Adapted solution of a degenerate backward spde, with applications

Jin Ma[a],[*],[1], Jiongmin Yong[b],[2]

[a]*Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395, USA*
[b]*Department of Mathematics, Fudan University, Shanghai 200433, China*

Abstract

In this paper we prove the existence and uniqueness, as well as the regularity, of the *adapted* solution to a class of degenerate linear backward stochastic partial differential equations (BSPDE) of parabolic type. We apply the results to a class of forward–backward stochastic differential equations (FBSDE) with random coefficients, and establish in a special case some explicit formulas among the solutions of FBSDEs and BSPDEs, including those involving Malliavin calculus. These relations lead to an adapted version of stochastic Feynman–Kac formula, as well as a stochastic Black–Scholes formula in mathematical finance. © 1997 Elsevier Science B.V.

AMS classification: 60H15; 35R60; 34F05; 93E20

Keywords: Degenerate backward stochastic partial differential equations; Adapted solutions; Forward–backward stochastic differential equations; Malliavin calculus; Feynman–Kac formula; Option pricing

1. Introduction

In this paper we study a class of (linear) backward stochastic partial differential equations (BSPDE for short) of the following type:

$$u(t,x,\cdot) = g(x,\cdot) + \int_t^T \{(\mathscr{L}u)(s,x,\cdot) + (\mathscr{M}q)(s,x,\cdot) + f(s,x,\cdot)\}\,ds$$

$$- \int_t^T \langle q(s,x,\cdot),\mathrm{d}W_s\rangle, \quad (t,x) \in [0,T] \times \mathbb{R}^n, \tag{1.1}$$

where $W = \{W_t: t \in [0,T]\}$ is a d-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathscr{F}, P; \{\mathscr{F}_t\}_{t \geqslant 0})$, with $\{\mathscr{F}_t\}_{t \geqslant 0}$ being the natural filtration generated

by W, augmented by all the P-null sets in $\mathscr{F}$; and for $\varphi \in C^2(\mathbb{R}^n)$, $\psi \in C^1(\mathbb{R}^n; \mathbb{R}^d)$:

$$(\mathscr{L}\varphi)(t,x,\omega) = \sum_{i,j=1}^{n} a^{ij}(t,x,\omega)\varphi_{x_i x_j} + \sum_{i=1}^{n} a^i(t,x,\omega)\varphi_{x_i} + a_0(t,x,\omega)\varphi;$$

$$(\mathscr{M}\psi)(t,x,\omega) = \sum_{k=1}^{n}\sum_{\ell=1}^{d} b^{k\ell}(t,x,\omega)\psi^\ell_{x_k} + \sum_{\ell=1}^{d} b^\ell(t,x,\omega)\psi^\ell.$$

(1.2)

We assume that $a^{ij}, a^i, a_0, b^{k\ell}$ and b^ℓ, $i,j,k=1,\ldots,n$, $\ell=1,\ldots,d$ (resp. g) are real-valued measurable random fields defined on $[0,T] \times \mathbb{R}^n \times \Omega$ (resp. $\mathbb{R}^n \times \Omega$), such that for fixed x, they are $\mathscr{F}_t$-progressively measurable (resp. $\mathscr{F}_T$-measurable). Further, we assume that $a^{ij} = a^{ji}$, $i,j = 1,2,\ldots,n$, and the following *parabolicity condition* holds:

$$(A^{ij}) \stackrel{\triangle}{=} \left(2a^{ij} - \sum_{\ell=1}^{d} b^{i\ell}b^{j\ell} \right) \geqslant 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n \quad \text{a.s.} \quad \omega \in \Omega,$$

(1.3)

where (c^{ij}) denotes a matrix. Our purpose is to find a pair of random fields $(u,q):[0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R} \times \mathbb{R}^d$, such that for each fixed $x \in \mathbb{R}^n$, $\{u(\cdot,x,\cdot), q(\cdot,x,\cdot)\}$ is a pair of adapted processes, and that (1.1) is satisfied almost surely; and to study the regularity of the solution pair (u,q) in the variable x.

We should point out here that the stochastic integral in (1.1) is a *forward* Itô integral and the solution pair is assumed to be adapted to the *forward* filtration $\{\mathscr{F}_t\}$. Therefore our BSPDE is an extension of the backward SDE initiated by Bismut (1973), later developed by Pardoux and Peng (1990); and is different from those presented by many other authors (e.g., Rozovskii, 1990; Krylov and Rozovskii, 1982; Kunita, 1990; Pardoux, 1979, etc.) in which no such adaptedness of the solutions was required. The BSPDEs of this kind were originally found useful in stochastic control theory. The works concerning the existence, uniqueness and regularity of the adapted solution to a BSPDE, mostly in the context of stochastic maximum principle for optimal control problems involving SPDEs (Zakai equation in particular), can be found in, for example, Bensoussan (1983, 1992), Hu and Peng (1991), Peng (1992) and Zhou (1992, 1993). However, most of the existing results essentially required the so-called *superparabolic* condition: to wit, the matrix (A^{ij}) in (1.3) is uniformly positive definite. We note that in Zhou (1992, 1993) the degenerate cases were discussed, but the results there required that the operator $\mathscr{M}$ to be bounded, which is unfortunately not the case we are interested in this paper, due to the special form of the BSPDE arising in our applications. As a matter of fact, it is this application (to be described in the next paragraph) that motivated the present paper; moreover, to our best knowledge, the existence, uniqueness and regularity of the adapted solution of a degenerate BSPDE with unbounded $\mathscr{M}$ has remained open so far.

The second aim of the paper is to apply the result of BSPDE to the study of a class of forward–backward SDEs (hereafter FBSDEs) with random coefficients, which has recently been found useful, apart from stochastic control theory, in mathematical finance. In the case when the coefficients of an FBSDE are deterministic, it was

proved in our previous work (Ma et al., 1994) that, under certain conditions, the FB-SDE has a unique *adapted* solution over an arbitrarily prescribed time duration; and more importantly, the backward and forward components of such adapted solutions can be related explicitly via a *classical* solution to a backward quasilinear parabolic PDE. Some earlier results concerning the solvability of an FBSDE can be found in Antonelli (1993) and Ma and Yong (1995), using different methods; and the applications of FBSDEs, especially in mathematical finance, can be found in Duffie et al. (1995) and Cvitanic and Ma (1996). However, when the coefficients and the terminal value g are allowed to be random, the problem becomes much more formidable. It turns out that in this case the corresponding PDE has to be replaced by a backward *quasilinear* SPDE similar to (1.1) with a strong degeneracy: i.e., the matrix $(A^{ij}) \equiv 0$, and the operator $\mathcal{M}$ is unbounded. We note that the existence and uniqueness of adapted solutions to a class of FBSDEs with random coefficients was studied recently by Hu and Peng (1997) and Peng and Wu (1996), under certain "monotonicity" conditions on the coefficients. In this paper we do not pursue the general solvability of such FBSDEs, instead we content ourselves with some simpler cases in which the structure of the adapted solution can be clearly seen, based on our results in BSPDEs. We establish in a special case the relation between an FBSDE and a BSPDE, which leads to an *adapted* version of stochastic Feynman–Kac formula, and later a generalized option pricing formula. The derivation of these formulas depends heavily on the method we have been using, namely, the Four-Step Scheme as presented in Ma et al. (1994).

Finally, we would like to point out that the main difficulty in deriving the satisfactory existence, uniqueness and regularity result for a degenerate BSPDE seems to be that the *best* a priori estimate of Krylov–Rozovskii (cf. Krylov and Rozovskii, 1982, or Zhou, 1993) for a degenerate second first-order differential operator, on which the existing method is heavily based, is not strong enough to guarantee the desired convergence of the finite-dimensional approximating sequences (see Zhou, 1993 for more discussion on this issue). In this paper, however, we shall derive an a priori estimate for the BSPDE directly without using the Krylov–Rozovskii estimate, which enables us to approximate the degenerate BSPDE by a sequence of nondegenerate ones for which the existence of adapted solutions is known; and derive the existence, uniqueness and regularity of the adapted solution to a degenerate BSPDE in a different way. For technical reasons, in this paper we deal only with the case when the coefficients (a^{ij}) and $(b^{k\ell})$ are independent of x. And we shall, hopefully, address the more general cases in our future publications.

This paper is organized as follows. In Section 2 we give some preliminaries. Sections 3–5 are devoted to the proof of the existence and uniqueness of the adapted solution to the degenerate BSPDE (1.1). In Sections 6 and 7 we discuss the relation between the BSPDE (1.1) and a class of FBSDEs. We relate their solutions explicitly through a set of formulas, including those involving Malliavin calculus; and as a corollary we derive a stochastic Feynman–Kac formula. We note that the relations involving the component q of the adapted solution to BSPDE (1.1) are new. Finally, in Section 8 we apply our result to an option pricing problem in mathematical finance, and derive (for the first time) a stochastic Black–Scholes formula.

2. Notation and preliminaries

Throughout this paper we assume that the time duration $[0, T]$ is fixed; and that $(\Omega, \mathscr{F}, P)$ is a complete probability space on which is defined a d-dimensional standard Brownian motion $W = \{W_t: t \in [0, T]\}$. We further assume that the filtration $\{\mathscr{F}_t\}_{t \geqslant 0}$ is generated by W, augmented by all the P-null set in $\mathscr{F}$, and thus satisfies the "usual hypotheses" (Protter, 1990).

For any integer $m \geqslant 0$, we denote by $C^m(\mathbb{R}^n; \mathbb{R}^{\ell})$ the set of functions from $\mathbb{R}^n$ to $\mathbb{R}^{\ell}$ that are continuously differentiable up to order m; by $C_b^m(\mathbb{R}^n; \mathbb{R}^{\ell})$ the set of those functions in $C^m(\mathbb{R}^d; \mathbb{R}^{d'})$ whose partial derivatives up to order m are uniformly bounded. If there is no danger of confusion, $C^m(\mathbb{R}^n; \mathbb{R}^{\ell})$ and $C_b^m(\mathbb{R}^n; \mathbb{R}^{\ell})$ will be abbreviated as C^m and C_b^m, respectively. We denote the inner product in an Euclidean space E by $\langle \cdot, \cdot \rangle$; and the norm in E by $|\cdot|$. With the notation $\partial_{x_i} = \partial/\partial_{x_i}$, $i = 1, \ldots, n$ and $\nabla = (\partial_{x_1}, \ldots, \partial_{x_n})^{\mathrm{T}}$, we shall denote, for any $\varphi \in C^1(\mathbb{R}^n)$, $\nabla\varphi = (\partial_{x_1}\varphi, \ldots, \partial_{x_n}\varphi)^{\mathrm{T}}$. If $\xi = (\xi^1, \ldots, \xi^n)$ is a vector field such that each $\xi^i \in C^1(\mathbb{R}^n)$, then we denote by $\nabla\xi$ the matrix $(\partial_{x_i}\xi^j)_{i,j}$; and by $\nabla \cdot \xi \overset{\triangle}{=} \sum_{i=1}^n \partial_{x_i}\xi^i$ the divergence of ξ.

For any multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, we denote $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ and $D^{\alpha} \overset{\triangle}{=} \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_n}^{\alpha_n}$. If $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ is another multi-index, then by $\beta \leqslant \alpha$ we mean that $\beta_i \leqslant \alpha_i$ for $1 \leqslant i \leqslant n$; and by $\beta < \alpha$ we mean that $\beta \leqslant \alpha$ and $|\beta| < |\alpha|$. For $m \geqslant 0$, we denote by $W^{m,p}(\mathbb{R}^n)$ the usual Sobolev space, and $H^m(\mathbb{R}^n) = W^{m,2}(\mathbb{R}^n)$. In the case of no ambiguity, we often write $H^m, W^{m,p}$, etc., instead of $H^m(\mathbb{R}^n)$, $W^{m,p}(\mathbb{R}^n)$, etc. for simplicity.

