國立台灣大學理學院數學所

碩士論文

Department of Mathematics

College of Science

National Taiwan University

Master Thesis

線性正倒向隨機微分方程與里卡蒂方程

Linear Forward-Backward Stochastic Differential Equations and a Riccati Type Equation

林柏佐

Po-Tso Lin

指導教授：姜祖恕、謝南瑞 教授

Advisor: Tzuu-Shuu Chiang, Narn-Rueih Shieh, Professor

中華民國101年6月
June 2012

臺灣大學碩士學位論文
口試委員會審定書

線性正倒向隨機微分方程與里卡蒂方程
Linear Forward-Backward Stochastic Differential Equations and a Riccati Type Equation

本論文係林柏佐君（R98221041）在國立臺灣大學數學系完成之碩士學位論文，於民國 101 年 6 月 29 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

（簽名）

（指導教授）

系主任（所長）（簽名）

誌謝

這兩年半來我覺得最幸運的事情就是碰到姜祖恕老師。從最開始找教授的不順利（由於下學期入學，大學也不是台大的學生），“陰錯陽差”最後找到老師收留我。這些日子以來我跟老師念過馬可夫過程、機率論、隨機擴散方程、大離差理論一直到論文裡的正倒向隨機微分方程。光念這些內容所花的精力就非常辛苦，我覺得老師真的是把很多時間都留給我，雖然不是都很懂這些內容，但是老師的確開啓我機率之門。

在我學術指導上最想感謝的就是許順吉老師以及參與中央大學讀書會的各位學長以及老師們，由於台大已經沒有機率相關的課程。所以能在讀書會碰到許順吉老師真的是很幸運，老師是個要求細節很嚴謹的人，而且常常可以指引我們留意到平常沒注意的細節。感謝固定參與讀書會的許玉生老師、趙一峰老師、謝仲老師，你們在讀書會所提的問題與解釋都對我們數學的思考有很大幫助。以及讀書會參與的學長須上苑、吳政訓、張凱君，你們的演講以及不辭辛勞的載我來回台北與中壢真的很感激。以及能夠與我討論很多機率隨機分析的同門：陳韋達、李旭唐、楊如琛等等，能認識你們與你們共同討論數學問題真的是非常開心的事情，也恭喜你們繼續升學往前邁進。

我要感謝在數學所期間，跟我討論課業的室友任宏、鴻儒、晉宏、育齡，以及念統計跟我無話不談的學弟昱廷。很感謝台大數學給了我們很紮實的訓練，讓我們能夠有嚴謹的數學專業知能。最後這一年，我遭遇到分手之後的情傷，受到許多人的幫忙；尤其姜老師除了關心我學業外，更在生活中鼓勵我，並分享他的經驗談。有幾位朋友我特別的感謝，陪我走過這一切；無話不談的老大政洋，與我一起成長聊天的艾咪，遇到困難時可以跟我一聊天就好幾個小時的筱薈，陪我度過感情與家人相處的問題。

感謝數學所期間指導我進入機率論的張志中老師，並且在老師的邀請下學習了一些隨機矩陣方面的知識，和讓我理解更多財務數學背景的彭柏堅老師。以及參與我口試的謝南瑞老師與韓傳祥老師，你們給我的意見，使得我的論文更加完整。

最後我最想感謝的就是我的家人，提供我無私的支持與奉獻，讓我能夠全心徜徉在知識的大海中。爸爸媽媽真的辛苦你們，把我一路拉拔到研究所，感謝哥哥分享過來人的經驗。沒有家人的支持，我無法做我想做的事情，我也順利完成兩份論文了！

中文摘要

本篇論文中我們探討線性正倒向隨機微分方程的可解性。我們在本篇論文中我們探討特殊情形時($\widehat{A} = O$)，可解性的充分必要條件。我們是針對Ma & Yong (2000)工作的延伸。最後我們提出了一個正向微分方程與倒向微分方程的關係，藉由解一個矩陣的常微分方程(里卡蒂方程)提出了類似的充分必要條件。

Abstract

In this paper we investigate the solvability of linear forward-backward stochastic differential equations (FBSDEs, for short). We give sufficient and necessary conditions of the solvability in linear forward-backward stochastic differential equations and prove it in a special case ($\widehat{A} = O$). These results are extensional work of Ma & Yong (2000). Then we introduce the relationship between forward equation and backward equation, we also can get similar sufficient and necessary conditions to solve linear forward-backward stochastic differential equations by solving a matrix ordinary differential equation (a Riccati type equation).

Contents

1. Introduction

The theory of *backward stochastic differential equations* (BSDEs, for short) was pioneered by Pardoux and Peng (1990). It became popular now, and it is an important field of stochastic analysis due to its connections with stochastic control, mathematical finance, and partial differential equations.

The main differences between a *stochastic differential equation* (SDE, for short) and a deterministic ordinary differential equation (ODE, for short) is that previous one can not reverse by "time" (the solution should be adapted). In this paper, we only consider the finite time horizon and a complete filtered probability space which is generated by *Brownian* filtration $\{\mathscr{F}_t\}_{t\geq 0}$ (Brownian motion is denoted by W_t). When we want to solve the *terminal value problem* as following:

$$\begin{cases} dY_t = 0, 0 \leq t \leq T, \\ Y_T = \xi, \end{cases}$$

where $\xi \in L^2_{\mathscr{F}_T}(\Omega;\mathbb{R})$. Since the only solution is $Y_t = \xi$, $0 \leq t \leq T$, which is not necessarily $\{\mathscr{F}_t\}_{t\geq 0}$-adapted unless ξ is a constant, the above Itô SDE, does not have a solution in general!

There are some ways to adjust this difficulty. We want to reformulate the *terminal value problem* of an SDE so that it may allow a solution which is $\{\mathscr{F}_t\}_{t\geq 0}$-adapted. A reasonable method of modifying the solution $Y_t = \xi$ is $\{\mathscr{F}_t\}_{t\geq 0}$-adapted, and it satisfies $Y_T = \xi$. We define

$$Y_t \triangleq E[\xi|\mathscr{F}_t], 0 \leq t \leq T.$$

An important tool in this derivation is the *Martingale Representation Theorem* (cf. e.g., Oksendal (2003), pp. 53-54). Since the above process is clearly a square integrable $\{\mathscr{F}_t\}_{t\geq 0}$-martingale, an application of the Martingale Representation Theorem leads to the following representation:

$$Y_t = Y_0 + \int_0^t Z_s dW_s, 0 \leq t \leq T, \text{ a.s.}$$

where $Z \in L^2_{\mathscr{F}}(0,T;\mathbb{R})$. Writing in a differential form we have

$$\begin{cases} dY_t = Z_t dW_t, 0 \le t \le T, \\ Y_T = \xi. \end{cases}$$

In other words, if we reformulate the above SDE, looking for not a single $\{\mathscr{F}_t\}_{t\ge 0}$-adapted process Y as a solution, we look for a pair (Y,Z), then finding a solution which is $\{\mathscr{F}_t\}_{t\ge 0}$-adapted becomes possible! Adding the extra component Z to the solution is the key factor that makes finding an *adapted* solution possible.

We like rewrite integral form as follows:

$$Y_t = Y_T - \int_t^T Z_s dW_s, 0 \le t \le T.$$

We would not distinguish the above integral forms; each of them is called a BSDE. We emphasize that the stochastic integral is the usual (forward) it$\widehat{o}$ integral.

We give an example in mathematical finance that have motivated the study of the forward-backward stochastic differential equations (FBSDEs, for short). We consider option pricing and contingent claim valuation. Consider a market given by one bond and one stock

$$\begin{cases} dB(t) = rB(t)dt, & \text{(bond)}; \\ dX_t = \mu(t)X_t dt + \sigma(t)X_t dW_t, & \text{(stock)}. \end{cases}$$

Now suppose that the agents sell the European option at price y and then invest it in the market, and we denote their total wealth at each time t by Y_t. Clearly, $Y_0 = y$. Assume that at each time t the agents invest part of their wealth, say π_t, into the stock, and invest the rest $Y_t - \pi_t$ into bond. If we assume that the portfolio is self-financing, then it can be easily shown that

$$\begin{aligned} dY_t &= (Y_t - \pi_t)dB_t + d\pi_t \\ &= r(Y_t - \pi_t)dt + \mu(t)\pi_t dt + \sigma(t)\pi_t dW_t \\ &= \{rY_t + \mu(t)\pi_t - r\pi_t\}dt + \sigma(t)\pi_t dW_t \\ &= [rY_t + \lambda(t)Z_t]dt + Z_t dW_t \end{aligned}$$

where $Z_t = \sigma(t)\pi_t$, and $\lambda(t) = \dfrac{\mu(t) - r}{\sigma(t)}$ (called the *market price of risk*). Suppose contingent claim $H = g(X_T) \in L^2_{\mathscr{F}_T}(\Omega; \mathbb{R})$, the purpose of the agent is to choose π with enough money to hedge the payoff H at time $t = T$, that is, $Y_T \geq H$ a.s. Such an investment, if it exists, is called a *hedging strategy against H*. The *fair price* of the contingent claim is the small initial income for which the hedging strategy exists. Otherwise, it is defined by

$$y^* = \inf\{Y_0 : \exists \pi \text{ such that } Y_T^\pi \geq H \text{ a.s.}\}$$

Now consider agents who are able to choose Z (whence π) by solving following FBSDE (a *decouped* FBSDE):

$$\begin{cases} dX_t = \mu(t)X_t dt + \sigma(t)X_t dW_t, \\ dY_t = [rY_t + \lambda(t)Z_t]dt + Z_t dW_t, \\ X_0 = x, Y_T = g(X_T), \end{cases}$$

An intuitive result is that if above FBSDE has an adapted solution (X, Y, Z), then we let $\pi_t = \dfrac{Z_t}{\sigma(t)}$, is the optimal hedging strategy and $y = Y_0$ is the fair price!

In the theory of ODE, for any first order linear ODE on finite duration with initial value problem, there exists a unique solution. But when we consider the boundary value problem, the uniqueness and existence of solution may not exist. The nonlinear FBSDEs are more complicated than linear ones. It is also an open problem now. We don't consider the nonlinear cases in our study.

Following the Chapter 2 in Ma & Yong (2000), we consider more general case. The Brownian motion takes value in $\mathbb{R}^d$, $d \in \mathbb{N}$ (not just one dimensional case). In this paper we want to solve the following linear coupled FBSDE (3):

$$\begin{cases} d\mathbf{X}_t = (A\mathbf{X}_t + B\mathbf{Y}_t)dt + \sum_{i=1}^{d}(A_1^i \mathbf{X}_t + B_1^i \mathbf{Y}_t + C_1^i \mathbf{Z}_t^i)dW_t^i, \\ d\mathbf{Y}_t = (\widehat{A}\mathbf{X}_t + \widehat{B}\mathbf{Y}_t)dt + \sum_{i=1}^{d} \mathbf{Z}_t^i dW_t^i, & 0 \leq t \leq T \\ \mathbf{X}_0 = \mathbf{0}, \mathbf{Y}_T = \mathbf{g}. \end{cases}$$

We want to answer the question: Does it have an adapted solution whenever $\mathbf{g} \in$

$L^2_{\mathscr{F}_T}(\Omega;\mathbb{R}^m)$? Is the solution unique (under a.s.) or not?

The study proceeds as follows. In Chapter 2, we give the Definitions and Notations throughout this thesis, and we introduce the general Itô Formula in matrix form. This is a very useful tool in the study. We formulate our problem and find the unique solution in the specific space.

In Chapter 3, we want to investigate the solvability of linear FBSDEs. In first section, we give some necessary conditions (29) and (30) to the solvability. To prove the argument, we use the *variation of constant formula* and introduce a linear operator $\mathcal{K}$. We use Burkholder-Davis-Gundy Inequality and Gronwall Inequality to estimate some upper bounds in this section and the following section. The next section, we deal with the sufficient conditions. We prove the range space of $\mathcal{K}$ is a closed subspace of $L^2_{\mathscr{F}_T}(\Omega,\mathbb{R}^m)$. Because our FBSDE (3) is a coupled FBSDE, it is easier to prove the pair $(\mathbf{X}_T,\mathbf{Y}_T)$ is the closed subspace of $L^2_{\mathscr{F}_T}(\Omega,\mathbb{R}^{n+m})$. But we only want to control the process at the terminal time T, it is much more difficult to prove the closeness of $R(\mathcal{K})$. The Lemma 3.5 in the book by Ma & Yong (2000), their argument to prove the closeness of $R(\mathcal{K})$ may be wrong. They claim the projection of closed subspace in Hilbert space is closed. We give one counterexample in the Remark 1 after Lemma 4. We use another methods to conquer this problem. Some estimations are needed, if we can prove $E[|\mathbf{X}_T|^2] \le CE[|\mathbf{Y}_T|^2]$, for some constant $C>0$ and it does not depend on $(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)$, then we can get result. We deal with the particular situation in the Appendix. Under the special case ($\widehat{A}=O$), the closeness of $R(\mathcal{K})$ can be proved. Our main result is Theorem 3.

In Chapter 4, we give some more precise assumption (the connection between forward equation and backward equation). Then we can derive the matrix-valued Riccati type equation. Firstly, we introduce a heuristic derivation. Secondly, by the theory in BSDE and condition (94), we can get the adapted solution that satisfies FBSDE (3). In the first section in chapter 3, we can obtain the necessary condition without any more assumption. Under the circumstance, it is also a sufficient condition. Also, we give the explicit expression of Riccati type equation P. Our major work is Theorem 5. We don't rely on any further more assumption, then the linear FBSDE(3) can be solved and derived the sufficient and necessary conditions.

At last, in Chapter 5 we conclude our work and give some possible extension in the future work and use the theory in some application fields like mathematical finance.

2. Definitions and Notations

Throughout this paper we let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, P)$ be an augmented filtered probability space on $C([0,\infty))$ which is defined a d-dimensional standard Brownian motion $\mathbf{W}_t$, such that $\mathscr{F}_t \triangleq \sigma(\mathscr{F}_t^{\mathbf{W}} \bigcup \mathscr{N})$, $\forall t \geq 0$, where the *natural filtration* $\mathscr{F}_t^{\mathbf{W}} \triangleq \sigma(\mathbf{W}_s; 0 \leq s \leq t)$, $\mathscr{F}_\infty^{\mathbf{W}} \triangleq \sigma\left(\bigcup_{t\geq 0} \mathscr{F}_t^{\mathbf{W}}\right)$, and the collection of *P-null sets* $\mathscr{N} \triangleq \{F \subset \Omega : \exists G \in \mathscr{F}_\infty^{\mathbf{W}}$ with $F \subset G, P(G) = 0\}$. That is, we only consider the *Brownian* filtration.

We list all the notations that will be frequently used throughout the paper, and give some definitions related to FBSDEs.

Let $M_{m\times n}(\mathbb{R})$ be the Banach space consisting of all $m \times n$ matrices with entries $\mathbb{R}$ and $M_n(\mathbb{R})$ be the set of all square matrices with order n over $\mathbb{R}$ with the operator norm $\|A\| = \sqrt{\rho(AA^T)}$, where $\rho(G)$ is the *spectral radius* of G and $\sigma(G)$ is the spectrum of G. Thus,

$$\rho(G) = \max_{\lambda\in\sigma(G)} |\lambda|.$$

Next, we let $T > 0$ be fixed. We denote

- for any sub-σ-algebra $\mathscr{G} \subset \mathscr{F}$, $L^2_{\mathscr{G}}(\Omega; \mathbb{R}^m)$ to be the set of all $\mathscr{G}$-measurable $\mathbb{R}^m$-valued square integrable random vectors;
- $L^2_{\mathscr{F}}(\Omega; L^2(0,T;\mathbb{R}^n))$ to be the set of all $\{\mathscr{F}_t\}_{t\geq 0}$-progressively measurable processes $\mathbf{X}$ values in $\mathbb{R}^n$ such that $E\left[\int_0^T |\mathbf{X}_t|^2 dt\right] < \infty$. The notation $L^2_{\mathscr{F}}(0,T;\mathbb{R}^n)$ is often used for simplicity, when there is no danger of confusion;
- $L^2_{\mathscr{F}}(\Omega; C([0,T];\mathbb{R}^n))$ to be the set of all $\{\mathscr{F}_t\}_{t\geq 0}$-progressively measurable continuous processes $\mathbf{X}$ taking values in $\mathbb{R}^n$ such that $E\left[\sup_{0\leq t\leq T} |\mathbf{X}_t|^2\right] < \infty$.

