Probab. Theory Relat. Fields **113**, 135–170 (1999)

Probability Theory and Related Fields

On linear, degenerate backward stochastic partial differential equations

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

[1] Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395, USA. e-mail: majin@math.purdue.edu
[2] Laboratory of Mathematics for Nonlinear Sciences and Department of Mathematics, Fudan University, Shanghai 200433, China.

Received: 24 September 1997 / Revised version: 3 June 1998

Abstract. In this paper we study the well-posedness and regularity of the *adapted solutions* to a class of linear, degenerate backward stochastic partial differential equations (BSPDE, for short). We establish new *a priori* estimates for the adapted solutions to BSPDEs in a general setting, based on which the existence, uniqueness, and regularity of adapted solutions are obtained. Also, we prove some comparison theorems and discuss their possible applications in mathematical finance.

Mathematics Subject Classification (1991): 60H15, 35R60, 34F05, 93E20

1. Introduction

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, P)$ be a complete filtered probability space on which is defined a d-dimensional standard Brownian motion

---

Key words: Degenerate backward stochastic partial differential equations, adapted solutions, comparison theorems

J. Ma is supported in part by the Office of Naval Research grant N00014-96-1-0262

J. Yong is supported in part by the NSFC under grant 79790130, the National Distinguished Youth Science Foundation of China under grant 19725106 and the Chinese State Education Commission Science Foundation. Part of this work was completed when this author was visiting the Department of Mathematics, Purdue University

$W = \{W(t) : t \in [0,T]\}$ such that $\{\mathscr{F}_t\}_{t\geq 0}$ is the natural filtration generated by W, augmented by all the P-null sets in $\mathscr{F}$. Consider the (linear) backward stochastic partial differential equation (BSPDE, for short) of the following form:

$$\begin{cases} du = \{-\frac{1}{2}\nabla\cdot(ADu) - \langle a, Du\rangle - cu - \nabla\cdot(Bq) - \langle b, q\rangle - f\}\,dt \\ \qquad + \langle q,\, dW(t)\rangle, \\ u\big|_{t=T} = g \ , \end{cases} \tag{1.1}$$

where

$$\begin{cases} D\varphi = (\partial_{x_1}\varphi, \partial_{x_2}\varphi, \ldots, \partial_{x_n}\varphi)^T, \quad \forall \varphi \in C^1(\mathbb{R}^n; \mathbb{R}), \\ \nabla\cdot\xi = \sum_{i=1}^n \partial_{x_i}\xi_i, \quad \forall \xi = (\xi_1, \ldots, \xi_n) \in C^1(\mathbb{R}^n; \mathbb{R}^n) \ , \end{cases}$$

and

$$\begin{cases} A : [0,T] \times \mathbb{R}^n \times \Omega \to S^n, \quad B : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^{n\times d}, \\ a : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n, \quad b : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^d, \\ c, f : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}, \quad g : \mathbb{R}^n \times \Omega \to \mathbb{R} \ , \end{cases}$$

are random fields (S^n is the set of all $n \times n$-symmetric matrices) satisfying appropriate measurability and regularity conditions. 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 $\{\mathscr{F}_t\}$-adapted processes, and that (1.1) is satisfied in some sense. We would also like to study the regularity of the solution pair (u, q) in the variable x, and we will try to fulfill all these tasks under the following "minimum" (parabolicity) condition:

$$A(t,x) - B(t,x)B(t,x)^T \geq 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \ \text{a.s.} \tag{1.2}$$

BSPDE of this kind was originally motivated by stochastic control theory, as an adjoint equation appeared in the Pontryagin's maximum principle when the controlled system is a stochastic PDE (Zakai equation, for example), (see, e.g., Bensoussan [1], Nagasa-Nisio [16] or Zhou [25]). In our previous work [15], we showed that such a BSPDE is in fact a natural replacement of the parabolic PDE appearing in the Feynman-Kac formula, when the associated "diffusion" process has random coefficients. As a consequence, we then derived a stochastic version of the celebrated Black-Scholes formula for option pricing, in the case when the parameters of the stochastic differential equation that describes the stock price are allowed to be random processes.

Motivations involving the BSPDEs and their relations with the adapted solution to forward-backward stochastic differential equations (FBSDE, for short) with random coefficients were also discussed in [15]. (For more information on FBSDEs and their applications in finance, one is referred to the references [3–5], [7], [10], [13], [14], [22], and [23].)

The aim of the present paper is to develop a more complete theory concerning the existence, uniqueness, and regularity of the adapted solutions to the linear BSPDE (1.1), and to remove the technical assumption that we had to impose in our previous work [15]. More precisely, we will not assume that the coefficients A and B in (1.1) are independent of x. We note that the main feature of the BSPDE under study is its degeneracy (see (1.2)), combined with the fact that the operator $(\mathscr{M}_q \stackrel{\Delta}{=} \nabla \cdot (Bq) + \langle b, q \rangle)$ is a unbounded operator. Thus the existing method (e.g., [1], [17] and [25]) of combining finite dimensional approximations and duality relation between the BSPDE and forward SPDEs does not seem to apply directly due to the lack of satisfactory estimates on the second-first order differential operator involved in (1.1). We shall therefore first establish a general a priori estimate for BSPDEs, and then prove the well-posedness as well as some regularity of the (adapted) solutions to the BSPDEs. Our proof will take a quite different route than the one in [15] in order to remove the technical condition there, and our estimates in fact more precisely reflect the essence of the BSPDE. We should point out that in general higher dimensional cases we still need some conditions on the random field B, which we call the "symmetry condition", to compensate the degeneracy. But such a new condition is more reasonable (it is trivial in one-dimensional case), and in fact contains all the existing results known up to date. In some higher dimensional cases when the degeneracy is not "deadly" (by which we mean, for example, the equality holds in (1.2)), we prove further that the symmetric condition can be removed and the problem is essentially solved completely.

Another topic we would like to address in this paper is the comparison theorem for the solution to BSPDEs, since it is one of the indispensable tools in the theory of backward SDEs, and has its own interest as well. We shall establish a fundamental inequality, which will lead to various comparison theorems as corollaries.

This paper is organized as follows. In Section 2 we give some preliminaries. In Section 3, we state the main theorem of the paper, whose proof will be given in Sections 4–6. In Section 7 we prove a set of comparison theorems, and in Section 8 we discuss the motivation of these comparison theorems and their potential applications in, for example, the study of *robustness* of the Black-Scholes formula in the sense of El Karoui-Jeanblanc-Shreve [6].

2. Preliminaries

For any integer $m \geq 0$, 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 $C^m(\mathbb{R}^n;\mathbb{R}^\ell)$ 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 $\mathbf{E}$ by $\langle\cdot,\cdot\rangle$; and the norm in $\mathbf{E}$ by $|\cdot|$. With the notation $\partial_{x_i} = \frac{\partial}{\partial x_i}$, $i = 1,\dots,n$ and $\nabla = (\partial_{x_1},\dots,\partial_{x_n})^T$, we shall denote

- $D\varphi = (\partial_{x_1}\varphi,\dots,\partial_{x_n}\varphi)^T$, if $\varphi \in C^1(\mathbb{R}^n)$;
- $\nabla\cdot\xi = \partial_{x_1}\xi_1 + \dots + \partial_{x_n}\xi_n$, if $\xi = (\xi_1,\dots,\xi_d) \in C^1(\mathbb{R}^n;\mathbb{R}^d)$;
- $\nabla\cdot\Phi = (\nabla\cdot\Phi_1,\dots,\nabla\cdot\Phi_d)$, if $\Phi = (\Phi_1,\dots,\Phi_d) \in C^1(\mathbb{R}^n;\mathbb{R}^{n\times d})$ (hence each $\Phi_i \in C^1(\mathbb{R}^n;\mathbb{R}^d)$, $i = 1,\dots,d$).

Let $\mathscr{A} \stackrel{\Delta}{=} \{\alpha = (\alpha_1,\dots,\alpha_n) \mid \alpha_i \geq 0, \text{integers}\}$ be the set of *multi-indices*. For any $\alpha \in \mathscr{A}$ and $x = (x_1,\dots,x_n) \in \mathbb{R}^n$, we denote

$$|\alpha| = \sum_{i=1}^n \alpha_i, \quad \partial^\alpha \stackrel{\Delta}{=} \partial_{x_1}^{\alpha_1}\partial_{x_2}^{\alpha_2}\dots\partial_{x_n}^{\alpha_n}, \quad x^\alpha = x_1^{\alpha_1}\dots x_n^{\alpha_n} \ .$$

Further, if $\beta = (\beta_1,\beta_2,\dots,\beta_n)$ is another multi-index, then by $\beta \leq \alpha$ we mean that $\beta_i \leq \alpha_i$ for $1 \leq i \leq n$; and by $\beta < \alpha$ we mean that $\beta \leq \alpha$ and $|\beta| < |\alpha|$.

We now define some other spaces that will be used in the paper. Let $m \geq 0$ be an integer, and $1 \leq p, r \leq \infty$ be any real numbers. Let X be a Banach space, and $\mathscr{G}$ be a sub-σ-field of $\mathscr{F}$. We denote

- by $H^m(\mathbb{R}^n;\mathbf{E}) = W^{m,2}(\mathbb{R}^n;\mathbf{E})$ the usual Sobolev space;
- by $L^p_{\mathscr{G}}(\Omega;X)$ the set of all X-valued, $\mathscr{G}$-measurable random variable η such that $E\|\eta\|_X^p < \infty$;
- by $L^p_{\mathscr{F}}(0,T;L^r(\Omega;X))$ the set of all $\{\mathscr{F}_t\}$-predictable X-valued processes $\varphi(t,\omega) : [0,T]\times\Omega \to X$ such that $\|\varphi\|_{L^p_{\mathscr{F}}(0,T;L^r(\Omega;X))} \stackrel{\Delta}{=} \left\{\int_0^T \left[E\|\varphi(t)\|_X^r\right]^{p/r} dt\right\}^{1/p} < \infty$.
- by $C_{\mathscr{F}}([0,T];L^r(\Omega;X))$ the set of all $\{\mathscr{F}_t\}$-predictable X-valued continuous processes $\varphi(t,\omega) : [0,T]\times\Omega \to X$ such that $\|\varphi\|_{C_{\mathscr{F}}([0,T];L^r(\Omega;X))} \stackrel{\Delta}{=} \max_{t\in[0,T]} \left[E\|\varphi(t)\|_X^r\right]^{1/r} < \infty$.

If $r = p$, we shall denote $L^p_{\mathscr{F}}(0,T;L^p(\Omega;X))$ by $L^p_{\mathscr{F}}(0,T;X)$ for simplicity. Also, if $\mathbf{E} = \mathbb{R}$, we denote $H^m(\mathbb{R}^n;\mathbb{R}) = H^m(\mathbb{R}^n)$ (or simply H^m).

The following fact concerning the differentiability of stochastic integrals with parameter is important for our purpose. Let $h \in L^2_{\mathscr{F}}(0,T;C^m_b(\mathbb{R}^n;\mathbb{R}^d))$. Then it can be shown (see, for example, [12, 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 $L^2_{\mathscr{F}}(0,T;C^{m-1})$ and it satisfies

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

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

Finally, if the coefficients A and B are differentiable, which we will always assume, then equation (1.1) (in *divergence form*) is equivalent to an equation of a *general form*. To wit, let D^2u be the Hessian of u and $Dq \overset{\Delta}{=} (Dq_1,\ldots,Dq_d) \overset{\Delta}{=} (\partial_{x_i} q_j)^{n,d}_{i,j=1}$; and note that

$$\begin{cases} \mathrm{tr}[AD^2u] = \nabla \cdot (ADu) - \langle \nabla \cdot A, Du\rangle; \\ \mathrm{tr}[B^T Dq] = \nabla \cdot (Bq) - \langle \nabla \cdot B, q\rangle \ , \end{cases} \tag{2.2}$$

then (1.1) is the same as

$$\begin{cases} du = \{-\frac{1}{2}\mathrm{tr}[AD^2u] - \langle \tilde{a}, Du\rangle - cu - \mathrm{tr}[B^T Dq] - \langle \tilde{b}, q\rangle - f\}\, dt \\ \qquad + \langle q, dW(t)\rangle, \\ u\big|_{t=T} = g \ , \end{cases} \tag{2.3}$$

where

$$\tilde{a} = a + \tfrac{1}{2}\nabla \cdot A; \qquad \tilde{b} = b + \nabla \cdot B \ , \tag{2.4}$$

Since (1.1) and (2.3) are equivalent, all the results for (1.1) can be automatically carried over to (2.3) and vice versa. To simplify discussion, we shall concentrate on (1.1) for well-posedness (§§3–6) and on (2.3) for comparison theorems (§7).

Now, we introduce the following definition.

Definition 2.1. *We say that the equation* (1.1) *is*

(i) *parabolic*, *if A and B satisfy* (1.2);
(ii) *super-parabolic*, *if there exists a constant* $\delta > 0$, *such that*

$$A(t,x) - B(t,x)B(t,x)^T \geq \delta I, \quad \text{a.e. } (t,x) \in [0,T] \times \mathbb{R}^n, \text{ a.s.} \tag{2.5}$$

(iii) *degenerate parabolic*, *if* (1.2) *holds and there exists a set* $G \subseteq [0, T] \times \mathbb{R}^n$ *of positive Lebesgue measure*, *such that*

$$\det[A(t,x) - B(t,x)B(t,x)^T] = 0, \quad \forall (t,x) \in G. \text{ a.s.} \tag{2.6}$$

We see that in Definition 2.1 only the coefficients A and B are involved and other coefficients are irrelevant. Thus the definitions apply to (2.3) automatically.

Note that if (1.1) is super-parabolic, then it is necessary that $A(t,x)$ is uniformly positive definite, i.e.,

$$A(t,x) \geq \delta I > 0, \quad \text{a.e.} \quad (t,x) \in [0,T] \times \mathbb{R}^n, \text{ a.s.} \tag{2.7}$$

However, (2.7) and (1.2) *do not* imply the super-parabolicity of (1.1). For example, if $A(t,x)$ satisfies (2.7) but

$$A(t,x) = B(t,x)B(t,x)^T, \quad \text{a.e.} \, (t,x) \in [0,T] \times \mathbb{R}^n, \text{ a.s.} \tag{2.8}$$

then (1.1) is always degenerate parabolic! We note that this is the case of special interest, because it is exactly what one has to consider for the Stochastic Feynman-Kac Formula, or for the BSPDEs related to an FBSDE with random coefficients via *Four Step Scheme* (see [15]).

