Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 123 (2013) 1616–1637

stochastic processes and their applications

www.elsevier.com/locate/spa

Semi-linear degenerate backward stochastic partial differential equations and associated forward–backward stochastic differential equations ☆

Kai Du, Qi Zhang *

School of Mathematical Sciences, Fudan University, Shanghai 200433, China

Received 4 September 2011; received in revised form 16 June 2012; accepted 3 January 2013
Available online 11 January 2013

Abstract

In this paper, we consider the Cauchy problem of semi-linear degenerate backward stochastic partial differential equations (BSPDEs) under general settings without technical assumptions on the coefficients. For the solution of semi-linear degenerate BSPDE, we first give a proof for its existence and uniqueness, as well as regularity. Then the connection between semi-linear degenerate BSPDEs and forward–backward stochastic differential equations (FBSDEs) is established, which can be regarded as an extension of the Feynman–Kac formula to the non-Markovian framework.

Keywords: Backward stochastic partial differential equations; Semi-linear degenerate equations; Forward–backward stochastic differential equations; Feynman–Kac formula

1. Introduction

BSPDEs were introduced by Bensoussan [2,3] as the adjoint equation of SPDE control systems. Since then BSPDEs appeared in a large amount of literature related to control theory as well as many other research fields. For example, in the study of stochastic maximum principle for

☆ This paper is supported by the National Natural Science Foundation of China with grant No. 11101090, the Doctoral Program of Higher Education of China with grant No. 20090071110001, the Specialized Research Fund for the Doctoral Program of Higher Education of China with grant No. 20090071120002 and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

* Corresponding author. Tel.: +86 21 55665549.
E-mail addresses: kdu@fudan.edu.cn (K. Du), qzh@fudan.edu.cn (Q. Zhang).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2013.01.005

stochastic PDEs or stochastic differential equations (SDEs) with partial information, the adjoint equations of Duncan–Mortensen–Zakai filtering equations are needed to solve first, which are actually BSPDEs. For this kind of discussions and applications, one can refer to [9,11,18,22,26], to name but a few. Moreover, by means of the classical duality argument, the controllability of stochastic evolution equations can be reduced to the observability estimate for BSPDEs, and this duality relation was utilized in e.g. [1,24]. Besides the applications in control theory, BSPDEs are also used to the stochastic process theory and mathematical finance, and we recommend the reader to see [5,8,14,15] for more details.

However, the solvability and the regularity of BSPDE, even for linear BSPDE, are tough problems due to the differential operators in the equation and its non-Markovian character. The recent work [7] by Du, Tang and Zhang made some progress and lifted the restrictions on the technical conditions for the Cauchy problem of linear degenerate BSPDEs. This work motivates us to consider the Cauchy problem of semi-linear degenerate BSPDEs under general settings. Actually, non-linear stochastic equations bear more application backgrounds without the exception of non-linear BSPDEs. For instance, Peng [21] discussed the Bellman dynamic principle for non-Markovian processes, whose corresponding backward stochastic Hamilton–Jacobi–Bellman equation is a fully non-linear BSPDE. Moreover, in many subjects of mathematical finance, such as imperfect hedging, portfolio choice, etc., non-linear BSPDEs appear as an important role and one can consult [16,17] for this aspect if interested.

Needless to say, more difficulties lay on the solvability of non-linear BSPDEs. In fact, the solvability of the solution to the fully non-linear BSPDE put forward in [21] is still an open problem, under general settings. Even for semi-linear BSPDE below we consider in this paper, only few work studied on it:

$$\begin{cases} du = -\big[\mathcal{L}u + \mathcal{M}q + f(t,x,u,q+u_x\sigma)\big]dt + q^k dW_t^k \\ u(T,x) = \varphi(x), \quad x \in \mathbb{R}^d, \end{cases} \tag{1.1}$$

where

$$\mathcal{L}u := a^{ij}u_{x^ix^j} + b^iu_{x^i} + cu \quad \text{and} \quad \mathcal{M}q := \sigma^{ik}q^k_{x^i} + \nu^k q^k.$$

In 2002, Hu, Ma and Yong considered the semi-linear BSPDE of the above form, under some specific settings and technical conditions in [10]. For instance, they only considered one-dimensional equation and the coefficients σ, ν were independent of x. One of our goal in this paper is to lift these restrictions and derive the existence, uniqueness and regularity of semi-linear degenerate BSPDE without technical assumptions. Also we would like to indicate that the similar regularity of solutions are obtained in this paper, but much weaker regularity requirements on the coefficients are needed in comparison with [10]. Besides, Tang [23] is also concerned with semi-linear degenerate BSPDEs by the method of stochastic flows, but this method causes a cost of assuming differentiability of higher orders in x on the coefficients.

Our another motivation is to establish the correspondence between semi-linear degenerate BSPDE and FBSDE. It is well known that, in the Markovian framework, the Feynman–Kac formula for semi-linear equations was established by Peng [20] and Pardoux–Peng [19]. This Feynman–Kac formula demonstrates a correspondence between semi-linear PDE and FBSDE whose coefficients are all Markov processes. But in the non-Markovian framework, FBSDE does not correspond to a deterministic PDE any more, but a BSPDE instead, by stochastic calculus. Certainly, as an extension of the Feynman–Kac formula, this kind of correspondence is basically important, whether in Mathematical finance research field or in a potential application to numerical calculus of BSPDE. To get the correspondence, one necessary step is to derive the continuity of a solution to FBSDE. Similar to [19], we utilize the Kolmogorov continuity theorem to prove it. But in our settings, no uniform Lipschitz conditions for $\varphi(x)$ and $f(s,x,0)$ with respect to x are assumed. Instead we suppose that $\varphi(\cdot)$ and $f(s,\cdot,0)$ belong to $W^{1,p}$ space and use the Sobolev embedding theorem to get the desired continuity.

Although [10,14] discussed the correspondence between BSPDE and FBSDE, our conditions are weaker but results are stronger in the solvable case, and thus can be applied to more equations. We expect that this kind of correspondence under our settings has independent interest in the areas of both SPDEs and backward stochastic differential equations (BSDEs).

The rest of this paper is organized as follows. In Section 2, we clarify all necessary notations and state the existing results used in this paper. In Section 3, we prove the existence, uniqueness and regularity of the solution to the semi-linear degenerate BSPDE. The correspondence between semi-linear degenerate BSPDEs and FBSDEs is established in Section 4.

2. Preliminaries

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space, among which the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ is generated by a d'-dimensional Wiener process $W = \{W_t; t \geq 0\}$ and all $\mathbb{P}$-null sets in $\mathscr{F}$.

The following notations will be used in this paper.

- For any multi-index $\gamma = (\gamma_1, \ldots, \gamma_d)$, we denote
$$D^\gamma = D_x^\gamma := \left(\frac{\partial}{\partial x^1}\right)^{\gamma_1}\left(\frac{\partial}{\partial x^2}\right)^{\gamma_2}\cdots\left(\frac{\partial}{\partial x^d}\right)^{\gamma_d}$$
and $|\gamma| = \gamma_1 + \cdots + \gamma_d$.
- For $n \in \mathbb{Z}^+, 0 < \alpha < 1$, denote by $C_0^\infty = C_0^\infty(\mathbb{R}^d)$ the set of infinitely differentiable real functions of compact support on $\mathbb{R}^d$, by $C^n = C^n(\mathbb{R}^d)$ the set of n times continuously differentiable functions on $\mathbb{R}^d$ such that
$$\|u\|_{C^n} := \sum_{|\gamma|\leq n} \sup_{x\in\mathbb{R}^d} |D^\gamma u(x)| < \infty,$$
and by $C^{n,\alpha} = C^{n,\alpha}(\mathbb{R}^d)$ the set of Hölder continuity functions on $\mathbb{R}^d$ such that
$$\|u\|_{C^{n,\alpha}} := \|u\|_{C^n} + \sum_{|\gamma|=n} \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|D^\gamma u(x) - D^\gamma u(y)|}{|x-y|^\alpha} < \infty,$$
where γ is a multi-index.
- For $p > 1$ and integer $m \geq 0$, we denote by $W^{m,p} = W^{m,p}(\mathbb{R}^d; \mathbb{R}^1)$ the Sobolev space of real functions on $\mathbb{R}^d$ with a finite norm
$$\|u\|_{m,p} := \left(\sum_{|\gamma|\leq m} \int_{\mathbb{R}^d} |D^\gamma u|^p dx\right)^{\frac{1}{p}}.$$
In particular, $W^{0,p} = L^p$. It is well known that $W^{m,2}$ is a Hilbert space and its inner product is denoted by $\langle\cdot,\cdot\rangle_m$.
- For $p > 1$ and integer $m \geq 0$, we denote by $[W^{m,p}]^{d'} = W^{m,p}(\mathbb{R}^d; \mathbb{R}^{d'})$ the Sobolev space of d' dimensional vector-valued functions on $\mathbb{R}^d$ with the norm $\|v\|_{m,p} = \left(\sum_{k=1}^{d'} \|v^k\|_{m,p}^p\right)^{1/p}$.

- Denote by $L^p_{\mathscr{F}}(0, T; W^{m,p})$ (resp. $L^p_{\mathscr{F}}(0, T; [W^{m,p}]^{d'})$) the space of all jointly measurable processes $u : \Omega \times [0, T] \longrightarrow W^{m,p}$ (resp. $u : \Omega \times [0, T] \longrightarrow [W^{m,p}]^{d'}$) such that u is $\mathscr{F}_t$-adapted and

$$\mathbb{E} \int_0^T \|u(t)\|^p_{m,p} dt < \infty.$$

- Denote by $C_{\mathscr{F}}([0, T]; W^{m,p})$ (resp. $C^w_{\mathscr{F}}([0, T]; W^{m,p})$) the space of all jointly measurable processes $u : \Omega \times [0, T] \longrightarrow W^{m,p}$ strongly (resp. weakly) continuous with respect to t on $[0, T]$ for a.s. ω, such that u is $\mathscr{F}_t$-adapted and

$$\mathbb{E} \sup_{t\in[0,T]} \|u(t)\|^p_{m,p} < \infty.$$

Here, when we talk about $u \in C^w_{\mathscr{F}}([0, T]; W^{m,p})$, it means that for any f in $(W^{m,p})^*$, the dual space of $W^{m,p}$, $f(u(t, \omega))$ is a.s. continuous with respect to t on $[0, T]$.

Throughout this paper the summation convention is in force for repeated indices. For the coefficients in the semi-linear BSPDE (1.1), we always assume that $a = (a^{ij})_{d\times d}$, $b = (b^1, \dots, b^d)$, c, $\sigma = (\sigma^{ik})_{d\times d'}$ and $\nu = (\nu^1, \dots, \nu^{d'})$ are jointly measurable and $\mathscr{F}_t$-adapted with values on the set of real symmetric $d \times d$ matrices, $\mathbb{R}^d$, $\mathbb{R}^1$, $\mathbb{R}^{d\times d'}$ and $\mathbb{R}^{d'}$, respectively; the real function $f(t, x, v, r)$ defined on $\Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^1 \times \mathbb{R}^{d'}$ is jointly measurable and $\mathscr{F}_t$-adapted for each (x, v, r) and continuous in (v, r) for each (ω, t, x); the real function φ is $\mathscr{F}_T \times \mathscr{B}(\mathbb{R}^d)$-measurable. Moreover, the following hypotheses will appear in the arguments.

($\mathbf{A}_m$) For a given constant $K_m \geq 0$ and a given integer $m \geq 0$, the functions b^i, c, ν^k and their derivatives with respect to x up to the order m, as well as a^{ij}, σ^{ik} and their derivatives up to the order $\max\{2, m\}$, are bounded by K_m.