Further, we define

$$\mathcal{M}[0,T] \triangleq L^2_{\mathscr{F}}(\Omega; C([0,T];\mathbb{R}^n)) \times L^2_{\mathscr{F}}(\Omega; C([0,T];\mathbb{R}^m)) \times [L^2_{\mathscr{F}}(0,T;\mathbb{R}^m)]^d. \tag{1}$$

The norm of this space is defined by

$$\|(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d)\| = \left\{ E\left[\sup_{0\le t\le T} |\mathbf{X}_t|^2 + \sup_{0\le t\le T} |\mathbf{Y}_t|^2 + \sum_{i=1}^{d} \int_0^T |\mathbf{Z}_t^i|^2 dt \right] \right\}^{\frac{1}{2}}, \quad (2)$$

$\forall(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0, T]$. It is clear that $\mathcal{M}[0, T]$ is a Banach space under norm (2).

We are going to study linear FBSDEs in any finite time duration. By deriving a sufficient and necessary condition of solvability, we obtain a reduction to a simple form of linear FBSDEs. We concentrate on the following FBSDE:

$$\begin{cases} d\mathbf{X}_t = (A\mathbf{X}_t + B\mathbf{Y}_t)dt + \sum_{i=1}^{d} (A_1^i \mathbf{X}_t + B_1^i \mathbf{Y}_t + C_1^i \mathbf{Z}_t^i) dW_t^i, \\ d\mathbf{Y}_t = (\widehat{A}\mathbf{X}_t + \widehat{B}\mathbf{Y}_t)dt + \sum_{i=1}^{d} \mathbf{Z}_t^i dW_t^i, \\ \mathbf{X}_0 = \mathbf{0}, \mathbf{Y}_T = \mathbf{g}. \end{cases} \quad 0 \le t \le T \quad (3)$$

In what follows, we will let

$$\begin{cases} A, A_1^i \in M_n(\mathbb{R}); B, B_1^i, C_1^i \in M_{n\times m}(\mathbb{R}), i = 1, \cdots, d \\ \widehat{A} \in M_{m\times n}(\mathbb{R}); \widehat{B} \in M_m(\mathbb{R}); \\ \mathbf{g} \in L^2_{\mathscr{F}_T}(\Omega; \mathbb{R}^m). \end{cases} \quad (4)$$

Definition 1. $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0, T]$ *is called an* ***adapted solution*** *of (3) if the following holds* $\forall t \in [0, T]$, *a.s.:*

$$\begin{cases} \mathbf{X}_t = \int_0^t (A\mathbf{X}_s + B\mathbf{Y}_s)ds + \sum_{i=1}^{d} \int_0^t (A_1^i \mathbf{X}_s + B_1^i \mathbf{Y}_s + C_1^i \mathbf{Z}_s^i) dW_s^i, \\ \mathbf{Y}_t = \mathbf{g} - \int_t^T (\widehat{A}\mathbf{X}_s + \widehat{B}\mathbf{Y}_s)ds - \sum_{i=1}^{d} \int_t^T \mathbf{Z}_s^i dW_s^i. \end{cases} \quad (5)$$

When (3) admits an adapted solution, we say that (3) is solvable.

By denoting

$$\begin{cases} \mathcal{A} = \begin{bmatrix} A & B \\ \widehat{A} & \widehat{B} \end{bmatrix}, \\ \mathcal{A}_1^i = \begin{bmatrix} A_1^i & B_1^i \\ O & O \end{bmatrix}, \mathcal{C}_1^i = \begin{bmatrix} C_1^i \\ I \end{bmatrix}, i = 1, \cdots, d \end{cases} \quad (6)$$

We can write (3) as follows:

$$\begin{cases} d\begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \end{bmatrix} = \mathcal{A}\begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \end{bmatrix} dt + \sum_{i=1}^{d}\left(\mathcal{A}_1^i \begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \end{bmatrix} + \mathcal{C}_1^i \mathbf{Z}_t^i\right) dW_t^i, \\ \mathbf{X}_0 = \mathbf{0}, \mathbf{Y}_T = \mathbf{g}. \end{cases} \quad (7)$$

We want to introduce the general Itô Formula in matrix form. If each of the processes $[A_t]_{jk}$, $[B_t^i]_{jk}$,$i = 1, \cdots, d$, $j = 1, \cdots, m$, $k = 1, \cdots, n$, and $[\widehat{A}_t]_{jk}$, and $[\widehat{B}_t^i]_{jk}$, $i = 1, \cdots, d$, $j = 1, \cdots, n$, $k = 1, \cdots, l$ are all in $L^2_{\mathscr{F}}(0, T; \mathbb{R})$ then we can form the following "matrix" Itô processes (in matrix notation)

$$\begin{cases} dX_t = A_t dt + \sum_{i=1}^{d} B_t^i dW_t^i \\ dY_t = \widehat{A}_t dt + \sum_{i=1}^{d} \widehat{B}_t^i dW_t^i \end{cases} \quad (8)$$

where $X_t = [X_t]_{jk}$, $A_t = [A_t]_{jk}$, $B_t^i = [B_t^i]_{jk}$, $i = 1, \cdots, d$, $j = 1, \cdots, m$, $k = 1, \cdots, n$, and $Y_t = [Y_t]_{jk}$, $\widehat{A}_t = [\widehat{A}_t]_{jk}$, and $\widehat{B}_t^i = [\widehat{B}_t^i]_{jk}$, $i = 1, \cdots, d$, $j = 1, \cdots, n$, $k = 1, \cdots, l$.

Lemma 1 (The General Itô Formula)**.** *Let X, Y be "matrices" Itô processes as (8). Then the process XY is also an Itô process, given by*

$$d(X_t Y_t) = (dX_t)Y_t + X_t dY_t + d\langle X, Y\rangle_t$$

where $\langle X, Y\rangle$ is their "generalized" ***cross-variation process*** *defined by*

$$d\langle X, Y\rangle_t \triangleq \sum_{i=1}^{d} B_t^i \widehat{B}_t^i dt, 0 \leq t \leq T.$$

Proof. Since

$$[X_t Y_t]_{jk} = \sum_{a=1}^{n} [X_t]_{ja}[Y_t]_{ak},$$

by Itô Formula

$$
\begin{aligned}
d[X_tY_t]_{jk} &= \sum_{a=1}^{n} d([X_t]_{ja}[Y_t]_{ak}) = \sum_{a=1}^{n} d[X_t]_{ja}[Y_t]_{ak} + [X_t]_{ja} d[Y_t]_{ak} + d\langle [X]_{ja}, [Y]_{ak}\rangle_t \\
&= \sum_{a=1}^{n} \Bigg([A_t]_{ja}[Y_t]_{ak} dt + \sum_{i=1}^{d} [B_t^i]_{ja}[Y_t]_{ak} dW_t^i \\
&\qquad + [X_t]_{ja}[\widehat{A}_t]_{ak} dt + \sum_{i=1}^{d} [X_t]_{ja}[\widehat{B}_t^i]_{ak} dW_t^i + \sum_{i=1}^{d} [B_t^i]_{ja}[\widehat{B}_t^i]_{ak} dt \Bigg) \\
&= \sum_{a=1}^{n} ([dX_t]_{ja}[Y_t]_{ak} + [X_t]_{ja}[dY_t]_{ak}) + \sum_{i=1}^{d} [B_t^i \widehat{B}_t^i]_{jk} dt \\
&= [(dX_t)Y_t]_{jk} + [X_t dY_t]_{jk} + [d\langle X, Y\rangle_t]_{jk} \\
&= [(dX_t)Y_t + X_t dY_t + d\langle X, Y\rangle_t]_{jk}, j = 1, \cdots, m, k = 1, \cdots, l.
\end{aligned}
$$

This completes the proof of the General Itô Formula. □

3. Solvability of Linear FBSDEs

In this chapter, we are going to present some solvability results for linear FBSDE (3). For convenience, we denote hereafter in this chapter that $H = L^2_{\mathscr{F}_T}(\Omega;\mathbb{R}^m)$ and $\mathcal{H} = [L^2_{\mathscr{F}}(0,T;\mathbb{R}^m)]^d$ (which are Hilbert spaces to which the final datum $\mathbf{g}$ and the processes $(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)$ belong, respectively).

3.1. Necessary Conditions

First of all, we let

$$\begin{cases} d\Phi_t = \mathcal{A}\Phi_t dt + \sum_{i=1}^{d} \mathcal{A}_1^i \Phi_t dW_t^i, t \geq 0, \\ \Phi_0 = I \end{cases} \tag{9}$$

where $\mathcal{A}$ and $\mathcal{A}_1^i$, $i = 1, \cdots, d$ are defined in (6).

Lemma 2. *If Φ is the solution of (9), then, Φ^{-1} exists and it satisfies the following linear SDE:*

$$\begin{cases} d\Phi_t^{-1} = -\Phi_t^{-1}\left[\mathcal{A} - \sum_{i=1}^{d} (\mathcal{A}_1^i)^2\right] dt - \Phi_t^{-1} \sum_{i=1}^{d} \mathcal{A}_1^i dW_t^i, t \geq 0, \\ \Phi_0^{-1} = I \end{cases} \tag{10}$$

Proof. Let us check that (10) is the SDE of Φ^{-1}, by the General Itô Formula,

$$\begin{aligned} d(\Phi_t\Phi_t^{-1}) &= d(\Phi_t)\Phi_t^{-1} + \Phi_t d(\Phi_t^{-1}) + d\langle \Phi, \Phi^{-1}\rangle_t \\ &= \mathcal{A}dt + \sum_{i=1}^{d} \mathcal{A}_1^i dW_t^i - \left[\mathcal{A} - \sum_{i=1}^{d} (\mathcal{A}_1^i)^2\right] dt - \sum_{i=1}^{d} \mathcal{A}_1^i dW_t^i - \sum_{i=1}^{d} (\mathcal{A}_1^i)^2 dt \\ &= O = dI. \end{aligned}$$

Due to (9) is the linear SDE and the Existence and Uniqueness Theorem for Stochastic Differential Equations, Φ^{-1} exists and it satisfies (10). □

Moreover, we consider the following General Itô Formula

$$
\begin{aligned}
d\left(\Phi_t^{-1}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}\right) &= d(\Phi_t^{-1})\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} + \Phi_t^{-1} d\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} + d\langle \Phi^{-1}, \begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\rangle_t \\
&= -\Phi_t^{-1}\left[\mathcal{A} - \sum_{i=1}^d (\mathcal{A}_1^i)^2\right]\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} dt - \Phi_t^{-1}\sum_{i=1}^d \mathcal{A}_1^i \begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} dW_t^i \\
&\quad + \Phi_t^{-1}\mathcal{A}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} dt + \Phi_t^{-1}\sum_{i=1}^d \left(\mathcal{A}_1^i \begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} + \mathcal{C}_1^i \mathbf{Z}_t^i\right) dW_t^i \\
&\quad - \Phi_t^{-1}\sum_{i=1}^d \left\{(\mathcal{A}_1^i)^2 \begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} + \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_t^i\right\} dt \\
&= -\Phi_t^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_t^i dt + \Phi_t^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_t^i dW_t^i
\end{aligned}
$$

So, $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0,T]$ is an adapted solution of (3) if and only if the following variation of constant formula holds:

$$
\Phi_t^{-1}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{y}\end{bmatrix} - \int_0^t \Phi_s^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_s^i ds + \int_0^t \Phi_s^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_s^i dW_s^i, 0 \le t \le T,
$$

or

$$
\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} = \Phi_t\begin{bmatrix}\mathbf{0}\\ \mathbf{y}\end{bmatrix} - \Phi_t\int_0^t \Phi_s^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_s^i ds + \Phi_t\int_0^t \Phi_s^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_s^i dW_s^i, 0 \le t \le T, \quad (11)
$$

for some $\mathbf{y} \in \mathbb{R}^m$ with the condition:

$$
\mathbf{g} = [\ O \quad I\ ]\left(\Phi_T\begin{bmatrix}\mathbf{0}\\ \mathbf{y}\end{bmatrix} - \Phi_T\int_0^T \Phi_s^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_s^i ds + \Phi_T\int_0^T \Phi_s^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_s^i dW_s^i\right). \quad (12)
$$

Let us introduce an operator $\mathcal{K} : \mathcal{H} \to H$ as follows:

$$
\mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d) = [\ O \quad I\ ]\left(-\Phi_T\int_0^T \Phi_s^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_s^i ds + \Phi_T\int_0^T \Phi_s^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_s^i dW_s^i\right). \quad (13)
$$

Then, for given $\mathbf{g} \in H$, finding an adapted solution to (3) is equivalent to the following: Find $\mathbf{y} \in \mathbb{R}^m$ and $(\mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{H}$ such that

$$
\mathbf{g} = [\ O \quad I\ ]\Phi_T\begin{bmatrix}O\\ I\end{bmatrix}\mathbf{y} + \mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d), \quad (14)
$$

and define $(\mathbf{X}, \mathbf{Y})$ by (11). Then $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0,T]$ is an adapted solution of (3). We now make some findings on Φ and $\mathcal{K}$. Let us first give the following lemma.

Lemma 3. *For any* $\mathbf{f} \in L_{\mathscr{F}}(0,T;\mathbb{R}^{n+m})$ *and* $\mathbf{h} \in L^2_{\mathscr{F}}(0,T;\mathbb{R}^{n+m})$*, it holds*

$$\begin{cases} E[\Phi_t] = e^{\mathcal{A}t}, \\ E\left[\Phi_t \int_0^t \Phi_s^{-1}\mathbf{f}_s ds\right] = \int_0^t e^{\mathcal{A}(t-s)} E[\mathbf{f}_s] ds, \\ E\left[\Phi_t \int_0^t \Phi_s^{-1}\mathbf{h}_s dW_s^i\right] = \int_0^t e^{\mathcal{A}(t-s)} \mathcal{A}_1^i E[\mathbf{h}_s] ds, i = 1, \cdots, d, \end{cases} \quad 0 \le t \le T. \tag{15}$$

Also, it holds that

$$E\left[\sup_{0\le t\le T} \|\Phi_t\|^{2k}\right] + E\left[\sup_{0\le t\le T} \|\Phi_t^{-1}\|^{2k}\right] < \infty, \forall k \ge 1 \tag{16}$$

Proof. We suppose first that (16) holds. Taking expectation in (9), we obtain

$$\begin{cases} dE[\Phi_t] = \mathcal{A}E[\Phi_t]dt, t \in [0,T], \\ \Phi_0 = I. \end{cases} \tag{17}$$

Thus,

$$E[\Phi_t] = e^{\mathcal{A}t}.$$

We have proved the first equality in (15). Let us prove the second equality in (15). Set

$$\xi_t = \Phi_t \int_0^t \Phi_s^{-1}\mathbf{f}_s ds, 0 \le t \le T. \tag{18}$$

By the General Itô Formula

$$\begin{aligned} d\xi_t &= (d\Phi_t)\int_0^t \Phi_s^{-1}\mathbf{f}_s ds + \mathbf{f}_t dt \\ &= (\mathcal{A}\xi_t + \mathbf{f}_t)dt + \sum_{i=1}^d \mathcal{A}_1^i \xi_t dW_t^i. \end{aligned}$$

Then ξ satisfies the following SDE:

$$\begin{cases} d\xi_t = (\mathcal{A}\xi_t + \mathbf{f}_t)dt + \sum\limits_{i=1}^d \mathcal{A}_1^i \xi_t dW_t^i, 0 \le t \le T \\ \xi_0 = \mathbf{0}. \end{cases} \tag{19}$$

Taking expectation in (19), we obtain

$$\begin{cases} dE[\xi_t] = (\mathcal{A}E[\xi_t] + E[\mathbf{f}_t])dt, 0 \le t \le T \\ E[\xi_0] = \mathbf{0}. \end{cases} \tag{20}$$

Hence,

$$\frac{d}{dt}e^{-\mathcal{A}t}E[\xi_t] = -\mathcal{A}e^{-\mathcal{A}t}E[\xi_t] + \mathcal{A}e^{-\mathcal{A}t}E[\xi_t] + e^{-\mathcal{A}t}E[\mathbf{f}_t] = e^{-\mathcal{A}t}E[\mathbf{f}_t],$$

i.e.

$$e^{-\mathcal{A}t}E[\xi_t] = \int_0^t e^{-\mathcal{A}s}E[\mathbf{f}_s]ds.$$

Thus,

$$E[\xi_t] = \int_0^t e^{\mathcal{A}(t-s)}E[\mathbf{f}_s]ds, 0 \le t \le T, \tag{21}$$

proving our claim.

