Forward Backward SDEs in Weak Formulation

Haiyang Wang* and Jianfeng Zhang†‡

August 7, 2018

Abstract

Although having been developed for more than two decades, the theory of forward backward stochastic differential equations is still far from complete. In this paper, we take one step back and investigate the formulation of FBSDEs. Motivated from several considerations, both in theory and in applications, we propose to study FBSDEs in weak formulation, rather than the strong formulation in the standard literature. That is, the backward SDE is driven by the forward component, instead of by the Brownian motion noise. We establish the Feyman-Kac formula for FBSDEs in weak formulation, both in classical and in viscosity sense. Our new framework is efficient especially when the diffusion part of the forward equation involves the Z-component of the backward equation.

Keywords. Forward backward SDEs, strong formulation, weak formulation, dynamic programming principle, stochastic maximum principle, quasilinear PDEs, path dependent PDEs, weak solution, viscosity solution, martingale problem.

2000 AMS Mathematics subject classification: 60H07, 60H30, 35R60, 34F05

*School of Mathematics and Statistics, Shandong Normal University, Jinan, 250014, China. Email: health.sea@163.com. This author is supported in part by the Distinguished Middle-Aged and Young Scientist Encourage and Reward Foundation of Shandong Province (ZR2017BA033).

†Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA. E-mail: jianfenz@usc.edu. This author is supported in part by NSF grant #1413717.

‡The second author would like to thank Daniel Lacker for very helpful discussion on Section 2.3.2.

1 Introduction

In the standard literature, a coupled FBSDE takes the following form:

$$\begin{cases} X_t = x + \int_0^t b(s,\Theta_s)ds + \int_0^t \sigma(s,\Theta_s)dB_s, \\ Y_t = g(X_T) + \int_t^T f(s,\Theta_s)ds - \int_t^T Z_s dB_s, \end{cases} \quad t \in [0,T], \quad \mathbb{P}_0\text{-a.s.} \tag{1.1}$$

where $\Theta := (X, Y, Z)$ is the solution triplet, B is a Brownian motion under the probability measure $\mathbb{P}_0$ and the coefficients b, σ, f, and g are $\mathbb{F}^B$-progressively measurable in all variables. There have been many publications on the subject, see e.g. Antonelli [1], Ma, Protter & Yong [17], Hu & Peng [14], Yong [30], Peng & Wu [23], Pardoux & Tang [22], Delarue [7], Zhang [32], Ma, Wu, Zhang & Zhang [18], as well as the monograph Ma & Yong [19]. However, the theory is still far from complete. The existing methods in the literature provide quite different sets of sufficient conditions, and the unified approach proposed in [18] works only in one dimensional case and the conditions there are rather technical. Even worse, many FBSDEs arising from applications do not fit in any existing works.

To understand the problem better, we take one step back and try to understand the formulation of the problem. Is (1.1) indeed the "right" formulation of the problem? As we will justify below, we feel the following alternative form, which we call FBSDEs in weak formulation, or simply weak FBSDEs, seems more appropriate in many situations:

$$\begin{cases} X_t = x + \int_0^t b(s,\Theta_s)ds + \int_0^t \sigma(s,\Theta_s)dB_s; \\ Y_t = g(X_T) + \int_t^T f(s,\Theta_s)ds - \int_t^T Z_s dX_s, \end{cases} \quad t \in [0,T], \quad \mathbb{P}_0\text{-a.s.} \tag{1.2}$$

To indicate the difference, we denote by $\Theta^S := (X^S, Y^S, Z^S)$ the solution to (1.1) and $\Theta^W := (X^W, Y^W, Z^W)$ the solution to (1.2), where the superscripts ^{S} and ^{W} stand for strong and weak, respectively. We note that in (1.2) the stochastic integration in the backward equation is against dX_t, not against the Brownian motion dB_t. In the case that

$$\text{the mapping } z \mapsto z\sigma(t,x,y,z) \text{ has an inverse function } \psi(t,x,y,z), \tag{1.3}$$

by denoting $\tilde{Z} := Z^W \sigma(t, \Theta_t^W)$ and thus $Z^W = \psi(t, X_t^W, Y_t^W, \tilde{Z}_t)$, one can easily check that $(X_t^W, Y_t^W, \tilde{Z}_t)$ is a solution to the following FBSDE in strong formulation:

$$\begin{cases} X_t = x + \int_0^t \tilde{b}(s,\Theta_s)ds + \int_0^t \tilde{\sigma}(s,\Theta_s)dB_s; \\ Y_t = g(X_T) + \int_t^T \tilde{f}(s,\Theta_s)ds - \int_t^T Z_s dB_s, \end{cases} \quad \mathbb{P}_0\text{-a.s.} \tag{1.4}$$

where, for $\theta := (t, x, y, z)$,

$$\tilde{b}(\theta) = b(t, x, y, \psi(\theta)), \quad \tilde{\sigma}(\theta) = \sigma(t, x, y, \psi(\theta)), \quad \tilde{f}(\theta) = f(t, x, y, \psi(\theta)) - \psi(\theta)\tilde{b}(\theta).$$

When $\sigma = \sigma(t, x, y)$ is independent of z and $\sigma > 0$, it is clear that $\psi(t, \theta) = \frac{z}{\sigma(t,x,y)}$. However, when σ depends on z, typically we do not have the inverse function ψ.

We justify the weak formulation (1.2) in four aspects. Firstly, in the option pricing and hedging theory, which is one of the main applications of BSDEs and FBSDEs, let S denote the stock price driven by a Brownian motion B. For a hedging portfolio h with wealth value V, the self financing condition gives $dV_t = [\cdots]dt + h_t dS_t$. Note that (S, V) here correspond to (X, Y) in FBSDE, and the stochastic integration in dV is against dS_t, not dB_t. In simple models like Black-Scholes model, S and B generate the same filtration, then such difference is not crucial and there is no problem for using the strong formulation. However, for superhedging problem in incomplete markets, for example, one has to use dS_t to superhedge, then the weak formulation is indeed more appropriate. In fact, in many practical applications, X is the state process we observe and B is the noise used to model the distribution of X. Note that one rationale of using Brownian motion is the central limit theorem, where the convergence is in distribution, in this case the value of B may even not exist physically. So in these applications the weak formulation is more appropriate.

Secondly, in Markovian setting and in the case $\sigma = \sigma(t, x, y)$, the FBSDE (1.1) is associated with the following quasilinear PDE with terminal condition $u(T, x) = g(x)$:

$$\partial_t u + \frac{1}{2}\sigma^2(t, x, u)\partial^2_{xx} u + b(t, x, u, \partial_x u\sigma(t, x, u))\partial_x u + f(t, x, u, \partial_x u\sigma(t, x, u)) = 0, \quad (1.5)$$

through the so called nonlinear Feynman-Kac formula:

$$Y^S_t = u(t, X^S_t), \quad Z^S_t = \partial_x u(t, X^S_t)\sigma(t, X^S_t, u(t, X^S_t)). \quad (1.6)$$

However, when σ depends on Z, the PDE will involve the inverse function ψ in (1.3) which typically does not exist. The weak formulation (1.2), instead, corresponds to the following more natural PDE even in the case $\sigma = \sigma(t, x, y, z)$:

$$\partial_t u + \frac{1}{2}\sigma^2(t, x, u, \partial_x u)\partial^2_{xx} u + f(t, x, u, \partial_x u) = 0, \quad (1.7)$$

and the nonlinear Feynman-Kac formula is also simpler:

$$Y^W_t = u(t, X^W_t), \quad Z^W_t = \partial_x u(t, X^W_t). \quad (1.8)$$

In particular, in the option pricing and hedging theory, the representation (1.8) means exactly that Z^W is the Delta-hedging portfolio h. The case that σ depends on Z indeed makes the difference between strong and weak formulations. For example, the following well known counterexample in strong formulation:

$$X_t = x + \int_0^t Z_s dB_s; \quad Y_t = X_T - \int_t^T Z_s dB_s, \tag{1.9}$$

has infinitely many solutions. However, the corresponding weak FBSDE is wellposed in the sense of Example 2.1 and Remark 2.2 below:

$$X_t = x + \int_0^t Z_s dB_s; \quad Y_t = X_T - \int_t^T Z_s dX_s, \tag{1.10}$$

Thirdly, as another major application, many FBSDEs arise from stochastic control problems through the stochastic maximum principle. However, the stochastic control problem typically does not have optimal control in strong formulation. Indeed, even the following simple problem may not have an optimal control in strong formulation:

$$X^\alpha_t = x + \int_0^t \alpha_s ds + B_t, \quad V_0 := \sup_{\alpha \in \mathcal{U}} \mathbb{E}^{\mathbb{P}_0}\Big[g(X^\alpha_\cdot) + \int_0^T f(t, \alpha_t) dt\Big], \tag{1.11}$$

The corresponding control problem in weak formulation:

$$\begin{array}{c} X_t := x + B_t, \quad B^\alpha_t := B_t - \int_0^t \alpha_s ds, \quad d\mathbb{P}^\alpha := e^{\int_0^t \alpha_s dB_s - \frac{1}{2}\int_0^t |\alpha_s|^2 ds} d\mathbb{P}_0, \\ V_0 := \sup_{\alpha\in\mathcal{U}} \mathbb{E}^{\mathbb{P}^\alpha}\Big[g(X_\cdot) + \int_0^T f(t, \alpha_t) dt\Big], \end{array} \tag{1.12}$$

has optimal control under mild and natural conditions. Consequently, the associated FBSDE will have weak solution but no strong solution. It is more natural and convenient to write the FBSDE in weak formulation when one studies weak solutions.

Fourthly, again for stochastic control problems, there are typically two approaches in the literature: the dynamic programming principle and the stochastic maximum principle. Both approaches lead to certain hamiltonians but the two hamiltonians for the same control problem look quite different. As we observe, if one uses weak FBSDE as the adjoint equation involved in the stochastic maximum principle, then the hamiltonian will coincide with the one derived from the dynamic programming principle. In this sense, the weak formulation provides an intrinsic connection between the two approaches.

After carrying out the above motivations in details, we define weak solutions for weak FBSDEs and the equivalent forward backward martingale problems. By utilizing the recently developed theory of path dependent PDEs, we establish the nonlinear Feynman-Kac formula for path dependent weak FBSDEs. That is, if the associate path dependent PDE has a classical solution, then the weak FBSDE has a (strong) solution.

Our main goal of this paper is to apply the viscosity solution method to establish the uniqueness of weak solution of the weak FBSDE. We shall follow the arguments in Ma, Zhang, & Zheng [21] and Ma & Zhang [20], which study weak solutions for FBSDEs in strong formulation in the case that σ is independent of z. Our arguments rely heavily on the regularity results for the PDE. Since such regularity results for the path dependent PDEs are not available in the literature, in this part we shall restrict to the Markovian case. The main idea is to study the so called nodal sets of the weak FBSDEs, whose upper and lower bounds provide viscosity subsolution and supersolution of the PDE. Then, provided the comparison principle for viscosity solutions of the PDE, we obtain the uniqueness of weak solutions to the weak FBSDE. We remark that, as in [21, 20], the problem is equivalent to the so called martingale problem, see also Costantini & Kurtz [5] for the application of viscosity solution methods on martingale problems in an abstract framework.

The rest of the paper is organized as follows. In Section 2 we motivate weak FBSDEs. In Section 3 we define weak solutions and establish the nonlinear Feynman-Kac formula, provided the path dependent PDE has a classical solution. In Section 4 we prove the existence and uniqueness of weak solutions for Markovian weak FBSDEs. Finally in Appendix we provide some counterexamples in control theory, which help to motivate the weak formulation, and provide some detailed arguments for the required regularities for the PDE.

2 Some motivations for weak FBSDEs

In this section we provide some heuristic motivations for weak FBSDE (1.2). To simplify the presentation, we restrict to Markovian case in one dimensional setting.

2.1 Applications in option pricing and hedging theory

Consider a financial market with a risky asset S and a risk free asset with interest rate $r = 0$ (for simplicity). Assume S satisfies the following SDE:

$$S_t = S_0 + \int_0^t \sigma(s, S_s) dB_s, \tag{2.1}$$

where B is a $\mathbb{P}_0$-Brownian motion (so we assume $\mathbb{P}_0$ is a risk neutral measure). Given a portfolio (λ, h) with value process $V_t = \lambda_t + h_t S_t$, the self-financing condition states that

$$dV_t = h_t dS_t. \tag{2.2}$$

Now given an European type of option with payoff ξ at terminal time T, we say a self-financing portfolio (λ, h) is a hedging portfolio if $V_T = \xi$, $\mathbb{P}$-a.s. This, combining with (2.2), leads to a backward SDE against dS_t:

$$V_t = \xi - \int_t^T h_s dS_s. \tag{2.3}$$

Then (2.1)-(2.3) become a decoupled weak FBSDE with solution $(X, Y, Z) = (S, V, h)$. We remark that BSDE (2.3) can be rewritten in strong formulation:

$$V_t = \xi - \int_t^T \tilde{h}_s dB_s, \quad \text{where} \quad \tilde{h}_t := h_t \sigma(t, S_t). \tag{2.4}$$

In particular, when $\sigma > 0$, (2.3) and (2.4) are equivalent. This is why many papers in the literature could use the strong formulation.

The situation is different, however, in incomplete markets. For example, consider the case that S is scalar but B is multi-dimensional. Then σ is a vector, and we assume that σ is Lipschitz in S so that (2.1) has a strong solution S. Assume further that we observe the noise B but can trade only S. Then ξ is in general $\mathbb{F}^B$-measurable. By the martingale representation theorem, BSDE (2.4) always admits a solution $(V, \tilde{h})$. However, since one cannot trade B, the process $\tilde{h}$ is not a legitimate trading portfolio. For practical purpose one has to solve the weak BSDE (2.3). In general $\tilde{h}$ may not be in the form of $h\sigma$, then in this case the strong BSDE (2.4) and the weak BSDE (2.3) are not equivalent and in general the weak BSDE (2.3) may not have a solution (V, h). One sensible resolution is to consider the super-hedging price:

$$V_0 := \inf\{y : \exists h \text{ such that } y + \int_0^T h_s dS_s \geq \xi, \ \mathbb{P}_0\text{-a.s.}\}. \tag{2.5}$$

This is in the sprit of the weak FBSDE. Indeed, one can formulate it as a reflected BSDE in weak formulation, which is beyond the scope of this paper and is left for future research.

An alternative explanation for the nonexistence of solution to the weak FBSDE in above situation is that X does not have martingale representation property for $\mathbb{F}^B$-martingales. In this case, for theoretical interest we may relax BSDE (2.3) by applying the extended martingale representation theorem, see e.g. Protter [25]:

$$V_t = \xi - \int_t^T h_s dS_s + N_T - N_t, \tag{2.6}$$

where N is an orthogonal martingale such that $N_0 = 0$ and $d\langle S, N\rangle_t = 0$. Then (2.6) will have a unique solution (V, h, N).

