le t \le T} E\{|Y_t - \widehat{Y}_t^\pi|^2\} + E\left\{\int_0^T |Z_t - \widehat{Z}_t^\pi|^2\,dt\right\} \le C(1 + |x|^2)|\pi|. \tag{6.6}$$

PROOF. By Theorem 5.3 and Corollary 4.4, it suffices to prove (6.4). We would also show simultaneously that, for each i, u_i^π and v_i^π are Lipschitz. (In fact it can be shown that u_i^π is uniformly Lipschitz, uniformly in π and i.) We proceed both by induction. For $i = n$ obviously (6.4) holds true and u_n^π and v_n^π are Lipschitz.

Assume that for i both (6.4) and the Lipschitz continuity hold true. First, by (5.1) obviously we have

$$\begin{aligned} Y_{t_{i-1}}^\pi = E\{&u_i^\pi(X_{t_0}^\pi, \dots, X_{t_i}^\pi) \\ &+ f(t_i, X_{t_i}^\pi, u_i^\pi(X_{t_0}^\pi, \dots, X_{t_i}^\pi), v_i^\pi(X_{t_0}^\pi, \dots, X_{t_i}^\pi))\Delta t_i | \mathcal{F}_{t_{i-1}}\}, \end{aligned}$$

which, combined with (6.1), implies that $Y_{t_{i-1}}^\pi = u_{i-1}^\pi(X_{t_0}^\pi, \dots, X_{t_{i-1}}^\pi)$.

Next, by the Lipschitz continuity assumption, one may apply Lemma 2.7. So Z_r^π is a martingale on $[t_{i-1}, t_i)$, therefore $Z_{t_{i-1}}^{\pi,1} = Z_{t_{i-1}}^\pi$. Moreover, following the formula of η_0 in Lemma 2.7, we get $Z^\pi_{t_{i-1}} = v^\pi_{i-1}(X^\pi_{t_0}, \dots, X^\pi_{t_{i-1}})$. This proves (6.4) for $i-1$.

Finally, by (6.2) and (6.3) one can easily check that u^π_{i-1} and v^π_{i-1} are also Lipschitz. This finishes the induction and thus proves the theorem. $\square$

REMARK 6.2. By (6.2) and (6.3), to calculate $(u^\pi_{i-1}, v^\pi_{i-1})$ from (u^π_i, v^π_i) one needs only approximate an expectation involving $W_{t_i} - W_{t_{i-1}}$, that is, a one-dimensional integral. We note that in Bally's method [2] one has to approximate integrals whose dimension is proportional to the partition size n.

There still remains, however, another type of high dimensionality: in (6.1) one has to compute the values of $\Phi(\hat{X}^\pi)$ for all possible choices of the high-dimensional vector $(X^\pi_{t_0}, \dots, X^\pi_{t_n})$. That is, if we discretize the real line $\mathbb{R}$ into M parts, then the number of values involved in the scheme is of the order $M^{d(n+1)}$, which would cause significant technical difficulties in implementation. We therefore impose the following restriction on Φ, motivated by applications in finance theory, so as to make our scheme implementable.

DEFINITION 6.3. Let $\mathbb{E}_1$ and $\mathbb{E}_2$ be two generic Euclidean spaces. A functional $\Phi : \mathbb{D}_{E_1}[0,T] \mapsto \mathbb{E}_2$, is called "constructible with construction φ" if for each $0 \le s < t \le T$ there exists a function $\Phi_t : \mathbb{D}_{E_1}[0,t] \mapsto \mathbb{E}_2$ and $\varphi_{s,t} : \mathbb{E}_2 \times \mathbb{D}_{E_1}[s,t] \mapsto \mathbb{E}_2$, such that for $0 \le s < t \le T$:

(i) $\Phi_T = \Phi$ and $\Phi_t(\mathbf{x}) = \Phi_t(\mathbf{x}\mathbb{1}_{[0,t)})$;
(ii) $\Phi_t(\mathbf{x}) = \varphi_{s,t}(\Phi_s(\mathbf{x}), \mathbf{x}\mathbb{1}_{[s,t)})$;
(iii) all $\varphi_{s,t}$ satisfy the L^∞-Lipschitz condition (2.2) with a uniformly Lipschitz constant.

We should point out here that since X is Markovian, by (ii) one can easily see that $\{(\Phi_\cdot(X), X_\cdot)\}$ is Markovian. Thus this definition essentially amounts to saying that we may add a state variable $\{\Phi_\cdot(X)\}$ such that $\{(\Phi_t(X), X_t, Y_t, Z_t)\}_{0 \le t \le T}$ is Markovian. Moreover, it is fairly easy to see that Φ also satisfies the L^∞-Lipschitz condition (2.2). [In fact, under assumptions (i) and (ii), Φ satisfies (2.2) if and only if (iii) holds true.]