(**P**) (*parabolicity*) For each $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$,

$$\big[2a^{ij}(t, x) - \sigma^{ik}\sigma^{jk}(t, x)\big]\xi^i\xi^j \geq 0, \quad \text{for arbitrary } \xi \in \mathbb{R}^d.$$

(**SP**) (*super-parabolicity*) There is a constant $\epsilon > 0$ such that for each $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$,

$$\big[2a^{ij}(t, x) - \sigma^{ik}\sigma^{jk}(t, x)\big]\xi^i\xi^j \geq \epsilon\delta_{ij}\xi^i\xi^j, \quad \text{for arbitrary } \xi \in \mathbb{R}^d,$$

where $\delta_{ij} = 1$ when $i = j$, otherwise $\delta_{ij} = 0$.

Definition 2.1. We call a pair function $(u, q) \in L^2_{\mathscr{F}}(0, T; W^{1,2}) \times L^2_{\mathscr{F}}(0, T; [W^{0,2}]^{d'})$ a (generalized) solution of BSPDE (1.1) if for each $\eta \in C_0^\infty$ and a.e. t,

$$\begin{aligned}\langle u(t), \eta\rangle_0 = \langle \varphi, \eta\rangle_0 &+ \int_t^T \langle \mathcal{L}u + \mathcal{M}q + f(s, x, u, q + u_x\sigma), \eta\rangle_0 ds \\ &- \int_t^T \langle q^k(s), \eta\rangle_0 dW^k_s \quad \mathbb{P}-\text{a.s.}\end{aligned} \tag{2.1}$$

Remark 2.1. According to the definition of $\mathcal{L}$, $\langle a^{ij}u_{x^ix^j}, \eta\rangle_0$ appears in (2.1) as the principal part of $\mathcal{L}u$, which is understood as

$$\langle a^{ij}u_{x^ix^j}, \eta\rangle_0 = -\langle a^{ij}u_{x^i}, \eta_{x^j}\rangle_0 - \langle a^{ij}_{x^j}u_{x^i}, \eta\rangle_0.$$

For convenience, we do a transform in Eq. (1.1) by setting

$$\widehat{q} = q + u_x \sigma.$$

Define $\alpha^{ij} = \frac{1}{2}\sigma^{ik}\sigma^{jk}$. Then Eq. (1.1) can be rewritten as the following form:

$$\begin{cases} du = -\big[\widehat{\mathcal{L}}u + \mathcal{M}\widehat{q} + f(t,x,u,\widehat{q})\big]dt + (\widehat{q} - u_x\sigma)dW_t \\ u(T,x) = \varphi(x), \quad x \in \mathbb{R}^d, \end{cases} \tag{2.2}$$

where

$$\widehat{\mathcal{L}}u = (a^{ij} - 2\alpha^{ij})u_{x^i x^j} + \widetilde{b}^i u_{x^i} + cu \quad \text{and} \quad \widetilde{b}^i = b^i - \sigma^{ik}_{x^j}\sigma^{jk} - \nu^k \sigma^{ik}.$$

It is clear that a function pair (u, q) satisfies (2.1) if and only if $(u, \widehat{q})$ satisfies the following

$$\begin{aligned} \langle u(t), \eta\rangle_0 = \langle \varphi, \eta\rangle_0 &+ \int_t^T \langle \widehat{\mathcal{L}}u(s) + \mathcal{M}\widehat{q}(s) + f(s,x,u(s),\widehat{q}(s)), \eta\rangle_0 ds \\ &- \int_t^T \langle \widehat{q}(s) - u_x\sigma(s), \eta\rangle_0 dW_s\ . \end{aligned} \tag{2.3}$$

To investigate semi-linear BSPDEs, we need some results about linear equations. In the linear case, f in Eq. (1.1) is taken to be independent of the last two variables, i.e.

$$f(t,x,v,r) = F(t,x),$$

and the corresponding linear BSPDE has a form as below:

$$\begin{cases} du = -\big[\mathcal{L}u + \mathcal{M}q + F\big]dt + q^k dW_t^k \\ u(T) = \varphi. \end{cases} \tag{2.4}$$

Then we have the following.

Theorem 2.1 (*Theorem 2.1 in Du–Tang–Zhang [7]*). *Let conditions* ($\mathbf{A}_m$) *and* (**P**) *be satisfied for given $m \geq 1$. If $F \in L^2_{\mathscr{F}}(0,T;W^{m,2})$ and $\varphi \in L^2_{\mathscr{F}_T}(\Omega;W^{m,2})$, then BSPDE (2.4) has a unique generalized solution (u,q) such that*

$$u \in C^w_{\mathscr{F}}([0,T];W^{m,2}) \quad \textit{and} \quad q + u_x\,\sigma \in L^2_{\mathscr{F}}(0,T;[W^{m,2}]^{d'}),$$

and for any integer $m_1 \in [0,m]$, we have the estimates

$$\begin{aligned} &\mathbb{E}\sup_{0\le t\le T}\|u(t)\|^2_{m_1,2} + \mathbb{E}\int_0^T \|(q+u_x\,\sigma)(t)\|^2_{m_1,2}\,dt \\ &\quad \le C\mathbb{E}\left(\|\varphi\|^2_{m_1,2} + \int_0^T \|F(t)\|^2_{m_1,2}\,dt\right). \end{aligned} \tag{2.5}$$

Here C is a generic constant which depends only on d, d', K_m, m and T.

In addition, if $F \in L^p_{\mathscr{F}}(0,T;W^{m,p})$ and $\varphi \in L^p_{\mathscr{F}_T}(\Omega;W^{m,p})$ for $p \geq 2$, then $u \in C^w_{\mathscr{F}}([0,T];W^{m,p})$, and for any integer $m_1 \in [0,m]$,

$$\mathbb{E}\sup_{0\le t\le T}\|u(t)\|^p_{m_1,p} \le Ce^{Cp}\mathbb{E}\left(\|\varphi\|^p_{m_1,p} + \int_0^T \|F(t)\|^p_{m_1,p}\,dt\right).$$

In the remaining part of this paper, we still use $C > 0$ as a generic constant only depending on given parameters, and when needed, a bracket will follow immediately after C to indicate what parameters C depends on.

However, (2.5) is not enough to obtain the estimates of the solution to the semi-linear equation, and we need more preparations. First let us see a lemma below.

Lemma 2.2. *Let conditions* $(\mathbf{A}_1)$ *and* $(\mathbf{P})$ *be satisfied, and* (u, q) *is the generalized solution of linear BSPDE (2.4). Then there exists a positive constant* $C = C(d, d', K_1, T)$ *such that for any positive number* $\lambda > C + 1$,

$$\mathbb{E}\int_0^T e^{\lambda t}\left(\|u\|_{1,2}^2 + \|q + u_x\sigma\|_{1,2}^2\right)dt \leq 2e^{\lambda T}\mathbb{E}\|\varphi\|_{1,2}^2 + \frac{2}{\lambda - C - 1}\mathbb{E}\int_0^T e^{\lambda t}\|F(t)\|_{1,2}^2 dt. \tag{2.6}$$

Proof. Take a small number $\varepsilon > 0$. Consider the following BSPDE:

$$\begin{cases} du^\varepsilon = -\big[(\varepsilon\Delta + \mathcal{L})u^\varepsilon + \mathcal{M}q^\varepsilon + F\big]dt + q^\varepsilon dW_t \\ u^\varepsilon(T) = \varphi. \end{cases} \tag{2.7}$$

Clearly, BSPDE (2.7) satisfies the super-parabolic condition (**SP**). In view of Theorem 2.3 in Du–Meng [6], Eq. (2.7) has a unique solution $(u^\varepsilon, q^\varepsilon)$ satisfying

$$u^\varepsilon \in L^2_{\mathscr{F}}(0, T; W^{2,2}) \cap C_{\mathscr{F}}([0, T]; W^{1,2}), \quad q^\varepsilon \in L^2_{\mathscr{F}}(0, T; [W^{1,2}]^{d'}).$$

Doing a similar transformation as in (2.2) with $\widehat{q}^\varepsilon = q^\varepsilon + u_x^\varepsilon\sigma$ and applying Itô's formula (c.f. [12]) to $e^{\lambda t}\left(\sum_{|\alpha|\leq 1}|D^\alpha u^\varepsilon|^2\right)$, we have

$$\begin{aligned} &\mathbb{E}\|u^\varepsilon(0)\|_{1,2}^2 - e^{\lambda T}\mathbb{E}\|\varphi\|_{1,2}^2 + \lambda\mathbb{E}\int_0^T e^{\lambda t}\|u^\varepsilon(t)\|_{1,2}^2 dt \\ &\quad = 2\sum_{|\alpha|\leq 1}\mathbb{E}\int_0^T e^{\lambda t}\langle D^\alpha u^\varepsilon, D^\alpha\big[(\varepsilon\Delta + \widehat{\mathcal{L}})u^\varepsilon + \mathcal{M}\widehat{q}^\varepsilon + F\big]\rangle_0 dt \\ &\qquad - \mathbb{E}\int_0^T e^{\lambda t}\|\widehat{q}^\varepsilon - u_x^\varepsilon\sigma\|_{1,2}^2 dt. \end{aligned} \tag{2.8}$$

From Lemma 3.1 in [7], we know that there exists a constant C depending only on d, d', K_1, T, but not ε, such that

$$\begin{aligned} &2\sum_{|\alpha|\leq 1}\langle D^\alpha u^\varepsilon, D^\alpha\big[(\varepsilon\Delta + \widehat{\mathcal{L}})u^\varepsilon + \mathcal{M}\widehat{q}^\varepsilon + F\big]\rangle_0 - \|\widehat{q}^\varepsilon - u_x^\varepsilon\sigma\|_{1,2}^2 \\ &\quad \leq -\frac{1}{2}\|\widehat{q}^\varepsilon(t)\|_{1,2}^2 + C\|u^\varepsilon(t)\|_{1,2}^2 + 2\langle u^\varepsilon, F\rangle_1. \end{aligned}$$

This along with (2.8) yields that

$$\begin{aligned} &\frac{1}{2}\mathbb{E}\int_0^T e^{\lambda t}\left(\|u^\varepsilon(t)\|_{1,2}^2 + \|\widehat{q}^\varepsilon(t)\|_{1,2}^2\right)dt \\ &\quad \leq e^{\lambda T}\mathbb{E}\|\varphi\|_{1,2}^2 + (C - \lambda + 1)\mathbb{E}\int_0^T e^{\lambda t}\|u^\varepsilon(t)\|_{1,2}^2 dt + 2\mathbb{E}\int_0^T e^{\lambda t}\langle u^\varepsilon, F\rangle_1(t)dt. \end{aligned} \tag{2.9}$$

Then taking $\lambda > C + 1$ and noting that

$$2\langle u^\varepsilon, F\rangle_1(t) \le (\lambda - C - 1)\|u^\varepsilon(t)\|_{1,2}^2 + \frac{1}{\lambda - C - 1}\|F(t)\|_{1,2}^2,$$

we obtain estimate (2.6) for $(u^\varepsilon, q^\varepsilon)$.

In view of the proof of Theorem 2.1 in [7], we know that there exists a subsequence $\{\varepsilon_n\} \downarrow 0$ such that $(u^{\varepsilon_n}, \widehat{q}^{\varepsilon_n})$ converges weakly to $(u, \widehat{q})$ in $L^2_{\mathscr{F}}(0, T; W^{1,2}) \times L^2_{\mathscr{F}}(0, T; [W^{1,2}]^{d'})$ as $n \to \infty$. Hence estimate (2.6) follows from the resonance theorem and the proof is complete. □

Remark 2.2. (i) Following the proof of Lemma 2.2, together with an application of Itô's formula to $e^{\lambda t}|u^\varepsilon|^2$, we can easily prove

$$\begin{aligned}&\mathbb{E}\int_0^T e^{\lambda t}\left(\|u\|_{0,2}^2 + \|q + u_x\sigma\|_{0,2}^2\right)dt\\ &\quad\le 2e^{\lambda T}\mathbb{E}\|\varphi\|_{0,2}^2 + \frac{2}{\lambda - C - 1}\mathbb{E}\int_0^T e^{\lambda t}\|F(t)\|_{0,2}^2 dt\end{aligned} \tag{2.10}$$

with the identical constant C in Lemma 2.2.

