Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Stochastic Processes and their Applications 115 (2005) 539–569
stochastic processes and their applications
www.elsevier.com/locate/spa

Representations and regularities for solutions to BSDEs with reflections

Jin Ma[a,*,1], Jianfeng Zhang[b]

[a]Department of Mathematics, Purdue University, 150 N, University Street, West Lafayette, IN 47907-1395, USA
[b]Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA

Received 19 September 2002; received in revised form 1 December 2003; accepted 20 May 2004
Available online 18 January 2005

Abstract

In this paper we study a class of backward stochastic differential equations with reflections (BSDER, for short). Three types of discretization procedures are introduced in the spirit of the so-called Bermuda Options in finance, so as to first establish a Feynman–Kac type formula for the martingale integrand of the BSDER, and then to derive the continuity of the paths of the martingale integrand, as well as the C^1-regularity of the solution to a corresponding obstacle problem. We also introduce a new notion of regularity for a stochastic process, which we call the "L^2-modulus regularity". Such a regularity is different from the usual path regularity in the literature, and we show that such regularity of the martingale integrand produces exactly the rate of convergence of a numerical scheme for BSDERs. Both numerical scheme and its rate of convergence are novel.

MSC: primary 60H10, secondary 34F05, 90A12

Keywords: Backward SDEs with reflections; Feynman–Kac formulae; Path regularities; Pseudo-approximations; L^2-modulus; Bermuda options; Rate of convergence

*Corresponding author. Tel.: +1 7654941973; fax: +1 7654940548.
E-mail addresses: majin@math.purdue.edu (J. Ma), jianfenz@usc.edu (J. Zhang).
[1]This author is supported in part by NSF grant #9971720.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2004.05.010

1. Introduction

In this paper we are interested in the following backward stochastic differential equations with reflection (BSDER for short) of the form:

$$\begin{aligned} X_t &= x + \int_0^t b(r, X_r)\,\mathrm{d}r + \int_0^t \sigma(r, X_r)\,\mathrm{d}W_r, \\ Y_t &= g(X_T) + \int_t^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_t^T \langle Z_r, \mathrm{d}W_r\rangle + K_T - K_t, \\ Y_t &\geqslant S_t, \quad t \geqslant 0; \quad \int_0^T [Y_t - S_t]\,\mathrm{d}K_t = 0, \end{aligned} \tag{1.1}$$

where b, σ, f, g are deterministic functions, W is a given Brownian motion, S_t is the reflecting barrier, and K is the reflecting process that keeps the solution Y from going below the barrier S at each time t. We will be particularly interested in the case $S_t = h(t, X_t)$, where h is a deterministic function satisfying $h(T, x) \leqslant g(x)$ for all $x \in \mathbb{R}^d$.

The BSDEs with reflections have been studied by many authors (see, for example, [6,5,11,13,19], to mention a few). One of the main motivations for studying such BSDER has been to use it to solve the hedging problem for American options (see, e.g. [7,13], or [14]). In that application the BSDER (1.1) is particularly useful, with Y_t being the option price at time t and Z_t being the hedging portfolio at time t.

The main purpose of this paper is two-fold. First, we would like to establish a Feynman–Kac type representation formula for the solutions to BSDER, especially for the martingale integrand Z. Second, we would like to utilize such representation to study various types of "regularities" of the martingale integrand Z. It turns out that our representation (for Z) is quite similar to the one in [16], which is the extension of the works of Fournié et al. [9,10]. The main feature of such representation is that it is independent of the derivatives of the functions f and g, although it in essence represents the gradient of the solution to an obstacle problem for a quasilinear PDE, in light of the existing nonlinear Feynman–Kac formula involving BSDERs (see, e.g. [7,13]). To our best knowledge, such a representation is new.

An apparent consequence of the representation is that the process Z is *continuous*. Unlike the situation for BSDEs without reflections (cf. [16]), such continuity is by no means obvious for BSDERs with only Lipschitz coefficients. For example, applying Tanaka–Itô formula to the process $Y_t = -|W_t|$, $t \geqslant 0$ (i.e., $h(t,x) = g(x) = -|x|$, and $X = W$), one obtains that $Z_t = -\mathrm{sgn}\{W_t\}$ and $K_t = L_t^0$, for $t \geqslant 0$, where L^0 is the local time of W at zero. Thus the process Z is *discontinuous*(!). We will nevertheless prove that the paths of Z are continuous under the assumption that function $h \in C^{1,2}([0,T])$. In fact, we will prove that the solution to the quasilinear obstacle problem is at least $C^{1/2,1}$ under such assumption, and it is our hope that the further development of these representations will pave the way to study regularity of solutions to the corresponding variational inequalities probabilistically.

At this point we would like to point out that the method of our previous works [16,15] (or of [9,10]), that is, via integration by parts formula in Malliavin Calculus, does not work well with the reflecting process K. We thus take a slight detour by introducing three discrete versions of (1.1): the first approximation follows the idea of the "*Bermuda Option*"-approximation in finance, which is the basis for the other two; the second one removes the reflection part, and leads to the representation theorems as well as the regularity results; and the third one can be considered as a numerical method for (1.1). We should note that although the idea of approximating general American option by Bermuda options is commonly used in numerical finance (cf., e.g. [1,3,12,4]), the numerical analysis for BSDERs is yet fully explored. Our study of the rate of convergence for the numerical scheme involves a new notion of regularity for stochastic processes, which we will call the "*L^2-modulus regularity*" in this paper, and it is different in nature from the usual path regularity as we will show by examples in the appendix. Our result on the rate of convergence of the numerical scheme is built upon the L^2-modulus regularity of the martingale integrand Z. To our best knowledge, such a rate of convergence is new. We should remark that after this paper was finished we were informed of a recent work by Bouchard and Touzi [2]. Although in the BSDER case, only a simplified version (no Z in the generator) was considered there and no rate of convergence was given in the general case, we feel that their method of computing the conditional expectations indeed provided an excellent complementary aspect of our scheme. We shall elaborate this point more in Section 7.

This paper is organized as follows. We give preliminaries in Section 2, and introduce the first discretized BSDER in Section 3. In Section 4 we prove a Feynman–Kac-type representation theorem; and use it to study the regularity of the obstacle problem in Section 5. In Section 6 we establish the L^2-modulus regularity for the martingale integrand Z and in Section 7 we use these results to study the rate of convergence of a numerical scheme for BSDER. Finally, the two examples given in the appendix are for better understanding the notion of L^2-modulus regularity.

2. Preliminaries

Throughout the paper, we assume that $(\Omega, \mathscr{F}, P; \{\mathscr{F}_t\})$ is a complete, filtered probability space on which is defined a standard d-dimensional Brownian motion W. We assume that the filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ is generated by the Brownian motion W, with usual augmentation. Thus it satisfies the *usual hypotheses*.

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

(A1) b and σ are continuous, differentiable and with uniformly bounded derivatives w.r.t. variable x; and $\sigma\sigma^T \geqslant \delta I_d$, where $\delta > 0$, and I_d is the $d \times d$ identity matrix.

(A2) f and g are continuous; and are uniformly Lipschitz w.r.t. variables (x, y, z).

(A3) $h \in C^{1,2}([0, T] \times \mathbb{R}^d)$ with all derivatives being uniformly bounded.

To simplify notations we will use a generic constant $L>0$ to denote all the bounds in (A1)–(A3), and we will also assume that

$$\sup_{0\leqslant t\leqslant T}[|b(t,0)|+|\sigma(t,0)|+|f(t,0,0,0)|+|h(t,0)|]+|g(0)|\leqslant L. \tag{2.1}$$

Finally, to simplify the presentation we shall discuss only the case when $d=1$. But all the results in this paper can be extended to the case $d>1$ without any significant difficulty. Thus in the rest of the paper we consider the following BSDER:

$$\begin{aligned} X_t &= x+\int_0^t b(r,X_r)\,\mathrm{d}r+\int_0^t \sigma(r,X_r)\,\mathrm{d}W_r,\\ Y_t &= g(X_T)+\int_t^T f(r,X_r,Y_r,Z_r)\,\mathrm{d}r-\int_t^T Z_r\,\mathrm{d}W_r+K_T-K_t,\\ Y_t &\geqslant h(t,X_t),\quad t\geqslant 0;\qquad \int_0^T [Y_t-h(t,X_t)]\,\mathrm{d}K_t=0. \end{aligned} \tag{2.2}$$

Recall that an (adapted) solution to BSDER (2.2) is a quadruple of processes (X,Y,Z,K) such that (2.2) holds almost surely. We will often denote $\Theta\triangleq(X,Y,Z)$ for notational simplicity.

It is well known that (2.2) is closely related to the following obstacle problem:

$$\begin{aligned} &\min(u-h(t,x),-u_t-\tfrac{1}{2}\sigma^2(t,x)u_{xx}-b(t,x)u_x-f(t,x,u,u_x\sigma))=0,\\ &u(T,x)=g(x). \end{aligned} \tag{2.3}$$

In fact, denoting $(X^{t,x},Y^{t,x},Z^{t,x},K^{t,x})$ to be the solution to (2.2) over subinterval $[t,T]\subseteq[0,T]$, with $X_t=x$ a.s., then the function $u(\cdot,\cdot)$ defined by

$$u(t,x)\triangleq Y_t^{t,x}=E\left\{g(X_T^{t,x})+\int_t^T f(r,\Theta_r^{t,x})\,\mathrm{d}r+K_T^{t,x}\right\} \tag{2.4}$$

is the unique viscosity solution to the obstacle problem (cf. e.g., Ma–Cvitanic [13]). On the other hand, by Markovian properties of (2.2) we have $Y_s^{t,x}=u(s,X_s^{t,x}), s\in[t,T]$.

Our discretization begins from the following standard Euler scheme for the forward SDE in (2.2): denote for any $x\in\mathbb{R}$ and $0\leqslant s\leqslant t\leqslant T$,

$$X_t(s,x)\triangleq x+b(s,x)(t-s)+\sigma(s,x)(W_t-W_s). \tag{2.5}$$

Let $\pi:0=t_0<\cdots<t_n=T$ be any partition of $[0,T]$, and define

$$X_{t_0}^{\pi,0}\triangleq x;\quad X_t^{\pi,0}\triangleq X_t(t_{i-1},X_{t_{i-1}}^{\pi,0});\quad t\in(t_{i-1},t_i]. \tag{2.6}$$

We shall denote X^π to be the *piecewise constant* version of $X^{\pi,0}$, that is, $X_t^\pi\triangleq X_{t_{i-1}}^{\pi,0}$, $t\in[t_{i-1},t_i)$. Using the process $X^{\pi,0}$ we define also the corresponding BSDER:

$$\begin{aligned} Y_t^{\pi,0} &= g(X_T^{\pi,0})+\int_t^T f(r,\Theta_r^{\pi,0})\,\mathrm{d}r-\int_t^T Z_r^{\pi,0}\,\mathrm{d}W_r+K_T^{\pi,0}-K_t^{\pi,0},\\ Y_t^{\pi,0} &\geqslant h(t,X_t^{\pi,0}),\quad t\geqslant 0;\qquad \int_0^T[Y_t^{\pi,0}-h(t,X_t^{\pi,0})]\,\mathrm{d}K_t^{\pi,0}=0, \end{aligned} \tag{2.7}$$

We now collect some results regarding the solution (Θ, K) and its discretized counterpart $(\Theta^{\pi,0}, K^{\pi,0})$. Most of these results are well-known, so we only list them for ready references. We note in all the estimates C is a generic constant depending only on T, L, and δ, which may vary from line to line. We will also use C_p to denote constants which may depend on p as well. We begin by some standard estimates for forward SDEs.

Lemma 2.1. *Assume* (A1). *Then for any* $p \geqslant 2$, *there exists a constant* $C_p > 0$, *such that*

(i) $E\{\sup_{0\leqslant t\leqslant T}[|X_t|^p + |X_t^{\pi,0}|^p + |X_t^{\pi}|^p]\} \leqslant C_p(1+|x|^p)$,
(ii) $E\{\sup_{0\leqslant t\leqslant T}|X_t^{\pi,0} - X_t|^p\} + \sup_{0\leqslant t\leqslant T} E\{|X_t^{\pi} - X_t|^p\} \leqslant C_p(1+|x|^p)|\pi|^{p/2}$,
(iii) $E\{[|X_t - X_s|^p + |X_t^{\pi,0} - X_s^{\pi,0}|^p]\} \leqslant C_p(1+|x|^p)|t-s|^{p/2}$.

Next, we give some standard estimates for BSDERs (see, e.g. [6]).

Lemma 2.2. *Assume* (A1)–(A3), *and let* (Θ, K) *and* $(\Theta^{\pi,0}, K^{\pi,0})$ *be the solutions to* (2.2) *and* (2.7), *respectively*.

(i) *There exists some universal constant* $C>0$, *such that*

$$E\left\{\sup_{0\leqslant t\leqslant T}|Y_t|^2 + \int_0^T |Z_t|^2\,\mathrm{d}t + K_T^2\right\} \leqslant C(1+|x|^2),$$
$$E\left\{\sup_{0\leqslant t\leqslant T}|Y_t^{\pi,0}|^2 + \int_0^T |Z_t^{\pi,0}|^2\,\mathrm{d}t + |K_T^{\pi,0}|^2\right\} \leqslant C(1+|x|^2). \tag{2.8}$$

(ii)
$$E\left\{\sup_{0\leqslant t\leqslant T}[|Y_t^{\pi,0} - Y_t|^2 + |K_t^{\pi,0} - K_t|^2] + \int_0^T |Z_t^{\pi,0} - Z_t|^2\right\} \leqslant C(1+|x|^2)\sqrt{|\pi|}.$$

(iii) *Denote*

$$k_t \triangleq |\partial_t h(t,X_t)| + |[(\partial_x h)b](t,X_t)| + \tfrac{1}{2}|[(\partial_{xx}h)\sigma^2](t,X_t)| + |f(t,\Theta_t)|,$$
$$k_t^{\pi,0} \triangleq |\partial_t h(t,X_t^{\pi,0})| + |(\partial_x h)(t,X_t^{\pi,0})b(t_{i-1},X_{t_{i-1}}^{\pi,0})| + \tfrac{1}{2}|(\partial_{xx}h)(t,X_t^{\pi,0})\sigma^2(t_{i-1},X_{t_{i-1}}^{\pi,0})| + |f(t,\Theta_t^{\pi,0})|, \quad t\in[t_{i-1},t_i). \tag{2.9}$$

Then, *in the sense of random measures*, *one has*

$$0\leqslant \mathrm{d}K_t \leqslant k_t\,\mathrm{d}t, \quad 0\leqslant \mathrm{d}K_t^{\pi,0} \leqslant k_t^{\pi,0}\,\mathrm{d}t, \quad \forall t\in[0,T], \text{ a.s.}$$

The following comparison theorem for the solutions to BSDERs will be useful in our discussion. We should note that the lemma is stated in a rather general form in which the process K's are not necessarily the reflecting processes(!). We refer to [14] for the proof.