Finally, for any $1 \leqslant p$, $r \leqslant \infty$, any Banach space X, and any sub-σ-field $\mathscr{G} \subseteq \mathscr{F}$, we denote

- by $L_{\mathscr{G}}^p(\Omega, X)$ the set of all X-valued, $\mathscr{G}$-measurable random variable φ such that $E\|\varphi\|_X^p < \infty$; and we simply write $L^p(\Omega; X) = L_{\mathscr{F}}^p(\Omega; X)$;
- by $L_{\mathscr{F}}^p(0, T; L^r(\Omega; X))$ the set of all $\{\mathscr{F}_t\}$-predictable X-valued processes $\varphi(t, \cdot, \omega)$: $[0, T] \times \Omega \to X$ such that $\|\varphi\|_{L_{\mathscr{F}}^p(0,T;L^r(\Omega;X))} \overset{\triangle}{=} \{\int_0^T [E\|\varphi(t)\|_X^r]^{p/r} \, dt\}^{1/p} < \infty$.

Similarly, we can define the space $C_{\mathscr{F}}([0, T]; L^r(\Omega; X))$, etc.. If $p = r$, we shall denote $L_{\mathscr{F}}^p(0, T; L^p(\Omega; X))$ by $L_{\mathscr{F}}^p(0, T; X)$ for simplicity. In particular, if $X = \mathbb{R}^n$ and $p = r = 2$, we denote $L_{\mathscr{F}}^2(0, T)^n = L_{\mathscr{F}}^2(0, T; \mathbb{R}^n) = L_{\mathscr{F}}^2(0, T; L^2(\Omega; \mathbb{R}^n))$, which is the set of all $\{\mathscr{F}_t\}$-adapted square integrable processes taking values in $\mathbb{R}^n$.

To conclude this section, we recall a useful fact. Let $h \in L_{\mathscr{F}}^2(0, T; C_b^m(\mathbb{R}^n; \mathbb{R}^d))$. Then it can be shown (see, for example, Kunita, 1990, Exercise 3.1.5) that the stochastic integral with parameter: $\int_0^t \langle h(s, x, \cdot), dW_s \rangle$ has a modification that belongs to $L_{\mathscr{F}}^2(0, T; C^{m-1})$ and it satisfies

$$D^{\alpha} \int_0^t \langle h(s, x, \cdot), dW_s \rangle = \int_0^t \langle D^{\alpha} h(s, x, \cdot), dW_s \rangle \quad \text{for } |\alpha| = 1, 2, \ldots, m-1. \tag{2.1}$$

Consequently, if $h \in L_{\mathscr{F}}^2(0, T; C_b^{\infty})$, then $\int_0^{\cdot} \langle h(s, \cdot, \cdot), dW_s \rangle \in L_{\mathscr{F}}^2(0, T; C^{\infty})$; and (2.1) holds for all multi-index α.

3. A linear degenerate BSPDE

The main result of this paper concerns the following BSPDE: for $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$u(t, x, \cdot) = g(x, \cdot) + \int_t^T \{(\mathscr{L}u)(s, x, \cdot) + (\mathscr{M}q)(s, x, \cdot) + f(s, x, \cdot)\} \, \mathrm{d}s$$

$$- \int_t^T \langle q(s, x, \cdot), \mathrm{d}W_s \rangle, \tag{3.1}$$

where the random differential operators $\mathscr{L}$ and $\mathscr{M}$ are given as (in vector forms):

$$(\mathscr{L}u)(t, x, \omega) = \nabla \cdot (A(t, \omega)\nabla u(t, x, \omega)) + \langle a(t, x, \omega), \nabla u(t, x, \omega) \rangle$$

$$+ a_0(t, x, \omega)u(t, x, \omega), \tag{3.2}$$

$$(\mathscr{M}q)(t, x, \omega) = \mathrm{tr}\{B(t, \omega)^T \nabla q(t, x, \omega)\} + \langle b(t, x, \omega), q(t, x, \omega) \rangle;$$

and

$$A = (a^{ij}) : [0, T] \times \Omega \to \mathbb{R}^{n \times n},$$

$$B = (b^{k\ell}) : [0, T] \times \Omega \to \mathbb{R}^{n \times d},$$

$$a = (a^1, \dots, a^n) : [0, T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n,$$

$$b = (b^1, \dots, b^d) : [0, T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^d, \tag{3.3}$$

$$a_0, \ f : [0, T] \times \mathbb{R}^n \times \Omega \to \mathbb{R},$$

$$g : \mathbb{R}^n \times \Omega \to \mathbb{R}.$$

In the sequel, the dependence of all the functions on ω will be suppressed for the simplicity of notation. We shall make the following assumptions: for an integer $m \geqslant 0$.

(A1)$_m$ functions A, a, a_0, B, b, f and g satisfy: $A \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^{n \times n})$; $B \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^{n \times d})$; $a \in L^2_{\mathscr{F}}(0, T; C_b^m(\mathbb{R}^n; \mathbb{R}^n))$; $b \in L^2_{\mathscr{F}}(0, T; C_b^m(\mathbb{R}^n; \mathbb{R}^d))$; $a_0 \in L^2_{\mathscr{F}}(0, T; C_b^m(\mathbb{R}^n))$; $f \in L^2_{\mathscr{F}}(0, \ T; C_b^m(\mathbb{R}^n) \cap H^m(\mathbb{R}^n))$; $g \in L^2_{\mathscr{F}_T}(\Omega; C_b^m(\mathbb{R}^n) \cap H^m(\mathbb{R}^n))$. Further, the partial derivatives of a, a_0, b, f and g in x up to order m are bounded uniformly in (t, ω). by a constant $K > 0$;

(A2) it holds that $2A(t) - B(t)B(t)^T \geqslant 0, \forall t \in [0, T]$ a.s.

We remark here that the assumption (A2) allows the degeneracy of the operator $\mathscr{L}$ in the sense of stochastic PDEs. It is such a degeneracy and our intention of finding the *adapted* solutions to be defined below that distinguish the BSPDE (3.1) from the existing ones in the literature, as we pointed out in Section 1.

Definition 3.1. (i) A pair of random fields $\{(u(t, x; \omega), q(t, x; \omega)), \ (t, x, \omega) \in [0, T] \times \mathbb{R}^n \times \Omega\}$ is called a *classical solution* of (3.1), if $u \in C_{\mathscr{F}}([0, T]; L^2(\Omega; C^2(\mathbb{R}^n)))$ and $q \in L^2_{\mathscr{F}}(0, T; C^1(\mathbb{R}^n; \mathbb{R}^d))$, such that (3.1) is satisfied for all $(t, x) \in [0, T] \times \mathbb{R}^n$, almost surely.

(ii) A pair of random fields $\{(u(t, x; \omega), q(t, x; \omega)), \ (t, x, \omega) \in [0, T] \times \mathbb{R}^n \times \Omega\}$ is called a *strong solution* of (3.1), if $u \in C_{\mathscr{F}}([0, T]; L^2(\Omega; H^2(\mathbb{R}^n)))$ and $q \in L^2_{\mathscr{F}}(0, T; H^1(\mathbb{R}^n; \mathbb{R}^d))$, such that (3.1) holds for all $t \in [0, T]$, a.e. $x \in \mathbb{R}^n$, almost surely.

(iii) A pair of random fields $\{(u(t,x;\omega),q(t,x;\omega)),\ (t,x,\omega)\in[0,T]\times\mathbb{R}^n\times\Omega\}$ is called a *weak solution* of (3.1), if $u\in C_{\mathcal{F}}([0,T];L^2(\Omega;H^1(\mathbb{R}^n)))$ and $q\in L^2_{\mathcal{F}}(0,T;L^2(\mathbb{R}^n;\mathbb{R}^d))$, such that $\forall\varphi\in H^1(\mathbb{R}^n)$, $\forall t\in[0,T]$, it holds almost surely that ($Q_t\overset{\triangle}{=}[t,T]\times\mathbb{R}^n$)

$$\int_{\mathbb{R}^n}u(t,x)\varphi(x)\,\mathrm{d}x$$

$$=\int_{\mathbb{R}^n}g(x)\varphi(x)\,\mathrm{d}x+\int_{Q_t}\{-\langle A(s)\nabla u(s,x),\nabla\varphi(x)\rangle$$

$$+\langle a(s,x),\nabla u(s,x)\rangle\varphi(x)+a_0(s,x)u(s,x)\varphi(s,x)-\langle\nabla\varphi(x),B(s)q(s,x)\rangle$$

$$+\langle b(s,x),q(s,x)\rangle\varphi(x)+f(s,x)\varphi(x)\}\,\mathrm{d}x\,\mathrm{d}s-\int_t^T\left\langle\int_{\mathbb{R}^n}q(s,x)\varphi(x)\,\mathrm{d}x,\mathrm{d}W_s\right\rangle.$$

$$(3.4)$$

The main theorem of this paper is stated below. The proof of the theorem will be carried out in the following two sections.

Theorem 3.2. *Let* $(A1)_m$ *and* (A2) *hold with* $m\geqslant1$. *Then,* (3.1) *admits a unique weak solution* (u,q) *with*

$$\{2A(\cdot)-B(\cdot)B(\cdot)^T\}^{1/2}\nabla u(\cdot,\cdot)\in L^2_{\mathcal{F}}(0,T;H^m(\mathbb{R}^n;\mathbb{R}^n)),$$
$$q(\cdot,\cdot)+B(\cdot)^T\nabla u(\cdot,\cdot)\in L^2_{\mathcal{F}}(0,T;H^m(\mathbb{R}^n;\mathbb{R}^d)),\tag{3.5}$$

and the following estimate holds:

$$\max_{t\in[0,T]}E\|u(t,\cdot)\|^2_{H^m}+E\int_0^T\|q(t,\cdot)\|^2_{H^{m-1}}\,\mathrm{d}t$$

$$+E\int_{[0,T]\times\mathbb{R}^d}\left\{\sum_{|\alpha|\leqslant m}|D^\alpha q(t,x)+B(t)^T\nabla(D^\alpha u(t,x))-b(t,x)D^\alpha u(t,x)|^2\right.$$

$$\left.+\sum_{|\alpha|\leqslant m}\langle(2A(t)-B(t)B(t)^T)\nabla(D^\alpha u(t,x)),\nabla(D^\alpha u(t,x))\rangle\right\}\,\mathrm{d}x\,\mathrm{d}t$$

$$\leqslant C\{\|f\|^2_{L^2_{\mathcal{F}}(0,T;H^m)}+\|g\|^2_{L^2_{\mathcal{F}_T}(\Omega;H^m)}\},\tag{3.6}$$

where the constant C *depends only on* m, T *and* K.

Furthermore, if $m\geqslant2$, *the weak solution* (u,q) *becomes the unique strong solution; and if* $m>2+d/2$, *then* (u,q) *is the unique classical solution.*

Remark 3.3. We note that in the case either $A(t)-B(t)B(t)^T\geqslant\delta I$, a.e. $t\in[0,T]$, for some $\delta>0$; or $A(t)\geqslant0$, $B(t)=0$, a.e. $t\in[0,T]$, one can easily show that (3.6) leads to similar estimates given in Zhou (1993).

Remark 3.4. The square root of the left-hand side of (3.6) is a norm, under which the set of all processes $(u,q)\in C_{\mathcal{F}}([0,T];H^m)\times L^2_{\mathcal{F}}(0,T;H^{m-1})$ with (3.5) being true is a Banach space. We will denote this space by $\mathcal{H}^m$.

4. A priori estimates

In order to prove Theorem 3.2, we first provide an *a priori* estimate for the solutions of (3.1). We begin by assuming that the coefficients a, a_0, b, f and g of the equation (3.1) are infinitely differentiable in x with all the partial derivatives up to order m being bounded by a constant $K_m > 0$ for all $m \geqslant 0$. We also assume that (u, q) is an adapted solution of (3.1) such that

$$
\begin{aligned}
&u \in C_{\mathscr{F}}([0, T]; L^2(\Omega; C^\infty(\mathbb{R}^n) \cap H^{m+1}(\mathbb{R}^n))), \\
&q \in L^2_{\mathscr{F}}(0, T; C^\infty(\mathbb{R}^n, \mathbb{R}^d) \cap H^m(\mathbb{R}^n; \mathbb{R}^d)).
\end{aligned}
\tag{4.1}
$$

We remark here that a random field u (resp. q) satisfying (4.1) can be roughly described as one that is continuous (resp. square-integrable) in t, square-integrable in ω, and infinitely differentiable in x, with all the partial derivatives up to order $m + 1$ (resp. m) being square-integrable on $\mathbb{R}^n$.

