Probab. Theory Relat. Fields (2011) 151:475–507
DOI 10.1007/s00440-010-0305-8

On weak solutions of forward–backward SDEs

Jin Ma · **Jianfeng Zhang**

Received: 1 June 2009 / Revised: 19 February 2010 / Published online: 25 June 2010

Abstract In this paper we continue exploring the notion of *weak solution* of forward–backward stochastic differential equations (FBSDEs) and associated *forward–backward martingale problems* (FBMPs). The main purpose of this work is to remove the constraints on the martingale integrands in the uniqueness proofs in our previous work (Ma et al. in Ann Probab 36(6):2092–2125, 2008). We consider a general class of non-degenerate FBSDEs in which all the coefficients are assumed to be essentially only bounded and uniformly continuous, and the uniqueness is proved in the space of all the square integrable adapted solutions, the standard solution space in the FBSDE literature. A new notion of *semi-strong* solution is introduced to clarify the relations among different definitions of weak solution in the literature, and it is in fact instrumental in our uniqueness proof. As a by-product, we also establish some a priori estimates of the second derivatives of the solution to the decoupling quasilinear PDE.

Keywords Forward–backward stochastic differential equations · Weak solution · Forward–backward martingale problems · Viscosity solutions

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

J. Ma is supported in part by NSF grants #0835051 and #0806017. J. Zhang is supported in part by NSF grant #0631366.

J. Ma (✉) · J. Zhang
Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA
e-mail: jinma@usc.edu

J. Zhang
e-mail: jianfenz@usc.edu

1 Introduction

The theory of backward stochastic differential equations (BSDE) and forward–backward stochastic differential equations (FBSDE) has been explored quite extensively in the past two decades since the seminal work of Pardoux and Peng [26] appeared in 1990. We refer to [1,9,13,17,19,21,27–29,32,33], and the book [23] for the wellposedness of various forms of such equations. All these works, however, are in the realm of "strong solutions".

The notion of weak solutions of FBSDEs, first proposed by Antonelli and Ma [2], has become an important branch of research in the theory of BSDEs/FBSDEs. In a sequence of three papers [5–7], Buckdhan et al. studied the weak solution for BSDEs where a forward component is implicitly given. In these works, the standard issues such as relations between pathwise uniqueness and uniqueness in law, as well as the Yamada–Watanabe type results regarding the weak solution and strong solution were discussed. These issues were further explored recently by Kurtz [18], in a more general framework. Despite all the efforts, however, the uniqueness of the weak solution remains open. We should mention that the notion of weak solution was also studied by Delarue–Guatteri [11], where the existence and uniqueness of weak solutions was established for a class of Markovian FBSDEs. However, since in that framework the coefficients are assumed to be Lipschitz continuous in terms of the backward components, the solution is therefore "weak" only in the forward component, and the fundamental nature of the weak solution for BSDEs, especially the uniqueness, was essentially avoided.

In our previous work [24] we studied weak solutions for fully coupled FBSDEs and introduced the notion of *Forward–Backward Martingale Problem* (FBMP, for short), and proved the equivalence between the two notions. We first established an existence result for general FBSDEs with possibly path dependent coefficients, under certain tightness conditions. We then verified, in the one dimensional and non-degenerate Markovian case, that the tightness conditions do hold for FBSDE with bounded and uniformly continuous coefficients, which leads to the existence of weak solution. To prove the uniqueness (in law) of weak solutions, we investigated a variation of the notion of "nodal set" in [22], and argued that the weak solution is unique if the comparison principle holds for the viscosity solution of the related quasi-linear PDE. However, this result only applies to those solutions whose component Z is bounded in a certain sense.

In this paper we continue exploring the well-posedness of non-degenerate Markovian FBSDEs with bounded and uniformly continuous coefficients. We shall introduce a new notion of *semi-strong* solution so as to clarify the relations between the different weak solutions in the literature. For example, we note that the strong solution defined in [7] is actually semi-strong under our definition. Next, we establish some a priori estimates for the derivatives of the solutions to the corresponding decoupling PDE, from which we obtain the existence of semi-strong solutions for FBSDEs with arbitrary dimension, extending the existence result in [24] to high dimensional cases. In the case when the backward component is scalar we further prove that, starting from any point in the "nodal set" one can construct a semi-strong solution whose component Z is locally bounded in the variable t.

The main goal of this paper is to remove the boundedness constraint on Z in the uniqueness result of [24]. As in [24], we shall prove that the upper bound of the "nodal set" is a viscosity sub-solution to the related PDE and the lower bound is a viscosity super-solution. Then the comparison principle of viscosity solutions implies the uniqueness of the weak solutions. To remove the boundedness constraint on the component Z, we show that for any weak solution whose initial value is in the "nodal set", we can actually construct a semi-strong solution starting from the same initial value and with locally bounded Z. From this fact we then obtain the same estimates and eventually prove the uniqueness among all weak solutions.

The rest of the paper is organized as follows. In Sect. 2 we give the preliminaries, introduce various types of solutions, and illustrate their difference by examples. In Sect. 3 we give some a priori estimates for the solutions to the decoupling PDE, and prove the existence result. In Sect. 4 we study the properties of nodal set. In particular, we show that starting from any point in the nodal set there exists a semi-strong solution whose component Z is locally bounded. Finally, we prove the uniqueness result in Sect. 5, and give the complete proofs of the a priori estimates for the PDE in Sect. 6.

2 Preliminaries

We first recall some basic notions from our previous work [24]. For a given finite time horizon $[0, T]$, we say that a quintuple $(\Omega, \mathcal{F}, P, \mathbf{F}, W)$ is a *standard set-up* if $(\Omega, \mathcal{F}, P)$ is a complete probability space; $\mathbf{F} \stackrel{\triangle}{=} \{\mathcal{F}_t\}_{t\in[0,T]}$ is a filtration satisfying the *usual hypothesis* (see, e.g. [30]); and W is an $\mathbf{F}$-Brownian motion. In particular, if $\mathbf{F} = \mathbf{F}^W$, the filtration generated by the Brownian motion with the usual augmentation, then we say that the set-up is *Brownian*.

We consider the following forward–backward SDE on a standard set-up:

$$\begin{cases} X_t = x + \int_0^t b(s, (X)_s, Y_s, Z_s)ds + \int_0^t \sigma(s, (X)_s, Y_s)dW_s \\ Y_t = g((X)_T) + \int_t^T f(s, (X)_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s. \end{cases} \tag{2.1}$$

Here $(X_t, Y_t, Z_t, W_t) \in \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times k} \times \mathbb{R}^k$, and the coefficients b, σ, f, and g are functions with appropriate dimensions. We note in particular that the coefficient b is a *progressively measurable function* defined on $[0, T] \times \mathbb{C}([0, T], \mathbb{R}^d) \times \mathbb{R}^m \times \mathbb{R}^{m\times k}$ with values in $\mathbb{R}^d$; and $(X)_t$ denotes the path of X up to time t. More precisely, for each $t \in [0, T]$, and $(y, z) \in \mathbb{R}^m \times \mathbb{R}^{m\times k}$, the mapping $\mathbf{x} \mapsto b(t, (\mathbf{x})_t, y, z)$ is measurable with respect to the σ-field $\mathcal{B}_t(\mathbb{C}([0, T]; \mathbb{R}^d))$, where $\mathcal{B}_t(\mathbb{C}([0, T]; \mathbb{R}^d)) \stackrel{\triangle}{=} \sigma\{\mathbf{x}(t \wedge \cdot) : \mathbf{x} \in \mathbb{C}([0, T]; \mathbb{R}^d)\}$ (cf. e.g. [14]). The other coefficients satisfy similar measurability.

We now recall the following definition of the *weak solution*, and also define the definitions of the *semi-strong* as well as *strong* solutions for the FBSDE (2.7).

Definition 2.1 A standard set-up $(\Omega, \mathcal{F}, P, \mathbf{F}, W)$ along with a triplet of processes (X, Y, Z) defined on this set-up is called a weak solution of (2.7) if

(i) the processes X, Y are continuous, and all processes X, Y, Z are $\mathbf{F}$-adapted;

(ii) denoting $\varphi_s \stackrel{\triangle}{=} \varphi(s, (X)_s, Y_s, Z_s)$ for $\varphi = b, \sigma, f$, it holds that

$$P\left\{\int_0^T\Big(|b_t| + |\sigma_t|^2 + |f_t| + |Z_t|^2\Big)\,dt + |g(X_T)| < \infty\right\} = 1.$$

(iii) (X, Y, Z) verifies (2.7) P-a.s.
Moreover, a weak solution is called "semi-strong" if $\mathbf{F} = \mathbf{F}^{X,W}$. It is called a "strong solution" if $\mathbf{F} = \mathbf{F}^W$.

Remark 2.2 (i) We note that if the FBSDE is "decoupled", that is, the forward equation does not depend on the backward component, then there is no significant difference between a strong solution and a semi-strong solution, as long as the process W remains a Brownian motion under $\mathbf{F}^{X,W}$. This is because as far as the backward SDE is concerned, the probability space can be determined a priori once X and W are given. We refer to [7] for weak solutions to BSDEs reflecting such reasonings.

(ii) In the coupled FBSDE case, however, the problem becomes quite different. Because of the mutual influence between the forward and backward components, the filtration generated by the solution (X, Y, Z, W) can be significantly bigger than $\mathbf{F}^W$, $\mathbf{F}^X$, or even $\mathbf{F}^{X,W}$, and the probability space cannot be predetermined even for the semi-strong solutions.

□

The following examples show the differences among weak, strong, and semi-strong solutions.

Example 2.3 (Semi-strong vs. strong) Consider a decoupled FBSDE (2.7) in which $b \equiv 0$, $\sigma = \operatorname{sgn}(x)$, and $f = f(x, y, z)$ is bounded and measurable in x but uniform Lipschitz in (y, z). Since the forward SDE is exactly the same as the *Tanaka example* (see, e.g. [31]):

$$X_t = x + \int_0^t \operatorname{sgn}(X_s)\,dW_s, \qquad t \geq 0, \tag{2.2}$$

it is well-known that the filtration $\mathbf{F}^X$ is strictly bigger than $\mathbf{F}^W$, hence $\mathbf{F}^{X,W} = \mathbf{F}^X$ is strictly bigger than $\mathbf{F}^W$. Furthermore, note that whenever a (weak) solution (X, W) of (2.2) is given on some probability space, $(X, \mathbf{F}^X)$ is also a Brownian motion. Thus for any duration $[0, T]$ and under very mild technical conditions, the FBSDE is always solvable on this probability space with solution being $\mathbf{F}^X$-adapted, see e.g. [19]. Namely, it is semi-strong but not strong. □

Remark 2.4 We note that in general the inclusion $\mathbf{F}^W \subseteq \mathbf{F}^X$ may not be true. A simple example would be to modify the FBSDE above so that $T = 2$, and $\sigma(t,x) = \operatorname{sgn}(x)\mathbf{1}_{[0,1]}(t)$. Then for $t \in [0,1]$ one still has $\mathcal{F}_t^{X,W} = \mathcal{F}_t^X$ which contains $\mathcal{F}_t^W$ strictly. But for $t \in (1,2]$, $\sigma \equiv 0$, and $X_t \equiv X_1$, hence $\mathcal{F}_t^X \equiv \sigma(X_t, t \le 1)$. Thus $\mathcal{F}_t^{X,W} = \mathcal{F}_t^W \vee \sigma(X_t, t \le 1)$ strictly contains both $\mathcal{F}_t^W$ and $\mathcal{F}_t^X$ for $t \in (1,2]$. Consequently, $\mathbf{F}^X$ and $\mathbf{F}^W$ are mutually non-inclusive, and are included strictly in $\mathbf{F}^{X,W}$. □

Example 2.5 (Weak vs. semi-strong) Note that in [24] we actually proved that the semi-strong solution always exists whenever the coefficients are bounded and uniformly continuous. Thus in order to find a weak solution that is not semi-strong one must consider the case where the solutions are actually non-unique.

Let us still consider the decoupled FBSDE (2.7), this time with $T = 1$, $g = 0$, and $f = f(y) \triangleq 2[\sqrt{|y|} \wedge 1]$. Note that the ODE

$$Y_t = \int_t^1 f(Y_s)ds, \qquad t \in [0,1] \tag{2.3}$$

has two solutions: $\underline{Y} \equiv 0$ and $\bar{Y}_t \triangleq (1-t)^2$. Denote $\underline{Z} \triangleq \bar{Z} \triangleq 0$. Clearly, for any weak solution $(\Omega, \mathcal{F}, P, \mathbf{F}^{W,X}, W, X)$ of the forward SDE (2.2), adding $(\underline{Y}, \underline{Z})$ and $(\bar{Y}, \bar{Z})$ we obtain two semi-strong solutions to the FBSDE (2.7).

We now try to construct a weak solution that is *not* semi-strong. Let $\mathbf{F}$ be any filtration that is strictly bigger than $\mathbf{F}^{W,X}$, but W is still an $\mathbf{F}$-Brownian motion, and choose $Z^0 \in L^2(\mathbf{F}; [0,1])$ that is not $\mathbf{F}^{W,X}$-adapted. Let Y^0 be an $\mathbf{F}$-adapted solution to the following "randomized" ODE with initial condition $y \in (0,1)$:

$$Y_t^0 = y - \int_0^t f(Y_r^0)dr + \int_0^t Z_r^0 dW_r, \qquad t \in [0,1], \tag{2.4}$$

Define $\tau_1 \triangleq \inf\{t : Y_t^0 = \underline{Y}_t\}$, $\tau_2 \triangleq \inf\{t : Y_t^0 = \bar{Y}_t\}$. Note that $\underline{Y}_0 < Y_0 < \bar{Y}_0$, $\underline{Y}_1 = \bar{Y}_1 = 0$, and all the processes are continuous, then $0 < \tau_1 \wedge \tau_2 \le 1$. Define

$$Y_t \triangleq Y_t^0 1_{\{t \le \tau_1 \wedge \tau_2\}} + \underline{Y}_t 1_{\{t > \tau_1, \tau_1 \le \tau_2\}} + \bar{Y}_t 1_{\{t > \tau_2, \tau_1 > \tau_2\}}, \quad Z_t \triangleq Z_t^0 1_{\{t \le \tau_1 \wedge \tau_2\}}.$$

Then one can check that $(\Omega, \mathcal{F}, P, \mathbf{F}, W, X, Y, Z)$ is a weak solution, and obviously (Y, Z) is $\mathbf{F}$-adapted but not $\mathbf{F}^{X,W}$-adapted, and thus the solution is not semistrong.

