Chin. Ann. Math. Ser. B
46(6), 2025, 875–910
DOI: 10.1007/s11401-025-0051-y

Chinese Annals of Mathematics, Series B
© The Editorial Office of CAM and
Springer-Verlag Berlin Heidelberg 2025

Probabilistic Interpretation for a System of Quasilinear Parabolic Partial Differential-Algebraic Equations: The Classical Solution*

Zhen WU[1] Bing XIE[2] Zhiyong YU[3]

Abstract In the present paper, by introducing a family of coupled forward-backward stochastic differential equations (FBSDEs for short), a probabilistic interpretation for a system consisting of m second order quasilinear (and possibly degenerate) parabolic partial differential equations and $(m \times d)$ algebraic equations is given in the sense of the classical solution. For solving the problem, an L^p-estimate ($p > 2$) for coupled FBSDEs on large time durations in the monotonicity framework is established, and a new method to analyze the regularity of solutions to FBSDEs is introduced. The new method avoids the use of Kolmogorov's continuity theorem and only employs L^2-estimates and L^4-estimates to obtain the desired regularity.

Keywords Forward-backward stochastic differential equation, Monotonicity condition, Parabolic partial differential equation, Classical solution
2020 MR Subject Classification 60H10, 35K59, 35C99

1 Introduction

It is classical that a system of first order semilinear partial differential equations (PDEs for short) can be solved via the method of characteristic curves (see Courant and Hilbert [4]). The well known Feynman-Kac formula provides a probabilistic interpretation for a kind of linear second order PDEs of elliptic or parabolic types. With the help of the theory of backward stochastic differential equations (BSDEs for short), researchers have given probabilistic interpretations for some semilinear second order PDEs, see Peng [23], Pardoux and Peng [19], Barles, Buckdahn and Pardoux [2], Darling and Pardoux [5], Pardoux, Pradeilles and Rao [20], Pardoux [18], Kobylanski [11], Zhang and Zhao [33], Pardoux and Răşcanu [21], and so on. Along this line, the next natural problem arises: Which kind of PDEs' probabilistic interpretation should be given by the coupled forward-backward stochastic differential equations (FBSDEs for short)

Manuscript received September 16, 2022. Revised May 17, 2023.
[1]School of Mathematics, Shandong University, Jinan 250100, China. E-mail: wuzhen@sdu.edu.cn
[2]School of Mathematics and Statistics, Shandong University, Weihai 264209, Shandong, China.
E-mail: xiebing@sdu.edu.cn
[3]Corresponding author. School of Mathematics, Shandong University, Jinan 250100, China.
E-mail: yuzhiyong@sdu.edu.cn
*This work was supported by the National Key Research and Development Program of China (No. 2023YFA1009200), the National Natural Science Foundation of China (Nos. 12521001, 12271304, 62561160159), the Natural Science Foundation of Shandong Province (No. ZR2024ZD35) and the Taishan Scholars Climbing Program of Shandong Province (No. TSPD20210302).

and in what sense (see Peng [24])? Until now, there have been a few results on this problem. Pardoux and Tang [22] introduced a kind of coupled FBSDEs and provided a probabilistic interpretation for quasilinear parabolic PDEs in the sense of the viscosity solution. Recently, Feng, Wang and Zhao [8] studied the probabilistic interpretation for a system of quasilinear parabolic and elliptic PDEs in the senses of both classical solutions and Sobolev space weak solutions.

We notice that, in [8, 22], the authors considered the case σ in FBSDEs (1.2) does not depend on the variable z. If σ depends on z, the corresponding quasilinear PDE should be combined with some algebraic equations to form a system. We call it a system of partial differential-algebraic equations (PDAEs for short). Wu and Yu [27] introduced a system of second order quasilinear (and possibly degenerate) parabolic PDAEs, and the issue of a probabilistic interpretation for it was studied in the sense of the viscosity solution. It should also be noticed that, due to the nature of the viscosity solution, the dimension of PDEs in the system was restricted to be 1 in [27]. In the present paper, we continue to investigate the probabilistic interpretation for the PDAE system with multidimensional PDEs in the sense of the classical solution.

Precisely, the following PDAE system will be considered in this paper:

$$\begin{cases} \partial_t u(t,x) + (\mathcal{L}u)(t,x,u(t,x),v(t,x)) + g(t,x,u(t,x),v(t,x)) = 0, \\ \hfill (t,x) \in [0,T] \times \mathbb{R}^n, \\ v(t,x) = \nabla u(t,x)\sigma(t,x,u(t,x),v(t,x)), \quad (t,x) \in [0,T] \times \mathbb{R}^n, \\ u(T,x) = \Phi(x), \quad x \in \mathbb{R}^n, \end{cases} \tag{1.1}$$

where (u,v) is a pair of unknown functions. In (1.1), $u = (u^1, u^2, \cdots, u^m)^{\mathrm{T}}$ takes values in $\mathbb{R}^m$ and its gradient is denoted by

$$\nabla u = \begin{pmatrix} \nabla u^1 \\ \nabla u^2 \\ \vdots \\ \nabla u^m \end{pmatrix} = \begin{pmatrix} \frac{\partial u^1}{\partial x_1} & \frac{\partial u^1}{\partial x_2} & \cdots & \frac{\partial u^1}{\partial x_n} \\ \frac{\partial u^2}{\partial x_1} & \frac{\partial u^2}{\partial x_2} & \cdots & \frac{\partial u^2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial u^m}{\partial x_1} & \frac{\partial u^m}{\partial x_2} & \cdots & \frac{\partial u^m}{\partial x_n} \end{pmatrix},$$

$\mathcal{L}u = (Lu^1, Lu^2, \cdots, Lu^m)^{\mathrm{T}}$ and L is an infinitesimal operator defined by

$$(L\phi)(t,x,y,z) := \frac{1}{2}\sum_{i,j=1}^{n} (\sigma\sigma^{\mathrm{T}})_{ij}(t,x,y,z)\frac{\partial^2 \phi}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{n} b_i(t,x,y,z)\frac{\partial \phi}{\partial x_i}(t,x)$$

for any smooth function ϕ. Besides the PDAE system (1.1), we shall also introduce a family of coupled FBSDEs parameterized by the initial pairs $(t,x) \in [0,T] \times \mathbb{R}^n$ as follows:

$$\begin{cases} \mathrm{d}X_s^{t,x} = b(s,\Theta_s^{t,x})\mathrm{d}s + \sigma(s,\Theta_s^{t,x})\mathrm{d}W_s, \quad s \in [t,T], \\ -\mathrm{d}Y_s^{t,x} = g(s,\Theta_s^{t,x})\mathrm{d}s - Z_s^{t,x}\mathrm{d}W_s, \quad s \in [t,T], \\ X_t^{t,x} = x, \quad Y_T^{t,x} = \Phi(X_T^{t,x}), \end{cases} \tag{1.2}$$

where we denote $\Theta^{t,x} = (X^{t,x}, Y^{t,x}, Z^{t,x})$ for simplicity. By FBSDEs (1.2), under suitable conditions, we shall define a pair of functions taking values in $\mathbb{R}^m$ and $\mathbb{R}^{m\times d}$, respectively:

$$u(t,x) := Y_t^{t,x}, \quad v(t,x) := Z_t^{t,x}, \quad (t,x) \in [0,T] \times \mathbb{R}^n. \tag{1.3}$$

The main result of this paper (see Theorem 4.1) is that under certain assumptions (u,v) defined by (1.3) is the unique classical solution to the PDAE system (1.1).

There exist three fundamental methods to investigate coupled FBSDEs on arbitrarily large time intervals: The method of contraction mapping (see Pardoux and Tang [22]), the four-step scheme approach (see Ma, Protter and Yong [15]) and the method of continuation (see Hu and Peng [10], Yong [30] and Peng and Wu [25]). Some recent developments on FBSDEs can be found in Yong [31] and Ma, Wu, Zhang and Zhang [16]. Compared with the other two methods, the third method has the advantage of dealing with the case: σ depends on Z. As we mentioned before, in this case, some algebraic equations will be involved into the corresponding probabilistic interpretation problem, which is just the feature of this paper. Moreover, the third method is also good at dealing with possibly degenerate diffusion coefficients σ. Due to these reasons, in this paper, we shall work in the related monotonicity framework (see Assumption (H3) in the next section) which is required by the method of continuation. Especially, we shall start with a standard result (see Lemma 2.2) on FBSDEs in this framework.

As an elementary analysis tool, the L^p-theory (including L^p-solutions and the related L^p-estimates) ($p \geq 2$) of FBSDEs will play a key role in the analysis of the probabilistic interpretation. In the literature, when $p = 2$, the L^2-theory of FBSDEs within monotonicity framework is standard (see Lemma 2.2). In comparison, when $p > 2$, the results are rare. For the L^p-results on small intervals, one can refer to Delarue [6], Li and Wei [12–13] and Xie and Yu [28]. On large intervals, Ma, Wu, Zhang and Zhang [16] provided an L^p-result for 1-dimensional FBSDEs. Feng, Wang and Zhao [8] established an L^p-result when both b and σ are independent of z. Recently, Hu, Ji and Xue [9] gave another L^p-result. In these results, the Lipschitz constants of both b and σ with respect to z are assumed to be very small.

In the present paper, we shall establish an L^p-result on a large time interval in the monotonicity framework (see Theorem 2.1). When they researched the L^2-theory of FBSDEs in [6, 14, 16], the authors used the following idea: "Splicing" a sequence of results on small intervals yields a corresponding result on a large interval. In this paper, we shall adopt this idea to investigate the L^p-theory with $p > 2$. In detail, we shall first establish L^p-results on small intervals. Then by the classical standard L^2-theory, we "splice" them to obtain a desired L^p-result on a large interval. However, due to the high degree of coupling of FBSDEs, even on small intervals, the L^p-result needs some additional assumptions (see Assumption $(\text{H2})^p$). It is clear that this L^p-result has a wide range of potential applications. For example, besides the issue of the probability interpretation, the L^p-theory is also necessary in the study of general maximum principle for controlled coupled FBSDEs (see [9]). The L^p-result with $p > 2$ can be regarded as a contribution of this paper.

Regularity analysis for the solutions to FBSDEs (1.2) is an important part of the probabilistic interpretation of the PDAE system (1.1). The classical method makes use of Kolmogorov's continuity theorem as well as L^p-estimates (for all $p \geq 2$) (see [8, 19]). For the case of the coupled FBSDEs in the monotonicity framework studied in this paper, the classical method will still work due to the establishment of the L^p-theory. However, as we mentioned in the above paragraph, some additional assumptions must be imposed to ensure the feasibility of L^p-estimates. Moreover, we notice that, with different $p > 2$, Assumption $(\text{H2})'^p$ (depending on p) is in fact a series of assumptions. In order to weaken assumptions, we introduce a new method to analyze regularity in this paper. The new method employs Lebesgue's dominated convergence theorem to prove the desired convergence many times, instead of the use of Kolmogorov's continuity theorem, then only L^2-estimates and L^4-estimates are involved in our analysis (see Theorem 3.1). The new method can also be applied to other probability interpretation problems, and can be regarded as another contribution of this paper.

The rest of this paper is organized as follows. In Section 2, we establish an L^p-result for coupled FBSDEs in the monotonicity framework. We also recall some elementary properties of the function u from [27]. Section 3 is devoted to the regularity analysis for the solutions to FBSDEs (1.2) including the Malliavin's differentiability of $\Theta^{t,x} = (X^{t,x}, Y^{t,x}, Z^{t,x})$ and the second-order continuous differentiability of u with respect to x. In Section 4, we prove that the function u is continuous differentiable with respect to t, and (u, v) is the unique classical solution to the PDAE system (1.1).

2 L^p-theory in Monotonicity Framework

In this paper, we work with a finite time horizon $T > 0$, a d-dimensional standard Brownian motion $(W_s)_{s\in[0,T]}$, a completed probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a filtration $\mathbb{F} = (\mathcal{F}_s)_{s\in[0,T]}$ which is a natural one generated by the progressively measurable processes on $\Omega \times [0, T]$. For simplicity, we omit all dependence in ω of any random variable or stochastic process in the notations.

We denote by $\mathbb{R}^k$ the k-dimensional Euclidean space with the inner product (resp. norm) $\langle \cdot,\ \cdot \rangle$ (resp. $|\cdot|$), $\mathbb{R}^{k\times l}$ the collection of $(k \times l)$ matrices with the inner product (resp. norm) $\langle z,\ \overline{z} \rangle = \text{tr}\ (z\overline{z}^{\text{T}})$ (resp. $|z| = \sqrt{\text{tr}\ (zz^{\text{T}})}$) for any $z, \overline{z} \in \mathbb{R}^{k\times l}$, where the superscript T denotes the transpose of vectors or matrices. For any $p, q \in [1, \infty)$ and any given $\mathbb{F}$-stopping time τ, we introduce some Banach (or Hilbert, in case of $L^2_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^k)$ or $L^2_{\mathbb{F}}(\tau,T;\mathbb{R}^k)$) spaces of random variables or stochastic processes as follows:

(1) $L^p_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^k)$ is the space of $\mathbb{R}^k$-valued $\mathcal{F}_\tau$-measurable random variables ξ such that

$$\|\xi\|_{L^p_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^k)} := (\mathbb{E}[|\xi|^p])^{\frac{1}{p}} < \infty.$$

(2) $L^p_{\mathbb{F}}(\Omega;C([\tau,T];\mathbb{R}^k))$ is the space of $\mathbb{R}^k$-valued $\mathbb{F}$-progressively measurable processes φ such that for almost all $\omega\in\Omega$, $s\mapsto\varphi(s,\omega)$ is continuous and

$$\|\varphi\|_{L^p_{\mathbb{F}}(\Omega;C([\tau,T];\mathbb{R}^k))} := \Big(\mathbb{E}\Big[\sup_{s\in[\tau,T]}|\varphi_s|^p\Big]\Big)^{\frac{1}{p}} < \infty.$$

(3) $L^p_{\mathbb{F}}(\Omega;L^q(\tau,T;\mathbb{R}^k))$ is the space of $\mathbb{R}^k$-valued $\mathbb{F}$-progressively measurable processes φ such that

$$\|\varphi\|_{L^p_{\mathbb{F}}(\Omega;L^q(\tau,T;\mathbb{R}^k))} := \Big(\mathbb{E}\Big[\Big(\int_\tau^T|\varphi_s|^q\mathrm{d}s\Big)^{\frac{p}{q}}\Big]\Big)^{\frac{1}{p}} < \infty.$$

When $p=q$, we denote $L^p_{\mathbb{F}}(\tau,T;\mathbb{R}^k) := L^p_{\mathbb{F}}(\Omega;L^p(\tau,T;\mathbb{R}^k))$.

Moreover, the following Banach spaces are also introduced:

(1) $M^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{n+m+m\times d}) := L^p_{\mathbb{F}}(\Omega;C([\tau,T];\mathbb{R}^n)) \times L^p_{\mathbb{F}}(\Omega;C([\tau,T];\mathbb{R}^m)) \times L^p_{\mathbb{F}}(\Omega;L^2(\tau,T;\mathbb{R}^{m\times d}))$. For any $\Theta=(X,Y,Z)\in M^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{n+m+m\times d})$, its norm is given by

$$\|\Theta\|_{M^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{n+m+m\times d})} := \Big\{\mathbb{E}\Big[\sup_{s\in[\tau,T]}|X_s|^p + \sup_{s\in[\tau,T]}|Y_s|^p + \Big(\int_\tau^T|Z_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big]\Big\}^{\frac{1}{p}}.$$

(2) $\mathcal{M}^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{m+n+n\times d}) := L^p_{\mathbb{F}}(\Omega;L^1(\tau,T;\mathbb{R}^m)) \times L^p_{\mathbb{F}}(\Omega;L^1(\tau,T;\mathbb{R}^n)) \times L^p_{\mathbb{F}}(\Omega;L^2(\tau,T;\mathbb{R}^{n\times d}))$. For any $\gamma=(g,b,\sigma)\in\mathcal{M}^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{m+n+n\times d})$, its norm is given by

$$\|\gamma\|_{\mathcal{M}^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{m+n+n\times d})} := \Big\{\mathbb{E}\Big[\Big(\int_\tau^T|g_s|\mathrm{d}s\Big)^p + \Big(\int_\tau^T|b_s|\mathrm{d}s\Big)^p + \Big(\int_\tau^T|\sigma_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big]\Big\}^{\frac{1}{p}}.$$

In the case without causing confusion, we sometimes omit the value spaces in the notations for simplicity. For example, $M^p_{\mathbb{F}}(\tau,T;\mathbb{R}^{n+m+m\times d})$ is sometimes abbreviated as $M^p_{\mathbb{F}}(\tau,T)$.

2.1 L^p-results for coupled FBSDEs

Let us have two mappings: $\Phi:\Omega\times\mathbb{R}^n\to\mathbb{R}^m$ and $\gamma=(g,b,\sigma)$ where

$$\begin{aligned} &g:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\to\mathbb{R}^m,\\ &b:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\to\mathbb{R}^n,\\ &\sigma:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\to\mathbb{R}^{n\times d}. \end{aligned}$$

We assume that, for any $\theta=(x,y,z)\in\mathbb{R}^{n+m+m\times d}$, the random variable $\Phi(x)$ and the stochastic process $\gamma(\cdot,\theta)=(g(\cdot,\theta),b(\cdot,\theta),\sigma(\cdot,\theta))$ are $\mathcal{F}_T$-measurable and $\mathbb{F}$-progressively measurable, respectively. Moreover, we assume the following Lipschitz condition.

(H1) Φ and γ are uniformly Lipschitz continuous with respect to x and θ, respectively. For convenience, the Lipschitz constant of Φ with respect to x is denoted by L_x, the Lipschitz constant of σ with respect to z is denoted by L_z, and the other Lipschitz constants are denoted by L.

Let $\tau \in [0,T]$ be an $\mathbb{F}$-stopping time and $\zeta \in L^2_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^n)$. We introduce a coupled FBSDE as follows:

$$\begin{cases} \mathrm{d}X_s^{\tau,\zeta} = b(s,\Theta_s^{\tau,\zeta})\mathrm{d}s + \sigma(s,\Theta_s^{\tau,\zeta})\mathrm{d}W_s, \quad s\in[\tau,T], \\ -\mathrm{d}Y_s^{\tau,\zeta} = g(s,\Theta_s^{\tau,\zeta})\mathrm{d}s - Z_s^{\tau,\zeta}\mathrm{d}W_s, \quad s\in[\tau,T], \\ X_\tau^{\tau,\zeta} = \zeta, \quad Y_T^{\tau,\zeta} = \Phi(X_T^{\tau,\zeta}), \end{cases} \tag{2.1}$$

in which we denote $\Theta^{\tau,\zeta} = (X^{\tau,\zeta}, Y^{\tau,\zeta}, Z^{\tau,\zeta})$. The special case (1.2) we are most interested in is that the initial pair (τ,ζ) is deterministic, i.e., $\tau = t \in [0,T]$ and $\zeta = x \in \mathbb{R}^n$.