Let us now turn to the notion of solutions to (1.1). In what follows, we denote $B_R = \{x \in \mathbb{R}^n \mid |x| < R\}$ for any $R > 0$.

Definition 2.2. *Let* $\{(u(t,x;\omega), q(t,x;\omega)), (t,x,\omega) \in [0,T] \times \mathbb{R}^n \times \Omega\}$ *be a pair of random fields.* (u,q) *is called*

(i) *an adapted classical solution* of (1.1), *if*

$$\begin{cases} u \in C_{\mathscr{F}}([0,T]; L^2(\Omega; C^2(\bar{B}_R))), \\ q \in L^2_{\mathscr{F}}(0,T; C^1(\bar{B}_R; \mathbb{R}^d)), \end{cases} \quad \forall R > 0 \ , \tag{2.9}$$

such that the following holds for all $(t,x) \in [0,T] \times \mathbb{R}^n$, *almost surely*:

$$\begin{aligned} u(t,x) = g(x) + \int_t^T \big\{ &\tfrac{1}{2} \nabla \cdot [A(s,x)Du(s,x)] + \langle a(s,x), Du(s,x) \rangle \\ &+ c(s,x)u(s,x) + \nabla \cdot [B(s,x)q(s,x)] \\ &+ \langle b(s,x), q(s,x) \rangle + f(s,x) \big\} \, ds \\ &- \int_t^T \langle q(s,x), dW(s) \rangle \ . \end{aligned} \tag{2.10}$$

(ii) *an adapted strong solution of* (1.1), *if*

$$\begin{cases} u \in C_{\mathscr{F}}([0,T]; L^2(\Omega; H^2(B_R))), \\ q \in L^2_{\mathscr{F}}(0,T; H^1(B_R; \mathbb{R}^d)), \end{cases} \quad \forall R > 0 \ , \tag{2.11}$$

such that (2.10) *holds for all* $t \in [0, T]$, *a.e.* $x \in \mathbb{R}^n$, *almost surely*.

(iii) *an adapted weak solution of* (1.1) *if*

$$\begin{cases} u \in C_{\mathscr{F}}([0,T]; L^2(\Omega; H^1(B_R))), \\ q \in L^2_{\mathscr{F}}(0,T; L^2(B_R; \mathbb{R}^d)), \end{cases} \quad \forall R > 0 \ , \tag{2.12}$$

such that for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ *and all* $t \in [0, T]$, *it holds almost surely that*

$$\begin{aligned} &\int_{\mathbb{R}^n} u(t,x)\varphi(x)\, dx - \int_{\mathbb{R}^n} g(x)\varphi(x)\, dx \\ &\quad = \int_t^T \int_{\mathbb{R}^n} \big\{ -\tfrac{1}{2}\langle A(s,x)Du(s,x), D\varphi(x)\rangle + \langle a(s,x), Du(s,x)\rangle \varphi(x) \\ &\qquad\qquad + c(s,x)u(s,x)\varphi(s,x) - \langle B(s,x)q(s,x), D\varphi(x)\rangle \\ &\qquad\qquad + \langle b(s,x), q(s,x)\rangle \varphi(x) + f(s,x)\varphi(x) \big\}\, dx\, ds \\ &\qquad\qquad - \int_t^T \Big\langle \int_{\mathbb{R}^n} q(s,x)\varphi(x)\, dx, dW(s) \Big\rangle \ . \end{aligned} \tag{2.13}$$

It is clear that for (1.1), an adapted classical solution is an adapted strong solution; and an adapted strong solution is an adapted weak solution. The following result tells the reverse implications which will be useful later. To simplify the statement we borrow the assumption "$(\mathrm{H})_m$" from the next section (see (3.1)).

Proposition 2.3. *Suppose that the assumption* $(H)_m$, (3.1), *holds for* $m = 0$. *Then an adapted weak solution* (u, q) *of* (1.1) *satisfying* (2.11) *is an adapted strong solution of* (1.1). *If further* (2.9) *holds, then* (u, q) *is an adapted classical solution of* (1.1).

Proof. Let (u, q) be an adapted weak solution of (1.1) such that (2.11) holds. Then using integration by parts in (2.13) we have

$$\begin{aligned} &\int_{\mathbb{R}^n} \{u(t,x) - g(x)\}\varphi(x)\, dx \\ &\quad = \int_{\mathbb{R}^n} \Bigg\{ \int_t^T \big\{ \tfrac{1}{2}\nabla\cdot[A(s,x)Du(s,x)] + \langle a(s,x), Du(s,x)\rangle \\ &\qquad\qquad + c(s,x)u(s,x) + \nabla\cdot[B(s,x)q(s,x)] + \langle b(s,x), q(s,x)\rangle \\ &\qquad\qquad + f(s,x)\big\}\, ds - \int_t^T \langle q(s,x), dW(s)\rangle \Bigg\} \varphi(x)\, dx \ . \end{aligned} \tag{2.14}$$

Since the above is true for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, (2.10) follows. Namely (u, q) is an adapted strong solution. The other assertion is obvious. □

3. Well-posedness of linear BSPDEs

In this section we state the main results concerning the well-posedness for BSPDE (1.1). The proofs of these results will be carried out in later sections. To begin with, let us introduce the following main assumption. Let $m \geq 1$ be an integer.

$(\mathrm{H})_m$ Functions $\{A, B, a, b, c\}$ satisfy the following:

$$\begin{cases} A \in L^\infty_{\mathscr{F}}(0,T; C_b^{m+1}(\mathbb{R}^n; S^n)), & B \in L^\infty_{\mathscr{F}}(0,T; C_b^{m+1}(\mathbb{R}^n; \mathbb{R}^{n\times d})), \\ a \in L^\infty_{\mathscr{F}}(0,T; C_b^{m}(\mathbb{R}^n; \mathbb{R}^n)), & b \in L^\infty_{\mathscr{F}}(0,T; C_b^{m}(\mathbb{R}^n; \mathbb{R}^d)), \\ c \in L^\infty_{\mathscr{F}}(0,T; C_b^{m}(\mathbb{R}^n)) \ . \end{cases} \tag{3.1}$$

We note that $(\mathrm{H})_m$ implies that the partial derivatives of A and B in x up to order $(m+1)$, and those of a, b and c up to order m are uniformly bounded in (t, x, ω) by a common constant $K_m > 0$.

Theorem 3.1. *Let the parabolicity condition* (1.2) *hold and* $(H)_m$ *hold for some* $m \geq 1$. *Suppose further that the coefficient* $B(t,x)$ *satisfies the following "symmetry condition"*: for $1 \leq i \leq n$,

$$\left[B(\partial_{x_i} B^T)\right]^T = B(\partial_{x_i} B^T), \quad a.e.\ (t,x) \in [0,T] \times \mathbb{R}^n, \quad a.s., \ . \tag{3.2}$$

Then for any random fields f *and* g *satisfying*

$$f \in L^2_{\mathscr{F}}(0,T; H^m(\mathbb{R}^n)), \quad g \in L^2_{\mathscr{F}_T}(\Omega; H^m(\mathbb{R}^n)) \ , \tag{3.3}$$

BSPDE (1.1) *admits an adapted weak solution*; *and it is unique in the class of random fields* (u, q) *such that* $u \in C_{\mathscr{F}}([0,T]; L^2(\Omega; H^1(\mathbb{R}^n)))$ *and* $q \in L^2_{\mathscr{F}}(0,T; L^2(\mathbb{R}^n; \mathbb{R}^d))$. *Moreover*, *the adapted weak solution* (u,q) *satisfies the following estimate*:

$$\begin{aligned} &\max_{t\in[0,T]} E\|u(t,\cdot)\|^2_{H^m} + E\int_0^T \|q(t,\cdot)\|^2_{H^{m-1}}\,dt \\ &\quad + \sum_{|\alpha|\leq m} E\int_0^T\int_{\mathbb{R}^n} \left\{\langle (A - BB^T)D(\partial^\alpha u), D(\partial^\alpha u)\rangle + |B^T[D(\partial^\alpha u)] \right. \\ &\quad \left. + \partial^\alpha q|^2\right\}\,dx\,dt \leq CE\left\{\int_0^T \|f(t,\cdot)\|^2_{H^m}\,dt + \|g\|^2_{H^m}\right\} \ , \end{aligned} \tag{3.4}$$

where the constant $C > 0$ *only depends on* m, T *and* K_m.

Furthermore, if $m \geq 2$, *the weak solution* (u, q) *becomes the unique adapted strong solution of* (1.1); *and if* $m > 2 + n/2$, *then* (u, q) *is the unique adapted classical solution of* (1.1).

The symmetry condition (3.2) is a technical condition that will play a very important role in deriving the fundamental a priori estimates in Lemma 5.1 to follow. However, we would like to point out here that such a condition is not necessary in the proof of the uniqueness of adapted weak (and hence strong and classical) solutions, as we shall see in 4. Several examples of B satisfying condition (3.2) are listed below:

- $d = n = 1$ (B is a scalar);
- B is independent of x;
- $B(t, x) = \varphi(t, x) B_0(t)$, where φ is a scalar-valued random field.

The following result shows that the symmetry condition (3.2) can be removed if the parabolicity condition (1.2) is slightly strengthened.

Theorem 3.2. *Suppose* (1.2) *holds and* $(H)_m$ *with* $m \geq 1$ *is in force. Suppose further that for some* $\varepsilon_0 > 0$, *either*

$$A - BB^T \geq \varepsilon_0 BB^T \geq 0, \quad a.e.\ (t, x) \in [0, T] \times \mathbb{R}^n, a.s. \quad , \tag{3.5}$$

or

$$A - BB^T \geq \varepsilon_0 \sum_{|\alpha|=1} (\partial^\alpha B)(\partial^\alpha B^T) \geq 0, \qquad a.e.\ (t, x) \in [0, T] \times \mathbb{R}^n, a.s. \tag{3.6}$$

Then the conclusion of Theorem 3.1 *remains true*; *and in the case* (3.5) *holds*, *the estimate* (3.4) *can be improved to the following*:

$$\begin{aligned} &\max_{t \in [0,T]} E\|u(t, \cdot)\|^2_{H^m} + E \int_0^T \|q(t, \cdot)\|^2_{H^m}\, dt \\ &\quad + \sum_{|\alpha| \leq m} E \int_0^T \int_{\mathbb{R}^n} \langle AD(\partial^\alpha u), D(\partial^\alpha u) \rangle \, dx\, dt \\ &\quad \leq CE \left\{ \int_0^T \|f(t, \cdot)\|^2_{H^m}\, dt + \|g\|^2_{H^m} \right\} \ , \end{aligned} \tag{3.7}$$

where the constant $C > 0$ *depends only on* m, T, K_m *and* ε_0.

In addition to (3.5) *or* (3.6), *if A is uniformly positive definite, i.e.,* (2.7) *holds for some $\delta > 0$ (this is the case if* (1.1) *is super-parabolic, i.e.,* (2.5) *holds*), *then* (3.7) *can further be improved to the following*:

$$\max_{t\in[0,T]} E\|u(t,\cdot)\|^2_{H^m} + E\int_0^T \{\|u(t,\cdot)\|^2_{H^{m+1}} + \|q(t,\cdot)\|^2_{H^m}\}\,dt \\ \le CE\left\{\int_0^T \|f(t,\cdot)\|^2_{H^m}\,dt + \|g\|^2_{H^m}\right\} \ . \tag{3.8}$$

Remark. The conditions (3.5) (or (3.6)), even together with (2.7), is still weaker than the super-parabolicity condition (2.5). For example, if $n > d$ and B is an $(n \times d)$ matrix, then BB^T is always degenerate. We can easily find an A such that (3.5) (or (3.6)) and (2.7) hold but (2.5) fails.

In Theorems 3.1 and 3.2 we have assumed that f and g are square integrable in $x \in \mathbb{R}^n$ globally. This will exclude the cases in which, say, f and g have polynomial growth, which is often seen in applications. In the rest of this section we would like to relax such a restriction, by using a more or less standard method. Note that if (u, q) is an adapted classical solution of (1.1), and if we let $\lambda > 0$, denote $\langle x\rangle \stackrel{\Delta}{=} \sqrt{|x|^2 + 1}$, and set

$$v(t,x) = \mathrm{e}^{-\lambda\langle x\rangle}u(t,x); \quad p(t,x) = \mathrm{e}^{-\lambda\langle x\rangle}q(t,x), \quad (t,x) \in [0,T] \times \mathbb{R}^n \ , \tag{3.9}$$

then a direct computation shows that (v, p) satisfies the following BSPDE (compare with (1.1)):

$$\begin{cases} dv = \{-\frac{1}{2}\nabla\cdot(ADv) - \langle\bar{a}, Dv\rangle - \bar{c}v - \nabla\cdot(Bp) - \langle\bar{b}, p\rangle - \bar{f}\}\,dt \\ \qquad + \langle p, dW(t)\rangle, \\ v\big|_{t=T} = \bar{g} \ , \end{cases} \tag{3.10}$$

with

$$\begin{cases} \bar{a} = a + \lambda A\left[\frac{x}{\langle x\rangle}\right], \quad \bar{b} = b + \lambda B^T\left[\frac{x}{\langle x\rangle}\right], \\ \bar{c} = c + \dfrac{\lambda^2}{2}\left\langle A\left[\frac{x}{\langle x\rangle}\right], \left[\frac{x}{\langle x\rangle}\right]\right\rangle + \dfrac{\lambda}{2}\nabla\cdot\left(A\left[\frac{x}{\langle x\rangle}\right]\right) + \lambda\left\langle a, \left[\frac{x}{\langle x\rangle}\right]\right\rangle, \\ \bar{f} = \mathrm{e}^{-\langle\lambda\rangle}f(t,x), \quad \bar{g} = \mathrm{e}^{-\langle\lambda\rangle}g(x) \ . \end{cases} \tag{3.11}$$

Conversely, if (v, p) is an adapted classical solution of (3.10), then (u, q) determined by (3.9) is an adapted classical solution of (1.1).

Clearly, the same equivalence between (1.1) and (3.10) holds for adapted strong and weak solutions, respectively.

On the other hand, from (3.11) we see easily that $\{A, B, a, b, c\}$ satisfies $(\mathrm{H})_m$ if and only if $\{A, B, \bar{a}, \bar{b}, \bar{c}\}$ satisfies $(\mathrm{H})_m$. This is because $\|\partial^\alpha \langle x \rangle\|_\infty \leq C_\alpha$ for any multiindex α. Hence, from Theorems 3.1 and 3.2, we can derive the following result.

