A Functional Approach to Backward Stochastic Dynamics

Gechun Liang
Exeter College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Michaelmas 2010

Acknowledgments

I am deeply indebted to my both supervisors Zhongmin Qian and Terry Lyons for their steady encouragement and uncountable suggestions during my DPhil study. In the past three years I benefited a lot from the numerous meetings with them and attending their seminars and workshops. Zhongmin is the one who guided me into the stochastic analysis, and into the area of backward stochastic differential equations in particular. His “working 16 hours a day without break” and his diligence in work always push me forwards. Terry is the one who introduced me into a brand new and exciting area: rough path theory. His broad knowledge in mathematics makes him able to answer almost every stochastic analysis question I asked him. During the preparation of our joint papers and my DPhil thesis, I benefited to a non-measurable extent from the advice, support and expertise of them.

My special thanks go to Lishang Jiang, who was my supervisor in Tongji University where I finished my master degree. It is him who guided me into mathematics after I graduated from financial department. His experience in teaching and his expertise in partial differential equations and financial mathematics helped me succeed in transferring from a finance undergraduate to a mathematics graduate. I am also grateful to the examiners of my transfer thesis in Oxford, Michael Monoyios and Xunyu Zhou, for their helpful comments and suggestions. During the past few years I learned a lot from the collaborations with my coauthors: Matteo Casserini (ETH), Vicky Henderson (Oxford), Jianwei Lin (Tongji), Eva Lütkebohmert (Freiburg), Xuemin Ren (Tongji) and Yajun Xiao (Freiburg). I want to thank them as well. I am very grateful to my DPhil thesis examiners, Thaleia Zariphopoulou and Huaizhong Zhao, for their detail comments and suggestions, and especially to Thaleia, for her very careful reading and the correction of my misuse of plurals. Thanks are also extended to Oxford-Man Institute, who fully sponsors my DPhil study in Oxford.

Last but not least, I am extremely thankful to my mother for her understanding and supporting, especially from April 2006 when my father passed away. Without her encouragement, I would never live and study abroad. This thesis should be dedicated to my parents.

Abstract

In this thesis, we consider a class of stochastic dynamics running backwards, so called backward stochastic differential equations (BSDEs) in the literature. We demonstrate BSDEs can be reformulated as functional differential equations defined on path spaces, and therefore solving BSDEs is equivalent to solving the associated functional differential equations. With such observation we can solve BSDEs on general filtered probability space satisfying the usual conditions, and in particular without the requirement of the martingale representation. We further solve the above functional differential equations numerically, and propose a numerical scheme based on the time discretization and the Picard iteration. This in turn also helps us solve the associated BSDEs numerically.

In the second part of the thesis, we consider a class of BSDEs with quadratic growth (QBSDEs). By using the functional differential equation approach introduced in this thesis and the idea of the Cole-Hopf transformation, we first solve the scalar case of such QBSDEs on general filtered probability space satisfying the usual conditions. For a special class of QBSDE systems (not necessarily scalar) in Brownian setting, we do not use such Cole-Hopf transformation at all, and instead introduce the weak solution method, which is to use the strong solutions of forward backward stochastic differential equations (FBSDEs) to construct the weak solutions of such QBSDE systems. Finally we apply the weak solution method to a specific financial problem in the credit risk setting, where we modify the Merton's structural model for credit risk by using the idea of indifference pricing. The valuation and the hedging strategy are characterized by a class of QBSDEs, which we solve by the weak solution method.

Contents

Chapter 1

Introduction

1.1 Backward stochastic differential equations (BSDEs)

Stochastic differential equations (SDEs) may be considered as dynamic systems perturbed by random signals which are often modeled by Brownian motion. One important class of SDEs considered in the literature are equations of Itô's type such as

$$dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t, \tag{1.1}$$

where $B = (B^1, \cdots, B^d)^T$ is a d-dimensional Brownian motion (The superscript T denotes matrix transposition). $a : R^+ \times R^n \to R^n$ and $\sigma : R^+ \times R^n \to R^{n\times d}$ are the coefficients satisfying certain regularity conditions. Itô gave the meaning of the solutions to SDE (1.1) by developing a theory of stochastic integration with respect to Brownian motion (so called Itô's stochastic integration), and obtained the solutions by specifying the data at a starting time $T > 0$.

Itô's stochastic integration requires that the solution X must be adapted to the underlying filtration generated by Brownian motion B (see [42] and [71]). We denote this filtration by $\{\mathcal{F}_t\}$. SDE (1.1) has to be interpreted as an integral equation:

$$X_t = X_T + \int_T^t a(s, X_s)ds + \int_T^t \sigma(s, X_s)dB_s.$$

If $t \geq T$, the right hand side of the above equality is $\mathcal{F}_t$-measurable, so it is possible to solve the equation forwards. However if $t \leq T$, then the right hand side of the above equality is only $\mathcal{F}_T$-measurable, and it is not always possible to solve SDE (1.1) backwards. Therefore SDE (1.1) might not be the correct form if one wants to solve the equations backwards.

On the other hand, nonlinear partial differential equations (PDEs) may give us some hints on how to propose a correct form of SDEs running backwards. Inddeed, let us suppose $\Phi(t, x)$ is a classical solution to the following reaction-diffusion equation:

$$\begin{cases} \partial_t \Phi + \frac{1}{2}\Delta\Phi + f(t, \Phi) = 0, \\ \Phi(T, x) = \phi(x), \qquad \text{for } (t, x) \in [0, T] \times R^d \end{cases} \tag{1.2}$$

with $f : [0,T] \times R^n \to R^n$ and $\phi : R^d \to R^n$ satisfying certain regularity conditions. By applying Itô's formula to the process $Y_t = \Phi(t, B_t)$, we have

$$Y_t = \phi(B_T) - \int_t^T \left\{ \partial_t \Phi(s, B_s) + \frac{1}{2} \Delta \Phi(s, B_s) \right\} ds - \int_t^T dM_s$$

with

$$M_t = \int_0^t \nabla \Phi(s, B_s) dB_s.$$

Using that $\Phi(t, x)$ satisfies PDE (1.2), we further obtain that

$$Y_t = \phi(B_T) + \int_t^T f(s, Y_s)\, ds - \int_t^T dM_s.$$

This tells us that if we want to run the equation backwards, in order to force the solutions of SDE being $\mathcal{F}_t$-adapted, we need a martingale term to correct the equation. (It is M in our case.) On the other hand, because we are in a Brownian filtration setting, according to martingale representation theorem (see [42] and [71]), M can further be written as a stochastic integral of a predictable representation process, denoted as Z, with respect to the Brownian motion B. In this way, we can also include Z_t in the coefficient f, which then becomes $f(t, Y_t, Z_t)$. A solution to a SDE running backwards should be a pair of processes (Y, M) or (Y, Z).

In the fundamental paper by Pardoux and Peng [62], the following type of SDEs were proposed and solved: on a Brownian filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$, for a given $\mathcal{F}_T$-measurable random variable ξ, they considered the equation:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s, \tag{1.3}$$

which is called a *backward stochastic differential equation*, or *BSDE* for short. Here, the coefficient $f(t, y, z)$ is called the *driver* of the BSDE, and it is assumed to be Lipschitz continuous in $(y, z) \in R^n \times R^{n \times d}$. A solution to BSDE (1.3) is a pair of $\mathcal{F}_t$-adapted processes (Y, Z). Next we give a short literature review of BSDEs.

(1) Development of backward stochastic differential equations

Bismut [7] [8] [9] firstly discovered and derived a class of SDEs with given terminal data when he used the Pontryagin maximum principle studying stochastic control problems. The equations in [7] [8] [9], which are linear, were extended and developed to a general nonlinear case by Pardoux and Peng [62] in the early 1990s. BSDE has become one of active areas in stochastic analysis since then.

Lots of efforts have been put to relax the conditions on the parameters of BSDEs. For example, Lepeltier and San Martin [52] relaxed the Lipschitz condition on the driver and worked on BSDEs with linear growth condition. Darling and Pardoux [20] extended the BSDE from constant terminal time to a stopping time. Various subsequent advances were summarized in the monograph [28], while the recent paper by El Karoui et al [26] is an updated summary of main results of BSDEs.

(2) Forward-backward stochastic differential equations

It is natural to consider BSDEs coupled with SDEs, which are called forward-backward stochastic differential equations (FBSDEs). Antonelli [1] is one of the first to consider FBSDEs. However the driver of his FBSDEs does not involve the martingale representation part Z, and he could only solve the equations with the time horizon small enough by employing the *method of contraction mapping*. FBSDEs were further investigated by Ma et al [54], who developed the *four-step-scheme* to solve the equations. They required the coefficients to be smooth enough in order to utilize the PDE theory. Hu and Peng [39], Yong [75] and Peng and Wu [68] adapted the *method of continuity* from the PDE theory, and solved FBSDEs with monotone coefficients. In [22] Delarue combined the first two methods and solved FBSDEs globally with Lipschitz continuous assumptions on the coefficients. Ma and Yong [55] is the first book to comprehensively study FBSDEs and their applications.

(3) Backward stochastic differential equations on general filtered probability space

Traditionally, BSDEs are always assumed to be defined on a probability space with Brownian filtration. Tang and Li [74] are one of the first to study BSDEs with random jumps, and Barles et al [4] discovered the connection between BSDEs with random jumps and a family of parabolic partial integral differential equations. Rong [72] proved the existence and uniqueness of this class of BSDEs under non-Lipschitz continuous assumption on the coefficients. The existence and uniqueness of the solutions to a more general class of BSDEs on the filtered potability space satisfying the *usual conditions* was proved by the author with Lyons and Qian in [49].

(4) Backward stochastic differential equations with quadratic growth

If the driver $f(t, y, z)$ of BSDEs has at most quadratic growth with respect to z, the nature of such stochastic equations completely changes. Such kind of BSDEs are usually called QBSDEs in the literature. QBSDEs were firstly solved by Kobylanski [45] based on the Cole-Hopf transformation adapted from the PDE theory. Her results were substantially developed and generalized by Briand and Hu in [11] and [12], where they extended to the equations with unbounded terminal value and convex driver, respectively. Later Morlais [60] generalized these results to the equations driven by martingales instead of Brownian motion. However most of the existing results of QBSDEs are only for the scalar case and heavily depend on the Cole-Hopf transformation. In [50] the author with Lyons and Qian firstly attempted to solve QBSDE systems by using the weak solution method, which does not employ the Cole-Hopf transformation.

(5) Reflected backward stochastic differential equations

El Karoui et al [27] introduced an obstacle to BSDEs such that the solution always stays above such obstacle. This so called reflected BSDE was further developed to double reflected barriers with Lipschitz continuous coefficients by Cvitanić and Karatzas [19], and by Hamadene et al [33] who relaxed the Lipschitz continuous assumption.

(6) Weak solutions of backward stochastic differential equations

Most of the existing literature concentrates on the solutions of BSDEs in the strong sense, i.e. the equations must be solved on *any given* probability space. One of the first attempts to introduce weak solutions for BSDEs was given by Buckdahn et al [13]. Buckdahn and Engelbert [14] further proved the uniqueness of weak solutions. However the driver of these BSDEs does not involve the martingale representation part Z. On the other hand, the notion of weak solutions for FBSDEs was introduced by Antonelli and Ma [2], and further developed by Ma et al in [57] [58] by employing the martingale problem approach. Recently the weak solution method was further investigated by the author with Lyons and Qian in [50] by considering a nonlinear version of Girsanov's transformation.

(7) Numerical solutions of backward stochastic differential equations

The early work of numerical solutions of BSDEs was basically based on the finite difference method for PDEs, which is not so surprising because of the intrinsic link between BSDEs and PDEs; see, for example, [24] and [53]. In [56], Ma and Zhang proved the path regularity of the martingale representation part of the solutions to BSDEs, which motivated Zhang [81] and Bouchard and Touzi [10] to develop the numerical algorithms for BSDEs based on the time discretization, which is in the same spirit of numerical solutions to SDEs. For further developments in this direction see [5] [18] [32] as well as the author's recent paper [16] with Casserini, where we considered the time discretization of functional differential equations associated to BSDEs.

(7) Applications of backward stochastic differential equations

The theory of backward stochastic differential equations has been found many applications in nonlinear partial differential equations, stochastic control, and mathematical finance.

(7.1) Connection to nonlinear partial differential equations

By using the solution of BSDEs, Peng [67] firstly derived a probabilistic representation (*Feynman-Kac representation*) for the solution of quasi-linear PDEs; see, also, Pardoux and Peng [63]. The connection of BSDEs with random jumps to a class of parabolic partial integral differential equations was discovered by Barles et al [4]. Recently Cheridito et al [17] represented the solutions of fully nonlinear PDEs by introducing a new class of BSDEs with second-order, so called second-order BSDEs. On the other hand, the connection between stochastic PDEs and a class of BSDEs with double noise, so called backward double stochastic differential equations (BDSDEs), was established by Pardoux and Peng [64]. This connection was further investigated by Bally and Matoussi [3], and was substantially developed by Zhao and Zhang [78] and [79].

(7.2) Connection to stochastic control problems

To derive the maximum principle as a necessary condition for optimal control problems, it is well known the associated adjoint equations must satisfy certain backward equations. For stochastic control problems, the corresponding adjoint equations are stochastic rather than deterministic. Indeed, Peng [66] established a general stochastic maximum principle by considering both first order and second order adjoint equations. The connection of stochastic control and BSDEs was also discussed by Quenez [70]. On the other hand, Kohlmann and Zhou [44] interpreted BSDEs as some stochastic control problems. A systematic account of BSDEs with their applications in stochastic control problems is summarized in the book [76] by Yong and Zhou.

(7.3) Applications to mathematical finance

In [25] Duffie and Epstein discovered a class of nonlinear BSDEs in their study of recursive utilities. Later El Karoui et al [30] applied BSDEs to option pricing problems and provided a general framework for the application of BSDEs in finance. In order to deal with constraints on the portfolios, Rouge and El Karoui [73] introduced a class of QBSDEs into finance. Hu et al [38] further studied this class of BSDEs in a more general setting. The application of BSDEs in credit risk modeling was studied by Bielecki and Jeanblanc [6], while El Karoui et al [29] used reflected BSDEs to study American option pricing problems. Finally, Cvitanić and Karatzas [19] applied such BSDEs to analyze stochastic differential games.

1.2 Main contribution of this thesis

1.2.1 Two central ideas of the thesis

The work herein is based on two key ideas. To ease the presentation, we first demonstrate them in a simple setting.

(1) Functional differential equation approach

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ be a filtered probability space satisfying the *usual conditions*. In this space, consider a semimartingale, say $(Y_t)_{t\in[0,T]}$, with the following decomposition:

$$Y_t = M_t - V_t, \quad \text{for } t \in [0,T],$$

where M is local martingale with respect to $\{\mathcal{F}_t\}$ and V is a finite variation process. If we further know the terminal data of Y, say $Y_T = \xi$ for some $\mathcal{F}_T$-measurable random variable ξ, then we also have $\xi = M_T - V_T$. Therefore,

$$\begin{cases} Y_t = E[\xi + V_T|\mathcal{F}_t] - V_t, \\ M_t = E[\xi + V_T|\mathcal{F}_t], \qquad \text{for } t \in [0,T]. \end{cases} \tag{1.4}$$

The above relationship indicates that given the terminal data of a semimartingale, we can always represent this semimartingale and its martingale part in terms of the associated finite variation part, i.e. Y *and* M *are affine functionals of* V. We will often write Y and M as $Y(V)$ and $M(V)$, when we want to emphasize their dependency on V.

Before we apply the above fundamental observation to a specific BSDE setting, let us look at such relationship from the *potential theory* point of view. For this, we first recall that for a given domain $D \subset R^d$, a real-valued l.s.c function f is called superharmonic in D if it satisfies

$$\int_{\partial B(0,r)} f(x+y)\sigma_r(dy) \le f(x), \quad \text{for } x \in D \text{ and } r < dist(x, \partial D),$$

where σ_r is the measure on the surface of the ball $\partial B(0,r)$, normalized to have total mass 1. The function f is called harmonic if the equality holds. If f is superharmonic in D, then there exists a unique positive Borel measure on D such that the following Riesz decomposition holds:

$$f(x) = h(x) + G_D\mu(x), \quad \text{for } x \in D,$$

where h is harmonic in D and $G_D\mu$ is the Green's potential of μ on D, i.e. $G_D\mu(x) = \int_D G(x,y)\mu(dy)$ with $G(x,y)$ as the Green's function of the Laplace equation in R^d. Given the boundary data of such superharmonic function f, by the above Riesz decomposition, the Green's potential $G_D\mu$ is often used to study f.

The probabilistic counterpart of the above Riesz decomposition is the Doob-Meyer decomposition. Indeed, for a supermartingale $(Y_t)_{t\in[0,T]}$ with càdlàg sample paths, the Doob-Meyer decomposition says that there exists a unique increasing predictably measurable process V starting from $V_0 = 0$ such that M defined by $M_t = Y_t + V_t$ for $t \in [0,T]$ is a martingale. The above relationship (1.4) we just established tells us exactly the same thing as in the potential theory: *given the terminal data of a supermartingale (which is a semimartingale) defined on* $[0,T]$*, we can study such supermartingale by investigating the increasing predictably measurable process* V.

Now we come back to the specific BSDE setting (1.3). We first observe that a solution to BSDE (1.3) must be a special semimartingale, which admits a unique canonical decomposition. Therefore it is natural to study BSDE (1.3) by investigating the equation that the finite variation part satisfies. In the thesis we establish the equivalence between BSDE (1.3) and the following equation:

$$V_t = \int_0^t f(s, Y(V)_s, Z(V)_s)ds \tag{1.5}$$

with

$$\begin{cases} Y(V)_t = E[\xi + V_T|\mathcal{F}_t] - V_t, \\ \displaystyle\int_t^T Z(V)_s dB_s = \xi + V_T - E[\xi + V_T|\mathcal{F}_t]. \end{cases}$$

Note that (1.5) is a functional differential equation defined on the path space. Due to this reason, we will call our approach the *functional differential equation approach*, or the *functional approach* for short, which explains the title of the thesis.

The functional differential equation approach may be made independent of the use of martingale representation, provided that one is willing to replace the predictable representation Z by a general functional of V, thus free us from the requirement of Brownian filtration setting. We may consider the correction martingale term M, rather than its predictable representation Z, as the part of solutions. BSDE (1.3) is reformulated as

$$\begin{cases} dY_t = -f(t, Y_t, \mathbf{L}(M)_t)dt + dM_t, \\ Y_T = \xi \end{cases} \tag{1.6}$$

with $\mathbf{L}$ as a prescribed operator defined on the martingale space. The martingale representation is then a special case of such operator $\mathbf{L}$. BSDE (1.6) is the representative BSDE that we will investigate in this thesis.

(2) Weak solution method

The solution to SDE (1.1) defined by Itô requires the solution exists on *any given* probability space. It is also called the *strong solution*. Modern theory of stochastic processes has been widely regarded as the theory of probability measures in Polish spaces (i.e. separable and complete) since the fundamental work of Prohorov in 1950s. If we regard a stochastic process as a probability distribution on sample path spaces, then there is no need to solve the equations on *any given* probability space, and we can certainly choose a convenient one to work with. Therefore another explanation of the solution to SDE (1.1) includes not only the stochastic process itself but also the associated probability space, and such solution is usually called the *weak solution*.

For BSDEs, we can also introduce the definition of weak solutions analogous to that of SDEs, which is nothing surprising. However, the essential point is how to utilize such a weak solution method. In this thesis, we mainly apply such method to solve a special class of BSDEs whose driver $f(t, y, z)$ has at most quadratic growth with respect to z. In the literature, such kind of BSDEs are called QBSDEs, as is was discussed in the last section.

Let us consider the following simple QBSDE, which is however sufficient to demonstrate our essential idea of the weak solution method:

$$\begin{cases} dY_t = -\left\{A_t(Z_t^1)^2 + B_t(Z_t^2)^2\right\} dt + Z_t^1 dW_t^1 + Z_t^2 dW_t^2, \\ Y_T = \phi(W_T^1, W_T^2), \end{cases} \tag{1.7}$$

where $W = (W^1, W^2)$ is a Brownian motion. Herein, $(A_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$ are two bounded and deterministic functions, and $\phi(x^1, x^2)$ is Lipschitz continuous in $(x^1, x^2) \in R^2$ and uniformly bounded. Note that unless $A_t = B_t$, the Cole-Hopf transformation does not help us to deduce the closed form solutions, though we know the solutions to (1.7) must exist by Kobylanski [45]. Nevertheless we can still approximate the solution by the Cole-Hopf transformation. In the author's paper [36] with Henderson, based on the Cole-Hopf transformation, we introduced the fractional-step (*prediction and correction*) from numerical analysis to obtain the analytical approximation of the solutions to QBSDEs which have the similar form to that of (1.7).

Next, we look at (1.7) from another point of view. The basic idea is to use the strong solution of an associated FBSDE to construct the weak solution of BSDE (1.7). Let us start with a Brownian motion $B = (B^1, B^2)$ on $(\Omega, \mathcal{F}, \mathbf{P})$ with the filtration $\{\mathcal{F}_t\}$ satisfying the *usual conditions*. Consider the system

$$\begin{cases} dY_t = Z_t^1 dB_t^1 + Z_t^2 dB_t^2, \\ Y_T = \phi(X_T^1, X_T^2), \\ dX_t^1 = A_t Z_t^1 dt + dB_t^1, \\ dX_t^2 = B_t Z_t^2 dt + dB_t^2, \\ (X_0^1, X_0^2) = (0, 0). \end{cases} \tag{1.8}$$

Note that FBSDE (1.8) is linear, so the coefficients are Lipschitz continuous. We can use the functional differential equation approach to solve the above (1.8), and obtain (Y, Z^1, Z^2). Using such solution (Y, Z^1, Z^2), we define a new probability measure $\mathbf{P}$ by

$$\frac{d\mathbf{P}}{d\mathbf{Q}} = \mathscr{E}(N),$$

where $\mathscr{E}(N)$ is the Doléans-Dade exponential of N:

$$N = -\int_0^\cdot A_s Z_s^1 dB_s^1 - \int_0^\cdot B_s Z_s^2 dB_s^2.$$

Under the new probability measure $\mathbf{P}$, by Girsanov's theorem, $W = (W^1, W^2)$ given by

$$\begin{cases} W_t^1 = B_t^1 - [B^1, N]_t = B_t^1 + \displaystyle\int_0^t A_s Z_s^1 ds \\ W_t^2 = B_t^2 - [B^2, N]_t = B_t^2 + \displaystyle\int_0^t B_s Z_s^2 ds \end{cases}$$

is a Brownian motion. Under the new probability measure $\mathbf{P}$ and with the new Brownian motion W, let us rewrite the backward equation in (1.8) as

$$\begin{aligned} dY_t &= Z_t^1(dW_t^1 - A_t Z_t^1 dt) + Z_t^2(dW_t^2 - B_t Z_t^2 dt) \\ &= -\left\{A_t (Z_t^1)^2 + B_t (Z_t^2)^2\right\} dt + Z_t^1 dW_t^1 + Z_t^2 dW_t^2 \end{aligned}$$

with $Y_T = \phi(W_T^1, W_T^2)$. Therefore, we used the strong solution of FBSDE (1.8) to construct the weak solution of BSDE (1.7). Based on this weak solution method, we study the following QBSDE system in the thesis:

$$\begin{cases} dY_t = -h(t, Y_t, Z_t)dt - Z_t f(t, Y_t, Z_t)dt + Z_t dW_t, \\ Y_T = \phi(W_T), \end{cases} \tag{1.9}$$

where $W = (W^1, \cdots, W^d)^T$ is a d-dimensional Brownian motion starting from $\mathbf{P}^x\{W_0 = x\} = 1$ for $x \in R^d$. The coefficients $h(t, y, z)$ and $f(t, y, z)$ are assumed to be Lipschitz continuous in $(y, z) \in R^n \times R^{n \times d}$. The terminal data $\phi(x)$ is Lipschitz continuous in $x \in R^d$ and uniformly bounded.

1.2.2 Thesis outline and future work

Since each chapter has its own introduction, we will not repeat them again here. Next, we only outline the main structure of the thesis and highlight future research directions.

Chapter 2 is the core of this thesis. Firstly, we demonstrate our functional differential equation approach by solving BSDEs on general filtered probability space. For the special linear case, we derive closed form solution. Specific examples of BSDEs on general filtered probability space which are not covered in the literature are also provided. Finally, we apply this approach to solve a family of FBSDEs.

In Chapter 3 we consider the numerical solutions of the associated functional differential equations based on time discretization. We solve a discretized time version of the functional differential equation, and prove the convergence rate of the corresponding functional difference equation to the functional differential equation. Finally we provide a numerical scheme based on Picard iteration.

In Chapter 4 we study QBSDEs using our functional differential approach. We prove the existence of solutions to scalar QBSDEs on a general filtered probability space by using the idea of the Cole-Hopf transformation and the functional differential equation approach. For a special class of QBSDE systems (not necessarily scalar) in a Brownian setting, we use the strong solutions of FBSDEs to construct the weak solutions of such QBSDE systems. Finally, we discuss the relationship between the weak and strong solutions of QBSDEs.

In Chapter 5 we apply the weak solution method and concentrate on its application in finance. In order to apply the weak solution method, we introduce the weak formulation of optimal portfolio problems. We then modify the classical Merton's structural model for credit risk by indifference pricing, and deduce an associated QBSDE as the characterization of the corporate bond price and the hedging strategy. The weak solution method is then employed to solve such QBSDE.

Finally, we provide in the Appendix some stochastic analysis tools particularly important to us and used throughout the whole thesis. They include, among others, special semimartingales, continuous BMO-martingales, Jacod-Yor's martingale representation and regular conditional probabilities.

In the PDE theory the regularity theory plays a central role. The regularity theory for BSDEs should also be expected to be very important. Specific question includes: given the terminal data which is driven by a SDE rather than a general random variable, can we say something about the path regularity of the martingale representation part of the solutions, and space regularity of the solutions? Both questions have positive answers under certain conditions. However in this thesis all of the regularity results are only cited from the existing literature as lemmas without proof. Specifically they are Lemma 2.15, 3.3, 3.4 and 4.14. Therefore the regularity theory of BSDEs could be and should be one of the future work for the author.

Chapter 2

A Functional Approach to BSDEs

2.1 Introduction

In this chapter, we present a new formulation of BSDEs and give a new approach to solve them. Traditionally, the martingale (or predictable) representation plays a crucial role in formulating and solving BSDEs, and this restricts the class of BSDEs one can solve, because the underlying filtered probability space is implicitly required to preserve the martingale representation property. A well known technique to prove the existence and uniqueness result is "applying Itô's formula to Y_t^2", which was pioneered by Pardoux and Peng in their fundamental paper [63].

We will show that BSDEs can be reformulated as functional differential equations defined on the path spaces. Using this fundamental idea, we can solve a class of BSDEs defined on filtered probability space satisfying the *usual conditions*, and in particular without the requirement of the martingale representation property. Solving such general BSDEs is equivalent to solving the associated functional differential equations on the path spaces.

We start with defining the spaces that we will work on. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ be a filtered probability space satisfying the *usual conditions*. For given $T > 0$, all of our processes are defined and considered only within the interval $[0, T]$. We define the following spaces:

(1) $\mathcal{C}([0,T]; R^d)$: the space of continuous and $\mathcal{F}_t$-adapted processes $(V_t)_{t\in[0,T]}$ valued in R^d such that $\sup_{t\in[0,T]} |V_t| \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$, and endowed with the norm:

$$||V||_{\mathcal{C}[0,T]} = E\left\{ \sup_{t\in[0,T]} |V_t|^2 \right\}^{1/2}.$$

(2) $\mathcal{H}^2([0,T]; R^d)$: the space of continuous and square integrable martingales $(M_t)_{t\in[0,T]}$ valued in R^d, and endowed with $\mathcal{H}^2$-norm:

$$||M||_{\mathcal{H}^2[0,T]} = E\left\{ [M,M]_T \right\}^{1/2}.$$

By the Burkholder-Davis-Gundy inequality, $||\cdot||_{\mathcal{H}^2[0,T]}$ is equivalent to the norm $||\cdot||_{\mathcal{C}[0,T]}$ for martingales, and is also equivalent to L_T^2-norm by Doob's inequality. See Appendix A.2 for further details.

(3) $\mathcal{S}([0,T];R^d)$: the space of $\mathcal{F}_t$-adapted processes $(Y_t)_{t\in[0,T]}$ with càdlàg samples path valued in R^d such that $\sup_{t\in[0,T]}|Y_t| \in L^2(\Omega,\mathcal{F}_T,\mathbf{P})$, and endowed with the norm $||\cdot||_{\mathcal{C}[0,T]}$.
(4) $\mathcal{M}^2([0,T];R^d)$: the space of square integrable martingales with càdlàg sample path valued in R^d together with the $\mathcal{H}^2$-norm, which is also equivalent to $||\cdot||_{\mathcal{C}[0,T]}$ and L^2_T-norm.
(5) $H^2([0,T];R^d)$: the space of predictably measurable processes Z endowed with the norm:

$$||Z||_{H^2[0,T]} = E\left\{\int_0^T |Z_s|^2 ds\right\}^{1/2}.$$

An example in $H^2([0,T];R^d)$ is the $\mathcal{F}_t$-adapted process $(Z_t)_{t\in[0,T]}$ with càglàd samples paths such that $||Z||_{H^2[0,T]} < \infty$. However, in general, there is no path regularity in $H^2([0,T];R^d)$.

2.1.1 Main results of this chapter

We focus on solving BSDE of the following form in a fixed duration $[0,T]$:

$$\begin{cases} dY_t = -f(t,Y_t,\mathbf{L}(M)_t)dt + dM_t, \\ Y_T = \xi \end{cases} \tag{2.1}$$

on a filtered probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbf{P})$ satisfying the *usual conditions*, where $f:[0,T]\times\Omega\times R^n\times R^m \to R^n$, given as the driver, is a Borel measurable function, $\xi \in L^2(\Omega,\mathcal{F}_T,\mathbf{P})$ is a R^n-valued random variable given as the terminal data, and $\mathbf{L}$ is a prescribed operator, possibly nonlinear, defined on the martingale space $\mathcal{M}^2([0,T];R^n)$ and valued in $H^2([0,T];R^m)$ (or in $\mathcal{S}([0,T];R^m)$). The detail definition of $\mathbf{L}$ will be presented in Section 2.2.

Definition 2.1 *An adapted solution to BSDE (2.1) is a pair of processes $(Y,M) \in \mathcal{S}([0,T];R^n)\times \mathcal{M}^2([0,T];R^n)$, where Y is a R^n-valued special semimartingale with M as the martingale part of its canonical decomposition, and satisfies*

$$Y_t = \xi + \int_t^T f(s,Y_s,\mathbf{L}(M)_s)ds - \int_t^T dM_s, \quad \textit{for } t\in[0,T],\ a.s.. \tag{2.2}$$

BSDE (2.1) is said to have a unique adapted solution if for any two adapted solutions (Y,M) and $(\bar{Y},\bar{M})$, it must hold that

$$\mathbf{P}\{(Y_t,M_t) = (\bar{Y}_t,\bar{M}_t), t\in[0,T]\} = 1.$$

We will often omit the term *adapted* if no confusion may arise. Now we describe our idea of how to reformulate BSDE (2.1) in terms of a functional differential equation. Note that if (Y,M) is a solution, then Y admits its canonical decomposition $Y_t = M_t - V_t$, where V is its finite variation part, which is predictably measurable and $V_0 = 0$. In particular we have $\xi = M_T - V_T$. A simple algebraic calculation gives the following crucial relationship:

$$\begin{cases} Y_t = E[\xi + V_T|\mathcal{F}_t] - V_t, \\ M_t = E[\xi + V_T|\mathcal{F}_t], \qquad \text{for } t\in[0,T],\ a.s.. \end{cases} \tag{2.3}$$

Basically, (2.3) tells us that as long as we know the terminal data of a semimartingale Y, we can always represent this semimartingale Y and its martingale part M in terms of its finite variation part V, together with the terminal data ξ. When we want to emphasize the dependency of Y and M on V, we will write $Y(V)$ and $M(V)$. The above fundamental observation motivates us to consider the equation that the finite variation part V satisfies:

$$V_t = \int_0^t f(s, Y(V)_s, \mathbf{L}(M(V))_s)ds, \quad \text{for } t \in [0, T], \ a.s. \tag{2.4}$$

with $Y(V)$ and $M(V)$ given by (2.3). Note that (2.4) is a functional differential equation defined on the path space. We will consider its solutions in the path space $\mathcal{C}([0, T]; R^n)$. Its uniqueness can be defined in an analogous way to Definition 2.1.

The next lemma shows that the solvability of BSDE (2.1) is equivalent to functional differential equation (2.4) with the given initial value $V_0 = 0$. Because of this, we will call our approach to solve BSDEs the *functional differential equation approach*, or the *functional approach* for short. This also explains the title of the thesis.

Lemma 2.2 *BSDE (2.1) admits a unique adapted solution* $(Y, M) \in \mathcal{S}([0, T]; R^n) \times \mathcal{M}^2([0, T]; R^n)$ *iff functional differential equation (2.4) admits a unique solution* $V \in \mathcal{C}([0, T]; R^n)$.

Proof. If (Y, M) is a solution to (2.1), by the canonical decomposition $Y_t = M_t - V_t$, we obtain the relationship (2.3). The integral equation (2.2) can be rewritten as

$$M_t - V_t = \xi + \int_t^T f(s, Y_s, \mathbf{L}(M)_s)dt - \int_t^T dM_s.$$

By taking conditional expectation with respect to $\mathcal{F}_t$ on both sides, we obtain

$$\begin{aligned} M_t - V_t &= E[\xi|\mathcal{F}_t] + E\left[\int_t^T f(s, Y_s, \mathbf{L}(M)_s)ds|\mathcal{F}_t\right] \\ &= E[\xi|\mathcal{F}_t] + E\left[\int_0^T f(s, Y_s, \mathbf{L}(M)_s)ds|\mathcal{F}_t\right] - \int_0^t f(s, Y_s, \mathbf{L}(M)_s)ds. \end{aligned}$$

By the uniqueness of the canonical decomposition of Y, and by identifying the martingale parts and finite variation parts of both sides, V must satisfy (2.4). Since $V_t = M_t - Y_t$, we also have

$$||V||_{\mathcal{C}[0,T]} \leq ||M||_{\mathcal{C}[0,T]} + ||Y||_{\mathcal{C}[0,T]} < \infty.$$

On the other hand, if V is a solution to (2.4), the first equation in (2.3) can be rewritten as

$$\begin{aligned} Y_t &= E[\xi + V_T|\mathcal{F}_t] - V_t \\ &= M_t - \int_0^t f(s, Y_s, \mathbf{L}(M)_s)ds, \end{aligned}$$

and thus (Y, M) must satisfy (2.2). By the relationship (2.3) again, we have

$$\begin{aligned} ||M||_{\mathcal{C}[0,T]} &= E\left\{\sup_{t\in[0,T]} |E[\xi + V_T|\mathcal{F}_t]|^2\right\}^{1/2} \\ &\leq E\left\{\sup_{t\in[0,T]} E[|\xi + V_T|^2|\mathcal{F}_t]\right\}^{1/2} \\ &\leq 2\sqrt{E|\xi + V_T|^2} \\ &\leq 2\sqrt{E|\xi|^2} + 2||V||_{\mathcal{C}[0,T]} < \infty, \end{aligned}$$

and

$$||Y||_{\mathcal{C}[0,T]} \leq ||M||_{\mathcal{C}[0,T]} + ||V||_{\mathcal{C}[0,T]} < \infty.$$

To prove the uniqueness, if (2.1) admits a unique solution, and suppose V and $\bar{V}$ are two solutions of (2.4), then (Y, M) given by $Y = Y(V)$, $M = M(V)$ and $(\bar{Y}, \bar{M})$ given by $Y = Y(\bar{V})$, $M = M(\bar{V})$ are two solutions to (2.1), which is a contraction.

On the other hand, if (2.4) admits a unique solution, and suppose (Y, M) and $(\bar{Y}, \bar{M})$ are two solutions of (2.1), denote V and $\bar{V}$ as the finite variation parts of the canonical decomposition of Y and $\bar{Y}$ respectively. But then V and $\bar{V}$ are two solutions to (2.4), which is a contraction. ■

By combining Lemma 2.2 with Theorem 2.5 in this chapter, we obtain the following theorem which is the main result of this chapter:

Theorem 2.3 *If the driver f satisfies the Lipschitz condition 2.1, the operator $\mathbf{L}$ satisfies the compatibility condition 2.2 (See Section 2.2 for the conditions 2.1 and 2.2), and $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$, then BSDE (2.1) admits a unique adapted solution $(Y, M) \in \mathcal{S}([0,T]; R^n) \times \mathcal{M}^2([0,T]; R^n)$.*

The chapter is organized as follows. Section 2.2 is devoted to the solvability of functional differential equation (2.4). In Section 2.2.1, we prove the existence and uniqueness of solutions to (2.4). Our strategy of the proof is we first solve the equation locally by fixed point arguments, and then we obtain the global solution by shifting the paths of V.

Section 2.3 contains some applications of our functional differential equation approach. In Section 2.3.1 we derive the closed form solutions for linear functional differential equations. In Section 2.3.2 we provide some specific examples of the operator $\mathbf{L}$, and we will see that the classical BSDE in the literature is just a special form of our BSDE (2.1). Finally in Section 2.3.3 we extend our approach to the BSDEs coupled with forward SDEs (FBSDEs).

The chapter is mainly adapted from the author's paper [49] with Lyons and Qian, except Section 2.3.3, which is based on the author's paper [50] with Lyons and Qian.

2.2 BSDEs on general filtered probability space

2.2.1 Solving functional differential equations

In this subsection, we solve BSDE (2.1) by considering functional differential equation (2.4). We first impose some conditions on the driver f and the operator $\mathbf{L}$.

Condition 2.1 *(Lipschitz) For any $(y, z) \in R^n \times R^m$, the driver $f(t, y, z)$ is predictably measurable with $f(\cdot, 0, 0) \in H^2([0, T]; R^n)$. Moreover there exists a constant C_1 such that*

$$|f(t, y, z) - f(t, \bar{y}, \bar{z})| \leq C_1(|y - \bar{y}| + |z - \bar{z}|)$$

for $t \in [0, T]$, $y, \bar{y} \in R^n$ and $z, \bar{z} \in R^m$, a.s..