Let α be any multi-index, $|\alpha| \leqslant m$. In this case, by using the fact (2.1) we can apply the operator D^α to both sides of (3.1) to obtain (suppressing ω):

$$
\begin{aligned}
(D^\alpha u)(t, x) = (D^\alpha g)(x) + \int_t^T D^\alpha \{ (\mathscr{L}u)(s, x) + (\mathscr{M}q)(s, x) + f(s, x) \} \, ds \\
- \int_t^T \langle (D^\alpha q)(s, x), \, dW_s \rangle, \quad (t, x) \in [0, T] \times \mathbb{R}^n.
\end{aligned}
\tag{4.2}
$$

Next, for each $n \in \mathbb{N}$, let $\varphi_n \in C^1(\mathbb{R})$ such that

$$
\begin{aligned}
&0 \leqslant \varphi_n(\rho) \leqslant 2\rho \wedge n, \\
&\varphi_n(\rho) \uparrow 2\rho, \quad \varphi'_n(\rho) \to 2, \quad \forall \rho \in \mathbb{R}, \; n \to \infty.
\end{aligned}
$$

Define $\Phi_n(\rho) = \int_0^\rho \varphi_n(r) \, dr$. It is clear that Φ_n is C^2 and $\Phi_n(\rho) \uparrow \rho^2$, $\forall \rho \in \mathbb{R}$. Now, by Itô's formula we have, for each $n \in \mathbb{N}$, that

$$
\begin{aligned}
\Phi_n((D^\alpha g)(x)) &- \Phi_n((D^\alpha u)(t, x)) \\
&= -\int_t^T \varphi_n((D^\alpha u)(s, x)) D^\alpha \{ (\mathscr{L}u)(s, x) + (\mathscr{M}q)(s, x) + f(s, x) \} \, ds \\
&\quad + \frac{1}{2} \int_t^T \varphi'_n(D^\alpha u(s, x)) \, |D^\alpha q|^2(s, x) \, ds + \int_t^T \varphi_n((D^\alpha u)(s, x)) \langle D^\alpha q(s, x), \, dW_s \rangle.
\end{aligned}
\tag{4.3}
$$

Since φ_n is bounded, the last term above is a martingale. We can first take expectation, and integrate with respect to x over $\mathbb{R}^n$, then let $n \to \infty$ by using the Dominated

Monotone Convergence Theorem to obtain that (recall $Q_t = [t, T] \times \mathbb{R}^n$ and note (4.2))

$$\Theta \stackrel{\triangle}{=} E \int_{\mathbb{R}^n} \{(D^\alpha g)^2(x) - (D^\alpha u)^2(t, x)\} \, dx$$

$$= -2E \int_{Q_t} (D^\alpha u)(s, x) D^\alpha \{(\mathscr{L} u)(s, x) + (\mathscr{M} q)(s, x) + f(s, x)\} \, dx \, ds$$

$$+ E \int_{Q_t} |D^\alpha q|^2(s, x) \, dx \, ds. \tag{4.4}$$

We shall now do a detailed analysis of the right-hand side of (4.4). To begin with, let us denote by I the $d \times d$ identity matrix, and define

$$\Lambda = \begin{pmatrix} I & B^{\mathrm{T}} & -b \\ B & 2A & -Bb \\ -b^{\mathrm{T}} & -b^{\mathrm{T}} B^{\mathrm{T}} & 1 + |b|^2 \end{pmatrix}, \quad \eta = \begin{pmatrix} D^\alpha q \\ \nabla(D^\alpha u) \\ (D^\alpha u) \end{pmatrix}. \tag{4.5}$$

Then a direct computation shows that

$$\eta^{\mathrm{T}} \Lambda \eta = |D^\alpha q|^2 + \langle 2A \nabla(D^\alpha u), \nabla(D^\alpha u) \rangle + (D^\alpha u)^2(1 + |b|^2)$$

$$+ 2\langle B(D^\alpha q), \nabla(D^\alpha u) \rangle - 2(D^\alpha u)\langle b, D^\alpha q \rangle - 2(D^\alpha u)\langle Bb, \nabla(D^\alpha u) \rangle. \tag{4.6}$$

On the other hand, we have

$$\eta^{\mathrm{T}} \Lambda \eta = \eta^{\mathrm{T}} \begin{pmatrix} I & 0 & 0 \\ B & I & 0 \\ -b^{\mathrm{T}} & 0 & 1 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & 2A - BB^{\mathrm{T}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I & B^{\mathrm{T}} & -b \\ 0 & I & 0 \\ 0 & 0 & 1 \end{pmatrix} \eta$$

$$= |D^\alpha q + B^{\mathrm{T}} \nabla(D^\alpha u) - (D^\alpha u)b|^2 + \langle (2A - BB^{\mathrm{T}}) \nabla(D^\alpha u), \nabla(D^\alpha u) \rangle + (D^\alpha u)^2. \tag{4.7}$$

Therefore, recalling the definition of the operators $\mathscr{L}$ and $\mathscr{M}$ (see (3.2)), and using standard techniques of integration by parts (note that all the coefficients together with their partial derivatives in x up to order m are bounded and (4.1) holds), we see that (4.4) can be written as follows (suppressing all variables):

$$\Theta = E \int_{Q_t} \{|D^\alpha q|^2 - 2(D^\alpha u)[\nabla \cdot \{A \nabla(D^\alpha u)\} + \mathrm{tr}\{B \nabla(D^\alpha q)\}$$

$$+ D^\alpha(\langle a, \nabla u \rangle) + D^\alpha(a_0 u) + D^\alpha \langle b, q \rangle + D^\alpha f]\} \, dx \, ds$$

$$= E \int_{Q_t} \left\{ |D^\alpha q|^2 + \langle 2A \nabla(D^\alpha u), \nabla(D^\alpha u) \rangle + 2\langle B(D^\alpha q), \nabla(D^\alpha u) \rangle \right.$$

$$- 2(D^\alpha u)\langle a, \nabla(D^\alpha u) \rangle - 2a_0(D^\alpha u)^2 - 2(D^\alpha u)\langle b, D^\alpha q \rangle$$

$$- 2 \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta}(D^\alpha u)\{\langle D^{\alpha-\beta} a, \nabla(D^\beta u) \rangle + (D^{\alpha-\beta} a_0)(D^\beta u)$$

$$\left. + \langle D^{\alpha-\beta} b, D^\beta q \rangle\} - 2(D^\alpha f)(D^\alpha u) \right\} \, dx \, ds, \tag{4.8}$$

where $C_{\alpha\beta} > 0$ are some constants depending only on α and β. Comparing the right-hand side of (4.6) with the integrand on the right-hand side of (4.8), and noting that

$$E \int_{Q_t} (D^\alpha u) \langle Bb - a, \nabla(D^\alpha u) \rangle \, dx \, ds = \frac{1}{2} E \int_{Q_t} \langle Bb - a, \nabla[(D^\alpha u)^2] \rangle \, dx \, ds$$

$$= -\frac{1}{2} E \int_{Q_t} \nabla \cdot (Bb - a)(D^\alpha u)^2 \, dx \, ds,$$

we see that (4.8) (hence (4.4)) becomes

$$E \int_{\mathbb{R}^n} \{ (D^\alpha g)^2(x) - (D^\alpha u)^2(t, x) \} \, dx$$

$$= E \int_{Q_t} \left\{ \eta^\mathsf{T} \Lambda \eta - \nabla \cdot (Bb - a)(D^\alpha u)^2 - (2a_0 + 1 + |b|^2)(D^\alpha u)^2 \right.$$

$$- 2 \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta} \{ \langle D^{\alpha-\beta}a, \nabla D^\beta u \rangle (D^\alpha u) + (D^{\alpha-\beta}a_0)(D^\beta u)(D^\alpha u)$$

$$\left. + \langle D^{\alpha-\beta}b, D^\beta q \rangle (D^\alpha u) \} - 2(D^\alpha f)(D^\alpha u) \right\} \, dx \, ds. \tag{4.9}$$

We now replace the quadratic form $\eta^\mathsf{T} \Lambda \eta$ by (4.6), then from (4.9) we obtain that

$$E \int_{\mathbb{R}^n} (D^\alpha u(t, x))^2 \, dx + E \int_{Q_t} \{ |D^\alpha q + B^\mathsf{T} \nabla(D^\alpha u) - (D^\alpha u)b|^2$$

$$+ \langle (2A - BB^\mathsf{T}) \nabla(D^\alpha u), \nabla(D^\alpha u) \rangle \} \, dx \, ds$$

$$= E \int_{\mathbb{R}^n} (D^\alpha u(t, x))^2 \, dx + E \int_{Q_t} [\eta^\mathsf{T} \Lambda \eta - (D^\alpha u)^2] \, dx \, ds$$

$$= E \int_{\mathbb{R}^n} (D^\alpha g)^2 \, dx + E \int_{Q_t} \left\{ (|b|^2 + 2a_0 + \nabla \cdot (Bb - a))(D^\alpha u)^2 \right.$$

$$+ 2 \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta} \{ \langle D^{\alpha-\beta}a, \nabla D^\beta u \rangle (D^\alpha u) + (D^{\alpha-\beta}a_0)(D^\beta u)(D^\alpha u)$$

$$\left. + \langle D^{\alpha-\beta}b, D^\beta q \rangle (D^\alpha u) \} + 2(D^\alpha f)(D^\alpha u) \right\} \, dx \, ds. \tag{4.10}$$

With the help of (4.10), we can now prove the main result of this section.

Proposition 4.1. *Let $m \in \mathbb{N}$. Suppose that the coefficients a, a_0 and b_0 are infinitely differentiable in x, and their partial derivatives in x up to order m are uniformly bounded by an absolute constant $K > 0$. Suppose also that (u, q) is a classical solution of (3.1), such that both u and q are C^∞ in x and*

$$\max_{t \in [0, T]} E \| u(t, \cdot) \|_{H^{m+1}}^2 + E \int_0^T \| q(t, \cdot) \|_{H^m}^2 \, dt < \infty. \tag{4.11}$$

Then the estimate (3.6) holds.

Proof. First, by (4.11), we see that (4.10) is true for all multi-indices α with $|\alpha| \leqslant m$ (if $|\alpha| = 0$, the term $\sum_{0 \leqslant \beta < \alpha} \cdots$ does not show up). We now add up (4.10) for all $|\alpha| \leqslant m$, and recalling the norms of H^m and $L^2_{\mathscr{F}}(0, T; H^m)$, we obtain

$$
E\|u(t, \cdot)\|^2_{H^m} + E \int_{Q_t} \left\{ \sum_{|\alpha| \leqslant m} |D^\alpha q + B^{\mathrm{T}} \nabla(D^\alpha u) - (D^\alpha u)b|^2 \right.
$$

$$
\left. + \sum_{|\alpha| \leqslant m} \langle (2A - BB^{\mathrm{T}}) \nabla(D^\alpha u), \nabla(D^\alpha u) \rangle \right\} \mathrm{d}x \, \mathrm{d}s
$$

$$
\leqslant E\|g\|^2_{H^m} + E \int_{Q_t} \left\{ \left(|b|^2 + 2|a_0| + |\nabla \cdot (Bb - a)| \right) \sum_{|\alpha| \leqslant m} (D^\alpha u)^2 \right.
$$

$$
+ 2 \sum_{|\alpha| \leqslant m} \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta} \{ |D^{\alpha-\beta}a||\nabla D^\beta u||D^\alpha u| + |D^{\alpha-\beta}a_0||D^\beta u||D^\alpha u|
$$

$$
\left. + |D^{\alpha-\beta}b||D^\beta q||D^\alpha u| \} + 2 \sum_{|\alpha| \leqslant m} |D^\alpha f||D^\alpha u| \right\} \mathrm{d}x \, \mathrm{d}s.
$$

$$
\leqslant E\|g\|^2_{H^m} + E \int_{Q_t} \left\{ (|b|^2 + 2|a_0| + |\nabla \cdot (Bb - a)| + 1) \sum_{|\alpha| \leqslant m} (D^\alpha u)^2 \right.
$$

$$
+ \left(\sum_{|\alpha| \leqslant m} \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta} \left\{ 2|D^{\alpha-\beta}a| + |D^{\alpha-\beta}a_0| + \frac{1}{\varepsilon}|D^{\alpha-\beta}b|^2 \right\} \right) \left(\sum_{|\alpha| \leqslant m} |D^\alpha u| \right)^2
$$

$$
\left. + \varepsilon \sum_{|\alpha| \leqslant m} \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta} |D^\beta q|^2 + \sum_{|\alpha| \leqslant m} |D^\alpha f|^2 \right\} \mathrm{d}x \, \mathrm{d}s. \tag{4.12}
$$

