A Probabilistic Approach for Discontinuous Coefficients FBSDEs and PDEs

Yunxi Xu

Supervisor: Kihun Nam

Co-supervisor: Fima Klebaner

A thesis submitted for the degree of

Master of Philosophy

at Monash University in 2021

School of Mathematics
Monash University
Melbourne, Australia
2021

Copyright notice

Abstract

This thesis aims at analyzing the solvability of the multidimensional deterministic coefficients forward backward stochastic differential equations (FBSDEs) with the following form

$$\begin{aligned} dX_t &= b(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t)dW_t; \quad X_0 = x \\ dY_t &= -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t; \qquad Y_T = h(X_T). \end{aligned} \tag{1}$$

Particularly, we allow all the coefficients to be discontinuous in x. By developing a pure probabilistic method, we provide many sets of sufficient conditions to prove the existence and uniqueness results for FBSDE (1).

Meanwhile, when the coefficients are smooth enough, the deterministic coefficients FBSDEs have a strong connection with the quasilinear parabolic partial differential equations (PDEs) of the form

$$\mathcal{L}u(t,x) + \partial_x u(t,x) \cdot b(t,x,u,\partial_x u \cdot \sigma) + f(t,x,u,\partial_x u \cdot \sigma) = 0 \quad u(T,x) = h(x), \tag{2}$$

where $\mathcal{L} := \partial_t + \frac{1}{2}\sum_{i,j=1}^{m}(\sigma\sigma^\intercal)_{ij} \cdot \partial^2_{x_i x_j}$. We are also interested in the solvability of above PDEs with discontinuous coefficients. However, when the coefficients are discontinuous in x, additional requirements are needed to obtain sufficient regularity of u.

In this thesis we will introduce a probabilistic approach related to Girsanov transform such that we could apply results from stochastic differential equations (SDEs) and backward stochastic differential equations (BSDEs) to help us provide the existence and uniqueness results for FBSDE (1). The core idea of our approach is to interplay between decoupled FBSDEs and coupled FBSDEs. Instead of solving FBSDE (1) directly, we start with solving a decoupled FBSDE. After that, we apply the Girsanov theorem to transform the decoupled FBSDE into a new FBSDE with the same form as FBSDE (1) but under a different filtration. At last, we verify that the new FBSDE has a (unique) strong solution such that the same result holds to FBSDE (1).

Furthermore, we will show that, under certain conditions, the solvability of FBSDE (1) would imply the existence of a solution for PDE (2). Since our coefficients may be discontinuous, one can not expect the solutions for PDE (2) are in the Sobolev space as usual. One can view our thesis as a study of parabolic PDEs with measurable coefficients with a slightly weaker class of solutions than the Sobolev solutions.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Yunxi Xu
14 10 2021

Publications during enrolment

Chapter 3 are based off joint work with Kihun Nam. The manuscript, titled "Coupled FBSDEs with Measurable Coefficients and its Application to Parabolic PDEs", is currently on the arXiv under the same name.

Acknowledgements

Undoubtedly, the person I want to thank the most is my supervisor Kihun Nam, who brings me to the real mathematics world and guides me to be a researcher. He is always patient with me and shares with me his ideas and knowledge. It is a real pleasure to have a such kind, talented, and knowledgeable person as my supervisor.

Also, I want to thank my dear friend Tao Chen, a PhD candidate at Sichuan University. It is never easy for me when I transfer from finance to mathematics. He really helps me a lot in comprehending mathematical concepts and reading research papers.

Of course I want to thank School of Mathematics and the Centre for Quantitative Finance and Investment Strategies for giving me this precious opportunity. In particular, I want to thank Fima Klebaner for being my co-supervisor.

I sincerely thank Dr Zhou Zhou and Dr Hyungbin Park for taking their precious time to mark my thesis and providing useful amendment suggestions.

At last I want to thank my family and my soul mate Jia Feng. Without their understanding and support I can not do anything.

Frequently Used Notation and Definition

Probability space

Let I be an index set with a total order. We denote $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ a filtered probability space on which is defined an $\mathbb{R}^d$-dimensional Brownian motion W, such that its augmented filtration $\mathbb{F} := (\mathcal{F}_t)_{t \in I}$ is generated by W. For a stochastic process X, we let $\mathbb{F}^X$ be the augmented filtration generated by X, and without specific explanation, we always consider $\mathbb{F} = \mathbb{F}^W$. We identify random variables that are equal $\mathbb{P}$-almost surely, so as the equality and inequality in the $\mathbb{P}$-almost sure sense.

Vectors and Matrices

For a matrix $X \in \mathbb{R}^{n \times d}$, we denote X^i the ith row of X and X^{ij} to be the component at row i and column j. We consider a vector as a column matrix and denote $X^\intercal$ the transpose of X.

Banach spaces

We denote $|X| := \sqrt{\mathrm{Tr}(XX^\intercal)}$ as the Euclidean norm for a matrix X.

We define the $\mathbb{L}^p$ space for the random variable X with

$$\|X\|_p := (\mathbb{E}|X|^p)^{\frac{1}{p}} < \infty; \quad \textit{for} \quad p < \infty$$
$$\|X\|_\infty := \operatorname*{ess\,sup}_{\omega} |X(\omega)| < \infty; \quad \textit{for} \quad p = \infty.$$

As usual, we assume Borel σ-algebra on Euclidean spaces, and we let $\mathbb{S}^p(E)$ and $\mathbb{H}^p(E)$ be the spaces of the adapted E-valued process X with

$$\|X\|_{\mathbb{S}^p} := \left\| \sup_{t \in [0,T]} \mathbb{E}|X_t| \right\|_p < \infty,$$

and

$$\|X\|_{\mathbb{H}^p} := \mathbb{E}\left[\int_0^T |X_t|^2 dt \right]^{\frac{p}{2}} < \infty; \quad \textit{for} \quad p < \infty$$
$$\|X\|_{\mathbb{H}^\infty} := \operatorname*{ess\,sup}_{(t,\omega)} |X_t(\omega)| < \infty; \quad \textit{for} \quad p = \infty.$$

Borel algebra

For a Banach space E, we let $\mathcal{B}(E)$ be the Borel algebra on E.

Contents

Chapter 1

Introduction

1.1 Introduction to Stochastic Differential Equations and Backward Stochastic Differential Equations

To better understand our main topic *forward backward stochastic differential equations* (FBSDEs), it is inevitable to know some backgrounds about forward *stochastic differential equations* (SDEs) and *backward stochastic differential equations* (BSDEs). So the first section is a brief introduction about SDEs and BSDEs with related topics. After that we will introduce the background for FBSDEs such as the usual setups, their applications, previous results, etc. In the later part, we will show that the solvability of FBSDEs is equivalent to the solvability of parabolic *partial differential equations* (PDEs) under smooth enough conditions. That would lead our interests from FBSDEs to quasilinear PDEs.

1.1.1 What are SDEs and BSDEs?

Forward SDEs, which arise when a random noise is introduced into *ordinary differential equations* (ODEs), usually recognize as the following form

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \tag{1.1}$$

where W is a d-dimensional Brownian motion; $x \in \mathbb{R}^n$ stands for the initial state; $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ is often called as the *drift* coefficient of the state process X; and $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ is known as the *diffusion* coefficient. We call X the solution of SDE (1.1) if X_t is adapted to the filtration $\mathcal{F}_t$, $\int_0^t (|b(s, X_s)| + |\sigma(s, X_s)|^2)\, ds < \infty$ for all $0 \le t < \infty$, and equation (1.1) holds.

Generally speaking, SDEs are applied to describe the evolution of a system that contains randomness. They are useful for explaining many phenomenons arise in finance, physics, biology, etc. For instance, Black–Scholes model, a well-known financial model governing the price evolution of European-style options, has an inseparable relation with SDEs. For more applications about SDEs, we refer the book Klebaner (2012) for an extensive list of references.

A classical form of a BSDE is

$$Y_t = \xi - \int_t^T f(s, Y_s, Z_s)ds + \int_t^T Z_s dW_s, \tag{1.2}$$

where the terminal condition ξ is a $\mathbb{R}^n$-valued $\mathcal{F}_T$-measurable random variable; the *driver* $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}^n$ is a $\mathcal{P} \bigotimes \mathcal{B}(\mathbb{R}^n) \bigotimes \mathcal{B}(\mathbb{R}^{n\times d})$-measurable function, where $\mathcal{P}$ denotes the σ-algebra of $\mathcal{F}_t$ progressively measurable subsets of $\Omega \times [0,T]$. Similar to SDEs, we call a pair (Y,Z) taking value in $\mathbb{R}^n \times \mathbb{R}^{n\times d}$, a solution of the BSDE (1.2) if the predictable processes Y and Z satisfy $\int_0^t (|f(s,Y_s,Z_s)| + |Z_s|^2)\,ds < \infty$ for all $0 \le t \le T$, and equation (1.2) holds.

Usually we call a BSDE is n-dimensional if $n \ge 2$ and one-dimensional if $n = 1$; a *Markovian* BSDE if the driver of the BSDE is deterministic; and a *quadratic* BSDE if the driver has at most quadratic growth in z.

The classical BSDE was first introduced by Pardoux and Peng (1990) and further discussed in El Karoui et al. (1997). The first paper provided the general well-posedness result for BSDEs that is when $\xi \in \mathbb{L}^2(\mathbb{R}^n)$, $f(\cdot,0,0) \in \mathbb{H}^2(\mathbb{R}^n)$, and f is uniformly Lipschitz. In the second paper the authors introduced the theory of contingent claim valuation and its relation to BSDEs. They pointed out that the well-known models such as Merton model, Black–Scholes model, etc., can be expressed in terms of BSDEs. Also the second paper collected and generalized many previous results done by the authors. For instance, Pardoux and Peng (1992), in which they introduced a solution Y of one Markovian BSDE becomes a viscosity solution of a quasilinear parabolic PDE and provided a set of sufficient conditions that guarantees the solution obtained by one BSDE to be a solution of the corresponding PDE; Peng (1992), in which the author firstly introduced the comparison theorem for one-dimensional BSDEs; etc. So we would highly recommend El Karoui et al. (1997) for ones who are interested in BSDEs and related topics. Meanwhile, BSDEs are also useful for the theory of stochastic optimal control. Bismut (1973) was the first paper introduced BSDE and applied it to analyze the optimal control problem for the linear case. Later, Hamadene and Lepeltier (1995) applied BSDE results to obtain the existence of optimal strategy for stochastic zero-sum differential games, and El-Karoui and Hamadène (2003) applied BSDE to study risk-sensitive control problem.

One should notice that there are numerous generalizations of the classical SDEs and BSDEs, for instance, the coefficients may depend on the path of the stochastic processes which leads to path-dependent SDEs and BSDEs; also, the coefficients may depend on the distribution of the processes, in this sense, the generalization including McKean-Vlasov and mean-field SDEs and BSDEs.

1.1.2 Well-posedness Theory for Classical SDEs

Since our project aims at analyzing the solvability of deterministic coefficients FBSDEs, in this subsection we would only focus on the well-posedness results for deterministic coefficients SDEs.

As one may expect, when the coefficients are locally Lipschitz and under linear growth condition, SDE (1.1) is uniquely solvable. The existence is carried out by the Picard iterations, which is a successive approximations technique, and for one-dimensional case, the uniqueness can be easily obtained via Gronwall's lemma, which we will introduce in chapter 2. However, the Lipschitz continuity condition is not always satisfied for applications in practice. The study of SDEs with non-Lipschitz coefficients has received a lot of attention in recent years. Generally speaking, there are mainly two branches for well-posedness results for SDEs (1.1). One is when σ is a constant, and another is when σ is a deterministic function.

When σ is a constant, Zvonkin (1974) provided the existence of a unique solution for one-dimensional (1.1) where the drift coefficient is bounded and measurable. Later, Veretennikov (1980) generalized Zvonkin's work to the multidimensional case. In these papers, the authors applied the similar idea that they firstly employ the estimates of solutions of parabolic partial differential equations to construct weak solutions. Then a pathwise uniqueness argument is applied to ensure a unique strong solution. In recent times, Meyer-Brandis and Proske (2010) developed a new technique for the construction of strong solutions for such bounded drift SDEs under a certain symmetry condition. This method is based on Malliavin calculus and white noise analysis. Menoukeu-Pamen et al. (2013) developed that approach and derived more general results by relaxing the symmetry condition on the drift term. Recently Menoukeu-Pamen and Mohammed (2019) generalized their previous work by proving the well-posedness of SDE with the drift coefficient is time dependent and has spatial linear growth.

When σ is a deterministic function, Zvonkin (1974) and Veretennikov (1980) also provided well-posedness results under such case when the drift coefficient b is bounded and the diffusion coefficient σ is a non-degenerate Lipschitz function. In Le Gall (1984), the author obtained existence and uniqueness result for one-dimensional SDE (1.1) with a bounded drift term and a measurable diffusion term. When b is a unbounded function, Gyongy and Martinez (2001) provided that if σ is a non-degenerate Lipschitz function and there exist a non-negative constant C and a positive valued function $F \in \mathbb{L}^{n+1}(\mathbb{R}^+ \times \mathbb{R}^n)$ such that $|b(t,x)| \leq C + F(t,x)$, then there exists of a unique solution for SDE (1.1). Roughly speaking, their method is to construct a solution via approximation and then prove the path-wise uniqueness of the solution. Therefore, by the well-known result of Yamada and Watanabe, which we will share in chapter 2, they obtain the existence of a unique strong solution. Later, Zhang (2005) studied the similar SDEs with a non-Lipschitz diffusion coefficient.

1.1.3 Well-posedness Theory for Markovian BSDEs

As our interest lies in deterministic coefficients cases, in this subsection we will introduce the well-posedness theory related to Markovian BSDEs.

El Karoui et al. (1997) showed the existence of a unique solution $(Y, Z) \in \mathbb{S}^p(\mathbb{R}^n) \times \mathbb{H}^p(\mathbb{R}^{n \times d})$ for $p > 1$ when the driver f is uniformly Lipschitz. They obtained the result by applying Banach fixed point theorem and a priori estimates technique. Pardoux (1999) relaxed the Lipszhitz condition to allow the driver is non-Lipschitz in y. Later, Hamadène (2003) further relaxed the Lipschitz condition of the driver to uniform continuity condition with linear growth. An interesting result was introduced by Kobylanski (2000). In this paper the author introduced the first well-posedness result for quadratic BSDEs. Her method is based on exponential change of variable and a priori estimates to deduce the existence of a solution for quadratic BSDEs. After that she applied comparison theorem introduced by El Karoui et al. (1997) and then obtained the uniqueness result for the one-dimensional quadratic BSDEs. A particular type of quadratic BSDEs, called the diagonally quadratic BSDEs, it stands for the BSDEs with the driver f such that $|f^i(t,y,z)| \leq C(1 + |y| + |z^i|^2)$ for some constant C. In such case, Hu and Tang (2016) provided the sufficient conditions for the well-posedness result, and notice that it is the first result on the general solvability for multidimensional quadratic BSDEs. They borrowed the idea from Cheridito and Nam (2015) to construct a global solution from local solutions and uniform a priori estimates.

1.2 Introduction to Forward Backward Stochastic Differential Equations

In this section we will introduce the FBSDEs with their applications and development history, and in the later part we will show that the solutions of FBSDEs are somehow equivalent to the solutions of parabolic PDEs.

1.2.1 FBSDEs and Their Applications

FBSDEs have received strong attention in recent years because of their interesting structure and usefulness in various practical applications. The most common *fully coupled* FBSDEs are of the following form

$$\begin{aligned} X_t &= x + \int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s)dW_s \\ Y_t &= h(X_T) - \int_t^T g(s, X_s, Y_s, Z_s)ds + \int_t^T Z_s dW_s, \end{aligned} \tag{1.3}$$

where W is a d-dimensional Brownian motion; $(b, \sigma, g) : \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^n$ are $\mathcal{P} \bigotimes \mathcal{B}(\mathbb{R}^m) \bigotimes \mathcal{B}(\mathbb{R}^n) \bigotimes \mathcal{B}(\mathbb{R}^{n\times d})$-measurable functions; $x \in \mathbb{R}^m$, and $h : \Omega \times \mathbb{R}^m \to \mathbb{R}^n$ such that $\omega \mapsto h(\omega, x)$ is $\mathcal{F}_T$-measurable for all $x \in \mathbb{R}^m$, are initial condition and terminal condition, respectively. We call the adapted processes $(X, Y, Z) \in \mathbb{H}^2(\mathbb{R}^m) \times \mathbb{H}^2(\mathbb{R}^n) \times \mathbb{H}^2(\mathbb{R}^{n\times d})$ a solution if FBSDE (1.3) holds for any $t \in [0,T]$.

As the second equation of (1.3) is a BSDE, FBSDEs share similar applications with BSDEs. For instance, FBSDEs are also useful for contingent claim valuation such as option pricing. Consider a security market contains only one stock and one risk-free bond. We assume they are subject to the following equations

$$\begin{aligned} dP_t^0 &= r_t P_t^0 dt, \quad \text{(bond)}; \\ dP_t &= b_t P_t dt + \sigma_t P_t dW_t, \quad \text{(stock)}, \end{aligned}$$

where r is the risk-free rate; b, σ are the appreciation rate and volatility rate of the stock, respectively. Now we assume an agent sells one option at price y and then invest it in such market. At time t, we denote his total wealth is Y_t, and π_t as his portion of wealth into the stock. For simplicity, we do not consider his consumption during the time. In this sense, we could obtain the following

$$dY_t = (r_t Y_t + \theta_t Z_t)dt + Z_t dW_t; \quad Y_0 = y$$

where $Z_t = \pi_t \sigma_t$, and the risk premium process $\theta_t = \frac{b_t - r_t}{\sigma_t}$. Our objective as the agent is to hedge the payoff $\xi = h(P_T)$ at time T. In other words, we are looking for $y^* = \inf\{y = Y_0;\ \exists \pi \text{ such that } Y_T = \xi\}$. Therefore, our problem can be translated into an FBSDE problem

$$\begin{aligned} dP_t &= b_t P_t dt + \sigma_t P_t dW_t, \quad P_0 = p; \\ dY_t &= (r_t Y_t + \theta_t Z_t)dt + Z_t dW_t, \quad Y_T = h(P_T). \end{aligned} \tag{1.4}$$

Usually we call such FBSDE a *decoupled* FBSDE, as the coefficients of the forward equation do not depend on Y and Z. If FBSDE (1.4) is solvable, we can obtain $\pi = \sigma^{-1} Z$ as the optimal portfolio, and $Y_0 = y^*$ as the option price.

Decoupled FBSDEs, came along with BSDEs, was fisrtly introduced by Bismut (1973) as a dual problem of stochastic control which corresponds to the Pontryagin principle. The coupled FBSDEs were firstly introduced by Antonelli (1993) as an extension of the earlier theory of BSDEs. The fully coupled FBSDEs were discussed in the book Ma et al. (1999). This book not only contains many important results and approaches about fully coupled FBSDEs, but also provides tremendous applications arising in mathematical finance or in stochastic control problems.

As we mentioned before, fully coupled FBSDE is a strong tool for studying the optimization problem of a stochastic control system. For instance, we consider a control system is

$$dX_t = b(t, X_t, v_t)dt + \sigma(t, X_t, v_t)dW_t, \quad X_0 = x,$$

where v is an admissible control process, i.e. an $\mathcal{F}_t$-adapted square integrable process taking values in a given subset $\mathcal{A}$ of $\mathbb{R}$. For simplicity, we consider everything is one-dimensional. A common control problem is to minimize some cost functional

$$J(v) = \mathbb{E} \int_0^T f(t, X_t, v_t)dt + g(X_T)$$

over the set of admissible controls. This optimization problem is the so-called stochastic maximum principle which may be considered as a natural generalization of the well-known Pontryagin's maximum principle to situations with uncertainty. This principle tells us that, under certain reasonable assumptions imposed on the coefficients, then, necessarily, there exists a pair of square integrable adapted processes (Y, Z), such that the triple of processes (X, Y, Z) satisfies the stochastic Hamiltonian system, which is an FBSDE, where the Hamiltonian H is defined by

$$H(t, x, y, z) = \inf_{v \in \mathcal{A}} b(t, x, v)y + \sigma(t, x, v)z + f(t, x, v).$$

In this sense, we obtain the following FBSDE (the stochastic Hamiltonian system)

$$\begin{aligned} dX_t &= H_y(t, X_t, Y_t, Z_t)dt + H_z(t, X_t, Y_t, Z_t)dW_t, \quad X_0 = x \\ dY_t &= -H_x(t, X_t, Y_t, Z_t)dt + Z_t dW_t, \quad Y_T = g_x(X_T). \end{aligned}$$

For details about fully coupled FBSDEs and the Hamiltonian, we will discuss them in Appendix D.