The following are some easy examples of constructible functionals.

EXAMPLE 6.4. (i) The functional $\Phi : \mathbf{x} \mapsto \int_0^T \mathbf{x}(t)\,dt$, $\mathbf{x} \in \mathbb{D}^d[0,T]$, is constructible. Indeed, if we define

$$\Phi_t(\mathbf{x}) \triangleq \int_0^t \mathbf{x}(r)\,dr, \qquad \varphi_{s,t}(a, \mathbf{x}) \triangleq a + \int_s^t \mathbf{x}(r)\,dr,$$

then Φ is constructible with construction φ.

(ii) The functional $\Phi:\mathbf{x}\mapsto \sup_{0\le t\le T}\mathbf{x}(t)$, $\mathbf{x}\in\mathbb{D}^1[0,1]$, is also constructible, with

$$\Phi_t(\mathbf{x})=\sup_{0\le r\le t}\mathbf{x}(r),\qquad \varphi_{s,t}(a,\mathbf{x})\stackrel{\triangle}{=}a\vee\sup_{s\le r\le t}\mathbf{x}(r).$$

In the sequel, by a slight abuse of notations we denote

$$\varphi_{s,t}(a,x)\stackrel{\triangle}{=}\varphi_{s,t}\big(a,x\mathbb{1}_{[s,t)}\big)\qquad\forall\,(a,x)\in\mathbb{R}^k\times\mathbb{R}^d.$$

Now we can simplify the numerical scheme (6.1) as follows.

THEOREM 6.5. *Assume that in BSDEs* (2.1) *and* (5.1) $\xi=g(\Phi(X),X_T)$ *and* $\xi^\pi=g(\Phi(\hat X^\pi),\hat X^\pi_T)$, *respectively, and denote* (Y,Z) *and* (Y^π,Z^π) *to be the corresponding solutions. Assume also that all the conditions in Theorem* 3.1 *hold true and assume further that* g *is Lipschitz continuous, and* $\Phi:\mathbb{D}^d[0,T]\mapsto\mathbb{R}^k$ *is constructible with construction* φ. *Define, for* $\forall\,(a,x)\in\mathbb{R}^{d+k}$,

$$\begin{aligned}
&u^\pi_n(a,x)\stackrel{\triangle}{=}g(a,x),\qquad v^\pi_n(a,x)\stackrel{\triangle}{=}0,\\
&U^\pi_i(a,x,\omega)\stackrel{\triangle}{=}u^\pi_i\big(\varphi_{t_{i-1},t_i}(a,x),X_{t_i}(t_{i-1},x)\big)\\
&\qquad\qquad+f\Big(t_i,X_{t_i}(t_{i-1},x),u^\pi_i\big(\varphi_{t_{i-1},t_i}(a,x),X_{t_i}(t_{i-1},x)\big),\\
&\qquad\qquad\qquad v^\pi_i\big(\varphi_{t_{i-1},t_i}(a,x),X_{t_i}(t_{i-1},x)\big)\Big)\Delta t_i,\\
&u^\pi_{i-1}(a,x)\stackrel{\triangle}{=}E\{U^\pi_i(a,x,\omega)\},\\
&v^\pi_{i-1}(a,x)\stackrel{\triangle}{=}\frac{1}{\Delta t_i}E\{U^\pi_i(a,x,\omega)[W_{t_i}-W_{t_{i-1}}]\}.
\end{aligned}\tag{6.7}$$

Then all the results in Theorem 6.1 *hold true.*

We note that if we discretize the real line $\mathbb{R}$ into M parts, then for each i, the number of grid points for a is M^k and that for x is M^d. Therefore, the total number of values involved in scheme (6.7) is of the order $M^{d+k}(n+1)$, rather than $M^{d(n+1)}$ as in scheme (6.4).