Here we have used the inequality $2ab \leqslant \frac{1}{\varepsilon}|a|^2 + \varepsilon|b|^2$. Now recall that all the coefficients a, a_0, B and b as well as their partial derivatives in x up to order m are uniformly bounded by a constant $K > 0$, and note that there exists a constant $C_1 > 0$, depending only on m, such that

$$
\sum_{|\alpha| \leqslant m} \sum_{0 \leqslant \beta < \alpha} C_{\alpha\beta} |D^\beta q|^2 \leqslant C_1 \sum_{|\beta| \leqslant m-1} |D^\beta q|^2;
$$

$$
\left(\sum_{|\alpha| \leqslant m} |D^\alpha u| \right)^2 \leqslant C_1 \sum_{|\alpha| \leqslant m} |D^\alpha u|^2.
$$

Thus, there exists a constant $C(\varepsilon) > 0$, such that the right-hand side of (4.12) is dominated by

$$E\|g\|_{H^m}^2 + C(\varepsilon)E\int_t^T \|u(s,\cdot)\|_{H^m}^2 \, \mathrm{d}s$$

$$+ \varepsilon C_1 E \int_{Q_t} \sum_{|\beta| \leq m-1} |D^\beta q|^2 \, \mathrm{d}x \, \mathrm{d}s + \|f\|_{L_{\mathscr{F}}^2(0,T;H^m)}^2. \tag{4.13}$$

Now note that for $|\beta| < m$, one has

$$E\int_{Q_t} |D^\beta q|^2 \, \mathrm{d}x \, \mathrm{d}s \leq 3E \int_{Q_t} \{|b|^2 |D^\beta u|^2 + |B|^2 |\nabla(D^\beta u)|^2$$

$$+ |D^\beta q + B^{\mathrm{T}} \nabla(D^\beta u) - (D^\beta u)b|^2\} \, \mathrm{d}x \, \mathrm{d}s. \tag{4.14}$$

Summing up both sides of (4.14) with respect to $|\beta| \leq m - 1$, we have

$$E\int_{Q_t} \sum_{|\beta| \leq m-1} |D^\beta q|^2 \, \mathrm{d}x \, \mathrm{d}s \leq 6K^2 E \int_t^T \|u(s,\cdot)\|_{H^m}^2 \, \mathrm{d}s$$

$$+ 3E \int_{Q_t} \sum_{|\alpha| \leq m} |D^\alpha q + B^{\mathrm{T}}(\nabla D^\alpha u) - (D^\alpha u)b|^2 \, \mathrm{d}x \, \mathrm{d}s. \tag{4.15}$$

Plugging (4.15) into (4.13), we see that the right-hand side of (4.12) is now dominated by

$$E\|g\|_{H^m}^2 + (C(\varepsilon) + 6\varepsilon C_1 K^2)E \int_t^T \|u(s,\cdot)\|_{H^m}^2 \, \mathrm{d}s + \|f\|_{L_{\mathscr{F}}^2(0,T;H^m)}$$

$$+ 3\varepsilon C_1 \sum_{|\alpha| \leq m} E \int_{Q_t} |D^\alpha q + B^{\mathrm{T}}(\nabla D^\alpha u) - (D^\alpha u)b|^2 \, \mathrm{d}x \, \mathrm{d}s. \tag{4.16}$$

Now let us replace the right-hand side of (4.12) by (4.16) with the choice that $\varepsilon = 1/6C_1$, and $C_2 = C(1/6C_1) + K^2$, then after a slight rearrangement of terms we obtain from (4.12) that

$$E\|u(t,\cdot)\|_{H^m}^2 + \frac{1}{2} \sum_{|\alpha| \leq m} E \int_{Q_t} \{|D^\alpha q + B^{\mathrm{T}} \nabla(D^\alpha u) - (D^\alpha u)b|^2$$

$$+ \langle (2A - BB^{\mathrm{T}})\nabla(D^\alpha u), \nabla(D^\alpha u)\rangle\} \, \mathrm{d}x \, \mathrm{d}s$$

$$\leq E\|g\|_{H^m}^2 + C_2 E \int_t^T \|u(s,\cdot)\|_{H^m}^2 \, \mathrm{d}s + \|f\|_{L_{\mathscr{F}}^2(0,T;H^m)}^2. \tag{4.17}$$

Applying Gronwall's inequality, one can easily derive, with a suitable choice of the constant C which depends only on m, T and K, that

$$\max_{0 \leqslant t \leqslant T} E\|u(t, \cdot)\|_{H^m}^2 + E \int_{[0,T] \times \mathbb{R}^n} \left\{ \sum_{|\alpha| \leqslant m} |D^\alpha q + B^{\mathrm{T}} \nabla(D^\alpha u) - (D^\alpha u)b|^2 \right.$$

$$\left. + \sum_{|\alpha| \leqslant m} \langle (2A - BB^{\mathrm{T}}) \nabla(D^\alpha u), \nabla(D^\alpha u) \rangle \right\} \mathrm{d}x \, \mathrm{d}s$$

$$\leqslant C \{ E\|g\|_{H^m}^2 + \|f\|_{L_{\mathscr{F}}^2(0,T;H^m)}^2 \}. \tag{4.18}$$

Finally, note that by (4.15) we have

$$E \int_0^T \|q(t, \cdot)\|_{H^{m-1}}^2 \, \mathrm{d}t \leqslant 6K^2 T \max_{t \in [0,T]} E\|u(t, \cdot)\|_{H^m}^2$$

$$+ 3E \int_{[0,T] \times \mathbb{R}^n} \sum_{|\alpha| \leqslant m} |D^\alpha q + B^{\mathrm{T}} \nabla(D^\alpha u) - (D^\alpha u)b|^2 \, \mathrm{d}x \, \mathrm{d}t. \tag{4.19}$$

By (4.18) we see that both terms on the right-hand side above can be dominated by $C'\{E\|g\|_{H^m}^2 + \|f\|_{L_{\mathscr{F}}^2(0,T;H^m)}^2\}$, with some absolute constant $C' > 0$. Combining this with (4.18), we derive (3.6). $\square$

5. Proof of Theorem 3.2

Let us first assume that the coefficients a, a_0, b, g and f are all infinitely differentiable in x. For any $\varepsilon > 0$, we consider the following approximate problem:

$$du^\varepsilon(t, x) = -\{(\mathscr{L}_\varepsilon u^\varepsilon)(t, x) + (\mathscr{M} q^\varepsilon)(t, x) + f(t, x)\} \, \mathrm{d}t + \langle q^\varepsilon(t, x) \, \mathrm{d}W_t \rangle,$$

$$(t, x) \in [0, T] \times \mathbb{R}^n, \tag{5.1}$$

$$u^\varepsilon(T, x) = g(x), \quad x \in \mathbb{R}^n,$$

where

$$(\mathscr{L}_\varepsilon u)(t, x) = \nabla \cdot \{(A(t) + \varepsilon I) \nabla u(t, x)\} + \langle a(t, x), \nabla u(t, x) \rangle + a_0(t, x)u(t, x). \tag{5.2}$$

We note that (5.1) is a nondegenerate BSPDE. Then, using the result of (Zhou, 1992), we know that there exists a unique classical solution $(u^\varepsilon, q^\varepsilon)$ to (5.1), such that both u^ε and q^ε are infinitely differentiable in x, thanks to the Sobolev Imbedding Theorem; and (4.11) holds for any $m \geqslant 1$. Therefore, we can apply Proposition 4.1 to get

$$\max_{t \in [0,T]} E\|u^\varepsilon(t, \cdot)\|_{H^m}^2 + E \int_0^T \|q^\varepsilon(t, \cdot)\|_{H^{m-1}}^2 \, \mathrm{d}t$$

$$+ E \int_{[0,T] \times \mathbb{R}^n} \left\{ \sum_{|\alpha| \leqslant m} |D^\alpha q^\varepsilon + B^{\mathrm{T}} \nabla(D^\alpha u^\varepsilon) - (D^\alpha u^\varepsilon)b|^2 \right.$$

$$+ \sum_{|\alpha| \leqslant m} \langle (2A + 2\varepsilon I - BB^{\mathrm{T}})\nabla(D^{\alpha}u^{\varepsilon}), \nabla(D^{\alpha}u^{\varepsilon}) \rangle \Bigg\} \, \mathrm{d}x \, \mathrm{d}t$$

$$\leqslant C\{ \|f\|^2_{L^2_{\mathcal{F}}(0,T;H^m)} + \|g\|^2_{L^2_{\mathcal{F}_t}(\Omega;H^m)}\}, \tag{5.3}$$

with the constant C depending only on m, T and K (independent of $\varepsilon > 0$). Now, for any $\varepsilon, \varepsilon' > 0$, we have

$$d(u^{\varepsilon} - u^{\varepsilon'})(t,x) = -\{\mathcal{L}(u^{\varepsilon} - u^{\varepsilon'})(t,x) + \mathcal{M}(q^{\varepsilon} - q^{\varepsilon'})(t,x)$$
$$+ \varepsilon \Delta u^{\varepsilon}(t,x) - \varepsilon' \Delta u^{\varepsilon'}(t,x)\} \, \mathrm{d}t + \langle q^{\varepsilon}(t,x) - q^{\varepsilon'}(t,x) \, \mathrm{d}W(t)\rangle,$$
$$(t,x) \in [0,T] \times \mathbb{R}^n,$$
$$(u^{\varepsilon} - u^{\varepsilon'})(T,x) = 0, \quad x \in \mathbb{R}^n, \tag{5.4}$$

where Δ is the Laplacian operator. Let $m \geqslant 3$ and apply Proposition 4.1 again. Noting that for any multi-index $\alpha, \langle (2A - BB^{\mathrm{T}})\nabla(D^{\alpha}u), \nabla(D^{\alpha}u)\rangle \geqslant 0, \ \forall(t,x)$, a.s., we obtain that

$$\max_{t \in [0,T]} E\|(u^{\varepsilon} - u^{\varepsilon'})(t,\cdot)\|^2_{H^{m-2}(\mathbb{R}^d)} + E \int_0^T \|(q^{\varepsilon} - q^{\varepsilon'})(t,\cdot)\|^2_{H^{m-3}(\mathbb{R}^n)} \, \mathrm{d}t$$

$$\leqslant C(\varepsilon\|u^{\varepsilon}\|^2_{L^2_{\mathcal{F}}(0,T;H^m)} + \varepsilon'\|u^{\varepsilon'}\|^2_{L^2_{\mathcal{F}}(0,T;H^m)})$$

$$\leqslant C(\varepsilon + \varepsilon')\{\|f\|^2_{L^2_{\mathcal{F}}(0,T;H^m)} + \|g\|^2_{L^2_{\mathcal{F}_T}(\Omega;H^m)}\}. \tag{5.5}$$

Therefore, by letting $m \geqslant 4$, we can find a random field (u,q) with $u \in L^2_{\mathcal{F}}(0,T;H^{m-2})$ and $q \in L^2_{\mathcal{F}}(0,T;H^{m-3})$, such that

$$\lim_{\varepsilon \to 0} \left\{ \max_{t \in [0,T]} \|u^{\varepsilon} - u\|^2_{L^2_{\mathcal{F}}(0,T;H^{m-2})} + E \int_0^T \|q^{\varepsilon}(t,\cdot) - q(t,\cdot)\|^2_{H^{m-3}} \, \mathrm{d}t \right\} = 0.$$

It is easily seen, by passing to the limits in (5.1), that (u,q) is an *adapted strong solution* to (3.1) satisfying (3.6). The uniqueness follows easily from the estimate (3.6).