Just like SDEs and BSDEs, we call an FBSDE a *Markovian* FBSDE if the coefficients are deterministic functions, and based on the different dependency of the coefficients, we also have path-dependent, McKean-Vlasov, and mean-field FBSDEs.

At last of this subsection we want to mention that, unlike SDEs and BSDEs, we do not expect the general existence and uniqueness results for an FBSDE, even the coefficients are under really good conditions e.g. uniformly Lipschitz conditions. A simple example can be constructed from a two-point boundary value problem for a system of one-dimensional linear ODEs which is *not* solvable:

$$\begin{aligned} \dot{X}(t) &= AX(t), \quad X(0) = x, \\ \dot{Y}(t) &= BY(t), \quad Y(T) = CX(T), \end{aligned}$$

where A, B and C are constants. Then for any properly defined σ and σ', the following FBSDE

$$\begin{aligned} dX_t &= AX_t dt + \sigma(t, X_t, Y_t, Z_t)dW_t, \quad X_0 = x \\ dY_t &= BY_t dt + \sigma'(t, X_t, Y_t, Z_t)dW_t, \quad Y_T = CX_T, \end{aligned}$$

does not admit any solutions. It can be easily proved by the contradiction that if the above FBSDE is solvable, then $(\mathbb{E}X, \mathbb{E}Y)$ should be a solution of the ODE system.

1.2.2 Brief History of Development for FBSDEs

The first result for coupled FBSDE was introduced in Antonelli (1993). In this paper the author obtained the well-posedness result for uniformly Lipschitz coefficients FBSDEs over a "small" duration by the fixed point theorem, and constructed a counterexample showing that for coupled FBSDEs, large time duration might lead to non-solvability.

Later, two methods were established to study the fully coupled FBSDEs on an arbitrarily given time interval. The first method, named "Four-Step Scheme", which combines PDE methods and methods of probability, was introduced in Ma et al. (1994). The method is inspired from an observation that the solvability of a Markovian FBSDE is equivalent to a "solution" of some parabolic system. In this paper they proved the existence and the uniqueness for fully coupled FBSDEs on an arbitrarily given time interval, but they required the diffusion coefficients are non-degenerate. Meanwhile, since in general, the PDE approach can not be used to deal with the case when the coefficients themselves are randomly disturbed, this method should only work for Markovian FBSDEs. But recently, Ma et al. (2015) proved the well-posedness for random coefficients FBSDEs. Following the main point of the Four-Step Scheme, they found a function u, which is the so-called *decoupling field*, such that $Y_t = u(t, X_t)$, but with u is a random field.

Another important method, named "Method of Continuation", was first introduced in Hu and Peng (1995) for Markovian FBSDEs and generalized by Yong (1997) to allow the coefficients are random. Later Peng and Wu (1999), and Pardoux and Tang (1999) borrowed the spirit of the method of continuation to study the well-posedness property for fully coupled FBSDEs. Recently in Yong (2010), the author applied the method of continuation to study another type of fully coupled FBSDEs. Such FBSDEs are with mixed initial-terminal conditions, that means FBSDEs are of the form (1.3) but with different boundary conditions $X_0 = h'(X_T, Y_0)$ and $Y_T = h(X_T, Y_0)$. The main assumption of this method is the so-called "monotonicity conditions" on the coefficients, which is usually not easy to verify. Generally speaking, by denoting $(X_t, Y_t, Z_t) = \theta_t$, the method of continuation considers the following FBSDE

$$\begin{aligned}
&dX_t = (b^0(t) + (1-\alpha)b^1(t,\theta_t) + \alpha b^2(t,\theta_t))dt + (\sigma^0(t) + (1-\alpha)\sigma^1(t,\theta_t) + \alpha\sigma^2(t,\theta_t))dW_t,\\
&dY_t = (g^0(t) + (1-\alpha)g^1(t,\theta_t) + \alpha g^2(t,\theta_t))dt + Z_t dW_t,\\
&X_0 = x, \qquad Y_T = h^0(T) + (1-\alpha)h^1(X_T) + \alpha h^2(X_T),
\end{aligned}$$

where $(b^1, \sigma^1, g^1, h^1)$ and $(b^2, \sigma^2, g^2, h^2)$ are linked by a direct *bridge*. For any well defined square integrable functions $(b^0, \sigma^0, g^0, h^0)$ and $x \in \mathbb{R}^m$, the essence of the method of continuation is to prove that there exists a fixed step-length $\epsilon_0 > 0$, such that for some $\alpha \in [0, 1)$ above FBSDE is uniquely solvable. Then the same conclusion holds for α being replaced by $\alpha + \epsilon \leq 1$ with $\epsilon \in [0, \epsilon_0]$.

Since FBSDEs are helpful for the research of stochastic control problems, recently the McKean-Vlasov, and mean-field FBSDEs attracted a lot of attention. Such FBSDEs aim at studying optimal control of McKean–Vlasov dynamics or mean-field games. By denoting $\mathcal{L}(X)$, the law (also called the distribution) of X, a common McKean–Vlasov type stochastic optimal control would be like:

Assume the stochastic dynamics satisfying a stochastic differential equation of the form

$$dX_t = b(t, X_t, \mathcal{L}(X_t), v_t)dt + \sigma(t, X_t, \mathcal{L}(X_t), v_t)dW_t, \quad X_0 = x,$$

and our goal is to minimize the following cost functional

$$J(v) = \mathbb{E}\int_0^T f(t, X_t, \mathcal{L}(X_t), v_t)dt + g(\mathcal{L}(X_T), X_T)$$

over the set of admissible controls.

As for mean-field game, the terminology mean-field is borrowed from statistical physics, and the goal of the theory is to derive effective equations for the optimal behavior of any single player when the size of the population grows unboundedly. In this sense, for an N-players game (with N to be very large), one can expect the empirical measure $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ will not be much affected by the deviation of one single player, and for all practical purposes, one should be able to assume that the empirical measure μ_t^N is approximately equal to its limit μ_t. Therefore, one goal of mean-field game would reduce to search for a *deterministic* flow of measures μ_t such that the law of the optimally controlled process is in fact μ_t. Namely the dynamic constraint and the cost functional are

$$dX_t = b(t, X_t, \mu_t, v_t)dt + \sigma(t, X_t, \mu_t, v_t)dW_t, \quad X_0 = x,$$

and

$$J(v) = \mathbb{E}\int_0^T f(t, X_t, \mu_t, v_t)dt + g(\mu_T, X_T),$$

with $\mathcal{L}(X_t) = \mu_t, \forall t \in [0, T]$.

In this direction, Carmona and Delarue (2015) studied optimal control problem of McKean–Vlasov dynamics and mean-field game with corresponding FBSDEs. For more details about the related topics, we refer book Carmona (2016) and references therein.

Back to our topic, the FBSDEs with discontinuous coefficients have been studied before. When FBSDE (1.3) is decoupled, which means b and σ does not depend on y and z, El Karoui et al. (1997) provided the well-posedness for such FBSDEs when $g(t, x, y, z)$ is Lipschitz with respect to (y, z); Meanwhile, Hamadene et al. (1997) provided the existence of a solution when $g(t, x, y, z)$ is continuous with respect to (y, z). For coupled cases, Carmona et al. (2013) studied the FBSDEs with discontinuous terminal condition h for carbon allowance pricing. Luo et al. (2020), which has similar results to ours, provided existence results for FBSDE (1.3) when σ is a constant, and h is bounded. Chen et al. (2018) provided well-posedness results for the coupled FBSDEs with b is a step function in y.

1.2.3 FBSDEs and Parabolic PDEs

At last of this section, we want to point out that, for Markovian FBSDEs, they have a strong connection with quasilinear parabolic PDEs. We consider a parabolic PDE is of the form

$$\begin{aligned} \mathcal{L}u(t,x) + \partial_x u(t,x) \cdot \tilde{g}(t, x, u, \partial_x u \cdot \sigma(t,x)) + f(t, x, u, \partial_x u \cdot \sigma) = 0 \\ u(T, x) = h(x), \end{aligned} \tag{1.5}$$

where $\mathcal{L} := \partial_t + \frac{1}{2}\sum_{i,j=1}^m (\sigma\sigma^\intercal)_{ij} \cdot \partial^2_{x_i x_j}$. Particularly, we assume all the coefficients are under *good enough* conditions, such that PDE (1.5) and the following SDE are uniquely solvable

$$dX_t = \tilde{g}(t, X_t, u(t, X_t), \partial_x u(t, X_t) \cdot \sigma(t, X_t))dt + \sigma(t, X_t)dW_t, \quad X_0 = x.$$

Then we apply Itô's formula on $u(t, X_t)$ and we could obtain

$$du(t, X_t) = \partial_t u(t, X_t)dt + \partial_x u(t, X_t)dX_t + \frac{1}{2}\partial^2_{xx} u(t, X_t)d[X_t],$$

where $[X]$ denotes the quadratic variation of process X.

Based on PDE (1.5) and above equation, we can deduce that

$$\begin{aligned} du(t, X_t) = &- f(t, X_t, u(t, X_t), \sigma(t, X_t) \cdot \partial_x u(t, X_t))dt \\ &+ \sigma(t, X_t) \cdot \partial_x u(t, X_t)dW_t; \quad u(T, X_T) = h(X_T). \end{aligned}$$

Therefore, (X, Y, Z) with $Y_t = u(t, X_t)$ and $Z_t = \sigma(t, X_t) \cdot \partial_x u(t, X_t)$ is the solution of an FBSDE which is of the form to our interest throughout the thesis,

$$\begin{aligned} dX_t &= \tilde{g}(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t)dW_t; \quad X_0 = x \\ dY_t &= -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t; \quad Y_T = h(X_T). \end{aligned} \tag{1.6}$$

Also, above observation motivates the development of "Four-Step Scheme" and again we want to mention that usually we call function u the *decoupling field* of the FBSDE.

Parabolic PDEs with measurable coefficients have been studied broadly in PDE literature; see Maugeri et al. (2000), Kim and Krylov (2007) and references therein. Most of the previous literature on PDE with discontinuous coefficients focused on the viscosity solution or the class of solutions in the Sobolev space. The viscosity solution cannot be applied if $n > 1$, because the comparison principle does not hold in general. On the other hand, the decoupling field u does not have to be in the Sobolev space $W^{1,2}_p([0, T] \times \mathbb{R}^m)$; we only require $u(t, X_t)$ to be an Itô process. Under certain conditions, Chitashvili and Mania (1996), and Mania and Tevzadae (2001) proved that $u \in \mathrm{Dom}(\mathcal{L})$ for an operator $\mathcal{L}$ defined by the closure of $\partial_t + \frac{1}{2}\sigma\sigma^\intercal\partial^2_{xx}$, is the necessary and sufficient condition for $Y_t = u(t, X_t)$ to be an Itô process.

1.3 Thesis Structure

In this thesis we will start with some important theory related to our topic. Since our goal is to provide well-posedness for FBSDE (1.6), in chapter 2 we start with providing the definition of unique solution. When one differential equation has a unique solution, we would sometimes call it uniquely solvable or well-posed. Later, as we introduced before, our method is inspired from Girsanov transform, the related topics are also discussed. Notice that (F)BSDEs have a close relationship with Feynman-Kac formula, and from Feynman-Kac formula, we introduce Kolmogorov's backward equation and Fokker-Planck equation, which are useful for presenting our results. Moreover, since our coefficients are allowed to be irregular, to make the solutions of PDEs be well-defined, we introduce the generalized derivatives and the suitable spaces. At last of the chapter, we will introduce a few SDEs and BSDEs' results in details, which are very helpful to us.

The chapter 3 aims at presenting our results. We provide sufficient conditions for the existence and uniqueness of a strong solution for (1.6) with deterministic measurable coefficients, which can be discontinuous, when $\sigma(t, x)$ is uniformly non-degenerate. Our results generalize Luo et al. (2020) in terms of the growth of coefficients and non-constant σ using the simpler technique. In particular, we prove the existence of measurable function u that satisfies (1.5) in a weak sense. The main technique is the decoupling of the FBSDE (1.6) with the Girsanov transform, and then verified that it is actually a strong solution using the existence of the Markovian representation $(Y_t, Z_t) = (u(t, X_t), d(t, X_t))$. While the Markovian representation resembles the decoupling field in the four step scheme, which is defined through PDE, the existence of u and d stems from a purely probabilistic argument based on Çinlar et al. (1980). In this sense, our method can be seen as a probabilistic generalization of the four step scheme for measurable coefficients. On the other hand, we can also view our thesis as a study of parabolic PDEs with measurable coefficients with a slightly weaker class of solutions than the Sobolev solutions. At last, two applications are shared. One is about optimal control of the spread of the infectious disease. Since the medical resources are usually limited, it is common that the running cost of one patient is not smooth. In this sense, FBSDEs with discontinuous coefficients would be helpful to deal with such optimal control problem. Another application is about the carbon allowance pricing. The natural property of such derivative is its terminal condition is discontinuous. Particularly, we construct an FBSDE model and by solving the FBSDE, one can obtain the allowance price when the firms apply their optimal production strategies.

Chapter 2

Elements of the General Theory

In this chapter we will introduce the general theory related to our topic.

2.1 Definition of Unique Solution and Related Topics

In this section we aims at introducing the definitions of weak and strong solutions, and pathwise uniqueness and uniqueness in law. Important results and examples are also shared. Arguments in this section are borrowed from Karatzas and Shreve (1991).

2.1.1 Strong and Weak Solution

We consider a classical SDE of the following form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{2.1}$$

or written componentwise as

$$dX_t^i = b_i(t, X_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t, X_t)dW_t^j; \quad 1 \leq i \leq d$$

where b_i and σ_{ij} are Borel measurable functions and $1 \leq i \leq d$, $1 \leq j \leq r$, from $[0, \infty) \times \mathbb{R}^d$ into $\mathbb{R}$; W is an r-dimensional Brownian motion while X is a suitable stochastic process with continuous sample path valued in $\mathbb{R}^d$.

Definition 2.1.1. We call X a *strong solution* of SDE (2.1) on the given filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with initial condition ξ if the following conditions are satisfied

- X_t is $\mathcal{F}_t$-adapted,
- $\mathbb{P}(X_0 = \xi) = 1$,
- $\mathbb{P}\left(\int_0^t |b_i(s, X_s)| + |\sigma_{ij}(s, X_s)|^2 ds < \infty\right) = 1$, and
- the integral version of (2.1)

  $$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$$

  holds a.s.

Definition 2.1.2. A *weak solution* of (2.1) is a triple (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,\infty)}$, where

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and for each $t \in [0, \infty)$, $\mathcal{F}_t$ is a sub-σ-algebra of $\mathcal{F}$,
- X_t is $\mathcal{F}_t$-adapted,
- $\mathbb{P}\left(\int_0^t |b_i(s, X_s)| + |\sigma_{ij}(s, X_s)|^2 ds < \infty\right) = 1$, and
- the integral version of (2.1)

  $$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$$

  holds a.s.

Remark 2.1.3. As one can see, the biggest difference between strong solutions and weak solutions is that weak solutions are allowed to be adapted with other filtrations. Intuitively speaking, a strong solution X as an "output" of a stochastic system is described by the coefficients (b, σ), while the "input" is W and initial condition ξ. In other words, when both W and ξ are given, their specification should determine X in an unambiguous way. However, for weak solution, the filtration $\mathbb{F}$ does not necessarily to be the augmentation of the filtration $\sigma(\xi) \vee \mathcal{F}_t^W, 0 \leq t < \infty$. In other words, the value of X_t is not necessarily given by a measurable function of the path of W and the initial condition ξ.

Remark 2.1.4. It can be easily seen that the existence of a strong solution does imply the existence of a weak solution but the existence of a weak solution does not imply the existence of a strong solution.

2.1.2 Two Notions of Uniqueness

Definition 2.1.5. We call *pathwise uniqueness* holds for SDE (2.1) if whenever two weak solutions (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{F}$ and $(\tilde{X}, W)$, $(\Omega, \mathcal{F}, \mathbb{P})$, $\tilde{\mathbb{F}}$ with same Brownian motion on the same probability space and with the same initial condition, X and $\tilde{X}$ are indistinguishable, i.e.,

$$\mathbb{P}\left(X_t = \tilde{X}_t; \forall 0 \leq t < \infty\right) = 1.$$

Definition 2.1.6. We call *uniqueness in law* holds for SDE (2.1) if for any two weak solutions (X, W), $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{F}$ and $(\tilde{X}, \tilde{W})$, $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, $\tilde{\mathbb{F}}$ with the same initial distribution, the two processes X and $\tilde{X}$ have the same distribution.

Example 2.1.7. In the book Cherny and Engelbert (2004), the authors collected ten examples in *section 1.3*. They provided examples about non-solvable SDEs, no strong solution SDEs, solvable but not uniquely solvable SDEs, etc. Moreover, some of the examples are classical in the SDEs' field such as Tanaka equation, Tsirelson equation, (both equations are classical examples about SDEs with weak solutions but no strong solutions) etc. We highly recommend the readers to have a look at that section and references therein to have a better view in the definitions we introduced above.

2.1.3 Results of Yamada and Watanabe

The following results are introduced in Yamada and Watanabe (1971). Generally speaking, there are two important results.

Proposition 2.1.8. Pathwise uniqueness implies uniqueness in law.

From above proposition, a remarkable corollary can be obtained.

Corollary 2.1.9. Weak existence and pathwise uniqueness imply strong existence.

Since the results are too classical in the SDEs' field, the proofs can be found in many places, and we omit them here.

2.1.4 Summary of the Section

With all the definitions and results introduced, we can have the following relations, which are widely applied and important.

- Strong existence $\Rightarrow$ weak existence
- Pathwise uniqueness $\Rightarrow$ uniqueness in law
- Weak existence + pathwise uniqueness $\iff$ strong existence + uniqueness in law = so called "unique solution"

2.2 Girsanov Theorem

Girsanov theorem was introduced by Girsanov (1960) and Cameron and Martin (1944). We let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be an augmented filtered probability space and a d-dimensional Brownian motion W is defined on it. In particular, we do not require the filtration $\mathbb{F}$ is generated from the Brownian motion W. Let X be a $\mathbb{R}^d$ valued square integrable adapted process and we define

$$\mathcal{E}_t(X) := \exp\left(\sum_{i=1}^{d}\int_0^t X_s^i dW_s^i - \frac{1}{2}\int_0^t |X|^2 ds\right).$$

Remark 2.2.1. Sometimes we write $M. := \int_0^{\cdot} X_s dW_s$ as a continuous local martingale, and we denote

$$\mathcal{E}_t(M) := \exp\left(M_t - \frac{1}{2}[M]_t\right).$$

Theorem 2.2.2. Assume $\mathcal{E}(X)$ is a martingale. Then a process B,

$$B_t^i := W_t^i - \int_0^t X_s^i ds,$$

for each fixed $T \in [0, \infty)$, the process $(B_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion in $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}_T)$.

As one can see, the only condition required is $\mathcal{E}(X)$ is a martingale. Here we introduce three common sufficient conditions that the theorem holds.

- Beneš Condition (Beneš (1971)). Consider a progressively measurable process X

$$X_t(\omega) = b(t, W_\cdot(\omega)); \quad 0 \le t < \infty,$$

where b is a progressively measurable function on $C[0,\infty)^d$, which is a space of continuous mappings from $[0,\infty)$ to $\mathbb{R}^d$. If for each $0 \le T < \infty$ and some constant K_T depending on T, such that

$$|b(t,x)| \le K_T(1 + x_t^*), \quad 0 \le t \le T,$$

where $x_t^* := \max_{0\le s\le t} |x_s|$, then $\mathcal{E}(X)$ is a martingale.