Lemma 2.3. *Assume that* (Y^i, Z^i), $i = 1, 2$ *are the solutions to the following BSDEs,*

$$Y_t^i = Y_T^i + \int_t^T \tilde{f}(s, Y_s^i, Z_s^i)\,\mathrm{d}s - \int_t^T Z_s^i\,\mathrm{d}W_s + K_T^i - K_t^i, \quad t \in [0,T], \;\; i = 1, 2,$$

where $\tilde{f} : \Omega \times [0,T] \times \mathbb{R}^2 \to \mathbb{R}$, *and* K^1 *and* K^2 *are two adapted processes with bounded variation paths. Suppose further that*

(i) $\tilde{f}$ *is uniformly Lipschitz continuous on* (y, z), *uniformly in* (ω, t);
(ii) $Y_T^1 \leqslant Y_T^2$, *a.s.*; *and*
(iii) *in the sense of random* (*signed*) *measure*, $d(K_t^1 - K_t^2) \leqslant 0$, $\forall t$, *a.s.*,

then, $Y_t^1 \leqslant Y_t^2$, $\forall t$, *a.s.*

Finally, let us review some results concerning the BSDEs *without* reflections on which our "Bermuda-option-approximation" depends heavily. Consider the following (decoupled) forward–backward SDEs:

$$\begin{aligned} X_t &= x + \int_0^t b(r, X_r)\,\mathrm{d}r + \int_0^t \sigma(r, X_r)\,\mathrm{d}W_r, \\ Y_t &= g(X_T) + \int_t^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \quad t \in [0,T]. \end{aligned} \tag{2.10}$$

By a slight abuse of notations, we still denote the solution of BSDE (2.10) by $\Theta = (X, Y, Z)$. For $0 \leqslant t < r \leqslant T$, we define the following process:

$$N_r^t \triangleq \frac{1}{r-t} \int_t^r \sigma^{-1}(s, X_s) \nabla X_s\,\mathrm{d}W_s [\nabla X_t]^{-1}, \tag{2.11}$$

where ∇X is the solution to the following linear SDE:

$$\nabla X_t = 1 + \int_0^t \partial_x b(s, X_s) \nabla X_s\,\mathrm{d}s + \int_0^t \partial_x \sigma(s, X_s) \nabla X_s\,\mathrm{d}W_s. \tag{2.12}$$

It is clear that $E\{N_r^t | \mathscr{F}_t\} = 0$.

Lemma 2.4. *Assume* (A1). *Then for any* $p \geqslant 2$, *there exist constants* C *and* C_p *such that*

(i) $$E\{|N_r^t|^p\} \leqslant \frac{C_p}{(r-t)^{p/2}},$$

(ii) $$E\{|N_{r_1}^t - N_{r_2}^t|^2\} \leqslant C \frac{|r_1 - r_2|}{(r_1 - t)(r_2 - t)},$$

(iii) $$(E\{|N_r^{t_1} - N_r^{t_2}|^2\} \leqslant C \frac{|t_1 - t_2|}{(r - t_1)(r - t_2)},$$

(iv) *if φ is a Lebesgue measurable function with polynomial growth, then for any $0 \leqslant t < r \leqslant T$, it holds that*

$$E\{\varphi(X_T)N_T^t|\mathscr{F}_t\} = E\{\varphi(X_T)N_r^t|\mathscr{F}_t\}. \tag{2.13}$$

In particular, if $\varphi \in C^1$, then

$$E\{\varphi(X_T)N_r^t|\mathscr{F}_t\} = E\{\partial_x\varphi(X_T)\nabla X_T|\mathscr{F}_t\}[\nabla X_t]^{-1}. \tag{2.14}$$

Proof. Estimates (i)–(iii) can be proved by straightforward calculation. To see (iv), we first assume that $\varphi \in C^1$. Applying the integration by parts formula of Malliavin Calculus, one can prove (2.14) and (2.13) fairly easily (see [16]).

For general φ, one can choose $\varphi_n \in C^1$ such that φ_n have the same polynomial growth and $\varphi_n(x) \to \varphi(x)$ for a.s. x. Noting that X_T has density (see e.g. [17]), we have $\varphi_n(X_T) \to \varphi(X_T)$, a.s. The result then follows from the Dominated Convergence Theorem. $\square$

Again let us denote $\Theta^{t,x} \triangleq (X^{t,x}, Y^{t,x}, Z^{t,x})$ to be the solution to (2.10) over subinterval $[t,T] \subseteq [0,T]$, with $X_t = x$ a.s. Let $N_\cdot^t(t,x)$ be the process $N_\cdot^t$ with X being replaced by $X^{t,x}$. To distinguish it from the solution to the obstacle problem (2.3) let us now define

$$u^0(t,x) \triangleq Y_t^{t,x}, \quad v^0(t,x) \triangleq Z_t^{t,x}\sigma^{-1}(t,x). \tag{2.15}$$

Then, under assumptions (A1) and (A2), one has (cf., e.g. [16,15,18])

$$\begin{aligned} u^0(t,x) &= E\left\{g(X_T^{t,x}) + \int_t^T f(r,\Theta_r^{t,x})\,\mathrm{d}r\right\}, \\ v^0(t,x) &= E\left\{g(X_T^{t,x})N_T^t(t,x) + \int_t^T f(r,\Theta_r^{t,x})N_r^t(t,x)\,\mathrm{d}r\right\}. \end{aligned} \tag{2.16}$$

Furthermore, under (A1) and (A2), one can show that $v \in C([0,T) \times \mathbb{R})$ and it is uniformly bounded; and the following relations hold: for all $(t,x) \in [0,T] \times \mathbb{R}^n$,

$$v^0(t,x) = \partial_x u^0(t,x), \tag{2.17}$$

$$Y_s^{t,x} = u^0(s, X_s^{t,x}), \quad \forall s \in [t,T], \text{ a.s.}, \tag{2.18}$$

$$\begin{aligned} Z_s^{t,x} &= v^0(s,X_s^{t,x})\sigma(s,X_s^{t,x}) \\ &= E\left\{g(X_T^{t,x})N_T^s(t,x) + \int_s^T f(r,\Theta_r^{t,x})N_r^s(t,x)\,\mathrm{d}r|\mathscr{F}_s\right\}\sigma(s,X_s^{t,x}); \\ &\qquad \forall s \in [t,T], \text{ a.s.}, \end{aligned} \tag{2.19}$$

As a consequence of the results above, one has the following estimates.

Lemma 2.5. *Assume* (A1)–(A3), *and let* (X, Y, Z) *denote the solution to* (2.10). *Then for any* $p \geqslant 2$, *there exists a constant* C_p *such that, for all* $s, t \in [0, T]$,

$$E\left\{\sup_{0 \leqslant t \leqslant T}[|Y_t|^p + |Z_t|^p]\right\} \leqslant C_p(1+|x|^p); \quad E\{|Y_t - Y_s|^p\} \leqslant C_p(1+|x|^p)|t-s|^{p/2}.$$

Proof. The estimate for $E\{\sup_t |Y_t|^p\}$ is standard. By (2.17) and (2.18), we see that $|Z_t| \leqslant C|\sigma(t, X_t)|$, then it follows that $E\{\sup_t |Z_t|^p\} \leqslant C_p(1+|x|^p)$. Finally, note that $Y_s - Y_t = \int_s^t f(r, \Theta_r)\,\mathrm{d}r - \int_s^t Z_r\,\mathrm{d}W_r$, the last inequality follows from the Burkholder–Davis–Gundy inequality and the estimates above. $\square$

To end this section we give a *discrete backward Gronwall Inequality*.

Lemma 2.6. *Let* $\pi : 0 = t_0 < \cdots < t_n = T$, *and denote* $\Delta t_i = t_i - t_{i-1}$. *Suppose that* $\{a_i, b_i\}_{i=1}^n$ *satisfies that* $a_i \geqslant 0$, $b_i \geqslant 0$, *and that* $a_{i-1} \leqslant (1 + C\,\Delta t_i)a_i + b_i$, $\forall i \geqslant 1$, *then*

$$\max_{0 \leqslant i \leqslant n} a_i \leqslant e^{CT}\left[a_n + \sum_{i=1}^n b_i\right]. \tag{2.20}$$

Proof. Note that $1 + C\,\Delta t_i \leqslant e^{C\,\Delta t_i}$. By backward induction one easily gets that

$$a_i \leqslant e^{C(T-t_i)}\left[a_n + \sum_{j=i+1}^n b_j\right], \quad i = 0, 1, \ldots, n$$

and (2.20) follows immediately. $\square$

3. Pseudo-discretization

In this section we introduce our first discretization of the BSDER (2.2). We note that although this discretization is a far cry from the true discrete version of the BSDER, which will be introduced in the end of this paper, it is nevertheless a fundamental step upon which all our discussion will be based. The main idea is similar to the so-called "Bermuda-option-approximation" in finance.

Let $\pi : 0 = t_0 < \cdots < t_n = T$ be any partition of $[0, T]$, and let X be the solution of the forward SDE in (2.2) and $X^{\pi,0}$ be its Euler discretization. We define recursively the processes $Y^{\pi,1}$ and $(\tilde{Y}^{\pi,1}, Z^{\pi,1}, K^{\pi,1})$ as follows:

- $Y^{\pi,1}_{t_n} = g(X^{\pi,0}_T)$;
- for $i = n, n-1, \ldots, 1$, and $t \in [t_{i-1}, t_i)$, let $(\tilde{Y}^{\pi,1}, Z^{\pi,1})$ be the solution of BSDE:

$$\tilde{Y}^{\pi,1}_t = Y^{\pi,1}_{t_i} + \int_t^{t_i} f(r, X^{\pi,0}_r, \tilde{Y}^{\pi,1}_r, Z^{\pi,1}_r)\,\mathrm{d}r - \int_t^{t_i} Z^{\pi,1}_r\,\mathrm{d}W_r, \tag{3.1}$$

- for each i and $t \in [t_{i-1}, t_i)$, define $Y^{\pi,1}_t \triangleq \tilde{Y}^{\pi,1}_t \vee h(t, X^{\pi,0}_t)$,
- let $K^{\pi,1}_0 \triangleq 0$, and for $t \in (t_{i-1}, t_i]$, define $K^{\pi,1}_t \equiv K^{\pi,1}_{t_i} \triangleq \sum_{j=1}^i (Y^{\pi,1}_{t_{j-1}} - \tilde{Y}^{\pi,1}_{t_{j-1}})$.

Clearly, $K^{\pi,1}_{t_i} \in \mathscr{F}_{t_{i-1}}$ for each i, thus $K^{\pi,1}$ is $\{\mathscr{F}_t\}$-predictable. Observe also that (3.1) leads to

$$Y^{\pi,1}_{t_{i-1}} = Y^{\pi,1}_{t_i} + \int_{t_{i-1}}^{t_i} f(r, \tilde{\Theta}^{\pi,1}_r)\,\mathrm{d}r - \int_{t_{i-1}}^{t_i} Z^{\pi,1}_r\,\mathrm{d}W_r + K^{\pi,1}_{t_i} - K^{\pi,1}_{t_{i-1}}, \tag{3.2}$$

where $\tilde{\Theta}^{\pi,1} \triangleq (X^{\pi,0}, \tilde{Y}^{\pi,1}, Z^{\pi,1})$. Since in (3.2), $\tilde{\Theta}^{\pi,1}$ contains $\tilde{Y}^{\pi,1}$ instead of $Y^{\pi,1}$, and $\tilde{Y}^{\pi,1}$ is obtained by solving a (non-discretized) BSDE (3.1), (3.2) is by no means a truly discretized version of (2.2). We henceforth call it a "pseudo-discretization". We also note that, all the results in this section should hold true if we replace $X^{\pi,0}$ by X in (3.1) and the definition of $Y^{\pi,1}$ (which is important in next section!). The following comparison results on $Y^{\pi,1}$'s are useful.

Lemma 3.1. *Assume* (A1)–(A3). *Then* $\tilde{Y}^{\pi,1}_t \leqslant Y^{\pi,1}_t \leqslant Y^{\pi,0}_t$.

Proof. By definition the first inequality is obvious. We prove by induction that $Y^{\pi,1}_t \leqslant Y^{\pi,0}_t$ on each interval $(t_{i-1}, t_i]$. Clearly, $Y^{\pi,1}_{t_n} = Y^{\pi,0}_{t_n}$. Assume that $Y^{\pi,1}_{t_i} \leqslant Y^{\pi,0}_{t_i}$, then by (2.7) and (3.1) and applying Lemma 2.3 we obtain that $\tilde{Y}^{\pi,1}_t \leqslant Y^{\pi,0}_t$, for $\forall t \in [t_{i-1}, t_i)$. Note that $Y^{\pi,0}_t \geqslant h(t, X^{\pi,0}_t)$, thus $Y^{\pi,1}_t = \tilde{Y}^{\pi,1}_t \vee h(t, X^{\pi,0}_t) \leqslant Y^{\pi,0}_t$, for $\forall t \in [t_{i-1}, t_i)$, completing the induction step, whence the lemma. $\square$

We now construct a sequence of functions $u^{\pi,1}_i$ using the pseudo-discretization defined above. Recall $X_\cdot(t_{i-1}, x)$ defined in (2.5). We define $u^{\pi,1}_n(x) \triangleq g(x)$; and for $i = n, \ldots, 1$, let $(\tilde{Y}, \tilde{Z})$ denote the solution to BSDE

$$\tilde{Y}_t = u^{\pi,1}_i(X_{t_i}(t_{i-1}, x)) + \int_t^{t_i} f(r, X_r(t_{i-1}, x), \tilde{Y}_r, \tilde{Z}_r)\,\mathrm{d}r - \int_t^{t_i} \tilde{Z}_r\,\mathrm{d}W_r \tag{3.3}$$

for $t \in [t_{i-1}, t_i]$; and then define $u^{\pi,1}_{i-1}(x) \triangleq \tilde{Y}_{t_{i-1}} \vee h(t_{i-1}, x)$. One can easily check that $Y^{\pi,1}_{t_i} = u^{\pi,1}_i(X^{\pi,0}_{t_i})$. Furthermore, we have

Lemma 3.2. *Assume* (A1)–(A3). *Then*

(i) $u^{\pi,1}_i$ *are uniformly Lipschitz continuous, uniformly in* π *and* i;
(ii) *for any* $p \geqslant 1$, *there exist constants* C *and* C_p, *independent of* π, *such that, for* $\forall i$ *and* $\forall t \in [t_{i-1}, t_i)$,

$$\begin{aligned} &E\left\{\sup_{t_{i-1} \leqslant t \leqslant t_i} [|\tilde{Y}^{\pi,1}_t|^p + |Z^{\pi,1}_t|^p]\right\} \leqslant C_p(1 + |x|^p), \\ &E\{|\tilde{Y}^{\pi,1}_t - Y^{\pi,1}_{t_i}|^2\} \leqslant C(1 + |x|^2)|t_i - t|. \end{aligned} \tag{3.4}$$

Proof. First we assume (i) holds true. We prove (ii) by modifying Lemma 2.5. Consider (3.3) as a special case of (2.10) over $[t_{i-1}, t_i]$, that is, the forward diffusion has constant coefficients $\tilde{b} \triangleq b(t_{i-1}, x), \tilde{\sigma} \triangleq \sigma(t_{i-1}, x)$. Applying (2.19), one can easily prove that $|\tilde{Z}_t| \leqslant C|\sigma(t_{i-1}, x)|$, which implies that $|Z^{\pi,1}_t| \leqslant C|\sigma(t_{i-1}, X^{\pi,0}_{t_{i-1}})|$. Then one can prove (ii) in the same manner as one does in Lemma 2.5.

