

Jin Ma · J. Zhang

Path regularity for solutions of backward stochastic differential equations

Received: 26 May 2000 / Revised version: 1 December 2000 /
Published online: 19 December 2001 – © Springer-Verlag 2001

Abstract. In this paper we study the path regularity of the adapted solutions to a class of backward stochastic differential equations (BSDE, for short) whose terminal values are allowed to be functionals of a forward diffusion. Using the new representation formula for the adapted solutions established in our previous work [7], we are able to show, under the minimum Lipschitz conditions on the coefficients, that for a fairly large class of BSDEs whose terminal values are functionals that are either Lipschitz under the L^∞ -norm or under the L^1 -norm, then there exists a version of the adapted solution pair that has at least càdlàg paths. In particular, in the latter case the version can be chosen so that the paths are in fact continuous.

1. Introduction

Let $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\})$ be a complete filtered probability space on which is defined a d -dimensional Brownian motion W ; and suppose that the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is generated by the Brownian motion W with the usual augmentation. The celebrated Martingale Representation Theorem (see, e.g., [3], [12]) states that every square integrable martingale must be of the form of a stochastic integral against the Brownian motion W . In particular, for any \mathcal{F}_T -measurable random variable ξ such that $E|\xi|^2 < \infty$, there exists an $\{\mathcal{F}_t\}$ -predictable process Z with $E \int_0^T |Z_t|^2 dt < \infty$ such that

$$E\{\xi|\mathcal{F}_t\} = E\{\xi\} + \int_0^t \langle Z_s, dW_s \rangle, \quad t \in [0, T], \quad (1.1)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in \mathbb{R}^n . Now our question is: *what can we say about the path regularity of the process Z ?* The answer to this question is quite indefinite, as the following examples show:

(A) $\xi = W_T$. Then $Z_t \equiv 1, \forall t \in [0, T]$;

J. Ma, J. Zhang: Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395, USA. e-mail: majin@math.purdue.edu; jfzhang@math.purdue.edu.

The first author is supported in part by NSF grant no. 9971720. The second author is supported in part by NSF grant no. 9971720.

Mathematics Subject Classification (2000): Primary 60H10; Secondary 34F05, 90A12

Key words or phrases: Backward SDE's – Adapted solutions – Path regularities – Pseudo-paths – Meyer-Zheng topology – Quasimartingales

- (B) $\xi = \max_{0 \leq t \leq T} W_t$. Then by the Clark-Ocone formula, $Z_t = E\{D_t \xi | \mathcal{F}_t\} = E\{1_{[0, \tau]}(t) | \mathcal{F}_t\}$, where D is the ‘‘Malliavin derivative’’ and τ is the a.s. unique point where W attains its maximum (cf. e.g., [9]).
- (C) $\xi = \int_0^T h_s dW_s$, where h is any $\{\mathcal{F}_t\}$ -predictable process such that $E \int_0^T |h_s|^2 ds < \infty$, then $Z_t \equiv h_t, \forall t \in [0, T]$.

While the example (C) shows that one may not conclude anything in general, examples (A) and (B) do show that in some cases Z could indeed be ‘‘piecewise continuous’’, especially when ξ is of the form as a functional of the Brownian path W . In fact, it has long been conjectured that the process Z should have a version that has càdlàg paths whenever $\xi = \Phi(W)_T$, where $\Phi : C([0, T]) \mapsto \mathbb{R}$ is a ‘‘nice’’ functional of W .

The above problem has been observed in a more general setting: instead of (1.1) let us consider a *backward stochastic differential equation* (BSDE) initiated by Pardoux-Peng (1990) [10]:

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r) dr - \int_t^T \langle Z_r, dW_r \rangle, \quad t \in [0, T], \quad (1.2)$$

where $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ is some appropriate measurable function. It is easy to check that (1.1) corresponds to the special case of (1.2) with $f \equiv 0$, and $Y_t = E\{\xi | \mathcal{F}_t\}, t \in [0, T]$. A well-investigated case of (1.2) is the following extension of Example (A): $\xi = g(X_T)$, where g is a function and $f = f(t, X_t, Y_t, Z_t)$, where X is a diffusion given by the SDE:

$$X_t = x + \int_0^t b(X_r) dr + \int_0^t \sigma(X_r) dW_r, \quad t \in [0, T]. \quad (1.3)$$

In this case, Pardoux and Peng proved in one of their seminal works [11] that, among other things, the process Z has continuous paths provided that the coefficients of (1.2) and (1.3) are sufficiently smooth (in particular, f and g are C^3 in their spatial variables). In fact, viewing (1.2) and (1.3) as a special (decoupled) case of the so-called *forward-backward* SDEs, it was shown in Ma-Protter-Yong [5] that the process Z has a more explicit expression:

$$Z_t = u_x(t, X_t) \sigma(X_t), \quad t \in [0, T], \quad (1.4)$$

where u is the classical solution to a quasilinear PDE. Consequently Z must be continuous, verifying the result of [11]. However, we should note that all these results require rather heavy smoothness of the coefficients f and g . In a recent paper Jacod-Méléard-Protter [2] studied, among other things, the explicit form and path regularity of Z , as a ‘‘martingale representation’’ problem when W is allowed to be any Lévy process and g is continuously differentiable. But the results there are only valid when $f \equiv 0$ and g is at most the ‘‘discrete functional’’ as is defined in this paper. In fact, to our best knowledge, so far there has been no regularity results for the process Z , even in the Brownian case, when $f (\neq 0)$ and g are only Lipschitz continuous.

In this paper we shall prove that the conjecture on the path regularity of the process Z is indeed true for a large class of BSDEs with functional terminal condition and Lipschitz generators. More precisely, we show that if the terminal condition of a BSDE is of the form $\xi = \Phi(X)_T$, where Φ is Lipschitz continuous (not necessarily bounded!) under the sup-norm, then the process Z must have a càdlàg version; and only requirement on the generator f is that it is Lipschitz continuous in all spatial variables. Moreover, if the functional g is Lipschitz under the L^1 -norm, we show that Z will even have an a.s. continuous version. The significance of such a result lies in that it will enable one to put the solution pair (Y, Z) in a *canonical* path space, such as the well-known D -space with Skorohod topology, which opens the door to many further studies on BSDEs, especially to those concerning the solutions in a *weak* sense, both theoretically or numerically. Our result relies heavily on the representation formula that we established in our previous paper [7]; and a key device we use to prove the convergence of our approximation scheme is the Meyer-Zheng topology on pseudo-paths of stochastic processes (cf. [8]).

The rest of the paper is organized as follows. In section 2 we give some preliminary results, including the notion of Meyer-Zheng pseudo-path topology. In section 3 we study the case when the functional Φ depends only on finitely many points of X and establish a crucial estimate on the conditional variation of Z . In section 4 we study the case where the functional Φ is Lipschitz under the sup-norm; and in section 5 we extend the result to the so-called “integral Lipschitz” case and prove a much stronger result.

2. Preliminaries

Throughout this paper we assume that (Ω, \mathcal{F}, P) is a complete probability space on which is defined a d -dimensional Brownian motion $W = (W_t)_{t \geq 0}$. Let $\mathbf{F} \triangleq \{\mathcal{F}_t\}_{t \geq 0}$ denote the natural filtration generated by W , augmented by the P -null sets of \mathcal{F} ; and let $\mathcal{F} = \mathcal{F}_\infty$. We shall denote \mathbb{E} to be a generic Euclidean space (or $\mathbb{E}_1, \mathbb{E}_2, \dots$, if different spaces are used simultaneously); and regardless of their dimensions we denote $\langle \cdot, \cdot \rangle$ and $|\cdot|$ to be the inner product and norm in all \mathbb{E} 's, respectively. We should point out here that although most of the vectors in this paper are considered as *column vectors*, we sometimes require certain multi-dimensional process to be of *row vector* form for notational convenience. For instance, throughout this paper we denote $\partial_x = (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_d})$. Thus if $\varphi : \mathbb{R}^d \mapsto \mathbb{R}$ is differentiable, then $\partial_x \varphi \triangleq (\partial_{x_1} \varphi, \dots, \partial_{x_d} \varphi)$ will be a *row vector*! Also, if $\psi = (\psi^1, \dots, \psi^d)^T : \mathbb{R}^d \mapsto \mathbb{R}^d$ is differentiable, then $\partial_x \psi \triangleq (\partial_{x_i} \psi^j)_{i,j=1}^d$ is a matrix, with $\partial_x \psi^j$, $j = 1, \dots, d$ being its row vectors.

We inherit the notations from [7] for the following spaces that will be frequently used in the sequel: let \mathcal{X} denote a generic Banach space,

- $L^0([0, T]; \mathcal{X})$ is the space of all Borel measurable functions $\varphi : [0, T] \mapsto \mathcal{X}$;
- $C([0, T]; \mathcal{X})$ is the space of all continuous functions $\varphi : [0, T] \mapsto \mathcal{X}$;

- for integers k and ℓ , $C^{k,\ell}([0, T] \times \mathbb{E}; \mathbb{E}_1)$ is the space of all \mathbb{E}_1 -valued functions $\varphi(t, e)$, $(t, e) \in [0, T] \times \mathbb{E}$, such that they are k -times continuously differentiable in t and ℓ -times continuously differentiable in e .
- $C_b^{k,\ell}([0, T] \times \mathbb{E}; \mathbb{E}_1)$ is the space of those $\varphi \in C^{k,\ell}([0, T] \times \mathbb{E}; \mathbb{E}_1)$ such that all the partial derivatives are uniformly bounded;
- $W^{1,\infty}(\mathbb{E}; \mathbb{E}_1)$ is the space of all measurable functions $\psi : \mathbb{E} \mapsto \mathbb{E}_1$, such that for some constant $K > 0$ it holds that $\|\psi(x) - \psi(y)\|_{\mathbb{E}_1} \leq K\|x - y\|_{\mathbb{E}}$, $\forall x, y \in \mathbb{E}$;
- for any sub- σ -field $\mathcal{G} \subseteq \mathcal{F}_T$ and $0 \leq p < \infty$, $L^p(\mathcal{G}; \mathbb{E})$ denotes all \mathbb{E} -valued, \mathcal{G} -measurable random variable ξ such that $E|\xi|^p < \infty$. Moreover, $\xi \in L^\infty(\mathcal{G}; \mathbb{R}^d)$ means it is \mathcal{G} -measurable and is bounded;
- for $0 \leq p < \infty$, $L^p(\mathbf{F}, [0, T]; \mathcal{X})$ is the space of all \mathcal{X} -valued, \mathbf{F} -adapted processes ξ satisfying $E \int_0^T \|\xi_t\|_{\mathcal{X}}^p dt < \infty$; and also, $\xi \in L^\infty(\mathbf{F}, [0, T]; \mathcal{X})$ means it is a process uniformly bounded in (t, ω) ;
- $C(\mathbf{F}, [0, T] \times \mathbb{E}; \mathbb{E}_1)$ is the space of all \mathbb{E}_1 -valued, continuous random field $\varphi : \Omega \times [0, T] \times \mathbb{E} \mapsto \mathbb{E}_1$, such that for fixed $e \in \mathbb{E}$, $\varphi(\cdot, \cdot, e)$ is an \mathbf{F} -adapted process.

To simplify notation we often denote: $C([0, T] \times \mathbb{E}; \mathbb{E}_1) = C^{0,0}([0, T] \times \mathbb{E}; \mathbb{E}_1)$; $C^{k,\ell}([0, T] \times \mathbb{E}) = C^{k,\ell}([0, T] \times \mathbb{E}; \mathbb{R})$; $C(\mathbf{F}, [0, T] \times \mathbb{E}) = C(\mathbf{F}, [0, T] \times \mathbb{E}; \mathbb{R})$ and $W^{1,\infty}(\mathbb{E}) = W^{1,\infty}(\mathbb{E}; \mathbb{R})$, ..., etc.

The main object of this paper is the following system of SDEs:

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t; \\ dY_t = -f(t, X_t, Y_t, Z_t)dt + \langle Z_t, dW_t \rangle, \\ X_0 = x, Y_T = \xi \end{cases} \tag{2.1}$$

where $x \in \mathbb{R}^n$ and $\xi \in L^2(\mathcal{F}_T, \mathbb{R})$, and the negative sign in front of the drift of the second equation is for convenience in the future. Note that the second SDE in (2.1) is specified for a *terminal value*, it is thus called a *backward SDE*. The solution to (2.1) is defined as a triplet $(X, Y, Z) \in L^2(\mathbf{F}; C([0, T]; \mathbb{R}^n)) \times L^2(\mathbf{F}; C([0, T])) \times L^2(\mathbf{F}, [0, T]; \mathbb{R}^d)$, where the process Z is called the *martingale part* of the solution. To simplify notation from now on we make the convention that the martingale part of the solution to the FBSDE (2.1), the process $Z \triangleq (Z^1, \dots, Z^d)$, is a *row vector*. Thus, the stochastic integral in (2.1) can now be simply written as $\int Z_s dW_s$.

Throughout this paper we shall make use of the following *Standing Assumptions*:

(A1) $n = d$. The functions $\sigma \in C_b^{0,1}([0, T] \times \mathbb{R}^d; \mathbb{R}^{d \times d})$, $b \in C_b^{0,1}([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$; and all the partial derivatives of b and σ are uniformly bounded by a common constant $K > 0$. Further, there exists a constant $c > 0$, such that

$$\xi^T \sigma(t, x) \sigma^T(t, x) \xi \geq c|\xi|^2, \quad \forall t \in [0, T] \forall x, \xi \in \mathbb{R}^d. \tag{2.2}$$

(A2) The function $f \in C([0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d) \cap L^0([0, T]; W^{1,\infty}(\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d))$. Furthermore, we denote the Lipschitz constant of f by the same $K > 0$ as in (A1); and assume that $\sup_{0 \leq t \leq T} \left\{ |b(t, 0)| + |\sigma(t, 0)| + |f(t, 0, 0, 0)| \right\} \leq K$.

The following results are either standard or slight variations of the well-known results in the SDE and the backward SDE literature. As in [7], we state them for ready references.