2.2 Nonlinear Feynman-Kac formula

As is well known, in the case $\sigma = \sigma(t, x, y)$, the strong FBSDE (1.1) is associated with the quasilinear PDE (1.5) via the nonlinear Feynman-Kac formula (1.6). The problem becomes tricky when $\sigma = \sigma(t,x,y,z)$ because the PDE will involve the inverse function ψ in (1.3), which typically does not exist. The weak FBSDE (1.2) is associated with the quasilinear PDE (1.7), which is more natural at least in the following aspects:

- σ may depend on z and the PDE does not involve the inverse function ψ in (1.3).
- The component Z of the solution corresponds to $\partial_x u$ directly, rather than $\partial_x u\,\sigma$. In particular, in the application to the option pricing and hedging theory, the Z in weak formulation corresponds directly to the Delta-hedging portfolio.
- The PDE is more natural in the sense that the coefficients σ, f depend directly on $\partial_x u$, instead of $\partial_x u \sigma$.
- It is more convenient to study weak solutions of the weak FBSDE, which is closely related to the viscosity solution of the PDE (1.7), than that of the strong FBSDE.

To see the advantage of the weak formulation more directly in the case that σ depends on z, let's consider the counterexample (1.9). It is well known that (1.9) has infinitely many solutions. Indeed, for any $Z \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{P}_0)$, $X_t := Y_t := x + \int_0^t Z_s dB_s$ is a solution to (1.9). However, the weak FBSDE (1.10) is wellposed in the following sense.

Example 2.1. *The weak FBSDE* (1.10) *has a unique solution such that* $Z \in \mathcal{Z} := \{Z \in \mathbb{L}^4(\mathbb{P}_0) : Z \neq 0\}$.

Note that $Z \in \mathbb{L}^4(\mathbb{P}_0)$ implying $\mathbb{E}^{\mathbb{P}_0}[\int_0^T |Z_t|^2 d\langle B\rangle_t + \int_0^T |Z_t|^2 d\langle X\rangle_t] < \infty$, and thus X, Y are $\mathbb{P}_0$-martingales. We shall comment on the requirement $Z \neq 0$ in Remark 2.2 below.

Proof It is clear that

$$X_t = Y_t = x + B_t, \qquad Z_t = 1 \tag{2.7}$$

is a solution to (1.10). We next show that it's the unique solution such that $Z \in \mathcal{Z}$.

For any (t,x,y) and $Z \in \mathbb{L}^4(\mathbb{P}_0)$, denote

$$X_s^{t,x,Z} := x + \int_t^s Z_r dB_r, \quad Y_s^{t,x,y,Z} := y + \int_t^s Z_r dX_r^{t,x,Z} = y + \int_t^s |Z_r|^2 dB_r,$$

and define

$$\begin{aligned} \overline{u}(t,x) &:= \inf\Big\{y : \exists Z \in \mathbb{L}^4(\mathbb{P}_0) \text{ such that } Y_T^{t,x,y,Z} \geq X_T^{t,x,Z},\ \mathbb{P}_0\text{-a.s.}\Big\}; \\ \underline{u}(t,x) &:= \sup\Big\{y : \exists Z \in \mathbb{L}^4(\mathbb{P}_0) \text{ such that } Y_T^{t,x,y,Z} \leq X_T^{t,x,Z},\ \mathbb{P}_0\text{-a.s.}\Big\}. \end{aligned} \tag{2.8}$$

Note that both $X^{t,x,Z}$ and $Y^{t,x,y,Z}$ are $\mathbb{P}_0$-martingales, then $Y_T^{t,x,y,Z} \geq X_T^{t,x,Z}$, $\mathbb{P}_0$-a.s. implies $y = \mathbb{E}^{\mathbb{P}_0}[Y_T^{t,x,y,Z}] \geq \mathbb{E}^{\mathbb{P}_0}[X_T^{t,x,Z}] = x$, and thus $\overline{u}(t,x) \geq x$. Similarly, $\underline{u}(t,x) \leq x$ and thus $\underline{u}(t,x) \leq x \leq \overline{u}(t,x)$. On the other hand, for any solution (X,Y,Z) to (1.2), by the definition of $\overline{u}(t,x)$ and $\underline{u}(t,x)$ we see that $\overline{u}(t,X_t) \le Y_t \le \underline{u}(t,X_t)$. Thus $\overline{u}(t,x) = \underline{u}(t,x) = u(t,x) := x$ and $Y_t = u(t,X_t) = X_t$. This implies further that $Z_t = |Z_t|^2$. Since $Z \neq 0$, we see that $Z = 1$ and hence (2.7) is the unique solution. ∎

Remark 2.2. (i) If we allow $Z = 0$, then the solution is not unique. Indeed, for any Z satisfying $Z = |Z|^2$ (namely Z takes values 0 and 1), it is clear that $X_t = Y_t = x + \int_0^t Z_s dB_s$ is a solution to weak FBSDE (1.2). However, we note that even in this case, the relationship $Y_t = X_t$ still holds, and the decoupling function $u(t,x) = x$ is still unique. Moreover, without surprise, $u(t,x) = x$ is a solution to the PDE (1.7) corresponding to $b = 0, \sigma = z, f = 0$:

$$\partial_t u + \frac{1}{2}|\partial_x u|^2 \partial_{xx}^2 u = 0, \quad u(T,x) = x.$$

(ii) When $Z = 0$, this is exactly the case that X has degenerate diffusion coefficient σ. As we will see in the paper, the nondegeneracy of σ is crucial.

(iii) As we mentioned in (i), even if we allow $Z = 0$, the decoupling function $u(t,x) = x$ is still unique. However, when X can be degenerate, $Y_t = u(t,X_t)$ and $dY_t = Z_t dX_t$ do not imply $Z_t = \partial_x u(t,X_t) = 1$. That's why the uniqueness fails in this degenerate case. ∎

To avoid the degeneracy issue, we may modify the example as follows.

Example 2.3. *Let $\sigma > 0$ be bounded such that the fixed point set $\mathcal{N} := \{z : \sigma(z) = z\} \neq \emptyset$.*

(i) For any $\mathcal{N}$-valued $Z \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{P}_0)$, $Y_t := X_t := x + \int_0^t Z_s dB_s$ is a solution to the following strong FBSDE:

$$X_t = x + \int_0^t \sigma(Z_s) dB_s, \quad Y_t = X_T - \int_t^T Z_s dB_s.$$

(ii) The corresponding weak FBSDE

$$X_t = x + \int_0^t \sigma(Z_s) dB_s, \quad Y_t = X_T - \int_t^T Z_s dX_s.$$

has a unique solution

$$Y_t := X_t := x + \sigma(1) B_t, \quad Z_t := 1.$$

Here the uniqueness holds for $Z \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{P}_0)$.

Proof (i) is obvious, and (ii) follows the same arguments as in Example 2.1. In particular, the weak BSDE can be rewritten as:

$$X_t = x + \int_0^t \sigma(Z_s) dB_s, \quad Y_t = X_T - \int_t^T Z_s \sigma(Z_s) dB_s,$$

and then we see that $Z = 1$ is the unique fixed point of: $\sigma(z) = z\sigma(z)$, thanks to the nondegeneracy of σ. Moreover, since σ is bounded, then $\mathbb{E}^{\mathbb{P}_0}[\int_0^T |Z_t|^2 d\langle X\rangle_t] < \infty$ for any $Z \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{P}_0)$, so the uniqueness holds for $Z \in \mathbb{L}^2(\mathbb{F}^B, \mathbb{P}_0)$. ∎

2.3 Connections with stochastic control theory

2.3.1 Stochastic control in strong formulation

Consider a simple stochastic control problem in strong formulation:

$$V_0 := \sup_{\alpha\in\mathcal{A}} V_0^\alpha, \quad \text{where} \tag{2.9}$$

$$X_t^\alpha := \int_0^t b(s,\alpha_s)ds + \int_0^t \sigma(s,\alpha_s)dB_s,\ V_0^\alpha := \mathbb{E}^{\mathbb{P}_0}\Big[g(X_T^\alpha) + \int_0^T f(t,\alpha_t)dt\Big].$$

Here the admissible controls α are $\mathbb{F}^B$-progressively measurable. Note that $V_0^\alpha = Y_0^\alpha$, where

$$Y_t^\alpha = g(X_T^\alpha) + \int_t^T f(s,\alpha_s)ds - \int_t^T Z_s^\alpha dB_s. \tag{2.10}$$

We first use the stochastic maximum principle to derive an associated FBSDE. Let $\Delta\alpha$ be given such that $\alpha + \varepsilon\Delta\alpha \in \mathcal{A}$ for any $\varepsilon \in [0,1]$. Denote

$$\nabla X^{\alpha,\Delta\alpha} := \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[X^{\alpha+\varepsilon\Delta\alpha} - X^\alpha], \quad \nabla V_0^{\alpha,\Delta\alpha} := \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[V_0^{\alpha+\varepsilon\Delta\alpha} - V_0^\alpha].$$

One can easily see that

$$\begin{aligned}\nabla X_t^{\alpha,\Delta\alpha} &= \int_0^t b'(s,\alpha_s)\Delta\alpha_s ds + \int_0^t \sigma'(s,\alpha_s)\Delta\alpha_s dB_s,\\ \nabla V_0^{\alpha,\Delta\alpha} &= \mathbb{E}^{\mathbb{P}_0}\Big[\partial_x g(X_T^\alpha)\nabla X_T^{\alpha,\Delta\alpha} + \int_0^T f'(t,\alpha_t)\Delta\alpha_t dt\Big],\end{aligned}$$

where b', σ', f' refer to the derivatives with respect to α. Introduce an adjoint BSDE:

$$\tilde{Y}_t^\alpha = \partial_x g(X_T^\alpha) - \int_t^T \tilde{Z}_s^\alpha dB_s. \tag{2.11}$$

By applying Itô formula on $\tilde{Y}_t^\alpha \nabla X_t^{\alpha,\Delta\alpha}$ we obtain

$$\nabla V_0^{\alpha,\Delta\alpha} = \mathbb{E}^{\mathbb{P}_0}\Big[\int_0^T [\tilde{Y}_t^\alpha b'(t,\alpha_t) + \tilde{Z}_t^\alpha \sigma'(t,\alpha_t) + f'(t,\alpha_t)]\Delta\alpha_t dt\Big].$$

Now assume $\alpha^* \in \mathcal{A}$ is an interior point of $\mathcal{A}$ and is an optimal control. Then $\nabla V_0^{\alpha^*,\Delta\alpha} \le 0$ for arbitrary $\Delta\alpha$. This implies

$$\tilde{Y}_t^{\alpha^*} b'(t,\alpha_t^*) + \tilde{Z}_t^{\alpha^*}\sigma'(t,\alpha_t^*) + f'(t,\alpha_t^*) = 0. \tag{2.12}$$

Assume further that (2.12) determines an α^*: $\alpha_t^* = I(t, \tilde{Y}_t^{\alpha^*}, \tilde{Z}_t^{\alpha^*})$ for a function I. Then combining (2.9)-(2.11), we obtain the following coupled FBSDE in strong formulation:

$$\begin{cases} X_t = \int_0^t b(s, I(s,\tilde{Y}_s,\tilde{Z}_s))ds + \int_0^t \sigma(s, I(s,\tilde{Y}_s,\tilde{Z}_s))dB_s;\\ Y_t = g(X_T) + \int_t^T f(s, I(s,\tilde{Y}_s,\tilde{Z}_s))ds - \int_t^T Z_s dB_s;\\ \tilde{Y}_t = \partial_x g(X_T) - \int_t^T \tilde{Z}_s dB_s. \end{cases} \tag{2.13}$$

However, the above FBSDE is typically not covered by the existing methods in the literature, especially since σ depends on $\tilde{Z}$. We remark that all the existing works on weak solutions of (strong) FBSDEs do not allow σ depending on Z, see e.g. Antonelli & Ma [2], Delarue & Guatteri [8], Ma, Zhang & Zheng [21], and Ma & Zhang [20].

We thus turn to weak FBSDE for which we can study weak solutions more conveniently. Rewrite the adjoint BSDE (2.11) in the spirit of weak formulation:

$$\hat{Y}_t^\alpha = \partial_x g(X_T^\alpha) + \int_t^T b(s,\alpha_s)\hat{Z}_s^\alpha ds - \int_t^T \hat{Z}_s^\alpha dX_s^\alpha. \tag{2.14}$$

One can easily see that its solution is: again assuming $\sigma > 0$,

$$\hat{Y}_t^\alpha := \tilde{Y}_t^\alpha, \quad \hat{Z}_t^\alpha := \tilde{Z}_t^\alpha \sigma^{-1}(t,\alpha_t), \tag{2.15}$$

and the optimality condition (2.12) becomes

$$\hat{Y}_t^{\alpha^*} b'(t,\alpha_t^*) + \hat{Z}_t^{\alpha^*}\sigma\sigma'(t,\alpha_t^*) + f'(t,\alpha_t^*) = 0. \tag{2.16}$$

Assume the above determines an optimal α^*: $\alpha_t^* = \hat{I}(t,\hat{Y}_t^{\alpha^*},\hat{Z}_t^{\alpha^*})$ for a function $\hat{I}$. Then (2.13) becomes a (multidimensional) FBSDE in weak formulation:

$$\begin{cases} X_t = \int_0^t b(s,\hat{I}(s,\hat{Y}_s,\hat{Z}_s))ds + \int_0^t \sigma(s,\hat{I}(s,\hat{Y}_s,\hat{Z}_s))dB_s; \\ Y_t = g(X_T) + \int_t^T [f(s,\hat{I}(s,\hat{Y}_s,\hat{Z}_s)) + b(s,\hat{I}(s,\hat{Y}_s,\hat{Z}_s))Z_s]ds - \int_t^T Z_s dX_s; \\ \hat{Y}_t = \partial_x g(X_T) + \int_t^T b(s,\hat{I}(s,\hat{Y}_s,\hat{Z}_s))\hat{Z}_s ds - \int_t^T \hat{Z}_s dX_s. \end{cases} \tag{2.17}$$

Remark 2.4. When the weak FBSDE (2.17) has no strong solution, but only weak solution, the stochastic optimization problem (2.9) in strong formulation still does not have optimal control. To obtain the existence of optimal control, it is more appropriate to study the optimization problem in weak formulation, see Subsection 2.3.3 below. ■

2.3.2 Consistency with dynamic programming principle

As is well known, another standard approach for stochastic control problem is the dynamic programming principle, which focuses more on the value function. Assume the control α takes values in A. Then $V_0 = u(0,0)$, where u satisfies the following HJB equation:

$$\partial_t u + H(t,\partial_x u,\partial_{xx}^2 u) = 0, \quad u(T,x) = g(x), \tag{2.18}$$
$$\text{where } H(t,z,\gamma) := \sup_{\alpha\in A}\Big[\frac{1}{2}\sigma^2(t,\alpha)\gamma + b(t,\alpha)z + f(t,\alpha)\Big].$$