(ii) If we further assume that $(\mathbf{A}_2)$ holds, $F \in L^2_{\mathscr{F}}(0, T; W^{2,2})$ and $\varphi \in L^2_{\mathscr{F}_T}(\Omega; W^{2,2})$, then we can similarly deduce, by applying Itô's formula to $e^{\lambda t}\left(\sum_{|\alpha|\le 2}|D^\alpha u^\varepsilon|^2\right)$, that there exists a positive constant $C = C(d, d', K_2, T)$ such that for any positive number $\lambda > C + 1$,

$$\begin{aligned}&\mathbb{E}\int_0^T e^{\lambda t}\left(\|u\|_{2,2}^2 + \|q + u_x\sigma\|_{2,2}^2\right)dt\\ &\quad\le 2e^{\lambda T}\mathbb{E}\|\varphi\|_{2,2}^2 + \frac{2}{\lambda - C - 1}\mathbb{E}\int_0^T e^{\lambda t}\|F(t)\|_{2,2}^2\, dt.\end{aligned} \tag{2.11}$$

3. Existence, uniqueness and regularity of solutions to semi-linear BSPDEs

We make a further hypothesis on the function f in BSPDE (1.1):

(F) the function $f(t, x, v, r)$ satisfies
- (1) for arbitrary (ω, t, x, v, r), f_x, f_v and f_r exist;
- (2) $f(\cdot, \cdot, 0, 0) \in L^2_{\mathscr{F}}(0, T; W^{1,2})$;
- (3) there exists a constant $L > 0$ such that for each (ω, t, x),

$$\begin{aligned}&|f(t, x, v_1, r_1) - f(t, x, v_2, r_2)| + \|f_x(t, x, v_1, r_1) - f_x(t, x, v_2, r_2)\|\\ &\quad\le L(|v_1 - v_2| + \|r_1 - r_2\|), \quad \text{for arbitrary } v_1, v_2 \in \mathbb{R},\ r_1, r_2 \in \mathbb{R}^{d'}.\end{aligned}$$

Obviously, f_v and f_r are bounded by the constant L.

First we give the proof for the existence and uniqueness of solutions to semi-linear BSPDEs.

Theorem 3.1. *Let conditions* $(\mathbf{A}_1)$, $(\mathbf{P})$ *and* $(\mathbf{F})$ *be satisfied. Suppose* $\varphi \in L^2_{\mathscr{F}_T}(\Omega; W^{1,2})$, *then BSPDE (1.1) has a unique solution* (u, q) *such that*

$$u \in C^w_{\mathscr{F}}([0, T]; W^{1,2}), \quad q + u_x\sigma \in L^2_{\mathscr{F}}(0, T; [W^{1,2}]^{d'}).$$

Moreover, there exists a constant $C = C(d, d', K_1, T, L)$ *such that*

$$\mathbb{E} \sup_{t\in[0,T]} \|u(t)\|_{1,2}^2 + \mathbb{E}\int_0^T \|q + u_x\sigma\|_{1,2}^2(t)dt \le C\mathbb{E}\left(\|\varphi\|_{1,2}^2 + \int_0^T \|f(t,\cdot,0,0)\|_{1,2}^2 dt\right). \tag{3.1}$$

Proof. We mainly use the Picard iteration in the proof of this theorem.

Step 1. Define a successive sequence by setting

$$(u_0, q_0) = (0, 0)$$

and $\{(u_n, q_n)\}_{n\ge1}$ to be the unique solution of the following equations:

$$\begin{cases} du_n = -\left[\mathcal{L}u_n + \mathcal{M}q_n + f(t, x, u_{n-1}, q_{n-1} + u_{n-1,x}\sigma)\right]dt + q_n^k dW_t^k \\ u_n(T) = \varphi. \end{cases} \tag{3.2}$$

The solvability of Eq. (3.2) is indicated by Theorem 2.1 since one can easily check that

$$f(\cdot,\cdot,u_{n-1}, q_{n-1} + u_{n-1,x}\sigma) \in L^2_{\mathscr{F}}(0, T; W^{1,2})$$

by virtue of condition (**F**). Then we obtain a sequence $\{(u_n, \widehat{q}_n)\}_{n\ge0} \subset C^w_{\mathscr{F}}([0, T]; W^{1,2}) \times L^2_{\mathscr{F}}(0, T; [W^{1,2}]^{d'})$, where

$$\widehat{q}_n = q_n + u_{n,x}\sigma.$$

Step 2. For the sequence $\{(u_n, \widehat{q}_n)\}_{n\ge0}$ defined in Step 1, we prove that a subsequence converges weakly in $L^2_{\mathscr{F}}(0, T; W^{1,2}) \times L^2_{\mathscr{F}}(0, T; [W^{1,2}]^{d'})$. First noticing condition (**F**), we have that for each integer $n \ge 1$, there exists a positive constant $\widetilde{C}$ depending only on L such that

$$\mathbb{E}\int_0^T e^{\lambda t}\|f(t,\cdot,u_{n-1},\widehat{q}_{n-1})\|_{1,2}^2 dt \le \widetilde{C}\mathbb{E}\left[\int_0^T e^{\lambda t}\|f(t,\cdot,0,0)\|_{1,2}^2 dt + \int_0^T e^{\lambda t}\left(\|u_{n-1}\|_{1,2}^2 + \|\widehat{q}_{n-1}\|_{1,2}^2\right)dt\right]. \tag{3.3}$$

If we denote the constant C in (2.6) and (2.10) by C_1, then taking $\lambda_0 = 4\widetilde{C} + C_1 + 1$, we can prove a claim that for each $n \ge 0$,

$$\mathbb{E}\int_0^T e^{\lambda_0 t}\left(\|u_n\|_{1,2}^2 + \|\widehat{q}_n\|_{1,2}^2\right)dt \le 4\mathbb{E}\left(e^{\lambda_0 T}\|\varphi\|_{1,2}^2 + \int_0^T e^{\lambda_0 t}\|f(t,\cdot,0,0)\|_{1,2}^2 dt\right). \tag{3.4}$$

To prove it, the mathematical induction is used. Assume that (3.4) is true for $n - 1$. Applying Lemma 2.2 to Eq. (3.2), by (3.3) we have

$$\mathbb{E}\int_0^T e^{\lambda_0 t}\left(\|u_n\|_{1,2}^2 + \|\widehat{q}_n\|_{1,2}^2\right)dt \le 2e^{\lambda_0 T}\mathbb{E}\|\varphi\|_{1,2}^2 + \frac{2}{\lambda_0 - C_1 - 1}\mathbb{E}\int_0^T e^{\lambda_0 t}\|f(t,\cdot,u_{n-1},\widehat{q}_{n-1})\|_{1,2}^2 dt$$

$$\begin{aligned}
&\leq 2e^{\lambda_0 T}\mathbb{E}\|\varphi\|_{1,2}^2 + \frac{2\widetilde{C}}{\lambda_0 - C_1 - 1}\mathbb{E}\left[\int_0^T e^{\lambda_0 t}\|f(t,\cdot,0,0)\|_{1,2}^2 dt\right.\\
&\quad\left. + \int_0^T e^{\lambda_0 t}\left(\|u_{n-1}\|_{1,2}^2 + \|\widehat{q}_{n-1}\|_{1,2}^2\right)dt\right]\\
&\leq 4\mathbb{E}\left(e^{\lambda_0 T}\|\varphi\|_{1,2}^2 + \int_0^T e^{\lambda_0 t}\|f(t,\cdot,0,0)\|_{1,2}^2 dt\right).
\end{aligned}$$

By (3.4), we immediately know that $\{(u_n,\widehat{q}_n)\}_{n\geq 0}$ is uniformly bounded with the norm of $L^2_{\mathscr{F}}(0,T;W^{1,2})\times L^2_{\mathscr{F}}(0,T;[W^{1,2}]^{d'})$. Hence there exist a subsequence $\{n'\}$ and a function pair

$$(\widetilde{u},\widetilde{\widehat{q}})\in L^2_{\mathscr{F}}(0,T;W^{1,2})\times L^2_{\mathscr{F}}(0,T;[W^{1,2}]^{d'})$$

such that as $n'\to\infty$,

$$(u_{n'},\widehat{q}_{n'})\rightharpoonup(\widetilde{u},\widetilde{\widehat{q}})\quad\text{weakly in } L^2_{\mathscr{F}}(0,T;W^{1,2})\times L^2_{\mathscr{F}}(0,T;[W^{1,2}]^{d'}).$$

Step 3. We then prove the strong convergence of $\{(u_n,\widehat{q}_n)\}_{n\geq 0}$ in $L^2_{\mathscr{F}}(0,T;W^{0,2})\times L^2_{\mathscr{F}}(0,T;[W^{0,2}]^{d'})$. In view of (2.10) and condition (**F**), taking $\lambda=\lambda_1=8L^2+C_1+1$ and $n\geq 1$ we have

$$\begin{aligned}
&\mathbb{E}\int_0^T e^{\lambda_1 t}\left(\|u_{n+1}-u_n\|_{0,2}^2 + \|\widehat{q}_{n+1}-\widehat{q}_n\|_{0,2}^2\right)dt\\
&\quad\leq \frac{2}{\lambda_1 - C_1 - 1}\mathbb{E}\int_0^T e^{\lambda_1 t}\|f(t,\cdot,u_n,\widehat{q}_n) - f(t,\cdot,u_{n-1},\widehat{q}_{n-1})\|_{0,2}^2 dt\\
&\quad\leq \frac{1}{2}\mathbb{E}\int_0^T e^{\lambda_1 t}\left(\|u_n-u_{n-1}\|_{0,2}^2 + \|\widehat{q}_n-\widehat{q}_{n-1}\|_{0,2}^2\right)dt,
\end{aligned}$$

which implies that $\{(u_n,\widehat{q}_n)\}_{n\geq 0}$ is a Cauchy sequence in the space $L^2_{\mathscr{F}}(0,T;W^{0,2})\times L^2_{\mathscr{F}}(0,T;[W^{0,2}]^{d'})$. Actually $\{(u_n,\widehat{q}_n) : n\geq 1\}$ is also a Cauchy sequence with the norm $\mathbb{E}\int_0^T\|\cdot\|_{0,2}^2 dt$ due to the norm equivalence between $\sqrt{\mathbb{E}\int_0^T e^{\lambda_1 t}\|\cdot\|_{0,2}^2 dt}$ and $\sqrt{\mathbb{E}\int_0^T\|\cdot\|_{0,2}^2 dt}$ in $L^2_{\mathscr{F}}(0,T;W^{0,2})\times L^2_{\mathscr{F}}(0,T;[W^{0,2}]^{d'})$. We denote the strong limit of $\{(u_n,\widehat{q}_n)\}_{n\geq 0}$ by $(u,\widehat{q})$. Recalling the subsequence $\{n'\}$ in Step 2, we know that $\{(u_{n'},\widehat{q}_{n'})\}$ converges strongly to $(u,\widehat{q})$ in $L^2_{\mathscr{F}}(0,T;W^{0,2})\times L^2_{\mathscr{F}}(0,T;[W^{0,2}]^{d'})$. By the uniqueness of the limit, we have

$$(u,\widehat{q})=(\widetilde{u},\widetilde{\widehat{q}})\in L^2_{\mathscr{F}}(0,T;W^{1,2})\times L^2_{\mathscr{F}}(0,T;[W^{1,2}]^{d'}).$$

Step 4. Next we prove that $(u,\widehat{q})$ is a solution of BSPDE (1.1) to complete the existence proof. For this, we need verify that $(u,\widehat{q})$ satisfies (2.3). First we know that

$$\begin{aligned}
\langle u_{n'}(t),\eta\rangle_0 &= \langle\varphi,\eta\rangle_0 + \int_t^T\langle\widehat{\mathcal{L}}u_{n'}(s) + \mathcal{M}\widehat{q}_{n'}(s) + f(s,x,u_{n'-1}(s),\widehat{q}_{n'-1}(s)),\eta\rangle_0 ds\\
&\quad - \int_t^T\langle\widehat{q}_{n'}(s) - u_{n',x}\sigma(s),\eta\rangle_0 dW_s.
\end{aligned}\tag{3.5}$$

Since $(u_{n'-1},\widehat{q}_{n'-1})$ converges strongly to $(u,\widehat{q})$ in $L^2_{\mathscr{F}}(0,T;W^{0,2})\times L^2_{\mathscr{F}}(0,T;[W^{0,2}]^{d'})$ as $n'\to\infty$, by condition (**F**) it follows that, as $n'\to\infty$,

$$\mathbb{E}\int_0^T\|f(t,\cdot,u_{n'-1},\widehat{q}_{n'-1}) - f(t,\cdot,u,\widehat{q})\|_{0,2}^2(t)dt \longrightarrow 0.$$

Hence, for any $\eta \in C_0^\infty$, all terms of (3.5) converge weakly to the corresponding terms of (2.3) in $L^2_{\mathscr{F}}(0,T;\mathbb{R}^1)$ since the operators of Lebesgue integration and stochastic integration are continuous in $L^2_{\mathscr{F}}(0,T;\mathbb{R}^1)$. Therefore, $(u,\widehat{q})$ is a generalized solution of (2.2). Setting $q=\widehat{q}-u_x\sigma$, we know that (u,q) is a generalized solution of BSPDE (1.1).