Lemma 2.1. *Suppose that $\tilde{b} \in C(\mathbf{F}, [0, T] \times \mathbb{R}^d; \mathbb{R}^d) \cap L^0(\mathbf{F}, [0, T]; W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d))$ and $\tilde{\sigma} \in C(\mathbf{F}, [0, T] \times \mathbb{R}^d; \mathbb{R}^{d \times d}) \cap L^0(\mathbf{F}, [0, T]; W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^{d \times d}))$, with a common Lipschitz constant $K > 0$. Assume further that $\tilde{b}(t, 0) = 0$, $\tilde{\sigma}(t, 0) = 0$, and for any $h^0 \in L^2(\mathbf{F}, [0, T]; \mathbb{R}^d)$, $h^1 \in L^2(\mathbf{F}, [0, T]; \mathbb{R}^{d \times d})$, let X be the solution of FSDE:*

$$X_t = x + \int_0^t [\tilde{b}(s, X_s) + h_s^0] ds + \int_0^t [\tilde{\sigma}(s, X_s) + h_s^1] dW_s; \tag{2.3}$$

Then for any $p \geq 2$, there exists a constant $C > 0$ depending only on p and T and the Lipschitz constant K , such that

$$E \left\{ \sup_{0 \leq t \leq T} |X_t|^p \right\} \leq C \left\{ |x|^p + E \int_0^T [|h_t^0|^p + |h_t^1|^p] dt \right\}. \tag{2.4}$$

Lemma 2.2. *Assume that $\tilde{f} \in C(\mathbf{F}, [0, T] \times \mathbb{R} \times \mathbb{R}^d) \cap L^0(\mathbf{F}, [0, T]; W^{1,\infty}(\mathbb{R} \times \mathbb{R}^d))$ with a uniform Lipschitz constant $K > 0$; and assume further that $\tilde{f}(\omega, s, 0, 0) = 0$, P -a.e. $\omega \in \Omega$. For any $\xi \in L^2(\mathcal{F}_T; \mathbb{R})$ and $h \in L^2(\mathbf{F}, [0, T])$, let (Y, Z) be the adapted solution to the BSDE:*

$$Y_t = \xi + \int_t^T [\tilde{f}(s, Y_s, Z_s) + h_s] ds - \int_t^T Z_s dW_s. \tag{2.5}$$

Then, there exists a constant $C > 0$ depending only on T and the Lipschitz constant K , such that

$$E \left\{ \int_0^T |Z_t|^2 dt \right\} \leq CE \left\{ |\xi|^2 + \int_0^T |h_t|^2 dt \right\}. \tag{2.6}$$

Moreover, for all $p \geq 2$, there exists a constant $C_p > 0$ that might depend on p , such that

$$E \left\{ \sup_{0 \leq t \leq T} |Y_t|^p \right\} \leq C_p E \left\{ |\xi|^p + \int_0^T |h_t|^p dt \right\}. \tag{2.7}$$

To conclude this section we introduce the notions of *pseudo-path topology* and *quasimartingales* (cf. Dellacherie-Meyer [1] or Meyer-Zheng [8]), adjusted to our setting. To begin with, let us denote $\mathbb{D} \triangleq \mathbb{D}([0, T]) \subset L^0([0, T])$ to be the space of all càdlàg function on $[0, T]$. For any $w \in L^0([0, T])$, we define the *pseudo-path* of w to be a probability measure on $[0, T] \times \bar{\mathbb{R}}$:

$$P^w(A) \triangleq \frac{1}{T} \int_0^T 1_A(t, w(t)) dt, \quad \forall A \in \mathcal{B}([0, T] \times \bar{\mathbb{R}}). \tag{2.8}$$

It can be shown that the mapping $\psi : w \mapsto P^w$ is 1-1 on $L^0([0, T])$. Thus we can identify all $w \in L^0([0, T])$ with its pseudo-path; and we denote all pseudo-paths by Ψ . In particular, using the mapping ψ the space \mathbb{D} can then be imbedded into the compact space $\overline{\mathcal{P}}$ of all probability laws on the compact space $[0, T] \times \overline{\mathbb{R}}$ (with the Prohorov metric). Clearly, in this sense

$$\mathbb{D} \subset \Psi \subset \overline{\mathcal{P}}. \tag{2.9}$$

The induced topology on Ψ and \mathbb{D} are known as the *pseudo-path topology* or sometimes called *Meyer-Zheng topology*. The following characterization of the Meyer-Zheng topology is worth noting.

Lemma 2.3. (Meyer-Zheng [8, Lemma 1]). *The pseudo-path topology on Ψ is equivalent to the convergence in measure.*

Furthermore, it is known that (see, e.g., [8]) Ψ is a Polish space; and \mathbb{D} is a Borel set in $\overline{\mathcal{P}}$. Consequently, we have

$$\mathcal{B}(\mathbb{D}) = \mathbb{D} \cap \mathcal{B}(\Psi) \stackrel{\Delta}{=} \{A \cap \mathbb{D} : A \in \mathcal{B}(\Psi)\}.$$

We now make the following observation. Denote $\mathcal{M}(\mathbb{D})$ to be the space of all probability measures on \mathbb{D} , and $\mathcal{M}(\Psi)$ be that of Ψ . Then, any probability measure $P \in \mathcal{M}(\mathbb{D})$ induces a probability measure $\widehat{P} \in \mathcal{M}(\Psi)$ by:

$$\widehat{P}(A) = P(A \cap \mathbb{D}), \quad \forall A \in \mathcal{B}(\Psi). \tag{2.10}$$

In this sense we then have $\mathcal{M}(\mathbb{D}) \subset \mathcal{M}(\Psi)$.

The most significant application of the Meyer-Zheng topology is a tightness result for *quasimartingales*, which we now briefly describe. Let X be an \mathbf{F} -adapted, càdlàg process defined on $[0, T]$, such that $E|X_t| < \infty$ for all $t \geq 0$. For any partition $\pi : 0 = t_0 < t_1 < \dots < t_n \leq T$, let us define

$$V_T^\pi(X) \stackrel{\Delta}{=} \sum_{0 \leq i < n} E\{|E\{X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}\}| + E|X_{t_n}|\}, \tag{2.11}$$

and define the *conditional variation* of X by $V_T(X) \stackrel{\Delta}{=} \sup_\pi V_T^\pi(X)$. If $V_T(X) < \infty$, then X is called a *quasimartingale*¹. We have the following result.

Lemma 2.4. (Meyer-Zheng [8]). *Let $\{P_n\}_{n \geq 1} \subset \mathcal{M}(\mathbb{D})$, such that under each P_n the coordinate process $X_t(\omega) = \omega(t)$, $t \in [0, T]$, $\omega \in \mathbb{D}$, is a quasimartingale. Assume that $V_n(X)$, $n \geq 1$, the conditional variation of X under P_n 's, are uniformly bounded in n . Then there exists a subsequence $\{P_{n_k}\}$ which converges weakly on \mathbb{D} to a law $P^* \in \mathcal{M}(\mathbb{D})$, and X is a quasimartingale under P^* .*

¹ We should note that the quasimartingale in [8] is defined on $[0, \infty]$. However, it is fairly easy to check that if X is a quasimartingale on $[0, T]$ as is defined above, then the process $\widehat{X}_t = X_t I_{[0, T)}(t) + X_T I_{[T, \infty)}(t)$, $t \in [0, \infty]$ is a quasimartingale in the sense of [8]. Furthermore, the conditional variation $V_T(X)$ defined here, although looks slightly different, is exactly the same as $V(\widehat{X})$ defined in [8]. In other words, our quasimartingale is a ‘‘local’’ version of that in [8].

3. Discrete functional case revisited

Before we begin our investigation, let us first recall a path regularity result we derived in [7]. Since this result is only valid when the terminal value ξ takes the form $\xi = g(X_{t_0}, \dots, X_{t_n})$, where $g \in C(\mathbb{R}^{d(n+1)})$, which is a functional depending on a discrete set of variables, we call it a *discrete functional case* in the sequel. In this section we shall establish some further properties of the adapted solution to BSDEs with such terminal values.

Let us assume that the standing assumptions (A1) and (A2) hold, and that $g \in C_b^1(\mathbb{R}^{d(n+1)})$ and $f \in C^{0,1}([0, T] \times \mathbb{R}^{2d+1})$. Let $\pi : 0 = t_0 < t_1 < \dots < t_n = T$ be a given partition of $[0, T]$, and let (Y, Z) be the adapted solution to the following BSDE:

$$Y_t = g(X_{t_0}, \dots, X_{t_n}) + \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0, T]. \quad (3.1)$$

Then from Theorem 5.1 of [7] and its proof we know that the martingale part Z has a càdlàg version, and that on each subinterval $[t_{i-1}, t_i]$, Z is continuous, such that the following representation holds:

$$Z_t = \nabla^i Y_t (\nabla X_t)^{-1} \sigma(t, X_t), \quad t \in [t_{i-1}, t_i]. \quad (3.2)$$

Here in (3.2), ∇X satisfies the following *variational equation*: for $k = 1, \dots, d$,

$$\nabla_k X_t = e_k + \int_0^t \partial_x b(r, X_r) \nabla_k X_r dr + \sum_{j=1}^d \int_0^t [\partial_x \sigma^j(r, X_r)] \nabla_k X_r dW_r^j, \quad (3.3)$$

and for $i = 1, \dots, n$, $(\nabla^i Y, \nabla^i Z)$ satisfies the following BSDE on $[t_{i-1}, T]$:

$$\begin{aligned} \nabla^i Y_t &= \sum_{j \geq i} \partial_j g \nabla X_{t_j} + \int_t^T [f_x(r) \nabla X_r + f_y(r) \nabla^i Y_r + f_z(r) \nabla^i Z_r] dr \\ &\quad - \left\{ \int_t^T \nabla^i Z_r dW_r \right\}^T, \quad t \in [t_{i-1}, T], \end{aligned} \quad (3.4)$$

where $e_k = (0, \dots, 1, \dots, 0)^T$ is the k -th coordinate vector of \mathbb{R}^d ; $\sigma^j(\cdot)$ is the j -th column of the matrix $\sigma(\cdot)$; and with $\Theta(\cdot) \triangleq (X, Y, Z)$,

$$\begin{cases} \partial_j g = (\partial_{x_j^1} g(X_{t_0}, \dots, X_{t_n}), \dots, \partial_{x_j^d} g(X_{t_0}, \dots, X_{t_n})); \\ (f_x(r), f_y(r), f_z(r)) = (\partial_x f(r, \Theta(r)), \partial_y f(r, \Theta(r)), \partial_z f(r, \Theta(r))). \end{cases}$$

The main purpose of this paper is to generalize the path regularity result of above to the case where the terminal value of the BSDE (3.1) takes the general form: $Y_T = \Phi(X)$, where $\Phi : C([0, T]; \mathbb{R}^d) \mapsto \mathbb{R}$ is some functional on the path space of X . Our plan is to approximate a general function Φ by a sequence of discrete functionals, and try to prove that the paths of the martingale part of the solution under study is a limit of the sequence of corresponding solutions of

BSDEs with discrete functional terminal on the path space(!), from which the path regularity will follow.

To this end, we need some further properties on the adapted solution to the BSDE (3.1). For a fixed partition $\pi : 0 = t_0 < t_1 < \dots < t_n = T$, we denote

$$\nabla^\pi Y_s = \sum_{i=1}^n \nabla^i Y_s 1_{[t_{i-1}, t_i)}(s) + \nabla^n Y_{T-1} 1_{\{T\}}(s), \quad s \in [0, T]. \tag{3.5}$$

Then $\nabla^\pi Y$ is a càdlàg process. The following result is essential.

Theorem 3.1. *Assume (A1) and (A2); and that g and f are both continuously differentiable in the spatial variable (x, y, z) with uniformly bounded partial derivatives. Assume further that for some constant $L > 0$, it holds for all $x = (x_0, \dots, x_n) \in \mathbb{R}^{d(n+1)}$, $y = (y_0, \dots, y_n) \in \mathbb{R}^{n+1}$ that*

$$\sum_{i=0}^n |\partial_{x_i} g(x)y_i| \leq L \max_i |y_i|. \tag{3.6}$$

Then, there exists a constant $C > 0$, depending on T, K and L , but independent of the partition π , such that

$$\sum_{i=1}^n E \left\{ \left| E \left\{ \nabla^\pi Y_{t_{i-1}} - \nabla^\pi Y_{t_i} \middle| \mathcal{F}_{t_{i-1}} \right\} \right| \right\} + E\{|\nabla^\pi Y_T|\} \leq C. \tag{3.7}$$

Proof. We begin by making the following convention: in what follows we denote $C > 0$ to be a generic constant depending only on constants K in (A1) and (A2); L in (3.6), and the time duration $T > 0$, which is allowed to vary from line to line.