Next we give conditions for the operator $\mathbf{L}$. Generally speaking, $\mathbf{L}$ must behave like a local operator and only depend on the increment of martingales. As it is pointed out in [49], $\mathbf{L}$ must satisfy *Lipschitz continuity*, *differential property* and *local-in-time property.*

Definition 2.4 *For any subinterval $[t_1, t_2] \subset [0, T]$, define an operator $\mathbf{L}_{[t_1,t_2]} : \mathcal{M}^2([t_1, t_2]; R^n) \to H^2([t_1, t_2]; R^m)$ (or $\mathcal{S}([t_1, t_2]; R^m)$) such that*
(1) *$\mathbf{L}_{[t_1,t_2]}$ is Lipschitz continuous, i.e. there exits a constant C_2 such that*

$$||\mathbf{L}_{[t_1,t_2]}(M) - \mathbf{L}_{[t_1,t_2]}(\bar{M})||_{H^2[t_1,t_2]} \leq C_2||M - \bar{M}||_{C[t_1,t_2]}$$

(or

$$||\mathbf{L}_{[t_1,t_2]}(M) - \mathbf{L}_{[t_1,t_2]}(\bar{M})||_{C[t_1,t_2]} \leq C_2||M - \bar{M}||_{C[t_1,t_2]})$$

for M, $\bar{M} \in \mathcal{M}^2([t_1, t_2]; R^n)$;
(2) *for any $\mathcal{F}_{t_1}$-measurable random variable V_{t_1},*

$$\mathbf{L}_{[t_1,t_2]}(M)_t = \mathbf{L}_{[t_1,t_2]}(M - V_{t_1})_t, \quad \textit{for } t \in [t_1, t_2].$$

The operator $\mathbf{L}$ is defined as following, which is based on the operator $\mathbf{L}_{[t_1,t_2]}$ defined above,

Condition 2.2 *(Compatibility) For a given partition $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ with the mesh:*

$$|\pi| := \max_{1 \leq i \leq N} |t_i - t_{i-1}| \leq \frac{1}{16C_1^2(1 + C_2)^2} \wedge 1,$$

the operator $\mathbf{L} : \mathcal{M}^2([0, T]; R^n) \to H^2([0, T]; R^m)$ (or $\mathcal{S}([0, T]; R^m)$) satisfies the following condition: for $i = 1, \cdots, N$,

$$\mathbf{L}(M)_t = \mathbf{L}_{[t_{i-1},t_i]}(M)_t, \quad \textit{for } t \in [t_{i-1}, t_i].$$

Therefore the operator $\mathbf{L}$ may be better written as $\mathbf{L}^\pi$, if one wants to emphasize the dependency of $\mathbf{L}$ on a particular partition π, and functional differential equation (2.4) in fact means

$$\begin{aligned} V_t = & \int_0^t f(s, Y(V)_s, \mathbf{L}(M(V))_s) ds \\ = & \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} f\left(s, Y(V)_s, \mathbf{L}_{[t_{j-1},t_j]}(M(V))_s\right) ds \\ & + \int_{t_{i-1}}^{t} f\left(s, Y(V)_s, \mathbf{L}_{[t_{i-1},t_i]}(M(V))_s\right) ds, \quad \text{for } t \in [t_{i-1}, t_i], \ a.s.. \end{aligned} \tag{2.5}$$

Some specific examples of $\mathbf{L}$ are provided in Section 2.3.2. Next we state our main theorem of this section, and the rest of the section is devoted to its proof.

Theorem 2.5 *If the driver f satisfies Lipschitz condition 2.1, the operator $\mathbf{L}$ satisfies compatibility condition 2.2, and $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$, then functional differential equation (2.4) admits a unique solution $V \in \mathcal{C}([0,T]; R^n)$.*

2.2.1.1 Local solution by fixed point argument

We consider functional differential equation (2.4) on the interval $[\tau, T]$ with $T - \tau \leq |\pi|$. In other words,

$$V_t = \int_\tau^t f(s, Y(V)_s, \mathbf{L}(M(V))_s) ds, \quad \text{for } t \in [\tau, T], \ a.s. \tag{2.6}$$

with

$$\begin{cases} Y(V)_t = E[\xi + V_T | \mathcal{F}_t] - V_t, \\ M(V)_t = E[\xi + V_T | \mathcal{F}_t]. \end{cases}$$

The nonlinear mapping defined by (2.6) is denoted by $\mathbb{L}$, and we will prove $\mathbb{L}$ admits a unique fixed point in $\mathcal{C}([\tau, T]; R^n)$. Indeed, for any V,$\bar{V} \in \mathcal{C}([\tau, T]; R^n)$, we have

$$\begin{aligned} ||\mathbb{L}(V) - \mathbb{L}(\bar{V})||_{\mathcal{C}[\tau,T]} = & E\left\{ \sup_{t\in[\tau,T]} |\mathbb{L}(V)_t - \mathbb{L}(\bar{V})_t|^2 \right\}^{1/2} \\ = & E\left\{ \sup_{t\in[\tau,T]} \left| \int_\tau^t f(s, Y(V)_s, \mathbf{L}(M(V))_s) - f(s, Y(\bar{V})_s, \mathbf{L}(M(\bar{V}))_s) ds \right|^2 \right\}^{1/2} \\ \leq & \sqrt{E\left(\int_\tau^T |f(s, Y(V)_s, \mathbf{L}(M(V))_s) - f(s, Y(\bar{V})_s, \mathbf{L}(M(\bar{V}))_s)| ds \right)^2} \\ \leq & \sqrt{T-\tau} \sqrt{E \int_\tau^T \left| f(s, Y(V)_s, \mathbf{L}(M(V))_s) - f(s, Y(\bar{V})_s, \mathbf{L}(M(\bar{V}))_s) \right|^2 ds} \\ \leq & C_1 \sqrt{T-\tau} \sqrt{E \int_\tau^T \left(|Y(V)_s - Y(\bar{V})_s| + |\mathbf{L}(M(V))_s - \mathbf{L}(M(\bar{V}))_s| \right)^2 ds} \\ \leq & C_1 \sqrt{T-\tau} \sqrt{E \int_\tau^T |Y(V)_s - Y(\bar{V})_s|^2 ds} \\ & + C_1 \sqrt{T-\tau} \sqrt{E \int_\tau^T |\mathbf{L}(M(V))_s - \mathbf{L}(M(\bar{V}))_s|^2 ds}. \end{aligned}$$

Note that

$$\begin{aligned}
\sqrt{E\int_\tau^T |Y(V)_s - Y(\bar V)_s|^2 ds} &= \sqrt{E\int_\tau^T \left|E[V_T-\bar V_T|\mathcal{F}_s]-(V_s-\bar V_s)\right|^2 ds} \\
&\le \sqrt{E\int_\tau^T \left|E[V_T-\bar V_T|\mathcal{F}_s]\right|^2 ds} + \sqrt{E\int_\tau^T \left|V_s-\bar V_s\right|^2 ds} \\
&\le \sqrt{E\int_\tau^T \left|V_T-\bar V_T\right|^2 ds} + \sqrt{E\int_\tau^T \left|V_s-\bar V_s\right|^2 ds} \\
&\le 2\sqrt{T-\tau}||V-\bar V||_{\mathcal{C}[\tau,T]} \\
&\le 2||V-\bar V||_{\mathcal{C}[\tau,T]}.
\end{aligned}$$

Since $s\in[\tau,T]$, by compatibility condition 2.2, $\mathbf{L}(M)_s = \mathbf{L}_{[\tau,T]}(M)_s$. If $\mathbf{L}_{[\tau,T]}$ valued in $H^2([\tau,T];R^m)$,

$$\begin{aligned}
\sqrt{E\int_\tau^T |\mathbf{L}(M(V))_s - \mathbf{L}(M(\bar V))_s|^2 ds} &\le C_2 E\left\{\sup_{t\in[\tau,T]} |M(V)_t - M(\bar V)_t|^2\right\}^{1/2} \\
&= C_2 E\left\{\sup_{t\in[\tau,T]} |E[V_T-\bar V_T|\mathcal{F}_t]|^2\right\}^{1/2} \\
&\le 2C_2\sqrt{E|V_T-\bar V_T|^2} \\
&\le 2C_2||V-\bar V||_{\mathcal{C}[\tau,T]}.
\end{aligned}$$

If $\mathbf{L}_{[\tau,T]}$ valued in $\mathcal{S}([\tau,T];R^m)$, in analogy to the above estimate, we obtain

$$\begin{aligned}
\sqrt{E\int_\tau^T |\mathbf{L}(M(V))_s - \mathbf{L}(M(\bar V))_s|^2 ds} &\le \sqrt{T-\tau}E\left\{\sup_{t\in[\tau,T]} |\mathbf{L}(M(V))_t - \mathbf{L}(M(\bar V))_t|^2\right\}^{1/2} \\
&\le 2C_2\sqrt{T-\tau}||V-\bar V||_{\mathcal{C}[\tau,T]} \\
&\le 2C_2||V-\bar V||_{\mathcal{C}[\tau,T]}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
||\mathbb{L}(V)-\mathbb{L}(\bar V)||_{\mathcal{C}[\tau,T]} &\le C_1\sqrt{T-\tau}(2+2C_2)||V-\bar V||_{\mathcal{C}[\tau,T]} \\
&\le \frac{1}{2}||V-\bar V||_{\mathcal{C}[\tau,T]}.
\end{aligned}$$

Thus $\mathbb{L}$ is a contraction mapping on $\mathcal{C}([\tau,T];R^n)$, so it admits a unique fixed point in $\mathcal{C}([\tau,T];R^n)$. The above result is summarized as the following lemma:

Lemma 2.6 *If the driver f satisfies Lipschitz condition 2.1, the operator $\mathbf{L}$ satisfies compatibility condition 2.2, and $\xi\in L^2(\Omega,\mathcal{F}_T,\mathbf{P})$, then functional differential equation (2.6) admits a unique solution $(V_t)_{t\in[\tau,T]}\in\mathcal{C}([\tau,T];R^n)$ provided $T-\tau\le|\pi|$.*

2.2.1.2 Global solution by "shifting the paths"

We extend to the global solution of (2.4) in this subsection. We make the partition π of $[0, T]$ according to the definition of operator $\mathbf{L}$ (See compatibility condition 2.2). Then for $i = 1, \cdots, N$ and $t \in [t_{i-1}, t_i]$, $\mathbf{L}(M)_t = \mathbf{L}_{[t_{i-1},t_i]}(M)_t$.

We start with the last interval $[t_{N-1}, t_N]$ together with the data ξ, and consider the following functional differential equation:

$$V_t = \int_{t_{N-1}}^{t} f\left(s, Y(N)_s, \mathbf{L}_{[t_{N-1},t_N]}(M(N))_s\right) ds, \quad \text{for } t \in [t_{N-1}, t_N],\ a.s.$$

with

$$\begin{cases} Y(N)_t = E[\xi + V_T | \mathcal{F}_t] - V_t, \\ M(N)_t = E[\xi + V_T | \mathcal{F}_t]. \end{cases}$$

By Lemma 2.6, there exists a unique solution $V(N) \in \mathcal{C}([t_{N-1}, t_N]; R^n)$. We also obtain processes $Y(N)_t$ and $M(N)_t$ for $t \in [t_{N-1}, t_N]$.

In general for $1 \le i \le N-1$, we repeat the same argument to each interval $[t_{i-1}, t_i]$ together with the data $Y(i+1)_{t_i}$, and consider the following functional differential equation:

$$V_t = \int_{t_{i-1}}^{t} f\left(s, Y(i)_s, \mathbf{L}_{[t_{i-1},t_i]}(M(i))_s\right) ds, \quad \text{for } t \in [t_{i-1}, t_i],\ a.s.$$

with

$$\begin{cases} Y(i)_t = E[Y(i+1)_{t_i} + V_{t_i} | \mathcal{F}_t] - V_t, \\ M(i)_t = E[Y(i+1)_{t_i} + V_{t_i} | \mathcal{F}_t]. \end{cases}$$

Again by Lemma 2.6, there exists a unique solution $V(i) \in \mathcal{C}([t_{i-1}, t_i]; R^n)$. We also obtain $Y(i)_t$ and $M(i)_t$ for $t \in [t_{i-1}, t_i]$.

Note however that naturally $V(i)$ for $i = 1, \cdots, N$ are not the real solutions of (2.4) on the corresponding interval $[t_{i-1}, t_i]$, because they start with the initial value $V(i)_{t_{i-1}} = 0$, but (2.4) starts from $V_{t_{i-1}} \equiv V_{t_{i-1}}$ on $[t_{i-1}, t_i]$. We need to shift the paths of $V(i)$ accordingly in order to match the starting point of the equation (2.4). For $t \in [t_{i-1}, t_i]$, following the convention $\sum_{j=1}^{0} = 0$, we define

$$\begin{cases} \bar{V}_t = V(i)_t + \sum_{j=1}^{i-1} V(j)_{t_j}, \\ \bar{Y}_t = Y(i)_t, \\ \bar{M}_t = M(i)_t + \sum_{j=1}^{i-1} V(j)_{t_j}. \end{cases} \tag{2.7}$$

Lemma 2.7 *The finite variation process $\bar{V}$ together with $(\bar{Y}, \bar{M})$ defined by (2.7) satisfy the relationship (2.3). Moreover $\bar{V}$ satisfies functional differential equation (2.4).*

Proof. Note that (2.3) can be rewritten via the following iteration form:

$$\begin{cases} Y(V)_s = E[Y(V)_{t_i} + V_{t_i} | \mathcal{F}_s] - V_s, \\ M(V)_s = E[Y(V)_{t_i} + V_{t_i} | \mathcal{F}_s] \end{cases} \tag{2.8}$$

where $s \in [t_{i-1}, t_i]$ and $i = 1, \cdots, N$, with the terminal data $Y(V)_{t_N} = \xi$.

By the definition of $\bar{Y}$, we have $\bar{Y}_{t_N} = Y(N)_{t_N} = \xi$. For $i = 1, \cdots, N$ and $s \in [t_{i-1}, t_i]$,

$$\begin{aligned}
\bar{Y}_s &= Y(i)_s \\
&= E[Y(i+1)_{t_i} + V(i)_{t_i} | \mathcal{F}_s] - V(i)_s \\
&= E[Y(i+1)_{t_i} + V(i)_{t_i} + \sum_{j=1}^{i-1} V(j)_{t_j} | \mathcal{F}_s] - V(i)_s - \sum_{j=1}^{i-1} V(j)_{t_j} \\
&= E[\bar{Y}_{t_i} + \bar{V}_{t_i} | \mathcal{F}_s] - \bar{V}_s,
\end{aligned}$$

where we used the fact $\sum_{j=1}^{i-1} V(j)_{t_j}$ is $\mathcal{F}_s$-measurable. Likewise, we obtain

$$\begin{aligned}
\bar{M}_s &= M(i)_s + \sum_{j=1}^{i-1} V(j)_{t_j} \\
&= E[Y(i+1)_{t_i} + V(i)_{t_i} | \mathcal{F}_s] + \sum_{j=1}^{i-1} V(j)_{t_j} \\
&= E[Y(i+1)_{t_i} + V(i)_{t_i} + \sum_{j=1}^{i-1} V(j)_{t_j} | \mathcal{F}_s] \\
&= E[\bar{Y}_{t_i} + \bar{V}_{t_i} | \mathcal{F}_s].
\end{aligned}$$

On the other hand, by the equation (2.5), (2.4) is equivalent to the following form:

$$\begin{aligned}
V_t = &\sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} f\left(s, Y(V)_s, \mathbf{L}_{[t_{j-1}, t_j]}(M(V))_s\right) ds \\
&+ \int_{t_{i-1}}^{t} f\left(s, Y(V)_s, \mathbf{L}_{[t_{i-1}, t_i]}(M(V))_s\right) ds, \quad \text{for } t \in [t_{i-1}, t_i],
\end{aligned}$$

where $(Y(V), M(V))$ is given by the iteration form (2.8), so we only need to verify $\bar{V}$ satisfies the above equation. Indeed, by the definition of $\bar{V}$, for $i = 1, \cdots, N$ and $t \in [t_{i-1}, t_i]$, we obtain

$$\begin{aligned}
\bar{V}_t &= \sum_{j=1}^{i-1} V(j)_{t_j} + V(i)_t \\
&= \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} f\left(s, Y(j)_s, \mathbf{L}_{[t_{j-1}, t_j]}(M(j))_s\right) ds + \int_{t_{i-1}}^{t} f\left(s, Y(i)_s, \mathbf{L}_{[t_{i-1}, t_i]}(M(i))_s\right) ds \\
&= \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} f\left(s, \bar{Y}_s, \mathbf{L}_{[t_{j-1}, t_j]}(\bar{M})_s\right) ds + \int_{t_{i-1}}^{t} f\left(s, \bar{Y}_s, \mathbf{L}_{[t_{i-1}, t_i]}(\bar{M})_s\right) ds,
\end{aligned}$$

where we used the conditions on the operator $\mathbf{L}$. This completes the proof. ∎

Lemma 2.8 *If functional differential equation (2.4) admits a solution in $\mathcal{C}([0, T]; R^n)$, then the solution must be unique.*

Proof. Suppose that V^1 and V^2 are two different solutions to (2.4) with the same terminal data ξ. Then, on the last interval $[t_{N-1}, t_N]$, V^1 satisfies

$$V_t^1 = V_{t_{N-1}}^1 + \int_{t_{N-1}}^{t} f(s, Y(V^1)_s, \mathbf{L}(M(V^1))_s) ds, \quad \text{for } t \in [t_{N-1}, t_N],$$

while V^2 satisfies

$$V_t^2 = V_{t_{N-1}}^2 + \int_{t_{N-1}}^t f(s, Y(V^2)_s, \mathbf{L}(M(V^2))_s)ds, \quad \text{for } t \in [t_{N-1}, t_N].$$

If we define $\bar{V}_t^1 = V_t^1 - V_{t_{N-1}}^1$ and $\bar{V}_t^2 = V_t^2 - V_{t_{N-1}}^2$, then by the compatibility condition 2.2 on the operator $\mathbf{L}$, both $\bar{V}^1$ and $\bar{V}^2$ satisfy

$$V_t = \int_{t_{N-1}}^t f(s, Y(V)_s, \mathbf{L}(M(V))_s)ds, \quad \text{for } t \in [t_{N-1}, t_N],$$

which admits a unique solution $V \in \mathcal{C}([t_{N-1}, t_N]; R^n)$ by Lemma 2.6, so $\bar{V}_t^1 = \bar{V}_t^2$ for $t \in [t_{N-1}, t_N]$, and

$$Y(V^1)_{t_{N-1}} = Y(\bar{V}^1)_{t_{N-1}} = Y(\bar{V}^2)_{t_{N-1}} = Y(V^2)_{t_{N-1}}.$$

Thus on the interval $[t_{N-1}, t_N]$, V^1 and V^2 are two solutions to (2.4) but still with the same data $Y(V^1)_{t_{N-1}} = Y(V^2)_{t_{N-1}}$. We continue this procedure until to $[0, t_1]$. Since $V_0^1 = V_0^2 = 0$, $\bar{V}_t^1 = \bar{V}_t^2$ then implies $V_t^1 = V_t^2$ for $t \in [0, t_1]$, which is a contradiction. ■

2.3 Some examples of BSDEs

2.3.1 Linear functional differential equations

We first consider the following linear functional differential equation whose driver does not depend on the operator $\mathbf{L}$:

$$V_t = \int_0^t \{A[E(\xi|\mathcal{F}_s) + E(V_T|\mathcal{F}_s) - V_s] + C\}ds, \tag{2.9}$$

where A and C are two constants, and $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$. By Lemma 2.2, the above functional differential equation is equivalent to the following linear BSDE:

$$\begin{cases} dY_t = -[AY_t + C]dt + dM_t, \\ Y_T = \xi. \end{cases}$$

Corollary 2.9 *There exists a unique solution $V \in \mathcal{C}([0,T]; R^n)$ to functional differential equation (2.9), which is given by*

$$V_t = \int_0^t e^{A(T-s)}[AE(\xi|\mathcal{F}_s) + C]ds.$$

Moreover, the unique adapted solution $(Y, M) \in \mathcal{S}([0,T]; R^n) \times \mathcal{M}^2([0,T]; R^n)$ to the associated linear BSDE is given by

$$Y_t = e^{A(T-t)}E(\xi|\mathcal{F}_t) + \int_t^T e^{A(T-s)}Cds$$

and

$$M_t = e^{A(T-t)}E(\xi|\mathcal{F}_t) + \int_0^T e^{A(T-s)}Cds + \int_0^t e^{A(T-s)}AE(\xi|\mathcal{F}_s)ds.$$

Proof. The basic idea is the variation of constants. First, consider the linear equation:

$$\begin{cases} d\bar{V}_t - A\bar{V}_t dt = (AE[\xi|\mathcal{F}_t] + C)dt, \\ V_0 = 0, \end{cases}$$

the solution of which is

$$\bar{V}_t = \int_0^t e^{\int_s^t Adu}(AE[\xi|\mathcal{F}_s] + C)ds.$$

Now suppose that the solution to (2.9) has the form:

$$V_t = g(t)\int_0^t e^{\int_s^t Adu}(AE[\xi|\mathcal{F}_s] + C)ds$$

Differentiate V_t against t yields

$$\begin{aligned} \frac{dV_t}{dt} = & \left(\frac{dg(t)}{dt} + Ag(t)\right) e^{\int_0^t Adu} \int_0^t e^{-\int_0^s Adu} AE[\xi|\mathcal{F}_s]ds \\ & + g(t)AE[\xi|\mathcal{F}_t] \\ & + \left(\frac{dg(t)}{dt} + Ag(t)\right) e^{\int_0^t Adu} \int_0^t e^{-\int_0^s Adu} Cds + g(t)C. \end{aligned} \tag{2.10}$$

On the other hand, substituting V_t into the right hand side of (2.9) and comparing it with (2.10), we obtain the system of equations:

$$\frac{dg(t)}{dt} + Ag(t) = Ag(T)e^{\int_t^T Adu} - Ag(t),$$

$$g(t) = 1 + g(T)e^{\int_0^T Adu}\int_t^T e^{-\int_0^s Adu}Ads,$$

and

$$\begin{aligned} & \left(\frac{dg(t)}{dt} + Ag(t)\right) e^{\int_0^t Adu} \int_0^t e^{-\int_0^s Adu} Cds + g(t)C \\ = & Ag(T)e^{\int_0^T Adu} \int_0^T e^{-\int_0^s Adu}Cds - Ag(t)e^{\int_0^t Adu}\int_0^t e^{-\int_0^s Adu}Cds + C. \end{aligned}$$

Solving the above system, we obtain

$$g(t) = e^{\int_t^T Adu}\left(e^{\int_t^T Adu} - \int_t^T Ae^{\int_t^s Adu}ds\right) = e^{\int_t^T Adu}.$$

Therefore the solution to (2.9) is $V_t = e^{\int_t^T Adu}\bar{V}_t$, which is obviously in $\mathcal{C}([0,T];R^n)$. Now given the expression of V, (Y,M) can be obtained from the relationship (2.3). ■

The second linear case we consider is the linear BSDE with deterministic terminal data. Then the BSDE reduces to an ODE, which is an important fact we will rely on when we use the comparison principle in Chapter 4.

Corollary 2.10 *Let $(A_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$ be two bounded and deterministic functions, and let $(C_t)_{t\in[0,T]}$ be a square integrable and deterministic function. If the operator* **L**, *besides satisfying* *compatibility condition 2.2, sends a constant martingale to zero, and if the terminal data is given as a constant vector $\alpha \in R^n$, then the following BSDE*

$$\begin{cases} dY_t = -[A_t Y_t + B_t \mathbf{L}(M)_t + C_t]dt + dM_t, \\ Y_T = \alpha \in R^n \end{cases} \tag{2.11}$$

admits a unique adapted solution (Y, M). Moreover, the solution Y can be obtain by solving the ODE:

$$X_t = \alpha + \int_t^T (A_s X_s + C_s)ds$$

whose unique solution is give by $X_t = \alpha \exp(\int_t^T A_s ds) + \int_t^T C_s \exp(\int_t^s A_u du)ds$, and the solution M is a constant martingale given by $M \equiv X_0$.

Proof. First it is easy to verify (X, M) is an adapted solution to BSDE (2.11). By the existence and uniqueness theorem 2.3, (X, M) is in fact the only candidate of the adapted solutions. ■

2.3.2 BSDEs associated with operator L

In this subsection we provide some specific examples of the operator $\mathbf{L} = \mathbf{L}^\pi$ (The superscript π is to emphasize the dependency on the partition π). We first observe that since $\mathbf{L}^\pi$ is constructed based on $\mathbf{L}_{[t_{i-1},t_i]}$ for $i = 1, \cdots, N$ by the compatibility condition 2.2, we only need to specify $\mathbf{L}_{[t_{i-1},t_i]}$.

Case (1) Define $\mathbf{L}^1_{[t_{i-1},t_i]} : \mathcal{M}^2([t_{i-1}, t_i]; R^n) \to \mathcal{M}^2([t_{i-1}, t_i]; R^n) \subset \mathcal{S}([t_{i-1}, t_i]; R^n)$ by

$$\mathbf{L}^1_{[t_{i-1},t_i]}(M)_t = M_t - M_{t_{i-1}}$$

for $M \in \mathcal{M}([0, T]; R^n)$ and $t \in [t_{i-1}, t_i]$. Then $\mathbf{L}^1_{[t_{i-1},t_i]}$ obviously satisfies Definition 2.4 with $C_2 = 2$. So functional differential equation (2.4) (or (2.5)) turns out to be

$$\begin{aligned} V_t = & \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} f\left(s, Y(V)_s, M(V)_s - M(V)_{t_{j-1}}\right) ds \\ & + \int_{t_{i-1}}^{t} f\left(s, Y(V)_s, M(V)_s - M(V)_{t_{i-1}}\right) ds, \quad \text{for } t \in [t_{i-1}, t_i],\ a.s., \end{aligned}$$

which is equivalent to the BSDE:

$$\begin{aligned} Y_t = \xi + & \sum_{j=i+1}^{N} \int_{t_{j-1}}^{t_j} f\left(s, Y_s, M_s - M_{t_{j-1}}\right) ds \\ & + \int_{t}^{t_i} f\left(s, Y_s, M_s - M_{t_{i-1}}\right) ds - \int_t^T dM_s, \quad \text{for } t \in [t_{i-1}, t_i],\ a.s.. \end{aligned}$$

Case (2) Define $\mathbf{L}^2_{[t_{i-1},t_i]} : \mathcal{M}^2([t_{i-1}, t_i]; R^n) \to \mathcal{C}([t_{i-1}, t_i]; R^m) \subset \mathcal{S}([t_{i-1}, t_i]; R^m)$ by

$$\mathbf{L}^2_{[t_{i-1},t_i]}(M)_t = \int_{t_{i-1}}^{t} g(M_s - M_{t_{i-1}})ds$$

for $M \in \mathcal{M}^2([0,T];R^n)$, $t \in [t_{i-1}, t_i]$, and $g : R^n \to R^m$ being a Lipschitz continuous function with Lipschitz constant C_3. Then, $\mathbf{L}^2_{[t_{i-1},t_i]}$ satisfies Definition 2.4. Indeed for M, $\bar{M} \in \mathcal{M}^2([0,T];R^n)$,

$$E\left\{\sup_{t\in[t_{i-1},t_i]}\left|\int_{t_{i-1}}^t g(M_s - M_{t_{i-1}})ds - \int_{t_{i-1}}^t g(\bar{M}_s - \bar{M}_{t_{i-1}})ds\right|^2\right\}^{1/2}$$
$$\leq C_3\sqrt{E\left(\int_{t_{i-1}}^{t_i}|(M_s - \bar{M}_s) - (M_{t_{i-1}} - \bar{M}_{t_{i-1}})|ds\right)^2}$$
$$\leq C_3\sqrt{|\pi|}\sqrt{E\int_{t_{i-1}}^{t_i}|(M_s - \bar{M}_s) - (M_{t_{i-1}} - \bar{M}_{t_{i-1}})|^2 ds}$$
$$\leq 2C_3|\pi|E\left\{\sup_{t\in[t_{i-1},t_i]}|M_t - \bar{M}_t|^2\right\}^{1/2},$$

so we can choose $C_2 = 2C_3|\pi|$. The functional differential equation (2.4) (or (2.5)) turns out to be

$$V_t = \sum_{j=1}^{i-1}\int_{t_{j-1}}^{t_j} f\left(s, Y(V)_s, \int_{t_{j-1}}^s g(M(V)_u - M(V)_{t_{j-1}})du\right)ds$$
$$+ \int_{t_{i-1}}^t f\left(s, Y(V)_s, \int_{t_{i-1}}^s g(M(V)_u - M(V)_{t_{i-1}})du\right)ds, \quad \text{for } t \in [t_{i-1}, t_i],\ a.s.,$$

which is equivalent to the following BSDE:

$$Y_t = \xi + \sum_{j=i+1}^{N}\int_{t_{j-1}}^{t_j} f\left(s, Y_s, \int_{t_{j-1}}^s g(M_u - M_{t_{j-1}})du\right)ds$$
$$+ \int_t^{t_i} f\left(s, Y_s, \int_{t_{i-1}}^s g(M_u - M_{t_{i-1}})du\right)ds - \int_t^T dM_s, \quad \text{for } t \in [t_{i-1}, t_i],\ a.s..$$

Case (3) For $M \in \mathcal{M}^2([0,T];R^n)$, we define $\mathbf{M}^i_t = M_t - M_{t_{i-1}}$ for $t \in [t_{i-1}, t_i]$, and $\mathbf{M}^i_t = 0$ for $t \in [0, t_{i-1}]$. So $\mathbf{M}^i$ is a shift of M, and

$$\sup_{t\in[0,t_i]}|\mathbf{M}^i_t|^2 = \sup_{t\in[t_{i-1},t_i]}|M_t - M_{t_{i-1}}|^2 \leq 2\sup_{t\in[t_{i-1},t_i]}|M_t|^2.$$

Define $\mathbf{L}^3_{[t_{i-1},t_i]} : \mathcal{M}^2([t_{i-1},t_i];R^n) \to \mathcal{S}([t_{i-1},t_i];R^n)$ by

$$\mathbf{L}^3_{[t_{i-1},t_i]}(M)_t = \sqrt{[\mathbf{M}^i, \mathbf{M}^i]_t}$$

for $M \in \mathcal{M}^2([0,T];R^n)$, $t \in [t_{i-1}, t_i]$, and $[\cdot,\cdot]$ denotes the quadratic variation process. Then $\mathbf{L}^3_{[t_{i-1},t_i]}$ satisfies Definition 2.4. Indeed, for M, $\bar{M} \in \mathcal{M}^2([0,T];R^n)$, according to Kunita-Watanabe's inequality, we have

$$[M, \bar{M}]_t \leq \sqrt{[M,M]_t}\sqrt{[\bar{M},\bar{M}]_t}.$$

It follows that

$$
\begin{aligned}
\left|\sqrt{[M,M]_t}-\sqrt{[\bar M,\bar M]_t}\right| &= \left|\frac{[M,M]_t-[\bar M,\bar M]_t}{\sqrt{[M,M]_t}+\sqrt{[\bar M,\bar M]_t}}\right| \\
&= \left|\frac{[M-\bar M,M]_t+[\bar M,M-\bar M]_t}{\sqrt{[M,M]_t}+\sqrt{[\bar M,\bar M]_t}}\right| \\
&\le \sqrt{[M-\bar M,M-\bar M]_t},
\end{aligned}
$$

and thus

$$
\begin{aligned}
& E\left\{\sup_{t\in[t_{i-1},t_i]}\left|\sqrt{[\mathbf{M}^i,\mathbf{M}^i]_t}-\sqrt{[\bar{\mathbf{M}}^i,\bar{\mathbf{M}}^i]_t}\right|^2\right\}^{1/2} \\
\le\; & E\left\{\sup_{t\in[t_{i-1},t_i]}[\mathbf{M}^i-\bar{\mathbf{M}}^i,\mathbf{M}^i-\bar{\mathbf{M}}^i]_t\right\}^{1/2} \\
=\; & \sqrt{E[\mathbf{M}^i-\bar{\mathbf{M}}^i,\mathbf{M}^i-\bar{\mathbf{M}}^i]_{t_i}} \\
\le\; & C_4 E\left\{\sup_{t\in[0,t_i]}|\mathbf{M}^i_t-\bar{\mathbf{M}}^i_t|^2\right\}^{1/2} \\
\le\; & 2C_4 E\left\{\sup_{t\in[t_{i-1},t_i]}|M_t-\bar M_t|^2\right\}^{1/2},
\end{aligned}
$$

where C_4 is the constant of Burkholder-Davis-Gundy inequality. Therefore we can choose $C_2 = 2C_4$, and functional differential equation (2.4) (or (2.5)) reduces to be

$$
\begin{aligned}
V_t = & \sum_{j=1}^{i-1}\int_{t_{j-1}}^{t_j} f\left(s, Y(V)_s, \sqrt{[\mathbf{M}^j(V),\mathbf{M}^j(V)]_s}\right) ds \\
& + \int_{t_{i-1}}^{t} f\left(s, Y(V)_s, \sqrt{[\mathbf{M}^i(V),\mathbf{M}^i(V)]_s}\right) ds, \quad \text{for } t\in[t_{i-1},t_i],\ a.s.,
\end{aligned}
$$

where $\mathbf{M}^i(V)_s$ is defined in an obvious way: $\mathbf{M}^i(V)_s = E[\xi+V_T|\mathcal{F}_s]-E[\xi+V_T|\mathcal{F}_{t_{i-1}}]$ for $s\in[t_{i-1},t_i]$, and $\mathbf{M}^i(V)_s = 0$ for $s\in[0,t_{i-1}]$. The above functional differential equation is equivalent to the following BSDE:

$$
\begin{aligned}
Y_t = \xi + & \sum_{j=i+1}^{N}\int_{t_{j-1}}^{t_j} f\left(s, Y_s, \sqrt{[\mathbf{M}^j,\mathbf{M}^j]_s}\right) ds \\
& + \int_t^{t_i} f\left(s, Y_s, \sqrt{[\mathbf{M}^i,\mathbf{M}^i]_s}\right) ds - \int_t^T dM_s, \quad \text{for } t\in[t_{i-1},t_i],\ a.s..
\end{aligned}
$$

Case (4) The following case is closely related to classical BSDEs considered in the literature. Let $X = (X^1,\cdots,X^d)^T$ be a d-dimensional orthogonal martingale with each of its components $X^j \in \mathcal{M}^2([0,T];R)$ on $(\Omega,\mathcal{F},\mathbf{P})$ with the natural filtration $\{\mathcal{F}^X_t\}$ after augumentation. Let $\mathbf{P}$ be an extremal point in $\mathcal{M}^2(\{X\})$. Then by Theorem A.9, $\mathcal{S}(\{X\}) = \mathcal{M}^2([0,T];R)$. That is for any $\mathcal{F}^X_t$-adapted martingale $M = (M^1,\cdots,M^n)^T \in \mathcal{M}^2([0,T];R^n)$, there exists a $R^{n\times d}$-valued predictably measurable process Z with

$$E\{\sum_{i=1}^{n}\sum_{j=1}^{d}\int_0^T |Z_s^{ij}|^2 d[X^j,X^j]_s\} < \infty$$

such that

$$M_t = \int_0^t Z_s dX_s, \quad \text{for } t \in [0,T].$$

We further assume the measure induced by the quadratic variation process $[X^j, X^j]_t$ is equivalent to the Lebesgue measure dt. This, in turn, implies

$$E\{\sum_{i=1}^{n}\sum_{j=1}^{d}\int_0^T |Z_s^{ij}|^2 ds\} \sim E\{\sum_{i=1}^{n}\sum_{j=1}^{d}\int_0^T |Z_s^{ij}|^2 d[X^j,X^j]_s\} < \infty.$$

Define $\mathbf{L}^4_{[t_{i-1},t_i]} : \mathcal{M}^2([t_{i-1},t_i];R^n) \to H^2([t_{i-1},t_i];R^{n\times d})$ by the martingale representation:

$$\mathbf{L}^4_{[t_{i-1},t_i]}(M)_t = Z_t$$

with Z being the martingale representation of M given by

$$\int_{t_{i-1}}^{t_i} Z_t dX_t = M_{t_i} - M_{t_{i-1}}.$$

Note that $\mathbf{L}^4_{[t_{i-1},t_i]}$ itself is a local operator, i.e. $\mathbf{L}^4_{[t_{i-1},t_i]}(M)_t$ depends only on $(M_s)_{s\in[t,t+\epsilon)}$ for any small $\epsilon > 0$, and in particular is independent of the interval $[t_{i-1}, t_i]$. Therefore the operator $\mathbf{L}$ does not depend on any particular partition π, and moreover, the functional differential equation (2.4) (or (2.5)) turns out to be

$$\begin{aligned} V_t = & \sum_{j=1}^{i-1}\int_{t_{j-1}}^{t_j} f\left(s, Y(V)_s, Z(V)_s\right) ds \\ & + \int_{t_{i-1}}^{t} f\left(s, Y(V)_s, Z(V)_s\right) ds \\ = & \int_0^t f\left(s, Y(V)_s, Z(V)_s\right) ds, \quad \text{for } t \in [0,T], \ a.s., \end{aligned}$$

where $Z(V)_s$ is defined obviously as

$$\int_t^T Z(V)_s dX_s = \xi + V_T - E[\xi + V_T|\mathcal{F}_t].$$

The associated BSDE is:

$$\begin{aligned} Y_t = & \ \xi + \sum_{j=i+1}^{N}\int_{t_{j-1}}^{t_j} f\left(s, Y_s, Z_s\right) ds \\ & + \int_t^{t_i} f\left(s, Y_s, Z_s\right) ds - \int_t^T dM_s \\ = & \ \xi + \int_t^T f\left(s, Y_s, Z_s\right) ds - \int_t^T dM_s, \quad \text{for } t \in [0,T], \ a.s.. \end{aligned}$$

We finish this subsection by summarizing the above examples of $\mathbf{L}$ in the following Corollary, but for a Brownian filtration setting.

Corollary 2.11 *Let $B = (B^1, \cdots, B^d)^T$ be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbf{P})$ with the natural filtration $\{\mathcal{F}_t^B\}$ after augumentation. If the coefficients satisfy the conditions in Theorem 2.3, then the following four types of BSDEs admits a unique adapted solution $(Y, M) \in \mathcal{C}([0,T]; R^n) \times \mathcal{H}^2([0,T]; R^n)$ with M being represented by $M = \int_0^\cdot Z_s dB_s$,*

Case (1)

$$Y_t = \xi + \sum_{j=i+1}^{N} \int_{t_{j-1}}^{t_j} f\left(s, Y_s, M_s - M_{t_{j-1}}\right) ds + \int_t^{t_i} f\left(s, Y_s, M_s - M_{t_{i-1}}\right) ds - \int_t^T dM_s, \quad \text{for } t \in [t_{i-1}, t_i], \ a.s..$$

Case (2)

$$Y_t = \xi + \sum_{j=i+1}^{N} \int_{t_{j-1}}^{t_j} f\left(s, Y_s, \int_{t_{j-1}}^{s} g(M_u - M_{t_{j-1}}) du\right) ds + \int_t^{t_i} f\left(s, Y_s, \int_{t_{i-1}}^{s} g(M_u - M_{t_{i-1}}) du\right) ds - \int_t^T dM_s, \quad \text{for } t \in [t_{i-1}, t_i], \ a.s..$$

Case (3)

$$Y_t = \xi + \sum_{j=i+1}^{N} \int_{t_{j-1}}^{t_j} f\left(s, Y_s, \sqrt{[\mathbf{M}^j, \mathbf{M}^j]_s}\right) ds + \int_t^{t_i} f\left(s, Y_s, \sqrt{[\mathbf{M}^i, \mathbf{M}^i]_s}\right) ds - \int_t^T dM_s, \quad \text{for } t \in [t_{i-1}, t_i], \ a.s..$$

Case (4)

$$Y_t = \xi + \int_t^T f\left(s, Y_s, Z_s\right) ds - \int_t^T dM_s, \quad \text{for } t \in [0, T], \ a.s..$$

In particular, for Case (4), we will call the martingale representation Z, rather than the martingale M, the solution of BSDEs.