For completeness we show that (2.4) has an $\mathbf{F}$-adapted solution Y^0. Let $\{f_n\}_{n \ge 0}$ be the smooth mollifiers of f such that $f_n \uparrow f$ as $n \to \infty$. Denote Y^n to be the (unique) strong solution to the following randomized ODE on the space $(\Omega, \mathcal{F}, P)$:

$$Y_t^n = y - \int_0^t f_n(Y_r^n)dr + \int_0^t Z_r^0 dW_r, \qquad t \in [0, 1].$$

Applying the comparison theorem, one sees that Y^n's are decreasing, thus there exists an $\mathbf{F}$-adapted process Y^0 such that $Y^n \downarrow Y^0$. By bounded convergence theorem and the uniform continuity of f one then shows that Y^0 solves (2.4). $\square$

We now recall the notion of the "*Forward–Backward Martingale Problem*". We begin by recalling the "canonical set-up". Define,

$$\Omega^1 \stackrel{\triangle}{=} \mathbb{C}([0, T]; \mathbb{R}^d); \qquad \Omega^2 \stackrel{\triangle}{=} \mathbb{C}([0, T]; \mathbb{R}^m); \quad \Omega \stackrel{\triangle}{=} \Omega^1 \times \Omega^2. \tag{2.5}$$

Here Ω^1 denotes the path space of the forward component X and Ω^2 the path space of the backward component Y of the FBSDE, respectively. Next, we define the canonical filtration by $\mathcal{F}_t \stackrel{\triangle}{=} \mathcal{F}_t^1 \otimes \mathcal{F}_t^2$, $0 \le t \le T$, where $\mathcal{F}_t^i \stackrel{\triangle}{=} \sigma\{\omega^i(r \wedge t) : r \ge 0\}$, $i = 1, 2$. We denote $\mathcal{F} \stackrel{\triangle}{=} \mathcal{F}_T$ and $\mathbf{F} \stackrel{\triangle}{=} \{\mathcal{F}_t\}_{0 \le t \le T}$.

In what follows we denote the generic element of Ω by $\omega = (\omega^1, \omega^2)$, and denote the canonical processes on $(\Omega, \mathcal{F})$ by

$$\mathbf{x}_t(\omega) \stackrel{\triangle}{=} \omega^1(t), \quad \text{and} \quad \mathbf{y}_t(\omega) \stackrel{\triangle}{=} \omega^2(t), \quad t \ge 0.$$

Finally, let $\mathcal{P}(\Omega)$ be the space of all the probability measures defined on $(\Omega, \mathcal{F})$, endowed with the Prohorov metric. We should note that for every $\mathbb{P} \in \mathcal{P}(\Omega)$, the term "filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; \mathbf{F})$" should always mean that $\mathcal{F} = \overline{\mathcal{F}}^{\mathbb{P}}$ and $\mathbf{F} = \overline{\mathbf{F}}^{\mathbb{P}}$. That is, $\mathcal{F}$ is completed and $\mathbf{F}$ is augmented, by the probability $\mathbb{P}$, respectively. We will not repeat this point in the future.

For $\varphi = b, f$, we denote $\hat{\varphi}(t, (\mathbf{x})_t, y, z) = \varphi(t, (\mathbf{x})_t, y, z\sigma(t, (\mathbf{x})_t, y))$, and let $a = \sigma\sigma^*$. We give the following definition for a *Forward–Backward Martingale Problem.*

Definition 2.6 Let b, σ, f, and g be given. For any $x \in \mathbb{R}^d$, a solution to the forward–backward martingale problem with coefficients (b, σ, f, g) ($\text{FBMP}_{x,T}(b, \sigma, f, g)$ for short) is a pair $(\mathbb{P}, \mathbf{z})$, where $\mathbb{P} \in \mathcal{P}(\Omega)$, and $\mathbf{z}$ is a $\mathbb{R}^{m\times d}$-valued predictable process defined on the filtered canonical space $(\Omega, \mathcal{F}, \mathbb{P}; \mathbf{F})$, such that following properties hold:

(i) the processes

$$M_{\mathbf{x}}(t) \stackrel{\triangle}{=} \mathbf{x}_t - \int_0^t \hat{b}(r, (\mathbf{x})_r, \mathbf{y}_r, \mathbf{z}_r)dr \;\text{ and }\; M_{\mathbf{y}}(t) \stackrel{\triangle}{=} \mathbf{y}_t + \int_0^t \hat{f}(r, (\mathbf{x})_r, \mathbf{y}_r, \mathbf{z}_r)dr \tag{2.6}$$

are both $(\mathbb{P}, \mathbf{F})$-martingales for $t \in [0, T]$;

(ii) $[M_{\mathbf{x}}^i, M_{\mathbf{x}}^j](t) = \int_0^t a_{ij}(r, (\mathbf{x})_r, \mathbf{y}_r)dr,\ t \in [0, T],\ i, j = 1, \ldots n;$

(iii) $M_{\mathbf{y}}(t) = \int_0^t \mathbf{z}_r dM_{\mathbf{x}}(r),\ t \in [0, T].$

(iv) $\mathbb{P}\{\mathbf{x}_0 = x\} = 1$ and $\mathbb{P}\{\mathbf{y}_T = g((\mathbf{x})_T)\} = 1.$

We note that by (iii) we imply that the quadratic variation of $M_{\mathbf{y}}$ is absolutely continuous with respect to the quadratic variation of $M_{\mathbf{x}}$, thus in the definition we require implicitly

$$\mathbb{P}\left\{\int_0^T |\mathbf{z}_t a(t, (\mathbf{x})_t, \mathbf{y}_t)\mathbf{z}_t^*|_{\mathbb{R}^{m\times m}} dt < \infty\right\} = 1.$$

In [24] we showed that, when σ is non-degenerate, the FBSDE (2.1) has a weak solution if and only if the FBMP has a solution (with $a = \sigma\sigma^*$), with the relation $Z_t = \mathbf{z}_t\sigma(t, (\mathbf{x})_t, \mathbf{y}_t)$, at least in distribution. In this case we shall also call $(\mathbb{P}, Z)$ a solution to the FBMP.

In this paper we shall mainly concentrate on the following Markovian type FBSDE:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s)dW_s \\ Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \end{cases} \quad t \in [0, T]. \tag{2.7}$$

It is well known that this FBSDE is associated with the following system of PDEs:

$$\begin{cases} u_t^i + \frac{1}{2}\mathrm{tr}\,(u_{xx}^i\sigma(t, x, u)\sigma^*(t, x, u)) + u_x^i b(t, x, u, u_x\sigma(t, x, u)) \\ \qquad\qquad + f^i(t, x, u, u_x\sigma(t, x, u)) = 0, \qquad i = 1, \ldots, m; \\ u(T, x) = g(x). \end{cases} \tag{2.8}$$

Here and in the sequel u_x^i is understood as row vectors. The FBSDE (2.7) and the PDE (2.8) are related via the following nonlinear Feynman–Kac formula:

$$Y_t = u(t, X_t), \quad Z_t = u_x(t, X_t)\sigma(t, X_t, u(t, X_t)). \tag{2.9}$$

In this case the forward SDE in (2.7) becomes

$$X_t = x + \int_0^t \tilde{b}(s, X_s)ds + \int_0^t \tilde{\sigma}(s, X_s)dW_s, \tag{2.10}$$

where $\tilde{b}(t,x) \triangleq b(t,x,u(t,x),u_x(t,x)\sigma(t,x,u(t,x)))$ and $\tilde{\sigma}(t,x) \triangleq \sigma(t,x,u(t,x))$. The following definition, which is a variation of the "k-weak solution" in [24], is important.

Definition 2.7 We say that a pair $(\mathbb{P}, Z)$, where $\mathbb{P} \in \mathcal{P}(\Omega)$ and Z is a predictable, square-integrable process defined on $(\Omega, \mathcal{F}, \mathbb{P}; \mathbf{F})$, is a "weak solution at (t,x,y)" if the following hold:

(i) $W_s \triangleq \int_t^s \sigma^{-1}(r,\mathbf{x}_r,\mathbf{y}_r)[d\mathbf{x}_r - b(r,\mathbf{x}_r,\mathbf{y}_r,Z_r)dr]$ is a $\mathbb{P}$-Brownian Motion for $s \ge t$;

(ii) $\mathbb{P}\{\mathbf{x}_t = x,\ \mathbf{y}_t = y\} = 1$;

(iii) $\mathbf{y}_s = y - \int_t^s f(r,\mathbf{x}_r,\mathbf{y}_r,Z_r)dr + \int_t^s Z_r dW_r$, $s \in [t,T]$, $\mathbb{P}$-a.s.

(iv) $\mathbb{P}\{\mathbf{y}_T = g(\mathbf{x}_T)\} = 1$;

We note that for a weak solution at (t,x,y), there is no requirement for $(\mathbb{P}, Z)$ over $[0,t)$. So $(\mathbf{x},\mathbf{y},Z)$ may have arbitrary distribution over $[0,t)$.

We end this section by stating some *Standing Assumptions:*

(**H1**) The coefficients (b,σ,f,g) are measurable and bounded by a common constant K.

(**H2**) $k = d$, and the function σ satisfies

$$[\sigma\sigma^*](t,x,y) \ge \frac{1}{K} I_d, \qquad \forall (t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^m. \tag{2.11}$$

(**H3**) The coefficients (b,f,g) are uniformly continuous in (x,y,z), uniformly in $t \in [0,T]$. Moreover, σ is uniformly continuous in (t,x,y).

In this paper we also need an extra assumption on the modulus of continuity of the coefficient σ, in order to obtain an uniform PDE estimate in Theorem 4.1 below.

(**H4**) There exists a continuous, increasing function $\rho : \mathbb{R}_+ \mapsto \mathbb{R}_+$, satisfying

$$\rho(0) = 0; \quad \text{and} \quad \int_{0+} \frac{\rho(t)}{t} dt < \infty, \tag{2.12}$$

such that for all $(t,x_i,y_i) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^m$, $i = 1,2$, it holds that

$$|\sigma(t,x_1,y_1) - \sigma(t,x_2,y_2)| \le \rho(|x_1 - x_2| + |y_1 - y_2|). \tag{2.13}$$

Remark 2.8 There are typically two situations where strong well-posedness may fail and weak solutions are in order: either the coefficients have "bad" growth conditions or they have "bad" regularity conditions. In this paper we shall focus mainly on the regularity of coefficients, and the assumption (**H1**) is merely for technical convenience. We believe that the cases of unbounded coefficients (e.g. linear growth with (y,z) or even quadratic growth in z) can be treated by combining various methods in the literature (see, e.g., Delarue–Guatteri [11] for f having quadratic growth in Z). But we prefer to leave such discussion to a separate work, in order to keep our main focus. □

3 Existence of semi-strong solutions

We begin by pointing out that in [24] the existence result was established only in the one dimensional case. The constraint there was due to an application of Nash's result [25]. In this section we shall prove a more general existence result by using the following a priori estimate on the solution to PDE (2.8).

Theorem 3.1 *Assume* (H1)–(H3)*, and that the coefficients b, σ, f, g are smooth with bounded derivatives. Let u be the unique classical solution to* (2.8)*. Then for any $\delta > 0$ and $\alpha \in (0, 1)$, there exist constants C, C_δ, and $C_{\delta,\alpha}$, depending only on the bounds in* (H1)–(H3)*, the time duration $T > 0$, the dimensions d, m, and the constants δ, α when necessary, but not on the derivatives of the coefficients, such that the following estimates hold*:

(i) $|u(t, x)| \le C$*, for all* $(t, x) \in [0, T] \times \mathbb{R}^d$*;*
(ii) $|u_x(t, x)| \le C_\delta$*, for all* $(t, x) \in [0, T - \delta] \times \mathbb{R}^d$*;*
(iii) *for all* $t_1, t_2 \in [0, T - \delta]$ *and* $x_1, x_2 \in \mathbb{R}^d$*,*

$$|u_x(t_1, x_1) - u_x(t_2, x_2)| \le C_{\delta,\alpha}\left[|x_1 - x_2|^\alpha + |t_1 - t_2|^{\frac{\alpha}{2}}\right]. \tag{3.1}$$

(iv) *Moreover, if $K_g \triangleq \|g_x\|_\infty + \|g_{xx}\|_\infty < \infty$, then (ii) and (iii) hold for $\delta = 0$, with constants C_0 and $C_{0,\alpha}$ there depending on K_g as well.*

The proof is quite technical and mainly analytic, we thus postpone it to Sect. 6. We now have the following existence result.

Theorem 3.2 *Assume* (H1)–(H3)*. There exists a function u satisfying* (i)–(iii) *in Theorem 3.1 and the FBSDE* (2.7) *admits a semi-strong solution such that* (2.9) *holds true. In particular, for any $\delta > 0$, $|Z_t| \le C_\delta$ for $t \in [0, T - \delta]$.*

Moreover, if the coefficient b is also independent of Z, and b and σ are uniformly Lipschitz continuous in (x, y), then the semi-strong solution is actually strong.