- Novikov Condition (Novikov (1972)). Let X be a square integrable measurable adapted process satisfying

$$\mathbb{E}\exp\left(\frac{1}{2}\int_0^t |X_s|^2 ds\right) < \infty, \quad \forall t \in [0,\infty).$$

Then $\mathcal{E}(X)$ is a martingale. One can easily obtain that Beneš condition would imply Novikov condition; see corollary 5.16, chapter 3, Karatzas and Shreve (1991).

- Kazamaki's Condition (Kazamaki (1977)). One should realize that above condition also implies $\int_0^\cdot X_s dW_s$ is a local martingale. In fact, $\int_0^\cdot X_s dW_s$ is a local martingale is enough to guarantee $\mathcal{E}(X)$ is a martingale.

If we denote $M_\cdot = \int_0^\cdot X_s dW_s$ a continuous local martingale, applying Itô formula with $\mathcal{M} = M - \frac{1}{2}[M]$ and $f(m) = e^m$ we obtain

$$\mathcal{E}_t(M) = 1 + \int_0^t \mathcal{E}_s(M) dM_s.$$

In fact, $\mathcal{E}(M)$ is a martingale if and only if $\mathbb{E}(\mathcal{E}_t(M)) = 1$. The related discussion can be found in Section 3.5.D, Karatzas and Shreve (1991).

One need to notice that we would like to have a single measure $\mathbb{Q}$ defined on $\mathcal{F}_\infty$ such that when $\mathbb{Q}$ is restricted to $\mathcal{F}_T$, it is equivalent to $\mathbb{P}_T$ for all T. However, such a measure does not exist in general. So we introduce the corollary below which is a "more common version" of Girsanov theorem.

Corollary 2.2.3. We consider the filtration $\mathbb{F} = \mathbb{F}^W$ is the augmented filtration generated by the Brownian motion W. If X is a square integrable predictable process, and any one of above conditions holds, then for each fixed $T \in [0,\infty)$ with

$$B_t := W_t - \int_0^t X_s ds$$

the process $(B_t)_{t\in[0,T]}$ is a Brownian motion on $(\Omega, \mathcal{F}, (\mathcal{F}_t^W)_{t\in[0,T]}, \mathbb{Q})$ with

$$\frac{d\mathbb{Q}(A)}{d\mathbb{P}(A)} = \exp\left(\int_0^T X_t dW_t - \frac{1}{2}\int_0^T X_t^2 dt\right),$$

where $A \in \mathcal{F}_T^W$.

Remark 2.2.4. For Theorem 2.2, X is not necessarily adapted to the augmented filtration generated by the Brownian W, which would imply B_t is not necessarily adapted to $\mathcal{F}_t^W$; e.g. Tsirelson's example (Rogers and Williams (2000) page 155).

2.3 BMO Martingales

Another important topic closely related to Girsanov theorem is bounded mean oscillation (BMO) martingales. We call a continuous, uniformly integrable martingale $(M_t, \mathcal{F}_t)$ with $M_0 = 0$ is in the class BMO if

$$\|M\|_{BMO} = \sup_{\tau} \|\mathbb{E}\left([M]_T - [M]_\tau | \mathcal{F}_\tau\right)^{\frac{1}{2}}\|_\infty < \infty$$

where the supremum is taken over all stopping times $\tau \in [0, T]$.

A useful theorem about BMO martingales is as follows

Theorem 2.3.1. (Theorem 2.3, 2.4, Kazamaki (2006)) If $M \in BMO$, then $\mathcal{E}(M)$ is a uniformly integrable martingale. Meanwhile,

$$\sup_{\tau} \left\| \mathbb{E}\left(\left(\frac{\mathcal{E}_\tau(M)}{\mathcal{E}_\infty(M)}\right)^{\frac{1}{p-1}} \middle| \mathcal{F}_\tau \right) \right\|_\infty < \infty,$$

for all $1 < p < \infty$, and all stopping times τ.

Of course, BMO martingales share many other important properties, for instance, BMO martingales can be characterized as the dual of hardy space of $p = 1$; the John-Nirenberg inequality is of fundamental importance for many BMO martingales studies. For more information about BMO martingales, we refer the book Kazamaki (2006).

2.4 Feynman-Kac Formula and Related Topics

In this section we assume all the coefficients are under *good enough* conditions and *deterministic*.

Let's have a look at the following parabolic PDE

$$\partial_t u(t, x) + \mathcal{L}_t u + f(t, x) = 0, \quad u(T, x) = h(x), \tag{2.2}$$

where $\mathcal{L}$ is the second order Dynkin operator

$$\mathcal{L}_t u := b(t, x)\partial_x u + \frac{1}{2}\mathrm{Tr}(\sigma\sigma^\intercal(t, x)\partial_{xx} u).$$

Thus, there exists a unique solution u to PDE (2.2), which can be represented by the Feynman-Kac formula

$$u(t, x) = \mathbb{E}\left(\int_t^T f\left(s, X_s^{(t,x)}\right) ds + h\left(X_T^{(t,x)}\right)\right), \quad (t, x) \in [0, T] \times \mathbb{R}^m, \tag{2.3}$$

where $X_s^{(t,x)}$ is the solution to the following SDE

$$dX_s = b(s, X_s)ds + \sigma(s, X_s)dW_s, \quad X_t = x, \quad s \geq t. \tag{2.4}$$

Notice that when u is smooth enough, then the Feynman-Kac formula (2.3) can be derived from Itô formula. Moreover, if we define

$$Y_t := u(t, X_t), \quad Z_t := \sigma^\intercal(t, X_t)\partial_x u(t, X_t), \quad 0 \leq t \leq T, \tag{2.5}$$

then by applying Itô formula to $u(s, X_s)$ between t and T, with u satisfying PDE (2.2), we have the Markovian BSDE

$$Y_t = h(X_T) + \int_t^T f(s, X_s)ds - \int_t^T Z_s dW_s.$$

Remark 2.4.1. We consider a SDE

$$dX_s = b(s, X_s)ds + \sigma(s, X_s)dW_s, \tag{2.6}$$

with the coefficients are under good enough condition and deterministic. Then based on Itô formula, for any C^2 function (i.e., twice continuously differentiable) we have

$$\phi(X_t) = \phi(X_0) + \int_0^t \mathcal{L}_s\phi(X_s)ds + \int_0^t \partial_x\phi(X_s)\sigma(s, X_s)dW_s.$$

Taking expectations on both sides and denoting by μ_t the distribution of X_t, we can obtain that

$$\langle \phi, \mu_t \rangle = \langle \phi, \mu_0 + \int_0^t \mathcal{L}_s^* \mu_s ds \rangle$$

where $\langle \phi, \mu_t \rangle := \int_{\mathbb{R}^d} \phi(x)\mu(dx)$ and $\mathcal{L}_s^*$ is the adjoint operate of $\mathcal{L}_s$ that is $\langle \mathcal{L}_t\phi, \mu_t \rangle = \langle \phi, \mathcal{L}_t^*\mu_t \rangle$. Since above equation holds for all ϕ, we can deduce that μ_t is a solution of the following equation

$$\frac{d}{dt}\mu_t = \mathcal{L}_t^*\mu_t$$

with initial condition μ_0. Notice that above equation is often known as Fokker–Planck equation. Moreover, if μ_t has a density, that is $\mu_t(dx) = p(t, x)dx$. Then by Itô formula it can be shown that

$$-\partial_t p(t, x) + \frac{1}{2}\partial_{xx}(\sigma^2(t, x)p) - \partial_x(b(t, x)p) = 0.$$

In this sense, we could borrow arguments from Aronson (1967) and results thereafter that provided certain boundedness results, which are known as Aronson-like bounds, for the probability density function p.

Remark 2.4.2. When (2.2) with $f = 0$, it is often known as Kolmogorov's backward equation. In particular, we write $u(t, x) = \int_{\mathbb{R}^m} h(y)p(y, s, x, t)dy$. Notice that we replace T with s here for better looking and one can consider s as a time variable. It can be shown that $\mu(A, s, x, t) = \int_A p(y, s, x, t)dy$ i.e. $\mu(A, s, x, t) = \mathbb{P}(X_s \in A | X_t = x)$ as a transition function defines a Markov process X such that it satisfies SDE (2.6). Moreover, for any bounded function $\phi(t, x) \in C^{1,2}$, the following holds

$$\int_{\mathbb{R}^m} \phi(s, y)\mu(dy, s, x, t) - \phi(t, x) = \int_t^s \int_{\mathbb{R}^m} (\partial_u + \mathcal{L}_u)\phi(u, y)\mu(dy, u, x, t)du$$

for all $t \in [0, s]$, $x \in \mathbb{R}^m$. Detailed discussion can be found in section 5.8, Klebaner (2012).

Remark 2.4.3. Notice that PDE (2.2) can be generalized to

$$\partial_t u + \mathcal{L}_t u + f(t, x, u, \sigma\partial_x u) = 0, \quad u(T, x) = h(x). \tag{2.7}$$

In this case, the corresponding BSDE would be like

$$Y_t = h(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \tag{2.8}$$

with the solution (Y, Z) of the same form to (2.5). An issue arises from the fact that in general, the function u obtained from BSDE does not smooth, and usually, people deal with this problem by the notion of viscosity solution. For instance, Pardoux and Peng (1992) proved that the function u obtained from the globally Lipschitz BSDE (2.8) is a unique viscosity solution to the PDE (2.7). However, as one can see from chapter 3, our thesis demonstrates the relation between the solutions of (F)BSDEs and PDEs in another perspective.

2.5 Generalized Derivatives and Related Topics

For a function $u \in C^{1,2}$, we denote $\mathcal{L} := \partial_t + \frac{1}{2}\sum_{i,j=1}^{m}(\sigma\sigma^\intercal)_{ij} \cdot \partial^2_{x_i x_j}$ and its gradient $\nabla := (\partial_{x_1}, \cdots, \partial_{x_m})$. When u (2.5) is a solution of PDE (2.2), it not necessarily belongs to $C^{1,2}$, especially when the coefficients are not smooth enough. Thus, when $u \notin C^{1,2}$, we need the following definitions such that generalized derivatives can be well defined. This section's arguments are based on the results from Krylov (1980) and Chitashvili and Mania (1996).

Definition 2.5.1. Let $\mu(ds, dy) := p(s, y)dsdy$, where p is the transition density corresponding to SDE $dX_t = \sigma(t, X_t)dW_t$ with $X_0 = x \in \mathbb{R}^m$. We say u belongs to $V^{\mathcal{L}}_{\mu}(loc)$, if there exists a sequence of functions $(u_n)_{n\geq 1} \subset C^{1,2}$, a sequence of bounded measurable domains $D_1 \subset D_2 \subset \cdots$ with $(0, x) \in D_1$ and $\cup_{n\in\mathbb{N}} D_n = [0, T] \times \mathbb{R}^m$, and a measurable locally μ-integrable function $\mathcal{L}u$ such that

- $\tau_k := \{t > 0 : (t, X_t) \notin D_k\}$ are stopping times with $\tau_n \nearrow T$.
- For each $k \geq 1$,

$$\sup_{s\leq\tau_k} |u^n(s, X_s) - u(s, X_s)| \xrightarrow{n\to\infty} 0 \qquad \text{a.s.}$$

$$\iint_{D_k} |\mathcal{L}u^n(s, X_s) - \mathcal{L}u(s, X_s)|\mu(ds, dx) \xrightarrow{n\to\infty} 0.$$

Then, we define the $\mathcal{L}$-derivative of u by $\mathcal{L}u$. Moreover, if $u \in V^{\mathcal{L}}_{\mu}(loc)$, then there exists $\nabla u(t, x)$ such that

$$\iint_{D_k} |\nabla u^n(s, X_s) - \nabla u(s, X_s)|^2 \mu(ds, dx) \xrightarrow{n\to\infty} 0.$$

We define ∇u to be the generalized gradient of u.

Definition 2.5.2. Similarly, we define $W^{1,2}_p(D)$. We let D an open set $D \subset \mathbb{R}^{1+m}$ and we call $C^{1,2}(\overline{D})$ the set of functions $u(t, x)$ such that u is continuously differentiable with respect to t and twice continuously differentiable with respect to x in D, and $u(t, x)$ as well as all these derivatives of $u(t, x)$ have extensions continuous in D.

We call a function $u \in W^{1,2}_p(D)$ if there exists a sequence of functions $u^n \in C^{1,2}(\overline{D})$, such that $\forall (t, x) \in \overline{D}$

$$\sup_{(t,x)\in\overline{D}} |u^n(t, x) - u(t, x)| \xrightarrow{n\to\infty} 0$$

$$\|u^n(t, x) - u^m(t, x)\| \xrightarrow{n,m\to\infty} 0$$

where $\|u\| := \sup_{(t,x)\in\overline{D}} |u(t, x)| + \|\partial_t u\|_{L^p} + \|\partial_x u\|_{L^p} + \|\partial^2_{xx} u\|_{L^p}$. Notice that the second inequality implies the existence of measurable functions u_t, u_i, u_{ij} to which $u^n_t, u^n_{x_i}, u^n_{x_i x_j}$ (partial derivatives) converge in $\mathbb{L}^p(D)$. Namely assuming $\phi \in C^\infty_0(D)$, (a smooth function with compact support), integration by parts, and by denoting $D = D_t \times D_x$, where D_t is the domain of variable t, and D_x is the domain of variable x, we have

$$\int_{D_x} \phi u^n_{x_i} dx = -\int_{D_x} \phi_{x_i} u^n dx.$$

Let $n \to \infty$,

$$\int_{D_x} \phi u_i dx = -\int_{D_x} \phi_{x_i} u dx.$$

While the others can be obtained similarly. In particular, we denote $W_p^{1,2}(loc)$ the class of functions defined on $[0,T] \times \mathbb{R}^m$ which belongs to $W_p^{1,2}(D)$ for every bounded open domain $D \in [0,T] \times \mathbb{R}^m$.

The following theorem is crucial for our result.

Theorem 2.5.3. (Theorem 1, Chitashvili and Mania (1996)) Consider SDE (2.6), with b is bounded, and σ is non-degenerate and $\sigma\sigma^\intercal$ is continuous. Then the process $(f(t, X_t))_{t\in[0,T]}$ is an Itô process if and only if $f \in V_\mu^{\mathcal{L}}(loc)$ and it admits the decomposition

$$f(t, X_t) = f(0, X_0) + \int_0^t \nabla f(s, X_s) dX_s + \int_0^t \mathcal{L} f(s, X_s) ds.$$

Our results could prove the well-posedness of BSDE (2.8) and provide the existence of the decoupling field u such that $Y_t = u(t, X_t)$. Since Y is an Itô process, then we can deduce that u is a unique solution of PDE (2.7) directly from above theorem.

Particularly, in the last chapter, we provided that under a set of stricter conditions, the unique solution u is in the space $W_p^{1,2}(D)$. Here we want to share some important properties about space $W_p^{1,2}(D)$.

Remark 2.5.4. $W_p^{1,2}(D)$ is not only a completion of space $C^{1,2}(D)$ but also equivalent to the usual Sobolev space for continuous functions with a sufficient regular domain. (See Krylov (1980).)

Corollary 2.5.5. Given b is bounded, and σ is non-degenerate and $\sigma\sigma^\intercal$ is continuous. For all $p \geq m+1$,

$$W_p^{1,2}(D) \subset V_\mu^{\mathcal{L}}(D),\ W_p^{1,2}(loc) \subset V_\mu^{\mathcal{L}}(loc)$$

for every bounded domain $D \in \mathcal{B}(\mathbb{R}^{1+m})$.

Proof. Without loss of generality, we consider the initial condition $x = 0$. Since the probability density of X satisfies a parabolic equation introduced in Remark 2.4.1, then from Aronson (1967) we know that

$$p(s, y) \leq C_1 s^{\frac{-m}{2}} \exp\left(-C_2 \frac{y^2}{s}\right)$$

where C_1, C_2 are two positive constants depending on the bounded value of b, σ and terminal time T. Thus by Hölder's inequality, for each $p \geq m+1$, we have

$$\int\int_D |u(s,x)| p(s,x) ds dx < C \left(\int\int_D |u(s,x)|^p ds dx\right)^{\frac{1}{p}}, \tag{2.9}$$

for some constant C. For a generic $u \in W_p^{1,2}(D)$, by Definition 2.5.1, Definition 2.5.2, and denoting $D = D_t \times D_x$, where D_t is the domain of variable t, and D_x is the domain of variable x, we have that there exists a sequence of functions $u^n \in C^{1,2}(\overline{D})$ such that

$$\sup_{(t,x)\in\overline{D}} |u^n(t,x) - u(t,x)| \xrightarrow{n\to\infty} 0 \implies \sup_{t\in D_t} |u^n(t, X_t) - u(t, X_t)| \xrightarrow{n\to\infty} 0$$

where $X_t \in D_x$. Moreover, according to (2.9), Definition 2.5.1, and Definition 2.5.2, we have for each $X_t \in D_x$, there exists a constant C, such that

$$\iint_D |\mathcal{L}u^n(t,X_t) - \mathcal{L}u^m(t,X_t)|\mu(dt,dx) \leq C\|u^n(t,X_t) - u^m(t,X_t)\| \xrightarrow{n,m\to\infty} 0,$$

where norm $\|\|$ was defined in Definition 2.5.2. While above formula also implies the existence of $\mathcal{L}u$. Therefore, we have $u \in V_\mu^{\mathcal{L}}(D)$. Similar arguments can be made for the localized case. □

2.6 Previous Results about SDEs and BSDEs

Since our results heavily depend on the previous results of SDEs and BSDEs', for the readers convenience, in the last section we will present some detailed results that we applied. As usual, we consider the coefficients are deterministic and the probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space on which is defined an $\mathbb{R}^d$-dimensional Brownian motion W, such that its augmented filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ is generated by W. For a terminal time $T \in [0,\infty)$, we consider SDEs in this section with the form

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t \quad X_0 = x \in \mathbb{R}^m, \tag{2.10}$$

where $(b,\sigma) : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^{m\times d}$, and BSDEs

$$Y_t = \xi - \int_t^T f(s,Y_s,Z_s)ds + \int_t^T Z_s dW_s, \tag{2.11}$$

where ξ is a square integrable $\mathcal{F}_T$-adapted random variable and $f : [0,T]\times\mathbb{R}^n\times\mathbb{R}^{n\times d} \to \mathbb{R}^n$. Notice that sometimes we would assume $\xi := h(X_T)$ with $h : \mathbb{R}^m \to \mathbb{R}^n$ and f depends on x. In this sense, equations (2.10) and (2.11) would form a so-called decoupled FBSDE. Please notice that for the decoupled FBSDEs, since in most cases X can be obtained from the SDEs, usually the BSDEs' results can be easily generalized to the decoupled FBSDEs'. Particularly, since we are considering the Markovian cases, we are interested in the existence of *decoupling field* that is a deterministic measurable function u such that $Y_t = u(t,X_t)$.

2.6.1 SDEs' Results

In Veretennikov (1980), the author provided if b is bounded, and σ is Lipschitz and uniformly non-degenerate that is, there exists a constant $\varepsilon > 0$ such that

$$\varepsilon^{-1}|x'|^2 \leq (x')^\intercal(\sigma\sigma^\intercal)(t,x)x' \leq \varepsilon|x'|^2$$

for all $x' \in \mathbb{R}^m$ and $(t,x) \in [0,T]\times\mathbb{R}^m$, then (2.10) is uniquely solvable.

Gyongy and Martinez (2001) provided the following conditions:

- $b \in \mathbb{L}^{2m+2}_{\mathrm{loc}}([0,T]\times\mathbb{R}^m)$, where subscript loc stands for locally integrable, and there exist a non-negative function $F \in \mathbb{L}^{m+1}([0,T]\times\mathbb{R}^m)$ and $K \geq 0$ such that
$$|b(t,x)| \leq K + F(t,x).$$
- σ is uniformly non-degenerate and locally Lipschitz, which is for every $R > 0$ there is a constant L_R such that
$$|\sigma(t,x) - \sigma(t,y)| \leq L_R|x-y|$$
for all $t \in [0,T]$, $x,y \in \mathbb{R}^m$ with $|x| \leq R$ and $|y| \leq R$.