First note that (3.6) implies that $|\partial_{x_n} g| \leq L$. Thus by (3.5) we have

$$E\{|\nabla^\pi Y_T|\} \leq LE\{|\nabla X_T|\} \leq C.$$

Next, since for each i $(\nabla^i Y, \nabla^i Z)$ satisfies a linear BSDE (3.4), let (γ^0, ζ^0) and (γ^j, ζ^j) , $j = 1, \dots, n$ be the adapted solutions of the BSDEs

$$\begin{aligned} \gamma_t^0 &= \int_t^T \left[f_x(r) \nabla X_r + f_y(r) \gamma_r^0 + f_z(r) \zeta_r^0 \right] dr - \left\{ \int_t^T \zeta_r^0 dW_r \right\}^T; \\ \gamma_t^j &= \partial_j g \nabla X_{t_j} + \int_t^T \left[f_y(r) \gamma_r^j + f_z(r) \zeta_r^j \right] dr - \left\{ \int_t^T \zeta_r^j dW_r \right\}^T, \end{aligned} \tag{3.8}$$

respectively, we have the following decomposition:

$$\nabla^i Y_s = \gamma_s^0 + \sum_{j \geq i} \gamma_s^j, \quad s \in [t_{i-1}, t_i). \tag{3.9}$$

Therefore, using (3.5) and (3.9) we see that for each i ,

$$\begin{aligned} \nabla^\pi Y_{t_{i-1}} - \nabla^\pi Y_{t_i} &= \nabla^i Y_{t_{i-1}} - \nabla^{i+1} Y_{t_i} = \left(\gamma_{t_{i-1}}^0 + \sum_{j \geq i} \gamma_{t_{i-1}}^j \right) - \left(\gamma_{t_i}^0 + \sum_{j \geq i+1} \gamma_{t_i}^j \right) \\ &= [\gamma_{t_{i-1}}^0 - \gamma_{t_i}^0] + \gamma_{t_i}^i + \sum_{j \geq i} \left[\gamma_{t_{i-1}}^j - \gamma_{t_i}^j \right]. \end{aligned} \tag{3.10}$$

Now let us denote the first term of the left hand side of (3.7) by I and show that $I \leq C$ as well. First note that

$$\begin{aligned} I &\leq E \left\{ \sum_{i=1}^n \left| E \left\{ \gamma_{t_{i-1}}^0 - \gamma_{t_i}^0 \middle| \mathcal{F}_{t_{i-1}} \right\} \right| \right\} + E \left\{ \sum_{i=1}^n \left| E \left\{ \gamma_{t_i}^i \middle| \mathcal{F}_{t_{i-1}} \right\} \right| \right\} \\ &\quad + E \left\{ \sum_{i=1}^n \left| \sum_{j \geq i} E \left\{ \gamma_{t_{i-1}}^j - \gamma_{t_i}^j \middle| \mathcal{F}_{t_{i-1}} \right\} \right| \right\} \\ &= I_1 + I_2 + I_3, \end{aligned} \quad (3.11)$$

where $I_i, i = 1, 2, 3$ are defined in the obvious way. We now estimate I_1 – I_3 separately. First, by definition (3.8) we have

$$\begin{aligned} I_1 &= E \left\{ \sum_{i=1}^n \left| E \left\{ \int_{t_{i-1}}^{t_i} \left[f_x(r) \nabla X_r + f_y(r) \gamma_r^0 + f_z(r) \zeta_r^0 \right] dr \middle| \mathcal{F}_{t_{i-1}} \right\} \right| \right\} \\ &\leq \sum_{i=1}^n E \left\{ \int_{t_{i-1}}^{t_i} \left| f_x(r) \nabla X_r + f_y(r) \gamma_r^0 + f_z(r) \zeta_r^0 \right| dr \right\} \\ &= E \left\{ \int_0^T \left| f_x(r) \nabla X_r + f_y(r) \gamma_r^0 + f_z(r) \zeta_r^0 \right| dr \right\} \\ &\leq CE \left\{ \int_0^T \left(1 + |\nabla X_r|^2 + |\gamma_r^0|^2 + |\zeta_r^0|^2 \right) dr \right\}. \end{aligned} \quad (3.12)$$

Applying Lemmas 2.1 and 2.2 to SDEs (3.3) and (3.8) we conclude that $I_1 \leq C$.

To estimate I_2 let us modify the BSDE (3.8) slightly. We define for any $\eta \in L^1(\mathbf{F}, [0, T])$ and $\theta \in L^2(\mathbf{F}, [0, T]; \mathbb{R}^d)$ (viewed as *row vector!*),

$$\Lambda_t^s(\eta) \triangleq \exp \left\{ \int_s^t \eta(r) dr \right\}; \quad s, t \in [0, T]; \quad (3.13)$$

$$\mathcal{E}_t^s(\theta) \triangleq \exp \left\{ \int_s^t \theta(r) dW_r - \frac{1}{2} \int_s^t |\theta(r)|^2 dr \right\}, \quad s, t \in [0, T]. \quad (3.14)$$

($\mathcal{E}^t(\theta)$ is known as the *Doléan-Dade stochastic exponential* of θ .) Then it is easily checked that, for any $p > 0$, one has

$$[\mathcal{E}_t^s(\theta)]^p = \mathcal{E}_t^s(p\theta) \Lambda_t^s \left(\frac{p(p-1)}{2} |\theta|^2 \right); \quad (3.15)$$

and

$$[\mathcal{E}_t^s(\theta)]^{-1} = \mathcal{E}_t^s(-\theta) \Lambda_t^s(|\theta|^2). \quad (3.16)$$

In particular, we denote, for $s, t \in [0, T]$,

$$\Lambda_t^s = \Lambda_t^s(-f_y); \quad M_t^s = \mathcal{E}_t^s(f_z); \quad (3.17)$$

and if there is no danger of confusion, we denote $\Lambda = \Lambda^0$ and $M = M^0$. Since f_z is uniformly bounded, by Girsanov's Theorem (see, e.g., [3]) we know that M is a

P -martingale on $[0, T]$, and $\tilde{W}_t \triangleq W_t - \int_0^t f_z(r)dr$, $t \in [0, T]$ is an \mathbf{F} -Brownian motion on the new probability space $(\Omega, \mathcal{F}, \tilde{P})$, where \tilde{P} is defined by $\frac{d\tilde{P}}{dP} = M_T$.

Now we define

$$\tilde{\gamma}_t^i \triangleq \gamma_t^i \Lambda_t^{-1}, \quad \tilde{\zeta}_t^i \triangleq \zeta_t^i \Lambda_t^{-1}, \quad t \in [0, T]. \quad (3.18)$$

Then, using integration by parts and equation (3.8) we have

$$\tilde{\gamma}_t^i = \xi_i - \left\{ \int_t^T \tilde{\zeta}_r^i d\tilde{W}_r \right\}^T, \quad t \in [0, T], \quad (3.19)$$

where $\xi_i \triangleq \partial_i g \nabla X_{t_i} \Lambda_T^{-1}$. Therefore, by the Bayes rule (see, e.g., [3, Lemma 3.5.3]) we have, for $t \in [0, T]$,

$$\gamma_t^i = \tilde{\gamma}_t^i \Lambda_t = \tilde{E}\{\xi_i | \mathcal{F}_t\} \Lambda_t = E\{M_T \xi_i | \mathcal{F}_t\} M_T^{-1} \Lambda_t = E\{M_T^t \Lambda_t \xi_i | \mathcal{F}_t\}. \quad (3.20)$$

Now, recall the definition of ξ^i 's and the boundedness of f_y (whence both Λ and Λ^{-1}), we obtain that

$$\begin{aligned} I_2 &= \sum_{i=1}^n E \left| E\{\gamma_{t_i}^i | \mathcal{F}_{t_{i-1}}\} \right| \leq \sum_{i=1}^n E \{|M_T^{t_i} \Lambda_{t_i} \xi_i|\} \\ &= E \left\{ \sum_{i=1}^n |\partial_i g \nabla X_{t_i} \Lambda_T^{-1} M_T^{t_i} \Lambda_{t_i}| \right\} \leq E \left\{ \sum_{i=1}^n |\partial_i g \nabla X_{t_i} M_T^{t_i} [\Lambda_T^{t_i}]^{-1}| \right\} \\ &\leq CE \left\{ \sup_{0 \leq t \leq T} |M_T^t \nabla X_t| \right\} \leq C. \end{aligned} \quad (3.21)$$

Here we have used the assumption (3.6), as well as the fact that both M and ∇X are continuous, square-integrable processes.

The estimate for I_3 is a little more involved. First, from (3.20) we see that

$$\begin{aligned} E\{\gamma_{t_{i-1}}^j - \gamma_{t_i}^j | \mathcal{F}_{t_{i-1}}\} &= E\{\xi_j [M_T^{t_{i-1}} \Lambda_{t_{i-1}} - M_T^{t_i} \Lambda_{t_i}] | \mathcal{F}_{t_{i-1}}\} \\ &= E\{\xi_j M_T^{t_i} (\Lambda_{t_{i-1}} - \Lambda_{t_i}) | \mathcal{F}_{t_{i-1}}\} + E\{\xi_j \Lambda_{t_{i-1}} (M_T^{t_{i-1}} - M_T^{t_i}) | \mathcal{F}_{t_{i-1}}\}. \end{aligned}$$

Thus

$$\begin{aligned} I_3 &= E \left\{ \sum_{i=1}^n \left| \sum_{j \geq i} E \left\{ \gamma_{t_{i-1}}^j - \gamma_{t_i}^j | \mathcal{F}_{t_{i-1}} \right\} \right| \right\} \\ &\leq \sum_{i=1}^n \sum_{j \geq i} E \{ |\xi_j| |M_T^{t_i}| |\Lambda_{t_{i-1}} - \Lambda_{t_i}| \} \\ &\quad + \sum_{i=1}^n E \left\{ \left| \sum_{j \geq i} E \{ \xi_j \Lambda_{t_{i-1}} (M_T^{t_{i-1}} - M_T^{t_i}) | \mathcal{F}_{t_{i-1}} \} \right| \right\} \\ &= I_3^1 + I_3^2. \end{aligned}$$

Again, using the boundedness of f_y we have

$$\begin{aligned} |\Lambda_{t_{i-1}} - \Lambda_{t_i}| &= |\Lambda_{t_{i-1}}| \left| \exp \left\{ - \int_{t_{i-1}}^{t_i} f_y(r) dr \right\} - 1 \right| \leq C \left| \int_{t_{i-1}}^{t_i} f_y(r) dr \right| \\ &\leq C(t_i - t_{i-1}). \end{aligned} \quad (3.22)$$

Moreover, by assumption (3.6) we have

$$\sum_{j \geq i} |\xi_j| \leq \sum_{j=0}^n |\partial_j g \nabla X_{t_j} \Lambda_T^{-1}| \leq C \max_{0 \leq i \leq n} |\nabla X_{t_i}| \leq C \max_{0 \leq s \leq T} |\nabla X_s|. \quad (3.23)$$

Thus, combining (3.22) and (3.23) we get

$$\begin{aligned} I_3^1 &= \sum_{i=1}^n E \left\{ |M_T^{t_i}| |\Lambda_{t_{i-1}} - \Lambda_{t_i}| \left(\sum_{j \geq i} |\xi_j| \right) \right\} \\ &\leq C \sum_{i=1}^n E \left\{ |M_T^{t_i}| \sup_{0 \leq t \leq T} |\nabla X_t| \right\} (t_i - t_{i-1}) \\ &\leq CE \left\{ \sup_{0 \leq t \leq T} |M_T^t|^2 + \sup_{0 \leq t \leq T} |\nabla X_t|^2 \right\} \sum_{i=1}^n (t_i - t_{i-1}) \leq C. \end{aligned} \quad (3.24)$$

We now turn to I_3^2 . Define $\tilde{M}_t = \mathcal{E}_t^0(-f_z)$, $t \in [0, T]$ (compare to M in (3.17)!). Again, the boundedness of f_z renders \tilde{M} a P -martingale, and by (3.16) we have

$$M_t^{-1} = \tilde{M}_t \Lambda_t(|f_z|^2) \triangleq \tilde{M}_t \tilde{\Lambda}_t, \quad t \in [0, T], \quad (3.25)$$

where $\tilde{\Lambda}_t \triangleq \Lambda_t(|f_z|^2)$. Now by definition of (3.17) we have

$$M_T^{t_{i-1}} - M_T^{t_i} = M_T \{M_{t_{i-1}}^{-1} - M_{t_i}^{-1}\} = M_T \{\tilde{M}_{t_{i-1}} \tilde{\Lambda}_{t_{i-1}} - \tilde{M}_{t_i} \tilde{\Lambda}_{t_i}\}.$$

Thus we see that the I_3^2 can be written as

$$\begin{aligned} I_3^2 &= \sum_{i=1}^n E \left\{ \left| \sum_{j \geq i} E \{ \xi_j M_T \Lambda_{t_{i-1}} (\tilde{M}_{t_{i-1}} \tilde{\Lambda}_{t_{i-1}} - \tilde{M}_{t_i} \tilde{\Lambda}_{t_i}) | \mathcal{F}_{t_{i-1}} \} \right| \right\} \\ &\leq \sum_{i=1}^n E \left\{ \left| \sum_{j \geq i} E \{ \xi_j M_T \Lambda_{t_{i-1}} \tilde{\Lambda}_{t_{i-1}} (\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i}) | \mathcal{F}_{t_{i-1}} \} \right| \right\} \\ &\quad + \sum_{i=1}^n E \left\{ \left| \sum_{j \geq i} E \{ \xi_j M_T \Lambda_{t_{i-1}} \tilde{M}_{t_i} (\tilde{\Lambda}_{t_{i-1}} - \tilde{\Lambda}_{t_i}) | \mathcal{F}_{t_{i-1}} \} \right| \right\}. \end{aligned} \quad (3.26)$$

Similar to (3.24) we can show that

$$\sum_{i=1}^n E \left\{ \left| \sum_{j \geq i} E \{ \xi_j M_T \Lambda_{t_{i-1}} \tilde{M}_{t_i} (\tilde{\Lambda}_{t_{i-1}} - \tilde{\Lambda}_{t_i}) | \mathcal{F}_{t_{i-1}} \} \right| \right\} \leq C, \quad (3.27)$$

thanks to the boundedness of f_z . Thus it remains to prove that the first term on the right hand side of (3.26) is bounded as well. To see this we note that the boundedness of f_z and (3.15) imply that $M_T \in L^p(\Omega)$ and $\nabla X \in L^p(\mathbf{F}; C([0, T]; \mathbb{R}^{d \times d}))$ for all $p \geq 2$. Therefore for each $p \geq 1$, we use (3.6) to get

$$E \left\{ \sum_{j=1}^n |M_T \xi_j| \right\}^p \leq C E \left\{ |M_T|^p \sup_{0 \leq t \leq T} |\nabla X_t|^p \right\} \leq C. \tag{3.28}$$

In particular, for each j , $M_T \xi_j \in L^2(\mathcal{F}_T)$. Let us now define P -martingales

$$\Gamma_t^j = E\{M_T \xi_j | \mathcal{F}_t\}, \quad t \in [0, T], \quad j = 0, 1, \dots, n. \tag{3.29}$$