The third one we can set

$$\zeta_t = \Phi_t \int_0^t \Phi_s^{-1}\mathbf{h}_s dW_s^i, 0 \le t \le T. \tag{22}$$

By the General Itô Formula

$$\begin{aligned} d\zeta_t &= (d\Phi_t)\int_0^t \Phi_s^{-1}\mathbf{h}_s dW_s^i + \mathbf{h}_t dW_t^i + d\langle \Phi, \int_0^\cdot \Phi_s^{-1}\mathbf{h}_s dW_s^i\rangle_t \\ &= (\mathcal{A}\zeta_t + \mathcal{A}_1^i\mathbf{h}_t)dt + \sum_{i=1}^d \mathcal{A}_1^i\zeta_t dW_t^i + \mathbf{h}_t dW_t^i. \end{aligned}$$

Then ζ satisfies the following SDE:

$$\begin{cases} d\zeta_t = (\mathcal{A}\zeta_t + \mathcal{A}_1^i\mathbf{h}_t)dt + \sum_{i=1}^d \mathcal{A}_1^i\zeta_t dW_t^i + \mathbf{h}_t dW_t^i, t \in [0,T], \\ \zeta_0 = \mathbf{0}. \end{cases} \tag{23}$$

Taking expectation in (23), we obtain

$$\begin{cases} dE[\zeta_t] = (\mathcal{A}E[\zeta_t] + \mathcal{A}_1^i E[\mathbf{h}_t])dt, 0 \le t \le T, \\ E[\zeta_0] = \mathbf{0}. \end{cases} \tag{24}$$

Hence,

$$\frac{d}{dt}e^{-\mathcal{A}t}E[\zeta_t] = -\mathcal{A}e^{-\mathcal{A}t}E[\zeta_t] + \mathcal{A}e^{-\mathcal{A}t}E[\zeta_t] + e^{-\mathcal{A}t}\mathcal{A}_1^i E[\mathbf{h}_t] = e^{-\mathcal{A}t}\mathcal{A}_1^i E[\mathbf{h}_t],$$

i.e.

$$e^{-\mathcal{A}t}E[\zeta_t] = \int_0^t e^{-\mathcal{A}s}\mathcal{A}_1^i E[\mathbf{h}_s]ds.$$

Thus,

$$E[\zeta_t] = \int_0^t e^{\mathcal{A}(t-s)}\mathcal{A}_1^i E[\mathbf{h}_s]ds, 0 \le t \le T, \tag{25}$$

proving our claim.

Now we prove (16). For any $\xi_0 \in \mathbb{R}^{n+m}$, process $\xi_t = \Phi_t \xi_0$ satisfies the following SDE:

$$\begin{cases} d\xi_t = \mathcal{A}\xi_t dt + \sum_{i=1}^{d} \mathcal{A}_1^i \xi_t dW_t^i, 0 \le t \le T, \\ \xi_0 = \xi_0. \end{cases} \tag{26}$$

Then $E\left[\sup_{0\le t\le T} |\xi_t|^{2k}\right] < \infty$. Since $|a+b|^k \le (2\max\{|a|,|b|\})^p \le 2^p(|a|^p + |b|^p)$ $\forall a, b \in \mathbb{R}^{n+m}, p \ge 1$, we see that

$$\begin{aligned} E\left[\sup_{0\le r\le t} |\xi_r|^{2k}\right] &\le E\left[\sup_{0\le r\le t} (|\xi_r - \xi_0| + |\xi_0|)^{2k}\right] \le 2^{2k}\left(E\left[\sup_{0\le r\le t} |\xi_r - \xi_0|^{2k}\right] + |\xi_0|^{2k}\right) \\ &\le 2^{2k}\left\{E\left[\sup_{0\le r\le t}\left(\left|\xi_r - \xi_0 - \int_0^r \mathcal{A}\xi_s ds\right| + \left|\int_0^r \mathcal{A}\xi_s ds\right|\right)^{2k}\right] + |\xi_0|^{2k}\right\} \\ &\le 2^{2k}\left\{2^{2k} E\left[\sup_{0\le r\le t}\left(\left|\xi_r - \xi_0 - \int_0^r \mathcal{A}\xi_s ds\right|^{2k} + \left|\int_0^r \mathcal{A}\xi_s ds\right|^{2k}\right)\right] + |\xi_0|^{2k}\right\} \end{aligned}$$

since $\xi_r - \xi_0 - \int_0^r \mathcal{A}\xi_s ds = \sum_{i=1}^{d} \int_0^r \mathcal{A}_1^i \xi_s dW_s^i$ is a martingale with respect to $\mathscr{F}_r$, by the Burkholder-Davis-Gundy Inequality, $\exists K > 0$ (depending only on k), and Hölder Inequality such that

$$\begin{aligned} & E\left[\sup_{0\le r\le t}\left|\xi_r - \xi_0 - \int_0^r \mathcal{A}\xi_s ds\right|^{2k}\right] \le KE\left[\left(\sum_{i=1}^{d}\int_0^t |\mathcal{A}_1^i \xi_s|^2 ds\right)^k\right] \\ \le & KE\left[\sum_{i=1}^{d}\left(\int_0^t |\mathcal{A}_1^i \xi_s|^2 ds\right)^k \left(\sum_{i=1}^{d} 1\right)^{k-1}\right] = Kd^{k-1}\sum_{i=1}^{d} E\left[\left(\int_0^t |\mathcal{A}_1^i \xi_s|^2 ds\right)^k\right] \\ \le & Kd^{k-1}\sum_{i=1}^{d} E\left[\int_0^t |\mathcal{A}_1^i \xi_s|^{2k} ds \left(\int_0^t 1 ds\right)^{k-1}\right] = K(td)^{k-1}\sum_{i=1}^{d} E\left[\int_0^t |\mathcal{A}_1^i \xi_s|^{2k} ds\right] \\ \le & K(td)^{k-1}\sum_{i=1}^{d} \|\mathcal{A}_1^i\|^{2k} E\left[\int_0^t |\xi_s|^{2k} ds\right], 0 \le t \le T. \end{aligned}$$

Using Hölder Inequality again, we obtain

$$E\left[\sup_{0\le r\le t}|\xi_r|^{2k}\right]\le 2^{2k}\left\{2^{2k}\left(K(td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}E\left[\int_0^t|\xi_s|^{2k}ds\right]+E\left[\sup_{0\le r\le t}\left|\int_0^r\mathcal{A}\xi_s ds\right|^{2k}\right]\right)+|\xi_0|^{2k}\right\}$$

$$\le 2^{2k}\left\{2^{2k}\left(K(td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}E\left[\int_0^t|\xi_s|^{2k}ds\right]+E\left[\sup_{0\le r\le t}\int_0^r|\mathcal{A}\xi_s|^{2k}ds\left(\int_0^r 1ds\right)^{2k-1}\right]\right)+|\xi_0|^{2k}\right\}$$

$$\le 2^{2k}\left\{2^{2k}\left(K(td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}E\left[\int_0^t|\xi_s|^{2k}ds\right]+E\left[\sup_{0\le r\le t}r^{2k-1}\int_0^r|\mathcal{A}\xi_s|^{2k}ds\right]\right)+|\xi_0|^{2k}\right\}$$

$$\le 2^{2k}\left\{2^{2k}\left(K(td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}E\left[\int_0^t|\xi_s|^{2k}ds\right]+t^{2k-1}\|\mathcal{A}\|^{2k}E\left[\int_0^t|\xi_s|^{2k}ds\right]\right)+|\xi_0|^{2k}\right\}$$

$$\le 2^{2k}\left\{2^{2k}\left(K(Td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}+T^{2k-1}\|\mathcal{A}\|^{2k}\right)E\left[\int_0^t\sup_{0\le r\le s}|\xi_r|^{2k}ds\right]+|\xi_0|^{2k}\right\},$$

and by Gronwall Inequality, we can show that

$$E\left[\sup_{0\le r\le t}|\xi_r|^{2k}\right]\le e^{2^{4k}\left(K(Td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}+T^{2k-1}\|\mathcal{A}\|^{2k}\right)t}2^{2k}|\xi_0|^{2k}$$

and

$$E\left[\sup_{0\le t\le T}|\xi_t|^{2k}\right]\le e^{2^{4k}\left(K(Td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}+T^{2k-1}\|\mathcal{A}\|^{2k}\right)T}2^{2k}|\xi_0|^{2k}, k\ge 1 \tag{27}$$

for some constant $K>0$. Thus, the first term on the left hand side of (16)

$$E\left[\sup_{0\le t\le T}\|\Phi_t\|^{2k}\right]=E\left[\sup_{0\le t\le T}\sup_{\xi_0\neq\mathbf{0}}\frac{|\xi_t|^{2k}}{|\xi_0|^{2k}}\right]\le 2^{2k}e^{2^{4k}\left(K(Td)^{k-1}\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^{2k}+T^{2k-1}\|\mathcal{A}\|^{2k}\right)T}$$

is finite. Similarly, one can prove that the second term is finite as well. □

Now, we let $\mathbf{Y}_T=\mathcal{K}(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)$ and

$$\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}=-\Phi_t\int_0^t\Phi_s^{-1}\sum_{i=1}^{d}\mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_s^i ds+\Phi_t\int_0^t\Phi_s^{-1}\sum_{i=1}^{d}\mathcal{C}_1^i\mathbf{Z}_s^i dW_s^i, 0\le t\le T,$$

then $(\mathbf{X},\mathbf{Y})$ fulfills the following SDE:

$$\begin{cases} d\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}=\mathcal{A}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}dt+\sum_{i=1}^{d}\left(\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_t^i\right)dW_t^i,\\ \mathbf{X}_0=\mathbf{0},\mathbf{Y}_0=\mathbf{0}.\end{cases}$$

From Itô Formula,

$$d(|\mathbf{X}_t|^2+|\mathbf{Y}_t|^2)=2\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}\cdot d\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}+d\langle\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\rangle_t$$

$$=\left(2\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}\mathcal{A}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}+\sum_{i=1}^{d}\left|\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_t^i\right|^2\right)dt$$

$$+2\sum_{i=1}^{d}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}\cdot\left(\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_t^i\right)dW_t^i,$$

taking expectation in both side,

$$E[|\mathbf{X}_t|^2+|\mathbf{Y}_t|^2]=E\left[\int_0^t 2\begin{bmatrix}\mathbf{X}_s\\ \mathbf{Y}_s\end{bmatrix}\mathcal{A}\begin{bmatrix}\mathbf{X}_s\\ \mathbf{Y}_s\end{bmatrix}+\sum_{i=1}^{d}\left|\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_s\\ \mathbf{Y}_s\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_s^i\right|^2 ds\right]$$

$$\leq 2\left(\|\mathcal{A}\|+\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^2\right)E\left[\int_0^t|\mathbf{X}_s|^2+|\mathbf{Y}_s|^2ds\right]+2\max_{i=1,\cdots,d}\|\mathcal{C}_1^i\|\sum_{i=1}^{d}E\left[\int_0^t|\mathbf{Z}_s^i|^2ds\right].$$

By Gronwall Inequality,

$$E[|\mathbf{X}_t|^2+|\mathbf{Y}_t|^2]\leq 2\max_{i=1,\cdots,d}\|\mathcal{C}_1^i\|e^{2t\left(\|\mathcal{A}\|+\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^2\right)}\sum_{i=1}^{d}E\left[\int_0^t|\mathbf{Z}_s^i|^2ds\right],$$

hence

$$E[|\mathcal{K}(\mathbf{Z}_1,\cdots,\mathbf{Z}_d)|^2]=E[|\mathbf{Y}_T|^2]\leq 2\max_{i=1,\cdots,d}\|\mathcal{C}_1^i\|e^{2T\left(\|\mathcal{A}\|+\sum_{i=1}^{d}\|\mathcal{A}_1^i\|^2\right)}\sum_{i=1}^{d}E\left[\int_0^T|\mathbf{Z}_s^i|^2ds\right],$$

so $\mathcal{K}:\mathcal{H}\to H$ is a bounded linear operator. Now, if (3) admits an adapted solution, by taking expectation in (12) and using (15), we obtain

$$E[\mathbf{g}]=[\begin{array}{cc}O & I\end{array}]\left\{e^{\mathcal{A}T}\begin{bmatrix}O\\ I\end{bmatrix}\mathbf{y}-\sum_{i=1}^{d}\int_0^T e^{\mathcal{A}(T-s)}\mathcal{A}_1^i\mathcal{C}_1^iE[\mathbf{Z}_s^i]ds+\sum_{i=1}^{d}\int_0^T e^{\mathcal{A}(T-s)}\mathcal{A}_1^i\mathcal{C}_1^iE[\mathbf{Z}_s^i]ds\right\},$$

$$=[\begin{array}{cc}O & I\end{array}]e^{\mathcal{A}T}\begin{bmatrix}O\\ I\end{bmatrix}\mathbf{y}\qquad(28)$$

for some $\mathbf{y}\in\mathbb{R}^m$. This leads to the following necessary condition for the solvability of (3).

Theorem 1. *Suppose (3) is solvable* $\forall\mathbf{g}\in H$. *Then*

$$\det\left([\begin{array}{cc}O & I\end{array}]e^{\mathcal{A}T}\begin{bmatrix}O\\ I\end{bmatrix}\right)\neq 0.\qquad(29)$$

Proof. Since $R(E) = \mathbb{R}^m$ and $\exists y \in \mathbb{R}^m$ such that (28) holds. So we can easily get $R\left(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}T} \begin{bmatrix} O \\ I \end{bmatrix}\right) = \mathbb{R}^m$. Then (29) holds. $\square$

Let us now present another necessary condition for the solvability of (3).