If above conditions are satisfied, then SDE (2.10) is uniquely solvable.

When $m = d = 1$, Le Gall (1984) proved the unique solvability of SDE (2.10). They introduced the following two conditions:

- There exists a strictly increasing function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\int_{\mathbb{R}^+} \frac{du}{\theta(u)} = \infty$ and $(\sigma(t,x) - \sigma(t,y))^2 \le \theta(|x-y|)$ for all t, x, y.
- σ is strictly positive and there exists a strictly increasing function $\theta' : \mathbb{R} \to \mathbb{R}$ such that $(\sigma(t,x) - \sigma(t,y))^2 \le |\theta'(x) - \theta'(y)|$ for all t, x, y.

When σ, b are bounded functions, σ is non-degenerate and satisfies one of above conditions, then SDE (2.10) has a unique solution.

The last result about SDEs' was introduced by Menoukeu-Pamen and Mohammed (2019). They provided a well-posedness result for SDE (2.10) when σ is a constant. (2.10) under such case is uniquely solvable if there exists a constant C, such that $|b(t,x)| \le C(1 + |x|)$ for all $(t,x) \in [0,T] \times \mathbb{R}^m$.

2.6.2 BSDEs' Results

El Karoui et al. (1997) introduced the first general result for BSDEs. It proved that BSDE (2.11) is uniquely solvable if ξ is square integrable and $\mathcal{F}_T$-measurable, and f is uniformly Lipschitz. Sometimes people would call the (ξ, f) standard parameter if they satisfy these conditions. In particular, the authors provided the well-known comparison theorem for one-dimensional BSDEs with standard parameters. Let (ξ^1, f^1) and (ξ^2, f^2) be two standard parameters of BSDEs, and (Y^1, Z^1), (Y^2, Z^2) be the associated solutions. If

$$\xi^1 \ge \xi^2, \quad \text{and} \quad f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2) \ge 0,$$

then we have for any $t \in [0,T]$, $Y_t^1 \ge Y_t^2$.

Hamadene et al. (1997) considered decoupled FBSDEs and assume b and σ are uniformly Lipschitz and under linear growth. (The conditions imply the well-posedness of SDE (2.10).) Particularly, for $p \in [1, \infty)$, $(x, y, z) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$, and some constant C, they assume $|g(x)| \le C(1 + |x|^p)$, $|f(t,x,y,z)| \le C(1 + |x|^p + |y| + |z|)$, and $\forall (s,x) \in [0,T] \times \mathbb{R}^m$, $f(s,x,\cdot,\cdot)$ is continuous. If all these assumptions are satisfied, then the decoupled FBSDE has a solution. Particularly, inspired from the Feynman-Kac formula we introduced in Section 2.3, they showed that there exist measurable and deterministic functions u, d such that $Y_t = u(t, X_t)$ and $Z_t = d(t, X_t)$.

As we mentioned in the introduction, Kobylanski (2000) was the first paper proved the well-posedness of quadratic BSDE (2.11). She considered the BSDE is one-dimensional and $f(t,y,z) = a_0(t,y,z)y + f_0(t,y,z)$. Let $a, b, c \in \mathbb{R}$ and α be a continuous increasing function. If for all $(t,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d$,

$$a \le a_0(t,y,z) \le b, \quad \text{and} \quad |f_0(t,y,z)| \le c + \alpha(|y|)|z|^2$$

then BSDE (2.11) has a solution. Since the BSDE is one-dimensional, the classical comparison theorem introduced in El Karoui et al. (1997) helps the author to obtain the boundedness of solution Y, where the boundary depends on a, b, c, and the uniqueness of BSDE (2.11) under a stricter set of conditions. Particularly, due to her novel approach, the way of obtaining the decoupling field is slightly different from the usual ones. Since Y is bounded, then one can assume $|f_0(t,y,z)| \le C(1 + |z|^2)$ for some $C \in \mathbb{R}^+$. An exponential change is given by $y = e^{2CY}$. By showing y is a solution of another set of parameters' decoupled FBSDE with a decoupling field u, one can obtain $Y_t = \frac{\ln u(t,X_t)}{2C}$. (The decoupled FBSDE was discussed in section 3 of the paper.)

Recently, another paper provided results for multidimensional quadratic BSDEs. Hu and Tang (2016) studied BSDE (2.11) of the following form

$$Y_t^i = \xi^i - \int_t^T f^i(s, Z_s^i) + h^i(s, Y_s, Z_s)ds + \int_t^T Z_s^i dW_s, \quad i = 1, \cdots, n.$$

They provided the following conditions:

- For all $f^i : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $i = 1, \cdots, n$, $z, z_1, z_2 \in \mathbb{R}^d$, and some constant C, they have the following quadratic growth and locally Lipschitz continuity in the last variable

$$|f^i(s,z)| \leq C|z|^2$$
$$|f^i(s,z_1) - f^i(s,z_2)| \leq C(1 + |z_1| + |z_2|)|z_1 - z_2|.$$

- $\xi = (\xi^1, \cdots, \xi^n)^\intercal$ is uniformly bounded.

- Denote $h = (h^1, \cdots, h^n)^\intercal$. For $(s, y, z) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n\times d}$ and $(y', z') \in \mathbb{R}^n \times \mathbb{R}^{n\times d}$, h satisfies

$$|h(s,y,z)| \leq C(1 + |y|), \quad |h(s,y,z) - h(s,y',z')| \leq C(|y - y'| + |z - z'|).$$

If all above conditions are satisfied, then BSDE (2.11) would have a unique solution. Particularly, they proved that if (Y, Z) is a solution, then Y is bounded and $Z \cdot W$ is a BMO martingale.

The last result we want to share was introduced by Mu and Wu (2015). They focus on the BSDE of the form

$$Y_t^i = h^i(X_T^{0,a}) - \int_t^T f^i(s, X_s^{0,a}, Y_s^1, \cdots, Y_s^n, Z_s^1, \cdots, Z_s^n)ds + \int_t^T Z_s^i dW_s,$$

with the forward process satisfies

$$X_s^{t,x} = x + \int_t^s \sigma(u, X_u^{t,x})dW_u \quad s \in [t,T],$$
$$X_s^{t,x} = x \quad s \in [0,t].$$

Particularly, they assume for each $(t, x, y^1, \cdots, y^n, z^1, \cdots, z^n) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times m}$ there exist a constant C and $r \geq 0$ such that for each $i = 1, \cdots, n$

$$|f^i(t, x, y^1, \cdots, y^n, z^1, \cdots, z^n)| \leq C(1 + |x|)|z^i| + C(1 + |x|^r + |y^i|),$$

and for each $(t,x) \in [0,T] \times \mathbb{R}^m$, $f^i(t, x, y^1, \cdots, y^n, z^1, \cdots, z^n)$ is continuous. Moreover, they assume terminal condition h is under polynomial growth with respect to x that is $|h^i(t,x)| \leq C(1 + |x|^r)$, with $r \geq 0$. Meanwhile, they assume σ is non-degenerate and Lipschitz. If all these assumptions are satisfied, they proved the existence of a solution for the decoupled FBSDE along with the existence of the decoupling field.

Chapter 3

Discontinuous Coefficients FBSDEs with Their Applications and Semilinear PDEs

In the last chapter we will present our main results of the thesis. As we introduced in chapter 1, we are interested in the solvability of FBSDE (1.6) with all the coefficients are discontinuous with respect to the forward process. Particularly, we let $\tilde{g}$ be of the form $\tilde{g}(t, x, y, z) = b(t, x) + \sigma(t, x)g(t, x, y, z)$ and thus we have

$$\begin{aligned} dX_t &= (b(t, X_t) + \sigma(t, X_t)g(t, X_t, Y_t, Z_t))dt + \sigma(t, X_t)dW_t; \quad X_0 = x \\ dY_t &= -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t; \quad Y_T = h(X_T). \end{aligned}$$

This chapter is organized as follows: in the first section we will introduce our main results about above FBSDEs (Theorem 3.1.3) and their corresponding PDEs' (Corollary 3.1.6). Followed by the next section we will demonstrate our approach and prove the results. At last we will provide two applications that are closely related to discontinuous coefficients FBSDEs.

3.1 Existence and Uniqueness Results

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space with an n-dimensional Brownian motion W and its augmented filtration $\mathbb{F}$ generated by W. For a function $u \in C^{0,1}$, we denote $\nabla := (\partial_{x_1}, \cdots, \partial_{x_m})$. Let

$$\begin{aligned} (b, \sigma) &: [0, T] \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^{m\times n} \\ (f, g) &: [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \times \mathbb{R}^{d\times n} \to \mathbb{R}^d \times \mathbb{R}^n \\ h &: \mathbb{R}^m \to \mathbb{R}^d. \end{aligned}$$

be (jointly) measurable functions. Unless otherwise stated, we assume the following conditions:

- σ is uniformly nondegenerate, that is, there exists a constant $\varepsilon > 0$ such that

  $$\varepsilon^{-1}|x'|^2 \le (x')^\intercal(\sigma\sigma^\intercal)(t, x)x' \le \varepsilon|x'|^2$$

  for all $x' \in \mathbb{R}^m$ and $(t, x) \in [0, T] \times \mathbb{R}^m$.

- There exists a positive constant κ such that,

$$|b(t,0)| + \sup_{|x-x'|\leq 1} |b(t,x) - b(t,x')| \leq \kappa$$

for all $t \in [0,T], x, x' \in \mathbb{R}^m$.

Remark 3.1.1. Under the non-degeneracy assumption on σ, we have

$$b(t,x) + \sigma(t,x)g(t,x,y,z) = \tilde{b}(t,x) + \sigma(t,x)\tilde{g}(t,x,y,z),$$

where $\tilde{b} = b - k$ and $\tilde{g} = \sigma^\intercal(\sigma\sigma^\intercal)^{-1}k + g$. This adds flexibility to the conditions described below.

Remark 3.1.2. Note that $b(t,x)$ can exhibit linear growth in x. For example, let $b(t,x) = b_0(t,x) + b_1(t,x)$ where $\sup_{t\in[0,T]} |b_0(t,0) + b_1(t,0)| = 1$, $b_0(t,x)$ is Hölder continuous in x, and $b_1(t,x)$ is bounded.

We use the following short-hand notations for different conditions on the coefficients, where $\overline{f}(t,x,y,z) := f(t,x,y,z) + zg(t,x,y,z)$, C and r are nonnegative constants, $\theta : \mathbb{R} \to \mathbb{R}_+$ is a strictly increasing function, and $\rho_r : \mathbb{R}_+ \to \mathbb{R}_+$ is a nondecreasing function with $\rho_r \equiv 0$ for $r > 0$:

(F1) $|b(t,x)| \leq C$ and $\sigma(t,x)$ is locally Lipschitz with respect to x.

(F2) $|b(t,x)| \leq C$, $m = n = 1$, and either

(i) $\int \frac{du}{\theta(u)} = \infty$ and $|\sigma(t,x) - \sigma(t,y)|^2 \leq \theta(|x-y|)$, or

(ii) θ is bounded and $|\sigma(t,x) - \sigma(t,y)|^2 \leq |\theta(x) - \theta(y)|$.

(F3) $\sigma(t,x)$ is a constant matrix.

(B1) $|h(x)| \leq C(1 + |x|^r)$, $\overline{f}(t,x,y,z)$ is continuous in (y,z), and

$$\begin{aligned}
|f(t,x,y,z)| &\leq C(1 + |x|^r + |y| + |z|) \\
|g(t,x,y,z)| &\leq C(1 + \rho_r(|y|)).
\end{aligned}$$

(B2) $|h(x)| \leq C$, $\overline{f}^i(t,x,y,z) = \tilde{f}^i(t,x,z^i) + \hat{f}^i(t,x,y,z)$ such that

$$\begin{aligned}
|\hat{f}(t,x,y,z)| &\leq C(1 + |y|) \\
|\tilde{f}(t,x,z)| &\leq C|z|^2 \\
|g(t,x,y,z)| &\leq C(1 + \rho_r(|y|)) \\
|\hat{f}(s,x,y,z) - \hat{f}(s,x,y',z')| &\leq C(|y-y'| + |z-z'|), \forall y, y' \in \mathbb{R}^d, z, z' \in \mathbb{R}^{d\times n} \\
|\tilde{f}(s,x,z_1) - \tilde{f}(s,x,z_2)| &\leq C(1 + |z_1| + |z_2|)|z_1 - z_2|, \forall z_1, z_2 \in \mathbb{R}^{d\times n}.
\end{aligned}$$

(B3) $d = 1$, $|h(x)| \leq C$, $\overline{f}(t,x,y,z)$ is continuous with respect to (y,z), and

$$\begin{aligned}
|f(t,x,y,z)| &\leq C(1 + |y| + |z|^2) \\
|g(t,x,y,z)| &\leq C(1 + \rho_r(|y|)).
\end{aligned}$$

(B4) $|h(x)| \leq C(1 + |x|^r)$, $\overline{f}(t,x,y,z)$ is continuous in (y,z), and

$$\begin{aligned}
|f^i(t,x,y,z)| &\leq C(1 + |x|^r + |y^i|) \text{ for all } i = 1, 2, ..., d \\
|g(t,x,y,z)| &\leq C(1 + |x| + \rho_r(|y|)).
\end{aligned}$$

(U1) $\overline{f}(t,x,y,z)$ is globally Lipschitz continuous with respect to (y,z), or

(U2) $d=1$, $|h(x)|\le C$, $\overline{f}(t,x,y,z)$ is differentiable with respect to (y,z), and for any $M,\varepsilon>0$, there exist $l_M, l_\varepsilon \in L^1([0,T];\mathbb{R}_+), k_M \in L^2([0,T];\mathbb{R}_+)$, and $C_M>0$ such that $\overline{f}$ satisfies

$$\begin{aligned}|\overline{f}(t,x,y,z)| &\le l_M(t)+C_M|z|^2\\ |\partial_z\overline{f}(t,x,y,z)| &\le k_M(t)+C_M|z|\\ |\partial_y\overline{f}(t,x,y,z)| &\le l_\varepsilon(t)+\varepsilon|z|^2\end{aligned}$$

for all $(t,x,y,z)\in[0,T]\times\mathbb{R}^m\times[-M,M]\times\mathbb{R}^{1\times n}$.

Theorem 3.1.3. Assume that there exist nonnegative constants $C,r>0$, a strictly increasing function $\theta:\mathbb{R}\to\mathbb{R}_+$, and a nondecreasing function $\rho_r:\mathbb{R}_+\to\mathbb{R}_+$ with $\rho_r\equiv 0$ for $r>0$ that satisfies either of the following conditions:

- one of (F1), (F2), (F3) and one of (B1), (B2), (B3) hold for any $(t,x,y,z)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^d\times\mathbb{R}^{d\times n}$.
- (F3) and (B4) hold for any $(t,x,y,z)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^d\times\mathbb{R}^{d\times n}$.

Then, FBSDE

$$\begin{aligned}dX_t &= (b(t,X_t)+\sigma(t,X_t)g(t,X_t,Y_t,Z_t))dt+\sigma(t,X_t)dW_t; && X_0=x\\ dY_t &= -f(t,X_t,Y_t,Z_t)dt+Z_tdW_t; && Y_T=h(X_T)\end{aligned} \tag{3.1}$$

has a strong solution in $\mathbb{H}^2(\mathbb{R}^m)\times\mathbb{H}^2(\mathbb{R}^d)\times\mathbb{H}^2(\mathbb{R}^{d\times n})$. In particular, if $r=0$, then the FBSDE has a strong solution (X,Y,Z) such that $\operatorname{ess\,sup}_{\omega\in\Omega}\sup_{t\in[0,T]}|Y_t(\omega)|<\infty$. In addition, if either (U1), (U2), or (B2) holds, then the solution is unique.

Proof. The proof is given in Section 3.2. □

Under the assumption of Theorem 3.1.3, we have measurable functions $u:[0,T]\times\mathbb{R}^m\to\mathbb{R}^d$ and $d:[0,T]\times\mathbb{R}^m\to\mathbb{R}^{d\times n}$ such that $Y_t=u(t,X_t), Z_t=d(t,X_t)$ for almost every $(t,\omega)\in[0,T]\times\Omega$. To state that u is a solution of a parabolic PDE, let us define $V^{\mathcal{L}}_\mu(loc)$, the class of $\mathcal{L}$-differentiable functions.

Definition 3.1.4 (Chitashvili and Mania (1996))**.** Let $\mu(ds,dy):=p(0,x,s,y)dsdy$, where p is the transition density corresponding to SDE $dX_t=\sigma(t,X_t)dW_t$. Further, for a function $f\in C^{1,2}$, we define

$$\mathcal{L}f:=\partial_t f+\frac{1}{2}\sum_{i,j=1}^m(\sigma\sigma^\intercal)_{ij}(t,x)\partial^2_{x_ix_j}f.$$

We say u belongs to $V^{\mathcal{L}}_\mu(loc)$, if there exists a sequence of functions $(u_n)_{n\ge1}\subset C^{1,2}$, a sequence of bounded measurable domains $D_1\subset D_2\subset\cdots$ with $(0,x)\in D_1$ and $\cup_{n\in\mathbb{N}}D_n=[0,T]\times\mathbb{R}^m$, and a measurable locally μ-integrable function $\mathcal{L}u$ such that

- $\tau_k:=\{t>0:(t,X_t)\notin D_k\}$ are stopping times with $\tau_n\nearrow T$.
- For each $k\ge1$,

$$\begin{aligned}\sup_{s\le\tau_k}|u^n(s,X_s)-u(s,X_s)| &\xrightarrow{n\to\infty}0 \qquad \text{a.s.}\\ \iint_{D_k}|\mathcal{L}u^n(s,X_s)-\mathcal{L}u(s,X_s)|\mu(ds,dx) &\xrightarrow{n\to\infty}0.\end{aligned}$$

Then, we define the $\mathcal{L}$-derivative of u by $\mathcal{L}u$. Moreover, if $u \in V_\mu^{\mathcal{L}}(loc)$, then there exists $\nabla u(t,x)$ such that

$$\iint_{D_k} |\nabla u^n(s,X_s) - \nabla u(s,X_s)|^2 \mu(ds,dx) \xrightarrow{n\to\infty} 0.$$

We define ∇u to be the generalized gradient of u.

For an open set $D \subset \mathbb{R}^{1+m}$, let $W_p^{1,2}(D)$ be the completion of $C^{1,2}(D)$ with respect to the norm

$$\|u\| := \sup_{(t,x)\in\overline{D}} |u(t,x)| + \|\partial_t u\|_{L^p} + \|\partial_x u\|_{L^p} + \|\partial^2_{xx} u\|_{L^p}.$$

We define $W_p^2(D)$ similarly. Note that $W_p^{1,2}(D)$ is equivalent to the usual Sobolev space for continuous u if D has smooth boundary and $p \geq m+1$: see p47 of Krylov (1980).

Proposition 3.1.5. For $p \geq m+1$,

$$W_p^{1,2}(D) \subset V_\mu^{\mathcal{L}}(loc)$$

for any bounded measurable domain $D \subset \mathbb{R}^m$.

Proof. See Corollary 2.5.5. □

We have the following corollary.

Corollary 3.1.6. Assume the existence conditions in Theorem 3.1.3. In addition, we assume that $\sigma\sigma^\intercal$ is continuous. Then, there exists $u \in V_\mu^{\mathcal{L}}(loc)$ that satisfies

$$\mathcal{L}u(t,x) + \nabla u(t,x)(b(t,x) + \sigma(t,x)g(t,x,u,\nabla u\sigma)) + f(t,x,u,\nabla u\sigma) = 0; \qquad u(T,x) = h(x). \tag{3.2}$$

If the uniqueness condition in Theorem 3.1.3 holds, then there is a unique $u \in V_\mu^{\mathcal{L}}(loc)$ satisfying (3.2). In addition, assume the following conditions:

$$\begin{aligned} &(\sigma\sigma^\intercal)(t,x) \text{ is uniformly continuous with respect to } x \text{ for each } t \in [0,T] \\ &b(t,x) \text{ is unifomly bounded in } (t,x) \\ &\int_{[0,T]\times\mathbb{R}^m} \sup_{z\in\mathbb{R}^{d\times n}} (|f(t,x,y,z)|^p)\, dtdx \leq C(1+\rho_r(|y|)) \text{ and } |g(t,x,y,z)| \leq C(1+\rho_r(|y|)) \\ &h \in W_p^2(\mathbb{R}^m) \end{aligned} \tag{3.3}$$

then $u \in W_p^{1,2}([0,T)\times\mathbb{R}^m)$.