Since \tilde{M} is a P -martingale as well, it is easily checked that,

$$\begin{aligned} & \sum_{i=1}^n E \left| \sum_{j \geq i} E\{\xi_j M_T \Lambda_{t_{i-1}} \tilde{\Lambda}_{t_{i-1}} (\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i}) | \mathcal{F}_{t_{i-1}}\} \right| \\ &= \sum_{i=1}^n E \left\{ |\Lambda_{t_{i-1}} \tilde{\Lambda}_{t_{i-1}}| \left| \sum_{j \geq i} E\{\xi_j M_T (\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i}) | \mathcal{F}_{t_{i-1}}\} \right| \right\} \\ &\leq C E \left\{ \sum_{i=1}^n \left| E \left\{ \left[\sum_{j \geq i} (\Gamma_{t_i}^j - \Gamma_{t_{i-1}}^j) \right] (\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i}) | \mathcal{F}_{t_{i-1}} \right\} \right| \right\} \\ &\leq C E \left\{ \sum_{i=1}^n \left| \sum_{j \geq i} (\Gamma_{t_i}^j - \Gamma_{t_{i-1}}^j) \right|^2 + \sum_{i=1}^n (\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i})^2 \right\}. \end{aligned} \tag{3.30}$$

Now, using Itô's formula one shows that the exponential martingale \tilde{M} satisfies

$$E(\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i})^2 = E \int_{t_{i-1}}^{t_i} |f_z(r) \tilde{M}_r|^2 dr \leq C(t_i - t_{i-1}),$$

we have

$$\sum_{i=1}^n E |(\tilde{M}_{t_{i-1}} - \tilde{M}_{t_i})|^2 \leq C. \tag{3.31}$$

On the other hand, since

$$\begin{aligned} & \sum_{i=1}^n E \left| \sum_{j \geq i} (\Gamma_{t_i}^j - \Gamma_{t_{i-1}}^j) \right|^2 \\ &= \sum_{i=1}^n \sum_{j_1, j_2 \geq i} E \int_{t_{i-1}}^{t_i} d[\Gamma^{j_1}, \Gamma^{j_2}]_r = \sum_{j_1, j_2=1}^n \sum_{i \leq j_1 \wedge j_2} E \int_{t_{i-1}}^{t_i} d[\Gamma^{j_1}, \Gamma^{j_2}]_r \\ &= \sum_{j_1, j_2} E \int_0^{t_{j_1} \wedge t_{j_2}} d[\Gamma^{j_1}, \Gamma^{j_2}]_r = \sum_{j_1, j_2} E \{ \langle \Gamma_{t_{j_1} \wedge t_{j_2}}^{j_1}, \Gamma_{t_{j_1} \wedge t_{j_2}}^{j_2} \rangle \} \\ &\leq 2 \sum_{j_1 \leq j_2} E \{ \langle \Gamma_{t_{j_1}}^{j_1}, \Gamma_{t_{j_1}}^{j_2} \rangle \}. \end{aligned} \tag{3.32}$$

Let us define, for each j , a positive martingale associated to Γ^j :

$$|\Gamma_t^j| = E\{|M_T \xi_j| | \mathcal{F}_t\}, \quad t \in [0, T],$$

then $|\Gamma_t^j| \leq |\Gamma_t^j|$, $t \in [0, T]$, $j = 1, \dots, n$, such that

$$\begin{aligned} 2 \sum_{j_1 \leq j_2} E\{\langle \Gamma_{t_{j_1}}^{j_1}, \Gamma_{t_{j_1}}^{j_2} \rangle\} &\leq 2 \sum_{j_1 \leq j_2} E\{|\Gamma_{t_{j_1}}^{j_1}| |\Gamma_{t_{j_1}}^{j_2}|\} \\ &= 2 \sum_{j_1=1}^n E\left\{|\Gamma_{t_{j_1}}^{j_1}| \sum_{j_2=j_1}^n |\Gamma_{t_{j_1}}^{j_2}|\right\} \leq 2 \sum_{j_1=1}^n E\left\{|\Gamma_{t_{j_1}}^{j_1}| \sum_{j_2=1}^n |\Gamma_{t_{j_1}}^{j_2}|\right\} \\ &= 2 \sum_{j_1=1}^n E\left\{|M_T \xi_{j_1}| \sum_{j_2=1}^n |\Gamma_{t_{j_1}}^{j_2}|\right\} \leq 2 \sum_{j_1=1}^n E\{|M_T \xi_{j_1}| \sup_{0 \leq t \leq T} \sum_{j_2=1}^n |\Gamma_t^{j_2}|\} \\ &\leq E\left\{\left(\sum_{j=1}^n |M_T \xi_j|\right)^2 + \left(\sup_{0 \leq t \leq T} \sum_{j=1}^n |\Gamma_t^j|\right)^2\right\} \leq CE\left\{\sum_{j=1}^n |M_T \xi_j|\right\}^2. \end{aligned} \quad (3.33)$$

Here for the last inequality above we used Doob’s inequality (applied to the martingale $\sum_{j=1}^n |\Gamma_t^j|$). Now by (3.28) we see that (3.33) and (3.32) yield that

$$\sum_{i=1}^n E\left|\sum_{j \geq i} (\Gamma_{t_i}^j - \Gamma_{t_{i-1}}^j)\right|^2 \leq CE\left\{\sum_{j=1}^n |M_T \xi_j|\right\}^2 \leq C. \quad (3.34)$$

Now plugging (3.34) and (3.31) into (3.30), then combining with (3.27) and (3.26) we obtain that $I_3^2 \leq C$. This, together with (3.24), shows that $I_3 \leq C$, and hence $I \leq C$. The proof is now complete. \square

Remark 3.2. We should point out that the generic constant C in (3.7) is independent of n and the choice of the partition π . This will be crucial in our future discussion.

4. L^∞ -Lipschitz case

We are now ready to study the path regularity of the adapted solution to an BSDE:

$$Y_t = \xi + \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0, T], \quad (4.1)$$

where ξ is an element in $L^2(\mathcal{F}_T; \mathbb{R})$. Throughout this paper we consider only the case where the terminal value ξ is of the form $\xi = \Phi(X)$, where Φ is some functional from $C([0, T]; \mathbb{R}^d)$ to \mathbb{R} and X is a diffusion, characterized by the following SDE:

$$X_t = x + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \quad t \in [0, T]. \quad (4.2)$$

In this section we are interested in the case where Φ satisfies the following L^∞ -Lipschitz condition:

There exists a constant $L > 0$ such that

$$|\Phi(\mathbf{x}^1) - \Phi(\mathbf{x}^2)| \leq L \sup_{0 \leq s \leq T} |\mathbf{x}^1(s) - \mathbf{x}^2(s)|, \quad \forall \mathbf{x}^1, \mathbf{x}^2 \in C([0, T]; \mathbb{R}^d). \quad (4.3)$$

Our first step is to approximate a functional Φ satisfying (4.3) by a sequence of discrete functionals satisfying (3.6). We proceed as follows. For any partition $\pi : 0 = t_0 < t_1 < \dots < t_n = T$, we define a mapping $\varphi_\pi : C([0, T]; \mathbb{R}^d) \mapsto C([0, T]; \mathbb{R}^d)$ by $\mathbf{x} \mapsto \varphi_\pi(\mathbf{x}) \triangleq \mathbf{x}_\pi$, where

$$\mathbf{x}_\pi(t) \triangleq \frac{1}{t_{i+1} - t_i} [(t_{i+1} - t)\mathbf{x}(t_i) + (t - t_i)\mathbf{x}(t_{i+1})], \quad t \in [t_i, t_{i+1}]. \quad (4.4)$$

Denote $|\pi| = \max_i |t_{i+1} - t_i|$ to be the mesh size of the partition π . Then, using the uniform continuity of \mathbf{x} it is easy to see that $\lim_{|\pi| \rightarrow 0} \sup_{0 \leq t \leq T} |\mathbf{x}_\pi(t) - \mathbf{x}(t)| = 0$. Next, for the given functional Φ we define a new functional Φ_π as

$$\Phi_\pi(\mathbf{x}) \triangleq \Phi(\mathbf{x}_\pi), \quad \forall \mathbf{x} \in C([0, T]; \mathbb{R}^d). \quad (4.5)$$

then by assumption (4.3), one has

$$\begin{aligned} \lim_{|\pi| \rightarrow 0} |\Phi_\pi(\mathbf{x}) - \Phi(\mathbf{x})| &\leq L \lim_{|\pi| \rightarrow 0} \sup_{0 \leq t \leq T} |\mathbf{x}_\pi(t) - \mathbf{x}(t)| = 0, \\ \forall \mathbf{x} \in C([0, T]; \mathbb{R}^d). \end{aligned} \quad (4.6)$$

Now let X be the solution to (4.2), and denote $\xi^\pi \triangleq \Phi_\pi(X)$. Then (4.6) implies that $\xi^\pi \rightarrow \Phi(X)$, P -a.s., as $|\pi| \rightarrow 0$. Moreover, if we denote $X^\pi(\omega) \triangleq \varphi_\pi(X)(\omega)$, then (4.3) leads to

$$|\Phi_\pi(X)| \leq C \left\{ |\Phi(0)| + \sup_{0 \leq s \leq T} |X_s^\pi| \right\} \leq C \left\{ |\Phi(0)| + \sup_{0 \leq s \leq T} |X_s| \right\}.$$

Thus, by the Dominated Convergence Theorem we see that

$$\lim_{|\pi| \rightarrow 0} E|\Phi_\pi(X) - \Phi(X)|^2 = 0. \quad (4.7)$$

Consequently, if one denotes (Y^π, Z^π) to be the adapted solution of (4.1) with $\xi = \Phi(X)$ being replaced by $\xi^\pi \triangleq \Phi_\pi(X)$, then the standard stability result of BSDE tells us that

$$E \left\{ \sup_{0 \leq t \leq T} |Y_t^\pi - Y_t|^2 + \int_0^T |Z_t^\pi - Z_t|^2 dt \right\} \rightarrow 0, \quad \text{as } |\pi| \rightarrow 0. \quad (4.8)$$

To construct the desired family of discrete functionals, we make a further reduction. For the given partition π we define a mapping $\psi_\pi : C([0, T]; \mathbb{R}^d) \mapsto \mathbb{R}^{d(n+1)}$ by

$$\psi_\pi(\mathbf{x}) = (\mathbf{x}(t_0), \mathbf{x}(t_1), \dots, \mathbf{x}(t_n)), \quad \forall \mathbf{x} \in C([0, T]; \mathbb{R}^d). \quad (4.9)$$

Denote $C_\pi([0, T]; \mathbb{R}^d) = \{\mathbf{x}_\pi : \mathbf{x} \in C([0, T]; \mathbb{R}^d)\}$, then $C_\pi([0, T]; \mathbb{R}^d)$ is a subspace of $C([0, T]; \mathbb{R}^d)$, and ψ_π is a 1 - 1 correspondence between $C_\pi([0, T]; \mathbb{R}^d)$ and $\mathbb{R}^{d(n+1)}$. We have the following lemma.

Lemma 4.1. *Suppose that $\Phi \in L^0(C([0, T]; \mathbb{R}^d); \mathbb{R})$ satisfies the L^∞ -Lipschitz condition (4.3). 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 $\pi \in \Pi$, $g_\pi \in C_b^\infty(\mathbb{R}^{d(n+1)})$ and satisfies (3.6), with constant L being the same as that in (4.3), where $n = \#\pi - 1$, and $\#\pi$ denotes the number of partition points in π .*

(ii) *for any $\mathbf{x} \in C([0, T]; \mathbb{R}^d)$, it holds that*

$$\lim_{|\pi| \rightarrow 0} |g_\pi(\psi_\pi(\mathbf{x})) - \Phi_\pi(\mathbf{x})| = 0. \tag{4.10}$$

Proof. Let Φ and $\pi \in \Pi$ be given. Define $G_\pi \triangleq \Phi \circ \psi_\pi^{-1}$, and denote $n = \#\pi - 1$. Then it is easily checked that G_π is a mapping from $\mathbb{R}^{d(n+1)}$ to \mathbb{R} , such that

$$G_\pi(x(t_0), x(t_1), \dots, x(t_n)) = G_\pi(\psi_\pi(\mathbf{x})) = \Phi(\mathbf{x}_\pi), \quad \forall \mathbf{x} \in C([0, T]; \mathbb{R}^d). \tag{4.11}$$

Since Φ satisfies (4.3), one has

$$|G_\pi(x_0, x_1, \dots, x_n) - G_\pi(y_0, y_1, \dots, y_n)| \leq L \max_{0 \leq i \leq n} |x_i - y_i|. \tag{4.12}$$

That is, G_π is (uniform) Lipschitz continuous with Lipschitz constant L being the same as that in (4.3).

Now let $\phi \in C_0^\infty(\mathbb{R}^{d(n+1)})$ be such that $\phi \geq 0$ and $\int_{\mathbb{R}^{d(n+1)}} \phi(z) dz = 1$. For fixed π and $\varepsilon > 0$ we define

$$G_\pi^\varepsilon(x) = \int_{\mathbb{R}^{d(n+1)}} G_\pi(x - \varepsilon z) \phi(z) dz,$$

Then $G_\pi^\varepsilon \in C_b^\infty(\mathbb{R}^{d(n+1)})$, such that $\sup_{x \in \mathbb{R}^{d(n+1)}} |G_\pi^\varepsilon(x) - G_\pi(x)| \rightarrow 0$, as $\varepsilon \rightarrow 0$. Next, for each $\pi \in \Pi$ choose $\varepsilon(\pi)$ such that

$$\sup_{(x_0, x_1, \dots, x_n)} |G_\pi^{\varepsilon(\pi)}(x_0, x_1, \dots, x_n) - G_\pi(x_0, x_1, \dots, x_n)| < |\pi|, \tag{4.13}$$

and define $g_\pi = G_\pi^{\varepsilon(\pi)}$. Then, clearly $g_\pi \in C_b^\infty(\mathbb{R}^{d(n+1)})$; and by definitions of $G_\pi^{\varepsilon(\pi)}$, G_π , and (4.11), for any $\mathbf{x} \in C([0, T]; \mathbb{R}^d)$ we have

$$\begin{aligned} |g_\pi(\psi_\pi(\mathbf{x})) - \Phi_\pi(\mathbf{x})| &= |G_\pi^{\varepsilon(\pi)}(\psi_\pi(\mathbf{x})) - G_\pi(\psi_\pi(\mathbf{x}))| \\ &\leq \sup_{x \in \mathbb{R}^{d(n+1)}} |G_\pi^{\varepsilon(\pi)}(x) - G_\pi(x)| \leq |\pi|. \end{aligned}$$

Namely (4.10) holds, proving (ii).