Theorem 2. *Suppose (3) is solvable* $\forall \mathbf{g} \in H$*. Then,*

$$\det(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}t} \mathcal{C}_1^i) > 0, \forall t \in [0, T], i = 1, \cdots, d. \tag{30}$$

Consequently, if

$$\widehat{T} = \min_{i=1,\cdots,d} \inf\{T > 0 : \det(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}T} \mathcal{C}_1^i) = 0\} < \infty, \tag{31}$$

then, for any $T > \widehat{T}$, $\exists \mathbf{g} \in H$, *such that (3) is not solvable.*

Proof. Suppose $\exists s_0 \in [0, T)$ and some $j \in \{1, \cdots, d\}$, such that

$$\det(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}(T-s_0)} \mathcal{C}_1^j) = 0. \tag{32}$$

Note that $s_0 < T$ has to be true. Then $\exists \eta \in \mathbb{R}^m$, $|\eta| = 1$, such that

$$\eta^T \begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}(T-s_0)} \mathcal{C}_1^j = \mathbf{0}^T. \tag{33}$$

We are going to prove that $\forall \epsilon > 0$ with $s_0 + \epsilon < T$, $\exists \mathbf{g} \in L^2_{\mathscr{F}_{s_0+\epsilon}}(\Omega; \mathbb{R}^m) \subset H$, such that (3) has no adapted solutions. To this end, we let $\beta : [0, T] \to \mathbb{R}$ be a Lebesgue measurable function such that

$$\begin{cases} \beta(s) = \pm 1, \forall s \in [0, s_0 + \epsilon]; \beta(s) = 0, \forall s \in (s_0 + \epsilon, T]; \\ m(\{s \in [s_0, s_k] : \beta(s) = 1\}) = \frac{s_k - s_0}{2}, \\ m(\{s \in [s_0, s_k] : \beta(s) = -1\}) = \frac{s_k - s_0}{2}, k \in \mathbb{N}, \end{cases} \tag{34}$$

where m means Lebesgue measure, for some sequence $s_k \searrow s_0$ and $s_k \leq s_0 + \epsilon$. Next, we define

$$\zeta_t = \sum_{i=1}^{d} \int_0^t \beta(s) dW_s^i, 0 \leq t \leq T \tag{35}$$

and take $\mathbf{g} = \zeta_T \eta \in L^2_{\mathscr{F}_{s_0+\epsilon}}(\Omega; \mathbb{R}^m) \subset H$. Suppose (3) admits an adapted solution $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0, T]$ for this $\mathbf{g}$. Then $\exists \mathbf{y} \in \mathbb{R}^m$, by (7) and the General Itô

Formula

$$
\begin{aligned}
d\left(e^{-\mathcal{A}t}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}\right) &= -\mathcal{A}e^{-\mathcal{A}t}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}dt + e^{-\mathcal{A}t}d\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}\\
&= -\mathcal{A}e^{-\mathcal{A}t}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}dt + e^{-\mathcal{A}t}\left\{\mathcal{A}\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix}dt + \sum_{i=1}^{d}\left(\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_t^i\right)dW_t^i\right\}\\
&= e^{-\mathcal{A}t}\sum_{i=1}^{d}\left(\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_t^i\right)dW_t^i,
\end{aligned}
$$

with integral form

$$
e^{-\mathcal{A}T}\begin{bmatrix}\mathbf{X}_T\\ \zeta_T\eta\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{y}\end{bmatrix} + \sum_{i=1}^{d}\int_0^T e^{-\mathcal{A}s}\left(\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_s\\ \mathbf{Y}_s\end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_s^i\right)dW_s^i,
$$

we have

$$
\zeta_T\eta = [\ O\quad I\ ]\begin{bmatrix}\mathbf{X}_T\\ \zeta_T\eta\end{bmatrix} = [\ O\quad I\ ]\left\{e^{\mathcal{A}T}\begin{bmatrix}\mathbf{0}\\ \mathbf{y}\end{bmatrix} + \sum_{i=1}^{d}\int_0^T e^{\mathcal{A}(T-s)}\left(\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_s\\ \mathbf{Y}_s\end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_s^i\right)dW_s^i\right\}, \tag{36}
$$

Multiplying η^T from left to (36) gives the following:

$$
\zeta_T = \alpha + \sum_{i=1}^{d}\int_0^T \gamma_s^i + \psi^i(s)\cdot\mathbf{Z}_s^i dW_s^i, \tag{37}
$$

where

$$
\begin{cases}
\alpha = \eta^T[\ O\quad I\ ]e^{\mathcal{A}T}\begin{bmatrix}\mathbf{0}\\ \mathbf{y}\end{bmatrix} \in \mathbb{R}\\
\gamma_s^i = \eta^T[\ O\quad I\ ]e^{\mathcal{A}(T-s)}\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_s\\ \mathbf{Y}_s\end{bmatrix} \in L^2_{\mathscr{F}}(\Omega; C([0,T])), i = 1,\cdots,d,\\
\psi^i(s) = \{\eta^T[\ O\quad I\ ]e^{\mathcal{A}(T-s)}\mathcal{C}_1^i\}^T \text{ is analytic, } i = 1,\cdots,d, \psi^j(s_0) = \mathbf{0}.
\end{cases} \tag{38}
$$

Let us denote

$$
\theta_t = \alpha + \sum_{i=1}^{d}\int_0^t \gamma_s^i + \psi^i(s)\cdot\mathbf{Z}_s^i dW_s^i, 0 \le t \le T. \tag{39}
$$

Then, it follows that

$$
\begin{cases}
d(\theta_t - \zeta_t) = \sum_{i=1}^{d}\gamma_t^i + \psi^i(t)\cdot\mathbf{Z}_t^i - \beta(t)dW_t^i, 0 \le t \le T\\
\theta_T - \zeta_T = 0.
\end{cases} \tag{40}
$$

By Itô Formula,

$$d(\theta_t - \zeta_t)^2 = 2(\theta_t - \zeta_t)d(\theta_t - \zeta_t) + d\langle \theta - \zeta \rangle_t$$
$$= 2(\theta_t - \zeta_t)\sum_{i=1}^{d} \gamma_t^i + \psi^i(t) \cdot \mathbf{Z}_t^i - \beta(t)dW_t^i + \sum_{i=1}^{d}[\gamma_t^i + \psi^i(t) \cdot \mathbf{Z}_t^i - \beta(t)]^2 dt,$$

we have

$$0 = E[(\theta_t - \zeta_t)^2] + \sum_{i=1}^{d} E\left[\int_t^T [\gamma_s^i + \psi^i(s) \cdot \mathbf{Z}_s^i - \beta(s)]^2 ds\right], 0 \le t \le T. \quad (41)$$

Thus,

$$\beta(s) - \gamma_s^i = \psi^i(s) \cdot \mathbf{Z}_s^i, \text{ a.e. } s \in [0, T], \text{ a.s. } \forall i = 1, \cdots, d, \quad (42)$$

which yields

$$\int_{s_0}^{s_k} E[[\beta(s) - \gamma_s^j]^2]ds = \int_{s_0}^{s_k} E[[\psi^j(s) \cdot \mathbf{Z}_s^j]^2]ds, \forall k \in \mathbb{N}. \quad (43)$$

Now, we observe and use Cauchy-Schwarz Inequality that (note $\gamma^j \in L^2_{\mathscr{F}}(\Omega; C([0,T]))$ and (34))

$$\int_{s_0}^{s_k} E[[\beta(s) - \gamma_s^j]^2]ds = \int_{s_0}^{s_k} E[[\beta(s) - \gamma_s^j]^2] + E[(\gamma_s^j - \gamma_{s_0}^j)^2] - E[(\gamma_s^j - \gamma_{s_0}^j)^2]ds$$
$$\ge \frac{1}{2}\int_{s_0}^{s_k} E[[\beta(s) - \gamma_{s_0}^j]^2]ds - \int_{s_0}^{s_k} E[(\gamma_s^j - \gamma_{s_0}^j)^2]ds$$
$$\ge \frac{s_k - s_0}{4} E[(1 - \gamma_{s_0}^j)^2 + (1 + \gamma_{s_0}^j)^2] - o(s_k - s_0), k \in \mathbb{N}. \quad (44)$$

(Since $\{\int_{s_0}^t (\gamma_s^j - \gamma_{s_0}^j)^2 ds\}_{t \ge s_0}$ is a submartingale with respect to $\{\mathscr{F}_t\}_{t \ge 0}$ and $\{\int_{s_0}^t (\gamma_s^j - \gamma_{s_0}^j)^2 ds\}_{t \ge s_0}$ is a continuous process. By the Theorem 3.13 in Karatzas & Shreve (1998), $\left\{E\left[\int_{s_0}^t (\gamma_s^j - \gamma_{s_0}^j)^2 ds\right]\right\}_{t \ge s_0}$ is continuous in $t \in [s_0, s_k]$. By the Mean Value Theorem for Integrals, $\exists c \in (s_0, s_k)$ such that $\frac{1}{s_k - s_0}\int_{s_0}^{s_k} E[(\gamma_s^j - \gamma_{s_0}^j)^2]ds = E[(\gamma_c^j - \gamma_{s_0}^j)^2] \to 0$ as $s_k \searrow s_0$.) On the other hand, since ψ^j is analytic with $\psi^j(s_0) = 0$, we must have

$$\psi^j(s) = (s - s_0)\widetilde{\psi}(s), 0 \le s \le T, \quad (45)$$

for some $\widetilde{\psi}$ which is analytic and hence bounded on $[0, T]$. Hence, we assume that $|\widetilde{\psi}(s)| \le K$, $\forall s \in [0, T]$. Consequently,

$$\int_{s_0}^{s_k} E[[\psi^j(s) \cdot \mathbf{Z}_s^j]^2]ds = (s_k - s_0)^2 \int_{s_0}^{s_k} E[[\widetilde{\psi}(s) \cdot \mathbf{Z}_s^j]^2]ds$$
$$\le K^2 (s_k - s_0)^2 \int_{s_0}^{s_k} E[|\mathbf{Z}_s^j|^2]ds. \quad (46)$$

Hence, (43)-(44) and (46) imply

$$\begin{aligned}&\frac{s_k - s_0}{4}E[(1-\gamma^j_{s_0})^2 + (1+\gamma^j_{s_0})^2] - o(s_k - s_0)\\ \le& K^2(s_k - s_0)^2\int_{s_0}^{s_k} E[|\mathbf{Z}^j_s|^2]ds, \forall k \in \mathbb{N}.\end{aligned} \tag{47}$$

Divide $s_k - s_0$ and let $s_k \searrow s_0$, we get

$$\frac{1}{4}E[(1-\gamma^j_{s_0})^2 + (1+\gamma^j_{s_0})^2] \le 0$$

This is impossible. Finally, noting the fact that $\det([\ O \quad I\ ]\mathcal{C}^i_1) = \det(I) = 1$, $\forall i = 1, \cdots, d$, and $[\ O \quad I\ ]e^{\mathcal{A}t}\mathcal{C}^i_1$ is analytic $\forall i = 1, \cdots, d$. We obtain (30). The final assertion is clear. □

It is not clear if the drift coefficient also contains some $\mathbf{Z}^i$ terms since the assumption with no $\mathbf{Z}^i$ terms is crucial in the proof.

3.2. Criteria for Solvability

We will use Gronwall Inequality many times, but we now introduce the special case.

Proposition 1 (Gronwall Inquality). *Let $v(t) : [0, T] \to \mathbb{R}$ such that*

$$v(t) \le C + A\int_t^T v(s)ds \text{ for } 0 \le t \le T$$

for some constants C and $A \ge 0$. Prove that

$$v(t) \le Ce^{A(T-t)} \text{ for } 0 \le t \le T. \tag{48}$$

Proof. If $A = 0$, the result is clear. We may assume $A > 0$. Define $w(t) = \int_t^T v(s)ds$. Then $-w'(t) \le C + Aw(t)$. Consider $f(t) = -w(t)e^{At}$,

$$f'(t) = -w'(t)e^{At} - Aw(t)e^{At} = [-w'(t) - Aw(t)]e^{At} \le Ce^{At},$$

so

$$-w(T)e^{AT} + w(t)e^{At} = w(t)e^{At} \le C\int_t^T e^{As}ds = \frac{C}{A}(e^{AT} - e^{At}).$$

Deduce that

$$w(t) \le \frac{C}{A}[e^{A(T-t)} - 1] \tag{49}$$

Use (49) to deduce

$$v(t) \le C + Aw(t) \le Ce^{A(T-t)}$$

$\square$

Let us now present some results on the operator $\mathcal{K}$ (see (13) for definition) which will lead to some sufficient conditions for solvability of linear FBSDEs.

Lemma 4. *Let $\widehat{A} = O$. Then the range $R(\mathcal{K})$ of $\mathcal{K}$ is closed in H.*

Proof. Let us denote $H_0 = L^2_{\mathscr{F}_T}(\Omega;\mathbb{R}^n)$ and $\widehat{H} = H_0 \times H \equiv L^2_{\mathscr{F}_T}(\Omega;\mathbb{R}^{n+m})$. Define

$$\widehat{\mathcal{K}}(\mathbf{Z}^1,\cdots,\mathbf{Z}^d) = -\Phi_T\int_0^T \Phi_s^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_s^i ds + \Phi_T\int_0^T \Phi_s^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_s^i dW_s^i, (\mathbf{Z}^1,\cdots,\mathbf{Z}^d) \in \mathcal{H}. \tag{50}$$

Then, $\widehat{\mathcal{K}}$ is a bounded linear operator and $\mathcal{K} = [\ O \quad I\ ]\widehat{\mathcal{K}}$. We claim that the range $R(\widehat{\mathcal{K}})$ of $\widehat{\mathcal{K}}$ is closed in $\widehat{H}$. To show this, let us take any convergence sequence

$$\begin{bmatrix} \mathbf{X}_T^{(k)} \\ \mathbf{Y}_T^{(k)} \end{bmatrix} \equiv \widehat{\mathcal{K}}(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)_k \to \zeta, \text{ in } \widehat{H}, \tag{51}$$

let

$$\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} = -\Phi_t\int_0^t \Phi_s^{-1}\sum_{i=1}^d \mathcal{A}_1^i\mathcal{C}_1^i\mathbf{Z}_{s,k}^i ds + \Phi_t\int_0^t \Phi_s^{-1}\sum_{i=1}^d \mathcal{C}_1^i\mathbf{Z}_{s,k}^i dW_s^i, (\mathbf{Z}^1,\cdots,\mathbf{Z}^d)_k \in \mathcal{H},$$

where $(\mathbf{X}^{(k)},\mathbf{Y}^{(k)})$ is the solution of the following:

$$\begin{cases} d\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} = \mathcal{A}\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} dt + \sum_{i=1}^d \left(\mathcal{A}_1^i \begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_{t,k}^i\right) dW_t^i, \\ \begin{bmatrix} \mathbf{X}_0^{(k)} \\ \mathbf{Y}_0^{(k)} \end{bmatrix} = \mathbf{0} \end{cases} \tag{52}$$

Then, by Itô Formula,

$$\begin{aligned} d(|\mathbf{X}_t^{(k)}|^2 + |\mathbf{Y}_t^{(k)}|^2) &= 2\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} \cdot d\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} + d\langle \begin{bmatrix} \mathbf{X}^{(k)} \\ \mathbf{Y}^{(k)} \end{bmatrix} \rangle_t \\ &= 2\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} \cdot \left\{ \mathcal{A}\begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} dt + \sum_{i=1}^d \left(\mathcal{A}_1^i \begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_{t,k}^i\right) dW_t^i \right\} \\ &\quad + \sum_{i=1}^d \left| \mathcal{A}_1^i \begin{bmatrix} \mathbf{X}_t^{(k)} \\ \mathbf{Y}_t^{(k)} \end{bmatrix} + \mathcal{C}_1^i\mathbf{Z}_{t,k}^i \right|^2 dt, \end{aligned}$$

we have

$$E\left[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2+\sum_{i=1}^d\int_t^T\left|\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_s^{(k)}\\ \mathbf{Y}_s^{(k)}\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_{s,k}^i\right|^2ds\right]$$
$$=E\left[|\mathbf{X}_T^{(k)}|^2+|\mathbf{Y}_T^{(k)}|^2-2\int_t^T\begin{bmatrix}\mathbf{X}_s^{(k)}\\ \mathbf{Y}_s^{(k)}\end{bmatrix}\cdot\mathcal{A}\begin{bmatrix}\mathbf{X}_s^{(k)}\\ \mathbf{Y}_s^{(k)}\end{bmatrix}ds\right]. \tag{53}$$

