University of Puerto Rico
Río Piedras Campus
Faculty of Natural Sciences
Department of Mathematics

Markovian switching systems: conditional McKean-Vlasov backward and forward-backward equations and their applications

By

Esteban J. Rolón Gutiérrez

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN MATHEMATICS AT THE UNIVERSITY OF PUERTO RICO, RÍO PIEDRAS CAMPUS

January 12, 2024

APPROVED BY THE DOCTORAL DISSERTATION COMMITTEE
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN MATHEMATICS
AT THE UNIVERSITY OF PUERTO RICO

ADVISOR:

Son L. Nguyen, Ph.D
Florida Institute of Technology

READERS:

George Yin, Ph.D
University of Connecticut

María E. Pérez, Ph.D
University of Puerto Rico, Río Piedras

Luis R. Pericchi, Ph.D
University of Puerto Rico, Río Piedras

Alejandro Vélez, Ph.D
University of Puerto Rico, Río Piedras

Abstract of Ph.D Dissertation Presented to the Graduate School
of the University of Puerto Rico, Río Piedras Campus in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in Mathematics

Markovian switching systems: conditional McKean-Vlasov backward and forward-backward equations and their applications

By

Esteban J. Rolón Gutiérrez

January 2024

Chair: Son L. Nguyen Ph.D.
Department: Mathematics

In recent years, forward-backward stochastic differential equations (FBSDE) have been extensively studied because of their numerous applications in many areas such as control and game theory, mathematical economics, and mathematical finance. Due to the pressing need of treating large-scale systems, there has been increasing effort of dealing with mean-field interactions, systems with mean-field interactions, and related control problems, and games. To deal with large-scale switching systems, the mean-field types of FBSDEs with Markovian switching naturally come into play when one needs to treat the mean-field control problems. In this work we derive useful estimates for the solutions of the backward stochastic differential equations (BSDE) with Markovian switching. We also work on the FBSDEs with regime-switching and FBSDEs with mean-field and regime-switching, providing sufficient conditions for the existence and uniqueness of the solutions. Then we consider a nonzero-sum game problem with N players in which the dynamics and cost functionals of each player depend on conditional mean-field terms and a regime-switching process, presenting conditions on the coefficients such that a Nash equilibrium point of the differential game exists and the

relationship of the existence of the Nash equilibrium point and the solution of the conditional mean-field FBSDE with regime switching.

ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Son Luu Nguyen for his guidance during the term of my PhD's degree at the University of Puerto Rico, Río Piedras Campus. Without his valuable supervision, this work would not have been completed.

I would also like to express my appreciation to Dr. María Eglée Pérez for helping during this process and to Dr. George Yin for his helps with my first paper.

I am grateful to Dr. Alejandro Vélez-Santiago, Dr. Lin Shan, and Dr. Luis Pericchi for their helps and for serving in my thesis committee.

I am in debt with the Department of Mathematics, University of Puerto Rico, Río Piedras Campus for the great opportunity to complete my PhD degree.

Finally, I am thankful to my friend Isaías Justinianes and my family for their constant supports.

The research of Esteban Rolón Gutiérrez was supported (as a student) in part by the Puerto Rico Science, Technology and Research Trust under Agreement Number 2022-00072. This content is only the responsibility of the authors and does not necessarily represent the official views of the Puerto Rico Science, Technology and Research Trust.

TABLE OF CONTENTS

page

LIST OF ABBREVIATIONS AND SYMBOLS

$\mathcal{F}$	A σ-algebra on a given set Ω.
$(\Omega, \mathcal{F})$	A measurable space.
$\mathbb{P}$	Probability measure
μ	$= \mathbb{P}_{(X \big\vert \mathcal{F})}$
ν	$= \mathbb{P}_{(X,Y \big\vert \mathcal{F})}$
$(\Omega, \mathcal{F}, \mathbb{P})$	A probability space.
$\mathcal{B}(\mathbb{R}^d)$	The Borel σ-algebra.
B	Is a Borel set.
$\{\mathcal{F}_t\}_{t\geqslant 0}$	A filtration on $(\Omega, \mathcal{F})$ at time t.
$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})_{t\geqslant 0})$	A filtered probability space.
$\\|\psi\\|_2^2$	$= \mathbb{E}\left[\int_0^T \lvert\psi_t\rvert^2 dt\right]$
$\mathcal{L}^p(\mathbb{R}^d)$	$= \{\xi : \Omega \to \mathbb{R}^d, \mathcal{F}\text{-measurable}, \mathbb{E}\lvert\xi\rvert^p < \infty\}, \quad p\geqslant 1$
$\mathcal{S}^2(0,T;\mathbb{R}^d)$	$= \left\{\varphi : [0,T] \times \Omega \to \mathbb{R}^d,\ \mathcal{F}\text{-adapted càdlàg process}, \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} \lvert\varphi_t\rvert^2\right] < \infty\right\}$
$\mathcal{L}^0(0,T;\mathbb{R}^d)$	$= \left\{\psi : [0,T] \times \Omega \to \mathbb{R}^d,\ \mathcal{F}\text{-progressively measurable process}\right\}$
$\mathcal{L}^2(0,T;\mathbb{R}^d)$	$= \left\{\psi \in \mathcal{L}^0(0,T;\mathbb{R}^d) : \\|\psi\\|_2^2 = \mathbb{E}\left[\int_0^T \lvert\psi_t\rvert^2 dt\right] < \infty\right\}$
$C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+; \mathbb{R})$	The family of all real-valued functions $V(x,t)$ such that it is twice continuously differentiable in x and once in t.
W_t	Brownian motion at time t.
$\mathcal{M}$	$= \{1, 2, \ldots, m_0\}$ is the state space of the markov chain.
Q	A generator matrix
α_t	A continuous-time Markov chain with state space $\mathcal{M}$ and generator matrix Q.
$\mathcal{M}^2(0,T;\mathbb{R}^d)$	$= \Big\{\lambda = \big(\lambda_{i_0 j_0} : i_0, j_0 \in \mathcal{M}\big)$ such that $\lambda_{i_0 j_0} \in \mathcal{L}^0(0,T;\mathbb{R}^d)$, $\lambda_{i_0 i_0} \equiv 0$, and $\sum_{i_0,j_0\in\mathcal{M}} \mathbb{E}\int_0^T \big\lvert\lambda_{i_0 j_0}(t)\big\rvert^2 d\big[M_{i_0 j_0}\big](t) < \infty\Big\}$
$\int_0^t \lambda_s \bullet dM_s$	$=\sum_{i_0,j_0\in\mathcal{M}} \int_0^t \lambda_{i_0 j_0}(s) dM_{i_0 j_0}(s)$
$\lambda_t \bullet dM_t$	$= \sum_{i_0,j_0\in\mathcal{M}} \lambda_{i_0 j_0}(t) dM_{i_0 j_0}(t)$
SDE	Stochastic Differential Equation
BSDE	Backward Stochastic Differential Equation
FBSDE	Forward-Backward Stochastic Differential Equation
$\mathcal{P}(\mathbb{R}^d)$	The set of all probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.
$\mathcal{P}_p(\mathbb{R}^d)$	$= \{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} \lvert x\rvert^p \mu(dx) < \infty\}$
$[X]_t$	The quadratic variation of X at time t.
$[X,Y]_t$	The covariation of X and Y at time t.
$\mathbb{1}$	The usual zero-one indicator function.

$\big[M_{i_0j_0}\big](t)$	$= \sum_{0\leqslant s\leqslant t} \mathbb{1}\big(\alpha(s_-) = i_0\big)\mathbb{1}\big(\alpha(s) = j_0\big)$
$\big\langle M_{i_0j_0}\big\rangle(t)$	$= \int_0^t q_{i_0j_0}\mathbb{1}\big(\alpha(s_-) = i_0\big)ds$
$W_2(\cdot,\cdot)$	The 2-Wasserstein distance
θ	$= (x, y, z)$
Θ	$= (X, Y, Z)$
$\\|\theta_1 - \theta_2\\|$	$= \big\|x_1 - x_2\big\| + \big\|y_1 - y_2\big\| + \big\\|z_1 - z_2\big\\|$
$\langle x, y\rangle$	The dot product of x and y
$[A, B]$	$= \sum_{i=1}^d \langle A_i, B_i\rangle$, where A_i and B_i, $i = 1, 2, \ldots, d$, are the i-th column of $d \times d$ matrices A and B.
$\Psi\big(t, \theta_1, \theta_2, i_0\big)$	$= \big\langle f\big(t, \theta_1, i_0\big) - f\big(t, \theta_2, i_0\big), y_1 - y_2\big\rangle + \big\langle g\big(t, \theta_1, i_0\big) - g\big(t, \theta_2, i_0\big).x_1 - x_2\big\rangle$ $+ \big[\sigma\big(t, \theta_1, i_0\big) - \sigma\big(t, \theta_2, i_0\big), z_1 - z_2\big].$

INTRODUCTION

In the new era, numerous complex and large-scale systems come into play. A common feature of such systems is the inclusion of both continuous dynamics and discrete events, and their interactions. The discrete events cannot be described by the usual stochastic differential equations, but can be recast as stochastic systems driven by pure jump processes [38, 56, 57, 58, 59]. In responding to the increasing needs from modeling, analysis, and computation, this dissertation focuses on a class of hybrid systems known as Markov switching diffusions where the switching between discrete events is driven by a continuous-time Markov chain. Because of the switching, some well-known results for systems without the switching do not carry over. For example, as explained in [57, Section 5.6, pp. 229-233] by putting two equations (e.g., linear equations) together switching back and forth, even if both equations are stable, the resulting switched system may be unstable. Thus, the intuition we have does not always work. This indicates that we cannot quickly come to the conclusion regarding the corresponding properties.

In recent years, FBSDEs have been extensively studied because of their numerous applications in many areas such as control and game theory [45, 62], mathematical economics [23], and mathematical finance [18, 29]. There are three main approaches for the wellposedess of the FBSDEs, each of which has its own advantages and disadvantages. In [11] and then [43], the method of contraction mapping was studied, which works well for small time durations. In [34], the four-step-scheme was first introduced to establish the existence and uniqueness of solutions of FBSDEs under non-degenerate condition of the forward equation and some regularity requirements of the coefficients (see also [20, 64]). In [25], the existence and uniqueness of solutions of FBSDEs are proved under monotonicity condition without non-degeneracy condition of the forward equation. The monotonicity condition is then remarkably weakened and developed in subsequent works [24, 44, 60]. For the progress and related works on FBSDEs, we refer the reader to [35, 37] and the references therein.