2.3.3 BSDEs coupled with forward SDEs

In this subsection, we apply our functional differential equation approach to the BSDEs coupled with forward SDEs. In particular, we consider the FBSDE:

$$\begin{cases} dY_t = -h(t, Y_t, Z_t)dt + Z_t dB_t, \\ Y_T = \phi(X_T), \\ dX_t = f(t, Y_t, Z_t)dt + dB_t, \\ X_0 = x \end{cases} \tag{2.12}$$

defined on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ satisfying the *usual conditions*, on which $B = (B^1, \cdots, B^d)^T$ is a d-dimensional Brownian motion. We impose the following conditions on the coefficients:

Condition 2.3 *(Lipschitz) The coefficients* $h : [0,T]\times R^n\times R^{n\times d} \to R^n$*,* $f : [0,T]\times R^n\times R^{n\times d} \to R^d$ *and* $\phi : R^d \to R^n$ *are continuous. Moreover* h*,* f *and* ϕ *are Lipschitz continuous, i.e.*

$$\begin{cases} |h(t,y,z) - h(t,\bar{y},\bar{z})| \le C_5(|y-\bar{y}| + |z-\bar{z}|), \\ |f(t,y,z) - f(t,\bar{y},\bar{z})| \le C_5(|y-\bar{y}| + |z-\bar{z}|), \\ |\phi(x) - \phi(\bar{x})| \le C_6|x-\bar{x}|, \end{cases}$$

and ϕ *is uniformly bounded,*

$$\sup_{x\in R^d} |\phi(x)| \le C_6,$$

for $t \in [0,T]$*,* $y,\bar{y} \in R^n$*,* $z,\bar{z} \in R^{n\times d}$ *and* $x,\bar{x} \in R^d$*.*

As we already discussed in Chapter 1, there are usually three ways to solve FBSDEs: *method of contraction mapping*, *four-step scheme*, and *method of continuity.* However most of them either solve the equations locally or assume some regularity on the coefficients (e.g. smoothness and monotonicity). In [22] Delarue combined the first two methods and solved FBSDEs globally with Lipschitz continuous assumptions on the coefficients. The aim of this subsection is to apply our functional differential equation approach to solve FBSDEs. This approach can help us really work out the solution, and we believe it will benefit especially numerical solutions of FBSDEs, because a usual obstacle to numerically solve FBSDEs is one needs to solve (Y,Z) backwards and X forwards at the same time. But by introducing an associated functional differential equation, we can solve all the components (Y,Z,X) in one direction altogether.

The main result of this subsection is the following theorem:

Theorem 2.12 *If the coefficients satisfy Lipschitz condition 2.3, then there exists at least one adapted solution* $(Y,Z) \in \mathcal{C}([0,T];R^n) \times H^2([0,T];R^{n\times d})$*, where* Y *is uniformly bounded, together with the forward process* $X \in \mathcal{C}([0,T];R^n)$ *of FBSDE (2.12).*

Note that the Lipschitz condition 2.3 is only sufficient but not necessary. For example, h and f can also depend on the forward process X, but needing some additional boundedness conditions on X. On the other hand, the solution to (2.12) is in fact unique. However, we do not attempt to such generalization, because the conclusion in Theorem 2.12, which be will be employed in Section 4.3 of Chapter 4, is enough for our purpose. Before proving Theorem 2.12, we recall the following counterexample from Antonelli [1], which is in the spirit of Sturm-Liouville problems:

Proposition 2.13 *If* $T = \frac{3}{4}\pi$ *and* $x \ne 0$*, then the following FBSDE does not admit any adapted solution:*

$$\begin{cases} dY_t = -X_t dt + Z_t dB_t, \\ Y_T = -X_T, \\ dX_t = Y_t dt + dB_t, \\ X_0 = x. \end{cases}$$

Proof. Suppose that the above FBSDE admits an adapted solution (Y, Z, X). By taking the expectation, $y(t) := E(Y_t)$ and $x(t) := E(X_t)$ for $t \in [0, T]$ satisfy the equations:

$$\begin{cases} dy(t) = -x(t)dt, \\ y(T) = -x(T), \\ dx(t) = y(t)dt, \\ x(0) = x. \end{cases}$$

Therefore, $x(t)$ satisfies the Sturm-Liouville equation:

$$\begin{cases} \ddot{x}(t) + x(t) = 0, \\ x(0) - x = 0, \quad \dot{x}(T) + x(T) = 0. \end{cases}$$

The general solution must be of the form $x(t) = C_7 \cos t + C_8 \sin t$, so $x(0) = C_7$ and $\dot{x}(T) = -C_7 \sin T + C_8 \cos T$. If $T = \frac{3}{4}\pi$, by the boundary conditions, we obtain that $x = 0$, which is a contradiction. ■

To solve FBSDE (2.12), we consider the functional differential equation on $[\tau, T]$:

$$V_t = \int_\tau^t h(s, Y(V, X)_s, Z(V, X)_s)ds, \tag{2.13}$$

together with the forward process X:

$$X_t = x + \int_\tau^t f(s, Y(V, X)_s, Z(V, X)_s)ds + \int_\tau^t dB_s, \tag{2.14}$$

where

$$Y(V, X)_t = E[\phi(X_T) + V_T|\mathcal{F}_t] - V_t,$$
$$\int_\tau^T Z(V, X)_s dB_s = \phi(X_T) + V_T - E[\phi(X_T) + V_T|\mathcal{F}_\tau].$$

If the functional differential equation (2.13) (2.14) can be solved with $\tau = 0$, then analogous to the proof of Lemma 2.2, it is easy to verify (Y, Z, X) is a solution to (2.12). For the notation's simplicity, let's denote $\Psi = (V, X)^T$ and $F = (h, f)^T$. Then (2.13) (2.14) are simplified to:

$$\Psi_t = \chi_x + \int_\tau^t F(s, Y(\Psi)_s, Z(\Psi)_s)ds + \int_\tau^t \chi_1 dB_s, \tag{2.15}$$

where $\chi_x \in R^{n+d}$ with the first n components being 0 and next d components being $x \in R^d$. If the solution to (2.15) exists for $\tau = 0$, then the solution exists globally.

Lemma 2.14 *If the coefficients satisfy Lipschitz condition 2.3, and τ satisfies*

$$\sqrt{T - \tau} \leq \frac{1}{8C_5(1 + C_6)} \wedge 1,$$

then functional differential equation (2.15) admits a unique solution $\Psi \in \mathcal{C}([\tau, T]; R^{n+d})$.

Proof. We denote the mapping defined by (2.15) by $\mathbb{L}$. We first show that $\mathbb{L} : \mathcal{C}([\tau, T]; R^{n+d}) \to \mathcal{C}([\tau, T]; R^{n+d})$. Indeed for $\Psi \in \mathcal{C}([\tau, T]; R^{n+d})$,

$$\begin{aligned}
&||\mathbb{L}(\Psi)||_{\mathcal{C}[\tau,T]} \\
&\leq |x| + \sqrt{E\left(\int_\tau^T |F(s, Y(\Psi)_s, Z(\Psi)_s)|ds\right)^2} + \sqrt{E \sup_{\tau\leq t\leq T} |\int_\tau^t \chi_1 dB_s|^2} \\
&\leq |x| + \sqrt{T-\tau}\sqrt{E\left(\int_\tau^T |F(s, Y(\Psi)_s, Z(\Psi)_s)|^2 ds\right)} + 2\sqrt{E|\int_\tau^T \chi_1 dB_s|^2} \\
&\leq |x| + (C_5\sqrt{T-\tau} + 2d)\sqrt{\int_\tau^T (s^2 \vee 1)ds} \\
&\quad + C_5\sqrt{T-\tau}\sqrt{E\int_\tau^T |Y(\Psi)_s|^2 ds} + C_5\sqrt{T-\tau}\sqrt{E\int_\tau^T |Z(\Psi)_s|^2 ds}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\sqrt{E\int_\tau^T |Y(\Psi)_s|^2 ds} \\
&\leq \sqrt{E\int_\tau^T \{E[\phi(X_T)|\mathcal{F}_s]\}^2 ds} + \sqrt{E\int_\tau^T \{E[V_T|\mathcal{F}_s]\}^2 ds} + \sqrt{E\int_\tau^T |V_s|^2 ds} \\
&\leq \sqrt{\int_\tau^T E|\phi(X_T)|^2 ds} + \sqrt{\int_\tau^T E|V_T|^2 ds} + \sqrt{\int_\tau^T E|V_s|^2 ds} \\
&\leq \sqrt{T-\tau}(2 + C_6)||\Psi||_{C[\tau,T]},
\end{aligned}$$

and by Itô's isometry,

$$\begin{aligned}
\sqrt{E\int_\tau^T |Z(\Psi)_s|^2 ds} &= \sqrt{E\left(\int_\tau^T Z(\Psi)_s dB_s\right)^2} \\
&\leq \sqrt{E|\phi(X_T)|^2} + \sqrt{E|V_T|^2} + \sqrt{E\left\{E[\phi(X_T)|\mathcal{F}_\tau]\right\}^2} \\
&\quad + \sqrt{E\left\{E[V_T|\mathcal{F}_\tau]\right\}^2} \\
&\leq (2 + 2C_6)||\Psi||_{C[\tau,T]}.
\end{aligned}$$

Therefore, $||\mathbb{L}(\Psi)||_{\mathcal{C}[\tau,T]} < \infty$. Likewise, for $\Psi, \Psi' \in \mathcal{C}([\tau, T]; R^{n+d})$, we have

$$\begin{aligned}
&||\mathbb{L}(\Psi) - \mathbb{L}(\Psi')||_{\mathcal{C}[\tau,T]} \\
&\leq C_5\sqrt{T-\tau}\left(\sqrt{T-\tau}(2 + C_6) + 2 + 2C_6\right)||\Psi - \Psi'||_{\mathcal{C}[\tau,T]} \\
&\leq \frac{1}{2}||\Psi - \Psi'||_{\mathcal{C}[\tau,T]}
\end{aligned}$$

by the condition on τ. Hence the operator $\mathbb{L}$ defined by (2.15) is a contraction mapping on $\mathcal{C}([\tau, T]; R^{n+d})$. ■

Next, we extend to the global solution on $[0, T]$ based on Lemma 2.14. To do this, we pursue the bounded solutions of FBSDE (2.12). First, by the Markov property of equations, there exists a measurable function Φ such that $Y_t = \Phi(t, X_t)$. By inspecting the proof for Lemma 2.14, we see that the crucial step to extend to the global solution of (2.15) is that one needs a uniform estimate for the gradient of Φ. We recall a regularity result from Delarue [23]:

Lemma 2.15 *(Delarue [23]) Under Lipschitz condition 2.3 on the coefficients, there exists a measurable Φ such that $Y_t = \Phi(t, X_t)$. Moreover, there exists a constant C_9 depending on the Lipschitz constants C_5 and C_6, the bound of the terminal data, the dimension n and d, and the terminal time T such that*

$$|\Phi(t,x)|,\ |\nabla_x \Phi(t,x)| \le C_9, \quad \text{for } (t,x) \in [0,T] \times R^d.$$

Using such constant C_9, we consider a partition of $[0,T]$ by $\pi : 0 = t_0 < t_1 < \cdots < t_N = T$, with the mesh $|\pi| = \max_{1 \le i \le N} |t_i - t_{i-1}|$ such that

$$\sqrt{|\pi|} = \frac{1}{8C_5(1 + C_9)} \wedge 1.$$

We start with $[t_{N-1}, t_N]$ and consider $\Psi(N) = (V(N), X(N)^{t_{N-1},x})^T$ such that

$$\Psi(N)_t = \chi_x + \int_{t_{N-1}}^{t} F(s, Y(N)_s, Z(N)_s) ds + \int_{t_{N-1}}^{t} \chi_1 dB_s$$

with

$$\begin{aligned} Y(N)_t &= E[\phi(X(N)_T^{t_{N-1},x}) + V(N)_T | \mathcal{F}_t] - V(N)_t, \\ \int_{t_{N-1}}^{T} Z(N)_s dB_s &= \phi(X(N)_T^{t_{N-1},x}) + V(N)_T \\ &\quad - E[\phi(X(N)_T^{t_{N-1},x}) + V(N)_T | \mathcal{F}_{t_{N-1}}], \end{aligned}$$

where we used the superscripts (t_{N-1}, x) to emphasize $X(N)^{t_{N-1},x}$ starting from $X(N)_{t_{N-1}}^{t_{N-1},x} = x$. By Lemma 2.14, there exists a unique solution:

$$\Psi(N) = (V(N), X(N)^{t_{N-1},x})^T \in \mathcal{C}([t_{N-1}, T]; R^{n+d}),$$

and we also get $(Y(N), Z(N))$. Moreover, there exists a measurable function Φ_{N-1} such that $Y(N)_{t_{N-1}} = \Phi_{N-1}(t_{N-1}, x)$ and by Lemma 2.15,

$$|\nabla_x \Phi_{N-1}(t_{N-1}, x)| \le C_9, \quad \text{for } x \in R^d.$$

In general on $[t_{i-1}, t_i]$ for $1 \le i \le N-1$, we consider $\Psi(i) = (V(i), X(i)^{t_{i-1},x})^T$ such that

$$\Psi(i)_t = \chi_x + \int_{t_{i-1}}^{t} F(s, Y(i)_s, Z(i)_s) ds + \int_{t_{i-1}}^{t} \chi_1 dB_s$$

with

$$\begin{aligned} Y(i)_t &= E[\Phi_i(t_i, X(i)_{t_i}^{t_{i-1},x}) + V(i)_{t_i} | \mathcal{F}_t] - V(i)_t, \\ \int_{t_{i-1}}^{t_i} Z(i)_s dB_s &= \Phi_i(t_i, X(i)_{t_i}^{t_{i-1},x}) + V(i)_{t_i} \\ &\quad - E[\Phi_i(t_i, X(i)_{t_i}^{t_{i-1},x}) + V(i)_{t_i} | \mathcal{F}_{t_{i-1}}]. \end{aligned}$$

Again by Lemma 2.14, we obtain that there exists a unique solution:

$$\Psi(i) = (V(i), X(i)^{t_{i-1},x})^T \in \mathcal{C}([t_{i-1}, t_i]; R^{n+d})$$

and we get $(Y(i), Z(i))$ as well. Moreover there exists a measurable function Φ_{i-1} such that $Y(i)_{t_{i-1}} = \Phi_{i-1}(t_{i-1}, x)$, and by Lemma 2.15,

$$|\nabla_x \Phi_{i-1}(t_{i-1}, x)| \le C_9, \quad \text{for } x \in R^d.$$

We stress that $(V(i), X(i)^{t_{i-1},x})$ for $1 \le i \le N$ are not the real solutions to (2.15) on the corresponding time interval $[t_{i-1}, t_i]$, because they start from

$$(V(i)_{t_{i-1}}, X(i)_{t_{i-1}}^{t_{i-1},x}) = (0, x).$$

Therefore, in order to match the starting points for the solution to (2.15) on each time interval $[t_{i-1}, t_i]$, we need to shift the paths of $(V(i), X(i)^{t_{i-1},x})$ accordingly, which is in the same spirit of Lemma 2.7.

Lemma 2.16 *If the coefficients satisfy Lipschitz condition 2.3, then the global solution $\Psi = (V, X)^T$ to (2.15) exists and is constructed as follows: for $1 \le i \le N$,*

$$V_t = V(i)_t + \sum_{j=1}^{i-1} V(j)_{t_j} \quad \textit{for } t_{i-1} \le t \le t_i,$$

where we follow the convention $\sum_{j=1}^{0} = 0$, and

$$X_t = \begin{cases} X(1)_t^{t_0,x} & \textit{for } t_0 \le t \le t_1; \\ X(2)_t^{t_1,X_{t_1}} & \textit{for } t_1 \le t \le t_2; \\ \cdots & \\ X(N)_t^{t_{N-1},X_{t_{N-1}}} & \textit{for } t_{N-1} \le t \le t_N \end{cases}$$

with (Y, Z) being constructed as $(Y_t, Z_t) = (Y(i)_t, Z(i)_t)$ for $t_{i-1} \le t \le t_i$.

Proof. Since the proof is analogous to that of Lemma 2.7, we only present the main steps. To this end, we need to show $\Psi = (V, X)^T$ with (Y, Z) satisfying (2.15) for $\tau = 0$. Indeed, for $t \in [t_{N-1}, t_N]$, by the definition of (V, X), we have

$$\begin{aligned} V_t - V_{t_{N-1}} &= V(N)_t + \sum_{j=1}^{N-1} V(j)_{t_j} - V(N)_{t_{N-1}} - \sum_{j=1}^{N-1} V(j)_{t_j} \\ &= \int_{t_{N-1}}^{t} h(s, Y(N)_s, Z(N)_s) ds, \end{aligned}$$

and

$$\begin{aligned} X_t - X_{t_{N-1}} &= X(N)_t^{t_{N-1},X_{t_{N-1}}} - X_{t_{N-1}} \\ &= \int_{t_{N-1}}^{t} f(s, Y(N)_s, Z(N)_s) ds + \int_{t_{N-1}}^{t} dB_s, \end{aligned}$$

where

$$
\begin{aligned}
Y(N)_t &= E\left[\phi\left(X(N)_T^{t_{N-1},X_{t_{N-1}}}\right) + V(N)_T|\mathcal{F}_t\right] - V(N)_t \\
&= E\left[\phi(X_T) + V(N)_T + \sum_{j=1}^{N-1} V(j)_{t_j}|\mathcal{F}_t\right] \\
&\quad - V(N)_t - \sum_{j=1}^{N-1} V(j)_{t_j} \\
&= E[\phi(X_T) + V_T|\mathcal{F}_t] - V_t,
\end{aligned}
$$

and

$$
\begin{aligned}
\int_{t_{N-1}}^{T} Z(N)_s dB_s &= \phi(X(N)_T^{t_{N-1},X_{t_{N-1}}}) + V(N)_T \\
&\quad - E\left[\phi\left(X(N)_T^{t_{N-1},X_{t_{N-1}}}\right) + V(N)_T|\mathcal{F}_{t_{N-1}}\right] \\
&= \phi(X_T) + V(N)_T + \sum_{j=1}^{N-1} V(j)_{t_j} \\
&\quad - E\left[\phi(X_T) + V(N)_T + \sum_{j=1}^{N-1} V(j)_{t_j}|\mathcal{F}_{t_{N-1}}\right] \\
&= \phi(X_T) + V_T - E[\phi(X_T) + V_T|\mathcal{F}_{t_{N-1}}].
\end{aligned}
$$

Hence (V, X) with (Y, Z) defined in the lemma satisfy (2.15) on $[t_{N-1}, t_N]$.

In general for $1 \leq i \leq N-1$, by the backward induction, it is easy to verify that $\Psi_t = (V_t, X_t)$ with (Y_t, Z_t) also satisfying (2.15) for $t \in [t_{i-1}, t_i]$. ■

Chapter 3

Time Discretization of Functional Differential Equations

3.1 Introduction

In this chapter, we continue our study of functional differential equations introduced in the last chapter and consider their numerical solutions. Since we already proved in Lemma 2.2 of the last chapter that solving BSDEs is equivalent to solving the associated functional differential equations, numerically solving functional differential equations also provides us with a numerical scheme for the solutions of the associated BSDEs. Specifically, we consider the following functional differential equation on $(\Omega, \mathcal{F}, \mathbf{P})$ with $B = (B^1, \cdots, B^d)^T$ being a Brownian motion with its natural filtration $\{\mathcal{F}_t\}$ after augmentation: (The superscript T denotes matrix transposition.)

$$V_t = \int_0^t f(s, Y(V)_s, Z(V)_s)ds, \quad \text{for } t \in [0, T] \tag{3.1}$$

with

$$\begin{aligned} Y(V)_t &= E[\phi(X_T) + V_T|\mathcal{F}_t] - V_t, \\ M(V)_t &= E[\phi(X_T) + V_T|\mathcal{F}_t], \\ \int_t^T Z(V)_s dB_s &= M(V)_T - M(V)_t, \end{aligned}$$

where X is driven by SDE:

$$\begin{cases} dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t, & \text{for } t \in [0, T], \\ X_0 = x. \end{cases}$$

As Lemma 2.2 in Chapter 2 shows, the above functional differential equation (3.1) is equivalent to the classical BSDE considered in the literature:

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt + Z_t dB_t, \\ Y_T = \phi(X_T). \end{cases} \tag{3.2}$$

The conditions imposed on the coefficients are given as following:

Condition 3.1 *(Lipschitz) The coefficients* $f : [0,T] \times R \times R^d \to R$, $\phi : R^m \to R$, $a : [0,T] \times R^m \to R^m$ *and* $\sigma : [0,T] \times R^m \to R^{m\times d}$ *are continuous. Moreover there exists a constant* C_1 *such that*

$$|f(t,y,z) - f(\bar{t},\bar{y},\bar{z})| \le C_1(\sqrt{t-\bar{t}} + |y-\bar{y}| + |z-\bar{z}|)$$

for $t, \bar{t} \in [0,T]$, $y, \bar{y} \in R$ *and* $z, \bar{z} \in R^d$. ϕ *is Lipschitz continuous in* $x \in R^m$ *with Lipschitz constant* C_2, *and* a, σ *are Lipschitz continuous in* $x \in R^m$ *with Lipschitz constant* C_3.

One may wonder why we do not consider the general form of functional differential equation (2.4) in Chapter 2. Actually at the moment we are unable to provide the numerical solutions of the general form of functional differential equation (2.4) yet. Due to the operator **L**, (2.4) is not defined locally, i.e. we can not solve it only based on the information within an infinitesimal time interval. On the other hand, existing numerical methods are mainly based on time discretization, which implicitly requires the equations must be defined locally. This is the main reason we can only consider (3.1) at the moment.

We start by defining the space that we will work on. Let $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ be a partition of $[0,T]$, and let k be an integer between 0 and N.

(1) $\mathcal{C}^\pi([t_k,t_N];R)$: the space of predictably measurable random sequence $(V^\pi_{t_i})_{k\le i\le N}$, i.e. $V^\pi_{t_i}$ is $\mathcal{F}_{t_{i-1}}$-measurable. $(V^\pi_{t_i})_{k\le i\le N}$ is valued in R and equipped with the following norm:

$$||V^\pi||_{\mathcal{C}^\pi([t_k,t_N];R)} = \max_{k\le i\le N} \sqrt{E|V^\pi_{t_i}|^2}.$$

3.1.1 Main results of this chapter

We first propose a discrete-time version of (3.1). To this end, let $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ be a partition of $[0,T]$ with the mesh:

$$|\pi| := \max_{1\le i\le N} |t_i - t_{i-1}|,$$

and denote $\Delta t_i = t_i - t_{i-1}$ and $\Delta B_{t_i} = B_{t_i} - B_{t_{i-1}}$ for $1 \le i \le N$.

Define the difference equation for the differential equation (3.1) as follows:

$$V^\pi_{t_i} = V^\pi_{t_{i-1}} + f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})\Delta t_i \tag{3.3}$$

for $1 \le i \le N$ and $V^\pi_0 = 0$, where

$$\begin{cases} Y(V^\pi)_{t_{i-1}} = & E[\phi(X_T) + V^\pi_T|\mathcal{F}_{t_{i-1}}] - V^\pi_{t_{i-1}}, \\ Z(V^\pi)_{t_{i-1}} = & \dfrac{1}{\Delta t_i} E\left[(\phi(X_T) + V^\pi_T)(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}\right]. \end{cases}$$

If the difference equation (3.3) admits a unique solution, then it defines a sequence of predictably measurable random variables $(V^\pi_{t_i})_{0\le i\le N}$. The later arguments will rely on the predictability of the sequence $(V^\pi_{t_i})_{0\le i\le N}$. The Functional difference equation (3.3) is the main object that we will investigate in this chapter.

For our analysis, we also need a continuous-time approximation of V. For $t \in [t_{i-1}, t_i]$, we use the linear interpolation to define V_t^π, i.e.

$$V_t^\pi = V_{t_{i-1}}^\pi + f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})(t - t_{i-1}).$$

Based on such V_t^π, we further define the corresponding continuous-time approximation $(Y(V^\pi), Z(V^\pi))$ of the solution to BSDE (3.2) as

$$\begin{aligned} Y(V^\pi)_t &= E[\phi(X_T) + V_T^\pi | \mathcal{F}_t] - V_t^\pi, \\ M(V^\pi)_t &= E[\phi(X_T) + V_T^\pi | \mathcal{F}_t], \\ \int_{t_{i-1}}^{t_i} Z(V^\pi)_s dB_s &= M(V^\pi)_{t_i} - M(V^\pi)_{t_{i-1}}, \qquad \text{for } t \in [t_{i-1}, t_i]. \end{aligned} \tag{3.4}$$

Our main result of this chapter is the following theorem about the convergence rate of V^π to V.

Theorem 3.1 *If the coefficients satisfy Lipschitz condition 3.1, then there exits a constant C_4, only depending on the Lipschitz constants C_1, C_2 and C_3, the time T, the initial value x and the dimensions m and d, such that*

$$\sup_{t \in [0,T]} E\{|V_t^\pi - V_t|^2\} \leq C_4 |\pi|$$

We call such kind of constant a generic constant, i.e. $C = C(C_1, C_2, C_3, T, x, m, d)$. Given the above convergence rate result, we only need to solve the associated functional difference equation (3.3), the solution of which will be used as the discrete-time approximation of the solution to functional differential equation (3.1). When the mesh of the partition goes to zero, the discrete-time approximation converges to the solution of functional difference equation (3.1). The way to solve (3.3) is essentially a discrete-time version of the approach we presented in the last chapter. Since such approach is constructive, it also provides us with a numerical scheme as a by-product. The second main result of this chapter is the solvability of functional difference equation (3.3), which is stated as the following theorem:

Theorem 3.2 *If the coefficients satisfy Lipschitz condition 3.1, by choosing $|\pi|$ small enough, functional difference equation (3.3) admits a unique solution $(V_{t_i}^\pi)_{0 \leq i \leq N} \in \mathcal{C}^\pi([t_0, t_N], R)$.*

The chapter is organized as follows. Section 3.2 is devoted to the proof of Theorem 3.1. In Section 3.2.1, we recall some regularity results from the existing literature of numerical solutions to BSDEs. Section 3.2.2 gives the proof of Theorem 3.1.

We prove the existence and uniqueness Theorem 3.2 for functional difference equation (3.3) in Section 3.3. The strategy of the proof is analogous to the proof of Theorem 2.5 in Chapter 2, i.e. we first solve the equation locally by fixed point arguments, and then we obtain the global solution by shifting the paths of V^π. We conclude this chapter by proposing a numerical scheme for (3.3) in Section 3.4 based on Picard iteration.

This chapter is mainly adapted from the author's paper with Casserini [16].

3.2 Discretizing functional differential equations

3.2.1 Some preliminary results

In [10] Bouchard and Touzi proposed the following numerical scheme to solve BSDE (3.2). For a given partition $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ of $[0, T]$, they defined the following discrete-time approximation of BSDE (3.2):

$$\begin{aligned} \mathbf{Y}^\pi_T &= \phi(X_T), \\ \mathbf{Y}^\pi_{t_{i-1}} &= E[\mathbf{Y}^\pi_{t_i}|\mathcal{F}_{t_{i-1}}] + f(t_{i-1}, \mathbf{Y}^\pi_{t_{i-1}}, \mathbf{Z}^\pi_{t_{i-1}})\Delta t_i, \\ \mathbf{Z}^\pi_{t_{i-1}} &= \frac{1}{\Delta t_i} E[\mathbf{Y}^\pi_{t_i}(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}], \qquad \text{for } 1 \le i \le N. \end{aligned}$$

For $t \in [t_{i-1}, t_i]$, they further defined the continuous-time approximation $(\mathbf{Y}^\pi_t, \mathbf{Z}^\pi_t)$ of the solution to BSDE (3.2) by the linear interpolation:

$$\begin{aligned} \mathbf{Y}^\pi_{t_i} &= E[\mathbf{Y}^\pi_{t_i}|\mathcal{F}_{t_{i-1}}] + \int_{t_{i-1}}^{t_i} \mathbf{Z}^\pi_s dB_s, \\ \mathbf{Y}^\pi_t &= \mathbf{Y}^\pi_{t_{i-1}} - f(t_{i-1}, \mathbf{Y}^\pi_{t_{i-1}}, \mathbf{Z}^\pi_{t_{i-1}})(t - t_{i-1}) + \int_{t_{i-1}}^{t} \mathbf{Z}^\pi_s dB_s. \end{aligned}$$

In [10] the following convergence rate of $(\mathbf{Y}^\pi, \mathbf{Z}^\pi)$ to the solution (Y, Z) of BSDE (3.2) was proved:

Lemma 3.3 *(Bouchard and Touzi [10]) If the coefficients satisfy Lipschitz condition 3.1, then there exists a generic constant C_5 such that*

$$\sup_{t\in[0,T]} E|\mathbf{Y}^\pi_t - Y_t|^2 + E\int_0^T |\mathbf{Z}^\pi_s - Z_s|^s ds \le C_5|\pi|.$$

Next we recall a result which is in fact motivated from proving the path regularity of the martingale representation Z. In [80], Zhang introduced an auxiliary process $\bar{Z}^\pi$ in order to give the best H^2-approximation of the martingale representation Z at the partition points: for $i = 1, \cdots, N$,

$$\bar{Z}^\pi_{t_{i-1}} = \frac{1}{\Delta t_i} E\left[\int_{t_{i-1}}^{t_i} Z_s ds|\mathcal{F}_{t_{i-1}}\right]. \tag{3.5}$$

Therefore we have

$$E\int_{t_{i-1}}^{t_i} |Z_s - \bar{Z}^\pi_{t_{i-1}}|^2 ds = \inf_{\eta\in L^2(\Omega,\mathcal{F}_{t_{i-1}},\mathbf{P})} E\int_{t_{i-1}}^{t_i} |Z_s - \eta|^2 ds$$

Lemma 3.4 *(Zhang [80]) If the coefficients satisfy Lipschitz condition 3.1, then there exists a generic constant C_5 such that*

$$\sup_{t\in[t_{i-1},t_i]} E|Y_{t_{i-1}} - Y_t|^2 + \sum_{k=1}^N E\int_{t_{k-1}}^{t_k} |\bar{Z}^\pi_{t_{k-1}} - Z_s|^2 ds \le C_5|\pi|.$$

3.2.2 Proof of the convergence rate

In this subsection we prove Theorem 3.1. The proof depends on the following lemmas:

Lemma 3.5 *The pair $(Y(V^\pi), Z(V^\pi))$ defined by (3.4) as the continuous-time approximation of the solution to BSDE (3.2) coincides with the solution pair of the discretization algorithm for BSDE (3.2) proposed in [10], i.e.*

$$\begin{cases} Y(V^\pi)_t = \mathbf{Y}_t^\pi, & \text{for } t \in [0,T] \text{ a.s.,} \\ Z(V^\pi)_t = \mathbf{Z}_t^\pi, & \text{for a.e. } t \in [0,T] \text{ a.s..} \end{cases}$$

Proof. What we need to show is that $(Y(V^\pi), Z(V^\pi))$ also satisfies the discretization algorithm defined by Bouchard and Touzi [10]. We first show that $Y(V^\pi)_{t_i}$ and $\mathbf{Y}_{t_i}^\pi$ coincide at all the partition points for $i = 0, 1, \cdots, N$ and $Z(V^\pi)_t = \mathbf{Z}_t^\pi$ for *a.e.* $t \in [0,T]$, *a.s.*. Clearly,

$$Y(V^\pi)_T = E[\phi(X_T) + V_T|\mathcal{F}_T] - V_T = \phi(X_T),$$

and by the predictability of the sequence $(V_{t_i}^\pi)$, for $i = 1, \cdots, N$,

$$\begin{aligned} Y(V^\pi)_{t_{i-1}} &= E[\phi(X_T) + V_T^\pi|\mathcal{F}_{t_{i-1}}] - V_{t_{i-1}}^\pi \\ &= E[Y(V^\pi)_{t_i} + V_{t_i}^\pi|\mathcal{F}_{t_{i-1}}] - V_{t_{i-1}}^\pi \\ &= E[Y(V^\pi)_{t_i}|\mathcal{F}_{t_{i-1}}] + f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})\Delta t_i. \end{aligned}$$

Using again the $\mathcal{F}_{t_{i-1}}$-measurability of $V_{t_i}^\pi$, we have

$$\begin{aligned} \int_{t_{i-1}}^{t_i} Z(V^\pi)_s dB_s &= E[\phi(X_T) + V_T^\pi|\mathcal{F}_{t_i}] - E[\phi(X_T) + V_T^\pi|\mathcal{F}_{t_{i-1}}] \\ &= E[Y(V^\pi)_{t_i} + V_{t_i}^\pi|\mathcal{F}_{t_i}] - E[Y(V^\pi)_{t_i} + V_{t_i}^\pi|\mathcal{F}_{t_{i-1}}] \\ &= Y(V^\pi)_{t_i} - E[Y(V^\pi)_{t_i}|\mathcal{F}_{t_{i-1}}]. \end{aligned}$$

Next we show that in fact $Y(V^\pi)_t = \mathbf{Y}_t^\pi$ for $t \in [0,T]$, *a.s.*. Indeed,

$$\begin{aligned} Y(V^\pi)_t &= E[\phi(X_T) + V_T^\pi|\mathcal{F}_t] - V_t^\pi \\ &= E[\phi(X_T) + V_T^\pi|\mathcal{F}_t] - E[\phi(X_T) + V_T^\pi|\mathcal{F}_{t_{i-1}}] \\ &\quad + E[\phi(X_T) + V_T^\pi|\mathcal{F}_{t_{i-1}}] - V_{t_{i-1}}^\pi - f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})(t - t_{i-1}) \\ &= \int_{t_{i-1}}^t Z(V^\pi)_s dB_s + Y(V^\pi)_{t_{i-1}} - f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})(t - t_{i-1}), \end{aligned}$$

which complete the proof. ■

Therefore, by the above lemma and Lemma 3.3, we have

$$\sup_{t\in[0,T]} E|Y(V^\pi)_t - Y_t|^2 + E\int_0^T |Z(V^\pi)_s - Z_s|^2 ds \le C_5|\pi|. \tag{3.6}$$

Next, we also introduce an auxiliary process in order to obtain the best H^2-approximation of the approximation process $Z(V^\pi)$ at the partition points: for $1 \le i \le N$,

$$\bar{Z}(V^\pi)_{t_{i-1}} = \frac{1}{\Delta t_i} E\left[\int_{t_{i-1}}^{t_i} Z(V^\pi)_s ds | \mathcal{F}_{t_{i-1}}\right]. \tag{3.7}$$

Lemma 3.6 *For $1 \le i \le N$, $Z(V^\pi)$ and $\bar{Z}(V^\pi)$ coincide at the partition points:*

$$Z(V^\pi)_{t_{i-1}} = \bar{Z}(V^\pi)_{t_{i-1}}.$$

Proof. By Itô's isometry, for $1 \le i \le N$,

$$E\left[\int_{t_{i-1}}^{t_i} Z(V^\pi)_s ds | \mathcal{F}_{t_{i-1}}\right] = E\left[\int_{t_{i-1}}^{t_i} Z(V^\pi)_s dB_s (\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right].$$

Recalling the definition of $Z(V^\pi)_s$, we have $\int_{t_{i-1}}^{t_i} Z(V^\pi)_s dB_s = M(V^\pi)_{t_i} - M(V^\pi)_{t_{i-1}}$, and by the tower property of conditional expectation,

$$\begin{aligned}
\bar{Z}(V^\pi)_{t_{i-1}} &= \frac{1}{\Delta t_i} E\left[\left(M(V^\pi)_{t_i} - M(V^\pi)_{t_{i-1}}\right)(\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right] \\
&= \frac{1}{\Delta t_i} E\left[M(V^\pi)_{t_i} (\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right] \\
&= \frac{1}{\Delta t_i} E\left[E[\phi(X_T) + V_T^\pi | \mathcal{F}_{t_i}](\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right] \\
&= \frac{1}{\Delta t_i} E\left[(\phi(X_T) + V_T^\pi)(\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right] \\
&= Z(V^\pi)_{t_{i-1}}.
\end{aligned}$$

■

Before continuing the proof, we clarify several notations which may cause some confusion. In this chapter, we use $Z(V)$ to represent the real solution of BSDE (3.2); $\bar{Z}(V)^\pi$ to represent the best H^2-approximation of $Z(V)$; $Z(V^\pi)$ to represent the discrete-time approximation of $Z(V)$, and $\bar{Z}(V^\pi)$ to represent the best H^2-approximation of $Z(V^\pi)$.

Next for $s \in [t_{i-1}, t_i]$, we compare the real solution pair $(Y(V)_s, Z(V)_s)$ defined by (3.1) and its approximation $(Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})$ defined by (3.3).