Proof. The existence of $u \in V_\mu^{\mathcal{L}}(loc)$ satisfying (3.2) is an immediate consequence of Theorem 1 in Chitashvili and Mania (1996).

Assume conditions (3.3). When $r=0$, conditions (B1)–(B4) implies the boundedness of u according to the same argument in the proof of Proposition 3.2.7. On the other hand, if $r>0$, then $\rho_r \equiv 0$. Therefore, without losing generality, we can assume that $|g(t,x,y,z)| \leq C$ and

$$\|f(\cdot,\cdot,u,\nabla u\sigma)\|_{L^p} \leq \left(\int_{[0,T]\times\mathbb{R}^m} \sup_{y,z\in\mathbb{R}^d\times\mathbb{R}^{d\times n}} (|f(t,x,y,z)|^p) dtdx \right)^{1/p} < \infty.$$

As we have a measurable function u and ∇u, let us define

$$L := \mathcal{L} + \sum_{i=1}^{m} (b + \sigma g)^i \partial_{x_i}$$
$$F(t,x) := f(t,x,u(t,x),(\nabla u \sigma)(t,x)) - Lh(x).$$

Then, $\tilde{u}(t,x) := u(t,x) - h(x)$ solves the following PDE

$$L\tilde{u} = F; \qquad \tilde{u}(T,x) = 0.$$

Note that the above PDE is linear parabolic with measurable $F \in L^p([0,T] \times \mathbb{R}^m)$, $b + \sigma g$ is bounded, and $(\sigma\sigma^\intercal)(t,\cdot) \in VMO(\mathbb{R}^m)$. (One can find the definition of VMO space at the beginning of section 2 in Krylov (2007).) Therefore, it satisfies the condition of Theorem 2.1 of Krylov (2007), and $\tilde{u} \in W_p^{1,2}([0,T) \times \mathbb{R}^m)$ is a unique solution. As a result, $u = \tilde{u} + h$ is the unique solution of (3.2) and $u \in W_p^{1,2}([0,T) \times \mathbb{R}^m)$. □

3.2 Proof of Theorem 3.1.3

3.2.1 Measure Change of FBSDE

In this subsection, we provide sufficient conditions that guarantee the existence of a strong solution under the Girsanov transform. We neither assume the non-degeneracy of σ nor the boundedness of $|b(t,0)| + \sup_{|x-x'|\leq 1} |b(t,x) - b(t,x')|$. Instead, we assume the following conditions:

(H1) The SDE

$$dF_t = b(t,F_t)dt + \sigma(t,F_t)dW_t \tag{3.4}$$

has a unique strong solution.

(H2) For the strong solution F obtained in (H1), there exist Borel measurable functions $(u,d) : [0,T] \times \mathbb{R}^m \to \mathbb{R}^d \times \mathbb{R}^{d\times n}$ such that $U_t = u(t,F_t)$ and $V_t = d(t,F_t)$ is a strong solution of BSDE

$$dU_t = -f(t,F_t,U_t,V_t) - V_t g(t,F_t,U_t,V_t)dt + V_t dW_t; \qquad U_T = h(F_T). \tag{3.5}$$

(H3) For (F,U,V) in (H1) and (H2), the process

$$\mathcal{E}\left(\int_0^{\cdot} g(s,F_s,U_s,V_s)^\intercal dW_s\right)$$

is a martingale on $[0,T]$.

(H4) For u,d in (H3), the forward SDE

$$d\tilde{F}_t = \left(b(t,\tilde{F}_t) + \sigma(t,\tilde{F}_t) g(t,\tilde{F}_t,u(t,\tilde{F}_t),d(t,\tilde{F}_t))\right) dt + \sigma(t,\tilde{F}_t)dW_t; \qquad \tilde{F}_0 = x$$

has a (pathwise) unique strong solution $\tilde{F}$.

Lemma 3.2.1. Assume (H1)–(H4). Then, the FBSDE

$$dX_t = (b(t, X_t) + \sigma(t, X_t)g(t, X_t, Y_t, Z_t))\,dt + \sigma(t, X_t)dW_t; \qquad X_0 = x \tag{3.6}$$

$$dY_t = -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t; \qquad Y_T = h(X_T) \tag{3.7}$$

has a strong solution (X, Y, Z) that satisfies (H3) and $(Y, Z) = (u(t, X_t), d(t, X_t))$. In addition, if BSDE (3.5) has a unique strong solution, then (3.6)–(3.7) has a unique strong solution (X, Y, Z) such that $\mathcal{E}\left(-\int_0^\cdot g(s, X_s, Y_s, Z_s)^\intercal dW_s\right)$ is a martingale on $[0, T]$.

Proof. By (H3), if we define

$$B_t = W_t - \int_0^t g(s, F_s, U_s, V_s)ds,$$

then B is a $\tilde{\mathbb{P}}$-Brownian motion, where $\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_t = \mathcal{E}_t\left(\int_0^\cdot g(s, F_s, U_s, V_s)^\intercal dW_s\right)$ and the FBSDE (3.4)–(3.5) becomes

$$dF_t = (b(t, F_t) + \sigma(t, F_t)g(t, F_t, u(t, F_t), d(t, F_t)))\,dt + \sigma(t, F_t)dB_t; \qquad F_0 = x$$

$$dU_t = -f(t, F_t, U_t, V_t)dt + V_t dB_t; \qquad Y_T = h(F_T)$$

by (H1) and (H2). Note that F is a strong solution by the pathwise uniqueness assumption on (H4). Therefore, F is adapted to the augmented filtration generated by B, and so do $(U_t = u(t, F_t) : t \in [0, T])$ and $(V_t = d(t, F_t) : t \in [0, T])$. As a result, (F, U, V) solves the FBSDE and is adapted to the filtration generated by the underlying Brownian motion B. This implies that, for the unique strong solution X of the SDE in (H4), $(X, u(\cdot, X.), d(\cdot, X.))$ is a strong solution of (3.6)–(3.7).

On the other hand, let (X, Y, Z) and $(\tilde{X}, \tilde{Y}, \tilde{Z})$ be strong solutions of (3.6)–(3.7) such that

$$\mathcal{E}\left(-\int_0^\cdot g(s, X_s, Y_s, Z_s)^\intercal dW_s\right) \text{ and } \mathcal{E}\left(-\int_0^\cdot g(s, \tilde{X}_s, \tilde{Y}_s, \tilde{Z}_s)^\intercal dW_s\right)$$

are martingales on $[0, T]$. Then, by the Girsanov transform, for

$$B_t = W_t + \int_0^t g(s, X_s, Y_s, Z_s)ds \qquad \text{and} \qquad \tilde{B}_t = W_t + \int_0^t g(s, \tilde{X}_s, \tilde{Y}_s, \tilde{Z}_s)ds,$$

both (X, Y, Z, B) and $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{B})$ are weak solutions of (3.4)–(3.5). As the (3.4) enjoys the pathwise uniqueness, $X = \tilde{X}$ almost surely for a given Brownian motion W. As the BSDE (3.5) has a unique solution, we obtain $(Y, Z) = (\tilde{Y}, \tilde{Z})$ almost surely for a given Brownian motion W. □

Remark 3.2.2. It is also possible to construct a solution of the decoupled FBSDE using a solution of coupled FBSDE. This technique can be used to study multidimensional quadratic BSDE (see Section 2 of Cheridito and Nam (2015)).

Remark 3.2.3. Lemma 3.2.1 can be extended to the case where σ also depends on Y, Z as well. In this case, the transformed FBSDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t; \qquad X_0 = x$$

$$dY_t = -\left(f(t, X_t, Y_t, Z_t) + Z_t g(t, X_t, Y_t, Z_t)\right)dt + Z_t dW_t; \qquad Y_T = h(X_T)$$

still has coupling through σ; therefore, it does not simplify the problem. As a result, we cannot obtain the existence results such as Theorem 3.1.3.

3.2.2 Verification of (H1)

In this subsection, we prove that (H1) is satisfied under either (F1), (F2), or (F3). In addition, we analyze the solution F of (3.4), which will be used in the subsequent subsections. Let us use the following definition introduced by Hamadene et al. (1997).

Definition 3.2.4. Consider a class of SDEs

$$dX_s^{(t,x)} = b(s, X_s^{(t,x)})ds + \sigma(s, X_s^{(t,x)})dW_s; \qquad X_t^{(t,x)} = x \in \mathbb{R}^m \tag{3.8}$$

defined on $[t, T]$. We say that the coefficients (b, σ) satisfy the L^2-domination condition if the following conditions are satisfied:

- For each $(t, x) \in [0, T] \times \mathbb{R}^m$, the SDE (3.8) has a unique strong solution $X^{(t,x)}$. We denote $\mu_s^{(t,x)}$ as the law of $X_s^{(t,x)}$, that is, $\mu_s^{(t,x)} := \mathbb{P} \circ (X_s^{(t,x)})^{-1}$.
- For any $t \in [0, T], a \in \mathbb{R}^m, \mu_t^{(0,a)}$-almost every $x \in \mathbb{R}^m$, and $\delta \in (0, T - t]$, there exists a function $\phi_t : [t, T] \times \mathbb{R}^m \to \mathbb{R}_+$ such that
  - for all $k \geq 1$, $\phi_t \in L^2([t + \delta, T] \times [-k, k]^m; \mu_s^{(0,a)}(d\xi)ds)$
  - $\mu_s^{(t,x)}(d\xi)ds = \phi_t(s, \xi)\mu_s^{(0,a)}(d\xi)ds$

Proposition 3.2.5. If (H1) holds, (b, σ) satisfies the L^2-domination condition and $\mathbb{E}|F_t|^2$ is bounded uniformly for $t \in [0, T]$.

Proof. First, let us assume b is bounded as in (F1) or (F2). Since $\sigma\sigma^\intercal$ is bounded, by Theorem 1 of Aronson (1967), there are constants K and λ which only depend on m, T, ε and C that satisfies

$$K^{-1}(s - t_0)^{-m/2} \exp\left(-\frac{\lambda^{-1}|\xi - x_0|^2}{s - t_0}\right) \leq \frac{d\mu_s^{(t_0,x_0)}}{d\xi} \leq K(s - t_0)^{-m/2} \exp\left(-\frac{\lambda|\xi - x_0|^2}{s - t_0}\right)$$

for any $(t_0, x_0), (s, \xi) \in (0, T) \times \mathbb{R}^m$ with $s > t_0$. Note that $\mu_s^{(t,x)}(d\xi) = \phi_t(s, \xi)\mu_s^{(0,a)}(d\xi)$ where

$$\phi_t(s, \xi) := \frac{d\mu_s^{(t,x)}}{d\xi} \left(\frac{d\mu_s^{(0,a)}}{d\xi}\right)^{-1} \leq K^2 \left(\frac{s}{s - t}\right)^{m/2} \exp\left(-\frac{\lambda|\xi - x|^2}{s - t} + \frac{\lambda^{-1}|\xi - a|^2}{s}\right),$$

$\phi_t \in L^2([t + \delta, T] \times [-k, k]^m; \mu_s^{(0,a)}(d\xi)ds)$ for all $k \geq 1$. Therefore, the L^2-domination condition holds. For the case of (F3), we can use similar argument based on Theorem 1.2 of Menozzi et al. (2021).

On the other hand, note that there exists a constant K such that $|b(t, x)| \leq K(1+|x|)$. Then, there exist non-negative constants K_1 and K_2 that satisfy

$$\mathbb{E}|F_t|^2 \leq K_1 \left(|F_0|^2 + \mathbb{E}\int_0^t |b(s, F_s)|^2 ds + \mathbb{E}\int_0^T |\sigma(s, F_s)|^2 ds\right) \leq K_2 \left(1 + \int_0^t \mathbb{E}|F_s|^2 ds\right)$$

By Grönwall's inequality, we have $\sup_{t\in[0,T]} \mathbb{E}|F_t|^2 < \infty$. $\square$

Proposition 3.2.6. If either (F1), (F2), or (F3) holds, then (3.4) has a unique strong solution F, $\mathbb{E}|F_t|^2$ is bounded uniformly for $t \in [0, T]$, and (b, σ) satisfies the L^2-domination condition.

Proof. It is easy to verify that (3.4) has a unique strong solution if either (F1) or (F2) holds (see Gyongy and Martinez (2001) and Le Gall (1984)).

On the other hand, assume that (F3) holds. As a symmetric matrix $A := \sigma\sigma^\intercal$ has a nonzero determinant, A has orthonormal eigenvectors $\{\xi_1, \cdots, \xi_m\}$ with the corresponding strictly positive eigenvalues $\{\lambda_1, \cdots, \lambda_m\}$. Let

$$E := \left(\frac{1}{\sqrt{\lambda_j}} \xi_j : j = 1, 2, \ldots, m \right).$$

Then, $E^\intercal A E$ becomes an identity matrix. Therefore, by Lévy characterization, we know that $B := E^\intercal \sigma W$ is a $\mathbb{P}$-Brownian motion. Note that Menoukeu-Pamen and Mohammed (2019) shows that the forward SDE

$$dP_t = E^\intercal b(t, (E^\intercal)^{-1} P_t)dt + dB_t; \qquad P_0 = E^\intercal x$$

has a unique strong solution because $b(t,x)$ has a linear growth in x, as pointed out in Remark 3.1.2. Therefore, $\tilde{X} := (E^\intercal)^{-1} P$ is a unique solution of (3.4).

Therefore, (H1) is satisfied under one of the assumptions (F1), (F2), or (F3). The remainder of our claim is proved by Proposition 3.2.5. □

3.2.3 Verification of (H2) and (H3)

In this subsection, we will always assume (H1).

Proposition 3.2.7. (B1) implies (H2) and (H3).

Proof. First, let us assume $r > 0$, which implies $\rho_r \equiv 0$. Then, (H3) holds automatically because g is bounded. Note that $\tilde{X}$ obtained by (3.4) is a Markov process because the corresponding Martingale problem is well posed. By Propositions 3.2.5 and B.0.1, all of the conditions in Remark 27.3 of Hamadene et al. (1997) are satisfied. As $\overline{f}$ exhibits linear growth in (y,z), Theorem 27.2 of Hamadene et al. (1997) proves (H2).

On the other hand, if $r = 0$, then h and $f(t,x,0,0)$ are bounded by C. Then, we will show that $U_t = u(t,F_t)$ is uniformly bounded by $e^{\frac{1}{2}aT}\sqrt{C^2+T}$. If so, (H3) is automatically satisfied. Moreover, because $\overline{f}$ exhibits linear growth in (y,z), (H2) holds by the previous argument again.

For a positive constant N, let

$$P_N : [0,T] \times \mathbb{R}^m \times \mathbb{R}^d \times \mathbb{R}^{d\times n} \ni (t,x,y,z) \mapsto \left(t, x, \frac{Ny}{|y| \vee N}, z\right) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^d \times \mathbb{R}^{d\times n}$$

and $f_N := f \circ P_N, g_N := g \circ P_N$, and $H_N(t,x,y,z) := f_N(t,x,y,z) + zg_N(t,x,y,z)$. Then, by (B1), there exists a constant $\tilde{C} > 0$ such that

$$|H_N(t,x,y,z)| \le \tilde{C}(1 + |z|).$$

Then, by the same argument for $r > 0$, the FBSDE

$$\begin{aligned} dF_t &= b(t,F_t)dt + \sigma dW_t; & F_0 &= x \\ dU_t &= -H_N(t,F_t,U_t,V_t)dt + V_t dW_t; & U_T &= h(F_T) \end{aligned} \tag{3.9}$$

has a strong solution (F,U,V) such that there exist Borel measurable functions (u,d) : $[0,T] \times \mathbb{R}^m \to \mathbb{R}^d \times \mathbb{R}^{d\times n}$ such that $U_t = u(t,F_t)$ and $V_t = d(t,F_t)$ $dt \otimes d\mathbb{P}$-almost everywhere. By Itô formula, we have

$$e^{at}|U_t|^2 = e^{aT}|h(F_T)|^2 + \int_t^T e^{as}\left(2U_s^\intercal f_N(s,F_s,U_s,V_s) - |V_s|^2 - a|U_s|^2\right)ds - \int_t^T 2e^{as}U_s^\intercal V_s d\tilde{W}_s$$

where $a = 2C(C+1)$ and

$$\tilde{W}_t = W_t - \int_0^t g_N(s,F_s,U_s,V_s)ds.$$

Note that $\tilde{W}$ is a Brownian motion under some measure $\tilde{\mathbb{P}}$ because g_N is bounded. In addition, by using the inequality $2Cxy \le C^2x^2 + y^2$, we have

$$\begin{aligned}&2U_t^\intercal f_N(t,F_t,U_t,V_t) - |V_t|^2 - a|U_t|^2\\ &\le 2C|U_t|(1+|U_t|+|V_t|) - |V_t|^2 - a|U_t|^2 \le (2C^2+2C-a)|U_t|^2 + 1 \le 1.\end{aligned}$$

Therefore, if we denote $\tilde{\mathbb{E}}$ as the expectation with respect to $\tilde{\mathbb{P}}$, we obtain

$$\begin{aligned}|U_t|^2 &= e^{-at}\tilde{\mathbb{E}}\left[e^{aT}|h(F_T)|^2 + \int_t^T e^{as}\left(2U_s^\intercal f_N(s,F_s,U_s,V_s) - |V_s|^2 - a|U_s|^2\right)ds\right]\\ &\le e^{a(T-t)}\left(C^2+T-t\right).\end{aligned}$$

Therefore, U is uniformly bounded, independent of the choice of N. If we set

$$N \ge e^{\frac{1}{2}aT}\sqrt{C^2+T},$$

the solution of (3.9) is the solution of (3.5) and $|U_t| \le N$. This proves the claim. $\square$

Proposition 3.2.8. (B2) implies (H2) and (H3).

Proof. Note that by Hu and Tang (2016), BSDE (3.5) has a unique solution (U,V) such that U is bounded. In this case, $V \cdot W$ is a BMO martingale. Therefore, without loss of generality, we can assume that $\hat{f}$ is a bounded Lipschitz function and $|g(t,x,y,z)| \le C$. Then, (H3) is satisfied.

Now, we only need to prove the existence of measurable functions u and d such that $(U_t,V_t) = (u(t,F_t),d(t,F_t))$. Let $U_t^{(0)} = u_0(t,F_t) = 0$ for all $t \in [0,T]$ and $V_t^{(0)} = d_0(t,F_t) = 0$, and we define

$$U_t^{(k+1),i} = h^i(F_T) + \int_t^T \hat{f}^i(s,F_s,U_s^{(k)},V_s^{(k)}) + \tilde{f}^i(s,F_s,V_s^{(k+1),i})ds - \int_t^T V^{(k+1),i}dW_t.$$

Then, as shown in the proof of Proposition 3.2.9, there exist measurable functions $u_k : [0,T]\times\mathbb{R}^m \to \mathbb{R}^d$ and $d_k : [0,T]\times\mathbb{R}^m \to \mathbb{R}^{d\times n}$ such that $U_t^{(k)} = u_k(t,F_t)$ and $V_t^{(k)} = d_k(t,F_t)$. As $U^{(k)} \to U$ in $\mathbb{S}^\infty$ and $V^{(k)}\cdot W \to V\cdot W$ in BMO by Hu and Tang (2016), if we let $u^i(t,x) := \limsup_{k\to\infty} u_k^i(t,x)$ and $d^{ij}(t,x) := \limsup_{k\to\infty} d_k^{ij}(t,x)$, where $u = (u^i)_{1\le i\le d}$ and $d = (d^{ij})_{1\le i\le d,1\le j\le n}$, we have

$$\begin{aligned}u^i(t,F_t) &= (\limsup_{k\to\infty} u_k^i)(t,F_t) = \limsup_{k\to\infty}(u_k^i(t,F_t)) = \lim_{k\to\infty} U_t^{(k),i} = U_t^i\\ d^{ij}(t,F_t) &= (\limsup_{k\to\infty} d_k^{ij})(t,F_t) = \limsup_{k\to\infty}(d_k^{ij}(t,F_t)) = \lim_{k\to\infty} V_t^{(k),ij} = V_t^{ij}\end{aligned}$$

in $dt\otimes d\mathbb{P}$-everywhere sense. Therefore, (H2) holds. $\square$

Proposition 3.2.9. (B3) implies (H2) and (H3).