Since the pioneering works [32, 26, 27] and because of the pressing need of treating large-scale systems, there has been increasing effort of dealing with mean-field interactions, systems with mean-field interactions (in which the coefficients also depend on the their distributions), and related control problems, and games. The stochastic maximum principles for both mean-field games and mean-field control problems naturally lead to a class of mean-field type FBSDEs (see [12, 16, 1, 61] and references therein). To study the well-posedness of this new class of FBSDEs, the approaches in [44] and [24] are extended in the recent works [13] and [19].

In contrast to the vast literature on FBSDEs, such equations with Markovian switching have not received as needed attention. Although BSDEs with Markovian switching were studied in [22, 33] and were used to formulate stochastic recursive control problems [65], to the best of our knowledge, there is no available well-posedness result even for the FBSDEs with Markovian switching. To deal with large-scale switching systems, the mean-field types of FBSDEs with Markovian switching naturally come into play when one needs to treat the mean-field control problems. Nevertheless, a main issue we encountered was that the associated limit mean-field measure was not known for the systems involving both mean-field interactions and random switching. To settle this issue, in the recent work of [9], they showed that the mean-field limit measure is not deterministic, but a conditional (random) measure that is a solution of a system of McKean-Vlasov stochastic differential equations. It is worth mentioning that conditional mean-field also appears in such problems as mean-field games and control with common noise [30, 46], major-minor mean-field games [39, 41], mean-field games with leader-follower [52], and filtering for McKean-Vlasov SDEs or mean-field control with partial-observations [14, 50]. In our setting, the conditioning is taken with respect to the past information generated by the switching process-the Markov chain. Thanks to this conditional measure, it enabled us to obtain maximum principles of such switching diffusion systems in [8, 40].

Continuing our study, in this dissertation we devote our attentions to a number of important issues. We begin by developing different approaches to examine well-posedness of FBSDEs with Markovian switching and then mean-field type FBSDEs with Markovian switching. The appearance of the switching process leads to two main differences from the FBSDEs in [11, 24, 25, 44, 60] and mean-field type FBSDEs in [13, 16, 19].

First, apart from the Brownian motions, the backward equations are also driven by martingales associated to the Markov switching process whose quadratic variations are random, not deterministic as those of Brownian motion. Second, the mean-field terms in the mean-field type FBSDEs with Markovian switching are represented by conditional (random) distributions of the processes involved given the history of the switching process. These differences make the estimates needed in the analysis more complex. It, in turn, requires us developing several new supporting results. We end the dissertation by treating the nonzero-sum conditional mean-field games using the new wellposedness results for FBSDEs. Conditions on the coefficients of the nonzero-sum conditional mean-field linear-quadratic stochastic differential game with regime switching and open-loop strategies are provide to guarantee a Nash equilibrium point for any (not necessarily small) time duration.

The rest of the dissertation is arranged as follows. Chapter 2 concentrates on forward–backward stochastic differential equations with Markov switching diffusions with non-degeneracy of the diffusion matrix. We prove a result of existence and uniqueness of solutions in two steps. The first part, section 1, studies the problem of existence and uniqueness over a small enough time duration. The second one, section 2, explains, by using the connection with a system of PDEs and the local result, how we can deduce the existence and uniqueness of a solution (under a non-degeneracy assumption) over an arbitrarily prescribed time duration. This approach relaxes the regularity assumptions required on the coefficients by the

Four-Step scheme. Section 3 gives some related PDE in the weak sense. Chapter 3 concentrates on Markov switching diffusions, developing the continuation method and monotonicity conditions to examine the well-posedness of FBSDEs with Markovian switching. The emphasis is on the corresponding backward stochastic differential equations, and forward-backward stochastic differential equations. Chapter 4, section 1 is devoted to stochastic differential equations with Markov switching with mean-field interactions and obtains certain results that tie up the conditional mean-field measure with that of the underlying systems. Finally, chapter 5 deals with linear (in continuous state) mean-field games.

Compared with the existing works, the main technical challenges stem from the appearance of the switching process and the conditional mean-field term used. The appearance of the switching process leads to two main differences from the FBSDEs in [11, 24, 25, 44, 60] and mean-field type FBSDEs in [13, 16, 19]. First, apart from the Brownian motions, the backward equations are also driven by martingales associated to the Markov switching process whose quadratic variations are random, not deterministic as those of Brownian motions. Second, the mean-field terms in the mean-field type FBSDEs with Markovian switching are represented by conditional (random) distributions of the processes involved given the history of the switching process. These differences make the estimates needed in the analysis more complicated. It, in turn, requires us developing several new supporting results.

CHAPTER 1
Stochastic Differential Equations

This chapter is devoted to introducing preliminaries on probability, conditional expectations, several important stochastic processes such as martingales, Brownian motions and Markov chains, stochastic integrals, several classes of stochastic differential equations, and Itô's formula. In order to study common properties of stochastic processes, stochastic integrals driven by martingales and some useful inequalities are also presented.

1 Probability Space and Conditional Expectations

In this section, we recall basic definitions and theorems probability theory needed to further our study of stochastic calculus.

1.1 Preliminaries on Probability Space

A *σ-algebra* $\mathcal{F}$ on a given set Ω is a family of subsets of Ω with the following properties

(i) $\varnothing \in \mathcal{F}$

(ii) A $\in \mathcal{F} \implies A^C \in \mathcal{F}$ where A^C is the complement of $\mathcal{F}$ in Ω

(iii) $A_1, A_2, ... \in \mathcal{F} \implies \cup_{i=1}^{\infty} A_i \in \mathcal{F}$

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*. A *probability measure* is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that

(a) $\mathbb{P}(\varnothing) = 0$ and $\mathbb{P}(\Omega) = 1$

(b) For disjoint $\{A_i\}_{i\geqslant 1} \subset \mathcal{F}$ we have $\mathbb{P}\Big(\cup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Finally we denote $(\Omega, \mathcal{F}, \mathbb{P})$ as a *probability space*. Any $A \in \mathcal{F}$ is called an *event*. Two events A_1, $A_2 \in \mathcal{F}$ are said to be *independent* if $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$. Two σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ of $\mathcal{F}$ are independent if any event $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$ are independent for all $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$. An *$\mathcal{F}$-measurable set* is a subset A that belongs to Ω. A complete probability space is a probability space such that $\mathcal{F}$ contains all subsets of Ω with $\mathbb{P}$ outer measure zero. If the event A occurs such that $\mathbb{P}(A) = 1$, we say A holds $\mathbb{P}$-a.s. (almost surely). A function $X : \Omega \to \mathbb{R}^d$ is called $\mathcal{F}$-measurable if

$$X^{-1}(A) := \{\omega \in \Omega; X(\omega) \in A\} \in \mathcal{F}, \text{for all open sets } A \in \mathbb{R}^d.$$

For any function X, denote the σ-algebra $\sigma(X)$ generated by X as the smallest σ- algebra on Ω containing all the sets $X^{-1}(A)$, where $A \subset \mathcal{F}$ open. When $\Omega = \mathbb{R}^n$, we call $\mathcal{B}(\mathbb{R}^n)$ the *Borel σ-algebra* and $B \in \mathcal{B}(\mathbb{R}^n)$ *Borel sets.* For a measurable space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, a $\mathcal{B}(\mathbb{R}^n)$-measurable function is called a *Borel measurable function.* For a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a *random variable* X is an $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^d$. A family of random variables is independent if the σ-algebras generated by them are independent, and a random variable X is independent of a σ-algebra $\mathcal{G}$ if the σ-algebra generated by X is independent of $\mathcal{G}$. Additionally, the *law* (or distribution) of the real-valued random variable X is the pushforward measure μ_X, defined

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\omega : X(\omega) \in B).$$

for some $B \in \mathcal{B}(\mathbb{R}^n)$. Two random values $X : \Omega \to \mathbb{R}^n$ and $Y : \Omega \to \mathbb{R}^m$ are independent if and only if

$$\mathbb{P}(\omega : X(\omega) \in A, Y(\omega) \in B) = \mathbb{P}(\omega : X(\omega) \in A)\mathbb{P}(\omega : Y(\omega) \in B)$$

for all $A \in \mathcal{B}(\mathbb{R}^n), B \in \mathcal{B}(\mathbb{R}^m)$.

If $\int_\Omega |X(\omega)| d\mathbb{P}(\omega) < \infty$, (hence integrable) The *expectation* of X is defined as

$$\mathbb{E}[X] = \int_\Omega X(\omega) d\mathbb{P}(\omega)$$

For an $\mathbb{R}^n$-valued random variable X, the law induced by X on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ allows for the expectation to be written

$$\mathbb{E}[X] = \int_{\mathbb{R}^d} x d\mathbb{P}_x(x)$$

A *stochastic process* is a collection $\{X_t\}_{t\geqslant 0}$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}^d$. Fixing t gives the random variable $X_t(\omega) \in \mathbb{R}^n$, while fixing ω gives a function $X_t(\omega) \in \mathbb{R}^n$. We will commonly denote the stochastic process $\{X_t\}_{t\geqslant 0}$ as X_t or $X(t)$.

Definition 1.1. *(i) A filtration on* $(\Omega, \mathcal{F})$ *is an increasing family* $\{\mathcal{F}_t\}_{t\geqslant 0}$ *of* σ*-algebras* $\mathcal{F}_t \subseteq \mathcal{F}$*; that is, if* $0\leqslant s < t$ *then* $\mathcal{F}_s \subseteq \mathcal{F}_t$*.*

(ii) A process $\{X_t\}_{t\geqslant 0}$ *is called adapted to the filtration* $\{\mathcal{F}_t\}_{t\geqslant 0}$ *if* X_t *is* $\mathcal{F}_t$*-measurable for all* t*.*

A filtration is *right continuous* if $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ for all $t \geqslant 0$. A filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})_{t\geqslant 0})$ is said to *satisfy the usual conditions* if the space is complete, the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ is right-continuous, and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets in $\mathcal{F}$. A stochastic process $X_t(\omega)$ is said to be adapted to the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ or $\{\mathcal{F}_t\}$-*adapted* if for each $t\geqslant 0$, $X_t(\omega)$ is an $\mathcal{F}_t$-measurable random variable. A stochastic process X_t is said to be *continuous* if for almost all $\omega \in \Omega$, $X_t(\omega)$ is continuous on $t \geqslant 0$. A stochastic process X_t on $\mathbb{R}$ is called a *càdlàg process* if

(i) X_t is right-continuous

(ii) For almost all $\omega \in \Omega$, the left hand limit $\lim_{s\to t^-} X_s(\omega)$ exists and is finite for all $t > 0$.