It is known that, due to the nature of the equation, only with the uniform Lipschitz condition (H1) for the coefficients (Φ,γ), FBSDE (2.1) does not necessarily have an adapted solution on a large enough duration (a counterexample can be found in Antonelli [1]). Next, we shall provide some preliminary results for FBSDEs on small durations. For this aim, let us introduce two constants and an assumption.

Let $p > 2$ be given. By the theories of stochastic differential equations (SDEs for short) and backward stochastic differential equations (BSDEs for short), there exist a pair of constants $C_F = C_F(p,T,L) > 0$ (see [32, Theorems 3.3.1 and 3.4.3]) and $C_B = C_B(p,T,L) > 0$ (see [32, Theorems 4.3.1 and 4.4.4]) such that, for any $x\in\mathbb{R}^n$, any $\xi \in L^p_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)$, any standard L^p-generator (b_0,σ_0) of SDE[1] with Lipschitz constant L, and any standard L^p-generator g_0 of BSDE[2] with Lipschitz constant L, the following estimates

$$\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,T]}|X_s|^p\Big] \le C_F\mathbb{E}^{\mathcal{F}_t}\Big\{|x|^p + \Big(\int_t^T |b_0(s,0)|\mathrm{d}s\Big)^p + \Big(\int_t^T |\sigma_0(s,0)|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big\}, \tag{2.2}$$

$$\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,T]}|Y_s|^p + \Big(\int_t^T |Z_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big] \le C_B\mathbb{E}^{\mathcal{F}_t}\Big\{|\xi|^p + \Big(\int_t^T |g_0(s,0,0)|\mathrm{d}s\Big)^p\Big\} \tag{2.3}$$

hold for all $t\in[0,T]$, where $\mathbb{E}^{\mathcal{F}_t}[\cdot] \equiv \mathbb{E}[\cdot|\mathcal{F}_t]$ denotes the conditional expectation operator with respect to $\mathcal{F}_t$, and $X \in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ is the unique solution to the following SDE:

$$\begin{cases} \mathrm{d}X_s = b_0(s,X_s)\mathrm{d}s + \sigma_0(s,X_s)\mathrm{d}W_s, \quad s\in[0,T], \\ X_0 = x \end{cases} \tag{2.4}$$

and $(Y,Z) \in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^m)) \times L^p_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^{m\times d}))$ is the unique solution to the following BSDE:

$$\begin{cases} -\mathrm{d}Y_s = g_0(s,Y_s,Z_s)\mathrm{d}s - Z_s\mathrm{d}W_s, \quad s\in[0,T], \\ Y_T = \xi. \end{cases} \tag{2.5}$$

[1] A standard L^p-generator of SDE (on the interval $[0,T]$) is a pair of mappings $b_0 : \Omega\times[0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma_0 : \Omega\times[0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times d}$ satisfying the following three conditions: (i) For any $x\in\mathbb{R}^n$, the stochastic processes $b_0(\cdot,x)$ and $\sigma_0(\cdot,x)$ are $\mathbb{F}$-progressively measurable; (ii) $b_0(\cdot,0) \in L^p_{\mathbb{F}}(\Omega;L^1(0,T;\mathbb{R}^n))$ and $\sigma_0(\cdot,0) \in L^p_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^{n\times d}))$; (iii) b_0 and σ_0 are uniformly Lipschitz continuous with respect to x.

[2] Similarly, a standard L^p-generator of BSDE (on the interval $[0,T]$) is a mapping $g_0 : \Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d} \to \mathbb{R}^m$ satisfying the following three conditions: (i) For any $(y,z)\in\mathbb{R}^m\times\mathbb{R}^{m\times d}$, the stochastic process $g_0(\cdot,y,z)$ is $\mathbb{F}$-progressively measurable; (ii) $g_0(\cdot,0,0) \in L^p_{\mathbb{F}}(\Omega;L^1(0,T;\mathbb{R}^m))$; (iii) g_0 is uniformly Lipschitz continuous with respect to (y,z).

(H2)$'^p$ For the given $p > 2$, the inequality $C_F(p,T,L) \cdot C_B(p,T,L) \cdot L_x^p \cdot L_z^p < 1$ holds.

Remark 2.1 The values of the constants $C_F(p,T,L) > 0$ and $C_B(p,T,L) > 0$ can be obtained by the theories of SDEs and BSDEs. More accurate estimation techniques of SDEs and BSDEs will lead to smaller values of C_F and C_B. Obviously, there are two special cases covered by Assumption (H2)$'^p$. One of them is $L_z = 0$, i.e., σ is independent of z, and the other one is $L_x = 0$, i.e., Φ degenerates to be an $\mathcal{F}_T$-measurable random variable.

Let $p > 2$ and $0 \le t < t' \le T$. Let $\Psi : L^p_{\mathcal{F}_{t'}}(\Omega;\mathbb{R}^n) \to L^p_{\mathcal{F}_{t'}}(\Omega;\mathbb{R}^m)$ be an operator satisfying the following condition: There exists a constant $L_x > 0$ such that

$$|\Psi[\xi] - \Psi[\overline{\xi}]| \le L_x|\xi - \overline{\xi}|, \quad \mathbb{P}\text{-a.s.}, \quad \forall\, \xi, \overline{\xi} \in L^p_{\mathcal{F}_{t'}}(\Omega;\mathbb{R}^n). \tag{2.6}$$

We notice that (2.6) can be regarded as a counterpart of Lipschitz continuity for functions in the operator case. Let $\zeta \in L^p_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$. We consider the following FBSDE:

$$\begin{cases} \mathrm{d}X_s = b(s,\Theta_s)\mathrm{d}s + \sigma(s,\Theta_s)\mathrm{d}W_s, \quad s \in [t,t'], \\ -\mathrm{d}Y_s = g(s,\Theta_s)\mathrm{d}s - Z_s\mathrm{d}W_s, \quad s \in [t,t'], \\ X_t = \zeta, \quad Y_{t'} = \Psi[X_{t'}], \end{cases} \tag{2.7}$$

where $\Theta = (X,Y,Z)$.

Lemma 2.1 *Let $p > 2$ and $0 \le t < t' \le T$. Let γ satisfy Assumption* (H1) *and Ψ satisfy* (2.6)*. Let Assumption* (H2)$'^p$ *hold. Let $\zeta \in L^p_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$, $\Psi[0] \in L^p_{\mathcal{F}_{t'}}(\Omega;\mathbb{R}^m)$ and $\gamma(\cdot,0) \in \mathcal{M}^p_{\mathbb{F}}(t,t';\mathbb{R}^{m+n+n\times d})$. Then, there exist a constant $\delta = \delta(p,T,L,L_x,L_z) > 0$ depending on p, T and the Lipschitz constants L, L_x, L_z, such that when the length of time duration $t' - t \le \delta$,* FBSDE (2.7) *with (Ψ,γ,ζ) admits a unique solution $\Theta = (X,Y,Z) \in M^p_{\mathbb{F}}(t,t';\mathbb{R}^{n+m+m\times d})$. Moreover, the following L^p-estimate holds:*

$$\begin{aligned} &\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,t']}|X_s|^p + \sup_{s\in[t,t']}|Y_s|^p + \Big(\int_t^{t'}|Z_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\ &\le \widetilde{C}_p\mathbb{E}^{\mathcal{F}_t}[|\zeta|^p + |\Psi[0]|^p + I(t,t';p)], \end{aligned} \tag{2.8}$$

where $\widetilde{C}_p = \widetilde{C}_p(p,T,L,L_x,L_z) > 0$ is a constant and

$$I(t,t';p) := \Big(\int_t^{t'}|g(s,0)|\mathrm{d}s\Big)^p + \Big(\int_t^{t'}|b(s,0)|\mathrm{d}s\Big)^p + \Big(\int_t^{t'}|\sigma(s,0)|^2\mathrm{d}s\Big)^{\frac{p}{2}}. \tag{2.9}$$

Proof For any $(y,z) \in L^p_{\mathbb{F}}(\Omega;C([t,t'];\mathbb{R}^m)) \times L^p_{\mathbb{F}}(\Omega;L^2(t,t';\mathbb{R}^{m\times d}))$, we introduce the following decoupled FBSDE:

$$\begin{cases} \mathrm{d}X_s = b(s,X_s,y_s,z_s)\mathrm{d}s + \sigma(s,X_s,y_s,z_s)\mathrm{d}W_s, \quad s \in [t,t'], \\ -\mathrm{d}Y_s = g(s,X_s,Y_s,Z_s)\mathrm{d}s - Z_s\mathrm{d}W_s, \quad s \in [t,t'], \\ X_t = \zeta, \quad Y_{t'} = \Psi[X_{t'}], \end{cases} \tag{2.10}$$

which admits a unique solution $(X, Y, Z) \in M_{\mathbb{F}}^p(t, t'; \mathbb{R}^{n+m+m\times d})$ by the classical theories of SDEs and BSDEs. We define a mapping from $L_{\mathbb{F}}^p(\Omega; C([t,t']; \mathbb{R}^m)) \times L_{\mathbb{F}}^p(\Omega; L^2(t,t'; \mathbb{R}^{m\times d}))$ into itself:

$$\mathscr{T} : (y, z) \mapsto (Y, Z). \tag{2.11}$$

Next, we shall prove that $\mathscr{T}$ is contractive when the time duration is small enough. Let $(y^i, z^i) \in L_{\mathbb{F}}^p(\Omega; C([t,t']; \mathbb{R}^m)) \times L_{\mathbb{F}}^p(\Omega; L^2(t,t'; \mathbb{R}^{m\times d}))$ $(i = 1, 2)$ be given. Let $(X^i, Y^i, Z^i) \in M_{\mathbb{F}}^p(t, t'; \mathbb{R}^{n+m+m\times d})$ be the corresponding solution to FBSDE (2.10). We denote

$$\widehat{y} = y^1 - y^2, \quad \widehat{z} = z^1 - z^2, \quad \widehat{X} = X^1 - X^2, \quad \widehat{Y} = Y^1 - Y^2, \quad \widehat{Z} = Z^1 - Z^2.$$

By the standard L^p-estimate for SDEs (see [32, Theorem 3.4.3]), we have

$$\begin{aligned}
\mathbb{E}\Big[\sup_{s\in[t,t']} |\widehat{X}_s|^p\Big] &\le C_F \mathbb{E}\Big\{\Big(\int_t^{t'} |b(s, X_s^2, y_s^1, z_s^1) - b(s, X_s^2, y_s^2, z_s^2)| \mathrm{d}s\Big)^p \\
&\quad + \Big(\int_t^{t'} |\sigma(s, X_s^2, y_s^1, z_s^1) - \sigma(s, X_s^2, y_s^2, z_s^2)|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big\} \\
&\le C_F \mathbb{E}\Big\{C(p, T, L, L_z, \varepsilon_1, \varepsilon_2)(t'-t)^{\frac{p}{2}}\Big[\sup_{s\in[t,t']} |\widehat{y}_s|^p + \Big(\int_t^{t'} |\widehat{z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\
&\quad + (1+\varepsilon_2)(1+\varepsilon_1)^{\frac{p}{2}} L_z^p \Big(\int_t^{t'} |\widehat{z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big\},
\end{aligned} \tag{2.12}$$

where $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ are two arbitrary positive numbers while $C(p, T, L, L_z, \varepsilon_1, \varepsilon_2) > 0$ is a constant which depends on p, T, L, L_z, ε_1 and ε_2. On the other hand, by the standard L^p-estimate for BSDEs (see [3, 10, 32, Theorem 4.4.4]), we also have

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{s\in[t,t']} |\widehat{Y}_s|^p + \Big(\int_t^{t'} |\widehat{Z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\
&\le C_B \mathbb{E}\Big\{|\Psi[X_{t'}^1] - \Psi[X_{t'}^2]|^p + \Big(\int_t^{t'} |g(s, X_s^1, Y_s^2, Z_s^2) - g(s, X_s^2, Y_s^2, Z_s^2)| \mathrm{d}s\Big)^p\Big\} \\
&\le C_B [L_x^p + L^p (t'-t)^p] \mathbb{E}\Big[\sup_{s\in[t,t']} |\widehat{X}_s|^p\Big].
\end{aligned} \tag{2.13}$$

Substituting (2.12) into (2.13), we have

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{s\in[t,t']} |\widehat{Y}_s|^p + \Big(\int_t^{t'} |\widehat{Z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\
&\le \mathbb{E}\Big\{\widetilde{C}(p, T, L, L_x, L_z, \varepsilon_1, \varepsilon_2)(t'-t)^{\frac{p}{2}}\Big[\sup_{s\in[t,t']} |\widehat{y}_s|^p + \Big(\int_t^{t'} |\widehat{z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\
&\quad + C_B C_F L_x^p (1+\varepsilon_2)(1+\varepsilon_1)^{\frac{p}{2}} L_z^p \Big(\int_t^{t'} |\widehat{z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big\}.
\end{aligned} \tag{2.14}$$

Due to Assumption $(\mathrm{H2})'^p$, we can select small enough constants $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that $C_B C_F L_x^p (1+\varepsilon_2)(1+\varepsilon_1)^{\frac{p}{2}} L_z^p < 1$. Then, we select a constant $\delta = \delta(p, T, L, L_x, L_z) > 0$ such that when $t' - t \le \delta$,

$$\mathbb{E}\Big[\sup_{s\in[t,t']}|\widehat{Y}_s|^p + \Big(\int_t^{t'}|\widehat{Z}_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big] \le \rho\mathbb{E}\Big[\sup_{s\in[t,t']}|\widehat{y}_s|^p + \Big(\int_t^{t'}|\widehat{z}_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big], \tag{2.15}$$

where $\rho \in (0,1)$ is a constant. In other words, the mapping $\mathscr{T}$ is contractive. Therefore it admits a unique fixed point denoted by (Y, Z), which together with the unique solution X to the following SDE:

$$X_s = \zeta + \int_t^s b(r, X_r, Y_r, Z_r)\mathrm{d}r + \int_t^s \sigma(r, X_r, Y_r, Z_r)\mathrm{d}W_r, \quad s \in [t, t'] \tag{2.16}$$

is just the unique solution to FBSDE (2.7).

The classical L^p-estimates for SDEs and BSDEs work again to obtain the estimate (2.8). Since the detailed proof is similar to the above one, we would like to omit it.

In order to ensure the solvability of FBSDE (2.1), we would like to give below a kind of monotonicity condition for the coefficients (Φ, γ), which was introduced by Hu and Peng [10] and Peng and Wu [25]. Let G be a given $(m \times n)$ full-rank matrix. We denote

$$A(t,\theta) = (-G^{\mathrm{T}}g(t,\theta),\ Gb(t,\theta),\ G\sigma(t,\theta)), \quad (t,\theta) \in [0,T] \times \mathbb{R}^{n+m+m\times d}.$$

(H3) There exist three nonnegative constants β_1, β_2 and μ_1 satisfying the following two conditions: (i) $\beta_1 > 0$, $\mu_1 > 0$ in the case of $m \ge n$, or $\beta_2 > 0$ in the case of $n \ge m$; (ii) for each $\theta = (x, y, z)$, $\overline{\theta} = (\overline{x}, \overline{y}, \overline{z})$,

$$\begin{aligned}&\langle \Phi(x) - \Phi(\overline{x}),\ G(x - \overline{x})\rangle \ge \mu_1|G(x - \overline{x})|^2,\\ &\langle A(t,\theta) - A(t,\overline{\theta}),\ \theta - \overline{\theta}\rangle \le -\beta_1|G(x-\overline{x})|^2 - \beta_2(|G^{\mathrm{T}}(y-\overline{y})|^2 + |G^{\mathrm{T}}(z-\overline{z})|^2).\end{aligned}$$

With the monotonicity condition (H3), we can get the L^2-results of FBSDEs. Similar to the symbol $I(t,T;p)$ in (2.9), we set

$$\begin{aligned}\widehat{I}(t,T;p) = {}&\Big(\int_t^T |g(s,\overline{\Theta}_s) - \overline{g}(s,\overline{\Theta}_s)|\mathrm{d}s\Big)^p + \Big(\int_t^T |b(s,\overline{\Theta}_s) - \overline{b}(s,\overline{\Theta}_s)|\mathrm{d}s\Big)^p\\ &+ \Big(\int_t^T |\sigma(s,\overline{\Theta}_s) - \overline{\sigma}(s,\overline{\Theta}_s)|^2\mathrm{d}s\Big)^{\frac{p}{2}}.\end{aligned} \tag{2.17}$$

Lemma 2.2 *Suppose the coefficients* (Φ, γ) *satisfy Assumptions* (H1) *and* (H3). *Let* $\Phi(0) \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^m)$ *and* $\gamma(\cdot, 0) \in \mathcal{M}^2_{\mathbb{F}}(0,T;\mathbb{R}^{m+n+n\times d})$. *Then, for any initial pair* $(t,\zeta) \in [0,T] \times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$, FBSDE (2.1) *with* (Φ,γ,ζ) *admits a unique solution*

$$\Theta^{t,\zeta} = (X^{t,\zeta}, Y^{t,\zeta}, Z^{t,\zeta}) \in M^2_{\mathbb{F}}(t,T;\mathbb{R}^{n+m+m\times d}).$$

Moreover, the following L^2*-estimate holds*:

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,T]}|X_s^{t,\zeta}|^2 + \sup_{s\in[t,T]}|Y_s^{t,\zeta}|^2 + \int_t^T |Z_s^{t,\zeta}|^2\mathrm{d}s\Big]\\ &\le C_2\mathbb{E}^{\mathcal{F}_t}[|\zeta|^2 + |\Phi(0)|^2 + I(t,T;2)],\end{aligned} \tag{2.18}$$

where $C_2 = C_2(T, L, L_x, L_z) > 0$ *is a constant. Furthermore, let* $(\overline{\Phi}, \overline{\gamma}, \overline{\zeta})$ *be another set of coefficients satisfying* (H1), *where* $\overline{\Phi}(0) \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^m)$, $\overline{\gamma}(\cdot, 0) \in \mathcal{M}^2_{\mathbb{F}}(t, T; \mathbb{R}^{m+n+n\times d})$ *and* $\overline{\zeta} \in L^2_{\mathcal{F}_t}(\Omega; \mathbb{R}^n)$. *Let* $\overline{\Theta} = (\overline{X}, \overline{Y}, \overline{Z}) \in M^2_{\mathbb{F}}(t, T; \mathbb{R}^{n+m+m\times d})$ *be a solution to* FBSDE (2.1) *with* $(\overline{\Phi}, \overline{\gamma}, \overline{\zeta})$. *Then*

$$\begin{aligned} &\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,T]} |X^{t,\zeta}_s - \overline{X}_s|^2 + \sup_{s\in[t,T]} |Y^{t,\zeta}_s - \overline{Y}_s|^2 + \int_t^T |Z^{t,\zeta}_s - \overline{Z}_s|^2 \mathrm{d}s\Big] \\ &\le C_2 \mathbb{E}^{\mathcal{F}_t}[|\zeta - \overline{\zeta}|^2 + |\Phi(\overline{X}_T) - \overline{\Phi}(\overline{X}_T)|^2 + \widehat{I}(t, T; 2)], \end{aligned} \tag{2.19}$$

where $\widehat{I}(t, T; 2)$ *is defined by* (2.17).