Theorem 3.3. *Let* $m \geq 1$ *and* $(H)_m$ *hold for* $\{A, B, a, b, c\}$. *Let* (1.2) *and* (3.2) *hold. Let* $\lambda > 0$ *such that*

$$e^{-\lambda\langle\cdot\rangle} f \in L^2_{\mathscr{F}}(0, T; H^m(\mathbb{R}^n)), \quad e^{-\lambda\langle\cdot\rangle} g \in L^2_{\mathscr{F}_T}(\Omega; H^m(\mathbb{R}^n)) \ . \tag{3.12}$$

Then BSPDE (1.1) *admits a unique adapted weak solution. Moreover, the weak solution* (u, q) *satisfies the following estimate*:

$$\begin{aligned} &\max_{t\in[0,T]} E\|e^{-\lambda\langle\cdot\rangle}u(t,\cdot)\|^2_{H^m} + E\int_0^T \|e^{-\lambda\langle\cdot\rangle}q(t,\cdot)\|^2_{H^{m-1}}\,dt \\ &\quad + \sum_{|\alpha|\leq m} E\int_0^T\!\!\int_{\mathbb{R}^n} \big\{\langle (A - BB^T)D[\partial^\alpha(e^{-\lambda\langle\cdot\rangle}u)], D[\partial^\alpha(e^{-\lambda\langle\cdot\rangle}u)]\rangle \\ &\quad + \big|B^T\{D[\partial^\alpha(e^{-\lambda\langle\cdot\rangle}u)]\} + \partial^\alpha(e^{-\lambda\langle\cdot\rangle}q)\big|^2\big\}\,dx\,dt \\ &\qquad \leq CE\bigg\{\int_0^T \|e^{-\lambda\langle\cdot\rangle}f(t,\cdot)\|^2_{H^m}\,dt + \|e^{-\lambda\langle\cdot\rangle}g\|^2_{H^m}\bigg\} \ , \end{aligned} \tag{3.13}$$

where the constant $C > 0$ *only depends on* m, T *and* K_m.

Furthermore, if $m \geq 2$, *the weak solution* (u, q) *becomes the unique adapted strong solution of* (1.1); *if* $m > 2 + n/2$, *then* (u, q) *is the unique adapted classical solution of* (1.1).

In the case that (3.2) *is replaced by* (3.5) *or* (3.6), *the above conclusion remains true*; *and if* (3.5) *holds*, *the estimate* (3.13) *can be improved to the following*:

$$\begin{aligned} &\max_{t\in[0,T]} E\|e^{-\lambda\langle\cdot\rangle}u(t,\cdot)\|^2_{H^m} + E\int_0^T \|e^{-\lambda\langle\cdot\rangle}q(t,\cdot)\|^2_{H^m}\,dt \\ &\quad + \sum_{|\alpha|\leq m} E\int_0^T\!\!\int_{\mathbb{R}^n} \langle AD[\partial^\alpha(e^{-\lambda\langle\cdot\rangle}u)], D[\partial^\alpha(e^{-\lambda\langle\cdot\rangle}u)]\rangle\,dx\,dt \\ &\qquad \leq CE\bigg\{\int_0^T \|e^{-\lambda\langle\cdot\rangle}f(t,\cdot)\|^2_{H^m}\,dt + \|e^{-\lambda\langle\cdot\rangle}g\|^2_{H^m}\bigg\} \ , \end{aligned} \tag{3.14}$$

Finally, if in addition to (3.5), (2.5) *holds for some* $\delta > 0$, *then* (3.14) *can be further improved to the following*:

$$\max_{t\in[0,T]} E\|e^{-\lambda\langle\cdot\rangle}u(t,\cdot)\|_{H^m}^2 + E\int_0^T \{\|e^{-\lambda\langle\cdot\rangle}u(t,\cdot)\|_{H^{m+1}}^2 + \|e^{-\lambda\langle\cdot\rangle}q(t,\cdot)\|_{H^m}^2\}\, dt$$
$$\le CE\left\{\int_0^T \|e^{-\lambda\langle\cdot\rangle}f(t,\cdot)\|_{H^{m-1}}^2\, dt + \|e^{-\lambda\langle\cdot\rangle}g\|_{H^m}^2\right\} \ . \tag{3.15}$$

Clearly, (3.12) means that f and g can have an exponential growth as $|x| \to \infty$. This is good enough for many applications.

4. Uniqueness of adapted solutions

In this section we are going to establish the uniqueness of adapted weak (whence strong and classical) solutions to our BSPDEs. We first note that since the equation is linear, the uniqueness of weak solutions among those that satisfy (3.4) is immediate. To wit, if (u^i, q^i), $i = 1,2$ are two adapted solutions that both satisfy (3.4) with $m = 1$, say, then $(\hat{u}, \hat{q}) \stackrel{\Delta}{=} (u^1 - u^2, q^1 - q^2)$ will be a solution to the homogeneous BSPDE ($f = g = 0$). Thus it follows from (3.4) that $\hat{u} = 0$ and $\hat{q} = 0$. The purpose of this section, however, is to prove uniqueness without using estimate (3.4), since one does not know a priori that all the weak solutions would satisfy (3.4). We also note that the conditions for uniqueness are much weaker than those for existence (for instance, the symmetry condition (3.2) is not needed), since only the *weak solutions* will be considered, thanks to the discussion right before Proposition 2.3. For notational convenience, we denote

$$\mathscr{L}u \stackrel{\Delta}{=} \tfrac{1}{2}\nabla\cdot[ADu] + \langle a, Du\rangle + cu, \qquad \mathscr{M}q \stackrel{\Delta}{=} \nabla\cdot[Bq] + \langle b, q\rangle \ . \tag{4.1}$$

Then equation (1.1) is the same as the following:

$$\begin{cases} du = -\{\mathscr{L}u + \mathscr{M}q + f\}\, dt + \langle q, dW(t)\rangle, & (t,x) \in [0,T]\times \mathbb{R}^n, \\ u|_{t=T} = g \ . \end{cases} \tag{4.2}$$

the main result of this section is the following theorem.

Theorem 4.1. *Suppose that the parabolicity condition* (1.2) *holds*; *and that* (3.3) *and* $(H)_m$ *hold for* $m = 1$. *Then* (4.2) *admits at most one adapted weak solution* (u, q) *in the class*

$$\begin{cases} u \in C_{\mathscr{F}}([0,T]; L^2(\Omega; H^1(\mathbb{R}^n))), \\ q \in L^2_{\mathscr{F}}(0,T; L^2(\mathbb{R}^n; \mathbb{R}^d)) \ . \end{cases} \tag{4.3}$$

To prove the above uniqueness theorem, we need some preparations. First let us recall the *Gelfand triple* $H^1(\mathbb{R}^n) \hookrightarrow L^2(\mathbb{R}^n) \hookrightarrow H^{-1}(\mathbb{R}^n)$. Here, $H^{-1}(\mathbb{R}^n)$ is the dual space of $H^1(\mathbb{R}^n)$, and the embeddings are dense and continuous. We denote the duality paring between $H^1(\mathbb{R}^n)$ and $H^{-1}(\mathbb{R}^n)$ by $\langle\cdot,\cdot\rangle_0$, and the inner product and the norm in $L^2(\mathbb{R}^n)$ by $(\cdot,\cdot)_0$ and $|\cdot|_0$, respectively. Then, by identifying $L^2(\mathbb{R}^n)$ with its dual $L^2(\mathbb{R}^n)^*$, we have the following:

$$\langle\psi,\varphi\rangle_0 = (\psi,\varphi)_0 = \int_{\mathbb{R}^n} \psi(x)\varphi(x)\,dx, \quad \forall\psi \in L^2(\mathbb{R}^n), \quad \varphi \in H^1(\mathbb{R}^n) \ , \tag{4.4}$$

and

$$\begin{cases} \displaystyle\sum_{i=1}^n \partial_i\psi_i \in H^{-1}(\mathbb{R}^n), \quad \forall\psi_i \in L^2(\mathbb{R}^n), \quad 1 \le i \le n, \\ \displaystyle\Big\langle \sum_{i=1}^n \partial_i\psi_i, \varphi \Big\rangle_0 \stackrel{\Delta}{=} -\int_{\mathbb{R}^n} \psi_i(x)\partial_i\varphi(x)\,dx, \quad \forall\varphi \in H^1(\mathbb{R}^n) \ . \end{cases} \tag{4.5}$$

Next, let (u,q) be an adapted weak solution of (4.2) satisfying (4.3). Note that in (4.3), the integrability of (u,q) in x is required to be global. By (4.4)–(4.5), we see that $\mathscr{L}u + \mathscr{M}q \in L^2_{\mathscr{F}}(0,T;H^{-1}(\mathbb{R}^n))$, and consequently (2.13) holds for any $\varphi \in H^1(\mathbb{R}^n)$ (not just $C_0^\infty(\mathbb{R}^n)$). Namely,

$$\begin{cases} d(u,\varphi)_0 = -\langle \mathscr{L}u + \mathscr{M}q + f, \varphi\rangle_0\,dt + \langle (q,\varphi)_0, dW(t)\rangle, \quad t \in [0,T], \\ (u,\varphi)_0\big|_{t=T} = (g,\varphi)_0 \ . \end{cases} \tag{4.6}$$

Here, $(q,\varphi)_0 \stackrel{\Delta}{=} ((q_1,\varphi)_0,\ldots,(q_d,\varphi)_0)$ and $q = (q_1,\ldots,q_d)$. In the sequel we shall say that (4.2) holds in $H^{-1}(\mathbb{R}^n)$ if (4.6) holds for all $\varphi \in H^1(\mathbb{R}^n)$.

The following form of Itô's formula can be found in [17].

Lemma 4.2. *Let* $\xi \in L^2_{\mathscr{F}}(0,T;H^{-1}(\mathbb{R}^n))$ *and* (u,q) *satisfy* (4.3), *such that*

$$du = \xi\,dt + \langle q, dW(t)\rangle, \quad t \in [0,T] \quad .$$

Then for $t \in [0,T]$ *it holds that*

$$\begin{aligned} |u(t)|_0^2 = |u(0)|_0^2 &+ \int_0^t \left\{2\langle\xi(s),u(s)\rangle_0 + |q(s)|_0^2\right\} ds \\ &+ 2\int_0^t \langle (q(s),u(s))_0, dW(s)\rangle \ . \end{aligned} \tag{4.7}$$

Proof of Theorem 4.1. We need only show that if (u,q) is any adapted weak solution of (4.2) with $f = g \equiv 0$ and (4.3) holds, then $(u,q) = 0$.

To this end, applying Lemma 4.2, and noting that $\mathscr{L}u + \mathscr{M}q \in L^2_{\mathscr{F}}(0, T; H^{-1}(\mathbb{R}^n))$, we have, for $t \in [0, T]$,

$$\begin{aligned} E|u(t)|_0^2 &= E\int_t^T \left\{2\langle \mathscr{L}u(s) + \mathscr{M}q(s), u(s)\rangle_0 - |q(s)|_0^2\right\} ds \\ &= E\int_t^T\int_{\mathbb{R}^n} \{-\langle ADu, Du\rangle + \langle a, D(u^2)\rangle + 2cu^2 \\ &\quad - 2\langle q, B^T Du\rangle + 2\langle bu, q\rangle - |q|^2\}\, dx\, ds \\ &= E\int_t^T\int_{\mathbb{R}^n} \{-\langle (A - BB^T)Du, Du\rangle - |q + B^T Du - bu|^2 \\ &\quad + [b^2 + 2c - \nabla\cdot(a + Bb)]u^2\}\, ds \\ &\le C\int_t^T E|u(s)|_0^2\, ds \ . \end{aligned} \tag{4.8}$$

Thus Gronwall's inequality leads to $E|u(t)|_0^2 = 0$, $t \in [0, T]$, i.e., $u = 0$. By (4.8) again, we must also have $q = 0$. This proves the uniqueness of adapted weak solutions to (4.2). □

5. Existence of adapted solutions

In this section we prove the existence part of Theorem 3.1. We shall start with two technical lemmas that will be useful in the proof. The first Lemma is fundamental, but its proof is rather technical and lengthy. We therefore postpone its proof to the next section in order not to disturb our discussion.

Lemma 5.1. *Let the parabolicity condition* (1.2) *and the symmetry condition* (3.2) *hold. Let* $(H)_m$ *hold for some* $m \ge 1$. *Then, there exists a constant* $C > 0$, *such that for any* $u \in C_0^\infty(\mathbb{R}^n)$ *and* $q \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^d)$, *it holds that*

$$\begin{aligned} &\int_{\mathbb{R}^n}\left\{\sum_{|\alpha|\le m}\{\langle (A - BB^T)D(\partial^\alpha u), D(\partial^\alpha u)\rangle + |B^T D(\partial^\alpha u) + \partial^\alpha q|^2\}\right. \\ &\qquad \left. + \sum_{|\alpha|\le m-1} |\partial^\alpha q|^2\right\} dx \\ &\qquad \le C\int_{\mathbb{R}^n}\sum_{|\alpha|\le m}\left\{ -2(\partial^\alpha u)\partial^\alpha(\mathscr{L}u + \mathscr{M}q) + |\partial^\alpha q|^2 + |\partial^\alpha u|^2\right\} dx, \\ &\qquad\qquad \text{a.e.}\, t \in [0, T], \ \text{a.s.} \end{aligned} \tag{5.1}$$

The above remains true if (3.5) *or* (3.6) *holds instead of* (3.2); *and in the case* (3.6) *holds, the above can be replaced by the following*:

$$\int_{\mathbb{R}^n} \Big\{ \sum_{|\alpha|\le m} \langle AD(\partial^\alpha u), D(\partial^\alpha u)\rangle + \sum_{|\alpha|\le m} |\partial^\alpha q|^2 \Big\}\, dx$$

$$\le C \int_{\mathbb{R}^n} \sum_{|\alpha|\le m} \Big\{ -2(\partial^\alpha u)\partial^\alpha(\mathscr{L}u + \mathscr{M}q) + |\partial^\alpha q|^2 + |\partial^\alpha u|^2 \Big\}\, dx,$$

$$\text{a.e. } t \in [0,T], \text{ a.s.} \tag{5.2}$$

In addition to (3.5), *if* $A(t,x)$ *is uniformly positive definite, then* (5.2) *can further be improved to the following*:

$$\int_{\mathbb{R}^n} \Big\{ \sum_{|\alpha|\le m+1} |\partial^\alpha u|^2 + \sum_{|\alpha|\le m} |\partial^\alpha q|^2 \Big\}\, dx$$

$$\le C \int_{\mathbb{R}^n} \sum_{|\alpha|\le m} \big\{ -2(\partial^\alpha u)\partial^\alpha(\mathscr{L}u + \mathscr{M}q) + |\partial^\alpha q|^2 + |\partial^\alpha u|^2 \big\}\, dx,$$

$$\text{a.e. } t \in [0,T], \text{ a.s.} \tag{5.3}$$

Lemma 5.2. *Let* $F \in H^m(\mathbb{R}^n)$ *and* $G \in H^m(\mathbb{R}^n)^n$, *such that*

$$(F,\varphi)_m = (G, D\varphi)_m, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^n) \ . \tag{5.4}$$