Thus it suffices to prove (i). We first prove by induction that for each i, $u_i^{\pi,1}$ is Lipschitz. Clearly $u_n^{\pi,1}=g$ is Lipschitz. Assume $u_i^{\pi,1}$ is also Lipschitz, and denote its Lipschitz constant as L_i. We show that $u_{i-1}^{\pi,1}$ is also Lipschitz.

Without loss of generality, we assume that $L_i \geqslant 1$ and that $L_i \geqslant \sup_{t,x} |\partial_x h(t,x)|$. For any $x_1, x_2 \in \mathbb{R}$ and $t \in [t_{i-1}, t_i]$, denote $(\tilde{Y}^j, \tilde{Z}^j)$ to be the solution to (3.3) with initial value x_j, $j=1,2$; and denote

$$
\begin{aligned}
&\Delta x \triangleq x_1 - x_2; \quad \Delta X_t \triangleq X_t(t_{i-1}, x_1) - X_t(t_{i-1}, x_2); \quad \Delta \tilde{Y}_t \triangleq \tilde{Y}_t^1 - \tilde{Y}_t^2; \\
&\Delta \tilde{Z}_t \triangleq \tilde{Z}_t^1 - \tilde{Z}_t^2.
\end{aligned}
$$

Then it is standard to check that $E\{|\Delta X_t|^2\} \leqslant (1 + C\Delta t_i)|\Delta x|^2$. Applying Itô's formula to $|\Delta \tilde{Y}|^2$ along with (3.3), and using some by now standard technique in BSDEs, one shows that

$$
\begin{aligned}
&E\left\{ |\Delta \tilde{Y}_t|^2 + \int_t^{t_i} |\Delta \tilde{Z}_r|^2 \, \mathrm{d}r \right\} \\
&\quad \leqslant E\left\{ L_i^2 |\Delta X_{t_i}|^2 + C \int_t^{t_i} |\Delta \tilde{Y}_r|^2 \, \mathrm{d}r + \frac{1}{2} \int_t^{t_i} [|\Delta X_r|^2 + |\Delta \tilde{Z}_r|] \, \mathrm{d}r \right\}.
\end{aligned} \tag{3.5}
$$

In the above $C>0$ is again a generic constant depending on L and T, which is allowed to vary from line to line. Now note that $L_i \geqslant 1$, thus

$$
E\left\{ |\Delta \tilde{Y}_t|^2 + \frac{1}{2} \int_t^{t_i} |\Delta \tilde{Z}_r|^2 \, \mathrm{d}r \right\} \leqslant E\left\{ (1 + C\Delta t_i) L_i^2 |\Delta x|^2 + C \int_t^{t_i} |\Delta \tilde{Y}_r|^2 \, \mathrm{d}r \right\}.
$$

Now applying the Gronwall inequality we deduce that $E\{|\Delta \tilde{Y}_t|^2\} \leqslant (1 + C\Delta t_i) L_i^2 |\Delta x|^2$ with a generic constant C. Now note that for any real numbers a_1, a_2, b_1, b_2 one has $|a_1 \vee b_1 - a_2 \vee b_2| \leqslant |a_1 - a_2| \vee |b_1 - b_2|$, the definition of $u^{\pi,1}$ and $Y^{\pi,1}$ then imply that

$$
\begin{aligned}
|u_{i-1}^{\pi,1}(x_1) - u_{i-1}^{\pi,1}(x_2)|^2 &\leqslant |\tilde{Y}_{t_{i-1}}^1 - \tilde{Y}_{t_{i-1}}^2|^2 \vee |h(t_{i-1}, x_1) - h(t_{i-1}, x_2)|^2 \\
&\leqslant [(1 + C\Delta t_i) L_i^2 |\Delta x|^2] \vee [L_i^2 |\Delta x|^2] \\
&= (1 + C\Delta t_i) L_i^2 |\Delta x|^2.
\end{aligned} \tag{3.6}
$$

Here in the last inequality we used the assumption that $L_i \geqslant \sup |\partial_x h|$. Thus we conclude that $u_{i-1}^{\pi,1}$ is also Lipschitz.

To finish the proof, we need only show that the Lipschitz constant for each $u_{i-1}^{\pi,1}$, denoted by L_{i-1}, can be chosen to be independent of π and i. But from (3.6) we see that $L_{i-1}^2 \leqslant (1 + C\Delta t_i) L_i^2$. Thus the 0a follows from Lemma 2.6. □

The main result of this section is the following theorem.

Theorem 3.3. *Assume* (A1)–(A3). *Then the following estimate holds*:

$$
\begin{aligned}
&E\left\{ \sup_{0 \leqslant t \leqslant T} [|\tilde{Y}_t^{\pi,1} - Y_t^{\pi,0}|^2 + |K_t^{\pi,1} - K_t^{\pi,0}|^2] + \int_0^T |Z_t^{\pi,1} - Z_t^{\pi,0}|^2 \, \mathrm{d}t \right\} \\
&\quad \leqslant C(1 + |x|^4)|\pi|.
\end{aligned}
$$

Proof. We first prove that under assumptions (A1)–(A3), it holds that

$$\max_{0\leqslant i\leqslant n} E\{|\tilde{Y}^{\pi,1}_{t_i} - Y^{\pi,0}_{t_i}|^2\} + E\left\{\int_0^T |Z^{\pi,1}_t - Z^{\pi,0}_t|^2\right\} \leqslant C(1+|x|^4)|\pi|. \tag{3.7}$$

To this end, let us denote $\Delta Y_t \triangleq Y^{\pi,0}_t - \tilde{Y}^{\pi,1}_t$ and $\Delta Z_t \triangleq Z^{\pi,0}_t - Z^{\pi,1}_t$. Recall (2.7), (3.1) and apply Itô's formula, we have, for $t \in (t_{i-1}, t_i)$,

$$\begin{aligned} d|\Delta Y_t|^2 &= 2\Delta Y_t \Delta Z_t \, \mathrm{d}W_t - 2\Delta Y_t[f(t,\Theta^{\pi,0}_t) - f(t,\tilde{\Theta}^{\pi,1}_t)]\,\mathrm{d}t \\ &\quad - 2\Delta Y_t \, \mathrm{d}K^{\pi,0}_t + |\Delta Z_t|^2 \, \mathrm{d}t. \end{aligned}$$

Since $2|\Delta Y_t||f(t,\Theta^{\pi,0}_t) - f(t,\tilde{\Theta}^{\pi,1}_t)| \leqslant C|\Delta Y_t|^2 + \frac{1}{2}|\Delta Z_t|^2$ for some constant $C>0$, denoting $\Lambda_t = e^{Ct}$ and applying Lemma 3.1 one shows that

$$\begin{aligned} \Lambda_t|\Delta Y_t|^2 + \frac{1}{2}\int_t^{t_i} \Lambda_r\,|\Delta Z_r|^2\,\mathrm{d}r \leqslant \Lambda_{t_i}\,|\Delta Y_{t_i}|^2 - 2\int_t^{t_i} \Lambda_r\,\Delta Y_r\,\Delta Z_r\,\mathrm{d}W_r \\ + 2\int_t^{t_i} \Lambda_r\,\Delta Y_r\,\mathrm{d}K^{\pi,0}_r. \end{aligned} \tag{3.8}$$

Next, applying Itô's formula to $h(\cdot, X^{\pi,0}_\cdot)$, using (A3), and applying Lemma 2.2 we have

$$\begin{aligned} \Delta Y_r \, \mathrm{d}K^{\pi,0}_r &= [h(r,X^{\pi,0}_r) - \tilde{Y}^{\pi,1}_r]\,\mathrm{d}K^{\pi,0}_r \\ &= \left[h(r,X^{\pi,0}_r) - E\left\{Y^{\pi,1}_{t_i} + \int_r^{t_i} f(s,\tilde{\Theta}^{\pi,1}_s)\,\mathrm{d}s|\mathscr{F}_r\right\}\right]\mathrm{d}K^{\pi,0}_r \\ &\leqslant E\left\{h(r,X^{\pi,0}_r) - h(t_i,X^{\pi,0}_{t_i}) - \int_r^{t_i} f(s,\tilde{\Theta}^{\pi,1}_s)\,\mathrm{d}s|\mathscr{F}_r\right\}\mathrm{d}K^{\pi,0}_r \\ &\leqslant CE\left\{\int_r^{t_i} [1 + |X^{\pi,0}_{t_{i-1}}|^2 + |X^{\pi,0}_s| + |\tilde{Y}^{\pi,1}_s| + |Z^{\pi,1}_s|]\,\mathrm{d}s|\mathscr{F}_r\right\}k^{\pi,0}_r\,\mathrm{d}r, \end{aligned}$$

Plugging this into (3.8) and combining it with (3.4) one further shows that

$$\begin{aligned} &E\left\{\Lambda_t|\Delta Y_t|^2 + \frac{1}{2}\int_t^{t_i} \Lambda_r|\Delta Z_r|^2\,\mathrm{d}r\right\} \\ &\quad\leqslant \Lambda_{t_i}E\left\{|\Delta Y_{t_i}|^2 + C\int_t^{t_i}\int_r^{t_i} |k^{\pi,0}_r|^2\,\mathrm{d}s\,\mathrm{d}r\right. \\ &\qquad \left. + C\int_t^{t_i}\int_r^{t_i} [1 + |X^{\pi,0}_{t_{i-1}}|^4 + |X^{\pi,0}_s|^2 + |\tilde{Y}^{\pi,1}_s|^2 + |Z^{\pi,1}_s|^2]\,\mathrm{d}s\,\mathrm{d}r\right\} \\ &\quad\leqslant \Lambda_t E\left\{(1 + C\,\Delta t_i)|\Delta Y_{t_i}|^2 + C(1+|x|^4)|\Delta t_i|^2 + C_0\,\Delta t_i \int_t^{t_i} |\Delta Z_r|^2\,\mathrm{d}r\right\}. \end{aligned}$$

Choosing π such that $|\pi| \leqslant 1/4C_0$ and noting that $\Lambda_r \geqslant \Lambda_t$ for $r \geqslant t$, we deduce that

$$E\left\{|\Delta Y_t|^2 + \frac{1}{4}\int_t^{t_i} |\Delta Z_r|^2\,\mathrm{d}r\right\} \leqslant E\{(1 + C\,\Delta t_i)|\Delta Y_{t_i}|^2 + C(1+|x|^4)|\Delta t_i|^2\}. \tag{3.9}$$

Since $\Delta Y_{t_n}=0$, setting $t=t_{i-1}$ in (3.9) we can then apply Lemma 2.6 to get $\max_{0\leqslant i\leqslant n} E\{|\Delta Y_{t_i}|^2\}\leqslant C(1+|x|^4)|\pi|$. Putting this back into (3.9) we obtain that

$$\sup_{0\leqslant t\leqslant T} E\{|\Delta Y_t|^2\}+E\left\{\int_0^T |\Delta Z_t|^2\,\mathrm{d}t\right\}\leqslant C(1+|x|^4)|\pi|. \tag{3.10}$$

Thus (3.7) follows. Furthermore, combining (3.8) and (3.10), it is standard to show (using the Burkholder–Davis–Gundy inequality) that $E\{\sup_{0\leqslant t\leqslant T}|\Delta Y_t|^2\}\leqslant C(1+|x|^4)|\pi|$.

It remains to check the estimate for K. But by (2.2), (3.2) and (3.1) we have

$$\begin{aligned} K_t^{\pi,0} &= Y_0^{\pi,0}-Y_t^{\pi,0}-\int_0^t f(r,\Theta_r^{\pi,0})\,\mathrm{d}r+\int_0^t Z_r^{\pi,0}\,\mathrm{d}W_r,\\ K_t^{\pi,1} &= Y_0^{\pi,1}-\tilde{Y}_t^{\pi,1}-\int_0^t f(r,\tilde{\Theta}_r^{\pi,1})\,\mathrm{d}r+\int_0^t Z_r^{\pi,1}\,\mathrm{d}W_r. \end{aligned} \tag{3.11}$$

for $t\in(t_{i-1},t_i]$. Therefore,

$$\begin{aligned}\Delta K_t\triangleq K_t^{\pi,0}-K_t^{\pi,1} = Y_0^{\pi,0}-Y_0^{\pi,1}-\Delta Y_t-\int_0^t[f(r,\Theta_r^{\pi,0})-f(r,\tilde{\Theta}_r^{\pi,1}]\,\mathrm{d}r\\ +\int_0^t \Delta Z_r\,\mathrm{d}W_r.\end{aligned}$$

Applying the Burkholder–Davis–Gundy inequality again one can easily prove that

$$E\left\{\sup_{0\leqslant t\leqslant T}|\Delta K_t|^2\right\}\leqslant C(1+|x|^4)|\pi|.$$

The proof is now complete. □

The following result is a direct consequence of Lemma 2.2(ii) and Theorem 3.3.

Corollary 3.4. *Assume* (A1)–(A3). *Then*

$$E\left\{\sup_{0\leqslant t\leqslant T}[|\tilde{Y}_t^{\pi,1}-Y_t|^2+|K_t^{\pi,1}-K_t|^2]+\int_0^T|Z_t^{\pi,1}-Z_t|^2\,\mathrm{d}t\right\}\leqslant C(1+|x|^4)\sqrt{|\pi|}.$$

To conclude this section we give the following estimates of the solution of the original BSDER (2.2), which are interest in their own rights.

Theorem 3.5. *Assume* (A1)–(A3). *Then*

(i) *for any* $p\geqslant 1$, *there exist constants* C *and* C_p, *such that for a.s.* $t\in[0,T]$, *and any* $t_1,t_2\in[0,T]$, *it holds that*

$$\begin{aligned} &E\{|Z_t|^p\}\leqslant C_p(1+|x|^p);\quad E\{|k_t|^p\}\leqslant C_p(1+|x|^{2p}),\\ &E\{|Y_{t_1}-Y_{t_2}|^2\}\leqslant C(1+|x|^4)|t_1-t_2|. \end{aligned} \tag{3.12}$$

(ii) *For any* $0\leqslant t_1<t_2\leqslant T$ *and any* x_1,x_2, *it holds that*

$$|u(t_1,x_1)-u(t_2,x_2)|\leqslant C[(1+|x_1|^2)\sqrt{t_2-t_1}+|x_1-x_2|].$$

Proof. (i) Recalling (3.4) and Corollary 3.4, and applying Fatou's Lemma, one has $E\{|Z_t|^p\} \leqslant C_p(1+|x|^p)$, for a.s. t. Then the estimate for k_t follows immediately. Moreover, for $t_1 < t_2$, by (2.2) we have

$$Y_{t_1} - Y_{t_2} = \int_{t_1}^{t_2} f(r, \Theta_r)\,\mathrm{d}r - \int_{t_1}^{t_2} Z_r\,\mathrm{d}W_r + \int_{t_1}^{t_2} \mathrm{d}K_r.$$

One can easily prove (3.12) now.