By now, the results in Lemma 2.2 are standard. Specifically, the unique solvability for FBSDE (2.1) can be found in Peng and Wu [25, Theorem 2.6], and the pair of L^2-estimates are from Yong [30] (see also Wu [26]).

With the help of Lemma 2.2, for any $t \in [0, T]$, we define an operator $u(t, [\cdot]) : L^2_{\mathcal{F}_t}(\Omega; \mathbb{R}^n) \to L^2_{\mathcal{F}_t}(\Omega; \mathbb{R}^m)$ as follows:

$$u(t, [\zeta]) := Y^{t,\zeta}_t, \quad \zeta \in L^2_{\mathcal{F}_t}(\Omega; \mathbb{R}^m). \tag{2.20}$$

A couple of immediate consequences of (2.18)–(2.19) are obtained:

$$\begin{cases} \text{(i) } |u(t, [\zeta])|^2 \le C_2\{|\zeta|^2 + \mathbb{E}^{\mathcal{F}_t}[|\Phi(0)|^2 + I(t, T; 2)]\}, \\ \text{(ii) } |u(t, [\zeta]) - u(t, [\overline{\zeta}])| \le \sqrt{C_2}|\zeta - \overline{\zeta}|, \end{cases} \tag{2.21}$$

where the constant C_2 is the same one given in (2.18).

Moreover, in order to use Lemma 2.1, we need a stronger condition $(\mathrm{H2})^p$ to replace $(\mathrm{H2})'^p$.

$\mathbf{(H2)}^p$ For the given $p > 2$, the inequality $C_F(p, T, L) \cdot C_B(p, T, L) \cdot C_2^{\frac{p}{2}} \cdot L_z^p < 1$ holds where $C_2 = C_2(T, L, L_x, L_z) > 0$ is given in (2.18) of Lemma 2.2.

Remark 2.2 Historically, for the L^p-estimates of coupled FBSDEs on small durations, Delarue [6] for the first time obtained a result when σ is independent of z. In 2014, Li and Wei [12–13] established L^p-estimates when the Lipschitz constant of σ with respect to z is small enough. In 2020, Xie and Yu [28] gave the L^p-estimates of FBSDEs (1.2) on small durations as $\Phi(\cdot)$ is linear. To our knowledge, even on small intervals, the L^p-estimates can only be obtained in the case of

$$\text{the Lipschitz constant of } \sigma(\cdot,\cdot) \text{ with respect to } z \text{ is small enough,} \tag{2.22}$$

or the terminal conditions are linear. Moreover, being "small enough" in (2.22) does not obtain a quantitative characterization. Clearly, condition (2.22) is a special case of $(\mathrm{H2})^p$.

On large time durations, in 2015, Ma et al. [16] firstly provided an L^p-estimate for 1-dimensional FBSDEs. Some other restrictive assumptions were also required. In 2018, Hu et al. [9] gave an L^p-result when the Lipschitz constants of both b and σ with respect to (y, z)

are very small. In 2018, Feng et al. [8] established an L^p-result when both b and σ are independent of z. In 2023, Xie and Yu [29] also established an L^p-result for the linear FBSDEs, without condition (2.22). So far, in the case of nonlinearity, (2.22) cannot be overcome.

Even more frustrating is that we cannot determine the necessity of condition $(\mathrm{H2})^p$ for L^p-estimates. Therefore, we propose a further work: Without using condition $(\mathrm{H2})^p$, prove the L^p-estimates of FBSDEs. Alternatively, provide an example to illustrate the necessity of condition $(\mathrm{H2})^p$. Note: This further work is proposed by the referee.

Now, with the help of $u(\cdot,[\cdot])$ as well as Lemma 2.1, we obtain the following theorem.

Theorem 2.1 *Let $p > 2$ and the coefficients (Φ,γ) satisfy Assumptions* (H1), $(\mathrm{H2})^p$ *and* (H3). *Let $\Phi(0) \in L^p_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)$ and $\gamma(\cdot,0) \in \mathcal{M}^p_{\mathbb{F}}(0,T;\mathbb{R}^{m+n+n\times d})$. Then, for any initial pair $(t,\zeta) \in [0,T] \times L^p_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$,* FBSDE (2.1) *with (Φ,γ,ζ) admits a unique solution*

$$\Theta^{t,\zeta} = (X^{t,\zeta}, Y^{t,\zeta}, Z^{t,\zeta}) \in M^p_{\mathbb{F}}(t,T;\mathbb{R}^{n+m+m\times d}).$$

Moreover, the following L^p-estimate holds:

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,T]}|X^{t,\zeta}_s|^p + \sup_{s\in[t,T]}|Y^{t,\zeta}_s|^p + \Big(\int_t^T |Z^{t,\zeta}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big]\\&\le C_p\mathbb{E}^{\mathcal{F}_t}[|\zeta|^p + |\Phi(0)|^p + I(t,T;p)],\end{aligned}\tag{2.23}$$

where $C_p = C_p(p,T,L,L_x,L_z) > 0$ is a constant and $I(t,T;p)$ is defined by (2.9). *Furthermore, let $(\overline{\Phi},\overline{\gamma},\overline{\zeta})$ be another set of coefficients satisfying* (H1), *where $\overline{\Phi}(0) \in L^p_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)$, $\overline{\gamma}(\cdot,0) \in \mathcal{M}^p_{\mathbb{F}}(t,T;\mathbb{R}^{m+n+n\times d})$ and $\overline{\zeta} \in L^p_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$. Let $\overline{\Theta} = (\overline{X},\overline{Y},\overline{Z}) \in M^p_{\mathbb{F}}(t,T;\mathbb{R}^{n+m+m\times d})$ be a solution to* FBSDE (2.1) *with $(\overline{\Phi},\overline{\gamma},\overline{\zeta})$. Then*

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t,T]}|X^{t,\zeta}_s - \overline{X}_s|^p + \sup_{s\in[t,T]}|Y^{t,\zeta}_s - \overline{Y}_s|^p + \Big(\int_t^T |Z^{t,\zeta}_s - \overline{Z}_s|^2 \mathrm{d}s\Big)^{\frac{p}{2}}\Big]\\&\le C_p\mathbb{E}^{\mathcal{F}_t}[|\zeta - \overline{\zeta}|^p + |\Phi(\overline{X}_T) - \overline{\Phi}(\overline{X}_T)|^p + \widehat{I}(t,T;p)],\end{aligned}\tag{2.24}$$

where $\widehat{I}(t,T;p)$ is defined by (2.17).

Proof Let $p > 2$. By Lemma 2.1, the existence and uniqueness of the L^p-solution and L^p-estimate hold true on small durations. For the given large duration $[t,T]$, we split it with points $t = t_0 < t_1 < \cdots < t_N = T$, and let the mesh size of the partition

$$\max_{0\le k\le N-1}(t_{k+1} - t_k) \le \delta,$$

where the constant δ is given in Lemma 2.1. On the one hand, from the viewpoint of Lemma 2.2, on each small duration $[t_k, t_{k+1}]$ with $k = 0,1,2,\cdots,N-1$, the L^2-solution $\Theta^{t,\zeta}$ satisfies the following FBSDE (notice the definition of $u(\cdot,[\cdot])$):

$$\begin{cases} \mathrm{d}X_s = b(s,\Theta_s)\mathrm{d}s + \sigma(s,\Theta_s)\mathrm{d}W(s), & s\in[t_k,t_{k+1}],\\ -\,\mathrm{d}Y_s = g(s,\Theta_s)\mathrm{d}s - Z_s\mathrm{d}W_s, & s\in[t_k,t_{k+1}],\\ X_{t_k} = X^{t,\zeta}_{t_k}, \quad Y_{t_{k+1}} = u(t_{k+1},[X_{t_{k+1}}]), \end{cases} \tag{2.25}$$

where $\Theta=(X,Y,Z)$. On the other hand, we change our viewpoint from Lemma 2.2 to Lemma 2.1, and reanalyze the sequence of FBSDEs (2.25). In fact, firstly on $[t_0,t_1]$, we notice that $X^{t,\zeta}_{t_0}=\zeta\in L^p_{\mathcal{F}_{t_0}}(\Omega;\mathbb{R}^n)$. Moreover, from (2.21)(i) and Hölder's inequality,

$$\begin{aligned} |u(t_1,[0])|^p &\le (C_2\{\mathbb{E}^{\mathcal{F}_{t_1}}[|\Phi(0)|^2 + I(t_1,T;2)]\})^{\frac{p}{2}}\\ &\le K\mathbb{E}^{\mathcal{F}_{t_1}}[|\Phi(0)|^p + I(t_1,T;2)^{\frac{p}{2}}]\\ &\le K\mathbb{E}^{\mathcal{F}_{t_1}}[|\Phi(0)|^p + I(t_1,T;p)]. \end{aligned} \tag{2.26}$$

Then

$$\mathbb{E}[|u(t_1,[0])|^p] \le K\mathbb{E}[|\Phi(0)|^p + I(t,T;p)] < \infty.$$

With the help of (2.21)(ii), we know that the restriction $u(t_1,[\cdot])|_{L^p_{\mathcal{F}_{t_1}}(\Omega;\mathbb{R}^n)}$ is an operator from $L^p_{\mathcal{F}_{t_1}}(\Omega;\mathbb{R}^n)$ into $L^p_{\mathcal{F}_{t_1}}(\Omega;\mathbb{R}^m)$. Then by Lemma 2.1, $\{\Theta^{\tau,\zeta}_s,\ s\in[t_0,t_1]\}$ belongs exactly to $M^p_{\mathbb{F}}(t_0,t_1;\mathbb{R}^{n+m+m\times d})$. This procedure can be carried out for $k=1,2,\cdots,N-1$ one by one to yield $\{\Theta^{\tau,\zeta}_s,\ s\in[t,T]\}$ belongs exactly to $M^p_{\mathbb{F}}(t,T;\mathbb{R}^{n+m+m\times d})$. We have proved the existence and uniqueness of the L^p-solution for FBSDE (2.1) on the whole duration $[t,T]$.

We turn to prove the L^p-estimate (2.23). For any $k=0,1,2,\cdots,N-1$, by (2.8), we have

$$\begin{aligned} &\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t_k,t_{k+1}]}|X_s|^p + \sup_{s\in[t_k,t_{k+1}]}|Y_s|^p + \Big(\int_{t_k}^{t_{k+1}}|Z_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big]\\ &\le C_p\mathbb{E}^{\mathcal{F}_t}[|X_{t_k}|^p + |u(t_{k+1},[0])|^p + I(t_k,t_{k+1};p)]. \end{aligned} \tag{2.27}$$

Similar to (2.26),

$$|u(t_{k+1},[0])|^p \le K\mathbb{E}^{\mathcal{F}_{t_{k+1}}}[|\Phi(0)|^p + I(t_{k+1},T;p)].$$

Substituting it into (2.27) leads to

$$\begin{aligned} &\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t_k,t_{k+1}]}|X_s|^p + \sup_{s\in[t_k,t_{k+1}]}|Y_s|^p + \Big(\int_{t_k}^{t_{k+1}}|Z_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big]\\ &\le K\mathbb{E}^{\mathcal{F}_t}[|X_{t_k}|^p + |\Phi(0)|^p + I(t,T;p)]. \end{aligned} \tag{2.28}$$

We notice that the first item $\mathbb{E}^{\mathcal{F}_t}[|X_{t_k}|^p]$ on the right hand side of the above inequality depends on $\mathbb{E}^{\mathcal{F}_t}[|X_{t_{k-1}}|^p]$, and then depends on $\mathbb{E}^{\mathcal{F}_t}[|X_{t_{k-2}}|^p],\cdots,\mathbb{E}^{\mathcal{F}_t}[|X_{t_0}|^p]=|\zeta|^p$, in the following sense:

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_t}[|X_{t_k}|^p] &\le \mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t_{k-1},t_k]}|X_s|^p\Big]\\ &\le K\mathbb{E}^{\mathcal{F}_t}[|X_{t_{k-1}}|^p + |\Phi(0)|^p + I(t,T;p)]\\ &\le K\mathbb{E}^{\mathcal{F}_t}[|X_{t_{k-2}}|^p + |\Phi(0)|^p + I(t,T;p)]\\ &\le \cdots \le K\mathbb{E}^{\mathcal{F}_t}[|\zeta|^p + |\Phi(0)|^p + I(t,T;p)]. \end{aligned}$$

Therefore, (2.28) is reduced to

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_t}\Big[\sup_{s\in[t_k,t_{k+1}]}|X_s|^p+\sup_{s\in[t_k,t_{k+1}]}|Y_s|^p+\Big(\int_{t_k}^{t_{k+1}}|Z_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big]\\&\le K\mathbb{E}^{\mathcal{F}_t}[|\zeta|^p+|\Phi(0)|^p+I(t,T;p)].\end{aligned}\tag{2.29}$$

By summing up the above inequalities with k from 0 to $N-1$, we obtain the desired L^p-estimate (2.23) on the large duration $[t,T]$.

The other L^p-estimate (2.24) can be regarded as a consequence of the first L^p-estimate (2.23). In detail, we denote $\widehat{\zeta}=\zeta-\overline{\zeta}$, $\widehat{\Theta}=\Theta-\overline{\Theta}=(X-\overline{X},Y-\overline{Y},Z-\overline{Z})$ and

$$\widehat{f}(s,x,y,z)=f(s,x+\overline{X}_s,y+\overline{Y}_s,z+\overline{Z}_s)-\overline{f}(s,\overline{X}_s,\overline{Y}_s,\overline{Z}_s),\quad s\in[t,T]$$

with $f=g,b,\sigma$. Then $\widehat{\Theta}=(\widehat{X},\widehat{Y},\widehat{Z})$ satisfies the following FBSDE

$$\begin{cases}\mathrm{d}\widehat{X}_s=\widehat{b}(s,\widehat{X}_s,\widehat{Y}_s,\widehat{Z}_s)\mathrm{d}s+\widehat{\sigma}(s,\widehat{X}_s,\widehat{Y}_s,\widehat{Z}_s)\mathrm{d}W_s,\quad s\in[t,T],\\-\mathrm{d}\widehat{Y}_s=\widehat{g}(s,\widehat{X}_s,\widehat{Y}_s,\widehat{Z}_s)\mathrm{d}s-\widehat{Z}_s\mathrm{d}W_s,\quad s\in[t,T],\\\widehat{X}_t=\widehat{\zeta},\quad \widehat{Y}_T=\widehat{\Phi}(\widehat{X}_T).\end{cases}\tag{2.30}$$

Clearly, $(\widehat{\Phi},\widehat{\gamma},\widehat{\zeta})$ satisfies all conditions in Theorem 2.1. Hence, the application of estimate (2.23) to the above FBSDE yields (2.24). We complete the proof.

2.2 Function u and its elementary properties

In this paper, we shall use the following smoothness assumption.

(H4) Φ and γ are deterministic functions. Moreover, γ is continuous with respect to t. For any $t\in[0,T]$, $(\Phi(\cdot),\gamma(t,\cdot))$ are of class C^2 and all the partial derivatives of order less than or equal to 2 are bounded on $[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}$.

Obviously, (H4) implies (H1) and the boundedness of $\gamma(\cdot,0)$. Under Assumption (H4), the triple of solutions to FBSDE (1.2) is independent of the past information $\mathcal{F}_t$, i.e., $\Theta^{t,x}=(X^{t,x},Y^{t,x},Z^{t,x})$ is $\mathbb{F}^t$-adapted, where $\mathbb{F}^t=(\mathcal{F}^t_s)_{s\in[t,T]}$ is the natural filtration generated by $(W_s-W_t)_{s\in[t,T]}$ and augmented by all $\mathbb{P}$-null sets. In particular, $Y^{t,x}_t$ is deterministic. Now, we recall the operator $u(\cdot,[\cdot])$ which is defined by (2.20). Obviously, in the Markovian framework (H4),

$$u(t,x):=Y^{t,x}_t,\quad (t,x)\in[0,T]\times\mathbb{R}^n\tag{2.31}$$

provides a function from $[0,T]\times\mathbb{R}^n$ to $\mathbb{R}^m$. Now, let us recall a result from [27, Lemma 2.4 or Proposition 2.6].

Proposition 2.1 *Let Assumptions* (H3) *and* (H4) *hold. Then,*

$$u(t,\zeta)=u(t,[\zeta])=Y^{t,\zeta}_t\quad \textit{for any } (t,\zeta)\in[0,T]\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n).\tag{2.32}$$

This result means that the value of the composite function is equal to the value of the operator.

For the convenience of notation, we shall define $\Theta_s^{t,x}$ for all $(s,t) \in [0,T]^2$ by setting $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}) \equiv (X_t^{t,x}, Y_t^{t,x}, 0)$ for $s < t$. The following corollary shows the continuity of the function u. It can be regarded as an improved version of [27, Proposition 2.5].

Corollary 2.1 *Let Assumptions* (H3) *and* (H4) *hold. Then there exists a constant $C > 0$ such that, for any (t,x), $(t',x') \in [0,T] \times \mathbb{R}^n$,*

$$\|\Theta^{t,x} - \Theta^{t',x'}\|^2_{M^2_{\mathbb{F}}(0,T)} \leq C\{|x-x'|^2 + (1+|x|^2+|x'|^2)|t'-t|\}. \tag{2.33}$$

Consequently, the function u is Lipschitz continuous in x uniformly for any given $t \in [0,T]$, and is $\frac{1}{2}$-Hölder continuous in t for any given $x \in \mathbb{R}^n$.

Proof Without loss of generality, we assume that $t \leq t'$. By the L^2-estimate, we have

$$\begin{aligned}\|\Theta^{t,x} - \Theta^{t',x'}\|^2_{M^2_{\mathbb{F}}(t,T)} \leq C_2\Big\{&|x-x'|^2 + \Big(\int_t^{t'} |g(s,x',Y_{t'}^{t',x'},0)|\mathrm{d}s\Big)^2 \\ &+ \Big(\int_t^{t'} |b(s,x',Y_{t'}^{t',x'},0)|\mathrm{d}s\Big)^2 + \int_t^{t'} |\sigma(s,x',Y_{t'}^{t',x'},0)|^2\mathrm{d}s\Big\}.\end{aligned}$$

With the help of Assumption (H4) and (2.21)(i), we obtain (2.33) and finish the proof.

At the end of this section, we recall a definition and a property of the function u from [27, Remark 2.9 and Proposition 2.8]. They will be used to deal with the algebraic equation in the PDAE system (1.1).

Definition 2.1 *Let $G \in \mathbb{R}^{m\times n}$ be a matrix, $\nu \geq 0$ be a constant and $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. If*

$$\langle f(x) - f(\overline{x}),\ G(x-\overline{x})\rangle \geq \nu|x-\overline{x}|^2 \quad \text{for any } x, \overline{x} \in \mathbb{R}^n, \tag{2.34}$$

then we call f G-monotonic with ν. Moreover, when $\nu > 0$, we also call f strictly G-monotonic with ν.