Lemma 3.7 *For $1 \le i \le N$, there exists a generic constant C_5 such that*

$$E\int_{t_{i-1}}^{t_i} |Y(V)_s - Y(V^\pi)_{t_{i-1}}|^2 ds \le C_5 |\pi|^2,$$

and

$$\begin{aligned}
&E\int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_{t_{i-1}}|^2 ds \\
&\le 2\left\{E\int_{t_{i-1}}^{t_i} |Z(V)_s - \bar{Z}(V)^\pi_{t_{i-1}}|^2 ds + E\int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_s|^2 ds\right\}.
\end{aligned}$$

Proof. To prove the first inequality we note that for $s \in [t_{i-1}, t_i]$,

$$\begin{aligned}&E|Y(V)_s - Y(V^\pi)_{t_{i-1}}|^2\\ &\leq 2\left\{E|Y(V)_s - Y(V)_{t_{i-1}}|^2 + E|Y(V)_{t_{i-1}} - Y(V^\pi)_{t_{i-1}}|^2\right\}.\end{aligned}$$

By Lemma 3.4,

$$E|Y(V)_s - Y(V)_{t_{i-1}}|^2 \leq C_5|\pi|,$$

and by (3.6),

$$E|Y(V)_{t_{i-1}} - Y(V^\pi)_{t_{i-1}}|^2 \leq C_5|\pi|,$$

the first conclusion then follows immediately. To prove the second inequality, we note that

$$\begin{aligned}&E\int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_{t_{i-1}}|^2 ds\\ &\leq 2\left\{E\int_{t_{i-1}}^{t_i} |Z(V)_s - \bar{Z}(V)^\pi_{t_{i-1}}|^2 ds + E\int_{t_{i-1}}^{t_i} |\bar{Z}(V)^\pi_{t_{i-1}} - Z(V^\pi)_{t_{i-1}}|^2 ds\right\}.\end{aligned}$$

By Lemma 3.6, $Z(V^\pi)_{t_{i-1}} = \bar{Z}(V^\pi)_{t_{i-1}}$, so the second term of right hand side of the above inequality reduces to

$$\begin{aligned}&E\int_{t_{i-1}}^{t_i} |\bar{Z}(V)^\pi_{t_{i-1}} - \bar{Z}(V^\pi)_{t_{i-1}}|^2 ds\\ &= E\int_{t_{i-1}}^{t_i} \left|\frac{1}{\Delta t_i} E\left[\int_{t_{i-1}}^{t_i} Z(V)_s - Z(V^\pi)_s ds|\mathcal{F}_{t_{i-1}}\right]\right|^2 ds\\ &\leq \frac{1}{\Delta t_i} E\left(\int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_s| ds\right)^2\\ &\leq E\int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_s|^2 ds,\end{aligned}$$

where we used Jensen's inequality and Hölder's inequality for the last two inequalities. ■

We are now ready to provide the proof of Theorem 3.1:

Proof of Theorem 3.1. For $1 \leq i \leq N$ and $t \in [t_{i-1}, t_i]$, we have

$$\begin{aligned}&E\{|V_t - V_t^\pi|^2\} - E\{|V_{t_{i-1}} - V_{t_{i-1}}^\pi|^2\}\\ &= 2E\int_{t_{i-1}}^{t} (V_s - V_s^\pi)\left(f(s, Y(V)_s, Z(V)_s) - f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})\right) ds\\ &\leq E\int_{t_{i-1}}^{t} |V_s - V_s^\pi|^2 ds\\ &\quad + E\int_{t_{i-1}}^{t} \left|f(s, Y(V)_s, Z(V)_s) - f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})\right|^2 ds\\ &\leq E\int_{t_{i-1}}^{t} |V_s - V_s^\pi|^2 ds + 3C_1\int_{t_{i-1}}^{t} (s - t_{i-1}) ds + 3C_1 E\int_{t_{i-1}}^{t} |Y(V)_s - Y(V^\pi)_{t_{i-1}}|^2 ds\\ &\quad + 3C_1 E\int_{t_{i-1}}^{t} |Z(V)_s - Z(V^\pi)_{t_{i-1}}|^2 ds, \qquad (3.8)\end{aligned}$$

where we used the Lipschitz property of the driver f for the last inequality. Note that

$$\int_{t_{i-1}}^{t} (s - t_{i-1}) ds \leq |\pi| \Delta t_i \leq |\pi|^2,$$

and by Lemma 3.7,

$$\begin{aligned} &E \int_{t_{i-1}}^{t} |Y(V)_s - Y(V^\pi)_{t_{i-1}}|^2 ds + E \int_{t_{i-1}}^{t} |Z(V)_s - Z(V^\pi)_{t_{i-1}}|^2 ds \\ &\leq C_5 |\pi|^2 + 2 \left\{ E \int_{t_{i-1}}^{t_i} |Z(V)_s - \bar{Z}(V)^\pi_{t_{i-1}}|^2 ds + E \int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_s|^2 ds \right\}. \end{aligned}$$

Plugging in the above estimates in (3.8), we obtain

$$\begin{aligned} &E\{|V_t - V^\pi_t|^2\} \leq E \int_{t_{i-1}}^{t} |V_s - V^\pi_s|^2 \, ds + E[|V_{t_{i-1}} - V^\pi_{t_{i-1}}|^2] \\ &+ C_6 \left(|\pi|^2 + E \int_{t_{i-1}}^{t_i} |Z(V)_s - \bar{Z}(V)^\pi_{t_{i-1}}|^2 ds + E \int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_s|^2 ds \right). \end{aligned}$$

By Gronwall's inequality,

$$\begin{aligned} &E\{|V_t - V^\pi_t|^2\} \leq e^{t - t_{i-1}} \Big\{ E[|V_{t_{i-1}} - V^\pi_{t_{i-1}}|^2] \\ &+ C_6 \left(|\pi|^2 + E \int_{t_{i-1}}^{t_i} |Z(V)_s - \bar{Z}(V)^\pi_{t_{i-1}}|^2 ds + E \int_{t_{i-1}}^{t_i} |Z(V)_s - Z(V^\pi)_s|^2 ds \right) \Big\} \end{aligned} \tag{3.9}$$

for $t \in [t_{i-1}, t_i]$. In particular, the above is true for $t = t_i$. Iterating the above inequality (3.9) and using the fact that $V_0 = V^\pi_0 = 0$, we obtain

$$\begin{aligned} &E\{|V_{t_i} - V^\pi_{t_i}|^2\} \\ &\leq e^{\sum_{k=1}^{i} \Delta t_k} C_6 \left(i|\pi|^2 + \sum_{k=1}^{i} E \int_{t_{k-1}}^{t_k} |Z(V)_s - \bar{Z}(V)^\pi_{t_{k-1}}|^2 ds \right. \\ &\left. + \sum_{k=1}^{i} E \int_{t_{k-1}}^{t_k} |Z(V)_s - Z(V^\pi)_s|^2 ds \right). \end{aligned}$$

By Lemma 3.4,

$$\sum_{k=1}^{N} E \int_{t_{k-1}}^{t_k} |Z(V)_s - \bar{Z}(V)^\pi_{t_{k-1}}|^2 ds \leq C_5 |\pi|,$$

and by (3.6),

$$\sum_{k=1}^{N} E \int_{t_{k-1}}^{t_k} |Z(V)_s - Z(V^\pi)_s|^2 ds = E \int_0^T |Z(V)_s - Z(V^\pi)_s|^2 ds \leq C_5 |\pi|,$$

so there exists a generic constant C_7 such that

$$E\{|V_{t_i} - V^\pi_{t_i}|^2\} \leq C_7 |\pi|.$$

Finally, plugging in the above estimate into (3.9), we obtain

$$\sup_{0 \leq t \leq T} E\{|V_t - V^\pi_t|^2\} \leq e^{|\pi|} \left\{ C_7 |\pi| + C_6 (|\pi|^2 + C_5 |\pi| + C_5 |\pi|) \right\} \leq C_8 |\pi|,$$

which competes the proof. ■

3.3 Solving functional difference equations

In this section we prove Theorem 3.2. The solution of functional difference equation (3.3) will be used as the discrete-time approximation of the solution to functional differential equation (3.1). The way to solve (3.3) is essentially a discrete-time version of the approach we presented in the last chapter.

We consider the solution of functional difference equation (3.3) in the space $\mathcal{C}^\pi([t_k, t_N], R)$. The following proof of the existence and uniqueness theorem 3.2 is noting but a discrete-time version of the proof for Theorem 2.5 in Chapter 2. The strategy is we first prove the local existence by fixed point arguments, and then extend to the global solution by shifting the path of V^π. We first give the following two auxiliary estimates.

Lemma 3.8 *For any given partition π, and any given number k, $0 \le k \le N-1$, if V^π and $\bar{V}^\pi$ are two solutions of functional difference equation (3.3), then*

$$\begin{aligned}
&\sqrt{E\left(\sum_{j=k+1}^{N} |Y(V^\pi)_{t_{j-1}} - Y(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2} \\
&\le 2|\pi|(N-k)\max_{k\le i\le N}\sqrt{E|V^\pi_{t_i} - \bar{V}^\pi_{t_i}|^2},
\end{aligned}$$

and

$$\begin{aligned}
&\sqrt{E\left(\sum_{j=k+1}^{N} |Z(V^\pi)_{t_{j-1}} - Z(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2} \\
&\le \sqrt{|\pi|}(N-k)\max_{k\le i\le N}\sqrt{E|V^\pi_{t_i} - \bar{V}^\pi_{t_i}|^2}.
\end{aligned}$$

Proof. To prove the first estimate, we use Minkowski's inequality to obtain

$$\begin{aligned}
&\sqrt{E\left(\sum_{j=k+1}^{N} |Y(V^\pi)_{t_{j-1}} - Y(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2} \\
&\le \sum_{j=k+1}^{N}\sqrt{E\left(|Y(V^\pi)_{t_{j-1}} - Y(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2} \\
&\le |\pi|\sum_{j=k+1}^{N}\sqrt{E\left\{E(V^\pi_T - \bar{V}^\pi_T|\mathcal{F}_{t_{j-1}}) + V^\pi_{t_{j-1}} - \bar{V}^\pi_{t_{j-1}}\right\}^2} \\
&\le |\pi|\sum_{j=k+1}^{N}\sqrt{E\left\{E(V^\pi_T - \bar{V}^\pi_T|\mathcal{F}_{t_{j-1}})\right\}^2} + |\pi|\sum_{j=k+1}^{N}\sqrt{E(V^\pi_{t_{j-1}} - \bar{V}^\pi_{t_{j-1}})^2} \\
&\le 2|\pi|(N-k)\max_{k\le i\le N}\sqrt{E|V^\pi_{t_i} - \bar{V}^\pi_{t_i}|^2},
\end{aligned}$$

where we used Jensen's inequality for the last inequality. For the second estimate, we have

$$\sqrt{E\left(\sum_{j=k+1}^{N}|Z(V^\pi)_{t_{j-1}}-Z(\bar V^\pi)_{t_{j-1}}|\Delta t_j\right)^2}$$
$$\le \sum_{j=k+1}^{N}\sqrt{E\left(|Z(V^\pi)_{t_{j-1}}-Z(\bar V^\pi)_{t_{j-1}}|\Delta t_j\right)^2}.$$

Recalling from the definition of $Z(V^\pi)_{t_{j-1}}$ that $Z(V^\pi)_{t_{j-1}}=E\left[(\phi(X_T)+V_T^\pi)(\Delta B_{t_j})^T|\mathcal{F}_{t_{j-1}}\right]/\Delta t_j$, and moreover, by Itô's isometry, we have

$$\sum_{j=k+1}^{N}\sqrt{E\left(|Z(V^\pi)_{t_{j-1}}-Z(\bar V^\pi)_{t_{j-1}}|\Delta t_j\right)^2}$$
$$=\sum_{j=k+1}^{N}\sqrt{E\left\{E[(V_T^\pi-\bar V_T^\pi)(\Delta B_{t_j})^T|\mathcal{F}_{t_{j-1}}]\right\}^2}$$
$$\le \sum_{j=k+1}^{N}\sqrt{E\left\{(V_T^\pi-\bar V_T^\pi)^2\Delta_{t_j}\right\}}$$
$$\le \sqrt{|\pi|}(N-k)\max_{k\le i\le N}\sqrt{E|V_{t_i}^\pi-\bar V_{t_i}^\pi|^2},$$

which completes the proof. ■

Given the above two estimates we prove the local existence of the difference equation (3.3).

Lemma 3.9 *Given the partition π with $\sqrt{|\pi|}\le \frac{1}{6C_1}\wedge 1$, define the integer k by*

$$k=N-\frac{1}{6C_1\sqrt{\pi}}.$$

Then the difference equation (3.3) starting from $V_{t_k}^\pi=0$, i.e.

$$V_{t_i}^\pi=V_{t_{i-1}}^\pi+f(t_{i-1},Y(V^\pi)_{t_{i-1}},Z(V^\pi)_{t_{i-1}})\Delta t_i$$

for $k+1\le i\le N$ and $V_{t_k}^\pi=0$, where

$$\begin{cases} Y(V^\pi)_{t_{i-1}} = & E[\phi(X_T)+V_T^\pi|\mathcal{F}_{t_{i-1}}]-V_{t_{i-1}}^\pi,\\ Z(V^\pi)_{t_{i-1}} = & \dfrac{1}{\Delta t_i}E\left[(\phi(X_T)+V_T^\pi)(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}\right]\end{cases}$$

admits a unique solution in $\mathcal{C}^\pi([t_k,t_N],R)$.

Proof. We first note that on $[t_k,t_N]$, by iterating the difference equation (3.3), we can define the following mapping on $\mathcal{C}^\pi([t_k,t_N],R)$:

$$\mathbb{L}^\pi(V^\pi)_{t_i}:=\sum_{j=k+1}^{i}f(t_{j-1},Y(V^\pi)_{t_{j-1}},Z(V^\pi)_{t_{j-1}})\Delta t_j. \tag{3.10}$$

For any V^π and $\bar{V}^\pi$ in $\mathcal{C}^\pi([t_k, t_N], R)$, by (3.10) we have

$$
\begin{aligned}
&\max_{k\le i\le N}\sqrt{E|\mathbb{L}^\pi(V^\pi)_{t_i}-\mathbb{L}^\pi(\bar{V}^\pi)_{t_i}|^2}\\
&\le \sqrt{E\left(\sum_{j=k+1}^{N}|f(t_{j-1},Y(V^\pi)_{t_{j-1}},Z(V^\pi)_{t_{j-1}})-f(t_{j-1},Y(\bar{V}^\pi)_{t_{j-1}},Z(\bar{V}^\pi)_{t_{j-1}})|\Delta t_j\right)^2}\\
&\le C_1\sqrt{E\left(\sum_{j=k+1}^{N}|Y(V^\pi)_{t_{j-1}}-Y(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j+\sum_{j=k+1}^{N}|Z(V^\pi)_{t_{j-1}}-Z(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2}\\
&\le C_1\sqrt{E\left(\sum_{j=k+1}^{N}|Y(V^\pi)_{t_{j-1}}-Y(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2}\\
&\quad +C_1\sqrt{E\left(\sum_{j=k+1}^{N}|Z(V^\pi)_{t_{j-1}}-Z(\bar{V}^\pi)_{t_{j-1}}|\Delta t_j\right)^2},
\end{aligned}
$$

where we used the Lipschitz property of f in the last but one inequality. Then, by Lemma 3.8 and the definition of k, we further have

$$
\begin{aligned}
&\max_{k\le i\le N}\sqrt{E|\mathbb{L}^\pi(V^\pi)_{t_i}-\mathbb{L}^\pi(\bar{V}^\pi)_{t_i}|^2}\\
&\le 3C_1\sqrt{\pi}(N-k)\max_{k\le i\le N}\sqrt{E|V^\pi_{t_i}-\bar{V}^\pi_{t_i}|^2}\\
&\le \frac{1}{2}\max_{k\le i\le N}\sqrt{E|V^\pi_{t_i}-\bar{V}^\pi_{t_i}|^2}.
\end{aligned}
$$

Therefore the mapping $\mathbb{L}^\pi : \mathcal{C}^\pi([t_k, t_N], R) \to \mathcal{C}^\pi([t_k, t_N], R)$ defined by the difference equation (3.10) is a contracting mapping, and thus admits a unique fixed point. We conclude that the functional difference equation (3.3) starting from $V^\pi_{t_k} = 0$ admits a unique solution in $\mathcal{C}^\pi([t_k, t_N], R)$. ■

The fact that k is independent of the terminal data ϕ allows us to extend the solution to the whole interval $[t_0, t_N]$. We are now in the position to prove Theorem 3.2. We rewrite functional difference equation (3.3) as follows:

$$
\begin{aligned}
V^\pi_{t_i} &= V_{t_{i-1}} + f(t_{i-1}, Y(V^\pi)_{t_{i-1}}, Z(V^\pi)_{t_{i-1}})\Delta t_i\\
&= \sum_{j=1}^{i} f(t_{j-1}, Y(V^\pi)_{t_{j-1}}, Z(V^\pi)_{t_{j-1}})\Delta t_j
\end{aligned}
\tag{3.11}
$$

with Y^π and Z^π being defined iteratively,

$$
\begin{cases}
Y(V^\pi)_{t_{i-1}} &= E[Y(V^\pi)_{t_i} + V^\pi_{t_i}|\mathcal{F}_{t_{i-1}}] - V^\pi_{t_{i-1}},\\
Z(V^\pi)_{t_{i-1}} &= \dfrac{1}{\Delta t_i} E\left[(Y(V^\pi)_{t_i} + V^\pi_{t_i})(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}\right]\\
&= \dfrac{1}{\Delta t_i} E\left[Y(V^\pi)_{t_i}(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}\right]
\end{cases}
$$

for $1 \le i \le N$ and $V^\pi_0 = 0$.

Proof of Theorem 3.2. We choose the partition π and the integer k according to Lemma 3.9, and in particular we let $\sqrt{|\pi|} = \frac{1}{6C_1} \wedge 1$ and $k = N-1$.

We start with the difference equation (3.11) on the last interval $[t_{N-1}, t_N]$ with $V^\pi_{t_{N-1}} = 0$. By Lemma 3.9, (3.11) admits a unique solution $V(N)^\pi \in \mathcal{C}^\pi([t_{N-1}, t_N]; R)$, and, in particular, we obtain $V(N)^\pi_{t_N}$ and

$$Y^\pi_{t_{N-1}} = E\left[\phi(X_T) + V(N)^\pi_{t_N} | \mathcal{F}_{t_{N-1}}\right],$$

which is used as the data to solve the difference equation on $[t_{N-2}, t_{N-1}]$.

In general, for $1 \le i \le N-1$, we consider the difference equation on the interval $[t_{i-1}, t_i]$:

$$V^\pi_{t_i} = f(t_{i-1}, Y^\pi_{t_{i-1}}, Z^\pi_{t_{i-1}})\Delta t_i \tag{3.12}$$

with $V^\pi_{t_{i-1}} = 0$, where

$$\begin{cases} Y^\pi_{t_{i-1}} &= E[Y^\pi_{t_i} + V^\pi_{t_i} | \mathcal{F}_{t_{i-1}}], \\ Z^\pi_{t_{i-1}} &= \dfrac{1}{\Delta t_i} E\left[\left(Y^\pi_{t_i} + V^\pi_{t_i}\right)(\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right] \\ &= \dfrac{1}{\Delta t_i} E\left[Y^\pi_{t_i} (\Delta B_{t_i})^T | \mathcal{F}_{t_{i-1}}\right]. \end{cases}$$

By Lemma 3.9 again, (3.12) admits a unique solution $V(i)^\pi \in \mathcal{C}^\pi([t_{i-1}, t_i]; R)$, and, in particular, we obtain $V(i)^\pi_{t_i}$ and

$$Y^\pi_{t_{i-1}} = E[Y^\pi_{t_i} + V(i)^\pi_{t_i} | \mathcal{F}_{t_{i-1}}],$$

the later of which is used as the data to solve the difference equation on $[t_{i-2}, t_{i-1}]$.

Next we shift the paths of $V(i)^\pi_{t_i}$ to construct the global solution of the difference equation (3.11).

$$\bar{V}^\pi_{t_i} = \sum_{j=1}^{i} V(j)^\pi_{t_j}, \quad \text{for } 1 \le i \le N \tag{3.13}$$

with $\bar{Y}^\pi_{t_{i-1}} = Y^\pi_{t_{i-1}}$ and $\bar{Z}^\pi_{t_{i-1}} = Z^\pi_{t_{i-1}}$. If we can further prove such $\bar{V}^\pi$ with $(\bar{Y}^\pi, \bar{Z}^\pi)$ satisfy functional difference equation (3.11) (or (3.3)), and moreover (3.11) admits a unique solution, then we are done. In fact, for $i = N$, we have

$$\bar{V}^\pi_{t_N} = \bar{V}^\pi_{t_{N-1}} + f(t_{N-1}, Y^\pi_{t_{N-1}}, Z^\pi_{t_{N-1}})\Delta t_N$$

with

$$\begin{aligned} Y^\pi_{t_{N-1}} &= E[\phi(X_T) + V(N)^\pi_{t_N} | \mathcal{F}_{t_{N-1}}] \\ &= E\left[\phi(X_T) + V(N)^T_{t_N} + \sum_{j=1}^{N-1} V(j)^\pi_{t_j} | \mathcal{F}_{t_{N-1}}\right] - \sum_{j=1}^{N-1} V(j)^\pi_{t_j} \\ &= E[\phi(X_T) + \bar{V}^\pi_{t_N} | \mathcal{F}_{t_{N-1}}] - \bar{V}^\pi_{t_{N-1}}, \end{aligned}$$

and

$$
\begin{aligned}
Z^{\pi}_{t_{N-1}} &= \frac{1}{\Delta t_i} E\left[\left(\phi(X_T) + V(N)^{\pi}_{t_N}\right)(\Delta B_{t_N})^T | \mathcal{F}_{t_{N-1}}\right] \\
&= \frac{1}{\Delta t_i} E\left[\left(\phi(X_T) + V(N)^{\pi}_{t_N} + \sum_{j=1}^{N-1} V(j)^{\pi}_{t_j}\right)(\Delta B_{t_N})^T | \mathcal{F}_{t_{N-1}}\right] \\
&\quad - \frac{1}{\Delta t_i} E\left[\left(\sum_{j=1}^{N-1} V(j)^{\pi}_{t_j}\right)(\Delta B_{t_N})^T | \mathcal{F}_{t_{N-1}}\right] \\
&= \frac{1}{\Delta t_i} E\left[\left(\phi(X_T) + \bar{V}^{\pi}_{t_N}\right)(\Delta B_{t_N})^T | \mathcal{F}_{t_{N-1}}\right].
\end{aligned}
$$

Therefore, $\bar{V}^{\pi}_{t_N}$ with $(\bar{Y}^{\pi}_{t_{N-1}}, \bar{Z}^{\pi}_{t_{N-1}})$ satisfy (3.11).

In general, for $1 \le i \le N-1$, by the backward induction, it is easy to verify $\bar{V}^{\pi}_{t_i}$ with $(\bar{Y}^{\pi}_{t_{i-1}}, \bar{Z}^{\pi}_{t_{i-1}})$ also satisfy (3.11).■

We finish this section by proving the uniqueness of the solutions to (3.11), which is in the same spirit of Lemma 2.8 in Chapter 2.

Lemma 3.10 *If functional difference equation (3.11) (or (3.3)) admits a solution in $\mathcal{C}^{\pi}([t_0, t_N]; R)$, then the solution must be unique.*

Proof. Suppose $(V^{\pi,1}_{t_i})_{0\le i\le N}$ and $(V^{\pi,2}_{t_i})_{0\le i\le N}$ are two different solutions to (3.11) with the same terminal data $\phi(X_T)$. For $k = 1,\ 2$, we define

$$
\bar{V}^{\pi,k}_0 = 0; \quad \bar{V}^{\pi,k}_{t_i} = V^{\pi,k}_{t_i} - V^{\pi,k}_{t_{i-1}} \quad \text{for } i = 1, \cdots, N
$$

Then on the last interval $[t_{N-1}, t_N]$, since $(V^{\pi,k}_{t_i})_{0\le i\le N}$ satisfies

$$
V^{\pi,k}_{t_N} = V^{\pi,k}_{t_{N-1}} + f(t_{N-1}, Y(V^{\pi,k})_{t_{N-1}}, Z(V^{\pi,k})_{t_{N-1}})\Delta t_N,
$$

with

$$
\begin{cases}
Y(V^{\pi,k})_{t_{N-1}} &= E[\phi(X_T) + V^{\pi,k}_{t_N} | \mathcal{F}_{t_{N-1}}] - V^{\pi,k}_{t_{N-1}}, \\
Z(V^{\pi,k})_{t_{N-1}} &= \dfrac{1}{\Delta t_N} E\left[\phi(X_T)(\Delta B_{t_N})^T | \mathcal{F}_{t_{N-1}}\right],
\end{cases}
$$

we deduce that $\bar{V}^{\pi,k}_{t_N}$ satisfies

$$
\bar{V}^{\pi,k}_{t_N} = f(t_{N-1}, Y(\bar{V}^{\pi,k})_{t_{N-1}}, Z(\bar{V}^{\pi,k})_{t_{N-1}})\Delta t_N
$$

with

$$
\begin{cases}
Y(\bar{V}^{\pi,k})_{t_{N-1}} &= E[\phi(X_T) + \bar{V}^{\pi,k}_{t_N} | \mathcal{F}_{t_{N-1}}], \\
Z(\bar{V}^{\pi,k})_{t_{N-1}} &= \dfrac{1}{\Delta t_N} E\left[\phi(X_T)(\Delta B_{t_N})^T | \mathcal{F}_{t_{N-1}}\right],
\end{cases}
$$

which admits a unique solution by Lemma 3.9, so $\bar{V}^{\pi,1}_{t_N} = \bar{V}^{\pi,2}_{t_N}$, and $Y(V^{\pi,1})_{t_{N-1}} = Y(V^{\pi,2})_{t_{N-1}}$.

Next on the last but one interval $[t_{N-2}, t_{N-1}]$, since the terminal data $Y(V^{\pi,1})_{t_{N-1}} = Y(V^{\pi,2})_{t_{N-1}}$, we can repeat the above argument to deduce that $\bar{V}^{\pi,1}_{t_{N-1}} = \bar{V}^{\pi,2}_{t_{N-1}}$, and therefore $Y(V^{\pi,1})_{t_{N-2}} = Y(V^{\pi,2})_{t_{N-2}}$.

We continue this procedure until to the interval $[0, t_1]$. Since $V^{\pi,1}_0 = V^{\pi,2}_0 = 0$, $\bar{V}^{\pi,1}_{t_1} = \bar{V}^{\pi,2}_{t_1}$ then implies $V^{\pi,1}_{t_1} = V^{\pi,2}_{t_1}$, which is a contradiction. ■

3.4 A numerical scheme by Picard iteration

Since the proof of Theorem 3.2 is constructive, it also provides us with a numerical scheme as a by-product. In this section, we propose a numerical scheme to functional differential equation (3.1) based on Picard iteration. Let $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ be a partition of $[0, T]$. For $1 \le i \le N$, let P_i represent the total number of Picard iterations on the time interval $[t_{i-1}, t_i]$.

Step (1) In the last interval $[t_{N-1}, t_N]$, we let $V(N)^{\pi,1}_{t_N} = 0$, and for $1 \le q \le P_N$, define

$$V(N)^{\pi,q+1}_{t_N} = f(t_{N-1}, Y^{\pi,q}_{t_{N-1}}, Z^{\pi,q}_{t_{N-1}})\Delta t_N, \tag{3.14}$$

where

$$\begin{cases} Y^{\pi,q}_{t_{N-1}} &= E[\phi(X_T)|\mathcal{F}_{t_{N-1}}] + V(N)^{\pi,q}_{t_N}, \\ Z^{\pi,q}_{t_{N-1}} &= \dfrac{1}{\Delta t_i} E\left[\phi(X_T)(\Delta B_{t_N})^T|\mathcal{F}_{t_{N-1}}\right], \end{cases}$$

from which we obtain $V(N)^{\pi,P_N+1}_{t_N}$ with $(Y^{\pi,P_N}_{t_{N-1}}, Z^{\pi,P_N}_{t_{N-1}})$. Note that we need to compute 2 times conditional expectations to achieve this, and in fact $Z^{\pi,P_N}_{t_{N-1}} = Z^{\pi,q}_{t_{N-1}}$ for $1 \le q \le P_N$.

Step (i) In $[t_{i-1}, t_i]$ for $1 \le i \le N-1$, we let $V(i)^{\pi,1}_{t_i} = 0$, and for $1 \le q \le P_i$, define

$$V(i)^{\pi,q+1}_{t_i} = f(t_{i-1}, Y^{\pi,q}_{t_{i-1}}, Z^{\pi,q}_{t_{i-1}})\Delta t_i, \tag{3.15}$$

where

$$\begin{cases} Y^{\pi,q}_{t_{i-1}} &= E[Y^{\pi,P_{i+1}}_{t_i}|\mathcal{F}_{t_{i-1}}] + V(i)^{\pi,q}_{t_i}, \\ Z^{\pi,q}_{t_{i-1}} &= \dfrac{1}{\Delta t_i} E\left[Y^{\pi,P_{i+1}}_{t_i}(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}\right], \end{cases}$$

from which we obtain $V(i)^{\pi,P_i+1}_{t_i}$ with $(Y^{\pi,P_i}_{t_{i-1}}, Z^{\pi,P_i}_{t_{i-1}})$. Note that, to achieve this, we need to compute 2 times conditional expectations as well.

Step (N+1) Letting $P(i) = \sum_{j=1}^{i} P_j$, we define

$$\bar{V}^{\pi,P(i)}_{t_i} = \sum_{j=1}^{i} V(j)^{\pi,P_j+1}_{t_j}, \quad \text{for } 1 \le i \le N \tag{3.16}$$

with $\bar{Y}^{\pi,P_i}_{t_{i-1}} = Y^{\pi,P_i}_{t_{i-1}}$ and $\bar{Z}^{\pi,P_i}_{t_{i-1}} = Z^{\pi,P_i}_{t_{i-1}}$ as the approximation of functional difference equation (3.3).

Therefore we see that the above numerical algorithm needs to compute $2 \times N$ times conditional expectations, and in order to obtain $\bar{V}^{\pi,P(i)}_{t_i}$ we need to do $P(i)$ times Picard iterations. The main result of this section is the following theorem about the convergence rate of Picard sequence $\bar{V}_{t_i}^{\pi,P(i)}$ to $V_{t_i}^{\pi}$:

Theorem 3.11 *If the coefficients satisfy Lipschitz condition 3.1, by choosing π such that $\sqrt{|\pi|} \leq \frac{1}{6C_1} \wedge 1$, and defining the number of Picard iterations as*

$$P_i = -3\log_2(\sqrt{|\pi|}) + N - i + 1$$

for $1 \leq i \leq N$, then the approximation sequence (3.16) converges to the solution V^{π} of (3.3) with the convergence rate:

$$\max_{1\leq i\leq N} \sqrt{E|\bar{V}_{t_i}^{\pi,P(i)} - V_{t_i}^{\pi}|^2} \leq C_9\sqrt{|\pi|}$$

for a generic constant C_9.

The proof uses the following lemmas.

Lemma 3.12 *Let the partition π be such that $\sqrt{|\pi|} \leq \frac{1}{6C_1} \wedge 1$. Then the Picard sequence $V(i)_{t_i}^{\pi,P_i+1}$ defined by (3.14) and (3.15) has the convergence rate:*

$$\sqrt{E|V(i)_{t_i}^{\pi,P_i+1} - V(i)_{t_i}^{\pi}|^2} \leq \frac{1}{2^{P_i-1}}\left\{C_1T + \frac{1}{3}\sqrt{E|Y_{t_i}^{\pi,P_{i+1}}|^2}\right\},$$

for $1 \leq i \leq N-1$, and

$$\sqrt{E|V(N)_{t_N}^{\pi,P_N+1} - V(N)_{t_N}^{\pi}|^2} \leq \frac{1}{2^{P_N-1}}\left\{C_1T + \frac{1}{3}\sqrt{E|\phi(X_T)|^2}\right\},$$

where $V(i)_{t_i}^{\pi}$ is given by (3.12).

Proof. First by Lemma 3.9, for any $V(i)^{\pi}$ and $\bar{V}(i)^{\pi}$ in $\mathcal{C}^{\pi}([t_{i-1},t_i],R)$ given by (3.12),

$$\sqrt{E|\mathbb{L}^{\pi}(V(i)^{\pi})_{t_i} - \mathbb{L}^{\pi}(\bar{V}(i)^{\pi})_{t_i}|^2} \leq \frac{1}{2}\sqrt{E|V(i)_{t_i}^{\pi} - \bar{V}(i)_{t_i}^{\pi}|^2},$$

and, therefore, for $n \geq 1$, we have

$$\begin{aligned}\sqrt{E|V(i)_{t_i}^{\pi,P_i+1} - V(i)_{t_i}^{\pi,P_i+1+n}|^2} &\leq \sum_{j=1}^{n}\sqrt{E|V(i)_{t_i}^{\pi,P_i+j} - V(i)_{t_i}^{\pi,P_i+1+j}|^2} \\ &\leq \sum_{j=1}^{n}\frac{1}{2^{P_i+j-1}}\sqrt{E|V(i)_{t_i}^{\pi,1} - V(i)_{t_i}^{\pi,2}|^2}. \end{aligned} \tag{3.17}$$

Note that $V(i)_{t_i}^{\pi,1} = 0$, and

$$V(i)_{t_i}^{\pi,2} = f(t_{i-1}, Y_{t_{i-1}}^{\pi,1}, Z_{t_{i-1}}^{\pi,1})\Delta t_i$$

with

$$\begin{cases} Y_{t_{i-1}}^{\pi,1} &= E[Y_{t_i}^{\pi,P_{i+1}}|\mathcal{F}_{t_{i-1}}], \\ Z_{t_{i-1}}^{\pi,1} &= \dfrac{1}{\Delta t_i}E\left[Y_{t_i}^{\pi,P_{i+1}}(\Delta B_{t_i})^T|\mathcal{F}_{t_{i-1}}\right]. \end{cases}$$

Therefore using the same arguments as before we obtain

$$\begin{aligned}
&\sqrt{E|V(i)^{\pi,1}_{t_i} - V(i)^{\pi,2}_{t_i}|^2} \\
&\le C_1\sqrt{E\left|t_{i-1}\Delta t_i\right|^2} + C_1\sqrt{E\left(|Y^{\pi,1}_{t_{i-1}}|\Delta t_i\right)^2} + C_1\sqrt{E\left(|Z^{\pi,1}_{t_{i-1}}|\Delta t_i\right)^2} \\
&\le C_1|\pi|T + C_1|\pi|\sqrt{E|Y^{\pi,P_{i+1}}_{t_i}|^2} + C_1\sqrt{|\pi|}\sqrt{E|Y^{\pi,P_{i+1}}_{t_i}|^2} \\
&\le C_1 T + \frac{1}{3}\sqrt{E|Y^{\pi,P_{i+1}}_{t_i}|^2}.
\end{aligned}$$

Letting $n \to \infty$ in (3.17), we obtain the convergence rate of $V(i)^{\pi,P_i+1}_{t_i}$ for $1 \le i \le N-1$. The proof for the case $i = N$ follows along the similar arguments. ■

Lemma 3.13 *Let the partition π be such that $\sqrt{|\pi|} \le \frac{1}{6C_1} \wedge 1$, for $1 \le i \le N-1$. Then Picard sequence $Y^{\pi,P_{i+1}}_{t_i}$ defined by (3.15) has the upper bound:*

$$\sqrt{E|Y^{\pi,P_{i+1}}_{t_i}|^2} \le \left\{C_1 T + 2||Y^\pi||_{\mathcal{C}^\pi([t_0,t_N];R)}\right\} 2^{N-i}.$$

Proof. We first consider the case $i = N-1$. Since

$$\begin{aligned}
Y^{\pi,P_N}_{t_{N-1}} &= E[\Phi(X_T)|\mathcal{F}_{t_{N-1}}] + V(N)^{\pi,P_N}_{t_N}, \\
Y^{\pi}_{t_{N-1}} &= E[\Phi(X_T)|\mathcal{F}_{t_{N-1}}] + V(N)^{\pi}_{t_N},
\end{aligned}$$

we have,

$$\begin{aligned}
\sqrt{E|Y^{\pi,P_N}_{t_{N-1}}|^2} &\le \sqrt{E|Y^{\pi,P_N}_{t_{N-1}} - Y^\pi_{t_{N-1}}|^2} + \sqrt{E|Y^\pi_{t_{N-1}}|^2} \\
&= \sqrt{E|V(N)^{\pi,P_N}_{t_N} - V(N)^\pi_{t_N}|^2} + \sqrt{E|Y^\pi_{t_{N-1}}|^2} \\
&\le C_1 T + \sqrt{E|\phi(X_T)|^2} + \sqrt{E|Y^\pi_{t_{N-1}}|^2} \\
&\le C_1 T + 2||Y^\pi||_{\mathcal{C}^\pi([t_0,t_N];R)},
\end{aligned}$$

where we used Lemma 3.12 in the last inequality. In general, for $1 \le i \le N-2$, we have

$$\begin{aligned}
&\sqrt{E|Y^{\pi,P_{i+1}}_{t_i}|^2} \\
&\le \sqrt{E|Y^{\pi,P_{i+1}}_{t_i} - Y^\pi_{t_i}|^2} + \sqrt{E|Y^\pi_{t_i}|^2} \\
&= \sqrt{E\left|E(Y^{\pi,P_{i+2}}_{t_{i+1}}|\mathcal{F}_{t_i}) + V(i+1)^{\pi,P_{i+1}}_{t_{i+1}} - E(Y^\pi_{t_{i+1}}|\mathcal{F}_{t_i}) - V(i+1)^\pi_{t_{i+1}}\right|^2} \\
&\quad + \sqrt{E|Y^\pi_{t_i}|^2} \\
&\le \sqrt{E|Y^{\pi,P_{i+2}}_{t_{i+1}} - Y^\pi_{t_{i+1}}|^2} + \sqrt{E|V(i+1)^{\pi,P_{i+1}}_{t_{i+1}} - V(i+1)^\pi_{t_{i+1}}|^2} + \sqrt{E|Y^\pi_{t_i}|^2} \\
&\le \sqrt{E|Y^{\pi,P_{i+2}}_{t_{i+1}} - Y^\pi_{t_{i+1}}|^2} + C_1 T + \sqrt{E|Y^{\pi,P_{i+2}}_{t_{i+1}}|^2} + \sqrt{E|Y^\pi_{t_i}|^2},
\end{aligned}$$

where we used Lemma 3.12 again in the last inequality. Therefore we obtain

$$\begin{aligned}
\sqrt{E|Y^{\pi,P_{i+1}}_{t_i}|^2} &\le 2\sqrt{E|Y^{\pi,P_{i+2}}_{t_{i+1}}|^2} + C_1 T + \sqrt{E|Y^\pi_{t_{i+1}}|^2} + \sqrt{E|Y^\pi_{t_i}|^2} \\
&\le 2\sqrt{E|Y^{\pi,P_{i+2}}_{t_{i+1}}|^2} + C_1 T + 2||Y^\pi||_{\mathcal{C}^\pi([t_0,t_N];R)}.
\end{aligned}$$

By iterating the above inequality, we deduce

$$\sqrt{E|Y_{t_i}^{\pi,P_{i+1}}|^2} \leq \left\{C_1T + 2||Y^\pi||_{\mathcal{C}^\pi([t_0,t_N];R)}\right\} 2^{N-i}.$$

■

We are now ready to provide the proof of Theorem 3.11.

Proof of Theorem 3.11. By the definition of $\bar{V}^{\pi,P(i)}$ and $\bar{V}^\pi$, we have

$$\begin{aligned}
\max_{1\leq i\leq N} \sqrt{E|\bar{V}_{t_i}^{\pi,P(i)} - \bar{V}_{t_i}^\pi|^2} &= \max_{1\leq i\leq N} \sqrt{E\left|\sum_{j=1}^{i} V(j)_{t_j}^{\pi,P_j+1} - V(j)_{t_j}^\pi\right|^2} \\
&= \sqrt{E\left|\sum_{j=1}^{N} V(j)_{t_j}^{\pi,P_j+1} - V(j)_{t_j}^\pi\right|^2} \\
&\leq N \max_{1\leq j\leq N} \sqrt{E|V(j)_{t_j}^{\pi,P_j+1} - V(j)_{t_j}^\pi|^2}.
\end{aligned}$$

By Lemma 3.12 and Lemma 3.13, we obtain

$$\begin{aligned}
&\sqrt{E|V(j)_{t_j}^{\pi,P_j+1} - V(j)_{t_j}^\pi|^2} \\
&\leq \frac{1}{2^{P_j-1}}\left\{C_1T + \frac{1}{3}\left\{C_1T + 2||Y^\pi||_{\mathcal{C}^\pi([t_0,t_N];R)}\right\} 2^{N-i}\right\}.
\end{aligned}$$

Note that by (3.6), we have

$$\begin{aligned}
||Y^\pi||_{\mathcal{C}^\pi([t_0,t_N];R)} &\leq \sup_{t\in[0,T]} \sqrt{E|Y_t^\pi|^2} \\
&\leq \sup_{t\in[0,T]} \sqrt{E|Y_t^\pi - Y_t|^2} + \sup_{t\in[0,T]} \sqrt{E|Y_t|^2} \\
&\leq \sqrt{C_5|\pi|} + \sup_{t\in[0,T]} \sqrt{E|Y_t|^2}.
\end{aligned}$$

Therefore, there exists a generic constant C_9 such that

$$\sqrt{E|V(j)_{t_j}^{\pi,P_j+1} - V(j)_{t_j}^\pi|^2} \leq C_9\frac{2^{N-j}}{2^{P_j-1}} = C_9|\pi|^{\frac{3}{2}},$$

and we easily conclude. ■

Chapter 4

BSDEs with Quadratic Growth (QBSDEs)

4.1 Introduction

In this chapter, we consider a class of BSDEs whose driver $f(t, y, z)$ has at most quadratic growth with respect to z. In the literature such class of BSDEs are usually called QBSDEs for short. QBSDEs are intrinsically related to a class of quasilinear PDEs whose nonlinear terms have quadratic growth with respect to the gradient of the solutions. A typical example is the so called Burger's equations in fluid mechanics, to which the Cole-Hopf transformation can be applied to remove such special kind of nonlinearity. It is Kobylanski [45] who introduced the idea of the Cole-Hopf transformation into QBSDEs and proved the existence of solutions to QBSDEs. QBSDEs have found numerous applications in mathematical finance. For example, they appear naturally when one wants to derive the value function for exponential utility maximization, to use the idea of indifference pricing to hedge contingent claims written on non-tradeable assets, and to consider the risk measure. In the next chapter, we will further discuss their applications in fiance, and in particular in the context of credit risk modeling by indifference pricing.

We will prove the existence of solutions to scalar QBSDEs on a general filtered probability space by using the idea of functional differential equation approach presented in Chapter 2 and the idea of the Cole-Hopf transformation introduced by Kobylanski [45]. For a special class of QBSDE systems (not necessarily scalar) in a Brownian setting, we will pursue another direction different from the existing method for QBSDEs. Namely we do not use the Cole-Hopf transformation at all and do not assume the underlying probability space and Brownian motion as *any given.* Instead, we work with weak solutions of QBSDEs, which is the second subject investigated in this chapter. The advantage of the weak solution method, as we will see in Section 4.3, is that it allows us to really work out the solutions rather than proving the existence only.