We note that (recall $\mathcal{C}_1^i=\begin{bmatrix}C_1^i\\ I\end{bmatrix}$) and use the Inequality of Arithmetic and Geometric Means to derive $\alpha a^2+\frac{1}{\alpha}b^2\geq 2|ab|\ \forall\alpha>0$, we get

$$\left|\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t^{(k)}\\ \mathbf{Y}_t^{(k)}\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_{t,k}^i\right|^2$$
$$=[I+(C_1^i)^TC_1^i]\mathbf{Z}_{t,k}^i\cdot\mathbf{Z}_{t,k}^i+\left|\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t^{(k)}\\ \mathbf{Y}_t^{(k)}\end{bmatrix}\right|^2+2(\mathcal{C}_1^i)^T\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t^{(k)}\\ \mathbf{Y}_t^{(k)}\end{bmatrix}\cdot\mathbf{Z}_{t,k}^i$$
$$\geq|\mathbf{Z}_{t,k}^i|^2-\|A_1^i\|^2[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2]-2\|C_1^i\|\|A_1^i\|[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2]^{\frac{1}{2}}|\mathbf{Z}_{t,k}^i|$$
$$\geq|\mathbf{Z}_{t,k}^i|^2-\|A_1^i\|^2[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2]-2\|C_1^i\|^2\|A_1^i\|^2[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2]-\frac{1}{2}|\mathbf{Z}_{t,k}^i|^2$$
$$\geq\frac{1}{2}|\mathbf{Z}_{t,k}^i|^2-C_i[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2],\forall i=1,\cdots,d \tag{54}$$

for some constant $C_i>0$. Thus, (53) implies

$$E\left[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2+\sum_{i=1}^d\int_t^T|\mathbf{Z}_{s,k}^i|^2ds\right]$$
$$\leq E\left[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2+2\sum_{i=1}^d\int_t^T C_i[|\mathbf{X}_s^{(k)}|^2+|\mathbf{Y}_s^{(k)}|^2]+\left|\mathcal{A}_1^i\begin{bmatrix}\mathbf{X}_t^{(k)}\\ \mathbf{Y}_t^{(k)}\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_{t,k}^i\right|^2ds\right]$$
$$\leq 2E\left[|\mathbf{X}_T^{(k)}|^2+|\mathbf{Y}_T^{(k)}|^2-2\int_t^T\begin{bmatrix}\mathbf{X}_s^{(k)}\\ \mathbf{Y}_s^{(k)}\end{bmatrix}\cdot\mathcal{A}\begin{bmatrix}\mathbf{X}_s^{(k)}\\ \mathbf{Y}_s^{(k)}\end{bmatrix}ds+\sum_{i=1}^dC_i\int_t^T|\mathbf{X}_s^{(k)}|^2+|\mathbf{Y}_s^{(k)}|^2ds\right]$$
$$\leq 2\left(1+2\|\mathcal{A}\|+\sum_{i=1}^dC_i\right)E\left[|\mathbf{X}_T^{(k)}|^2+|\mathbf{Y}_T^{(k)}|^2+\int_t^T|\mathbf{X}_s^{(k)}|+|\mathbf{Y}_s^{(k)}|^2+\sum_{i=1}^d\int_s^T|\mathbf{Z}_{r,k}^i|^2drds\right],0\leq t\leq T. \tag{55}$$

Using Gronwall Inequality (Proposition 1), we obtain

$$\begin{aligned} &E\left[|\mathbf{X}_t^{(k)}|^2+|\mathbf{Y}_t^{(k)}|^2+\sum_{i=1}^{d}\int_t^T|\mathbf{Z}_{s,k}^i|^2ds\right] \\ \leq &2\left(1+2\|\mathcal{A}\|+\sum_{i=1}^{d}C_i\right)e^{2\left(1+2\|\mathcal{A}\|+\sum_{i=1}^{d}C_i\right)(T-t)}E\left[|\mathbf{X}_T^{(k)}|^2+|\mathbf{Y}_T^{(k)}|^2\right],0\leq t\leq T. \end{aligned} \tag{56}$$

From the convergence (51) and (56), we see that $(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)_k$ is bounded in $\mathcal{H}$. Since $\mathcal{H}$ is a Hilbert space (hence is a reflexive Banach space), the bounded set is weakly sequentially compact. Thus, we may assume that $\exists(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)_{k_l}\rightharpoonup(\widetilde{\mathbf{Z}}^1,\cdots,\widetilde{\mathbf{Z}}^d)$ in $\mathcal{H}$. Then it is easy to see that $\widehat{K}(\widetilde{\mathbf{Z}}^1,\cdots,\widetilde{\mathbf{Z}}^d)=\zeta$, proving the closeness of $R(\widehat{\mathcal{K}})$.

We take any convergence in $R(\mathcal{K})$

$$\mathbf{Y}_T^{(k)}=\mathcal{K}(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)_k\to\eta \text{ in } H$$

and $\{\mathbf{X}_T^{(k)}\}_{k\in\mathbb{N}}$ is defined by (51). By Lemma 7 in Appendix, $\exists C>0$ with

$$E[|\mathbf{X}_T^{(k)}-\mathbf{X}_T^{(l)}|^2]\leq CE[|\mathbf{Y}_T^{(k)}-\mathbf{Y}_T^{(l)}|^2]$$

Since $\{\mathbf{Y}_T^{(k)}\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $R(\mathcal{K})$, then $\{(\mathbf{X}_T^{(k)},\mathbf{Y}_T^{(k)})\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $R(\widehat{\mathcal{K}})$. Due to the closeness in $R(\widehat{\mathcal{K}})$, $\exists(\mathbf{X},\mathbf{Y})\in R(\widehat{\mathcal{K}})$ with $\{(\mathbf{X}_T^{(k)},\mathbf{Y}_T^{(k)})\}_{k\in\mathbb{N}}\to(\mathbf{X},\mathbf{Y})$ in $R(\widehat{\mathcal{K}})$. We get the result $\mathbf{Y}\equiv\eta$ a.s. Hence, $R(\mathcal{K})$ is closed. □

Remark 1. *In Ma & Yong (2000), pp. 41. They claim $R(\mathcal{K})$ is closed by following procedures. First we know that $R(\widehat{\mathcal{K}})$ is a Hilbert space with the induced inner product from that of $\widehat{H}$. In this space, we define an orthogonal projection $P_H:\widehat{H}\to\widehat{H}$ by the following:*

$$P_H\begin{bmatrix}\xi\\ \eta\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \eta\end{bmatrix},\forall\begin{bmatrix}\xi\\ \eta\end{bmatrix}\in\widehat{H}\equiv H_0\times H.$$

Then the space

$$P_H(R(\widehat{\mathcal{K}}))=\{\mathbf{0}\}\times R(\mathcal{K})$$

is closed in $R(\widehat{\mathcal{K}})$ and so in $\widehat{H}$. Hence, $R(\mathcal{K})$ is closed in H.

The argument is not correct. We give a counterexample. Let $\widehat{H}$ *be an infinitely dimensional separable Hilbert space, and let* $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ *be an orthonormal basis. Let* $H_0 = span\{\mathbf{e}_{2n-1}\}_{n\in\mathbb{N}}$ *and* $H = span\{\mathbf{e}_{2n}\}_{n\in\mathbb{N}}$. *It is clear that* $\widehat{H} = H_0 \oplus H$. *Let*

$$\mathbf{f}_n = \frac{n\mathbf{e}_{2n-1} + \mathbf{e}_{2n}}{\sqrt{n^2+1}}, n \in \mathbb{N},$$

and let $V = \overline{span\{\mathbf{f}_n\}}_{n\in\mathbb{N}}$. *We want to claim* $P_H(V)$ *is not closed in* H. *Take*

$$\mathbf{y}_n = \sum_{m=1}^{n} \frac{\mathbf{e}_{2m}}{m}, \mathbf{z}_n = \sum_{m=1}^{n} \frac{\sqrt{m^2+1}}{m}\mathbf{f}_m, n \in \mathbb{N},$$

then $\mathbf{z}_n \in V$, $\forall n \in \mathbb{N}$.

$$P_H(\mathbf{z}_n) = \sum_{m=1}^{n} \frac{\sqrt{m^2+1}}{m} \frac{\mathbf{e}_{2m}}{\sqrt{m^2+1}} = \mathbf{y}_n.$$

Hence $\mathbf{y}_n \in P_H(V)$, *and*

$$\mathbf{y}_n \to \sum_{n=1}^{\infty} \frac{\mathbf{e}_{2n}}{n} = \mathbf{y} \in H,$$

but $\mathbf{y}$ *does not sit in* $P_H(V)$.

If $\mathbf{y} \in P_H(V)$, $\exists\{a_n\}_{n\in\mathbb{N}}$ *with* $\sum_{n=1}^{\infty} a_n^2 < \infty$ *such that (projection is a continuous mapping)*

$$\mathbf{y} = P_H(\sum_{n=1}^{\infty} a_n\mathbf{f}_n) = \sum_{n=1}^{\infty} a_n P_H(\mathbf{f}_n) = \sum_{n=1}^{\infty} \frac{a_n}{\sqrt{n^2+1}}\mathbf{e}_{2n}.$$

Hence,

$$a_n = \frac{\sqrt{n^2+1}}{n} \to 1 \neq 0 \text{ as } n \to \infty.$$

It is a contradiction.

The following result gives some more information for the operator $\mathcal{K}$.

Lemma 5. *Let (30) hold and* $\widehat{A} = O$. *Then*

$$R(\mathcal{K}) = \{\eta \in L^2_{\mathscr{F}_T}(\Omega, \mathbb{R}^m) : E[\eta] = \mathbf{0}\} \triangleq N(E), \tag{57}$$

$$N(\mathcal{K}) = \{(\mathbf{0}, \cdots, \mathbf{0})\}. \tag{58}$$

Proof. First of all, by Lemma 4, we see that $R(\mathcal{K})$ is closed. Also, by (13) and Lemma 3,

$$\forall \eta \in R(\mathcal{K}), \exists(\mathbf{Z}^1, \cdots, \mathbf{Z}^d) \text{ such that } \mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d) = \eta,$$

and

$$E[\eta] = E[\mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d)] = [\ O \quad I\ ]E\left[-\Phi_T \int_0^T \Phi_s^{-1} \sum_{i=1}^d \mathcal{A}_1^i \mathcal{C}_1^i \mathbf{Z}_s^i ds + \Phi_T \int_0^T \Phi_s^{-1} \sum_{i=1}^d \mathcal{C}_1^i \mathbf{Z}_s^i dW_s^i\right]$$

$$= [\ O \quad I\ ] \sum_{i=1}^d - \int_0^T e^{\mathcal{A}(T-s)} \mathcal{A}_1^i \mathcal{C}_1^i E[\mathbf{Z}_s^i] ds + \int_0^T e^{\mathcal{A}(T-s)} \mathcal{A}_1^i \mathcal{C}_1^i E[\mathbf{Z}_s^i] ds = \mathbf{0};$$

therefore, $\eta \in N(E)$ and $R(\mathcal{K}) \subset N(E)$. Thus, to show (57), it suffice to show that

$$N(E) \bigcap R(\mathcal{K})^{\perp} = \{\mathbf{0}\}. \tag{59}$$

If (59) holds, then

$$H = R(\mathcal{K}) \oplus R(\mathcal{K})^{\perp} \subset N(E) \oplus R(\mathcal{K})^{\perp} \subset H,$$

hence $H = R(\mathcal{K}) \oplus R(\mathcal{K})^{\perp} = N(E) \oplus R(\mathcal{K})^{\perp}$. Therefore, if $\eta \in N(E)$, then $\exists ! \mathbf{u} \in R(\mathcal{K}), \mathbf{w} \in R(\mathcal{K})^{\perp}$ such that $\eta = \mathbf{u} + \mathbf{w}$, i.e. $\eta - \mathbf{u} = \mathbf{w} \in N(E) \bigcap R(\mathcal{K})^{\perp} = \{\mathbf{0}\}$, so we get $\eta = \mathbf{u} \in R(\mathcal{K})$ and $N(E) = R(\mathcal{K})$.

We now prove (59). Take $\eta \in N(E) \bigcap R(\mathcal{K})^{\perp}$. Suppose

$$0 = E[\eta \cdot \mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d)], \forall (\mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{H}. \tag{60}$$

The above holds $\forall (\mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{H}$, let $0 < \delta < T$ and take

$$\mathbf{Z}_s^i = \begin{cases} (\mathcal{C}_1^j)^T e^{\mathcal{A}^T (T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j 1_{[T-\delta, T]}(s), & i = j \\ \mathbf{0}, & i \neq j \end{cases}, 0 \leq s \leq T. \tag{61}$$

Then $\overline{\mathbf{X}}_s = \mathbf{0}$, $\overline{\mathbf{Y}}_s = \mathbf{0}$, $\forall s \in [0, T-\delta)$. And

$$\begin{bmatrix} \overline{\mathbf{X}}_t \\ \overline{\mathbf{Y}}_t \end{bmatrix} = -\Phi_t \int_0^t \Phi_s^{-1} \mathcal{A}_1^j \mathcal{C}_1^j \mathbf{Z}_s^j ds + \Phi_t \int_0^t \Phi_s^{-1} \mathcal{C}_1^j \mathbf{Z}_s^j dW_s^j, 0 \leq t \leq T. \tag{62}$$

Then, we have

$$\begin{cases} d \begin{bmatrix} \overline{\mathbf{X}}_t \\ \overline{\mathbf{Y}}_t \end{bmatrix} = \mathcal{A} \begin{bmatrix} \overline{\mathbf{X}}_t \\ \overline{\mathbf{Y}}_t \end{bmatrix} dt + \left(\mathcal{A}_1^j \begin{bmatrix} \overline{\mathbf{X}}_t \\ \overline{\mathbf{Y}}_t \end{bmatrix} + \mathcal{C}_1^j \mathbf{Z}_t^j \right) dW_t^j, \\ \begin{bmatrix} \overline{\mathbf{X}}_0 \\ \overline{\mathbf{Y}}_0 \end{bmatrix} = \mathbf{0}. \end{cases} \tag{63}$$

By Itô Formula,

$$
\begin{aligned}
d(|\overline{\mathbf{X}}_t|^2+|\overline{\mathbf{Y}}_t|^2) &= 2\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}\cdot d\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}+d\langle\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}\rangle_t \\
&= 2\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}\cdot\left\{\mathcal{A}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}dt+\left(\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_t^i\right)dW_t^j\right\} \\
&\quad+\left|\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_t^j\right|^2dt,
\end{aligned}
$$

using Cauchy-Schwarz Inequality

$$
\begin{aligned}
E[|\overline{\mathbf{X}}_t|^2+|\overline{\mathbf{Y}}_t|^2] &= E\left[\int_0^t 2\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix}\cdot\mathcal{A}\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix}+\left|\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_s^j\right|^2ds\right] \\
&\le 2\left(\|\mathcal{A}\|+\|\mathcal{A}_1^j\|^2\right)E\left[\int_0^t|\overline{\mathbf{X}}_s|^2+|\overline{\mathbf{Y}}_s|^2ds\right]+2\|\mathcal{C}_1^j\|^2E\left[\int_0^t|\mathbf{Z}_s^j|^2ds\right],
\end{aligned}
$$

and Gronwall Inequality, we obtain

$$
E[|\overline{\mathbf{X}}_t|^2+|\overline{\mathbf{Y}}_t|^2]\le 2\|\mathcal{C}_1^j\|^2e^{2(\|\mathcal{A}\|+\|\mathcal{A}_1^j\|^2)t}E\left[\int_0^t|\mathbf{Z}_s^j|^2ds\right],0\le t\le T. \tag{64}
$$