Proposition 2.2 *Let Assumptions* (H3) *and* (H4) *hold. Then, there exists a constant $\nu \geq 0$ such that for any $t \in [0,T]$, the function $u(t,\cdot)$ defined by* (2.31) *is G-monotonic with ν, where the matrix G is the same one appearing in Assumption* (H3)*. Moreover, when $\beta_1 > 0$, $\mu_1 > 0$, $\beta_2 \geq 0$ and $m \geq n$ in Assumption* (H3)*, the constant ν is strictly greater than* 0.

3 Regularity of Solutions to FBSDEs

The following lemma is helpful for our analysis below.

Lemma 3.1 *For any $t \in [0,T]$, assume that $(\Phi(\cdot), \gamma(t,\cdot))$ are of class C^1. Then the following two statements are equivalent.*

(i) *(Φ, γ) satisfies the monotonicity assumption* (H3)*.*

(ii) *For any* $\theta_0 = (x_0, y_0, z_0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$*, denote*

$$\begin{cases} \widetilde{\Phi}(x) := \dfrac{\mathrm{d}\Phi}{\mathrm{d}x}(x_0)x, \\ \widetilde{f}(t,\theta) := \nabla f(t,\theta_0)\theta := \dfrac{\partial f}{\partial x}(t,\theta_0)x + \dfrac{\partial f}{\partial y}(t,\theta_0)y + \dfrac{\partial f}{\partial z}(t,\theta_0)z, \end{cases}$$

where $f = g, b, \sigma$*. Then* $(\widetilde{\Phi}, \widetilde{\gamma}) = (\widetilde{\Phi}, \widetilde{g}, \widetilde{b}, \widetilde{\sigma})$ *satisfies the monotonicity condition* (H3) *with the same constants* β_1*,* β_2 *and* μ_1*.*

Proof We only prove the part about Φ and $\widetilde{\Phi}$. The same technique can also be used to prove the other part about γ and $\widetilde{\gamma}$.

We first prove that Statement (i) implies Statement (ii). Let us begin with the following inequality: For any x_0 and $\overline{x}_0$,

$$\langle \Phi(\overline{x}_0) - \Phi(x_0), G(\overline{x}_0 - x_0)\rangle \geq \mu_1 |G(\overline{x}_0 - x_0)|^2.$$

Now, for any vector $K \in \mathbb{R}^n$ and any real number $\delta > 0$, we select $\overline{x}_0 = x_0 + \delta K$, then

$$\delta\langle \Phi(x_0 + \delta K) - \Phi(x_0),\ GK\rangle \geq \delta^2 \mu_1 |GK|^2.$$

By the smoothness condition of Φ, by Lagrange's differential mean value theorem, for each $\delta > 0$, there exists an $\alpha \in (0,1)$ depending on δ, such that

$$\Big\langle \frac{\mathrm{d}\Phi}{\mathrm{d}x}(x_0 + \alpha\delta K)K,\ GK\Big\rangle \geq \mu_1 |GK|^2.$$

Letting $\delta \to 0^+$, we get

$$\Big\langle \frac{\mathrm{d}\Phi}{\mathrm{d}x}(x_0)K,\ GK\Big\rangle \geq \mu_1 |GK|^2.$$

For any $x, \overline{x} \in \mathbb{R}^n$, replacing K in the above inequality with $x - \overline{x}$, we obtain

$$\Big\langle \frac{\mathrm{d}\Phi}{\mathrm{d}x}(x_0)(x - \overline{x}),\ G(x - \overline{x})\Big\rangle \geq \mu_1 |G(x - \overline{x})|^2,$$

which is the monotonicity of $\widetilde{\Phi}$.

Next we shall employ a framework of reduction to absurdity to prove that Statement (ii) also implies Statement (i). Assuming that the monotonicity condition of Φ does not hold, i.e., there exist x, $\overline{x} \in \mathbb{R}^n$ such that

$$\langle \Phi(x) - \Phi(\overline{x}),\ G(x - \overline{x})\rangle < \mu_1 |G(x - \overline{x})|^2.$$

Lagrange's differential mean value theorem works again to yield that there exists an $\alpha \in (0,1)$ such that

$$\Big\langle \frac{\mathrm{d}\Phi}{\mathrm{d}x}(\overline{x} + \alpha(x - \overline{x}))(x - \overline{x}),\ G(x - \overline{x})\Big\rangle < \mu_1 |G(x - \overline{x})|^2.$$

Clearly, the above inequality shows that when $x_0 = \overline{x} + \alpha(x - \overline{x})$, the monotonicity condition of function $\widetilde{\Phi}$ does not hold at x and $\overline{x}$, which contradicts Statement (ii).

3.1 Malliavin derivative of solutions to FBSDEs

First of all, we recall some notions about Malliavin derivatives on Wiener space from Nualart [17] and El Karoui, Peng and Quenez [7].

(1) Let $C_b^k(\mathbb{R}^p;\mathbb{R}^q)$ denote the set of functions of class C^k from $\mathbb{R}^p$ into $\mathbb{R}^q$ whose partial derivatives of order less than or equal to k are bounded.

(2) Let $\mathcal{S}$ denote the set of random variables ξ of the form $\xi=\varphi(W(h^1),\cdots,W(h^k))$, where $\varphi\in C_b^\infty(\mathbb{R}^k;\mathbb{R})$, $h^1,\cdots,h^k\in L^2(0,T;\mathbb{R}^d)$ and $W(h^i)=\int_0^T\langle h_s^i,\ \mathrm{d}W_s\rangle$.

(3) For any $\xi\in\mathcal{S}$, its Malliavin derivative is defined by the following d-dimensional process

$$D_\varrho\xi=\sum_{j=1}^k\frac{\partial\varphi}{\partial x_j}(W(h^1),\cdots,W(h^k))h_\varrho^j,\quad \varrho\in[0,T].$$

The ith component of $D_\varrho\xi$ is denoted by $D_\varrho^i\xi$, $i=1,2,\cdots,d$. The $(1,2)$-norm of ξ is defined by

$$\|\xi\|_{\mathbb{D}_{1,2}}=\Big\{\mathbb{E}\Big[|\xi|^2+\int_0^T|D_\varrho\xi|^2\mathrm{d}\varrho\Big]\Big\}^{\frac{1}{2}}.$$

It is known that the operator D has a closed extension to the space $\mathbb{D}_{1,2}$, which is the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{\mathbb{D}_{1,2}}$.

(4) Let $\mathbb{L}_{1,2}^a(0,T;\mathbb{R}^n)$ be the set of $\mathbb{R}^n$-valued $\mathbb{F}$-progressively measurable processes $\{\phi(t,\omega),\ t\in[0,T];\ \omega\in\Omega\}$ satisfying

(i) for a.e. $t\in[0,T]$, $\phi(t,\cdot)\in(\mathbb{D}_{1,2})^n$.

(ii) $(t,\omega)\to D_\varrho\phi(t,\omega)$ admits an $\mathbb{F}$-progressively measurable version.

(iii)

$$\begin{aligned}\|\phi\|^2_{\mathbb{L}_{1,2}^a(0,T;\mathbb{R}^n)}:&=\|\phi\|^2_{L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)}+\|D_\cdot\phi(\cdot)\|^2_{\widetilde{\mathbb{L}_{1,2}^a}}\\ :&=\mathbb{E}\int_0^T|\phi(s)|^2\mathrm{d}s+\mathbb{E}\int_0^T\int_0^T|D_\varrho\phi(s)|^2\mathrm{d}\varrho\mathrm{d}s<\infty.\end{aligned}$$

The space $\mathbb{L}_{1,2}^a(0,T;\mathbb{R}^n)$ is closed under the norm $\|\cdot\|_{\mathbb{L}_{1,2}^a(0,T;\mathbb{R}^n)}$.

We now show that under Assumptions (H3) and (H4), the L^2-solution to an FBSDE is differentiable in Malliavin's sense and the derivative is a solution to a linear FBSDE. This result generalizes the one stated by Pardoux and Peng [19] in the decoupled case. For simplicity of notation, we restrict ourselves to the case $d=1$ in this subsection, and suppress (t,x) in $\Theta^{t,x}=(X^{t,x},Y^{t,x},Z^{t,x})$. We shall combine the method used in [7, 19] (for BSDEs) with the method of continuation to establish the differentiability in Malliavin's sense. We split it into two cases according to the signs of β_1, β_2 and μ_1 (see the monotonicity assumption (H3)).

First case $\beta_1>0$, $\mu_1>0$, $\beta_2\geq 0$ and $n\leq m$.

For any $\phi,\psi\in L^2_{\mathbb{F}}(t,T;\mathbb{R}^n)\cap\mathbb{L}_{1,2}^a(t,T;\mathbb{R}^n)$, $\kappa\in L^2_{\mathbb{F}}(t,T;\mathbb{R}^m)\cap\mathbb{L}_{1,2}^a(t,T;\mathbb{R}^m)$ and $\xi\in$

$L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)\cap(\mathbb{D}_{1,2})^m$, we introduce the following family of FBSDEs parameterized by $\alpha\in[0,1]$:

$$\begin{cases} \mathrm{d}X_s^\alpha = [\alpha b(s,\Theta_s^\alpha)+\phi_s]\mathrm{d}s + [\alpha\sigma(s,\Theta_s^\alpha)+\psi_s]\mathrm{d}W_s, \\ -\mathrm{d}Y_s^\alpha = [(1-\alpha)\beta_1 GX_s^\alpha + \alpha g(s,\Theta_s^\alpha)+\kappa_s]\mathrm{d}s - Z_s^\alpha \mathrm{d}W_s, \\ X_t^\alpha = x, \quad Y_T^\alpha = (1-\alpha)\mu_1 GX_T^\alpha + \alpha\Phi(X_T^\alpha)+\xi, \end{cases} \tag{3.1}$$

where $\Theta^\alpha := (X^\alpha, Y^\alpha, Z^\alpha)$. It is clear that, when $\alpha=0$, FBSDE (3.1) is in a decoupled form and then the Malliavin's differentiability has been established due to the work of El Karoui, Peng and Quenez [7]; when $\alpha=1$ and (ϕ,ψ,κ,ξ) vanish, FBSDE (3.1) coincides with FBSDE (1.2). We shall show that there exists a fixed step-length $\delta_0>0$, such that if for some $\alpha_0\in[0,1)$, the solution to (3.1) is differentiable in Malliavin's sense for any (ϕ,ψ,κ,ξ), then the same conclusion holds for α_0 replaced by $\alpha_0+\delta<1$ with $\delta\in[0,\delta_0]$. Once this has been proved, we can increase the parameter α step by step and finally reach $\alpha=1$, which gives the Malliavin's differentiability of the solution to FBSDE (1.2). This method was originally introduced by Hu and Peng [10] for dealing with the L^2-solvability of coupled FBSDEs, which is called the method of continuation.

We have the following continuation lemma.

Lemma 3.2 *Let Assumptions* (H3) *and* (H4) *hold. Then there exists an absolute constant* $\delta_0>0$*, such that if for some* $\alpha_0\in[0,1)$*, the solution to* FBSDE (3.1) *is differentiable in Malliavin's sense for any* ϕ, $\psi\in L^2_{\mathbb{F}}(t,T;\mathbb{R}^n)\cap\mathbb{L}^a_{1,2}(t,T;\mathbb{R}^n)$, $\kappa\in L^2_{\mathbb{F}}(t,T;\mathbb{R}^m)\cap\mathbb{L}^a_{1,2}(t,T;\mathbb{R}^m)$ *and* $\xi\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)\cap(\mathbb{D}_{1,2})^m$*, then the same is true for* $\alpha=\alpha_0+\delta$ *with* $\delta\in[0,\delta_0]$, $\alpha_0+\delta\le 1$.

Proof Let $\delta_0>0$ be undetermined and $\delta\in[0,\delta_0]$. We introduce the following sequence defined by $\Theta^0 := (X^0,Y^0,Z^0)\equiv(0,0,0)$ and

$$\begin{cases} \mathrm{d}X_s^{k+1} = [\alpha_0 b(s,\Theta_s^{k+1})+\delta b(s,\Theta_s^k)+\phi_s]\mathrm{d}s \\ \qquad\qquad + [\alpha_0\sigma(s,\Theta_s^{k+1})+\delta\sigma(s,\Theta_s^k)+\psi_s]\mathrm{d}W_s, \\ -\mathrm{d}Y_s^{k+1} = [(1-\alpha_0)\beta_1 GX_s^{k+1}+\alpha_0 g(s,\Theta_s^{k+1})-\delta\beta_1 GX_s^k+\delta g(s,\Theta_s^k)+\kappa_s]\mathrm{d}s \\ \qquad\qquad - Z_s^{k+1}\mathrm{d}W_s, \\ X_t^{k+1} = x, \quad Y_T^{k+1} = (1-\alpha_0)\mu_1 GX_T^{k+1}+\alpha_0\Phi(X_T^{k+1})-\delta\mu_1 GX_T^k+\delta\Phi(X_T^k)+\xi, \end{cases} \tag{3.2}$$

where $\Theta^k := (X^k,Y^k,Z^k)$. It is easy to verify that the coefficients of the above FBSDE (3.2) satisfy the monotonicity conditions. Then, by applying the L^2-estimate, we have

$$\|\Theta^{k+1}-\Theta^k\|^2_{M^2_{\mathbb{F}}(t,T)} \le C\delta^2\|\Theta^k-\Theta^{k-1}\|^2_{M^2_{\mathbb{F}}(t,T)}.$$

We note that the constant $C>0$ appearing in the inequality is independent of α_0 and δ. Hence, if we choose $\delta_1>0$ such that $C\delta_1^2\le\frac{1}{4}$, then, for any $\delta\in[0,\delta_1]$, it turns out that Θ^k is a Cauchy sequence in $M^2_{\mathbb{F}}(t,T;\mathbb{R}^{n+m+m})$. By passing to the limit in FBSDEs (3.2), we see that the limit of Cauchy sequence Θ^k solves FBSDE (3.1) for $\alpha=\alpha_0+\delta$. In other words, Θ^k converges in $M^2_{\mathbb{F}}(t,T;\mathbb{R}^{n+m+m})$ to $\Theta^{\alpha_0+\delta}$ as $k\to\infty$.

By our assumption, the solution to FBSDE (3.1) is differentiable in the Malliavin's sense for $\alpha = \alpha_0$ and for any Malliavin differentiable $(\phi, \psi, \kappa, \xi)$. From (3.2), it is clear that $\Theta^k \in \mathbb{L}^a_{1,2}(t,T;\mathbb{R}^n) \times \mathbb{L}^a_{1,2}(t,T;\mathbb{R}^m) \times \mathbb{L}^a_{1,2}(t,T;\mathbb{R}^m)$ recursively. Moreover, the Malliavin derivative $D_\varrho \Theta^{k+1} := (D_\varrho X^{k+1}, D_\varrho Y^{k+1}, D_\varrho Z^{k+1})$ is a solution to the following linear FBSDE:

$$\begin{cases} D_\varrho X^{k+1}_s = [\alpha_0 \sigma(\varrho, \Theta^{k+1}_\varrho) + \delta\sigma(\varrho, \Theta^k_\varrho) + \psi_\varrho] \\ \qquad + \displaystyle\int_\varrho^s \{\alpha_0 \nabla b(r, \Theta^{k+1}_r) D_\varrho \Theta^{k+1}_r + \delta \nabla b(r, \Theta^k_r) D_\varrho \Theta^k_r + D_\varrho \phi_r\} \mathrm{d}r \\ \qquad + \displaystyle\int_\varrho^s \{\alpha_0 \nabla \sigma(r, \Theta^{k+1}_r) D_\varrho \Theta^{k+1}_r + \delta \nabla \sigma(r, \Theta^k_r) D_\varrho \Theta^k_r + D_\varrho \psi_r\} \mathrm{d}W_r, \\ D_\varrho Y^{k+1}_s = \Big[(1-\alpha_0)\mu_1 G D_\varrho X^{k+1}_T + \alpha_0 \dfrac{\mathrm{d}\Phi}{\mathrm{d}x}(X^{k+1}_T) D_\varrho X^{k+1}_T - \delta\mu_1 G D_\varrho X^k_T \\ \qquad + \delta \dfrac{\mathrm{d}\Phi}{\mathrm{d}x}(X^k_T) D_\varrho X^k_T + D_\varrho \xi\Big] + \displaystyle\int_s^T \{\alpha_0 \nabla g(r, \Theta^{k+1}_r) D_\varrho \Theta^{k+1}_r \\ \qquad + \delta \nabla g(r, \Theta^k_r) D_\varrho \Theta^k_r + (1-\alpha_0)\beta_1 G D_\varrho X^{k+1}_r - \delta\beta_1 G D_\varrho X^k_r + D_\varrho \kappa_r\} \mathrm{d}r \\ \qquad - \displaystyle\int_s^T D_\varrho Z^{k+1}_r \mathrm{d}W_r. \end{cases} \tag{3.3}$$

We notice that, due to Lemma 3.1, the monotonicity condition is satisfied for the above FBSDEs (3.3). Moreover, Lemma 3.1 works once again to ensure the unique solvability of the following FBSDE:

$$\begin{cases} X^\varrho_s = [(\alpha_0+\delta)\sigma(\varrho, \Theta^{\alpha_0+\delta}_\varrho) + \psi_\varrho] + \displaystyle\int_\varrho^s \{(\alpha_0+\delta)\nabla b(r, \Theta^{\alpha_0+\delta}_r)\Theta^\varrho_r + D_\varrho \phi_r\} \mathrm{d}r \\ \qquad + \displaystyle\int_\varrho^s \{(\alpha_0+\delta)\nabla\sigma(r, \Theta^{\alpha_0+\delta}_r)\Theta^\varrho_r + D_\varrho \psi_r\} \mathrm{d}W_r, \\ Y^\varrho_s = \Big[(1-\alpha_0-\delta)\mu_1 G X^\varrho_T + (\alpha_0+\delta)\dfrac{\mathrm{d}\Phi}{\mathrm{d}x}(X^{\alpha_0+\delta}_T) X^\varrho_T + D_\varrho \xi\Big] \\ \qquad + \displaystyle\int_s^T \{(\alpha_0+\delta)\nabla g(r, \Theta^{\alpha_0+\delta}_r)\Theta^\varrho_r + (1-\alpha_0-\delta)\beta_1 G X^\varrho_r + D_\varrho \kappa_r\} \mathrm{d}r \\ \qquad - \displaystyle\int_s^T Z^\varrho_r \mathrm{d}W_r, \end{cases} \tag{3.4}$$

where $\Theta^\varrho := (X^\varrho, Y^\varrho, Z^\varrho)$.