We will commonly denote such left limits as X_{t-}. The smallest σ-algebra on $\mathbb{R}_+ \times \Omega$ where every left-continuous process is a measurable function of (t, ω) is denoted $\mathcal{P}$. A stochastic process X_t is said to be *predictable* if it is $\mathcal{P}$-measurable when regarded as function of (t, ω). An $\mathbb{R}^d$-valued stochastic process X_t is said to be $\mathcal{F}$*-progressively measurable* if for all $t \in [0, T]$, the map $(s, \omega) \mapsto X_s(\omega)$ is $\mathcal{B}[0, \infty) \times \mathcal{F}_t/\mathcal{B}(\mathbb{R}^d)$-measurable. A random variable $\tau : \Omega \to [0, \infty]$ is called a *stopping time* with respect to filtration $\{\mathbb{F}\}_t$ if $\{\tau(\omega)\leqslant t\} \in \mathcal{F}_t$ for all $t\geqslant 0$.

1.2 Conditional Expectations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a random variable $X : \Omega \to \mathbb{R}^d$ such that $\mathbb{E}[|X|] < \infty$. For a σ-algebra $\mathcal{G} \subset \mathcal{F}$, we define the conditional expectation of X given $\mathbb{F}$ as follows:

Definition 1.2. *The conditional expectation of X given $\mathcal{G}$, $\mathbb{E}[X|\mathcal{G}]$ is the a.s. unique function from Ω to $\mathbb{R}^d$ satisfying*

(i) $\int_{\mathcal{A}} \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_{\mathcal{A}} X d\mathbb{P}$, *for all* $\mathcal{A} \in \mathcal{G}$

(ii) $\mathbb{E}[X|\mathcal{G}]$ *is* $\mathcal{G}$*-measurable.*

Additionally, for $B \in \mathcal{F}$, we define the *conditional probability* of B given $\mathcal{F}$ as $\mathbb{P}[B|\mathcal{F}] = \mathbb{E}[\mathbb{1}_B|\mathcal{F}]$.

We list a few important properties of conditional expectation

Lemma 1.3. *Let Y be a $\mathcal{F}_t$-measurable random variable with $\mathcal{F}_s \subset \mathcal{F}_t$ and $\mathcal{G} \subset \mathcal{F}$. Then,*

(a) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$

(b) For a $\mathcal{G}$-measurable random variable X, $\mathbb{E}[X|\mathcal{G}] = X$, a.s.

(c) For $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}[\alpha X + \beta Y|\mathcal{F}] = \alpha \mathbb{E}[X|\mathcal{F}] + \beta \mathbb{E}[Y|\mathcal{F}]$

(d) For an X independent of $\mathcal{F}$, $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$

(e) $\mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{F}] = \mathbb{E}[X|\mathcal{G}]$

(f) If $\phi : \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}[|\phi(X)|] < \infty$ then

$$\phi(\mathbb{E}[X|\mathcal{F}]) \leqslant \mathbb{E}[\phi(X)|\mathcal{F}].$$

We conclude this section recalling the definitions of Markov kernels and that of regular conditional distributions.

Definition 1.4. *Let $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$ be measurable spaces. A map $\kappa : \Omega_1 \times \mathcal{A}_2 \to [0, \infty]$ is called a Markov kernel if*

(i) $\omega_1 \mapsto \kappa(\omega_1, A_2)$ is $\mathcal{A}_1$-measurable for any $A_2 \in \mathcal{A}_2$.

(ii) $A_2 \mapsto \kappa(\omega_1, A_2)$ is a probability measure on $(\Omega_2, \mathcal{A}_2)$ for any $\omega_1 \in \Omega_1$.

Definition 1.5. *Let X be a random variable taking values in the measurable space $(E, \mathcal{E})$, and let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra. A Markov kernel $\kappa_{X,\mathcal{F}}$ from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ is called a regular conditional distribution of X given $\mathcal{G}$ if*

$$\kappa_{X,\mathcal{F}}(\omega, A) = \mathbb{P}(X \in A|\mathcal{G})(\omega)$$

for $\mathbb{P}$-almost all $\omega \in \Omega$ and for all $A \in \mathcal{E}$.

When $\mathcal{F} = \sigma(X)$ for a random variable X, we denote $\kappa_{Y,\sigma(X)}(X^{-1}(x), A)$ as the regular conditional distribution *of Y given X.*

2 Martingales, Brownian Motions, and Markov Chains

For a positive integer d, vectors $x, y \in \mathbb{R}^d$, denote by $\langle x, y\rangle$ their dot product and by $x^\intercal$ the transpose of x. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geqslant 0})$ be a fixed probability space satisfying the usual conditions.

2.1 $\mathcal{L}^p(\mathbb{R}^d)$ Spaces

Denote

$$\mathcal{L}^p(\mathbb{R}^d) = \{\xi : \Omega \to \mathbb{R}^d, \mathcal{F}\text{-measurable}, \mathbb{E}|\xi|^p < \infty\}, \quad p \geqslant 1,$$

$$\mathcal{S}^2(0,T;\mathbb{R}^d) = \left\{\varphi : [0,T] \times \Omega \to \mathbb{R}^d,\ \mathcal{F}\text{-adapted càdlàg process}, \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |\varphi_t|^2\right] < \infty\right\},$$

and

$$\mathcal{L}^0(0,T;\mathbb{R}^d) = \left\{\psi : [0,T] \times \Omega \to \mathbb{R}^d,\ \mathcal{F}\text{-progressively measurable process}\right\},$$
$$\mathcal{L}^2(0,T;\mathbb{R}^d) = \left\{\psi \in \mathcal{L}^0(0,T;\mathbb{R}^d) : \|\psi\|_2^2 = \mathbb{E}\left[\int_0^T |\psi_t|^2 dt\right] < \infty\right\}.$$

It can be shown that $\mathcal{L}^2(0,T;\mathbb{R}^d)$ is a Hilbert spaces; see [21, Lemma A.2.5].

2.2 Martingales

This subsection is devoted to the definitions and basic properties of martingales, supermartingales, submartingales. First, we have the following definitions.

Definition 1.6. *An n-dimensional stochastic process $\{U_t\}_{t\geqslant 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ and measure $\mathbb{P}$ if* *(i)* U_t *is* $\{\mathcal{F}_t\}_{t\geqslant 0}$*-adapted,*

(ii) $\mathbb{E}[|U_t|] < \infty$ *for all* t*,*

(iii) $\mathbb{E}[U_s \big| \mathcal{F}_t] = U_t$ *for all* $s\geqslant t\geqslant 0$*.*

A supermartingale is defined similarly but replace (iii) by

$$\mathbb{E}[U_s \big| \mathcal{F}_t]\leqslant U_t \quad \text{for all} \quad s\geqslant t\geqslant 0.$$

A submartingale is defined similarly but replace (iii) by

$$\mathbb{E}[U_s \big| \mathcal{F}_t]\geqslant U_t \quad \text{for all} \quad s\geqslant t\geqslant 0.$$

Further, we define another important class of martingale.

Definition 1.7. *A process* $\{V_t\}$ *is said to be a local martingale if there exists a sequence of stopping times* τ_k *such that* $V_t^{\tau_k}$ *is a martingale, with* $\tau_k \to \infty$ *a.s. increasing.*

Of course, every martingale is a local martingale, and additionally every bounded local martingale is a martingale. We now look at the conditions needed to ensure convergence of martingales.

Theorem 1.8. *Let* U *be a right-continuous supermartingale such that* $\sup_{0\leqslant t\leqslant\infty} \mathbb{E}\{|U_t|\} < \infty$*. Then* $V = \lim_{t\to\infty} U_t$ *exists a.s., and* $\mathbb{E}[|V|] < \infty$*.*

Next, we have the Doob-Meyer decompsotion theorem.

Theorem 1.9. *The Doob-Meyer decomposition expresses a submartingale in continuous time as the unique way as the sum of a martingale and a nondecreasing predictible process. That is, if* U_t *is a submartingale, then*

$$U_t = M_t - A_t, \quad 0\leqslant t\leqslant T,$$

where M_t *is a* $\mathcal{F}$*-martingale and* A_t *is a nondecreasing process.*

To proceed, we mention the definitions of quadratic variations (optional and predictable quadratic variations) and covariations which will be needed for important estimates in the next chapters.

Definition 1.10. *(i) If $X_t(\cdot) : \Omega \to \mathbb{R}$ is a continuous stochastic process, then the quadratic variation process of X_t, $[X]_t$ is defined by*

$$[X]_t(\omega) = \lim_{\Delta t_k \to 0} \sum_{t_k \leqslant t} |X_{t_{k+1}}(\omega) - X_{t_k}(\omega)|^2 \text{ (limit in probability)}$$

where $0 = t_1 < t_2 < \ldots < t_n = t$ and $\Delta t_k = t_{k+1} - t_k$.

(ii) More generally, the covariation of two processes X and Y is

$$[X,Y]_t(\omega) = \lim_{\Delta t_k \to 0} \sum_{t_k \leqslant t} (X_{t_{k+1}}(\omega) - X_{t_k}(\omega))(Y_{t_{k+1}}(\omega) - Y_{t_k}(\omega)) \text{ (limit in probability).}$$

The predictable quadratic variation is sometimes used for locally square integrable martingales. This is written as $\langle U_t \rangle$, and is defined to be the unique right-continuous and increasing predictable process starting at zero such that $U^2 - \langle U \rangle$ is a local martingale. Its existence follows from the Doob–Meyer decomposition theorem and, for continuous local martingales, it is the same as the quadratic variation.

Definition 1.11. *A square integrable martingale is called purely discontinuous martingale if it is (strongly) orthogonal to all the square integrable martingales with continuous trajectories.*

According to Theorem 3 in Section 1.5 in [51], any square integrable martingale can be decomposed into the sum of a martingale with continuous trajectory and null at 0 and a purely discontinuous martingale. The (optional) quadratic variation of a purely discontinuous martingale is a pure jump process.

Now we are in a position to state the Burkholder-Davis-Gundy inequality, which is frequently used throughout the dissertation.

Theorem 1.12 (Burkholder-Davis-Gundy inequality)**.** *For any $1 \leqslant p < \infty$ there exist positive constants c_p, C_p such that, for all local martingales X with $X_0 = 0$ and stopping times τ,* *the following inequality holds*

$$c_p \mathbb{E}\left[[X]_\tau^{p/2}\right] \leqslant \mathbb{E}\left[\sup_{s\leqslant\tau} |X_s|^p\right] \leqslant C_p \mathbb{E}\left[[X]_\tau^{p/2}\right].$$

Furthermore, for continuous local martingales, this statement holds for all $0 < p < \infty$*.*

Finally, let us mention some important martingale convergence theorems.