We start by introducing the spaces that we will work on. Besides the spaces we have introduced in Chapter 2, we further define the following space:

(1) $\mathcal{C}^\infty([0,T];R^d)$: the space of continuous and $\mathcal{F}_t$-adapted processes $(Y_t)_{t\in[0,T]}$ valued in R^d such that $\sup_{t\in[0,T]}|V_t| \in L^\infty(\Omega,\mathcal{F}_T,\mathbf{P})$.

4.1.1 Main results of this chapter

In the first part of this chapter, we consider the following scalar BSDE in $[0,T]$:

$$\begin{cases} dY_t = -f(t,Y_t,\mathbf{L}(M)_t)dt + dM_t, \\ Y_T = \xi \end{cases} \tag{4.1}$$

on a filtered probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbf{P})$ satisfying the *usual conditions*, where $f:[0,T]\times\Omega\times R\times R^m \to R$ satisfies quadratic growth condition 4.1 to be given in the following, $\xi \in L^\infty(\Omega,\mathcal{F}_T,\mathbf{P})$, and $\mathbf{L}$ satisfies compatibility condition 4.2, which is a stronger version of compatibility condition 2.2 introduced in Chapter 2. The solution to (4.1) is defined in the same manner as Definition 2.1 in Chapter 2.

Condition 4.1 *(Quadratic growth) For any $(y,z) \in R\times R^m$, the driver $f(t,y,z)$ is predictably measurable with $f(\cdot,0,0) \in H^2([0,T];R)$. Moreover $f(t,\cdot,\cdot)$ is continuous for $t\in[0,T]$, a.s., and there exist constants C_1 and C_2 such that*

$$|f(t,y,z)| \le C_1(1+|y|) + C_2|z|^2$$

for $t \in [0,T]$, $y \in R$ and $z\in R^m$, a.s..

Definition 4.1 *For any subinterval $[t_1,t_2]\subset[0,T]$, define an operator $\mathbf{L}_{[t_1,t_2]}:\mathcal{H}^2([t_1,t_2];R)\to H^2([t_1,t_2];R^m)$ such that*

(1) *$\mathbf{L}_{[t_1,t_2]}$ is a linear and bounded operator, i.e. for $\alpha,\bar\alpha\in R$ and M, $\bar M\in\mathcal{H}^2([t_1,t_2];R)$,*

$$\mathbf{L}_{[t_1,t_2]}(\alpha M+\bar\alpha\bar M)_t = \alpha\mathbf{L}_{[t_1,t_2]}(M)_t + \bar\alpha\mathbf{L}_{[t_1,t_2]}(\bar M)_t, \quad \textit{for } t\in[t_1,t_2],$$

and there exists a constant C_3 such that

$$||\mathbf{L}_{[t_1,t_2]}(M)||_{H^2[t_1,t_2]} \le C_3||M||_{\mathcal{C}[t_1,t_2]};$$

(2) *for any $\mathcal{F}_{t_1}$-measurable random variable V_{t_1},*

$$\mathbf{L}_{[t_1,t_2]}(M)_t = \mathbf{L}_{[t_1,t_2]}(M-V_{t_1})_t, \quad \textit{for } t\in[t_1,t_2];$$

(3) *for any nonnegative $\mathcal{F}_t$-adapted process $(G_t)_{t\in[t_1,t_2]}$,*

$$\int_{t_1}^{t_2} G_s|\mathbf{L}_{[t_1,t_2]}(M)_s|^2ds = \int_{t_1}^{t_2} G_s d[M,M]_s.$$

Clearly Definition 4.1 is a stronger version of Definition 2.4 in Chapter 2. The reason we introduce the extra condition (3) is to handle the quadratic growth term in the driver of QBSDEs. We further define the operator $\mathbf{L}$ based on the operator $\mathbf{L}_{[t_1,t_2]}$ defined above:

Condition 4.2 *(Compatibility) For a given partition $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ with the mesh:*

$$|\pi| := \max_{1\leq i\leq N} |t_i - t_{i-1}| \leq \frac{1}{16C_1^2(1+C_2)^2} \wedge 1,$$

the operator $\mathbf{L} : \mathcal{H}^2([0,T];R) \to H^2([0,T];R^m)$ satisfies the condition: for $i = 1, \cdots, N$,

$$\mathbf{L}(M)_t = \mathbf{L}_{[t_{i-1},t_i]}(M)_t, \quad \textit{for } t \in [t_{i-1}, t_i].$$

Then by (3) in Definition 4.1, for any nonnegative $\mathcal{F}_t$-adapted process $(G_t)_{t\in[0,T]}$, we have

$$\begin{aligned}\int_0^T G_t|\mathbf{L}(M)_t|^2 dt &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} G_t|\mathbf{L}_{[t_{i-1},t_i]}(M)_t|^2 dt\\ &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} G_t d[M,M]_t = \int_0^T G_t d[M,M]_t.\end{aligned}$$

Therefore an adapted solution to BSDE (4.1) is a pair of processes $(Y, M) \in \mathcal{C}^\infty([0,T];R) \times \mathcal{H}^2([0,T];R)$, where Y is a real-valued continuous special semimartingale with M as the martingale part of its canonical decomposition, such that

$$\begin{aligned}Y_t &= \xi + \int_t^T f(s, Y_s, \mathbf{L}(M)_s)\, ds - \int_t^T dM_s\\ &= \xi + \sum_{j=i+1}^N \int_{t_{j-1}}^{t_j} f\left(s, Y_s, \mathbf{L}_{[t_{j-1},t_j]}(M)_s\right) ds\\ &\quad + \int_t^{t_i} f\left(s, Y_s, \mathbf{L}_{[t_{i-1},t_i]}(M)_s\right) ds - \int_t^T dM_s, \quad \text{for } t \in [t_{i-1}, t_i],\ a.s..\end{aligned}$$

A typical example of the operator $\mathbf{L}$ is the Brownian martingale representation which was presented in Corollary 2.11, i.e. $\mathbf{L}(M)_t = Z_t$. Then the above integral equation reduces to the classical BSDE in the literature:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T dM_s, \quad \text{for } t \in [0,T],\ a.s..$$

Our first main result of this chapter is the following existence theorem for BSDE (4.1):

Theorem 4.2 *If the driver f satisfies quadratic growth condition 4.1, the operator $\mathbf{L}$ satisfies compatibility condition 4.2, and the terminal data $\xi \in L^\infty(\Omega, \mathcal{F}_T, \mathbf{P})$, say $|\xi| \leq \alpha$, a.s., then BSDE (4.1) admits at least one adapted solution $(Y, M) \in \mathcal{C}^\infty([0,T];R) \times \mathcal{H}^2([0,T];R)$.*

The proof is adapted from Kobylanski [45], which uses the idea of the Cole-Hopf transformation. To illustrate this idea in a simple setting, let us first consider the following special case:

$$Y_t = \xi + \int_t^T |\mathbf{L}(M)_s|^2 ds - \int_t^T dM_s.$$

By making the transformation $X_t = e^{2Y_t}$, and using the conditions on the operator $\mathbf{L}$, we have

$$X_t = e^{2\xi} - \int_t^T 2e^{2Y_t} dM_t,$$

which implies $X_t = E[e^{2\xi}|\mathcal{F}_t]$. Hence the solution is given by $Y_t = \frac{1}{2}\log E[e^{2\xi}|\mathcal{F}_t]$ with M being the martingale part of the canonical decomposition of Y.

So far, all the discussions are based on the assumption that the equations must be solved on *any given* probability space. Nevertheless this requirement sometimes might be too strong. In the second part of the thesis we will pursue the weak solutions for a special class of QBSDEs. Specifically, we consider the following QBSDE system (not necessarily scalar):

$$\begin{cases} dY_t = -h(t, Y_t, Z_t)dt - Z_t f(t, Y_t, Z_t)dt + Z_t dW_t, \\ Y_T = \phi(W_T), \end{cases} \tag{4.2}$$

where $W = (W^1, \cdots, W^d)^T$ is a d-dimensional Brownian motion starting from $\mathbf{P}^x\{W_0 = x\} = 1$ for $x \in R^d$. The coefficients satisfy the following quadratic growth condition which is the same as Lipschitz condition 2.3 in Chapter 2. However we call it quadratic growth condition here due to the structure of the equation.

Condition 4.3 *(Quadratic growth) The coefficients $h : [0,T] \times R^n \times R^{n\times d} \to R^n$, $f : [0,T] \times R^n \times R^{n\times d} \to R^d$ and $\phi : R^d \to R^n$ are continuous. Moreover h and f are Lipschitz continuous in $(y, z) \in R^n \times R^{n\times d}$ with Lipschitz constant C_4, ϕ is Lipschitz continuous in $x \in R^d$ with Lipschitz constant C_5, and ϕ is uniformly bounded, say $\sup_{x\in R^d} |\phi(x)| \le C_5$.*

Because of the term with the coefficient f, the equation has at most quadratic growth, i.e. there exists a constant C_6 such that for any $y \in R^n$ and $z \in R^{n\times d}$,

$$|zf(t, y, z)| \le C_6|z|(t + |y| + |z|).$$

Note that the quadratic growth term in (4.2) is more special than the usual one considered in the literature. However this special structure is enough to cover the most of QBSDE examples known in mathematical finance, at least with some extra conditions added, as we will see in the next chapter. Moreover, most of the existing results for QBSDEs are only for the scalar case $n = 1$. In this thesis we extend to the high-dimensional QBSDE systems.

On the other hand, one may wonder why the terminal data has the special form $\phi(W_T)$. This is only for the presentation's simplicity. The results can be extended without difficulty to the case $\phi(X_T)$ where X is driven by SDE:

$$\begin{cases} dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t, \quad \text{for } t \in [0,T], \\ X_0 = x. \end{cases}$$

with the coefficients a and σ satisfying certain regularity conditions.

In the following, we introduce the notion of weak solutions to the above BSDE (4.2), and when we want to emphasize the difference, we will call the solutions we have previously considered the *strong solutions*.

Definition 4.3 *A weak solution to BSDE (4.2) is a triple* $(\Omega, \mathcal{F}, \mathbf{P}^x)$, $\{\mathcal{F}_t\}$ *and* $(Y, Z^{\mathbf{P}^x}, W)$ *such that*

(1) $(\Omega, \mathcal{F}, \mathbf{P}^x)$ *is a complete probability space with the filtration* $\{\mathcal{F}_t\}$ *satisfying the usual conditions;*

(2) *under such filtered probability space,* $Y, Z^{\mathbf{P}^x}$ *and* W *are* $\mathcal{F}_t$*-adapted. Moreover* Y *is a special semimartingale,* $Z^{\mathbf{P}^x}$ *is the predictable representation of* Y *under* $\mathbf{P}^x$*, and* W *is a Brownian motion starting from* $\mathbf{P}^x\{W_0 = x\} = 1$*;*

(3) *the increments* $\{W_u - W_t : t \leq u \leq T\}$ *are independent of the* σ*-algebra* $\mathcal{F}_t$*;*

(4) *the following integral equation holds:*

$$Y_t = \phi(W_T) + \int_t^T h(s, Y_s, Z_s^{\mathbf{P}^x})ds + \int_t^T Z_s^{\mathbf{P}^x} f(s, Y_s, Z_s^{\mathbf{P}^x})ds - \int_t^T Z_s^{\mathbf{P}^x} dW_s, \quad \text{for } t \in [0, T], \text{ a.s..} \tag{4.3}$$

Before we proceed, we make some comments on the above definition. Firstly, the predictable representation $Z^{\mathbf{P}^x}$ means it is the predictable representation for the martingale part of the canonical decomposition of the special semimartingale Y, and we use the superscript $\mathbf{P}^x$ to emphasize the dependency of the predictable representation on the probability measure $\mathbf{P}^x$.

Secondly, we should mention our definition of weak solutions is more related to Buckdahn et al [13]. The filtration $\{\mathcal{F}_t\}$ plays an important role here. If $\mathcal{F}_t = \mathcal{F}_t^W$, i.e. the filtration only generated by the Brownian motion W augmented by the $\mathbf{P}^x$-null sets in $\mathcal{F}$, the solution turns to be a *strong solution*. In Ma and Zhang [57], such solution is also called a *semi-strong solution*. Actually the smallest filtration for weak solutions is the filtration $\{\mathcal{F}_t^{W,Y,Z}\}$ generated by W, Y, Z and satisfying the *usual conditions*.

Thirdly, (3) automatically holds given the Brownian motion W and the filtration $\{\mathcal{F}_t\}$. In fact such condition simply means $\{\mathcal{F}_t\}$ consists, additionally to $\{\mathcal{F}_t^W\}$, only of independent experiments. In Buckdahn et al [13], such condition is formulated in terms of martingales, i.e. any $\mathcal{F}_t^W$-martingale must also be an $\mathcal{F}_t$-martingale. In Kurtz [46], such kind of condition is also called the *compatibility constraint*. (3) is extremely useful when we want to identify which weak solutions are really *strong solutions* to QBSDEs.

In the following we will suppress the superscript x of $\mathbf{P}^x$ for the notation's simplicity, if no confusion may arise. We describe our idea formally before presenting the existence theorem for BSDE (4.2). The basic idea is using a *strong solution* of FBSDE to construct a weak solution of QBSDE. Let us start with a Brownian motion B on $(\Omega, \mathcal{F}, \mathbf{Q})$ with the filtration $\{\mathcal{F}_t\}$ satisfying the *usual conditions* and consider the following FBSDE:

$$\begin{cases} dX_t = f(t, Y_t, Z_t^{\mathbf{Q}})dt + dB_t, \\ X_0 = x, \\ dY_t = -h(t, Y_t, Z_t^{\mathbf{Q}})dt + Z_t^{\mathbf{Q}} dB_t, \\ Y_T = \phi(X_T). \end{cases} \tag{4.4}$$

Suppose FBSDE (4.4) admits a *strong solution* $(X, Y, Z^{\mathbf{Q}})$. Then define a new probability measure $\mathbf{P}$ by

$$\frac{d\mathbf{P}}{d\mathbf{Q}} = \mathscr{E}(N),$$

where $\mathscr{E}(N)$ is the Doléans-Dade exponential of N with

$$N = -\int_0^{\cdot} \langle f(s, Y_s, Z_s^{\mathbf{Q}}), dB_s \rangle_d.$$

Here, $\langle \cdot, \cdot \rangle_d$ denotes the inner product in R^d. Under the new probability measure $\mathbf{P}$, by Girsanov's theorem, B has the decomposition:

$$\begin{aligned} B &= (B - [B, N]) + [B, N] \\ &= \left(B + \int_0^{\cdot} f(s, Y_s, Z_s^{\mathbf{Q}}) ds \right) - \int_0^{\cdot} f(s, Y_s, Z_s^{\mathbf{Q}}) ds, \end{aligned}$$

where $B - [B, N] = B + \int_0^{\cdot} f(s, Y_s, Z_s^{\mathbf{Q}}) ds$ is a martingale under $\mathbf{P}$, and furthermore by Levy's characterization, it is in fact a Brownian motion under $\mathbf{P}$. We further define W by $W = x + B - [B, N]$. Under the probability measure $\mathbf{P}$ and with the new Brownian motion W, let us rewrite the backward equation in FBSDE (4.4) as

$$dY_t = -h(t, Y_t, Z_t^{\mathbf{Q}}) dt - Z_t^{\mathbf{Q}} f(t, Y_t, Z_t^{\mathbf{Q}}) dt + Z_t^{\mathbf{Q}} dW_t$$

with $Y_T = \phi(W_T)$. If we can prove $Z^{\mathbf{Q}} = Z^{\mathbf{P}}$, then triple $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ and $(Y, Z^{\mathbf{P}}, W)$ is just one weak solution we want to find.

There are mainly three steps needed to be verified. The first step is to establish the invariance property of the predictable representation under the change of probability measure, i.e. $Z^{\mathbf{Q}} = Z^{\mathbf{P}}$; the second step is of course the solvability of FBSDE (4.4); the last step is to prove that the Doléans-Dade exponential $\mathscr{E}(N)$ is a uniform-integrable martingale, in order to guarantee $\mathbf{P}$ is an equivalent probability measure. Our second main result of this chapter is the following theorem:

Theorem 4.4 *If the coefficients satisfy quadratic growth condition 4.3, then BSDE (4.2) admits at least one weak solution* $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ *and* $(Y, Z^{\mathbf{P}}, W)$.

The chapter is organized as follows. Section 4.2 is devoted to the proof of Theorem 4.2. In section 4.2.1 we do some preparations by proving a version of comparison principle for BSDEs on general filtered probability space, which is then employed to prove the existence result for BSDEs with linear growth coefficients. With these preparations, we prove Theorem 4.2 in Section 4.2.2.

Section 4.3 is mainly about the weak solutions of BSDE (4.2). We prove the existence theorem 4.4 in Section 4.3.1. For the special case $n = 1$, i.e. the scalar version of BSDE (4.2), we use Girsanov's theorem reversely to prove the pathwise uniqueness of solutions. The celebrated Yamada-Watanabe theorem then implies the weak solution we constructed is indeed the *strong solution* for the case $n = 1$. The above results are presented in Section 4.3.2.

Section 4.2 is based on [52] by Lepeltier and San Martin and [45] by Kobylanski, and adapted in our setting of general filtered probability space. Section 4.3 is adapted from author's paper [50] with Lyons and Qian.

4.2 Existence of solutions to QBSDEs

4.2.1 BSDEs with linear growth coefficients

The arguments of in the section 4.2 heavily depend on the comparison principle for BSDEs on general filtered probability space, so we first provide a version of comparison principle suitable to later use.

The condition (3) in Definition 4.1 (or Condition 4.2) is in fact only used to control the quadratic growth term in the driver of QBSDEs. Next we give a weaker version of the condition (3), which will be used very often later.

(3)' in Definition 4.1 (or Condition 4.2): *For any nonnegative $\mathcal{F}_t$-adapted process $(G_t)_{t\in[t_1,t_2]}$, there exist constants C_7 and C_8 such that*

$$C_7\int_{t_1}^{t_2} G_s|\mathbf{L}_{[t_1,t_2]}(M)_s|^2ds \leq \int_{t_1}^{t_2} G_s d[M,M]_s \leq C_8\int_{t_1}^{t_2} G_s|\mathbf{L}_{[t_1,t_2]}(M)_s|^2ds,$$

and therefore the operator $\mathbf{L}$ also enjoys the same property: for any nonnegative $\mathcal{F}_t$-adapted process $(G_t)_{t\in[0,T]}$,

$$C_7\int_0^T G_s|\mathbf{L}(M)_s|^2ds \leq \int_0^T G_s d[M,M]_s \leq C_8\int_0^T G_s|\mathbf{L}(M)_s|^2ds.$$

Lemma 4.5 *(Comparison principle) Let $(f^1,\xi^1,\mathbf{L})$ and $(f^2,\xi^2,\mathbf{L})$ be two parameter sets for BSDE (4.1), where f^1 is Lipschitz continuous in $(y,z)\in R\times R^m$ with Lipschitz constant C_9 and f^2 is such that the associated BSDE admits an adapted solution, ξ^1 and ξ^2 are in $L^2(\Omega,\mathcal{F}_T,\mathbf{P})$, and $\mathbf{L}$ satisfies compatibility condition 4.2 with (3) replaced by $(3)'$. Let (Y^1,M^1) and (Y^2,M^2) be the corresponding solutions of the associated BSDEs. If $\delta\xi := \xi^2-\xi^1\leq 0$, a.s., and*

$$f^2(t,Y_t^2,\mathbf{L}(M^2)_t)-f^1(t,Y_t^2,\mathbf{L}(M^2)_t)\leq 0, \quad \text{for } t\in[0,T], \text{ a.s.},$$

then $\delta Y_t := Y_t^2-Y_t^1\leq 0$, for $t\in[0,T]$, a.s..

Proof. The proof is a standard use of Itô-Tanaka's formula. We apply Itô-Tanaka's formula to $e^{\alpha t}(\delta Y_t^+)^2$ for some α, which is to be determined:

$$\begin{aligned} e^{\alpha t}(\delta Y_t^+)^2 &= e^{\alpha T}(\delta\xi^+)^2 - \int_t^T e^{\alpha s}d(\delta Y_s^+)^2 - \int_t^T \alpha e^{\alpha s}(\delta Y_s^+)^2ds \\ &= e^{\alpha T}(\delta\xi^+)^2 - \int_t^T 2e^{\alpha s}\delta Y_s^+ d(\delta Y_s^+) - \int_t^T e^{\alpha s}d[\delta Y^+,\delta Y^+]_s - \int_t^T \alpha e^{\alpha s}(\delta Y_s^+)^2ds. \end{aligned}$$

By denoting L^0 the local time of δY in 0, we have

$$\begin{aligned} d(\delta Y_s^+) =& 1_{\{\delta Y_s>0\}} d(\delta Y_s) + \frac{1}{2} dL_s^0 \\ =& 1_{\{\delta Y_s>0\}} \left\{ f^1(s, Y_s^1, \mathbf{L}(M^1)_s) - f^2(s, Y_s^2, \mathbf{L}(M^2)_s) \right\} ds + 1_{\{\delta Y_s>0\}} d(\delta M_s) + \frac{1}{2} dL_s^0. \end{aligned}$$

Note that the measure dL_t^0 is carried by the set $\{t : Y_t = 0\}$. Therefore,

$$\begin{aligned} e^{\alpha t}(\delta Y_t^+)^2 = e^{\alpha T}(\delta \xi^+)^2 - \int_t^T 2e^{\alpha s}\delta Y_s^+ \left\{ f^1(s, Y_s^1, \mathbf{L}(M^1)_s) - f^2(s, Y_s^2, \mathbf{L}(M^2)_s) \right\} ds \\ - \int_t^T 2e^{\alpha s}\delta Y_s^+ d(\delta M_s) - \int_t^T e^{\alpha s} 1_{\{\delta Y_s>0\}} d[\delta M, \delta M]_s - \int_t^T \alpha e^{\alpha s}(\delta Y_s^+)^2 ds. \quad (4.5) \end{aligned}$$

By compatibility condition 4.2 on the operator $\mathbf{L}$, we have

$$\int_t^T e^{\alpha s} 1_{\{\delta Y_s>0\}} d[\delta M, \delta M]_s \geq C_7 \int_t^T e^{\alpha s} 1_{\{\delta Y_s>0\}} |\delta \mathbf{L}(M)_s|^2 ds,$$

and by the Lipschitz property of f^1,

$$\begin{aligned} f^1(s, Y_s^1, \mathbf{L}(M^1)_s) - f^2(s, Y_s^2, \mathbf{L}(M^2)_s) &\geq f^1(s, Y_s^1, \mathbf{L}(M^1)_s) - f^1(s, Y_s^2, \mathbf{L}(M^2)_s) \\ &\geq -C_9\{|\delta Y_s| + |\delta \mathbf{L}(M)_s|\}. \end{aligned}$$

Hence taking expectation on both sides of (4.5), we have

$$\begin{aligned} E[e^{\alpha t}(\delta Y_t^+)^2] \leq& e^{\alpha T} E[(\delta \xi^+)^2] \\ &+ \int_t^T e^{\alpha s} E \left\{ 2C_9 \delta Y_s^+ [|\delta Y_s| + |\delta \mathbf{L}(M)_s|] - C_7 1_{\{\delta Y_s>0\}} |\delta \mathbf{L}(M)_s|^2 - \alpha(\delta Y_s^+)^2 \right\} ds. \end{aligned}$$

Choosing α large enough, say $\alpha > 2C_9 + C_9^2/C_7$, we obtain

$$2C_9 \delta Y_s^+ \{|\delta Y_s| + |\delta \mathbf{L}(M)_s|\} - C_7 1_{\{\delta Y_s>0\}} |\delta \mathbf{L}(M)_s|^2 - \alpha(\delta Y_s^+)^2 \leq 0,$$

and noting that $E[(\delta \xi^+)^2] = 0$, we obtain that $\delta Y_t \leq 0$. ■

As the first application of the above comparison principle, we prove the existence result for BSDEs with linear growth coefficients:

Theorem 4.6 *If we assume that the following conditions are satisfied: (i) for any $(y, z) \in R \times R^m$, the driver $f(t, y, z)$ is predictable measurable with $f(\cdot, 0, 0) \in H^2([0, T]; R)$; moreover $f(t, \cdot, \cdot)$ is continuous, for $t \in [0, T]$, a.s., and there exists a constant C_{10} such that*

$$|f(t, y, z)| \leq C_{10}(1 + |y| + |z|)$$

for $t \in [0, T]$, $y \in R$ and $z \in R^m$, a.s.; (ii) the operator $\mathbf{L}$ satisfies compatibility condition 4.2 with (3) *replaced by* $(3)'$*; (iii) the terminal data $\xi \in L^\infty(\Omega, \mathcal{F}_T, \mathbf{P})$, say $|\xi| \leq \alpha$, a.s.. Then, BSDE (4.1) admits at least one adapted solution $(Y, M) \in \mathcal{C}^\infty([0, T]; R) \times \mathcal{H}^2([0, T]; R)$.*

Proof. The proof is adapted from Lepeltier and San Martin [52], which is based on the approximation of the driver f by its inf-convolution: for $n \geq C_{10}$, let

$$f^n(t,y,z) = \inf_{(\bar{y},\bar{z})\in Q^{1+m}} \left\{ f(t,\bar{y},\bar{z}) + n|y-\bar{y}| + n|z-\bar{z}| \right\}.$$

We can show that: f^n is still with linear growth, $|f^n(t,y,z)| \leq C_{10}(1+|y|+|z|)$, f^n is Lipschitz continuous in $(y,z) \in R \times R^m$ with Lipschitz constant n, and f^n is monotone increasing in n. Moreover, $f^n(t,y^n,z^n) \to f(t,y,z)$ as $(y^n,z^n) \to (y,z)$ for $t \in [0,T]$. We refer to [52] for the proof.

We deduce from the existence and uniqueness theorem 2.3 in Chapter 2 that the BSDE:

$$Y_t = \xi + \int_t^T f^n(s, Y_s, \mathbf{L}(M)_s)ds - \int_t^T dM_s \tag{4.6}$$

admits a unique adapted solution (Y^n, M^n). Moreover, by the comparison principle lemma 4.5, Y_n is increasing and bounded, i.e.

$$\underline{Y}_t \leq Y_t^n \leq Y_t^{n+1} \leq \bar{Y}_t, \quad \text{for } t \in [0,T], \ a.s.,$$

where $\underline{Y}$ is the solution to the linear BSDE:

$$Y_t = -\alpha - \int_t^T C_{10}(1+|Y_s| + |\mathbf{L}(M)_s|)ds - \int_t^T dM_s,$$

and $\bar{Y}$ is the solution to the linear BSDE:

$$Y_t = \alpha + \int_t^T C_{10}(1+|Y_s| + |\mathbf{L}(M)_s|)ds - \int_t^T dM_s.$$

By the uniqueness of solutions to linear BSDEs (See Corollary 2.10), it is easy to deduce that

$$\begin{aligned}
\underline{Y}_t &= -\alpha e^{C_{10}(T-t)} - (e^{C_{10}(T-t)} - 1) \geq -(\alpha+1)e^{C_{10}T}, \\
\bar{Y}_t &= \alpha e^{C_{10}(T-t)} + (e^{C_{10}(T-t)} - 1) \leq (\alpha+1)e^{C_{10}T}.
\end{aligned}$$

Therefore, there exists a limit process Y which is bounded, and by Dominated Convergence Theorem, the sequence Y^n converges to Y under H^2-norm:

$$\lim_{n\to\infty} E\int_0^T |Y_t^n - Y_t|^2 dt = 0.$$

Next, we prove the convergence of M^n in $\mathcal{H}^2([0,T];R)$. For $n,p \geq C_{10}$, we apply Itô's formula to some auxiliary function $\Phi(Y_t^n - Y_t^p)$:

$$\begin{aligned}
&\Phi(Y_0^n - Y_0^p) \\
=&\int_0^T \Phi'(Y_t^n - Y_t^p)\{f^n(t, Y_t^n, \mathbf{L}(M^n)_t) - f^p(t, Y_t^p, \mathbf{L}(M^p)_t)\}dt \\
&- \int_0^T \Phi'(Y_t^n - Y_t^p)d(M_t^n - M_t^p) - \int_0^T \frac{1}{2}\Phi''(Y_t^n - Y_t^p)d[M^n - M^p, M^n - M^p]_t.
\end{aligned}$$

We can choose $\Phi(x) = x^2$, and take expectation on both sides of the above equality, we obtain

$$
\begin{aligned}
&||M^n - M^p||^2_{\mathcal{C}[0,T]} \\
&\le 2E\int_0^T (Y_t^n - Y_t^p)\{f^n(t, Y_t^n, \mathbf{L}(M^n)_t) - f^p(t, Y_t^p, \mathbf{L}(M^p)_t)\}dt \\
&\le 2\sqrt{E\int_0^T |Y_t^n - Y_t^p|^2 dt}\sqrt{E\int_0^T \left|f^n(t, Y_t^n, \mathbf{L}(M^n)_t) - f^p(t, Y_t^p, \mathbf{L}(M^p)_t)\right|^2 dt}.
\end{aligned}
$$

By the linear growth condition on f^n, and uniform boundedness of Y^n and $\mathbf{L}(M^n)$ under H^2-norm (independent of n), the second integral term of the above inequality is bounded by a constant C_{11} which is independent of n and p. Therefore

$$
||M^n - M^p||^2_{\mathcal{C}[0,T]} \le 2C_{11}||Y^n - Y^p||_{H^2[0,T]},
$$

i.e. M^n is a Cauchy sequence, and thus converges to $M \in \mathcal{H}^2([0,T]; R)$.

Finally we show that the convergence of Y_t^n to Y_t is in fact uniform in $t \in [0,T]$, *a.s.*, so the limit process $(Y_t)_{t\in[0,T]}$ is continuous. Indeed, since $M^n \to M$ in $\mathcal{H}^2([0,T]; R)$,

$$
\lim_{n\to\infty} \sup_{t\in[0,T]} |M_t^n - M_t| = 0, \quad \text{in probability,}
$$

and, thus, there exists a subsequence M_t^{nk} converging to M_t uniformly in $t \in [0,T]$, *a.s.*. On the other hand, since $\mathbf{L}(M^n) \to \mathbf{L}(M)$ and $Y^n \to Y$ under H^2-norm, there exist subsequences, denoted as $\mathbf{L}(M^{nk})$ and Y^{nk} such that

$$
\mathbf{L}(M^{nk})_t \to \mathbf{L}(M)_t, \quad \text{and} \quad Y_t^{nk} \to Y_t, \quad \text{for } a.e.\ t \in [0,T],\ a.s..
$$

Then by the property of f^{nk} and Dominated Convergence Theorem,

$$
\int_0^T |f^{nk}(s, Y_s^{nk}, \mathbf{L}(M^{nk})_s)|ds \to \int_0^T |f(s, Y_s, \mathbf{L}(M)_s)|ds, \quad a.s.,
$$

and Scheffé's lemma implies

$$
\lim_{k\to\infty} \int_0^T |f^{nk}(s, Y_s^{nk}, \mathbf{L}(M^{nk})_s) - f(s, Y_s, \mathbf{L}(M)_s)|ds = 0, \quad a.s..
$$

Therefore,

$$
\sup_{t\in[0,T]} |Y_t^{nk} - Y_t| \le \int_0^T |f^{nk}(s, Y_s^{nk}, \mathbf{L}(M^{nk})_s) - f(s, Y_s, \mathbf{L}(M)_s)|ds + \sup_{t\in[0,T]} |\int_t^T d(M_s^{nk} - M_s)|,
$$

from which we deduce $Y_t^{nk} \to Y_t$ uniformly in $t \in [0,T]$, *a.s.*, so $(Y_t)_{t\in[0,T]}$ is continuous. On the other hand, Y_t^n is monotone increasing, so Dini's theorem yields that the pointwise convergence of Y_t^n to Y_t is in fact uniformly in $t \in [0,T]$, *a.s.*. Then, we can pass the limit in (4.6) and deduce that (Y, M) is an adapted solution to (4.1). ■

We finish this subsection by verifying the uniform boundedness of Y^n and $\mathbf{L}(M^n)$ under H^2-norm:

Lemma 4.7 *Both Y^n and $\mathbf{L}(M^n)$ are uniformly bounded under H^2-norm, i.e.*

$$\sup_n ||Y^n||_{H^2[0,T]} < \infty; \quad \sup_n ||\mathbf{L}(M^n)||_{H^2[0,T]} < \infty.$$

Proof. The first assertion is obvious. To prove the second one, we apply Itô's formula to $(Y_t^n)^2$ which yields that

$$(Y_t^n)^2 = \xi^2 + \int_t^T 2Y_s^n f^n(s, Y_s^n, \mathbf{L}(M^n)_s)ds - \int_t^T 2Y_s^n dM_s^n - \int_t^T d[M^n, M^n]_s.$$

By taking expectation on both sides with $t = 0$, and using the linear growth condition on f^n, we obtain

$$\begin{aligned}
E\int_0^T d[M^n, M^n]_s &\leq E[\xi^2] + 2C_{10}E\int_0^T |Y_s^n|^2 ds + 2C_{10}E\int_0^T |Y_s^n|(1 + |\mathbf{L}(M^n)_s|)ds \\
&\leq E[\xi^2] + 2C_{10}E\int_0^T |Y_s^n|^2 ds \\
&\quad + \frac{C_{10}}{\lambda^2}E\int_0^T |Y_s^n|^2 ds + 2C_{10}\lambda^2 T + 2C_{10}\lambda^2 E\int_0^T |\mathbf{L}(M^n)_s|^2 ds,
\end{aligned}$$

where we used the elementary inequality $2ab \leq a^2/\lambda^2 + \lambda^2 b^2$ in the last step. On the other hand, by the compatibility condition 4.2 on the operator $\mathbf{L}$,

$$E\int_0^T d[M^n, M^n]_s \geq C_7 E\int_0^T |\mathbf{L}(M^n)_s|^2 ds.$$

Therefore, we can choose λ small enough such that $C_7 - 2C_{10}\lambda^2 > 0$, and obtain

$$E\int_0^T |\mathbf{L}(M^n)_s|^2 ds \leq \frac{E[\xi^2] + (2 + 1/\lambda^2)C_{10}||Y^n||_{H^2[0,T]}^2 + 2C_{10}\lambda^2 T}{C_7 - 2C_{10}\lambda^2}.$$

■

4.2.2 Cole-Hopf transformation for QBSDEs

In this subsection, we prove Theorem 4.2 by mainly adapting the techniques of Kobylanski (2000). The basic idea is to transform BSDE (4.1), which is hard to handle, to a new type of BSDE, which is relatively easy to handle. To this end, let

$$X_t = e^{2C_2 Y_t},$$

then by Itô's formula BSDE (4.1) becomes

$$\begin{aligned}
X_t &= e^{2C_2\xi} + \int_t^T 2C_2 X_s \left(f\left(s, \frac{\ln X_s}{2C_2}, \mathbf{L}(M)_s\right) ds - C_2 d[M,M]_s \right) - \int_t^T 2C_2 X_s dM_s \\
&= e^{2C_2\xi} + \int_t^T 2C_2 X_s \left(f\left(s, \frac{\ln X_s}{2C_2}, \mathbf{L}(M)_s\right) - C_2|\mathbf{L}(M)_s|^2 \right) ds - \int_t^T 2C_2 X_s dM_s,
\end{aligned}$$

where we used compatibility condition 4.2 on the operator $\mathbf{L}$ for the last equality. We further truncate the above BSDE by considering:

$$X_t = e^{2C_2\xi} + \int_t^T 2C_2\psi(X_s) \left(f\left(s, \frac{\ln \psi(X_s)}{2C_2}, \mathbf{L}(M)_s\right) - C_2|\mathbf{L}(M)_s|^2 \right) ds - \int_t^T 2C_2\psi(X_s) dM_s, \tag{4.7}$$

where $\psi : R \to [A-\epsilon, B+\epsilon]$ is a mollifier given by

$$\psi(x) = \begin{cases} B+\epsilon & \text{if } \ x \geq B+\epsilon \\ x & \text{if } \ A \leq x \leq B \\ A-\epsilon & \text{if } \ x \leq A-\epsilon \\ C^\infty & \text{otherwise} \end{cases}$$

with A and B being two positive constants to be determined. In fact, they will be chosen as the lower and upper bounds of X later, so $\psi(X_t) = X_t$ for $t \in [0,T]$.

If we further define

$$Z_t = \int_0^t 2C_2 \psi(X_s) dM_s, \quad \text{for } t \in [0,T],$$

since M is the martingale part of the canonical decomposition of Y, $M_0 = Y_0$, and moreover,

$$M_t = Y_0 + \int_0^t \frac{1}{2C_2\psi(X_s)} dZ_s.$$

Thus M is an affine functional of Z. Next we will show $\mathbf{L}(M) = \mathbf{L}(M(Z))$, as an operator defined on Z, also satisfies compatibility condition 4.2 but with (3) replaced by (3)′. We first verify for $i = 1, \cdots, N$, $\mathbf{L}_{[t_{i-1},t_i]}$ only depends on $(Z_t)_{t\in[t_{i-1},t_i]}$. By (2) in Definition 4.1, for $M \in \mathcal{H}^2([t_{i-1},t_i];R)$,

$$\begin{aligned} \mathbf{L}_{[t_{i-1},t_i]}(M) &= \mathbf{L}_{[t_{i-1},t_i]}(M - M_{t_{i-1}}) \\ &= \mathbf{L}_{[t_{i-1},t_i]}\left(Y_0 + \int_0^\cdot \frac{1}{2C_2\psi(X_s)} dZ_s - Y_0 - \int_0^{t_{i-1}} \frac{1}{2C_2\psi(X_s)} dZ_s\right) \\ &= \mathbf{L}_{[t_{i-1},t_i]}\left(\int_{t_{i-1}}^\cdot \frac{1}{2C_2\psi(X_s)} dZ_s\right). \end{aligned}$$

Therefore we can define an operator $\mathbf{L}^*_{[t_{i-1},t_i]} : \mathcal{H}^2([t_{i-1},t_i];R) \to H^2([t_{i-1},t_i];R^m)$ by

$$\mathbf{L}^*_{[t_{i-1},t_i]}(Z)_t = \mathbf{L}_{[t_{i-1},t_i]}\left(\int_{t_{i-1}}^\cdot \frac{1}{2C_2\psi(X_s)} dZ_s\right)_t, \quad \text{for } t \in [t_{i-1},t_i]. \tag{4.8}$$

Proposition 4.8 *The operator* $\mathbf{L}^*_{[t_{i-1},t_i]}$ *defined above satisfies Definition 4.1 but with* (3) *replaced by* (3)′.

Proof. For any $(Z_t)_{t\in[t_{i-1},t_i]} \in \mathcal{H}^2([t_{i-1},t_i];R)$, we verify (1) (2) and (3)′ in Definition 4.1.

(1) $\mathbf{L}^*_{[t_{i-1},t_i]}$ is obviously linear since $\mathbf{L}_{[t_{i-1},t_i]}$ is linear. To verify it is also a bounded operator, we observe that

$$\begin{aligned} ||\mathbf{L}^*_{[t_{i-1},t_i]}(Z)||^2_{H^2[t_{i-1},t_i]} &= E\int_{t_{i-1}}^{t_i} |\mathbf{L}^*_{[t_{i-1},t_i]}(Z)_t|^2 dt \\ &= E\int_{t_{i-1}}^{t_i} \left|\mathbf{L}_{[t_{i-1},t_i]}\left(\int_{t_{i-1}}^\cdot \frac{1}{2C_2\psi(X_s)} dZ_s\right)_t\right|^2 dt \\ &= E\int_{t_{i-1}}^{t_i} d\left[\int_{t_{i-1}}^\cdot \frac{1}{2C_2\psi(X_s)} dZ_s, \int_{t_{i-1}}^\cdot \frac{1}{2C_2\psi(X_s)} dZ_s\right]_t \\ &= E\int_{t_{i-1}}^{t_i} \frac{1}{4C_2^2\psi(X_t)^2} d[Z,Z]_t \\ &\leq \frac{1}{4C_2^2(A-\epsilon)^2}||Z||^2_{\mathcal{C}[t_{i-1},t_i]}. \end{aligned}$$

(2) It is obvious $\mathbf{L}^*_{[t_{i-1},t_i]}(Z - V_{t_{i-1}}) = \mathbf{L}^*_{[t_{i-1},t_i]}(Z)$ for any $\mathcal{F}_{t_1}$-measurable random variable V_{t_1}.