Then

$$(F,\varphi)_0 = (G, D\varphi)_0, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^n) \ .$$

Proof. Let $\mathscr{S} \overset{\Delta}{=} \mathscr{S}(\mathbb{R}^n)$ be the set of all $\varphi \in C^\infty(\mathbb{R}^n)$, such that

$$\Phi_{\alpha,\beta}(\varphi) \overset{\Delta}{=} \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta \varphi(x)| < \infty, \quad \forall \alpha, \beta \ .$$

Under the family of semi-norms $\Phi_{\alpha,\beta}$, $\mathscr{S}$ is a Fréchet space. Also, $C_0^\infty(\mathbb{R}^n)$ is a dense subset of $\mathscr{S}$. Thus, (5.4) holds for all $\varphi \in \mathscr{S}$. Next, it is known that (see [8, p. 161]) the Fourier transformation $\varphi \mapsto \hat{\varphi}$ is an isomorphism of $\mathscr{S}$ onto itself. Applying Parseval's formula to (5.4), we obtain

$$\int_{\mathbb{R}^n} [\hat{F}(\xi) - \langle \hat{G}(\xi), \xi\rangle] \left(\sum_{|\alpha|\le m} |\xi^\alpha|^2 \right) \hat{\varphi}(\xi)\, d\xi = 0, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^n) \ . \tag{5.5}$$

Now, for any $\psi \in C_0^\infty(\mathbb{R}^n)$, we have $\hat{\psi}(\xi)(\sum_{|\alpha|\le m} |\xi^\alpha|^2)^{-1} \in \mathscr{S}$. Thus, there exists a $\varphi \in \mathscr{S}$, such that $\hat{\varphi}(\xi) = \hat{\psi}(\xi)(\sum_{|\alpha|\le m} |\xi^\alpha|^2)^{-1}$. Combining this with (5.5) and using Parseval's formula again we obtain that $(F, \psi)_0 = (G, D\psi)_0$, for all $\psi \in C_0^\infty(\mathbb{R}^n)$, proving the lemma. □

Proof of Theorem 3.1. We start with the case of weak solution. Let $\{\varphi_k\}_{k\geq 1} \subset C_0^\infty(\mathbb{R}^n)$ be an orthonormal basis for the Hilbert space $H^m \equiv H^m(\mathbb{R}^n)$, and we denote the inner product in H^m by

$$(\varphi, \psi)_m \equiv \int_{\mathbb{R}^n} \sum_{|\alpha|\leq m} (\partial^\alpha \varphi)(\partial^\alpha \psi)\, dx, \quad \forall \varphi, \psi \in H^m \ . \tag{5.6}$$

Denote $|\varphi|_m = (\varphi, \varphi)_m^{1/2}$; and as a usual convention, $H^0 \equiv L^2(\mathbb{R}^n)$. When $q = (q_1, \ldots, q_d)$, $p = (p_1, \ldots, p_d) \in (H^m)^d$, we denote $(q, p)_m = \sum_{i=1}^d (q_i, p_i)_m$.

Let $k \geq 1$ be fixed. Consider the following linear BSDE:

$$\begin{cases} du^{kj} = -\left\{\sum_{i=1}^k \left[(\mathscr{L}\varphi_i, \varphi_j)_m u^{ki} + \langle(\mathscr{M}\varphi_i, \varphi_j)_m, q^{ki}\rangle\right] + (f, \varphi_j)_m\right\} dt \\ \qquad\qquad + \langle q^{kj}, dW(t)\rangle, \\ u^{kj}(T) = (g, \varphi_j)_m, \quad 1 \leq j \leq k \ . \end{cases}$$

It is by now standard (see e.g., [18]) that there exists a unique adapted solution $u^{kj}(\cdot) \in C_{\mathscr{F}}([0, T]; \mathbb{R})$, and $q^{kj}(\cdot) \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^d)$, $1 \leq j \leq k$. Let

$$u^k(t, x, \omega) = \sum_{j=1}^k u^{kj}(t, \omega)\varphi_j(x), \quad q^k(t, x, \omega) = \sum_{j=1}^k q^{kj}(t, \omega)\varphi_j(x) \ ,$$

for $(t, x, \omega) \in [0, T] \times \mathbb{R}^n \times \Omega$. Then we see that for any fixed $(t, \omega) \in [0, T] \times \Omega$, $u^k(t, \cdot, \omega) \in C_0^\infty(\mathbb{R}^n)$, and $q^k(t, \cdot, \omega) \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^d)$. Further, if we denote $P_k : H^m \to \mathrm{span}\{\varphi_1, \ldots, \varphi_k\} \stackrel{\Delta}{=} H_k^m$, to be the orthogonal projection in H^m, then obviously we have $(P^k)^* = P^k$, and $P_k u^k = u^k$, $k \geq 1$, since u^k (as well as $q_1^k, \ldots, q_d^k$ with $q^k = (q_1^k, \ldots, q_d^k)$) are H_k^m-valued processes. Now let $f^k = P_k f$ and $g^k = P_k g$, then clearly the following holds:

$$\begin{cases} du^k = \left\{-P_k[\mathscr{L}u^k + \mathscr{M}q^k] - f^k\right\} dt + \langle q^k, dW(t)\rangle, \\ u^k\big|_{t=T} = g^k \ . \end{cases} \tag{5.7}$$

We now derive the estimate for (u^k, q^k). By Lemma 5.1, we have

$$\begin{aligned} &\int_{\mathbb{R}^n} \Bigg\{ \sum_{|\alpha|\leq m} \{\langle(A - BB^T)D(\partial^\alpha u^k), D(\partial^\alpha u^k)\rangle. + |B^T D(\partial^\alpha u^k) + \partial^\alpha q^k|^2\} \\ &\qquad + \sum_{|\alpha|\leq m-1} |\partial^\alpha q^k|^2 \Bigg\} dx \\ &\qquad \leq C \int_{\mathbb{R}^n} \sum_{|\alpha|\leq m} \left\{ -2(\partial^\alpha u^k)\partial^\alpha(\mathscr{L}u^k + \mathscr{M}q^k) + |\partial^\alpha q^k|^2 + |\partial^\alpha u^k|^2 \right\} dx \ . \end{aligned} \tag{5.8}$$

On the other hand, applying Itô's formula to $|\partial^\alpha u^k|^2$ and using integration by parts we have from (5.7) that

$$
\begin{aligned}
&E\int_{\mathbb{R}^n}\sum_{|\alpha|\le m}\left\{|\partial^\alpha g^k(x)|^2-|\partial^\alpha u^k(t,x)|^2\right\}dx\\
&=E\int_t^T\int_{\mathbb{R}^n}\sum_{|\alpha|\le m}\left\{2(\partial^\alpha u^k)\partial^\alpha\left[P_k\left(-\mathscr{L}u^k-\mathscr{M}q^k\right)-f^k\right]+|\partial^\alpha q^k|^2\right\}dx\,ds\\
&=E\int_t^T\left\{-2(u^k,P_k(\mathscr{L}u^k+\mathscr{M}q^k)+f^k)_m+|q^k|_m^2\right\}ds\\
&=E\int_t^T\left\{-2(u^k,\mathscr{L}u^k+\mathscr{M}q^k+f^k)_m+|q^k|_m^2\right\}ds\\
&=E\int_t^T\int_{\mathbb{R}^n}\sum_{|\alpha|\le m}\Big\{-2(\partial^\alpha u^k)\partial^\alpha(\mathscr{L}u^k+\mathscr{M}q^k)\\
&\qquad\qquad+|\partial^\alpha q^k|^2-2(\partial^\alpha u^k)(\partial^\alpha f^k)\Big\}dx\,ds\\
&\ge\frac{1}{C}E\int_{\mathbb{R}^n}\Big\{\sum_{|\alpha|\le m}\left\{\langle(A-BB^T)D(\partial^\alpha u^k),D(\partial^\alpha u^k)\rangle+|B^{TD}(\partial^\alpha u^k)+\partial^\alpha q^k|^2\right\}\\
&\qquad+\sum_{|\alpha|\le m-1}|\partial^\alpha q^k|^2\Big\}-\sum_{|\alpha|\le m}\left\{|\partial^\alpha u^k|^2+2(\partial^\alpha u^k)(\partial^\alpha f^k)\right\}dx\,ds\ . \qquad (5.9)
\end{aligned}
$$

We remark here that for the first equality in the above we assumed without loss of generality that the expectation of the stochastic integral appearing in the Itô's formula is zero. In general one can always "localize" to such a case via a sequence of stopping times, and then pass to the limit, thanks to the Dominated Convergence Theorem. Moreover, the third equality in (5.9) is due to the fact that $(P^k)^*=P^k$ and $P^ku^k=u^k$. Combining (5.8) and (5.9), and modifying the constant C if necessary, we obtain

$$
\begin{aligned}
&\int_{\mathbb{R}^n}\Bigg\{\sum_{|\alpha|\le m}\{\langle (A-BB^T)D(\partial^\alpha u^k), D(\partial^\alpha u^k)\rangle \\
&\qquad +|B^T D(\partial^\alpha u^k)+\partial^\alpha q^k|^2\} + \sum_{|\alpha|\le m-1}|\partial^\alpha q^k|^2\Bigg\}\,dx \\
&\le C\Bigg\{E|g^k|_m^2 - E|u^k(t)|_m^2 + E\sum_{|\alpha|\le m}\int_t^T\int_{\mathbb{R}^n}\big[2(\partial^\alpha u^k)(\partial^\alpha f^k) \\
&\qquad + |\partial^\alpha u^k|^2\big]\,dx\,ds\Bigg\} \\
&\le C\Bigg\{E|g^k|_m^2 - E|u^k(t)|_m^2 + \int_t^T E\big[|u^k(s)|_m^2 + |f^k(s)|_m^2\big]\,ds\Bigg\}\ . \qquad (5.10)
\end{aligned}
$$

By Gronwall's inequality, we have

$$
\begin{aligned}
&\max_{t\in[0,T]} E|u^k(t)|_m^2 + E\int_0^T |q^k(t)|_{m-1}^2\,dt \\
&+\sum_{|\alpha|\le m} E\times\int_0^T\int_{\mathbb{R}^n}\{\langle (A-BB^T)D(\partial^\alpha u^k), D(\partial^\alpha u^k)\rangle \\
&+\left|B^T[D(\partial^\alpha u^k)]+\partial^\alpha q^k\right|^2\}\,dx\,dt \\
&\quad\le CE\Bigg\{\int_0^T |f^k(t)|_m^2\,dt + |g^k|_m^2\Bigg\} \le CE\Bigg\{\int_0^T |f(t)|_m^2\,dt + |g|_m^2\Bigg\}\ . \qquad (5.11)
\end{aligned}
$$

Note that the constant $C>0$ in (5.11) only depends on T, m and K_m, and the right side is independent of k, so we may conclude that

$$
\begin{cases} u^k \to u, & \text{weak}^* \text{ in } L^\infty_{\mathscr{F}}(0,T;L^2(\Omega;H^\ell)), \quad 0\le \ell\le m, \\ q^k\to q, & \text{weakly in } L^2_{\mathscr{F}}(0,T;H^\ell)^d, \quad 0\le\ell\le m-1\ , \end{cases} \qquad (5.12)
$$

and for any $|\alpha|\le m$,

$$
\begin{cases} (A-BB^T)^{1/2}D(\partial^\alpha u^k)\to (A-BB^T)^{1/2}D(\partial^\alpha u), \\ B^T[D(\partial^\alpha u^k)]+\partial^\alpha q^k \to B^T[D(\partial^\alpha u)]+\partial^\alpha q, \quad \text{weakly in } L^2_{\mathscr{F}}(0,T;H^0)\ . \end{cases}
$$

Now by Mazur's theorem, (u,q) is in fact a strong limit of certain sequence of the convex combinations of (u^k,q^k)'s. Note again that the right side of (5.11) is independent of k, we can take limit in (5.11) (along the sequence of the convex combinations if necessary) to show that (u,q) also satisfies the estimate (5.11), or in other words, (3.4). It remains to prove that (u, q) is a weak solution of (4.2). To this end, let us take $\rho \in H^1(0, T)$ such that $\rho(0) = 0$, $\rho(T) = 1$, $0 \le \rho(t) \le 1$, for all $t \in [0, T]$. Next, let $\ell > 0$ be fixed and $k \ge \ell$. For any $\varphi \in H_\ell^m \subset C_0^\infty(\mathbb{R}^n)$, from (5.7) and the fact $P_k\varphi = \varphi$, we have

$$(g^k, \varphi)_m = \int_0^T \{\dot{\rho}(t)(u^k(t), \varphi)_m - \rho(t)(\mathscr{L}u^k(t) + \mathscr{M}q^k(t) + f^k(t), \varphi)_m\}\, dt + \int_0^T \rho(t)\langle (q^k(t), \varphi)_m,\, dW(t)\rangle \ .$$

By the definition of $\mathscr{L}$ and $\mathscr{M}$, using integration by parts, we obtain

$$\begin{aligned}(g^k, \varphi)_m = \int_0^T &\{\dot{\rho}(t)(u^k(t), \varphi)_m - \rho(t)\big[-\tfrac{1}{2}(A(t)Du^k(t) + B(t)q^k(t), D\varphi)_m \\ &+ (\langle a(t), Du^k(t)\rangle + c(t)u^k(t) + \langle b(t), q^k(t)\rangle + f^k(t), \varphi)_m\big]\}\, dt \\ &+ \int_0^T \rho(t)\langle (q^k(t), \varphi)_m,\, dW(t)\rangle, \qquad \text{a.s.} \end{aligned} \tag{5.13}$$