In the general case when only Assumptions (A1)$_m$ and (A2) are satisfied, we adopt the standard technique of "smoothing coefficients" (see, for example, Rozovskii, 1990). Namely, we first apply an "averaging operator" on both sides of (3.1) so that the coefficients become infinitely differentiable, and therefore obtain the corresponding approximating solutions which are all infinitely differentiable. Then, together with the *a priori* estimates derived in Section 4, we can show that the approximating solution converges to the strong solution (or weak solution) of (3.1), provided the Assumptions (A1)$_m$ and (A2) are in force. Also, estimate (3.6) holds true since we are taking the limits in the space $\mathcal{H}^m$ (see Remark 3.4 for the definition of $\mathcal{H}^m$). Finally, the classical solution can be obtained by a simple application of the well-known Sobolev imbedding theorem. For the detailed arguments, one is referred to, e.g., Rozovskii (1990). This completes the proof of Theorem 3.2. □

Discussion. Recall that in (A1)$_m$ we have assumed both g and f are bounded with bounded derivatives. This is sometimes too restrictive in applications, and we would

like to consider the following relaxation: suppose that there exists some $\psi \in C^\infty$, $\psi(x) > 0$, $\forall x \in \mathbb{R}$, such that

$$|g(x)|, |f(t,x)| \leqslant C\psi(x), \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n. \tag{5.6}$$

Define

$$v(t,x) = \psi(x)^{-1}u(t,x), \qquad p(t,x) = \psi(x)^{-1}q(t,x). \tag{5.7}$$

Then, multiplying Eq. (3.1) by $\psi(x)^{-1}$, and noting that $\nabla u = \psi\nabla v + (\nabla\psi)v$, $\nabla q = \psi\nabla p + (\nabla\psi)p$, we obtain from some simple computation that

$$v(t,x) = \widehat{g}(x) + \int_t^T \{\nabla \cdot (A\nabla v) + \langle \widehat{a}, \nabla v \rangle + \widehat{a}_0 v + \mathrm{tr}(B\nabla p) + \widehat{b}p + \widehat{f}\}\,\mathrm{d}s$$
$$- \int_t^T p\,\mathrm{d}W_s, \tag{5.8}$$

where

$$\widehat{g}(x) = \frac{g(x)}{\psi(x)}, \qquad \widehat{f}(t,x) = \frac{f(t,x)}{\psi(x)},$$
$$\widehat{a}(t,x) = a(t,x) + \frac{2A(t)\nabla\psi(x)}{\psi(x)},$$
$$\widehat{a}(t,x) = a_0(t,x) + \frac{\nabla \cdot (A(t)\nabla\psi(x)) + \langle a(t,x), \nabla\psi(x)\rangle}{\psi(x)}, \tag{5.9}$$
$$\widehat{b}(t,x) = b(t,x) + \frac{B(t)\nabla\psi(x)}{\psi(x)}.$$

Therefore, if the new coefficients $\widehat{a}, \widehat{a}_0, \widehat{b}, \widehat{g}$ and $\widehat{f}$ satisfy conditions $(A1)_m$ and $(A2)$, BSPDE (5.8) will have a unique strong solution (v, p), which in turn shows that (3.1) has a unique strong solution (u, q) such that (u, q) and (v, p) are related by (5.7). Thus, we have proved the following theorem that slightly generalizes Theorem 3.2.

Theorem 5.1. *Suppose that there exists a function $\psi \in C^\infty$ such that the corresponding coefficients $\widehat{a}, \widehat{a}_0, \widehat{b}, \widehat{g}$ and $\widehat{f}$ defined by (5.9) satisfy conditions $(A1)_m$ and $(A2)$ for some $m \geqslant 1$. Then, (3.1) admits a unique weak solution (u, q) with (3.5) being true. Furthermore, the weak solution (u, q) becomes the unique strong solution if $m \geqslant 2$; and it becomes the unique classical solution if $m > 2 + d/2$. Finally, the estimate (3.6) holds when the functions u, q, f and g are replaced by $\psi^{-1}u, \psi^{-1}q, \widehat{f}$ and $\widehat{g}$.* $\square$

To conclude this section we point out that if $(A1)_m$ holds for A, a, a_0, B and b except for f and g, (5.6) holds with $\Psi(\cdot) = |\nabla\psi(\cdot)|/\psi(\cdot) \in C_b^m(\mathbb{R}^n)$, then, $(A1)_m$ holds for $\widehat{a}, \widehat{a}_0, \widehat{b}, \widehat{f}$ and $\widehat{g}$. Some obvious examples for functions ψ with desired properties are, e.g.,

$$\psi(x) = 1 + |x|^{2m} \quad \text{or} \quad \psi(x) = e^{K\sqrt{1+|x|^2}}, \ x \in \mathbb{R}^n.$$

6. Relations between BSPDEs and FBSDEs

In this section we discuss how to use the adapted solution of a BSPDE to obtain the adapted solution of a forward–backward SDE (FBSDE) with random coefficients, based on the *Four-Step Scheme* designed in Ma et al. (1994). We shall start from a general discussion that will at least reveal our motivation; and then we content ourselves with a special case for which the results in the previous sections apply.

Let us consider an FBSDE with random coefficients:

$$
\begin{aligned}
X_t &= x + \int_0^t \bar{b}(s, X_s, Y_s, Z_s, \cdot)\,\mathrm{d}s + \int_0^t \sigma(s, X_s, Y_s, \cdot)\,\mathrm{d}W_s, \\
Y_t &= g(X_T, \cdot) + \int_t^T \widehat{b}(s, X_s, Y_s, Z_s, \cdot)\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s.
\end{aligned}
\tag{6.1}
$$

We assume that the coefficients $\bar{b}$, $\widehat{b}$ and σ are random fields defined on $[0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \times \Omega$ such that for fixed (x, y, z), the processes $\bar{b}(\cdot, x, y, z, \cdot)$, $\widehat{b}(\cdot, x, y, z, \cdot)$ and $\sigma(\cdot, x, y, \cdot)$ are predictable. Also, we assume for the moment that for fixed t and ω, the functions b, $\widehat{b}$ and σ are smooth in x, y and z. Finally, we assume that for fixed x, $g(x, \cdot)$ is an $\mathcal{F}_T$- measurable random variable, and it is smooth in x for fixed ω.

Our purpose is to find an "adapted solution" (X, Y, Z) to Eq. (6.1). In light of our previous work (Ma et al., 1994), we proceed by the following heuristic argument. Supposing that (X, Y, Z) is an adapted solution to (6.1), we denote $\widetilde{b}(t) = \bar{b}(t, X_t, Y_t, Z_t)$ and $\widetilde{\sigma}(t) = \sigma(t, X_t, Y_t)$. Suppose that there is a random field $\{u(t, x, \omega) : (t, x, \omega) \in [0, T] \times \mathbb{R}^n \times \Omega\}$, which takes the form (suppressing ω):

$$
u(t, x) = u(x, 0) + \int_0^t p(s, x)\,\mathrm{d}s + \int_0^t \langle q(t, x), \mathrm{d}W_s \rangle,
$$

where $p \in L^2_{\mathscr{F}}(0, T; C^k)$ and $q \in L^2_{\mathscr{F}}(0, T; C^{\ell})$ with $k \geqslant 3$, $\ell \geqslant 2$, such that Y and X are related as $Y_t = u(t, X_t, \cdot)$, $\forall t \in [0, T]$, a.s. Then, by applying a generalized Itô-Ventzell formula (see [Kunita, 1990, Theorem 3.3.1]) from t to T, together with some computation using the definition of the stochastic integral based on semimartingales with spatial parameter x, one shows that

$$
\begin{aligned}
u(t, X_t) = u(T, X_T) &- \int_t^T \Big\{ p(s, X_s) + \tfrac{1}{2}\mathrm{tr}\{\widetilde{\sigma}\widetilde{\sigma}^{\mathrm{T}}(s)(D^2 u)(s, X_s)\} \\
&+ \langle \widetilde{b}(s), \nabla u(s, X_s) \rangle + \mathrm{tr}\{\widetilde{\sigma}^{\mathrm{T}}(s)\nabla q(s, X_s)\} \Big\}\,\mathrm{d}s \\
&- \int_t^T \langle q(s, X_s) + \widetilde{\sigma}^{\mathrm{T}}(s)\nabla u(s, X_s)\,\mathrm{d}W_s \rangle,
\end{aligned}
\tag{6.2}
$$

where $(D^2 u)(t, x)$ denotes the matrix $(u_{x_i x_j}(t, x))$. Now noting that $Y_t = u(t, X_t, \cdot)$, we can compare (6.2) with (6.1) to obtain that

$$
\begin{aligned}
Z_t &= q(t, X_t) + \sigma^{\mathrm{T}}(t, X_t, u(t, X_t))\nabla u(t, X_t); \\
p(t, X_t) &= -\Big\{\widehat{b}(t, X_t, u(t, X_t), Z_t) + \tfrac{1}{2}\mathrm{tr}\{\sigma\sigma^{\mathrm{T}}(t, X_t, u(t, X_t))(D^2 u)(t, X_t)\} \\
&\quad + \langle \bar{b}(t, X_t, u(t, X_t), Z_t), \nabla u(t, X_t) \rangle + \mathrm{tr}\{\sigma^{\mathrm{T}}(t, X_t, u(t, X_t))\nabla q(t, X_t)\}\Big\}, \\
u(T, X_T) &= g(X_T).
\end{aligned}
\tag{6.3}
$$

Combining the above, we have the following modified Four Step Scheme:

Step 1: Define a function $z(t, x, y, w, q) = q + \sigma^T(t, x, y)w$.

Step 2: Using the function $z(\cdots)$ defined above, solve the quasilinear BSPDE:

$$
\begin{aligned}
u(t,x) = g(x) + \int_t^T &\{\widehat{b}(s, x, u(s, x), z(s, x, u(s, x), \nabla u(s, x), q(s, x))) \\
&+ \tfrac{1}{2}\mathrm{tr}\{\sigma\sigma^T(s, x, u(s, x))(D^2 u)(s, x)\} \\
&+ \langle \overline{b}(s, x, u(s, x), z(s, x, u(s, x), \nabla u(t, x), q(t, x))), \nabla u(s, x)\rangle \\
&+ \mathrm{tr}\{\sigma^T(s, x, u(s, x))\nabla q(s, x)\}\}\mathrm{d}s - \int_t^T \langle q(s, x)\,\mathrm{d}W_s\rangle,
\end{aligned} \tag{6.4}
$$

and denote the (adapted) solution by (u, q).

Step 3: Using the solution (u, q) from above, define

$$
\widetilde{b}(t, x, \omega) = \overline{b}(t, x, u(t, x, \omega), z(t, x, u(t, x, \omega), \nabla u(t, x, \omega), q(t, x, \omega)), \omega);
$$

$$
\widetilde{\sigma}(t, x, \omega) = \sigma(t, x, u(t, x, \omega), \omega).
$$

Then we solve the forward SDE:

$$
X_t = x + \int_0^t \widetilde{b}(s, X_s, \cdot)\mathrm{d}s + \int_0^t \widetilde{\sigma}(s, X_s, \cdot)\mathrm{d}W_s.
$$

Step 4: Define $Y_t = u(t, X_t, \cdot)$ and $Z_t = q(t, X_t, \cdot) + \sigma^T(t, X_t, u(t, X_t, \cdot), \cdot)\nabla u(t, X_t, \cdot)$, then one shows that (X, Y, Z) is the adaptd solution to (6.1).

It is clear that the most difficult part in the Four-Step Scheme is Step 2, in which we have to find a (classical) adapted solution to a quasilinear BSPDE. Notice that in this case we have (using the corresponding notations in the previous sections)

$$
A(t, x, \omega) = \tfrac{1}{2}\sigma\sigma^T(t, x, u(t, x, \omega), \omega); \qquad B(t, x, \omega) = \sigma(t, x, u(t, x, \omega), \omega),
$$

hence $2A(t, x, \omega) - B(t, x, \omega)B^T(t, x, \omega) \equiv 0$. Namely, the BSPDE is *degenerate*; and the coefficient in front of ∇q is $\sigma^T \neq 0$, which means that the operator $\mathcal{M}$ in the BSPDE is *unbounded*. Thus the problem becomes very difficult in general. Also, we should point out that although the solvability of BSPDEs might imply the solvability of the FBSDE, it does not guarantee the uniqueness of the adapted solution of the latter. We refer the interested readers to Duffie et al. (1995) and Ma et al. (1994) for the issue of a "nodal solution" of an FBSDE (a solution that can be obtained by Four-Step Scheme). The following is a simplified case where the Four-Step Scheme does go through, thanks to the results of previous sections. We note that our attention is not the well-posedness of the FBSDE, but rather the relation between the adapted solutions of the two equations.