Moreover, since $(u,\widehat{q})$ is obtained, we regard $f(t,x,u,q+u_x\sigma)$ as the known coefficient and $(u,\widehat{q})$ as the solution of linear BSPDE with given $f(t,x,u,q+u_x\sigma)$. By condition (**F**), $f(t,x,u,q+u_x\sigma) \in L^2_{\mathscr{F}}(0,T;W^{m,2})$. Then we get from Theorem 2.1 that $u \in C^w_{\mathscr{F}}([0,T];W^{1,2})$ and (3.1) follows.

Step 5. We finally deduce the uniqueness of solution to semi-linear BSPDE. Assume that (u_1,q_1) and (u_2,q_2) are two generalized solutions to BSPDE (1.1). Set $\widehat{q_i}=q_i+u_{i,x}\sigma$, $i=1,2$. Noticing (2.10) and taking $\lambda=\lambda_1$ again, by condition (**F**) we have

$$\begin{aligned}&\mathbb{E}\int_0^T e^{\lambda_1 t}\left(\|u_1-u_2\|^2_{0,2}+\|\widehat{q}_1-\widehat{q}_2\|^2_{0,2}\right)dt\\ &\quad\le \frac{1}{2}\mathbb{E}\int_0^T e^{\lambda_1 t}\left(\|u_1-u_2\|^2_{0,2}+\|\widehat{q}_1-\widehat{q}_2\|^2_{0,2}\right)dt.\end{aligned}$$

The uniqueness of solution immediately follows, which completes the proof of Theorem 3.1. □

In the remaining part of this section, the regularity of solution to semi-linear BSPDE is explored. We consider a simpler form of BSPDE (1.1) with $f(t,x,v,r)$ independent of r:

$$\begin{cases} du=-[\mathcal{L}u+\mathcal{M}q+f(t,x,u)]dt+q^k dW_t^k\\ u(T,x)=\varphi(x), \quad x\in\mathbb{R}^d.\end{cases} \tag{3.6}$$

For BSPDE (3.6), condition (**F**) is simplified as follows:

($\mathbf{F}'$) the function $f(t,x,v)$ satisfies
- (1) for arbitrary (ω,t,x,v), f_x and f_v exist;
- (2) $f(\cdot,\cdot,0)\in L^2_{\mathscr{F}}(0,T;W^{1,2})$;
- (3) there exists a constant $L>0$ such that for each (ω,t,x),
$$|f(t,x,v_1)-f(t,x,v_2)|+\|f_x(t,x,v_1)-f_x(t,x,v_2)\|\le L|v_1-v_2|,$$
for arbitrary $v_1,v_2\in\mathbb{R}$.

Obviously, f_v is bounded by the constant L.

We know from Theorem 3.1 that under conditions ($\mathbf{A}_1$), (**P**) and ($\mathbf{F}'$), if $\varphi\in L^2_{\mathscr{F}_T}(\Omega;W^{1,2})$, BSPDE (3.6) has a unique solution $(u,q)\in C^w_{\mathscr{F}}([0,T];W^{1,2})\times L^2_{\mathscr{F}}(0,T;[W^{0,2}]^{d'})$. Moreover, some regularity results for BSPDE (3.6) can be obtained.

Theorem 3.2. *We assume that conditions* ($\mathbf{A}_1$), (**P**) *and* ($\mathbf{F}'$) *are satisfied, and for* $p\ge 2$, $f(\cdot,\cdot,0)\in L^p_{\mathscr{F}}(0,T;W^{1,p})$ *and* $\varphi\in L^p_{\mathscr{F}_T}(\Omega;W^{1,p})$, *then* $u\in C^w_{\mathscr{F}}([0,T];W^{1,p})$ *and there exists a constant* $C=C(d,d',K_1,T,L,p)$ *such that*

$$\mathbb{E}\sup_{t\le T}\|u(t)\|^p_{1,p}\le Ce^{Cp}\mathbb{E}\left(\|\varphi\|^p_{1,p}+\int_0^T\|f(t,\cdot,0)\|^p_{1,p}\,dt\right). \tag{3.7}$$

Proof. By condition ($\mathbf{F}'$), it is easy to see that for arbitrary $v\in W^{1,p}$

$$\begin{aligned}\|f(t,\cdot,v)\|^p_{1,p}&=\|f(t,\cdot,v)\|^p_{0,p}+\|f_x(t,\cdot,v)\|^p_{0,p}\\ &\le C(p)(\|f(t,\cdot,0)\|^p_{1,p}+L^p\|v\|^p_{1,p}).\end{aligned} \tag{3.8}$$

To avoid heavy notation, we set

$$M_1 = \mathbb{E}\left(\|\varphi\|_{1,p}^p + \int_0^T \|f(t,\cdot,0)\|_{1,p}^p \, dt \right).$$

Similar to arguments in Theorem 3.1, we define a recursive sequence $\{(u_n, q_n)\}_{n\geq 1}$ as follows:

$$\begin{cases} du_n = -\left[\mathcal{L}u_n + \mathcal{M}q_n + f(t,x,u_{n-1})\right] dt + q_n \, dW_t \\ u_n(T) = \varphi. \end{cases}$$

If $u_{n-1} \in L^p_{\mathscr{F}}(0,T;W^{1,p})$, by (3.8) $f(\cdot,\cdot,u_{n-1}) \in L^p_{\mathscr{F}}(0,T;W^{1,p})$, thus $u_n \in C^w_{\mathscr{F}}([0,T];W^{1,p})$ follows immediately from Theorem 2.1. By setting $u_0 = 0$, we know from mathematical induction that $\{u_n\}_{n\geq 0} \subset C^w_{\mathscr{F}}([0,T];W^{1,p})$. Furthermore, by the estimate in Theorem 2.1 and (3.8), we have for arbitrary $t \in [0,T]$, $n \geq 1$,

$$\begin{aligned} \mathbb{E}\|u_n(t)\|_{1,p}^p &\leq C\,\mathbb{E}\left(\|\varphi\|_{1,p}^p + \int_t^T \|f(s,\cdot,u_{n-1})\|_{1,p}^p \, ds \right) \\ &\leq C \int_t^T \mathbb{E}\|u_{n-1}(s)\|_{1,p}^p \, ds + CM_1, \end{aligned}$$

where C is independent of n. A simple calculation leads to

$$\mathbb{E}\|u_n(t)\|_{1,p}^p \leq CM_1 \sum_{k=0}^{n-1} \frac{1}{k!} C^k (T-t)^k \leq CM_1 e^{C(T-t)}. \tag{3.9}$$

Hence there exist a subsequence $\{n'\}$ and a function $u \in L^2_{\mathscr{F}}(0,T;W^{1,2})$ such that as $n' \to \infty$, $u_{n'}$ converges weakly to u in $L^2_{\mathscr{F}}(0,T;W^{1,2})$. By the Banach–Saks Theorem, we can construct a sequence u^k from finite convex combinations of $u_{n'}$ such that u^k and u^k_x converge to u and u_x for a.e. $t \in [0,T] x \in \mathbb{R}^d$ a.s., respectively. Due to the norm itself is convex, (3.9) implies

$$\mathbb{E}\int_0^T \|u^k(t)\|_{1,p}^p \, dt \leq CM_1(e^{CT} - 1).$$

By the Fatou Lemma, it turns out that

$$\mathbb{E}\int_0^T \|u(t)\|_{1,p}^p \, dt \leq CM_1(e^{CT} - 1).$$

Regarding u as the solution of linear BSPDE with given coefficient $f(t,x,u)$, by (3.8) and Theorem 2.1 we obtain (3.7). $\square$

From the proof of Theorem 3.2 and Corollary 2.3 in [7], it is not hard to derive the following corollary.

Corollary 3.3. *Let conditions* $(\mathbf{A}_1)$, $(\mathbf{P})$ *and* $(\mathbf{F}')$ *be satisfied. If* $f(\cdot,\cdot,0) \in L^\infty_{\mathscr{F}}(0,T;W^{1,\infty})$, $\varphi \in L^\infty_{\mathscr{F}_T}(\Omega;W^{1,\infty})$, *then* $u \in L^\infty_{\mathscr{F}}(0,T;W^{1,\infty})$, *i.e.*

$$\|u\|_{L^\infty_{\mathscr{F}}(0,T;W^{1,\infty})} \leq C(d,d_0,K_1,T,L,f(\cdot,\cdot,0),\varphi) \triangleq C_\infty.$$

With the help of the Sobolev embedding theorem, it is not hard to deduce the corollary below.

Corollary 3.4. *Under the conditions in Theorem* 3.2 *with* $p > 2$ *replaced by* $p > d$, $u(t,x)$ *is jointly continuous on* (t,x) *a.s.*

Proof. For $R \in \mathbb{N}$, take a nonnegative function $\rho \in C_0^\infty(\mathbb{R}^d)$ such that $\rho(x) = 1$ on the closed ball $\bar{B}_R \triangleq \{x \in \mathbb{R}^d : |x| \le R\}$ and $\rho(x) = 0$ on $\{x \in \mathbb{R}^d : |x| \ge R+1\}$, then it is clear that $u\rho \in C^w_{\mathscr{F}}\left([0,T]; W^{1,p}(\bar{B}_{R+1})\right)$. Since the embedding from $W^{1,p}(\bar{B}_{R+1})$ to $C^\alpha\left(\bar{B}_{R+1}\right)$ is compact, $u\rho \in C_{\mathscr{F}}\left([0,T]; C^\alpha(\bar{B}_{R+1})\right)$ which implies that $u(t,x)\rho(x)$ is a.s. jointly continuous with respect to (t,x) on $\bar{B}_{R+1}$. Thus u is a.s. jointly continuous with respect to (t,x) on $\bar{B}_R$. By the arbitrariness of R, the corollary follows. □

Based on Theorem 3.2, we explore the regularity of solution to BSPDE (3.6).

Theorem 3.5. *We assume that*

(1) *conditions* ($\mathbf{A}_2$) *and* ($\mathbf{P}$) *hold, and* $\varphi \in L^2_{\mathscr{F}_T}(\Omega; W^{2,2}) \cap L^\infty_{\mathscr{F}_T}(\Omega; W^{1,\infty})$;
(2) *for arbitrary* (ω,t,x,v), $f_x, f_v, f_{xx}, f_{xv}, f_{vv}$ *exist;*
(3) $f(\cdot,\cdot,0) \in L^2_{\mathscr{F}}(0,T; W^{2,2}) \cap L^\infty_{\mathscr{F}}(0,T; W^{1,\infty})$;
(4) f_v, f_{xv}, f_{vv} *are bounded by* L;
(5) *for arbitrary* (ω,t,x,v), $|f_{xx}(t,x,v)| \le |f_{xx}(t,x,0)| + L|v|$.