Proof We fix a Brownian set-up $(\Omega, \mathcal{F}, P)$. Let $(b_n, \sigma_n, f_n, g_n)$ be the standard smooth mollifiers of (b, σ, f, g) such that they satisfy (H1)–(H3) uniformly and converge to (b, σ, f, g) uniformly. Let u^n denote the unique classical solution to (2.8) with coefficients $(b_n, \sigma_n, f_n, g_n)$, and (X^n, Y^n, Z^n) the unique strong solution to (2.7) with coefficients $(b_n, \sigma_n, f_n, g_n)$ on $(\Omega, \mathcal{F}, P)$. By Theorem 3.1, u^n satisfy (i)–(iii) in Theorem 3.1 uniformly. Applying the Arzela–Ascoli Theorem we see that, possibly along a subsequence, u^n converges to some function u uniformly and u also satisfies (i)–(iii) in Theorem 3.1. Now following the arguments in [24] Theorem 3.1, in particular noting that (3.1) implies that Z^n satisfies the key condition in [24]:

$$\lim_{\varepsilon\to 0}\sup_n E\left\{\int_0^{T-\delta} |Z_t^n - Z_t^{n,\varepsilon}|^2 dt\right\} = 0, \quad Z_t^{n,\varepsilon} \triangleq \frac{1}{\varepsilon}\int_{(t-\varepsilon)^+}^{t} Z_s^n ds, \quad \forall \delta > 0,$$

we know that (X^n, Y^n, Z^n) converges to (X, Y, Z) weakly on any subinterval $[0, T-\delta]$. Thus, following a standard "diagonalization" scheme and using the continuity of

X and Y we can show that (X, Y, Z) is a weak solution to (2.7) on $[0, T]$ and that (2.9) holds true. This implies that (X, Y, Z) is in fact a semi-strong solution. Moreover, by (2.9) and Theorem 3.1 (ii) we conclude that Z_t is bounded for $t \le T - \delta$.

To prove the last part of the Theorem, we assume the coefficient b is also independent of Z. Now by Theorem 3.1 (ii), we see that $\tilde{b}$ and $\tilde{\sigma}$ in SDE (2.10) are uniform Lipschitz continuous on any closed subinterval of $[0, T)$. Therefore, a standard argument would show that the forward SDE (2.10) will have a *strong* solution X on $[0, T]$, which in turn will guarantee the existence of the strong solution to the FBSDE. Namely the aforementioned semi-strong solution is actually a strong one. □

4 The properties of the "nodal set"

In this section we try to characterize the set of all weak solutions in terms of their "initial data", inspired by the so-called "nodal set" introduced in [22]. The results presented here will be important for the uniqueness proof in next section. To begin with, we give an a priori estimate for the second derivative of the solution to PDE (2.8), whose proof is again postponed to Sect. 6 in order not to disturb the discussion.

Theorem 4.1 *Assume* (H1)–(H4) *and the coefficients b, σ, f, g are smooth with bounded derivatives. Let u be the unique classical solution to* (2.8). *Let*

$$K_0 \triangleq \|b_x\|_\infty + \|b_y\|_\infty + \|b_z\|_\infty + \|f_x\|_\infty + \|f_y\|_\infty + \|f_z\|_\infty + \|g_x\|_\infty + \|g_{xx}\|_\infty.$$

Then, there exists a constant $C > 0$, depending only on the bounds in (H1)–(H4)*, the time duration $T > 0$, the dimensions d, m, as well as K_0, but not on the derivatives of σ, such that*

$$|u_{xx}(t, x)| \le C; \qquad \forall (t, x) \in [0, T] \times \mathbb{R}^d.$$

We now define the "nodal sets".

Definition 4.2 (i) Let $\mathcal{O}(t, x, y)$ denote the space of weak solutions (X, Y, Z) at (t, x, y), in the sense of Definition 2.7.

(ii) $\mathcal{O}(t, x) \triangleq \{y : \mathcal{O}(t, x, y) \text{ is not empty}\}$.

(iii) $\mathcal{O} \triangleq \{(t, x, y) : y \in \mathcal{O}(t, x)\}$.

By Theorem 3.2, for the function u there we have $u(t, x) \in \mathcal{O}(t, x)$ for any (t, x).

In order to study the uniqueness, from now on we always assume $m = 1$. Then (2.8) becomes the following PDE:

$$u_t + h(t, x, u, u_x, u_{xx}) = 0, \quad u(T, x) = g(x); \tag{4.1}$$

where

$$\begin{aligned} h(t, x, y, z, \gamma) \triangleq\ & \frac{1}{2} \mathrm{tr}\,(\gamma \sigma \sigma^*(t, x, y)) + z b(t, x, y, z\sigma(t, x, y)) \\ & + f(t, x, y, z\sigma(t, x, y)). \end{aligned} \tag{4.2}$$

We first note that, in this case the function u in Theorem 3.2 is in fact a viscosity solution to PDE (4.1). The following result is the key for the uniqueness.

Theorem 4.3 *Assume* (H1)–(H4).

(i) *For any* $y \in \mathcal{O}(t,x)$, *there exists a semi-strong solution* $(X, Y, Z) \in \mathcal{O}(t,x,y)$ *such that for any* $\delta > 0$, $|Z_s| \le C_\delta$ *for* $s \in [t, T-\delta]$.
(ii) *For any* (t,x), $\mathcal{O}(t,x)$ *is a bounded closed interval.*

To prove the theorem, we need the following comparison result for FBSDEs.

Theorem 4.4 *Assume* (H1)–(H4). *Then for each* $n \in \mathbb{N}$ *there exist functions* $(b^n, \sigma^n, f^n, g^n)$ *such that*

(i) $(b^n, \sigma^n, f^n, g^n)$ *are smooth and satisfy* (*H1*)–(*H4*) *uniformly*;
(ii) $(b^n, \sigma^n, f^n, g^n)$ *converges to* (b, σ, f, g) *uniformly*;
(iii) *For any weak solution* (X, Y, Z) *of* (2.7) *and any* n, *it holds that* $Y_0^n \ge Y_0$, *where* (X^n, Y^n, Z^n) *is the unique strong solution to* (2.7) *with coefficients* $(b^n, \sigma^n, f^n, g^n)$.

Proof For any n, and any $c_{n,1} > 0$, $c_{n,2} > 0$, we mollify the coefficients (b, σ, f, g) to get $(b^n, \sigma^n, f^n, g^n)$ so that

$$\frac{1}{n} \le f^n - f \le \frac{2}{n}; \quad 0 \le g^n - g \le \frac{1}{n}; \quad |b^n - b| \le c_{n,1}; \quad |\sigma^n - \sigma| \le c_{n,2};$$

and

$$|f_x^n| + |f_y^n| + |f_z^n| + |g_x^n| \le Cn, \quad |g_{xx}^n| \le Cn^2, \quad |b_x^n| + |b_y^n| + |b_z^n| \le \frac{C}{c_{n,1}}.$$

Let u^n be the classical solution to PDE (4.1) with coefficients $(b^n, \sigma^n, f^n, g^n)$. Then $Y_0^n = u^n(0,x)$. Now, by Theorem 3.1 (i) and (iv) we have

$$|u^n| \le C; \quad |u_x^n| \le C_{n,0}, \tag{4.3}$$

where $C_{n,0}$ depends on n but not on $c_{n,1}, c_{n,2}$. Moreover, by Theorem 4.1, we have

$$|u_{xx}^n| \le C_{n,1}, \tag{4.4}$$

where $C_{n,1}$ depends on n and $c_{n,1}$, but not on $c_{n,2}$.

For any weak solution (X, Y, Z) of (2.7), denote

$$\Delta Y_t^n \stackrel{\triangle}{=} u^n(t, X_t) - Y_t; \quad \Delta Z_t^n \stackrel{\triangle}{=} u_x^n(t, X_t)\sigma(t, X_t, Y_t) - Z_t.$$

For $\varphi = b, \sigma, f, g$, denote $\Delta\varphi^n \stackrel{\triangle}{=} \varphi^n - \varphi$. To simplify notations in the rest of the proof we assume that $d = k = 1$ as well. The general cases are the same, except for some tedious notational differences. Also, for notational simplicity in what follows we let α_t denote a generic *bounded* process whose bound is independent of n. Applying Itô's formula we have (often suppressing variables when context is clear):

$$
\begin{aligned}
&d\Delta Y_t^n\\
&\quad = \left\{u_t^n + u_x^n b(t, X_t, Y_t, Z_t) + \frac{1}{2}u_{xx}^n \sigma^2(t, X_t, Y_t) + f(t, X_t, Y_t, Z_t)\right\} dt\\
&\qquad + \Delta Z_t^n dW_t\\
&\quad = -\left[\frac{1}{2}u_{xx}^n[\sigma^n]^2(t, X_t, u^n) + u_x^n b^n(t, X_t, u^n, u_x^n\sigma^n(t, X_t, u^n))\right.\\
&\qquad\qquad \left. + f^n(t, X_t, u^n, u_x^n\sigma^n(t, X_t, u^n))\right] dt\\
&\qquad + \left[u_x^n b(t, X_t, Y_t, Z_t) + \frac{1}{2}u_{xx}^n\sigma^2(t, X_t, Y_t) + f(t, X_t, Y_t, Z_t)\right] dt + \Delta Z_t^n dW_t\\
&\quad = \Big\{[\alpha_t u_{xx}^n \Delta\sigma^n(t, X_t, Y_t) + \alpha_t u_{xx}^n \bar{\sigma}_y^n \Delta Y_t^n]\\
&\qquad - u_x^n([\Delta b^n(t, X_t, Y_t, Z_t) + \bar{b}_y^n \Delta Y_t^n + \bar{b}_z^n \Delta Z_t^n + \bar{b}_z^n u_x^n \Delta\sigma^n + \bar{b}_z^n u_x^n \bar{\sigma}_y^n \Delta Y_t^n]\\
&\qquad - [\Delta f^n(t, X_t, Y_t, Z_t) + \bar{f}_y^n \Delta Y_t^n + \bar{f}_z^n \Delta Z_t^n + \bar{f}_z^n u_x^n \Delta\sigma^n + \bar{f}_z^n u_x^n \bar{\sigma}_y^n \Delta Y_t^n]\Big\} dt\\
&\qquad + \Delta Z_t^n dW_t\\
&\quad = -[\alpha_t^n + \beta_t^n \Delta Y_t^n + \gamma_t^n \Delta Z_t^n]dt + \Delta Z_t^n dW_t.
\end{aligned}
$$

In the above we used the fact: $\varphi(x) - \varphi(y) = [\int_0^1 \varphi_x(y + \theta(x-y))d\theta](x-y)$, for any C^1 function φ, and denoted $\bar{\sigma}_y^n$ to be the process $\bar{\sigma}_y^n(t) \triangleq \int_0^1 \sigma_y^n(t, X_t, Y_t + \theta\Delta Y_t^n)d\theta$ (similarly for $\bar{b}_y^n$, $\bar{b}_z^n$, $\bar{f}_y^n$, …, etc., with appropriate modifications); and

$$
\begin{aligned}
\alpha_t^n &\triangleq [-\alpha_t u_{xx}^n + (u_x^n)^2\bar{b}_z^n + \bar{f}_z^n u_x^n]\Delta\sigma^n(t, X_t, Y_t) + [u_x^n \Delta b^n + \Delta f^n](t, X_t, Y_t, Z_t);\\
\beta_t^n &\triangleq -\alpha_t u_{xx}^n \bar{\sigma}_y^n + u_x^n \bar{b}_y^n + (u_x^n)^2 \bar{b}_z^n \bar{\sigma}_y^n + \bar{f}_y^n + \bar{f}_z^n u_x^n \bar{\sigma}_y^n;\\
\bar{\gamma}_t^n &\triangleq \bar{u}_x^n \bar{b}_z^n + \bar{f}_z^n.
\end{aligned}
$$

Note that, by virtue of (4.3) and (4.4),

$$
\alpha_t^n \geq -C_{n,1}c_{n,2} - C_{n,0}c_{n,1} - \frac{C_{n,0}^2}{c_{n,1}}c_{n,2} + \frac{1}{n} - nC_{n,0}c_{n,2},
$$

we can first choose $c_{n,1}$ and then choose $c_{n,2}$ such that

$$
C_{n,0}c_{n,1} \leq \frac{1}{2n}; \qquad \left[C_{n,1} + \frac{C_{n,0}^2}{c_{n,1}} + nC_{n,0}\right]c_{n,2} \leq \frac{1}{2n}.
$$

Then $\alpha_t^n \geq 0$ and thus

$$
d\Delta Y_t^n \leq -[\beta_t^n \Delta Y_t^n + \gamma_t^n \Delta Z_t^n]dt + \Delta Z_t^n dW_t.
$$

Moreover, since $\Delta Y^n_T = \Delta g^n(X_T) \ge 0$, we get $\Delta Y^n_t \ge 0$. In particular,

$$Y^n_0 - Y_0 = u^n(0,x) - Y_0 = \Delta Y^n_0 \ge 0.$$

This proves the theorem. □

Proof of Theorem 4.3 (i) Without loss of generality, let us assume $t = 0$. Let x be fixed, and let $(b^{n,0}, \sigma^{n,0}, f^{n,0}, g^{n,0})$ be the mollifiers constructed in Theorem 4.4. We similarly construct mollifiers $(b^{n,1}, \sigma^{n,1}, f^{n,1}, g^{n,1})$ that approximate (b, σ, f, g) from below. Denote, for $\varphi = b, \sigma, f, g$ and $\alpha \in [0,1]$,

$$\varphi^{n,\alpha} \triangleq (1-\alpha)\varphi^{n,0} + \alpha\varphi^{n,1}.$$

Let $(X^{n,\alpha}, Y^{n,\alpha}, Z^{n,\alpha})$ be the unique strong solution to the FBSDE (2.7) with coefficients $(b^{n,\alpha}, \sigma^{n,\alpha}, f^{n,\alpha}, g^{n,\alpha})$. By Theorem 4.4, for any $y \in \mathcal{O}(0,x)$, $Y^{n,1}_0 \le y \le Y^{n,0}_0$. It is readily seen that, for fixed n, $Y^{n,\alpha}_0$ is continuous in α. Thus one can find α_n such that $Y^{n,\alpha_n}_0 = y$. Let u^n be the classical solution to the PDE (2.8) with coefficients $(b^{n,\alpha_n}, \sigma^{n,\alpha_n}, f^{n,\alpha_n}, g^{n,\alpha_n})$. Then u^n satisfies the estimates in Theorem 3.1 uniformly, and $u^n(0,x) = y$. Possibly along a subsequence, we have $u^n \to u$ for some u satisfying by the estimates in Theorem 3.1. Now following the same arguments in Theorem 3.2, there exists a semi-strong solution (X, Y, Z) such that (2.9) holds and $|Z_t| \le C_\delta$ for ant $t \le T - \delta$. It is clear that $u(0,x) = y$, then (X, Y, Z) is a semi-strong solution at $(0, x, y)$.