Assuming u is sufficiently smooth and FBSDE (2.13) is wellposed. By Yong & Zhou [31] Chapter 5, Theorem 4.1 we have

$$Y_t = u(t, X_t),\ Z_t = \partial_x u(t, X_t)\sigma(t, \alpha^*_t),\ \tilde{Y}_t = \partial_x u(t, X_t),\ \tilde{Z} = \partial^2_{xx} u(t, X_t)\sigma(t, \alpha^*_t), \quad (2.19)$$

where $\alpha^*_t = I(t, \tilde{Y}_t, \tilde{Z}_t)$ is the optimal control. On the other hand, notice that the optimality condition (2.12) can be viewed as the first order condition of

$$\tilde{H}(t, \tilde{y}, \tilde{z}) := \sup_{\alpha \in A} \big[\tilde{y} b(t, \alpha) + \tilde{z}\sigma(t, \alpha) + f(t, \alpha)\big].$$

However, we have the following discrepancy which has already been noticed in [31]:

$$\begin{aligned} \tilde{H}(t, X_t, \tilde{Y}_t, \tilde{Z}_t) &= \partial_x u(t, X_t) b(t, \alpha^*_t) + \partial^2_{xx} u(t, X_t)\sigma^2(t, \alpha^*_t) + f(t, X_t, \alpha^*_t) \\ &\neq H(t, X_t, \partial_x u(t, X_t), \partial^2_{xx} u(t, X_t)). \end{aligned} \quad (2.20)$$

This discrepancy is due to the fact that $\tilde{Z}$ involves $\sigma(t, \alpha)$ and thus twisted the optimization in the Hamiltonian. It will disappear if we consider the weak FBSDE (2.17). Indeed, in this case the optimality condition (2.16) can be viewed as the first order condition of

$$\hat{H}(t, \hat{y}, \hat{z}) := \sup_{\alpha \in A} \big[\hat{y} b(t, \alpha) + \frac{1}{2}\hat{z}\sigma^2(t, \alpha) + f(t, \alpha)\big]. \quad (2.21)$$

Similar to (2.19) we have the correspondence for the solution to weak FBSDE (2.17):

$$Y_t = u(t, X_t), \quad Z_t = \partial_x u(t, X_t), \quad \hat{Y}_t = \partial_x u(t, X_t), \quad \hat{Z} = \partial^2_{xx} u(t, X_t). \quad (2.22)$$

Then we have the desired identity:

$$\begin{aligned} \hat{H}(t, \hat{Y}_t, \hat{Z}_t) &= \partial_x u(t, X_t) b(t, \alpha^*_t) + \frac{1}{2}\partial^2_{xx} u(t, X_t)\sigma^2(t, \alpha^*_t) + f(t, \alpha^*_t) \\ &= H(t, \partial_x u(t, X_t), \partial^2_{xx} u(t, X_t)). \end{aligned} \quad (2.23)$$

Remark 2.5. (i) It is clear that $\hat{H} = H$, with the correspondence $\hat{y} = z, \hat{z} = \gamma$. This is reflected in (2.22). In particular, we have $\hat{Y}_t = Z_t$ in this model.

(ii) The derivation of (2.16) requires the differentiation of the coefficients b, σ, f in α. However, such differentiation is not needed for the optimization of the Hamiltonian in (2.21). In fact, one may determine $\hat{I}$ by the optimal arguments in (2.21), and then formally derive the same FBSDE (2.17). These arguments are in the line of dynamic programming principle, rather than stochastic maximum principle. ■

2.3.3 Stochastic drift control under weak formulation

To understand the weak FBSDE (2.17) better, we consider a special case that

$$\sigma = 1.$$

The general case with diffusion control will involve the second order BSDE introduced in Soner, Touzi, & Zhang [27]. In this case (2.16) becomes

$$\hat{Y}_t^{\alpha^*} b'(t, \alpha_t^*) + f'(t, \alpha_t^*) = 0. \tag{2.24}$$

Then the optimal control takes the form $\alpha_t^* = \hat{I}(t, \hat{Y}_t^{\alpha^*})$ and thus FBSDE (2.17) becomes

$$\begin{cases} X_t = \int_0^t b(s, \hat{I}(s, \hat{Y}_s))ds + B_t; \\ Y_t = g(X_T) + \int_t^T [f(s, \hat{I}(s, \hat{Y}_s)) + b(s, \hat{I}(s, \hat{Y}_s))Z_s]ds - \int_t^T Z_s dX_s; \\ \hat{Y}_t = \partial_x g(X_T) + \int_t^T b(s, \hat{I}(s, \hat{Y}_s))\hat{Z}_s ds - \int_t^T \hat{Z}_s dX_s. \end{cases} \tag{2.25}$$

Recalling (2.22) that $\hat{Y} = Z$, the second equation in (2.25) is equivalent to

$$Y_t = g(X_T) + \int_t^T [f(s, \hat{I}(s, Z_s)) + b(s, \hat{I}(s, Z_s))Z_s]ds - \int_t^T Z_s dX_s. \tag{2.26}$$

Moreover, note that (2.24) is the first order condition of the following optimization problem:

$$f^*(t, z) := \sup_{\alpha \in A} [zb(t, \alpha) + f(t, \alpha)]. \tag{2.27}$$

Then, together with (2.25) and under appropriate technical conditions, (2.26) leads to

$$\begin{cases} X_t = \int_0^t I(s, Z_s)ds + B_t; \\ Y_t = g(X_T) + \int_t^T f^*(s, Z_s)ds - \int_t^T Z_s dX_s. \end{cases} \tag{2.28}$$

The FBSDE (2.28) can be understood a lot easier if we use weak formulation for the control problem:

$$\bar{V}_0 := \sup_{\alpha \in \mathcal{A}} \bar{V}_0^\alpha := \sup_{\alpha \in \mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\Big[g(X_T) + \int_0^T f(t, \alpha_t)dt\Big] \tag{2.29}$$

$$\text{where} \quad X_t := B_t, \quad d\mathbb{P}^\alpha := \exp\Big(\int_0^T \alpha_t dB_t - \frac{1}{2}\int_0^T |\alpha_s|^2 dt\Big)d\mathbb{P}_0.$$

Note that $\bar{V}_0^\alpha = \bar{Y}_0^\alpha$, where, $B_t^\alpha := B_t - \int_0^t \alpha_s ds$ is a $\mathbb{P}^\alpha$-Brownian motion and

$$\begin{aligned} \bar{Y}_t^\alpha &= g(X_T) + \int_t^T f(s, \alpha_s)ds - \int_t^T \bar{Z}_s^\alpha dB_s^\alpha \\ &= g(X_T) + \int_t^T [f(s, \alpha_s) + \alpha_s \bar{Z}_s^\alpha]ds - \int_t^T \bar{Z}_s^\alpha dX_s. \end{aligned} \tag{2.30}$$

Consider the BSDE

$$\bar{Y}_t = g(X_T) + \int_t^T f^*(s, \bar{Z}_s)ds - \int_t^T \bar{Z}_s dX_s, \quad \mathbb{P}_0\text{-a.s.} \tag{2.31}$$

By comparison of BSDE we see immediately that $\bar{V}_0 = \bar{Y}_0$ and $\alpha_t^* := \hat{I}(t, \bar{Z}_t)$ is an optimal control of (2.29), for the same $\hat{I}$ in (2.28). Now together with the definition of B^α and $X = B$, we may rewrite (2.31) as

$$\begin{cases} X_t = \int_0^t \hat{I}(s, \bar{Z}_s)ds + B_t^{\alpha^*}, \\ \bar{Y}_t = g(X_T) + \int_t^T f^*(s, \bar{Z}_s)ds - \int_t^T \bar{Z}_s dX_s. \end{cases} \tag{2.32}$$

In the spirit of weak solution as we will introduce in the next section, this is equivalent to (2.28). So in this sense, the weak FBSDE (2.28), or the more general one (2.17), is more in the spirit of weak formulation.

Remark 2.6. (i) Recall (2.9) with $\sigma = 1$ and (2.29). Note that formally $(B^\alpha, B, \mathbb{P}^\alpha)$ in weak formulation corresponds to $(B, X^\alpha, \mathbb{P}_0)$ in strong formulation. However, for fixed α, note that in general α has different distributions under $\mathbb{P}^\alpha$ and under $\mathbb{P}_0$, so $V_0^\alpha \neq \bar{V}_0^\alpha$. But nevertheless, under appropriate conditions, their optimal values are equal: $V_0 = \bar{V}_0$. See more discussions along this line in Zhang [33] Chapter 9.

(ii) There are many situations that the optimal control in weak formulation exists but that in strong formulation does not. See some examples in Appendix.

(iii) The difference between strong formulation and weak formulation becomes more crucial when one considers zero sum stochastic differential games, see Hamadene & Lepeltier [13] and Pham & Zhang [24]. ■

3 Weak solutions of FBSDEs and Feynman-Kac formula

Our objective is the following weak FBSDE:

$$\begin{cases} X_t = x + \int_0^t b(s, X_\cdot, Y_s, Z_s)ds + \int_0^t \sigma(s, X_\cdot, Y_s, Z_s)dB_s, \\ Y_t = g(X_\cdot) + \int_t^T f(s, X_\cdot, Y_s, Z_s)ds - \int_t^T Z_s dX_s + N_T - N_t, \end{cases} \quad \mathbb{P}_0\text{-a.s.} \tag{3.1}$$

Here (B, X, Y) take values in $\mathbb{R}^{d_0} \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, and all other processes and functions have appropriate dimensions. The coefficients b, σ, f, g may depend on the paths of X, among them b, σ, f are $\mathbb{F}^X$-progressively measurable in all variables, and g is $\mathcal{F}_T^X$-measurable.

Given a probability space $(\Omega, \mathbb{F}, \mathbb{P})$, let $\mathbb{L}^0(\mathbb{F})$ denote the set of $\mathbb{F}$-progressively measurable processes with appropriate dimensions. For $p, q \geq 1$, denote

$$\mathbb{L}^{p,q}(\mathbb{F},\mathbb{P}) := \Big\{X \in \mathbb{L}^0(\mathbb{F}) : \mathbb{E}^{\mathbb{P}}\Big[\Big(\int_0^T |X_t|^p dt\Big)^{\frac{q}{p}}\Big] < \infty\Big\}, \quad \mathbb{L}^p(\mathbb{F},\mathbb{P}) := \mathbb{L}^{p,p}(\mathbb{F},\mathbb{P});$$
$$\mathbb{S}^p(\mathbb{F},\mathbb{P}) := \Big\{X \in \mathbb{L}^0(\mathbb{F}) : X \text{ is continuous, } \mathbb{P}\text{-a.s. and } \mathbb{E}^{\mathbb{P}}\big[\sup_{0\le t\le T} |X_t|^p\big] < \infty\Big\}.$$

Throughout this paper, we shall assume

Assumption 3.1. *(i) b, σ are bounded.*

(ii) $f(t, \mathbf{x}, 0, 0)$, $g(\mathbf{x})$ have polynomial growth in $\|\mathbf{x}\| := \sup_{0\le t\le T} |\mathbf{x}_t|$, and f is uniformly Lipschitz continuous in (y, z).

(iii) $\sigma\sigma^\top \geq c_0^2 I_{d_1}$ as $d_1 \times d_1$-matrice, for some constant $c_0 > 0$.

Remark 3.2. (i) For strong FBSDEs, typically the coefficients may depend on B. For weak FBSDEs, both for practical considerations and for theoretical reasons, it is more natural that the coefficients depend on X. However, in a more general setting, for example in the incomplete market with observable noise as in Subsection 2.1, we may allow the coefficients to depend on both X and B. The problem will become harder in this case. In this paper we restrict to the case that the coefficients do not depend on B.

(ii) As explained in Section 2.1, the presence of N is due to the fact that X may not satisfy the martingale representation property. ■

3.1 Definitions

We introduce the following types of solutions. Recall $\Theta = (X, Y, Z)$.

Definition 3.3. *(i) We say a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$ and a quintuple of $\mathbb{F}$-progressively measurable processes (B, Θ, N) is a weak solution of the weak FBSDE* (3.1) *if B is a $\mathbb{P}$-Brownian motion, $N \in \mathbb{S}^2(\mathbb{F},\mathbb{P})$ is a $\mathbb{P}$-martingale orthogonal to X with $N_0 = 0$, $X, Y \in \mathbb{S}^2(\mathbb{F},\mathbb{P})$, $Z \in \mathbb{L}^2(\mathbb{F},\mathbb{P})$, and* (3.1) *holds $\mathbb{P}$-a.s.*

(ii) We say a weak solution is semi-strong if (Y, Z) are $\mathbb{F}^X$-progressively measurable.

(iii) We say a weak solution is strong if $N = 0$ and Θ is $\mathbb{F}^B$-progressively measurable.

Given our conditions, all weak solutions actually have stronger integrability.

Lemma 3.4. *Let Assumption 3.1 hold and $(B, \Theta, N, \mathbb{P})$ be a weak solution to* (3.1)*. Then*

$$\mathbb{E}^{\mathbb{P}}\Big[\sup_{0\le t\le T}\big[|X_t|^p + |Y_t|^p + |N_t|^p\big] + \Big(\int_0^T |Z_t|^2 dt\Big)^{\frac{p}{2}}\Big] < \infty, \text{ for any } p \geq 1.$$

Proof By the boundedness of b, σ, the estimate for X is obvious. Since $f(t, \mathbf{x}, 0, 0)$ and $g(\mathbf{x})$ have polynomial growth, we have $\mathbb{E}^{\mathbb{P}}\Big[|g(X.)|^p + \int_0^T |f(t, X.)|^p dt\Big] < \infty$. Now by the uniform Lipschitz continuity of f in (y, z), the rest estimates follows from standard BSDE arguments, see e.g. El Karoui & Huang [11]. ■

As in Stroock & Varadahn [28], weak solutions are closely related to martingale problems. Motivated by Ma, Zhang & Zheng [21] which studies strong FBSDE with σ independent of z, we introduce the following forward-backward martingale problem.

Definition 3.5. Let $\Omega := C([0,T], \mathbb{R}^{d_1}) \times C([0,T], \mathbb{R}^{d_2})$ be the canonical space, (X, Y) the canonical processes, and $\mathbb{F} = \mathbb{F}^{X,Y}$ the natural filtration. We say $(\mathbb{P}, Z)$ is a solution to the forward-backward martingale problem of (3.1) if:

(i) $\mathbb{P}(X_0 = x) = \mathbb{P}(Y_T = g(X.)) = 1$ and $X, Y \in \mathbb{S}^2(\mathbb{F}, \mathbb{P})$, $Z \in \mathbb{L}^2(\mathbb{F}, \mathbb{P})$.