Then (3.6) *has a unique generalized solution* (u,q) *satisfying*

$$u \in C^w_{\mathscr{F}}([0,T]; W^{2,2}) \cap L^\infty_{\mathscr{F}}(0,T; W^{1,\infty}) \quad \textit{and} \quad q + u_x\sigma \in L^2_{\mathscr{F}}(0,T; [W^{2,2}]^{d'}).$$

Proof. First of all, our assumptions satisfy the conditions in Theorem 3.2 and Corollary 3.3; thus (3.6) has a unique solution (u,q) satisfying

$$u \in C^w_{\mathscr{F}}([0,T]; W^{1,2}) \cap L^\infty_{\mathscr{F}}(0,T; W^{1,\infty}) \quad \text{and} \quad q + u_x\sigma \in L^2_{\mathscr{F}}(0,T; [W^{1,2}]^{d'}).$$

To get a better regularity, for arbitrary $\delta > 0$, we consider the non-degenerate BSPDE below:

$$\begin{cases} du^\delta = -\left[(\delta\Delta + \mathcal{L})u^\delta + \mathcal{M}q^\delta + f(t,x,u^\delta)\right]dt + q^\delta\, dW_t \\ u^\delta(T) = \varphi. \end{cases}$$

By Theorem 3.1 we know that above BSPDE has a unique solution $(u^\delta, q^\delta) \in L^2_{\mathscr{F}}(0,T; W^{1,2}) \times L^2_{\mathscr{F}}(0,T; [W^{0,2}]^{d'})$, which together with condition (4) leads to a fact that $f(t,x,u^\delta) \in L^2_{\mathscr{F}}(0,T; W^{1,2})$. Regarding $f(t,x,u^\delta)$ as a given coefficient and using Theorem 2.3 in [6] for non-degenerate linear BSPDE, we can get a better regularity of solution, i.e. $(u^\delta, q^\delta) \in L^2_{\mathscr{F}}(0,T; W^{3,2}) \times L^2_{\mathscr{F}}(0,T; [W^{2,2}]^{d'})$. Then $q^\delta + u^\delta_x\sigma \in L^2_{\mathscr{F}}(0,T; [W^{2,2}]^{d'})$, and by (2.11) there exists a positive constant $C_2 = C_2(d,d',K_2,T)$ such that for any positive number $\lambda > C_2 + 1$,

$$\begin{aligned} &\mathbb{E}\int_0^T e^{\lambda t}\left(\|u^\delta\|^2_{2,2} + \|q^\delta + u^\delta_x\sigma\|^2_{2,2}\right)dt \\ &\quad \le 2e^{\lambda T}\mathbb{E}\|\varphi\|^2_{2,2} + \frac{2}{\lambda - C_2 - 1}\mathbb{E}\int_0^T e^{\lambda t}\|f(t,\cdot,u^\delta)\|^2_{2,2}\,dt. \end{aligned} \tag{3.10}$$

Also, by Corollary 3.3 we have $|u^\delta_x| \le C_\infty$, so it follows from conditions (2), (4) and (5) that

$$\begin{aligned} |f(t,x,u^\delta)| &\le |f(t,x,0)| + L|u^\delta|, \\ |\{f(t,x,u^\delta)\}_x| &\le |f_x(t,x,0)| + L(|u^\delta| + |u^\delta_x|), \\ |\{f(t,x,u^\delta)\}_{xx}| &\le |f_{xx}(t,x,u^\delta)| + 2|f_{xv}(t,x,u^\delta)|\cdot|u^\delta_x| + |f_{vv}(t,x,u^\delta)|\cdot|u^\delta_x|^2 \\ &\quad + |f_v(t,x,u^\delta)|\cdot|u^\delta_{xx}| \\ &\le |f_{xx}(t,x,0)| + L|u^\delta| + (2+C_\infty)L|u^\delta_x| + L|u^\delta_{xx}|. \end{aligned}$$

Hence,

$$\|f(t,\cdot,u^\delta)\|_{2,2}^2 \le C(L,C_\infty)\left(\|f(t,\cdot,0)\|_{2,2}^2 + \|u^\delta\|_{2,2}^2\right).$$

Putting this estimate into (3.10), we immediately get

$$\begin{aligned}&\mathbb{E}\int_0^T e^{\lambda t}\left(\|u^\delta\|_{2,2}^2 + \|q^\delta + u_x^\delta\sigma\|_{2,2}^2\right) dt\\ &\quad \le 2e^{\lambda T}\mathbb{E}\|\varphi\|_{2,2}^2 + \frac{2\,C(L,C_\infty)}{\lambda - C_2 - 1}\mathbb{E}\int_0^T e^{\lambda t}\left(\|f(t,\cdot,0)\|_{2,2}^2 + \|u^\delta\|_{2,2}^2\right) dt.\end{aligned}$$

Then taking $\lambda = 4\,C(L,C_\infty) + C_2 + 1$ in above and setting

$$M_2 = \mathbb{E}\left(\|\varphi\|_{2,2}^2 + \int_0^T \|f(t,\cdot,0)\|_{2,2}^2\, dt\right),$$

we obtain the uniformly bounded estimate for $(u^\delta, q^\delta + u_x^\delta\sigma)$ in $L^2_{\mathscr{F}}(0,T;W^{2,2}) \times L^2_{\mathscr{F}}(0,T;[W^{2,2}]^{d'})$, i.e.

$$\mathbb{E}\int_0^T \left(\|u^\delta\|_{2,2}^2 + \|q^\delta + u_x^\delta\sigma\|_{2,2}^2\right) dt \le 4e^{\lambda T} M_2, \tag{3.11}$$

where λ is independent of δ. So we can get a sequence $\{\delta_n\} \downarrow 0$ and $(\widehat{u},\widehat{r}) \in L^2_{\mathscr{F}}(0,T;W^{2,2}) \times L^2_{\mathscr{F}}(0,T;[W^{2,2}]^{d'})$ such that $(u^n, r^n) \triangleq (u^{\delta_n}, q^{\delta_n} + u_x^{\delta_n}\sigma)$ converges weakly to $(\widehat{u},\widehat{r})$ in $L^2_{\mathscr{F}}(0,T;W^{2,2}) \times L^2_{\mathscr{F}}(0,T;[W^{2,2}]^{d'})$. The weak convergence of u^{δ_n} to $\widehat{u}$ in $L^2_{\mathscr{F}}(0,T;W^{2,2})$ also implies the weak convergence of $u_x^{\delta_n}\sigma$ to $\widehat{u}_x\sigma$ in $L^2_{\mathscr{F}}(0,T;W^{1,2})$. Hence $q^{\delta_n} = r^n - u_x^{\delta_n}\sigma$ converges weakly to $\widehat{q} \triangleq \widehat{r} - \widehat{u}_x\sigma$ in $L^2_{\mathscr{F}}(0,T;[W^{1,2}]^{d'})$.

Next we show that $\{u^{\delta_n}\}$ is a Cauchy sequence in $L^2_{\mathscr{F}}(0,T;W^{0,2})$. If so, the strong convergence of u^{δ_n} to $\widehat{u}$ in $L^2_{\mathscr{F}}(0,T;W^{0,2})$ follows and it is easy to see that $(\widehat{u},\widehat{q})$ is the unique solution to (3.6) referring to the arguments as in Theorem 3.1.

To prove that $\{u^{\delta_n}\}$ is a Cauchy sequence in $L^2_{\mathscr{F}}(0,T;W^{0,2})$, we set

$$u^{n,m} = u^{\delta_n} - u^{\delta_m}, \qquad q^{n,m} = q^{\delta_n} - q^{\delta_m}.$$

Obviously, $(u^{n,m}, q^{n,m})$ satisfies equations as follows:

$$\begin{cases} du^{n,m} = -\big\{(\delta_n\,\Delta + \mathcal{L})\,u^{n,m} + \mathcal{M}q^{n,m} + f(t,x,u^{\delta_n}) - f(t,x,u^{\delta_m})\\ \qquad\quad +(\delta_n - \delta_m)\,\Delta u^{\delta_m}\big\}\,dt + q^{n,m}\,dW_t\\ u^{n,m}(T) = 0.\end{cases}$$

By (2.5) in the case of $m_1 = 0$ and (3.11), for $t \in [0,T]$, we have

$$\begin{aligned}\mathbb{E}\|u^{n,m}(t)\|_{0,2}^2 &\le C\,\mathbb{E}\left\{\int_t^T \|f(s,\cdot,u^n) - f(s,\cdot,u^m)\|_{0,2}^2\, ds\right.\\ &\qquad\qquad \left. + (\delta_n - \delta_m)\int_s^T \|\Delta u^m(s)\|_{0,2}^2\, ds\right\}\\ &\le C\,\mathbb{E}\int_t^T \|u^{n,m}(s)\|_{0,2}^2\, ds + (\delta_n - \delta_m)\,CM_2,\end{aligned}$$

where the constant C is independent of δ_n, δ_m. Therefore, we can apply the Gronwall inequality and take $n, m \to \infty$ to deduce that $\{u^n\}$ is a Cauchy sequence in $L^2_{\mathscr{F}}(0, T; W^{0,2})$. The proof of Theorem 3.5 is complete. □

Remark 3.1. (i) Theorems 3.2 and 3.5 improve much in many aspects in comparison with Theorems 3.2 and 5.1 in Hu–Ma–Yong [10]. For example, our result includes the multi-dimensional equation and the coefficients σ, ν in BSPDE can depend on x (actually, all the coefficients in our setting are a function of (ω, t, x)). Also the regularity condition of coefficients in Theorem 3.5 is weaker than that in Theorem 5.1 in [10]. Needless to say, all these improvements are not trivial.

(ii) Denote by $D_x^i D_v^j f,\ i, j \in \mathbb{Z}^+$ the derivative of f which is i order with respect to x and j order with respect to v. For $m \geq 1$, if we assume

(1) conditions $(\mathbf{A}_m)$ and $(\mathbf{P})$ hold, and $\varphi \in L^2_{\mathscr{F}_T}(\Omega; W^{m,2}) \cap L^\infty_{\mathscr{F}_T}(\Omega; W^{m-1,\infty})$,
(2) for arbitrary (ω, t, x, v), all $D_x^i D_v^j f$ exist, where $0 \leq i, j \leq m$ and $i + j > 0$,
(3) $f(\cdot, \cdot, 0) \in L^2_{\mathscr{F}}(0, T; W^{m,2}) \cap L^\infty_{\mathscr{F}}(0, T; W^{m-1,\infty})$,
(4) all $D_x^i D_v^j f$ are bounded by L, where $0 \leq i \leq m-1, 0 \leq j \leq m$ and $i + j > 0$,
(5) for arbitrary (ω, t, x, v), $|D_x^m f(t, x, v)| \leq |D_x^m f(t, x, 0)| + L|v|$,

then from the argument of Theorem 3.5, it is not hard to prove that (3.6) has a unique generalized solution (u, q) satisfying

$$u \in C^w_{\mathscr{F}}([0, T]; W^{m,2}) \cap L^\infty_{\mathscr{F}}(0, T; W^{m-1,\infty}) \quad \text{and}$$
$$q + u_x\sigma \in L^2_{\mathscr{F}}(0, T; [W^{m,2}]^{d'}).$$

4. Connection between BSPDEs and FBSDEs

In this section, we study the connection between semi-linear BSPDEs and FBSDEs. This kind of connection is established in a non-Markovian framework and can be regarded as an extension of the Feynman–Kac formula for semi-linear PDEs and BSDEs (c.f. [19,20]).