(ii) First, by Lemma 3.1 and then by Theorem 3.3 and Lemma 2.2 we know

$$\begin{aligned}|u_0^{\pi,1}(x) - u(0,x)| &= |Y_0^{\pi,1} - Y_0| \leqslant |Y_0^{\pi,1} - Y_0^{\pi,0}| + |Y^{\pi,0} - Y_0| \\ &\leqslant |\tilde{Y}_0^{\pi,1} - Y_0^{\pi,0}| + |Y^{\pi,0} - Y_0| \leqslant C(1+|x|^4)\sqrt{|\pi|}.\end{aligned}$$

Then by Lemma 3.2(i), $u(0,x)$ is Lipschitz continuous on x, with a Lipschitz constant C depending only on T and L. Repeat the same arguments for BSDER (2.2) over $[t, T]$, we have

$$|u(t,x_1) - u(t,x_2)| \leqslant C|x_1 - x_2|, \quad \forall t \in [0,T]. \tag{3.13}$$

Moreover, for any $t_1 < t_2$, by (i) we have

$$\begin{aligned}|u(t_1,x) - u(t_2,x)| &\leqslant E\{|u(t_1,x) - u(t_2, X_{t_2}^{t_1,x})| + |u(t_2, X_{t_2}^{t_1,x}) - u(t_2,x)|\} \\ &\leqslant E\{|Y_{t_1}^{t_1,x} - Y_{t_2}^{t_1,x}| + C|X_{t_2}^{t_1,x} - x|\} \leqslant C(1+|x|^2)\sqrt{t_2 - t_1},\end{aligned}$$

which, combined with (3.13), proves (ii). □

4. Representation formulae

In this section we present the first main result of the paper: the representation formula for the martingale integrand Z. We begin by modifying our fundamental (pseudo) discretization investigated in the last section, so that it is "closer" to the original equations (2.2). We proceed as follows. For $i = n, n-1, \ldots, 1$, and $t \in [t_{i-1}, t_i)$, let $(\tilde{Y}^{\pi,2}, Z^{\pi,2})$ be the solution of BSDE:

$$\tilde{Y}_t^{\pi,2} = Y_{t_i} + \int_t^{t_i} f(r, X_r, \tilde{Y}_r^{\pi,2}, Z_r^{\pi,2})\,\mathrm{d}r - \int_t^{t_i} Z_r^{\pi,2}\,\mathrm{d}W_r; \tag{4.1}$$

Analogous to previous section we denote $\tilde{\Theta}^{\pi,2} \triangleq (X, \tilde{Y}^{\pi,2}, Z^{\pi,2})$, and define

$$\begin{aligned}&Y_{t_n}^{\pi,2} \triangleq g(X_T); \quad Y_t^{\pi,2} \triangleq \tilde{Y}_t^{\pi,2} \vee h(t, X_t), \quad t \in [0,T), \\ &K_0^{\pi,2} \triangleq 0; \quad K_t^{\pi,2} \equiv K_{t_i}^{\pi,2} \triangleq \sum_{j=1}^{i} (Y_{t_{j-1}}^{\pi,2} - \tilde{Y}_{t_{j-1}}^{\pi,2}), \quad \forall t \in (t_{i-1}, t_i].\end{aligned}$$

We note that (4.1) differs from (3.1) in two ways. First, we replaced $X^{\pi,0}$ by the original diffusion X; Second, the terminal value of the BSDE (4.1) is the true solution Y_{t_i} of (2.2), rather than the approximate value $Y_{t_i}^{\pi,2}$. Note that, since $Y_{t_i} = u(t_i, X_{t_i})$

and $u(t_i,\cdot)$ are uniformly Lipschitz continuous(!), thanks to Theorem 3.5(ii), one can still apply Lemma 2.5 and the representation results (2.16)–(2.19) to BSDE (4.1).

The following Lemma can be proved by using similar (in fact easier) arguments as those of Lemma 3.2 and Theorem 3.3. We state only the result without the proof.

Lemma 4.1. *Assume* (A1)–(A3). *Then it holds that*

(i) $$E\left\{\sup_{t_{i-1}\leqslant t\leqslant t_i}[|\tilde{Y}^{\pi,2}_t|^p+|Z^{\pi,2}_t|^p]\right\}\leqslant C_p(1+|x|^p),\quad \forall i;$$

(ii) $$E\left\{\sup_{0\leqslant t\leqslant T}[|\tilde{Y}^{\pi,2}_t-Y_t|^2+|Y^{\pi,2}_t-Y_t|^2]+\int_0^T|Z^{\pi,2}_t-Z_t|^2\,\mathrm{d}t\right\}\leqslant C(1+|x|^4)|\pi|.$$

The main result of this section is the following.

Theorem 4.2. *Assume* (A1)–(A3). *Let* $(\Theta,K)=(X,Y,Z,K)$ *be the solution to BSDER* (2.2). *Then, the martingale integrand* Z *can be written as*

$$Z_t=E\left\{g(X_T)N^t_T+\int_t^T f(r,\Theta_r)N^t_r\,\mathrm{d}r+\int_t^T N^t_r\,\mathrm{d}K_r|\mathscr{F}_t\right\}\sigma(t,X_t),\quad \forall t\in[0,T).\tag{4.2}$$

Proof. First we note that $Y_{t_i}=u(t_i,X_{t_i})$. For $\forall t\in[t_{i-1},t_i)$, by (4.1) and (2.19) we have

$$Z^{\pi,2}_t=E\left\{Y_{t_i}N^t_{t_i}+\int_t^{t_i}f(r,\tilde{\Theta}^{\pi,2}_r)N^t_r\,\mathrm{d}r|\mathscr{F}_t\right\}\sigma(t,X_t).\tag{4.3}$$

Denote $\tilde{Z}_t$ as the right side of (4.2). In light of estimate Lemma 4.1(ii), it suffices to prove that

$$\lim_{|\pi|\to 0}E\left\{\int_0^T|Z^{\pi,2}_t-\tilde{Z}_t|\sigma^{-1}(t,X_t)\,\mathrm{d}t\right\}=0.\tag{4.4}$$

To this end, we note that, for $\forall t\in[t_{i-1},t_i)$,

$$\begin{aligned}\tilde{Z}_t\sigma^{-1}(t,X_t)&=E\left\{g(X_T)N^t_T+\int_t^{t_i}f(r,\Theta_r)N^t_r\,\mathrm{d}r+\int_t^{t_i}N^t_r\,\mathrm{d}K_r\right.\\&\qquad\left.+\sum_{j=i}^{n-1}\left[\int_{t_j}^{t_{j+1}}f(r,\Theta_r)N^t_r\,\mathrm{d}r+\int_{t_j}^{t_{j+1}}N^t_r\,\mathrm{d}K_r\right]\Bigg|\mathscr{F}_t\right\}\\&=E\left\{\int_t^{t_i}f(r,\Theta_r)N^t_r\,\mathrm{d}r+\int_t^{t_i}N^t_r\,\mathrm{d}K_r+I_1(t)+I_2(t)|\mathscr{F}_t\right\},\end{aligned}\tag{4.5}$$

where

$$I_1(t) \triangleq \sum_{j=i}^{n-1}\left[\int_{t_j}^{t_{j+1}} f(r,\Theta_r)(N_r^t - N_{t_j}^t)\,\mathrm{d}r + \int_{t_j}^{t_{j+1}} (N_r^t - N_{t_j}^t)\,\mathrm{d}K_r\right],$$

$$I_2(t) \triangleq g(X_T)N_T^t + \sum_{j=i}^{n-1}\left[\int_{t_j}^{t_{j+1}} f(r,\Theta_r)\,\mathrm{d}r + (K_{t_{j+1}} - K_{t_j})\right]N_{t_j}^t.$$

Recalling (2.2), we see that $E\{I_2(t)|\mathscr{F}_t\} = E\{Y_T N_T^t + \sum_{j=i}^{n-1}(Y_{t_j} - Y_{t_{j+1}})N_{t_j}^t|\mathscr{F}_t\}$. Since $Y_{t_j} = u(t_j, X_{t_j})$, we can apply Lemma 2.4(iv) to get

$$E\{I_2(t)|\mathscr{F}_t\} = E\left\{Y_T N_{t_i}^t + \sum_{j=i}^{n-1}(Y_{t_j} - Y_{t_{j+1}})N_{t_i}^t|\mathscr{F}_t\right\} = E\{Y_{t_i}N_{t_i}^t|\mathscr{F}_t\}.$$

Therefore, (4.3) and (4.5) lead to

$$(\tilde{Z}_t - Z_t^{\pi,2})\sigma^{-1}(t,X_t) = E\left\{\int_t^{t_i} N_r^t\,\mathrm{d}K_r + I_1(t) + \int_t^{t_i}[f(r,\Theta_r) - f(r,\tilde{\Theta}_r^{\pi,2})]N_r^t\,\mathrm{d}r|\mathscr{F}_t\right\}.$$

By Lemmas 2.2, 2.4, 4.1, and Theorem 3.5 one can easily get that

$$E\{|\tilde{Z}_t - Z_t^{\pi,2}|\sigma^{-1}(t,X_t)\} \leqslant C(1+|x|^2)\sqrt{|\pi|} + E\{|I_1(t)|\}. \tag{4.6}$$

It remains to estimate $I_1(t)$. By Lemmas 2.2, 2.4, and Theorem 3.5 again we have

$$\begin{aligned}
E\int_0^T |I_1(t)|\,\mathrm{d}t &\leqslant E\left\{\sum_{i\leqslant j}\int_{t_{i-1}}^{t_i}\int_{t_j}^{t_{j+1}}[|f(r,\Theta_r)| + |k_r|]|N_r^t - N_{t_j}^t|\,\mathrm{d}r\,\mathrm{d}t\right\}\\
&\leqslant C(1+|x|^2)\sum_{i\leqslant j}\int_{t_{i-1}}^{t_i}\int_{t_j}^{t_{j+1}}\sqrt{\frac{r-t_j}{(t_j-t)(r-t)}}\,\mathrm{d}r\,\mathrm{d}t\\
&\leqslant C(1+|x|^2)\sum_{j=1}^{n-1}\int_{t_j}^{t_{j+1}}\sqrt{r-t_j}\,\mathrm{d}r\int_0^{t_j}\frac{\mathrm{d}t}{\sqrt{(t_j-t)(r-t)}}\\
&\leqslant C(1+|x|^2)\sum_{j=1}^{n-1}\int_{t_j}^{t_{j+1}}\sqrt{r-t_j}\int_0^{\frac{t_j}{r-t_j}}\frac{\mathrm{d}t'}{\sqrt{t'(1+t')}}\,\mathrm{d}r.
\end{aligned} \tag{4.7}$$

Here the last inequality is due to a change of variable $t' \triangleq (t_j - t)/(r - t_j)$. Note that $\int_0^{t_j/(r-t_j)}\{t'(1+t')\}^{-1/2}\,\mathrm{d}t' \leqslant C(1 + |\log(r-t_j)|)$, assuming $|\pi| \leqslant \frac{1}{2}$ we have

$$\begin{aligned}
E\int_0^T |I_1(t)|\,\mathrm{d}t &\leqslant C(1+|x|^2)\sum_{j=1}^{n-1}\int_{t_j}^{t_{j+1}}\sqrt{r-t_j}(1+|\log(r-t_j)|)\,\mathrm{d}r\\
&\leqslant C(1+|x|^2)\sqrt{|\pi|}\,\log\frac{1}{|\pi|}.
\end{aligned}$$

This, together with (4.6), proves (4.4), whence the theorem. $\square$

Corollary 4.3. *Assume* (A1)–(A3). *Then for any* $0 \leqslant s < t \leqslant T$,

$$|E\{Z_t\sigma^{-1}(t,X_t)\} - E\{Z_s\sigma^{-1}(s,X_s)\}| \leqslant C(1+|x|^2)\sqrt{t-s}(1+|\log(t-s)|).$$

Proof. By Theorem 4.2, we have

$$\begin{aligned}
&|E\{Z_t\sigma^{-1}(t,X_t)\} - E\{Z_s\sigma^{-1}(s,X_s)\}| \\
&\quad \leqslant |E\{g(X_T)N_T^t\} - E\{g(X_T)N_T^s\}| + E\left\{\left|\int_t^T f(r,\Theta_r)N_r^t\,\mathrm{d}r\right.\right. \\
&\qquad \left.\left. - \int_s^T f(r,\Theta_r)N_r^s\,\mathrm{d}r\right|\right\} + E\left\{\left|\int_t^T N_r^t\,\mathrm{d}K_r - \int_s^T N_r^s\,\mathrm{d}K_r\right|\right\} \\
&\quad = I_1 + I_2 + I_3,
\end{aligned} \tag{4.8}$$

where $I_i, i = 1,2,3$ are defined in an obvious way. We shall estimate I_i's separately.

First, if $g \in C^1$, then by Cauchy–Schwartz inequality and (2.14) we have

$$I_1 = |E\{g'(X_T)\nabla X_T([\nabla X_t]^{-1} - [\nabla X_s]^{-1})\}| \leqslant C\sqrt{t-s}. \tag{4.9}$$

For Lipschitz continuous g, by standard approximation, we see that (4.9) still holds.

Next, recalling Lemma 2.4 and Theorem 3.5, we have

$$\begin{aligned}
I_2 &\leqslant E\left\{\int_t^T |f(r,\Theta_r)||N_r^t - N_r^s|\,\mathrm{d}r + \int_s^t |f(r,\Theta_r)||N_r^s|\,\mathrm{d}r\right\} \\
&\leqslant C(1+|x|)\left[\int_t^T \sqrt{\frac{t-s}{(r-t)(r-s)}}\,\mathrm{d}r + \int_s^t \frac{\mathrm{d}r}{\sqrt{r-s}}\right] \\
&\leqslant C(1+|x|)\left[\sqrt{t-s}\int_0^{(T-t)/(t-s)} \frac{\mathrm{d}r}{\sqrt{r(r+1)}} + \sqrt{t-s}\right] \\
&\leqslant C(1+|x|)\sqrt{t-s}(1+|\log(t-s)|).
\end{aligned} \tag{4.10}$$

Analogous to (4.10), one can prove that $I_3 \leqslant C(1+|x|^2)\sqrt{t-s}(1+|\log(t-s)|)$, which, combined with (4.8)–(4.10), proves the corollary. □

5. Regularity of the obstacle problem

In this section we study the regularity of the solution to the obstacle problem (2.3). Comparing representations (2.16), (2.19), and (4.2) let us define:

$$v(t,x) \triangleq E\left\{g(X_T^{t,x})N_T^t(t,x) + \int_t^T f(r,\Theta_r^{t,x})N_r^t(t,x)\,\mathrm{d}r + \int_t^T N_r^t(t,x)\,\mathrm{d}K_r^{t,x}\right\}, \tag{5.1}$$

where $N_r^t(t,x)$ is the N_r^t corresponding to $X_\cdot^{t,x}$. Then, by Markovian properties we see that $Z_s^{t,x} = v(s,X_s^{t,x})\sigma(s,X_s^{t,x})$, $s \in [t,T]$. We have the following properties of v.

Theorem 5.1. *Assume* (A1)–(A3). *Then on* $[0,T)\times\mathbb{R}^n$, *it holds that*

(i) $v(\cdot,\cdot)$ *is bounded*,
(ii) $v(\cdot,\cdot)$ *is continuous*.
(iii) $v(\cdot,\cdot)=\partial_x u(\cdot,\cdot)$, *where u is the viscosity solution to the obstacle problem* (2.3).

Moreover, if we assume further that g is differentiable, then (i)–(iii) *hold true on* $[0,T]$ *by defining* $v(T,x)\triangleq g'(x)$.