Theorem 1.13 (Doob's martingale convergence theorem I)**.** *Let* U_t *be a right-continuous supermartingale with the property that*

$$\sup_{t>0} \mathbb{E}[U_t^-] < \infty,$$

where $U_t^- = \max(-U_t, 0)$*. Then the pointwise limit*

$$U(\omega) = \lim_{t\to\infty} U_t(\omega)$$

exists for a.a. ω *and* $\mathbb{E}[U^-] < \infty$*.*

Theorem 1.14 (Doob's martingale convergence theorem II)**.** *Let* U_t *be a right-continuous supermartingale. Then the following are equivalent:*

(i) $\{U_t\}_{t\geqslant 0}$ *is uniformly integrable.*

(ii) There exists $U \in \mathcal{L}^1(\mathbb{P})$ *such that* $U_t \to U$ *a.e.* $(\mathbb{P})$ *and* $U_t \to U$ *in* $\mathcal{L}^1(\mathbb{P})$*; that is,*

$$\int |U_t - U| d\mathbb{P} \to 0 \text{ as } t \to \infty.$$

2.3 Brownian Motions

On $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, a (one-dimensional) process $W(\cdot)$ is called a (standard) Brownian motion if it satisfies the following conditions

(i) W_t is almost surely continuous. That is, $\mathbb{P}\big(\omega : W_t(\omega)$ is continuous in $t\big) = 1$.

(ii) W_t has stationary, independent increments.

(iii) W_t is a Gaussian process. That is, $W_t - W_s \sim \mathcal{N}(0, t-s)$ for $0 \leqslant s \leqslant t$.

Note that W_t has independent increments means that for all $0 \leqslant t_1 < \ldots < t_k < \infty$, the random variables

$$W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_k} - W_{t_{k-1}} \text{ are independent.}$$

Equivalently, $W_s - W_t$ is independent of $\mathcal{F}_t$ for any $s > t \geqslant 0$. In addition, W_t has stationary increments means that the process $\{W_{t+h} - W_t\}_{h \geqslant 0}$ has the same distribution for all $t \geqslant 0$. Furthermore, the quadratic variations of a one-dimensional Brownian motion W_t is $[W]_t = t$ a.s.

A d-dimensional process $W(\cdot) = \big(W_1(\cdot), W_2(\cdot), \ldots, W_d(\cdot)\big)^\intercal$ is a d-dimensional Brownian motion if each $W_i(\cdot)$, $1 \leqslant i \leqslant d$, is a standard one-dimensional Brownian motion and all d components $\{W_1(\cdot), W_2(\cdot), \ldots, W_d(\cdot)\}$ are independent.

Due to the property of independent increments, every d-dimensional Brownian motion W_t is a martingale with respect to the σ-algebras $\mathcal{F}_t$ generated by $\{W_s \mid s \leqslant t\}$. More precisely, we have

$$\mathbb{E}\big(W_t \big| \mathcal{F}_s\big) = W_s, \quad \text{for all} \quad 0 \leqslant s \leqslant t.$$

2.4 Markov Chains

Let $\mathcal{M} = \{1, 2, \ldots, m_0\}$ be a finite set and μ a probability measure on $\mathcal{M}$. A *transition matrix* $P(t) = (p_{i_0 j_0}(t))_{i_0, j_0 \in \mathcal{M}}$ is a $m_0 \times m_0$ matrix that satisfies the following conditions for every $i_0, j_0 \in \mathcal{M}$ and $s, t \geqslant 0$,

$$\begin{aligned}
&\text{(i)} \quad p_{i_0 j_0}(t) \geqslant 0,\\
&\text{(ii)} \quad \sum_{j_0 \in \mathcal{M}} p_{i_0 j_0}(t) = 1,\\
&\text{(iii)} \quad \sum_{k_0 \in \mathcal{M}} p_{i_0 k_0}(s) p_{k_0 j_0}(t) = p_{i_0 j_0}(s+t).
\end{aligned}$$

A transition matrix P is called *standard* if $\lim_{t \to 0_+} p_{i_0 j_0}(t) = \delta_{i_0 j_0}$. In addition, matrix P is called *measurable* if $p_{i_0 j_0}(\cdot)$ is a measurable function in $(0, \infty)$. Note that if $P(t)$ satisfies the first two conditions, it is called a *stochastic matrix*. The last condition is often referred to as the *Chapman-Kolmogorov equation*. We shall always assume that the transition matrix P is standard and measurable.

Definition 1.15. *A continuous-time process $\alpha(\cdot)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the state space $\mathcal{M}$ is called a Markov chain with initial distribution μ and transition matrix $P(t)$ if*

(i) $\mathbb{P}(\alpha_0 = i_0) = \mu(i_0)$, and

(ii) for any $j_0, i_0, i_1, \ldots, i_n \in \mathcal{M}$ and $s > t > t_1 > \ldots > t_n \geqslant 0$,

$$\mathbb{P}(\alpha_s = j_0 | \alpha_t = i_0, \alpha_{t_k} = i_k, 1 \leqslant k \leqslant n) = \mathbb{P}(\alpha_s = j_0 | \alpha_t = i_0) = p_{i_0 j_0}(s-t).$$

It can be shown that $P(t)$ is differentiable at 0 (See the book by Kai Lai Chung [2], Sections II.2-3.) In other words, for any $i_0 \neq j_0 \in \mathcal{M}$, the limits

$$-p'_{i_0 i_0}(0) = \lim_{t \to 0_+} \frac{1 - p_{i_0 i_0}(t)}{t}, \qquad p'_{i_0 j_0}(0) = \lim_{t \to 0_+} \frac{p_{i_0 j_0}(t)}{t}$$

exist and finite. Denote $Q = (q_{i_0 j_0}) = P'(0)$ then P satisfies the following initial value linear ordinary differential equation

$$\begin{cases} \dfrac{dP(t)}{dt} &= P(t)Q, \quad t \geqslant 0, \\ P(0) &= I_{m_0} \end{cases}$$

which leads to the explicit presentation

$$P(t) = e^{tQ} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n, \qquad t \geqslant 0.$$

Since $P(t)$ is a stochastic matrix, it is clear that the matrix $Q = \left(q_{i_0 j_0}\right)_{i_0, j_0 \in \mathcal{M}}$ satisfies the following properties for any $i_0 \neq j_0 \in \mathcal{M}$

$$\text{(i)} \quad q_{i_0 j_0} \geqslant 0,$$

$$\text{(ii)} \quad \sum_{j_0 \in \mathcal{M}} q_{i_0 j_0} = 0.$$

Note that beside the Definition 1.15, we can also define a Makov chain in two other ways using a matrix Q satisfied above conditions (which is called the *generator* of the Markov chain), or using the jump chain and holding times. It is obvious that for a Markov chain with the generator matrix Q, we have

$$\mathbb{P}(\alpha(t+h)=j_0|\alpha(t)=i_0)=\begin{cases} q_{i_0j_0}t+o(h), & \text{if } i_0\neq j_0,\\ 1+q_{i_0i_0}t+o(h), & \text{if } i_0=j_0. \end{cases}$$

Martingales Associated with a Markov Chain.

Definition 1.16. *For a Markov chain $\alpha(\cdot)$ with state space $\mathcal{M}$ and generator matrix Q, associated with each pair $(i_0,j_0)\in\mathcal{M}\times\mathcal{M}$ satisfying $i_0\neq j_0$, define the process*

$$M_{i_0j_0}(t)=\big[M_{i_0j_0}\big](t)-\big\langle M_{i_0j_0}\big\rangle(t) \tag{1.2.1}$$

where

$$\big[M_{i_0j_0}\big](t)=\sum_{0\leqslant s\leqslant t}\mathbb{1}\big(\alpha(s_-)=i_0\big)\mathbb{1}\big(\alpha(s)=j_0\big),\qquad \big\langle M_{i_0j_0}\big\rangle(t)=\int_0^t q_{i_0j_0}\mathbb{1}\big(\alpha(s_-)=i_0\big)ds,$$

and $\mathbb{1}$ denotes the usual zero-one indicator function.

It follows from [22] that the process $M_{i_0j_0}(t)$, $0\leqslant t\leqslant T$ is a discontinuous and *square integrable martingale* with respect to $\mathcal{F}_t^\alpha$, which is null at the origin. The processes $[M_{i_0j_0}](t)$ and $\langle M_{i_0j_0}\rangle(t)$ are the optional and predictable quadratic variations, respectively. For convenience, we define

$$M_{i_0i_0}(t)=\big[M_{i_0i_0}\big](t)=\big\langle M_{i_0i_0}\big\rangle(t)=0,\qquad i_0\in\mathcal{M}.$$

From the definition of optional quadratic covariations we have the following orthogonality relation

$$\big[M_{i_0j_0},W\big]=0,\quad \big[M_{i_0j_0},M_{p_0q_0}\big]=0 \text{ when } (i_0,j_0)\neq(p_0,q_0), \tag{1.2.2}$$

where $\big[M_{i_0j_0}, W\big]$ and $\big[M_{i_0j_0}, M_{p_0q_0}\big]$ are the optional quadratic covariations of the pairs $(M_{i_0j_0}, W)$ and $(M_{i_0j_0}, M_{p_0q_0})$, respectively; see [22, 33]. Denote

$$\mathcal{M}^2(0,T;\mathbb{R}^d) = \Big\{\lambda = \big(\lambda_{i_0j_0} : i_0, j_0 \in \mathcal{M}\big) \text{ such that } \lambda_{i_0j_0} \in \mathcal{L}^0(0,T;\mathbb{R}^d),\ \lambda_{i_0i_0} \equiv 0,$$
$$\text{and} \sum_{i_0,j_0\in\mathcal{M}} \mathbb{E}\int_0^T \big|\lambda_{i_0j_0}(t)\big|^2 d\big[M_{i_0j_0}\big](t) < \infty\Big\}.$$

For a collection of $\mathcal{F}$-progressively measurable functions $\lambda_t = \big(\lambda_{i_0j_0}(t)\big)_{i_0,j_0\in\mathcal{M}}$, $t \geqslant 0$, we denote

$$\int_0^t \lambda_s \bullet dM_s = \sum_{i_0,j_0\in\mathcal{M}} \int_0^t \lambda_{i_0j_0}(s) dM_{i_0j_0}(s) \quad \text{and} \quad \lambda_t \bullet dM_t = \sum_{i_0,j_0\in\mathcal{M}} \lambda_{i_0j_0}(t) dM_{i_0j_0}(t).$$

It can be shown that $\mathcal{M}^2(0,T;\mathbb{R}^d)$ is a Hilbert space; see [21, Lemma A.2.5].

3 Stochastic Integrals with Martingales

In this section, we expand upon what we have discussed so far to define the stochastic integrals needed to make sense of the SDEs we will see. We begin with the standard stochastic integral, then the case with a continuous markov chain, and finally with jump. With these martingale and stochastic integral connections established, we are able to work with such stochastic differential equations. This section closely follows Watanabe [28], chapter II, which can be consulted for a details on the concepts which follow.