A Special Case. Suppose that in the FBSDE (6.1) we have

$$
\overline{b}(t, x, y, z, \omega) = \overline{b}(t, x, \omega),
$$

$$
\sigma(t, x, y, \omega) = \sigma(t, \omega),
$$

$$
\widehat{b}(t, x, y, z, \omega) = \widehat{b}_1(t, x, \omega)y + \widehat{b}_2(t, x, \omega)z + \widehat{b}_3(t, x, \omega).
$$

In other words, we consider the decoupled FBSDE with random coefficients and terminal condition:

$$X_t = x + \int_0^t \overline{b}(s, X_s)\, ds + \int_0^t \sigma(s)\, dW_s;$$
$$Y_t = g(X_T) + \int_t^T [\widehat{b}_1(s, X_s)Y_s + \widehat{b}_2(s, X_s)Z_s + \widehat{b}_3(s, X_s)]\, ds - \int_t^T \langle Z_s, dW_s \rangle. \tag{6.5}$$

Suppose for the moment that $n = d = 1$. Then following the Four-Step Scheme described above, we see that by setting $z(t, w, q, \omega) = q + w\sigma(t, \omega)$, the BSPDE (6.4) becomes (suppresssing variables in the integrals):

$$u(t, x) = g(x) + \int_t^T \{ \tfrac{1}{2}\sigma^2 u_{xx} + \overline{b}u_x + \sigma q_x + u\widehat{b}_1 + (q + u_x\sigma)\widehat{b}_2 + \widehat{b}_3 \}\, ds - \int_t^T q\, dW_s$$

$$= g(x) + \int_t^T \{ \tfrac{1}{2}\sigma^2 u_{xx} + (\overline{b} + \sigma\widehat{b}_2)u_x + \widehat{b}_1 u + \sigma q_x + \widehat{b}_2 q + \widehat{b}_3 \}\, ds - \int_t^T q\, dW_s \tag{6.6}$$

Now, if we define the operators:

$$(\mathscr{L}\varphi)(t, x) = \tfrac{1}{2}\sigma^2(t)\varphi_{xx} + (\overline{b}(t, x) + \sigma(t)\widehat{b}_2(t, x))\varphi_x + \widehat{b}_1(t, x)\varphi,$$

$$(\mathscr{M}\psi)(t, x) = \sigma(t)\psi_x + \widehat{b}_2(t, x)\psi,$$

and $f(t, x) = \widehat{b}_3(t, x)$, then (6.6) becomes

$$u(t, x) = g(x) + \int_t^T [(\mathscr{L}u)(s, x) + (\mathscr{M}q)(s, x) + f(s, x)]\, ds - \int_t^T q(s, x)\, dW(s), \tag{6.7}$$

which is exactly the BSPDE (3.1) with $n = d = 1$. Returning to the higher-dimensional case, we see that the (6.6) is the same as (3.1) with

$$A(t) = \tfrac{1}{2}\sigma(t)\sigma(t)^{\mathrm{T}}; \quad B(t) = \sigma(t); \quad a(t, x) = \overline{b}(t, x) + \sigma(t)\widehat{b}_2(t, x)$$

$$a_0(t, x) = \widehat{b}_1(t, x); \quad b(t, x) = \widehat{b}_2(t, x); \quad f(t, x) = \widehat{b}_3(t, x). \tag{6.8}$$

Thus, we have the following theorem.

Theorem 6.1. *Suppose that the random functions (fields) $\overline{b}$, σ, $\widehat{b}_1$, $\widehat{b}_2$, $\widehat{b}_3$ and g are such that the corresponding functions A, B, a, a_0, b, f, g defined by (6.8) satisfy either the conditions of Theorem 3.2 or that of Theorem 5.1 with $m \geqslant 2 + d/2$. Then the FBSDE (6.5) has a unique adapted solution (X, Y, Z) such that*

$$Y_t = u(t, X_t, \cdot); \qquad Z_t = q(t, X_t, \cdot) + \sigma^{\mathrm{T}}(t)\nabla u(t, X_t, \cdot) \quad \text{a.e. } t \in [0, T] \text{ a.s.} \tag{6.9}$$

where (u, q) is the unique adapted classical solution to the BSPDE (6.6).

In particular, if all the coefficients in (6.5) are deterministic, then the unique solution of (6.6) is deterministic. More precisely, it is given by $(u, 0)$, where u is the classical solution to a backward PDE, derived from (6.6) by setting $q \equiv 0$.

Proof. Note that since the FBSDE (6.5) is decoupled, it is well known (see Pardoux and Peng (1990)) that it possesses a unique adapted solution. Next, by Theorem 3.2 (or Theorem 5.1), we see that under our assumption, the BSPDE (6.6) has a unique classical solution, therefore the Four-Step Scheme goes through (note that in this case Step 3 is trivial). Consequently, (6.9) must hold by the uniqueness of the FBSDE and the construction of the Four-Step Scheme.

To see the second assertion, we note that by setting $q \equiv 0$, the BSPDE (6.6) becomes a backward PDE, which, under our assumptions on the coefficients, possesses a unique classical solution u since Theorem 3.2 holds regardless of the data being deterministic or random. Thus by the uniqueness of the BSPDE, the pair $(u, 0)$ must be the only solution to the BSPDE. The theorem is proved. □

We should point out here that the real issue in Theorem 6.1 is the relation between the solutions (X, Y, Z) and (u, q). Such a relation, especially the one between Y and X, will lead to a stochastic Feynman–Kac formula, as stated in the following corollary. Let us now consider the FBSDE (6.5) on the interval $[s, T]$ with $X_s = x$, a.s. for some $0 \leqslant s \leqslant T$. Denote the solution to such an FBSDE by $(X^{s,x}, Y^{s,x}, Z^{s,x})$.

Corollary 6.2. (Stochastic Feynman–Kac Formula). *Suppose that the conditions of Theorem 6.1 are in force, with $\widehat{b}_2 \equiv 0$. Let (u, q) be the classical solution to the corresponding BSPDE (6.6). Then, for any $(s, x) \in [0, T] \times \Omega$ and a.e. $\omega \in \Omega$, $u(s, x, \omega)$ has the following representation*:

$$u(s, x, \omega) = u(s, X_s^{s,x}, \omega) = Y_s^{s,x}$$
$$= E \left\{ e^{\int_s^T \widehat{b}_1(r, X_r^{s,x}) \, dr} g(X_T^{s,x}) + \int_s^T e^{\int_s^r \widehat{b}_1(u, X_u^{s,x}) \, du} \widehat{b}_3(r, X_r^{s,x}) \, dr \,\Big|\, \mathscr{F}_s \right\}. \quad (6.10)$$

Proof. Since in this case the backward SDE in (6.5) is linear in Y and the drift is independent of Z, thus solution $Y_t^{s,x}$ can be written as, for $s \leqslant t \leqslant T$,

$$Y_t^{s,x} = e^{\int_t^T \widehat{b}_1(r, X_r^{s,x}) \, dr} g(X_T^{s,x}) + \int_t^T e^{\int_t^r \widehat{b}_1(u, X_u^{s,x}) \, du} [\widehat{b}_3(r, X_r^{s,x}) \, dr - Z_r^{s,x} \, dW_r],$$

by taking the conditional expectation on both sides above, noting the adaptedness of the solution $(X^{s,x}, Y^{s,x}, Z^{s,x})$ and setting $s = t$, we see that (6.10) follows from (6.9). □

Remark 6.3. Except for our restriction on the coefficient σ, Corollary 6.2 extends the classical Feynman–Kac formula and (partially) the one in Pardoux and Peng (1992) in which all the coefficients were assumed to be deterministic. We note that the Feynman–Kac formulas involving backward stochastic integrals or Malliavin calculus/Skorohod integrals were studied by many authors, e.g., Krylov and Rozovskii (1982), Pardoux (1979), and Ocone and Pardoux (1993). But with the *adaptedness* requirement, our version is different.

7. A look via Malliavin calculus

The relation between an FBSDE and a BSPDE can be further studied by using Malliavin calculus. In this section, we shall establish such a relation for (6.5) and (6.6), based on the results of Pardoux and Peng (1992) and El Karoui, Peng and Quenez (1997). In order to keep the paper to a proper length, we refer the reader to the recent book of Nualart (1995) for all undefined notation concerning derivation on Wiener Space, and here we give only some less standard ones which will be useful in the discussion. To simplify notation, throughout this section we assume $n = d = 1$, the higher-dimensional cases can be obtained analogously.

Let $D : L^2(\Omega) \to L^2([0, T] \times \Omega)$ be the Malliavin derivation operator, and $\mathbb{D}_{1,p}$, $p \geqslant 2$ be the set of all $\xi \in L^2(\Omega)$ such that

$$\|\xi\|_{1,p} = \|\xi\|_{L^p(\Omega)} + \|\|D\xi\|_{L^2([0,T])}\|_{L^p(\Omega)} < \infty. \tag{7.1}$$

For any ℓ-dimensional random vector η, and any random field $f : [0, T] \times \mathbb{R}^\ell \times \Omega$, we shall distinguish $D.(f(t, \eta(\omega), \omega))$ from $D.f(t, x, \omega)|_{x=\eta(\omega)}$ by use of the notation (see also Ocone and Pardoux (1993): $[D.f](t, \eta(\omega), \omega) := D.f(t, x, \omega)|_{x=\eta(\omega)}$. Further, for $p \geqslant 2$, we define $\mathbb{L}_{1,p} = L^p(0, T; \mathbb{D}_{1,p})$; and we denote $\mathbf{L}_{1,p}^a$ to be the set of all elements $u \in \mathbb{L}_{1,p}$, such that $u.(\cdot)$ is progressively measurable, and

$$\|u\|_{\mathbf{L}_{1,p}^a}^p \overset{\triangle}{=} E\left\{ \int_0^T |u_t|^p \, \mathrm{d}t + \int_0^T \int_0^T |D_s u_t|^p \, \mathrm{d}s \, \mathrm{d}t \right\} < \infty. \tag{7.2}$$

Finally, we give a definition that is an adaptation of Definition 2.1 in Ocone and Pardoux (1993).

Definition 7.1. We say that a random field $\varphi = \{\varphi(x, \omega): x \in \mathbb{R}, \omega \in \Omega\}$ satisfies hypothesis (B) with moment p, if φ is a measurable function such that
(1) $\varphi(\cdot, \omega) \in C^1(\mathbb{R})$ almost surely, and there exists a constant $K > 0$ such that

$$|\nabla \varphi(x, \omega)| \leqslant K, \quad \forall (x, \omega) \in \mathbb{R}^d \times \Omega;$$

(2) $\varphi(x, \cdot) \in \mathbb{D}_{1,p}$ for every x; and the map $(s, x, \omega) \mapsto D_s g(x, \omega)$ admits a measurable version that is continuous in x for a.e. (s, ω). Further, there exists a process $c \in L^p([0, T] \times \Omega)$ and a constant $\beta > 0$, such that for all x and a.s. $\omega \in \Omega$, it holds that

$$|D_s \varphi(x, \omega)| \leqslant c(s, \omega)(1 + |x|^\beta), \quad \forall x \in \mathbb{R} \text{ a.e. } s \in [0, T] \text{ a.s. } \omega \in \Omega.$$

Let us now consider again the FBSDE (6.5) and BSPDE (6.6) with $n = d = 1$. Besides the assumption of Theorem 6.1 on the coefficients $\overline{b}, \sigma, \widehat{b}_1 - \widehat{b}_3$ and g, we further assume that they satisfy the following:

Assumption 7.2. (i) For fixed $x \in \mathbb{R}$, $\overline{b}(\cdot, x, \cdot)$, $\sigma(\cdot, \cdot)$, $\widehat{b}_1(\cdot, x, \cdot)$, $\widehat{b}_2(\cdot, x, \cdot)$, $\widehat{b}_3(\cdot, x, \cdot)$ $\in \mathbf{L}_{1,8}^a$.