(3)$'$ For any nonnegative $\mathcal{F}_t$-adapted process $(G_t)_{t\in[t_{i-1},t_i]}$,

$$\int_{t_{i-1}}^{t_i} G_t|\mathbf{L}^*_{[t_{i-1},t_i]}(Z)_t|^2 dt = \int_{t_{i-1}}^{t_i} G_t \left| \mathbf{L}_{[t_{i-1},t_i]} \left(\int_{t_{i-1}}^{\cdot} \frac{1}{2C_2\psi(X_s)} dZ_s \right)_t \right|^2 dt$$
$$= \int_{t_{i-1}}^{t_i} \frac{G_t}{4C_2^2\psi(X_t)^2} d[Z,Z]_t.$$

By the boundedness of ψ and nonnegativity of G, we obtain

$$4C_2^2(A-\epsilon)^2 \int_{t_{i-1}}^{t_i} G_t|\mathbf{L}^*_{[t_{i-1},t_i]}(Z)_t|^2 dt \le \int_{t_{i-1}}^{t_i} G_t d[Z,Z]_t \le 4C_2^2(B+\epsilon)^2 \int_{t_{i-1}}^{t_i} G_t|\mathbf{L}^*_{[t_{i-1},t_i]}(Z)_t|^2 dt. \tag{4.9}$$

■

Based on the operator $\mathbf{L}^*_{[t_{i-1},t_i]}$, we can further define an operator $\mathbf{L}^* : \mathcal{H}^2([0,T];R) \to H^2([0,T];R^m)$ in the same manner as compatibility condition 4.2. With all the above preparations, we rewrite BSDE (4.7) as

$$X_t = e^{2C_2\xi} + \int_t^T 2C_2\psi(X_s) \left(f\left(s, \frac{\ln\psi(X_s)}{2C_2}, \mathbf{L}^*(Z)_s \right) - C_2|\mathbf{L}^*(Z)_s|^2 \right) ds - \int_t^T dZ_s. \tag{4.10}$$

Note that the driver of BSDE (4.10), denoted as $F(t, X_t, \mathbf{L}^*(Z)_t)$, satisfies the following condition:

$$- 2C_2\psi(X_t)\left(C_1 + C_1\frac{\ln\psi(X_t)}{2C_2} \right) - 4C_2^2(B+\epsilon)|\mathbf{L}^*(Z)_t|^2$$
$$\le F(t, X_t, \mathbf{L}^*(Z)_t) \le \; 2C_2\psi(X_t)\left(C_1 + C_1\frac{\ln\psi(X_t)}{2C_2} \right), \quad \text{for } t \in [0,T], \; a.s..$$

Next we approximate the driver $F(t, X_t, \mathbf{L}^*(Z)_t)$ of BSDE (4.10) by using the sequence:

$$F^n(t, X_t, \mathbf{L}^*(Z)_t) = (1 - \varphi_n(\mathbf{L}^*(Z)_t)) \left(2C_2\psi(X_t)\left(C_1 + C_1\frac{\ln\psi(X_t)}{2C_2} \right) \right)$$
$$+ \varphi_n(\mathbf{L}^*(Z)_t) F(t, X_t, \mathbf{L}^*(Z)_t),$$

where $\varphi_n : R^m \to [0,1]$ is a mollifier given by

$$\varphi_n(x) = \begin{cases} 1 & \text{if } |x| \le n-1 \\ 0 & \text{if } |x| \ge n \\ C^\infty & \text{otherwise} \end{cases}$$

The approximation driver $F^n(t, X_t, \mathbf{L}^*(Z)_t)$ has the following properties:

Proposition 4.9 *The approximation driver $F^n(t, X_t, \mathbf{L}^*(Z)_t)$ is monotone decreasing, bounded and continuous (so satisfying the conditions in Theorem 4.6). Moreover, for each $\omega \in \Omega$,*

$$\lim_{n\to\infty} F^n(t, X_t, \mathbf{L}^*(Z)_t) = F(t, X_t, \mathbf{L}^*(Z)_t), \quad \textit{for } t \in [0,T].$$

Proof. To prove F^n monotone decreasing, we note that for given $n \ge 1$, if $|\mathbf{L}^*(Z)_t| \le n-1$,

$$F^n(t, X_t, \mathbf{L}^*(Z)_t) - F^{n+1}(t, X_t, \mathbf{L}^*(Z)_t) = F(t, X_t, \mathbf{L}^*(Z)_t) - F(t, X_t, \mathbf{L}^*(Z)_t) = 0.$$

If $n-1 < |\mathbf{L}^*(Z)_t| < n$,

$$\begin{aligned}
&F^n(t, X_t, \mathbf{L}^*(Z)_t) - F^{n+1}(t, X_t, \mathbf{L}^*(Z)_t) \\
&= (1 - \varphi_n(\mathbf{L}^*(Z)_t)) \left(2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) \right) \\
&\quad + \varphi_n(\mathbf{L}^*(Z)_t) F(t, X_t, \mathbf{L}^*(Z)_t) - F(t, X_t, \mathbf{L}^*(Z)_t) \\
&= (1 - \phi(\mathbf{L}^*(Z)_t)) \left(2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) - F(t, X_t, \mathbf{L}^*(Z)_t) \right) \geq 0.
\end{aligned}$$

If $|\mathbf{L}^*(Z)_t| \geq n$,

$$\begin{aligned}
&F^n(t, X_t, \mathbf{L}^*(Z)_t) - F^{n+1}(t, X_t, \mathbf{L}^*(Z)_t) \\
&= 2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) - (1 - \phi_{n+1}(\mathbf{L}^*(Z)_t)) \left(2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) \right) \\
&\quad - \varphi_{n+1}(\mathbf{L}^*(Z)_t) F(t, X_t, \mathbf{L}^*(Z)_t) \\
&= \varphi_{n+1}(\mathbf{L}^*(Z)_t) \left(2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) - F(t, X_t, \mathbf{L}^*(Z)_t) \right) \geq 0.
\end{aligned}$$

To prove the boundedness of F^n, we note that for given $n \geq 1$, if $|\mathbf{L}^*(Z)_t| \leq n-1$,

$$\begin{aligned}
|F^n(t, X_t, \mathbf{L}^*(Z)_t)| &= |F(t, X_t, \mathbf{L}^*(Z)_t)| \\
&\leq 2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) + 4C_2^2(B+\epsilon)|\mathbf{L}^*(Z)_t|^2 \\
&\leq 2C_2(B+\epsilon) \left(C_1 + C_1 \frac{\ln(B+\epsilon)}{2C_2} \right) + 4C_2^2(B+\epsilon)(n-1)^2.
\end{aligned}$$

If $n-1 < |\mathbf{L}^*(Z)_t| < n$,

$$\begin{aligned}
|F^n(t, X_t, \mathbf{L}^*(Z)_t)| &\leq 2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) + |F(t, X_t, \mathbf{L}^*(Z)_t)| \\
&\leq 4C_2(B+\epsilon) \left(C_1 + C_1 \frac{\ln(B+\epsilon)}{2C_2} \right) + 4C_2^2(B+\epsilon)n^2.
\end{aligned}$$

If $\mathbf{L}^*(Z)_t \geq n$,

$$\begin{aligned}
|F^n(t, X_t, \mathbf{L}^*(Z)_t)| &= 2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) \\
&\leq 2C_2(B+\epsilon) \left(C_1 + C_1 \frac{\ln(B+\epsilon)}{2C_2} \right).
\end{aligned}$$

The continuity follows from the conditions on the driver F and the definition of mollifiers ϕ_n and ψ. It follows immediately from the definition of F^n that the sequence F^n converges to F pointwise. ■

By the above Proposition, $F^n(t, X_t, \mathbf{L}^*(Z)_t)$ satisfies the condition:

$$\begin{aligned}
&- 2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right) - 4C_2^2(B+\epsilon)|\mathbf{L}^*(Z)_t|^2 \\
&\leq F^n(t, X_t, \mathbf{L}^*(Z)_t) \leq \; 2C_2\psi(X_t) \left(C_1 + C_1 \frac{\ln \psi(X_t)}{2C_2} \right), \quad \text{for } t \in [0,T], \; a.s..
\end{aligned}$$

Employing the approximation driver F^n, we use the following auxiliary sequence of BSDEs, which are easy to handle, to approximate BSDE (4.10),

$$X_t = e^{2C_2\xi} + \int_t^T F^n(s, X_s, \mathbf{L}^*(Z)_s)ds - \int_t^T dZ_s. \tag{4.11}$$

By Theorem 4.6, for each n, BSDE (4.11) admits an adapted solution (X^n, Z^n). Moreover, by comparison principle Lemma 4.5,

$$\underline{X}_t \le X_t^{n+1} \le X_t^n \le \overline{X}_t, \quad \text{for } t \in [0,T], \ a.s.,$$

where $\underline{X}$ is the solution of the BSDE:

$$X_t = e^{-2C_2\alpha} - \int_t^T 2C_2\psi(X_s)\left(C_1 + C_1\frac{\ln\psi(X_s)}{2C_2}\right) + 4C_2^2(B+\epsilon)|\mathbf{L}^*(Z)_s|^2 ds - \int_t^T dZ_s,$$

and $\overline{X}$ is the solution of the BSDE:

$$X_t = e^{2C_2\alpha} + \int_t^T 2C_2\psi(X_s)\left(C_1 + C_1\frac{\ln\psi(X_s)}{2C_2}\right) ds - \int_t^T dZ_s.$$

By Corollary 2.10 in Chapter 2, $\underline{X}$ and $\overline{X}$ can in fact be obtained by solving the ODEs:

$$X_t = e^{-2C_2\alpha} - \int_t^T 2C_2X_s\left(C_1 + C_1\frac{\ln X_s}{2C_2}\right) ds,$$

and

$$X_t = e^{2C_2\alpha} + \int_t^T 2C_2X_s\left(C_1 + C_1\frac{\ln X_s}{2C_2}\right) ds,$$

whose solutions are given by

$$\underline{X}_t = e^{2C_2[(1-\alpha)e^{-C_1(T-t)}-1]} \ge e^{2C_2[(1-\alpha)e^{-C_1T}-1]},$$
$$\bar{X}_t = e^{2C_2[(1+\alpha)e^{C_1(T-t)}-1]} \le e^{2C_2[(1+\alpha)e^{C_1T}-1]}.$$

We then choose the lower bound $A = e^{2C_2[(1-\alpha)e^{-C_1T}-1]}$ and the upper bound $B = e^{2C_2[(1+\alpha)e^{C_1T}-1]}$. Then $\psi(X_t) = X_t$ for $t \in [0,T]$. There exists a limit process X which is bounded, and by dominated convergence theorem,

$$\lim_{n\to\infty} E\int_0^T |X_t^n - X_t|^2 dt = 0.$$

The rest of the proof follows from arguments similar to the ones used for the proof of Theorem 4.6, provided we have the following result about the convergence of the sequence Z^n.

Lemma 4.10 *The sequence Z^n is Cauchy in $\mathcal{H}^2([0,T];R)$, and therefore, there exists a limit process denoted by $Z \in \mathcal{H}^2([0,T];R)$.*

Proof. First we prove that $\mathbf{L}^*(Z^n)$ is bounded in $H^2([0,T];R^m)$, which will give us a weak convergence limit. The idea is the same as Lemma 4.7. Applying Itô's formula to $(X_t^n)^2$, we have

$$(X_t^n)^2 = e^{4C_2\xi} + \int_t^T 2X_s^n F^n(s, X_s^n, \mathbf{L}^*(Z^n)_s)ds - \int_t^T 2X_s^n dZ_s^n - \int_t^T d[Z^n, Z^n]_s.$$

For $t = 0$, taking expectation on both sides, and using nonnegativity of X_s^n and the upper bound of F^n, we obtain

$$E\int_0^T d[Z^n, Z^n]_s \le E[e^{4C_2\xi}] + 4C_2 E\int_0^T (X_s^n)^2 \left(C_1 + C_1 \frac{\ln X_s^n}{2C_2}\right) ds.$$

Because $|X_s^n| \le \bar{X}_s \le B$ for $s \in [0, T]$, *a.s.*, and $|\xi| \le \alpha$, *a.s.*, the right hand side of the above inequality is uniformly bounded. Moreover, by the compatibility condition 4.2 on the operator $\mathbf{L}^*$ (recalling (4.9) in Proposition 4.8), we have

$$E\int_0^T |\mathbf{L}^*(Z^n)_t|^2 dt \le \frac{1}{4C_2^2(A-\epsilon)^2} E\int_0^T d[Z^n, Z^n]_t.$$

Therefore, there exists a subsequence, still denoted by $\mathbf{L}^*(Z^n)$, which converges weakly to a limit process $\mathbf{L}^*(Z)$ in $H^2([0,T]; R^m)$.

Next we prove $\mathbf{L}^*(Z^n)$ in fact converges to $\mathbf{L}^*(Z)$ in $H^2([0,T]; R^m)$. The idea is to find an appropriate auxiliary function Ψ, which is in the same spirit of Theorem 4.6. For $n \le p$, applying Itô's formula to some auxiliary function $\Psi(X_t^n - X_t^p)$, and taking expectation, we have

$$\begin{aligned} &E[\Psi(X_0^n - X_0^p)] + E\int_0^T \frac{1}{2}\Psi''(X_t^n - X_t^p) d[Z^n - Z^p, Z^n - Z^p]_t \\ =&E\int_0^T \Psi'(X_t^n - X_t^p)\left\{F^n(t, X_t^n, \mathbf{L}^*(Z^n)_t) - F^p(t, X_t^p, \mathbf{L}^*(Z^p)_t)\right\} dt \end{aligned} \tag{4.12}$$

If $\Psi'' \ge 0$, by the compatibility condition 4.2 on the operator $\mathbf{L}^*$ ((4.9) in Proposition 4.8), we have

$$\begin{aligned} \int_0^T \frac{1}{2}\Psi''(X_t^n - X_t^p) d[Z^n - Z^p, Z^n - Z^p]_t &\ge 4C_2^2(A-\epsilon)^2 \int_0^T \frac{1}{2}\Psi''(X_t^n - X_t^p)|\mathbf{L}^*(Z^n - Z^p)_t|^2 dt \\ &= C_{11}\int_0^T \frac{1}{2}\Psi''(X_t^n - X_t^p)|\mathbf{L}^*(Z^n - Z^p)_t|^2 dt. \end{aligned}$$

By using the upper bound of F^n and the lower bound of F^p, we have

$$\begin{aligned} &F^n(t, X_t^n, \mathbf{L}^*(Z^n)_t) - F^p(t, X_t^p, \mathbf{L}^*(Z^p)_t) \\ \le\ & 2C_2 X_t^n \left(C_1 + C_1\frac{\ln X_t^n}{2C_2}\right) + 2C_2 X_t^p \left(C_1 + C_1\frac{\ln X_t^p}{2C_2}\right) + 4C_2^2(B+\epsilon)|\mathbf{L}^*(Z^p)_t|^2 \\ \le\ & C_{12} + C_{13}|\mathbf{L}^*(Z^p)_t|^2 \end{aligned}$$

with $C_{13} = 4C_2^2(B+\epsilon)$, so if $\Psi' \ge 0$, again by the compatibility condition 4.2 on the operator $\mathbf{L}^*$,

$$\begin{aligned} &\int_0^T \Psi'(X_t^n - X_t^p)\left\{C_{12} + C_{13}|\mathbf{L}^*(Z^p)_t|^2\right\} dt \\ \le &\int_0^T \Psi'(X_t^n - X_t^p)\left\{C_{12} + 3C_{13}|\mathbf{L}^*(Z^n - Z^p)_t|^2 + 3C_{13}|\mathbf{L}^*(Z^n - Z)_t|^2 + 3C_{13}|\mathbf{L}^*(Z)_t|^2\right\} dt. \end{aligned}$$

Therefore we obtain

$$\begin{aligned} &E[\Psi(X_0^n - X_0^p)] + E\int_0^T (\frac{1}{2}C_{11}\Psi'' - 3C_{13}\Psi')(X_t^n - X_t^p)|\mathbf{L}^*(Z^n - Z^p)_t|^2 dt \\ \le\ & E\int_0^T 3C_{13}\Psi'(X_t^n - X_t^p)|\mathbf{L}^*(Z^n - Z)_t|^2 dt \\ &+ E\int_0^T \Psi'(X_t^n - X_t^p)\{C_{12} + 3C_{13}|\mathbf{L}^*(Z)_t|^2\} dt. \end{aligned} \tag{4.13}$$

If we have the uniform boundedness of $\left\{(\frac{1}{2}C_{11}\Psi'' - 3C_{13}\Psi')^{1/2}(X_t^n - X_t^p)\right\}_{t\in[0,T]}$ in terms of p, since $\mathbf{L}^*(Z^n - Z^p)$ converges weakly to $\mathbf{L}^*(Z^n - Z)$ in $H^2([0,T];R^m)$ as $p \to \infty$,

$$\left\{(\frac{1}{2}C_{11}\Psi'' - 3C_{13}\Psi')^{1/2}(X_t^n - X_t^p)\mathbf{L}^*(Z^n - Z^p)_t\right\}_{t\in[0,T]} \quad \text{converges weakly to}$$

$$\left\{(\frac{1}{2}C_{11}\Psi'' - 3C_{13}\Psi')^{1/2}(X_t^n - X_t)\mathbf{L}^*(Z^n - Z)_t\right\}_{t\in[0,T]} \quad \text{in } H^2([0,T];R^m) \text{ as } p \to \infty.$$

Hence using that $||u|| \leq \liminf_{p\to\infty} ||u^p||$ if u^p converges weakly to u as $p \to \infty$, we deduce that

$$\begin{aligned} & E\int_0^T (\frac{1}{2}C_{11}\Psi'' - 3C_{13}\Psi')(X_t^n - X_t)|\mathbf{L}^*(Z^n - Z)_t|^2 dt \\ \leq & \liminf_{p\to\infty} E\int_0^T (\frac{1}{2}C_{11}\Psi'' - 3C_{13}\Psi')(X_t^n - X_t^p)|\mathbf{L}^*(Z^n - Z^p)_t|^2 dt. \end{aligned} \tag{4.14}$$

On the other hand, if we have the uniform boundedness of $\{\Psi'(X_t^n - X_t^p)\}_{t\in[0,T]}$ in terms of p, Dominated Convergence Theorem then implies that

$$\begin{aligned} & \lim_{p\to\infty} E\int_0^T \Psi'(X_t^n - X_t^p)\left\{3C_{13}|\mathbf{L}^*(Z^n - Z)_t|^2 + C_{12} + 3C_{13}|\mathbf{L}^*(Z)_t|^2\right\} dt \\ = & E\int_0^T \Psi'(X_t^n - X_t)\left\{3C_{13}|\mathbf{L}^*(Z^n - Z)_t|^2 + C_{12} + 3C_{13}|\mathbf{L}^*(Z)_t|^2\right\} dt. \end{aligned} \tag{4.15}$$

If $\Psi \geq 0$, we let $p \to \infty$ in (4.13) and use the inequalities (4.14) and (4.15),

$$\begin{aligned} & E\int_0^T (\frac{1}{2}C_{11}\Psi'' - 6C_{13}\Psi')(X_t^n - X_t)|\mathbf{L}^*(Z^n - Z)_t|^2 dt \\ \leq & E\int_0^T \Psi'(X_t^n - X_t)\{C_{12} + 3C_{13}|\mathbf{L}^*(Z)_t|^2\} dt. \end{aligned}$$

Now we can choose the auxiliary function:

$$\Psi(x) = \exp\left(\frac{12C_{13}}{C_{11}}x\right) - \frac{12C_{13}}{C_{11}}x - 1, \quad \text{for } x \geq 0,$$

so that $\frac{1}{2}C_{11}\Psi'' - 6C_{13}\Psi' = 72C_{13}^2/C_{11}$, and $\Psi, \Psi', \Psi'' \geq 0$. (Recall $C_{11} = 4C_2^2(A-\epsilon)^2$, $C_{13} = 4C_2^2(B+\epsilon)$.) Therefore $\mathbf{L}^*(Z^n)$ converges to $\mathbf{L}^*(Z)$ in $H^2([0,T];R^m)$. The convergence of Z^n in $\mathcal{H}^2([0,T];R)$ then follows from the compatibility condition 4.2. ■

4.3 Weak solutions of QBSDEs

4.3.1 Existence of weak solutions

In this subsection, we mainly consider weak solutions of BSDE (4.2), and prove the existence Theorem 4.4. Different notions of uniqueness are also discussed.

As we already discussed in Section 4.1.1, there are mainly three steps needed to be verified for the solving procedure. The first step is to establish the invariance property of the predictable representation under the change of probability measure; the second step is the solvability of FBSDE (4.4); the last step is to prove that the Doléans-Dade exponential $\mathscr{E}(N)$ is a uniform-integrable martingale in order to guarantee $\mathbf{P}$ is an equivalent probability measure. We verify these three steps next.

4.3.1.1 Invariance property of martingale representation

The following proposition follows almost trivially but crucial to our results. It states that the predictable representation of a special semimartingale is invariant under the equivalent change of probability measure.

Proposition 4.11 *Let B be a Brownian motion on $(\Omega, \mathcal{F}, \mathbf{Q})$ with the filtration $\{\mathcal{F}_t\}$ satisfying the usual conditions. Let $Z^{\mathbf{Q}}$ be the predictable representation of a special semimartingale Y under $\mathbf{Q}$. If define an equivalent probability measure $\mathbf{P}$ by $\frac{d\mathbf{P}}{d\mathbf{Q}} = \mathscr{E}(N)$ for some uniform integrable martingale $\mathscr{E}(N)$, then $Z_t^{\mathbf{P}} = Z_t^{\mathbf{Q}}$ for a.e. $t \in [0,T]$, a.s.*

Proof. Under the probability measure $\mathbf{Q}$, Y has the canonical decomposition $Y = M - V$ with M being a local martingale and V being a finite variation process, and moreover, M admits the martingale representation:

$$M_t - M_0 = \int_0^t Z_s^{\mathbf{Q}} dB_s, \quad \text{for } t \in [0,T], \ a.s. \tag{4.16}$$

for some predictably measurable process $Z^{\mathbf{Q}}$.

By Girsanov's theorem, Y is still a special semimartingale under $\mathbf{P}$ but with the canonical decomposition $Y = \bar{M} - \bar{V}$, where $\bar{M} = M - [M,N]$ is a martingale, and $\bar{V} = \bar{M} - Y$ is a finite variation process. We also have $\bar{B} = B - [B,N]$ as a Brownian motion under $\mathbf{P}$. Hence under $\mathbf{P}$, (4.16) becomes

$$\bar{M}_t - \bar{M}_0 + [M,N]_t = \int_0^t Z_s^{\mathbf{Q}} d\bar{B}_s + \int_0^t Z_s^{\mathbf{Q}} d[B,N]_s, \quad \text{for } t \in [0,T], \ a.s..$$

Identifying the martingale parts and finite variation parts of the above equality, we must have

$$\bar{M}_t - \bar{M}_0 = \int_0^t Z_s^{\mathbf{Q}} d\bar{B}_s, \quad \text{for } t \in [0,T], \ a.s..$$

On the other hand, under $\mathbf{P}$, we also have

$$\bar{M}_t - \bar{M}_0 = \int_0^t Z_s^{\mathbf{P}} d\bar{B}_s, \quad \text{for } t \in [0,T], \ a.s.$$

for some predictably measurable process $Z^{\mathbf{P}}$, so

$$\int_0^T |Z_s^{\mathbf{P}} - Z_s^{\mathbf{Q}}|^2 ds = 0, \quad a.s.,$$

we easily conclude. ■

Since the martingale representation usually determines the hedging (or replicating) strategy in derivative valuation, a direct consequence of Proposition 4.11 is that the hedging strategy is independent of the choice of equivalent (martingale) probability measures. Due to Proposition 4.11, we will not emphasize the dependency of the martingale representation on the probability measure, and will simply write it as Z from now on.

4.3.1.2 Constructing weak solutions by strong solutions of FBSDEs

In this subsection we use *strong solutions* of FBSDEs to construct weak solutions of QBSDEs. We start with a Brownian motion B on $(\Omega, \mathcal{F}, \mathbf{Q})$ with the filtration $\{\mathcal{F}_t\}$ satisfying the *usual conditions* and consider FBSDE (4.4), i.e.

$$\begin{cases} dX_t = f(t, Y_t, Z_t)dt + dB_t, \\ X_0 = x, \\ dY_t = -h(t, Y_t, Z_t)dt + Z_t dB_t, \\ Y_T = \phi(X_T). \end{cases} \tag{4.17}$$

The solvability of FBSDE (4.17) has already been presented in Section 2.3.3 of Chapter 2. We recall Theorem 2.12 of Chapter 2 as the following lemma for our convenience.

Lemma 4.12 *If the coefficients satisfy quadratic growth condition 4.3, then there exists at least one strong solution* $(Y, Z) \in \mathcal{C}^\infty([0,T]; R^n) \times H^2([0,T]; R^{n\times d})$ *together with the forward process* $X \in \mathcal{C}([0,T]; R^n)$ *of FBSDE (4.17).*

Based on the *strong solution* (Y, Z), we define a new probability measure $\mathbf{P}$ by

$$\frac{d\mathbf{P}}{d\mathbf{Q}} = \mathscr{E}(N),$$

where $\mathscr{E}(N)$ is the Doléans-Dade exponential of N with

$$N = -\int_0^\cdot \langle f(s, Y_s, Z_s), dB_s\rangle_d,$$

where $\langle \cdot, \cdot \rangle_d$ denotes the inner product in R^d. By Lemma 4.13 in the next subsection, we know $\mathbf{P}$ is indeed an equivalent probability measure. Under the new probability measure $\mathbf{P}$, by Girsanov's theorem, B has the following decomposition:

$$\begin{aligned} B &= (B - [B, N]) + [B, N], \\ &= \left(B + \int_0^\cdot f(s, Y_s, Z_s)ds\right) - \int_0^\cdot f(s, Y_s, Z_s)ds, \end{aligned}$$

where $B - [B, N] = B + \int_0^\cdot f(s, Y_s, Z_s)ds$ is a martingale under $\mathbf{P}$, and furthermore by Levy's characterization, it is in fact a Brownian motion under $\mathbf{P}$. We further define W by $W = x + B - [B, N]$. Under the probability measure $\mathbf{P}$ and with the new Brownian motion W, we rewrite the backward equation in FBSDE (4.17):

$$dY_t = -h(t, Y_t, Z_t)dt - Z_t f(t, Y_t, Z_t)dt + Z_t dW_t$$

with $Y_T = \phi(W_T)$. Therefore the triple $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ and (Y, Z, W) is just one weak solution we want to find.

4.3.1.3 Uniformly integrability of stochastic exponential

In this subsection we verify that the Doléans-Dade exponential $\mathscr{E}(N)$ is a uniform-integrable martingale. To prove this we need an appropriate martingale space. It turns out the correct martingale space is BMO-martingale space. For the definition of BMO-martingales and further details, we refer to Appendix A.2.

Lemma 4.13 *If the coefficients satisfy quadratic growth condition 4.3, then*

$$N = -\int_0^{\cdot} \langle f(s, Y_s, Z_s), dB_s\rangle_d$$

is a BMO-martingale under $\mathbf{Q}$*, and therefore the Doléans-Dade exponential $\mathscr{E}(N)$ is a uniform-integrable martingale under* $\mathbf{Q}$ *by Theorem A.5.*

Proof. For any stopping time $\tau \leq T$, by Itô's isometry and the linear growth condition $|f(t, y, z)| \leq C_4(t + |y| + |z|)$, we obtain

$$\begin{aligned} &\sup_\tau E\left[|N_T - N_\tau|^2|\mathcal{F}_\tau\right] \\ =\ &\sup_\tau E\left[\int_\tau^T |f(t, Y_t, Z_t)|^2 dt \mid \mathcal{F}_\tau\right] \\ \leq\ &C_4^2 T^3 + 3C_4^2 \sup_\tau E\left[\int_\tau^T |Y_t|^2 dt \mid \mathcal{F}_\tau\right] + 3C_4^2 \sup_\tau E\left[\int_\tau^T |Z_t|^2 dt \mid \mathcal{F}_\tau\right], \quad a.s.. \end{aligned} \tag{4.18}$$

Since Y is uniformly bounded, we only need to control the last term of (4.18). By applying Itô's formula to $(Y_t)^2$ and taking the conditional expectation on $\mathcal{F}_\tau$, we obtain

$$\begin{aligned} &(Y_\tau)^2 + E\left[\int_\tau^T |Z_t|^2 dt \mid \mathcal{F}_\tau\right] \\ =\ &E[\phi(X_T)^2|\mathcal{F}_\tau] + 2E\left[\int_\tau^T \langle Y_t, h(t, Y_t, Z_t)\rangle_n dt \mid \mathcal{F}_\tau\right] \\ \leq\ &M^2 + \lambda^2 E\left[\int_\tau^T |Y_t|^2 dt \mid \mathcal{F}_\tau\right] + \frac{1}{\lambda^2} E\left[\int_\tau^T |h(t, Y_t, Z_t)|^2 dt \mid \mathcal{F}_\tau\right] \\ \leq\ &M^2 + \lambda^2 E\left[\int_\tau^T |Y_t|^2 dt \mid \mathcal{F}_\tau\right] + \frac{C_4^2 T^3}{\lambda^2} + \frac{3C_4^2}{\lambda^2} E\left[\int_\tau^T |Y_t|^2 dt \mid \mathcal{F}_\tau\right] \\ &+ \frac{3C_4^2}{\lambda^2} E\left[\int_\tau^T |Z_t|^2 dt \mid \mathcal{F}_\tau\right], \quad a.s., \end{aligned}$$

where we used the elementary inequality $2ab \leq \lambda^2 a^2 + b^2/\lambda^2$. By choosing λ large enough such that $1 - 3C_4^2/\lambda^2 > 0$, and by the unform boundedness of Y, we deduce that there exists a constant C_{14} independent of τ such that

$$\sup_\tau E\left[\int_\tau^T |Z_t|^2 dt \mid \mathcal{F}_\tau\right] \leq C_{14}, \quad a.s.,$$

and the conclusion follows by plugging the above estimate into (4.18). ■

4.3.2 Pathwise uniqueness and strong solutions for $n = 1$

As in the classical theory of SDEs there are several notions of uniqueness for solutions, there are also different notions of uniqueness for the solutions to BSDEs. In this subsection, we discuss different notions of uniqueness for BSDE (4.2) and their relationship.

4.3.2.1 Canonical setup for solutions of QBSDEs

In this subsection we interpret weak solutions to BSDE (4.2) as the canonical setup for stochastic processes. For reasons which will become clear later on, we impose the following condition on the coefficients, which is a stronger version of the quadratic growth condition 4.3.

Condition 4.4 *(One-dimension) Quadratic growth condition 4.3 is assumed to be satisfied. Moreover, (i) $n = 1$, i.e. BSDE (4.2) is a scalar BSDE; (ii) $F = F(t, z)$ with $F = (h, f)^T$, i.e. both of the coefficients h and f only depend on t and z; (iii) $f^j = f^j(t, z^j)$ for $j = 1, \cdots, d$, i.e. there is no mixture terms of z in f.*

If we regard weak solutions as the canonical setup for stochastic processes, then the weak solution we constructed in the last subsection for BSDE (4.2) can be regarded as a probability distribution on some sample space $\mathbf{E}^T$. (See Appendix A.4 for $\mathbf{E}^T$.) Usually we need certain topological structure on sample path space. From Definition 4.3 for weak solutions to BSDE (4.2), we know that Y and W must be continuous. However it is not obvious at all that the predictable representation Z has any path regularity. Fortunately, under the One-dimensional Condition 4.4, Imkeller and Dos Reis [40] already did this job for us:

Lemma 4.14 *(Imkeller and Dos Reis [40]) Under the One-dimensional Condition 4.4, there is a continuous modification of Z for the solutions to BSDE (4.2).*

Therefore we will choose such continuous version of Z from now on and our sample path space will then be the space of continuous functions. By $\mathbf{W}^m$ we denote the space of continuous functions $C([0, T]; R^m)$. Define the coordinate mapping $X_t : \mathbf{W}^m \to R^m$ by

$$X_t(x) = x_t, \quad \text{for } x \in \mathbf{W}^m,$$

and on $\mathbf{W}^m$, define the following σ-algebras:

$$\mathcal{B}_t^X = \sigma(x_s : s \leq t); \quad \mathcal{B}_{\hat{t}}^X = \sigma(x_u - x_t : t \leq u \leq T);$$

and $\mathcal{B}^X = \vee_{t\in[0,T]} \mathcal{B}_t^X$. Obviously we have the relationship $\mathcal{B}^X = \mathcal{B}_t^X \vee \mathcal{B}_{\hat{t}}^X$ for any $t \in [0, T]$.

Now we can interpret weak solutions to BSDE (4.2) as the canonical setup for stochastic processes. If the triple $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ and (Y, Z, W) is a weak solution to BSDE (4.2), we define the image measure of $\mathbf{P}$ under the mapping $(Y, Z, W) : \Omega \to \mathbf{W}^{1+d} \times \mathbf{W}^d$ defined by

$$\bar{\omega} \mapsto (Y(\bar{\omega}), Z(\bar{\omega}), W(\bar{\omega})), \quad \text{for } \bar{\omega} \in \Omega.$$

In other words, $\mathbf{P}_{(Y,Z,W)} = \mathbf{P} \circ (Y, Z, W)^{-1}$. We let

$$(\mathbf{E}^T, \mathcal{E}^T) = \left(\mathbf{W}^{1+d} \times \mathbf{W}^d, \overline{\mathcal{B}^Y \otimes \mathcal{B}^Z \otimes \mathcal{B}^W}^{\mathbf{P}_{(Y,Z,W)}} \right),$$

where $\overline{\mathcal{B}^Y \otimes \mathcal{B}^Z \otimes \mathcal{B}^W}^{\mathbf{P}_{(Y,Z,W)}}$ denotes the completion of $\mathcal{B}^Y \otimes \mathcal{B}^Z \otimes \mathcal{B}^W$ under $\mathbf{P}_{(Y,Z,W)}$. On $(\mathbf{E}^T, \mathcal{E}^T, \mathbf{P}_{(Y,Z,W)})$, we further define the filtration $\{\mathcal{E}_t^T\}$ which is generated by $\sigma((y_s, z_s), \omega_s : s \leq t)$ augmented by the $\mathbf{P}_{(Y,Z,W)}$-null sets in $\mathcal{E}^T$. Here we use $((y, z), w)$ to denote a generic element in the sample path space $\mathbf{E}^T$. Then the triple $(\mathbf{E}^T, \mathcal{E}^T, \mathbf{P}_{(Y,Z,W)})$, $\{\mathcal{E}_t^T\}$ and $((y, z), w)$ is the weak solution to BSDE (4.2), and under $\mathbf{P}_{(Y,Z,W)}$,

$$y_t = \phi(w_T) + \int_t^T h(s, z_s) ds + \int_t^T \sum_{j=1}^d z_s^j f^j(s, z_s^j) ds - \int_t^T \sum_{j=1}^d z_s^j dw_s^j.$$

We can also introduce regular conditional probability, which is the main machinery we will employ in the next subsection. We define the projection $\Pi_W : \mathbf{E}^T \to \mathbf{W}^d$ by

$$\Pi_W((y, z), w) = w, \quad \text{for } ((y, z), w) \in \mathbf{E}^T.$$

Then the image measure of $\mathbf{P}_{(Y,Z,W)}$ under the projection Π_W is a Wiener measure on $(\mathbf{W}^d, \mathcal{B}^W)$, i.e. $\mathbf{P}_W = \mathbf{P}_{(Y,Z,W)} \circ \Pi_W^{-1}$. Because both $\mathbf{E}^T$ and $\mathbf{W}^d$ are continuous function spaces, they are Polish under the uniform topology. By Theorem A.12 there exists a unique regular conditional probability for $\mathcal{E}^T$ given the projection Π_W, which is a function $\mathbf{Q}\{\cdot|\omega\} : \mathcal{E}^T \times \mathbf{W}^d \to [0, 1]$ such that

(1) for $\omega \in \mathbf{W}^d$, $\mathbf{Q}\{\cdot|\omega\}$ is a probability measure on $(\mathbf{E}^T, \mathcal{E}^T)$;

(2) for $A \in \mathcal{E}^T$, the map $\omega \mapsto \mathbf{Q}\{A|\omega\}$ is $\mathcal{B}^W$-measurable;

(3) for $A \in \mathcal{E}^T$ and $B \in \mathcal{B}^W$, we have

$$\mathbf{P}_{(Y,Z,W)}(A \cap B) = \int_B \mathbf{Q}\{A|\omega\} \mathbf{P}_W(d\omega);$$

and moreover, $\mathbf{Q}\{\cdot|\omega\}$ concentrates on the set $\{((y, z), w) \in \mathbf{E}^T : \Pi_W((y, z), w) = w\}$, i.e. there exists a $\mathbf{P}_W$-null set N such that

$$\mathbf{Q}\left\{((y, z), w) \in E^T : \Pi_W((y, z), w) = \omega | \omega\right\} = \mathbf{Q}\left\{\mathbf{W}^{1+d} \times \{\omega\} | \omega\right\} = 1 \quad \text{for } \omega \in \mathbf{W}^d \backslash N.$$

Based on the canonical setup for weak solutions of BSDE (4.2), we can further define *strong solutions* of BSDE (4.2) on such canonical setup. Herein we use the term *strong solutions* in order to emphasize the difference between weak solutions and the solutions we have previously considered.

Definition 4.15 *We say a function* $\Phi = (\Phi^Y, \Phi^Z) : \mathbf{W}^d \to \mathbf{W}^{1+d}$ *is a strong solution to BSDE (4.2) if*

$$\omega \mapsto \Phi(\omega), \quad \text{for } \omega \in \mathbf{W}^d$$

is $\overline{\mathcal{B}_t^W}^{\mathbf{P}_W} / \mathcal{B}_t^Y \otimes \mathcal{B}_t^Z$*-measurable, where* $\overline{\mathcal{B}_t^W}^{\mathbf{P}_W}$ *denotes the completion of* $\mathcal{B}_t^W$ *under* $\mathbf{P}_W$*, and moreover, on any given filtered probability space* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ *satisfying the usual conditions with a Brownian motion* W*, the pair defined by* $(Y_t, Z_t) = (\Phi^Y(W_t), \Phi^Z(W_t))$ *satisfies the integral equation (4.3).*

4.3.2.2 Yamada-Watanabe theorem

For a SDE, the celebrated Yamada-Watanabe theorem states that the weak existence and pathwise uniqueness of the solutions to a SDE implies the existence of a *strong solution*, which is regarded as one of the most powerful tools to prove a SDE possesses a *strong solution*. As Kurtz [46] pointed out: *strong solution is a consequence of measurable selection, and such result has little to do with the equation, but really a consequence of the convexity of collections of the probability distributions of solutions.* If the *compatibility constraint* (See (3) in Definition 4.3) is satisfied, we will further have the adapteness of solutions.

We will adapt the Yamada-Watanabe theorem for our setting in this subsection. First we give two notions of uniqueness for BSDE (4.2).

Definition 4.16 *(i) The weak solution to (4.2) is called unique in law if for any two weak solutions* $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$, (Y, Z, W) *and* $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$, $\{\bar{\mathcal{F}}_t\}$, $(\bar{Y}, \bar{Z}, \bar{W})$, *the probability distributions of* (Y, Z) *and* $(\bar{Y}, \bar{Z})$ *are equal. i.e.* $\mathbf{P}_{(Y,Z)} = \bar{\mathbf{P}}_{(\bar{Y},\bar{Z})}$.