First give a SDE whose coefficients may be non-Markovian:

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})dW_r, & s \geq t, \\ X_s^{t,x} = x, \quad 0 \leq s < t, \end{cases} \tag{4.1}$$

where $W = \{W_t; t \geq 0\}$ is a d'-dimensional Wiener process. We always assume that b, σ satisfy $(\mathbf{A}_1)$. Then we present a BSDE whose coefficients depend on the solution of the above SDE:

$$Y_s^{t,x} = \varphi(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x})dr - \int_s^T Z_r^{t,x} dW_r. \tag{4.2}$$

Thus (4.1) and (4.2) constitute a FBSDE system.

Remark 4.1. (i) Given $p > d$, note that

$$\mathbb{E}\int_0^T |f(s, X_s^{t,x}, 0)|^p ds \leq \mathbb{E}\int_0^T \|f(s, \cdot, 0)\|^p_{L^\infty} ds \leq C\mathbb{E}\int_0^T \|f(s, \cdot, 0)\|^p_{W^{1,p}} ds.$$

Hence, if $f(\cdot, \cdot, 0) \in L^p_{\mathscr{F}}(0, T; W^{1,p}), \varphi \in L^p_{\mathscr{F}_T}(\Omega; W^{1,p})$, and for any $\omega \in \Omega, s \in [0, T]$, $f(s, x, y)$ satisfies the uniformly Lipschitz condition with respect to y, we can first prove that there exists a unique solution $(Y_s^{t,x}, Z_s^{t,x})_{s\in[t,T]}$ to BSDE (4.2) in a standard way. Then, by the results of Proposition 3.2 in [4], we know that the solution of BSDE (4.2) satisfies

$$\mathbb{E}\sup_{t\le s\le T}|Y_s^{t,x}|^p ds+\mathbb{E}\int_t^T|Y_s^{t,x}|^{p-2}|Z_s^{t,x}|^2ds+\left(\mathbb{E}\int_t^T|Z_s^{t,x}|^2ds\right)^{\frac{p}{2}}<\infty. \tag{4.3}$$

(ii) For $s\in[0,t]$, (4.2) is equivalent to the following BSDE:

$$Y_s^x=Y_t^{t,x}+\int_s^t f(r,x,Y_r^x)dr-\int_s^t Z_r^x dW_r.$$

As stated in (i), in view of $Y_t^{t,x}\in L^p_{\mathscr{F}_T}(\Omega;L^p)$, the above equation has a unique solution $(Y_s^x,Z_s^x)_{s\in[0,t]}$. To unify the notation, we define $(Y_s^{t,x},Z_s^{t,x})=(Y_s^x,Z_s^x)$ when $s\in[0,t)$.

Our purpose is to investigate the connection between FBSDE (4.1), (4.2) and the following BSPDE:

$$\begin{cases} du=-\big[a^{ij}u_{x^ix^j}+b^iu_{x^i}+\sigma^{ik}q^k_{x^i}+f(t,x,u)\big]dt+q^kdW_t^k\\ u(T,x)=\varphi(x),\quad x\in\mathbb{R}^d,\end{cases} \tag{4.4}$$

where $a^{ij}=\frac{1}{2}\sigma^{ik}\sigma^{jk}$.

As for the case that f in BSDE (4.2) involves nonlinear $Z_r^{t,x}$, correspondingly f in the above semi-linear degenerate BSPDE should involve nonlinear $q+u_x\sigma$. However, for this kind of BSPDE a high regularity of u is still an unsolved problem in our settings, which is necessary to establish the correspondence between semi-linear BSPDE and FBSDE as our method below shows. So here we assume that f in BSDE (4.2) only involves nonlinear $Y_r^{t,x}$.

We begin with the linear case that $f(t,x,y,z)=c(t,x)y+\nu^k(t,x)z^k+F(t,x)$ and in this case the involving BSDE has a form like below:

$$\begin{aligned} Y_s^{t,x}&=\varphi(X_T^{t,x})+\int_s^T\big[c(r,X_r^{t,x})Y_r^{t,x}+\nu(r,X_r^{t,x})Z_r^{t,x}+F(r,X_r^{t,x})\big]dr\\ &\quad-\int_s^T Z_r^{t,x}dW_r. \end{aligned} \tag{4.5}$$

The corresponding linear BSPDE is as follows:

$$\begin{cases} du=-\big[\mathcal{L}u+\mathcal{M}q+F\big]dt+q^kdW_t^k\\ u(T,x)=\varphi(x),\quad x\in\mathbb{R}^d,\end{cases} \tag{4.6}$$

where $\mathcal{L}$, $\mathcal{M}$ are defined as in (1.1).

Referring to Lemma 4.5.6 in [13], we first give a useful lemma.

Lemma 4.1. *Under condition* $(\mathbf{A}_1)$*, for* $p\ge 1, t',t\in[0,T]$*, the stochastic flow defined by* (4.1) *satisfies*

$$\begin{aligned} &\mathbb{E}\sup_{s\in[0,T]}|X_s^{t',x'}-X_s^{t,x}|^{2p}\\ &\quad\le C(p,T)\left(1+|x|^{2p}+|x'|^{2p}\right)\left(|x'-x|^{2p}+|t'-t|^p\right)\quad a.s. \end{aligned}$$

The following proposition borrows ideas from [19,25]. Although $F(s,x)$ is not Lipschitz continuous on x, we still can derive the continuity of $Y_t^{t,x}$ due to the Hölder continuity of $F(s,x)$ on x.

Proposition 4.2. *Let conditions* ($\mathbf{A}_1$) *be satisfied. For a given* $p > 2d+2$, *suppose* $F \in L^p_{\mathscr{F}}(0,T;W^{1,p})$ *and* $\varphi \in L^p_{\mathscr{F}_T}(\Omega;W^{1,p})$. *If* $(Y_s^{t,x})_{s\in[t,T]}$ *is the solution of BSDE* (4.5), *then for* $t \in [0,T]$, $x \in \mathbb{R}^d$, $(t,x) \longrightarrow Y_t^{t,x}$ *is a.s. continuous.*

Proof. By Remark 4.1, we know that BSDE (4.5) has a unique solution $(Y_s^{t,x}, Z_s^{t,x})_{s\in[0,T]}$ and it satisfies (4.3). For $t,t' \in [0,T]$, $x,x' \in \mathbb{R}^d$, $s \geq 0$, $0<\beta<1$, assume without loss of any generality that $\beta p > 2d+2$ and $|x-x'| \leq 1$. Applying Itô's formula to $\mathrm{e}^{\beta pKs}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p}$, we have

$$\begin{aligned}
&\mathrm{e}^{\beta pKs}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p} + \beta pK\int_s^T \mathrm{e}^{\beta pKr}|Y_r^{t',x'} - Y_r^{t,x}|^{\beta p}dr\\
&\quad + \frac{\beta p(\beta p-1)}{2}\int_s^T \mathrm{e}^{\beta pKr}|Y_r^{t',x'} - Y_r^{t,x}|^{\beta p-2}|Z_r^{t',x'} - Z_r^{t,x}|^2dr\\
&\leq \mathrm{e}^{\beta pKT}|\varphi(X_T^{t',x'}) - \varphi(X_T^{t',x'})|^{\beta p} + (\beta p + 2K_1^2\beta p)\int_s^T \mathrm{e}^{\beta pKr}|Y_r^{t',x'} - Y_r^{t,x}|^{\beta p}dr\\
&\quad + \frac{\beta p}{2}\int_s^T \mathrm{e}^{\beta pKr}|Y_r^{t',x'} - Y_r^{t,x}|^{\beta p-2}|c(r,X_r^{t',x'})Y_r^{t',x'} - c(r,X_r^{t,x})Y_r^{t,x}|^2dr\\
&\quad + \frac{\beta p}{8K_1^2}\int_s^T \mathrm{e}^{\beta pKr}|Y_r^{t',x'} - Y_r^{t,x}|^{\beta p-2}|\nu(r,X_r^{t',x'})Z_r^{t',x'} - \nu(r,X_r^{t,x})Z_r^{t,x}|^2dr\\
&\quad + \beta p\int_s^T \mathrm{e}^{\beta pKr}|Y_r^{t',x'} - Y_r^{t,x}|^{\beta p-2}|F(r,X_r^{t',x'}) - F(r,X_r^{t,x})|^2dr\\
&\quad - \frac{\beta p}{2}\int_s^T \mathrm{e}^{\beta pKr}(Y_r^{t',x'} - Y_r^{t,x})^{\beta p-2}(Y_r^{t',x'} - Y_r^{t,x})(Z_r^{t',x'} - Z_r^{t,x})dW_r.
\end{aligned} \tag{4.7}$$

We need to deal with those terms involving the solutions of coupled SDEs on the right hand side of (4.7), so as to follow the procedure in the proof of Proposition 3.2 in [4] to deduce the continuity dependence of the solution of BSDE under expectation. First,

$$\begin{aligned}
&\mathbb{E}\int_t^T \mathrm{e}^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p-2}|c(s,X_s^{t',x'})Y_s^{t',x'} - c(s,X_s^{t,x})Y_s^{t,x}|^2ds\\
&\leq 2\mathbb{E}\int_t^T \mathrm{e}^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p-2}\big(|c(s,X_s^{t',x'})|^2|Y_s^{t',x'} - Y_s^{t,x}|^2\\
&\quad + |c(s,X_s^{t',x'}) - c(s,X_s^{t,x})|^2|Y_s^{t,x}|^2\big)ds\\
&\leq 2K_1^2\mathbb{E}\int_t^T \mathrm{e}^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p}ds + \varepsilon\mathbb{E}\int_t^T \mathrm{e}^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p}ds\\
&\quad + C\left(\mathbb{E}\int_t^T |Y_s^{t,x}|^p ds\right)^{\beta}\left(\mathbb{E}\int_t^T |X_s^{t',x'} - X_s^{t,x}|^{\frac{\beta p}{1-\beta}}ds\right)^{1-\beta},
\end{aligned} \tag{4.8}$$

where $\varepsilon > 0$ is a generic constant which can be taken sufficiently small. Similarly, it follows that

$$\begin{aligned}
&\mathbb{E}\int_t^T e^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p-2}|\nu(s, X_s^{t',x'})Z_s^{t',x'} - \nu(s, X_s^{t,x})Z_s^{t,x}|^2 ds \\
&\le 2K_1^2\mathbb{E}\int_t^T e^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p-2}|Z_s^{t',x'} - Z_s^{t,x}|^2 ds \\
&\quad + 2K_1^2\mathbb{E}\int_t^T e^{\beta pKr}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p-2}|X_s^{t',x'} - X_s^{t,x}|^2|Z_s^{t,x}|^2 ds. \qquad (4.9)
\end{aligned}$$

Noticing that the $C^{0,\alpha}$ norm is controlled by the $W^{1,p}$ norm in view of the Sobolev embedding theorem with $\alpha = 1 - \frac{d}{p} < 1$, we have

$$\begin{aligned}
&\mathbb{E}\int_t^T e^{\beta pKr}|F(s, X_s^{t',x'}) - F(s, X_s^{t,x})|^{\beta p} ds \\
&\le \mathbb{E}\int_t^T e^{\beta pKr}\|F(s,\cdot)\|_{C^{0,\alpha}}^{\beta p}|X_s^{t',x'} - X_s^{t,x}|^{\alpha\beta p} ds \\
&\le \left(\mathbb{E}\int_t^T e^{\beta pKr}\|F(s,\cdot)\|_{W^{1,p}}^{p} ds\right)^{\beta}\left(\mathbb{E}\int_t^T e^{\beta pKr}|X_s^{t',x'} - X_s^{t,x}|^{\frac{\alpha\beta p}{1-\beta}} ds\right)^{1-\beta} \\
&\le C\left(\mathbb{E}\int_t^T |X_s^{t',x'} - X_s^{t,x}|^{\frac{\alpha\beta p}{1-\beta}} ds\right)^{1-\beta}. \qquad (4.10)
\end{aligned}$$