(ii) First, since f and g are bounded, thus the component Y of the solution to the BSDE in (2.7) must be bounded, uniformly for all initial state x. Thus $\mathcal{O}(0,x)$ is bounded. Next, for any $y_1, y_2 \in \mathcal{O}(0,x)$ with $y_1 < y_2$, and any $y \in [y_1, y_2]$, we have $Y^{n,1}_0 \le y_1 \le y \le y_2 \le Y^{n,0}_0$. Then there exists α_n such that $Y^{n,\alpha_n}_0 = y$. Following the same arguments as in (i) we know $y \in \mathcal{O}(0,x)$. That is, the set $\mathcal{O}(0,x)$ must be an interval. Finally, let $\bar{y}$ denote the right end of $\mathcal{O}(0,x)$. Since $Y^{n,1}_0 \le y \le Y^{n,0}_0$ for any $y \in \mathcal{O}(0,x)$, it holds that $Y^{n,1}_0 \le \bar{y} \le Y^{n,0}_0$. Then the similar arguments imply that $\bar{y} \in \mathcal{O}(0,x)$. Similarly, $\mathcal{O}(0,x)$ is also closed at the left end. That is, $\mathcal{O}(0,x)$ is a closed bounded interval, proving the theorem. □

4.1 The decoupled case

We should note that although the proof of Theorem 4.3 does not depend on (H4) directly, the assumption is important in proving Theorem 4.4 (actually Theorem 4.1). On the other hand, it is well known that the comparison result holds for decoupled FBSDEs, without assuming (H4). To conclude this section we shall prove some stronger results in the simplified decoupled case.

We begin by assuming that b, σ are independent of (y, z) and satisfy (H1)–(H3). For any (t, x), let X be the unique (in law) weak solution to the forward SDE in (2.7) with $X_t = x$ (see [31] for existence and uniqueness of X). Then it follows from (H2) that $\mathbf{F}^W \subseteq \mathbf{F}^X$. We note that under these conditions in general the distribution of X_T may be singular to the Lebesgue measure (see [12] for a counterexample).

We first prove a comparison result for BSDEs.

Lemma 4.5 (i) *Assume* (H1)–(H3) *and FBSDE* (2.7) *is decoupled. If f is uniformly Lipschitz continuous in (y, z), then FBSDE* (2.7) *has a semi-strong solution (X, Y, Z) and it is the unique (in law) weak solution.*

(ii) *Assume further that $(\tilde{f}, \tilde{g})$ also satisfy* (H1)–(H3) *and $\tilde{f} \le f$ and $\tilde{g} \le g$. Let $(X, \tilde{Y}, \tilde{Z})$ be an arbitrary weak solution to FBSDE* (2.7) *with coefficients $(b, \sigma, \tilde{f}, \tilde{g})$ and assume that (X, Y, Z) and $(X, \tilde{Y}, \tilde{Z})$ are on the same probability space with the same process X (and W). Then $\tilde{Y}_t \le Y_t$, for all $t \in [0, T]$, $\mathbb{P}$-a.s.*

Proof (i) The existence of a semi-strong solution follows from Theorem 3.2, and the uniqueness of weak solution follows immediately from the comparison result at (ii).

(ii) Let (X, Y, Z) be the semi-strong solution to FBSDE (2.7) with coefficients (b, σ, f, g) constructed in Theorem 3.2. Denote

$$\Delta Y_t \triangleq \tilde{Y}_t - Y_t, \quad \Delta Z_t \triangleq \tilde{Z} - Z, \quad \Delta f \triangleq \tilde{f} - f, \quad \Delta g \triangleq \tilde{g} - g.$$

Since f is uniformly Lipschitz continuous in (y, z), there exist bounded processes α, β such that

$$\Delta Y_t = (\Delta g)(X_T) + \int_t^T [(\Delta f)(s, X_s, \tilde{Y}_s, \tilde{Z}_s) + \alpha_s \Delta Y_s + \beta_s \Delta Z_s] ds - \int_t^T \Delta Z_s dW_s.$$

Since $\Delta f \le 0$, $\Delta g \le 0$, by standard arguments in BSDE literature we obtain that $\Delta Y_t \le 0$, for all $t \in [0, T]$, $\mathbb{P}$-a.s. □

We emphasize that in the proof of (ii) above we did not assume the uniqueness of semi-strong solution in (i), so there is no danger of cycle proof here.

Theorem 4.6 *Assume* (H1)–(H3) *and FBSDE* (2.7) *is decoupled. Then there exist two functions $\bar{u}$ and $\underline{u}$ satisfying Theorem* 3.1 (i)–(iii) *such that*

(i) $\mathcal{O}(t, x) = [\underline{u}(t, x), \bar{u}(t, x)]$.

(ii) *Let X be the unique (in law) weak solution to the FSDE in* (2.7)*, and $(\bar{Y}, \bar{Z})$ and $(\underline{Y}, \underline{Z})$ be defined by* (2.9) *corresponding to $\bar{u}$ and $\underline{u}$, respectively. Then $(X, \bar{Y}, \bar{Z})$ and $(X, \underline{Y}, \underline{Z})$ are two semi-strong solutions to FBSDE* (2.7).

(iii) *For any $y \in \mathcal{O}(t, x)$ and any $(X, Y, Z) \in \mathcal{O}(t, x, y)$, we have $\underline{Y}_s \le Y_s \le \bar{Y}_s$.*

(iv) *For any $y \in \mathcal{O}(t, x)$, there exists a semi-strong solution $(X, Y, Z) \in \mathcal{O}(t, x, y)$ such that for any $\delta > 0$, $|Z_s| \le C_\delta$ for $s \in [t, T - \delta]$.*

Proof Let $\bar{f}_n, \bar{g}_n$ be standard smooth molifiers of f, g such that $\bar{f}_n \downarrow f, \bar{g}_n \downarrow g$. By Lemma 4.5 (i) there exist $\bar{u}^n$ satisfying Theorem 3.1 (i)–(iii) and the corresponding semi-strong solution $(\bar{Y}^n, \bar{Z}^n)$. Then by the arguments in Theorem 3.2 we can assume, without loss of generality, that $\bar{u}^n \to \bar{u}, \bar{u}^n_x \to \bar{u}_x$ for some function $\bar{u}$ satisfying Theorem 3.2 (i)–(iii). By the comparison result in Lemma 4.5 (ii), one can easily see that $\bar{u}(t, x)$ is an upper bound of $\mathcal{O}(t, x)$ and $\bar{Y}^n$ is decreasing in n. Then $\bar{Y}^n_t \downarrow \bar{Y}_t = \bar{u}(t, X_t)$. Similarly, by using smooth molifiers $\underline{f}_n, \underline{g}_n$ such that

$\underline{f}_n \uparrow f$, $\underline{g}_n \uparrow g$, we may obtain $\underline{u}$ and the corresponding $(\underline{Y}, \underline{Z})$. Then following the arguments in Theorem 4.3 one can easily prove the results. □

5 Uniqueness of weak solutions

We now turn our attention to the key issue of the paper: the uniqueness of the solution to FBMP. Again, we shall assume $m = 1$. In [24] we proved the uniqueness among the class of solutions whose component Z is bounded in some sense. We aim to remove this constraint here. Throughout this section we assume (H1)–(H4).

Recall from Sect. 2 the canonical space $\Omega \stackrel{\triangle}{=} \mathbb{C}([0,T];\mathbb{R}^d) \times \mathbb{C}([0,T];\mathbb{R})$. Let $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ denote the filtration generated by the canonical processes, which we shall denote by $(\mathbf{x}, \mathbf{y})$. We recall that for any given probability measure $\mathbb{P} \in \mathcal{P}(\Omega)$ and any $t < T$, there exists a *regular conditional probability distribution* (r.c.p.d. for short) of $\mathbb{P}$ given $\mathcal{F}_t$, denoted by $\mathbb{P}^{\omega}_t$, $\omega \in \Omega$, in the sequel (see e.g. [31]). Furthermore, we can choose a version of $\mathbb{P}^{\omega}_t$ so that $\mathbb{P}^{\omega}_t \in \mathcal{P}(\Omega)$ for all $\omega \in \Omega$. In what follows we will always take such a version without further specification.

Recall Definition 4.2. Let $\bar{\mathcal{O}}$ denote the closure of $\mathcal{O}$ (we note that it is not clear to us whether or not $\mathcal{O}$ is Lebesgue measurable!). Clearly, for the viscosity solution u to (4.1) constructed in Theorem 3.2, we have $(t, x, u(t,x)) \in \mathcal{O}$ for any $(t,x) \in [0,T] \times \mathbb{R}^d$. Now define two functions on $(t,x) \in [0,T] \times \mathbb{R}^d$:

$$\underline{u}(t,x) \stackrel{\triangle}{=} \inf\{y : (t,x,y) \in \bar{\mathcal{O}}\}; \quad \bar{u}(t,x) \stackrel{\triangle}{=} \sup\{y : (t,x,y) \in \bar{\mathcal{O}}\}. \tag{5.1}$$

We remark that, when the FBSDE is decoupled, $\underline{u}, \bar{u}$ in (5.1) are the same as those in Theorem 4.6.

Lemma 5.1 *For some constant C_0, it holds*

$$-C_0 \leq \underline{u}(t,x) \leq u(t,x) \leq \bar{u}(t,x) \leq C_0; \quad \underline{u}(T,x) = \bar{u}(T,x) = g(x). \tag{5.2}$$

Proof First, by Theorem 4.3 (ii) where the bound is independent of (t,x), we know $\underline{u}(t,x), \bar{u}(t,x)$ are bounded.

Second, for any $(T,x,y) \in \bar{\mathcal{O}}$, assume $(t_n, x_n, y_n) \in \mathcal{O}$ and $(t_n, x_n, y_n) \to (T,x,y)$. Let $(\mathbb{P}^n, Z^n)$ be a weak solution at (t_n, x_n, y_n) and W^n be the corresponding $\mathbb{P}^n$-Brownian motion. Then

$$\begin{cases} \mathbf{x}_T = x_n + \displaystyle\int_{t_n}^{T} b(s, \mathbf{x}_s, \mathbf{y}_s, Z_s) ds + \int_{t_n}^{T} \sigma(s, \mathbf{x}_s, \mathbf{y}_s) dW^n_s; \\ y_n = g(\mathbf{x}_T) + \displaystyle\int_{t_n}^{T} f(s, \mathbf{x}_s, \mathbf{y}_s, Z_s) ds - \int_{t_n}^{T} Z_s dW^n_s; \end{cases} \quad \mathbb{P}^n\text{-a.s.}$$

Thus, applying standard arguments we have

$$
\begin{aligned}
|y_n - g(x_n)|^2 &= \left| E^{\mathbb{P}^n} \left\{ g(\mathbf{x}_T) - g(x_n) + \int_{t_n}^{T} f(s, \mathbf{x}_s, \mathbf{y}_s, Z_s) ds \right\} \right|^2 \\
&\le 2E^{\mathbb{P}^n} \left\{ |g(\mathbf{x}_T) - g(x_n)|^2 + \left| \int_{t_n}^{T} f(s, \mathbf{x}_s, \mathbf{y}_s, Z_s) ds \right|^2 \right\} \\
&\le C E^{\mathbb{P}^n} \left\{ \left| g\left(x_n + \int_{t_n}^{T} b(s, \mathbf{x}_s, \mathbf{y}_s, Z_s) ds + \int_{t_n}^{T} \sigma(s, \mathbf{x}_s, \mathbf{y}_s) dW_s^n \right) \right. \right. \\
&\qquad \left. \left. - g(x_n) \right|^2 \right\} + C|T - t_n|^2.
\end{aligned}
$$

Now note that by (H1) g is bounded and uniformly continuous, a standard argument using Chebyshev's inequality and the boundedness of σ, one shows easily that $\lim_{n\to\infty} |y_n - g(x_n)| = 0$. To wit, $y = g(x)$. □

Moreover, since $\bar{\mathcal{O}}$ is a closed set, we have $(t, x, \underline{u}(t, x)) \in \bar{\mathcal{O}}$ and $(t, x, \bar{u}(t, x)) \in \bar{\mathcal{O}}$. By the same arguments as in [24], we have

Lemma 5.2 *$\underline{u}$ is lower semi-continuous and $\bar{u}$ is upper semi-continuous.*

Our main result of this section is the following theorem.

Theorem 5.3 *Assume* (H1)–(H4)*, and b and f are uniformly continuous in t. Then, $\underline{u}$ and $\bar{u}$ are viscosity supersolution and subsolution, respectively, of the PDE* (4.1).