If we denote

$$\begin{cases} F(x, \omega) = g^k - \displaystyle\int_0^T \{\dot{\rho}(t)u^k(t) - \rho(t)\big[\langle a(t), Du^k(t)\rangle + c(t)u^k(t) \\ \qquad\qquad + \langle b(t), q^k(t)\rangle + f^k(t)\big]\}\, dt - \displaystyle\int_0^T \langle \rho(t)q^k(t),\, dW(t)\rangle, \\ G(x, \omega) = \displaystyle\int_0^T \rho(t)[\tfrac{1}{2}A(t)Du^k(t) + B(t)q^k(t)]\, dt \ , \end{cases}$$

then (5.13) reads $(F, \varphi)_m = (G, D\varphi)_m$, for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, a.s.. So by Lemma 5.2 we must have $(F, \varphi)_0 = (G, D\varphi)_0$, for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, a.s., which means

$$(g^k, \varphi)_0 = \int_0^T \{\dot{\rho}(t)(u^k(t), \varphi)_0 - \rho(t)\langle \mathscr{L}u^k(t) + \mathscr{M}q^k(t) + f^k(t), \varphi\rangle_0\}\, dt + \int_0^T \rho(t)\langle (q^k(t), \varphi)_0,\, dW(t)\rangle, \qquad \text{a.s.} \tag{5.14}$$

We are now going to let $k \to \infty$ in (5.14) in order to obtain the similar equality for (u, q). Note that by (5.12) with $\ell = 1$ for u^k and $\ell = 0$ for q^k, together with the convergence of (f^k, g^k) to (f, g), we can pass to the limit in (5.14) weakly in $L^2(\Omega)$ for all terms except the last term which involves the Itô integral. To treat this last term, we define $K : L^2_{\mathscr{F}}(0, T; H^0)^d \to L^2(\Omega)$ by

$$Kp = \int_0^T \rho(t)\langle (p(t), \varphi)_0, dW(t)\rangle, \quad \forall p \in L^2_{\mathscr{F}}(0,T;H^0)^d \ .$$

Then it is easy to check that K is a bounded linear operator. Thus, for any $\eta \in L^2(\Omega)$, one has

$$\begin{aligned} E\bigg(\eta \int_0^T \rho(t)\langle (q^k(t) - q(t), \varphi)_0, dW(t)\rangle\bigg) &= (\eta, K(q^k - q))_{L^2(\Omega)} \\ &= (K^*\eta, q^k - q)_{L^2_{\mathscr{F}}(0,T;H^0)^d} \to 0 \ . \end{aligned}$$

Hence, it follows from (5.14) that

$$\begin{aligned} (g,\varphi)_0 = &\int_0^T \{\dot\rho(t)(u(t),\varphi)_0 - \rho(t)\langle \mathscr{L}u(t) + \mathscr{M}q(t) + f(t), \varphi\rangle_0\}\, dt \\ &+ \int_0^T \rho(t)\langle (q(t),\varphi)_0, dW(t)\rangle, \qquad \text{a.s.} \end{aligned} \tag{5.15}$$

Now, fixed any $t \in (0,T)$. For $\varepsilon > 0$ small enough we let

$$\rho_\varepsilon(s) = \begin{cases} 0, & s \le t - \varepsilon/2, \\ \frac{1}{2} + \frac{s-t}{\varepsilon}, & t - \varepsilon/2 < s < t + \varepsilon/2, \\ 1, & s \ge t + \varepsilon/2 \ . \end{cases}$$

Replacing $\rho = \rho_\varepsilon$ in (5.15) and letting $\varepsilon \to 0$, we obtain that, for all $\varphi \in C_0^\infty(\mathbb{R}^n)$,

$$\begin{aligned} (g,\varphi)_0 = &(u(t),\varphi)_0 - \int_t^T \langle \mathscr{L}u(t) + \mathscr{M}q(t) + f(t), \varphi\rangle_0\, dt \\ &+ \int_t^T \langle (q(t),\varphi)_0, dW(t)\rangle \ , \end{aligned}$$

almost surely. This means that (u,q) is an adapted weak solution of (4.2).

Now suppose that $m \ge 2$. From (3.4) we see that (2.11) holds and thus, by Proposition 2.3, (u,q) becomes an adapted strong solution. If one has further that $m > 2 + n/2$, then by Sobolev's embedding theorem, (2.9) holds and therefore (u,q) becomes an adapted classical solution, thanks again to Proposition 2.3. The proof is now complete. □

Proof of Theorem 3.2. We assume now that (3.5) or (3.6) holds instead of (3.2). Since (5.1) still holds by Lemma 5.1, all the proof that we have presented above remains true. Further, if (3.5) holds, then we have (5.2), which leads to estimate (3.7); in addition, if $A(t,x)$ is uniformly positive definite, a little more computation would yield (3.8). In fact, in this case we can use integration by parts in the last step of (5.9) to get

$$E\sum_{|\alpha|\le m}\int_t^T\int_{\mathbb{R}^n} 2(\partial^\alpha u^k)(\partial^\alpha f^k)\,dx\,ds \le CE\int_t^T \left\{|u^k(s)|_{m+1}^2 + |f^k(s)|_{m-1}^2\right\} ds \ .$$

We leave the details of the proof to the interested readers. □

6. Proof of Lemma 5.1

We now proof the fundamental lemma of this paper: Lemma 5.1. Let $\ell \overset{\Delta}{=} |\alpha| \le m$. For any $u \in C_0^\infty(\mathbb{R}^n)$ and $q \in C_0^\infty(\mathbb{R}^n;\mathbb{R}^d)$, by definition of $\mathscr{L}$ and $\mathscr{M}$ (see (4.1)), and differentiation, we have

$$\begin{aligned}
\mathscr{I}^\alpha &\overset{\Delta}{=} \int_{\mathbb{R}^n}\left\{-2(\partial^\alpha u)\partial^\alpha(\mathscr{L}u+\mathscr{M}q)+|\partial^\alpha q|^2\right\}dx\\
&= \int_{\mathbb{R}^n}\Big\{-2(\partial^\alpha u)\partial^\alpha\Big[\tfrac{1}{2}\nabla\cdot(ADu)+\langle a,Du\rangle+cu\\
&\qquad\qquad + \nabla\cdot[Bq]+\langle b,q\rangle\Big]+|\partial^\alpha q|^2\Big\}dx\\
&= \int_{\mathbb{R}^n}\Big\{-2(\partial^\alpha u)\Big[\tfrac{1}{2}\nabla\cdot[AD(\partial^\alpha u)]+\langle a,D(\partial^\alpha u)\rangle\\
&\qquad + c(\partial^\alpha u)+\nabla\cdot[B(\partial^\alpha q)]+\langle b,\partial^\alpha q\rangle\Big]+|\partial^\alpha q|^2-2(\partial^\alpha u)\\
&\qquad \times\sum_{0\le\beta<\alpha} C_{\alpha\beta}\Big[\tfrac{1}{2}\nabla\cdot[(\partial^{\alpha-\beta}A)D(\partial^\beta u)]+\langle\partial^{\alpha-\beta}a,D(\partial^\beta u)\rangle\\
&\qquad +(\partial^{\alpha-\beta}c)(\partial^\beta u)+\nabla\cdot[(\partial^{\alpha-\beta}B)(\partial^\beta q)]+\langle\partial^{\alpha-\beta}b,\partial^\beta q\rangle\Big]\Big\}dx\\
&\equiv \mathscr{I}_0^\alpha+\mathscr{I}_1^\alpha+\mathscr{I}_2^\alpha+\mathscr{I}_3^\alpha\ ,
\end{aligned}$$

where $C_{\alpha\beta}$ is a positive integer depending on α and β, and

$$\begin{cases}
\mathscr{I}_0^\alpha = \int_{\mathbb{R}^n}\Big\{(-2\partial^\alpha u)\Big[\tfrac{1}{2}\nabla\cdot[AD(\partial^\alpha u)]+\langle a,D(\partial^\alpha u)\rangle+c(\partial^\alpha u)\\
\qquad\qquad +\nabla\cdot[B(\partial^\alpha q)]+\langle b,\partial^\alpha q\rangle\Big]+|\partial^\alpha q|^2\Big\}dx,\\
\mathscr{I}_1^\alpha = -2\int_{\mathbb{R}^n}\sum_{0\le\beta<\alpha}C_{\alpha\beta}(\partial^\alpha u)\Big[\tfrac{1}{2}\nabla\cdot[(\partial^{\alpha-\beta}A)D(\partial^\beta u)]\\
\qquad\qquad +\langle\partial^{\alpha-\beta}a,D(\partial^\beta u)\rangle+(\partial^{\alpha-\beta}c)(\partial^\beta u)+\langle\partial^{\alpha-\beta}b,\partial^\beta q\rangle\Big]dx,\\
\mathscr{I}_2^\alpha = -2\int_{\mathbb{R}^n}\sum_{\substack{0\le\beta<\alpha\\|\beta|<|\alpha|-1}}C_{\alpha\beta}(\partial^\alpha u)\,\nabla\cdot[(\partial^{\alpha-\beta}B)(\partial^\beta q)]\,dx,\\
\mathscr{I}_3^\alpha = -2\int_{\mathbb{R}^n}\sum_{\substack{0\le\beta<\alpha\\|\beta|=|\alpha|-1}}C_{\alpha\beta}(\partial^\alpha u)\,\nabla\cdot[(\partial^{\alpha-\beta}B)(\partial^\beta q)]\,dx\ .
\end{cases}\tag{6.5}$$

We note that in the case $\ell = 0$, $\mathscr{I}_1^\alpha$, $\mathscr{I}_2^\alpha$ and $\mathscr{I}_3^\alpha$ are all absent. To estimate $\mathscr{I}_i^\alpha$, $i = 0, 1, 2, 3$, we first note that A and B are C_b^{m+1} in x, which immediately yields that

$$|\mathscr{I}_1^\alpha| + |\mathscr{I}_2^\alpha| \le C\big(|u|_\ell^2 + |q|_{\ell-1}^2\big) \ . \tag{6.6}$$

Furthermore, using integration by parts, we have

$$\begin{aligned}
\mathscr{I}_0^\alpha &= \int_{\mathbb{R}^n} \Big\{ \langle AD(\partial^\alpha u), D(\partial^\alpha u)\rangle + 2\langle \partial^\alpha q, B^T D(\partial^\alpha u)\rangle + |\partial^\alpha q|^2 \\
&\qquad - \langle a, D[(\partial^\alpha u)^2]\rangle - 2c(\partial^\alpha u)^2 - 2\langle b(\partial^\alpha u), \partial^\alpha q\rangle \Big\}\, dx \\
&= \int_{\mathbb{R}^n} \Big\{ \langle (A - BB^T) D(\partial^\alpha u), D(\partial^\alpha u)\rangle + |B^T D(\partial^\alpha u)|^2 \\
&\qquad + |\partial^\alpha q|^2 + 2\langle \partial^\alpha q, B^T D(\partial^\alpha u)\rangle - 2\langle b(\partial^\alpha u), \partial^\alpha q\rangle \\
&\qquad - 2\langle B^T D(\partial^\alpha u), b(\partial^\alpha u)\rangle + [\nabla\cdot(a - Bb) - 2c](\partial^\alpha u)^2 \Big\}\, dx \ .
\end{aligned} \tag{6.7}$$

Also, we note that for $\beta < \alpha$ with $|\beta| = |\alpha| - 1$, doing integration by parts again we have

$$\begin{aligned}
&-\int_{\mathbb{R}^n} (\partial^\alpha u)\, \nabla\cdot[(\partial^{\alpha-\beta}B)(\partial^\beta q)]\, dx = \int_{\mathbb{R}^n} (\partial^\beta u)\, \nabla\cdot\{\partial^{\alpha-\beta}[(\partial^{\alpha-\beta}B)(\partial^\beta q)]\}\, dx \\
&\quad = -\int_{\mathbb{R}^n} \langle D(\partial^\beta u), (\partial^{\alpha-\beta}B)\partial^\alpha q + (\partial^{2(\alpha-\beta)}B)\partial^\beta q\rangle\, dx \\
&\quad = -\int_{\mathbb{R}^n} \{\langle (\partial^{\alpha-\beta}B^T) D(\partial^\beta u), \partial^\alpha q\rangle + \langle D(\partial^\beta u), (\partial^{2(\alpha-\beta)}B)\partial^\beta q)\rangle\}\, dx \ .
\end{aligned} \tag{6.8}$$

Thus, it follows that

$$\begin{aligned}
\mathscr{I}_0^\alpha + \mathscr{I}_3^\alpha = \int_{\mathbb{R}^n} \Bigg\{ & \langle (A - BB^T) D(\partial^\alpha u), D(\partial^\alpha u)\rangle \\
&+ \Bigg| B^T D(\partial^\alpha u) + \partial^\alpha q - b(\partial^\alpha u) - \sum_{\substack{0\le\beta<\alpha \\ |\beta|=|\alpha|-1}} C_{\alpha\beta}(\partial^{\alpha-\beta}B^T) D(\partial^\beta u) \Bigg|^2 \\
&- |b(\partial^\alpha u)|^2 - \Bigg| \sum_{\substack{0\le\beta<\alpha \\ |\beta|=|\alpha|-1}} C_{\alpha\beta}(\partial^{\alpha-\beta}B^T) D(\partial^\beta u) \Bigg|^2 \\
&+ 2 \sum_{\substack{0\le\beta<\alpha \\ |\beta|=|\alpha|-1}} C_{\alpha\beta}\langle (\partial^{\alpha-\beta}B^T) D(\partial^\beta u), B^{TD}(\partial^\alpha u)\rangle
\end{aligned}$$

$$
\begin{aligned}
&-2\sum_{\substack{0\le\beta<\alpha\\|\beta|=|\alpha|-1}} C_{\alpha\beta}\langle(\partial^{\alpha-\beta}B^T)D(\partial^\beta u), b(\partial^\alpha u)\rangle\\
&\quad+[\nabla\cdot(a-Bb)-2c](\partial^\alpha u)^2\\
&-2\sum_{\substack{0\le\beta<\alpha\\|\beta|=|\alpha|-1}} C_{\alpha\beta}\langle D(\partial^\beta u),(\partial^{2(\alpha-\beta)}B)(\partial^\beta q)\rangle\Bigg\}\,dx\ .
\end{aligned}
\tag{6.9}
$$