By the General Itô Formula again,

$$
\begin{aligned}
d\left(e^{-\mathcal{A}t}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}\right) &= -\mathcal{A}e^{-\mathcal{A}t}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}dt+e^{-\mathcal{A}t}d\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix} \\
&= -\mathcal{A}e^{-\mathcal{A}t}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}dt+e^{-\mathcal{A}t}\left\{\mathcal{A}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}dt+\left(\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_t^j\right)dW_t^j\right\} \\
&= e^{-\mathcal{A}t}\left(\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_t^j\right)dW_t^j,
\end{aligned}
$$

with integral form

$$
e^{-\mathcal{A}t}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}=\int_0^t e^{-\mathcal{A}s}\left(\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_s^j\right)dW_s^j.
$$

Also, we have

$$
\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}=\int_0^t e^{\mathcal{A}(t-s)}\left(\mathcal{A}_1^j\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix}+\mathcal{C}_1^j\mathbf{Z}_s^j\right)dW_s^j,0\le t\le T. \tag{65}
$$

Since $E[\eta] = \mathbf{0}$ and $\eta \in H$, by the Itô Representation Theorem, $\exists(\zeta^1, \cdots, \zeta^d) \in \mathcal{H}$, such that

$$\eta = \sum_{i=1}^{d} \int_0^T \zeta_s^i dW_s^i. \tag{66}$$

Then, from (60) and (65), we have

$$\begin{aligned} 0 = E[\eta \cdot \mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d)] &= E\left[\eta \cdot [\; O \quad I \;] \begin{bmatrix} \overline{\mathbf{X}}_T \\ \overline{\mathbf{Y}}_T \end{bmatrix}\right] \\ &= E\left[\int_0^T \zeta_s^j \cdot [\; O \quad I \;] e^{\mathcal{A}(T-s)} \left(\mathcal{A}_1^j \begin{bmatrix} \overline{\mathbf{X}}_s \\ \overline{\mathbf{Y}}_s \end{bmatrix} + \mathcal{C}_1^j \mathbf{Z}_s^j\right) ds\right]. \end{aligned} \tag{67}$$

This yields

$$\begin{aligned} &E\left[\int_0^T (\mathcal{C}_1^j)^T e^{\mathcal{A}^T(T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j \cdot \mathbf{Z}_s^j ds\right] \\ =&- E\left[\int_0^T (\mathcal{A}_1^j)^T e^{\mathcal{A}^T(T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j \cdot \begin{bmatrix} \overline{\mathbf{X}}_s \\ \overline{\mathbf{Y}}_s \end{bmatrix} ds\right]. \end{aligned} \tag{68}$$

Consequently, (64) and (68) result in (use Tonelli Theorem and Cauchy-Schwarz Inequality)

$$\begin{aligned} &E\left[\int_{T-\delta}^T \left|(\mathcal{C}_1^j)^T e^{\mathcal{A}^T(T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j\right|^2 ds\right] \\ \le& \|\mathcal{A}_1^j\| e^{\|\mathcal{A}\|T} E\left[\int_{T-\delta}^T |\zeta_s^j| (|\overline{\mathbf{X}}_s|^2 + |\overline{\mathbf{Y}}_s|^2)^{\frac{1}{2}} ds\right] \\ \le& \sqrt{2} \|\mathcal{A}_1^j\| \|\mathcal{C}_1^j\| e^{(2\|\mathcal{A}\| + \|\mathcal{A}_1^j\|^2)T} \int_{T-\delta}^T (E[|\zeta_s^j|^2])^{\frac{1}{2}} \left(\int_{T-\delta}^s E[|\mathbf{Z}_r^j|^2] dr\right)^{\frac{1}{2}} ds \\ \le& \sqrt{2} \|\mathcal{A}_1^j\| \|\mathcal{C}_1^j\|^2 e^{(3\|\mathcal{A}\| + \|\mathcal{A}_1^j\|^2)T} \int_{T-\delta}^T (E[|\zeta_s^j|^2])^{\frac{1}{2}} \left(\int_{T-\delta}^s E[|\zeta_r^j|^2] dr\right)^{\frac{1}{2}} ds \end{aligned} \tag{69}$$

By (30), we obtain

$$\begin{aligned} &\int_{T-\delta}^T E[|\zeta_s^j|^2] ds \le K \int_{T-\delta}^T E\left[\left|\mathcal{C}_1^T e^{\mathcal{A}^T(T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j\right|^2\right] ds \\ \le& \sqrt{2} \|\mathcal{A}_1^j\| \|\mathcal{C}_1^j\|^2 e^{(3\|\mathcal{A}\| + \|\mathcal{A}_1^j\|^2)T} K \int_{T-\delta}^T (E[|\zeta_s^j|^2])^{\frac{1}{2}} \left(\int_{T-\delta}^s E[|\zeta_r^j|^2] dr\right)^{\frac{1}{2}} ds \\ \le& \frac{1}{2} \int_{T-\delta}^T E[|\zeta_s^j|^2] ds + \|\mathcal{A}_1^j\|^2 \|\mathcal{C}_1^j\|^4 e^{2(3\|\mathcal{A}\| + \|\mathcal{A}_1^j\|^2)T} K^2 \int_{T-\delta}^T \int_{T-\delta}^s E[|\zeta_r^j|^2] dr ds \end{aligned} \tag{70}$$

Thus, it follows that

$$\int_{T-\delta}^{T} E[|\zeta_s^j|^2]ds \le 2\|\mathcal{A}_1^j\|^2\|\mathcal{C}_1^j\|^4 e^{2(3\|\mathcal{A}\|+\|\mathcal{A}_1^j\|^2)T}K^2\delta \int_{T-\delta}^{T} E[|\zeta_s^j|^2]ds, \tag{71}$$

with $K > 0$ being an absolute constant (independent of δ). Therefore, for $\delta > 0$ small, we must have

$$\zeta_s^j = \mathbf{0}, \text{ a.e. } s \in [T-\delta, T], \text{ a.s.} \tag{72}$$

This together with (68) implies that

$$\begin{aligned} &E\left[\int_0^{T-\delta} (\mathcal{C}_1^j)^T e^{\mathcal{A}^T(T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j \cdot \mathbf{Z}_s^j ds\right] \\ =& -E\left[\int_0^{T-\delta} (\mathcal{A}_1^j)^T e^{\mathcal{A}^T(T-s)} \begin{bmatrix} O \\ I \end{bmatrix} \zeta_s^j \cdot \begin{bmatrix} \overline{\mathbf{X}}_s \\ \overline{\mathbf{Y}}_s \end{bmatrix} ds\right]. \end{aligned} \tag{73}$$

Then, thanks to (30), we can continue the above procedure to conclude that (72) holds over $[0,T]$ $\forall j = 1, \cdots, d$.

If the above argument is not true, $\exists j = 1, \cdots, d, s_0^j \in (0,T)$ such that $\zeta_s^j = \mathbf{0}$, a.e. $s \in [s_0^j, T]$, a.e., and $\forall \epsilon_j \in (0, s_0^j)$, $\exists A_{\epsilon_j} \subset [s_0^j - \epsilon, s_0^j)$ with $m(A_{\epsilon_j}) > 0$ such that

$$\zeta_s^j \neq 0, \forall s \in A_{\epsilon_j} \text{ a.s.},$$

but from above similar procedure, we can find some $\delta_j > 0$ such that

$$\zeta_s^j = \mathbf{0}, \text{ a.e. } s \in [s_0^j - \delta_j, s_0^j], \text{ a.s.},$$

it is a contradiction, and hence it follows from (66) that $\eta \equiv \mathbf{0}$ a.s. This proves (59).

We now prove (58). Suppose $\mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d) = \mathbf{0}$. Again, we let $(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$ be defined by

$$\begin{bmatrix} \overline{\mathbf{X}}_t \\ \overline{\mathbf{Y}}_t \end{bmatrix} = -\Phi_t \int_0^t \Phi_s^{-1} \sum_{i=1}^d \mathcal{A}_1^i \mathcal{C}_1^i \mathbf{Z}_s^i ds + \Phi_t \int_0^t \Phi_s^{-1} \sum_{i=1}^d \mathcal{C}_1^i \mathbf{Z}_s^i dW_s^i, 0 \le t \le T.$$

Like (65), we have

$$\begin{bmatrix} \overline{\mathbf{X}}_t \\ \overline{\mathbf{Y}}_t \end{bmatrix} = \int_0^t e^{\mathcal{A}(t-s)} \sum_{i=1}^d \left(\mathcal{A}_1^i \begin{bmatrix} \overline{\mathbf{X}}_s \\ \overline{\mathbf{Y}}_s \end{bmatrix} + \mathcal{C}_1^i \mathbf{Z}_s^i \right) dW_s^i, 0 \le t \le T.$$

Then, $\forall(\zeta^1,\cdots,\zeta^d)\in\mathcal{H}$ and above, we have

$$
\begin{aligned}
0 &= E\left[\left\{\sum_{i=1}^{d}\int_0^T \zeta_s^i dW_s^i\right\}\cdot\mathcal{K}(\mathbf{Z}^1,\cdots,\mathbf{Z}^d)\right]\\
&= \sum_{i=1}^{d} E\left[\int_0^T \zeta_s^i\cdot[\ O\quad I\ ]e^{\mathcal{A}(T-s)}\left(\mathcal{A}_1^i\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_s^i\right)ds\right]. \qquad (74)
\end{aligned}
$$

This implies that

$$
[\ O\quad I\ ]e^{\mathcal{A}(T-s)}\left(\mathcal{A}_1^i\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix}+\mathcal{C}_1^i\mathbf{Z}_s^i\right)=\mathbf{0},\ \text{a.e. } s\in[0,T],\ \text{a.s.}, i=1,\cdots,d \qquad (75)
$$

By (30), we easily see that

$$
s\mapsto\mathcal{B}^i(s)\triangleq\{[\ O\quad I\ ]e^{\mathcal{A}(T-s)}\mathcal{C}_1^i\}^{-1}[\ O\quad I\ ]e^{\mathcal{A}(T-s)}\mathcal{A}_1^i, i=1,\cdots,d
$$

is analytic and hence bounded over $[0,T]$. From (75), we obtain

$$
\mathbf{Z}_s^i=-\mathcal{B}^i(s)\begin{bmatrix}\overline{\mathbf{X}}_s\\ \overline{\mathbf{Y}}_s\end{bmatrix},\ \text{a.e. } s\in[0,T],\ \text{a.s.} i=1,\cdots,d \qquad (76)
$$

Then $(\overline{\mathbf{X}},\overline{\mathbf{Y}})$ is the solution of

$$
\begin{cases}
d\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}=\mathcal{A}\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}dt+\sum_{i=1}^{d}[\mathcal{A}_1^i-\mathcal{C}_1^i\mathcal{B}^i(t)]\begin{bmatrix}\overline{\mathbf{X}}_t\\ \overline{\mathbf{Y}}_t\end{bmatrix}dW_t,\\
\begin{bmatrix}\overline{\mathbf{X}}_0\\ \overline{\mathbf{Y}}_0\end{bmatrix}=\mathbf{0}.
\end{cases} \qquad (77)
$$

Hence, we must have $(\overline{\mathbf{X}},\overline{\mathbf{Y}})\equiv(\mathbf{0},\mathbf{0})$ a.s., which yields $\mathbf{Z}^i\equiv\mathbf{0}$ a.s., $i=1,\cdots,d$ due to (76). This proves (58). □

A consequence of the above is the following.

Theorem 3. *Linear FBSDE (3) (with $\widehat{A}=O$) is solvable $\forall\mathbf{g}\in H$ if and only if (29) and (30) hold. In this case, the adapted solution to (3) is unique (for any given $\mathbf{g}\in H$).*

Proof. Theorem 1 and 2 tell us that (29) and (30) are necessary conditions. We now prove the sufficiency. First of all, for any $\mathbf{g}\in H$, we can find $\mathbf{y}\in\mathbb{R}^m$, such that (28) holds (by (29)). Then we have

$$
\mathbf{g}-[\ O\quad I\ ]\Phi_T\begin{bmatrix}O\\ I\end{bmatrix}\mathbf{y}\in N(E). \qquad (78)
$$

Next, by (57), $\exists(\mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{H}$, such that

$$\mathbf{g} - [\ O \quad I\ ]\Phi_T \begin{bmatrix} O \\ I \end{bmatrix} \mathbf{y} = \mathcal{K}(\mathbf{Z}^1, \cdots, \mathbf{Z}^d). \tag{79}$$

For any pair $(\mathbf{y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathbb{R}^m \times \mathcal{H}$, we define $(\mathbf{X}, \mathbf{Y})$ by (11). Then one can easily check that $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0, T]$ is an adapted solution of (3). The uniqueness follows easily from (30) and (58). $\square$

The above result gives a complete solution to the solvability of linear FBSDE (3) without any $\mathbf{Z}^i$ term in drift coefficient. It is also a problem when we consider the general linear FBSDE case.

4. A Riccati Type Equation

First, we consider the following BSDE:

$$\begin{cases} d\mathbf{Y}_t = \mathbf{h}(t, \mathbf{Y}_t, \mathbf{Z}_t^1, \cdots, \mathbf{Z}_t^d)dt + \sum_{i=1}^{d} \mathbf{Z}_t^i dW_t^i, 0 \le t \le T \\ \mathbf{Y}_T = \xi, \end{cases} \tag{80}$$

where $\xi \in L^2_{\mathscr{F}_T}(\Omega, \mathbb{R}^m)$ and $\mathbf{h} \in L^2_{\mathscr{F}}(0, T; W^{1,\infty}(\mathbb{R}^m \times (\mathbb{R}^m)^d; \mathbb{R}^m))$ i.e., $\mathbf{h} : [0, T] \times \mathbb{R}^m \times (\mathbb{R}^m)^d \times \Omega \to \mathbb{R}^m$, such that $(t, \omega) \mapsto \mathbf{h}(t, \mathbf{y}, \mathbf{z}^1, \cdots, \mathbf{z}^d; \omega)$ is $\{\mathscr{F}_t\}_{t\ge 0}$-progressively measurable $\forall (\mathbf{y}, \mathbf{z}^1, \cdots, \mathbf{z}^d) \in \mathbb{R}^m \times (\mathbb{R}^m)^d$ with $\mathbf{h}(t, \mathbf{0}, \mathbf{0}, \cdots, \mathbf{0}; \omega) \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^m)$ and $\exists L > 0$,

$$|\mathbf{h}(\mathbf{y}, \mathbf{z}^1, \cdots, \mathbf{z}^d) - \mathbf{h}(\overline{\mathbf{y}}, \overline{\mathbf{z}}^1, \cdots, \overline{\mathbf{z}}^d)| \le L \left(|\mathbf{y} - \overline{\mathbf{y}}| + \sum_{i=1}^{d} |\mathbf{z}^i - \overline{\mathbf{z}}^i| \right),$$

$$\forall \mathbf{y}, \mathbf{z}^1, \cdots, \mathbf{z}^d, \overline{\mathbf{y}}, \overline{\mathbf{z}}^1, \cdots, \overline{\mathbf{z}}^d \in \mathbb{R}^m, \text{ a.e. } t \in [0, T], \text{ a.s.} \tag{81}$$

Denote

$$\mathcal{N}[0, T] \triangleq L^2_{\mathscr{F}}(\Omega; C(0, T; \mathbb{R}^m)) \times [L^2_{\mathscr{F}}(0, T; \mathbb{R}^m)]^d \tag{82}$$

and

$$\|(\mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d)\|_{\mathcal{N}[0,T]} \triangleq \left(E\left[\sup_{0\le t\le T} |\mathbf{Y}_t|^2 + \sum_{i=1}^{d} \int_0^T |\mathbf{Z}_t|^2 dt \right] \right)^{\frac{1}{2}} \tag{83}$$

Then, $\mathcal{N}[0, T]$ is a Banach space under norm (83).