Remark 5.4 We note that the uniform continuity of b and f in t is merely for technical reasons due to our methodology. We believe that in general one can extend the result by considering the semi-continuous Hamiltonian (cf. e.g. [4, 15]) and arguing, for example, that $\underline{u}$ is a viscosity *supersolution* of $u_t + \underline{h}(t, x, u, u_x, u_{xx}) = 0$ and $\bar{u}$ is a viscosity *subsolution* of $u_t + \bar{h}(t, x, u, u_x, u_{xx}) = 0$, where, with h being defined in (4.2),

$$
\underline{h}(t, x, y, z, \gamma) \stackrel{\triangle}{=} \varliminf_{t'\to t} h(t', x, y, z, \gamma), \quad \bar{h}(t, x, y, z, \gamma) \stackrel{\triangle}{=} \varlimsup_{t'\to t} h(t', x, y, z, \gamma).
$$

But this would require that some of the arguments be fine tuned. Due to the length of the paper, we do not pursue such a generality here. □

Proof of Theorem 5.3 Without loss of generality we only check $\underline{u}$. The arguments are similar to those in [24]. For any $(t_0, x_0) \in [0, T) \times \mathbb{R}^d$, let $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^d)$ be such that $y_0 \triangleq \underline{u}(t_0, x_0) = \varphi(t_0, x_0)$ and $\underline{u}(t, x) \geq \varphi(t, x)$, for all $(t, x) \in [0, T] \times \mathbb{R}^d$. We shall prove that

$$[\mathcal{L}\varphi](t_0, x_0, \varphi(t_0, x_0)) \leq 0. \tag{5.3}$$

where, recalling (4.2) again,

$$[\mathcal{L}\varphi](t, x, y) \triangleq \varphi_t(t, x) + h(t, x, y, \varphi_x(t, x), \varphi_{xx}(t, x)).$$

To prove (5.3), let us first recall (5.2). Without loss of generality we may assume that $\varphi(t, x) = -C_0 - 1$ for x outside of a compact set. Then φ, φ_t, φ_x, and φ_{xx} are all bounded and uniformly continuous. Next we fix $\delta_0 > 0$ such that $t_0 < T - \delta_0$. Note that $(t_0, x_0, y_0) = (t_0, x_0, \underline{u}(t_0, x_0)) \in \bar{\mathcal{O}}$, so for each n there exists a $(t_n, x_n, y_n) \in \mathcal{O}$ such that

$$|t_n - t_0| + |x_n - x_0| + |y_n - y_0| \leq \frac{1}{n}. \tag{5.4}$$

Without loss of generality we assume $t_n < T - \delta_0$. Applying Theorem 4.3, we can find a weak solution at (t_n, x_n, y_n), denoted by $(\mathbb{P}^n, Z^n)$, such that

$$|Z_t^n| \leq C_{\delta_0}, \quad \forall t \in [t_n, T - \delta_0]. \tag{5.5}$$

Let W^n denote the corresponding $\mathbb{P}^n$-Brownian motion. For $t \in (t_n, T)$, it is readily seen that $(\mathbb{P}_t^{n,\omega}, Z^n)$ is a weak solution at $(t, \mathbf{x}_t, \mathbf{y}_t)$, $\mathbb{P}^n$-a.s. $\omega \in \Omega$. In other words, we must have $(t, \mathbf{x}_t, \mathbf{y}_t) \in \mathcal{O}$, $\mathbb{P}^n$-a.s, and consequently $\mathbf{y}_t \geq \underline{u}(t, \mathbf{x}_t) \geq \varphi(t, \mathbf{x}_t)$, $\mathbb{P}^n$-a.s. , $\forall t \geq t_n$.

Now let us denote

$$\Delta Y_t \triangleq \varphi(t, \mathbf{x}_t) - \mathbf{y}_t; \quad \Delta Z_t^n \triangleq \varphi_x(t, \mathbf{x}_t)\sigma(t, \mathbf{x}_t, \mathbf{y}_t) - Z_t^n.$$

Also, for any $\varepsilon > 0$, let b_ε, f_ε be smooth mollifiers of b, f such that

$$\|b_\varepsilon - b\|_\infty \leq \varepsilon; \quad \|f_\varepsilon - f\|_\infty \leq \varepsilon; \quad \|\partial_z b_\varepsilon\|_\infty \leq \frac{C}{\varepsilon}; \quad \|\partial_z f_\varepsilon\|_\infty \leq \frac{C}{\varepsilon}.$$

Denote

$$\begin{aligned}
\alpha_t^{n,\varepsilon} \triangleq\ & \varphi_x(t, \mathbf{x}_t)[b(t, \mathbf{x}_t, \mathbf{y}_t, Z_t^n) - b(t, \mathbf{x}_t, \mathbf{y}_t, \varphi_x \sigma(t, \mathbf{x}_t, \mathbf{y}_t))] \\
& - \varphi_x(t, \mathbf{x}_t)[b_\varepsilon(t, \mathbf{x}_t, \mathbf{y}_t, Z_t^n) - b_\varepsilon(t, \mathbf{x}_t, \mathbf{y}_t, \varphi_x \sigma(t, \mathbf{x}_t, \mathbf{y}_t))] \\
& + [f(t, \mathbf{x}_t, \mathbf{y}_t, Z_t^n) - f(t, \mathbf{x}_t, \mathbf{y}_t, \varphi_x \sigma(t, \mathbf{x}_t, \mathbf{y}_t))] \\
& - [f_\varepsilon(t, \mathbf{x}_t, \mathbf{y}_t, Z_t^n) - f_\varepsilon(t, \mathbf{x}_t, \mathbf{y}_t, \varphi_x \sigma(t, \mathbf{x}_t, \mathbf{y}_t))]; \\
\beta_t^{n,\varepsilon} \triangleq\ & \int_0^1 [\varphi_x(t, \mathbf{x}_t)\partial_z b_\varepsilon(t, \mathbf{x}_t, \mathbf{y}_t, Z_t^n + \theta \Delta Z_t^n) + \partial_z f_\varepsilon(t, \mathbf{x}_t, \mathbf{y}_t, Z_t^n + \theta \Delta Z_t^n)] d\theta.
\end{aligned}$$

Then, it holds that

$$|\alpha_t^{n,\varepsilon}| \le C\varepsilon, \qquad |\beta_t^{n,\varepsilon}| \le \frac{C}{\varepsilon}; \tag{5.6}$$

where $C > 0$ may depend on φ. Furthermore, applying Itô's formula and using the definition of $\mathcal{L}\varphi$, $\alpha^{n,\varepsilon}$, and $\beta^{n,\varepsilon}$ we have

$$\begin{aligned} d\Delta Y_t =& \left[\varphi_t + \varphi_x b(t,\mathbf{x}_t,\mathbf{y}_t,Z_t^n) + \frac{1}{2}\mathrm{tr}\,(\sigma\sigma^*(t,\mathbf{x}_t,\mathbf{y}_t)\varphi_{xx}) + f(t,\mathbf{x}_t,\mathbf{y}_t,Z_t^n)\right]dt \\ &+\Delta Z_t^n dW_t^n \\ =& [\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)dt + \varphi_x(t,\mathbf{x}_t)[b(t,\mathbf{x}_t,\mathbf{y}_t,Z_t^n) - b(t,\mathbf{x}_t,\mathbf{y}_t,\varphi_x\sigma(t,\mathbf{x}_t,\mathbf{y}_t))]dt \\ &+ [f(t,\mathbf{x}_t,\mathbf{y}_t,Z_t^n) - f(t,\mathbf{x}_t,\mathbf{y}_t,\varphi_x\sigma(t,\mathbf{x}_t,\mathbf{y}_t))]dt + \Delta Z_t^n dW_t^n \\ =& [\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)dt + \alpha_t^{n,\varepsilon}dt - \langle\, \beta_t^{n,\varepsilon}, \Delta Z_t^n\,\rangle\, dt + \Delta Z_t^n dW_t^n. \end{aligned}$$

Now let us denote

$$\Gamma_t^{n,\varepsilon} \stackrel{\triangle}{=} \exp\left\{\int_{t_n}^t \beta_s^{n,\varepsilon} dW_s^n - \frac{1}{2}\int_{t_n}^t |\beta_s^{n,\varepsilon}|^2 ds\right\}, \qquad t \in [t_n, T]$$

One easily checks that, by denoting $E^n \stackrel{\triangle}{=} E^{\mathbb{P}^n}$,

$$\Gamma_{t_n}^{n,\varepsilon} = 1, \quad \Gamma_t^{n,\varepsilon} > 0, \quad E^n\{\Gamma_t^{n,\varepsilon}\} = 1, \quad \text{and} \quad E^n\{|\Gamma_t^{n,\varepsilon}|^2\} \le C_\varepsilon, \quad \forall t \ge t_n. \tag{5.7}$$

Moreover, applying Itô's formula again we have, for $t \in [t_n, T]$,

$$d(\Gamma_t^{n,\varepsilon}\Delta Y_t) = \Gamma_t^{n,\varepsilon}[\mathcal{L}\varphi]dt + \Gamma_t^{n,\varepsilon}\alpha_t^{n,\varepsilon}dt + \Gamma_t^{n,\varepsilon}[\Delta Z_t^n + \beta_t^{n,\varepsilon}\Delta Y_t]dW_t^n; \tag{5.8}$$

Now, for n large and for any $\delta \in (\frac{1}{n}, T-\delta_0-t_0)$, choose $t \stackrel{\triangle}{=} t_0+\delta \in (t_n, T-\delta_0]$ (see (5.4)), we deduce from (5.8) that

$$0 \ge E^n\left\{\Gamma_{t_0+\delta}^{n,\varepsilon}\Delta Y_{t_0+\delta}\right\} = E^n\left\{\Delta Y_{t_n} + \int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}\{[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t) + \alpha_t^{n,\varepsilon}\}dt\right\}.$$

Therefore, using (5.6) and (5.7) we get

$$\begin{aligned} E^n\left\{\int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)dt\right\} &\le -E^n\left\{\Delta Y_{t_n} + \int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}\alpha_t^{n,\varepsilon}dt\right\} \\ &\le E^n\left\{|y_n - y_0| + |\varphi(t_0,x_0) - \varphi(t_n,x_n)| + \int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}|\alpha_t^{n,\varepsilon}|dt\right\} \end{aligned}$$

$$\le CE^n\left\{\frac{1}{n}+\varepsilon\int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}dt\right\}\le C\left[\varepsilon+\frac{1}{n\delta-1}\right]$$
$$\times(t_0+\delta-t_n), \tag{5.9}$$

By (H1) and (H3), $\mathcal{L}\varphi$ is uniformly continuous in (t,x,y). Let ρ_φ denote the modulus of continuity of $\mathcal{L}\varphi$, and write

$$\Delta_n[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)=\mathcal{L}\varphi(t,\mathbf{x}_t,\mathbf{y}_t)-\mathcal{L}\varphi(t_n,x_n,y_n).$$

We see that (5.9) yields

$$\mathcal{L}\varphi(t_0,x_0,y_0)\le\mathcal{L}\varphi(t_n,x_n,y_n)+\rho_\varphi\left(\frac{1}{n}\right)$$
$$=E^n\left\{\frac{1}{t_0+\delta-t_n}\int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}\mathcal{L}\varphi(t_n,x_n,y_n)dt\right\}+\rho_\varphi\left(\frac{1}{n}\right) \tag{5.10}$$
$$=E^n\left\{\frac{1}{t_0+\delta-t_n}\int_{t_n}^{t_0+\delta}\Gamma_t^{n,\varepsilon}\{[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)-\Delta_n[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)\}dt\right\}+\rho_\varphi\left(\frac{1}{n}\right)$$
$$\le C\varepsilon+\frac{C}{n\delta-1}+\rho_\varphi\left(\frac{1}{n}\right)+\frac{1}{t_{0,n}^\delta}E^n\left\{\int_{t_n}^{t_0+\delta}|\Gamma_t^{n,\varepsilon}\Delta_n[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)|dt\right\},$$

where $t_{0,n}^\delta\overset{\triangle}{=}t_0+\delta-t_n$. To estimate the last term on the right hand side above we first apply Cauchy-Schwartz inequality and the estimate (5.7) to get

$$E^n\int_{t_n}^{t_0+\delta}|\Gamma_t^{n,\varepsilon}\Delta_n[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)|dt\le C_\varepsilon\left\{\sup_{t_n\le t\le t_0+\delta}E^n\{|\Delta_n[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)|^2\}\right\}^{\frac{1}{2}}t_{0,n}^\delta. \tag{5.11}$$

Note that, for any $\eta>0$ and $t\in[t_n,t_0+\delta]$, we apply the Chebychev inequality to get

$$E^n\left\{|\Delta_n[\mathcal{L}\varphi](t,\mathbf{x}_t,\mathbf{y}_t)|^2\right\}$$
$$\le C\rho_\varphi^2(t_{0,n}^\delta)+C\rho_\varphi^2(\eta)+CP^n\left(|\mathbf{x}_t-x_n|+|\mathbf{y}_t-y_n|\ge\eta\right)$$
$$\le C\left[\rho_\varphi^2(t_{0,n}^\delta)+\rho_\varphi^2(\eta)+\frac{1}{\eta^2}E^n\left\{|\mathbf{x}_t-x_n|^2+|\mathbf{y}_t-y_n|^2\right\}\right]$$

$$\leq C\left[\rho_\varphi^2(t_{0,n}^\delta)+\rho_\varphi^2(\eta)+\frac{1}{\eta^2}E^n\left\{\left|\int_{t_n}^t b(s,\mathbf{x}_s,\mathbf{y}_s,Z_s^n)ds\right|^2\right.\right.$$

$$\left.\left.+\int_{t_n}^t \mathrm{tr}\,(\sigma\sigma^*(s,\mathbf{x}_s,\mathbf{y}_s))ds+\left|\int_{t_n}^t f(s,\mathbf{x}_s,\mathbf{y}_s,Z_s^n)ds\right|^2+\int_{t_n}^t |Z_s^n|^2ds\right\}\right]$$

$$\leq C\left[\rho_\varphi^2(t_{0,n}^\delta)+\rho_\varphi^2(\eta)+\frac{1}{\eta^2}[C+C_{\delta_0}^2]t_{0,n}^\delta\right], \tag{5.12}$$

thanks to (5.5). Combining (5.11) and (5.12), we see that (5.10) now becomes

$$\mathcal{L}\varphi(t_0,x_0,y_0)\leq C\varepsilon+\frac{C}{n\delta-1}+\rho_\varphi\left(\frac{1}{n}\right)$$

$$+C_{\varepsilon,\delta_0}\left[\rho_\varphi(t_{0,n}^\delta)+\rho_\varphi(\eta)+\frac{\sqrt{t_{0,n}^\delta}}{\eta}\right]. \tag{5.13}$$

Now fix ε and η, choose $\delta\triangleq\frac{1}{\sqrt{n}}$, and let $n\to\infty$.

$$\mathcal{L}\varphi(t_0,x_0,y_0)\leq C\varepsilon+C_{\varepsilon,\delta_0}\rho_\varphi(\eta).$$

Finally, letting $\eta\to 0$ and then $\varepsilon\to 0$, we obtain (5.3). That is, $\underline{u}$ is a viscosity supersolution, proving the theorem. □

We now give the definition of the uniqueness for FBMP.