3.1 Stochastic Integrals

We first begin with the original formulation of the stochastic integral courtesy of K. Itô himself, as described by Watanabe [28]. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geqslant 0})$ be a fixed probability space satisfying the usual conditions. Let W_t be a one dimensional $\mathcal{F}_t$-Brownian motion. To begin, we introduce the the following spaces

Definition 1.17. *Let $\mathbb{L}^2$ be the space of all real-valued measurable processes $\{\phi_t\}_{t\geqslant 0}$ on Ω which are adapted to the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ such that for every $T > 0$*

$$||\phi||^2_{2,T} = \mathbb{E}\left[\int_0^T \phi^2(s,\omega)ds\right] < \infty.$$

We consider $\phi_1 = \phi_2$ if $||\phi_1 - \phi_2||_{2,T} = 0$ for every $T > 0$. Further we define for every $\phi \in \mathbb{L}^2$

$$||\phi||_2 = \sum_{n=1}^{\infty} \frac{1}{2^n}(||\phi||_{2,n} \wedge 1)$$

Note that $||\phi - \phi_2||_2$ defines a metric on $\mathbb{L}$, which is complete. Additionally, note that we can always take a $\phi \in \mathbb{L}^2$ such that ϕ is predictable without loss of generality. Further we define another class of processes:

Definition 1.18. *Let $\mathbb{L}^0$ be the space of processes $\phi \in \mathbb{L}^2$ that satisfies the following properties:*

(i) There exists a sequence of real numbers $0 = t_0 < t_1 < ... < t_n < ... < \infty$

(ii) There exists a sequence of random variables $\{f_i\}_{i=0}^{\infty}$ such that f_i is $\mathcal{F}_{t_i}$-measurable, with $\sup_i ||f_i||_\infty < \infty$ and

$$\phi(t, \omega) = \begin{cases} f_0(\omega), & t = 0 \\ f_i(\omega), & t \in (t_i, t_{i+1}], \quad i = 0, 1, ... \end{cases} \tag{1.3.1}$$

It is known that $\mathbb{L}^0$ is dense in $\mathbb{L}^2$ with the metric $|| \cdot ||_2$. See Watanabe [28], lemma 1.1 from chapter II, for the proof. Next, we define the space of all square integrable martingales.

Definition 1.19. *Let $\mathbb{M}^2$ be the space of all square integrable martingales $\{X_t\}_{t \geqslant 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to filtration $(\mathcal{F}_t)_{t \geqslant 0}$ and $X_0 = 0$ almost surely. Let $\mathbb{M}_c^2 = \{X \in \mathbb{M}^2 : t \mapsto X_t$ is continuous a.s.$\}$.*

Definition 1.20. *For $X \in \mathbb{M}^2$, we set*

$$||X||_T^a = \mathbb{E}[X_T^2]^{\frac{1}{2}}, \quad T > 0$$

and

$$||X||^b = \sum_{n=1}^{\infty} \frac{1}{2^n}(||X||_t^a \wedge 1).$$

It is known that $\mathbb{M}^2$ is a complete metric space with $||X - Y||^b$, for $X, Y \in \mathbb{M}^2$. Moreover, $\mathbb{M}_c^2$ is a closed subspace of $\mathbb{M}^2$. Finally, we can define the stochastic integral. Let $W(t)$ be an $\mathcal{F}_t$-Brownian motion on our probability space, and $\phi \in \mathbb{L}^0$. By, (1.3.1), we then set

$$I(\phi)(t,\omega) = \sum_{i=0}^{n-1} f_i(\omega)(W(t_{i+1},\omega) - W(t_i,\omega)) + f_n(\omega)(W(t,\omega) - W(t_n,\omega))$$

for $t_n \leqslant t \leqslant t_{n+1}$, $n = 0, 1, \ldots$. Note we can then express $I(\phi)$ as the infinite sum

$$I(\phi)(t) = \sum_{i=0}^{\infty} f_i(W(t \wedge t_{i+1}) - W(s \wedge t_i))$$

For $s \leqslant t$, we have

$$\mathbb{E}[f_i(W(t \wedge t_{i+1}) - W(s \wedge t_i))|\mathcal{F}_s] = f_i(W(t \wedge t_{i+1}) - W(s \wedge t_i))$$

Thus $I(\phi)(t) \in \mathbb{M}_c^2$. Additionally,

$$\mathbb{E}[I(\phi)(t)^2] = \sum \mathbb{E}[f_i(t \wedge t_{i+1}) - (t \wedge t_i))] = \mathbb{E}\left[\int_0^t \phi^2(s,\omega)ds\right]$$

Hence

$$||I(\phi)||_T^a = ||\phi||_{2,T} \tag{1.3.2}$$

$$||I(\phi)||^b = ||\phi||_2 \tag{1.3.3}$$

Now, for $\phi \in \mathbb{L}^2$, we know through $\mathbb{L}^0$ being dense in $\mathbb{L}^2$ with respect to $||\cdot||_2$ that there is a $\phi_n \in \mathbb{L}^0$ such that $||\phi - \phi_n||_2 \to 0$. We also know $I(\phi_n)$ is a Cauchy sequence in $\mathbb{M}^2$ since by (3.3) we have $||I(\phi)_n - I(\phi)_m||^b = ||\phi_n - \phi_m||_2$ and hence through the completeness of $\mathbb{M}^2$, it converges to a unique limit which we denote $I(\phi) \in \mathbb{M}_c^2$.

Definition 1.21. *$I(\phi) \in \mathbb{M}_c^2$ as defined above is called the stochastic integral of $\phi \in \mathbb{L}^2$ with respect to the Brownian motion W_t, with representation $I(\phi)(t) = \int_0^t \phi_s dW_s$.*

Thus the stochastic integral is defined as a stochastic process itself, and one should now note for a fixed t we also call the random variable $I(\phi)(t)$ a stochastic integral. Further, for an m-dimensional F_t-Brownian motion W_t, and $\phi_t^1, ..., \phi_t^m \in \mathbb{L}^2$, we can define the stochastic integrals $\int_0^t \phi_s^i dW_s^m$ for $i = 1, 2, .., r$. We refrain from listing the properties of the stochastic integral until we have defined it for martingales.

We now expand the stochastic integral to that of local martingales. The framework for this formulation is quite similar to the construction of the previous stochastic integral, so we instruct the reader to persue the construction in its entirety from Watanabe [28], page 51. However, due to the definition of semimartingales (to be seen) requiring that of local martingales, we will list the required definitions. Similar to Definition 1.17 we have

Definition 1.22. *Let $\mathbb{L}^2_{loc}$ be the space of all real-valued measurable processes $\{\phi\}_{t\geqslant 0}$ on Ω which are adapted to the filtration $\{\mathcal{F}_t\}$ such that for each $T > 0$*
$\int_0^T \phi^2(s,\omega)ds < \infty \quad a.s.$

Similarly, we write $\phi_1 = \phi_2$ if $\int_0^T |\phi_1(t,\omega) - \phi_2(t,\omega)|^2 dt = 0 \quad a.s.$. Again, we take $\phi \in \mathbb{L}^2_{loc}$ as a predictable process without loss of generality. Recall the definition of a *local* martingale per Definition 1.7.

Definition 1.23. *Let $\mathbb{M}^2_c$ be the space of locally square integrable $\{\mathcal{F}_t\}_{t\geqslant 0}$ martingales $\{X_t\}_{t\geqslant 0}$ with $X_0 = 0$. Let $\mathbb{M}^2_{c,loc} = \{X \in \mathbb{M}^2_c : t \mapsto X_t$ is continuous a.s.$\}$*

As stated, the stochastic integral $I(\phi) \in \mathbb{M}^2_{loc}$ is defined in a similar manner to the previous.

Now we look to another formulation of the stochastic integral, which is constructed using $M \in \mathbb{M}^2$ martingales rather than solely Brownian motion.

3.2 Stochastic Integrals with Martingales

In this section, we wish to define the stochastic integral $\int_0^t \phi(s)dM(s)$ where $M \in \mathbb{M}^2$. This is concurrent with our original formulation when M is an $\{\mathcal{F}_t\}_{t\geqslant 0}$-Brownian motion.

Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geqslant 0})$ be a fixed probability space satisfying the usual conditions. Let $M \in \mathbb{M}^2$ with $\langle M \rangle$ as its corresponding quadratic variation. We proceed to begin to define the stochastic integral with respect to martingales in the same manner as Section 3.1, first introducing the following similar spaces

Definition 1.24. *$\mathbb{L}^2(M)$ is the space of real-valued $\{\mathcal{F}_t\}_{t\geqslant 0}$-predictable processes such that for every $T > 0$,*

$$(||\phi||_{2,T}^M)^2 = \mathbb{E}\left[\int_0^T \phi^2(s,\omega)d\langle M\rangle(s)\right] < \infty \tag{1.3.4}$$

Similarly as the previous section we set

$$||\phi||_2^M = \sum_{n=1}^{\infty} \frac{1}{2^n}(||\phi||_{2,n}^M \wedge 1) \tag{1.3.5}$$

Similar to the previous section, it can be shown that $\mathbb{L}^0$ is dense in $\mathbb{L}^2(M)$ in the metric $||\cdot||_2^M$. The stochastic integral with respect to a martingale is thus defined in the same manner as Section 3.1, for a $\phi \in \mathbb{L}^0$ defined by a similar function as (1.3.1), and setting

$$I^M(\phi)(t,\omega) = \sum_{i=0}^{n-1} f_i(\omega)(M(t_{i+1},\omega) - M(t_i,\omega)) + f_n(\omega)(M(t,\omega) - M(t_n,\omega))$$

for $t_n \leqslant t \leqslant t_{n+1}$, $n = 0, 1, ...$ By the same isometry for $||I^M(\phi)||^b = ||\phi||_2^M$ we can take a similar conclusion to Definition 1.21.

Definition 1.25. *$I^M(\phi) \in \mathbb{M}^2$ is called the stochastic integral of $\phi \in \mathbb{L}^2(M)$ with respect to $M \in \mathbb{M}^2$. We denote $I^M(\phi) = \int_0^t \phi(s)dM(s)$*

Note that if $M \in \mathbb{M}_c^2$, then $I^M(\phi) \in \mathbb{M}_c^2$. Additionally, the case when M is an $\mathcal{F}_t$-Brownian motion is the stochastic integral as defined by Definition 1.21. Next there are some properties of the stochastic integrals.