Note that

$$
\begin{aligned}
&\int_{\mathbb{R}^n}\Bigg|B^TD(\partial^\alpha u)+\partial^\alpha q-b(\partial^\alpha u)-\sum_{\substack{0\le\beta<\alpha\\|\beta|=|\alpha|-1}}C_{\alpha\beta}(\partial^{\alpha-\beta}B^T)D(\partial^\beta u)\Bigg|^2\,dx\\
&\ge\tfrac12\int_{\mathbb{R}^n}\left|B^TD(\partial^\alpha u)+\partial^\alpha q\right|^2dx\\
&\quad-\tfrac12\int_{\mathbb{R}^n}\Bigg|b(\partial^\alpha u)+\sum_{\substack{0\le\beta<\alpha\\|\beta|=|\alpha|-1}}C_{\alpha\beta}(\partial^{\alpha-\beta}B^T)D(\partial^\beta u)\Bigg|^2\,dx\\
&\ge\tfrac12\int_{\mathbb{R}^n}\left|B^TD(\partial^\alpha u)+\partial^\alpha q\right|^2dx-C|u|_\ell^2\ .
\end{aligned}
\tag{6.10}
$$

Combining (6.6), (6.9)–(6.10) yields

$$
\begin{aligned}
\mathscr{I}^\alpha=\mathscr{I}_0^\alpha+\mathscr{I}_1^\alpha+\mathscr{I}_2^\alpha+\mathscr{I}_3^\alpha\ge&\int_{\mathbb{R}^n}\Big\{\langle(A-BB^T)D(\partial^\alpha u),D(\partial^\alpha u)\rangle.\\
&+\tfrac12\,|B^TD(\partial^\alpha u)+\partial^\alpha q|^2\Big\}\,dx\ -C(|u|_\ell^2+|q|_{\ell-1}^2)\\
&+2\int_{\mathbb{R}^n}\sum_{\substack{0\le\beta<alpha\\|\beta|=|\alpha|-1}}C_{\alpha\beta}\langle(\partial^{\alpha-\beta}B^T)D(\partial^\beta u),B^TD(\partial^\alpha u)\rangle\,dx\ .
\end{aligned}
\tag{6.11}
$$

To estimate the last term in the above, we use the symmetry condition (3.2). For $\beta<\alpha$, $|\beta|=|\alpha|-1$, we have

$$
\begin{aligned}
&\int_{\mathbb{R}^n}\langle(\partial^{\alpha-\beta}B^T)D(\partial^\beta u),B^TD(\partial^\alpha u)\rangle\,dx=\int_{\mathbb{R}^n}\langle(B\partial^{\alpha-\beta}B^T)D(\partial^\beta u),D(\partial^\alpha u)\rangle\,dx\\
&\quad=\int_{\mathbb{R}^n}\tfrac12\big[\partial^{\alpha-\beta}\langle(B\partial^{\alpha-\beta}B^T)D(\partial^\beta u),D(\partial^\beta u)\rangle\\
&\qquad\qquad-\langle\partial^{\alpha-\beta}(B\partial^{\alpha-\beta}B^T)D(\partial^\beta u),D(\partial^\beta u)\rangle\big]\,dx\\
&\quad=-\tfrac12\textstyle\int_{\mathbb{R}^n}\langle\partial^{\alpha-\beta}(B\partial^{\alpha-\beta}B^T)D(\partial^\beta u),D(\partial^\beta u)\rangle\,dx\ge-C|u|_\ell^2\ .
\end{aligned}
\tag{6.12}
$$

Then (6.11) becomes

$$
\begin{aligned}
\mathscr{I}^\alpha&=\mathscr{I}_0^\alpha+\mathscr{I}_1^\alpha+\mathscr{I}_2^\alpha+\mathscr{I}_3^\alpha\\
&\ge\int_{\mathbb{R}^n}\Big\{\langle(A-BB^T)D(\partial^\alpha u),D(\partial^\alpha u)\rangle\\
&\quad+\frac12|B^TD(\partial^\alpha u)+\partial^\alpha q|^2\Big\}\,dx\ -C(|u|_\ell^2+|q|_{\ell-1}^2)\ .
\end{aligned}
\tag{6.13}
$$

Now, we sum (6.13) up for all $|\alpha| \le \ell$ to get the following:

$$\Psi_\ell \overset{\Delta}{=} \sum_{|\alpha|\le\ell} \int_{\mathbb{R}^n} \Big\{ -2(\partial^\alpha u)\partial^\alpha(\mathscr{L}u + \mathscr{M}q) + |\partial^\alpha q|^2 \Big\}\, dx$$
$$\ge \frac{1}{2} \sum_{|\alpha|\le\ell} \int_{\mathbb{R}^n} \Big\{ \langle (A - BB^T) D(\partial^\alpha u), D(\partial^\alpha u)\rangle + |B^T D(\partial^\alpha u) + \partial^\alpha q|^2 \Big\}\, dx$$
$$- C(|u|_\ell^2 + |q|_{\ell-1}^2) \ . \tag{6.14}$$

Thus, it follows that

$$\Phi_\ell \overset{\Delta}{=} \sum_{|\alpha|\le\ell} \int_{\mathbb{R}^n} \Big\{ \langle (A - BB^T) D(\partial^\alpha u), D(\partial^\alpha u)\rangle + |B^T D(\partial^\alpha u) + \partial^\alpha q|^2 \Big\}\, dx$$
$$\le C\Big(\Psi_\ell + |u|_\ell^2 + |q|_{\ell-1}^2 \Big) \ . \tag{6.15}$$

Note that

$$|\partial^\alpha q|^2 \le 2|B^T D(\partial^\alpha u) + \partial^\alpha q|^2 + 2|B^T D(\partial^\alpha u)|^2 \ . \tag{6.16}$$

Using the parabolicity condition (1.2) and the definition of Φ_ℓ (see (6.15)), we have

$$|q|_{\ell-1}^2 \le C(\Phi_{\ell-1} + |u|_\ell^2) \ . \tag{6.17}$$

Consequently, from (6.15) and (6.17), we obtain

$$\Phi_\ell \le C(\Psi_\ell + \Phi_{\ell-1} + |u|_\ell^2), \quad 1 \le \ell \le m \ . \tag{6.18}$$

On the other hand, for $\ell = 0$ (i.e., $\alpha = 0$), we have

$$\int_{\mathbb{R}^n} \Big\{ -2u(\mathscr{L}u + \mathscr{M}q) + |q|^2 \Big\}\, dx$$
$$= \int_{\mathbb{R}^n} \Big\{ -2u\Big[\frac{1}{2}\nabla\cdot[ADu] + \langle a, Du\rangle + cu + \nabla\cdot[Bq] + \langle b, q\rangle \Big] + |q|^2 \Big\}\, dx$$
$$= \int_{\mathbb{R}^n} \Big\{ \langle (A - BB^T) Du, Du\rangle + |B^T Du|^2 + |q|^2 + 2\langle q, B^T Du\rangle - 2\langle bu, q\rangle$$
$$- 2\langle B^T Du, bu\rangle + [\nabla\cdot(a + Bb) - 2c]u^2 \Big\}\, dx$$
$$\ge \int_{\mathbb{R}^n} \{ \langle (A - BB^T) Du, Du\rangle + \big|B^T Du + q - bu\big|^2 \}\, dx - C|u|_0^2$$
$$\ge \frac{1}{2} \int_{\mathbb{R}^n} \{ \langle (A - BB^T) Du, Du\rangle + \big|B^T Du + q\big|^2 \}\, dx - C|u|_0^2 \ . \tag{6.19}$$

This implies

$$\Phi_0 \le C\Big\{ \int_{\mathbb{R}^n} \{ -2u(\mathscr{L}u + \mathscr{M}q) + |q|^2 \}\, dx + C|u|_0^2 \Big\} \ . \tag{6.20}$$

Hence, it follows from (6.18) and (6.20) that

$$\Phi_m \le C(\Psi_m + |u|_m^2) \ , \tag{6.21}$$

which is the same as (5.1).

In the case that (3.5) holds (and (3.2) is not assumed), we need to replace (6.12) by another proper estimate. To this end, we use the following estimate:

$$\begin{aligned} &\int_{\mathbb{R}^n} \langle (\partial^{\alpha-\beta} B^T) D(\partial^\beta u), B^T D(\partial^\alpha u)\rangle \, dx \\ &\ge -\varepsilon \int_{\mathbb{R}^n} |B^T D(\partial^\alpha u)|^2 \, dx - C \int_{\mathbb{R}^n} |D(\partial^\beta u)|^2 \, dx \ , \end{aligned} \tag{6.22}$$

for small enough $\varepsilon > 0$ to get

$$\begin{aligned} &\int_{\mathbb{R}^n} \big\{ \langle (A - BB^T) D(\partial^\alpha u), D(\partial^\alpha u)\rangle \\ &\quad + 2 \sum_{\substack{0 \le \beta < \alpha \\ |\beta| = |\alpha| - 1}} C_{\alpha\beta} \langle (\partial^{\alpha-\beta} B^T) D(\partial^\beta u), B^T D(\partial^\alpha u)\rangle \\ &\ge \frac{\varepsilon_0}{2} \int_{\mathbb{R}^n} \langle (A - BB^T) D(\partial^\alpha u), D(\partial^\alpha u)\rangle \, dx - C|u|_\ell^2 \ . \end{aligned} \tag{6.23}$$

Then, we still have (6.15) and finally have (6.21) which is the same as (5.1).

In the case (3.6) holds (without assuming (3.2)), we use the following estimate: (note $\beta < \alpha$)

$$\begin{aligned} &\int_{\mathbb{R}^n} \langle (\partial^{\alpha-\beta} B^T) D(\partial^\beta u), B^T D(\partial^\alpha u)\rangle \, dx \\ &= \int_{\mathbb{R}^n} \langle (\partial^{\alpha-\beta} B^T) D(\partial^\beta u), \partial^{\alpha-\beta}[B^T D(\partial^\beta u)] - (\partial^{\alpha-\beta} B^T) D(\partial^\beta u)\rangle \, dx \\ &= -\int_{\mathbb{R}^n} \big\{ \langle (\partial^{2(\alpha-\beta)} B^T) D(\partial^\beta u), B^T D(\partial^\beta u)\rangle \\ &\qquad + \langle (\partial^{(\alpha-\beta)} B^T) D(\partial^\alpha u), B^T D(\partial^\beta u)\rangle + |(\partial^{\alpha-\beta} B^T) D(\partial^\beta u)|^2 \big\} \, dx \\ &\ge -\varepsilon \int_{\mathbb{R}^n} \big| (\partial^{\alpha-\beta} B^T) D(\partial^\alpha u) \big|^2 \, dx - C|u|_\ell^2 \ , \end{aligned} \tag{6.24}$$

for $\varepsilon > 0$ small enough to obtain (6.23) and finally to obtain (6.21).

Note that in the case (3.5) holds, we have (3.8). Then, (5.2) follows from (5.1) easily. Finally, if in addition, (2.7) also holds, then, (5.3) follows from (5.2). This completes the proof of Lemma 5.1. □

7. Comparison theorems

In this section we present some comparison theorems for the adapted solutions of different BSPDEs. For convenience, we consider BSPDEs of form (2.3), for which we have the well-posedness under proper conditions (see the discussion at the end of §3). Let us denote (compare (4.1))

$$\begin{cases} \mathscr{L}u \stackrel{\Delta}{=} \frac{1}{2}\mathrm{tr}[AD^2u] - \langle a, Du\rangle - cu, & \mathscr{M}q \stackrel{\Delta}{=} \mathrm{tr}[B^T Dq] - \langle b, q\rangle, \\ \bar{\mathscr{L}}u \stackrel{\Delta}{=} \frac{1}{2}\mathrm{tr}[\bar{A}D^2\bar{u}] - \langle \bar{a}, D\bar{u}\rangle - \bar{c}\bar{u}, & \bar{\mathscr{M}}q \stackrel{\Delta}{=} \mathrm{tr}[\bar{B}^T D\bar{q}] - \langle \bar{b}, \bar{q}\rangle \ . \end{cases} \tag{7.1}$$

Consider the following BSPDEs:

$$\begin{cases} du = -\{\mathscr{L}u + \mathscr{M}q + f\}\,dt + \langle q, dW(t)\rangle, \quad (t,x) \in [0,T] \times \mathbb{R}^n, \\ u\big|_{t=T} = g \ . \end{cases} \tag{7.2}$$

$$\begin{cases} d\bar{u} = -\{\bar{\mathscr{L}}\bar{u} + \bar{\mathscr{M}}\bar{q} + \bar{f}\}\,dt + \langle \bar{q}, dW(t)\rangle, \quad (t,x) \in [0,T] \times \mathbb{R}^n, \\ \bar{u}\big|_{t=T} = \bar{g} \ . \end{cases} \tag{7.3}$$

Note that (7.2) and (4.2) are a little different in that the latter is in a *divergence form*, while the former is in a *general form*. Throughout this section, we assume that the parabolicity condition (1.2), the symmetry condition (3.2) and $(\mathrm{H})_m$ (for some $m \geq 2$) hold for $\{A, B, a, b, c\}$, and $\{\bar{A}, \bar{B}, \bar{a}, \bar{b}, \bar{c}\}$. Then by the discussion at the end of §3, we know that Theorems 3.1–3.3 also hold for (7.2) and (7.3). Namely, for any pairs (f, g) and $(\bar{f}, \bar{g})$ satisfying (3.12), there exist unique adapted strong solutions (u, q) and $(\bar{u}, \bar{q})$ to (7.2) and (7.3), respectively, satisfying estimates of form (3.13). We hope to establish some comparisons between u and $\bar{u}$ in various cases.

Our main result is the following theorem. For $a \in \mathbb{R}$, we denote $a^- = -(a \wedge 0)$, as usual.