(ii) The functions $\overline{b}, \widehat{b}_1, \widehat{b}_2$, and $\widehat{b}_3$ and their first order partial derivatives in x are uniformly bounded and Lipschitz continuous in x, uniform in (t, ω), with a common constant $K > 0$. Further, $\overline{b}$ is C^2 in x, and there exists a version of $D.\widehat{b}$ that is C^1 in x.

(iii) There exist a random field $\{c_1(s,t,\omega)\}$ satisfying $E\int_0^T\int_0^T|c_1(s,t,\omega)|^{p_1}\,\mathrm{d}s\,\mathrm{d}t<\infty$, and a process $c_2\in L^{p_1}([0,T]\times\Omega)$ for some $p_1>4$, such that for all s,t,x,y and a.s. $\omega\in\Omega$, it holds that

$$|D_2\overline{b}(t,0,\omega)|+|\nabla D_s\overline{b}(t,x,\omega)|+|D_s\sigma(t,0,\omega)|+|D_s\widehat{b}_1(t,x,\omega)|\leqslant c_1(s,t,\omega);$$

$$|D_s\widehat{b}_2(t,x,\omega)|+|D_s\widehat{b}_3(t,x,\omega)|\leqslant c_2(s,\omega).$$

(iv) The function $g(x,\cdot)$ satisfies the hypothesis (B) with $p>4$ and $\beta=1$.

Theorem 7.3. *Suppose that the assumptions of Theorem 6.1 and Assumptions 7.2 hold. Let (X,Y,Z) be the adapted solution to the FBSDE (6.5) and (u,q) be the adapted classical solution to (6.6). Then the process $u(\cdot,X,\cdot)\in\mathbb{D}_{1,2}$; and for $t\in[0,T]$, it holds that*

$$D_t u(t,X_t,\cdot)=D_t Y_t=Z_t=q(t,X_t,\cdot)+\sigma(t,\cdot)\nabla u(t,X_t,\cdot)\quad a.s. \tag{7.3}$$

Moreover, if for each $t\in[0,T]$, $u(t,\cdot,\cdot)$ satisfies hypothesis (B) with $p>2$ (see Definition 7.1), then one has

$$q(t,X_t,\cdot)=[D_t u](t,X_t,\cdot),\quad\forall t\in[0,T]\ a.s. \tag{7.4}$$

Remark 7.4. We note that when the function g and the coefficients $\overline{b},\sigma,\widehat{b}_1$ and $\widehat{b}_2$ are all deterministic, the relation $Z_t=D_t Y_t$ was proved by Pardoux and Peng (1992). El Karoui et al. (1997) generalized the result to the random coefficient case, but essentially assumed that $\widehat{b}_1$ and $\widehat{b}_2$ are independent of X. Theorem 7.3 is a further generalization of the latter, with the drift coefficient of the backward SDE being linear in y and z; and the connection with solutions of BSPDE is new.

Proof of Theorem 7.3. Since the proof is merely technical, and many estimates are more or less standard in the context of Malliavin calculus, we shall give only a sketch without going through all the details. First, it is clear that the first and last equalities in (7.3) are the direct consequence of Theorem 6.1, therefore we need only show that $D_t Y_t=Z_t$, $\forall t\in[0,T]$. Note that if X is the solution to the forward SDE in (6.5), then by the uniform Lipschitz property of the function $\overline{b}$ and the boundness of σ, and following the proof of Lemma 3.16 in Ocone and Pardoux (1993), one shows that

$$X\in\mathbf{L}_{1,4+\varepsilon}^a\cap L^p(\Omega;C[0,T]), \tag{7.5}$$

for all $p>1$ and all $\varepsilon\geqslant0$ such that $4+\varepsilon<p_1$, where $p_1>0$ is the constant appeared in Assumption 7.2(iii). Further, for fixed $s\in[0,T]$, the process D_sX. satisfies a linear SDE:

$$D_sX_t=1_{\{s\leqslant t\}}\left\{\sigma(s)+\int_s^t\{\nabla\overline{b}(r,X_r,\cdot)D_sX_r+[D_s\overline{b}](r,X_r,\cdot)\}\,\mathrm{d}r\right.$$

$$\left.+\int_s^t D_s\sigma(r,\cdot)\,\mathrm{d}W_r\right\},\quad 0\leqslant s,\ t\leqslant T. \tag{7.6}$$

Thus, by standard arguments one derives from (7.6) that, for any $2 \leqslant p' < p_1 \wedge 8$,

$$E \left\{ \sup_{t \in [0,T]} |D_s X_t|^{p'} \right\} \leqslant CE \left\{ |\sigma(s)|^{p'} + \int_0^T (|D_s \sigma(r)|^{p'} + |[D_s \bar{b}](r, X_r)|^{p'}) \, dr \right\}.$$

(7.7)

Since $\mathbf{L}_{1,8}^a \subset \mathbf{L}_{1,p'}^a$, using Assumption 7.2(i)–(iii), Definition (7.2), and Hölder's inequality, it can be shown that the right side of (7.7) is finite. Consequently, we have

$$E \int_0^T \sup_{t \in [0,T]} |D_s X_t|^{p'} \, ds < \infty \quad \text{for all } 2 \leqslant p' < p_1 \wedge 8.$$

(7.8)

Now, let us define a random variable $\xi = g(X_T, \cdot)$, and a random field

$$f(t, \omega, y, z) = \hat{b}_1(t, X_t(\omega), \omega)y + \hat{b}_2(t, X_t(\omega), \omega)z + \hat{b}_3(t, X_t(\omega), \omega),$$

for $(t, y, z, \omega) \in [0, T] \times \mathbb{R}^2 \times \Omega$. Then by Assumption 7.2(ii)–(iv) and (7.5),

$$E|\xi|^p = E|g(X_T, \cdot)|^p \leqslant K^p E(1 + |X_T|)^p < \infty, \quad \forall p > 1;$$

(7.9)

and

$$E \left\{ \int_0^T |f(t, 0, 0)|^2 \, dt \right\}^2 = E \left\{ \int_0^T |\hat{b}_3(t, X_t, \cdot)|^2 \, dt \right\}^2 < \infty.$$

Thus by Theorem 5.1 in El Karoui et al. (1997), the unique solution (Y, Z) of the backward SDE in (6.5) satisfies

$$E \left\{ \sup_{0 \leqslant t \leqslant T} |Y_t|^p \right\} + E \left\{ \int_0^T |Z_t|^2 \, dt \right\}^{p/2} < \infty, \quad \forall p > 1.$$

(7.10)

Further, by (7.5) and (7.8), $X_T \in \mathbb{D}_{1,4+\varepsilon}$ for any $\varepsilon \geqslant 0$ such that $4 + \varepsilon < p_1$ (recall Definition (7.1)), and it has finite moments of all orders. Applying Lemma 2.2 in Ocone and Pardoux (1993) again we obtain that $g(X_T(\cdot), \cdot) \in \mathbb{D}^{1,k}$ for any $2 \leqslant k < p_1 \wedge p$ with $p > 4$, and

$$D_s \xi = D_s(g(X_T, \cdot)) = \nabla g(X_T, \cdot) D_s X_T + [D_s g](X_T, \cdot).$$

(7.11)

Now, using hypothesis (B)–(2), and Hölder's inequality, one shows that both terms on the right side of (7.11) belong to $L^4(\Omega \times [0, T])$. Consequently, one has

$$E \int_0^T |D_s \xi|^4 \, ds < \infty.$$

(7.12)

Finally, following a similar argument using Assumption 7.2(ii) and (iii), and applying Lemma 2.2 of Ocone and Pardoux (1993) to get that for each $s, t \in [0, T]$,

$$[D_s f](t, Y_t, Z_t) = D_s(\hat{b}_1(t, X_t)y + \hat{b}_2(t, X_t)z)|_{y = Y_t, z = Z_t}$$

$$= \nabla \hat{b}_1(t, X_t)(D_s X_t)Y_t + [D_s \hat{b}_1](t, X_t)Y_t + \nabla \hat{b}_2(t, X_t)(D_s X_t)Z_t + [D_s \hat{b}_2](t, X_t)Z_t$$

$$+ \nabla \hat{b}_3(t, X_t)(D_s X_t) + [D_s \hat{b}_3](t, X_t).$$

Again, analysing right-hand side above term by term, using Assumption 7.2(ii), Hölder's inequality (7.10), and the fact that $X \in \mathbf{L}_{4+\varepsilon}^q$, $\forall 4 + \varepsilon < p_1$, one shows that

$$E \int_0^T \left\{ \int_0^T |[D_s f](t, Y_t, Z_t)|^2 \, \mathrm{d}t \right\}^2 \mathrm{d}s < \infty. \tag{7.13}$$

Now by using (7.12) and (7.13), together with Assumption 7.2, we can apply Proposition 5.4 in El Karoui et al. (1997) to obtain that $Y_t \in \mathbb{D}_{1,2}$ such that $D_t Y_t = Z_t$, $\forall t \in [0, T]$, a.s., whence (7.3).

To see the second part of the theorem, we note that if u satisfies hypothesis (B), then by Lemma 2.2 of Ocone and Pardoux (1993) and (7.6), we have

$$D_t u(t, X_t) = \nabla u(t, X_t) D_t X_t + [D_t u](t, X_t) = \nabla u(t, X_t) \sigma(t) + [D_t u](t, X_t).$$

Comparing this with (7.3), we have $q(t, X_t) = [D_t u](t, X_t)$. This completes the proof of the theorem. $\square$

8. An application in option pricing

In this section we apply our result to a problem in option pricing. For the detailed formulation of an option pricing problem (or the problem of hedging a contingent claim), we refer the reader to the book of Karatzas and Shreve (1988); or the expository paper of El Karoui et al. (1997) for the formulation involving backward SDEs. Here we give only a brief description.

Let us consider a financial market consisting of 1 bond and d stocks. For notational simplicity, we assume $d = 1$. Suppose that their prices per share at any time t are described by the following differential equations, respectively,

$$\begin{aligned}
\mathrm{d}P_t^0 &= r(t)P_t^0 \, \mathrm{d}t; \quad \text{(bond)}, \\
\mathrm{d}P_t &= P_t[b(t) \, \mathrm{d}t + \sigma(t) \, \mathrm{d}W_t], \quad \text{(stock)},
\end{aligned} \tag{8.1}$$

where r is the *interest rate*, b is the *appreciation rate*, and σ is the *volatility*. We assume that r, b and σ are bounded, progressively measurable processes and σ is bounded away from zero. Furthermore, if we denote by Y the *wealth process* and π the *portfolio* process, that is, the amount of money that the investor puts in the stock at time t (therefore what he puts in the bond is $Y_t - \pi_t$), then the process Y (with no consumption) satisfies the following SDE:

$$\mathrm{d}Y_t = [Y_t r(t) + \pi_t(b(t) - r(t))] \, \mathrm{d}t + \pi_t \sigma(t) \, \mathrm{d}W_t. \tag{8.2}$$

Finally, a *contingent claim* is by definition some given $\mathscr{F}_T$-measurable random variable ξ satisfying some integrability conditions. We shall assume that ξ is of the form $\xi(\omega) = g(P_T(\omega), \omega)$, where $g : \mathbb{R} \times \Omega \to \mathbb{R}$ is such that it is jointly measurable; and for each fixed $x, g(x, \cdot)$ is $\mathscr{F}_T$-measurable. Note that if $g = g(x)$ is deterministic, then we have an *option*; while if $g \equiv g(\omega)$ is independent of x, then we have a general contingent claim. Moreover, a hedging price (or fair price) at $t = 0$ is defined to be the smallest initial endowment Y_0 such that there exists a portfolio process π for which the corresponding wealth process satisfies: $Y_t \geqslant 0$, $\forall t \in [0, T]$, a.s., and $Y_T \geqslant g(P_T, \cdot)$, a.s.