Definition 5.5 We say that the forward–backward martingale problem (FBMP) of (2.7) has unique solution if $(\mathbb{P}^i, Z^i)$, $i=1,2$ are two solutions to the FBMP such that $\mathbb{P}^i(\mathbf{x}_0=x)=1$, $i=1,2$, then $\mathbb{P}^1=\mathbb{P}^2$, and $Z^1=Z^2$, $dt\otimes d\mathbb{P}^1$ $(=dt\otimes d\mathbb{P}^2)$-a.e.

We have the following result.

Theorem 5.6 *Assume* (H1)–(H4)*, b and f are uniformly continuous in t, and the comparison theorem for viscosity solution to the PDE* (4.1) *holds true, then the FBMP has a unique solution and consequently the uniqueness in law of weak solutions to FBSDE* (2.7) *holds true.*

Proof Since the proof is very similar to that of [24, Theorem 5.6], we shall give only a sketch, and point out the differences.

Suppose that $(\mathbb{P}^i, Z^i)$, $i=1,2$, are two weak solutions at $(0,x,y_1)$ and $(0,x,y_2)$, respectively. Following exactly the same arguments as in [24, Theorem 5.6], one shows that $\mathbf{y}_t=u(t,\mathbf{x}_t)$, for all $t\in[0,T]$, $\mathbb{P}^i$-a.s., $i=1,2$, thanks to Theorem 5.3. In particular, this implies that $y_1=y_2$.

Note that, unlike [24], the coefficient b is not 0 in the current case. We therefore modify the arguments as follows. First, for fixed $i = 1, 2$, denote

$$W_t^i \stackrel{\triangle}{=} \int_0^t \sigma^{-1}(s, \mathbf{x}_s, \mathbf{y}_s) d\mathbf{x}_s - \int_0^t \theta_s^i ds, \ t \in [0, T], \qquad \mathbb{P}^i\text{-a.s.},$$

where $\theta_s^i \stackrel{\triangle}{=} [\sigma^{-1} b](s, \mathbf{x}_s, \mathbf{y}_s, Z_s^i)$. By Definition 2.7-(i), W^i is a $\mathbb{P}^i$-Brownian motion. We next define a new probability measure:

$$\frac{d\widehat{\mathbb{P}}^i}{d\mathbb{P}^i} \stackrel{\triangle}{=} \exp\left\{ - \int_0^T \langle \theta_s^i, dW_s^i \rangle - \frac{1}{2} \int_0^T |\theta_s^i|^2 ds \right\}.$$

Then by Girsanov Theorem we know that $\widehat{\mathbb{P}}^i$ is a probability measure equivalent to $\mathbb{P}^i$ (hence $\mathbf{y}_t = u(t, \mathbf{x}_t)$, $\widehat{\mathbb{P}}^i$-a.s.), and the process

$$\widehat{W}_t^i \stackrel{\triangle}{=} W_t^i + \int_0^t \theta_s^i ds, \ t \in [0, T], \quad \mathbb{P}^i\text{-a.s.},$$

is a $\widehat{\mathbb{P}}^i$-Brownian motion. Further, note that

$$d\mathbf{x}_t = \sigma(t, \mathbf{x}_t, \mathbf{y}_t)[dW_t^i + \theta_t^i dt] = \sigma(t, \mathbf{x}_t, u(t, \mathbf{x}_t)) d\widehat{W}_t^i, \ \ \widehat{\mathbb{P}}^i\text{-a.s.}$$

That is, $(\widehat{W}^i, \mathbf{x}, \widehat{\mathbb{P}}^i)$, $i = 1, 2$, are weak solutions to the forward SDE (2.10) with $b \equiv 0$. It then follows from the uniqueness of weak solution to a forward SDE (see, e.g. [31]) that the $\widehat{\mathbb{P}}^1 \circ (\widehat{W}^1, \mathbf{x})^{-1} = \widehat{\mathbb{P}}^2 \circ (\widehat{W}^2, \mathbf{x})^{-1}$ (here $\mathbb{P} \circ X^{-1}$ denotes the distribution of X under $\mathbb{P}$). Recall that $\mathbf{y}_t = u(t, \mathbf{x}_t)$, $\widehat{\mathbb{P}}^i$-a.s., $i = 1, 2$, this in turn leads to that $\widehat{\mathbb{P}}^1 \circ (\widehat{W}^1, \mathbf{x}, \mathbf{y})^{-1} = \widehat{\mathbb{P}}^2 \circ (\widehat{W}^2, \mathbf{x}, \mathbf{y})^{-1}$. In particular, since $(\mathbf{x}, \mathbf{y})$ are the canonical process, we have $\widehat{\mathbb{P}}^1 = \widehat{\mathbb{P}}^2$ and $\widehat{W}^1 = \widehat{W}^2$, $\widehat{\mathbb{P}}^1$-a.s. and $\widehat{\mathbb{P}}^2$-a.s.

Now by Definition 2.7-(iii), as the density of quadratic covariation $\langle \mathbf{y}, W^i \rangle = \langle \mathbf{y}, \widehat{W}^i \rangle$ under $\widehat{\mathbb{P}}^i$, $i = 1, 2$, we have $Z^1 = Z^2$, $dt \otimes d\widehat{\mathbb{P}}^1$ $(= dt \otimes d\widehat{\mathbb{P}}^2)$-a.e. But by definition of θ^i's, this means that $\theta^1 = \theta^2$, $dt \otimes d\widehat{\mathbb{P}}^1$ $(= dt \otimes d\widehat{\mathbb{P}}^2)$-a.e. Moreover, since for each i,

$$\frac{d\mathbb{P}^i}{d\widehat{\mathbb{P}}^i} \stackrel{\triangle}{=} \exp\left\{ \int_0^T \langle \theta_s^i, dW_s^i \rangle + \frac{1}{2} \int_0^T |\theta_s^i|^2 ds \right\} = \exp\left\{ \int_0^T \langle \theta_s^i, d\widehat{W}_s^i \rangle - \frac{1}{2} \int_0^T |\theta_s^i|^2 ds \right\},$$

we conclude that $\mathbb{P}^1 = \mathbb{P}^2$. Finally, by the equivalence of $\widehat{\mathbb{P}}^i$ and $\mathbb{P}^i$ again, we have $Z^1 = Z^2$, $dt \otimes d\mathbb{P}^1$ $(= dt \otimes d\mathbb{P}^2)$-a.e. $\square$

Remark 5.7 We should note that under the assumptions of Theorem 5.6, the unique weak solution of FBSDE (2.7) is actually semi-strong, thanks to Theorem 3.2. Moreover, if the forward SDE (2.10) has a strong solution (e.g., if b is independent of Z), then the unique weak solution is actually strong. □

Remark 5.8 As it was mentioned in [24, Remark 5.7], the comparison principle of viscosity solutions is in general a rather delicate issue, and should be checked case by case. In the FBSDE case the degree of difficulty increases when the coupling nature of the coefficients becomes stronger. For instance, when the diffusion coefficient σ depends on the backward component Z, there have been only very few results in the literature. For interested reader we refer to the standard references [8,16], and the recent paper [3] for some general discussions of the comparison principles for fully nonlinear PDEs. Since it is not the main purpose of this paper to address this issue, we assumed the comparison principle in Theorem 5.6 to simplify the discussion. We note that the examples in [24, Sect. 5] regarding the cases where the comparison theorem holds for the viscosity solutions of (2.8) are still valid in the current case. □

6 Proofs for the PDE estimates

In this section we shall complete several proofs in the previous sections by establishing some key a priori estimates for the solutions to (2.8). We should note that most of the arguments in these proofs are more or less standard in the PDE literature, but we have not been able to find the exact reference for the desired results. We therefore provide the detailed proofs for the sake of completeness.

We start with the Hölder estimates, which is standard in the literature, and can be found in, e.g. [10].

Lemma 6.1 *Assume* (H1) *and* (H2)*, and the coefficients are all smooth with bounded derivatives. Let u denote the unique classical solution to PDE* (2.8)*.*

(i) *For any $\delta > 0$, there exist constants $\alpha \in (0,1)$ and $C_\delta > 0$, depending only on T, K, the dimensions d, m, and C_δ depending on δ as well, but not on the derivatives of the coefficients, such that, for any $t_1, t_2 \in [0, T-\delta]$, $x_1, x_2 \in \mathbb{R}^d$,*

$$|u(t_1,x_1) - u(t_2,x_2)| \le C_{\delta,\alpha}[|t_1 - t_2|^{\frac{\alpha}{2}} + |x_1 - x_2|^\alpha]. \tag{6.1}$$

(ii) *Moreover, if g is Hölder-α_0 continuous for some $\alpha_0 \in (0,1)$, then there exists $\alpha \in (0, \alpha_0)$ such that* (6.1) *holds for $\delta = 0$, with the constant $C_{0,\alpha}$ may depend on the Hölder constant of g as well.*

Next, we state a result regarding the a priori estimates for the following (linear) parabolic PDE with $m = 1$:

$$\begin{cases} u_t + \frac{1}{2}\mathrm{tr}\,\{u_{xx}\sigma(t,x)\sigma^*(t,x)\} + u_x b(t,x) + f(t,x) = 0; \\ u(T,x) = g(x). \end{cases} \tag{6.2}$$

This result may have several variations, but we find that the following form, as the direct consequences of Lieberman [20, Theorem 7.22, p. 175 and Theorem 7.30, p.181], suits our needs the best.

Lemma 6.2 *Assume that the coefficients b, σ, f, g in* (6.2) *satisfy* (H1), (H2)*, that σ is uniformly continuous in (t, x, y), and that all the coefficients are smooth, and with bounded derivatives. Let u denote the unique classical solution to* (6.2)*.*

(i) *For any $\delta > 0$, there exists a constant C_δ, depending only on T, K, δ, the dimension d, and the uniform continuity of σ, but not on the derivatives of the coefficients, such that for any bounded domain $D \subset \mathbb{R}^d$ and any $p \geq 2$,*

$$\int_0^{T-\delta} \int_D [|u_x(t,x)|^p + |u_{xx}(t,x)|^p] dx dt \leq C_\delta^p |D|, \tag{6.3}$$

where $|D|$ denotes the Lebesgue measure of D.

(ii) *Furthermore, if in addition $g \in C_b^2(\mathbb{R}^d)$, then the estimate* (6.3) *holds for $\delta = 0$. But in this case the constant C_0 may depend on the bounds $\|g_x\|_\infty$ and $\|g_{xx}\|_\infty$ as well.*

We shall prove Theorems 3.1 and 4.1 in the following two subsections, respectively.

6.1 Proof of Theorem 3.1

We first prove (i). Since all the coefficients are assumed to be bounded and smooth, by the Four Step Scheme (cf. [23]) the FBSDE (2.7) has a unique strong solution (X, Y, Z). Note that

$$u(0,x) = Y_0 = E\left\{ g(X_T) + \int_0^T f(t, X_t, Y_t, Z_t) dt \right\}.$$

Then by (H1) we immediately conclude that $|u(0,x)| \leq C$, as well as $|u(t,x)| \leq C$, $\forall (t,x)$.