REMARK 6.6. In Theorem 6.5, if we assume that $\xi=g(X_T)$, then we may consider Φ as a constant functional. In this case, $\varphi_{s,t}$ are also constant functionals, thus (6.7) becomes

$$\begin{aligned}
&u^\pi_n(x)\stackrel{\triangle}{=}g(x),\qquad v^\pi_n(x)\stackrel{\triangle}{=}0,\\
&U^\pi_i(x,\omega)\stackrel{\triangle}{=}u^\pi_i\big(X_{t_i}(t_{i-1},x)\big)\\
&\qquad\qquad+f\Big(t_i,X_{t_i}(t_{i-1},x),u^\pi_i\big(X_{t_i}(t_{i-1},x)\big),v^\pi_i\big(X_{t_i}(t_{i-1},x)\big)\Big)\Delta t_i,\\
&u^\pi_{i-1}(x)\stackrel{\triangle}{=}E\{U^\pi_i(x,\omega)\},\qquad v^\pi_{i-1}(x)\stackrel{\triangle}{=}\frac{1}{\Delta t_i}E\{U^\pi_i(x,\omega)[W_{t_i}-W_{t_{i-1}}]\}.
\end{aligned}\tag{6.8}$$

We still have

$$\widehat{Y}^{\pi}_{t_i} = Y^{\pi}_{t_i} = u^{\pi}_i(\hat{X}^{\pi}_{t_i}), \qquad \widehat{Z}^{\pi}_{t_i} = Z^{\pi,1}_{t_i} = Z^{\pi}_{t_i} = v^{\pi}_i(\hat{X}^{\pi}_{t_i}).$$

Moreover, since $E\{|g(X_T) - g(\hat{X}^{\pi}_T)|^2\} \le C(1 + |x|^2)|\pi|$, we conclude that the rate of convergence becomes $C(1 + |x|^2)|\pi|$, which coincides with the result of [16].

Acknowledgments. The author thanks J. Ma for lots of useful discussion and the referee for his/her careful reading of the manuscript and many helpful suggestions.

REFERENCES

[1] ASMUSSEN, S., GLYNN, P. and PITMAN, J. (1995). Discretization error in simulation of one-dimensional reflecting Brownian motion. *Ann. Appl. Probab.* **5** 875–896.

[2] BALLY, V. (1995). An approximation scheme for BSDEs and applications to control and nonlinear PDEs. Prepublication 95-15 du Laboratoire de Statistique et Processus de l'Universite du Maine.

[3] BISMUT, J. M. (1973). Théorie probabiliste du contrôle des diffusions. *Mem. Amer. Math. Soc.* **176**.

[4] BJÖRK, T. (1998). *Arbitrage Theory in Continuous Time.* Oxford Univ. Press.

[5] BRIAND, P., DELYON, B. and MÉMIN, J. (2001). Donsker-type theorem for BSDEs. *Electron. Comm. Probab.* **6** 1–14.

[6] CHEVANCE, D. (1996). Ph.D. thesis, Univ. Provence.

[7] DAI, M. and JIANG, L. (2000). On path-dependent options. In *Mathematical Finance—Theory and Applications* (J. Yong and R. Cont, eds.) 290–316. Higher Education Press, China.

[8] DOUGLAS, J., Jr., MA, J. and PROTTER, P. (1996). Numerical methods for forward–backward stochastic differential equations. *Ann. Appl. Probab.* **6** 940–968.

[9] EL KAROUI, N. and MAZLIAK, L. (1997). *Backward Stochastic Differential Equations*. Longman, Harlow.

[10] EL KAROUI, N., PENG, S. and QUENEZ, M. C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–72.

[11] GIKHMAN, I. I. and SKOROHOD, A. V. (1972). *Stochastic Differential Equations*. Springer, New York.

[12] KARATZAS, I. and SHREVE, S. E. (1987). *Brownian Motion and Stochastic Calculus*. Springer, New York.

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

[14] MA, J., PROTTER, P., SAN MARTÍN, J. and TORRES, S. (2002). Numerical method for backward SDEs. *Ann. Appl. Probab.* **12** 302–316.

[15] 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.

[16] MA, J. and YONG, J. (1999). *Forward–Backward Stochastic Differential Equations and Their Applications. Lecture Notes in Math.* **1702**. Springer, Berlin.

[17] MA, J. and ZHANG, J. (2002). A representation theorem for backward SDE's with non-smooth coefficients. *Ann. Appl. Probab.* **12** 1390–1418.

[18] MA, J. and ZHANG, J. (2002). Path regularity for solutions of backward SDE's. *Probab. Theory Related Fields* **122** 163–190.

[19] NUALART, D. (1995). *The Malliavin Calculus and Related Topics*. Springer, New York.

[20] PARDOUX, E. and PENG, S. (1990). Adapted solutions of backward stochastic equations. *System and Control Lett.* **14** 55–61.

[21] PARDOUX, E. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Lecture Notes in Control and Inform. Sci.* **176** 200–217. Springer, Berlin.

[22] ZHANG, J. (2001). Some fine properties of backward stochastic differential equations. Ph.D. dissertation, Purdue Univ.

[23] ZHANG, Y. and ZHENG, W. (2001). Discretizing a backward stochastic differential equations. Preprint.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA 90089
USA
E-MAIL: jianfenz@usc.edu