We now show that g_π satisfies (3.6). Indeed, denoting $\delta_j(x) = \text{sgn}(\partial_j g_\pi(x) y_j)$, $x = (x_0, x_1, \dots, x_n)$ (same for $y, z \in \mathbb{R}^{d(n+1)}$), and by a slight abuse of notation, $y\delta \triangleq (y_0\delta_0(x), y_1\delta_1(x), \dots, y_n\delta_n(x))$ we have

$$\begin{aligned}
 \sum_{j=0}^n |\partial_j g_\pi(x_0, \dots, x_n) y_j| &= \sum_{j=0}^n \partial_j g_\pi(x_0, \dots, x_n) y_j \delta_j \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} (g_\pi(x + hy\delta) - g_\pi(x)) \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^{d(n+1)}} \left\{ G_\pi(x - \varepsilon(\pi)z + hy\delta) - G_\pi(x - \varepsilon(\pi)z) \right\} \phi(z) dz \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^{d(n+1)}} \left[\Phi \circ \psi_\pi^{-1}(x - \varepsilon(\pi)z + hy\delta) - \Phi \circ \psi_\pi^{-1}(x - \varepsilon(\pi)z) \right] \phi(z) dz \\
 &\leq \liminf_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^{d(n+1)}} L \sup_{0 \leq s \leq T} |[\psi_\pi^{-1}(x - \varepsilon(\pi)z + hy\delta) - \psi_\pi^{-1}(x - \varepsilon(\pi)z)](s)| \phi(z) dz \\
 &= \liminf_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^{d(n+1)}} L \max_{0 \leq j \leq n} |h y_j \delta_j| \phi(z) dz = L \max_j |y_j|. \tag{4.14}
 \end{aligned}$$

This proves (i), whence the lemma. □

We now give the main result of this section.

Theorem 4.2. *Assume (A1) and (A2). Assume that the terminal value of the BSDE (4.1) is of the form $\xi = \Phi(X)$, where X satisfies (4.2) and Φ satisfies (4.3). Let (Y, Z) be the (unique) adapted solution of (4.1), then Z admits a càdlàg version.*

Proof. Again, in the sequel we denote all constants that depend only on T, K in (A1) and (A2), and L in (4.3) by a generic one which is allowed to vary from line to line.

Let $\Pi = \{\pi\}$ be the family of all partitions of $[0, T]$; and let $\{g_\pi, \pi \in \Pi\}$ be the family of discrete functionals constructed in Lemma 4.1. Further, let $\{f^\varepsilon, \varepsilon > 0\}$ be a family of mollifiers of f , that is $f^\varepsilon \in C_b^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$ such that

$$\sup_{(t,x,y,z)} |f^\varepsilon(t, x, y, z) - f(t, x, y, z)| \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \tag{4.15}$$

Let $\varepsilon(\pi)$ be the one chosen in (4.13), and define $f^\pi = f^{\varepsilon(\pi)}$. Then it is clear that $f^\pi(t, x, y, z) \rightarrow f(t, x, y, z)$, as $|\pi| \rightarrow 0$, uniformly in (t, x, y, z) .

Now let us consider the following BSDE: for each $\pi : 0 = t_0 < t_1 < \dots < t_n = T$,

$$Y_t^\pi = \xi^\pi + \int_t^T f^\pi(r, X_r, Y_r^\pi, Z_r^\pi) dr - \int_t^T Z_r^\pi dW_r, \quad t \in [0, T], \tag{4.16}$$

where $\xi^\pi = g_\pi(X_{t_0}, \dots, X_{t_n})$ and X satisfies (4.2). Since each g_π satisfies (3.6) with the same constant $L > 0$ of (4.3), we see that

$$\begin{aligned}
 |\xi^\pi|^2 &= |g_\pi(X_{t_0}, \dots, X_{t_n})|^2 \leq 2 \left(|\Phi(0)|^2 + L^2 \max_{0 \leq i \leq n} |X_{t_i}|^2 \right) \\
 &\leq C \left(1 + \sup_{0 \leq t \leq T} |X_t|^2 \right). \tag{4.17}
 \end{aligned}$$

Further, similar to the derivation of (4.7) (recall the notations there), we now apply Lemma 4.1 to get, P -almost surely,

$$|\xi^\pi - \Phi(X)| \leq |g_\pi(\psi_\pi(X)) - \Phi_\pi(X)| + |\Phi_\pi(X) - \Phi(X)| \rightarrow 0, \quad \text{as } |\pi| \rightarrow 0. \tag{4.18}$$

Therefore, the Dominated Convergence Theorem leads to that $E|\xi^\pi - \Phi(X)|^2 \rightarrow 0$, as $|\pi| \rightarrow 0$. Now taking (4.15) into account and applying the standard stability result for BSDEs, we have that

$$E \left\{ \sup_{0 \leq t \leq T} |Y_t^\pi - Y_t|^2 + \int_0^T |Z_t^\pi - Z_t|^2 dt \right\} \rightarrow 0, \quad \text{as } |\pi| \rightarrow 0. \tag{4.19}$$

We now analyze the family $\{Z^\pi\}$. First, recall from the previous section (or [7]) we know that each Z^π has a càdlàg version, we will always take such version from now on. Second, each Z^π has the following representation:

$$Z_t^\pi = \nabla^\pi Y_t^\pi [\nabla X_t]^{-1} \sigma(t, X_t), \quad t \in [0, T], \tag{4.20}$$

where $\nabla^\pi Y^\pi$ is defined similar to (3.5), with Y being replaced by Y^π , and $\nabla^i Y^\pi$ satisfies the following BSDE:

$$\begin{aligned} \nabla^i Y_t^\pi &= \sum_{j \geq i} \partial_j g_\pi \nabla X_{t_j} + \int_t^T \left[\partial_x f_\pi(r) \nabla X_r + \partial_y f_\pi(r) \nabla^i Y_r^\pi + \partial_z f_\pi(r) \nabla^i Z_r^\pi \right] dr \\ &\quad - \left\{ \int_t^T \nabla^i Z_r^\pi dW_r \right\}^T, \quad t \in [t_{i-1}, T], \quad i = 1, \dots, n. \end{aligned} \tag{4.21}$$

We prove that the family $\{\nabla^\pi Y^\pi\}$ is tight. To this end, fix $\pi \in \Pi$, and let $\widehat{\pi} : 0 = s_0 < \dots < s_m = T$ be any partition of $[0, T]$. We shall estimate the conditional variation of $\nabla^\pi Y^\pi$ (see (2.11)):

$$V_{\widehat{\pi}}(\nabla^\pi Y^\pi) = \sum_{i=1}^m E\{ |E\{\nabla^\pi Y_{s_i}^\pi - \nabla^\pi Y_{s_{i-1}}^\pi | \mathcal{F}_{s_{i-1}}\}| + E\{ |\nabla^\pi Y_T^\pi| \} \}. \tag{4.22}$$

To begin with, we note that for any process A , $V_\pi(A) \leq V_{\pi'}(A)$ if $\pi \subseteq \pi'$, here the inclusion means all partition points of π are contained in π' . Indeed, for any $r_1 < r_2 < r_3$ one has

$$\begin{aligned} &E\{ |E\{A_{r_3} - A_{r_1} | \mathcal{F}_{r_1}\}| \} \\ &\leq E\{ |E\{A_{r_3} - A_{r_2} | \mathcal{F}_{r_1}\}| \} + E\{ |E\{A_{r_2} - A_{r_1} | \mathcal{F}_{r_1}\}| \} \\ &= E\{ |E\{E\{A_{r_3} - A_{r_2} | \mathcal{F}_{r_2}\} | \mathcal{F}_{r_1}\}| \} + E\{ |E\{A_{r_2} - A_{r_1} | \mathcal{F}_{r_1}\}| \} \\ &\leq E\{ |E\{A_{r_3} - A_{r_2} | \mathcal{F}_{r_2}\}| \} + E\{ |E\{A_{r_2} - A_{r_1} | \mathcal{F}_{r_1}\}| \}. \end{aligned} \tag{4.23}$$

Namely, the conditional variation increases as the partition gets finer. Therefore without loss of generality we may assume that $\pi \subseteq \widehat{\pi}$ (otherwise we simply consider $\pi \cup \widehat{\pi}$). To be more precise, let us assume that $t_i = s_{\ell_i}$, for $i = 0, \dots, n$. We now

recast the BSDE (4.16) as follows. Define a discrete functional $\tilde{g}_{\hat{\pi}} : \mathbb{R}^{d(m+1)} \mapsto \mathbb{R}$ by

$$\tilde{g}_{\hat{\pi}}(x_0, x_1, \dots, x_m) = g_{\pi}(x_{\ell_0}, \dots, x_{\ell_n}),$$

then $\tilde{g}_{\hat{\pi}}(X_{s_0}, \dots, X_{s_m}) = g_{\pi}(X_{t_0}, \dots, X_{t_n}) = \xi^{\pi}$, and (Y^{π}, Z^{π}) can be viewed as the solution of the BSDE

$$Y_s^{\pi} = \tilde{g}_{\hat{\pi}}(X_{s_0}, \dots, X_{s_m}) + \int_s^T f^{\pi}(r, X_r, Y_r^{\pi}, Z_r^{\pi}) dr - \int_s^T Z_r^{\pi} dW_r. \quad (4.24)$$

Furthermore, since $\partial_{x_j} \tilde{g}_{\hat{\pi}}(x_0, \dots, x_m) = 0$, if $j \notin \{\ell_0, \dots, \ell_n\}$, we have

$$\begin{aligned} \sum_{k=0}^m |\partial_k \tilde{g}_{\hat{\pi}}(x_0, \dots, x_m) y_k| &= \sum_{i=0}^n |\partial_i g_{\pi}(x_{\ell_0}, \dots, x_{\ell_n}) y_{\ell_i}| \leq L \max_i |y_{\ell_i}| \\ &\leq L \max_k |y_k|, \end{aligned}$$

thanks to Lemma 4.1. Thus $\tilde{g}_{\hat{\pi}}$ satisfies (3.6) as well. We now apply Theorem 3.1 to (Y^{π}, Z^{π}) (regarded as the solution to BSDE (4.24)!) to get that

$$\sum_{k=1}^m E \left\{ \left| E \left\{ \nabla^{\hat{\pi}} Y_{s_k}^{\pi} - \nabla^{\hat{\pi}} Y_{s_{k-1}}^{\pi} \middle| \mathcal{F}_{s_{k-1}} \right\} \right| \right\} + E |\nabla^{\hat{\pi}} Y_T^{\pi}| \leq C, \quad (4.25)$$

where $C > 0$ is a constant independent of the choice of partition $\hat{\pi}$,

$$\nabla^{\hat{\pi}} Y_t^{\pi} = \sum_{k=1}^m \widehat{\nabla}^k Y_t^{\pi} 1_{[s_{k-1}, s_k)}(t) + \partial_{x_m} \tilde{g}_{\hat{\pi}}(X_{s_0}, \dots, X_{s_m}) \nabla X_T 1_{\{T\}}(t), \quad t \in [0, T],$$

and

$$\begin{aligned} \widehat{\nabla}^k Y_t^{\pi} &= \sum_{j \geq k} \partial_j \tilde{g}_{\hat{\pi}} \nabla X_{s_j} + \int_t^T \left[\partial_x f_{\pi}(r) \nabla X_r + \partial_y f_{\pi}(r) \widehat{\nabla}^k Y_r^{\pi} + \partial_z f_{\pi}(r) \widehat{\nabla}^k Z_r^{\pi} \right] dr \\ &\quad - \left\{ \int_t^T \widehat{\nabla}^k Z_r^{\pi} dW_r \right\}^T, \quad t \in [s_{k-1}, T], \quad k = 1, \dots, m. \end{aligned} \quad (4.26)$$

Now note that $\partial_j \tilde{g}_{\hat{\pi}}(x_0, \dots, x_m) = 0$ for $j \notin \{\ell_0, \dots, \ell_n\}$. For any $[s_{k-1}, s_k) \subseteq [t_{i-1}, t_i)$ we have

$$\sum_{j \geq k} \partial_j \tilde{g}_{\hat{\pi}} \nabla X_{s_j} = \sum_{j \geq i} \partial_j g_{\pi} \nabla X_{t_j}.$$

Thus, by the uniqueness of the solution to BSDE (4.21) we have $\widehat{\nabla}^k Y_s^{\pi} = \nabla^i Y_s^{\pi}$, $\forall s \in [s_{k-1}, s_k) \subseteq [t_{i-1}, t_i)$. In other words, we have $\nabla^{\hat{\pi}} Y_t^{\pi} = \nabla^{\pi} Y_t^{\pi}$, $t \in [0, T]$, and (4.25) becomes $V_{\hat{\pi}}(\nabla^{\hat{\pi}} Y^{\pi}) \leq C$. Since C is independent of $\hat{\pi}$ and π , and both π and $\hat{\pi}$ are arbitrarily chosen, we obtain that $\sup_{\pi \in \Pi} V(\nabla^{\pi} Y^{\pi}) \leq C$. Consequently, all $\nabla^{\pi} Y^{\pi}$'s are quasi-martingales; and the family $\{\nabla^{\pi} Y^{\pi}\}$ is tight, thanks to Lemma 2.4.