Applying the L^2-estimate to Θ^ϱ leads to

$$\begin{aligned} \|\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)} \le C\mathbb{E}\Big\{ & 1 + |\Theta^{\alpha_0+\delta}_\varrho|^2 + |\psi_\varrho|^2 + |D_\varrho \xi|^2 \\ & + \int_\varrho^T [|D_\varrho \phi_r|^2 + |D_\varrho \psi_r|^2 + |D_\varrho \kappa_r|^2] \mathrm{d}r\Big\}. \end{aligned}$$

Hence,

$$\begin{aligned} \int_0^T \|\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)} \mathrm{d}\varrho \le C\mathbb{E}\Big\{ & 1 + \int_0^T |\Theta^{\alpha_0+\delta}_\varrho|^2 \mathrm{d}\varrho + \int_0^T |\psi_\varrho|^2 \mathrm{d}\varrho + \int_0^T |D_\varrho \xi|^2 \mathrm{d}\varrho \\ & + \int_0^T \int_0^T [|D_\varrho \phi_r|^2 + |D_\varrho \psi_r|^2 + |D_\varrho \kappa_r|^2] \mathrm{d}r \mathrm{d}\varrho\Big\}. \end{aligned}$$

With the help of the L^2-estimate of the solution $\Theta^{\alpha_0+\delta}$ to FBSDE (3.1),

$$\begin{aligned}\int_0^T \|\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\mathrm{d}\varrho &\le C\mathbb{E}\Big\{1+|\xi|^2+\int_0^T |D_\varrho\xi|^2\mathrm{d}\varrho+\int_0^T[|\phi_r|^2+|\psi_r|^2+|\kappa_r|^2]\mathrm{d}r\\ &\qquad+\int_0^T\int_0^T[|D_\varrho\phi_r|^2+|D_\varrho\psi_r|^2+|D_\varrho\kappa_r|^2]\mathrm{d}r\mathrm{d}\varrho\Big\}\\ &= C\{1+\|\xi\|^2_{\mathbb{D}_{1,2}}+\|\phi\|^2_{\mathbb{L}^a_{1,2}}+\|\psi\|^2_{\mathbb{L}^a_{1,2}}+\|\kappa\|^2_{\mathbb{L}^a_{1,2}}\}\le C. \end{aligned}\tag{3.5}$$

By applying the L^2-estimate to the difference of $D_\varrho\Theta^{k+1}$ and Θ^ϱ, we obtain

$$\|D_\varrho\Theta^{k+1}-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\le C(\alpha_0^2A^\varrho_{k+1}+\delta^2A^\varrho_k+\delta^2B^\varrho_k),\tag{3.6}$$

where

$$\begin{aligned}A^\varrho_k=\mathbb{E}\Big\{&|\sigma(\varrho,\Theta^k_\varrho)-\sigma(\varrho,\Theta^{\alpha_0+\delta}_\varrho)|^2+\Big|\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X^k_T)-\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X^{\alpha_0+\delta}_T)\Big|^2|X^\varrho_T|^2\\ &+\int_\varrho^T[|\nabla g(r,\Theta^k_r)-\nabla g(r,\Theta^{\alpha_0+\delta}_r)|^2+|\nabla b(r,\Theta^k_r)-\nabla b(r,\Theta^{\alpha_0+\delta}_r)|^2\\ &+|\nabla\sigma(r,\Theta^k_r)-\nabla\sigma(r,\Theta^{\alpha_0+\delta}_r)|^2]|\Theta^\varrho_r|^2\mathrm{d}r\Big\}\end{aligned}$$

and

$$\begin{aligned}B^\varrho_k=\mathbb{E}\Big\{&\Big|\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X^k_T)\Big|^2|D_\varrho X^k_T-X^\varrho_T|^2+|D_\varrho X^k_T-X^\varrho_T|^2+\int_\varrho^T|D_\varrho X^k_r-X^\varrho_r|^2\mathrm{d}r\\ &+\int_\varrho^T[|\nabla g(r,\Theta^k_r)|^2+|\nabla b(r,\Theta^k_r)|^2+|\nabla\sigma(r,\Theta^k_r)|^2]|D_\varrho\Theta^k_r-\Theta^\varrho_r|\mathrm{d}r\Big\}.\end{aligned}$$

With the help of (3.5), the fact that Θ^k converges to $\Theta^{\alpha_0+\delta}$ in $M^2_{\mathbb{F}}(0,T)$, and Lebesgue's dominated convergence theorem, we have

$$\lim_{k\to\infty}\int_0^T A^\varrho_k\mathrm{d}\varrho=0.$$

We also have

$$B^\varrho_k\le C\|D_\varrho\Theta^k-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}.$$

Then, (3.6) is deduced to

$$\|D_\varrho\Theta^{k+1}-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\le C\delta^2\|D_\varrho\Theta^k-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}+C(A^\varrho_{k+1}+A^\varrho_k).$$

Choose $\delta_0\in(0,\delta_1]$ such that $C\delta_0^2\le\frac{1}{4}$. Let $\delta\in[0,\delta_0]$. For any $\varepsilon>0$, there exists $N>0$ such that, for any $k\ge N$,

$$\int_0^T\|D_\varrho\Theta^{k+1}-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\mathrm{d}\varrho\le\varepsilon+\frac{1}{4}\int_0^T\|D_\varrho\Theta^k-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\mathrm{d}\varrho.$$

Thus, we obtain recursively, for every $k\ge N$,

$$\int_0^T\|D_\varrho\Theta^{k+1}-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\mathrm{d}\varrho\le\frac{4}{3}\varepsilon+\frac{1}{4^{k+1-N}}\int_0^T\|D_\varrho\Theta^N-\Theta^\varrho\|^2_{M^2_{\mathbb{F}}(\varrho,T)}\mathrm{d}\varrho.$$

Hence, it follows that $D_\varrho\Theta^k$ is a Cauchy sequence under the norm $\|\cdot\|_{\widetilde{\mathbb{L}^a_{1,2}}}$. Then, we have proved that Θ^k is a Cauchy sequence in $\mathbb{L}^a_{1,2}(0,T)$. Since $\mathbb{L}^a_{1,2}(0,T)$ is closed under the norm $\|\cdot\|_{\mathbb{L}^a_{1,2}(0,T)}$, the limit $\Theta^{\alpha_0+\delta}$ belongs to $\mathbb{L}^a_{1,2}(0,T)$ and a version of $D_\varrho\Theta^{\alpha_0+\delta}$ is given by Θ^ϱ. We complete the proof of the lemma.

Second case $\beta_1 \geq 0$, $\mu_1 \geq 0$, $\beta_2 > 0$ and $m \leq n$.

Instead of (3.1), we need to consider the following family of FBSDEs parameterized by $\alpha \in [0,1]$:

$$\begin{cases} \mathrm{d}X^\alpha_s = [(1-\alpha)\beta_2(-G^{\mathrm{T}}Y^\alpha_s) + \alpha b(s,\Theta^\alpha_s) + \phi_s]\mathrm{d}s \\ \qquad\qquad + [(1-\alpha)\beta_2(-G^{\mathrm{T}}Z^\alpha_s) + \alpha\sigma(s,\Theta^\alpha_s) + \psi_s]\mathrm{d}W_s, \\ -\mathrm{d}Y^\alpha_s = [\alpha g(s,\Theta^\alpha_s) + \kappa_s]\mathrm{d}s - Z^\alpha_s\mathrm{d}W_s, \\ X^\alpha_t = x, \quad Y^\alpha_T = \alpha\Phi(X^\alpha_T) + \xi. \end{cases} \tag{3.7}$$

Similar to Lemma 3.2, for the second case, we can prove the following lemma.

Lemma 3.3 *Let Assumptions* (H3) *and* (H4) *hold. Then there exists an absolute constant* $\delta_0 > 0$*, such that if for some* $\alpha_0 \in [0,1)$*, the solution to* FBSDE (3.7) *is differentiable in Malliavin's sense for any* ϕ*,* $\psi \in L^2_{\mathbb{F}}(t,T;\mathbb{R}^n) \cap \mathbb{L}^a_{1,2}(t,T;\mathbb{R}^n)$*,* $\kappa \in L^2_{\mathbb{F}}(t,T;\mathbb{R}^m) \cap \mathbb{L}^a_{1,2}(t,T;\mathbb{R}^m)$ *and* $\xi \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^m) \cap (\mathbb{D}_{1,2})^m$*, then the same is true for* $\alpha = \alpha_0 + \delta$ *with* $\delta \in [0,\delta_0]$*,* $\alpha_0 + \delta \leq 1$.

Theorem 3.1 *Let Assumptions* (H3) *and* (H4) *hold. Then the solution* $\Theta^{t,x}$ *to* FBSDE (1.2) *belongs to* $\mathbb{L}^a_{1,2}(0,T;\mathbb{R}^{n+m+m\times d})$*, and a version of* $\{D_\varrho\Theta^{t,x}_s := (D_\varrho X^{t,x}_s, D_\varrho Y^{t,x}_s, D_\varrho Z^{t,x}_s), s \in [\varrho,T]\}$ *is given by*

$$\begin{cases} D_\varrho X^{t,x}_s = \sigma(\varrho,\Theta^{t,x}_\varrho) + \displaystyle\int_\varrho^s \nabla b(r,\Theta^{t,x}_r)D_\varrho\Theta^{t,x}_r\mathrm{d}r + \int_\varrho^s \nabla\sigma(r,\Theta^{t,x}_r)D_\varrho\Theta^{t,x}_r\mathrm{d}W_r, \\ D_\varrho Y^{t,x}_s = \dfrac{\mathrm{d}\Phi}{\mathrm{d}x}(X^{t,x}_T)D_\varrho X^{t,x}_T + \displaystyle\int_s^T \nabla g(r,\Theta^{t,x}_r)D_\varrho\Theta^{t,x}_r\mathrm{d}r - \int_s^T D_\varrho Z^{t,x}_r\mathrm{d}W_r. \end{cases} \tag{3.8}$$

Moreover, $\{D_sY^{t,x}_s;\ s \in [t,T]\}$ *defined by* (3.8) *is a version of* $\{Z^{t,x}_s;\ s \in [t,T]\}$.

Proof For the first case, by Lemma 3.2, we can establish the Malliavin's differentiability for FBSDEs (3.1) with any (ϕ,ψ,κ,ξ) and $\alpha \in [0,1]$. In particular, (3.1) with $(\phi,\psi,\kappa,\xi) = 0$ and $\alpha = 1$, which is (1.2), is differentiable in Malliavin's sense, and (3.4) coincides with (3.8). For the second case, we consider (3.7) and use Lemma 3.3 to get the same conclusion.

Now the remaining thing is to prove $D_sY_s = Z_s$. The same as El Karoui, Peng and Quenez [7], we notice that

$$Y_s = Y_t - \int_t^s g(r,\Theta_r)\mathrm{d}r + \int_t^s Z_r\mathrm{d}W_r, \quad s \in [t,T].$$

Then, by [7, Lemma 5.1],

$$D_\varrho Y_s = Z_\varrho - \int_\varrho^s \nabla g(r,\Theta_r)D_\varrho\Theta_r\mathrm{d}r + \int_\varrho^s D_\varrho Z_r\mathrm{d}W_r, \quad s \in [\varrho,T].$$

By taking $s = \varrho$, we get $D_\varrho Y_\varrho = Z_\varrho$, then the proof is finished.

3.2 Continuous differentiability of u with respect to x

By Corollary 2.1, the function u defined by (2.31) is continuous in (t,x). Now we continue to study the continuous differentiability with respect to x under Assumptions (H3) and (H4). Let $(t,x)\in[0,T]\times\mathbb{R}^n$ be given. Let $\{e_1,e_2,\cdots,e_n\}$ denote an orthonormal basis of $\mathbb{R}^n$. For any $i=1,2,\cdots,n$ and $h\in\mathbb{R}\setminus\{0\}$, we define

$$\begin{cases}\triangle_h^i X_s^{t,x}:=h^{-1}[X_s^{t,x+he_i}-X_s^{t,x}],\\ \triangle_h^i Y_s^{t,x}:=h^{-1}[Y_s^{t,x+he_i}-Y_s^{t,x}], \qquad s\in[t,T].\\ \triangle_h^i Z_s^{t,x}:=h^{-1}[Z_s^{t,x+he_i}-Z_s^{t,x}],\end{cases}$$

Then $\triangle_h^i\Theta^{t,x}=(\triangle_h^i X^{t,x},\triangle_h^i Y^{t,x},\triangle_h^i Z^{t,x})$ satisfies the following FBSDE:

$$\begin{cases}\triangle_h^i X_s^{t,x}=e_i+\int_t^s\int_0^1\nabla b(r,\Theta_r^{t,x}+\lambda h\triangle_h^i\Theta_r^{t,x})\triangle_h^i\Theta_r^{t,x}\mathrm{d}\lambda\mathrm{d}r\\ \qquad\qquad+\int_t^s\int_0^1\nabla\sigma(r,\Theta_r^{t,x}+\lambda h\triangle_h^i\Theta_r^{t,x})\triangle_h^i\Theta_r^{t,x}\mathrm{d}\lambda\mathrm{d}W_r,\\ \triangle_h^i Y_s^{t,x}=\int_0^1\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x}+\lambda h\triangle_h^i X_T^{t,x})\triangle_h^i X_T^{t,x}\mathrm{d}\lambda-\int_s^T\triangle_h^i Z_r^{t,x}\mathrm{d}W_r\\ \qquad\qquad+\int_s^T\int_0^1\nabla g(r,\Theta_r^{t,x}+\lambda h\triangle_h^i\Theta_r^{t,x})\triangle_h^i\Theta_r^{t,x}\mathrm{d}\lambda\mathrm{d}r.\end{cases}\tag{3.9}$$

For any $i=1,2,\cdots,n$, we also denote by $\frac{\partial\Theta^{t,x}}{\partial x_i}:=\big(\frac{\partial X^{t,x}}{\partial x_i},\frac{\partial Y^{t,x}}{\partial x_i},\frac{\partial Z^{t,x}}{\partial x_i}\big)$ the unique solution to the following FBSDE:

$$\begin{cases}\frac{\partial X_s^{t,x}}{\partial x_i}=e_i+\int_t^s\nabla b(r,\Theta_r^{t,x})\frac{\partial\Theta_r^{t,x}}{\partial x_i}\mathrm{d}r+\int_t^s\nabla\sigma(r,\Theta_r^{t,x})\frac{\partial\Theta_r^{t,x}}{\partial x_i}\mathrm{d}W_r,\\ \frac{\partial Y_s^{t,x}}{\partial x_i}=\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x})\frac{\partial X_T^{t,x}}{\partial x_i}+\int_s^T\nabla g(r,\Theta_r^{t,x})\frac{\partial\Theta_r^{t,x}}{\partial x_i}\mathrm{d}r-\int_s^T\frac{\partial Z_r^{t,x}}{\partial x_i}\mathrm{d}W_r.\end{cases}\tag{3.10}$$

Equivalently, by setting $\nabla\Theta^{t,x}=\big(\frac{\partial\Theta^{t,x}}{\partial x_1},\frac{\partial\Theta^{t,x}}{\partial x_2},\cdots,\frac{\partial\Theta^{t,x}}{\partial x_n}\big)$ as well as setting $\nabla X^{t,x}$, $\nabla Y^{t,x}$ and $\nabla Z^{t,x}$ similarly, we collect the above FBSDEs from $i=1$ to n in the following form:

$$\begin{cases}\nabla X_s^{t,x}=I+\int_t^s\nabla b(r,\Theta_r^{t,x})\nabla\Theta_r^{t,x}\mathrm{d}r+\int_t^s\nabla\sigma(r,\Theta_r^{t,x})\nabla\Theta_r^{t,x}\mathrm{d}W_r,\\ \nabla Y_s^{t,x}=\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x})\nabla X_T^{t,x}+\int_s^T\nabla g(r,\Theta_r^{t,x})\nabla\Theta_r^{t,x}\mathrm{d}r-\int_s^T\nabla Z_r^{t,x}\mathrm{d}W_r,\end{cases}\tag{3.11}$$

where I denotes the $(n\times n)$ identity matrix. We shall later interpret $\nabla X_s^{t,x}$ (resp. $\nabla Y_s^{t,x}$, $\nabla Z_s^{t,x}$) as the matrix of first order partial derivatives of $X_s^{t,x}$ (resp. $Y_s^{t,x}$, $Z_s^{t,x}$) with respect to x.

Due to Lemma 3.1, the monotonicity condition is satisfied for FBSDEs (3.9) and (3.10) (or (3.11)). Then these FBSDEs admit unique L^2-solutions, and we can use the L^2-estimate to get

$$\|\triangle_h^i\Theta^{t,x}\|^2_{M^2_{\mathbb{F}}(t,T)}\le C_2,\quad\Big\|\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,T)}\le C_2,\quad i=1,2,\cdots,n\tag{3.12}$$

and

$$\Big\|\triangle_h^i\Theta^{t,x}-\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,T)}\le C(\mathrm{I}^\Phi+\mathrm{I}^g+\mathrm{I}^b+\mathrm{I}^\sigma),$$

where

$$\mathrm{I}^\Phi=\mathbb{E}\Big[\Big|\int_0^1\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x}+\lambda h\triangle_h^iX_T^{t,x})\mathrm{d}\lambda-\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x})\Big|^2\Big|\frac{\partial X_T^{t,x}}{\partial x_i}\Big|^2\Big],$$
$$\mathrm{I}^f=\mathbb{E}\int_t^T\Big|\int_0^1\nabla f(r,\Theta_r^{t,x}-\lambda h\triangle_h^i\Theta_r^{t,x})\mathrm{d}\lambda-\nabla f(r,\Theta_r^{t,x})\Big|^2\Big|\frac{\partial\Theta_r^{t,x}}{\partial x_i}\Big|^2\mathrm{d}r\quad\text{with } f=g,b,\sigma.$$

With the help of (3.12), from Lebesgue's dominated convergence theorem, we have

$$\lim_{h\to0}\Big\|\triangle_h^i\Theta^{t,x}-\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,T)}=0. \tag{3.13}$$

Then by the definition of partial derivatives, $\nabla\Theta^{t,x}$, of the unique solution to FBSDE (3.11), is the gradient of $\Theta^{t,x}$ with respect to x. In particular, when $s=t$, the function u is partially differentiable with respect to x, and

$$\nabla u(t,x)=\nabla Y_t^{t,x}.$$

Moreover, from (3.12), ∇u is bounded.