To prove (ii)–(iv), we first make some preparations. Denote

$$\begin{aligned} &\bar{b}(t,x) \stackrel{\triangle}{=} b(t,x,u(t,x),u_x(t,x)\sigma(t,x,u(t,x))); \quad \bar{\sigma}(t,x) \stackrel{\triangle}{=} \sigma(t,x,u(t,x)); \\ &\bar{f}^i(t,x) \stackrel{\triangle}{=} f^i(t,x,u(t,x),u_x(t,x)\sigma(t,x,u(t,x))), \quad i = 1,\dots,m. \end{aligned} \tag{6.4}$$

Then u^i is the solution to the following linear PDE:

$$\begin{cases} u_t^i + \dfrac{1}{2}\mathrm{tr}\,\{u_{xx}^i \bar{\sigma}(t,x)\bar{\sigma}^*(t,x)\} + u_x^i \bar{b}(t,x) + \bar{f}(t,x) = 0; \\ u^i(T,x) = g^i(x). \end{cases} \quad i = 1,\dots,m. \tag{6.5}$$

Since we shall prove the theorem for each i, without loss of generality we may assume $m = 1$ for simplicity. Denote

$$\bar{a}(t,x) \stackrel{\triangle}{=} \bar{\sigma}\bar{\sigma}^*(t,x); \quad \Gamma^{t_1}_{t_2} \stackrel{\triangle}{=} \int_{t_1}^{t_2} \bar{a}(s,0)ds; \quad p(\Gamma,x) \stackrel{\triangle}{=} \frac{1}{\sqrt{(2\pi)^d \det(\Gamma)}} e^{-\frac{1}{2}x^*\Gamma^{-1}x}. \tag{6.6}$$

Then

$$p_x(\Gamma,x) = -p(\Gamma,x)x^*\Gamma^{-1}. \tag{6.7}$$

Let $\eta \in C^\infty(\mathbb{R}^d)$ be such that $\eta(x) = 1$ for $|x| \le 1$ and $\eta(x) = 0$ for $|x| > 2$. Denote $v(t,x) \stackrel{\triangle}{=} u(t,x)\eta(x)$. Then

$$v_t(t,x) + \frac{1}{2}\text{tr}\ (v_{xx}(t,x)\bar{a}(t,0)) + \tilde{f}(t,x) = 0;$$

where

$$\begin{aligned}\tilde{f}(t,x) \stackrel{\triangle}{=}\ & \frac{1}{2}\text{tr}\ (u_{xx}(t,x)[\bar{a}(t,x) - \bar{a}(t,0)])\,\eta(x) - \text{tr}\ (u_x(t,x)\eta_x(x)\bar{a}(t,0)) \\ & -\frac{1}{2}\text{tr}\ (u(t,x)\eta_{xx}(x)\bar{a}(t,0)) - u_x(t,x)\eta(x)\bar{b}(t,x) + \bar{f}(t,x)\eta(x).\end{aligned}$$

It then follows from (i) of this theorem that

$$|\tilde{f}(t,y)| \le C[|u_{xx}(t,y)| + |u_x(t,y)| + 1]. \tag{6.8}$$

Furthermore, for $|x| < 1$ and $0 \le t_1 < t_2 \le T$, by standard arguments we have two representation formulas for u:

$$\begin{aligned}u(t_1,x) = v(t_1,x) &= \int_{\mathbb{R}^d} v(t_2,y)p(\Gamma^{t_1}_{t_2}, y-x)dy + \int_{t_1}^{t_2}\int_{\mathbb{R}^d} \tilde{f}(t,y)p(\Gamma^{t_1}_t, y-x)dydt \\ &= \int_{\mathbb{R}^d} v(t_2,y+x)p(\Gamma^{t_1}_{t_2}, y)dy + \int_{t_1}^{t_2}\int_{\mathbb{R}^d} \tilde{f}(t,y)p(\Gamma^{t_1}_t, y-x)dydt.\end{aligned}$$

Then

$$u_x(t_1,x) = \int_{\mathbb{R}^d} v(t_2,y)p(\Gamma_{t_2}^{t_1}, y-x)(y-x)^*[\Gamma_{t_2}^{t_1}]^{-1}dy$$
$$+\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \tilde f(t,y)p(\Gamma_t^{t_1}, y-x)(y-x)^*[\Gamma_t^{t_1}]^{-1}dydt; \tag{6.9}$$
$$= \int_{\mathbb{R}^d} v_x(t_2,y+x)p(\Gamma_{t_2}^{t_1}, y)dy$$
$$+\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \tilde f(t,y)p(\Gamma_t^{t_1}, y-x)(y-x)^*[\Gamma_t^{t_1}]^{-1}dydt. \tag{6.10}$$

We are now ready to prove (ii)–(iv). Fix $\delta > 0$. For notational simplicity, denote

$$\Gamma_t \triangleq \Gamma_t^0; \quad T_\delta \triangleq T-\delta; \quad D_\delta \triangleq [0,T_\delta]\times\mathbb{R}^d.$$

(ii) Without loss of generality, we prove the estimate only at point $(0,0)$. Set $t_1 = 0, t_2 = T_\delta, x = 0$ in (6.9), we get

$$u_x(0,0) = \int_{\mathbb{R}^d} v(T_\delta,y)p(\Gamma_{T_\delta},y)y^*\Gamma_{T_\delta}^{-1}dy + \int_{D_\delta}\tilde f(t,y)p(\Gamma_t,y)y^*\Gamma_t^{-1}dydt.$$

For $p \ge 2$ (to be specified later), let q be its conjugate. Since $|v(t,x)| \le C$, we have

$$|u_x(0,0)| \le \frac{C}{\sqrt{T_\delta}} + C\left[\int_{D_\delta}|\tilde f(t,y)|^p dydt\right]^{\frac{1}{p}}$$
$$\times\left[\int_{D_\delta} p(\Gamma_t,y)^q|y^*\Gamma_t^{-1}|^q dydt\right]^{\frac{1}{q}}. \tag{6.11}$$

Set $D \triangleq B_2(0)$, then $\tilde f(t,y) = 0$ for $y \notin D$. By Lemma 6.1 (i), $\bar\sigma$ is uniformly continuous on $\left[0, T-\frac{\delta}{2}\right]\times\mathbb{R}^d$. By considering $\frac{\delta}{2}$ instead of δ and recalling (6.8), we can apply Lemma 6.2 (i) on $\left[0, T-\frac{\delta}{2}\right]$ to get

$$\int_{D_\delta}|\tilde f(t,y)|^p dydt \le C\int_0^{T_\delta}\int_D [|u_{xx}|^p + |u_x|^p + 1]dydt \le C_\delta. \tag{6.12}$$

Moreover, by changing variable in an obvious way, we have

$$\int_{D_\delta} p(\Gamma_t, y)^q |y^*\Gamma_t^{-1}|^q dydt \le \int_{D_\delta} \frac{1}{\sqrt{\det(\Gamma_t)^q}} e^{-\frac{q}{2}|y|^2} |y^*\Gamma_t^{-1/2}|^q \sqrt{\det(\Gamma_t)} dydt$$
$$\le C_\delta \int_0^{T_\delta} t^{-\frac{dq+q-d}{2}} dt. \tag{6.13}$$

We can now choose $p > d + 2$, so that $q < \frac{d+2}{d+1}$, and that

$$|u_x(0, 0)| \le \frac{C}{\sqrt{T_\delta}} + C_\delta \left[(T_\delta)^{1-\frac{dq+q-d}{2}} \right]^{\frac{1}{q}} < \infty.$$

This proves the estimate at (0, 0).

(iii) We first prove the Hölder continuity in x. Again, without loss of generality, we shall only estimate $|u_x(0, x) - u_x(0, 0)|$ for $|x| < 1$. Recall (6.9). Following similar arguments as in (ii) one can easily show that, for $\delta < T/2$,

$$\begin{aligned}
|u_x(0, x) - u_x(0, 0)| \le\ & C \int_{\mathbb{R}^d} |v(T_\delta, y)| \Big| p(\Gamma_{T_\delta}, y - x)(y - x)^* \Gamma_{T_\delta}^{-1} \\
& - p(\Gamma_{T_\delta}, y) y^* \Gamma_{T_\delta}^{-1} \Big| dy \\
& + C \int_{D_\delta} |\tilde{f}(t, y)| \Big| p(\Gamma_t, y - x)(y - x)^* \Gamma_t^{-1} \\
& - p(\Gamma_t, y) y^* \Gamma_t^{-1} \Big| dydt \\
\le\ & C|x| + C_\delta \Bigg[\int_{D_\delta} \Big| p(\Gamma_t, y - x)(y - x)^* \Gamma_t^{-1} \\
& - p(\Gamma_t, y) y^* \Gamma_t^{-1} \Big|^q dydt \Bigg]^{\frac{1}{q}}.
\end{aligned}$$

For a constant $\lambda > 0$ which will be specified later, similar to (6.13) we get

$$
\begin{aligned}
&|u_x(0,x) - u_x(0,0)|^q \\
&\quad \le C|x|^q + C_\delta \int_0^\lambda \int_{\mathbb{R}^d} \Big[|p(\Gamma_t, y-x)(y-x)^* \Gamma_t^{-1}|^q + |p(\Gamma_t, y) y^* \Gamma_t^{-1}|^q \Big] dy dt \\
&\qquad + C_\delta \int_\lambda^{T_\delta} \int_{\mathbb{R}^d} \Big| p(\Gamma_t, y-x)(y-x)^* \Gamma_t^{-1} - p(\Gamma_t, y) y^* \Gamma_t^{-1} \Big|^q dy dt \\
&\quad \le C|x|^q + C_\delta \lambda^{1-\frac{dq+q-d}{2}} \\
&\qquad + C \int_\lambda^{T_\delta} \int_{\mathbb{R}^d} \left| \int_0^1 p(\Gamma_t, y-\theta x)[x^* \Gamma_t^{-1}(y-\theta x)(y-\theta x)^* - x^*] \Gamma_t^{-1} d\theta \right|^q dy dt \\
&\quad \le C|x|^q + C_\delta \lambda^{1-\frac{dq+q-d}{2}} + C_\delta \int_\lambda^{T_\delta} t^{-\frac{dq}{2}-q+\frac{d}{2}} dt |x|^q \\
&\quad \le C|x|^q + C_\delta \lambda^{1-\frac{dq+q-d}{2}} + C_\delta \lambda^{1-\frac{dq}{2}-q+\frac{d}{2}} |x|^q .
\end{aligned}
$$

Here we note that $-\frac{dq}{2} - q + \frac{d}{2} < -1$, thanks to the fact that $q > 1$. Choose λ to minimize the right side of above, that is, $\lambda = |x|^2$, then we get

$$
|u_x(0,x) - u_x(0,0)| \le C|x| + C_\delta |x|^{\frac{2+d}{q} - (1+d)} .
$$

Now for any $\alpha < 1$, choose $p = \frac{2+d}{1-\alpha}$. Thus $q = \frac{2+d}{1+d+\alpha}$ or $\frac{2+d}{q} - (1+d) = \alpha$. Then,

$$
|u_x(0,x) - u_x(0,0)| \le C|x| + C_\delta |x|^\alpha \le C_\delta |x|^\alpha ,
$$

proving the Hölder continuity in x. To show the Hölder continuity in t, we again assume without loss of generality that $t_1 = 0$, $t_2 = t$, and $x_1 = x_2 = 0$. Note that

$$
\int_{\mathbb{R}^d} p(\Gamma_t, y) y^* \Gamma_t^{-1} dy = [0, \ldots, 0], \quad \int_{\mathbb{R}^d} p(\Gamma_t, y) y^* \Gamma_t^{-1} y dy = 1.
$$

By (6.9) again we have

$$
\begin{aligned}
u_x(0,0) &= \int_{\mathbb{R}^d} v(t,y)p(\Gamma_t,y)y^*\Gamma_t^{-1}dy + \int_0^t\int_{\mathbb{R}^d} \tilde{f}(s,y)p(\Gamma_s,y)y^*\Gamma_s^{-1}dyds \\
&= \int_{\mathbb{R}^d} [v(t,y)-v(t,0)]p(\Gamma_t,y)y^*\Gamma_t^{-1}dy \\
&\quad + \int_0^t\int_{\mathbb{R}^d} \tilde{f}(s,y)p(\Gamma_s,y)y^*\Gamma_s^{-1}dyds \\
&= \int_{\mathbb{R}^d}\int_0^1 v_x(t,\theta y)d\theta y p(\Gamma_t,y)y^*\Gamma_t^{-1}dy + \int_0^t\int_{\mathbb{R}^d} \tilde{f}(s,y)p(\Gamma_s,y)y^*\Gamma_s^{-1}dyds.
\end{aligned}
$$

Then

$$
\begin{aligned}
u_x(0,0)-u_x(t,0) &= \int_{\mathbb{R}^d}\int_0^1 [v_x(t,\theta y)-v_x(t,0)]d\theta y p(\Gamma_t,y)y^*\Gamma_t^{-1}dy \\
&\quad + \int_0^t\int_{\mathbb{R}^d} \tilde{f}(s,y)p(\Gamma_s,y)y^*\Gamma_s^{-1}dyds.
\end{aligned}
$$

Therefore, for any $\alpha < 1$ and $q = \frac{2+d}{1+d+\alpha}$, we get

$$
\begin{aligned}
|u_x(0,0)-u_x(t,0)| &\le C\int_{\mathbb{R}^d}\int_0^1 |\theta y|^\alpha d\theta |y| p(\Gamma_t,y)|y^*\Gamma_t^{-1}|dy + Ct^{\frac{1}{q}-\frac{dq+q-d}{2q}} \\
&\le Ct^{\frac{\alpha}{2}}.
\end{aligned}
$$

(iv) We first estimate $u_x(t,x)$. Again we will do it only at $(0,0)$. Set $t_1 = 0$, $t_2 = T$, $x = 0$ in (6.10), we get

$$
u_x(0,0) = \int_{\mathbb{R}^d} g_x(y)p(\Gamma_T,y)dy + \int_0^T\int_{\mathbb{R}^d} \tilde{f}(t,y)p(\Gamma_t,y)y^*\Gamma_t^{-1}dydt.
$$

By Lemma 6.2 (ii) and following similar arguments in (ii) of this proof one can easily show that $|u_x(0,0)| \le C_0$. Similarly we can prove $|u_x(t,x)| \le C_0$.