(ii) The following two processes are $\mathbb{P}$-martingales:

$$M_t^X := X_t - \int_0^t b(s, X., Y_s, Z_s)ds,$$
$$M_t^Y := Y_t + \int_0^t f(s, X., Y_s, Z_s)ds - \int_0^t Z_s b(s, X., Y_s, Z_s)ds.$$

(iii) $d\langle M^X\rangle_t = \sigma\sigma^\top(t, X., Y_t, Z_t)dt$ and $d\langle M^Y, M^X\rangle_t = Z_t d\langle X\rangle_t$, $\mathbb{P}$-a.s. ■

Proposition 3.6. *Let Assumption 3.1 hold. Then a weak solution to FBSDE* (3.1) *is equivalent to a solution to the forward-backward martingale problem of* (3.1).

Proof Let $(\Omega, \mathbb{F}, \mathbb{P}, B, \Theta, N)$ be a weak solution to FBSDE (3.1). Note that $d\langle Y, X\rangle_t = Z_t d\langle X\rangle_t = Z_t\sigma\sigma^\top(t, X., Y_t, Z_t)dt$ and $\langle X\rangle, \langle Y\rangle$ are all $\mathbb{F}^{X,Y}$-progressively measurable. Since $\sigma\sigma^\top > 0$, then Z is also $\mathbb{F}^{X,Y}$-progressively measurable. Now by recasting everything into the canonical space of (X, Y), it is straightforward to verify that $(\mathbb{P}, Z)$ is a solution to the forward-backward martingale problem of (3.1).

To see the other direction, let $(\Omega, \mathbb{F}, X, Y)$ be the canonical setting in Definition 3.5 and $(\mathbb{P}, Z)$ a solution to the forward-backward martingale problem of (3.1). Note that Assumption 3.1 (iii) implies $d_0 \geq d_1$, and there exist orthogonal matrices $U \in \mathbb{R}^{d_1 \times d_1}$ and $V \in \mathbb{R}^{d_0 \times d_0}$ as well as $k_1, \cdots, k_{d_1} \neq 0$ such that

$$\sigma(t, X., Y_t, Z_t) = U_t\ [K_t, 0]\ V_t \quad \text{where } K \text{ is the diagonal matrix of } k_1, \cdots, k_{d_1},$$

and 0 refers to the $d_1 \times (d_0 - d_1)$-zero matrix. It is clear that U, V, K are $\mathbb{F}$-progressively measurable processes. Denote

$$\tilde{B}_t := \int_0^t K_s^{-1} U_s^\top dM_s^X.$$

Then $\tilde{B}$ is a continuous local martingale under $\mathbb{P}$ and

$$d\langle\tilde{B}\rangle_t = K_t^{-1}U_t^\top\sigma\sigma^\top U_t K_t^{-1}dt = K_t^{-1}U_t^\top U_t[K_t,0]VV^\top[K_t,0]^\top U_t^\top U_t K_t^{-1}dt = I_{d_1}dt.$$

By Levy's characterization theorem we see that $\tilde{B}$ is a $\mathbb{P}$-Brownian motion. Now let $\bar{B}$ be an $d_0 - d_1$-dimensional Brownian motion independent of $\mathbb{F}$, and let us extend $\mathbb{F}$ to $\hat{\mathbb{F}} := \mathbb{F}\vee\mathbb{F}^{\bar{B}}$, and still denote the probability measure as $\mathbb{P}$. Then $\hat{B} := [B^\top, \bar{B}^\top]^\top$ is a d_0-dimensional $\mathbb{P}$-Brownian motion. Thus

$$dM_t^X = U_t K_t d\tilde{B}_t = U_t[K_t,0]d\hat{B}_t = \sigma(t, X_\cdot, Y_t, Z_t)dB_t,$$

where $dB_t := V_t^\top d\hat{B}_t$ is also a d_0-dimensional $\mathbb{P}$-Brownian motion, since V is orthogonal. Now define

$$N_t := Y_0 - M_t^Y + \int_0^t Z_s dM_s^X. \tag{3.2}$$

Then N is a $\mathbb{P}$-martingale. Note that $d\langle X, N\rangle_t = -d\langle X, M^Y\rangle_t + Z_t d\langle X\rangle_t = 0$. Then $(\Omega, \hat{\mathbb{F}}, \mathbb{P}, B, X, Y, Z, N)$ is a weak solution to FBSDE (3.1). ■

Remark 3.7. Note that the martingale problem involves only $\sigma\sigma^\top$, not the σ itself. Then by Proposition 3.6 we may assume without loss of generality that

$$d_0 = d_1 =: d, \quad \sigma \text{ is symmetric and } \sigma \geq c_0 I_d. \tag{3.3}$$

In the rest of the paper this will be enforced. ■

Given (3.3), we have another equivalence result.

Proposition 3.8. *Let Assumption 3.1 hold. Then FBSDE* (3.1) *admits a weak solution if and only if* (3.1) *with coefficients* $(0, \sigma, f, g)$ *has a weak solution.*

Proof We assume without loss of generality that (3.3) holds. Let $(B, \Theta, N, \mathbb{P})$ be a weak solution to FBSDE (3.1) with coefficients (b, σ, f, g). Denote

$$\theta_t := -\sigma^{-1}b(t, X_\cdot, Y_t, Z_t),\ \tilde{B}_t := B_t - \int_0^t \theta_s ds,\ d\tilde{\mathbb{P}} := \exp\Big(\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t |\theta_s|^2 ds\Big)d\mathbb{P}.$$

Then θ is bounded, and thus it follows from Lemma 3.4 that (Θ, N) have the desired integrability under $\tilde{\mathbb{P}}$. Since B and N are orthogonal, by Girsanov theorem one can easily check that $(\tilde{B}, \Theta, N, \tilde{\mathbb{P}})$ is a weak solution to FBSDE (3.1) with coefficients $(0, \sigma, f, g)$. This proves the only if part. The if part can be proved similarly. ■

3.2 Path dependent PDEs

In this subsection we introduce the PPDE in the setting of Ekren, Touzi, & Zhang [9, 10]. Let $\Omega := C([0,T],\mathbb{R}^d)$ be the canonical space equipped with $\|\omega\| := \sup_{0\le t\le T}|\omega_t|$, X the canonical process, $\mathbb{F} := \mathbb{F}^X$ the natural filtration, and $\Lambda := [0,T]\times\Omega$ equipped with

$$\mathbf{d}((t,\omega),(t',\omega')) := |t-t'| + \sup_{0\le s\le T}|\omega_{t\wedge s} - \omega'_{t'\wedge s}|.$$

For some generic dimension m, let $C^0(\Lambda;\mathbb{R}^m)$ be the space of continuous functions $\Lambda \to \mathbb{R}^m$.

Next, let $\mathcal{P}$ denote the set of semimartingale measures $\mathbb{P}$ whose drift and diffusion characteristics are bounded, and $C^{1,2}(\Lambda;\mathbb{R})$ be the space of $u \in C^0(\Lambda;\mathbb{R})$ such that there exist $\partial_t u \in C^0(\Lambda;\mathbb{R})$, $\partial_\omega u \in C^0(\Lambda;\mathbb{R}^{1\times d})$ (row vector for convenience!), and symmetric $\partial^2_{\omega\omega}u \in C^0(\Lambda;\mathbb{R}^{d\times d})$ satisfying: for all $\mathbb{P}\in\mathcal{P}$, $u(t,X.)$ is a semimartingale and the following *functional Itô formula* holds:

$$du(t,X.) = \partial_t u(t,X.)dt + \partial_\omega u(t,X.)dX_t + \frac{1}{2}\partial^2_{\omega\omega}u(t,X.) : d\langle X\rangle_t, \quad \mathbb{P}\text{-a.s.} \tag{3.4}$$

For each $u \in C^{1,2}(\Theta;\mathbb{R})$, by [9] the path derivatives $\partial_t u, \partial_\omega u, \partial^2_{\omega\omega}u$ are unique. Moreover, we say $u = [u_1,\cdots,u_m]^\top \in C^{1,2}(\Lambda;\mathbb{R}^m)$ if each $u_i \in C^{1,2}(\Lambda;\mathbb{R})$ for $i = 1,\cdots,m$.

Denote $f = [f_1,\cdots,f_{d_2}]^\top$. The weak FBSDE (3.1) is closely related to the following system of PPDEs:

$$\begin{cases} \partial_t u_i + \frac{1}{2}\sigma\sigma^\top(t,\omega,u,\partial_\omega u) : \partial^2_{\omega\omega}u_i + f_i(t,\omega,u,\partial_\omega u) = 0; \\ u(T,\omega) = g(\omega), \end{cases} \qquad i = 1,\cdots,d_2. \tag{3.5}$$

3.3 Nonlinear Feynman-Kac formula

The following result is an extension of the four step scheme of Ma, Protter, & Yong [17].

Theorem 3.9. *Let Assumption 3.1 hold, and b,σ be uniformly Lipschitz continuous in $(\mathbf{x},y,z)$. Assume PPDE (3.5) has a classical solution $u \in C^{1,2}(\Lambda;\mathbb{R}^{d_2})$ such that $\partial_\omega u$, $\partial^2_{\omega\omega}u$ are bounded and $u,\partial_\omega u$ are uniformly Lipschitz continuous in ω. Then FBSDE (3.1) admits a strong solution and it holds that*

$$Y_t = u(t,X.), \quad Z_t = \partial_\omega u(t,X.). \tag{3.6}$$

Moreover, the solution is unique (in law) among all weak solutions.

Proof *Existence.* Set

$$\tilde{b}(t,\omega) := b(t,\omega,u(t,\omega),\partial_\omega u(t,\omega)), \quad \tilde{\sigma}(t,\omega) := \sigma(t,\omega,u(t,\omega),\partial_\omega u(t,\omega)).$$

Under our conditions, both $\tilde{b}(t,\omega)$ and $\tilde{\sigma}(t,\omega)$ are bounded and are uniformly Lipschitz continuous in ω. Thus, for any $x \in \mathbb{R}^{d_1}$, the following forward SDE

$$X_t = x + \int_0^t \tilde{b}(s, X.)ds + \int_0^t \tilde{\sigma}(s, X.)dB_s, \quad t \in [0,T], \tag{3.7}$$

has a (unique) strong solution. Define (Y, Z) by (3.6) and $N_t := 0$. By applying functional Itô's formula (3.4), we can easily verify (3.1), hence (X, Y, Z) is a strong solution of (3.1).

Uniqueness. For notational simplicity let's assume $d_2 = 1$. The multidimensional case can be proved similarly without any significant difficulty. Let $(B, \Theta, N, \mathbb{P})$ be an arbitrary weak solution of (3.1). We first claim that (3.6) holds. Indeed, denote

$$\tilde{Y}_t = u(t, X.), \quad \tilde{Z}_t = \partial_\omega u(t, X.), \quad \Delta Y_t := \tilde{Y}_t - Y_t, \quad \Delta Z_t := \tilde{Z}_t - Z_t.$$

Applying functional Itô formula (3.4) on $u(t, X.)$ and recalling (3.5), we have:

$$\begin{aligned}
& d\Delta Y_t = du(t, X.) + f(t, X., Y_t, Z_t)dt - Z_t dX_t + dN_t \\
& = \Big[\partial_t u(t, X.) + \frac{1}{2}\partial^2_{\omega\omega} u(t, X.) : \sigma\sigma^\top(t, X., Y_t, Z_t) + f(t, X., Y_t, Z_t)\Big]dt + \Delta Z_t dX_t + dN_t \\
& = -\Big[\frac{1}{2}\partial^2_{\omega\omega} u(t, X.) : \sigma\sigma^\top(t, X., \tilde{Y}_t, \tilde{Z}_t) + f(t, X., \tilde{Y}_t, \tilde{Z}_t)\Big]dt \\
& \qquad + \Big[\frac{1}{2}\partial^2_{\omega\omega} u(t, X.) : \sigma\sigma^\top(t, X., Y_t, Z_t) + f(t, X., Y_t, Z_t)\Big]dt + \Delta Z_t dX_t + dN_t \\
& = [\alpha_t \Delta Y_t + \beta_t \Delta Z_t]dt + \Delta Z_t \sigma(t, X., Y_t, Z_t)dB_t + dN_t,
\end{aligned}$$

where α, β are bounded. Note that $\Delta Y_T = 0$. Applying Itô formula on $|\Delta Y_t|^2$ and recalling Assumption 3.1 (iii) we have

$$\begin{aligned}
& \mathbb{E}\Big[|\Delta Y_t|^2 + c_0^2 \int_t^T |\Delta Z_s|^2 ds + tr(\langle N \rangle_T - \langle N \rangle_t)\Big] \\
\leq \; & \mathbb{E}\Big[|\Delta Y_t|^2 + \Big|\int_t^T \Delta Z_s \sigma(s, X., Y_s, Z_s)dB_s\Big|^2 + \mathrm{tr}\,(\langle N \rangle_T - \langle N \rangle_t)\Big] \\
= \; & \mathbb{E}\Big[\int_t^T 2\Delta Y_s[\alpha_s \Delta Y_s + \beta_s \Delta Z_s]ds\Big] \leq \mathbb{E}\Big[\int_t^T [C|\Delta Y_s|^2 + \frac{c_0^2}{2}|\Delta Z_s|^2]ds\Big]
\end{aligned}$$

Then by the standard BSDE arguments we have $|\Delta Y| = |\Delta Z| = 0$. This proves (3.6).

Now plug (3.6) into the forward SDE of (3.1), we see that X has to satisfy the SDE (3.7). By the uniqueness of (3.7) we see that X is unique, which, together with (3.6), implies further the uniqueness of Θ, hence that of N. ■

4 Wellposedness for Markovian weak FBSDEs

We now turn to weak solutions. We shall follow the approach in Ma, Zhang, & Zheng [21] and Ma & Zhang [20]. Our approach will rely heavily on viscosity solutions as well as the a priori estimates for the related PDE. We remark that all the results can be easily extended to path dependent case provided that the corresponding estimates can be established for PPDEs, which however are not available in the literature and are in general challenging. We thus restrict to Markovian case, and for the purpose of viscosity theory, we assume $d_2 = 1$. Moreover, by Proposition 3.8, we may assume without loss of generality that $b = 0$. That is, our objective of this section is the following weak FBSDE:

$$\begin{cases} X_t = x + \int_0^t \sigma(s, \Theta_s) dB_s; \\ Y_t = g(X_T) + \int_t^T f(s, \Theta_s) ds - \int_t^T Z_s dX_s + N_T - N_t, \end{cases} \quad \mathbb{P}_0\text{-a.s.} \tag{4.1}$$

In this case the PPDE (3.5) becomes a standard quasi-linear PDE:

$$\mathcal{L}u(t,x) := \partial_t u(t,x) + \frac{1}{2}\sigma\sigma^\top(t,x,u,\partial_x u) : \partial^2_{xx} u + f(t,x,u,\partial_x u) = 0, \; u(T,x) = g(x), \tag{4.2}$$

extending (1.7) to multidimensional case, and (3.6) becomes

$$Y_t = u(t, X_t), \quad Z_t = \partial_x u(t, X_t), \quad \mathbb{P}\text{-a.s.} \tag{4.3}$$

By Proposition 3.6, throughout this section, we shall assume

Assumption 4.1. *(i) $d := d_0 = d_1$, $d_2 = 1$, and σ, f, g are state dependent;*

(ii) $\sigma, f(t,x,0,0), g$ are bounded by C_0, and σ, f are continuous in t;

(iii) σ, f, g are uniformly Lipschitz continuous in (x,y,z) with Lipschitz constant L;

(iv) σ is symmetric and is uniformly nondegenerate: $\sigma \geq c_0 I_d$ for some $c_0 > 0$;

(v) Either $|\sigma(t,x,y,z_1) - \sigma(t,x,y,z_2)| \leq \frac{C_0}{1+|z_1|}|z_1 - z_2|$, or $d = 1$.