Proof. (i) We first prove that $|v(0,x)|\leqslant C$ for some generic constant C. To this end we define $u^{\pi,2}(0,x)\triangleq Y_0^{\pi,2}$ and $v^{\pi,2}(0,x)\triangleq Z_0^{\pi,2}\sigma^{-1}(0,x)$. Since $Y_{t_1}=u(t_1,X_{t_1})$ and $u(t_1,x)$ is Lipschitz on x, applying (2.17) and (2.19) we have

$$v^{\pi,2}(0,x)=E\left\{Y_{t_1}N^0_{t_1}+\int_0^{t_1}f(r,\tilde{\Theta}^{\pi,2}_r)N^0_r\,\mathrm{d}r\right\}.$$

On the other hand, by restricting BSDER (2.2) on $[0,t_1]$ with terminal value $Y_{t_1}=u(t_1,X_{t_1})$, and applying Theorem 4.2 on it, we have

$$v(0,x)=E\left\{Y_{t_1}N^0_{t_1}+\int_0^{t_1}f(r,\Theta_r)N^0_r\,\mathrm{d}r+\int_0^{t_1}N^0_r\,\mathrm{d}K_r\right\}.$$

Then by Lemmas 2.4, 4.1, and Theorem 3.5 one can easily get

$$\begin{aligned}|v^{\pi,2}(0,x)-v(0,x)|&\\ \leqslant E\left\{\int_0^{t_1}|f(r,\tilde{\Theta}^{\pi,2}_r)-f(r,\Theta_r)||N^0_r|\,\mathrm{d}r+\int_0^{t_1}|N^0_r|k_r\,\mathrm{d}r\right\}&\\ \leqslant C(1+|x|^2)\sqrt{|\pi|},&\end{aligned}$$

which implies that

$$\lim_{|\pi|\to 0}v^{\pi,2}(0,x)=v(0,x). \tag{5.2}$$

Furthermore, since $v^{\pi,2}$ has the same property as the function v defined in (2.15), we conclude that $|v^{\pi,2}(0,x)|\leqslant C$, for some universal constant C that is independent of π. Consequently, $|v(0,x)|\leqslant C$, for all x, as well.

For general $t\in(0,T)$ we can assume without loss of generality that all partitions will contain t as a partition point. Therefore, a line by line analogy of the above arguments would lead to that $|v(t,x)|\leqslant C$, for all $(t,x)\in[0,T)\times\mathbb{R}^n$.

The proof for (ii) is a little lengthy, so we differ it to the end and prove (iii) instead. But again, we need only prove $u_x(0,x)=v(0,x)$ since the same argument works for any t. But by (4.1), (2.17) and the definition of $u^{\pi,2}$ we see immediately that $u^{\pi,2}_x(0,x)=v^{\pi,2}(0,x)$. Thus for any x_1,x_2, we have

$$u^{\pi,2}(0,x_1)-u^{\pi,2}(0,x_2)=\int_{x_1}^{x_2}v^{\pi,2}(0,x)\,\mathrm{d}x. \tag{5.3}$$

Letting $|\pi| \to 0$ in (5.3), and applying (5.2), Lemma 4.1, and Dominated Convergence Theorem we get

$$u(0,x_1) - u(0,x_2) = \int_{x_1}^{x_2} v(0,x)\,\mathrm{d}x.$$

Since v is continuous by (ii), we obtain that $u_x(0,x) = v(0,x)$.

It remains to prove (ii). We first show that v is continuous in x. Again it suffices to show that $v(0,\cdot)$ is continuous. Thus let $x_n \to x$ and let (X^n, Y^n, Z^n, K^n) denote the solution to (2.2) with initial value x_n. By El Karoui [6] (Proposition 3.6) we have

$$E\left\{\sup_{0\leqslant t\leqslant T}[|\Delta X_t^n|^2 + |\Delta Y_t^n|^2 + |\Delta K_t^n|^2] + \int_0^T |\Delta Z_t^n|^2\,\mathrm{d}t\right\} \leqslant C(1+|x|)\sqrt{|x_n - x|}, \tag{5.4}$$

where $\Delta\Xi_t^n \triangleq \Xi_t^n - \Xi_t$, $\Xi = X, Y, Z, K$. Denote $N_r \triangleq N_r^0$ and $N_r^n \triangleq N_r^{n,0}$. Then

$$\begin{aligned}|v(0,x_n) - v(0,x)| \leqslant E\bigg\{ & |g(X_T^n)N_T^n - g(X_T)N_T| \\ & + \int_0^T |f(r,\Theta_r^n)N_r^n - f(r,\Theta_r)N_r]|\,\mathrm{d}r \\ & + \left|\int_0^T N_r^n\,\mathrm{d}K_r^n - \int_0^T N_r\,\mathrm{d}K_r\right|\bigg\}\end{aligned} \tag{5.5}$$

By (5.4) and standard arguments one can prove that

$$\lim_{n\to 0} E\left\{|g(X_T^n)N_T^n - g(X_T)N_T| + \int_0^T |f(r,\Theta_r^n)N_r^n - f(r,\Theta_r)N_r|\,\mathrm{d}r\right\} = 0. \tag{5.6}$$

To estimate $I_n \triangleq E\{|\int_0^T N_r^n\,\mathrm{d}K_r^n - \int_0^T N_r\,\mathrm{d}K_r|\}$, let π_m be a partition of $[0,T]$ with $|\pi_m| = T/m$. Denoting $\delta\xi_{t,t_{i-1}} = \xi_t - \xi_{t_{i-1}}$, $t \geqslant t_{i-1}$, $\xi = K, K^n, N$, we have

$$\begin{aligned} I_n \leqslant E\Bigg\{ & \left|\int_0^T \Delta N_r^n\,\mathrm{d}K_r^n\right| + \left|\int_0^T N_r\,\mathrm{d}K_r^n - \sum_{i=2}^m N_{t_{i-1}}\delta K_{t_i,t_{i-1}}^n\right| \\ & + \left|\sum_{i=2}^m N_{t_{i-1}}\,\Delta[\delta K_{t_i,t_{i-1}}^n]\right| + \left|\sum_{i=2}^m N_{t_{i-1}}\delta K_{t_i,t_{i-1}} - \int_0^T N_r\,\mathrm{d}K_r\right|\Bigg\} \\ \leqslant E\Bigg\{ & \int_0^T |\Delta N_r^n| k_r^n\,\mathrm{d}r + \sum_{i=2}^m \int_{t_{i-1}}^{t_i} |\delta N_{r,t_{i-1}}|[k_r^n + k_r]\,\mathrm{d}r \\ & + \int_0^{t_1} |N_r|[k_r^n + k_r]\,\mathrm{d}r + \left|\sum_{i=2}^m N_{t_{i-1}}\,\Delta[\delta K_{t_i,t_{i-1}}^n]\right|\Bigg\},\end{aligned}$$

where ΔN^n and $\Delta[\delta K^n]$ are defined as usual. Recalling Lemmas 2.2 and 2.4 we have

$$
\begin{aligned}
I_n &\leqslant C(1+|x|^2)\Bigg[\int_0^T \sqrt{E|\Delta N_r^n|^2}\,\mathrm{d}r + \sum_{i=2}^m \int_{t_{i-1}}^{t_i} \frac{\sqrt{r-t_{i-1}}}{t_{i-1}}\,\mathrm{d}r + \sqrt{t_1} \\
&\quad + E\left|\sum_{i=2}^m N_{t_{i-1}}\,\Delta[\delta K^n_{t_i,t_{i-1}}]\right|\Bigg] \\
&\leqslant C(1+|x|^2)\left[\int_0^T \sqrt{E|\Delta N_r^n|^2}\,\mathrm{d}r + \frac{\log m}{\sqrt{m}} + E\left|\sum_{i=2}^m N_{t_{i-1}}\,\Delta[\delta K^n_{t_i,t_{i-1}}]\right|\right].
\end{aligned}
$$

Thus, by virtue of (5.4), first letting $n \to \infty$ and then letting $m \to \infty$ we obtain $\lim_{n\to\infty} I_n = 0$. This, together with (5.5) and (5.6), implies that $\lim_{n\to 0} v(0,x_n) = v(0,x)$.

Now let us assume $(t_n,x_n) \to (t,x)$. For any $t^* > t$, we have $t_n < t^*$ for n large enough. Since $v(r,X_r) = Z_r\sigma^{-1}(r,X_r)$, applying Corollary 4.3, we have

$$
\begin{aligned}
|v(t_n,x_n) - v(t,x)| &\leqslant |E\{[Z_{t_n}^{t_n,x_n}\sigma^{-1}(t_n,x_n) - Z_{t^*}^{t_n,x_n}\sigma^{-1}(t^*,X_{t^*}^{t_n,x_n})] + [v(t^*,X_{t^*}^{t_n,x_n}) \\
&\quad - v(t^*,X_{t^*}^{t,x})] + [Z_{t^*}^{t,x}\sigma^{-1}(t^*,X_{t^*}^{t,x}) - Z_t^{t,x}\sigma^{-1}(t,x)]\}| \\
&\leqslant C(1+|x|^2)[\sqrt{t^*-t_n}(1+|\log(t^*-t_n)|) \\
&\quad + \sqrt{t^*-t}(1+|\log(t^*-t)|)] \\
&\quad + E\{|v(t^*,X_{t^*}^{t_n,x_n}) - v(t^*,X_{t^*}^{t,x})|\}.
\end{aligned}
$$

Note that $v(t^*,\cdot)$ is continuous and v is bounded, first sending $n\to\infty$ and then sending $t^* \downarrow t$ in the above we obtain that $\lim_{n\to\infty} v(t_n,x_n) = v(t,x)$, proving (ii).

Finally, if $g \in C^1$, by (2.14) we have

$$
v(t,x) = E\left\{g'(X_T^{t,x})\nabla X_T^{t,x} + \int_t^T f(r,\Theta_r^{t,x})N_r^t(t,x)\,\mathrm{d}r + \int_t^T N_r^t(t,x)\,\mathrm{d}K_r^{t,x}\right\}.
$$

It is then clear that $\lim_{t\uparrow T} v(t,x) = g'(x)$. Thus (i)–(iii) also hold true at $t = T$. $\square$

The following result is a direct consequence of Theorem 5.1.

Corollary 5.2. *Assume* (A1)–(A3). *Then Z is continuous on* $[0,T]$ *a.s., and*

$$
E\left\{\sup_{0\leqslant t\leqslant T} |Z_t|^p\right\} \leqslant C_p(1+|x|^p).
$$

Proof. By Theorem 5.1, it remains only to show that Z is continuous at $t = T$. Namely, we need to show that $\lim_{t\uparrow T} Z_t$ exists a.s. Noting that g is Lipschitz continuous, one can show that (see e.g. [17]) there exists $\xi \in \mathscr{F}_T$ such that $|\xi| \leqslant C$ and

$$
E\{g(X_T)N_T^t|\mathscr{F}_t\} = E\{\xi\nabla X_T[\nabla X_t]^{-1}|\mathscr{F}_t\}.
$$

Now by Theorem 4.2 we have

$$Z_t = E\left\{\xi \nabla X_T [\nabla X_t]^{-1} + \int_t^T f(r, \Theta_r) N_r^t \,\mathrm{d}r + \int_t^T N_r^t \,\mathrm{d}K_r | \mathscr{F}_t\right\} \sigma(t, X_t).$$

Then one can easily prove that $Z_{T-} = \xi\sigma(T, X_T)$. The proof is complete. $\square$

6. L^2-Modulus regularity

In this section we study a new property of the process Z, which we shall name as the "L^2-modulus regularity" in the sequel. Recall that for $\varphi \in C([0,T])$, the modulus of continuity of φ with "accuracy" δ is defined by $w_\delta(\varphi) \triangleq \sup_{0<|t-s|\leqslant\delta} |\varphi(t) - \varphi(s)|$. Moreover, if $\pi : 0 = t_0 < t_1 < \cdots < t_n = T$ is any partition of $[0,T]$, and let

$$w^\pi(\varphi) \triangleq \max_{1\leqslant i\leqslant n} \sup_{s\in[t_{i-1},t_i]} \sup_{t\in[t_{i-1},t_i]} |\varphi(t) - \varphi(s)|. \tag{6.1}$$

Then it is easily seen that $w_\delta(\varphi) = \sup_{|\pi|\leqslant\delta} w^\pi(\varphi)$. Heuristically, the "$L^2$-modulus of continuity" is a slight modification of w^π defined in (6.1). To be more precise, let $L^2_{\mathrm{loc}}(\Omega; C[0,T])$ be the space of all *continuous* process $\eta \in L^2(\Omega \times [0,T])$.

Definition 6.1. For any process $\eta \in L^2_{\mathrm{loc}}(\Omega; C[0,T])$, and any partition $\pi : 0 = t_0 < \cdots < t_n = T$, define the L^2-modulus of continuity of η with respect to π by

$$|||\eta|||_\pi^2 \triangleq \sum_{i=1}^n \sup_{s\in[t_{i-1},t_i]} E\left\{\int_{t_{i-1}}^{t_i} |\eta_t - \eta_s|^2 \,\mathrm{d}t\right\}.$$

We shall denote $\mathscr{L}^2(\pi) \triangleq \{\eta \in L^2_{\mathrm{loc}}(\Omega; C[0,T]) : |||\eta|||_\pi < \infty\}$.

The following theorem gives the first characterization of the space $\mathscr{L}^2(\pi)$.

Theorem 6.2. *For any partition π it holds that*

$$|||\eta|||_\pi^2 \leqslant 4T \sup_{0\leqslant t\leqslant T} E\{|\eta_t|^2\} \leqslant 4TE\left\{\sup_{0\leqslant t\leqslant T} |\eta_t|^2\right\} = 4T\|\eta\|^2_{L^2(\Omega;C[0,T])}. \tag{6.2}$$

Consequently, it holds that $L^2(\Omega; C[0,T]) \subseteq \mathscr{L}^2(\pi) \subseteq L^2_{\mathrm{loc}}(\Omega; C[0,T])$. Furthermore, if $\eta \in L^2(\Omega; C[0,T])$, then $\lim_{|\pi|\to 0} |||\eta|||_\pi = 0$.

Proof. Let $\eta \in L^2(\Omega; C[0,T])$ and $\pi : 0 = t_0 < t_1 < \cdots < t_n = T$. For any i and $s \in [t_{i-1}, t_i]$, we have $E\{|\eta_t - \eta_s|^2\} \leqslant 4\sup_{0\leqslant t\leqslant T} E\{|\eta_t|^2\}$. Thus by definition of the L^2-modulus,

$$\begin{aligned} |||\eta|||_\pi^2 &\leqslant \int_0^T 4 \sup_{0\leqslant t\leqslant T} E\{|\eta_t|^2\} \,\mathrm{d}t = 4T \sup_{0\leqslant t\leqslant T} E\{|\eta_t|^2\} \\ &\leqslant 4TE\left\{\sup_{0\leqslant t\leqslant T} |\eta_t|^2\right\} < \infty. \end{aligned}$$

That is, (6.2) holds, and $\eta \in \mathscr{L}^2(\pi)$. Thus $L^2(\Omega; C[0,T]) \subset \mathscr{L}^2(\pi)$. Note that $\mathscr{L}^2(\pi) \subset L^2_{\mathrm{loc}}(\Omega; C[0,T])$ is obvious by definition, the sequence of the inclusions follows.

Finally, if $\eta \in L^2(\Omega; C[0,T])$, then $\lim_{|\pi|\to 0} |\eta_t(\omega) - \eta_{s_i}(\omega)| \leqslant \lim_{|\pi|\to 0} w_{|\pi|}(\eta_\cdot(\omega)) = 0$, whenever $\eta_\cdot(\omega)$ is continuous. An easy application of the Dominated Convergence Theorem then leads to that $\lim_{|\pi|\to 0} |||\,\eta\,|||_\pi = 0$, proving the theorem. □

We remark that both inclusions in Theorem 6.2 are actually strict. In order not to disturb our discussion, however, we provide two examples in the appendix for interested readers.