Now denote $\tilde{Z}^\pi = \nabla^\pi Y^\pi$ and $\tilde{Z}_t = Z_t \sigma^{-1}(t, X_t) \nabla X_t$. Since \tilde{Z}^π satisfies (4.20) and $\nabla X \in L^p(\mathbf{F}; C([0, T]; \mathbb{R}^d))$ for all $p \geq 2$, using (4.19) and the Hölder inequality we have for any $1 < q < 2$,

$$\lim_{|\pi| \rightarrow 0} E \int_0^T |\nabla^\pi Y_t^\pi - \tilde{Z}_t|^q dt = \lim_{|\pi| \rightarrow 0} E \int_0^T |(Z_t^\pi - Z_t) \sigma^{-1}(t, X_t) \nabla X_t|^q dt = 0. \tag{4.27}$$

Therefore, we can find a sequence $\{\pi_k\}$ such that outside an exceptional P -null set, for all $\omega \in \Omega$, one has $\int_0^T |\tilde{Z}_t^{\pi_k}(\omega) - \tilde{Z}_t(\omega)|^q dt \rightarrow 0$, as $k \rightarrow \infty$. Thus, as functions in $L^0([0, T])$, $\tilde{Z}^{\pi_k}(\omega)$ converges to $\tilde{Z}(\omega)$ in measure. Applying Lemma 2.3, we see that, as Ψ -valued random variables \tilde{Z}^{π_k} converges to \tilde{Z} in Meyer-Zheng pseudo-path topology, P -a.s., and hence convergence in law. Denote the law of \tilde{Z}^{π_k} by P^k , and that of \tilde{Z} by P^0 .

On the other hand, since $\{\tilde{Z}^{\pi_k}\}$ are quasimartingales with uniformly bounded conditional variations, by Lemma 2.4 we know that, possibly along a subsequence, P^k converges weakly to a probability law $P^* \in \mathcal{M}(\mathbb{D})$. Let $\hat{P}^* \in \mathcal{M}(\Psi)$ be the extension of P^* in the sense of (2.10). The uniqueness of the weak limit then implies that $\hat{P}^*(A) = P^0(A), \forall A \in \mathcal{B}(\Psi)$. Since $\mathbb{D} \in \mathcal{B}(\Psi)$, from (2.10), the definition of P^* , and the equality above we see that

$$1 = P^*(\mathbb{D}) = \hat{P}^*(\mathbb{D}) = P^0(\mathbb{D}) = P\{\tilde{Z} \in \mathbb{D}\}.$$

In other words, \tilde{Z} , whence Z , has paths in \mathbb{D} almost surely. This proves the theorem. □

5. L^1 -Lipschitz case

In this section we consider BSDE (4.1) where Φ satisfies a stronger Lipschitz condition of functional type, which we shall call L^1 -Lipschitz condition. To be more precise, we assume:

There exists a constant $L > 0$, such that

$$|\Phi(\mathbf{x}^1) - \Phi(\mathbf{x}^2)| \leq L \int_0^T |\mathbf{x}^1(s) - \mathbf{x}^2(s)| ds, \quad \forall \mathbf{x}^1, \mathbf{x}^2 \in C([0, T]; \mathbb{R}^d). \tag{5.1}$$

Clearly, the L^1 -Lipschitz condition (5.1) implies the L^∞ -Lipschitz condition (4.3). Therefore if (Y, Z) is the solution to the BSDE (4.1) with Φ satisfying (5.1), then Z at least has a càdlàg version. The main purpose of this section is to prove the following stronger result.

Theorem 5.1. *Assume (A1) and (A2). Assume that the terminal value of the BSDE (4.1) is of the form $\xi = \Phi(X)$, where X satisfies (4.2), and Φ satisfies (5.1). Let (Y, Z) be the (unique) adapted solution of the BSDE (4.1), then Z has a continuous version.*

The proof of Theorem 5.1 is quite lengthy, we shall split it into several lemmas. We begin with some preparations. Let $\Pi = \{\pi\}$ be a family of partitions of $[0, T]$. For a given partition π , let (Y^π, Z^π) be the solution to BSDE (4.16). Recall the process $\tilde{Z}_t \triangleq Z_t \sigma^{-1}(t, X_t) \nabla X_t$. Since Z has a continuous version if and only if \tilde{Z} does, it would suffice to prove that \tilde{Z} has a continuous version, which we shall do in the sequel.

Let us first give a lemma which is a refinement of Lemma 4.1, under the condition (5.1). Recall the mappings φ_π (or \mathbf{x}_π) and ψ_π defined by (4.4) and (4.9), respectively, for a given partition $\pi \in \Pi$.

Lemma 5.2. *Assume that $\Phi : C([0, T]; \mathbb{R}^d) \mapsto \mathbb{R}$ satisfies the Integral Lipschitz condition (5.1). Then there exists a family of discrete functionals $\{g_\pi : \pi \in \Pi\}$ such that*

- (i) for each $\pi \in \Pi$, $g_\pi \in C_b^\infty(\mathbb{R}^{d(n+1)})$, where $n = \#\pi - 1$;
- (ii) for each $\pi \in \Pi$, $0 \leq s_1 < s_2 \leq T$, and $\mathbf{x}, \mathbf{y} \in C([0, T]; \mathbb{R}^d)$, it holds that

$$\sum_{\substack{s_1 < t_j \leq s_2 \\ t_j \in \pi}} |\partial_j g_\pi(\psi_\pi(\mathbf{x}_\pi)) \mathbf{y}(t_j)| \leq 2L \sup_{t \in [0, T]} |\mathbf{y}_t| (|s_2 - s_1| + |\pi|); \tag{5.2}$$

where L is the constant in (5.1);

- (iii) for any $\mathbf{x} \in C([0, T]; \mathbb{R}^d)$, it holds that

$$\lim_{|\pi| \rightarrow 0} |g_\pi(\psi_\pi(\mathbf{x}_\pi)) - \Phi(\mathbf{x}_\pi)| = 0. \tag{5.3}$$

Proof. Let Φ and $\pi \in \Pi$ be given. We construct the family $\{g_\pi\}$ as the same as those in Lemma 4.1. That is, $G_\pi \triangleq \Phi \circ \psi_\pi^{-1}$, and $g_\pi = G_\pi^{\varepsilon(\pi)}$, where $G_\pi^\varepsilon \in C_b^\infty(\mathbb{R}^{d(n+1)})$ is the mollifier of G_π , and $\varepsilon(\pi)$ is chosen so that (4.13) holds.

Since the condition (5.1) implies (4.3), (i) and (iii) follow from Lemma 4.1, and we need only check (ii). To do this let $0 \leq s_1 < s_2 \leq T$, and $\mathbf{x}, \mathbf{y} \in C([0, T]; \mathbb{R}^d)$ be given. Assume that $s_1 \in [t_{j_1-1}, t_{j_1})$ and $s_2 \in [t_{j_2-1}, t_{j_2})$, for some $1 \leq j_1 \leq j_2 \leq n$. Thus

$$|t_{j_2} - t_{j_1-1}| \leq (|s_2 - s_1| + 2|\pi|) \leq 2(|s_2 - s_1| + |\pi|). \tag{5.4}$$

Next, for each j we denote

$$\delta_j = \begin{cases} \text{sgn}[\partial_j g_\pi(\psi_\pi(\mathbf{x}_\pi)) \mathbf{y}(t_j)] & j_1 \leq j < j_2; \\ 0 & \text{otherwise,} \end{cases} \tag{5.5}$$

and $\delta_\pi \triangleq (\delta_0, \dots, \delta_n)$. Also, by a slight abuse of notation we denote $\psi_\pi(\mathbf{y}_\pi) \delta_\pi \triangleq (\mathbf{y}(t_0) \delta_0, \dots, \mathbf{y}(t_n) \delta_n)$, and $\bar{\mathbf{y}}_\pi^\delta \triangleq \psi_\pi^{-1}(\psi_\pi(\mathbf{y}_\pi) \delta_\pi)$. Notice that both ψ_π (whence ψ_π^{-1}) and φ_π are linear mappings, and that $\bar{\mathbf{y}}_\pi^\delta(s) = 0$, for $s \notin [t_{j_1-1}, t_{j_2}]$, then similar to (4.14) we have (with $\varepsilon = \varepsilon(\pi)$)

$$\begin{aligned}
\sum_{s_1 < t_j \leq s_2} |\partial_j g_\pi(\psi_\pi(\mathbf{x}_\pi))\mathbf{y}(t_j)| &= \sum_{j=0}^n \partial_j g_\pi(\psi_\pi(\mathbf{x}_\pi))\mathbf{y}(t_j)\delta_j \\
&= \lim_{h \rightarrow 0} \frac{1}{h} (g_\pi(\psi_\pi(\mathbf{x}_\pi) + h\psi_\pi(\mathbf{y}_\pi)\delta_\pi) - g_\pi(\psi_\pi(\mathbf{x}_\pi))) \\
&= \lim_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^{d(n+1)}} \left[\Phi(\mathbf{x}_\pi - \varepsilon\psi_\pi^{-1}(z) + h\bar{\mathbf{y}}_\pi^\delta) - \Phi(\mathbf{x}_\pi - \varepsilon\psi_\pi^{-1}(z)) \right] \phi(z) dz \\
&\leq \liminf_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^{d(n+1)}} L \int_0^T \left[h|\bar{\mathbf{y}}_\pi^\delta(s)| \right] ds \phi(z) dz \\
&= \int_{\mathbb{R}^{d(n+1)}} L \int_{t_{j-1}}^{t_{j_2}} |\psi_\pi^{-1}(\psi_\pi(\mathbf{y}_\pi)\delta_\pi)(s)| ds \phi(z) dz \leq L \max_j |\mathbf{y}(t_j)\delta_j| |t_{j_2} - t_{j_1-1}| \\
&\leq 2L \sup_{0 \leq t \leq T} |\mathbf{y}(t)| (|s_2 - s_1| + |\pi|),
\end{aligned}$$

thanks to (5.4). This proves (ii), whence the lemma. \square

We now take a closer look at the process $\tilde{Z}^\pi = \nabla^\pi Y^\pi$. Let us introduce some notations similar to those used in §3. Define

$$\begin{cases} \gamma_s^{\pi,0} = \int_s^T [\partial_x f_\pi(r) \nabla X_r + \partial_y f_\pi(r) \gamma_r^{\pi,0} + \partial_z f_\pi(r) \zeta_r^{\pi,0}] dr - \left\{ \int_s^T \zeta_r^{\pi,0} dW_r \right\}^T; \\ \gamma_s^{\pi,i} = \partial_{x_i} g_\pi \nabla X_{t_i} + \int_s^T [\partial_y f_\pi(r) \gamma_r^{\pi,i} + \partial_z f_\pi(r) \zeta_r^{\pi,i}] dr - \left\{ \int_s^T \zeta_r^{\pi,i} dW_r \right\}^T. \end{cases} \quad (5.6)$$

Then, using the linearity of the BSDE (4.21) we see that $\nabla^i Y^\pi$ can be written as

$$\nabla^i Y_s^\pi = \gamma_s^{\pi,0} + \sum_{j \geq i} \gamma_s^{\pi,j}, \quad s \in [t_{i-1}, T], \quad i = 1, \dots, n. \quad (5.7)$$

Now let us recall the “exponentials” Λ and \mathcal{E} defined by (3.13) and (3.14). We denote, for a given $\pi \in \Pi$, $\Lambda_t^\pi \triangleq \Lambda_t^0(\partial_y f_\pi)$ and $M_t^\pi \triangleq \mathcal{E}_t^0(\partial_z f_\pi)$, $t \in [0, T]$. Since $f_\pi \in C_b^{0,1}([0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$, we may assume without loss of generality that $(\partial_x f_\pi, \partial_y f_\pi, \partial_z f_\pi)$ is uniformly bounded. Thus by the Girsanov theorem, M^π is a martingale; and the process $W_t^\pi = W_t - \int_0^t [\partial_z f_\pi(r)]^T dr$, $t \in [0, T]$, is a Brownian motion on the new probability space $(\Omega, \mathcal{F}, P^\pi)$, where $\frac{dP^\pi}{dP} = M_T^\pi$.

Furthermore, by virtue of (3.15) and (3.16) we see that there exist constants $C, C_p > 0$ with $p \geq 1$ such that

$$\begin{cases} \sup_{0 \leq t \leq T} |\Lambda_t^\pi|^p \leq C_p, \quad p \geq 1; \\ E \left\{ \sup_{0 \leq t \leq T} |M_t^\pi|^p \right\} + E \left\{ \sup_{0 \leq t \leq T} |[M_t^\pi]^{-1}|^p \right\} \leq C_p, \quad p \geq 1; \\ |\Lambda_{s_1}^\pi - \Lambda_{s_2}^\pi| \leq C|s_1 - s_2|, \quad \forall s_1, s_2 \in [0, T]. \end{cases} \quad (5.8)$$

Now, using integration by parts we have

$$\Lambda_t^\pi \gamma_t^{\pi,0} = \int_t^T \Lambda_r^\pi \partial_x f_\pi(r) \nabla X_r dr - \left\{ \int_t^T \Lambda_r^\pi \zeta_r^{\pi,0} dW_r^\pi \right\}^T$$

$$\triangleq \Gamma_T^{\pi,0} - \Gamma_t^{\pi,0} - \left\{ \int_t^T \Lambda_r^\pi \zeta^{\pi,0} dW_r^\pi \right\}^T, \quad t \in [0, T], \quad (5.9)$$

where

$$\Gamma_t^{\pi,0} \triangleq \int_0^t \Lambda_r^\pi \partial_x f_\pi(r) \nabla X_r dr, \quad t \in [0, T]. \quad (5.10)$$

Taking conditional expectation $E^\pi\{\cdot|\mathcal{F}_t\}$ on both sides of (5.9) and using the Bayes rule we have

$$\Lambda_t^\pi \gamma_t^{\pi,0} = E^\pi\{\Gamma_T^{\pi,0}|\mathcal{F}_t\} - \Gamma_t^{\pi,0} = E\{M_T^\pi \Gamma_T^{\pi,0}|\mathcal{F}_t\}[M_t^\pi]^{-1} - \Gamma_t^{\pi,0}. \quad (5.11)$$

Similarly, for each i we have

$$\Lambda_t^\pi \gamma_t^{\pi,i} = E^\pi\{\Lambda_T^\pi \partial_i g_\pi \nabla X_{T_i}|\mathcal{F}_t\} = E\{M_T^\pi \Gamma_T^{\pi,i}|\mathcal{F}_t\}[M_t^\pi]^{-1},$$

where

$$\Gamma_t^{\pi,i} \triangleq \Lambda_T^\pi \partial_i g_\pi \nabla X_{T_i}. \quad (5.12)$$