Theorem 7.1. *Let* (1.2), (3.2) *and* $(H)_m$ *hold for* (7.2) *and* (7.3), *for some* $m \geq 2$. *Let* (f, g) *and* $(\bar{f}, \bar{g})$ *satisfy* (3.12) *with some* $\lambda \geq 0$. *Let* (u, q) *and* $(\bar{u}, \bar{q})$ *be adapted strong solutions of* (7.2) *and* (7.3), *respectively. Then for some* $\mu > 0$,

$$\begin{aligned} &E\int_{\mathbb{R}^n} e^{-\lambda\langle x\rangle}\Big\{[u(t,x) - \bar{u}(t,x)]^-\Big\}^2 dx \\ &\quad \leq e^{\mu(T-t)}E\int_{\mathbb{R}^n} e^{-\lambda\langle x\rangle}\Big\{[g(x) - \bar{g}(x)]^-\Big\}^2 dx + E\int_t^T e^{\mu(s-t)} \\ &\qquad \times \int_{\mathbb{R}^n} e^{-\lambda\langle x\rangle}\Big\{[(\mathscr{L} - \bar{\mathscr{L}})\bar{u}(s,x) + (\mathscr{M} - \bar{\mathscr{M}})\bar{q}(s,x) \\ &\qquad\quad + f(s,x) - \bar{f}(s,x)]^-\Big\}^2 dx\,ds, \ \forall t \in [0,T] \ , \end{aligned} \tag{7.4}$$

Proof. We first prove the case when $\bar{f} = \bar{g} = 0$ (consequently $\bar{u} = \bar{q} = 0$ by linearity). Also, we assume that (f, g) satisfies (3.12) with $\lambda = 0$. Define a function $\varphi : \mathbb{R} \to [0, \infty)$ as follows:

$$\varphi(r) = \begin{cases} r^2, & r \le -1, \\ (6r^3 + 8r^4 + 3r^5)^2, & -1 \le r \le 0, \\ 0, & r \ge 0 \ . \end{cases}$$

We can directly check that φ is C^2 and $\varphi(0) = \varphi'(0) = \varphi''(0) = 0$, $\varphi(-1) = 1$, $\varphi'(-1) = -2$, and $\varphi''(-1) = 2$. Next, for any $\varepsilon > 0$, we let $\varphi_\varepsilon(r) = \varepsilon^2 \varphi(\frac{r}{\varepsilon})$. One shows that

$$\begin{cases} \lim_{\varepsilon \to 0} \varphi_\varepsilon(r) = [r^-]^2, \quad \lim_{\varepsilon \to 0} \varphi'_\varepsilon(r) = -2r^-, \quad \text{uniformly;} \\ |\varphi''_\varepsilon(r)| \le C, \quad \forall \varepsilon > 0, \, r \in \mathbb{R}; \qquad \lim_{\varepsilon \to 0} \varphi''_\varepsilon(r) = \begin{cases} 2, & r < 0, \\ 0, & r > 0 \ . \end{cases} \end{cases}$$

Denote $\hat{a} = a - \frac{1}{2} \nabla \cdot A$ and $\hat{b} = b - \nabla \cdot B$, we have by (2.2) that

$$\begin{cases} \frac{1}{2} \operatorname{tr}[AD^2 u] + \langle a, Du \rangle = \frac{1}{2} \nabla \cdot [ADu] + \langle \hat{a}, Du \rangle, \\ \operatorname{tr}[B^T Dq] + \langle b, q \rangle = \nabla \cdot (Bq) + \langle \hat{b}, q \rangle \ . \end{cases}$$

Applying the Itô's formula to $\varphi_\varepsilon(u)$, and then integrating with respect to x, we obtain that (let $Q_t = [t, T] \times \mathbb{R}^n$)

$$\begin{aligned} & E \int_{\mathbb{R}^n} \varphi_\varepsilon(g(x))\, dx - E \int_{\mathbb{R}^n} \varphi_\varepsilon(u(t, x))\, dx \\ &= E \int_{Q_t} \Big\{ \varphi'_\varepsilon(u) \big[-\frac{1}{2} \nabla \cdot (ADu) - \nabla \cdot (Bq) - \langle \hat{a}, Du \rangle \\ &\qquad\qquad - cu - \langle \hat{b}, q \rangle - f \big] + \frac{1}{2} \varphi''_\varepsilon(u) |q|^2 \Big\}\, dx\, ds \\ &= E \int_{Q_t} \Big\{ \frac{1}{2} \varphi''_\varepsilon(u) \big[\langle ADu, Du \rangle + 2 \langle B^{TDu}, q \rangle + |q|^2 \big] \\ &\qquad\qquad - \varphi'_\varepsilon(u) \big[\langle \hat{a}, Du \rangle + cu + \langle \hat{b}, q \rangle + f \big] \Big\}\, dx\, ds \\ &= E \int_{Q_t} \Big\{ \frac{1}{2} \varphi''_\varepsilon(u) \big[\langle (A - BB^T) Du, Du \rangle + |B^T Du + q - \hat{b} u|^2 \big] \\ &\qquad\qquad + \frac{1}{2} \varphi''_\varepsilon(u) \big[-|\hat{b}|^2 u^2 + 2 \langle B^{TDu}, \hat{b} u \rangle + 2 \langle \hat{b} u, q \rangle \big] \\ &\qquad\qquad - \langle \hat{a}, D\varphi_\varepsilon(u) \rangle - \varphi'_\varepsilon(u) [cu + \langle \hat{b}, q \rangle + f] \Big\}\, dx\, ds \end{aligned}$$

$$\geq E\int_{Q_t}\Big\{-\frac{1}{2}\varphi''_\varepsilon(u)|\hat b|^2u^2+\langle B\hat b, D\int_0^u\varphi''_\varepsilon(r)r\,dr\rangle$$
$$+[\varphi''_\varepsilon(u)u-\varphi'_\varepsilon(u)]\langle\hat b,q\rangle+(\nabla\cdot\hat a)\varphi_\varepsilon(u)-\varphi'_\varepsilon(u)[cu+f]\Big\}\,dx\,ds\ . \tag{7.5}$$

Since $\int_0^u\varphi''_\varepsilon(r)r\,dr=\varphi'_\varepsilon(u)u-\varphi_\varepsilon(u)$, and $\lim_{\varepsilon\to0}[\varphi''_\varepsilon(u)u-\varphi'_\varepsilon(u)] = 2uI_{(u\leq0)}+2u^-=0$, letting $\varepsilon\to0$ in (7.5) we obtain

$$E\int_{\mathbb{R}^n}[g(x)^-]^2\,dx-E\int_{\mathbb{R}^n}[u(t,x)^-]^2\,dx\,ds$$
$$\geq E\int_{Q_t}\Big\{-I_{(u\leq0)}|\hat b|^2u^2-\nabla\cdot(B\hat b)[-2u^-u-|u^-|^2]$$
$$+(\nabla\cdot\hat a)|u^-|^2+2u^-[cu+f]\Big\}\,dx\,ds$$
$$\geq E\int_{Q_t}\Big\{(-|\hat b|^2-\nabla\cdot(B\hat b)+\nabla\cdot\hat a-2c)|u^-|^2-2u^-f^-\Big\}\,dx\,ds$$
$$\geq-\mu E\int_{Q_t}[u^-]^2\,dx\,ds-E\int_{Q_t}[f^-]^2\,dx\,ds\ ,$$

where $\mu\stackrel{\Delta}{=}\sup_{t,x,\omega}\big[-\nabla\cdot\hat a+\nabla\cdot(B\hat b)+|\hat b|^2+2c+1\big]<\infty$. Therefore, by Gronwall's inequality we obtain (7.4) for the case $\bar f=\bar g=0$ and $\lambda=0$. If $\lambda\neq0$, we can prove (7.4) (with $\bar u=\bar q=\bar f=\bar g=0$) by using transformation (3.9) and working on (v,p) for the transformed equations.

We now consider the general case when $\bar f\neq0$ and $\bar g\neq0$. It is clear that

$$\begin{cases} d(u-\bar u)=-\big\{\mathscr{L}(u-\bar u)+\mathscr{M}(q-\bar q)+(\mathscr{L}-\bar{\mathscr{L}})\bar u \\ \qquad\qquad +(\mathscr{M}-\bar{\mathscr{M}})\bar q+f-\bar f\big\}\,dt+\langle q-\bar q,\,dW(t)\rangle, \\ (u-\bar u)\big|_{t=T}=g-\bar g\ . \end{cases} \tag{7.6}$$

Then, using the previous arguments by replacing f by $(\mathscr{L}-\bar{\mathscr{L}})\bar u+(\mathscr{M}-\bar{\mathscr{M}})\bar q+f-\bar f$, we derive (7.4) immediately. The proof is now complete. □

Corollary 7.2. *Let the conditions of Theorem* 7.1 *hold. Then we have the following direct consequence*:

(i) *If* $g(x)-\bar g(x)\geq0$, $\forall x\in\mathbb{R}^n$, *a.s., and*

$$(\mathscr{L}-\bar{\mathscr{L}})\bar u(t,x)+(\mathscr{M}-\bar{\mathscr{M}})\bar q(t,x)+f(t,x)-\bar f(t,x)\geq0,\ \forall(t,x),\text{a.s.}\ , \tag{7.7}$$

then $u(t,x)\geq\bar u(t,x)$, $\forall(t,x)\in[0,T]\times\mathbb{R}^n$, *a.s.*.

(ii) *If* $\mathscr{L} = \bar{\mathscr{L}}$, $\mathscr{M} = \bar{\mathscr{M}}$, *and* $g(x) \geq \bar{g}(x)$, $f(t,x) \geq \bar{f}(t,x)$, $\forall (t,x)$, a.s., *then* $u(t,x) \geq \bar{u}(t,x)$, $\forall (t,x) \in [0,T] \times \mathbb{R}^n$, *a.s.*

(iii) *If* $g(x) \geq 0$, *and* $f(t,x) \geq 0$, $\forall (t,x)$, *a.s., then* $u(t,x) \geq 0$, *a.e.*$(t,x) \in [0,T] \times \mathbb{R}^n$, *a.s.*

Proof. (i) Under the assumption we have

$$E \int_{\mathbb{R}^n} e^{-\lambda \langle x \rangle} \left\{ [u(t,x) - \bar{u}(t,x)]^- \right\}^2 dx \leq 0, \quad \forall t \in [0,T] \ .$$

the conclusion follows immediately.

(ii) is a special case of (i).
(iii) follows from (i) by setting $\bar{f} = \bar{g} = 0$ and noting that $(\bar{u}, \bar{q}) = (0,0)$ is the unique solution of the corresponding BSPDE.

□

Let us now make an observation on Theorem 7.1. Suppose $(\bar{u}, \bar{q})$ is an adapted strong solution of (7.3). Then (7.7) gives a condition on A, B, a, b, c, f and g, such that the solution (u,q) of the equation (7.2) satisfies $u \geq \bar{u}$. The case when $\bar{u} = 0$ is very useful in the discussion of the convexity of the solution in the variable x.

The following corollary lists some sufficient condition for (7.7).

Corollary 7.3. *Suppose that the assumption of Theorem* 7.1 *hold.*

(i) *If* $\bar{A}$, $\bar{B}$, $\bar{a}$, $\bar{b}$ *and* $\bar{c}$ *are independent of* x. *Let* $\bar{f}$ *and* $\bar{g}$ *be convex in* x. *Let* $(\bar{u}, \bar{q})$ *be a strong solution of* (7.3). *Then*, $\bar{u}$ *is convex in* x *almost surely*.

(ii) *In addition to the assumptions in* (i), *assume that* $\bar{f}$ *and* $\bar{g}$ *are nonnegative. If* $\mathscr{M} = \bar{\mathscr{M}}$, *and for* $(t,x) \in [0,T] \times \mathbb{R}^n$, *it holds almost surely that*

$$\begin{cases} A(t,x) = \bar{A}(t) + A_0(t,x), \quad c(t,x) = \bar{c}(t) + c_0(t,x), \quad a(t,x) = \bar{a}(t,x), \\ f(t,x) = \bar{f}(t,x) + f_0(t,x), \quad g(x) = \bar{g}(x) + g_0(x) \ , \end{cases} \tag{7.8}$$

where

$$A_0(t,x) \geq 0, \quad c_0(t,x) \geq 0, \quad f_0(t,x) \geq 0, \quad g_0(x) \geq 0, \quad \forall (t,x), \text{ a.s. } . \tag{7.9}$$

Then $u(t,x) \geq \bar{u}(t,x)$, $\forall (t,x)$, a.s.

(iii) *If* $\bar{A}$, $\bar{B}$, $\bar{a}$, $\bar{b}$, $\bar{c}$, $\bar{f}$ *and* $\bar{g}$ *are deterministic, and let* $\bar{u}$ *be the solution of the following equation:*

$$\begin{cases} \bar{u}_t = -\bar{\mathscr{L}}u - \bar{f}, & (t,x) \in [0,T] \times \mathbb{R}^n, \\ \bar{u}\big|_{t=T} = \bar{g} \ . \end{cases}$$

Suppose that $\bar{u}(t,x)$ *is convex in* x. *Then for any solution* (u,q) *to the BSPDE* (7.2) *with* $\mathscr{L}$ *satisfying* (7.8)–(7.9), *it holds that* $u(t,x) \geq \bar{u}(t,x)$, $\forall (t,x)$, *a.s.*.

Proof. (i) We first assume that f and g are smooth enough in x. Then the corresponding solution $(\bar{u},\bar{q})$ of (7.3) is smooth enough in x. Now, for any $\eta \in \mathbb{R}^n$, we define

$$\begin{cases} v(t,x) = \langle D^2\bar{u}(t,x)\eta,\eta\rangle, \\ p(t,x) = (p^1(t,x),\ldots,p^d(t,x)), & \forall (t,x) \in [0,T] \times \mathbb{R}^n, \text{a.s.} \\ p^k(t,x) = \langle D^2\bar{q}^k(t,x)\eta,\eta\rangle, & 1 \leq k \leq d, \end{cases}$$

Then the following holds

$$\begin{cases} dv = [-\bar{\mathscr{L}}v - \bar{\mathscr{M}}p - \langle (D^2\bar{f})\eta,\eta\rangle]\,dt + \langle p,\, dW(t)\rangle, \\ v\big|_{t=T} = \langle (D^2\bar{g})\eta,\eta\rangle \ . \end{cases} \tag{7.23}$$

By Corollary 7.3 and the convexity of $\bar{f}$ and $\bar{g}$ (in x), we obtain

$$\langle D^2\bar{u}(t,x)\eta,\eta\rangle = v(t,x) \geq 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \quad \eta \in \mathbb{R}^n, \text{a.s.} \tag{7.24}$$

This implies the convexity of $\bar{u}(t,x)$ in x almost surely. In the case that $\bar{f}$ and $\bar{g}$ are not necessarily smooth enough, we may make approximation.

(ii) By Corollary 7.2 and part (i), $\bar{u}$ is convex and nonnegative. Thus,

$$(\mathscr{L} - \bar{\mathscr{L}})\bar{u}(t,x) = \tfrac{1}{2}\mathrm{tr}[A_0 D^2\bar{u}] + c_0\bar{u} \geq 0 \ .$$

The conclusion follows.