Similar to above, we can also get

$$\mathbb{E}e^{\beta pKr}|\varphi(X_T^{t',x'}) - \varphi(X_T^{t,x})|^{\beta p} \le C\left(\mathbb{E}|X_T^{t',x'} - X_T^{t,x}|^{\frac{\alpha\beta p}{1-\beta}}\right)^{1-\beta}. \qquad (4.11)$$

Then following the proof of Proposition 3.2 in [4], by (4.8)–(4.11), Lemma 4.1 and the fact that K can be taken sufficiently large K, we similarly deduce

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[0,T]}|Y_s^{t',x'} - Y_s^{t,x}|^{\beta p} \\
&\le C\left(\mathbb{E}|X_T^{t',x'} - X_T^{t,x}|^{\frac{\alpha\beta p}{1-\beta}}\right)^{1-\beta} + C\mathbb{E}\left(\int_0^T |X_s^{t',x'} - X_s^{t,x}|^{\frac{\beta p}{1-\beta}} ds\right)^{1-\beta} \\
&\quad + C\left(\mathbb{E}\sup_{s\in[0,T]}|X_s^{t,x,\varepsilon} - X_s^{t,x}|^{\frac{\beta p}{1-\beta}}\right)^{1-\beta}\left(\mathbb{E}\left(\int_0^T |Z_s^{t,x}|^2 ds\right)^{\frac{p}{2}}\right)^{\beta} \\
&\quad + C\left(\mathbb{E}\int_0^T |X_s^{t',x'} - X_s^{t,x}|^{\frac{\alpha\beta p}{1-\beta}} ds\right)^{1-\beta} \\
&\le C(p,T)\left(1 + |x|^p + |x'|^p\right)\left(|x'-x|^{\alpha\beta p} + |t'-t|^{\frac{\alpha\beta p}{2}}\right) \quad \text{a.s.}
\end{aligned}$$

Since $\beta p > 2d + 2$, by the Kolmogorov continuity theorem (see e.g. Theorem 1.4.1 in [13]) we know that $Y_s^{(\cdot,\cdot)}$ has a continuous modification for $t \in [0, T]$ and $x \in \bar{B}_R$ with the norm $\sup_{s\in[0,T]}|Y_s^{(\cdot,\cdot)}|$, where $\bar{B}_R$ is defined as in Corollary 3.4. In particular,

$$\lim_{\substack{t'\to t\\ x'\to x}} |Y_{t'}^{t',x'} - Y_{t'}^{t,x}| = 0.$$

Thus we have

$$\lim_{\substack{t'\to t\\x'\to x}} |Y_{t'}^{t',x'} - Y_t^{t,x}| \le \lim_{\substack{t'\to t\\x'\to x}} (|Y_{t'}^{t',x'} - Y_{t'}^{t,x}| + |Y_{t'}^{t,x} - Y_t^{t,x}|) = 0 \quad \text{a.s.}$$

The convergence of the second term follows from the continuity of $Y_s^{t,x}$ in s. That is to say $Y_t^{t,x}$ is a.s. continuous; therefore $Y_t^{t,x}$ is continuous with respect to $t \in [0, T]$ and $x \in \bar{B}_R$ on a full-measure set Ω^R for any $R \in \mathbb{N}$. Taking $\tilde{\Omega} = \bigcap_{R\in\mathbb{N}} \Omega^R$, we have $P(\tilde{\Omega}) = 1$. Since $\bigcup_{R\in\mathbb{N}} \bar{B}_R = \mathbb{R}^d$, for any $t \in [0, T]$ and $x \in \mathbb{R}^d$, there exists an R such that $x \in \bar{B}_R$. On the other hand, for any $\omega \in \tilde{\Omega}$, obviously $\omega \in \Omega^R$, $R = 1, 2, \cdots$. So $Y_t^{t,x}$ is continuous with respect to $t \in [0, T]$ and $x \in \mathbb{R}^d$ on $\tilde{\Omega}$. Proposition 4.2 is proved. □

Then we can get the correspondence between BSPDE and FBSDE in the linear case.

Theorem 4.3. *Let conditions* ($\mathbf{A}_1$) *and* (**P**) *be satisfied. For a given* $p > 2d + 2$, *suppose* $F \in L^p_{\mathscr{F}}(0, T; W^{1,p})$ *and* $\varphi \in L^p_{\mathscr{F}_T}(\Omega; W^{1,p})$, *then the solution* (u, q) *to BSPDE* (4.6) *satisfies*

$$u(s, X_s^{t,x}) = Y_s^{t,x} \quad \text{for all } s \in [t, T],\ x \in \mathbb{R}^d \text{ a.s.}, \tag{4.12}$$

where (X, Y, Z) *is the solution of FBSDE* (4.1), (4.5).

Proof. *Step* 1. First we smootherize all the coefficients in FBSDE (4.1), (4.5) and BSPDE (4.6). For this, take a nonnegative function $\rho \in C_0^\infty(\mathbb{R}^d; \mathbb{R}^1)$ such that $\int_{\mathbb{R}^d} \rho(x)dx = 1$. For arbitrary $\varepsilon > 0$ and a mapping $h : \mathbb{R}^d \longrightarrow \mathbb{R}^1$, we define h^ε by

$$h^\varepsilon(x) = \varepsilon^{-d} \rho\left(\frac{x}{\varepsilon}\right) * h(x) \quad \text{for } x \in \mathbb{R}^d.$$

Moreover, if h is a vector or matrix, we get the smootherized h^ε by smootherizing each element in h. In this way, we can smootherize all the coefficients and get three equations with smootherized coefficients:

$$\begin{cases} X_s^{t,x,\varepsilon} = x + \int_t^s b^\varepsilon(r, X_r^{t,x})dr + \int_t^s \sigma^\varepsilon(r, X_r^{t,x})dW_r, & s \ge t, \\ X_s^{t,x,\varepsilon} = x, \quad 0 \le s < t, \end{cases} \tag{4.13}$$

$$\begin{aligned} Y_s^{t,x,\varepsilon} \ &= \ \varphi^\varepsilon(X_T^{t,x,\varepsilon}) + \int_s^T \left[c^\varepsilon(r, X_r^{t,x,\varepsilon})Y_r^{t,x,\varepsilon} + \nu^\varepsilon(r, X_r^{t,x,\varepsilon})Z_r^{t,x,\varepsilon} + F^\varepsilon(r, X_r^{t,x,\varepsilon})\right]dr \\ &\quad - \int_s^T Z_r^{t,x,\varepsilon} dW_r \end{aligned} \tag{4.14}$$

and

$$\begin{cases} du^\varepsilon(t, x) = -\left[\mathcal{L}^\varepsilon u^\varepsilon(t, x) + \mathcal{M}^\varepsilon q^\varepsilon(t, x) + F^\varepsilon(t, x)\right]dt + q^\varepsilon(t, x)dW_t \\ u^\varepsilon(T, x) = \varphi^\varepsilon(x), \quad x \in \mathbb{R}^d, \end{cases} \tag{4.15}$$

where $X^\varepsilon, (Y^\varepsilon, Z^\varepsilon)$ and $(u^\varepsilon, q^\varepsilon)$ are the unique solutions of (4.13), (4.14) and (4.15), respectively. Due to the smooth coefficients, we know that all $X_s^{t,x,\varepsilon}$, $Y_s^{t,x,\varepsilon}$ and $u^\varepsilon(t, x)$ have a high regularity on variable x such that $u^\varepsilon(s, X_s^{t,x,\varepsilon}), Y_s^{t,x,\varepsilon} \in C^{1,2}([0, T] \times \mathbb{R}^d)$. By the Itô–Wentzell formula, it is not hard to deduce that $u^\varepsilon(s, X_s^{t,x,\varepsilon})$ is also a solution to BSDE (4.14). Due to the uniqueness of solution, we have $u^\varepsilon(s, X_s^{t,x,\varepsilon}) = Y_s^{t,x,\varepsilon}$ for a.e. $t \in [0, T]$, $x \in \mathbb{R}^d$ a.s., then the continuity with respect to t, x ensures that this equality is true for all $t \in [0, T]$, $x \in \mathbb{R}^d$ on a full measure set in Ω.

Step 2. We then prove that as for a.e. $x \in \mathbb{R}^d$,

$$\mathbb{E} \sup_{t \le s \le T} |Y_s^{t,x,\varepsilon} - Y_s^{t,x}|^2 \longrightarrow 0, \quad \text{as } \varepsilon \to 0. \tag{4.16}$$

First noting condition ($\mathbf{A}_1$) and the construction of convolution we have

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \mathbb{E} \int_t^T |b^\varepsilon(s, X_s^{t,x,\varepsilon}) - b(s, X_s^{t,x})|^2 ds \\
&\quad \le \lim_{\varepsilon \to 0} 2\mathbb{E} \int_t^T \left(\sup_{y \in \mathbb{R}^d} |b^\varepsilon(s,y) - b(s,y)|^2 + K_1^2 |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 \right) ds \\
&\quad = \lim_{\varepsilon \to 0} 2K_1^2 \mathbb{E} \int_t^T |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 ds.
\end{aligned}$$

A similar calculation leads to

$$\lim_{\varepsilon \to 0} \mathbb{E} \int_t^T |\sigma^\varepsilon(s, X_s^{t,x,\varepsilon}) - \sigma(s, X_s^{t,x})|^2 ds \le \lim_{\varepsilon \to 0} 2K_1^2 \mathbb{E} \int_t^T |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 ds.$$

Hence applying Itô's formula and the B–D–G inequality, we have

$$\lim_{\varepsilon \to 0} \mathbb{E} \sup_{t \le s \le T} |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 ds = 0,$$

and there exists a subsequence of $\{X_s^{t,x,\varepsilon}\}$, still denoted by $\{X_s^{t,x,\varepsilon}\}$, which satisfies

$$\lim_{\varepsilon \to 0} \sup_{t \le s \le T} |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 = 0 \quad \text{a.s.}$$

In the rest of arguments, we always consider this a.s. continuous subsequence.

In order to get (4.16), we need to deal with the following convergence:

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \mathbb{E} \int_t^T |c^\varepsilon(s, X_s^{t,x,\varepsilon}) Y_s^{t,x,\varepsilon} - c(s, X_s^{t,x}) Y_s^{t,x}|^2 ds \\
&\quad \le \lim_{\varepsilon \to 0} 2\mathbb{E} \int_t^T \Big(|c^\varepsilon(s, X_s^{t,x,\varepsilon})|^2 |Y_s^{t,x,\varepsilon} - Y_s^{t,x}|^2 \\
&\qquad + |c^\varepsilon(s, X_s^{t,x,\varepsilon}) - c(s, X_s^{t,x})|^2 |Y_s^{t,x}|^2 \Big) ds \\
&\quad \le \lim_{\varepsilon \to 0} 2K_1^2 \mathbb{E} \int_t^T |Y_s^{t,x,\varepsilon} - Y_s^{t,x}|^2 ds \\
&\qquad + 4\mathbb{E} \int_t^T \lim_{\varepsilon \to 0} \left(\sup_{y \in \mathbb{R}^d} |c^\varepsilon(s,y) - c(s,y)|^2 + K_1^2 |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 \right) |Y_s^{t,x}|^2 ds \\
&\quad = \lim_{\varepsilon \to 0} 2K_1^2 \mathbb{E} \int_t^T |Y_s^{t,x,\varepsilon} - Y_s^{t,x}|^2 ds.
\end{aligned} \tag{4.17}$$

Similarly, we have

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \mathbb{E} \int_t^T |\nu^\varepsilon(s, X_s^{t,x,\varepsilon}) Z_s^{t,x,\varepsilon} - \nu(s, X_s^{t,x}) Z_s^{t,x}|^2 ds \\
&\quad \le \lim_{\varepsilon \to 0} 2K_1^2 \mathbb{E} \int_t^T |Z_s^{t,x,\varepsilon} - Z_s^{t,x}|^2 ds.
\end{aligned} \tag{4.18}$$

By the Sobolev embedding theorem again, it yields that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\mathbb{E}\int_t^T |F^\varepsilon(s, X_s^{t,x,\varepsilon}) - F(s, X_s^{t,x})|^2 ds \\
&\quad \le \lim_{\varepsilon\to 0} 2\left(\mathbb{E}\int_t^T \|F(s,\omega)\|^2_{W^{1,p}} ds\right)^{\frac{1}{2}} \left(\mathbb{E}\int_t^T |X_s^{t,x,\varepsilon} - X_s^{t,x}|^{2\alpha} ds\right)^{\frac{1}{2}} \\
&\quad \le \lim_{\varepsilon\to 0} C\left(\mathbb{E}\int_t^T |X_s^{t,x,\varepsilon} - X_s^{t,x}|^2 ds\right)^{\frac{\alpha}{2}} = 0. \qquad (4.19)
\end{aligned}$$

Similarly, we also obtain

$$\lim_{\varepsilon\to 0}\mathbb{E}|\varphi^\varepsilon(X_T^{t,x,\varepsilon}) - \varphi(X_T^{t,x})|^2 \le \lim_{\varepsilon\to 0} C\left(\mathbb{E}|X_T^{t,x,\varepsilon} - X_T^{t,x}|^2\right)^{\frac{\alpha}{2}} = 0. \qquad (4.20)$$

In view of (4.17)–(4.20), a priori estimates of the solution of BSDE with sufficiently large K (see [4]) yields that (4.16) is true.