Next, we show that ∇u is continuous. Let (t,x), $(t',x')\in[0,T]\times\mathbb{R}^n$. For the case $t'\le t$, by applying the L^2-estimate on the interval $[t',T]$, we have

$$\begin{aligned}\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i}-\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(0,T)}&=\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i}-\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t',T)}\\&\le C(\mathrm{II}^\Phi+\mathrm{II}^g+\mathrm{II}^b+\mathrm{II}^\sigma+\mathrm{III}^g+\mathrm{III}^b+\mathrm{III}^\sigma),\end{aligned}$$

where

$$\mathrm{II}^\Phi=\mathbb{E}\Big[\Big|\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t',x'})-\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x})\Big|^2\Big|\frac{\partial X_T^{t,x}}{\partial x_i}\Big|^2\Big],$$
$$\mathrm{II}^f=\mathbb{E}\int_t^T|\nabla f(s,\Theta_s^{t',x'})-\nabla f(s,\Theta_s^{t,x})|^2\Big|\frac{\partial\Theta_s^{t,x}}{\partial x_i}\Big|^2\mathrm{d}s\quad\text{with } f=g,b,\sigma,$$
$$\mathrm{III}^f=\mathbb{E}\int_{t'}^t\Big|\frac{\partial f}{\partial x}(s,\Theta_s^{t',x'})e_i+\frac{\partial f}{\partial y}(s,\Theta_s^{t',x'})\frac{\partial Y_t^{t,x}}{\partial x_i}\Big|^2\mathrm{d}s\quad\text{with } f=g,b,\sigma.$$

Firstly, we estimate III^f. With the help of (3.12),

$$\mathrm{III}^f\le C\int_{t'}^t\Big(1+\Big|\frac{\partial Y_t^{t,x}}{\partial x_i}\Big|^2\Big)\mathrm{d}s\le C(t-t').$$

Secondly, we analyze II^Φ. From (2.33), we have

$$\mathbb{E}[|X_T^{t',x'}-X_T^{t,x}|^2]\le C\{|x'-x|+(1+|x|^2+|x'|^2)|t'-t|\}.$$

Then $X_T^{t',x'}$ converges to $X_T^{t,x}$ in $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ as $(t',x')\to(t^-,x)$. Consequently, $X_T^{t',x'}$ converges to $X_T^{t,x}$ in probability $\mathbb{P}$. With the help of (3.12) and Lebesgue's dominated convergence theorem, we get

$$\lim_{(t',x')\to(t^-,x)}\mathrm{II}^\Phi=0.$$

Finally, a similar analysis as II^{Φ} leads to

$$\lim_{(t',x')\to(t^-,x)} \mathrm{II}^f = 0.$$

In summary,

$$\lim_{(t',x')\to(t^-,x)} \Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i} - \frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(0,T)} = 0. \tag{3.14}$$

For the other case $t \le t'$, we do not apply the L^2-estimate on the whole interval $[t,T]$. Instead, we apply it only on the interval $[t',T]$, which leads to

$$\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i} - \frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t',T)} \le C\Big(\mathbb{E}\Big[\Big|\frac{\partial X^{t,x}_{t'}}{\partial x_i} - e_i\Big|^2\Big] + \mathrm{II}^{\Phi} + \widetilde{\mathrm{II}}^g + \widetilde{\mathrm{II}}^b + \widetilde{\mathrm{II}}^{\sigma}\Big),$$

where II^{Φ} is defined in the previous paragraph, and

$$\widetilde{\mathrm{II}}^f = \mathbb{E}\int_{t'}^T |\nabla f(s,\Theta^{t',x'}_s) - \nabla f(s,\Theta^{t,x}_s)|^2 \Big|\frac{\partial\Theta^{t,x}_s}{\partial x_i}\Big|^2 \mathrm{d}s \quad \text{with } f = g, b, \sigma.$$

For simplicity, we also introduce the notation

$$\widetilde{\mathrm{III}}^f = \mathbb{E}\int_t^{t'} |\nabla f(s,\Theta^{t,x}_s)|^2 \Big|\frac{\partial\Theta^{t,x}_s}{\partial x_i}\Big|^2 \mathrm{d}s \quad \text{with } f = g, b, \sigma.$$

By considering the forward equation in (3.10) on the interval $[t,t']$ and with the help of Burkholder-Davis-Gundy inequality, we get

$$\mathbb{E}\Big[\sup_{s\in[t,t']}\Big|\frac{\partial X^{t,x}_s}{\partial x_i} - e_i\Big|^2\Big] \le C(\widetilde{\mathrm{III}}^b + \widetilde{\mathrm{III}}^{\sigma}).$$

Then,

$$\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i} - \frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t',T)} \le C(\mathrm{II}^{\Phi} + \widetilde{\mathrm{II}}^g + \widetilde{\mathrm{II}}^b + \widetilde{\mathrm{II}}^{\sigma} + \widetilde{\mathrm{III}}^b + \widetilde{\mathrm{III}}^{\sigma}).$$

Similar to the previous paragraph, we can check that II^{Φ}, $\widetilde{\mathrm{II}}^f$ and $\widetilde{\mathrm{III}}^f$ tend to zero as $(t',x') \to (t^+,x)$. Therefore,

$$\lim_{(t',x')\to(t^+,x)} \Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i} - \frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t',T)} = 0. \tag{3.15}$$

On the interval $[t,t']$, we notice that $\frac{\partial\Theta^{t',x'}}{\partial x_i} \equiv \big(e_i, \frac{\partial Y^{t',x'}_{t'}}{\partial x_i}, 0\big)$. Burkholder-Davis-Gundy inequality works once again to yield

$$\begin{aligned}
&\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i} - \frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,t')} \\
&\le C\Big(\mathbb{E}\Big|\frac{\partial Y^{t,x}_{t'}}{\partial x_i} - \frac{\partial Y^{t',x'}_{t'}}{\partial x_i}\Big|^2 + \mathbb{E}\int_t^{t'}\Big|\frac{\partial Z^{t,x}_s}{\partial x_i}\Big|^2\mathrm{d}s + \widetilde{\mathrm{III}}^g + \widetilde{\mathrm{III}}^b + \widetilde{\mathrm{III}}^{\sigma}\Big) \\
&\le C\Big(\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i} - \frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t',T)} + \mathbb{E}\int_t^{t'}\Big|\frac{\partial Z^{t,x}_s}{\partial x_i}\Big|^2\mathrm{d}s + \widetilde{\mathrm{III}}^g + \widetilde{\mathrm{III}}^b + \widetilde{\mathrm{III}}^{\sigma}\Big).
\end{aligned}$$

From (3.12) and (3.15), we have

$$\lim_{(t',x')\to(t^+,x)} \Big\| \frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i} \Big\|^2_{M^2_{\mathbb{F}}(t,t')} = 0. \tag{3.16}$$

It is obvious that

$$\Big\| \frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i} \Big\|^2_{M^2_{\mathbb{F}}(0,T)} \le \Big\| \frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i} \Big\|^2_{M^2_{\mathbb{F}}(t',T)} + \Big\| \frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i} \Big\|^2_{M^2_{\mathbb{F}}(t,t')}.$$

Then, (3.15)–(3.16) imply

$$\lim_{(t',x')\to(t^+,x)} \Big\| \frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i} \Big\|^2_{M^2_{\mathbb{F}}(0,T)} = 0. \tag{3.17}$$

Combining (3.14) and (3.17), we obtain

$$\lim_{(t',x')\to(t,x)} \Big\| \frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i} \Big\|^2_{M^2_{\mathbb{F}}(0,T)} = 0. \tag{3.18}$$

In particular, ∇u is continuous in (t,x).

We summarize the above analysis as follows.

Lemma 3.4 *Let Assumptions* (H3) *and* (H4) *hold. Then* $u \in C^{0,1}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$. *Moreover, all the partial derivatives are bounded on* $[0,T]\times\mathbb{R}^n$.

3.3 Twice continuous differentiability of u with respect to x

It should be noticed that just the L^2-estimates of coupled FBSDEs are involved in our analysis of the above two subsections. However, when we consider the same issues on the second order derivatives of u, there are some additional non-homogeneous items appearing in the corresponding equations (compare (3.9) and (3.10) with (3.22) and (3.23) below), which brings us difficulties. Since the non-homogeneous items will appear in quadratic forms of $\frac{\partial \Theta^{t,x}}{\partial x_i}$ or/and $\triangle^i_h \Theta^{t,x}$ $(i = 1, 2, \cdots, n)$, it seems the L^2-estimates will not be enough for our analysis below. Due to this, we shall employ L^4-estimates by imposing Assumption $(\mathrm{H2})^4$ in this subsection. Besides Assumption $(\mathrm{H2})^4$, we shall also need the following assumption.

(H5) The coefficient σ depends linearly on z, i.e., σ is in the form: For any $t \in [0,T]$ and any $\theta = (x,y,z) \in \mathbb{R}^{n+m+m\times d}$,

$$\sigma(t,\theta) = \sigma_0(t,x,y) + \sum_{p=1}^{m}\sum_{q=1}^{d} \sigma^{pq}(t,x,y) z_{pq},$$

where σ^{pq} takes values in $\mathbb{R}^{m\times d}$ and z_{pq} is the element of matrix z located in the p-th row and the q-th column.

Consequently, (H5) implies that the Hessian matrix $\frac{\partial^2 \sigma}{\partial z^2} \equiv 0$. The following analysis in this subsection is under Assumptions $(\mathrm{H2})^4$, (H3), (H4) and (H5).

It is clear that in the present smoothness situation, the constant L_x (involved in Assumption (H2)^4) is the bound of $\frac{\mathrm{d}\Phi}{\mathrm{d}x}$, L_z is the bound of $(\sigma^{pq}(t,x,y))_{m\times d}$, and L is the bound of other first order partial derivatives of coefficients $\gamma=(g,b,\sigma)$ with respect to θ.

First of all, under Assumption (H2)^4, (H3) and (H4), replacing L^2-estimates with L^4-estimates, we can improve (3.12)–(3.13) and (3.18) to the following:

$$\|\triangle_h^i\Theta^{t,x}\|^4_{M^4_{\mathbb{F}}(t,T)}\le C_4,\quad \Big\|\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^4_{M^4_{\mathbb{F}}(t,T)}\le C_4, \tag{3.19}$$

$$\lim_{h\to 0}\Big\|\triangle_h^i\Theta^{t,x}-\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^4_{M^4_{\mathbb{F}}(t,T)}=0 \tag{3.20}$$

and

$$\lim_{(t',x')\to(t,x)}\Big\|\frac{\partial\Theta^{t',x'}}{\partial x_i}-\frac{\partial\Theta^{t,x}}{\partial x_i}\Big\|^4_{M^4_{\mathbb{F}}(0,T)}=0. \tag{3.21}$$

The proofs are similar to those appearing in the previous subsection, hence we omit them.

Now, similar to the previous subsection, for any $i,j=1,2,\cdots,n$ and $h\in\mathbb{R}\setminus\{0\}$, we introduce the notations

$$\begin{cases}\triangle_h^j\dfrac{\partial X_s^{t,x}}{\partial x_i}:=h^{-1}\Big[\dfrac{\partial X_s^{t,x+he_j}}{\partial x_i}-\dfrac{\partial X_s^{t,x}}{\partial x_i}\Big],\\ \triangle_h^j\dfrac{\partial Y_s^{t,x}}{\partial x_i}:=h^{-1}\Big[\dfrac{\partial Y_s^{t,x+he_j}}{\partial x_i}-\dfrac{\partial Y_s^{t,x}}{\partial x_i}\Big], \qquad s\in[t,T].\\ \triangle_h^j\dfrac{\partial Z_s^{t,x}}{\partial x_i}:=h^{-1}\Big[\dfrac{\partial Z_s^{t,x+he_j}}{\partial x_i}-\dfrac{\partial Z_s^{t,x}}{\partial x_i}\Big].\end{cases}$$

Then $\triangle_h^j\frac{\partial\Theta^{t,x}}{\partial x_i}=\big(\triangle_h^j\frac{\partial X^{t,x}}{\partial x_i},\triangle_h^j\frac{\partial Y^{t,x}}{\partial x_i},\triangle_h^j\frac{\partial Z^{t,x}}{\partial x_i}\big)$ satisfies the following FBSDE:

$$\begin{cases}\triangle_h^j\dfrac{\partial X_s^{t,x}}{\partial x_i}=\displaystyle\int_t^s\int_0^1\Big\langle D^2b(r,\Theta_r^{t,x}+\lambda h\triangle_h^j\Theta_r^{t,x})\triangle_h^j\Theta_r^{t,x},\ \frac{\partial\Theta_r^{t,x}}{\partial x_i}\Big\rangle\mathrm{d}\lambda\mathrm{d}r\\ \qquad+\displaystyle\int_t^s\int_0^1\Big\langle D^2\sigma(r,\Theta_r^{t,x}+\lambda h\triangle_h^j\Theta_r^{t,x})\triangle_h^j\Theta_r^{t,x},\ \frac{\partial\Theta_r^{t,x}}{\partial x_i}\Big\rangle\mathrm{d}\lambda\mathrm{d}W_r\\ \qquad+\displaystyle\int_t^s\nabla b(r,\Theta_r^{t,x+he_j})\triangle_h^j\frac{\partial\Theta_r^{t,x}}{\partial x_i}\mathrm{d}r+\int_t^s\nabla\sigma(r,\Theta_r^{t,x+he_j})\triangle_h^j\frac{\partial\Theta_r^{t,x}}{\partial x_i}\mathrm{d}W_r,\\ \triangle_h^j\dfrac{\partial Y_s^{t,x}}{\partial x_i}=\displaystyle\int_0^1\Big\langle\frac{\mathrm{d}^2\Phi}{\mathrm{d}x^2}(X_T^{t,x}+\lambda h\triangle_h^jX_T^{t,x})\triangle_h^jX_T^{t,x},\ \frac{\partial X_T^{t,x}}{\partial x_i}\Big\rangle\mathrm{d}\lambda\\ \qquad+\displaystyle\int_t^s\int_0^1\Big\langle D^2g(r,\Theta_r^{t,x}+\lambda h\triangle_h^j\Theta_r^{t,x})\triangle_h^j\Theta_r^{t,x},\ \frac{\partial\Theta_r^{t,x}}{\partial x_i}\Big\rangle\mathrm{d}\lambda\mathrm{d}r\\ \qquad+\displaystyle\frac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x+he_j})\triangle_h^j\frac{\partial X_T^{t,x}}{\partial x_i}+\int_s^T\nabla g(r,\Theta_r^{t,x+he_j})\triangle_h^j\frac{\partial\Theta_r^{t,x}}{\partial x_i}\mathrm{d}r\\ \qquad-\displaystyle\int_s^T\triangle_h^j\frac{\partial Z_r^{t,x}}{\partial x_i}\mathrm{d}W_r.\end{cases} \tag{3.22}$$

Here, for an $\mathbb{R}^n$-valued function $f(\cdot)=(f^1(\cdot),f^2(\cdot),\cdots,f^n(\cdot))^{\mathrm{T}}$, we use the notation $\langle D^2f(\cdot)\theta,\ \overline{\theta}\rangle$

to denote another $\mathbb{R}^n$-valued function in which the k-th component is $\langle D^2 f^k(\cdot)\theta,\ \overline{\theta}\rangle$ and

$$D^2 f^k(\cdot) = \begin{pmatrix} \dfrac{\partial^2 f^k}{\partial x^2}(\cdot) & \dfrac{\partial^2 f^k}{\partial y \partial x}(\cdot) & \dfrac{\partial^2 f^k}{\partial z \partial x}(\cdot) \\ \dfrac{\partial^2 f^k}{\partial x \partial y}(\cdot) & \dfrac{\partial^2 f^k}{\partial y^2}(\cdot) & \dfrac{\partial^2 f^k}{\partial z \partial y}(\cdot) \\ \dfrac{\partial^2 f^k}{\partial x \partial z}(\cdot) & \dfrac{\partial^2 f^k}{\partial y \partial z}(\cdot) & \dfrac{\partial^2 f^k}{\partial z^2}(\cdot) \end{pmatrix}.$$

Moreover, for an $\mathbb{R}^{n\times d}$-valued function $f(\cdot)$, the notation $\langle D^2 f(\cdot)\theta,\ \overline{\theta}\rangle$ has a similar meaning. Besides FBSDE (3.22), we also introduce another FBSDE as follows:

$$\begin{cases} \dfrac{\partial^2 X_s^{t,x}}{\partial x_j \partial x_i} = \displaystyle\int_t^s \Big\langle D^2 b(r,\Theta_r^{t,x})\frac{\partial \Theta_r^{t,x}}{\partial x_j},\ \frac{\partial \Theta_r^{t,x}}{\partial x_i}\Big\rangle \mathrm{d}r + \int_t^s \nabla b(r,\Theta_r^{t,x})\frac{\partial^2 \Theta_r^{t,x}}{\partial x_j \partial x_i}\mathrm{d}r \\ \qquad\qquad + \displaystyle\int_t^s \Big\langle D^2 \sigma(r,\Theta_r^{t,x})\frac{\partial \Theta_r^{t,x}}{\partial x_j},\ \frac{\partial \Theta_r^{t,x}}{\partial x_i}\Big\rangle \mathrm{d}W_r + \int_t^s \nabla \sigma(r,\Theta_r^{t,x})\frac{\partial^2 \Theta_r^{t,x}}{\partial x_j \partial x_i}\mathrm{d}W_r, \\ \dfrac{\partial^2 Y_s^{t,x}}{\partial x_j \partial x_i} = \Big\langle \dfrac{\mathrm{d}^2\Phi}{\mathrm{d}x^2}(X_T^{t,x})\dfrac{\partial X_T^{t,x}}{\partial x_j},\ \dfrac{\partial X_T^{t,x}}{\partial x_i}\Big\rangle + \dfrac{\mathrm{d}\Phi}{\mathrm{d}x}(X_T^{t,x})\dfrac{\partial^2 X_T^{t,x}}{\partial x_j \partial x_i} \\ \qquad\qquad + \displaystyle\int_t^s \Big\langle D^2 g(r,\Theta_r^{t,x})\frac{\partial \Theta_r^{t,x}}{\partial x_j},\ \frac{\partial \Theta_r^{t,x}}{\partial x_i}\Big\rangle \mathrm{d}r + \int_t^s \nabla g(r,\Theta_r^{t,x})\frac{\partial^2 \Theta_r^{t,x}}{\partial x_j \partial x_i}\mathrm{d}r \\ \qquad\qquad - \displaystyle\int_s^T \frac{\partial^2 Z_r^{t,x}}{\partial x_j \partial x_i}\mathrm{d}W_r. \end{cases} \tag{3.23}$$

As in the previous subsection, Assumption (H4) ensures that the Lipschitz condition holds true for FBSDEs (3.22) and (3.23). Moreover, with the help of Lemma 3.1 and Assumption (H3), the monotonicity condition is also satisfied for the above two FBSDEs. However, when we continue to check the corresponding square integrability conditions for the coefficients of FBSDEs (3.22) and (3.23), the following two items

$$\Big\langle \Big[\int_0^1 \frac{\partial^2 \sigma}{\partial z^2}(r,\Theta_r^{t,x} + \lambda h \triangle_h^j \Theta_r^{t,x})\mathrm{d}\lambda\Big]\triangle_h^j Z_r^{t,x},\ \frac{\partial Z_r^{t,x}}{\partial x_i}\Big\rangle, \quad r\in[t,T] \tag{3.24}$$

and

$$\Big\langle \frac{\partial^2 \sigma}{\partial z^2}(r,\Theta_r^{t,x})\frac{\partial Z_r^{t,x}}{\partial x_j},\ \frac{\partial Z_r^{t,x}}{\partial x_i}\Big\rangle, \quad r\in[t,T] \tag{3.25}$$

appear in the diffusion coefficients, which bring us a difficulty: For fixed $\omega\in\Omega$, we require that the above two items are square integrable with respect to the time variable $r\in[t,T]$. However, $\triangle_h^j Z^{t,x}$, $\frac{\partial Z^{t,x}}{\partial x_i}$ and $\frac{\partial Z^{t,x}}{\partial x_j}$ are also known to be only square integrable. Therefore, the square integrability requirements of (3.24)–(3.25) cannot be satisfied in general. To avoid this difficulty, we introduce Assumption (H5). Under Assumption (H5), the estimates (3.19) ensure that the corresponding square integrability conditions for the coefficients of FBSDEs (3.22) and (3.23) are satisfied. Therefore, FBSDEs (3.22) and (3.23) admit unique L^2-solutions and the corresponding L^2-estimates can be used in our analysis below.