Therefore (5.7) can be written as

$$\begin{aligned} \nabla^i Y_t^\pi &= [\Lambda_t^\pi]^{-1} [\Lambda_t^\pi \gamma_t^{\pi,0} + \sum_{j \geq i} \Lambda_t^\pi \gamma_t^{\pi,j}] \\ &= [\Lambda_t^\pi]^{-1} \left\{ E\{M_T^\pi [\Gamma_T^{\pi,0} + \sum_{j \geq i} \Gamma_{T_j}^{\pi,j}]|\mathcal{F}_t\} [M_t^\pi]^{-1} - \Gamma_t^{\pi,0} \right\} \\ &= E\{\Xi_i^\pi|\mathcal{F}_t\} [\Lambda_t^\pi M_t^\pi]^{-1} - [\Lambda_t^\pi]^{-1} \Gamma_t^{\pi,0}, \quad t \in [t_{i-1}, T], \end{aligned} \quad (5.13)$$

where

$$\Xi_i^\pi \triangleq M_T^\pi [\Gamma_T^{\pi,0} + \sum_{j \geq i} \Gamma_{T_j}^{\pi,j}], \quad i = 0, 1, \dots, n.$$

Consequently, we have

$$\begin{aligned} \nabla^\pi Y_t^\pi &= E\left\{ \sum_{i=1}^n \Xi_i^\pi 1_{[t_{i-1}, t_i)}(t) \middle| \mathcal{F}_t \right\} [\Lambda_t^\pi M_t^\pi]^{-1} - [\Lambda_t^\pi]^{-1} \Gamma_t^{\pi,0} \\ &= E\left\{ \Xi_i^\pi \middle| \mathcal{F}_t \right\} [\Lambda_t^\pi M_t^\pi]^{-1} - [\Lambda_t^\pi]^{-1} \Gamma_t^{\pi,0}, \quad t \in [0, T], \end{aligned} \quad (5.14)$$

where

$$\begin{aligned} \Xi_i^\pi &\triangleq \sum_{i=1}^n \Xi_i^\pi 1_{[t_{i-1}, t_i)}(t) = M_T^\pi [\Gamma_T^{\pi,0} + \sum_{i=1}^n \sum_{j \geq i} \Gamma_{T_j}^{\pi,j} 1_{[t_{i-1}, t_i)}(t)] \\ &= M_T^\pi [\Gamma_T^{\pi,0} + \sum_{j=1}^n \sum_{i \leq j} \Gamma_{T_j}^{\pi,j} 1_{[t_{i-1}, t_i)}(t)] \\ &= M_T^\pi [\Gamma_T^{\pi,0} + \sum_{j=1}^n \Gamma_{T_j}^{\pi,j} 1_{[0, t_j)}(t)], \quad t \in [0, T]. \end{aligned} \quad (5.15)$$

For notational convenience we shall again denote $C > 0$ to be a generic constant depending only on the constant K in (A1) and (A2), L in (5.1), and $T > 0$; and further we denote $\{\chi_\pi : \pi \in \Pi\}$ be a family of generic random variables that may depend on the partition π , such that for all $p \geq 2$,

$$\sup_{\pi \in \Pi} E|\chi_\pi|^p \leq C_p, \tag{5.16}$$

for some constant $C_p > 0$. Note that all C , C_p , and χ are allowed to vary from line to line. Moreover, from now on we shall fix a sequence $\{\pi_n\} \subset \Pi$ such that $\lim_{n \rightarrow \infty} |\pi_n| \rightarrow 0$; and denote $\Psi^n \triangleq \Psi^{\pi_n}$, where $\Psi = \Lambda, M, \tilde{M}, \Xi, \dots$, etc. Furthermore, we denote $\tilde{Z}^n = \nabla^{\pi_n} Y^{\pi_n}$, $f_n = f_{\pi_n}$, $\Gamma^{n,0} = \Gamma^{\pi_n,0}$, $\Gamma^{n,i} = \Gamma^{\pi_n,i}$, and $\chi_n = \chi_{\pi_n}$. We have the following lemma.

Lemma 5.3. *There exists a family of positive random variables $\{\chi_n\}_{n \geq 1}$ satisfying (5.16) such that for all stopping time $\bar{\tau} \in [0, T]$, it holds that $[M_{\bar{\tau}}^n]^{-1} \leq \chi_n$ and $|\Xi_{\bar{\tau}}^n| \leq \chi_n$, $n \geq 1$, P -a.s. Furthermore, for all $0 \leq s_1 < s_2 \leq T$,*

$$|[M_{s_1}^n]^{-1} - [M_{s_2}^n]^{-1}| + |\Xi_{s_1}^n - \Xi_{s_2}^n| \leq \chi_n(|s_1 - s_2|^{\frac{1}{3}} + |\pi_n|), \quad n \geq 1. \tag{5.17}$$

Proof. First, note that (5.8) implies that $[M_{\bar{\tau}}^n]^{-1} \leq \chi_n$. Second, for each n and any $p \geq 2$, by (5.10) and (5.12),

$$\left\{ \begin{aligned} E|\Gamma_T^{n,0}|^p &\leq C_p E \left\{ \int_0^T |\Lambda_r^n \partial_x f_n(r) \nabla X_r|^p dr \right\} \leq C_p E \left\{ \sup_{0 \leq t \leq T} |\nabla X_t|^p \right\}; \\ E \left\{ \sup_{t \in [0, T]} \left| \sum_{j \geq 1} \Gamma_{t_j}^{n,j} 1_{[0, t_j^n)}(t) \right|^p \right\} &\leq C_p E \left\{ \sum_{j \geq 1} |\partial_j g_\pi \nabla X_{t_j^n}|^p \right\} \\ &\leq C_p E \left\{ \sup_{t \in [0, T]} |\nabla X_t|^p \right\}, \end{aligned} \right. \tag{5.18}$$

thanks to the assumption (5.1) (whence (3.6)). Combining this with (5.8) we see that (5.15) yields $|\Xi_{\bar{\tau}}^n| \leq \chi_n$.

To estimate $|[M_{s_1}^n]^{-1} - [M_{s_2}^n]^{-1}|$ we recall from (3.25) that $[M_t^n]^{-1} = \tilde{M}_t^n \Lambda_t^n (|\partial_z f_n|^2)$, where $\tilde{M}^n \triangleq \mathcal{E}^0(-\partial_z f_n)$ is a P -martingale, thanks to the boundedness of $\partial_z f_n$. Now define

$$\tilde{M}_n^* = \sup_{0 \leq r_1 < r_2 \leq T} \frac{|\tilde{M}_{r_2}^n - \tilde{M}_{r_1}^n|}{(r_2 - r_1)^{\frac{1}{3}}}.$$

Using (3.15) it is easy to show that the exponential martingale \tilde{M}^n satisfies, for any $p \geq 1$, that $E\{|\tilde{M}_{r_2}^n - \tilde{M}_{r_1}^n|^{2p}\} \leq C|r_2 - r_1|^p$. Therefore, applying Theorem 1.2.1 of [13] one shows that $\tilde{M}_n^* \in L^p(\Omega)$ for all $p \geq 1$. Consequently,

$$\begin{aligned} &|[M_{s_1}^n]^{-1} - [M_{s_2}^n]^{-1}| \\ &\leq [|\Lambda_{s_1}^n (|\partial_z f_n|^2) - \Lambda_{s_2}^n (|\partial_z f_n|^2)| \tilde{M}_{s_1}^n] + |\Lambda_{s_2}^n (|\partial_z f_n|^2)| [\tilde{M}_{s_1}^n - \tilde{M}_{s_2}^n] \\ &\leq \chi_n[|s_1 - s_2| + |s_1 - s_2|^{\frac{1}{3}}] \leq \chi_n |s_1 - s_2|^{\frac{1}{3}}. \end{aligned}$$

Next, recalling the definitions (5.12) and (5.15), we apply Lemma 5.2 to get

$$\begin{aligned} |\Xi_{s_1}^n - \Xi_{s_2}^n| &\leq |M_T^n| \sum_{j=1}^n |\Lambda_T^n \partial_j g_n \nabla X_{t_j}| 1_{(s_1, s_2]}(t_j) \\ &\leq \chi_n \sup_{0 \leq t \leq T} |\nabla X_t| (|s_1 - s_2| + |\pi_n|) \leq \chi_n (|s_1 - s_2| + |\pi_n|). \end{aligned}$$

Combining the above we derive (5.17). □

Finally, we give a seemingly simple lemma to facilitate our argument in the proof of Theorem 5.1.

Lemma 5.4. *Let $\{\xi_n\}_{n \geq 1}, \{\eta_n\}_{n \geq 1} \subset L^1(\Omega)$ be two sequences such that*

- (i) $|\xi_n| \leq \eta_n, \forall n, P$ -a.s.;
 - (ii) $\lim_{n \rightarrow \infty} \xi_n = \xi$ and $\lim_{n \rightarrow \infty} \eta_n = \eta$, both weakly in $L^1(\Omega)$.
- Then it holds P -almost surely that $|\xi| \leq \eta$.*

Proof. Denote $D \triangleq \{\omega : |\xi| - \eta > 0\}$ and $\rho \triangleq \text{sgn}\{\xi\}$. Then $\rho 1_D \in L^\infty(\Omega)$, and $E\{|\xi|1_D\} = E\{\xi\rho 1_D\} = \lim_{n \rightarrow \infty} E\{\xi_n\rho 1_D\} \leq \liminf_{n \rightarrow \infty} E\{|\xi_n|1_D\} \leq \lim_{n \rightarrow \infty} E\{\eta_n 1_D\} = E\{\eta 1_D\}$.

That is, $E\{[|\xi| - \eta]1_D\} \leq 0$. By definition of the set D we see that $P(D) = 0$ must hold, proving the lemma. □

Proof of Theorem 5.1. As we pointed out before, we need only show that \tilde{Z} has a continuous version on $[0, T]$. Note that Z has a càdlàg version, so does \tilde{Z} . We will take such a version of \tilde{Z} from now on.

We first prove that \tilde{Z} is a.s. continuous on $[0, T_1]$, for all $T_1 < T$. Since \tilde{Z} is already càdlàg, we need only show that for all stopping time $\tau \in (0, T_1]$, it holds that $\tilde{Z}_{\tau-} = \tilde{Z}_\tau$ (cf. [14] or [1]). To this end, we first recall that (4.27) implies that for all $1 < q < 2$,

$$\int_0^T E\{|\tilde{Z}_r^n - \tilde{Z}_r|^q\} dr \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

thus for any stopping time $\bar{\tau}$ such that $0 < \bar{\tau} \leq T_1$, a.s., we have

$$\begin{aligned} E\left\{\int_0^{T-T_1} |\tilde{Z}_{\bar{\tau}+r}^n - \tilde{Z}_{\bar{\tau}+r}|^q dr\right\} &= E\left\{\int_{\bar{\tau}}^{T-(T_1-\bar{\tau})} |\tilde{Z}_r^n - \tilde{Z}_r|^q dr\right\} \\ &\leq E\left\{\int_0^T |\tilde{Z}_r^n - \tilde{Z}_r|^q dr\right\} \rightarrow 0, \quad \text{as } n \rightarrow \infty. \end{aligned}$$

In other words, for a.e. $r \in [0, T - T_1]$, one has

$$E\left\{|\tilde{Z}_{\bar{\tau}+r}^n - \tilde{Z}_{\bar{\tau}+r}|^q\right\} \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{5.19}$$

Next, we note that \mathbf{F} is a Brownian filtration, whence quasi-left continuous. Thus every stopping time $\tau > 0$ is accessible. To wit, there exists a sequence

of stopping times $\{\tau_k\}$ such that $\tau_k < \tau$, and $\tau_k \uparrow \tau$, as $k \rightarrow \infty$. Now setting $\bar{\tau} = \tau_0 \triangleq \tau$ and $\tau_k, k = 1, 2, \dots$, respectively. Taking away a countable union of null sets in $[0, T - T_1]$, we see that (5.19) should hold for $\tau_k, k = 0, 1, \dots$, for a.e. $r \in [0, T - T_1]$. Now let us choose $r_m \downarrow 0$ such that (5.19) holds for all k, m .

Since $\partial_y f_{\pi_n}$ and $\partial_z f_{\pi_n}$ are bounded, using definitions of Λ^n and $\Gamma^{n,0}$ one derives that, for all k and m ,

$$\begin{cases} |\Gamma_{\tau_k+r_m}^{n,0} - \Gamma_{\tau+r_m}^{n,0}| \leq C \sup_{0 \leq t \leq T} |\nabla X_t| |\tau - \tau_k|; \\ |[\Lambda_{\tau_k+r_m}^n]^{-1} - [\Lambda_{\tau+r_m}^n]^{-1}| \leq C |\tau_k - \tau|. \end{cases}$$

Thus, denoting $\rho(\eta, t) \triangleq \eta + E\{\eta | \mathcal{F}_t\}$, $(\eta, t) \in L^2(\Omega) \times [0, T]$, $\tilde{\Xi}_s^n \triangleq \Xi_s^n [\Lambda_s^n M_s^n]^{-1}$, and applying Lemma 5.3 we derive from (5.14) that

$$\begin{aligned} & \left| [\tilde{Z}_{\tau_k+r_m}^n - \tilde{Z}_{\tau+r_m}^n] - [E\{\tilde{\Xi}_{\tau+r_m}^n | \mathcal{F}_{\tau_k+r_m}\} - E\{\tilde{\Xi}_{\tau+r_m}^n | \mathcal{F}_{\tau+r_m}\}] \right| \\ & \leq E \left\{ |\tilde{\Xi}_{\tau_k+r_m}^n - \tilde{\Xi}_{\tau+r_m}^n| | \mathcal{F}_{\tau_k+r_m} \right\} + |[\Lambda_{\tau_k+r_m}^n]^{-1} \Gamma_{\tau_k+r_m}^{n,0} - [\Lambda_{\tau+r_m}^n]^{-1} \Gamma_{\tau+r_m}^{n,0}| \\ & \leq E \left\{ |\Xi_{\tau_k+r_m}^n - \Xi_{\tau+r_m}^n| |[\Lambda_{\tau_k+r_m}^n M_{\tau_k+r_m}^n]^{-1}| | \mathcal{F}_{\tau_k+r_m} \right\} \\ & \quad + E \left\{ |\Xi_{\tau+r_m}^n| |[\Lambda_{\tau_k+r_m}^n]^{-1} - [\Lambda_{\tau+r_m}^n]^{-1}| | [M_{\tau_k+r_m}^n]^{-1}| | \mathcal{F}_{\tau_k+r_m} \right\} \\ & \quad + E \left\{ |\Xi_{\tau+r_m}^n| |[\Lambda_{\tau+r_m}^n]^{-1}| | [M_{\tau_k+r_m}^n]^{-1} - [M_{\tau+r_m}^n]^{-1}| | \mathcal{F}_{\tau_k+r_m} \right\} \\ & \quad + |[\Lambda_{\tau_k+r_m}^n]^{-1}| |\Gamma_{\tau_k+r_m}^{n,0} - \Gamma_{\tau+r_m}^{n,0}| \\ & \quad + |[\Lambda_{\tau_k+r_m}^n]^{-1} - [\Lambda_{\tau+r_m}^n]^{-1}| |\Gamma_{\tau+r_m}^{n,0}| \\ & \leq \rho(\chi_n(|\tau_k - \tau|^{\frac{1}{3}} + |\pi_n|), \tau_k + r_m). \end{aligned} \quad (5.20)$$