We emphasize again that, by Propositions 3.6 and 3.8, we may allow $d_0 \neq d_1$ and (4.1) may depend on $b(t, X_\cdot, Y_t, Z_t)$ as well. Throughout this section, we use a generic constant $C > 0$ which depends only on T and C_0, c_0, L, d in Assumption 4.1.

Under the above assumption, we have the following regularity results for the PDE (4.2). The arguments are mainly from Ladyzenskaja, Solonnikov & Uralceva [16], and we sketch a proof in Appendix.

Theorem 4.2. *Let Assumption 4.1 hold. Assume further that σ, f, g are smooth with bounded derivatives. Then*

(i) PDE (4.2) has a classical solution $u \in C^{1,2}_b([0,T] \times \mathbb{R}^d)$.

(ii) There exists a constant $\alpha > 0$, depending only on T and C_0, c_0, L, d in Assumption 4.1 , but not on the derivatives of σ, f, g, such that, for any $\delta > 0$,

$$\begin{gathered} |u| \leq C, \quad |\partial_x u| \leq C, \quad |u(t_1,x) - u(t_2,x)| \leq C|t_1 - t_2|^{\frac{1}{2}}; \\ |\partial_x u(t_1,x_1) - \partial_x u(t_2,x_2)| \leq C_\delta\big[|x_1 - x_2|^\alpha + |t_1 - t_2|^{\frac{\alpha}{2}}\big] \quad 0 \leq t_1 < t_2 \leq T - \delta. \end{gathered} \tag{4.4}$$

where C_δ may depend on δ as well.

(iii) There exists a constant C_g, which depends on the same parameters T, C_0, c_0, L, d, as well as $\|\partial_{xx} g\|_\infty$, such that $|\partial_{xx} u| \le C_g$.

4.1 Existence

Theorem 4.3. *Let Assumption 4.1 hold. Then FBSDE* (3.1) *admits a bounded semi-strong solution Θ, and* (4.3) *holds where u is a viscosity solution of PDE* (4.2)*.*

Proof Let (σ_n, f_n, g_n) be a smooth mollifier of (σ, f, g) such that they satisfy Assumption 4.1 uniformly. Applying Theorem 4.2, let u_n be the classical solution to PDE (4.2) with coefficients (σ_n, f_n, g_n), and then $\{u_n\}_{n\ge 1}$ satisfy (4.4) uniformly, uniformly in n. Applying the Arzela-Ascoli theorem, possibly along a subsequence, u_n converges to a function u such that u satisfies (4.4) and the convergence of $(u_n, \partial_x u_n)$ to $(u, \partial_x u)$ is uniform. In particular, by the stability of viscosity solutions we see that u is a viscosity solution of PDE (4.2).

Next, by Proposition 3.6 and Theorem 3.9 the martingale problem (4.1) with coefficients (σ_n, f_n, g_n) has a solution $(\mathbb{P}_n, Z^n)$ such that $Y_t = u_n(t, X_t), Z_t^n = \partial_x u_n(t, X_t)$, $\mathbb{P}_n$-a.s. By Zheng [34], possibly along a subsequence, we see that $\mathbb{P}_n$ converges to some $\mathbb{P}$ weakly. By the uniform convergence of $(u_n, \partial_x u_n)$, we have $Z_t^n \to Z_t$ uniformly, and (4.3) holds. Moreover, it follows from (4.4) that (Y, Z) are bounded. Finally, by the uniform convergence, it is straightforward to verify that $(\mathbb{P}, Z)$ solves the martingale problem (4.1) with coefficients (σ, f, g). ■

4.2 Nodal sets

For any $t \in [0, T]$, we first extend Definition 3.3 to interval $[t, T]$.

Definition 4.4. *Let $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$. We say $(B, \Theta, N, \mathbb{P})$ is a weak solution of FBSDE* (4.1) *at (t, x, y) if they are processes on $[t, T]$ satisfying the requirements in Definition 3.3 on $[t, T]$ and $\mathbb{P}(X_t = x) = \mathbb{P}(Y_t = y) = 1$. Define semi-strong solution, strong solution, and martingale problem at (t, x, y) in an obvious sense.*

We next define the nodal sets.

Definition 4.5. *(i) For $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$, let $\mathcal{O}(t, x, y)$ denote the space of weak solutions of* (4.1) *at (t, x, y).*

(ii) $O := \{(t, x, y) : \mathcal{O}(t, x, y) \ne \emptyset\}$.

By Theorem 4.3, $(t, x, u(t, x)) \in O$ for any $(t, x) \in [0, T] \times \mathbb{R}^d$, where u is the viscosity solution of PDE (4.2) in Theorem 4.3. We remark that a priori the measurability of O is not clear. Nevertheless, let $\overline{O}$ denote the closure of O, and define

$$\underline{u}(t, x) := \inf\{y : (t, x, y) \in \overline{O}\}, \quad \overline{u}(t, x) := \sup\{y : (t, x, y) \in \overline{O}\}. \tag{4.5}$$

Then $\underline{u}$ and $\overline{u}$ are Borel measurable and $\underline{u} \le u \le \overline{u}$.

Proposition 4.6. *Let Assumption 4.1 hold. Then*

(i) $\overline{u}$ and $\underline{u}$ are bounded;

(ii) $\overline{u}$ is upper semi-continuous and $\underline{u}$ is lower semi-continuous;

(iii) $\overline{u}(T, x) = \underline{u}(T, x) = g(x)$.

Proof (i) For any $(t, x, y) \in O$ with corresponding weak solution $(\Theta, N, \mathbb{P})$, we have

$$Y_s = g(X_T) + \int_s^T f(r, \Theta_r)dr - \int_t^T Z_r dX_r + N_T - N_t.$$

Since g and $f(t, x, , 0, 0)$ are bounded by C_0 and f is uniformly Lipschitz continuous in (y, z), it follows from standard BSDE arguments that

$$\mathbb{E}^{\mathbb{P}}\Big[\sup_{t \le s \le T}[|Y_s|^2 + |N_s|^2] + \int_t^T |Z_s|^2 ds\Big] \le C. \tag{4.6}$$

In particular, $|y| = |Y_t| \le C$. This implies $|\underline{u}|, |\overline{u}| \le C$.

Since $\overline{O}$ is closed, (ii) is a direct consequence of the definitions of $\overline{u}$, $\underline{u}$. To see (iii), let $(T, x, y) \in \overline{O}$. By definition there exist $(t_n, x_n, y_n) \in O$ such that $t_n \uparrow T$ and $(x_n, y_n) \to (x, y)$. Let $(B^n, \Theta^n, \mathbb{P}^n)$ be a weak solution at (t_n, x_n, y_n). Then

$$\begin{aligned}
|y_n - g(x_n)|^2 &= \Big|\mathbb{E}^{\mathbb{P}_n}\Big[g(X_T^n) + \int_{t_n}^T f(s, \Theta_s^n)ds\Big] - g(x_n)\Big|^2 \\
&\le C\mathbb{E}^{\mathbb{P}_n}\Big[|X_T^n - x_n|^2 + (T - t_n)\int_{t_n}^T [1 + |Y_s|^2 + |Z_s^n|^2]ds\Big] \\
&\le C\mathbb{E}^{\mathbb{P}_n}\Big[\int_{t_n}^T |\sigma(s, \Theta_s^n)|^2 ds\Big] + C(T - t_n) \le C(T - t_n),
\end{aligned}$$

thanks to (4.6). Send $n \to \infty$, we see that $y = g(x)$. This proves (iii). ■

We have the following result improving Theorem 4.3, which is not used in this paper but is nevertheless interesting in its own right.

Theorem 4.7. *Let Assumption 4.1 hold. Then $(t, x, y) \in O$ if and only if $y \in [\underline{u}(t, x), \overline{u}(t, x)]$. Moreover, for any $(t, x, y) \in O$, there exists a semi-strong solution $(B, \Theta, N, \mathbb{P})$ at (t, x, y) such that $|Z| \le C$.*

Proof It is clear that $(t,x,y)\in O$ implies $y\in[\underline{u}(t,x),\overline{u}(t,x)]$. Then it suffices to show that, for any $y\in[\underline{u}(t,x),\overline{u}(t,x)]$, there exists a weak solution at (t,x,y) such that Z is bounded. We proceed in two steps.

Step 1. For any $n\ge 1$, let σ_n, f_n, g_n be smooth mollifiers of σ, f, g such that

$$|\sigma_n-\sigma|\le\varepsilon_n,\quad |f_n-f|\le\frac{1}{n},\quad |g_n-g|\le\frac{1}{n}, \tag{4.7}$$

for some small $\varepsilon_n>0$ which will be specified later. Denote

$$\overline{f}_n := f_n+\frac{2}{n},\quad \underline{f}_n := f_n-\frac{2}{n},\quad \overline{g}_n := g_n+\frac{1}{n},\quad \underline{g}_n := g_n-\frac{1}{n}.$$

By Theorem 4.2, the PDE (4.2) with coefficients $(\sigma_n,\overline{f}_n,\overline{g}_n)$ (resp. $(\sigma_n,\underline{f}_n,\underline{g}_n)$) has a classical solution $\overline{u}_n$ (resp. $\underline{u}_n$). We claim that, for any $(t,x,y)\in O$ and any n,

$$\underline{u}_n(t,x)\le y\le\overline{u}_n(t,x). \tag{4.8}$$

Without loss of generality we will prove only the right inequality at $t=0$. We shall follow similar arguments as in Theorem 3.9. Let $(B,\Theta,N,\mathbb{P})$ be a weak solution to FBSDE (4.1) at $(0,x,y)$ with coefficients (σ,f,g). Fix n and denote

$$\tilde{Y}_t := \overline{u}_n(t,X_t),\ \tilde{Z}_t := \partial_x\overline{u}_n(t,X_t),\ \tilde{\Theta} := (X,\tilde{Y},\tilde{Z}),\ \Delta Y_t := \tilde{Y}_t-Y_t,\ \Delta Z_t := \tilde{Z}_t-Z_t.$$

Apply Itô formula, we have

$$\begin{aligned} d\Delta Y_t &= \Big[\partial_t\overline{u}_n+\frac{1}{2}\partial^2_{xx}\overline{u}_n:\sigma^2(t,\Theta_t)+f(t,\Theta_t)\Big]dt+\Delta Z_t dX_t+dN_t\\ &= \Big[\frac{1}{2}\partial^2_{xx}\overline{u}_n:[\sigma^2(t,\Theta_t)-\sigma_n^2(t,\tilde{\Theta}_t)]+[f(t,\Theta_t)-\overline{f}_n(t,\tilde{\Theta}_t)]\Big]dt+\Delta Z_t dX_t+dN_t. \end{aligned}$$

By Theorem 4.2 (iii), there exists a constant C_n, which is independent of ε_n, such that $|\partial^2_{xx}\overline{u}_n|\le C_n$. Note that $\overline{f}_n-f=f_n+\frac{2}{n}-f\ge\frac{1}{n}$ and $|\sigma|\le C_0$. Then, for $\varepsilon_n\le\frac{1}{nC_0C_n}$, we have

$$\begin{aligned} d\Delta Y_t &\le \Big[\frac{1}{2}\partial^2_{xx}\overline{u}_n:[\sigma^2(t,\Theta_t)-\sigma^2(t,\tilde{\Theta}_t)]+[f(t,\Theta_t)-f(t,\tilde{\Theta}_t)]\Big]dt+\Delta Z_t dX_t+dN_t\\ &= \Big[\alpha^n_t\Delta Y_t+\beta^n_t\Delta Z_t\Big]dt+\Delta Z_t dX_t+dN_t, \end{aligned}$$

where $|\alpha^n|,|\beta^n|\le C_n$. Note further that $\Delta Y_T=\overline{g}_n(X_T)-g(X_T)=g_n(X_T)+\frac{1}{n}-g(X_T)\ge 0$. It is clear that $\Delta Y_0\ge 0$. This implies $0\le\tilde{Y}_0-Y_0=\overline{u}_n(0,x)-y$, proving (4.8).

Step 2. Let $y\in[\underline{u}(t,x),\overline{u}(t,x)]$. There exist $(\underline{t}_m,\underline{x}_m,\underline{y}_m)\in O$ and $(\overline{t}_m,\overline{x}_m,\overline{y}_m)\in O$ such that $(\underline{t}_m,\underline{x}_m,\underline{y}_m)\to(t,x,\underline{u}(t,x))$ and $(\overline{t}_m,\overline{x}_m,\overline{y}_m)\to(t,x,\overline{u}(t,x))$. Then, by (4.8),

$$\underline{u}_n(\underline{t}_m,\underline{x}_m)\le\underline{y}_m,\quad \overline{y}_m\le\overline{u}_n(\overline{t}_m,\overline{x}_m),\quad \text{for all } m,n.$$

Send $m \to \infty$, we obtain

$$\underline{u}_n(t,x) \le \underline{u}(t,x) \le y \le \overline{u}(t,x) \le \overline{u}_n(t,x), \quad \text{for all } n. \tag{4.9}$$

For any $n \ge 1$ and $\alpha \in [0,1]$, denote $\varphi_n^\alpha := \alpha \overline{\varphi}_n + [1-\alpha]\underline{\varphi}_n$ for $\varphi = f, g$, and let u_n^α be the classical solution of PDE (4.2) with coefficients $(\sigma_n, f_n^\alpha, g_n^\alpha)$. By the arguments in Theorem 4.2, it is clear that the mapping $\alpha \mapsto u_n^\alpha(0,x)$ is continuous. Since $u_n^0(t,x) = \underline{u}_n(t,x) \le y \le \overline{u}_n(t,x) = u_n^1(t,x)$, there exists $\alpha_n \in [0,1]$ such that $u_n^{\alpha_n}(t,x) = y$. For each $n \ge 1$, by Proposition 3.6 and Theorem 3.9 the martingale problem (4.1) at (t,x,y) with coefficients $(\sigma_n, f_n^{\alpha_n}, g_n^{\alpha_n})$ has a solution $(\mathbb{P}^n, Z^n)$ such that $Y_s = u_n^{\alpha_n}(s, X_s), Z_s^n = \partial_x u_n^{\alpha_n}(s, X_s)$, $t \le s \le T$, $\mathbb{P}^n$-a.s. Now following the arguments in Theorem 4.3 we see that, possibly following a subsequence, $\mathbb{P}^n \to \mathbb{P}$, $Z^n \to Z$, $u_n^{\alpha_n} \to u$, where $(\mathbb{P}, Z)$ is a solution to the martingale problem (4.1) at (t,x,y) with coefficients (σ, f, g) and u is a viscosity solution to PDE (4.2) with coefficients (σ, f, g). It is clear that $|Z_s| = |\partial_x u(s, X_s)| \le C$, $\mathbb{P}$-a.s. ■

4.3 Uniqueness

Theorem 4.8. *Let Assumption 4.1 hold. Then $\overline{u}$ (resp. $\underline{u}$) is a viscosity subsolution (resp. supersolution) of PDE* (4.2).