Proof. The existence of solution (F, U, V) for BSDE (3.5) can be seen in Kobylanski (2000). In particular, U is bounded and $V \cdot W$ is a BMO martingale (see Briand and Elie (2013)). Therefore, without loss of generality, we assume that

$$|f(t,x,y,z)| \leq C(1+|z|^2) \qquad \text{and} \qquad |g(t,x,y,z)| \leq C,$$

and therefore, (H3) holds.

On the other hand, by Kobylanski (2000), there is a sequence of measurable functions $\theta_k(t,x,y,z)$ such that

- $\theta_k(t,x,y,z)$ is uniformly Lipschitz in (y,z).
- For the solution $(Y^{(k)}, Z^{(k)})$ of BSDE

$$Y_t^{(k)} = \exp(Lh(F_t)) + \int_t^T \theta_k(s, F_s, Y_s^{(k)}, Z_s^{(k)}) - \int_t^T Z_s^{(k)} dW_s,$$

  we have

$$\lim_{k\to\infty} \frac{\log(Y_t^{(k)})}{2L} = U_t \text{ uniformly in } t$$
$$\lim_{k\to\infty} \frac{Z^{(k)}}{2LY^{(k)}} = V \text{ in } \mathbb{H}^2.$$

Here, L is a constant determined by coefficients h and $\overline{f}$.

From Proposition B.0.1, there are measurable functions $\tilde{u}_k : [0,T] \times \mathbb{R}^m \to \mathbb{R}$ and $\tilde{d}_k : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{1\times n}$ such that $Y_t^{(k)} = \tilde{u}_k(t, F_t)$ and $Z_t^{(k)} = \tilde{d}_k(t, F_t)$. Therefore, if we let $\tilde{u}(t,x) := \limsup_{k\to\infty} \tilde{u}_k(t,x)$ and $\tilde{d}^i(t,x) := \limsup_{k\to\infty} \tilde{d}_k^i(t,x)$, where $\tilde{d} = (\tilde{d}^i)_{1\leq i\leq n}$, then $U_t = u(t,F_t) := \frac{log(\tilde{u}(t,F_t))}{2L}$ and $V_t = d(t,F_t) := \frac{\tilde{d}(t,F_t)}{2L\tilde{u}(t,F_t)}$. This proves (H2). $\square$

Proposition 3.2.10. (F3) and (B4) imply (H2) and (H3).

Proof. For (H2), the proof is identical to that for Theorem 3.1 of Mu and Wu (2015). Let E be the matrix defined in the proof of Proposition 3.2.6. Then,

$$dP_t = E^\intercal b(t, (E^\intercal)^{-1} P_t)dt + E^\intercal \sigma dW_t; \qquad P_0 = E^\intercal x$$

for $P = E^\intercal F$. Note that there exists $C' > 0$ such that

$$\begin{aligned} |P_t| &\leq |E^\intercal x| + \int_0^t |E^\intercal b(s, (E^\intercal)^{-1} P_s)|ds + |E^\intercal \sigma W_t| \\ &\leq C' + |E^\intercal \sigma W_t| + C' \int_0^t |P_s|ds. \end{aligned}$$

By Gronwall's inequality, there exists a constant $C'' > 0$ such that

$$|P_t| \leq C'' \left(1 + \max_{s\in[0,t]} |E^\intercal \sigma W_s|\right).$$

This implies

$$|F_t| = |(E^\intercal)^{-1} P_t| \leq C''' \left(1 + \max_{s\in[0,t]} |W_s|\right)$$

for a constant $C''' > 0$. Therefore, by the Beneš condition, (H3) holds. $\square$

3.2.4 Verification of (H4)

Let us prove (H4) under the assumptions made in Theorem 3.1.3.

Proof. Note that if (H1) and either one of (B1), (B2) or (B3) hold, $g(t, F_t, U_t, V_t)$ is bounded because U_t is bounded. Therefore, we only need to prove that (H4) holds for $g_N(t, x, y, z) := g(t, x, Ny/(|y| \vee N), z)$. If either (F1) or (F2) holds, then $\tilde{b}(t, x) := b(t, x) + \sigma(t, x)g_N(t, x, u(t, x), d(t, x))$ is bounded because σ and g_N are bounded. Therefore, conditions (F1) or (F2) hold with $\tilde{b}$ instead of b. Likewise, if (F3) holds, then there exists a positive constant κ such that,

$$|\tilde{b}(t, 0)| + \sup_{|x-x'|\leq 1} |\tilde{b}(t, x) - \tilde{b}(t, x')| \leq \kappa$$

for all $t \in [0, T], x, x' \in \mathbb{R}^m$. By Proposition 3.2.6, (H4) holds.

On the other hand, assume (F3) and (B4) hold. Note that, if $r = 0$ in (B4), then $|U|$ is bounded; therefore, $|g(t, x, y, z)| \leq K(1 + |x|)$ for some K. If $r > 0$, $|g(t, x, y, z)| \leq K(1 + |x|)$ for $K = C$. Therefore, we have

$$|\tilde{b}(t, x)| \leq |b(t, x)| + |\sigma(t, x)||g(t, x, y, z)| \leq C'(1 + |x|)$$

for a non-negative constant C'. Again, by Proposition 3.2.6, (H4) holds. □

3.2.5 Uniqueness

Assume the conditions in Theorem 3.1.3. Let us prove that either (B2), (U1), or (U2) implies the uniqueness of the solution for (3.6)–(3.7).

Proof. When $f(t, x, y, z) + zg(t, x, y, z)$ is Lipschitz, the solution for (3.5) is unique by Pardoux and Peng (1990). If $r = 0$ and $d = 1$, Kobylanski (2000) proved the uniqueness of the solution for (3.5). The uniqueness of the solution under the condition (B2) was proved by Hu and Tang (2016). By applying Lemma 3.2.1, the uniqueness of the solution for (3.6)–(3.7) is proved. □

3.3 Applications of Discontinuous Coefficients FBSDEs

In this section, we provide simple applications of our main result to the optimal control of the spread of an infectious disease and the carbon market allowance pricing.

3.3.1 Controlling the Spread of an Infectious Disease

Let W be a one-dimensional Brownian motion, P be the number of infections, and α be the measures imposed by the policymaker to stop the spread. The admissible set for α is the set of non-negative adapted processes in $\mathbb{H}^2(\mathbb{R})$. For measurable functions $\theta : \mathbb{R} \to \mathbb{R}$ and positive constant σ, assume that P follows the dynamics

$$\frac{dP_t}{P_t} = \left(\theta(\log P_t) + \frac{1}{2}|\sigma|^2 - \alpha_t\right) dt + \sigma dW_t; \qquad P_0 = e^x. \tag{3.10}$$

The interpretation of the dynamics is straightforward. Assuming there is no randomness in the spread ($\sigma = 0$), $\theta(\log P)$ represents the exponent of the infection growth when there is no intervention ($\alpha = 0$). In summary, if policy α is introduced, the growth exponent will be reduced to $\theta(\log P) - \alpha$.

Let us define $X = \log P$. By Itô formula, (3.10) transforms to

$$dX_t = (\theta(X_t) - \alpha_t)\,dt + \sigma dW_t; \qquad\qquad X_0 = x.$$

Our objective as the policymaker is to minimize

$$J(\alpha) := \mathbb{E}\left[\int_0^T |\alpha_t|^2 + q(X_t)dt\right].$$

Here, $|\alpha_t|^2$ represents the running cost of the policy α, and $q : \mathbb{R} \to [0, \infty)$ is the cost incurred by the number of infections.

Remark 3.3.1. It is realistic to assume that q is a non-differentiable function, as it is the cost of infection. For example, consider that there is a capacity for medical services. If the infected patient number P exceeds a certain level, there will be a shortage of medical services, which will cost a lot more per additional patient.

Let the corresponding Hamiltonian and its minimizer be

$$\begin{aligned} H(t, x, y, z, \pi) &:= (\theta(x) - \pi)y + |\pi|^2 + q(x) \\ \arg\min_{\pi} H(t, x, y, z, \pi) &= \frac{(y \vee 0)}{2}. \end{aligned}$$

The following proposition is a version of Theorem 4.25 Carmona (2016) for convex, possibly not continuously differentiable θ and q. Here, ∂_+ denotes the right derivative. We introduce related topics in Appendix D.

Proposition 3.3.2. Assume that θ and q are convex Lipschitz functions, q is non-decreasing, and $\sigma > 0$. Let $(X, Y, Z) \in \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$ be the unique solution of FBSDE

$$\begin{aligned} dX_t &= \left(\theta(X_t) - \frac{(Y_t \vee 0)}{2}\right) dt + \sigma dW_t; \qquad X_0 = x \\ dY_t &= -\left(\partial_+ q(X_t) + \partial_+\theta(X_t)Y_t\right) dt + Z_t dW_t; \quad Y_T = 0 \end{aligned} \tag{3.11}$$

such that Y is bounded. Then, for $\alpha_t^* := (Y_t \vee 0)/2$, we have $J(\alpha^*) \leq J(\alpha)$ for any non-negative process $\alpha \in \mathbb{H}^2(\mathbb{R})$.

Before we prove the proposition, we need the following lemma.

Lemma 3.3.3. The following holds:

(i) (3.11) has a unique solution $(X, Y, Z) \in \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$ such that Y is bounded almost surely.

(ii) There exists a constant C such that $Y_t \in [0, C]$ for all t almost surely.

(iii) $\mathbb{E}\int_0^T (X_t - X_t^\alpha) Z_t dW_t = 0$ for any $\alpha \in \mathbb{H}^2(\mathbb{R})$, where

$$dX_t^\alpha = (\theta(X_t^\alpha) - \alpha_t)\,dt + \sigma dW_t; \qquad\qquad X_0 = x.$$

Proof. First, we prove (ii) under the assumption that (3.11) has a solution $(X, Y, Z) \in \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$. Note that there exists a constant C such that $\partial_+ q \in [0, C]$ and $\partial_+\theta \in [-C, C]$. By the comparison principle, $Y_t^d \le Y_t \le Y_t^u$ for all $t \in [0, T]$ almost surely, where

$$\begin{aligned} dY_t^u &= -\left(C + C|Y_t^u|\right) dt + Z_t^u dW_t; & Y_T^u &= 0 \\ dY_t^d &= C|Y_t^d| dt + Z_t^d dW_t; & Y_T^d &= 0. \end{aligned}$$

As $Y_t^u = e^{C(T-t)} - 1 \le e^{CT}$ and $Y^d \equiv 0$ almost surely, the claim is proved.

Next, we prove (i) by using the localization argument. For C_y, the bound of Y we obtained in (ii), we define φ be a smooth function on $\mathbb{R}$ satisfying

$$\varphi(y) = \begin{cases} y, & \text{if } y \in [0, C_y] \\ 0, & \text{if } y \in (-\infty, -1] \cup [C_y + 1, \infty) \end{cases}$$

and $|\varphi(y)| \le |y|$. Consider the FBSDE

$$\begin{aligned} d\tilde{X}_t &= \left(\theta(\tilde{X}_t) - \frac{\varphi(\tilde{Y}_t)}{2}\right) dt + \sigma dW_t; & X_0 &= x \\ d\tilde{Y}_t &= -\left(\partial_+ q(\tilde{X}_t) + \partial_+\theta(\tilde{X}_t)\varphi(\tilde{Y}_t)\right) dt + \tilde{Z}_t dW_t; & Y_T &= 0. \end{aligned} \tag{3.12}$$

Let $b(t, x) \equiv \theta(x), \sigma(t, x) :\equiv \sigma, g(t, x, y, z) := -\frac{\varphi(y)}{2}, f(t, x, y, z) := \partial_+ q(x) + \partial_+\theta(x)\varphi(\tilde{y})$, and $h \equiv 0$. As θ and q are Lipschitz, $\partial_+\theta$ and $\partial_+ q$ are bounded. Then, one can verify that the coefficients satisfy (F3), (B4), and (U2) with $r = 0$. Therefore, there exists a unique $(\tilde{X}, \tilde{Y}, \tilde{Z}) \in \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$ such that $\tilde{Y}$ is bounded. In particular, by the same comparison arguement we used in the proof of (ii), we obtain $\tilde{Y} \in [0, C_y]$. Therefore, $(\tilde{X}, \tilde{Y}, \tilde{Z})$ also solves (3.11). Therefore, we proved the existence of a solution. On the other hand, let $(X', Y', Z') \in \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$ be another solution to (3.11). Then, by part (ii), $Y_t' \in [0, C]$ for all t. As (X', Y', Z') also solves (3.12), we have $(X', Y', Z') = (\tilde{X}, \tilde{Y}, \tilde{Z}) = (X, Y, Z)$. This proves the uniqueness of a solution.
Now, let us prove (iii). As θ grows linearly, there exists $C > 0$ such that $|\theta(x)| \le C(1+|x|)$. Therefore,

$$\begin{aligned} |X_t^\alpha| &\le |x| + \int_0^t C(1 + |X_s^\alpha|) + |\alpha_s| ds + |\sigma W_t| \\ &\le |x| + CT + \int_0^T |\alpha_s| ds + C\int_0^t |X_s^\alpha| ds + |\sigma W_t|. \end{aligned}$$

Note that if we let $X_t^* := \sup_{0 \le u \le t} |X_u^\alpha|$, there exists another constant C' such that

$$\mathbb{E}|X_t^*|^2 \le C'\left(1 + \int_0^t \mathbb{E}|X_s^*|^2 ds\right)$$

as $\mathbb{E}\sup_{0 \le u \le T} |W_u|^2 < \infty$ and $\mathbb{E}\int_0^T |\alpha_s|^2 ds < \infty$. By Grönwall's inequality, $\mathbb{E}|X_T^*|^2 = \mathbb{E}\sup_{0 \le u \le T} |X_u^\alpha|^2 < \infty$. We obtain $\mathbb{E}\sup_{0 \le u \le T} |X_u|^2 < \infty$ by the same argument.

As

$$\begin{aligned}\mathbb{E}\sqrt{\left\langle \int_0^\cdot (X_t - X_t^\alpha) Z_t dW_t \right\rangle_T} &= \mathbb{E}\sqrt{\int_0^T |X_t - X_t^\alpha|^2 |Z_t|^2 dt} \\ &\leq \mathbb{E} \sup_{0\leq u\leq T} |X_u - X_u^\alpha| \sqrt{\int_0^T |Z_t|^2 dt} \\ &\leq \frac{1}{2}\mathbb{E} \sup_{0\leq u\leq T} |X_u - X_u^\alpha|^2 + \frac{1}{2}\mathbb{E}\int_0^T |Z_t|^2 dt < \infty,\end{aligned}$$

we prove the claim by the Burkholder-Davis-Gundy inequality. □

Proof of Proposition 3.3.2. For a given control α, let us denote the corresponding dynamics of the log of infection number as X', that is,

$$dX'_t = (\theta(X'_t) - \alpha_t)\, dt + \sigma dW_t; \qquad X_0 = x.$$

Note that, for $\alpha^* := (Y \vee 0)/2$,

$$\begin{aligned}&J(\alpha^*) - J(\alpha) = \mathbb{E}\int_0^T \left[|\alpha_t^*|^2 - |\alpha_t|^2 + q(X_t) - q(X'_t)\right] dt \\ &= \mathbb{E}\int_0^T \left[H(t, X_t, Y_t, Z_t, \alpha_t^*) - H(t, X'_t, Y_t, Z_t, \alpha_t)\right] dt - \mathbb{E}\int_0^T \left[\theta(X_t) - \theta(X'_t) - \alpha_t^* + \alpha_t\right] Y_t dt.\end{aligned}$$

Note that, by integration by parts, we have

$$\mathbb{E}\int_0^T \left[\theta(X_t) - \theta(X'_t) - \alpha_t^* + \alpha_t\right] Y_t dt = \mathbb{E}\int_0^T (X_t - X'_t)\left(\partial_+ q(X_t) + \partial_+\theta(X_t) Y_t\right) dt.$$

Therefore,

$$\begin{aligned}&J(\alpha^*) - J(\alpha) \\ &= \mathbb{E}\int_0^T \left[H(t, X_t, Y_t, Z_t, \alpha_t^*) - H(t, X'_t, Y_t, Z_t, \alpha_t)\right] - (X_t - X'_t)\left(\partial_+ q(X_t) + \partial_+\theta(X_t) Y_t\right) dt \\ &\leq \mathbb{E}\int_0^T \left[H(t, X_t, Y_t, Z_t, \alpha_t) - H(t, X'_t, Y_t, Z_t, \alpha_t)\right] - (X_t - X'_t)\left(\partial_+ q(X_t) + \partial_+\theta(X_t) Y_t\right) dt.\end{aligned}$$

Here, we used the fact that

$$H(t, X_t, Y_t, Z_t, \alpha_t^*) \leq H(t, X_t, Y_t, Z_t, \alpha_t)$$

for any non-negative process $\alpha \in \mathbb{H}^2(\mathbb{R})$. As $Y_t \geq 0$, $q(x) - q(x') \leq (x - x')\partial_+ q(x)$ and $\theta(x) - \theta(x') \leq (x - x')\partial_+\theta(x)$, we have

$$\begin{aligned}H(t, X_t, Y_t, Z_t, \alpha_t) - H(t, X'_t, Y_t, Z_t, \alpha_t) &= (\theta(X_t) - \theta(X'_t)) Y_t + q(X_t) - q(X'_t) \\ &\leq (X_t - X'_t)\left(\partial_+ q(X_t) + \partial_+\theta(X_t) Y_t\right).\end{aligned}$$

Therefore, $J(\alpha^*) \leq J(\alpha)$. □

3.3.2 Electricity Market with Carbon Emission Allowance

Let us provide an example of an FBSDE with measurable coefficients in the pricing of carbon emission allowance in the electricity market. We follow the example in Carmona et al. (2013) except that we assume there exists a cost c^i depending on the carbon emission abatement and the total cumulative carbon emission, whereas Carmona et al. (2013) assumes only the cost's dependency on the carbon emission abatement. Heuristically, if total cumulative carbon emission increases, the government will try to reduce the marginal cost of carbon emission abatement (e.g. Emission Reduction Fund) and the society will be urge to develop cost efficient green technologies.

For simplicity, let $\mathbb{P}$ be a risk-neutral measure and let the cumulative emission of the ith firm ($i = 1, 2, ..., N$) up to time t be E^i_t. Assume that E^is follow the dynamics

$$E^i_t = E^i_0 + \int_0^t \left(b^i\left(s, \overline{E}_s\right) - \xi^i_s\right) ds + \int_0^t \sigma^i(s, \overline{E}_s) dW_s,$$

where $\overline{E}_s := \sum_{j=1}^N E^j_s$. Here, b^i denotes the so-called *business-as-usual*, the rate of emission without carbon regulation. The process ξ^i is the instantaneous rate of abatement chosen by the firm. The firm controls its own abatement schedule ξ^i and the carbon emission allowance quantity θ^i, which is traded in the allowance market. Both control processes need to be $dt \otimes d\mathbb{P}$-square-integrable adapted processes, which is denote by $\mathcal{A}$. The firm's wealth is given by

$$X^i_T\left(\xi^i, \theta^i\right) = x^i + \int_0^T \theta^i_s dY_s - \int_0^T c^i(\xi^i_s, \overline{E}_s) ds - E^i_T Y_T$$

where x^i is the initial wealth, Y is the allowance price, $c^i : \mathbb{R}^2 \to \mathbb{R}$ is the cost occurred by the abatement ξ^i. We assume that $c^i(e, y)$ is jointly measurable and convex in x. Then, one can define

$$g^i(e, y) = \arg\min_x \left(c^i(x, e) - yx\right).$$

We assume that the utility of each firm is given by an increasing, strictly concave function $U : \mathbb{R} \to \mathbb{R}$, which satisfies the Inada conditions: $U'(-\infty) = +\infty$ and $U'(+\infty) = 0$.