To analyze (5.20) we observe that Lemma 5.3 implies that for any stopping time $\bar{\tau} \in (0, T]$, the sequence $\{\tilde{\Xi}_{\bar{\tau}}^n\}_{n \geq 1}$ is bounded (uniformly in $\bar{\tau}$) in $L^2(\Omega)$, thus it is weakly relatively compact in $L^2(\Omega)$, and so is in $L^1(\Omega)$. Consequently, possibly along a subsequence, may assume itself, it holds that

$$\begin{cases} \lim_{n \rightarrow \infty} \tilde{\Xi}_{\tau+r_m}^n = \tilde{\Xi}_m \in L^1(\Omega), & \text{weakly in } L^1(\Omega); \\ \lim_{m \rightarrow \infty} \tilde{\Xi}_m = \tilde{\Xi} \in L^1(\Omega), & \text{weakly in } L^1(\Omega). \end{cases} \quad (5.21)$$

An elementary calculation then shows that, for fixed k and m ,

$$\begin{cases} \lim_{n \rightarrow \infty} E\{\tilde{\Xi}_{\tau+r_m}^n | \mathcal{F}_{\tau_k+r_m}\} = E\{\tilde{\Xi}_m | \mathcal{F}_{\tau_k+r_m}\}, & \text{weakly in } L^1(\Omega); \\ \lim_{n \rightarrow \infty} E\{\tilde{\Xi}_{\tau+r_m}^n | \mathcal{F}_{\tau+r_m}\} = E\{\tilde{\Xi}_m | \mathcal{F}_{\tau+r_m}\}, & \text{weakly in } L^1(\Omega). \end{cases} \quad (5.22)$$

Similarly, since by (5.16) $\{\chi_n\}$ is also bounded in $L^2(\Omega)$, we can also conclude that $\chi_n \rightarrow \chi \in L^2(\Omega)$, weakly in $L^1(\Omega)$ (!), as $n \rightarrow \infty$. Therefore,

$$\begin{cases} \lim_{n \rightarrow \infty} \chi_n(|\tau_k - \tau|^{\frac{1}{3}} + |\pi_n|) = \chi |\tau_k - \tau|^{\frac{1}{3}}; \\ \lim_{n \rightarrow \infty} E\{\chi_n(|\tau_k - \tau|^{\frac{1}{3}} + |\pi_n|) | \mathcal{F}_{\tau_k+r_m}\} = E\{\chi |\tau_k - \tau|^{\frac{1}{3}} | \mathcal{F}_{\tau_k+r_m}\}, \end{cases}$$

both weakly in $L^1(\Omega)$. Let us now denote

$$A_{k,m}^n \triangleq [\tilde{Z}_{\tau_k+r_m}^n - \tilde{Z}_{\tau+r_m}^n] - [E\{\tilde{\Xi}_{\tau+r_m}^n | \mathcal{F}_{\tau_k+r_m}\} - E\{\tilde{\Xi}_{\tau+r_m}^n | \mathcal{F}_{\tau+r_m}\}];$$

$$B_{k,m}^n \triangleq \rho(\chi_n(|\tau_k - \tau|^{\frac{1}{3}} + |\pi_n|), \tau_k + r_m).$$

Then (5.20) shows that $|A_{k,m}^n| \leq B_{k,m}^n$, P -a.s. Further, by (4.27) (or (5.19)) and (5.22) we see that, as $n \rightarrow \infty$,

$$\begin{cases} A_{k,m}^n \rightarrow A_{k,m} \triangleq [\tilde{Z}_{\tau_k+r_m} - \tilde{Z}_{\tau+r_m}] - [E\{\tilde{\Xi}_m | \mathcal{F}_{\tau_k+r_m}\} - E\{\tilde{\Xi}_m | \mathcal{F}_{\tau+r_m}\}]; \\ B_{k,m}^n \rightarrow B_{k,m} \triangleq \rho(\chi|\tau_k - \tau|^{\frac{1}{3}}, \tau_k + r_m), \end{cases} \tag{5.23}$$

both weakly in $L^1(\Omega)$. Applying Lemma 5.4 we obtain that

$$|[\tilde{Z}_{\tau_k+r_m} - \tilde{Z}_{\tau+r_m}] - [E\{\tilde{\Xi}_m | \mathcal{F}_{\tau_k+r_m}\} - E\{\tilde{\Xi}_m | \mathcal{F}_{\tau+r_m}\}]| \leq \rho(\chi|\tau - \tau_k|^{\frac{1}{3}}, \tau_k+r_m), \quad P\text{-a.s.} \tag{5.24}$$

To complete the proof we need to send $m \rightarrow \infty$ in (5.24) and apply Lemma 5.4 again. To this end, for any $\phi \in L^\infty(\Omega)$ we let $\phi_0 = E\{\phi | \mathcal{F}_\tau\}$ and $\phi_m = E\{\phi | \mathcal{F}_{\tau+r_m}\}$. Then using the right-continuity of the filtration \mathbf{F} and the Dominated Convergence Theorem one has $\|\phi_m - \phi_0\|_{L^2(\Omega)} \rightarrow 0$, as $m \rightarrow \infty$. Note that $\{\tilde{\Xi}_m\}$ is bounded in $L^2(\Omega)$ and converges weakly in $L^1(\Omega)$ (see (5.21)), we see that for any $\phi \in L^\infty(\Omega)$, it holds that

$$\begin{aligned} & \left| E\{[E\{\tilde{\Xi}_m | \mathcal{F}_{\tau+r_m}\} - E\{\tilde{\Xi} | \mathcal{F}_\tau\}]\phi\} \right| = |E\{\tilde{\Xi}_m \phi_m - \tilde{\Xi} \phi_0\}| \\ & \leq |E\{[\tilde{\Xi}_m - \tilde{\Xi}]\phi_0\}| + |E\{\tilde{\Xi}_m[\phi_m - \phi_0]\}| \\ & \leq |E\{[\tilde{\Xi}_m - \tilde{\Xi}]\phi_0\}| + \|\tilde{\Xi}_m\|_{L^2(\Omega)} \|\phi_m - \phi_0\|_{L^2(\Omega)} \rightarrow 0, \quad \text{as } m \rightarrow \infty. \end{aligned}$$

That is, $E\{\tilde{\Xi}_m | \mathcal{F}_{\tau+r_m}\} \rightarrow E\{\tilde{\Xi} | \mathcal{F}_\tau\}$, weakly in $L^1(\Omega)$, as $m \rightarrow \infty$. Similarly, we have $E\{\tilde{\Xi}_m | \mathcal{F}_{\tau_k+r_m}\} \rightarrow E\{\tilde{\Xi} | \mathcal{F}_{\tau_k}\}$, and $E\{\chi|\tau - \tau_k|^{\frac{1}{3}} | \mathcal{F}_{\tau_k+r_m}\} \rightarrow E\{\chi|\tau - \tau_k|^{\frac{1}{3}} | \mathcal{F}_{\tau_k}\}$, weakly in $L^1(\Omega)$, as $m \rightarrow \infty$. Furthermore, we define for each integer $\ell \geq 1$ a set

$$\Omega_\ell \triangleq \left\{ \omega \in \Omega : \sup_{0 \leq r \leq r_1} [|\tilde{Z}_{\tau_k+r} - \tilde{Z}_{\tau_k}| + |\tilde{Z}_{\tau+r} - \tilde{Z}_\tau|] \leq \ell \right\},$$

where $r_1 \geq r_m \downarrow 0$, as $m \rightarrow \infty$. Then $\Omega_\ell \uparrow \Omega$, as $\ell \rightarrow \infty$, modulo a P -null set; and for each ℓ , Dominated Convergence Theorem yields that

$$1_{\Omega_\ell} A_{k,m} \rightarrow 1_{\Omega_\ell} \left\{ [\tilde{Z}_{\tau_k} - \tilde{Z}_\tau] - [E\{\tilde{\Xi} | \mathcal{F}_{\tau_k}\} - E\{\tilde{\Xi} | \mathcal{F}_\tau\}] \right\}, \quad \text{weakly in } L^1(\Omega).$$

(see (5.23) for definition of $A_{k,m}$). Since (5.24) implies that $|1_{\Omega_\ell} A_{k,m}| \leq \rho(\chi|\tau - \tau_k|^{\frac{1}{3}}, \tau_k + r_m)$, we can now send $m \rightarrow \infty$ in (5.24) and apply Lemma 5.4 again to get

$$\left| 1_{\Omega_\ell} \left\{ [\tilde{Z}_{\tau_k} - \tilde{Z}_\tau] - [E\{\tilde{\Xi} | \mathcal{F}_{\tau_k}\} - E\{\tilde{\Xi} | \mathcal{F}_\tau\}] \right\} \right| \leq \rho(\chi|\tau_k - \tau|^{\frac{1}{3}}, \tau_k), \quad P\text{-a.s.} \tag{5.25}$$

Finally, first letting $\ell \rightarrow \infty$ and then taking expectation and letting $k \rightarrow \infty$ on both sides of (5.25), using the fact that \mathbf{F} is *quasi-left continuous*, and applying Fatou's Lemma, we conclude that $E|\tilde{Z}_{\tau-} - \tilde{Z}_{\tau}| \leq 0$. That is, $\tilde{Z}_{\tau-} = \tilde{Z}_{\tau}$, P -a.s. Since τ is arbitrary, \tilde{Z} (whence Z) is continuous on $[0, T_1]$, for all $T_1 < T$. That is, Z is continuous on $[0, T)$. Defining $Z_T = Z_{T-}$, we see that Z is continuous on $[0, T]$. The proof is complete. \square

The following theorem is a direct consequence of Theorem 4.2 and Theorem 5.1.

Theorem 5.5. *Assume (A1) and (A2), and assume that the terminal value of BSDE (4.1) is of the form $\xi = \Phi(X)$ where, for some $0 \leq t_1 < t_2 \leq T$, Φ satisfies that*

$$|\Phi(x^1) - \Phi(x^2)| \leq L \left(\int_{t_1}^{t_2} |x^1(t) - x^2(t)| dt + \sup_{t \in [0, T] - (t_1, t_2)} |x^1(t) - x^2(t)| \right).$$

Then the martingale part of the solution to (4.1), Z , has a version that is càdlàg on $[0, T]$ and continuous in $[t_1, t_2)$.

Proof. By Theorem 4.2, Z is càdlàg. Restricting the stopping time τ in (t_1, t_2) and following the same argument as that of Theorem 5.1 one shows that Z is continuous in $[t_1, t_2)$. \square

In particular, we have the following result proved in our previous paper [7].

Corollary 5.6. *If in BSDE(4.1) the terminal value is of the form $\xi = g(X_T)$, where $g \in W^{1,\infty}(\mathbb{R}^d)$, then the martingale part of the solution to (4.1), Z , is continuous on $[0, T]$.*

Proof. By Theorem 5.5 we know that Z is càdlàg on $[0, T]$ and continuous in $[0, T)$. Letting $Z_T = Z_{T-}$, we see that Z is indeed continuous on $[0, T]$. \square

Acknowledgements. We would like to thank the anonymous referee for his/her careful reading of the manuscript and many helpful suggestions.

References

1. Dellacherie, C., Meyer, P.A.: *Probabilités et Potentiel*, Volumes A and B, Hermann, Paris, North Holland, Amsterdam (1982)
2. Jacod, J., Méléard, S., Protter, P.: *Explicit form and robustness of martingale representations*, Preprint (1999)
3. Karatzas, I., Shreve, S.E.: *Brownian Motion and Stochastic Calculus*, Springer (1987)
4. Ma, J., Protter, P., San Martin, J., Torres, S.: *Numerical method for backward SDEs*, Preprint (2000)
5. Ma, J., Protter, P., Yong, J.: *Solving forward-backward stochastic differential equations explicitly - a four step scheme*, *Probab. Theory Relat. Fields.*, **98**, 339–359 (1994)
6. Ma, J., Yong, J.: *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Math., **1702**, Springer (1999)
7. Ma, J., Zhang, J.: *Representation theorems for backward SDE's*, Preprint (1999)

8. Meyer, P., Zheng, W.: *Tightness criteria for laws of semimartingales*, *Ann. Inst. Henri Poincaré*, Vol. 20 **4**, 353–372 (1984)
9. Nualart, D.: *The Malliavin Calculus and Related Topics*, Springer (1995)
10. Pardoux, E., Peng S.: *Adapted solutions of backward stochastic equations*, *System and Control Letters*, **14**, 55–61 (1990)
11. Pardoux, E., Peng S.: *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, *Lecture Notes in CIS*, Springer, **176**, 200–217 (1992)
12. Protter, P.: *Integration and Stochastic Differential Equations*, Springer (1990)
13. Revuz, D., Yor, M.: *Brownian Motion and Continuous Martingales*, Springer (1991)
14. Rogers, L. C. G., Williams, D.: *Diffusions, Markov Processes, and Martingales, Volume 2: Itô Calculus*, John Wiley & Sons (1987)