Step 3. On the other hand, we can further prove

$$\mathbb{E}\int_0^T \|u^\varepsilon(t) - u(t)\|^2_{0,2} dt \longrightarrow 0.$$

Indeed, similar to inequality (2.9), it is not hard to prove

$$\begin{aligned}
&\mathbb{E}\int_0^T e^{\lambda t}\left\|u(t)^\varepsilon - u(t)\right\|^2_{0,2} dt \\
&\quad \le C e^{\lambda T}\mathbb{E}\left\|\varphi^\varepsilon - \varphi\right\|^2_{0,2} + C\mathbb{E}\int_0^T e^{\lambda t}\left\|F(t)^\varepsilon - F(t)\right\|^2_{0,2} dt \\
&\qquad + C\mathbb{E}\int_0^T\int_{\mathbb{R}^d} e^{\lambda t}(u^\varepsilon - u)\big[(a^{\varepsilon,ij} - a^{ij})u_{x^i x^j} + (\sigma^{\varepsilon,ik} - \sigma^{ik})q^k_{x^i} \\
&\qquad + (b^{\varepsilon,i} - b^i)u_{x^i} + (c^\varepsilon - c)u + (\nu^\varepsilon - \nu)q\big] dx dt,
\end{aligned}$$

where $a^{\varepsilon,ij} = \frac{1}{2}\sigma^{\varepsilon,ik}\sigma^{\varepsilon,jk}$, $a^{ij} = \frac{1}{2}\sigma^{ik}\sigma^{jk}$ and λ is a sufficiently large number. By condition $(\mathbf{A}_1)$ and constructions of smootherized coefficients, we can deduce for each (t, x, ω),

$$\begin{aligned}
&\left|D(a^{\varepsilon,ij} - a^{ij})\right| + \left|D(\sigma^{\varepsilon,ik} - \sigma^{ik})\right| + \left|a^{\varepsilon,ij} - a^{ij}\right| + \left|\sigma^{\varepsilon,ik} - \sigma^{ik}\right| \\
&\quad + \left|b^{\varepsilon,i} - b^i\right| + \left|c^\varepsilon - c\right| + \left|\nu^\varepsilon - \nu\right| \le C\varepsilon.
\end{aligned}$$

Thus, by integration by parts, it turns out that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\mathbb{E}\int_0^T e^{\lambda t}\left\|u^\varepsilon(t) - u(t)\right\|^2_{0,2} dt \\
&\quad \le \lim_{\varepsilon\to 0}\varepsilon C e^{\lambda T}\mathbb{E}\int_0^T\int_{\mathbb{R}^d}\big[|D\left(u^\varepsilon - u\right)|(|Du| + |q|) \\
&\qquad + |u^\varepsilon - u|(|Du| + |u| + |q|)\big] dx dt \\
&\quad \le \lim_{\varepsilon\to 0}\varepsilon C e^{\lambda T}\mathbb{E}\int_0^T\left(\left\|u^\varepsilon - u\right\|^2_{1,2} + \|u\|^2_{1,2} + \|q\|^2_{0,2}\right) dt
\end{aligned}$$

$$\leq \lim_{\varepsilon\to 0} \varepsilon C e^{\lambda T} \mathbb{E}\left[\|\varphi^\varepsilon\|_{1,2}^2 + \|\varphi\|_{1,2}^2 + \int_0^T \left(\|F^\varepsilon(t)\|_{1,2}^2 + \|F(t)\|_{1,2}^2 \right) dt \right] = 0.$$

Therefore, there exists a subsequence of $\{u^\varepsilon\}$, still denoted by $\{u^\varepsilon\}$, such that $u^\varepsilon(t,x) \longrightarrow u(t,x)$ as $\varepsilon \to 0$ for a.e. $t \in [0,T], x \in \mathbb{R}^d$ a.s., which implies that $Y_t^{t,x} = u(t,x)$ for a.e. $t \in [0,T], x \in \mathbb{R}^d$ a.s. in view of (4.16). Noticing Corollary 3.4, we know that $u(t,x)$ is continuous with respect to (t,x), which together with Proposition 4.2 leads to

$$u(t,x) = Y_t^{t,x} \quad \text{for all } t \in [0,T],\ x \in \mathbb{R}^d \text{ a.s.}$$

In particular,

$$u(s, X_s^{t,x}) = Y_s^{s,X_s^{t,x}} \quad \text{for all } s \in [t,T],\ x \in \mathbb{R}^d \text{ a.s.}$$

By the uniqueness of the solution of BSDE (4.14), (4.12) follows. □

Utilizing the connection between FBSDE and BSPDE in the linear case, we further study the same kind of connection in the semi-linear case.

Theorem 4.4. *Suppose that the conditions in Theorem 3.2 are satisfied; then we have a same kind of connection as (4.12) between the solution u to BSPDE (4.4) and the solution (X,Y) to FBSDE (4.1), (4.2).*

Proof. Let u be the solution of BSPDE (4.4) and $\hat{F}(t,x) = f(t,x,u(t,x))$. Obviously, $\hat{F}(t,x) \in L^p_{\mathscr{F}}(0,T;W^{1,p})$ and we regard BSPDE (4.4) as a linear equation with generator $\hat{F}$. By Theorem 4.3, we know

$$u(s, X_s^{t,x}) = \hat{Y}_s^{t,x} \quad \text{for all } s \in [t,T],\ x \in \mathbb{R}^d \text{ a.s.,}$$

where $(\hat{Y}_s^{t,x}, \hat{Z}_s^{t,x})_{s\in[t,T]}$ is the solution of the BSDE with generator $\hat{F}$ as follows:

$$\hat{Y}_s^{t,x} = \varphi(X_T^{t,x}) + \int_s^T \hat{F}(r, X_r^{t,x})dr - \int_s^T \hat{Z}_r^{t,x} dW_r.$$

By the definition of $\hat{F}$, we know that $(\hat{Y}_s^{t,x}, \hat{Z}_s^{t,x})$ is also the solution of BSDE (4.2). Since Remark 4.1, the solution of BSDE (4.2) is unique. Then Theorem 4.4 follows immediately. □

Acknowledgments

We would like to thank Prof. Shanjian Tang for constructive suggestions and the anonymous referees for their valuable comments.

References

[1] V. Barbu, A. Răşcanu, G. Tessitore, Carleman estimates and controllability of linear stochastic heat equations, Appl. Math. Optim. 47 (2003) 97–120.

[2] A. Bensoussan, Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions, Stochastics 9 (1983) 169–222.

[3] A. Bensoussan, Stochastic maximum principle for distributed parameter systems, J. Franklin Inst. 315 (1983) 387–406.

[4] Ph. Briand, B. Delyon, Y. Hu, E. Pardoux, L. Stoica, L^p solutions ofbackward stochastic differential equations, Stoch. Process. Appl. 108 (2003) 109–129.

[5] M. Chang, Y. Pang, J. Yong, Optimal stopping problem for stochastic differential equations with random coefficients, SIAM J. Control Optim. 48 (2009) 941–971.

[6] K. Du, Q. Meng, A revisit to W_2^n-theory of super-parabolic backward stochastic partial differential equations in $\mathbb{R}^d$, Stoch. Process. Appl. 120 (2010) 1996–2015.
[7] K. Du, S. Tang, Q. Zhang, $W^{m,p}$-Solution ($p \geq 2$) of linear degenerate backward stochastic partial differential equations in the whole space, J. Differential Equations (2013). http://dx.doi.org/10.1016/j.jde.2013.01.013.
[8] N. Englezos, I. Karatzas, Utility maximization with habit formation: dynamic programming and stochastic PDEs, SIAM J. Control Optim. 48 (2009) 481–520.
[9] U.G. Haussmann, The maximum principle for optimal control of diffusions with partial information, SIAM J. Control Optim. 25 (1987) 341–361.
[10] Y. Hu, J. Ma, J. Yong, On semi-linear degenerate backward stochastic partial differential equations, Probab. Theory Related Fields 123 (2002) 381–411.
[11] Y. Hu, S. Peng, Adapted solution of a backward semilinear stochastic evolution equation, Stoch. Anal. Appl. 9 (1991) 445–459.
[12] N.V. Krylov, B.L. Rozovskii, Stochastic evolution equations, J. Sov. Math. 16 (1981) 1233–1277.
[13] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge Univ. Press, 1990.
[14] J. Ma, J. Yong, Adapted solution of a degenerate backward SPDE with applications, Stoch. Process. Appl. 70 (1997) 59–84.
[15] J. Ma, J. Yong, On linear degenerate backward stochastic partial differential equations, Probab. Theory Related Fields 113 (1999) 135–170.
[16] M. Mania, R. Tevzadze, Backward stochastic PDE and imperfect hedging, Int. J. Theor. Appl. Finance 6 (2003) 663–692.
[17] M. Mania, T. Zariphopoulou, Stochastic partial differential equations and portfolio choice, in: C. Chiarella, A. Novikov (Eds.), Contemparary Quantitative Finance: Essays in Honour of Eckhard Platen, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 195–215.
[18] N. Nagasa, M. Nisio, Optimal controls for stochastic partial differential equations, SIAM J. Control Optim. 28 (1990) 186–213.
[19] E. Pardoux, S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, in: B.L. Rozuvskii, R.B. Sowers (Eds.), Stochastic Partial Differential Equations, in: Lect. Notes Control Inf. Sci., vol. 176, Springer, Berlin, Heidelberg, New York, 1992, pp. 200–217.
[20] S. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, Stochastics 37 (1991) 61–74.
[21] S. Peng, Stochastic Hamilton–Jacobi–Bellman equations, SIAM J. Control Optim. 30 (1992) 284–304.
[22] S. Tang, The maximum principle for partially observed optimal control of stochastic differential equations, SIAM J. Control Optim. 36 (1998) 1596–1617.
[23] S. Tang, Semi-linear systems of backward stochastic partial differential equations in $\mathbb{R}^n$, Chin. Ann. Math. 26B (2005) 437–456.
[24] S. Tang, X. Zhang, Null controllability for forward and backward stochastic parabolic equations, SIAM J. Control Optim. 48 (2009) 2191–2216.
[25] Q. Zhang, Stationary stochastic viscosity solutions of SPDEs, Stoch. Dyn. 11 (2011) 691–713.
[26] X. Zhou, On the necessary conditions of optimal controls for stochastic partial differential equations, SIAM J. Control Optim. 31 (1993) 1462–1478.