We now establish the L^2-modulus of the process Z. To begin with, we add the following strengthened assumptions on the coefficients.

(A4) b, σ, and f are uniformly Hölder-$\frac{1}{2}$ continuous in the variable t.

Theorem 6.3. *Assume* (A1)–(A4). *Assume also that $L\,\Delta t_i \geqslant |\pi|$ for all i, where $L>0$ is the generic constant in* (2.1). *Then there exists a constant $C>0$, independent of the partition π, such that* $|||\,Z\,|||_\pi^2 \leqslant C(1+|x|^4)\sqrt{|\pi|}$.

Proof. First, we denote $Z_T \triangleq \xi\sigma(T, X_T)$, where ξ is as in Corollary 5.2, and $M_t \triangleq E\{\sup_{0\leqslant s\leqslant T} |Z_s| |\mathscr{F}_t\}$. Also, for any i we denote $m^i_t \triangleq E\{Z_{t_i}|\mathscr{F}_t\}$ for $t\in[0,t_i]$. Then for $\forall t,s \in [t_{i-1}, t_i]$ we have

$$|Z_t - Z_s|^2 \leqslant 3\{|Z_t - m^i_t|^2 + |m^i_t - m^i_s|^2 + |m^i_s - Z_s|^2\}, \tag{6.3}$$

It is clear that

$$|Z_t - m^i_t|^2 \leqslant 2M_t|Z_t - m^i_t|; \quad |Z_s - m^i_s|^2 \leqslant 2M_s|Z_s - m^i_s|. \tag{6.4}$$

Further, note that m^i_s is an L^2-martingale for $0\leqslant s\leqslant t_i$, applying the martingale representation theorem we can write $m^i_s = a + \int_0^s \eta^i_r \,\mathrm{d}W_r,\ \ s\in[0,t_i]$ for some predictable process η^i. Thus, for $s,t\in[t_{i-1},t_i]$ one has

$$\begin{aligned} E\{|m^i_t - m^i_s|^2\} &= E\left|\int_s^t \eta^i_r\,\mathrm{d}W_r\right|^2 \leqslant E\int_{t_{i-1}}^{t_i} |\eta^i_r|^2\,\mathrm{d}r = E\{|Z_{t_i}|^2\} - E\{|m^i_{t_{i-1}}|^2\} \\ &= E\{|Z_{t_i}|^2 - |Z_{t_{i-1}}|^2\} + E\{|Z_{t_{i-1}}|^2 - |m^i_{t_{i-1}}|^2\} \\ &\leqslant E\{|Z_{t_i}|^2 - |Z_{t_{i-1}}|^2\} + 2E\{M_{t_{i-1}}|Z_{t_{i-1}} - m^i_{t_{i-1}}|\}, \end{aligned}$$

which, combined with (6.3) and (6.4), implies that

$$E\{|Z_t - Z_s|^2\} \leqslant 3E\{|Z_{t_i}|^2 - |Z_{t_{i-1}}|^2\} + C \sup_{t_{i-1}\leqslant r\leqslant t_i} E\{M_r|Z_r - m^i_r|\}.$$

Note that the right side of the above is independent of t and s. Therefore,

$$\begin{aligned} &\sup_{t_{i-1}\leqslant s\leqslant t_i} E\left\{\int_{t_{i-1}}^{t_i} |Z_t - Z_s|^2\,\mathrm{d}t\right\} \\ &\quad\leqslant |\pi| \sup_{t_{i-1}\leqslant t,\,s\leqslant t_i} E\{|Z_t - Z_s|^2\} \\ &\quad\leqslant 3|\pi| E\{|Z_{t_i}|^2 - |Z_{t_{i-1}}|^2\} + C|\pi| \sup_{t_{i-1}\leqslant t\leqslant t_i} E\{M_t|Z_t - m^i_t|\}. \end{aligned}$$

Recalling Definition 6.1 and then applying Corollary 5.2 we get that

$$
\begin{aligned}
|||\, Z \,|||_{\pi}^{2} &\leqslant 3|\pi| E\{|Z_T|^2 - |Z_0|^2\} + C|\pi| \sum_{i=1}^{n} \sup_{t_{i-1}\leqslant t\leqslant t_i} E\{M_t |Z_t - m_t^i|\} \\
&\leqslant C(1+|x|^2)|\pi| + C|\pi| \sum_{i=1}^{n} \sup_{t_{i-1}\leqslant t\leqslant t_i} E\{M_t |Z_t - m_t^i|\}.
\end{aligned} \tag{6.5}
$$

Now let us estimate $E\{M_t|Z_t - m_t^i|\}$. First note that for $i = 1, \dots, n$, the process $(\Theta_t, K_t)_{0\leqslant t\leqslant t_i}$ can also be considered as the solution to the BSDER (2.2) over $[0, t_i]$ with terminal value $Y_{t_i} = u(t_i, X_{t_i})$. Recalling Theorem 5.1(iii), (2.14) with $i<n$, and the proof of Corollary 5.2 with $i = n$, we derive from Theorem 4.2 that

$$
Z_t = E\left\{ Z_{t_i}\sigma_X^{-1}(t_i)\nabla X_{t_i}[\nabla X_t]^{-1} + \int_t^{t_i} f(r,\Theta_r) N_r^t \,\mathrm{d}r + \int_t^{t_i} N_r^t \,\mathrm{d}K_r \big| \mathscr{F}_t \right\} \sigma_X(t), \tag{6.6}
$$

for all $t \in [0, t_i]$. Here and in what follows $\sigma_X(t) \triangleq \sigma(t, X_t)$, $\forall t$, for simplicity. In particular, for $t \in [t_{i-1}, t_i]$, one can easily prove that

$$
\begin{aligned}
|Z_t - m_t^i| \leqslant E\{&|Z_{t_i}|\sigma_X^{-1}(t_i)\nabla X_{t_i}|[\nabla X_t]^{-1}\sigma_X(t) - [\nabla X_{t_i}]^{-1}\sigma_X(t_i)| \\
&+ \sigma_X(t)\int_t^{t_i} [|f(r,\Theta_r)| + k_r]|N_r^t| \,\mathrm{d}r | \mathscr{F}_t\}.
\end{aligned}
$$

Now by Hölder's inequality we have

$$
\begin{aligned}
&E\{M_t|Z_t - m_t^i|\} \\
&\quad\leqslant \sqrt{E\{M_t^2|Z_{t_i}|^2\sigma_X^{-2}(t_i)|\nabla X_{t_i}|^2\}}\sqrt{E\{|[\nabla X_t]^{-1}\sigma_X(t) - [\nabla X_{t_i}]^{-1}\sigma_X(t_i)|^2\}} \\
&\qquad + \int_t^{t_i} \sqrt{2E\{M_t^2\sigma_X^2(t)[|f(r,\Theta_r)|^2 + |k_r|^2]\}}\sqrt{E\{|N_r^t|^2\}}\,\mathrm{d}r \\
&\quad\leqslant C(1+|x|^2)\sqrt{E\{|[\nabla X_t]^{-1}\sigma_X(t) - [\nabla X_{t_i}]^{-1}\sigma_X(t_i)|^2\}} \\
&\qquad + C(1+|x|^4)\int_t^{t_i} \sqrt{E\{|N_r^t|^2\}}\,\mathrm{d}r,
\end{aligned} \tag{6.7}
$$

thanks to Corollary 5.2. Finally, by (A4) and Lemma 2.4, we have

$$
E\{|[\nabla X_t]^{-1}\sigma_X(t) - [\nabla X_{t_i}]^{-1}\sigma_X(t_i)|^2\} \leqslant C(1+|x|^2)|\pi| \quad \text{and} \quad \int_t^{t_i} \sqrt{E\{|N_r^t|^2\}}\,\mathrm{d}r \leqslant C\sqrt{|\pi|}.
$$

Thus (6.5), (6.7), and the assumption $L\,\Delta t_i \geqslant |\pi|$ lead to that

$$
|||\, Z \,|||_{\pi}^{2} \leqslant C(1+|x|^2)|\pi| + C|\pi| \sum_{i=1}^{n} (1+|x|^4)\sqrt{|\pi|} \leqslant C(1+|x|^4)\sqrt{|\pi|},
$$

proving the theorem. $\square$

The following L^2 estimate of the modulus of continuity of process Y is new.

Corollary 6.4. *Assume that all the conditions in Theorem* 6.3 *hold true. Then*

$$E\{\omega^{\pi}(Y)|^2\}\leqslant C(1+|x|^4)\sqrt{|\pi|}.$$

Proof. First note that we need only prove the following inequality:

$$E\left\{\max_{1\leqslant i\leqslant n}\sup_{t\in[t_{i-1},t_i)}|Y_t-Y_{t_{i-1}}|^2\right\}\leqslant C(1+|x|^4)\sqrt{|\pi|}. \tag{6.8}$$

To do this, again let us denote, for any process ξ and each i, $\delta\xi_{t,t_{i-1}}\triangleq\xi_t-\xi_{t_{i-1}}$, $t\geqslant t_{i-1}$. Now from (2.2) we can easily see that

$$\delta Y_{t,t_{i-1}}=-\int_{t_{i-1}}^{t}f(r,\Theta_r)\,\mathrm{d}r+Z_{t_{i-1}}\delta W_{t,,t_{i-1}}+\int_{t_{i-1}}^{t}\delta Z_{r,t_{i-1}}\,\mathrm{d}W_r-\delta K_{t,t_{i-1}}. \tag{6.9}$$

Thus, denoting $|\delta\xi|^{*,p}_{t_i,t_{i-1}}\triangleq\sup_{t\in[t_{i-1},t_i]}|\delta\xi_{t,t_{i-1}}|^p$ for any $p>0$ and any continuous process ξ as usual, we have

$$|\delta Y|^{*}_{t_i,t_{i-1}}\leqslant\int_{t_{i-1}}^{t_i}[|f(r,\Theta_r)|+k_r]\,\mathrm{d}r+|Z_{t_{i-1}}||\delta W|^{*}_{t_i,t_{i-1}}+\sup_{t_{i-1}\leqslant t<t_i}\left|\int_{t_{i-1}}^{t}\delta Z_{r,t_{i-1}}\,\mathrm{d}W_r\right|.$$

Now recall Lemma 2.2, Corollary 5.2, and Theorem 6.3, applying the Burkholder Inequality, and using the facts that $E\{\max_{1\leqslant i\leqslant n}|\delta W|^{*,4}_{t_i,t_{i-1}}\}\leqslant C|\pi|^2[1+(\log|\pi|)^2]$ (see, e.g. [18]), and that $\max_{1\leqslant i\leqslant n}a_i\leqslant\sum_{i=1}^{n}a_i$ for $a_i\geqslant 0$, we can easily obtain that

$$\begin{aligned}
&E\left\{\max_{1\leqslant i\leqslant n}|\delta Y|^{*,2}_{t_i,t_{i-1}}\right\}\\
&\quad\leqslant CE\left\{\sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i}[|f(r,\Theta_r)|+k_r]\,\mathrm{d}r\right)^2\right\}+\max_{0\leqslant i\leqslant n}\\
&\qquad\times|Z_{t_i}|^2\max_{1\leqslant i\leqslant n}|\delta W|^{*,2}_{t_i,t_{i-1}}+\sum_{i=1}^{n}\sup_{t_{i-1}\leqslant t<t_i}\left|\int_{t_{i-1}}^{t}\delta Z_{r,t_{i-1}}\,\mathrm{d}W_r\right|^2\\
&\quad\leqslant CE\left\{\sum_{i=1}^{n}\Delta t_i\int_{t_{i-1}}^{t_i}[1+|X_r|^4+|Y_r|^2+|Z_r|^2]\,\mathrm{d}r\right\}\\
&\qquad+CE\left\{\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}|\delta Z_{r,t_{i-1}}|^2\,\mathrm{d}r\right\}+C\sqrt{E\left\{\sup_{0\leqslant t\leqslant T}|Z_t|^4\right\}}\sqrt{E\left\{\max_{1\leqslant i\leqslant n}|\delta W|^{*,4}_{t_i,t_{i-1}}\right\}}\\
&\quad\leqslant C(1+|x|^4)\Big[|\pi|+\sqrt{|\pi|}+|\pi||\log|\pi||\Big]\leqslant C(1+|x|^4)\sqrt{|\pi|}.\qquad\square
\end{aligned}$$

7. A numerical scheme for BSDERs

In this section we introduce the last discretization of BSDER (2.2). The main feature of this discretization is that it consists of piecewise constant processes for all the components, and are all computable in theory. Therefore it can be considered as a numerical scheme for (2.2).

The main idea is the following. Let $\pi : 0 = t_0 < \cdots < t_n = T$ be any partition of $[0, T]$, and let $X^{\pi,0}$ to be the Euler approximation of the process X (see (2.6)) and X^π to be the corresponding piecewise constant process. We then define

- $Y^{\pi,3}_{t_n} \triangleq g(X^\pi_{t_n})$; $Z^{\pi,3}_{t_n} \triangleq 0$;
- for $i = n, n-1, \dots, 1$, and $t \in [t_{i-1}, t_i)$, let $(\tilde{Y}^{\pi,3}, Z^{\pi,3})$ be the solution of BSDE:
$$\tilde{Y}^{\pi,3}_t = Y^{\pi,3}_{t_i} + f(t_i, \Theta^{\pi,3}_{t_i})\,\Delta t_i - \int_t^{t_i} Z^{\pi,3}_r\,\mathrm{d}W_r, \tag{7.1}$$
where $\Theta^{\pi,3}_{t_i} \triangleq (X^\pi_{t_i}, Y^{\pi,3}_{t_i}, Z^{\pi,3}_{t_i})$;
- for each i, define $Y^{\pi,3}_{t_i} \triangleq \tilde{Y}^{\pi,3}_{t_i} \vee h(t_i, X^\pi_{t_i})$;
- let $K^{\pi,3}_0 = 0$, and for $t \in (t_{i-1}, t_i]$, define $K^{\pi,3}_t \equiv K^{\pi,3}_{t_i} \triangleq \sum_{j=1}^i (Y^{\pi,3}_{t_{j-1}} - \tilde{Y}^{\pi,3}_{t_{j-1}})$.

We should point out that the main difference between this scheme and the previous ones is that the BSDE (7.1) is linear, and $\Theta^{\pi,3}$ contains Y^π, rather than $\tilde{Y}^\pi$ as before.

Recall that X^π and $K^{\pi,3}$ are by definition piecewise constant processes. We denote $K^\pi \triangleq K^{\pi,3}$ and define two other piecewise constant processes (Y^π, Z^π) in an obvious way: $Y^\pi_t \triangleq Y^{\pi,3}_{t_{i-1}}$ and $Z^\pi_t \triangleq Z^{\pi,3}_{t_{i-1}}$, for $\forall t \in [t_{i-1}, t_i)$.