The corresponding optimization problem is to find a pair of $(\xi^i, \theta^i) \in \mathcal{A}$ that maximizes $\mathbb{E}U(X^i_T(\xi^i, \theta^i))$.

By the same argument in Proposition 1 of Carmona et al. (2013), we can deduce that $\xi^i_t = g^i(\overline{E}_t, Y_t)$ is the optimal control. Therefore, $(\overline{E}, Y)$ should solve the following FBSDE: for $b = \sum_{i=1}^N b^i, \sigma = \sum_{i=1}^N \sigma^i$, and $g = \sum_{i=1}^N g^i$,

$$\begin{aligned} d\overline{E}_t &= \left[b(t, \overline{E}_t) - g(\overline{E}_t, Y_t)\right] dt + \sigma(t, \overline{E}_t) dW_t; & \overline{E}_0 &\in \mathbb{R} \\ dY_t &= Z_t dW_t; & Y_T &= \lambda 1_{[\Lambda, \infty)}(\overline{E}_T). \end{aligned} \tag{3.13}$$

Here, the terminal condition of allowance is assumed to be an indicator function based on Carmona et al. (2010). As one can see from the following example, our main theorem 3.1.3 generalizes Theorem 1 of Carmona et al. (2013) as we allow all the coefficients to be discontinuous.

Example 3.3.4. Assume that there exists constants C, K and $\alpha^i \in (0, 1), i = 1, 2, ..., N$ such that, for all $(t, e, y) \in [0, T] \times \mathbb{R} \times \mathbb{R}$,

$$\begin{aligned} &|b(t, e)| \le C \text{ and } C^{-1} \le (\sigma\sigma^\intercal)(t, e) \le C \\ &c^i(x, e) = \frac{1}{2} x^2 \left(1 - \alpha^i \mathbf{1}_{e \ge K}\right). \end{aligned}$$

Furthermore, assume that $\sigma(t,x)$ is locally Lipschitz with respect to x. Then, (3.13) has a unique strong solution such that Y is bounded.

Proof. Note that

$$g(e,y) = \sum_{i=1}^{N} \frac{y}{1-\alpha^i \mathbf{1}_{e\geq K}}.$$

Then it is easy to check (F1) and (B1) are satisfied with $r=0$. Therefore, there exists a strong solution of (3.13) such that Y is bounded. Moreover, (U2) holds. The uniqueness result of Theorem 3.1.3 implies the uniqueness of a solution. □

In the Appendix C we provide some background about the carbon market and show how we achieve FBSDE (3.13) in details.

Bibliography

F. Antonelli. *Backward forward stochastic differential equations.* PhD thesis, Purdue University, 1993.

D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bulletin of the American Mathematical society*, 73(6):890–896, 1967.

G. Barles and E. Lesigne. Sde, bsde and pde. *Pitman Research Notes in Mathematics Series*, pages 47–82, 1997.

V. Beneš. Existence of optimal stochastic control laws. *SIAM Journal on Control*, 9(3): 446–472, 1971.

J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

P. Briand and R. Elie. A simple constructive approach to quadratic BSDEs with or without delay. *Stochastic processes and their applications*, 123(8):2921–2939, 2013.

R. H. Cameron and W. T. Martin. Transformations of Weiner integrals under translations. *Annals of Mathematics*, pages 386–396, 1944.

R. Carmona. *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications.* SIAM, 2016.

R. Carmona and F. Delarue. Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics. *The Annals of Probability*, 43(5):2647–2700, 2015.

R. Carmona, M. Fehr, J. Hinz, and A. Porchet. Market design for emission trading schemes. *Siam Review*, 52(3):403–452, 2010.

R. Carmona, F. Delarue, G.-E. Espinosa, and N. Touzi. Singular forward–backward stochastic differential equations and emissions derivatives. *The Annals of Applied Probability*, 23(3):1086–1128, 2013.

J.-F. Chassagneux, H. Chotai, and M. Muûls. *A Forward-Backward SDEs Approach to Pricing in Carbon Markets.* Springer, 2017.

J. Chen, J. Ma, and H. Yin. Forward-backward SDEs with discontinuous coefficients. *Stochastic Analysis and Applications*, 36(2):274–294, 2018.

P. Cheridito and K. Nam. Multidimensional quadratic and subquadratic BSDEs with special structure. *Stochastics An International Journal of Probability and Stochastic Processes*, 87(5):871–884, 2015.

A. S. Cherny and H.-J. Engelbert. *Singular Stochastic Differential Equations.* Springer, 2004.

R. Chitashvili and M. Mania. Generalized Itô formula and derivation of Bellman's equation. *K. Engelbert et alia. Stochastic processes and related topics (Siegmundsberg, 1994)*, pages 1–21, 1996.

E. Çinlar. *Probability and stochastics*, volume 261. Springer, 2011.

E. Çinlar, J. Jacod, P. Protter, and M. J. Sharpe. Semimartingales and Markov processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 54(2):161–219, 1980.

R. H. Coase. The problem of social cost. pages 87–137, 1960.

N. El-Karoui and S. Hamadène. Bsdes and risk-sensitive control, zero-sum and nonzero-sum game problems of stochastic functional differential equations. *Stochastic Processes and their Applications*, 107(1):145–169, 2003.

N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

I. V. Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3): 285–301, 1960.

I. Gyongy and T. Martinez. On Stochastic Differential Equations with Locally Unbounded Drift. *Czechoslovak Mathematical Journal*, 51:763–783, 2001.

S. Hamadène. Multidimensional backward stochastic differential equations with uniformly continuous coefficients. *Bernoulli*, pages 517–534, 2003.

S. Hamadene and J.-P. Lepeltier. Zero-sum stochastic differential games and backward equations. *Systems & Control Letters*, 24(4):259–263, 1995.

S. Hamadene, J.-P. Lepeltier, and S. Peng. BSDEs with continuous coefficients and stochastic differential games. *Pitman Research Notes in Mathematics Series*, pages 115–128, 1997.

Y. Hu and S. Peng. Solution of forward-backward stochastic differential equations. *Probability Theory and Related Fields*, 103(2):273–283, 1995.

Y. Hu and S. Tang. Multi-dimensional backward stochastic differential equations of diagonally quadratic generators. *Stochastic Processes and their Applications*, 126(4): 1066–1086, 2016.

I. Karatzas and S. Shreve. Brownian motion and stochastic calculus. *Graduate texts in mathematics*, 1991.

N. Kazamaki. On a problem of Girsanov. *Tohoku Mathematical Journal*, 29(4):597–600, 1977.

N. Kazamaki. *Continuous exponential martingales and BMO.* Springer, 2006.

D. Kim and N. Krylov. Parabolic equations with measurable coefficients. *Potential analysis*, 26(4):345–361, 2007.

F. C. Klebaner. *Introduction to stochastic calculus with applications*. Imperial College Press, 2012.

M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of probability*, pages 558–602, 2000.

N. V. Krylov. *Controlled diffusion processes*, volume 14. Springer Science & Business Media, 1980.

N. V. Krylov. Parabolic and elliptic equations with VMO coefficients. *Communications in Partial Differential Equations*, 32(3):453–475, 2007.

J.-F. Le Gall. One—dimensional stochastic differential equations involving the local times of the unknown process. pages 51–82, 1984.

P. Luo, O. Menoukeu-Pamen, and L. Tangpi. Strong solutions of forward-backward stochastic differential equations with measurable coefficients. *arXiv preprint arXiv:2001.07753*, 2020.

J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probability theory and related fields*, 98(3):339–359, 1994.

J. Ma, J.-M. Morel, and J. Yong. *Forward-backward stochastic differential equations and their applications*. Number 1702. Springer Science & Business Media, 1999.

J. Ma, Z. Wu, D. Zhang, and J. Zhang. On well-posedness of forward–backward SDEs—A unified approach. *The Annals of Applied Probability*, 25(4):2168–2214, 2015.

M. Mania and R. Tevzadae. Semimartingale functions of a class of diffusion processes. *Theory of Probability & Its Applications*, 45(2):337–343, 2001.

A. Maugeri, D. K. Palagachev, and L. Softova. Elliptic and parabolic equations with discontinuous coefficients. *Mathematical Research*, 2000.

O. Menoukeu-Pamen and S. E. Mohammed. Flows for singular stochastic differential equations with unbounded drifts. *Journal of Functional Analysis*, 277(5):1269–1333, 2019.

O. Menoukeu-Pamen, T. Meyer-Brandis, T. Nilssen, F. Proske, and T. Zhang. A variational approach to the construction and Malliavin differentiability of strong solutions of SDE's. *Mathematische Annalen*, 357(2):761–799, 2013.

S. Menozzi, A. Pesce, and X. Zhang. Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift. *Journal of Differential Equations*, 272:330–369, 2021.

T. Meyer-Brandis and F. Proske. Construction of strong solutions of SDE's via Malliavin calculus. *Journal of Functional Analysis*, 258(11):3922–3953, 2010.

R. Mu and Z. Wu. One kind of multiple dimensional Markovian BSDEs with stochastic linear growth generators. *Advances in Difference Equations*, 2015(1):1–15, 2015.

A. A. Novikov. On an identity for stochastic integrals. *Teoriya Veroyatnostei i ee Primeneniya*, 17(4):761–765, 1972.

É. Pardoux. Bsdes, weak convergence and homogenization of semilinear pdes. In *Nonlinear analysis, differential equations and control*, pages 503–549. Springer, 1999.

E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, pages 200–217. Springer, 1992.

E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasi-linear parabolic PDEs. *Probability Theory and Related Fields*, 114(2):123–150, 1999.

S. Peng. Stochastic Hamilton–Jacobi–Bellman equations. *SIAM Journal on Control and Optimization*, 30(2):284–304, 1992.

S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37(3): 825–843, 1999.

L. C. G. Rogers and D. Williams. *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, volume 2. Cambridge university press, 2000.

A. Y. Veretennikov. On the strong solutions of stochastic differential equations. *Theory of Probability & Its Applications*, 24(2):354–366, 1980.

T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155–167, 1971.

J. Yong. Finding adapted solutions of forward–backward stochastic differential equations: method of continuation. *Probability Theory and Related Fields*, 107(4):537–572, 1997.

J. Yong. Forward-backward stochastic differential equations with mixed initial-terminal conditions. *Transactions of the American Mathematical Society*, 362(2):1047–1096, 2010.

X. Zhang. Strong solutions of SDES with singular drift and Sobolev diffusion coefficients. *Stochastic Processes and their Applications*, 115(11):1805–1818, 2005.

A. K. Zvonkin. A transformation of the phase space of a diffusion process that removes the drift. *Mathematics of the USSR-Sbornik*, 22(1):129, 1974.

Appendix A

Frequently Used Inequalities

- Doob's (Martingale) Maximal Inequality. For a martingale M, $p > 1$, and $T > 0$, if $\mathbb{E}(|M_T|^p) \leq \infty$, then we denote $M_t^* := \sup_{t\in[0,T]} |M_t|$ and have the following two inequalities: for a constant $C > 0$

$$\mathbb{P}(|M_t^*| > C) \leq \frac{\mathbb{E}(|M_T|^p)}{C^p},$$

and

$$\mathbb{E}(|M_t^*|^p) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(|M_T|^p).$$

- Burkholder–Davis–Gundy Inequality. For a local martingale M with $M_0 = 0$, any $p \geq 1$, and a stopping time τ, there exist two constant C_p, c_p such that

$$c_p\mathbb{E}\left([M]_\tau^{\frac{p}{2}}\right) \leq \mathbb{E}(|M_\tau^*|^p) \leq C_p\mathbb{E}\left([M]_\tau^{\frac{p}{2}}\right),$$

where $M_t^* := \sup_{s\leq t} |M_s|$ and [ ] stands for quadratic variation.

- Gronwall Inequality. Let $u : [t_0, t_1] \to \mathbb{R}^+$ be continuous and non-negative, and suppose u obeys the inequality

$$u(t) \leq A + \int_{t_0}^{t} B(s)u(s)ds,$$

for all $t \in [t_0, t_1]$, where $A \geq 0$ and $B : [t_0, t_1] \to \mathbb{R}^+$ is continuous. Then we have

$$u(t) \leq A \exp\left(\int_{t_0}^{t} B(s)ds\right)$$

for all $t \in [t_0, t_1]$.

Appendix B

Markovian Solution of Decoupled FBSDE

In this appendix, we provide sufficient conditions for (H1) and (H3) under the assumption that the forward SDE (3.4) has a unique strong solution.

It is well known that a decoupled FBSDE with a Lipschitz BSDE driver has a unique solution if the forward SDE has a unique strong solution. The Markovian property of the solution has been proved in Theorem 4.1 of El Karoui et al. (1997) and Theorem 14.5 of Barles and Lesigne (1997) under the assumption that the forward SDE has Lipschitz coefficients. The following theorem slightly generalizes the existence and uniqueness results in the sense that we do not require $b(t,x)$ and $\sigma(t,x)$ to be Lipschitz with respect to x and we allow linear growth of $\overline{f}(t,x,y,z)$ with respect to (y,z).

Proposition B.0.1. Let $\overline{f}(t,x,y,z) := f(t,x,y,z)+zg(t,x,y,z)$ for jointly $\mathcal{B}$-measurable functions $(f,g) : [0,T]\times\mathbb{R}^m\times\mathbb{R}^d\times\mathbb{R}^{d\times n}\to\mathbb{R}^d\times\mathbb{R}^n$. Assume the following conditions: there exist $C>0, p\geq 2$, and $r\geq\frac{1}{2}$ such that

- The forward SDE (3.4) has a unique strong solution F and $\mathbb{E}\sup_{t\in[0,T]}|F_t|^{pr}\leq C$
- $|h(x)|\leq C(1+|x|^r)$ for all $x\in\mathbb{R}^m$
- $\left|\overline{f}(t,x,y,z)\right|\leq C(1+|x|^r+|y|+|z|)$ for all $(t,x,y,z)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^d\times\mathbb{R}^{d\times n}$
- $\left|\overline{f}(t,x,y,z)-\overline{f}(t,x,y',z')\right|\leq C\left(|y-y'|+|z-z'|\right)$ for all $(t,x,y,z),(t,x,y',z')\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^d\times\mathbb{R}^{d\times n}$

Then, (H2) holds. Moreover, the solution $(F,U,V)\in\mathbb{S}^{pr}(\mathbb{R}^m)\times\mathbb{S}^p(\mathbb{R}^d)\times\mathbb{H}^p(\mathbb{R}^{d\times n})$ is unique.

Remark B.0.2. We do not need the nondegeneracy of σ in this proposition.

Proof. Note that, $\mathbb{E}|g(F_T)|^p\leq C^p\mathbb{E}(1+|F_T|^r)^p\leq 2^{p-1}C^p\left(1+\mathbb{E}|F_T|^{pr}\right)<\infty$ and

$$\mathbb{E}\left[\left(\int_0^T|\overline{f}(t,F_t,0,0)|^2dt\right)^{p/2}\right]\leq C^p\mathbb{E}\left[\left(\int_0^T 2\left(1+|F_t|^{2r}\right)dt\right)^{p/2}\right]$$
$$\leq 2^{p/2}C^pT^{p/2}\left(1+\mathbb{E}\sup_{t\in[0,T]}|F_t|^{pr}\right)<\infty.$$

Therefore, from the classical result of Pardoux and Peng (1990), the BSDE (3.5) has a unique solution $(Y,Z)\in\mathbb{S}^p(\mathbb{R}^d)\times\mathbb{H}^p(\mathbb{R}^{d\times n})$.

For $k = 0, 1, 2, ...$, let us define Borel measurable functions $(u_k, d_k) : [0, T] \times \mathbb{R}^m \ni (t, x) \mapsto (u_k(t, x), d_k(t, x)) \in \mathbb{R}^d \times \mathbb{R}^{d\times n}$ as follows: let $u_0 \equiv 0, d_0 \equiv 0$, $Y_t^{(k)} = u_k(t, F_t)$ and $Z_t^{(k)} = d_k(t, F_t)$ $dt \otimes d\mathbb{P}$-almost everywhere for $(Y^{(k)}, Z^{(k)})$, which is the unique solution of

$$dY_t^{(k)} := -\overline{f}(t, F_t, u_{k-1}(t, F_t), d_{k-1}(t, F_t))dt + Z_t^{(k)}dB_t; \qquad Y_T^{(k)} = g(F_T).$$

The well-definedness of $(u_k, d_k)_{k=0,1,2,...}$ is proved in Lemma B.0.3.It is well known that $Y^{(k)} \to Y$ in $\mathbb{S}^p$ and $Z^{(k)} \to Z$ in $\mathbb{H}^p$. If we let $u^i(t, x) := \limsup_{k\to\infty} u_k^i(t, x)$ and $d^{ij}(t, x) := \limsup_{k\to\infty} d_k^{ij}(t, x)$, where $u = (u^i)_{1\le i\le d}$ and $d = (d^{ij})_{1\le i\le d, 1\le j\le n}$, we have

$$\begin{aligned}
u^i(t, P_t) &= (\limsup_{k\to\infty} u_k^i)(t, P_t) = \limsup_{k\to\infty}(u_k^i(t, P_t)) = \lim_{k\to\infty} Y_t^{(k),i} = Y_t^i\\
d^{ij}(t, P_t) &= (\limsup_{k\to\infty} d_k^{ij})(t, P_t) = \limsup_{k\to\infty}(d_k^{ij}(t, P_t)) = \lim_{k\to\infty} Z_t^{(k),ij} = Z_t^{ij}.
\end{aligned}$$

Therefore, the claim is proved. □

Now, let us prove that $(u_k, d_k)_{k=0,1,2,...}$ are well defined.

Lemma B.0.3. For all $k = 0, 1, 2...$, we have that (u_k, d_k) are well-defined. Moreover, $u_k(\cdot, F_\cdot) \in \mathbb{S}^p(\mathbb{R}^d)$ and $d_k(\cdot, F_\cdot) \in \mathbb{H}^p(\mathbb{R}^{d\times n})$.