Lemma 1.26. *The stochastic integral $I^m(\phi)$, $\phi \in \mathbb{L}^2(M)$, $M \in \mathbb{M}^2$ has the following properties:*

(a) $I^M(\phi)(0) = 0 \quad a.s.$

(b) For each $t > s \geqslant 0$,

$$\mathbb{E}[I^M(\phi)(t) - I^M(\phi)(s)|\mathcal{F}_s] = 0$$

(c) If $\phi, \psi \in \mathbb{L}^2(M)$ and $c_1, c_2 \in \mathbb{R}$, then

$$I^M(c_1\phi + c_2\psi)(t) = c_1 I^M(\phi)(t) + c_2 I^M(\psi)(t) \quad \textit{for each } t \geqslant 0 \quad a.s.$$

(d) If $\phi, \psi \in \mathbb{L}^2(M)$

$$\mathbb{E}[(I^M(\phi)(t) - I^M(\phi)(s))(I^M(\psi)(t) - I^M(\psi)(s))|\mathcal{F}_s] = \mathbb{E}\left[\int_s^t (\phi \cdot \psi)(r) d\langle M \rangle(r) \Big| \mathcal{F}_s\right] \quad a.s.$$

(e) $X = I^M(\phi)$ *is characterized as the unique* $X \in \mathbb{M}^2$ *such that*

$$\langle X, N \rangle(t) = \int_0^t \phi(r) d\langle M, N \rangle(r)$$

for each $N \in \mathbb{M}^2$ *and all* $t \geqslant 0$*.*

Proof. See Watanabe [28], pages 55-57. □

3.3 Itô's Formula

This section gives all of the relevant forms of Itô's formula for the conditional Mckean-Vlasov diffusion we study in Chapter 4. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geqslant 0})$ be a fixed probability space satisfying the usual conditions, with a given m-dimensional Brownian motion $W_t = (W_1, ..., W_m)^T$, $t \geqslant 0$ defined on the space.

Definition 1.27. *An n-dimensional* $\mathbb{R}^n$*-valued continuous and adapted process* $x(t) = (x_1(t), ..., x_n(t))^T$ *on* $t \geqslant 0$ *with form*

$$x(t) = x(0) + \int_0^t f(s) ds + \int_0^t \sigma(s) dW(s),$$

where $f = (f_1, \ldots, f_n)^T \in \mathbb{L}^1(\mathbb{R}_+, \mathbb{R}^n)$ *and* $\sigma = (\sigma_{ij})_{n \times m} \in \mathbb{L}^2(\mathbb{R}_+, \mathbb{R}^{n \times m})$ *is called an n-dimensional Itô process.*

Let $C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+; \mathbb{R})$ denote family of all real-valued functions $V(x,t)$ such that it is twice continuously differentiable in x and once in t.

Theorem 1.28. *Let* $x(t)$ *be an n-dimensional Itô process on* $t \geqslant 0$ *that satisfies Definition* 1.27*. Let* $V \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ : \mathbb{R})$*. Then,* $V(x(t), t)$ *is a real valued Itô process with its* *stochastic differential given by*

$$dV(x(t),t) = \left[V_t(x(t),t) + V_x(x(t),t)f(t) + \frac{1}{2}Tr(\sigma^T(t)V_{xx}(x(t),t)\sigma(t))\right]dt$$
$$+ V_x(x(t),t)\sigma(t)dW(t) \quad a.s.$$

Proof. See Oksendal [42], Thm 4.2.1. □

Now, we seek to formulate a version of the Itô formula in a similar respect when we have added a Markov chain to our problem. Let $(\alpha_s)_{s\geqslant 0}$ be a right-continuous Markov chain with finite state space $\mathcal{M} = \{1, 2, \dots, m_0\}$ and generator matrix $Q = (q_{i_0 j_0})_{i_0, j_0 \in \mathcal{M}}$ satisfying $q_{i_0 j_0 > 0}$ for $i_0 \neq j_0$ and $\sum_{j_0 \in \mathcal{M}} q_{i_0 j_0} = 0$. We shall assume the $\alpha(\cdot)$ is adapted to the Brownian motion W. Consider the process n-dimensional process X_t defined for each $n = 1, ..., N$

$$dX_t^n = b^n(t, X_t, \alpha_{t-})dt + \sum_{m=1}^{N} \sigma_{nm}(t, X_t, \alpha_{t-})dW_t^m \tag{1.3.6}$$

$$X_0^n = x_0^n, \quad a.s., \tag{1.3.7}$$

for some $x_0^n \in \mathbb{R}$. The following is a formulation of Donnelly [3] for the Itô formula of such a process

Theorem 1.29. *If* $V \in C^{1,3}([0,T] \times \mathbb{R}^n)$ *for each* $i = 1, ..., D$, *then*

$$V(t, X_t, \alpha_t) = V(0, X_0, \alpha_0) + \int_0^t LV(s, X_s, \alpha_{s-})ds$$
$$+ \sum_{n=1}^{N} \int_0^t \frac{\partial V}{\partial x_n}(s, X_s, \alpha_{s-}) \sum_{m=1}^{N} \sigma_{nm}(s, X_s, \alpha_{s-})dW_s^m$$
$$+ \sum_{j \neq i} \int_0^t (V(s, X_s, j) - V(s, X_s, i))dM_{ij}(t)$$

where

$$\begin{aligned} LV(t,x,i) =& \frac{\partial V}{\partial t}(t,x,i) + \sum_{n=1}^{N} \frac{\partial V}{\partial x_n}(t,x,i) b_n(t,x,i) \\ &+ \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \frac{\partial^2 V}{\partial x_n \partial x_m}(t,x,i) \sum_{l=1}^{N} \sigma_{nl}(t,x,i)\sigma_{ml}(t,x,i) \\ &+ \sum_{j=1}^{D} q_{ij}(V(t,x,j) - V(t,x,i)) \end{aligned}$$

for all $(t,x,i) \in [0,T] \times \mathbb{R}^n \times \mathcal{M}$*.*

Proof. See Protter [48], Thm 18, page 278. □

Now, we consider Itô's formula for a very general class of processes we call semimartingales. Semimartingales are the largest class of processes for which the Itô integral can be defined, and as such, they have a special place in the theory. Particularly in the conditional Mckean-Vlasov diffusions we study further on, we use a form of Itô's formula given for a function on a space of pushforward probability measures for a specific class of semimartingales.

Definition 1.30. *A process* x_t *such that* $x_t = x_0 + A_t + M_t$*, where* $x_0 \in \mathcal{F}_0$*,* $\{A_t\}_{t \geqslant 0}$ *is a continuous finite-variational process with* $A_0 = 0$ *that is adapted to* $\mathcal{F}_t$*, and* $\{M_t\} \in \mathbb{M}^2_{c,loc}$*, is called a a continuous semimartingale.*

Theorem 1.31. *If* $f(x) \in C^2(\mathbb{R})$*, then*

$$f(x_t) - f(x_0) = \int_0^t f'(x_s)dA_s + \int_0^t f'(x_s)dM_s + \frac{1}{2}\int_0^t f''(x_s)d\langle M\rangle_s$$

Proof. See Rong [49], Thm. 92. □

Proposition 1.32 (General Itô's Formula)**.** *If* $X = (X^1, X^2, \ldots, X^d)$ *is a d-dimensional semimartingale and* f *is a twice continuously differentiable real valued function on* $\mathbb{R}^d$ *then*

$f(X)$ is a semimartingale, and

$$\begin{aligned} f(X_t) =& f(X_0) + \sum_{i=1}^{d}\int_0^t f_i(X_{s-})dX_s^i + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t f_{i,j}(X_{s-})d[X^i, X^j]_s \\ &+ \sum_{s\leqslant t}\left(\Delta f(X_s) - \sum_{i=1}^{d} f_i(X_{s-})\Delta X_s^i - \frac{1}{2}\sum_{i,j=1}^{d} f_{i,j}(X_{s-})\Delta X_s^i \Delta X_s^j\right) \end{aligned}$$

where X_{t-} is the left limit in t, $\Delta X_t = X_t - X_{t-}$ are the jumps, $d[X]_t$ is the quadratic variation of X_t and $d[X,Y]_t$ is the quadratic covariation of X_t and Y_t, f_i is the first derivative of the ith element, and $f_{i,j}$ is the second derivative of the jth element with first derivative of the ith element. People often write $d[X]_t = (dX_t)^2$ and $d[X,Y]_t = (dX_t)(dY_t)$. This differs from the formula for continuous semimartingales by the use of the left limits X_{t-}, to ensure predictability, and the additional term summing over the jumps of X, which ensures that the jump of the right hand side at time t is $\Delta f(X_t)$.

4 Stochastic Differential Equations

This section gives a brief introduction into stochastic differential equations (SDEs) and the results which give us the existence and uniqueness of their solutions. Of particular note, we introduce the classical SDE, the case with an added jump process, and the case with Markovian switching. This section follows the material provided from Mao and Yuan [38], Ikeda and Watanabe [28], Yong and Zhou [62], Li and Zheng [33], Rong [49], Ma, Protter, and Yong [34], and Platen [47].

4.1 Stochastic Differential Equations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$, with $W_t = (W_1, ..., W_m)^T$, $t \geqslant 0$ an m-dimensional Brownian motion defined on the space. Let both functions $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^n$ and $\sigma : [t_0, t] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be Borel measurable. Consider the following classical stochastic differential equation

$$dX_t = f(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t_0 \leqslant t \leqslant T \tag{1.4.1}$$

with $X_{t_0} \equiv X_0 = \xi$.

Definition 1.33. *Let* $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t_0 \leqslant t \leqslant T}, \mathbb{P})$ *be given, with* W_t *a given* m*-dimensional* $\{\mathcal{F}_t\}_{t \geqslant 0}$*-Brownian motion space with filtration* $\{\mathcal{F}_t\}_{t \geqslant 0}$*, and* ξ $\mathcal{F}_0$*-measurable. An* $\mathbb{R}^d$*-valued stochastic process* $\{X_t\}_{t_0 \leqslant t \leqslant T}$ *is called a solution of* (1.4.1) *if it has the following properties*

a. $\{X_t\}$ *is continuous and adapted to* $\mathcal{F}_t$ *with* $X_0 = \xi \quad \mathbb{P}$*-a.s.,*

b. $\{f(t, X_t)\} \in \mathcal{L}^1(t_0, T; \mathbb{R}^d)$ *and* $\{g(t, X_t)\} \in \mathcal{L}^2(t_0, T; \mathbb{R}^{d \times d})$*,*

c. $X_t = X_0 + \int_{t_0}^t f(s, X_s)ds + \int_{t_0}^t \sigma(s, X_s)dW_s$ *holds with probability 1 for all* $t \in [t_0, T]$*.*

A solution is considered unique if $\mathbb{P}(X_t = \bar{X}_t \text{ for all } t_0 \leqslant t \leqslant T) = 1$ *where* $\bar{X}_t$ *is any other solution* $\{\bar{X}_t\}$*. Under the above conditions, we denote this as a strong solution.*

We now turn to the conditions which guarantee the existence and uniqueness for such a solution.