The next analysis is a bit complicated, but is very similar to the previous subsection. Then we would like to provide a brief derivation procedure. By virtue of (3.19), we have

$$\Big\|\triangle_h^j \frac{\partial \Theta^{t,x}}{\partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,T)} \le C, \quad \Big\|\frac{\partial^2 \Theta^{t,x}}{\partial x_j \partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,T)} \le C. \tag{3.26}$$

With the help of Lebesgue's dominated convergence theorem as well as (3.19)–(3.20) and (3.26), we obtain

$$\lim_{h\to 0}\Big\|\triangle_h^j \frac{\partial \Theta^{t,x}}{\partial x_i} - \frac{\partial^2 \Theta^{t,x}}{\partial x_j \partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(t,T)} = 0. \tag{3.27}$$

The above equation implies that the function u is twice partially differentiable with respect to x, and

$$\frac{\partial^2 u}{\partial x_j \partial x_i}(t,x) = \frac{\partial^2 Y_t^{t,x}}{\partial x_j \partial x_i}, \quad i,j = 1,2,\cdots,n.$$

Moreover, (3.26) implies $\frac{\partial^2 u}{\partial x_j \partial x_i}$ is bounded for any $i,j = 1,2,\cdots,n$.

In order to exhibit the treatment techniques on the non-homogeneous items appearing in this subsection, we would like to indicate how we treat the following "hard" term: When we analyze the "left continuity" of the second order partial derivatives, we need to prove

$$\lim_{(t',x')\to(t^-,x)} \mathbb{E}\int_{t'}^t \Big|\frac{\partial X_s^{t',x'}}{\partial x_j}\Big|^2 \Big|\frac{\partial Z_s^{t',x'}}{\partial x_i}\Big|^2 \mathrm{d}s = 0. \tag{3.28}$$

In fact, by Hölder's inequality,

$$\mathbb{E}\int_{t'}^t \Big|\frac{\partial X_s^{t',x'}}{\partial x_j}\Big|^2 \Big|\frac{\partial Z_s^{t',x'}}{\partial x_i}\Big|^2 \mathrm{d}s \le \Big\{\mathbb{E}\Big[\sup_{s\in[t',t]}\Big|\frac{\partial X_s^{t',x'}}{\partial x_j}\Big|^4\Big]\Big\}^{\frac{1}{2}} \Big\{\mathbb{E}\Big[\Big(\int_{t'}^t \Big|\frac{\partial Z_s^{t',x'}}{\partial x_i}\Big|^2 \mathrm{d}s\Big)^2\Big]\Big\}^{\frac{1}{2}}.$$

From (3.19), the first item on the right hand side of the above inequality is bounded by $C_4^{\frac{1}{2}}$. Moreover, we notice that on the time interval $[t',t]$, $\frac{\partial Z^{t,x}}{\partial x_i} \equiv 0$. Then

$$\begin{aligned}\mathbb{E}\Big[\Big(\int_{t'}^t \Big|\frac{\partial Z_s^{t',x'}}{\partial x_i}\Big|^2 \mathrm{d}s\Big)^2\Big] &= \mathbb{E}\Big[\Big(\int_{t'}^t \Big|\frac{\partial Z_s^{t',x'}}{\partial x_i} - \frac{\partial Z_s^{t,x}}{\partial x_i}\Big|^2 \mathrm{d}s\Big)^2\Big] \\ &= \Big\|\frac{\partial Z^{t',x'}}{\partial x_i} - \frac{\partial Z^{t,x}}{\partial x_i}\Big\|^4_{L^4_{\mathbb{F}}(\Omega;L^2(t',t))} \le \Big\|\frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i}\Big\|^4_{M^4_{\mathbb{F}}(t',t)}.\end{aligned}$$

Therefore,

$$\mathbb{E}\int_{t'}^t \Big|\frac{\partial X_s^{t',x'}}{\partial x_j}\Big|^2 \Big|\frac{\partial Z_s^{t',x'}}{\partial x_i}\Big|^2 \mathrm{d}s \le C_4^{\frac{1}{2}} \Big\|\frac{\partial \Theta^{t',x'}}{\partial x_i} - \frac{\partial \Theta^{t,x}}{\partial x_i}\Big\|^2_{M^4_{\mathbb{F}}(t',t)}.$$

Due to (3.21), we obtain (3.28).

We continue our analysis. By virtue of Lebesgue's dominated convergence theorem, Hölder's inequality, Burkholder-Davis-Gundy inequality as well as (3.19), (3.21) and (3.26), we successfully obtain

$$\lim_{(t',x')\to(t,x)}\Big\|\frac{\partial^2 \Theta^{t',x'}}{\partial x_j \partial x_i} - \frac{\partial^2 \Theta^{t,x}}{\partial x_j \partial x_i}\Big\|^2_{M^2_{\mathbb{F}}(0,T)} = 0. \tag{3.29}$$

In particular, $\frac{\partial^2 u}{\partial x_j \partial x_i}$ is continuous in (t,x) for any $i,j = 1,2,\cdots,n$.

In summary, we have the following lemma.

Lemma 3.5 *Let Assumptions* (H2)^4, (H3), (H4) *and* (H5) *hold. Then* $u \in C^{0,2}([0,T] \times \mathbb{R}^n; \mathbb{R}^m)$. *Moreover, all the partial derivatives are bounded on* $[0,T] \times \mathbb{R}^n$.

Remark 3.1 In this subsection, due to the non-homogeneous items appearing in FBSDEs (3.22) and (3.23), we impose Assumptions (H2)^4 and (H5). Instead of them, we can also introduce the following assumption.

(H6) All the coefficients (Φ, γ) depend linearly on θ.

Under Assumption (H6), all the non-homogeneous items in (3.22)–(3.23) (and then the corresponding difficulties) disappear. Moreover, all the conclusions in the rest of this paper are right. However, it is easy to understand that Assumption (H6) will lead to the second order partial derivatives of u to be zero. Therefore, the system of PDAEs (1.1) will degenerate to be a first order one.

4 Classical Solution to the System of PDAEs

In this section, we shall link the family of coupled FBSDEs (1.2) to the system (1.1) of PDAEs. In the first subsection, we shall work for the algebraic equation in the PDAE system (1.1), and in the second subsection, we shall consider the differential equation in (1.1).

4.1 Algebraic equations and the function v

The method used in the following Lemmas 4.1–4.2 is similar to the one in [27]. Since the issue of the viscosity solution was investigated in [27], then the dimension m was restricted to 1 there. In comparison, this subsection will focus on the multidimensional case, i.e., $m \geq 1$.

As a start, we give a property for smooth G-monotonic functions (see Definition 2.1).

Lemma 4.1 *Let* $f \in C^1(\mathbb{R}^n; \mathbb{R}^m)$ *be a* G*-monotonic function with* $\nu \geq 0$. *Then, for any matrix* $\mathcal{K} \in \mathbb{R}^{n \times d}$, *we have*

$$\langle \nabla f(x)\mathcal{K},\ G\mathcal{K} \rangle \geq \nu |\mathcal{K}|^2 \quad \text{for any } x \in \mathbb{R}^n.$$

Proof By the definition of G-monotonicity, for any vector x, $K \in \mathbb{R}^n$ and any positive number $\delta > 0$, we have

$$\langle f(x + \delta K) - f(x),\ G(\delta K) \rangle \geq \nu |\delta K|^2.$$

From Lagrange's differential mean value theorem, there exists an $\alpha \in (0,1)$ depending on δ, such that

$$\langle \nabla f(x + \alpha\delta K)(\delta K),\ G(\delta K) \rangle \geq \nu |\delta K|^2.$$

By dividing δ^2 on both sides, and then letting $\delta \to 0^+$, we have

$$\langle \nabla f(x)K,\ GK \rangle \geq \nu |K|^2. \tag{4.1}$$

Now, let $\mathcal{K} = (K_1, K_2, \cdots, K_d) \in \mathbb{R}^{n\times d}$ be a given matrix, where $K_j \in \mathbb{R}^n$ is the j-th volume vector of $\mathcal{K}$ ($j = 1, 2, \cdots, d$). Applying (4.1) to each K_j yields

$$\langle \nabla f(x)\mathcal{K},\ G\mathcal{K}\rangle = \sum_{j=1}^{d} \langle \nabla f(x)K_j,\ GK_j\rangle \geq \sum_{j=1}^{d} \nu|K_j|^2 = \nu|\mathcal{K}|^2.$$

We finish the proof.

For clarity, we extract the algebraic equation from the PDAE system (1.1) and rewrite it as follows:

$$v(t,x) = \nabla u(t,x)\sigma(t,x,u(t,x),v(t,x)). \tag{4.2}$$

Although the above algebraic equation can be considered for any given $u \in C^{0,1}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$, in this subsection, we are concerned about a special case where the function u appearing in (4.2) is given by (2.31). Since there are only the first order partial derivatives involved in (4.2), most of our results in this subsection will be under Assumptions (H3) and (H4). From Proposition 2.2, we know that u is G-monotonic. Moreover, Lemma 3.4 implies that $u \in C^{0,1}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$ and all its first order partial derivatives are bounded. The following lemma provides some preliminary results.

Lemma 4.2 *Let Assumptions* (H3) *and* (H4) *hold. Let u be defined by* (2.31).

(a) *Let (t,x), $(\overline{t},\overline{x}) \in [0,T]\times\mathbb{R}^n$ be given. If v and $\overline{v}$ satisfy the algebraic equations* (4.2) *with (t,x) and $(\overline{t},\overline{x})$, respectively, then*

$$|\overline{v} - v| \leq C|\nabla u(\overline{t},\overline{x})\sigma(\overline{t},\overline{x},u(\overline{t},\overline{x}),v) - \nabla u(t,x)\sigma(t,x,u(t,x),v)|,$$

where $C > 0$ is a constant independent of (t,x) and $(\overline{t},\overline{x})$.

(b) *For any $(t,x) \in [0,T]\times\mathbb{R}^n$, the algebraic equation* (4.2) *admits at most one solution.*

Proof (a) For simplicity, we use the denotations

$$\nabla u := \nabla u(t,x), \quad \nabla\overline{u} := \nabla\overline{u}(\overline{t},\overline{x}), \quad u := (t,x,u(t,x)), \quad \overline{u} := (\overline{t},\overline{x},u(\overline{t},\overline{x}))$$

in this part of proof. From the algebraic equations, we have

$$\overline{v} - v = \nabla\overline{u}[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)] + [\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)]. \tag{4.3}$$

By taking inner product with $G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]$, from Lemma 4.1 and the monotonicity condition of σ, we get

$$\begin{aligned}
-\beta_2|G^{\mathrm{T}}(\overline{v}-v)|^2 &\geq \langle \overline{v}-v,\ G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]\rangle \\
&= \langle \nabla\overline{u}[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)],\ G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]\rangle \\
&\quad + \langle \nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v),\ G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]\rangle \\
&\geq \nu|\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)|^2 \\
&\quad + \langle \nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v),\ G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]\rangle,
\end{aligned} \tag{4.4}$$

where $\nu \geq 0$ is the G-monotonicity constant of u (see Proposition 2.2). Next, similar to Subsection 3.1, according to the signs of β_1, β_2 and μ_1 in the monotonicity assumption (H3), we split our problem into two cases.

First case $\beta_1 > 0$, $\mu_1 > 0$, $\beta_2 \geq 0$ and $n \leq m$. In this case, by Proposition 2.2, we have $\nu > 0$. Then, (4.4) implies

$$\begin{aligned}|\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)|^2 \leq & -\frac{1}{\nu}\langle \nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v),\ G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]\rangle \\ & \leq \frac{|G|}{\nu}|\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)||\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)|.\end{aligned}$$

Therefore,

$$|\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)| \leq \frac{|G|}{\nu}|\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)|.$$

With the help of the above inequality, from (4.3), we deduce

$$\begin{aligned}|\overline{v} - v| &\leq |\nabla\overline{u}[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]| + |\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)| \\ &\leq \Big(1 + \frac{C_\nabla|G|}{\nu}\Big)|\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)|,\end{aligned}$$

where the constant $C_\nabla > 0$ is the bound of the gradient of the function u. We get the desired result in the first case.

Second case $\beta_1 \geq 0$, $\mu_1 \geq 0$, $\beta_2 > 0$ and $m \leq n$. In this case, the $(m \times m)$ matrix GG^{T} is positive definite, and we denote its minimum eigenvalue by $\lambda_{\min} > 0$. From (4.4),

$$\begin{aligned}\lambda_{\min}\beta_2|\overline{v} - v|^2 \leq \beta_2|G^{\mathrm{T}}(\overline{v} - v)|^2 \leq & -\langle \nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v),\ G[\sigma(\overline{u},\overline{v}) - \sigma(\overline{u},v)]\rangle \\ & \leq |G|L_z|\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)||\overline{v} - v|,\end{aligned}$$

where the constant $L_z > 0$ is the bound of $\frac{\partial\sigma}{\partial z}$. Therefore,

$$|\overline{v} - v| \leq \frac{|G|L_z}{\lambda_{\min}\beta_2}|\nabla\overline{u}\sigma(\overline{u},v) - \nabla u\sigma(u,v)|.$$

The result in the second case is proved.

(b) By letting $(\overline{t},\overline{x}) = (t,x)$ in the conclusion (a) of this lemma, we obtain the uniqueness of the algebraic equation.

With the previous preparations, now we give the main result of this subsection.

Proposition 4.1 *Let Assumptions* (H3) *and* (H4) *hold. For any* $(t,x) \in [0,T] \times \mathbb{R}^n$, *let* $\Theta^{t,x} = (X^{t,x}, Y^{t,x}, Z^{t,x})$ *be the unique solution to* FBSDE (1.2).

(a) *For any* (t,x), *the trajectories of the process* $Z^{t,x}$ *are continuous. Consequently, similar to the definition* (2.31) *of the function* u, *we can define another function* $v : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{m\times d}$ *as follows*:

$$v(t,x) = Z_t^{t,x}, \quad (t,x) \in [0,T] \times \mathbb{R}^n. \tag{4.5}$$

(b) *For any* (t,x), *the above defined* $v(t,x)$ *is the unique solution to the algebraic equation* (4.2).

(c) *The function* v *defined by* (4.5) *is continuous with respect to* $(t,x)\in[0,T]\times\mathbb{R}^n$.

(d) *Similar to the function* u, *the following Markovian property of the function* v *holds true*:

$$v(s,X_s^{t,x})=Z_s^{t,x},\quad s\in[t,T],\ (t,x)\in[0,T]\times\mathbb{R}^n. \tag{4.6}$$

(e) *There exists a constant* $C>0$ *such that*

$$|v(t,x)|\le C(1+|x|),\quad (t,x)\in[0,T]\times\mathbb{R}^n,$$

i.e., the function v *is linear growth with respect to* x.

(f) *Let Assumptions* (H2)[4] *and* (H5) *also hold. Then, there exists a constant* $C>0$ *such that*

$$|v(t,\overline{x})-v(t,x)|\le C(1+|x|)|\overline{x}-x|,\quad t\in[0,T],\ x,\overline{x}\in\mathbb{R}^n.$$

Consequently, the function v *is local Lipschitz continuous with respect to* x.

Proof (a) From Lemma 3.4, Theorem 3.1 and the Markovian property of the function u (see [27, Proposition 2.6 or Remark 2.7]), we derive

$$\begin{aligned}Z_s^{t,x}&=D_sY_s^{t,x}=D_su(s,X_s^{t,x})=\nabla u(s,X_s^{t,x})D_sX_s^{t,x}=\nabla u(s,X_s^{t,x})\sigma(s,\Theta_s^{t,x})\\&=\nabla u(s,X_s^{t,x})\sigma(s,X_s^{t,x},u(s,X_s^{t,x}),Z_s^{t,x}).\end{aligned} \tag{4.7}$$

Now, we fix $\omega\in\Omega$ and consider the corresponding trajectories of $X^{t,x}(\omega)$ and $Z^{t,x}(\omega)$. For any s, $s'\in[t,T]$, (4.7) means that $Z_s^{t,x}(\omega)$ and $Z_{s'}^{t,x}(\omega)$ are the solutions to the algebraic equations (4.2) with $(s,X_s^{t,x}(\omega))$ and $(s',X_{s'}^{t,x}(\omega))$, respectively. By Lemma 4.2(a),

$$\begin{aligned}|Z_{s'}^{t,x}(\omega)-Z_s^{t,x}(\omega)|\le C|&\nabla u(s',X_{s'}^{t,x}(\omega))\sigma(s',X_{s'}^{t,x}(\omega),u(s',X_{s'}^{t,x}(\omega)),Z_s^{t,x}(\omega))\\&-\nabla u(s,X_s^{t,x}(\omega))\sigma(s,X_s^{t,x}(\omega),u(s,X_s^{t,x}(\omega)),Z_s^{t,x}(\omega))|.\end{aligned}$$

Due to the continuity of σ, u, ∇u and $X^{t,x}(\omega)$, we get

$$\lim_{s'\to s}Z_{s'}^{t,x}(\omega)=Z_s^{t,x}(\omega).$$

Due to the arbitrariness of $s\in[t,T]$ and $\omega\in\Omega$, we prove the trajectory continuity of $Z^{t,x}$.

(b)–(d) After defining the function v by (4.5), we find that the conclusion (b) is an obvious consequence of (4.7) with $s=t$. Then, Lemma 4.2(a) works again to ensure the conclusion (c). Due to the conclusion (b) of this proposition, for any $(t,x)\in[0,T]\times\mathbb{R}^n$ and any $s\in[t,T]$, we have

$$v(s,X_s^{t,x})=\nabla u(s,X_s^{t,x})\sigma(s,X_s^{t,x},u(s,X_s^{t,x}),v(s,X_s^{t,x})),$$

where v is given by (4.5). From the uniqueness of the algebraic equation (4.2) (see Lemma 4.2(b)) as well as (4.7), we obtain the Markovian property of v.

(e) From the conclusion (c), $v(\cdot, 0)$ is bounded. With the help of Lemma 4.2(a), we calculate

$$\begin{aligned}|v(t,x)| &\leq |v(t,0)| + |v(t,x) - v(t,0)| \\ &\leq C + C|\nabla u(t,x)\sigma(t,x,u(t,x),v(t,0)) - \nabla u(t,0)\sigma(t,0,u(t,0),v(t,0))| \\ &\leq C + C|\nabla u(t,x)\sigma(t,x,u(t,x),v(t,0))| \\ &\leq C + C(1+|x|+|u(t,x)|+|v(t,0)|) \leq C(1+|x|).\end{aligned}$$

(f) Under Assumptions (H2)^4 and (H5), Lemma 4.2(a) works once again to lead

$$\begin{aligned}|v(t,\overline{x}) - v(t,x)| &\leq C|\nabla u(t,\overline{x}) - \nabla u(t,x)||\sigma(t,x,u(t,x),v(t,x))| \\ &\quad + C|\nabla u(t,\overline{x})||\sigma(t,\overline{x},u(t,\overline{x}),v(t,x)) - \sigma(t,x,u(t,x),v(t,x))| \\ &\leq C|\overline{x} - x|(1+|x|) + C(|\overline{x} - x|) \leq C(1+|x|)|\overline{x} - x|.\end{aligned}$$

The proof is completed.