We now analyze the computability of $(X^\pi_{t_i}, Y^\pi_{t_i}, Z^\pi_{t_i}, K^\pi_{t_i})$'s. But clearly it would suffice if we can show that $(Y^\pi_{t_i}, Z^\pi_{t_i})$ can be written as functions of $X^\pi_{t_i}$. We argue by backward induction. First, we have $Y^\pi_{t_n} = Y^{\pi,3}_{t_n} = g(X^\pi_{t_n})$ and $Z^\pi_{t_n} = Z^{\pi,3}_{t_n} = 0$. Now assume $Y^\pi_{t_i} = u^\pi_i(X^\pi_{t_i})$ and $Z^\pi_{t_i} = v^\pi_i(X^\pi_{t_i})$ for some functions u^π_i and v^π_i. We then define

$$\begin{aligned} &a^\pi_i(x, y) \triangleq x + b(t_{i-1}, x)\,\Delta t_i + \sigma(t_{i-1}, x)y, \\ &g^\pi_i(x, y) \triangleq u^\pi_i(a^\pi_i(x, y)) + f(t_i, a^\pi_i(x, y), u^\pi_i(a^\pi_i(x, y)), v^\pi_i(a^\pi_i(x, y)))\,\Delta t_i. \end{aligned} \tag{7.2}$$

Then, recall (2.5), we see that (7.1) can be written as

$$\tilde{Y}^{\pi,3}_t = g^\pi_i(X^\pi_{t_{i-1}}, W_{t_i} - W_{t_{i-1}}) - \int_t^{t_i} Z^{\pi,3}_r\,\mathrm{d}W_r, \quad t \in [t_{i-1}, t_i), \tag{7.3}$$

Now applying results (2.17)–(2.19) to BSDE (7.3) (note that the forward diffusion in (7.3) is simply $x + W_t$, hence the corresponding $N^t_r = (W_r - W_t)/(r - t)$), we see that, with $\Delta_i W \triangleq W_{t_i} - W_{t_{i-1}}$,

$$\tilde{u}^\pi_{i-1}(x) = E\{g^\pi_i(x, \Delta_i W)\}; \quad v^\pi_{i-1}(x) = \frac{1}{\Delta t_i} E\{g^\pi_i(x, \Delta_i W)\Delta_i W\} \tag{7.4}$$

and $u^\pi_{i-1}(x) = \tilde{u}^\pi_{i-1}(x) \vee h(t_{i-1}, x)$, it holds that $\tilde{Y}^{\pi,3}_{t_{i-1}} = \tilde{u}^\pi_{i-1}(X^\pi_{t_{i-1}})$, $Y^\pi_{t_{i-1}} = u^\pi_{i-1}(X^\pi_{t_{i-1}})$, and $Z^\pi_{t_{i-1}} = v^\pi_{i-1}(X^\pi_{t_{i-1}})$. This completes the inductional step. Since both $\tilde{u}^\pi$ and v^π in (7.4) are computable in theory by Monte-Carlo simulations, we substantiated our claim.

We should note that the numerical feasibility of our scheme can also be seen (indeed more clearly) from some recent publications. For example, if we denote $\xi^\pi_i \triangleq Y^\pi_{t_i} + f(t_i, \Theta^\pi_{t_i})\,\Delta t_i$, then by (7.4) we see that $\tilde{Y}^{\pi,3}_{t_{i-1}} = E\{\xi^\pi_i | \mathscr{F}_{t_{i-1}}\}$ and $Z^\pi_{t_{i-1}} = (1/\Delta t_i) E\{\xi^\pi_i\,\Delta_i W | \mathscr{F}_{t_{i-1}}\}$. Thus the problem is reduced to computing the conditional expectations. This problem was recently studied by Bouchard–Touzi [2], in which a feasible numerical method is introduced.

In the rest of the section we shall use our first pseudo-discretization and the result in L^2-modulus of continuity to investigate rate of convergence of the approximation we just designed. We first give a simple but useful inequality. The proof is straightforward and left to the readers:

$$|Y^{\pi,1}_{t_i} - Y^{\pi,3}_{t_i}| \leqslant |\tilde{Y}^{\pi,1}_{t_i} - \tilde{Y}^{\pi,3}_{t_i}|, \quad \forall i. \tag{7.5}$$

Our main result is the following.

Theorem 7.1. *Assume all the conditions in Theorem* 6.3 *hold true*; *and that* $|\pi| \leqslant 1$. *Then there exists a constant* $C > 0$, *independent of the partition* π, *such that*

$$E\left\{\max_{0\leqslant i\leqslant n}[|\tilde{Y}^{\pi,1}_{t_i} - \tilde{Y}^{\pi,3}_{t_i}|^2 + |K^{\pi,1}_{t_i} - K^{\pi,3}_{t_i}|^2] + \int_0^T |Z^{\pi,1}_t - Z^{\pi,3}_t|^2\,\mathrm{d}t\right\} \leqslant C(1+|x|^4)\sqrt{|\pi|}, \tag{7.6}$$

$$E\left\{\sup_{0\leqslant t\leqslant T}[|Y_t - Y^{\pi}_t|^2 + |K_t - K^{\pi}_t|^2] + \int_0^T |Z_t - Z^{\pi}_t|^2\,\mathrm{d}t\right\} \leqslant C(1+|x|^4)\sqrt{|\pi|}. \tag{7.7}$$

Proof. We first assume (7.6) to prove (7.7). Recall the notations $\delta\xi_{t,t_{i-1}}$ and $|\delta\xi|^*_{t_i,t_{i-1}}$ defined before. We have

$$\begin{aligned}\sup_{0\leqslant t\leqslant T}|Y_t - Y^{\pi}_t| &= \max_{1\leqslant i\leqslant n}\sup_{t_{i-1}\leqslant t<t_i}|Y_t - Y^{\pi,3}_{t_{i-1}}|\\ &\leqslant \max_{1\leqslant i\leqslant n}|\delta Y|^*_{t_i,t_{i-1}} + \max_{0\leqslant i\leqslant n}[|Y_{t_i} - Y^{\pi,0}_{t_i}| + |Y^{\pi,0}_{t_i} - Y^{\pi,1}_{t_i}|\\ &\quad + |Y^{\pi,1}_{t_i} - Y^{\pi,3}_{t_i}|].\end{aligned}$$

Now applying Lemma 3.1 and using (7.5), we can rewrite above as

$$\begin{aligned}\sup_{0\leqslant t\leqslant T}|Y_t - Y^{\pi}_t| \leqslant{}& \max_{1\leqslant i\leqslant n}|\delta Y|^*_{t_i,t_{i-1}} + \max_{0\leqslant i\leqslant n}[|Y_{t_i} - Y^{\pi,0}_{t_i}| + |Y^{\pi,0}_{t_i} - \tilde{Y}^{\pi,1}_{t_i}|\\ &+ |\tilde{Y}^{\pi,1}_{t_i} - \tilde{Y}^{\pi,3}_{t_i}|].\end{aligned}$$

It then follows from Corollary 6.4, Lemma 2.2(ii), Theorems 3.3, and (7.6) that

$$E\left\{\sup_{0\leqslant t\leqslant T}|Y_t - Y^{\pi}_t|^2\right\} \leqslant C(1+|x|^4)\sqrt{|\pi|}.$$

To see the estimate for Z in (7.7), we recall (7.3). Note that if the function $g^{\pi}_i(x,y)$ is differentiable in y, then by (2.19) and (2.14) we see that $Z^{\pi,3}_t = E\{\partial_y g^{\pi}_i(X^{\pi}_{t_{i-1}}, \delta W_{t_i,t_{i-1}})|\mathscr{F}_t\}$ is a martingale. In general we can also show that $Z^{\pi,3}_t$ is a martingale for $t \in [t_{i-1}, t_i)$ by approximating g^{π}_i by smooth functions. Consequently, $Z^{\pi,3}_{t_{i-1}} = (1/\Delta t_i)E\{\int_{t_{i-1}}^{t_i} Z^{\pi,3}_t\,\mathrm{d}t|\mathscr{F}_{t_{i-1}}\}$. Furthermore, if we denote

$\widehat{Z}^{\pi}_{t_{i-1}} \triangleq (1/\Delta t_i)E\{\int_{t_{i-1}}^{t_i} Z_t \,\mathrm{d}t|\mathscr{F}_{t_{i-1}}\}$, then one has

$$E\{|Z^{\pi,3}_{t_{i-1}} - \widehat{Z}^{\pi}_{t_{i-1}}|^2\} \leqslant \frac{1}{\Delta t_i} E\left\{\int_{t_{i-1}}^{t_i} |Z^{\pi,3}_t - Z_t|^2 \,\mathrm{d}t\right\}. \tag{7.8}$$

Moreover, noting that the mean minimizes the square error, we have

$$E\left\{\int_{t_{i-1}}^{t_i} |Z_t - \widehat{Z}^{\pi}_{t_{i-1}}|^2 \,\mathrm{d}t\right\} \leqslant E\left\{\int_{t_{i-1}}^{t_i} |Z_t - Z_{t_{i-1}}|^2 \,\mathrm{d}t\right\}. \tag{7.9}$$

Combining (7.8) and (7.9) we have

$$\begin{aligned}
E\left\{\int_0^T |Z_t - Z^{\pi}_t|^2 \,\mathrm{d}t\right\} &= E\left\{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |Z_t - Z^{\pi,3}_{t_{i-1}}|^2 \,\mathrm{d}t\right\} \\
&\leqslant 2E\left\{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} [|Z_t - \widehat{Z}^{\pi}_{t_{i-1}}|^2 + |\widehat{Z}^{\pi}_{t_{i-1}} - Z^{\pi,3}_{t_{i-1}}|^2] \,\mathrm{d}t\right\} \\
&\leqslant 2E\left\{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} [|Z_t - Z_{t_{i-1}}|^2 + |Z_t - Z^{\pi,3}_t|^2] \,\mathrm{d}t\right\} \\
&\leqslant CE\left\{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |Z_t - Z_{t_{i-1}}|^2 \,\mathrm{d}t + \int_0^T [|Z_t - Z^{\pi,1}_t|^2 \right. \\
&\qquad \left. + |Z^{\pi,1}_t - Z^{\pi,3}_t|^2] \,\mathrm{d}t\right\}.
\end{aligned}$$

This, together with Theorem 6.3, Corollary 3.4, and (7.6), leads to the estimate for Z in (7.7). Note that similar to Theorem 3.3 one can derive the estimate for K.

We now prove (7.6). First we denote $\tilde{\Delta}_{1,3} Y_t \triangleq \tilde{Y}^{\pi,1}_t - \tilde{Y}^{\pi,3}_t$ and $\Delta_{1,3}\xi_t \triangleq \xi^{\pi,1}_t - \xi^{\pi,3}_t$ for $\xi = Y, Z$. Then by the definition of the two approximations we have

$$\tilde{\Delta}_{1,3} Y_{t_{i-1}} + \int_{t_{i-1}}^{t_i} \Delta_{1,3} Z_r \,\mathrm{d}W_r = \Delta_{1,3} Y_{t_i} + \int_{t_{i-1}}^{t_i} [f(r, \tilde{\Theta}^{\pi,1}_r) - f(t_i, \Theta^{\pi,3}_{t_i})] \,\mathrm{d}r. \tag{7.10}$$

Squaring both sides above, taking expectations, and noting that $2ab \leqslant (\Delta t_i/\varepsilon)a^2 + (\varepsilon/\Delta t_i)b^2$ for $\forall \varepsilon > 0$, we get

$$\begin{aligned}
&E\left\{|\tilde{\Delta}_{1,3} Y_{t_{i-1}}|^2 + \int_{t_{i-1}}^{t_i} |\Delta_{1,3} Z_r|^2 \,\mathrm{d}r\right\} \\
&\quad = E\left\{\left(\int_{t_{i-1}}^{t_i} [f(r, \tilde{\Theta}^{\pi,1}_r) - f(t_i, \Theta^{\pi,3}_{t_i})] \,\mathrm{d}r\right)^2 \right. \\
&\qquad \left. + |\Delta_{1,3} Y_{t_i}|^2 + 2\Delta Y_{t_i} \int_{t_{i-1}}^{t_i} [f(r, \tilde{\Theta}^{\pi,1}_r) - f(t_i, \Theta^{\pi,3}_{t_i})] \,\mathrm{d}r\right\} \\
&\quad \leqslant E\left\{(1 + C\varepsilon^{-1} \Delta t_i)|\Delta_{1,3} Y_{t_i}|^2 + C(\varepsilon + \Delta t_i) \right. \\
&\qquad \left. \times \int_{t_{i-1}}^{t_i} [f(r, \tilde{\Theta}^{\pi,1}_r) - f(t_i, \Theta^{\pi,3}_{t_i})]^2 \,\mathrm{d}r\right\}.
\end{aligned} \tag{7.11}$$

Applying Lemma 2.1(iii) and Lemma 3.2(ii) we have

$$
\begin{aligned}
&E\left\{\int_{t_{i-1}}^{t_i}[f(r,\tilde{\Theta}_r^{\pi,1})-f(t_i,\Theta_{t_i}^{\pi,3})]^2\,\mathrm{d}r\right\}\\
&\quad\leqslant CE\left\{\int_{t_{i-1}}^{t_i}[(t_i-r)+|X_r^{\pi,0}-X_{t_i}^{\pi}|^2+|\tilde{Y}_r^{\pi,1}-Y_{t_i}^{\pi,3}|^2+|Z_r^{\pi,1}-Z_{t_i}^{\pi,3}|^2]\,\mathrm{d}r\right\}\\
&\quad\leqslant CE\left\{\int_{t_{i-1}}^{t_i}[|\pi|+|X_r^{\pi,0}-X_{t_i}^{\pi,0}|^2+|\tilde{Y}_r^{\pi,1}-Y_{t_i}^{\pi,1}|^2+|\Delta_{1,3}Y_{t_i}|^2\right.\\
&\qquad\left.+|Z_r^{\pi,1}-Z_r|^2+|\delta Z_{r,t_i}|^2+|Z_{t_i}-Z_{t_i}^{\pi,3}|^2]\,\mathrm{d}r\right\}\\
&\quad\leqslant CE\left\{(1+|x|^2)|\pi|^2+\Delta t_i(|\Delta_{1,3}Y_{t_i}|^2+|Z_{t_i}-Z_{t_i}^{\pi,3}|^2)\right.\\
&\qquad\left.+\int_{t_{i-1}}^{t_i}[|Z_r^{\pi,1}-Z_r|^2+|\delta Z_{r,t_i}|^2]\,\mathrm{d}r\right\}. \qquad (7.12)
\end{aligned}
$$

Recall that $Z^{\pi,3}$ is a martingale so that $Z_{t_i}^{\pi,3}=(1/\Delta t_{i+1})E\{\int_{t_i}^{t_{i+1}}Z_r^{\pi,3}\,\mathrm{d}r|\mathscr{F}_{t_i}\}$, and that $\Delta t_i\leqslant|\pi|\leqslant L\,\Delta t_{i+1}$, we have

$$
\begin{aligned}
E\{\Delta t_i|Z_{t_i}-Z_{t_i}^{\pi,3}|^2\}&\leqslant CE\left\{\int_{t_i}^{t_{i+1}}|Z_r^{\pi,3}-Z_{t_i}|^2\,\mathrm{d}r\right\}\\
&\leqslant CE\left\{\int_{t_i}^{t_{i+1}}[|\Delta_{1,3}Z_r|^2+|Z_r^{\pi,1}-Z_r|^2+|\delta Z_{r,t_i}|^2]\,\mathrm{d}r\right\}. \qquad (7.13)
\end{aligned}
$$