Theorem 1.34. *Let there be two constants* $C_1, C_2 > 0$ *such that*

(a) for all $x, y \in \mathbb{R}^d$ *and* $t \in [t_0, T]$

$$|f(t,x) - f(t,y)|^2 \wedge |\sigma(t,x) - \sigma(t,y)|^2 \leqslant C|x-y|^2;$$

(b) for all $(t, x) \in [t_0, T] \times \mathbb{R}^d$

$$|f(t,x)|^2 \wedge |\sigma(t,x)|^2 \leqslant C_2(1 + |x|)^2$$

Then there exists a unique solution X_t *to equation* (1.4.1) *and the solution belongs to* $\mathcal{L}^2([t_0, T]; \mathbb{R}^d)$*.*

Condition (a) is commonly referred to as the *Lipschitz* condition while (b) is referred to as the *linear growth* condition.

Proof. For a detailed proof of this result, see Mao and Yuan [38] page 82. □

4.2 Stochastic Differential Equations with Regime Switching

Now, we turn to a set of similar results for SDEs which include Markovian switching. Let $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t_0 \leqslant t \leqslant T}, \mathbb{P})$ be given, with W_t a given m-dimensional and $\mathcal{F}_t$-adapted Brownian motion. Let $(\alpha_s)_{s \geqslant 0}$ be a right-continuous Markov chain with finite state space $\mathcal{M} = \{1, 2, \ldots, m_0\}$ and generator matrix $Q = (q_{i_0 j_0})_{i_0, j_0 \in \mathcal{M}}$ satisfying $q_{i_0 j_0 > 0}$ for $i_0 \neq j_0$

and $\sum_{j_0\in\mathcal{M}} q_{i_0 j_0} = 0$. We shall assume the $\alpha(\cdot)$ is independent to the Brownian motion W. Consider the following SDE with Markovian switching

$$dX_t = f(t, X_t, \alpha_t)dt + \sigma(t, X_t, \alpha_t)dW_t, \quad t_0 \leqslant t \leqslant T \tag{1.4.2}$$

with $X_{t_0} \equiv X_0 = \xi$ and $\alpha_{t_0} = i_0$, where i_0 is an $\mathcal{M}$-valued $\mathcal{F}_{t_0}$-measurable random variable, and $f : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{M} \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{M} \to \mathbb{R}^{d\times m}$.

Theorem 1.35. *Let there be two constants $\bar{C}_1, \bar{C}_2 > 0$ such that*

(a) for all $x, y \in \mathbb{R}^d$ and $t \in [t_0, T]$ and $i \in \mathcal{M}$

$$|f(t, x, i) - f(t, y, i)|^2 \wedge |\sigma(t, x, i) - \sigma(t, y, i)|^2 \leqslant \bar{C}_1 |x - y|^2;$$

(b) for all $(t, x, i) \in [t_0, T] \times \mathbb{R}^d \times \mathcal{M}$

$$|f(t, x, i)|^2 \wedge |\sigma(t, x, i)|^2 \leqslant \bar{C}_2 (1 + |x|)^2$$

Then there exists a unique solution X_t to equation (1.4.2) *and the solution belongs to $\mathcal{L}^2([t_0, T]; \mathbb{R}^d)$.*

Proof. See Mao and Yuan [38], Thm. 3.13. □

5 Backward Stochastic Differential Equations

This section is devoted to introducing a class of stochastic differential equations with terminal conditions conditions called backward and forward-backward stochastic differential equations. We first consider such equations without switching process. For the cases with regime switching, the martingales associated with the Markov chain will be needed to formulate the backward equations.

5.1 Backward Stochastic Differential Equations

Let $g : [0,T] \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \Omega \to \mathbb{R}^q$ and $\xi \in \mathcal{L}^2_{\mathcal{F}_T}(\mathbb{R}^q)$. A double of functions $\big(Y_t, Z_t\big) \in \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p})$ is called a solution of the BSDE

$$\begin{cases} dY_t = g\big(t, Y_t, Z_t\big)dt + Z_t dW_t, \quad 0\leqslant t\leqslant T, \\ Y_T = \xi, \end{cases} \tag{1.5.1}$$

if they satisfy the following equation

$$Y_t = \xi - \int_t^T g\big(s, Y_s, Z_s\big)ds - \int_t^T Z_s dW_s, \qquad 0\leqslant t\leqslant T.$$

Theorem 1.36. *Let for any $(y,z), (\bar{y}, \bar{z}) \in \mathbb{R}^q \times \mathbb{R}^{q\times p}$, $g(t,y,z)$ is $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted with $g(\cdot,0,0) \in \mathcal{L}^2_{\mathcal{F}}(0,T;\mathbb{R}^q)$. Moreover, there exists an $L > 0$ such that*

$$|g(t,y,z) - g(t,\bar{y},\bar{z})|\leqslant L\{|y-\bar{y}| + |z-\bar{z}|\}, \ \forall t \in [0,T], \ \mathbb{P}\text{-}a.s.$$

Then for any given $\xi \in \mathcal{L}^2_{\mathcal{F}_T}(\mathbb{R}^q)$, the BSDE (1.5.1) *admits a unique adapted solution $\big(Y_t, Z_t\big) \in \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p})$.*

Proof. See Yong and Zhou [62], Chapter 7, Thm. 3.2. □

5.2 BSDEs with Regime Switching

Let $g : [0,T] \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \times \Omega \to \mathbb{R}^q$ and $\xi \in \mathcal{L}^2_{\mathcal{F}_T}(\mathbb{R}^q)$. A triple of functions $\big(Y_t, Z_t, \Lambda_t\big) \in \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p}) \times \mathcal{M}^2(0,T;\mathbb{R}^q)$ is called a solution of the BSDE

$$\begin{cases} dY_t = g\big(t, Y_t, Z_t, \alpha_t\big)dt + Z_t dW_t + \Lambda_t \bullet dM_t, \quad 0\leqslant t\leqslant T, \\ Y_T = \xi, \end{cases} \tag{1.5.2}$$

if they satisfy the following equation

$$Y_t = \xi - \int_t^T g\big(s, Y_s, Z_s, \alpha_s\big)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \qquad 0\leqslant t\leqslant T.$$

Theorem 1.37. *Given a pair (ξ, g) satisfying*

(a) $\mathbb{E}|\xi|^2 < \infty$,

(b) $g : \Omega \times [0,T] \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \to \mathbb{R}^q$ *such that*

(i) $g(t, y, z, i)$ *is* $\mathcal{F}_t$*-progressively measurable for all* y, z*,*

(ii) $g(t, 0, 0, i) \in \mathcal{L}^2_{\mathcal{F}}(\mathbb{R}^q)$

(iii) g *satisfies uniform Lipschitz condition in* (y, z)*, i.e* $\exists C > 0$ *such that*

$$|g(t, y_1, z_1, i) - g(t, y_2, z_2, i)|^2 \leqslant C\left(|y_1 - y_2| + |z_1 - z_2|\right),$$

$$\forall y_1, y_2 \in \mathbb{R}^q, z_1, z_2 \in \mathbb{R}^{q\times p} \ \mathbb{P} \otimes Leb \ a.e..$$

Then there exists a unique solution $\left(Y_t, Z_t, \Lambda_t\right) \in \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p}) \times \mathcal{M}^2(0,T;\mathbb{R}^q)$ *to the regime switching BSDE* (1.5.2)*.*

Proof. See Li and Zheng [33], Thm. 5.15. □

6 Forward-Backward Stochastic Differential Equations

This section continue to introducing another class of stochastic differential equations with combined initial-terminal conditions called forward-backward stochastic differential equations. Again, we first consider such equations without regime switching. As seen in the previous section, for the cases with regime switching, the martingales associated with the Markov chain will be needed to formulate the backward equations in the forward-backward systems. Different from the backward stochastic differential equations, the coupled forward-backward stochastic differential equations normally require complicated conditions for their well-posedness.

FBSDEs have been studied extensively since 1990s because of their numerous applications in many areas such as control and game theory [45, 62], mathematical economics [23], and mathematical finance [18, 29]. However, while in many situations the solvability of the original (applied) problems is essentially equivalent to the solvability of certain type of FBSDEs, these FBSDEs are often beyond the scope of any existing frameworks, especially when they are outside the Markovian paradigm, where the PDE tool becomes powerless. In fact, the balance between the regularity of the coefficients and the time duration, as well as the nondegeneracy (of the forward diffusion), has been a longstanding problem in the FBSDE literature, especially in a general non-Markovian framework.

6.1 Forward-Backward Stochastic Differential Equations

There are three main approaches for the wellposedess of the FBSDEs, each of which has its own advantages and disadvantages. In [11] and then [43], the method of contraction mapping was studied, which works well for small time durations. In [34], the four-step-scheme was first introduced to establish the existence and uniqueness of solutions of FBSDEs under non-degenerate condition of the forward equation and some regularity requirements of the coefficients (see also [20, 64]). In [25], the existence and uniqueness of solutions of FBSDEs are proved under monotonicity condition without non-degeneracy condition of the forward equation and use the continuation method. The monotonicity condition is then remarkably weakened and developed in subsequent works [24, 44, 60]. For the progress and related works on FBSDEs, we refer the reader to [35, 37] and the references therein. It is worth noting that these three methods do not cover each other.

Let the coefficient functions

$$
\begin{aligned}
&f : [0,T] \times \Omega \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \to \mathbb{R}^p,\\
&g : [0,T] \times \Omega \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \to \mathbb{R}^q,\\
&\sigma : [0,T] \times \Omega \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \to \mathbb{R}^{q\times p},\\
&h : \Omega \times \mathbb{R}^p \to \mathbb{R}^q
\end{aligned}
$$

be measurable functions with respect to the Borel σ-fields. We consider a measurable process $(X_t, Y_t, Z_t) \in \mathcal{S}^2(0,T;\mathbb{R}^d) \times \mathcal{S}^2(0,T;\mathbb{R}^d) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p})$ which is a solution of the problem

$$
\begin{cases}
X_t = \xi + \int_0^t f(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s)dW_s,\\
Y_t = h(X_T) - \int_t^T g(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \in [0,T].
\end{cases}
\tag{1.6.1}
$$

First, the *Method of Contraction Mapping*. This method, first used by Antonelli [11] and later detailed by Pardoux and Tang [43], works well when the duration T is relatively small.

Second, the *Method of Continuation*. This was a method that can treat non-Markovian FBSDEs with arbitrary duration, initiated by Hu and Peng [25] and Peng and Wu [44], and later developed by Yong [60]. The main assumption for this method is the so-called "monotonicity conditions" on the coefficients, which is restrictive in a different way. This method has been used widely in applications (see, e.g., Wu [54], Wu and Yu [55], Yu [63]) because of its pure probabilistic nature.

Third, the *Four Step Scheme*. This was the first solution method that removed restriction on the time duration for Markovian FBSDEs, initiated by Ma, Protter and Yong [34]; the trade-off is the requirement on the regularity of the coefficients so that a "decoupling" quasi-linear PDE has a classical solution.