4.2 Existence and uniqueness of the classical solution

Let us introduce a couple of spaces:

(1) $\mathbb{U}$ is a subspace of $C^{1,2}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$ in which the functions are of linear growth with respect to $x\in\mathbb{R}^n$.

(2) $\mathbb{V}$ is a subspace of $C([0,T]\times\mathbb{R}^n;\mathbb{R}^{m\times d})$ in which the functions are of linear growth and locally Lipschitz continuous with respect to $x\in\mathbb{R}^n$.

We first recall a result from [27, Theorem 3.1] which provides the uniqueness for (1.1).

Lemma 4.3 *Let Assumptions* (H3) *and* (H4) *hold. Let* $(\widetilde{u},\widetilde{v})\in\mathbb{U}\times\mathbb{V}$ *be a classical solution to the* PDAE *system* (1.2)*. Then* $\widetilde{u}$ *and* $\widetilde{v}$ *are uniquely determined by* (2.31) *and* (4.5)*, respectively.*

The following lemma collects some calculations which will be used in the proof of our main result (Theorem 4.1).

Lemma 4.4 *Let Assumptions* (H2)^4, (H3), (H4) *and* (H5) *hold. Let* $0\leq t'\leq r\leq t\leq T$ *and* $x\in\mathbb{R}^n$. *Then, there exists a constant* $C>0$ *independent of* t', r *and* x *such that*

$$\|X^{t',x}\|^2_{M^2_{\mathbb{F}}(t',T)} \leq C. \tag{4.8}$$

Moreover, we have the following convergence

$$\lim_{\substack{(t',r)\to(t,t)\\ t'\leq r\leq t}} \mathbb{E}[|\Theta_r^{t',x} - \Theta_t^{t,x}|^2] = 0 \tag{4.9}$$

and

$$\lim_{\substack{(t',r)\to(t,t)\\ t'\leq r\leq t}} \mathbb{E}[|f(r,\Theta_r^{t',x}) - f(t,\Theta_t^{t,x})|^2] = 0 \quad \text{with } f = g,b,\sigma. \tag{4.10}$$

Proof Obviously, (4.8) is a consequence of the standard L^2-estimate. Now, we calculate

$$\begin{aligned}\mathbb{E}[|X_r^{t',x}-x|^2] &\le 2\mathbb{E}\Big[\Big(\int_{t'}^{r}|b(s,\Theta_s^{t',x})|\mathrm{d}s\Big)^2\Big]+2\mathbb{E}\Big[\Big|\int_{t'}^{r}\sigma(s,\Theta_s^{t',x})\mathrm{d}W_s\Big|^2\Big]\\ &\le C\mathbb{E}\int_{t'}^{t}(1+|\Theta_s^{t',x}|^2)\mathrm{d}s.\end{aligned}$$

From the Markovian property and the linear growth property of (u,v),

$$\mathbb{E}[|X_r^{t',x}-x|^2]\le C\mathbb{E}\int_{t'}^{t}(1+|X_s^{t',x}|^2)\mathrm{d}s\le C\Big\{1+\mathbb{E}\Big[\sup_{s\in[t't]}|X_s^{t',x}|^2\Big]\Big\}(t'-t).$$

With the help of (4.8), we deduce from the above inequality that

$$\lim_{\substack{(t',r)\to(t,t)\\ t'\le r\le t}}\mathbb{E}[|X_r^{t',x}-x|^2]=0. \tag{4.11}$$

Next we are going to prove (4.9)–(4.10). Noticing Proposition 4.1(f), we calculate

$$\begin{aligned}\mathbb{E}[|\Theta_r^{t',x}-\Theta_t^{t,x}|^2]&=\mathbb{E}[|X_r^{t',x}-x|^2+|u(r,X_r^{t',x})-u(t,x)|^2+|v(r,X_r^{t',x})-v(t,x)|^2]\\ &\le C(1+|x|^2)\mathbb{E}[|X_r^{t',x}-x|^2]+2|u(r,x)-u(t,x)|^2+2|v(r,x)-v(t,x)|^2.\end{aligned}$$

By virtue of (4.11), we get the conclusion (4.9). Moreover,

$$\begin{aligned}&\mathbb{E}[|f(r,\Theta_r^{t',x})-f(t,\Theta_t^{t,x})|^2]\\ &\le 2\mathbb{E}[|f(r,\Theta_r^{t',x})-f(r,\Theta_t^{t,x})|^2]+2|f(r,\Theta_t^{t,x})-f(r,\Theta_t^{t,x})|^2\\ &\le C\mathbb{E}[|\Theta_r^{t',x}-\Theta_t^{t,x}|^2]+2|f(r,x,u(t,x),v(t,x))-f(t,x,u(t,x),v(t,x))|^2.\end{aligned}$$

(4.9) works to yield the conclusion (4.10). We finish the proof of this lemma.

We are now in the position to give the main result of this paper.

Theorem 4.1 *Under Assumptions* (H2)^4, (H3), (H4) *and* (H5), *the* PDAE *system* (1.1) *admits a unique solution in the space* $\mathbb{U}\times\mathbb{V}$. *Moreover, the unique pair of solutions* (u,v) *is defined by* (2.31) *and* (4.5). *Furthermore, all the first and second order partial derivatives of* u *with respect to* $x\in\mathbb{R}^n$ *are bounded on* $[0,T]\times\mathbb{R}^n$.

Proof Due to Lemmas 3.5, 4.3 and Proposition 4.1, the remaining thing is to prove that u defined by (2.31) is continuously differentiable with respect to $t\in[0,T]$ and solves the differential equation in system (1.1).

From Lemma 3.5, $u\in C^{0,2}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$. Let (t,x), $(t',x)\in[0,T]\times\mathbb{R}^n$. For the case $t'\ge t$, from Itô's formula and FBSDE (1.2), we have

$$\begin{aligned}u(t',x)-u(t,x)&=u(t',x)-u(t',X_{t'}^{t,x})+u(t',X_{t'}^{t,x})-u(t,x)\\ &=-\int_t^{t'}F^{t,x}(r,t')\mathrm{d}r+\int_t^{t'}[Z_r^{t,x}-\nabla u(t',X_r^{t,x})\sigma(r,\Theta_r^{t,x})]\mathrm{d}W_r,\end{aligned} \tag{4.12}$$

where $F^{t,x}(r,t')$ takes values in $\mathbb{R}^m$ and its k-th component $(k=1,2,\cdots,m)$ is given by

$$F_k^{t,x}(r,t') = \frac{1}{2}\sum_{i,j=1}^{n}(\sigma\sigma^{\mathrm{T}})_{ij}(r,\Theta_r^{t,x})\frac{\partial^2 u^k}{\partial x_i \partial x_j}(t',X_r^{t,x}) + \sum_{i=1}^{n} b_i(r,\Theta_r^{t,x})\frac{\partial u^k}{\partial x_i}(t',X_r^{t,x}) + g_k(r,\Theta_r^{t,x}).$$

After taking expectation on both sides of (4.12), the stochastic integral vanishes. By dividing $t'-t$, we have

$$\frac{u(t',x)-u(t,x)}{t'-t} = -\frac{1}{t'-t}\int_t^{t'}\mathbb{E}[F^{t,x}(r,t')]\mathrm{d}r. \tag{4.13}$$

From the continuity of $\gamma=(g,b,\sigma)$, D^2u, ∇u and $\Theta^{t,x}=(X^{t,x},Y^{t,x},Z^{t,x})$, it is easy to know that $F^{t,x}(r,t')$ is continuous with respect to (r,t'). By the boundedness of D^2u, ∇u and the linear growth of (u,v), we have

$$|F^{t,x}(r,t')| \le C[1+|\Theta_r^{t,x}|^2] \le C[1+|X_r^{t,x}|^2] \le C\Big[1+\sup_{r\in[t,T]}|X_r^{t,x}|^2\Big].$$

From Lebesgue's dominated convergence theorem,

$$\lim_{\substack{(t',r)\to(t,t)\\ t\le r\le t'}}\mathbb{E}[F^{t,x}(r,t')] = F^{t,x}(t,t). \tag{4.14}$$

This implies the convergence of (4.13) as $t'\to t^+$. By taking the limit on both sides of (4.13), we obtain

$$\partial_t^+ u(t,x) = -(\mathcal{L}u)(t,x,u(t,x),v(t,x)) - g(t,x,u(t,x),v(t,x)), \tag{4.15}$$

where $\partial_t^+ u(t,x)$ denotes the right derivative of u with respect to t at the point (t,x).

For the other case $t'\le t$, similar to (4.13), we obtain

$$\frac{u(t,x)-u(t',x)}{t-t'} = -\frac{1}{t-t'}\int_{t'}^{t}\mathbb{E}[F^{t',x}(r,t)]\mathrm{d}r. \tag{4.16}$$

Analogous to (4.14), we need the following convergence

$$\lim_{\substack{(t',r)\to(t,t)\\ t'\le r\le t}}\mathbb{E}[F^{t',x}(r,t)] = F^{t,x}(t,t). \tag{4.17}$$

For this aim, for any $k=1,2,\cdots,m$, we consider

$$\mathbb{E}[|F_k^{t',x}(r,t)-F_k^{t,x}(t,t)|] \le \frac{1}{2}\sum_{i,j=1}^{d}\Delta_{i,j,k}^{\sigma} + \sum_{i=1}^{n}\Delta_{i,k}^{b} + \Delta_k^g,$$

where

$$\Delta_k^g = \mathbb{E}[|g_k(r,\Theta_r^{t',x}) - g_k(t,\Theta_t^{t,x})|],$$
$$\Delta_{i,k}^b = \mathbb{E}\Big[\Big|b_i(r,\Theta_r^{t',x})\frac{\partial u^k}{\partial x_i}(t,X_r^{t',x}) - b_i(t,\Theta_t^{t,x})\frac{\partial u^k}{\partial x_i}(t,x)\Big|\Big]$$

and

$$\Delta^{\sigma}_{i,j,k} = \mathbb{E}\Big[\Big|(\sigma\sigma^{\mathrm{T}})_{ij}(r,\Theta_r^{t',x})\frac{\partial^2 u^k}{\partial x_i x_j}(t,X_r^{t',x}) - (\sigma\sigma^{\mathrm{T}})_{ij}(t,\Theta_t^{t,x})\frac{\partial^2 u^k}{\partial x_i x_j}(t,x)\Big|\Big].$$

(4.10) in Lemma 4.4 implies that $\Delta^g_k \to 0$ as $(t',r)\to(t,t)$. For simplicity, we omit the proof of $\Delta^b_{i,k}\to 0$, and then continue to consider the convergence of the "hardest" term $\Delta^{\sigma}_{i,j,k}$. In fact, by Hölder's inequality, we can deduce that

$$\begin{aligned}\Delta^{\sigma}_{i,j,k} \le{}& C\{\mathbb{E}[1+|x|^2+|X_r^{t',x}|^2]\}^{\frac12}\{\mathbb{E}[|\sigma(r,\Theta_r^{t',x})-\sigma(t,\Theta_t^{t,x})|^2]\}^{\frac12}\\ &+(\sigma\sigma^{\mathrm{T}})_{ij}(t,\Theta_t^{t,x})\mathbb{E}\Big[\Big|\frac{\partial^2 u^k}{\partial x_i\partial x_j}(t,X_r^{t',x})-\frac{\partial^2 u^k}{\partial x_i\partial x_j}(t,x)\Big|\Big].\end{aligned}$$

(4.8) and (4.10) in Lemma 4.4 imply the first item on the right hand side of the above inequality tends to 0 as $(t',r)\to(t,t)$. Moreover, (4.9) in Lemma 4.4 implies that $X_r^{t',x}$ converges to x in probability $\mathbb{P}$. From Lebesgue's dominated convergence theorem, the last item in the above inequality also tends to 0 as $(t',r)\to(t,t)$. In summary, we have proved (4.17). Therefore, taking the limit on both sides of (4.16) leads to

$$\partial_t^- u(t,x) = -(\mathcal{L}u)(t,x,u(t,x),v(t,x)) - g(t,x,u(t,x),v(t,x)), \tag{4.18}$$

where $\partial_t^- u(t,x)$ denotes the left derivative of u with respect to t at the point (t,x).

By combining (4.15) and (4.18), we obtain that $u\in C^{1,2}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$ and satisfies the differential equation in the PDAE system (1.1). The proof is completed.

Acknowledgements The authors wish to thank the editor and the anonymous referees for their helpful comments and criticisms which led to a much improved version of the manuscript.

Declarations

Conflicts of interest The authors declare no conflicts of interest.

References

[1] Antonelli, F., Backward-forward stochastic differential equations, *Ann. Appl. Probab.*, **3**(3), 1993, 777–793.

[2] Barles, G., Buckdahn, R. and Pardoux, É., Backward stochastic differential equations and integral-partial differential equations, *Stochastics Stochastics Rep.*, **60**(1–2), 1997, 57–83.

[3] Briand, P., Delyon, B., Hu, Y., et al., L^p solutions of backward stochastic differential equation, *Stochastic Process. Appl.*, **108**(1), 2003, 109–129.

[4] Courant, R. and Hilbert, D., Methods of mathematical physics, Vol. II, Partial differential equations, Interscience Publishers, New York-London, 1962.

[5] Darling, R. W. R. and Pardoux, É., Backwards SDE with random terminal time and applications to semilinear elliptic PDE, *Ann. Probab.*, **25**(3), 1997, 1135–1159.

[6] Delarue, F., On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case, *Stochastic Process. Appl.*, **99**(2), 2002, 209–286.

[7] El Karoui, N., Peng, S. and Quenez, M. C., Backward stochastic differential equations in finance, *Math. Finance*, **7**(1), 1997, 1–71.

[8] Feng, C., Wang, X. and Zhao, H., Quasi-linear PDEs and forward-backward stochastic differential equations: Weak solutions, *J. Differential Equations*, **264**(2), 2018, 959–1018.

[9] Hu, M., Ji, S. and Xue, X., A global stochastic maximum principle for fully coupled forward-backward stochastic systems, *SIAM J. Control Optim.*, **56**(6), 2018, 4309–4335.

[10] Hu, Y. and Peng, S., Solution of forward-backward stochastic differential equations, *Probab. Theory Related Fields*, **103**(2), 1995, 273–283.

[11] Kobylanski, M., Backward stochastic differential equations and partial differential equations with quadratic growth, *Ann. Probab.*, **28**(2), 2000, 558–602.

[12] Li, J. and Wei, Q., Optimal control problems of fully coupled FBSDEs and viscosity solutions of Hamilton-Jacobi-Bellman equations, *SIAM J. Control Optim.*, **52**(3), 2014, 1622–1662.

[13] Li, J. and Wei, Q., L^p estimates for fully coupled FBSDEs with jumps, *Stochastic Process. Appl.*, **124**(4), 2014, 1582–1611.

[14] Li, N. and Yu, Z., Forward-backward stochastic differential equations and linear-quadratic generalized Stackelberg games, *SIAM J. Control Optim.*, **56**(6), 2018, 4148–4180.

[15] Ma, J., Protter, P. and Yong, J., Solving forward-backward stochastic differential equations explicitly–a four step scheme, *Probab. Theory Related Fields*, **98**(3), 1994, 339–359.

[16] Ma, J., Wu, Z., Zhang, D. and Zhang, J., On well-posedness of forward-backward SDEs–a unified approach, *Ann. Appl. Probab.*, **25**(4), 2015, 2168–2214.

[17] Nualart, D., The Malliavin calculus and related topics, Second edition, Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.

[18] Pardoux, É., Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, Stochastic analysis and related topics, VI (Geilo, 1996), 79–127, Progr. Probab., **42**, Birkhäuser Boston, Boston, MA, 1998.

[19] Pardoux, É. and Peng, S., Backward stochastic differential equations and quasilinear parabolic partial differential equations, Stochastic partial differential equations and their applications (Charlotte, NC, 1991), 200–217, Lect. Notes Control Inf. Sci., **176**, Springer-Verlag, Berlin, 1992.

[20] Pardoux, É., Pradeilles, F. and Rao, Z., Probabilistic interpretation of a system of semi-linear parabolic partial differential equations, *Ann. Inst. H. Poincaré Probab. Statist.*, **33**(4), 1997, 467–490.

[21] Pardoux, É. and Răşcanu, A., Stochastic differential equations, backward SDEs, partial differential equations, Stochastic Modelling and Applied Probability, **69**, Springer-Verlag, Cham, 2014.

[22] Pardoux, É. and Tang, S., Forward-backward stochastic differential equations and quasilinear parabolic PDEs, *Probab. Theory Related Fields*, **114**(2), 1999, 123–150.

[23] Peng, S., Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, *Stochastics Stochastics Rep.*, **37**(1–2), 1991, 61–74.

[24] Peng, S., Open problems on backward stochastic differential equations, Control of distributed parameter and stochastic systems (Hangzhou, 1998), 265–273, Kluwer Acad. Publ., Boston, MA, 1999.

[25] Peng, S. and Wu, Z., Fully coupled forward-backward stochastic differential equations and applications to optimal control, *SIAM J. Control Optim.*, **37**(3), 1999, 825–843.

[26] Wu, Z., Adapted solutions of forward-backward stochastic differential equations and their parameter dependence, *Chinese J. Contemp. Math.*, **19**(1), 1998, 9–18.

[27] Wu, Z. and Yu, Z., Probabilistic interpretation for a system of quasilinear parabolic partial differential equation combined with algebra equations, *Stochastic Process. Appl.*, **124**(12), 2014, 3921–3947.

[28] Xie, B. and Yu, Z., An exploration of L^p-theory for forward-backward stochastic differential equations with random coefficients on small durations, *J. Math. Anal. Appl.*, **483**(2), 2020, 123642, 18 pp.

[29] Xie, B. and Yu, Z., L^p-estimate for linear forward-backward stochastic differential equations, *Acta Math. Sin., Engl. Series.*, (2023, to appear), https://doi.org/10.1007/s10114-023-1147-5.

[30] Yong, J., Finding adapted solutions of forward-backward stochastic differential equations: Method of continuation, *Probab. Theory Related Fields*, **107**(4), 1997, 537–572.

[31] Yong, J., Forward-backward stochastic differential equations with mixed initial-terminal conditions, *Trans. Amer. Math. Soc.*, **362**(2), 2010, 1047–1096.

[32] Zhang, J., Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory, Springer-Verlag, New York, 2017.

[33] Zhang, Q. and Zhao, H., Probabilistic representation of weak solutions of partial differential equations with polynomial growth coefficients, *J. Theoret. Probab.*, **25**(2), 2012, 396–423.