By using (6.10) and Lemma 6.2 (ii), similar to (iii) of this proof we can prove, for any $|x| \le 1$,

$$|u_x(0,x) - u_x(0,0)| \le C_\alpha |x|^\alpha.$$

Similar to (iii) we can prove the rest of the estimates, completing the proof. $\square$

6.2 Proof of Theorem 4.1

We first recall (6.4)–(6.6) and write

$$u^i_t(t,x) + \frac{1}{2}\mathrm{tr}\left(u^i_{xx}(t,x)\bar{a}(t,0)\right) + \hat{f}^i(t,x) = 0;$$

where

$$\hat{f}^i(t,x) \triangleq \frac{1}{2}\mathrm{tr}\left(u^i_{xx}(t,x)[\bar{a}(t,x) - \bar{a}(t,0)]\right) + u^i_x(t,x)\bar{b}(t,x) + \bar{f}^i(t,x).$$

Then u^i can be solved explicitly in the following form:

$$u^i(t,x) = \int\limits_{\mathbb{R}^d} g^i(y+x)p(\Gamma^t_T, y)dy + \int\limits_t^T\int\limits_{\mathbb{R}^d} \hat{f}^i(s,y)p(\Gamma^t_s, y-x)dyds.$$

Differentiating the equation above we obtain that

$$\begin{aligned}
u^i_x(t,x) &= \int\limits_{\mathbb{R}^d} g^i_x(y+x)p(\Gamma^t_T, y)dy \\
&\quad + \int\limits_t^T\int\limits_{\mathbb{R}^d} \hat{f}^i(s,y)p(\Gamma^t_s, y-x)(y-x)^*[\Gamma^t_s]^{-1}dyds; \\
u^i_{xx}(t,x) &= \int\limits_{\mathbb{R}^d} g^i_{xx}(y+x)p(\Gamma^t_T, y)dy \\
&\quad + \int\limits_t^T\int\limits_{\mathbb{R}^d} \hat{f}^i(s,y)p(\Gamma^t_s, y-x)[\Gamma^t_s]^{-1}[(y-x)(y-x)^*[\Gamma^t_s]^{-1} - I_d]dyds.
\end{aligned}$$

Setting $x=0$ we get

$$u^i_{xx}(t,0)=\int\limits_{\mathbb{R}^d} g^i_{xx}(y)p(\Gamma^t_T,y)dy$$
$$+\int\limits_t^T\int\limits_{\mathbb{R}^d}[\hat{f}^i(s,y)-\hat{f}^i(s,0)]p(\Gamma^t_s,y)[\Gamma^t_s]^{-1}[yy^*[\Gamma^t_s]^{-1}-I_d]dyds. \tag{6.14}$$

In the above we have used the fact that $\int_{\mathbb{R}^d} p(\Gamma^t_s,y)[yy^*[\Gamma^t_s]^{-1}-I_d]dy=0$. Next, denote

$$A_t\overset{\triangle}{=}\sup_i\sup_{x\in\mathbb{R}^d,s\in[t,T]}|u^i_{xx}(s,x)|. \tag{6.15}$$

For $t\le s\le T$, recalling Theorem 3.1 (iv) we have

$$\begin{aligned}|\hat{f}^i(s,y)-\hat{f}^i(s,0)|&\le\frac{1}{2}\left|\mathrm{tr}\left(u^i_{xx}(s,y)[\bar{a}(s,y)-\bar{a}(s,0)]\right)\right|\\&\quad+|u^i_x(s,y)\bar{b}(s,y)-u^i_x(s,0)\bar{b}(s,0)|+|\bar{f}^i(s,y)-\bar{f}^i(s,0)|\\&\le CA_t|\bar{a}(s,y)-\bar{a}(s,0)|+CA_t|y|+C|\bar{b}(s,y)-\bar{b}(s,0)|\\&\quad+|\bar{f}^i(s,y)-\bar{f}^i(s,0)|.\end{aligned} \tag{6.16}$$

Since $|\sigma|\le K$, without loss of generality, we may assume $\rho(y)=2K$ when $|y|\ge 1$. Then for $\alpha<1$, by Lemma 3.1 (iv) again, we have

$$\begin{aligned}&|\bar{a}(s,y)-\bar{a}(s,0)|=|\sigma\sigma^*(s,y,u(s,y))-\sigma\sigma^*(s,0,u(s,0))|\\&\quad\le C\rho(|y|+|u(s,y)-u(s,0)|)\le C\rho(C[|y|+|y|^\alpha])\le C\rho(C|y|^\alpha);\\&|\bar{b}(s,y)-\bar{b}(s,0)|\\&\quad=|b(s,y,u(s,y),u_x(s,y)\sigma(s,y,u(s,y)))\\&\qquad-b(s,0,u(s,0),u_x(s,0)\sigma(s,0,u(s,0)))|\\&\quad\le C\left[|y|+|y|^\alpha+A_t|y|+\rho(C|y|^\alpha)\right];\\&|\bar{f}^i(s,y)-\bar{f}^i(s,0)|\\&\quad=|f^i(s,y,u(s,y),u_x(s,y)\sigma(s,y,u(s,y)))\\&\qquad-f^i(s,0,u(s,0),u_x(s,0)\sigma(s,0,u(s,0)))|\\&\quad\le C\left[|y|+|y|^\alpha+A_t|y|+\rho(C|y|^\alpha)\right].\end{aligned}$$

Plug into (6.16) we have

$$|\hat{f}^i(s,y)-\hat{f}^i(s,0)|\le C\left[|y|+|y|^\alpha+\rho(C|y|^\alpha)\right][1+A_t].$$

Then by (6.14) we get

$$\begin{aligned}|u^i_{xx}(t,0)|\le \|g_{xx}\|_\infty + C[1+A_t]\int_t^T\int_{\mathbb{R}^d} p(\Gamma^t_s,y)[|[\Gamma^t_s]^{-1}yy^*[\Gamma^t_s]^{-1}|+|[\Gamma^t_s]^{-1}|]\\ \times\left[|y|+|y|^\alpha+\rho(C|y|^\alpha)\right]dyds.\end{aligned}$$

Applying a change of variable $y=[\Gamma^t_s]^{\frac{1}{2}}y'$, we get

$$\begin{aligned}|u^i_{xx}(t,0)|&\le \|g_{xx}\|_\infty + C[1+A_t]\int_t^T\int_{\mathbb{R}^d} e^{-\frac{1}{2}|y|^2}\frac{1+|y|^2}{s-t}\\ &\quad\times\left[\sqrt{s-t}|y|+(s-t)^{\frac{\alpha}{2}}|y|^\alpha+\rho(C(s-t)^{\frac{\alpha}{2}}|y|^\alpha)\right]dyds\\ &\le \|g_{xx}\|_\infty + C[1+A_t]\left[(T-t)^{\frac{\alpha}{2}}\right.\\ &\quad\left.+\int_t^T\int_{\mathbb{R}^d} e^{-\frac{1}{2}|y|^2}\frac{(1+|y|)^2}{s-t}\rho(C_1(s-t)^{\frac{\alpha}{2}}(1+|y|)^\alpha)\right]dyds.\end{aligned}$$

Now change variable $s=t+\left[\frac{s'}{C_1(1+|y|)^\alpha}\right]^{\frac{2}{\alpha}}$ we get

$$\begin{aligned}|u^i_{xx}(t,0)|&\le \|g_{xx}\|_\infty + C_2[1+A_t]\\ &\quad\times\left[(T-t)^{\frac{\alpha}{2}}+\int_{\mathbb{R}^d} e^{-\frac{1}{2}|y|^2}dy\int_0^{C_1(T-t)^{\frac{\alpha}{2}}(1+|y|)^\alpha}\frac{\rho(s)}{s}ds\right].\end{aligned}$$

By (H4), we can choose $\delta_0>0$ such that

$$\delta_0^{\frac{\alpha}{2}}\le\frac{1}{4C_2};\quad \int_{\mathbb{R}^d} e^{-\frac{1}{2}|y|^2}dy\int_0^{C_1\delta_0^{\frac{\alpha}{2}}(1+|y|)^\alpha}\frac{\rho(s)}{s}ds\le\frac{1}{4C_2}. \tag{6.17}$$

Then for $T - \delta_0 \le t \le T$, we have $|u^i_{xx}(t, 0)| \le \|g_{xx}\|_\infty + \frac{1}{2}[1 + A_t]$. In general, we can prove similarly that

$$\sup_x |u^i_{xx}(t, x)| \le \|g_{xx}\|_\infty + \frac{1}{2}[1 + A_t].$$

Moreover, since A_t is decreasing in t, we see that $A_t \le \|g_{xx}\|_\infty + \frac{1}{2}[1 + A_t]$. That is,

$$A_t \le 2\|g_{xx}\|_\infty + 1, \quad \forall t \in [T - \delta_0, T].$$

In particular, $A_{T-\delta_0} \le 2\|g_{xx}\|_\infty + 1$. We note that δ_0 is a constant which does not depend on g. Then following the same arguments over $[T - 2\delta_0, T - \delta_0]$ we get

$$A_{T-2\delta_0} \le 2A_{T-\delta_0} + 1 \le 4\|g_{xx}\|_\infty + 3.$$

Repeating the arguments for at most finitely many times, we conclude that $A_0 \le C$, proving the theorem. □

Acknowledgments The authors are grateful to the anonymous referee for his/her careful reading of the manuscript and many helpful suggestions, which helped to improve the quality of the paper greatly.

References

1. Antonelli, F.: Backward-forward stochastic differential equations. Ann. Appl. Probab. **3**(3), 777–793 (1993)
2. Antonelli, F., Ma, J.: Weak solutions of forward-backward SDE's. Stoch. Anal. Appl. **21**(3), 493–514 (2003)
3. Barles, G., Biton, S., Bourgoing, M., Ley, O.: Uniqueness results for quasilinear parabolic equations through viscosity solutions' methods. Calc. Var. **18**, 159–179 (2003)
4. Bouchard, B., Touzi, N.: Weak dynamic programming principle for viscosity solutions. (2009, preprint)
5. Buckdahn, R., Engelbert, H.-J.: A backward stochastic differential equation without strong solution. Teor. Veroyatn. Primen. **50**(2), 390–396 (2005); translation in Theory Probab. Appl. **50**(2), 284–289 (2006)
6. Buckdahn, R., Engelbert, H.-J.: On the continuity of weak solutions of backward stochastic differential equations. Teor. Veroyatn. Primen. **52**(1), 190–199 (2007); translation in Theory Probab. Appl. **52**(1), 152–160 (2008)
7. Buckdahn, R., Engelbert, H.-J., Rascanu, A.: On weak solutions of backward stochastic differential equations. Teor. Veroyatn. Primen. **49**(1), 70–108 (2004)
8. Crandall, M.G., Ishii, H., Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. (NS) **27**, 1–67 (1992)
9. Delarue, F.: On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. Stoch. Process. Appl. **99**, 209–286 (2002)
10. Delarue, F.: Estimates of the Solutions of a System of Quasi-linear PDEs. A Probabilistic Scheme. Séminaire de Probabilités XXXVII. Lecture Notes in Mathematics, vol. 1832, pp. 290–332. Springer, Berlin (2003)
11. Delarue, F., Guatteri, G.: Weak existence and uniqueness for forward-backward SDEs. Stoch. Process. Appl. **116**(12), 1712–1742 (2006)
12. Fabes, E., Kenig, C.: Examples of singular parabolic measures and singular transition probability densities. Duke Math. J. **48**, 845–856 (1981)
13. Hu, Y., Peng, S.: Solution of forward-backward stochastic differential equations. Probab. Theory Relat. Fields **103**(2), 273–283 (1995)

14. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. North Holland, Amstrerdam (1981)
15. Ishii, H.: Degenerate parabolic PDEs with discontinuities and generalized evolutions of surfaces. Adv. Differ. Equ. **1**(1), 51–72 (1996)
16. Ishii, H., Lions, P.-L.: Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. J. Differ. Equ. **83**, 26–78 (1990)
17. Kobylanski, M.: Backward stochastic differential equations and partial differential equations with quadratic growth. Ann. Probab. **28**(2), 558–602 (2000)
18. Kurtz, T.: The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. Electron. J. Probab. **12**, 951–965 (2007)
19. Lepeltier, M., San Martin, J.: Backward stochastic differential equations with nonLipschitz coefficients. Stat. Prob. Lett. **32**(4), 425–430 (1997)
20. Lieberman, G.: Second Order Parabolic Differential Equations. World Scientific Publishing Co., Inc., River Edge (1996)
21. Ma, J., Protter, P., Yong, J.: Solving forward-backward stochastic differential equations explicitly—a four step scheme. Probab. Theory Relat. Fields **98**, 339–359 (1994)
22. Ma, J., Yong, J.: Solvability of forward-backward SDEs and the nodal set of Hamilton-Jaccobi-Bellman equations. Chin. Ann. Math. **16B**, 279–298 (1995)
23. Ma, J., Yong, J.: Forward-backward stochastic differential equations and their applications. Lecture Notes in Mathematics, vol. 1702. Springer, New York (1999)
24. Ma, J., Zhang, J., Zheng, Z.: Weak solutions for forward-backward SDEs—a martingale problem approach. Ann. Probab. **36**(6), 2092–2125 (2008)
25. Nash, J.: Continuity of solutions of parabolic and elliptic equations. Am. J. Math. **80**(4), 931–954 (1958)
26. Pardoux, E., Peng, S.: Adapted solutions of backward stochastic equations. Syst. Control Lett. **14**, 55–61 (1990)
27. Pardoux, E., Peng, S.: Backward stochastic differential equations and quasilinear parabolic partial differential equations. Lecture Notes in CIS, vol. 176, pp. 200–217. Springer, Berlin (1992)
28. Pardoux, E., Tang, S.: Forward-backward stochastic differential equations and quasilinear parabolic PDEs. Probab. Theory Relat. Fields **114**(2), 123–150 (1999)
29. Peng, S., Wu, Z.: Fully coupled forward-backward stochastic differential equations and applications to optimal control. SIAM J. Control Optim. **37**(3), 825–843 (1999)
30. Protter, P.: Integration and Stochastic Differential Equations. Springer, Berlin (1990)
31. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Springer-Verlag, Berlin (1979)
32. Yong, J.: Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. Probab. Theory Relat. Fields **107**(4), 537–572 (1997)
33. Zhang, J.: The wellposedness of FBSDEs. Discrete Contin. Dyn. Syst. Ser. B **6**(4), 927–940 (2006)