In this chapter, we present another method. It will give a sufficient condition for the unique solvability of (3). We will obtain a Riccati type equation and a BSDE associated with (3). Let us now carry heuristic derivation.

Suppose $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0, T]$ is an adapted solution of (3). We assume that $\mathbf{X}$ and $\mathbf{Y}$ are related by

$$\mathbf{Y}_t = P(t)\mathbf{X}_t + \mathbf{p}_t, \forall t \in [0, T], \text{ a.s.} \tag{84}$$

where $P : [0, T] \to \mathbb{R}^{m\times n}$ is a deterministic matrix-valued function and $\mathbf{p} : [0, T]\times\Omega \to \mathbb{R}^m$ is an $\{\mathscr{F}_t\}_{t\ge 0}$-adapted process. We are going to derive the equations for P and $\mathbf{p}$. First of all, from (8) and the terminal condition in (3), we have

$$\mathbf{g} = P(T)\mathbf{X}_T + \mathbf{p}_T. \tag{85}$$

Let us impose

$$P(T) = O, \mathbf{p}_T = \mathbf{g}. \tag{86}$$

Since $\mathbf{g} \in L^2_{\mathscr{F}_T}(\Omega; \mathbb{R}^m)$ and $\mathbf{p}$ is required to be $\{\mathscr{F}_t\}_{t\geq 0}$-adapted, we should assume that $\mathbf{p}$ satisfies a BSDE:

$$\begin{cases} d\mathbf{p}_t = \alpha_t dt + \sum_{i=1}^{d} \mathbf{q}_t^i dW_t^i, 0 \leq t \leq T \\ \mathbf{p}_T = \mathbf{g} \end{cases} \tag{87}$$

with α, $\mathbf{q}^1, \cdots, \mathbf{q}^d \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^m)$ being undetermined. Next, by Itô Formula, we have

$$\begin{aligned} d\mathbf{Y}_t =& \dot{P}(t)\mathbf{X}_t dt + P(t)d\mathbf{X}_t + d\mathbf{p}_t \\ =& [\dot{P}(t)\mathbf{X}_t + P(t)(A\mathbf{X}_t + B\mathbf{Y}_t) + \alpha_t]dt \\ &+ \sum_{i=1}^{d} [P(t)(A_1^i \mathbf{X}_t + B_1^i \mathbf{Y}_t + C_1^i \mathbf{Z}_t^i) + \mathbf{q}_t^i] dW_t^i \\ =& \{[\dot{P}(t) + P(t)A + P(t)BP(t)]\mathbf{X}_t + P(t)B\mathbf{p}_t + \alpha_t\}dt \\ &+ \sum_{i=1}^{d} \{[P(t)A_1^i + P(t)B_1^i P(t)]\mathbf{X}_t + P(t)C_1^i \mathbf{Z}_t^i + P(t)B_1^i \mathbf{p}_t + \mathbf{q}_t^i\} dW_t^i \end{aligned} \tag{88}$$

Now compare (88) with the second equation in (3) (note (84)), we obtain that (drift coefficient)

$$[\dot{P}(t) + P(t)A + P(t)BP(t)]\mathbf{X}_t + P(t)B\mathbf{p}_t + \alpha_t = [\widehat{A} + \widehat{B}P(t)]\mathbf{X}_t + \widehat{B}\mathbf{p}_t, \tag{89}$$

and (diffusion coefficient)

$$[P(t)A_1^i + P(t)B_1^i P(t)]\mathbf{X}_t + P(t)C_1^i \mathbf{Z}_t^i + P(t)B_1^i \mathbf{p}_t + \mathbf{q}_t^i = \mathbf{Z}_t^i, i = 1, \cdots, d \tag{90}$$

By assuming $I - P(t)C_1^i$ to be invertible $\forall t \in [0, T], i = 1, \cdots, d$, we have from (90) that

$$\mathbf{Z}_t^i = [I - P(t)C_1^i]^{-1}\{[(P(t)A_1^i + P(t)B_1^i P(t)]\mathbf{X}_t + P(t)B_1^i \mathbf{p}_t + \mathbf{q}_t^i\}, i = 1, \cdots, d \tag{91}$$

Then, (89) can be written as

$$0 = [\dot{P}(t) + P(t)A + P(t)BP(t) - \widehat{A} - \widehat{B}P(t)]\mathbf{X}_t + [P(t)B - \widehat{B}]\mathbf{p}_t + \alpha_t. \tag{92}$$

Now, we introduce the following differential equation for $M_{m\times n}(\mathbb{R})$-valued function P:

$$\begin{cases} \dot{P}(t) + P(t)A + P(t)BP(t) - \widehat{A} - \widehat{B}P(t) = O, 0 \leq t \leq T \\ P(T) = O. \end{cases} \tag{93}$$

We refer to (93) as a *Riccati type equation*. Suppose (93) admits a solution P over $[0,T]$ such that

$$[I - P(t)C_1^i]^{-1} \text{ is bounded }, i = 1, \cdots, d, \forall t \in [0,T]. \tag{94}$$

Then, (92) gives

$$\alpha_t = [\widehat{B} - P(t)B]\mathbf{p}_t$$

Combining this with (87), we see that one should introduce the following BSDE:

$$\begin{cases} d\mathbf{p}_t = [\widehat{B} - P(t)B]\mathbf{p}_t dt + \sum_{i=1}^{d} \mathbf{q}_t^i dW_t^i, 0 \le t \le T \\ \mathbf{p}_T = \mathbf{g}. \end{cases} \tag{95}$$

When (93) admits solution P such that (94) holds, BSDE (95) admits a unique solution $(\mathbf{p}, \mathbf{q}^1, \cdots, \mathbf{q}^d) \in \mathcal{N}[0,T]$. In the form provided here, a proof can be found in several sources, for instance Ma & Yong (2000), pp. 15-16. From (84) and (91), the forward equation ($\mathbf{X}$):

$$\begin{cases} d\mathbf{X}_t = \{[A + BP(t)]\mathbf{X}_t + B\mathbf{p}_t\}dt \\ \qquad + \sum_{i=1}^{d} \Big(\{A_1^i + B_1^i P(t) + C_1^i[I - P(t)C_1^i]^{-1}[P(t)A_1^i + P(t)B_1^i P(t)]\}\mathbf{X}_t \\ \qquad + B_1^i \mathbf{p}_t + C_1^i[I - P(t)C_1^i]^{-1}[P(t)B_1^i \mathbf{p}_t + \mathbf{q}_t^i]\Big) dW_t^i, 0 \le t \le T \\ \mathbf{X}_0 = \mathbf{0} \end{cases}$$

Then we can define the following:

$$\begin{cases} \widetilde{A}(t) = A + BP(t), \\ \widetilde{A}_1^i(t) = A_1^i + B_1^i P(t) + C_1^i[I - P(t)C_1^i]^{-1}[P(t)A_1^i + P(t)B_1^i P(t)], i = 1, \cdots, d \\ \widetilde{\mathbf{b}}_t = B\mathbf{p}_t, \\ \widetilde{\sigma}_t^i = B_1^i \mathbf{p}_t + C_1^i[I - P(t)C_1^i]^{-1}[P(t)B_1^i \mathbf{p}_t + \mathbf{q}_t^i], i = 1, \cdots, d. \end{cases} \tag{96}$$

It is clear that $\widetilde{A}$ and $\widetilde{A}_1^i$ are time-dependent matrix-valued function and $\widetilde{\mathbf{b}}$ and $\widetilde{\sigma}^i$ are $\{\mathscr{F}_t\}_{t \ge 0}$-adapted processes. Further, under (94), by the Existence and Uniqueness Theorem for Stochastic Differential Equations, the following SDE admits a unique (strong) solution:

$$\begin{cases} d\mathbf{X}_t = [\widetilde{A}(t)\mathbf{X}_t + \widetilde{\mathbf{b}}_t]dt + \sum_{i=1}^{d} [\widetilde{A}_1^i(t)\mathbf{X}_t + \widetilde{\sigma}_t^i] dW_t^i, 0 \le t \le T, \\ \mathbf{X}_0 = \mathbf{0}. \end{cases} \tag{97}$$

The following Theorem gives a representation of the adapted solution of FBSDE (3).

Theorem 4. *Let (93) admits a solution P such that (94) holds. Then FBSDE (3) admits a unique solution $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0,T]$ which is determined by (97), (84) and (91).*

Proof. First of all, a direct computation (from above) shows that the process $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d)$ determined by (97), (84) and (91) is an adapted solution of (3). We now prove the uniqueness. Let $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d) \in \mathcal{M}[0,T]$ be any adapted solution of (3). Set

$$\begin{cases} \overline{\mathbf{Y}}_t = P(t)\mathbf{X}_t + \mathbf{p}_t, \\ \overline{\mathbf{Z}}_t^i = [I - P(t)C_1^i]^{-1}\{[P(t)A_1^i + P(t)B_1^i P(t)]\mathbf{X}_t + P(t)B_1^i\mathbf{p}_t + \mathbf{q}_t^i\}, i = 1, \cdots, d \end{cases} \tag{98}$$

where P and $(\mathbf{p}, \mathbf{q}^1, \cdots, \mathbf{q}^d)$ are (adapted) solutions of (93) and (95), respectively. Denote $\widehat{\mathbf{Y}} = \mathbf{Y} - \overline{\mathbf{Y}}$ and $\widehat{\mathbf{Z}}^i = \mathbf{Z}^i - \overline{\mathbf{Z}}^i$. By the Itô Formula,

$$\begin{aligned} d\overline{\mathbf{Y}}_t &= \dot{P}(t)\mathbf{X}_t dt + P(t)d\mathbf{X}_t + d\mathbf{p}_t \\ &= [\widehat{A} + \widehat{B}P(t) - P(t)A - P(t)BP(t)]\mathbf{X}_t dt \\ &\quad + P(t)\left[(A\mathbf{X}_t + B\mathbf{Y}_t)dt + \sum_{i=1}^d (A_1^i\mathbf{X}_t + B_1^i\mathbf{Y}_t + C_1^i\mathbf{Z}_t^i)dW_t^i\right] \\ &\quad + [\widehat{B} - P(t)B]\mathbf{p}_t dt + \sum_{i=1}^d \mathbf{q}_t^i dW_t^i \\ &= [\widehat{A}\mathbf{X}_t + P(t)B(\mathbf{Y}_t - \overline{\mathbf{Y}}_t) + \widehat{B}\overline{\mathbf{Y}}_t]dt \\ &\quad + \sum_{i=1}^d \{P(t)B_1^i(\mathbf{Y}_t - \overline{\mathbf{Y}}_t) + P(t)B_1^i[P(t)\mathbf{X}_t + \mathbf{p}_t] + P(t)(A_1^i\mathbf{X}_t + C_1^i\mathbf{Z}_t^i) + \mathbf{q}_t^i\}dW_t^i \\ &= [\widehat{A}\mathbf{X}_t + P(t)B(\mathbf{Y}_t - \overline{\mathbf{Y}}_t) + \widehat{B}\overline{\mathbf{Y}}_t]dt + \sum_{i=1}^d \{P(t)B_1^i(\mathbf{Y}_t - \overline{\mathbf{Y}}_t) + P(t)C_1^i\mathbf{Z}_t^i + [I - P(t)C_1^i]\overline{\mathbf{Z}}_t^i\}dW_t^i \\ &= [\widehat{A}\mathbf{X}_t + P(t)B(\mathbf{Y}_t - \overline{\mathbf{Y}}_t) + \widehat{B}\overline{\mathbf{Y}}_t]dt + \sum_{i=1}^d \{P(t)B_1^i(\mathbf{Y}_t - \overline{\mathbf{Y}}_t) + \overline{\mathbf{Z}}_t^i + P(t)C_1^i(\mathbf{Z}_t^i - \overline{\mathbf{Z}}_t^i)\}dW_t^i \end{aligned}$$

Then a direct computation shows that (compare to (3))

$$\begin{cases} d\widehat{\mathbf{Y}}_t = [\widehat{B} - P(t)B]\widehat{\mathbf{Y}}_t dt + \sum_{i=1}^d \{[I - P(t)C_1^i]\widehat{\mathbf{Z}}_t^i - P(t)B_1^i\widehat{\mathbf{Y}}_t\}dW_t^i, \\ \widehat{\mathbf{Y}}_T = \mathbf{0}. \end{cases} \tag{99}$$

By (94), We may set

$$\widetilde{\mathbf{Z}}_t^i = [I - P(t)C_1^i]\widehat{\mathbf{Z}}_t^i - P(t)B_1^i\widehat{\mathbf{Y}}_t, i = 1, \cdots, d \tag{100}$$

to get the following equivalent BSDE (of (99)):

$$\begin{cases} d\widehat{\mathbf{Y}}_t = [\widehat{B} - P(t)B]\widehat{\mathbf{Y}}_t dt + \sum_{i=1}^{d} \widetilde{\mathbf{Z}}_t^i dW_t^i, \\ \widehat{\mathbf{Y}}_T = \mathbf{0}. \end{cases} \tag{101}$$

It is clear that such a BSDE admits a unique adapted solution $(\widehat{\mathbf{Y}}, \widetilde{\mathbf{Z}}^1, \cdots, \widetilde{\mathbf{Z}}^d) \equiv (\mathbf{0}, \mathbf{0}, \cdots, \mathbf{0})$ a.s. (see Ma & Yong (2000), pp. 15-16). Consequently, $\widehat{\mathbf{Z}}^i \equiv \mathbf{0}$ a.s., $i = 1, \cdots, d$ (since $\widehat{\mathbf{Z}}_t^i = [I - P(t)C_1^i]^{-1}[\widetilde{\mathbf{Z}}_t^i + P(t)B_1^i\widehat{\mathbf{Y}}_t], i = 1, \cdots, d$). Hence, by (98), we obtain

$$\begin{cases} \mathbf{Y}_t = P(t)\mathbf{X}_t + \mathbf{p}_t, \\ \mathbf{Z}_t^i = [I - P(t)C_1^i]^{-1}\{[P(t)A_1^i + P(t)B_1^i P(t)]\mathbf{X}_t + P(t)B_1^i \mathbf{p}_t + \mathbf{q}_t^i\}, i = 1, \cdots, d. \end{cases} \tag{102}$$

This means that any adapted solution $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}^1, \cdots, \mathbf{Z}^d)$ of (3) must satisfy (102). Then, similar to the heuristic derivation above, we have that $\mathbf{X}$ has to be the solution of (97). Hence, we obtain the uniqueness. □

The following result tells us something more.

Proposition 2. *Let (93) admits a solution P such that (94) holds for $t \in [T_0, T]$ (with some $T_0 \geq 0$). Then, $\forall \widetilde{T} \in [0, T - T_0]$, linear FBSDE (3) is uniquely solvable on $[0, \widetilde{T}]$.*

Proof. Let

$$\widetilde{P}(t) = P(t + T - \widetilde{T}), 0 \leq t \leq \widetilde{T}. \tag{103}$$

Then $\widetilde{P}$ satisfies (93) with $[0, T]$ replaced by $[0, \widetilde{T}]$ and

$$[I - \widetilde{P}(t)C_1^i]^{-1} \text{ is bounded for } 0 \leq t \leq \widetilde{T}, i = 1, \cdots, d \tag{104}$$

Thus, Theorem 4.1 applies. □

The above proposition tells that if (93) admits a solution P satisfying (94), FBSDE (3) is uniquely solvable over any $[0, \widetilde{T}]$ (with $\widetilde{T} \leq T$). Then in this case, by Theorem 1, the solvability (3) of FBSDE over $[0, \widetilde{T}]$ admits a solution $\forall \mathbf{g} \in L^2_{\mathscr{F}_{\widetilde{T}}}(\Omega; \mathbb{R}^m)$, of which a necessary condition is

$$\det\left(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}t} \begin{bmatrix} O \\ I \end{bmatrix} \right) > 0, 0 \leq t \leq T. \tag{105}$$

Therefore, by Theorem 3, compare (105) and (29), we see that the solvability of Riccati type equation (93) is *only* sufficient condition for the solvability of (3).