Proof We shall prove the result only for $\overline{u}$. The result for $\underline{u}$ can be proved similarly.

Fix $(t_0, x_0) \in [0,T) \times \mathbb{R}^d$ and denote $y_0 := \overline{u}(t_0, x_0)$. Let $\varphi \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$ be a test function at (t_0, x_0), namely

$$[\varphi - \overline{u}](t_0, x_0) = 0 = \inf_{(t,x) \in [0,T] \times \mathbb{R}^d} [\varphi - \overline{u}](t,x).$$

Let $(t_n, x_n, y_n) \in O$ such that $(t_n, x_n, y_n) \to (t_0, x_0, y_0)$, and $(\mathbb{P}^n, Z^n)$ a weak solution to the martingale problem (4.1) at (t_n, x_n, y_n). Define N^n as in (3.2). By using regular conditional probability distribution, it is clear that $(t, X_t, Y_t) \in O$, $\mathbb{P}^n$-a.s. for $t_n \le t \le T$. Then, by the definition of $\overline{u}$, we have $Y_t \le \overline{u}(t, X_t) \le \varphi(t, X_t)$.

Now denote

$$\Delta Y_t := \varphi(t, X_t) - Y_t \ge 0, \quad \Delta Z_t^n := \partial_x \varphi(t, X_t) - Z_t^n,$$
$$\Theta_t^n := (X_t, Y_t, Z_t^n), \quad \tilde{\Theta}_t := (X_t, \varphi(t, X_t), \partial_x \varphi(t, X_t)).$$

Applying Itô formula we have, under $\mathbb{P}^n$,

$$
\begin{aligned}
d\Delta Y_t &= \Big[\partial_t\varphi(t,X_t) + \frac{1}{2}\partial^2_{xx}\varphi(t,X_t) : \sigma^2(t,\Theta^n_t)dt + f(t,\Theta^n_t)\Big]dt + \Delta Z^n_t dX_t + dN^n_t \\
&= \Big[\mathcal{L}\varphi(t,X_t) + \frac{1}{2}\partial^2_{xx}\varphi(t,X_t) : [\sigma^2(t,\Theta^n_t) - \sigma^2(t,\tilde{\Theta}_t)] + [f(t,\Theta^n_t) - f(t,\tilde{\Theta}_t)]\Big]dt \\
&\quad + \Delta Z^n_t dX_t + dN^n_t \\
&= \Big[\mathcal{L}\varphi(t,X_t) - \alpha^n_t\Delta Y_t - \Delta Z^n_t\sigma(t,\Theta^n_t)\beta^n_t\Big]dt + \Delta Z^n_t dX_t + dN^n_t,
\end{aligned}
$$

where $|\alpha^n|, |\beta^n| \le C$. Denote

$$
\Gamma^n_t := \exp\Big(\int_{t_n}^t \beta^n_s \cdot \sigma^{-1}(s,\Theta^n_s)dX_s + \int_{t_n}^t [\alpha^n_s - \frac{1}{2}|\beta^n_s|^2]ds\Big).
$$

Then

$$
d[\Gamma^n_t\Delta Y_t] = \Gamma^n_t\mathcal{L}\varphi(t,X_t)dt + \Gamma^n_t[\Delta Z^n_t + \Delta Y_t[\beta^n_s]^\top\sigma^{-1}(s,\Theta^n_s)]dX_t + \Gamma^n_t dN^n_t.
$$

Thus, for any $\delta > 0$ small,

$$
\begin{aligned}
0 &\le \mathbb{E}^{\mathbb{P}_n}[\Gamma^n_{t+\delta}\Delta Y_{t_n+\delta}] = \mathbb{E}^{\mathbb{P}_n}\Big[\Gamma^n_{t_n}\Delta Y_{t_n} + \int_{t_n}^{t_n+\delta}\Gamma^n_t\mathcal{L}\varphi(t,X_t)dt\Big] \\
&= \varphi(t_n,x_n) - y_n + \mathcal{L}\varphi(t_n,x_n)\delta + \mathbb{E}^{\mathbb{P}_n}\Big[\int_{t_n}^{t_n+\delta}[\Gamma^n_t\mathcal{L}\varphi(t,X_t) - \Gamma^n_{t_n}\mathcal{L}\varphi(t_n,X_{t_n}]dt\Big].
\end{aligned}
$$

Note that $\mathcal{L}\varphi$ is uniformly continuous, and since σ is bounded, one can easily show that

$$
\mathbb{E}^{\mathbb{P}_n}\Big[\big|\Gamma^n_t\mathcal{L}\varphi(t,X_t) - \Gamma^n_{t_n}\mathcal{L}\varphi(t_n,X_{t_n}\big|\Big] \le \rho(\delta), \quad t_n \le t \le t_n+\delta,
$$

for some modulus of continuity function ρ. Then

$$
0 \le \varphi(t_n,x_n) - y_n + \mathcal{L}\varphi(t_n,x_n)\delta + \delta\rho(\delta).
$$

Send $n \to \infty$, we have

$$
0 \le \mathcal{L}\varphi(t_0,x_0)\delta + \delta\rho(\delta).
$$

Divide both sides by δ and then send $\delta \to 0$, we obtain $\mathcal{L}\varphi(t_0,x_0) \ge 0$. ■

We remark that, in the case that σ is independent of z, [20] and [21] established similar results without requiring the uniform Lipschitz continuity of the coefficients, and thus the arguments there are more involved.

Our final result relies on the comparison principle for viscosity solutions of PDEs, for which we refer to the classical reference Crandall, Ishii, & Lions [6]. We say a PDE satisfies the comparison principle for viscosity solutions if: *for any upper semi-continuous viscosity subsolution u_1 and any lower semi-continuous viscosity supersolution u_2 with $u_1(T,\cdot) \le u_2(T,\cdot)$, we have $u_1 \le u_2$.*

Theorem 4.9. *Let Assumption 4.1 hold. Assume further that the comparison principle for the viscosity solutions of PDE* (4.2) *holds true. Then the weak solution to FBSDE* (4.1) *is unique (in law).*

Proof First, by the comparison principle, it follows from Theorem 4.8 that $\overline{u} = \underline{u} = u$, where u is the unique viscosity solution of the PDE (4.2) satisfying (4.4). Now let $(B, \Theta, N, \mathbb{P})$ be an arbitrary weak solution of FBSDE (4.1). Since $(t, X_t, Y_t) \in O$, $\mathbb{P}$-a.s., then $Y_t = u(t, X_t)$, $\mathbb{P}$-a.s.

Next, for any $\delta > 0$, $0 < t \leq T - \delta$, and any partition $0 = t_0 < \cdots < t_n = t$ with $t_{i+1} - t_i = h := \frac{t}{n}$, by (4.4) we have

$$
\begin{aligned}
& \Big|\sum_{i=0}^{n-1}[Y_{t_{i+1}} - Y_{t_i}][X_{t_{i+1}} - X_{t_i}] - \int_0^t \partial_x u(s, X_s) d\langle X\rangle_s\Big| \\
= & \Big|\sum_{i=0}^{n-1}[u(t_{i+1}, X_{t_{i+1}}) - u(t_i, X_{t_i})][X_{t_{i+1}} - X_{t_i}] - \int_0^t \partial_x u(s, X_s) d\langle X\rangle_s\Big| \\
\leq & \Big|\sum_{i=0}^{n-1}[u(t_{i+1}, X_{t_i}) - u(t_i, X_{t_i})][X_{t_{i+1}} - X_{t_i}]\Big| \\
& + \Big|\sum_{i=0}^{n-1}\Big[[u(t_{i+1}, X_{t_{i+1}}) - u(t_{i+1}, X_{t_i})][X_{t_{i+1}} - X_{t_i}] - \partial_x u(t_{i+1}, X_{t_i})[\langle X\rangle_{t_{i+1}} - \langle X\rangle_{t_i}]\Big]\Big| \\
& + \sum_{i=0}^{n-1}\Big|\int_{t_i}^{t_{i+1}}[\partial_x u(s, X_s) - \partial_x u(t_{i+1}, X_{t_i})] d\langle X\rangle_s\Big| \\
\leq & \Big|\sum_{i=0}^{n-1}[u(t_{i+1}, X_{t_i}) - u(t_i, X_{t_i})][X_{t_{i+1}} - X_{t_i}]\Big| \\
& + \Big|\sum_{i=0}^{n-1}\partial_x u(t_{i+1}, X_{t_i})\big[[X_{t_{i+1}} - X_{t_i}]^\top[X_{t_{i+1}} - X_{t_i}] - [\langle X\rangle_{t_{i+1}} - \langle X\rangle_{t_i}]\big]\Big| \\
& + C_\delta \sum_{i=0}^{n-1}|X_{t_{i+1}} - X_{t_i}|^{2+\alpha} + C_\delta \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}[h^{\frac{\alpha}{2}} + |X_{t_{i+1}} - X_s|^\alpha] ds.
\end{aligned}
$$

Since $X_t - X_s = \int_s^t \sigma(r, \Theta_r) dB_r$ and σ is bounded, one can easily show that

$$
\mathbb{E}^{\mathbb{P}}[|X_t - X_s|^p] \leq C_p|t - s|^{\frac{p}{2}}.
$$

Moreover, by the martingale property of X, we have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\Big[\Big|\sum_{i=0}^{n-1}[u(t_{i+1},X_{t_i})-u(t_i,X_{t_i})][X_{t_{i+1}}-X_{t_i}]\Big|\Big]^2\\
= &\mathbb{E}^{\mathbb{P}}\Big[\sum_{i=0}^{n-1}\Big|[u(t_{i+1},X_{t_i})-u(t_i,X_{t_i})][X_{t_{i+1}}-X_{t_i}]\Big|^2\Big]\\
\le &Ch\mathbb{E}^{\mathbb{P}}\Big[\sum_{i=0}^{n-1}|X_{t_{i+1}}-X_{t_i}|^2\Big]\le Ch\sum_{i=0}^{n-1}h=Ch;
\end{aligned}$$

and, applying Itô formula,

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\Big[\Big|\sum_{i=0}^{n-1}\partial_x u(t_{i+1},X_{t_i})\big[[X_{t_{i+1}}-X_{t_i}]^\top[X_{t_{i+1}}-X_{t_i}]-[\langle X\rangle_{t_{i+1}}-\langle X\rangle_{t_i}]\big]\Big|^2\Big]\\
= &\mathbb{E}^{\mathbb{P}}\Big[\Big|\sum_{i=0}^{n-1}\partial_x u(t_{i+1},X_{t_i})\int_{t_i}^{t_{i+1}}[X_s-X_{t_i}]^\top dX_s\Big|^2\Big]\\
= &\mathbb{E}^{\mathbb{P}}\Big[\sum_{i=0}^{n-1}\Big|\partial_x u(t_{i+1},X_{t_i})\int_{t_i}^{t_{i+1}}[X_s-X_{t_i}]^\top dX_s\Big|^2\Big]\\
\le &C\mathbb{E}^{\mathbb{P}}\Big[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}|X_s-X_{t_i}|^2ds\Big]\le\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}[s-t_i]ds\le Ch.
\end{aligned}$$

Then we have

$$\mathbb{E}^{\mathbb{P}}\Big[\Big|\sum_{i=0}^{n-1}[Y_{t_{i+1}}-Y_{t_i}][X_{t_{i+1}}-X_{t_i}]-\int_0^t\partial_x u(s,X_s)d\langle X\rangle_s\Big|^2\Big]\le Ch+C_\delta h^\alpha\le C_\delta h^\alpha.$$

Send $n\to\infty$ and thus $h\to 0$, note that

$$\sum_{i=0}^{n-1}[Y_{t_{i+1}}-Y_{t_i}][X_{t_{i+1}}-X_{t_i}]\to\langle Y,X\rangle_t=\int_0^t Z_s d\langle X\rangle_s,\quad \text{in } \mathbb{L}^2(\mathbb{P}),$$

then we have

$$\int_0^t Z_s d\langle X\rangle_s=\int_0^t\partial_x u(s,X_s)d\langle X\rangle_s,\quad \mathbb{P}\text{-a.s.},\quad 0\le t\le T-\delta.$$

Since σ is nondegenerate and t and δ are arbitrary, we obtain

$$Z_t=\partial_x u(t,X_t),\quad dt\times d\mathbb{P}\text{-a.s. on } [0,T)\times\Omega.$$

That is, (4.3) holds.

Now similar to the existence part of Theorem 3.9, denote

$$\tilde\sigma(t,x):=\sigma(t,x,u(t,x),\partial_x u(t,x)).$$

Then $\tilde{\sigma}$ is Hölder continuous and $(B, X, \mathbb{P})$ satisfies the SDE:

$$X_t = x + \int_0^t \tilde{\sigma}(s, X_s) dB_s, \quad \mathbb{P}\text{-a.s.}$$

By Stroock & Varadahn [28], the above SDE has a unique (in law) weak solution. This, together with (4.3), implies the uniqueness (in law) of $(B, \Theta, \mathbb{P})$. Finally, by (3.2), the joint law with N is also unique. ■

Remark 4.10. An alternative approach to prove the uniqueness is to consider the stochastic target problem, as in Soner and Touzi [26]. That is, in the spirit of (2.8), define

$$\overline{u}(t,x) := \inf\Big\{ y : \exists Z \text{ such that } Y_T^{t,x,y,Z} \geq g(X_T^{t,x,Z}),\ \mathbb{P}_0\text{-a.s.}\Big\}, \quad \text{where}$$
$$X_s^{t,x,y,Z} = x + \int_t^s \sigma(r, X_r^{t,x,y,Z}) dB_r,$$
$$Y_s^{t,x,y,Z} = y - \int_t^s f(r, X_r^{t,x,y,Z}, Y_r^{t,x,y,Z}, Z_r) dr + \int_t^s Z_r dX_r^{t,x,y,Z},$$

and define $\underline{u}$ similarly. The idea is to prove that $\overline{u}$ and $\underline{u}$ are viscosity solutions of the PDE. However, there are technical difficulties in establishing the regularity and the dynamic programming principle for these functions. We shall leave this possible approach to future research. ■

5 Appendix

5.1 Some counterexamples

In this subsection we provide two counterexamples related to the control problems in Section 2.3. In particular, they will show that the stochastic control problems in weak formulation have optimal controls, while the corresponding problems in strong formulation do not have optimal control. In the first example, we also show that the associated weak FBSDE has a weak solution, but no strong solution.