Proof. We prove this by mathematical induction. First, note that the claim holds true for $k = 0$. Assume that the claim holds for $k - 1 \ge 0$. It should be noted that $\mathbb{E}|g(F_T)|^p \le C^p\mathbb{E}(1 + |F_T|^r)^p \le 2^{p-1}C^p(1 + \mathbb{E}|F_T|^{pr}) < \infty$ and

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^T |\overline{f}(t, F_t, u_{k-1}(s, F_s), d_{k-1}(s, F_s))|^2 dt\right)^{p/2}\right]\\
&\le C^p\mathbb{E}\left[\left(\int_0^T 4\left(1 + |F_t|^{2r} + |u_{k-1}(s, F_s)|^2 + |d_{k-1}(s, F_s)|^2\right) dt\right)^{p/2}\right]\\
&\le 2^pC^pT^{p/2}\left(1 + \mathbb{E}\sup_{t\in[0,T]}|F_t|^{pr} + \mathbb{E}\sup_{t\in[0,T]}|u_{k-1}(t, F_t)|^p + \mathbb{E}\left[\left(\int_0^T |d_{k-1}(t, F_t)|^2 dt\right)^{p/2}\right]\right) < \infty.
\end{aligned}$$

Therefore, the BSDE

$$Y_t^{(k)} = g(F_T) + \int_t^T \overline{f}(s, F_s, u_{k-1}(s, F_s), d_{k-1}(s, F_s))ds - \int_t^T Z_s^{(k)}dB_s$$

has a unique solution such that $Y^{(k)} \in \mathbb{S}^p(\mathbb{R}^d)$ and $Z^{(k)} \in \mathbb{H}^p(\mathbb{R}^{d\times n})$. Note that, because $(t, F_t)_{t\ge0}$ is a Markov process, we know

$$\begin{aligned}
Y_t^{(k)} &= \mathbb{E}\left[g(F_T) + \int_t^T \overline{f}(s, F_s, u_{k-1}(s, F_s), d_{k-1}(s, F_s))ds\middle| \mathcal{F}_t\right]\\
&= \mathbb{E}\left[g(F_T) + \int_t^T \overline{f}(s, F_s, u_{k-1}(s, F_s), d_{k-1}(s, F_s))ds\middle| \mathcal{F}_t\right].
\end{aligned}$$

Moreover, by Proposition II.4.6 of Çinlar (2011), there exists a Borel measurable function $u_k : [0, T] \times \mathbb{R}^m \to \mathbb{R}^d$ such that $Y_t^{(k)} = u_k(t, F_t)$. On the other hand, note that

$$Y_t^{(k)} + \int_0^t \overline{f}(s, F_s, u_{k-1}(s, F_s), d_{k-1}(s, F_s))ds$$

is an additive martingale. By Theorem 6.27 of Çinlar et al. (1980), there exists a $\overline{\mathcal{B}}$-measurable function $\overline{d}_k : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{d\times n}$ such that

$$Y_t^{(k)} + \int_0^t \overline{f}(s, F_s, u_{k-1}(s, F_s), d_{k-1}(s, F_s))ds = Y_0^{(k)} + \int_0^t \overline{d}_k(s, F_s)dB_s.$$

Here $\overline{\mathcal{B}}$ is the σ-algebra of universally measurable sets. Let $G(t,\omega) := (t, F_t(\omega))$ and consider $\mu := (\lambda \otimes \mathbb{P}) \circ G^{-1}$, where λ is the Lebesgue measure on $[0,T]$. Then, μ is a finite measure on $[0,T] \times \mathbb{R}^m$; therefore, there exists a $\mathcal{B}$-measurable function $d_k : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{d\times n}$ such that

$$\mu\left(\left\{(t,x) \in [0,T] \times \mathbb{R}^m : \overline{d}_k(t,x) \neq d_k(t,x)\right\}\right) = 0.$$

This implies that $\overline{d}_k(t, F_t) = d_k(t, F_t)$ in $dt \otimes d\mathbb{P}$-almost everywhere sense. Therefore, the claim is proved. $\square$

Appendix C

Single Period Carbon Allowance Pricing

C.0.1 Problem Background

This problem arises from Climate Change, which is mainly caused by Green House Gases (GHGs), and one of the most well-known GHGs is carbon dioxide (CO_2). Due to the emissions of GHGs, the average surface temperature is much warmer than what it should be. To reduce the emissions of GHGs, one of the most efficient ways is to implement emissions trading markets globally.

The idea of building up emissions trading markets was triggered by the theory introduced by Coase (1960). The idea suggests that to reduce the emissions of GHGs in one economy, the governors should implement an emissions trading market, with setting a collective target on total emissions, or cap. Meanwhile, each emitter is given a part of total in the form of emission rights, which later are called the "allowance". If that target is set lower than what the total emissions would have been without the cap, then the system yields emission reductions. Notice that any participants in the market can trade their allowances, which means those who can reduce at low cost will do so, and therefore have an excess of allowances that they will be willing to sell. However, such market settings give us a critical question: what is the appropriate price for the emission allowance?

Before we discuss the question, one should realize that, the theoretical idea we introduced has been widely applied in reality. A popular type of emissions trading markets is carbon market, which is aiming at deducing the emissions of CO_2. From the signing of the Kyoto Protocol in 1996 (the first international carbon market system), the European Union (EU) started running a carbon market in 2005. Until now, tens of carbon markets have been or will be implemented worldwide, for instance, the world's biggest GHGs emitter, China, will start a national carbon market in the near future. For more information about emissions trading markets and related topic, we refer the book Chassagneux et al. (2017).

C.0.2 Mathematical Model for Carbon Market

To answer that question, we plan to apply FBSDEs to model a carbon market. We will consider a carbon market with N firms, and each of them will produce based on their own judgement. By solving an irregular coefficients FBSDE, we can obtain the total emissions of the market when the firms apply their optimal production strategies, and the allowance price in such equilibrium. Such FBSDE comes from a forward SDE for the aggregate emissions in the economy, and a BSDE for the allowance price. The following model is borrowed from Carmona et al. (2013).

The Model of Single Firm's Emission Dynamics

We firstly introduce an optimization problem of a single producer whose production generates emission of CO_2. We assume all firms in the market are small players such that their actions are negligible. Moreover, we assume that the price of the allowance $(Y_t)_{0 \le t \le T}$ is a martingale under the risk-neutral environment. Notice the notations may vary from section 3.3.2 and without loss of generality, we assume $\mathbb{P}$ to be a risk-neutral measure.

We denote the firm's total emissions up to time t is E_t, and

$$E_t = E_0 + \int_0^t (b_s - \xi_s) ds + \int_0^t \sigma_s dW_s, \tag{C.1}$$

where b stands for so-called *business-as-usual*, the rate of emission without carbon regulation, and ξ is the instantaneous rate of abatement chosen by the firm. Notice that each extra ton of abatement is more costly: changing the components is cheap, but changing the production machines is expensive. Therefore, we assume the cost of abatement $C(x, e) = c(x)g(e)$ is described by a strictly convex, continuously differentiable function $c(x) : \mathbb{R} \to \mathbb{R}$, which satisfies Inada-like conditions

$$c'(-\infty) = -\infty \quad \text{and} \quad c'(+\infty) = +\infty,$$

and an impact factor function $g(e) : \mathbb{R} \to \mathbb{R}^+/\{0\}$.

Remark C.0.1. One should notice that unlike Carmona et al. (2013), in which they assume the abatement cost $C(x, e) = c(x)$, we assume the abatement cost can be influenced from outside, for instance, Australian Emission Reduction Fund, which aims at reducing the firms' abatement costs under certain situations.

The firm controls its own abatement schedule ξ, and the quantity θ of the emission rights by trading in the allowance market. For those controls to be admissible, ξ and θ need only be progressively measurable processes and satisfy

$$\mathbb{E} \int_0^T [\xi_t^2 + \theta_t^2] dt < \infty.$$

We denote $\mathcal{A}$ to be the set of admissible controls (ξ, θ). Given the initial wealth x and an impact factor e, we obtain the terminal value X_T of the firm

$$X_T = X_T^{\xi,\theta} = x + \int_0^T \theta_t dY_t - \int_0^T C(\xi_t, e) dt - E_T Y_T.$$

Each of the terms on the right-hand side can be interpreted as: amount of trading in allowance market, abatement cost, and cost of the emission regulation.

We assume such firm's utility is given by an increasing, strictly concave function $U : \mathbb{R} \to \mathbb{R}$, which satisfys the Inada conditions

$$U'(-\infty) = +\infty \quad \text{and} \quad U'(+\infty) = 0.$$

Therefore, the corresponding optimization problem appears: find a pair of $(\xi^*, \theta^*) \in \mathcal{A}$, such that

$$V(x) = \sup_{(\xi,\theta) \in \mathcal{A}} \mathbb{E} U(X_T^{\xi,\theta}).$$

Proposition C.0.2. For a given $e \in \mathbb{R}$, the optimal abatement strategy of the firm is

$$\xi_t^* = (c')^{-1}\left(\frac{Y_t}{g(e)}\right).$$

Proof. Firstly, from integration by parts we can see that

$$\begin{aligned} E_T Y_T &= Y_T\left(E_0 + \int_0^T b_t dt + \int_0^T \sigma_t dW_t\right) - Y_T \int_0^T \xi_t dt \\ &= Y_T\left(E_0 + \int_0^T b_t dt + \int_0^T \sigma_t dW_t\right) - \int_0^T Y_t \xi_t dt - \int_0^T \left(\int_0^t \xi_s ds\right) dY_t. \end{aligned}$$

Therefore, we can denote $X_T = A_T^{\tilde{\theta}} + B_T^{\xi}$ with

$$A_T^{\tilde{\theta}} = \int_0^T \tilde{\theta}_t dY_t - Y_T\left(E_0 + \int_0^T b_t dt + \int_0^T \sigma_t dW_t\right),$$

where $\tilde{\theta}_t = \theta_t + \int_0^t \xi_s ds$, and

$$B_T^{\xi} = x - \int_0^T [C(\xi_t, e) - Y_t \xi_t] dt.$$

Notice that B^{ξ} and $A^{\tilde{\theta}}$ depend only on ξ and $\tilde{\theta}$ respectively, moreover, the set $\mathcal{A}$ of admissible controls is equivalently described by varying the couples (ξ, θ) or $(\xi, \tilde{\theta})$. Therefore,

$$\sup_{(\xi,\theta)\in\mathcal{A}} \mathbb{E}U(X_T^{\xi,\theta}) = \sup_{(\xi,\tilde{\theta})\in\mathcal{A}} \mathbb{E}U(A_T^{\tilde{\theta}} + B_T^{\xi}).$$

In this sense, one can perform the optimizations over $\tilde{\theta}$ and ξ separately. Since U is an increasing function, the quantity of B_T^{ξ} is maximized by choosing $\xi_t^* = (c')^{-1}\left(\frac{Y_t}{g(e)}\right)$. □

Remark C.0.3. In the Remark 2 of Carmona et al. (2013), the authors proved that in a complete market, for a given ξ^*, there exists a unique optimal investment strategy θ^*.

The Model of Carbon Market with N Firms

We now consider a risk-neutral carbon market with N firms. Particularly, we use a superscript ^{i} to emphasize the dependence upon the ith firm. In equilibrium, which means each firms in the market can apply their best production strategy, for the ith firm, we have

$$E_t^i = E_0^i + \int_0^t \left(b_s^i - (c^{i'})^{-1}\left(\frac{Y_s}{g^i(e)}\right)\right) ds + \int_0^t \sigma_s^i dW_s.$$

We denote the following

$$E_t = \sum_{i=1}^N E_t^i, \quad b_t = \sum_{i=1}^N b_t^i, \quad \sigma_t = \sum_{i=1}^N \sigma_t^i.$$

Moreover, we let the impact factor $e = E_t$ and denote

$$f(E_t, Y_t) = \sum_{i=1}^N (c^{i'})^{-1}\left(\frac{Y_t}{g^i(E_t)}\right).$$

To consist with our results, we further assume there exist deterministic functions $b_t = b(t, E_t)$ and $\sigma_t = \sigma(t, E_t)$, which are mappings $(b, \sigma) : [0, T] \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}$.

Moreover, based on martingale representation theorem, we know that there exists a progressively measurable process Z such that

$$dY_t = Z_t dW_t, \quad \text{and} \quad \mathbb{E} \int_0^T |Z_t|^2 dt < \infty.$$

Therefore, we could obtain a discontinuous coefficients FBSDE

$$\begin{aligned} dE_t &= (b(t, E_t) - f(E_t, Y_t))dt + \sigma(t, E_t)dW_t, \quad E_0 = x, \\ dY_t &= Z_t dW_t, \quad Y_T = \lambda \mathbf{1}_{[Y^{cap}, \infty)}(E_T), \end{aligned}$$

where the terminal condition states that any emissions excess the global emission target $Y^{cap} \in \mathbb{R}^+$ should be penalized by the fine λ.

As one can see, above FBSDE falls nicely into our framework. Our Theorem 3.1.3 can be applied for discontinuous b, σ, and f.

Appendix D

The Optimization Problem and the Hamiltonian System

Arguments in this appendix are borrowed from Carmona (2016).

To start with, we assume the actions or strategies are chosen from a measurable space $(A, \mathcal{A})$. Usually one can consider A as a subspace of Euclidean space $\mathbb{R}^n$. Particularly, we denote $\mathbb{A}$ the set of all the admissible actions or controls. In this sense, $\mathbb{A}$ will be the set of processes taking values in A which satisfy a set of admissibility conditions. We consider a augmented filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ and a m-dimensional Brownian motion is defined on it. We also denote by $\mathcal{P}$ the σ-field of $\mathbb{F}$-progressively measurable subsets of $\Omega \times [0, T]$.

We assume that the state of the system at time $t \in [0, T]$ is given by a d-dimensional Itô process satisfying

$$dX_t = b(t, X_t, \alpha_t)dt + \sigma(t, X_t, \alpha_t)dW_t, \quad X_0 = x' \in \mathbb{R}^d \tag{D.1}$$

where $(b, \sigma) : \Omega \times [0, T] \times \mathbb{R}^d \times A \to \mathbb{R}^d \times \mathbb{R}^{d\times m}$. Also, we need b and σ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{A}$-measurable, and assume SDE (D.1) can be uniquely solve for any given $\alpha \in \mathbb{A}$. (b and σ are Lipschitz and under linear growth with respect to x for all $t \in [0, T]$.)

An optimization problem aims at finding an admissible control $\alpha \in \mathbb{A}$ which minimizes a cost functional $J(\alpha)$. Usually the cost functional consists two parts: *terminal cost* and *running cost*. Terminal cost is often of the form $g(X_T)$, where $g : \Omega \times \mathbb{R}^d \to \mathbb{R}$ is $\mathcal{F}_T \times \mathcal{B}(\mathbb{R}^d)$-measurable. Meanwhile, running cost is often of the form $f(t, X_t, \alpha_t)$, where $f : \Omega \times [0, T] \times \mathbb{R}^d \times A \to \mathbb{R}$ is $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{A}$-measurable. Therefore, we define the cost functional J by

$$J(\alpha) = \mathbb{E}\left(\int_0^T f(t, X_t, \alpha_t)dt + g(X_T)\right).$$

The Hamiltonian of the system is a function of time t, possibly the random scenario ω, the state variable x, an action $\alpha \in \mathbb{A}$, and two new variables y and z (often called dual variables or covariables). Specifically, the corresponding Hamiltonian is the function $H : \Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \times A \to \mathbb{R}$ defined by

$$H(t, x, y, z, \alpha) = b(t, x, \alpha) \cdot y + \sigma(t, x, \alpha) \cdot z + f(t, x, \alpha),$$

where $\cdot$ denotes for the scalar product, that means $\sigma(t, x, \alpha) \cdot z = \mathrm{Tr}(\sigma(t, x, \alpha)^\intercal z)$.

For an admissible process α, we denote X be its corresponding controlled state process. We call adjoint processes associated with α for any solution (Y, Z) of the BSDE

$$dY_t = -\partial_x H(t, X_t, Y_t, Z_t, \alpha_t)dt + Z_t dW_t, \quad Y_T = \partial_x g(X_T). \tag{D.2}$$

This equation is called the adjoint equation associated with the admissible control α. Notice that when α and X are given, BSDE (D.2) is uniquely solvable since it is linear.

The related topics are introduced in chapter 4, Carmona (2016). Since the arguments and proofs are technical and tedious, here we only introduce the most important result.

Theorem D.0.1. (Necessary condition) Given the function $\alpha \mapsto J(\alpha)$ is Gâteaux differentiable, from BSDE (D.2) we could deduce

$$\begin{aligned}\frac{d}{d\varepsilon}J(\alpha+\varepsilon\beta)\Big|_{\varepsilon=0} &:= \lim_{\varepsilon\to 0}\frac{J(\alpha+\varepsilon\beta)-J(\alpha)}{\varepsilon}\\ &=\mathbb{E}\int_0^T \partial_\alpha H(t,X_t,Y_t,Z_t,\alpha_t)\beta_t dt,\end{aligned}$$

where β as the direction that can be thought of $\beta = \alpha' - \alpha$ for some other admissible control α'. And that could give us: if α^* is the optimal control, then for each $t \in [0,T]$ and $\forall \alpha \in A$,

$$H(t,X_t,Y_t,Z_t,\alpha_t^*) \le H(t,X_t,Y_t,Z_t,\alpha).$$

However, in most cases, we do not have the optimal strategy. So generally speaking, we are more interested in the following result. We assume:

- the terminal condition g is convex;
- for each $t \in [0,T]$, $H(t,x,y,z,\alpha)$ is convex with respect to (x,α).

In this sense, we would have

Theorem D.0.2. (Sufficient condition) If $H(t,X_t,Y_t,Z_t,\alpha_t^*) = \inf_{\alpha\in A} H(t,X_t,Y_t,Z_t,\alpha)$ a.s., then α^* is an optimal control, i.e., $J(\alpha^*) = \inf_{\alpha\in\mathbb{A}} J(\alpha)$. Here $(X_t,Y_t,Z_t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times m}$.

Proof. Let $\alpha' \in \mathbb{A}$ be a generic admissible control, and we denote X' the associated controlled process. By integration by parts and the convexity of the functions, we have

$$\begin{aligned}&\mathbb{E}g(X_T) - g(X_T')\\ &\quad\le \mathbb{E}\partial_x g(X_T)(X_T - X_T')\\ &\quad= \mathbb{E}Y_T(X_T - X_T')\\ &\quad= \mathbb{E}\int_0^T (X_t - X_t')dY_t + \int_0^T Y_t d[X_t - X_t'] + \int_0^T ((\sigma(t,X_t,\alpha_t^*) - \sigma(t,X_t',\alpha_t'))Z_t dt\\ &\quad= \mathbb{E}\int_0^T Y_t(b(t,X_t,\alpha_t^*) - b(t,X_t',\alpha_t')) + Z_t(\sigma(t,X_t,\alpha_t^*) - \sigma(t,X_t',\alpha_t'))\\ &\qquad - (X_t - X_t')\partial_x H(t,X_t,Y_t,Z_t,\alpha_t^*)dt.\end{aligned}$$

Similarly, we can deduce that

$$\begin{aligned}&\mathbb{E}\int_0^T f(t,X_t,\alpha_t^*) - f(t,X_t',\alpha_t')dt\\ &\quad= \mathbb{E}\int_0^T (H(t,X_t,Y_t,Z_t,\alpha_t^*) - H(t,X_t',Y_t',Z_t',\alpha_t'))\\ &\qquad - (Y_t(b(t,X_t,\alpha_t^*) - b(t,X_t',\alpha_t')) + Z_t(\sigma(t,X_t,\alpha_t^*) - \sigma(t,X_t',\alpha_t')))dt.\end{aligned}$$

Therefore, we can obtain

$$
\begin{aligned}
J(\alpha^*) - J(\alpha') &= \mathbb{E}g(X_T) - g(X'_T) + \mathbb{E}\int_0^T f(t, X_t, \alpha^*_t) - f(t, X'_t, \alpha'_t)dt \\
&\leq \mathbb{E}\int_0^T (H(t, X_t, Y_t, Z_t, \alpha^*_t) - H(t, X'_t, Y'_t, Z'_t, \alpha'_t)) - (X_t - X'_t)\partial_x H(t, X_t, Y_t, Z_t, \alpha^*_t)dt \\
&\leq 0.
\end{aligned}
$$

$\square$

Remark D.0.3. By the convexity of H, usually we could obtain $\alpha^* := \arg\min_{\alpha \in A} H(t, X_t, Y_t, Z_t, \alpha)$ by solving

$$\partial_\alpha H(t, X_t, Y_t, Z_t, \alpha) = 0.$$

Here $(X_t, Y_t, Z_t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times m}$. Particularly, under the Markovian settings, we would have a deterministic function u such that $\alpha^*_t = u(t, X_t, Y_t, Z_t)$.

Therefore, under the case of Remark D.0.3, by SDE (D.1) and BSDE (D.2), we can obtain a fully-coupled FBSDE

$$
\begin{aligned}
dX_t &= \partial_y H(t, X_t, Y_t, Z_t, u(t, X_t, Y_t, Z_t))dt + \partial_z H(t, X_t, Y_t, Z_t, u(t, X_t, Y_t, Z_t))dW_t, \quad X_0 = x' \\
dY_t &= -\partial_x H(t, X_t, Y_t, Z_t, u(t, X_t, Y_t, Z_t)) + Z_t dW_t, \quad Y_T = \partial_x g(X_T).
\end{aligned}
$$

If above FBSDE is uniquely solvable, then we can obtain $\alpha^* = u(t, X_t, Y_t, Z_t)$ as the unique optimal control to the optimization problem.