(ii) The weak solution to (4.2) is called pathwise unique if for any two weak solutions (Y, Z) *and* $(\bar{Y}, \bar{Z})$ *defined on the same probability space* $(\Omega, \mathcal{F}, \mathbf{P})$ *with the filtration* $\{\mathcal{F}_t\}$ *and the same Brownian motion* W, (Y, Z) *is a continuous modification of* $(\bar{Y}, \bar{Z})$, *i.e.*

$$(Y_t, Z_t) = (\bar{Y}_t, \bar{Z}_t), \quad \text{for } t \in [0, T], \ a.s..$$

Then main result of this subsection is the following version of Yamada-Watanabe theorem:

Theorem 4.17 *If the coefficients satisfy the One-dimensional Condition 4.4, then the weak solution to (4.2) is also a strong solution, and such solution is pathwise unique.*

First by employing Girsanov's theorem, we have the following pathwise uniqueness of the solutions, which proves the second conclusion of the above theorem.

Lemma 4.18 *If the coefficients satisfy the One-dimensional Condition 4.4, then the weak solution to (4.2) is pathwise unique.*

Proof. Suppose (Y, Z) and $(\bar{Y}, \bar{Z})$ are two weak solutions on $(\Omega, \mathcal{F}, \mathbf{P})$ with $\{\mathcal{F}_t\}$ and Brownian motion W. By applying Itô's formula to $e^{\alpha t}(Y_t - \bar{Y}_t)^2$ for some α to be determined, we obtain

$$\begin{aligned}
& e^{\alpha t}(Y_t - \bar{Y}_t)^2 \\
={}& -2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s) d(Y_s - \bar{Y}_s) - \int_t^T e^{\alpha s} d[Y - \bar{Y}, Y - \bar{Y}]_s - \int_t^T \alpha e^{\alpha s}(Y_s - \bar{Y}_s)^2 ds \\
={}& 2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s) \left\{ \sum_{j=1}^d (Z_s^j f^j(s, Z_s^j) - \bar{Z}_s^j f^j(s, \bar{Z}_s^j)) + (h(s, Z_s) - h(s, \bar{Z}_s)) \right\} ds \\
& - 2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s) \sum_{j=1}^d (Z_s^j - \bar{Z}_s^j) dW_s^j - \int_t^T e^{\alpha s} \sum_{j=1}^d |Z_s^j - \bar{Z}_s^j|^2 ds - \int_t^T \alpha e^{\alpha s}(Y_s - \bar{Y}_s)^2 ds.
\end{aligned} \tag{4.19}$$

Note that for $s \in [0,T]$, and $z^j, \bar{z}^j \in R$ for $j = 1, \cdots, d$,

$$
\begin{aligned}
& \left| z^j f^j(s, z^j) - \bar{z}^j f^j(s, \bar{z}^j) \right| \\
\leq & \left| z^j f^j(s, z^j) - \bar{z}^j f^j(s, z^j) \right| + \left| \bar{z}^j f^j(s, z^j) - \bar{z}^j f^j(s, \bar{z}^j) \right| \\
\leq & \, C_4(T + |z^j|)|z^j - \bar{z}^j| + C_4|\bar{z}^j||z^j - \bar{z}^j| \\
\leq & \, C_4(T + |z^j| + |\bar{z}^j|)|z^j - \bar{z}^j|.
\end{aligned}
$$

Now if we set

$$
\beta_s^j = \frac{Z_s^j f^j(s, Z_s^j) - \bar{Z}_s^j f^j(s, \bar{Z}_s^j)}{Z_s^j - \bar{Z}_s^j}, \quad \text{for } s \in [0,T],
$$

when $Z_s^j - \bar{Z}_s^j \neq 0$, and $\beta_s^j = 0$ for $s \in [0,T]$ otherwise, then, there exists a constant C_{15} such that $|\beta_s^j|^2 \leq C_{15}(1 + |Z_s^j|^2 + |\bar{Z}_s^j|^2)$. Using such β^j, we define a new probability measure $\mathbf{Q}$ by $\frac{d\mathbf{Q}}{d\mathbf{P}} = \mathscr{E}(N)$ where

$$
N = \sum_{j=1}^{d} \int_0^{\cdot} \beta_s^j dW_s^j,
$$

and under $\mathbf{Q}$ define a new Brownian motion B by $B = W - [W, N]$. Then under the probability measure $\mathbf{Q}$, (4.19) reduces to

$$
\begin{aligned}
& e^{\alpha t}(Y_t - \bar{Y}_t)^2 \\
= & - 2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s) \sum_{j=1}^{d} (Z_s^j - \bar{Z}_s^j)(dW_s^j - \beta_s^j ds) + 2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s)(h(s, Z_s) - h(s, \bar{Z}_s))ds \\
& - \int_t^T e^{\alpha s} \sum_{j=1}^{d} |Z_s^j - \bar{Z}_s^j|^2 ds - \int_t^T \alpha e^{\alpha s}(Y_s - \bar{Y}_s)^2 ds \\
= & - 2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s) \sum_{j=1}^{d} (Z_s^j - \bar{Z}_s^j) dB_s^j + 2\int_t^T e^{\alpha s}(Y_s - \bar{Y}_s)(h(s, Z_s) - h(s, \bar{Z}_s))ds \\
& - \int_t^T e^{\alpha s} \sum_{j=1}^{d} |Z_s^j - \bar{Z}_s^j|^2 ds - \int_t^T \alpha e^{\alpha s}(Y_s - \bar{Y}_s)^2 ds.
\end{aligned}
$$

By taking expectation under $\mathbf{Q}$ we have

$$
\begin{aligned}
E^{\mathbf{Q}}[e^{\alpha t}(Y_t - \bar{Y}_t)^2] = & \, E^{\mathbf{Q}} \left\{ \int_t^T e^{\alpha s} 2(Y_s - \bar{Y}_s)(h(s, Z_s) - h(s, \bar{Z}_s))ds \right\} \\
& - E^{\mathbf{Q}} \left\{ \int_t^T e^{\alpha s}|Z_s - \bar{Z}_s|^2 ds \right\} - E^{\mathbf{Q}} \left\{ \int_t^T \alpha e^{\alpha s}(Y_s - \bar{Y}_s)^2 ds \right\}.
\end{aligned}
$$

By the elementary inequality $2ab \leq \lambda^2 a^2 + b^2/\lambda^2$, we have

$$
2(Y_s - \bar{Y}_s)(h(s, Z_s) - h(s, \bar{Z}_s)) \leq \lambda^2 (Y_s - \bar{Y}_s)^2 + \frac{C_4^2}{\lambda^2}|Z_s - \bar{Z}_s|^2.
$$

Choosing $\lambda^2 = \alpha$ and $\alpha = 2C_4^2$, we obtain

$$
E^{\mathbf{Q}}[e^{2C_4^2 t}(Y_t - \bar{Y}_t)^2] \leq -\frac{1}{2} E^{\mathbf{Q}} \left\{ \int_t^T e^{2C_4^2 s}|Z_s - \bar{Z}_s|^2 ds \right\} \leq 0.
$$

Therefore, $Y_t = \bar{Y}_t$ for $t \in [0,T]$, *a.s.*, and $Z_t = \bar{Z}_t$ for *a.e.* $t \in [0,T]$, *a.s.*. Now the only step left is to verify $\mathscr{E}(N)$ is a uniformly-integrable martingale. By theorem A.5, we only need to verify N is a BMO-martingale under $\mathbf{P}$. Indeed, for any stopping time $\tau \le T$,

$$
\begin{aligned}
E\left\{|N_T - N_\tau|^2 | \mathcal{F}_\tau\right\} &= \sum_{j=1}^{d} E\left\{\int_\tau^T |\beta_s^j|^2 ds | \mathcal{F}_\tau\right\} \\
&\le C_{15} E\left\{\int_\tau^T (d + |Z_s|^2 + |\bar{Z}_s|^2) ds | \mathcal{F}_\tau\right\} \quad a.s..
\end{aligned}
$$

The way to control the integral term involving Z and $\bar{Z}$ has already been presented in the proof of Lemma 4.13. Therefore $\mathbf{Q}$ defined above is indeed an equivalent probability measure. ■

Now we turn to the proof of Theorem 4.17. The basic idea has already been presented in Section 4.3.2.1: to transfer the structure of weak solutions such that $\mathbf{W}^{1+d}$ becomes the sample path space for (Y, Z) and $\mathbf{W}^d$ that for W. What allows us to carry it through is regular conditional probability.

Proof of Theorem 4.17 We first prove pathwise uniqueness implies uniqueness in law. Let $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$, (Y, Z, W) and $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$, $\{\bar{\mathcal{F}}_t\}$, $(\bar{Y}, \bar{Z}, \bar{W})$ be two weak solutions. Let $\mathbf{W}^{1+d}$ and $\bar{\mathbf{W}}^{1+d}$ be two copies of $C([0,T]; R) \times C([0,T]; R^d)$. By using the regular conditional probability $\mathbf{Q}\{\cdot|\omega\}$ and $\bar{\mathbf{Q}}\{\cdot|\omega\}$, we define a probability measure π on the probability space $(\Theta, \mathcal{B}(\Theta))$ by

$$
\pi\left((dy, dz), (d\bar{y}, d\bar{z}), d\omega\right) = \mathbf{Q}\{(dy, dz)|\omega\}\bar{\mathbf{Q}}\{(d\bar{y}, d\bar{z})|\omega\}\mathbf{P}_W(d\omega)
$$

where

$$
(\Theta, \mathcal{B}(\Theta)) = \left(\mathbf{W}^{1+d} \times \bar{\mathbf{W}}^{1+d} \times \mathbf{W}^d, \overline{\mathcal{B}^Y \otimes \mathcal{B}^Z \otimes \mathcal{B}^{\bar{Y}} \otimes \mathcal{B}^{\bar{Z}} \otimes \mathcal{B}^W}^\pi\right)
$$

with $\overline{\mathcal{B}^Y \otimes \mathcal{B}^Z \otimes \mathcal{B}^{\bar{Y}} \otimes \mathcal{B}^{\bar{Z}} \otimes \mathcal{B}^W}^\pi$ denoting the completion of $\mathcal{B}^Y \otimes \mathcal{B}^Z \otimes \mathcal{B}^{\bar{Y}} \otimes \mathcal{B}^{\bar{Z}} \otimes \mathcal{B}^W$ under π. On $(\Theta, \mathcal{B}(\Theta), \pi)$, we further define the filtration $\{\mathcal{G}_t\}$ which is generated by $\sigma((y_s, z_s), (\bar{y}_s, \bar{z}_s), \omega_s : s \le t)$ augmented by the π-null sets in $\mathcal{B}(\Theta)$. Then under π and $\{\mathcal{G}_t\}$, ω is still a Wiener process. In fact by the *compatibility constraint* in Definition 4.3, we know $\mathcal{B}_t^Y \otimes \mathcal{B}_t^Z$ is independent of $\mathcal{B}_t^W$. It is obvious $\mathcal{B}_t^W$ is independent of $\mathcal{B}_t^W$. Hence for $A_t \in \mathcal{B}_t^Y \otimes \mathcal{B}_t^Z$,

$$
\mathbf{Q}\{A_t|\omega\} = \mathbf{Q}\{A_t|\omega_t\}.
$$

Likewise we also have $\bar{\mathbf{Q}}\{\bar{A}_t|\omega\} = \bar{\mathbf{Q}}\{\bar{A}_t|\omega_t\}$ for $\bar{A}_t \in \mathcal{B}_t^{\bar{Y}} \otimes \mathcal{B}_t^{\bar{Z}}$. Based on the above relationship, for $B_t \in \mathcal{B}_t^W$, $u \in [t,T]$ and $\xi \in R^d$, we obtain

$$
\begin{aligned}
E^\pi\left[e^{i\langle \xi, \omega_u - \omega_t\rangle_d} 1_{A_t} 1_{\bar{A}_t} 1_{B_t}\right] &= \int_{B_t} e^{i\langle \xi, \omega_u - \omega_t\rangle_d} \mathbf{Q}\{A_t|\omega\}\bar{\mathbf{Q}}\{\bar{A}_t|\omega\}\mathbf{P}_W(d\omega) \\
&= \int_{B_t} e^{i\langle \xi, \omega_u - \omega_t\rangle_d} \mathbf{Q}\{A_t|\omega_t\}\bar{\mathbf{Q}}\{\bar{A}_t|\omega_t\}\mathbf{P}_W(d\omega) \\
&= \int_{\mathbf{W}^d} e^{i\langle \xi, \omega_u - \omega_t\rangle_d} \mathbf{P}_W(d\omega) \int_{B_t} \mathbf{Q}\{A_t|\omega_t\}\bar{\mathbf{Q}}\{\bar{A}_t|\omega_t\}\mathbf{P}_W(d\omega) \\
&= E^\pi\left[e^{i\langle \xi, \omega_u - \omega_t\rangle_d}\right] \int_{B_t} \mathbf{Q}\{A_t|\omega\}\bar{\mathbf{Q}}\{\bar{A}_t|\omega\}\mathbf{P}_W(d\omega) \\
&= E^\pi\left[e^{i\langle \xi, \omega_u - \omega_t\rangle_d}\right] \pi\left(A_t \cap \bar{A}_t \cap B_t\right),
\end{aligned}
$$

and by the standard Dynkin arguments (see [42] and [71]), we deduce $\{\omega_u - \omega_t : t \le u \le T\}$ is independent of the σ-algebra $\mathcal{G}_t$.

Therefore $((y,z),\omega)$ and $((\bar{y},\bar{z}),\omega)$ are two weak solutions on the same filtered probability space $(\Theta, \mathcal{B}(\Theta), \{\mathcal{G}_t\}, \pi)$. Pathwise uniqueness means

$$\pi\left(\{(y,z),(\bar{y},\bar{z}),\omega\} \in \Theta : (y,z) = (\bar{y},\bar{z})\right) = 1, \tag{4.20}$$

so for any $A \in \mathcal{B}^Y \otimes \mathcal{B}^Z$, the probability distribution $\mathbf{P}\{\bar{\omega} \in \Omega : (Y(\bar{\omega}), Z(\bar{\omega})) \in A\}$ equals

$$\pi\left(\{(y,z),(\bar{y},\bar{z}),\omega\} \in \Theta : (y,z) \in A\right) = \pi\left(\{(y,z),(\bar{y},\bar{z}),\omega\} \in \Theta : (\bar{y},\bar{z}) \in A\right)$$

which is equal to the probability distribution $\bar{\mathbf{P}}\{\bar{\omega} \in \bar{\Omega} : (\bar{Y}(\bar{\omega}), \bar{Z}(\bar{\omega})) \in A\}$.

Now we prove the weak solution we constructed is in fact a strong solution. We, firstly, show $\mathbf{Q}\{\cdot|\omega\}$ and $\bar{\mathbf{Q}}\{\cdot|\omega\}$ assign full measure to the same singleton. By (4.20) and the definition of π, if we define a set $D = \{((y,z),(y,z)) : (y,z) \in \mathbf{W}^{1+d}\}$, then

$$\int_{\mathbf{W}^d} \int_{\mathbf{W}^{1+d}} \int_{\mathbf{W}^{1+d}} 1_D((y,z),(\bar{y},\bar{z}))\mathbf{Q}\{(dy,dz)|\omega\}\bar{\mathbf{Q}}\{(d\bar{y},d\bar{z})|\omega\}\mathbf{P}_W(d\omega) = 1.$$

Therefore, there exists a $\mathbf{P}_W$-null set $N \in \mathcal{B}^W$ such that

$$\int_{\mathbf{W}^{1+d}} \int_{\mathbf{W}^{1+d}} 1_D((y,z),(\bar{y},\bar{z}))\mathbf{Q}\{(dy,dz)|\omega\}\bar{\mathbf{Q}}\{(d\bar{y},d\bar{z})|\omega\} = 1, \quad \text{for } \omega \in \mathbf{W}^d \backslash N.$$

But this can only occur if there exists a $\overline{\mathcal{B}^W}^{\mathbf{P}_W}/\mathcal{B}^Y \otimes \mathcal{B}^Z$-measurable map $\Phi = (\Phi^Y, \Phi^Z) : \mathbf{W}^d \to \mathbf{W}^{1+d}$ such that

$$\mathbf{Q}\{(dy,dz)|\omega\} = \bar{\mathbf{Q}}\{(dy,dz)|\omega\} = \delta_{\Phi(\omega)}((dy,dz)), \quad \text{for } \omega \in \mathbf{W}^d \backslash N.$$

It then follows that $(y,z) = (\Phi^Y(\omega), \Phi^Z(\omega))$ for $\omega \in \mathbf{W}^d \backslash N$. Recalling that the mapping $\omega \mapsto \mathbf{Q}\{A_t|\omega\}$ is $\mathcal{B}_t^W$-measurable for $A_t \in \mathcal{B}_t^Y \otimes \mathcal{B}_t^Z$, then by the standard Dynkin arguments used earlier, Φ is in fact also $\overline{\mathcal{B}_t^W}^{\mathbf{P}_W}/\mathcal{B}_t^Y \otimes \mathcal{B}_t^Z$-measurable. On any given filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ satisfying the usual conditions with W being a Brownian motion on it, we also have

$$\begin{aligned}\Phi^Y(W_t) = \phi(W_T) &+ \int_t^T h(s, \Phi^Z(W_s))ds \\ &+ \int_t^T \sum_{j=1}^d \Phi^{Z,j}(W_s) f^j(s, \Phi^{Z,j}(W_s))ds - \int_t^T \sum_{j=1}^d \Phi^{Z,j}(W_s)dW_s^j.\end{aligned}$$

Hence Φ is a strong solution. ■

Chapter 5

QBSDEs and Credit Risk Modeling

5.1 Introduction

QBSDEs have found numerous applications in mathematical finance. For example, they appear naturally when one wants to derive the value function for exponential utility maximization, to use the idea of indifference pricing to hedge contingent claims written on non-tradeable assets, and to consider the risk measure. We refer to the following papers and the reference therein for the examples of QBSDEs in fiance: [6] [31] [36] [38] [60] and [73].

In this chapter we will further demonstrate the weak solution method to QBSDEs by considering an example coming from credit risk modeling. Specifically, we will modify the Merton's structural model for credit risk by indifference pricing. Then QBSDEs appear naturally when we want to characterize the indifference price and the associated hedging strategy. We emphasize the Cole-Hopf transformation does not help us deduce the closed form solutions in our case, though by Theorem 4.2 in the last chapter we know the solutions to our QBSDE must exist. The advantage of our weak solution method is that it allows us to really work out the solutions of QBSDEs rather than proving the existence only.

Since indifference pricing plays a central role in this chapter, we start with a short review of the indifference pricing literature. Indifference pricing is a pricing mechanism especially usefully when the financial market is incomplete. It originates from optimal portfolio problems in incomplete markets. Instead of choosing an equivalent martingale measure for the valuation of a derivative, the mechanism is to find the price at which the buyer (or writer) of such derivative is indifferent in terms of the maximum expected utilities of holding and of not holding that derivative. The approach was initiated by Hodges and Neuberger [37] and further developed by Davis [21]. In [77] Zariphopoulou made a crucial observation of the structure of the equations derived from optimal portfolio problems in incomplete markets, and introduced the Cole-Hopf transformation to solve such kind of problems. This is a crucial step to the development of indifference pricing. For example, by applying the Cole-Hopf transformation, Henderson [35] provided the solutions for power utilities and exponential utilities respectively, in which the pricing formula for exponential utility was derived and also expansions for the price under power utility. On the other hand, Musiela and Zariphopoulou [61] also solved the indifference price explicitly for the exponential utility. For a general overview of indifference pricing, we refer to the monograph [15].

5.1.1 Main results of this chapter

The main aim of this chapter is to demonstrate how the weak solution method introduced in the last chapter can be applied to mathematical finance. The discussion is somehow intuitive and formal, so none of this chapter's results are formulated as theorems.

In Section 5.2, we propose the weak formulation of optimal portfolio problems, which is in the same spirit of the weak formulation for stochastic control problems (See Yong and Zhou [76]). Roughly speaking, instead of solving optimal portfolio problems on *any given* probability space, we will choose a convenient probability space that we will work on. The reason why this might work is that our objective of optimal portfolio problems is to maximize the expectation of a certain random variable (the terminal wealth) that depends only on the probability distribution of the processes involved. Therefore if the state equations (the wealth equations) on different probability spaces have the same probability distribution, then we have more freedom in choosing a convenient probability space to work with. Under the weak formulation, we can use weak solutions of QBSDEs to characterize the value function and the optimal strategies of optimal portfolio problems.

In Section 5.3, we consider a specific financial problem in the setting of credit risk modeling. The classical Merton's structural model implicitly assumes the company's assets, as the underlying assets, are tradeable in the market. However in reality this can hardly be verified, as the company's asset value is only a concept on the company's balance sheet. We modify the classical Merton's model by indifference pricing. The indifference price and the hedging strategy of the corporate bond subject to credit risk are characterized by the weak solution of an associated QBSDE. We further apply our weak solution method to really work out the solution of such QBSDE.

This chapter is mainly a continuation of the author's transfer thesis [47]. Some of the results are also taken from the author's paper [48] with Jiang and [36] with Henderson. We should also mention that our work are influenced by Bielecki and Jeanblanc [6], though they considered the credit risk in a reduced-form setting.

5.2 Weak formulation of optimal portfolio problems

In this section we consider the weak formulation of optimal portfolio problems, which will be used later to model the credit risk of corporate bonds.

5.2.1 Basic assumptions and formulation

We first give the following assumptions for the optimal portfolio problem we want to consider.

Assumption 5.1 (*model uncertainty*)

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space which is to be determined, and $\{\mathcal{F}_t\}$ be its associated filtration satisfying the *usual conditions*, which is also to be determined.

Assumption 5.2 (*the market*)

The market is built with three assets: a risk-free bond with zero interest rate, a stock and a corporate bond. The latter of the two are issued by a company. The stock price satisfies the following SDE on the above *given* probability space:

$$\begin{cases} dS_t/S_t = \mu_t^S dt + \sigma_t^S dW_t^S + \bar{\sigma}_t^S d\bar{W}_t, \\ S_0 = s, \end{cases} \tag{5.1}$$

while the price of the company's asset value follows

$$\begin{cases} dV_t/V_t = \mu_t^V dt + \sigma_t^V dW_t^V + \bar{\sigma}_t^V d\bar{W}_t, \\ V_0 = v \end{cases} \tag{5.2}$$

on the same *given* probability space. The process $\mathbf{W} = (W^S, W^V, \bar{W})$ is a three-dimensional Brownian motion to be determined. (Note that here $\mathbf{W}$ is not the continuous function space as in Chapter 4.) The market coefficients, $\mu^i, \sigma^i, \bar{\sigma}^i$ for $i = S, V$, are bounded and deterministic functions. The stock can be traded in the market, while the company's assets are non-tradeable, since the company's asset value is only a concept on the company's balance sheet.

Assumption 5.3 (*the contract of the corporate bond*)

The corporate bond's maturity is $T > 0$. At time T, the bondholder (or called investor) will get the face value F if no default happens up to T. If the company is unable to honor its liability, the investor will take over the company and will get V_T. Therefore the payoff of this contract is $\min(V_T, F)$. We note that the payoff function is Lipschitz continuous and bounded.

Assumption 5.4 (*the investor*)

The investor has an exponential utility function depending on his/her terminal wealth, which has the form:

$$U(x) = -e^{-\gamma x}, \quad \text{for } x \in R,$$

where $\gamma \geq 0$ representing the degree of the investor's risk aversion.

Assumption 5.5 (*the trading strategy*)

The investor, with initial wealth x, invests in both the stock and the risk-free bond during the time period $[0, T]$. Let π be the amount of money invested in the stock. We assume π is taken from the following admissible set, which of course depends on the above *given* probability space.

$$\mathcal{A}_{ad} := \{\pi : [0,T] \times \Omega \to R : \ \pi \text{ is } \mathcal{F}_t\text{-adapted, self-financing and } ||\pi||_{H^2[0,T]} < \infty.\}$$

We call π an admissible trading strategy if $\pi \in \mathcal{A}_{ad}$. The adapteness of π means the investor cannot anticipate the future. π is self-financing means during the time period $[0,T]$, there is no extra money flow out of and into his/her portfolio, and gains and losses are only obtained by trading the stock. π is H^2-finite in order to guarantee that there is no arbitrage. The dynamics of the investor's wealth process, denoted by $X^x(\pi)$, follows

$$\begin{cases} dX_t^x(\pi) = \pi_t(\mu_t^S dt + \sigma_t^S dW_t^S + \bar{\sigma}_t^S d\bar{W}_t), \\ X_0^x(\pi) = x. \end{cases} \tag{5.3}$$

Assumption 5.6 (*the cost functional*)

The investor purchases the corporate bond with price C at time $t = 0$, and invests the rest of the wealth $x - C$ in the stock and the risk-free bond. During the investment horizon $[0,T]$, he/she continuously trades the stock and the risk-free bond by the admissible trading strategy π. At time $t = T$, he/she receives the total amount $X_T^{x-C}(\pi) + \min(V_T, F)$. The investor chooses the optimal trading strategy to maximize the following cost functional:

$$\sup_{\pi \in \mathcal{A}_{ad}} E^{\mathbf{P}} \left[-e^{-\gamma(X_T^{x-C}(\pi) + \min(V_T,F))} \right].$$

Here we use the superscript $\mathbf{P}$ to emphasize the expectation is taken under the probability measure $\mathbf{P}$, which is to be determined.

Now we can give the definition of weak admissible trading strategy and the corresponding weak formulation of optimal portfolio problem.

Definition 5.1 *A triple* $(\Omega, \mathcal{F}, \mathbf{P})$ $\{\mathcal{F}_t\}$ *and* $(\pi, \mathbf{W})$ *is called a weak admissible trading strategy if*
(1) $(\Omega, \mathcal{F}, \mathbf{P})$ *is a complete probability space with the filtration* $\{\mathcal{F}_t\}$ *satisfying the usual conditions;*
(2) $\mathbf{W}$ *is a Brownian motion, and the increment* $\{\mathbf{W}_u - \mathbf{W}_t : t \le u \le T\}$ *must be independent of* σ*-algebra* $\mathcal{F}_t$*;*
(3) π *is taken from the admissible set* $\mathcal{A}_{ad}$.

The set of all weak admissible trading strategies is denoted as $\mathcal{A}_{ad}^W$*, and a generic element in such weak admissible set* $\mathcal{A}_{ad}^W$ *is denoted as* Π*. The investor decides the optimal weak admissible trading strategy* Π *in order to maximize his/her cost functional:*

$$\sup_{\Pi \in \mathcal{A}_{ad}^W} E^{\mathbf{P}} \left[-e^{-\gamma(X_T^{x-C}(\pi) + \min(V_T,F))} \right]. \tag{5.4}$$

If the above optimal portfolio problem (5.4) can be solved under any given filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ *with the Brownian motion* $\mathbf{W}$*, then we will not emphasize whether the formulation is* *weak or not, and (5.4) is simply written as: (without emphasizing the superscript* $\mathbf{P}$*)*

$$\sup_{\pi \in \mathcal{A}_{ad}} E\left[-e^{-\gamma(X_T^{x-C}(\pi)+\min(V_T,F))}\right].$$

Under the weak formulation, the solutions to SDE (5.1) and (5.2) can certainly be explained as the weak solutions in the usual sense. However we need to explain (5.3) more carefully. The following explanation is taken from Yong and Zhou [76]. Because the coefficients of (5.3) depend on $\omega \in \Omega$ explicitly through the admissible trading strategy π, the weak solution to (5.3) does not make sense. On the other hand, (5.3) is not defined on *any given* probability space, so the *strong solution* cannot be defined as well. A solution to (5.3) is explained as follows: for a *given* filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ with the pair $(\pi, \mathbf{W})$ taken from the weak admissible set $\mathcal{A}_{ad}^W$, $X^{x-C}(\pi)$ satisfies the following integral equation:

$$X_t^{x-C}(\pi) = (x-C) + \int_0^t \pi_s \mu_s^S ds + \int_0^t \pi_s \left(\sigma_s^S dW_s^S + \bar{\sigma}_s^S d\bar{W}_s\right).$$

The motivation of introducing the weak formulation of optimal portfolio problems is more from mathematics rather than finance. Later we will employ martingale optimality principle to deduce an associate QBSDE as the characterization of the optimal portfolio, and we will further apply the weak solution method introduced in the last chapter to solve such QBSDE. So weak formulation serves as an auxiliary but effective mathematical model aiming at ultimately solving optimal portfolio problems. The reason why this might work is that the objective is to maximize the expectation of a certain random variable that depends only on the probability distribution of the processes involved. Therefore if the solutions to (5.3) on different probability spaces have the same probability distribution, then we have more freedom in choosing a convenient probability space to work with, and the probability space will be chosen from the weak solution of the associated QBSDE.

5.2.2 Characterization by martingale optimality principle

In order to find the value function for the optimal portfolio problem (5.4) and the optimal weak admissible trading strategy Π, we apply the martingale optimality principle. The basic idea behind is to find a stochastic process which is a martingale for the optimal control, but a supermartingale for any other control. To realize this, we resort to an auxiliary process $(Y_t)_{t\in[0,T]}$, which will be characterized by the weak solution of an associated QBSDE.

For a *given* filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ with a three-dimensional Brownian motion $\mathbf{W} = (W^S, W^V, \bar{W})$, all of which are to be determined, we want to construct a family of stochastic processes $(-e^{-\gamma(X_t^{x-C}(\pi)+Y_t)})_{t\in[0,T]}$, indexed by $\pi \in \mathcal{A}_{ad}$, such that

(1) the process $(-e^{-\gamma(X_t^{x-C}(\pi)+Y_t)})_{t\in[0,T]}$ is a supermartingale for any $\pi \in \mathcal{A}_{ad}$, and there exists an optimal $\pi^* \in \mathcal{A}_{ad}$ such that $(-e^{-\gamma(X_t^{x-C}(\pi^*)+Y_t)})_{t\in[0,T]}$ is a martingale;
(2) the auxiliary process $(Y_t)_{t\in[0,T]}$ has the terminal value $Y_T = \min(V_T, F)$.

If such auxiliary process $(Y_t)_{t\in[0,T]}$ and the optimal π^* exist, then we have

$$E^{\mathbf{P}}\left[-e^{-\gamma(X_T^{x-C}(\pi)+Y_T)}\right] \le -e^{-\gamma(x-C+Y_0)}, \quad \text{for any } \pi \in \mathcal{A}_{ad},$$

and

$$E^{\mathbf{P}}\left[-e^{-\gamma(X_T^{x-C}(\pi^*)+Y_T)}\right] = -e^{-\gamma(x-C+Y_0)}, \quad \text{for optimal } \pi^* \in \mathcal{A}_{ad}.$$

Therefore

$$\begin{aligned}\sup_{\pi\in\mathcal{A}_{ad}} E^{\mathbf{P}}\left[-e^{-\gamma(X_T^{x-C}(\pi)+Y_T)}\right] &= E^{\mathbf{P}}\left[-e^{-\gamma(X_T^{x-C}(\pi^*)+Y_T)}\right]\\ &= -e^{-\gamma(x-C+Y_0)}.\end{aligned}$$

Note that the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ and the Brownian motion $\mathbf{W}$ are still to be determined. Next we use the weak solution of an associated QBSDE to characterize the auxiliary processes Y and π^*, which also provides us with the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ and the Brownian motion $\mathbf{W}$.

Proposition 5.2 *Let $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ and $(Y, \mathbf{Z}, \mathbf{W})$ (with $\mathbf{Z} = (Z^V, \bar{Z})$) be the weak solution to the following QBSDE:*

$$Y_t = \min(V_T, F) - \int_t^T f_s ds - \int_t^T (Z_s^V dW_s^V + \bar{Z}_s d\bar{W}_s) \tag{5.5}$$

with

$$f_t = \frac{\gamma}{2}(Z_t^V)^2 + \frac{\gamma}{2}\frac{(\sigma_t^S)^2}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2}\left(\bar{Z}_t + \frac{\mu_t^S \bar{\sigma}_t^S}{\gamma(\sigma_t^S)^2}\right)^2 - \frac{(\mu_t^S)^2}{2\gamma(\sigma_t^S)^2}.$$

Then the value function of the optimal portfolio problem (5.4) is given by

$$\sup_{\Pi\in\mathcal{A}_{ad}^W} E^{\mathbf{P}}\left[-e^{-\gamma(X_T^{x-C}(\pi)+\min(V_T,F))}\right] = -e^{-\gamma(x-C+Y_0)},$$

and the optimal weak admissible trading strategy Π^* is the triple $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ and $(\pi^*, \mathbf{W})$ with

$$\pi_t^* = -\frac{\bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2}\bar{Z}_t + \frac{\mu_t^S}{\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]}. \tag{5.6}$$

Proof. On a *given* filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ with the Brownian motion $\mathbf{W}$, we suppose $(Y_t)_{t\in[0,T]}$ satisfies the following BSDE

$$Y_t = \min(V_T, F) - \int_t^T f_s ds - \int_t^T (Z_s^V dW_s^V + \bar{Z}_s d\bar{W}_s),$$

where the driver f is to be determined. By applying Itô's formula to $e^{-\gamma(X_t^{x-C}(\pi)+Y_t)}$, we obtain

$$\begin{aligned}&de^{-\gamma(X_t^{x-C}(\pi)+Y_t)}\\ &= e^{-\gamma(X_t^{x-C}(\pi)+Y_t)}\left\{-\gamma(dX_t^{x-C}(\pi) + dY_t) + \frac{\gamma^2}{2}d[X^{x-C}(\pi) + Y, X^{x-C}(\pi) + Y]_t\right\}\\ &= e^{-\gamma(X_t^{x-C}(\pi)+Y_t)}\left\{-\gamma\mu_t^S\pi_t - \gamma f_t + \frac{\gamma^2}{2}\left[(\sigma_t^S)^2\pi_t^2 + (\bar{\sigma}_t^S)^2\pi_t^2 + (Z_t^V)^2 + (\bar{Z}_t)^2 + 2\bar{\sigma}_t^S\bar{Z}_t\pi_t\right]\right\}dt\\ &\quad + \text{ martingale term.}\end{aligned}$$

Since $(-e^{-\gamma(X_t^{x-C}(\pi)+Y_t)})_{t\in[0,T]}$ is a supermartingale for any $\pi \in \mathcal{A}_{ad}$, and a martingale for optimal $\pi^* \in \mathcal{A}_{ad}$, we must have

$$-\gamma\mu_t^S\pi_t - \gamma f_t + \frac{\gamma^2}{2}\left[(\sigma_t^S)^2\pi_t^2 + (\bar{\sigma}_t^S)^2\pi_t^2 + (Z_t^V)^2 + (\bar{Z}_t)^2 + 2\bar{\sigma}_t^S\bar{Z}_t\pi_t\right] \geq 0, \quad \text{for any } \pi \in \mathcal{A}_{ad},$$

and

$$-\gamma\mu_t^S\pi_t^* - \gamma f_t + \frac{\gamma^2}{2}\left[(\sigma_t^S)^2(\pi_t^*)^2 + (\bar{\sigma}_t^S)^2(\pi_t^*)^2 + (Z_t^V)^2 + (\bar{Z}_t)^2 + 2\bar{\sigma}_t^S\bar{Z}_t\pi_t^*\right] = 0.$$

By rewriting the above inequalities, we need to find f_t and π^* such that

$$\begin{aligned}&\frac{\gamma^2}{2}[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]\left(\pi_t + \frac{\bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2}\bar{Z}_t - \frac{\mu_t^S}{\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]}\right)^2 \\ &+ \frac{\gamma^2}{2}(Z_t^V)^2 + \frac{\gamma^2}{2}\frac{(\sigma_t^S)^2}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2}\left(\bar{Z}_t + \frac{\mu_t^S\bar{\sigma}_t^S}{\gamma(\sigma_t^S)^2}\right)^2 - \frac{(\mu_t^S)^2}{2(\sigma_t^S)^2} - \gamma f_t \geq 0, \quad \text{for any } \pi \in \mathcal{A}_{ad},\end{aligned}$$

and equality holds for optimal π^*. Therefore, we must have

$$f_t = \frac{\gamma}{2}(Z_t^V)^2 + \frac{\gamma}{2}\frac{(\sigma_t^S)^2}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2}\left(\bar{Z}_t + \frac{\mu_t^S\bar{\sigma}_t^S}{\gamma(\sigma_t^S)^2}\right)^2 - \frac{(\mu_t^S)^2}{2\gamma(\sigma_t^S)^2},$$

and

$$\pi_t^* = -\frac{\bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2}\bar{Z}_t + \frac{\mu_t^S}{\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]}.$$

■

To finish this subsection, we use the martingale optimality principle to characterize the value function for the above optimal portfolio problem (5.4) but without the extra payoff $\min(V_T, F)$. The following is nothing new, but we present it here in order to compare the difference of constructing the auxiliary process Y with and without the extra payoff $\min(V_T, F)$.

We are still under the Assumptions 5.1-5.5. However the cost functional in Assumption 5.6 is replaced by the following

Assumption 5.6′ (*the cost functional*)

The investor does not purchase the corporate bond during the time period $[0, T]$. At time $t = 0$, the investor invests in both the stock and the risk-free bond, and during time period $[0, T]$ he/she continuously trades them based on the admissible trading strategy π. At time $t = T$, he/she gets the total amount $X_T^x(\pi)$. He/She wants to maximize the following cost functional:

$$\sup_{\pi\in\mathcal{A}_{ad}} E^{\mathbf{P}}\left[-e^{-\gamma X_T^x(\pi)}\right] \tag{5.7}$$

on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ with the Brownian motion $\mathbf{W}$, both of which are to be determined. If we use the weak formulation as in Definition 5.1, then (5.7) can be rewritten as

$$\sup_{\Pi\in\mathcal{A}_{ad}^W} E^{\mathbf{P}}\left[-e^{-\gamma X_T^x(\pi)}\right].$$

Next we also apply the martingale optimality principle to find the value function for the optimal portfolio problem (5.7). The idea is the same as before. For a *given* filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ with the Brownian motion $\mathbf{W}$, we want to construct a family of stochastic processes $(-e^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)})_{t\in[0,T]}$, indexed by $\pi \in \mathcal{A}_{ad}$, such that

(1) the process $(-e^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)})_{t\in[0,T]}$ is a supermartingale for any $\pi \in \mathcal{A}_{ad}$, and there exists an optimal $\bar{\pi}^* \in \mathcal{A}_{ad}$ such that $(-e^{-\gamma(X_t^{x-C}(\bar{\pi}^*)+\bar{Y}_t)})_{t\in[0,T]}$ is a martingale;
(2) the process $(\bar{Y}_t)_{t\in[0,T]}$ has the terminal value $\bar{Y}_T = 0$.