5.1.1 The case with drift control

In this case we shall consider an example with path dependence. We note that all the heuristic analysis in Section 2.3 can be easily extended to the path dependent case. We first recall a result due to Tsirel'son [29].

Lemma 5.1. *Let $t_n > 0$, $n \ge 1$, be strictly decreasing with $t_0 = T$ and $t_n \downarrow 0$, and $\theta(x) := x - [x]$ where $[x]$ is the largest integer in $(-\infty, x]$. Define the non-curtailing functional K:*

$$K(t, \mathbf{x}) := \theta(\frac{\mathbf{x}(t_n) - \mathbf{x}(t_{n+1})}{t_n - t_{n+1}}), \text{ for } t \in [t_n, t_{n-1}), \ \mathbf{x} \in C([0,T]). \tag{5.10}$$

Then the following path dependent SDE has no strong solution:

$$X_t = \int_0^t K(s, X_\cdot)ds + B_t. \tag{5.11}$$

We remark that K is bounded and thus SDE (5.11) has a unique (in law) weak solution, following the standard Girsanov Theorem. We also note that the above K is discontinuous. When K is state dependent, namely $K = K(t, X_t)$, the SDE could have a strong solution even when K is discontinuous, see Cherny & Engelbert [4] and Halidias & Kloeden [12] for some positive results.

Our example considers the following setting, with f depending on the paths of X:

$$b(t,\alpha) := \alpha, \quad \sigma := 1, \quad f(t, \mathbf{x}, \alpha) := -\frac{1}{2}|\alpha - K(t, \mathbf{x})|^2, \quad g := 0. \tag{5.12}$$

Example 5.2. *Let K be defined in* (5.10)*, and $\mathcal{A} := \mathbb{L}^2(\mathbb{F}^B, \mathbb{P}_0)$.*

(i) The optimization problem in weak formulation has an optimal control $\alpha^*_t := K(t, X_\cdot)$:

$$\bar{V}_0 := \sup_{\alpha\in\mathcal{A}} \bar{V}_0^\alpha := \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\Big[-\frac{1}{2}\int_0^T |\alpha_t - K(t, X_\cdot)|^2 dt\Big], \tag{5.13}$$

$$\text{where} \quad X_t := B_t, \quad d\mathbb{P}^\alpha := M_T^\alpha d\mathbb{P}_0 := \exp\Big(\int_0^T \alpha_t dB_t - \frac{1}{2}\int_0^T |\alpha_s|^2 dt\Big) d\mathbb{P}_0.$$

(ii) The optimization problem in strong formulation has no optimal control:

$$V_0 := \sup_{\alpha\in\mathcal{A}} V_0^\alpha, \quad \text{where} \tag{5.14}$$

$$X_t^\alpha := \int_0^t \alpha_s ds + B_t, \quad V_0^\alpha = \mathbb{E}^{\mathbb{P}_0}\Big[-\frac{1}{2}\int_0^T |\alpha_t - K(t, X^\alpha_\cdot)|^2 dt\Big].$$

Proof (i) Since $\bar{V}_0^\alpha \le 0$, it is obvious that $\bar{V}_0 \le 0$. Moreover, it is clear that $\bar{V}^{\alpha^*} = 0$ for $\alpha^*_t := K(t, X_\cdot) = K(t, B_\cdot)$, then $\bar{V}_0 = 0$ with optimal control α^*.

(ii) For each n, denote $t_i := \frac{iT}{n}$, $i = 0, \cdots, n$, and $\alpha^n_t := \sum_{i=1}^{n-1} \frac{T}{n}\int_{t_{i-1}}^{t_i} K(s, B_\cdot)ds \mathbf{1}_{[t_i, t_{i+1})}$. Recall (5.13) and note that

$$\bar{V}_0^\alpha = \mathbb{E}^{\mathbb{P}_0}\Big[M_T^\alpha\big[-\frac{1}{2}\int_0^T |\alpha_t - K(t, B_\cdot)|^2 dt\big]\Big].$$

It is clear that

$$\lim_{n\to\infty} \bar{V}_0^{\alpha^n} = \bar{V}_0^{\alpha^*} = 0. \tag{5.15}$$

Since α^n is piecewise constant, then $\mathbb{F}^B = \mathbb{F}^{B^{\alpha^n}}$, and thus there exists a piecewise constant process $\tilde{\alpha}^n$ such that $\alpha^n_t(B.) = \tilde{\alpha}^n_t(B^{\alpha^n}_{\cdot})$. That is,

$$B_t = \int_0^t \tilde{\alpha}^n_s(B^{\alpha^n}_{\cdot})ds + B^{\alpha^n}_t, \quad \mathbb{P}^{\alpha^n}\text{-a.s.}$$

Therefore, the $\mathbb{P}_0$-distribution of $(B, X^{\tilde{\alpha}^n}, \tilde{\alpha}^n(B))$ coincides with the $\mathbb{P}^{\alpha^n}$-distribution of $(B^{\alpha^n}, B, \alpha^n(B))$. This implies that $V_0^{\tilde{\alpha}^n} = \bar{V}_0^{\alpha^n}$. Then by (5.15) we see that $V_0 \geq \lim_{n\to\infty} V_0^{\tilde{\alpha}^n} = 0$. On the other hand, it is obvious that $V_0 \leq 0$. Then $V_0 = 0$.

Now if (5.14) has an optimal control $\tilde{\alpha}^*$, then $V_0^{\tilde{\alpha}^*} = 0$ and thus $\tilde{\alpha}^*_t = K(t, X^{\tilde{\alpha}^*}_{\cdot})$, $\mathbb{P}_0$-a.s. Thus $X^* := X^{\tilde{\alpha}^*}$ satisfies SDE (5.11). Since by definition $\tilde{\alpha}^*$ is $\mathbb{F}^B$-progressively measurable, we see that X^* is also $\mathbb{F}^B$-progressively measurable, and hence X^* is a strong solution of SDE (5.11). This contradicts with Lemma 5.1. ■

Remark 5.3. By extending the arguments to this case, one can (formally) show that the weak FBSDE (2.28) and the equivalent one (2.32) becomes

$$\begin{cases} X_t = \displaystyle\int_0^t [Z_s + K(s, X.)]ds + B_t; \\ Y_t = \displaystyle\int_t^T [\frac{1}{2}|Z_s|^2 + K(s, X.)Z_s]ds - \int_t^T Z_s dX_s. \end{cases} \tag{5.16}$$

This FBSDE has a weak solution: $Y = Z = 0$ and X is the weak solution to SDE (5.11). However, it does not have a strong solution such that $\int_0^t Z_s dB_s$ is a BMO martingale. We refer to Zhang [33] Chapter 7 for BMO martingales. Indeed, if there is such a solution, then by (5.16) we immediately have

$$Y_t = -\frac{1}{2}\int_t^T |Z_s|^2 ds - \int_t^T Z_s dB_s.$$

This implies that $Y = Z = 0$. Then X has to be a strong solution of SDE (5.11), contradicting with Lemma 5.1. ■

5.1.2 The case with diffusion control

We first recall a result due to Barlow [3]. Recall the function $\theta(x)$ in Lemma 5.1.

Lemma 5.4. *Let $\frac{\sqrt{2}}{2} < \lambda < 1$ and define*

$$\sigma_0(x) := 1 + \sum_{n=0}^{\infty} \lambda^n \eta\big(\theta(2^n x)\big), \quad \textit{where} \quad \eta(x) := x\mathbf{1}_{[0,\frac{1}{2})}(x) + (1-x)\mathbf{1}_{[\frac{1}{2},1)}(x). \tag{5.17}$$

Then the following SDE has a unique weak solution but no strong solution:

$$X_t = \int_0^t \sigma_0(X_s)dB_s, \quad \mathbb{P}_0\text{-a.s.} \tag{5.18}$$

Proof We first note that, although θ is discontinuous at integers, $\eta\circ\theta$ is actually Lipschitz continuous and periodic. Then σ_0 is uniformly continuous, and clearly $\sigma_0 \geq 1$. Thus it follows from Stroock & Varadahn [28] that (5.18) has a unique weak solution.

On the other hand, one may verify that σ_0 satisfies the hypotheses in [3] Theorem 1.3 with $\alpha = \beta = -\ln\lambda/\ln 2$. Then we see that (5.18) has no strong solution. ■

The next example considers the following setting with diffusion control:

$$b := 0,\ \sigma(t,\alpha) := \alpha,\ f(t,x,\alpha) := -\frac{1}{4}[|\alpha|^4 + |\sigma_0(x)|^4],\ g(x) := \int_0^x \int_0^\lambda [\sigma_0(r)]^2 dr d\lambda. \tag{5.19}$$

Note that in this case we need the weak formulation for diffusion control problems. We refer to Zhang [33] Chapter 9 for details.

Example 5.5. *Consider* (5.17) *and* (5.19) *with* $\lambda = \frac{3}{4}$*, and let the control set* $A := [1,2]$*.*

(i) The optimization problem in weak formulation has optimal control $\alpha_t^* := \sigma_0(X_t)$*:*

$$\bar V_0 := \sup_{\alpha\in\mathcal{A}} \bar V_0^\alpha := \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\Big[g(X_T) + \int_0^T f(t,X_t,\alpha_t)dt\Big], \tag{5.20}$$
$$\text{where } \mathbb{P}^\alpha \text{ is a weak solution of SDE: } X_t = \int_0^t \alpha_s(X_\cdot)dB_s.$$

(i) The optimization problem in strong formulation has no optimal control:

$$V_0 := \sup_{\alpha\in\mathcal{A}} V_0^\alpha := \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{\mathbb{P}_0}\Big[g(X_T^\alpha) + \int_0^T f(t,X_t^\alpha,\alpha_t)dt\Big]. \tag{5.21}$$
$$\text{where} \quad X_t^\alpha := \int_0^t \alpha_s(B_\cdot)dB_s.$$

Proof (i) By standard literature, $\bar V_0 = u(0,0)$, where u satisfies the HJB equation:

$$\partial_t u + \sup_{\alpha\in[1,2]} \Big[\frac{1}{2}\alpha^2\partial_{xx}^2 u - \frac{1}{4}\alpha^4\Big] - \frac{1}{4}|\sigma_0(x)|^4 = 0, \quad u(T,x) = g(x). \tag{5.22}$$

Note that the above PDE has a classical solution $u(t,x) = g(x)$. Then $\bar V_0 = g(0) = 0$. On the other hand, let $\alpha_t^*(X_\cdot) := \sigma_0(X_t)$ and $\mathbb{P}^* := \mathbb{P}^{\alpha^*}$ be the (unique) weak solution of SDE (5.18). Denote $Y_t := g(X_t) + \int_0^t f(s,X_s,\alpha_s^*)ds$ and note that $g''(x) = |\sigma_0(x)|^2$. Then applying Itô formula we have

$$\begin{aligned} dY_t &= \Big[\frac{1}{2}g''(X_t)|\sigma_0(X_t)|^2 + f(t,X_t,\alpha_t^*)\Big]dt + g'(X_t)\sigma_0(X_t)dB_t \\ &= \Big[\frac{1}{2}|\sigma_0(X_t)|^4 - \frac{1}{4}[|\alpha_t^*|^4 + |\sigma_0(X_t)|^4]\Big]dt + g'(X_t)\sigma_0(X_t)dB_t = g'(X_t)\sigma_0(X_t)dB_t. \end{aligned}$$

This is a $\mathbb{P}^*$-martingale. Then $\bar V_0 = Y_0 = \mathbb{E}^{\mathbb{P}^*}[Y_T] = \bar V_T^{\alpha^*}$. That is, α^* is an optimal control.

(ii) By standard literature we also have $V_0 = u(0,0) = g(0) = 0$. Assume by contradiction that (5.21) has an optimal control $\alpha^*(B_\cdot)$. Note that the optimal control for the Hamiltonian in (5.22) is $\sqrt{\partial^2_{xx}u(t,x)} = \sigma_0(x)$, then we must have $\alpha^*_t(B_\cdot) = \sigma_0(X^{\alpha^*}_t)$, $\mathbb{P}_0$-a.s.. Thus $X^* := X^{\alpha^*}$ satisfies SDE (5.18). Since by definition α^* is $\mathbb{F}^B$-progressively measurable, we see that X^* is also $\mathbb{F}^B$-progressively measurable, and hence X^* is a strong solution of SDE (5.18), contradicting with Lemma 5.4. ■

Remark 5.6. In this example, since σ_0 is not differentiable in x, then neither is f. Consequently, the stochastic maximum principle in Section 2.3.1 does not work. ■

5.2 Proof of Theorem 4.2

Following the arguments in Ladyzenskaja, Solonnikov & Uralceva [16], we prove the theorem in four steps.