In the case where $r(\cdot) \equiv r$, $b(\cdot) \equiv b$, and $\sigma(\cdot) \equiv \sigma$ are all constants and g is deterministic ($g(x) = (x - q)^+$ for the European call option), the celebrated *Black–Scholes Option Valuation Formula* (see, for example, Karatzas and Shreve, 1988) tells us that the fair price of the option at any time $t \in [0, T]$ is given by

$$Y_t = \tilde{E}\{e^{-r(T-t)}g(P_T) \mid \mathscr{F}_t\}, \tag{8.3}$$

Here $\tilde{E}$ is the expectation with respect to some *risk-neutral* probability measure (or "equivalent martingale measure"). Furthermore, if we denote $v(t, x)$ to be the (classical) solution to the backward PDE:

$$\begin{aligned}
&v_t + \tfrac{1}{2}\sigma^2 x^2 v_{xx} + rx v_x - rv = 0, \quad (t, x) \in [0, T) \times (0, \infty); \\
&v(T, x) = g(x),
\end{aligned} \tag{8.4}$$

then it holds that $Y_t = v(t, P_t)$, $\forall t \in [0, T]$, a.s. Using the theory of backward SDE, it can be shown (see, El Karoui et al., 1997) that, if we denote (Y, Z) to be the unique adapted solution of the backward SDE:

$$Y_t = g(P_T) - \int_t^T [rY_s + \sigma^{-1}(b - r)Z_s]\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s, \tag{8.5}$$

then Y coincides with that in (8.3); and the optimal hedging strategy is given by $\pi_t = \sigma^{-1}Z_t = v_x(t, P_t)$.

We note that the option (or contingent claim) valuation formula (8.3) has been proved to be valid for more general cases in which the coefficients r, b, σ and the terminal condition g are allowed to be random (see, e.g., Karatzas and Shreve, 1988 and El Karoui et al., 1997), but as far as the "Black–Scholes PDE" (8.4) is concerned, no significant progress has been made when the coefficients are random, since in such a case a PDE is no longer appropriate to handle the situation. In the rest of this section we apply our results in the previous sections to derive a new result in this regard.

Let us consider the price equation (8.1) with random coefficients r, b, σ; and we consider the general terminal value g as described at the beginning of the section. We allow further that r and b may depend on the stock price in a nonanticipating way. In other words, we assume that $r(t, \omega) = r(t, P_t(\omega), \omega)$; $b(t, \omega) = b(t, P_t(\omega), \omega)$, where for each fixed $p \in \mathbb{R}$, $r(\cdot, p, \cdot)$ and $b(\cdot, p, \cdot)$ are predictable processes. Thus we can write (8.1) and (8.5) as an FBSDE:

$$\begin{aligned}
P_t &= p + \int_0^t P_s b(s, P_s)\,\mathrm{d}s + \int_0^t P_s \sigma(s)\,\mathrm{d}W_s, \\
Y_t &= g(P_T) - \int_t^T [Y_s r(s, P_s)) + Z_s \theta(s, P_s)]\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s,
\end{aligned} \tag{8.6}$$

where θ is the so-called *risk premium* process defined by

$$\theta(t, P_t) = \sigma^{-1}(t)[b(t, P_t) - r(t, P_t)], \quad \forall t \in [0, t]; \tag{8.7}$$

and $Z_t \overset{\triangle}{=} \pi_t \sigma(t)$. We do a slight transformation so that the results in the previous sections can be applied. First we note that the process P can be written as an stochastic

exponential:

$$P_t = p \exp \left\{ \int_0^t \left[b(s, P_s) - \frac{1}{2} \sigma^2(s) \right] \mathrm{d}s + \int_0^t \sigma(s) \, \mathrm{d}W(s) \right\}. \tag{8.8}$$

Thus $P_t > 0$, $\forall t \in [0, T]$, a.s., provided $p > 0$. (In the higher dimensional case, we can use the same arguments as those in (Cvitanic and Ma (1996) to show that every component is positive, a.s.). Therefore, we can define a *log-price* process X by $X_t = \log P_t$, $\forall t$, and by Itô's formula we have

$$
\begin{aligned}
X_t &= x + \int_0^t \overline{b}(s, X_s) \, \mathrm{d}s + \int_0^t \sigma(s) \, \mathrm{d}W_s, \\
Y_t &= \overline{g}(X_T) - \int_t^T [Y_s \overline{r}(s, X_s) + Z_s \overline{\theta}(s, X_s)] \, \mathrm{d}s - \int_t^T Z_s \, \mathrm{d}W_s,
\end{aligned}
\tag{8.9}
$$

where $x = \log p$, and

$$
\begin{aligned}
\overline{b}(t, x, \omega) &= b(t, \mathrm{e}^x, \omega) - \tfrac{1}{2}\sigma^2(t, \omega); & \overline{g}(x, \omega) &= g(\mathrm{e}^x, \omega); \\
\overline{r}(t, x, \omega) &= r(t, \mathrm{e}^x, \omega); & \overline{\theta}(t, x, \omega) &= \theta(t, \mathrm{e}^x, \omega);
\end{aligned}
\tag{8.10}
$$

Note that (8.9) is exactly the same as (6.5), with $\widehat{b}_1$ begin replaced by $-\overline{r}$; $\widehat{b}_2$ by $-\overline{\theta}$; and g by $\overline{g}$. Therefore, we can now apply Theorem 6.1 to obtain the following result.

Theorem 8.1. (Stochastic Black–Scholes Formula). *Suppose that the random functions* $\overline{b}, \overline{r}, \overline{\theta}$ *and* $\overline{g}$ *defined in* (8.10) *satisfy the conditions of Theorem* 6.1. *Let the unique adapted solution of* (8.9) *be* (X, Y, Z). *Then the hedging price against the contingent claim* $\xi = g(P_T, \cdot)$ *at time any* $t \in [0, T]$ *is given by*

$$Y_t = \tilde{E}\{\mathrm{e}^{-\int_t^T \overline{r}(s, X_s) \, \mathrm{d}s} \overline{g}(X_T, \cdot) | \mathscr{F}_t\} = \tilde{E}\{\mathrm{e}^{-\int_t^T r(s, P_s) \, \mathrm{d}s} g(P_T, \cdot) | \mathscr{F}_t\}, \tag{8.11}$$

where $\tilde{E}\{\cdot | \mathscr{F}_t\}$ *is the conditional expectation with respect to the equivalent martingale measure* $\tilde{P}$ *defined by*

$$\frac{\mathrm{d}\tilde{P}}{\mathrm{d}P} = \exp \left\{ -\int_0^T \overline{\theta}(t, X_t) \, \mathrm{d}W_t - \frac{1}{2} \int_0^T |\overline{\theta}(t, X_t)|^2 \, \mathrm{d}t \right\}.$$

Furthermore, the backward SPDE

$$
\begin{aligned}
u(t, x) = \overline{g}(x) + \int_t^T \Big\{ &\frac{1}{2} \sigma^2(s) u_{xx}(s, x) + (\overline{b}(s, x) - \sigma(s) \overline{\theta}(s, x)) u_x(s, x) \\
&- \overline{r}(s, x) u(s, x) + \sigma(s) q_x(s, x) - q(s, x) \overline{\theta}(t, x) \Big\} \, \mathrm{d}s - \int_t^T q(s, x) \, \mathrm{d}W_s
\end{aligned}
\tag{8.12}
$$

has a unique adapted solution (u, q), *such that the log-price* X *and the wealth process* Y *are related by*

$$Y_t = u(t, X_t, \cdot), \quad \forall t \in [0, T] \ \text{a.s.} \tag{8.13}$$

Finally, the optimal hedging strategy π is given by, for all $t \in [0, T]$,

$$\pi_t = \sigma^{-1}(t)Z_t = \nabla u(t,X_t,\cdot) + \sigma(t)^{-1}q(t,X_t,\cdot) = \sigma^{-1}(t)D_t(u(t,X_t,\cdot)) \quad a.s., \quad (8.14)$$

where D is the Malliavin derivative operator. In particular, if u satisfies the hypothesis (B) in Section 7, then

$$q(t,X_t,\cdot) = [D_t u](t,X_t,\cdot), \quad \forall t \in [0,T] \ a.s. \tag{8.15}$$

Proof. First, it is by now well known that in this setting the hedging price at any time t is just Y_t (see El Karoui et al., 1997 or Cvitanic and Ma, 1996), and (8.11) holds. The rest of the theorem, including (8.13) and (8.14), is now a direct consequence of Theorem 6.1 and 7.3, the definition $Z_t = \sigma(t)\pi_t$, and the fact that $\sigma^{-1}(t)$ exists for any t. $\square$

Remark 8.2. In the case when all the coefficients are constants, we see from Theorem 6.1 that the unique solution of the BSPDE (8.12) is $(u, 0)$, where u is the classical solution to a backward PDE which, after a change of variable $x = \log x'$ and by setting $v(t,x') = u(t, \log x')$, becomes exactly the same as (8.4). Thus, by setting $q = 0$ in (8.13) and (8.14), we see that the Theorem 8.1 recovers the classical Black–Scholes formula.

References

Antonelli, F., 1993. Backward–forward stochastic differential equations. Ann. Appl. Prob. 3, 777–793.

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

Bensoussan, A., 1992. Stochastic Control of Partially Observable Systems. Cambridge University Press, Cambridge.

Bismut, J.M., 1973. Théorie Probabiliste du Contrôle des Diffusions. Mem. Amer. Math. Soc. 176, Providence, RI.

Cvitanic, J., Ma, J., 1996. Hedging options for a large investor and forward–backward SDEs. The Annals of Applied Probabil. 6, 2, 370–398.

Duffie, D., Ma, J., Yong, J., 1995. Black's consol rate conjecture. Ann. Appl. Probab. 5(2), 356–382.

El Karoui, N., Peng, S., Quenez, M.C., Backward stochastic differential equations in finance. Math. Finance, to appear.

Hu, Y., Peng, S., 1991. Adapted solution of a backward semilinear stochastic evolution equations. Stoch. Anal. Appl. 9, 445–459.

Hu, Y., Peng, S., Solution of forward–backward stochastic differential equations. Probab. Theory Related Fields, to appear.

Karatzas, I., Shreve, S., 1988. Brownian Motion and Stochastic Calculus. Springer, Berlin.

Krylov, N.V., Rozovskii, B.L., 1982. Stochastic partial differential equations and diffusion processes. Uspekhi Mat. Nauk 37(6), 75–95.

Kunita, H., Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Math. vol. 24, Cambridge University Press, Cambridge, 1990.

Ma, J., Yong, J., 1995. Solvability of forward–backward SDEs and the nodal set of HJB equations. Chin. Ann. Math. 16B(3), 279–298.

Ma, J., Protter, P., Yong, J., 1994. Solving forward–backward stochastic differential equations explicitly – a four step scheme. Probab. Theory Related Fields 98, 339–359.

Nualart, D., Malliavin Calculus and Related Topics. Springer, Berlin, 1995.

Nualart, D., Pardoux, E., 1988. Stochastic calculus with anticipating integrands. Probab. Theory Related Fields 78, 535–581.

Ocone, D., 1984. Malliavin's calculus and stochastic integral representation of functionals of diffusion processes. Stochastics 12, 161–185.

Ocone, D., Pardoux, E., 1993. A stochastic Feynman–Kac formula for anticipating SPDE's and application to nonlinear smoothing. Stochastic Stochastic Rep. 45, 79–126.

Pardoux, E., 1979. Stochastic partial differential equations and filtering of diffusion processes. Stochastics 3, 127–167.

Pardoux, E., Peng, S., 1990. Adapted solution of a backward stochastic differential equation. Systems Control Lett. 14, 55–61.

Pardoux, E., Peng, S., 1992. Backward stochastic differential equations and quasilinear parabolic partial differential equations. Lecture Notes in CIS, Springer, Berlin, vol. 176, 200–217.

Peng, S., 1992. Stochastic Hamilton–Jacobi–Bellman equations. SIAM J. Control Optim. 30, 284–304.

Peng, S., Wu, Z., Fully coupled forward–backward stochastic differential equations with arbitrarily large time duration. Preprint.

Protter, P., 1990. Stochastic Integration and Differential Equations. A New Approach. Springer, Berlin.

Rozovskii, B.L., 1990. Stochastic Evolution Systems: Linear Theory and Applications to Non-linear Filtering. Kluwer Academic Publishers, Boston.

Zhou, X.Y., 1992. A duality analysis on stochastic partial differential equations. J. Funct. Anal. 103, 275–293.

Zhou, X.Y., 1993. On the necessary conditions of optimal controls for stochastic partial differential equations. SIAM J. Control Optim. 31, 1462–1478.