If such an auxiliary process $(\bar{Y}_t)_{t\in[0,T]}$ and the optimal $\bar{\pi}^*$ exist, then we have

$$E^{\mathbf{P}}\left[-e^{-\gamma X_T^x(\pi)}\right] = E^{\mathbf{P}}\left[-e^{-\gamma(X_T^x(\pi)+\bar{Y}_T)}\right] \leq -e^{-\gamma(x+\bar{Y}_0)}, \quad \text{for any } \pi \in \mathcal{A}_{ad},$$

and

$$E^{\mathbf{P}}\left[-e^{-\gamma X_T^x(\bar{\pi}^*)}\right] = E^{\mathbf{P}}\left[-e^{-\gamma(X_T^x(\bar{\pi}^*)+\bar{Y}_T)}\right] = -e^{-\gamma(x+\bar{Y}_0)}, \quad \text{for optimal } \bar{\pi}^* \in \mathcal{A}_{ad}.$$

Therefore,

$$\begin{aligned}\sup_{\pi\in\mathcal{A}_{ad}} E^{\mathbf{P}}\left[-e^{-\gamma X_T^x(\pi)}\right] &= E^{\mathbf{P}}\left[-e^{-\gamma X_T^x(\bar{\pi}^*)}\right] \\ &= -e^{-\gamma(x+\bar{Y}_0)}.\end{aligned}$$

To construct such an auxiliary process $\bar{Y}$, because its terminal data is deterministic $\bar{Y}_T = 0$, by Corollary 2.10 in Chapter 2, $\bar{Y}$ should be expected to satisfy an ODE. We suppose

$$d\bar{Y}_t = \bar{f}_t dt \quad \text{with} \quad \bar{Y}_T = 0,$$

where $\bar{f}_t$ is to be determined. Applying Itô's formula to $e^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)}$, we obtain

$$\begin{aligned} & de^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)} \\ &= e^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)}\left\{-\gamma(dX_t^{x-C}(\pi)+d\bar{Y}_t)+\frac{\gamma^2}{2}d[X^{x-C}(\pi)+\bar{Y}, X^{x-C}(\pi)+\bar{Y}]_t\right\} \\ &= e^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)}\left\{-\gamma\mu_t^S\pi_t - \gamma\bar{f}_t + \frac{\gamma^2}{2}\left[(\sigma_t^S)^2\pi_t^2 + (\bar{\sigma}_t^S)^2\pi_t^2\right]\right\}dt \\ &\quad + \text{ martingale term.}\end{aligned}$$

Since $(-e^{-\gamma(X_t^{x-C}(\pi)+\bar{Y}_t)})_{t\in[0,T]}$ is supermartingale for any $\pi \in \mathcal{A}_{ad}$, and a martingale for optimal $\bar{\pi}^* \in \mathcal{A}_{ad}$, in analogy to the arguments in Proposition 5.2, we must have

$$\frac{\gamma^2}{2}[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]\left(\pi_t - \frac{\mu_t^S}{\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]}\right)^2 - \frac{(\mu_t^S)^2}{2[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} - \gamma\bar{f}_t \geq 0, \quad \text{for any } \pi \in \mathcal{A}_{ad},$$

with inequality holding for optimal $\bar{\pi}^*$. Therefore,

$$\bar{f}_t = -\frac{(\mu_t^S)^2}{2\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]},$$

and

$$\bar{\pi}_t^* = \frac{\mu_t^S}{\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]}. \tag{5.8}$$

In this case, we can solve $\bar{Y}$ explicitly on *any given* probability space, so there is no need to emphasize the weak formulation by the convention in Definition 5.1. In fact

$$Y_t = \int_t^T \frac{(\mu_s^S)^2}{2\gamma[(\sigma_s^S)^2 + (\bar{\sigma}_s^S)^2]} ds,$$

and the value function for the optimal portfolio problem (5.7) is: (without emphasizing the superscript $\mathbf{P}$)

$$\sup_{\pi \in \mathcal{A}_{ad}} E\left[-e^{-\gamma X_T^x(\pi)}\right] = -\exp\left\{-\gamma\left(x + \int_0^T \frac{(\mu_s^S)^2}{2\gamma[(\sigma_s^S)^2 + (\bar{\sigma}_s^S)^2]} ds\right)\right\}. \tag{5.9}$$

5.3 A modified structural model for credit risk

In this section we propose a modified Merton's structural model based on the idea of indifference pricing. The basic machinery we will employ is the weak formulation of optimal portfolio problems and the weak solution method for QBSDEs.

5.3.1 Revisit of Merton's model

The basis of structural models, which goes back to Merton [59], is built on the premise that there exists a fundamental process V interpreted as the value of the company's assets, which is in fact a concept on the company's balance sheet. The corporate bond is regarded as a contingent claim written on the company's assets, and the default is triggered if the company is under financial distress, i.e. the company's asset value V falls below its liability value. Therefore the no-arbitrage argument might be applied to model credit risk. For a detail introduction of Merton's model and the related literature, we refer to the author's transfer thesis [47], as well as the author's paper [51] with Ren.

There is an implicit, but critical, assumption in most of structural models, namely the uniqueness of equivalent martingale measure. (See Appendix A.3 for the definition of martingale measures.) Most of the literature on credit risk focuses on the pricing issues and postulates this assumption without questioning it. However such assumption seems to make sense only if investors can trade the company's assets in the market. Since it is hard to define a meaningful process for company's asset value, let alone trade it in the market, the no-arbitrage argument is not sufficient to determine a unique price. In other words, we cannot obtain the corporate bond price uniquely, because the credit risk, which arises from the collapse of the firm's asset value, cannot be hedged by the underlying. We want to relax the assumption of the uniqueness of equivalent martingale measure, and modify Merton's model by embedding it into the optimal portfolio problem in an incomplete market framework.

5.3.2 Modifying Merton's model by indifference pricing

We apply the idea of indifference pricing to Merton's model in this subsection. Roughly speaking, the indifference price of the corporate bond is such that the investor is indifferent with and without holding such corporate bond when he/she solves the optimal portfolio problem. We follow the notations in Section 5.2.

Definition 5.3 *The indifference price of the corporate bond issued by the company, denoted by C, is such that the value functions for the optimal portfolio problems (5.4) and (5.7) coincide. That is*

$$\sup_{\Pi \in \mathcal{A}_{ad}^W} E^{\mathbf{P}} \left[-e^{-\gamma(X_T^{x-C}(\pi)+\min(V_T,F))} \right] = \sup_{\pi \in \mathcal{A}_{ad}} E^{\mathbf{P}} \left[-e^{-\gamma X_T^x(\pi)} \right],$$

where Π is taken from the weak admissible set $\mathcal{A}_{ad}^W$, i.e. the corresponding probability space also needs to be determined, and π is taken from the admissible set $\mathcal{A}_{ad}$ on the probability space which is determined from Π^*.

The hedging strategy for such corporate bond is given by $\pi^* - \bar{\pi}^*$ on the probability space determined from Π^*, where π^* is given by (5.6) and $\bar{\pi}^*$ is given by (5.8).

For the properties of such indifference price and hedging strategy and how they can be related to the classical Merton's model, we refer to the author's transfer thesis [47]. In the following we concentrate on how to solve such indifference price C and the hedging strategy $\pi^* - \bar{\pi}^*$.

By Proposition 5.2, we know the value function of the optimal portfolio problem (5.4) has the form

$$-e^{-\gamma(x-C+Y_0)},$$

and by (5.9), the value function for the optimal portfolio problem (5.7) is

$$-\exp\left\{ -\gamma \left(x + \int_0^T \frac{(\mu_s^S)^2}{2\gamma[(\sigma_s^S)^2 + (\bar{\sigma}_s^S)^2]} ds \right) \right\}.$$

So the indifference price C is given by

$$C = Y_0 - \int_0^T \frac{(\mu_s^S)^2}{2\gamma[(\sigma_s^S)^2 + (\bar{\sigma}_s^S)^2]} ds,$$

and by (5.6) and (5.8), the hedging strategy $\pi^* - \bar{\pi}^*$ is given by

$$\pi^* - \bar{\pi}^* = -\frac{\bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2} \bar{Z}_t.$$

Hence constructing the indifference price of the corporate bond and its hedging strategy reduces to find the weak solution of BSDE (5.5). Before presenting how to solve BSDE (5.5), we emphasize that the Cole-Hopf transformation does not help us deduce the closed form solutions, though by Theorem 4.2 we know the solutions to BSDE (5.5) must exist. The advantage of our weak solution method is that it allows us to really work out the solution of BSDE (5.5) rather than proving the existence only.

We use the *strong solution* of the following FBSDE (5.10) to construct the weak solution of BSDE (5.5). We start with a Brownian motion $\mathbf{B} = (B^S, B^V, \bar{B})$ on $(\Omega, \mathcal{F}, \mathbf{Q})$ with the filtration $\{\mathcal{F}_t\}$ satisfying the *usual conditions*, and consider the following FBSDE:

$$\begin{cases} d\ln V_t &= \left\{\mu_t^V - \frac{1}{2}[(\sigma_t^V)^2 + (\bar{\sigma}_t^V)^2]\right\} dt - \sigma_t^V \frac{\gamma}{2} Z_t^V dt \\ & \quad -\bar{\sigma}_t^V \left\{ \frac{\gamma(\sigma_t^S)^2}{2[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} \bar{Z}_t + \frac{\mu_t^S \bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2} \right\} dt + \sigma_t^V dB_t^V + \bar{\sigma}_t^V d\bar{B}_t, \\ \ln V_0 &= \ln v, \\ dY_t &= -\frac{(\mu_t^S)^2}{2\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} dt + Z_t^V dB_t^V + \bar{Z}_t d\bar{B}_t, \\ Y_T &= \min(e^{\ln V_T}, F). \end{cases} \tag{5.10}$$

Note that the above FBSDE (5.10) is linear, and moreover, its coefficients satisfy Lipschitz condition 2.3. Therefore, by Theorem 2.12, there exist $(Y, \mathbf{Z}) \in \mathcal{C}^\infty([0,T]; R) \times H^2([0,T]; R^2)$ with the forward process $(\ln V_t)_{t\in[0,T]} \in \mathcal{C}([0,T]; R)$.

Based on the solution $(Y, \mathbf{Z})$, we define a new probability measure $\mathbf{P}$ by

$$\frac{d\mathbf{P}}{d\mathbf{Q}} = \mathscr{E}(N),$$

where $\mathscr{E}(N)$ is the Doléans-Dade exponential of N with

$$N = \int_0^\cdot \frac{\gamma}{2} Z_t^V dB_t^V + \int_0^\cdot \left\{ \frac{\gamma(\sigma_t^S)^2}{2[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} \bar{Z}_t + \frac{\mu_t^S \bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2} \right\} d\bar{B}_t.$$

By Lemma 4.13 in Chapter 4, we know that $\mathbf{P}$ is indeed an equivalent probability measure. Under the new probability measure $\mathbf{P}$, by Girsanov's theorem, $\mathbf{W} = \mathbf{B} - [\mathbf{B}, N]$ is a Brownian motion with

$$\begin{cases} W^S = B^S, \\ W^V = B^V - \int_0^\cdot \frac{\gamma}{2} Z_t^V dt, \\ \bar{W} = \bar{B} - \int_0^\cdot \left\{ \frac{\gamma(\sigma_t^S)^2}{2[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} \bar{Z}_t + \frac{\mu_t^S \bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2} \right\} dt. \end{cases}$$

Under the probability measure $\mathbf{P}$ and with the Brownian motion $\mathbf{W}$, we rewrite the backward equation in FBSDE (5.10):

$$\begin{aligned} dY_t =& -\frac{(\mu_t^S)^2}{2\gamma[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} dt \\ &+ Z_t^V \left(dW_t^V + \frac{\gamma}{2} Z_t^V dt \right) + \bar{Z}_t \left(d\bar{W}_t + \left\{ \frac{\gamma(\sigma_t^S)^2}{2[(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2]} \bar{Z}_t + \frac{\mu_t^S \bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2} \right\} dt \right) \\ =& f_t dt + Z_t^V dW_t^V + \bar{Z}_t d\bar{W}_t \end{aligned}$$

with $Y_T = \min(V_T, F)$, and rewrite the forward equation in FBSDE (5.10):

$$dV_t/V_t = \mu_t^V\, dt + \sigma_t^V\, dW_t^V + \bar{\sigma}_t^V\, d\bar{W}_t.$$

Therefore the triple $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ and $(Y, \mathbf{Z}, \mathbf{W})$ is just one weak solution we want to find.

We conclude this section by summarizing the above results as the following proposition:

Proposition 5.4 *Let Y and $\bar{Z}$ be constructed from the weak solution of BSDE (5.5). Then, the indifference price of the corporate bond offering payoff $\min(V_T, F)$ is given by*

$$Y_0 - \int_0^T \frac{(\mu_s^S)^2}{2\gamma[(\sigma_s^S)^2 + (\bar{\sigma}_s^S)^2]} ds,$$

and the associated hedging strategy is given by

$$-\frac{\bar{\sigma}_t^S}{(\sigma_t^S)^2 + (\bar{\sigma}_t^S)^2} \bar{Z}_t.$$

Appendix A

Preliminaries of Stochastic Analysis

In this appendix we recall some stochastic analysis tools that are particular important to us and used throughout the thesis. For the general theory and its details we refer to [42] by Karatzas and Shreve and [71] by Revuz and Yor .

A.1 Special semimartingales and canonical decomposition

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ be a filtered probability space which satisfies the *usual conditions*. That is $(\Omega, \mathcal{F}, \mathbf{P})$ is a complete probability space, $\mathcal{F}_0$ contains all the $\mathbf{P}$-null sets of $\mathcal{F}$ and $\mathcal{F}_t = \mathcal{F}_{t+}$, i.e. the filtration $\{\mathcal{F}_t\}$ is right continuous. Any martingale on such $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ has a modification whose sample paths are right continuous with left limits, or simply called Càdlàg. So by a martingale, we always mean a martingale which has Càdlàg sample paths. For given $T > 0$, all of our processes are defined and considered only within the interval $[0, T]$.

Recall an $\mathcal{F}_t$-adapted Càdlàg process Y is a semimartingale if there exists processes M and V with $V_0 = 0$ such that

$$Y_t = M_t - V_t, \tag{A.1}$$

where M is a local martingale and V is a finite variation process. If V is further predictably measurable, then Y is called a special semimartingale. The typical example of a special semimartingale is a semimartingale with bounded jumps, i.e. $|\Delta Y_t| \leq \epsilon < \infty$ for $t \in [0, T]$. In particular, a continuous semimartingale is a special semimartingale with both M and V being continuous.

For the special semimartingale Y, the above decomposition (A.1) is unique, which depends on the following fact: (See Protter [69])

Proposition A.1 *Let M be a local martingale with paths of finite variation. If M is predictably measurable, then M must be a constant, i.e. $M_t = M_0$ for $t \in [0, T]$.*

Now suppose the special semimartingale Y admits another decomposition $Y_t = N_t - A_t$ with N being a local martingale and A being a predictably measurable finite variation process, then

$$M_t - N_t = V_t - A_t.$$

Therefore, $V - A$ is a predictable measurable finite variation process which is also a local martingale. Since $V_0 - A_0 = 0$, by Proposition A.1, we must have $V_t = A_t$ for $t \in [0, T]$.

Definition A.2 *If Y is a special semimartingale, then the unique decomposition $Y_t = M_t - V_t$ with $V_0 = 0$ and V being predictably measurable is called the canonical decomposition.*

A.2 Continuous BMO-martingales

In this section we only consider continuous martingales. Let M be a continuous martingale. For $p \geq 1$, define the continuous martingale space $\mathcal{H}^p([0,T]; R^d)$ equipped with the $\mathcal{H}^p$-norm:

$$||M||_{\mathcal{H}^p[0,T]} = E\left\{[M,M]_T^{p/2}\right\}^{1/p}.$$

By the Burkholder-Davis-Gundy inequality, there exist constants C_1 and C_2 such that

$$C_1||M||_{\mathcal{H}^p[0,T]} \leq E\left\{\sup_{t\in[0,T]} |M_t|^p\right\}^{1/p} \leq C_2||M||_{\mathcal{H}^p[0,T]}.$$

If furthermore $p > 1$, by Doob's inequality, $|| \cdot ||_{\mathcal{H}^p[0,T]}$ is equivalent to the following L_T^p-norm:

$$||M||_{\mathcal{H}^p[0,T]} \sim E\left[|M_T|^p\right]^{1/p}.$$

For $p > 1$, $\mathcal{H}^p([0,T]; R^d)$ is the dual space of $\mathcal{H}^q([0,T]; R^d)$, where q is the conjugate of p, i.e. $1/p + 1/q = 1$. However for $p = 1$, the dual space of $\mathcal{H}^1([0,T]; R^d)$ is strictly larger than $\mathcal{H}^\infty([0,T]; R^d)$, the class of all continuous martingales with bounded quadratic variation. In fact the dual space of $\mathcal{H}^1([0,T]; R^d)$ is $BMO([0,T]; R^d)$. Let us first recall BMO-martingales.

Definition A.3 *Let $M \in \mathcal{H}^2([0,T]; R^d)$. Then M is called a BMO_2-martingale if there exists a constant C_3 such that for any stopping time $\tau \leq T$, we have*

$$E\left\{|M_T - M_\tau|^2 | \mathcal{F}_\tau\right\} \leq C_3^2 \quad a.s.,$$

and the smallest C_3 is defined to be the BMO_2-norm. We define $BMO_2([0,T]; R^d)$ space by

$$BMO_2([0,T]; R^d) = \{M \in \mathcal{H}^2([0,T]; R^d) : ||M||_{BMO_2[0,T]} < \infty\}.$$

Analogously, we can also define the $BMO_p([0,T]; R^d)$ space for any $p \geq 1$, and all BMO_p-norms turn out to be equivalent to each other:

Proposition A.4 *There exists a constant C_4 such that*

$$||M||_{BMO_1[0,T]} \leq ||M||_{BMO_p[0,T]} \leq C_4||M||_{BMO_1[0,T]}. \tag{A.2}$$

Proof. The first inequality simply follows from Jensen's inequality, while the second inequality is deduced from the following John-Nirenberg inequality: (See He et al [34])

$$E\left\{e^{|M_T-M_\tau|}|\mathcal{F}_\tau\right\} \leq \frac{1}{1-4||M||_{BMO_1[0,T]}} \quad a.s. \tag{A.3}$$

for any stopping time $\tau \leq T$, if $||M||_{BMO_1[0,T]} \leq \frac{1}{4}$.

Now we define $\tilde{M} = \frac{M}{8||M||_{BMO_1[0,T]}}$. Since

$$||\tilde{M}||_{BMO_1[0,T]} = \frac{1}{8} < \frac{1}{4},$$

we can apply (A.3) to $\tilde{M}$ and obtain

$$E\left\{e^{|\tilde{M}_T-\tilde{M}_\tau|}|\mathcal{F}_\tau\right\} \leq \frac{1}{1-4||\tilde{M}||_{BMO_1[0,T]}} = 2 \quad a.s..$$

So for any $p \geq 1$,

$$\frac{E\{|M_T - M_\tau|^p|\mathcal{F}_\tau\}}{p!(8||M||_{BMO_1[0,T]})^p} \leq 2 \quad a.s.,$$

i.e. $E\{|M_T - M_\tau|^p|\mathcal{F}_\tau\}^{1/p} \leq 8 \cdot 2^{1/p}(p!)^{1/p}||M||_{BMO_1[0,T]}$, $a.s..$ ■

Because of Proposition A.4, we will simply write BMO-martingales without specifying p in the thesis. Next, we turn to the duality between $\mathcal{H}^1([0,T];R^d)$ and $BMO([0,T];R^d)$, which in fact follows from the following Fefferman's inequality: (See He et al [34])

$$E\{[M,N]_T\} \leq C_5||N||_{BMO[0,T]}||M||_{\mathcal{H}^1[0,T]}.$$

We finish this section by a characterization result of BMO-martingales: (See Kazamaki [43])

Theorem A.5 *M is a continuous BMO-martingale iff its Doléans-Dade exponential $\mathcal{E}(M)$ is a uniformly integrable martingale satisfying the following reverse Hölder's inequality:*

$$E[\mathcal{E}(M)_T^p|\mathcal{F}_\tau] \leq C_6\mathcal{E}(M)_\tau^p \quad a.s.$$

for some $p > 1$, where $\tau \leq T$ is an arbitrary stopping time.

A.3 Jacod-Yor's martingale representation

It is well known that the martingale (or predictable) representation is intimately related to the geometry property of the probability measure **P**, i.e. **P** is an extremal point. We will recall the Jacod-Yor's martingale representation in this section. In analogy to $\mathcal{H}^2([0,T];R^d)$, we denote $\mathcal{M}^2([0,T];R^d)$ the space of R^d-valued square integrable martingales with Càdlàg sample paths.

Definition A.6 *A closed subspace $\mathcal{A}$ of $\mathcal{M}^2([0,T];R)$ is called a stable space if (i) for any $A \in \mathcal{F}_0$ and any martingale $M \in \mathcal{A}$, then $M.1_A \in \mathcal{A}$; (ii) for any stopping time $\tau \leq T$ and any martingale $M \in \mathcal{A}$, then $M_{\cdot\wedge\tau} \in \mathcal{A}$.*

The stable space is also stable under taking Itô's integrals: (See Protter [69])

Proposition A.7 *Let $\mathcal{A}$ be a closed subspace of $\mathcal{M}^2([0,T];R)$. Then, $\mathcal{A}$ is stable iff $\mathcal{A}$ is stable under taking Itô's integrals: for any orthogonal martingale $M = (M^1, \cdots, M^d)^T$ with $M^j \in \mathcal{A}$, and Z being predictably measurable such that*

$$E\{\sum_{j=1}^{d}\int_0^T |Z_s^j|^2 d[M^j, M^j]_s\} < \infty,$$

then $\int_0^\cdot Z_s dM_s \in \mathcal{A}$. (The superscript T denotes matrix transposition.)

For $\mathcal{A} \subset \mathcal{M}^2([0,T];R)$, we denote $\mathcal{S}(\mathcal{A})$ the stable subspace generated by $\mathcal{A}$, i.e. the intersection of all closed stable subspaces containing $\mathcal{A}$. We want to identify a condition such that

$$\mathcal{S}(\mathcal{A}) = \mathcal{M}^2([0,T];R), \tag{A.4}$$

which means any martingale $M \in \mathcal{M}^2([0,T];R)$ can be represented by the elements in $\mathcal{A}$. This turns out to be closely related to the property of $\mathbf{P}$. We first define the set of martingale measures associated with $\mathcal{A}$, denoted as $\mathcal{M}^2(\mathcal{A})$. A probability measure $\mathbf{Q} \in \mathcal{M}^2(\mathcal{A})$ if (i) $\mathbf{Q}$ is absolutely continuous with respect to $\mathbf{P}$; (ii) $\mathbf{Q} = \mathbf{P}$ on $\mathcal{F}_0$; (iii) for any $M \in \mathcal{A}$, M is also a square integrable martingale under $\mathbf{Q}$.

Definition A.8 *A probability measure $\mathbf{P} \in \mathcal{M}^2(\mathcal{A})$ is called an extremal point of $\mathcal{M}^2(\mathcal{A})$ if whenever $\mathbf{P} = \lambda\mathbf{Q}^1 + (1-\lambda)\mathbf{Q}^2$ for $\mathbf{Q}^1, \mathbf{Q}^2 \in \mathcal{M}^2(\mathcal{A})$ and $\lambda \in [0,1]$, then λ must be either 0 or 1.*

The following Jacod-Yor's martingale representation provides us with a condition such that (A.4) holds:

Theorem A.9 *(Jacod and Yor [41]) Let $\mathcal{A}$ be a subset of $\mathcal{M}^2([0,T];R)$ containing constant martingales. Then (A.4) holds iff $\mathbf{P}$ is an extremal point of $\mathcal{M}^2(\mathcal{A})$.*

A familiar example is the Brownian martingale representation. To see this, let $B = (B^1, \cdots, B^d)^T$ be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbf{P})$ with the natural filtration $\{\mathcal{F}_t^B\}$ after argumentation. Since $[B^i, B^j]_t = \delta_{ij}t$ under any $\mathbf{Q} \in \mathcal{M}^2(\{B\})$, $\mathcal{M}^2(\{B\}) = \{\mathbf{P}\}$, and the probability measure $\mathbf{P}$ is the trivial extremal point of $\mathcal{M}^2(\{B\})$. Therefore $\mathcal{S}(\{B\}) = \mathcal{M}^2([0,T];R)$, i.e. any Brownian martingale $M \in \mathcal{M}^2([0,T];R)$ can be represented by the Brownian motion B.

A.4 Canonical setup and regular conditional probabilities

To finish this appendix we recall the definitions of canonical setup and regular conditional probabilities. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space, and let $(\mathbf{E}, \mathcal{E})$ be a measurable space. Define $\mathbf{E}^T$ to be the set of all functions from $[0, T]$ to $\mathbf{E}$.

Let $(X_t)_{t\in[0,T]}$ be a stochastic process with the state space $(\mathbf{E}, \mathcal{E})$ on $(\Omega, \mathcal{F}, \mathbf{P})$. It can be regarded as a random variable with values in $\mathbf{E}^T$, which is called the sample path space. Indeed, for $t \in [0, T]$, we define the coordinate mapping $X_t : \mathbf{E}^T \to \mathbf{E}$ by

$$X_t(x) = x(t), \quad \text{for } x \in \mathbf{E}^T.$$

and the σ-algebra $\mathcal{E}^T$ by $\mathcal{E}^T = \sigma\{x(t) : t \in [0, T]\}$. We define the probability distribution of $(X_t)_{t\in[0,T]}$ as the image measure of $\mathbf{P}$ under the mapping $X : \Omega \to \mathbf{E}^T$ by $\omega \mapsto X(\omega)$ for $\omega \in \Omega$. That is $\mathbf{P}_X = \mathbf{P} \circ X^{-1}$.

Definition A.10 *Let $(X_t)_{t\in[0,T]}$ be a stochastic process given as above. Then $(\mathbf{E}^T, \mathcal{E}^T, \mathbf{P}_X)$ together with the coordinate mapping X_t (or the generic element $x \in \mathbf{E}^T$) is called the canonical setup for $(X_t)_{t\in[0,T]}$.*

One of the advantages of the above canonical setup is that it allows us to compare two different stochastic processes, possibly defined on different probability spaces. Next we recall regular conditional probabilities (or disintegration of probability measures).

Definition A.11 *Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space, and let X be a measurable map from this space to a measurable space $(\mathbf{E}, \mathcal{E})$ on which it induces an image measure of $\mathbf{P}$ by $\mathbf{P}_X = \mathbf{P} \circ X^{-1}$. Then a regular conditional probability for $\mathcal{F}$ given X is a function $\mathbf{Q}\{\cdot|x\} : \mathcal{F} \times \mathbf{E} \to [0, 1]$ such that*
(1) *for $x \in \mathbf{E}$, $\mathbf{Q}\{\cdot|x\}$ is a probability measure on $(\Omega, \mathcal{F})$;*
(2) *for $A \in \mathcal{F}$, $x \mapsto \mathbf{Q}\{A|x\}$ is $\mathcal{E}$-measurable;*
(3) *for $A \in \mathcal{F}$ and $B \in \mathcal{E}$, we have*

$$\mathbf{P}(A \cap X^{-1}(B)) = \int_B \mathbf{Q}\{A|x\}\mathbf{P}_X(dx).$$

We say the above regular conditional probability is unique if $\mathbf{Q}'\{\cdot|x\}$ is another such, then there exists a $\mathbf{P}_X$-null set N such that $\mathbf{Q}\{A|x\} = \mathbf{Q}'\{A|x\}$ for $A \in \mathcal{F}$ and $x \in \mathbf{E}\backslash N$. If both Ω and $\mathbf{E}$ have certain topological structure, then the regular conditional probabilities exist and are unique. (See Parthasarathy [65])

Theorem A.12 *If both Ω and $\mathbf{E}$ are Polish spaces (i.e. separable and complete), then there exists a unique regular conditional probability for $\mathcal{F}$ given X, and moreover $\mathbf{Q}\{\cdot|x\}$ concentrates on the set $\{\omega \in \Omega : X(\omega) = x\}$. That is there exists a $\mathbf{P}_X$-null set N such that*

$$\mathbf{Q}\{\omega \in \Omega : X(\omega) = x|x\} = 1, \quad \textit{for } x \in \mathbf{E}\backslash N.$$

Bibliography

[1] Antonelli, F., Backward-forward stochastic differential equations, *The Annals of Applied Probability*, 3(3), 1993, 777–793.

[2] Antonelli, F. and Ma. J, Weak solutions of forward-backward SDEs, *Stochastic Analysis and Applications*, 21(3), 2003, 493–514.

[3] Bally, V. and Matoussi, A., Weak solutions for SPDEs and backward doubly stochastic differential equations, *Journal of Theoretical Probability*, 14(1), 2001, 125–164.

[4] Barles, G., Buckdahn, R., and Pardoux, É., Backward stochastic differential equations and integral-partial differential equations, *Stochastics and Stochastics Reports*, 60(1-2), 1997, 57–83.

[5] Bender, C. and Denk, R., A forward scheme for backward SDEs, *Stochastic Processes and their Applications*, 117(12), 2007, 1793–1812.

[6] Bielecki, T. and Jeanblanc, M., Indifference pricing of defaultable claims, *Indifference pricing: theory and applications, Princeton University Press*, 2009, 211–240.

[7] Bismut, J. M., Analyse convexe et probabilitiés, *These, Faculté des Sciences de Paris*, Paris, 1973.

[8] Bismut, J. M., Théorie probabiliste du contrôle des diffusions, *Memoirs of the American Mathematical Society*, 4(167), 1976, 1–167.

[9] Bismut, J. M., An introductory approach to duality in optimal stochastic control, *SIAM Review*, 20(1), 1978, 62–78.

[10] Bouchard, B. and Touzi, N., Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations, *Stochastic Processes and their Applications*, 111(2), 2004, 175–206.

[11] Briand, P. and Hu, Y., BSDE with quadratic growth and unbounded terminal value, *Probability Theory and Related Fields*, 136(4), 2006, 604–618.

[12] Briand, P. and Hu, Y., Quadratic BSDEs with convex generators and unbounded terminal conditions, *Probability Theory and Related Fields*, 141(3-4), 2008, 543–567.

[13] Buckdahn, R., Engelbert, H.-J. and Răşcanu, A., On weak solutions of backward stochastic differential equations, *Rossiĭskaya Akademiya Nauk. Teoriya Veroyatnosteĭ i ee Primeneniya*, 49(1), 2004, 70–108.

[14] Buckdahn, R. and Engelbert, H.-J., On the continuity of weak solutions of backward stochastic differential equations, *Rossiĭskaya Akademiya Nauk. Teoriya Veroyatnosteĭ i ee Primeneniya*, 52(1), 2007, 190–199.

[15] Carmona, R. (editor), Indifference pricing, theory and applications, *Princeton University Press*, Princeton, NJ, 2009.

[16] Casserini, M. and Liang, G., Functional differential equations for coupled forward-backward stochastic dynamics and related numerical approximations, *working paper*, 2010.

[17] Cheridito, P., Soner, M., Touzi, N. and Victoir, N., Second order backward stochastic differential equations and fully non-linear parabolic PDEs, *to appear in Communications in Pure and Applied Mathematics.*

[18] Crisan, D. and Manolarakis, K., Solving backward stochastic differential equations using the cubature method, *working paper*, 2008.

[19] Cvitanić, J. and Karatzas, I., Backward stochastic differential equations with reflection and Dynkin games, *The Annals of Probability*, 24(4), 1996, 2024–2056.

[20] Darling, R. W. R. and Pardoux, E., Backwards SDE with random terminal time and applications to semilinear elliptic PDE, *The Annals of Probability*, 25(3), 1997, 1135–1159.

[21] Davis, M. H. A., Option pricing in incomplete markets, *Mathematics of Derivative Securities, Cambridge University Press*, 1997, 216–227.

[22] Delarue, F., On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case, *Stochastic Processes and Their Applications*, 99, 2002, 209–286.

[23] Delarue, F., Estimates of the solutions of a system of quasi-linear PDEs: A probabilistic scheme, *Séminaire de Probabilités XXXVII, Lecture Notes in Math., Spring*, 1832, 2003, 290–332.

[24] Douglas, Jr. J., Ma, J. and Protter, P., Numerical methods for forward-backward stochastic differential equations, *The Annals of Applied Probability*, 6(3), 1996, 940–968.

[25] Duffie, D. and Epstein, L., Stochastic differential utility, *Econometrica*, 60, 1992, 353–394.

[26] El Karoui, N., Hamadene, S. and Matoussi, A., BSDEs and applications, *Indifference pricing: theory and applications, Princeton University Press*, 2009, 267–320.

[27] El Karoui, N., Kapoudjian, C., Pardoux, É., Peng, S. and Quenez, M. C., Reflected solutions of backward SDE's, and related obstacle problems for PDE's, *The Annals of Probability*, 25(2), 1997, 702–737.

[28] El Karoui, N. and Mazliak, L. (editors), Backward stochastic differential equations, *Pitman Research Notes in Mathematics Series, (Paris, 1995–1996)*, 364, 1997.

[29] El Karoui, N., Pardoux, É. and Quenez, M. C., Reflected backward SDEs and American options, *Numerical methods in finance, Publ. Newton Inst.*, 13, 1997, 215–231.

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

[31] Frei, C., Malamud, S. and Schweizer, M., Convexity bounds for BSDE solutions, with applications to indifference valuation, *working paper*, 2009.

[32] Gobet, E., Lemor, J. and Warin, X., A regression-based Monte Carlo method to solve backward stochastic differential equations, *The Annals of Applied Probability*, 15(3), 2005, 2172–2202.

[33] Hamadene, S., Lepeltier, J.-P. and Matoussi, A., Double barrier backward SDEs with continuous coefficient, *Backward stochastic differential equations (Paris, 1995–1996),Pitman Res. Notes Math. Ser.*, 364, 1997, 161–175.

[34] He, S., Wang, J. and Yan, J. A., Semimartingale theory and stochastic calculus, *Kexue Chubanshe (Science Press)*, Beijing, 1992.

[35] Henderson, V., Valuation of claims on nontraded assets using utility maximization, *Mathematical Finance*, 12, 2002, 351–373.

[36] Henderson, V. and Liang, G., Counterparty risk of derivatives: a utility indifference pricing approach, *working paper*, 2010.

[37] Hodges, S. and Neuberger, A., Optimal replication of contingent claims under transaction costs, *Review of Future Markets*, 8, 1989, 222–239.

[38] Hu, Y., Imkeller, P. and Müller, M., Utility maximization in incomplete markets, *The Annals of Applied Probability*, 15(3), 2005, 1691–1712.

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

[40] Imkeller, P. and Dos Reis, G., Path regularity and explicit convergence rate for BSDE with truncated quadratic growth, *Stochastic Processes and Their Applications*, 120, 2010, 348–379.

[41] Jacod, J. and Yor, M., Étude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 38(2), 1977, 83–125.

[42] Karatzas, I. and Shreve, S. E., Brownian motion and stochastic calculus, (the second edition), *Springer-Verlag*, New York, 1991.

[43] Kazamaki, N., Continuous exponential martingales and BMO, *Lecture Notes in Mathematics (1579), Springer-Verlag*, Berlin, 1994.

[44] Kohlmann, M. and Zhou, X. Y., Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach, *SIAM Journal on Control and Optimization*, 38(5), 2000, 1392–1407.

[45] Kobylanski, M., Backward stochastic differential equations and partial differential equations with quadratic growth, *The Annals of Probability*, 28(2), 2000, 558–602.

[46] Kurtz, T., The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities, *Electronic Journal of Probability*, 12, 2007, 951–965.

[47] Liang, G., Credit risk in incomplete markets, *Transfer thesis*, Oxford University, 2008.

[48] Liang, G. and Jiang, L., A modified structural model for credit risk, *to appear in IMA Journal of Management Mathematics*.

[49] Liang, G., Lyons, T. and Qian, Z., Backward stochastic dynamics on a filtered probability space, *to appear in The Annals of Probability*.

[50] Liang, G., Lyons, T. and Qian, Z., A functional approach to FBSDEs and its applications in optimal portfolios *working paper*, 2010.

[51] Liang, G. and Ren, X., The credit risk and pricing of OTC options, *Asia-Pacific Financial Markets*, 14, 2007, 45–68.

[52] Lepeltier, J. P. and San Martin, J., Backward stochastic differential equations with continuous coefficient, *Statistics & Probability Letters*, 32(4), 1997, 425–430.

[53] Ma, J., Protter, P., San Martín, J. and Torres, S., Numerical method for backward stochastic differential equations, *The Annals of Applied Probability*, 12(1), 2002, 302–316.

[54] Ma, J., Protter, P. and Yong, J., Solving forward-backward stochastic differential equations explicitly–a four step scheme, *Probability Theory and Related Fields*, 98(3), 1994, 339–359.

[55] Ma, J. and Yong, J., Forward-backward stochastic differential equations and their applications, *Lecture Notes in Mathematics, Springer-Verlag*, Berlin, 1999.

[56] Ma, J. and Zhang, J., Path regularity for solutions of backward stochastic differential equations, *Probability Theory and Related Fields*, 122, 2002, 163–190.

[57] Ma, J. and Zhang, J., On weak solutions of forward-backward SDEs, *working paper*, 2009.

[58] Ma, J, Zhang, J. and Zheng, Z., Weak solutions for forward-backward SDEs: a martingale problem approach, *The Annals of Probability*, 36(6), 2008, 2092–2125.

[59] Merton, R. C., On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance*, 29, 1974, 449–470.

[60] Morlais, M.-A., Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem, *Finance and Stochastics*, 13(1), 2009, 121–150.

[61] Musiela, M. and Zariphopoulou, T., An example of indifference prices under exponential preferences, *Finance and Stochastics*, 8, 2004, 229–239.

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

[63] Pardoux, É. and Peng, S., Backward stochastic differential equations and quasilinear parabolic partial differential equations, *Stochastic partial differential equations and their applications (Charlotte, NC, 1991), Lecture Notes in Control and Inform. Sci.*, 176, 1992, 200–217.

[64] Pardoux, É. and Peng, S., Backward doubly stochastic differential equations and systems of quasilinear SPDEs, *Probability Theory and Related Fields*, 98(2), 1994, 209–227.

[65] Parthasarathy, K. R., Probability measures on metric spaces, *Academic Press Inc.*, 1967, New York.

[66] Peng, S. G., A general stochastic maximum principle for optimal control problems, *SIAM Journal on Control and Optimization*, 28(4), 1990, 966–979.

[67] Peng, S. G., Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, *Stochastics and Stochastics Reports*, 37(1-2), 1991, 61–74.

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

[69] Protter, P., Stochastic integration and differential equations, (the 2^{nd} edition), *Springer-Verlag*, Berlin, 2004.

[70] Quenez, M. C., Stochastic control and BSDEs, *Backward stochastic differential equations (Paris, 1995–1996),Pitman Res. Notes Math. Ser.*, 364, 1997, 83–100.

[71] Revuz, D. and Yor, M., Continuous martingales and Brownian motion, (the third edition), *Springer-Verlag*, Berlin, 1999.

[72] Rong, S., On solutions of a backward stochastic differential equations with jumps and application, *Stochastics Processes and Their Applications*, 66, 1997, 209–236.

[73] Rouge, R. and El Karoui, N., Pricing via utility maximization and entropy, *Mathematical Finance*, 10(2), 2000, 259–276.

[74] Tang, S. and Li, X., Maximum principle for optimal control of distributed parameter stochastic systems with random jumps, *Differential equations, dynamical systems, and control science*, 152, 1994, 867–890.

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

[76] Yong, J. and Zhou, X. Y., Stochastic controls: Hamiltonian systems and HJB equations, *Springer-Verlag*, New York, 1999.

[77] Zariphopoulou, T., A solution approach to valuation with unhedgeable risks, *Finance and Stochastics*, 5, 2001, 61–82.

[78] Zhang, Q. and Zhao, H., Stationary solutions of SPDEs and infinite horizon BDSDEs, *Journal of Functional Analysis*, 252(1), 2007, 171–219.

[79] Zhang, Q. and Zhao, H., Stationary solutions of SPDEs and infinite horizon BDSDEs with non-Lipschitz coefficients, *Journal of Differential Equations*, 248(5), 2010, 953–991.

[80] Zhang, J., Some fine properties of backward stochastic differential equations, *PhD thesis*, Purdue University, 2001

[81] Zhang, J., A numerical scheme for BSDEs, *The Annals of Applied Probability*, 14(1), 2004, 459–488.