Step 1. First, for $n \geq 1$, denote

$$\begin{gathered}O_n := \{x \in \mathbb{R}^d : |x| < n\}, \quad \partial O_n := \{x \in \mathbb{R}^d : |x| = n\},\\ Q_n := [0,T) \times O_n, \quad \partial Q_n := (\{T\} \times O_n) \cup ([0,T] \times \partial O_n),\\ g_n(t,x) := g(x)I_n(x) + [T-t]f(T,x,0,0) \quad \text{and thus} \quad \mathcal{L}g_n(T,x) = 0 \text{ for } x \in \partial O_n,\end{gathered}$$

where $I_n \in C_0^\infty(\mathbb{R}^d)$ satisfying $I_n(x) = 1$ for $|x| \leq n-1$ and $I_n(x) = 0$ for $|x| \geq n$. Next, for $k \geq 1$, define

$$\sigma_k(t,x,y,z) := [1 - I_k(z)]I_d + I_k(z)\sigma(t,x,y,z), \quad f_k(t,x,y,z) := I_k(z)f(t,x,y,z).$$

Now for $k, n \geq 1$, consider the following PDE on Q_n:

$$\begin{gathered}\partial_t u^k_n(t,x) + \frac{1}{2}\sigma^2_k(t,x,u^k_n,\partial_x u^k_n) : \partial^2_{xx} u^k_n + f_k(t,x,u^k_n,\partial_x u^k_n) = 0, \ (t,x) \in Q_n;\\ u^k_n(t,x) = g_n(t,x), \quad (t,x) \in \partial Q_n.\end{gathered} \tag{5.23}$$

One can check that (5.23) satisfies all the conditions in [16] Chapter VI, Theorem 4.1, with $m = 2$, $\varepsilon = 0$, $P(|z|) = 0$ for $|z| \geq k$, and $\mu_1 = \mu_1(k)$ depending on k in (4.6)-(4.10) there, and thus (5.23) has a classical solution $u^k_n \in C_b^{1+\frac{\beta}{2},2+\beta}(Q_n \cup \partial Q_n)$ for some $\beta > 0$ independent of (n,k). Moreover, following the arguments of the above theorem as well as that of [16] Chapter V, Theorem 6.1, we have

$$\|u^k_n\|_{C_b^{1+\frac{\beta}{2},2+\beta}(Q_n \cup \partial Q_n)} \leq C_k, \tag{5.24}$$

where C_k depends on T, c_0, C_0, L, d in Assumption 4.1, the derivatives of the coefficients σ, f, g, and the index k, but is uniform in n. Now fix k and send $n \to \infty$. Following the arguments of [16] Chapter V, Theorem 8.1 and using the uniform estimate (5.24), there exists $u^k \in C_b^{1+\frac{\beta}{2},2+\beta}([0,T]\times\mathbb{R}^d)$ such that

$$\begin{gathered}\partial_t u^k(t,x) + \frac{1}{2}\sigma_k^2(t,x,u^k,\partial_x u^k) : \partial^2_{xx}u^k + f_k(t,x,u^k,\partial_x u^k) = 0, (t,x)\in[0,T)\times\mathbb{R}^d;\\ u^k(T,x) = g(x), \quad x\in\mathbb{R}^d.\end{gathered} \tag{5.25}$$

Step 2. In this step we prove the first line of (4.4). Denote

$$\tilde\sigma_k(t,x) := \sigma_k\big(t,x,u^k(t,x),\partial_x u^k(t,x)\big), \quad \tilde f_k(t,x,y,z) := f_k\big(t,x,y,z\tilde\sigma_k^{-1}(t,x)\big).$$

By our conditions, $\tilde\sigma_k$ is uniformly Lipschitz continuous in x, with a Lipschitz constant possibly depending on k, and $\tilde f_k$ is uniformly Lipschitz continuous in (y,z), with Lipschitz constant uniform in k. By standard arguments for (strong) BSDEs, we see that

$$\begin{gathered} u^k(t,x) = \tilde Y_t^{k,t,x}, \quad \text{where}\\ \tilde X_s^{k,t,x} = x + \int_t^s \tilde\sigma_k(r,\tilde X_r^{k,t,x})dB_r,\\ \tilde Y_s^{k,t,x} = g(\tilde X_T^{k,t,x}) + \int_s^T \tilde f_k(r,\tilde X_r^{k,t,x},\tilde Y_r^{k,t,x},\tilde Z_r^{k,t,x})dr - \int_s^T \tilde Z_r^{k,t,x}dB_r.\end{gathered} \tag{5.26}$$

Since g and $\tilde f_k(t,x,0,0)$ are bounded by C_0. It is clear that

$$|u^k(t,x)| \le M_0 \quad \text{where } M_0 \text{ depends only on } T, L, C_0, c_0, \text{ and } d, \text{ but not on } k. \tag{5.27}$$

We next estimate $|\partial_x u|$ under Assumption 4.1 (v). Note that the first case there implies $|\partial_z\sigma(t,x,y,z)| \le \frac{C_0}{1+|z|}$. Applying [16] Chapter VI, Theorem 3.1 on PDE (5.23), with $m=2$, $\varepsilon = 0$, $P(|z|) = L$, and $\mu_1 = C_0$ in (3.2)-(3.6) there, and passing $n\to\infty$, we obtain

$$|\partial_x u^k(t,x)| \le M_1 \quad \text{where } M_1 \text{ depends only on } T, L, C_0, c_0, \text{ and } d, \text{ but not on } k. \tag{5.28}$$

In the second case that $d=1$, denote $v^k := \partial_x u^k$. Then v^k satisfies the following PDE:

$$\begin{gathered}\partial_t v^k + \tfrac{1}{2}\tilde\sigma_k^2\partial^2_{xx}v^k + \tilde b_k\partial_x v^k + \tilde c_k v^k + \partial_x f_k(t,x,u^k,v^k), \quad v^k(T,x) = \partial_x g(x),\\ \text{where} \quad \tilde b_k(t,x) := \tilde\sigma_k\partial_x\tilde\sigma_k(t,x) + \partial_z f_k(t,x,u^k,v^k), \quad \tilde c_k(t,x) := \partial_y f_k(t,x,u^k,v^k).\end{gathered}$$

Note that $|\partial_x f_k|, |\partial_y f_k|, |\partial_x g| \le L$, then one may easily verify (5.28) in this case too. Now let $k \ge M_1 + 1$, we see that $I_k(\partial_x u^k) = 1$ and thus $\varphi_k(t,x,u^k,\partial_x u^k) = \varphi(t,x,u^k,\partial_x u^k)$ for $\varphi = \sigma, f$. That is, u^k is a classical solution to the original PDE (4.2).

We finally prove the Hölder continuity of u in terms of t. Let k be large enough and omit the subscripts $_k$ and superscripts ^{k} in (5.26). Then we have the representation $u(t,x) = \tilde Y_t^{t,x}$, and $\tilde Y^{t,x}_s = u(s,\tilde X^{t,x}_s)$, $\tilde Z^{t,x}_s = \partial_x u\tilde\sigma(s,X^{t,x}_s)$ are bounded. For $0\le t_1<t_2\le T$ and $x\in\mathbb{R}^d$, by (5.28) we have

$$\begin{aligned}
&|u(t_1,x)-u(t_2,x)|^2 \le C\mathbb{E}\Big[|u(t_1,x)-u(t_2,\tilde X^{t_1,x}_{t_2})|^2 + |u(t_2,\tilde X^{t_1,x}_{t_2}) - u(t_2,x)|^2\Big]\\
&\le C\mathbb{E}\Big[|Y^{t_1,x}_{t_1} - Y^{t_1,x}_{t_2}|^2 + |\tilde X^{t_1,x}_{t_2} - \tilde X^{t_1,x}_{t_1}|^2\Big]\\
&\le C\mathbb{E}\Big[\int_{t_1}^{t_2}\big[|\tilde f(s,\tilde X^{t_1,x}_s,\tilde Y^{t_1,x}_s,\tilde Z^{t_1,x}_s)|^2 + |Z^{t_1,x}_s|^2 + |\tilde\sigma(s,\tilde X^{t_1,x}_s)|^2\big]ds\Big]\\
&\le C[t_2-t_1].
\end{aligned}$$

This implies the desired Hölder continuity.

Step 3. We now prove the second line of (4.4). We first notice that the C_k in (5.24) may depend on the derivatives of the coefficients and thus (5.24) does not lead to (4.4). Instead, for any k,n large, we see that u satisfies the following PDE on Q_n with u itself as the boundary condition:

$$\begin{aligned}
&\partial_t u + \tfrac12\sigma^2(t,x,I_k(u),I_k(\partial_x u)) : \partial^2_{xx}u + f(t,x,I_k(u),I_k(\partial_x u)) = 0, \quad (t,x)\in Q_n;\\
&u(t,x) = u(t,x), \quad (t,x)\in\partial Q_n.
\end{aligned}$$

Now apply [16] Chapter VI, Theorem 1.1, we have

$$\langle\partial_x u\rangle^\alpha_{[0,T-\delta]\times O_{n-1}} \le C_\delta. \tag{5.29}$$

Since n is arbitrary, this implies the second line of (4.4) immediately.

Step 4. We finally prove (iii). First, again by [16] Chapter VI, Theorem 1.1, we can improve (5.29) to

$$\langle\partial_x u\rangle^\alpha_{[0,T]\times\mathbb{R}^d} \le C_g. \tag{5.30}$$

Then u satisfies the following linear PDE:

$$\partial_t u + \frac12\hat\sigma^2(t,x) : \partial^2_{xx}u + \hat f(t,x) = 0, \quad u(T,x) = g(x), \tag{5.31}$$

where, for $\varphi = \sigma, f$, $\hat\varphi(t,x) = \varphi(t,x,u(t,x),\partial_x u(t,x))$ is uniformly Hölder continuous. Then the estimate of $\partial^2_{xx}u$ is a classical result, see e.g. Krylov [15], Theorem 8.9.2. ■

References

[1] Antonelli, Fabio *Backward-forward stochastic differential equations. Ann. Appl. Probab.* 3 (1993), no. 3, 777-793.

[2] Antonelli, Fabio; Ma, Jin *Weak solutions of forward-backward SDE's. Stochastic Anal. Appl.* 21 (2003), no. 3, 493-514.

[3] Barlow, M. T. *One-dimensional stochastic differential equations with no strong solution. J. London Math. Soc.* 26 (1982), no. 2, 335-347.

[4] Cherny, Alexander S.; Engelbert, Hans-Jurgen *Singular stochastic differential equations.* Lecture Notes in Mathematics, 1858. Springer-Verlag, Berlin, 2005. viii+128 pp.

[5] Costantini, Cristina; Kurtz, Thomas G. *Viscosity methods giving uniqueness for martingale problems. Electron. J. Probab.* 20 (2015), no. 67, 27 pp.

[6] Crandall, M.; Ishii, H.; Lions, P.-L. *User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. (N.S.)* 27 (1992), no. 1, 1-67.

[7] Delarue, Francois *On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. Stochastic Process. Appl.* 99 (2002), no. 2, 209-286.

[8] Delarue, F.; Guatteri, G. *Weak existence and uniqueness for forward-backward SDEs. Stochastic Process. Appl.* 116 (2006), no. 12, 1712-1742.

[9] Ekren, Ibrahim; Touzi, Nizar; Zhang, Jianfeng *Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part I. Ann. Probab.* 44 (2016), no. 2, 1212-1253.

[10] Ekren, Ibrahim; Touzi, Nizar; Zhang, Jianfeng *Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part II. Ann. Probab.* 44 (2016), no. 4, 2507-2553.

[11] El Karoui, N.; Huang, S.J. *A general result of existence and uniqueness of backward stochastic differential equations. Backward Stochastic Differential Equations*, (1997), 27–36, N. El Karoui and L. Mazliak, eds., Longman, Harlow.

[12] Halidias, N.; Kloeden, P. E. *A note on strong solutions of stochastic differential equations with a discontinuous drift coefficient*, *Journal of Applied Mathematics and Stochastic Analysis*, Volume 2006 (2006), Article ID 73257, 6 pages.

[13] Hamadene, S.; Lepeltier, J.-P. *Zero-sum stochastic differential games and backward equations. Systems Control Lett.* 24 (1995), no. 4, 259-263.

[14] Hu, Y.; Peng, S. *Solution of forward-backward stochastic differential equations. Probab. Theory Related Fields* 103 (1995), no. 2, 273-283.

[15] Krylov, Nicolai V. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. American Mathematical Society, 1996.

[16] Ladyzenskaja, O. A.; Solonnikov, V. A.; Uralceva, N. N. *Linear and quasilinear equations of parabolic type.* (Russian) Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I. 1968 xi+648 pp.

[17] Ma, Jin; Protter, Philip; Yong, Jiongmin *Solving forward-backward stochastic differential equations explicitly – a four step scheme. Probab. Theory Related Fields* 98 (1994), no. 3, 339-359.

[18] Ma, Jin; Wu, Zhen; Zhang, Detao; Zhang, Jianfeng *On well-posedness of forward-backward SDEs – a unified approach. Ann. Appl. Probab.* 25 (2015), no. 4, 2168-2214.

[19] Ma, Jin; Yong, Jiongmin *Forward-backward stochastic differential equations and their applications.* Lecture Notes in Mathematics, 1702. Springer-Verlag, Berlin, 1999. xiv+270 pp.

[20] Ma, Jin; Zhang, Jianfeng *On weak solutions of forward-backward SDEs. Probab. Theory Related Fields*. 151 (2011), no. 3-4, 475-507.

[21] Ma, Jin; Zhang, Jianfeng; Zheng, Ziyu *Weak solutions for forward-backward SDEs – a martingale problem approach. Ann. Probab.* 36 (2008), no. 6, 2092-2125.

[22] Pardoux, Etienne; Tang, Shanjian *Forward-backward stochastic differential equations and quasilinear parabolic PDEs. Probab. Theory Related Fields*. 114 (1999), no. 2, 123-150.

[23] Peng, Shige; Wu, Zhen *Fully coupled forward-backward stochastic differential equations and applications to optimal control. SIAM J. Control Optim.* 37 (1999), no. 3, 825-843.

[24] Pham, Triet; Zhang, Jianfeng *Two person zero-sum game in weak formulation and path dependent Bellman-Isaacs equation. SIAM J. Control Optim.* 52 (2014), no. 4, 2090-2121.

[25] Protter, Philip E. *Stochastic integration and differential equations.* Second edition. Version 2.1. Corrected third printing. Stochastic Modeling and Applied Probability, 21. Springer-Verlag, Berlin, 2005. xiv+419 pp.

[26] Soner, H. Mete; Touzi, Nizar *Stochastic target problems, dynamic programming and viscosity solutions*, *SIAM Journal on Control and Optimization*, 41 (2002), 404-424.

[27] Soner, H. Mete; Touzi, Nizar; Zhang, Jianfeng *Wellposedness of second order backward SDEs. Probab. Theory Related Fields* 153 (2012), no. 1-2, 149-190.

[28] Stroock, Daniel W.; Varadhan, S. R. Srinivasa *Multidimensional diffusion processes.* Reprint of the 1997 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2006. xii+338 pp.

[29] Tsirelson, Boris *An Example of a Stochastic Differential Equation Having no Strong Solution*. *Theory of Probability and Its Applications*. 20 (1975), no. 2, 416-418.

[30] Yong, Jiongmin *Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. Probab. Theory Related Fields*. 107 (1997), no. 4, 537-572.

[31] Yong, Jiongmin; Zhou, Xun Yu *Stochastic controls. Hamiltonian systems and HJB equations.* Applications of Mathematics (New York), 43. Springer-Verlag, New York, 1999. xxii+438 pp.

[32] Zhang, Jianfeng *The wellposedness of FBSDEs. Discrete Contin. Dyn. Syst. Ser. B* . 6 (2006), no. 4, 927-940 (electronic).

[33] Zhang, Jianfeng *Backward Stochastic Differential Equations – from linear to fully nonlinear theory*, Springer, New York, 2017.

[34] Zheng, W. A. *Tightness results for laws of diffusion processes application to stochastic mechanics. Ann. Inst. H. Poincare Probab. Statist.* 21 (1985), no. 2, 103-124.