We have seen that (105) is necessary condition for (93) having a solution P satisfying (94). The following result gives the inverse of this.

Theorem 5. *Let (105) hold. Then (93) admits a unique solution P which has the following representation:*

$$P(t) = -\left\{ [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} O \\ I \end{bmatrix} \right\}^{-1} [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} I \\ O \end{bmatrix}, 0 \leq t \leq T. \qquad (106)$$

Moreover, it holds

$$I - P(t)C_1^i = \left\{ [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} O \\ I \end{bmatrix} \right\}^{-1} [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} C_1^i \\ I \end{bmatrix}, 0 \leq t \leq T, i = 1, \cdots, d. \qquad (107)$$

Consequently, if in addition to (105), (30) holds, then (94) holds and the linear FBSDE (3) is uniquely solvable with the representation given by (97), (84) and (91).

Proof. We can easily check that (106) is a solution of (93). You can find in Ma & Yong (2000), pp. 49-50.

Uniqueness is obvious since (93) is a terminal value problem with the right hand side of the equation being locally Lipschitz.

$$\begin{aligned} I - P(t)C_1^i &= I + \left\{ [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} O \\ I \end{bmatrix} \right\}^{-1} [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} C_1^i \\ O \end{bmatrix} \\ &= \left\{ [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} O \\ I \end{bmatrix} \right\}^{-1} [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \left(\begin{bmatrix} O \\ I \end{bmatrix} + \begin{bmatrix} C_1^i \\ O \end{bmatrix} \right) \\ &= \left\{ [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} O \\ I \end{bmatrix} \right\}^{-1} [\ O \quad I\ ] e^{\mathcal{A}(T-t)} \begin{bmatrix} C_1^i \\ I \end{bmatrix}. \end{aligned}$$

Finally, an easy calculation shows (by (105), (30))

$$\det(I - P(t)C_1^i) = \frac{\det\left(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}(T-t)}\mathcal{C}_1^i\right)}{\det\left(\begin{bmatrix} O & I \end{bmatrix} e^{\mathcal{A}(T-t)} \begin{bmatrix} O \\ I \end{bmatrix}\right)} > 0, \forall t \in [0,T], i = 1, \cdots, d,$$

and $I - P(t)C_1^i$ is a continuous function on $[0,T]$, $\forall i = 1, \cdots, d$, hence (94) holds. Then we complete proof. □

5. Conclusion

This study proposed some extension of Ma & Yong (2000). We give the sufficient and necessary conditions of the linear FBSDE (3), and modify their work to prove the closeness of $R(\mathcal{K})$ (with $\widehat{A} = O$). Then we find the connection between a Riccati type equation and the linear FBSDE.

There are at least three more possible extensions of our method for future research. First, we can add nonzero $\mathbf{Z}_i$s term in drift coefficient and derive the sufficient and necessary conditions. Second, one can prove or give a counterexample of the closeness in general case (not just special case in this study). Third, someone can consider some special nonlinear cases (like quadratic form) and discuss the solvability of the FBSDE. This work may be can apply in the mathematical finance like mean-variance portfolio selection and consider in optimal control and LQ problem (like Zhou and Li (2000)) in the future.

Bibliography

Bjork, T. (2009). Arbitrage theory in continuous time (3rd ed.). New York: Oxford.

Karatzas, I., & Shreve, S., E. (1998). Brownian motion and stochastic calculus (2nd ed.). New York: Springer.

Karoui, N. E., Peng, S., & Quenez, M. C. (1997). Backward stochastic differential equations in finance. Mathematical Finance, Vol. 7., No. 1, 1-71.

Lax, P. D., (2002). Functional Analysis. Wiley.

Ma, J., & Yong, J. (2000). Forward-backward stochastic equations and their applications. Springer.

Oksendal, B. (2003). Stochastic differential equations (6th ed.) Springer.

Pham, H. (2000). Continuous-time stochastic control and optimization with Financial Application. Springer.

Zhou, X. Y., & Li, D. (2000). Continuous-time mean-variance portfolio selection: a stochastic LQ framework, Applied Mathematics Optimization, 19-33.

Appendix

Lemma 6. *Consider the following SDE:*

$$\begin{cases} d\mathbf{X}_t = (A_t\mathbf{X}_t + B_t\mathbf{Y}_t)dt + \sum_{i=1}^{d}(A_{1,t}^i\mathbf{X}_t + B_{1,t}^i\mathbf{Y}_t + C_{1,t}^i\mathbf{Z}_t^i)dW_t^i, \\ d\mathbf{Y}_t = \widehat{B}_t\mathbf{Y}_t dt + \sum_{i=1}^{d}\mathbf{Z}_t^i dW_t^i, & 0 \le t \le T, \\ \mathbf{X}_0 = \mathbf{0}, \mathbf{Y}_0 = \mathbf{0}, \end{cases}$$

where

$$\begin{cases} A, A_1^i : [0,T]\times\Omega \to M_n(\mathbb{R}) \\ B, B_1^i, C_1^i : [0,T]\times\Omega \to M_{n\times m}(\mathbb{R}) \\ \widehat{B} : [0,T]\times\Omega \to M_m(\mathbb{R}) \end{cases}$$

all the processes are $\{\mathscr{F}_t\}_{t\ge 0}$*-progressively measurable,* $\exists M \ge 0$ *with* $\|A_t\|, \|A_{1,t}^i\|, \|B_t\|, \|B_{1,t}^i\|, \|C_{1,t}^i\| \le M$*,* $i = 1\cdots, d$*, a.e.* $t \in [0,T]$*, a.s. And* $\exists\widehat{b} > 0$ *such that* $\|\widehat{B}_t\| > \widehat{b}$ *a.e.* $t \in [0,T]$*, a.s.*
Then there exists $c > 0$ *(independent of* $(\mathbf{Z}_1, \cdots, \mathbf{Z}^d)$*),* $\forall(\mathbf{Z}_1, \cdots, \mathbf{Z}^d) \in \mathcal{H}$ *such that*

$$E[|\mathbf{X}_T|^2] \le cE[|\mathbf{Y}_T|^2].$$

Proof. By the Itô Formula,

$$\begin{aligned} d|\mathbf{Y}_t|^2 &= 2\mathbf{Y}_t\cdot d\mathbf{Y}_t + d\langle\mathbf{Y}\rangle_t \\ &= \left(2\mathbf{Y}_t\cdot\widehat{B}_t\mathbf{Y}_t + \sum_{i=1}^{d}|\mathbf{Z}_t^i|^2\right)dt + 2\sum_{i=1}^{d}\mathbf{Y}_t\cdot\mathbf{Z}_t^i dW_t^i, \end{aligned}$$

hence,

$$\begin{aligned} E[|\mathbf{Y}_t|^2] &= E\left[\int_0^t 2\mathbf{Y}_s\cdot\widehat{B}_s\mathbf{Y}_s + \sum_{i=1}^{d}|\mathbf{Z}_s^i|^2 ds\right] \\ &\ge 2\widehat{b}E\left[\int_0^t|\mathbf{Y}_s|^2 ds\right] + \sum_{i=1}^{d}E\left[\int_0^t|\mathbf{Z}_s^i|^2 ds\right]. \end{aligned} \tag{108}$$

Similarly, by the Itô Formula,

$$\begin{aligned} d|\mathbf{X}_t|^2 &= 2\mathbf{X}_t\cdot d\mathbf{X}_t + d\langle\mathbf{X}\rangle_t \\ &= \left[2\mathbf{X}_t\cdot(A_t\mathbf{X}_t + B_t\mathbf{Y}_t) + \sum_{i=1}^{d}|A_{1,t}^i\mathbf{X}_t + B_{1,t}^i\mathbf{Y}_t + C_{1,t}^i\mathbf{Z}_t^i|^2\right]dt \\ &\quad + 2\sum_{i=1}^{d}\mathbf{X}_t\cdot(A_{1,t}^i\mathbf{X}_t + B_{1,t}^i\mathbf{Y}_t + C_{1,t}^i\mathbf{Z}_t^i)dW_t^i, \end{aligned}$$

taking expectation we get and use the inequality $\alpha a^2 + \frac{1}{\alpha} b^2 \geq 2|ab|\ \forall \alpha > 0$,

$$
\begin{aligned}
E[|\mathbf{X}_t|^2] &= E\left[\int_0^t 2\mathbf{X}_s \cdot (A_s\mathbf{X}_s + B_s\mathbf{Y}_s) + \sum_{i=1}^d |A_{1,s}^i\mathbf{X}_s + B_{1,s}^i\mathbf{Y}_s + C_{1,s}^i\mathbf{Z}_s^i|^2 ds\right] \\
&\leq 2ME\left[\int_0^t |\mathbf{X}_s|^2 ds\right] + \alpha M^2 E\left[\int_0^t |\mathbf{X}_s|^2 ds\right] + \frac{1}{\alpha} E\left[\int_0^t |\mathbf{Y}_s|^2 ds\right] \\
&\quad + 3M^2 \sum_{i=1}^d E\left[\int_0^t |\mathbf{X}_s|^2 + |\mathbf{Y}_s|^2 + |\mathbf{Z}_s^i|^2 ds\right] \\
&\leq M(2 + \alpha M + 3Md) E\left[\int_0^t |\mathbf{X}_s|^2 ds\right] + \left(\frac{1}{\alpha} + 3M^2 d\right) E\left[\int_0^t |\mathbf{Y}_s|^2 ds\right] \\
&\quad + 3M^2 \sum_{i=1}^d E\left[\int_0^t |\mathbf{Z}_s^i|^2 ds\right], \forall \alpha > 0.
\end{aligned}
$$

By Gronwall Inequality and (108),

$$
\begin{aligned}
E[|\mathbf{X}_t|^2] &\leq e^{M(2+\alpha M+3Md)t} \left\{ \left(\frac{1}{\alpha} + 3M^2 d\right) E\left[\int_0^t |\mathbf{Y}_s|^2 ds\right] + 3M^2 \sum_{i=1}^d E\left[\int_0^t |\mathbf{Z}_s^i|^2 ds\right] \right\} \\
&\leq e^{M(2+\alpha M+3Md)t} \left\{ \left(\frac{1}{\alpha} + 3M^2 d\right) \frac{1}{2\widehat{b}} \left(E\left[|\mathbf{Y}_t|^2\right] - \sum_{i=1}^d E\left[\int_0^t |\mathbf{Z}_s^i|^2 ds\right] \right) \right. \\
&\quad \left. + 3M^2 \sum_{i=1}^d E\left[\int_0^t |\mathbf{Z}_s^i|^2 ds\right] \right\} \\
&= e^{M(2+\alpha M+3Md)t} \left\{ \frac{1}{2\widehat{b}} \left(\frac{1}{\alpha} + 3M^2 d\right) E\left[|\mathbf{Y}_t|^2\right] \right. \\
&\quad \left. + \left[3M^2 - \frac{1}{2\widehat{b}} \left(\frac{1}{\alpha} + 3M^2 d\right)\right] \sum_{i=1}^d E\left[\int_0^t |\mathbf{Z}_s^i|^2 ds\right] \right\}, \forall \alpha > 0, t \in [0, T].
\end{aligned}
$$

Take α sufficient small, we obtain

$$
E[|\mathbf{X}_t|^2] \leq e^{M(2+\alpha M+3Md)t} \frac{1}{2\widehat{b}} \left(\frac{1}{\alpha} + 3M^2 d\right) E\left[|\mathbf{Y}_t|^2\right], \forall t \in [0, T].
$$

let $c = e^{M(2+\alpha M+3Md)T} \frac{1}{2\widehat{b}} \left(\frac{1}{\alpha} + 3M^2 d\right)$, we complete the argument. □

The following case is that we want to show in the Chapter 3.

Lemma 7. *Consider the following SDE:*

$$
\begin{cases}
d\mathbf{X}_t = (A\mathbf{X}_t + B\mathbf{Y}_t)dt + \sum_{i=1}^d (A_1^i\mathbf{X}_t + B_1^i\mathbf{Y}_t + C_1^i\mathbf{Z}_t^i)dW_t^i, \\
d\mathbf{Y}_t = \widehat{B}\mathbf{Y}_t dt + \sum_{i=1}^d \mathbf{Z}_t^i dW_t^i, & 0 \leq t \leq T, \\
\mathbf{X}_0 = \mathbf{0}, \mathbf{Y}_0 = \mathbf{0},
\end{cases}
$$

the matrices are defined by (4). Then there exists $C > 0$ *(independent of* $(\mathbf{Z}_1, \cdots, \mathbf{Z}^d)$*),* $\forall (\mathbf{Z}_1, \cdots, \mathbf{Z}^d) \in \mathcal{H}$ *such that*

$$E[|\mathbf{X}_T|^2] \leq CE[|\mathbf{Y}_T|^2]. \tag{109}$$

Proof. We use the General Itô Formula in $e^{-\widehat{B}t}\mathbf{Y}_t$,

$$d(e^{-\widehat{B}t}\mathbf{Y}_t) = -\widehat{B}e^{-\widehat{B}t}\mathbf{Y}_t dt + e^{-\widehat{B}t}d\mathbf{Y}_t = e^{-\widehat{B}t}\sum_{i=1}^{d}\mathbf{Z}_t^i dW_t^i.$$

Now, we consider $e^t e^{-\widehat{B}t}\mathbf{Y}_t$ and use the Itô Formula again,

$$\begin{aligned} d(e^t e^{-\widehat{B}t}\mathbf{Y}_t) &= e^t e^{-\widehat{B}t}\mathbf{Y}_t dt + e^t d(e^{-\widehat{B}t}\mathbf{Y}_t) \\ &= e^t e^{-\widehat{B}t}\mathbf{Y}_t dt + e^t e^{-\widehat{B}t}\sum_{i=1}^{d}\mathbf{Z}_t^i dW_t^i \end{aligned}$$

Let $\widehat{\mathbf{Y}}_t = e^t e^{-\widehat{B}t}\mathbf{Y}_t$, and $\widehat{\mathbf{Z}}_t^i = e^t e^{-\widehat{B}t}\mathbf{Z}_t^i$, we derive the SDE:

$$\begin{cases} d\mathbf{X}_t = (A\mathbf{X}_t + Be^{-t}e^{\widehat{B}t}\widehat{\mathbf{Y}}_t)dt + \sum_{i=1}^{d}(A_1^i\mathbf{X}_t + B_1^i e^{-t}e^{\widehat{B}t}\widehat{\mathbf{Y}}_t + C_1^i e^{-t}e^{\widehat{B}t}\widehat{\mathbf{Z}}_t^i)dW_t^i, \\ d\widehat{\mathbf{Y}}_t = \widehat{\mathbf{Y}}_t dt + \sum_{i=1}^{d}\widehat{\mathbf{Z}}_t^i dW_t^i, & 0 \leq t \leq T, \\ \mathbf{X}_0 = \mathbf{0}, \widehat{\mathbf{Y}}_0 = \mathbf{0}, \end{cases}$$

by Lemma 6, $\exists c > 0$ such that

$$E[|\mathbf{X}_T|^2] \leq cE[|\widehat{\mathbf{Y}}_T|^2] \leq ce^{2T}e^{2\|\widehat{B}\|T}E[|\mathbf{Y}_T|^2]$$

Take $C = ce^{2T}e^{2\|\widehat{B}\|T}$, we prove (109). □