(iii) Note that if all the coefficient are deterministic, then $(\bar{u},0)$ is the unique (adapted) strong solution of (7.3). Using the similar arguments as in (ii), and noting that $\bar{q} = 0$, we obtain the assertion readily. □

We remark that in Corollary 7.3-(iii), the operator $\mathscr{M}$ (or equivalently, the coefficients B and b) are arbitrary.

8. Discussions

The comparison theorems presented in the previous sections are strongly motivated by the problems in finance, especially by the issue of robustness of Black-Scholes formula (see [6] for a detailed exposition of the problem). Let us quickly recall some facts. The standard Black-Scholes PDE (see, for example, [11, p.379]) in a one dimensional version is of the following form:

$$\begin{cases} u_t + \frac{1}{2}x^2\sigma^2 u_{xx} + xru_x - ru = 0, & \text{on } [0,T) \times (0,\infty), \\ u(T,x) = h(x), & x \geq 0 \quad . \end{cases} \tag{8.1}$$

Here, r is the interest rate of the (non-risky) money market and σ is the volatility of the (risky) stock market. If both r and σ are constant, then the Black-Scholes option valuation formula tells us that for the given European contingent claim $h(S_T)$, where S_T is the stock price at time T, the value of the contingent claim at time $t \in [0,T)$ is given by

$$Y_t = u(t,S_t) = E\{e^{-r(T-t)}h(S_T)|\mathscr{F}_t\}, \quad 0 \leq t < T \quad . \tag{8.2}$$

The robustness of the Black-Schole formula concerns the following problem: suppose a practitioner's information leads him to a misspecified value of, say, volatility σ, and he calculates the option price according to this misspecified parameter and equation (8.1), and then tries to hedge the contingent claim, what will be the consequence? In [6] it is proved that under certain conditions (in particular that contingent claims have convex payoffs), then the value of European contingent claim is convex; and, if the misspecified volatility dominates (resp. is dominated by) the true volatility, which is assumed to be any $\{\mathscr{F}_t\}$-adapted process $\sigma(\cdot)$, then the contingent claim price corresponding to the misspecified volatility dominates (resp. is dominated by) the true contingent claim price. An interesting counterexample is also given in that paper, which shows that if the misspecified volatility is not the deterministic function of the stock price, the comparison may fail.

In what follows we give some discussions in this direction, in terms of our comparison theorems. We note that our discussion is not a "translation" of the existing results, but rather a new way of looking at the problem. However, some of our conclusions do not seem to be easy consequences of the methods presented in [7].

To begin with, we first recall that in [15] we proved that in the case where the parameters r and σ are allowed to be $\{\mathscr{F}_t\}$-adapted processes, the option price can still be written in the form (8.2), but with u

being replaced by an adapted solution to a BSPDE of the following form:

$$\begin{cases} du = \big\{ -\frac{1}{2}x^2\sigma^2 u_{xx} - xru_x + ru - x\sigma q_x - \theta q \big\}\, dt - q\, dW(t), \\ \qquad \text{on } [0,T) \times (0,\infty), \\ u(T,x) = h(x), \qquad x \geq 0 \ , \end{cases} \tag{8.3}$$

where $\theta = \sigma^{-1}(b - r)$ is so-called *risk premium* process. We remark that the well-posedness of (8.3) is not the direct consequence of the results in §3, because the coefficients of the BSPDE are linear in x (whence unbounded). However, a simple change of variable: $x = e^y$ would convert the equation to an equivalent BSPDE with random but bounded coefficients (see [15] or [4] for detailed calculations), so that the results of §3 can be applied. The interpretation of this mathematically well-known *Euler transformation* in finance terms is nothing but to consider the ''log-price'' instead of stock price itself, a common way to look at the ''geometric Brownian motion''. We also recall that if all the parameters: r, σ, and b are deterministic, then the solution of (8.3) is $(u, 0)$, where u is the solution of (8.1). Therefore we at times consider (8.1) as a special case of (8.3) and apply the results of BSPDEs to (8.1) as well.

0.1 Convexity of the European contingent claims. Assume that r and σ are stochastic processes, independent of the current stock price. Since the coefficients in (8.3) depend on x, we cannot apply Corollary 7.3 directly. But if we differentiate (8.3) twice (in x) and denote $v = u_{xx}$, $p = q_{xx}$, then we see that (v, p) satisfies the following (linear) BSPDE:

$$\begin{cases} dv = \big\{ -\frac{1}{2}x^2\sigma^2 v_{xx} - (2x\sigma^2 + xr)v_x - (\sigma^2 + r)v - x\sigma p_x \\ \qquad - (2\sigma - \theta)p \big\}\, dt - p\, dW(t); \\ v(T,x) = h''(x) \ . \end{cases} \tag{8.4}$$

Here again the well-posedness of (8.4) can be obtained by considering its equivalent form after the Euler transformation (since r and σ are independent of x). Moreover, applying Corollary 7.2-(iii) (to the transformed equation) it is easily seen that $v \geq 0$, whenever $h'' \geq 0$, and hence u is convex. Note that since we allow both r and σ to be random, this result seems to be new.

In fact, we can discuss more complicated situation by using the comparison theorem. For example, let us assume that both r and σ are *deterministic* functions of (t, x), and we assume that they are both C^2 for simplicity. Then as we pointed out before, (8.3) coincides with (8.1). Now differentiating (8.1) twice and denoting $v = u_{xx}$, we see that v satisfies the following PDE:

$$\begin{cases} 0 = v_t + \frac{1}{2}x^2\sigma^2 v_{xx} + \hat{a}xv_x + \hat{b}v + r_{xx}(xu_x - u), & \text{on } [0,T) \times (0,\infty), \\ v(T,x) = h''(x), & x \geq 0 \quad , \end{cases} \tag{8.5}$$

where

$$\hat{a} = 2\sigma^2 + 2x\sigma\sigma_x + r;$$
$$\hat{b} = \sigma^2 + 4x\sigma\sigma_x + (x\sigma_x)^2 + x^2\sigma\sigma_{xx} + 2xr_x + r \quad .$$

Now let us denote $V = xu_x - u$, then some computation shows that V satisfies the equation:

$$\begin{cases} 0 = V_t + \frac{1}{2}x^2\sigma^2 V_{xx} + \hat{c}xV_x + (xr_x - r)V, & \text{on } [0,T) \times (0,\infty), \\ V(T,x) = xh'(x) - h(x), & x \geq 0 \quad , \end{cases} \tag{8.6}$$

for some function $\hat{c}$ depending on $\hat{a}$ and $\hat{b}$ (whence r and σ).

Now we can apply Corollary 7.2 repeatedly (use Euler transformation if necessary) to get the following results: assume that h is convex, then

(i) if r is convex and $xh'(x) - h(x) \geq 0$, then u is convex.
(ii) if r is concave and $xh'(x) - h(x) \leq 0$, then u is convex.
(iii) if r is independent of x, then u is convex.

Indeed, if $xh'(x) - h(x) \geq 0$, then $V \geq 0$ by Corollary 7.2-(iii). This, together with the convexity of r and h, in turn shows that the solution v of (8.5) is non-negative. Namely, u is convex, proving (i). Part (ii) can be argued similarly. We note that when r is independent of x, (8.5) becomes a homogeneous equation, thus the convexity of h already implies that of u, thanks to Corollary 7.2-(iii) again. In fact, this special case can be proved more directly under weaker conditions. We refer the interested reader to [7, Theorem 5.2]. We also remark here that, modulo the smoothness conditions, one can check that a standard European call option satisfies $xh'(x) - h(x) \geq 0$, but European put option satisfies $xh'(x) - h(x) \leq 0$. In other words, in the case when r depends also on x, the convexity of the value of the European contingent claim depends further on the type of payoff function, which is quite different from the case (iii) where the convexity is the only requirement for h.

0.2 Robustness of Black-Scholes Formula. Let us first assume that the only misspecified parameter is the volatility, and denote it by $\sigma = \sigma(t,x)$, which is C^2 in x; and assume that the interest rate is deterministic and independent of the stock price. By the conclusion (iii) in the previous part we know that u is convex in x. Now let us assume that the true volatility is an $\{\mathscr{F}_t\}$-adapted process, denoted by $\hat{\sigma}$, satisfying

$$\hat{\sigma}(t) \geq \sigma(t,x), \quad \forall (t,x), \text{ a.s.} \tag{8.7}$$

Also let us denote the operators corresponding to (8.3) with r and $\hat{\sigma}$ by $(\hat{\mathscr{L}}, \hat{\mathscr{M}})$, and that to (8.1) by $(\mathscr{L}, \mathscr{M})$. Let the solution to (8.3) be $(\hat{u}, \hat{q})$, and that of (8.1) be $(u, 0)$. Note that both (8.1) and (8.3) are homogeneous equations, i.e., $\hat{f} = f = 0$, thus

$$(\hat{\mathscr{L}} - \mathscr{L})u + (\hat{\mathscr{M}} - \mathscr{M})q + \hat{f} - f = \tfrac{1}{2}x^2[\hat{\sigma}^2 - \sigma^2]u_{xx} \geq 0 \ , \tag{8.8}$$

because u is convex. Consequently one has $\hat{u}(t,x) \geq u(t,x)$, $\forall (t,x)$, a.s., thanks to Corollary 7.2-(ii).

Now let us assume that the inequality in (8.7) is reversed. Note that both (8.1) and (8.3) are linear and homogeneous, we see that $(-\hat{u}, -\hat{q})$ and $(-u, 0)$ also satisfy (8.3) and (8.1), respectively, with the terminal condition being replaced by $-h(x)$. But in this case (8.8) becomes

$$(\hat{\mathscr{L}} - \mathscr{L})(-u) = \tfrac{1}{2}x^2[\hat{\sigma}^2 - \sigma^2](-u_{xx}) \geq 0 \ ,$$

because u is convex, and $\hat{\sigma}^2 \leq \sigma^2$. Thus $-\hat{u} \geq -u$, namely $\hat{u} \leq u$.

Using the similar technique we can again discuss some more complicated situations. For example, let us allow the interest rate r to be misspecified as well, but in the form that it is convex in x, say. Assume that the payoff function h satisfies $xh'(x) - h(x) \geq 0$, and that $\hat{r}$ and $\hat{\sigma}$ are true interest rate and volatility such that they are $\{\mathscr{F}_t\}$-adapted random fields satisfying $\hat{r}(t,x) \geq r(t,x)$, and $\hat{\sigma}(t,x) \geq \sigma(t,x)$, $\forall (t,x)$. Then, using the notation as before, one shows that

$$(\hat{\mathscr{L}} - \mathscr{L})u = \tfrac{1}{2}x^2[\hat{\sigma}^2 - \sigma^2]u_{xx} + (\hat{r} - r)[xu_x - u] \geq 0 \ ,$$

because u is convex, and $xu_x - u = V \geq 0$, thanks to the arguments in the previous part. Consequently one has $\hat{u}(t,x) \geq u(t,x)$, $\forall (t,x)$, a.s.. Namely, we also derive a one-sided domination of the true values and misspecified values. One can of course try to get more comparison results, by combining the arguments before, but we prefer not to pursue any further, as this is not the main purpose of the paper.

Acknowledgement. We would like to thank the anonymous referee for the careful reading of the manuscript and helpful suggestions.

References

[1] Bensoussan, A.: Stochastic Control of Partially Observable Systems, Cambridge University Press, 1992

[2] Bergman,Y.Z., Grundy, B.D., Wiener, Z.: General Properties of Option Pricing, Preprint, 1996

[3] Buckdahn, R., Hu, Y.: Hedging contingent claims for a large investor in a incomplete market, Adv. in Appl. Probab. 30 (1998), 239–255

[4] Cvitanic, J., Ma, J.: Hedging Options for a Large Investor and Forward-Backward SDEs, Ann. Appl. Probab., Vol. 6, **2**, 370–398 (1996)

[5] Duffie, D., Ma, J., Yong, J.: Black's consol rate conjecture, Ann. Appl. Prob., **5**, 356–382 (1995)

[6] El Karoui, N., Jeanblanc-Picqué, M., Shreve, S.E.: Robustness of the Black and Scholes Formula, Mathematical Finance, to appear

[7] El Karoui, N., Peng, S., Quenez, M.C.: Backward Stochastic Differential Equations in Finance, Mathematical Finance 7, 1–72 (1997)

[8] Hörmander, L.: The Analysis of Linear Partial Differential Operators, I, Springer-Verlag, Berlin, 1983

[9] Hu, Y., Peng, S.: Adapted solution of a backward semilinear stochastic evolution equations, Stoch. Anal. & Appl., **9**, 445–459 (1991)

[10] Hu Y., and Peng, S.: Solution of forward-backward stochastic differential equations, Prob. Th. Rel. Fields, **103**, 273–283 (1995)

[11] Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus, Springer-Verlag, Berlin, 1988

[12] Kunita, H.: Stochastic Flows and Stochastic Differential Equations, Cambridge Studies in Advanced Math. **24**, Cambridge University Press, 1990

[13] Ma, J., Protter, P., Yong, J.: Solving Forward-Backward Stochastic Differential Equations Explicitly – A Four Step Scheme, Prob. Th. Rel. Fields, **98**, 339–359 (1994)

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

[15] Ma J., Yong, J.: Adapted solution of a degenerate backward SPDE, with applications, Stoch. Proc. & Appl. **70**, 59–84 (1997)

[16] Nagasa, N., Nisio, M.: Optimal controls for stochastic partial differential equations, SIAM J. Control Optim., **28**, 186–213 (1990)

[17] Pardoux, E.: Stochastic partial differential equations and filtering of diffusion processes, Stochastics, **3**, 127–167 (1979)

[18] Pardoux, E., Peng, S.: Adapted Solution of a Backward Stochastic Differential Equation, Syst. Control Lett. **14**, 55–61 (1990)

[19] Pardoux, E., Peng, S.: Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations, Lecture Notes in CIS, Springer-Verlag, Berlin, **176**, 200–217 (1992)

[20] Protter, P.: Stochastic Integration and Differential Equations, A New Approach, Springer-Verlag, Berlin, 1990

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

[22] Yong, J.: Finding adapted solution of forward-backward stochastic differential equations – method of continuation, Prob. Th. Rel. Fields, **107**, 537–572 (1997)

[23] Yong, J.: Linear forward-backward stochastic differential equations, Appl. Math. Optim., to appear
[24] Zhou, X.Y.: A duality analysis on stochastic partial differential equations, J. Funct. Anal., **103**, 275–293 (1992)
[25] Zhou, X.Y.: On the Necessary Conditions of Optimal Controls for Stochastic Partial Differential Equations, SIAM J. Control & Optim., **31**, 1462–1478 (1993)