Plug (7.13) into (7.12) we get

$$
\begin{aligned}
&E\left\{\int_{t_{i-1}}^{t_i}[f(r,\tilde{\Theta}_r^{\pi,1})-f(t_i,\Theta_{t_i}^{\pi,3})]^2\,\mathrm{d}r\right\}\\
&\quad\leqslant CE\left\{(1+|x|^2)|\pi|^2+\Delta t_i|\Delta_{1,3}Y_{t_i}|^2\right.\\
&\qquad\left.+\int_{t_{i-1}}^{t_{i+1}}[|Z_r^{\pi,1}-Z_r|^2+|\delta Z_{r,t_i}|^2]\,\mathrm{d}r+\int_{t_i}^{t_{i+1}}|\Delta_{1,3}Z_r|^2\,\mathrm{d}r\right\}. \qquad (7.14)
\end{aligned}
$$

We then plug (7.14) into (7.11) to get

$$
\begin{aligned}
&E\left\{|\tilde{\Delta}_{1,3}Y_{t_{i-1}}|^2+\int_{t_{i-1}}^{t_i}|\Delta_{1,3}Z_r|^2\,\mathrm{d}r\right\}\\
&\quad\leqslant E\left\{C_1(\varepsilon)|\Delta_{1,3}Y_{t_i}|^2+C_2(\varepsilon)\int_{t_i}^{t_{i+1}}|\Delta_{1,3}Z_r|^2\,\mathrm{d}r\right.\\
&\qquad\left.+C_2(\varepsilon)\int_{t_{i-1}}^{t_{i+1}}[|Z_r^{\pi,1}-Z_r|^2+|\delta Z_{r,t_i}|^2]\,\mathrm{d}r+C(1+|x|^2)|\pi|^2\right\},
\end{aligned}
$$

where $C_1(\varepsilon) = (1 + C_0\,\Delta t_i/\varepsilon)$ and $C_2(\varepsilon) = C_0(\varepsilon + \Delta t_i)$. Thus, letting $\varepsilon = 1/4C_0$, $|\pi| \leqslant 1/4C_0$, and using (7.5) we obtain that

$$
\begin{aligned}
&E\left\{|\tilde{\Delta}_{1,3} Y_{t_{i-1}}|^2 + \int_{t_{i-1}}^{t_i} |\Delta_{1,3} Z_r|^2\,\mathrm{d}r\right\}\\
&\quad\leqslant E\left\{(1 + C\,\Delta t_i)\left[|\tilde{\Delta}_{1,3} Y_{t_i}|^2 + \frac{1}{2}\int_{t_i}^{t_{i+1}} |\Delta_{1,3} Z_r|^2\,\mathrm{d}r\right]\right.\\
&\qquad\left. + C\int_{t_{i-1}}^{t_{i+1}} [|Z_r^{\pi,1} - Z_r|^2 + |\delta Z_{r,t_i}|^2]\,\mathrm{d}r + C(1 + |x|^2)|\pi|^2\right\}. \qquad (7.15)
\end{aligned}
$$

Now apply Lemma 2.6, Corollary 3.4, and Theorem 6.3, we have

$$
\begin{aligned}
\max_{0\leqslant i\leqslant n} E\{|\tilde{\Delta}_{1,3} Y_{t_i}|^2\} &\leqslant CE\left\{(1 + |x|^2)|\pi| + \int_0^T |Z_r^{\pi,1} - Z_r|^2\,\mathrm{d}r\right.\\
&\qquad\left. + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} [|\delta Z_{r,t_{i-1}}|^2 + |\delta Z_{r,t_i}|^2]\,\mathrm{d}r\right\}\\
&\leqslant C(1 + |x|^4)\sqrt{|\pi|}. \qquad (7.16)
\end{aligned}
$$

It then follows from (7.15) that $E\{\int_0^T |\Delta_{1,3} Z_r|^2\,\mathrm{d}r\} \leqslant C(1 + |x|^4)\sqrt{|\pi|}$.

We now estimate $E\{\max_{0\leqslant i\leqslant n} |\tilde{\Delta}_{1,3} Y_{t_i}|^2\}$. Again we square (7.10) to get

$$
\begin{aligned}
|\tilde{\Delta}_{1,3} Y_{t_{i-1}}|^2 &\leqslant |\tilde{\Delta}_{1,3} Y_{t_i}|^2 - 2\tilde{\Delta}_{1,3} Y_{t_{i-1}} \int_{t_{i-1}}^{t_i} \Delta_{1,3} Z_r\,\mathrm{d}W_r\\
&\quad + 2|\tilde{\Delta}_{1,3} Y_{t_i}| \int_{t_{i-1}}^{t_i} |f(r, \tilde{\Theta}_r^{\pi,1}) - f(t_i, \Theta_{t_i}^{\pi,3})|\,\mathrm{d}r\\
&\quad + \left|\int_{t_{i-1}}^{t_i} [f(r, \tilde{\Theta}_r^{\pi,1}) - f(t_i, \Theta_{t_i}^{\pi,3})]\,\mathrm{d}r\right|^2.
\end{aligned}
$$

Here we used (7.5). Noting that $Y_{t_n}^{\pi,1} = Y_{t_n}^{\pi,3}$, it is fairly easy to check recursively that

$$
\begin{aligned}
|\tilde{\Delta}_{1,3} Y_{t_{i-1}}|^2 &\leqslant -2\sum_{j=i}^n \int_{t_{j-1}}^{t_j} \tilde{\Delta}_{1,3} Y_{t_{j-1}}\,\Delta_{1,3} Z_r\,\mathrm{d}W_r\\
&\quad + |\pi| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |f(r, \tilde{\Theta}_r^{\pi,1}) - f(t_j, \Theta_{t_j}^{\pi,3})|^2\,\mathrm{d}r\\
&\quad + 2\max_{0\leqslant j\leqslant n} |\tilde{\Delta}_{1,3} Y_{t_j}| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |f(r, \tilde{\Theta}_r^{\pi,1}) - f(t_j, \Theta_{t_j}^{\pi,3})|\,\mathrm{d}r.
\end{aligned}
$$

Apply the Burkholder–Davis–Gundy inequality, we get

$$
\begin{aligned}
E\left\{\max_{0\leqslant i\leqslant n} |\tilde{\Delta}_{1,3} Y_{t_i}|^2\right\} &\leqslant CE\left\{\int_0^T |\Delta_{1,3} Z_r|^2\,\mathrm{d}r\right.\\
&\qquad\left. + \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |f(r, \tilde{\Theta}_r^{\pi,1}) - f(t_j, \Theta_{t_j}^{\pi,3})|^2\,\mathrm{d}r\right\}.
\end{aligned}
$$

Now recalling (7.14) one can easily show that $E\{\max_{0\leqslant i\leqslant n}|\tilde{\Delta}_{1,3}Y_{t_i}|^2\}\leqslant C(1+|x|^4)\sqrt{|\pi|}$.

Finally, note that $K^{\pi,3}_{t_i}=Y^{\pi,3}_0-Y^{\pi,3}_{t_i}-\sum_{j=1}^i f(t_j,\Theta^{\pi,3}_{t_j})\Delta t_j+\int_0^{t_i}Z^{\pi,3}_r\,\mathrm{d}W_r$, applying (3.11) we have

$$|\Delta_{1,3}K_{t_i}|\leqslant|\Delta_{1,3}Y_0|+|\Delta_{1,3}Y_{t_i}|+\sum_{j=1}^n\int_{t_{j-1}}^{t_j}|f(r,\tilde{\Theta}^{\pi,1}_r)-f(t_j,\Theta^{\pi,3}_{t_j})|\,\mathrm{d}r$$
$$+\left|\int_0^{t_i}\Delta_{1,3}Z_r\,\mathrm{d}W_r\right|.$$

By standard arguments one can prove that $E\{\max_{0\leqslant i\leqslant n}|\Delta_{1,3}K_{t_i}|^2\}\leqslant C(1+|x|^4)\sqrt{|\pi|}$. $\square$

Acknowledgements

We would like to thank the anonymous referee for the helpful suggestions and comments.

Appendix

We now show by examples that both inclusions in Theorem (6.2) are actually strict. These examples also indicate the fundamental difference between the space $\mathscr{L}^2(\pi)$ and the usual continuous path spaces.

Example 1. In this example we show that $\mathscr{L}^2(\pi)\backslash L^2(\Omega;C[0,T])\neq\emptyset$. In other words, we will find a process η so that for any π, $|||\,\eta\,|||_\pi\leqslant 4$, but $\|\eta\|_{L^2(\Omega;C[0,1])}=\infty$.

Consider $\Omega\triangleq\mathbb{N}$, $P\{n\}=2^{-n}$, for $n=1,2,\ldots$; and $T=1$. Let $\eta:\Omega\times[0,1]\mapsto\mathbb{R}$ be as follows:

$$\eta_t(n)\triangleq\begin{cases}2^{n/2}[1-|2^{n+2}t-3|], & 2^{-n-1}<t<2^{-n},\\ 0 & \text{otherwise}.\end{cases}$$

Then it is easily seen that for each $n\in\Omega$ $\eta_\cdot(n)$ is continuous, and $\sup_{0\leqslant t\leqslant T}|\eta_t(n)|=2^{n/2}<\infty$. Further, for any t, such that $2^{-n-1}\leqslant t<2^{-n}$, one has

$$E\{|\eta_t|^2\}=2^{-n}|\eta_t(n)|^2\leqslant 2^{-n}2^n=1.$$

Thus by Theorem 6.2 we get $|||\,\eta\,|||^2_\pi\leqslant 4\sup_{0\leqslant t\leqslant 1}E\{|\eta_t|^2\}\leqslant 4$, hence $\eta\in\mathscr{L}^2(\pi)$.

But on the other hand, since

$$E\left\{\sup_{0\leqslant t\leqslant T}|\eta_t|^2\right\}=\sum_{n=1}^\infty 2^{-n}\sup_{0\leqslant t\leqslant T}|\eta_t(n)|^2=\sum_{n=1}^\infty 1=\infty,$$

we see that $\eta\notin L^2(\Omega;C[0,T])$, as claimed.

Example 2. In this example we show that $L^2_{\mathrm{loc}}(\Omega;C[0,T])\backslash\mathscr{L}^2(\pi)\neq\emptyset$. Let us still consider the same probability space defined in Example 1, and let $T=1$. This time we define for each $n \in \Omega$,

$$\eta_t(n) \triangleq \begin{cases} 2^n[1 - 2^{2n}|2^n t - k|], & |2^n t - k| < 2^{-2n},\ k = 1, \dots, 2^n - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8.1}$$

We claim that this η belongs to $L^2_{\mathrm{loc}}(\Omega; C[0,1])$, but for any $\delta > 0$, there exists a partition π such that $|\pi| \leqslant \delta$ and $|||\,\eta\,|||_\pi = \infty$. Indeed, it is clear that for any n, $\eta_\cdot(n)$ is continuous. Moreover, a direct computation shows that

$$E\left\{\int_0^T |\eta_t|^2\,\mathrm{d}t\right\} = \sum_{n=1}^{\infty} 2^{-n} \int_0^T |\eta_t(n)|^2\,\mathrm{d}t = \sum_{n=1}^{\infty} 2^{-n} \sum_{k=1}^{2^n-1} \frac{2}{3} 2^{-n} \leqslant \frac{2}{3} \sum_{n=1}^{\infty} 2^{-n} = \frac{2}{3}.$$

Thus $\eta \in L^2_{\mathrm{loc}}(\Omega; C[0,1])$.

But on the other hand, for any m, let $\pi_m : 0 = t_0 < \cdots < t_{2^m} = 1$ be such that $t_i = i2^{-m}$, $i = 0, 1, \dots, 2^m$. Since for any $n \geqslant m$, each t_i must be a point of dyadics $\{k2^{-n}\}$. Namely, $2^n t_i - k = 0$ for some $0 \leqslant k \leqslant 2^n$, hence $\eta_{t_i}(n) \equiv 2^n$ for all $i = 1, \dots, 2^m$, whenever $n \geqslant m$. Now by definition of $|||\cdot|||_{\pi_m}$ and using the inequality $|a-b|^2 \geqslant \frac{1}{2}|a|^2 - |b|^2$ we see that

$$\begin{aligned} |||\,\eta\,|||^2_{\pi_m} &\geqslant E\left\{\sum_{i=1}^{2^m} \int_{t_{i-1}}^{t_i} |\eta_t - \eta_{t_i}|^2\,\mathrm{d}t\right\} \geqslant E\left\{\frac{1}{2}\sum_{i=1}^{2^m} |\eta_{t_i}|^2\,\Delta t_i - \int_0^1 |\eta_t|^2\,\mathrm{d}t\right\} \\ &\geqslant \sum_{n=m}^{\infty} 2^{-n} 2^{2n-1} - E\left\{\int_0^1 |\eta_t|^2\,\mathrm{d}t\right\} = \infty, \end{aligned}$$

proving the claim.

References

[1] L. Anderson, M. Broadie, A primal-dual simulation algorithm for pricing multi-dimensional American options, preprint, 2001.
[2] B. Bouchard, N. Touzi, Discrete Time Approximation and Monte-Carlo Simulation of BSDEs, preprint, 2003.
[3] M. Broadie, P. Glasserman, Pricing American-style securities using simulation, J. Econom. Dyn. Control 21 (1997) 1323–1352.
[4] E. Clément, D. Lamberton, P. Protter, An analysis of a least squares regression method for American option pricing, Finance Stochastics 6 (4) (2002) 449–471.
[5] J. Cvitanic, I. Karatzas, Backward SDE's with reflection and dynkin games, Ann. Probab. 24 (1996) 2024–2056.
[6] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, M. Quenez, Reflected solutions of backward SDE's, and related obstacle problems for PDE's, Ann. Probab. 25 (2) (1997) 702–737.
[7] N. El Karoui, E. Pardoux, M.C. Quenez, Reflected backward SDEs and American options, Numerical Methods in Finance, Cambridge University Press, Cambridge, 1997, pp. 215–231.
[9] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, N. Touzi, An application of Malliavin calculus to Monte Carlo methods in finance, Finance Stochastics 3 (4) (1999) 391–412.
[10] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, Applications of Malliavin calculus to Monte-Carlo methods in finance. II, Finance Stochastics 5 (2) (2001) 201–236.

[11] A. Gégout-Petit, E. Pardoux, Équations différentielles stochastiques rétrogrades réfléchies dans un convexe, Stochastics Stochastics Rep. 57 (1–2) (1996) 111–128.
[12] F.A. Longstaff, E.S. Schwartz, Valuing American options by simulation: a simple least-square approach, Rev. Financial Stud. 14 (2001) 113–148.
[13] J. Ma, J. Cvitanic, Reflected forward–backward SDEs and obstacle problems with boundary conditions, J. Appl. Math. Stochastic Anal. 14 (2) (2001) 113–138.
[14] J. Ma, J. Yong, Forward–backward stochastic differential equations and their applications, in: Lecture Notes in Mathematics, vol. 1702, Springer, Berlin, 1999.
[15] J. Ma, J. Zhang, Path regularity for solutions of backward SDE's, Probab. Theory Relat. Fields 122 (2001) 163–190.
[16] J. Ma, J. Zhang, Representation theorems for backward SDE's, Ann. Appl. Probab. 12 (4) (2002) 1390–1418.
[17] D. Nualart, The Malliavin Calculus and Related Topics, Springer, Berlin, 1995.
[18] J. Zhang, Some fine properties of backward stochastic differential equations, Ph.D. Thesis, Purdue University, 2001.
[19] Z. Zheng, Reflected BSDEs with random terminal time and applications, Part I–IV, preprints, 2002.