Theorem 1.38. *Let*

(a) The functions $f, g, \sigma, \hat{\sigma}$ and h are smooth functions taking values in $\mathbb{R}^p, \mathbb{R}^q, \mathbb{R}^{p\times p}, \mathbb{R}^{q\times p}$ and $\mathbb{R}^q$, respectively, and with first order derivatives in x, y, z being bounded by some constant $L > 0$.

(b) The function σ satisfies

$$\sigma(t,x,y)\sigma(t,x,y)^\top \geqslant \nu(|y|)I, \ \ \forall (t,x,y) \in [0,T] \times \mathbb{R}^p \times \mathbb{R}^q,$$

for some positive continuous function $\nu(\cdot)$.

(c) For each fixed $(t,x,y,z) \in [0,T] \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p}$, the linear map $\hat{\sigma}_z(t,x,y,z) \in \mathcal{L}(\mathbb{R}^{q\times p})$ (the space of all linear transforms on $\mathbb{R}^{q\times p}$) is invertible with the inverse $\hat{\sigma}_z(t,x,y,z)^{-1}$

satisfying

$$\|\hat{\sigma}_z(t,x,y,z)^{-1}\|_{\mathcal{L}(\mathbb{R}^{q\times p})} \leqslant \lambda(|y|),$$

$$(t,x,y,z) \in [0,T] \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p},$$

for some continuous function $\lambda(\cdot)$. *Moreover, for any* $(t,x,y) \in [0,T] \times \mathbb{R}^p \times \mathbb{R}^q$,

$$\{\hat{\sigma}(t,x,y,z) | x \in \mathbb{R}^{q\times p}\} = \mathbb{R}^{q\times p};$$

and there exists a positive continuous function $\kappa(\cdot)$, *such that*

$$\sup\{|z| \big| \hat{\sigma}(t,x,y,z) = O\} \leqslant \kappa(|y|), \forall (t,x,y) \in [0,T] \times \mathbb{R}^p \times \mathbb{R}^q.$$

(d) *There exists a function* μ *and constants* $C > 0$ *and* $\alpha \in (0,1)$, *such that* h *is bounded in* $C^{2+\alpha}(\mathbb{R}^q)$ *and for all* $(t,x,y,z) \in [0,T] \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{n\times m}$,

$$|\sigma(t,x,y)| \leqslant C,$$

$$|f(t,x,y,0)| \leqslant \mu(|y|),$$

$$|g(t,x,0,z)| \leqslant C.$$

Then the forward-backward SDE (1.6.1) *admits a unique adapted solution* (X,Y,Z).

Proof. See Ma, Protter and Yong [34], Thm. 4.1. □

6.2 FBSDEs with Regime Switching

In contrast to the vast literature on FBSDEs, such equations with Markovian switching have not received as needed attention. Although BSDEs with Markovian switching were studied in [22, 33] and were used to formulate stochastic recursive control problems [65], to the best of our knowledge, there is no available well-posedness result even for the FBSDEs with Markovian switching.

To proceed, let the coefficient functions

$$f : [0,T] \times \Omega \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \to \mathbb{R}^p,$$

$$g : [0,T] \times \Omega \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \to \mathbb{R}^q,$$

$$\sigma : [0,T] \times \Omega \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \to \mathbb{R}^{q\times p},$$

$$h : \Omega \times \mathbb{R}^p \times \mathcal{M} \to \mathbb{R}^q$$

be measurable functions with respect to the Borel σ-fields. We consider a measurable process $(X_t, Y_t, Z_t, \Lambda_t) \in \mathcal{S}^2(0,T;\mathbb{R}^p) \times \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p}) \times \mathcal{M}^2(0,T;\mathbb{R}^q)$ which is a solution of the problem

$$\begin{cases} X_t = \xi + \int_0^t f(s, X_s, Y_s, Z_s, \alpha_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s, \alpha_s)dW_s, \\ Y_t = h(X_T, \alpha_T) - \int_t^T g(s, X_s, Y_s, Z_s, \alpha_s)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \quad t \in [0,T]. \end{cases} \tag{1.6.2}$$

7 McKean-Vlasov Stochastic Differential Equations

The topic of weakly interacting systems has a long history beginning with the study of systems of interacting particles by the Austrian physicist Ludwig Boltzamann.

The more mathematically rigorous construction was introduced by Kac and expanded upon by Mckean.

We introduce the following construction for a linear Mckean-Vlasov process. Let $\{\mathcal{F}_t\}_{t\geqslant 0}$ be a filtered probability space with $\mathbb{R}^d$ valued independent Brownian motion $\{W_t\}_{t\geqslant 0}$ Let the function $f(\cdot,\cdot) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be Lipschitz continuous and bounded. Allow X_0 to be an $\mathcal{F}_0$-measurable, $\mathbb{R}^d$ valued random variable with distribution u_0. We introduce the equation

$$\begin{cases} dX_t = dW_t + \int f(X_t, y)u_t(dy)dt \\ X_{t=0} = X_0, \end{cases} \tag{1.7.1}$$

where $u_t(dy)$ is the law of X_t. This is the original specialized linear case as originally studies by Mckean and Sznitman. The following result can be found in Sznitman [53], Thm 1.1.

Theorem 1.39. *The SDE* (1.7.1) *has a unique strong solution.*

The proof ends up being a fixed point argument. With this result established, we introduce the interacting diffusion system,

$$\begin{cases} dX_t^{i,N} = dW_s^i + \frac{1}{N}\sum_{j=1}^N f(X_t^{i,N}, X_t^{j,N})dt, \quad i = 1, 2, ..., N \\ X_0^{i,N} = x_0^i. \end{cases}$$

The desired result is that each $X^{i,N}$ tends to a limit $\bar{X}^{i,N}$ as N goes to infinity. It turns out this $\bar{X}^{i,N}$ is none other than the nonlinear process as described by (1.7.1). Let $\bar{X}^i, i \geqslant 1$ be given as the solution of (using the previous theorem)

$$\bar{X}_t^i = x_0^i + W_0^i + \int_0^t \int f(\bar{X}_s^i, y) u_s(dy) ds, \tag{1.7.2}$$

where $u_s(dy)$ is the law of $\bar{X}_s^i$. Then we have the following result,

Theorem 1.40.

$$\sup_N \sqrt{N}\mathbb{E}\Big[\sup_{t \leqslant T} |X_t^{i,N} - \bar{X}_t^i|\Big] < \infty$$

The particular case detailed above is a special linear version used to introduce the concepts. The limit result of Theorem 1.40 is commonly found in the literature as *propagation of chaos*. A more generalized version of (1.7.1) where the coefficients depend on the law exists, commonly found in the literature as the *nonlinear Mckean-Vlasov process*. This is the general case that we study in our research hereafter, with the added complexities of Markov switching. To introduce the generalized version we first define the *Wasserstein metric*. Let $\mathcal{P}(\mathbb{R}^d)$ denote the set of all probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. $\{\mathbb{P}_t\}_{t \geqslant 0}$ are commonly denoted as *flows of probability measures* in the literature. For each $p \geqslant 1$, let $\mathcal{P}_p(\mathbb{R}^d)$, be the subset of $\mathcal{P}(\mathbb{R}^d)$ containing all measures with bounded p-moments, that is $\mathcal{P}_p(\mathbb{R}^d) = \{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty\}$. We endow $\mathcal{P}(\mathbb{R}^d)$ with the p-Wasserstein metric

$W_p(\cdot,\cdot)$ defined as follows:

$$W_p(\mu,\eta) = \inf\left\{\left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \pi(dxdy)\right)^{1/p} : \pi \in \Pi(\mu,\eta)\right\}, \quad \mu,\eta \in \mathcal{P}_p(\mathbb{R}^d), \tag{1.7.3}$$

where $\Pi(\mu,\eta) = \{\pi \in \mathcal{P}(\mathbb{R}^{2d}) : \pi(A\times\mathbb{R}^d) = \mu(A), \pi(\mathbb{R}^d \times B) = \eta(B),\ \forall A, B \in \mathcal{B}(\mathbb{R}^d)\}$.

Consider the Mckean-Vlasov diffusion

$$dX_t = f(X_t,\mathbb{P}_t)dt + \sigma(X_t,\mathbb{P}_t)dW_t \tag{1.7.4}$$

where W_t is a Brownian motion and $\mathbb{P}_t$ is the law of X_t. Naturally, one might see this equation and think of the notion of differentiation along flows of probability measures. While there is much relevant research which provides answers to this question, Pham [45] derived a form of Itô's formula for flows of measures on a class of semimartingales. Furthermore, regarding the existence and uniqueness of (1.7.4),

Theorem 1.41. *Let $\mathbb{P}_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Assume that for b and σ there exists $C > 0$ such that all $x, y \in \mathbb{R}^d$ and for all $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^d)$ it holds that*

$$|f(x,\mu) - f(y,\nu)| + |\sigma(x,\mu) - \sigma(y,\nu)| \leqslant C(|x-y| + W_2(\mu,\nu)),$$

where W_2 denotes the Wasserstein metric with $p = 2$. Then for any $T > 0$ the SDE (1.7.4) *has a unique strong solution on $[0,T]$.*

The proof for this result can be found in Carmona [15].

With these results established, we are ready to move on to a similar set of equations, this time dependent upon the *conditional law* with added Markovian switching.

CHAPTER 2
Local Solutions to FBSDEs with Regime Switching

Let

$$f : [0,T] \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \to \mathbb{R}^p,$$

$$g : [0,T] \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M} \to \mathbb{R}^q,$$

$$\sigma : [0,T] \times \mathbb{R}^p \times \mathbb{R}^q \times \mathcal{M} \to \mathbb{R}^{q\times p},$$

$$h : \mathbb{R}^p \times \mathcal{M} \to \mathbb{R}^q$$

be measurable functions with respect to the Borel σ-fields. We consider a measurable process $(X_t, Y_t, Z_t, \Lambda_t) \in \mathcal{S}^2(0,T;\mathbb{R}^p) \times \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p}) \times \mathcal{M}^2(0,T;\mathbb{R}^q)$ solution of the problem

$$\begin{cases} X_t &= \xi + \int_0^t f(s, X_s, Y_s, Z_s, \alpha_s)ds + \int_0^t \sigma(s, X_s, Y_s, \alpha_s)dW_s, \\ Y_t &= h(X_T, \alpha_T) + \int_t^T g(s, X_s, Y_s, Z_s, \alpha_s)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \\ t &\in [0,T]. \end{cases} \tag{2.0.1}$$

1 Existence and Uniqueness of Local Solutions

Assumption (A) We say that the functions f, g, h, and σ satisfy Assumption (A) if there exist constants K and L such that

(A1) For all $t \in [0,T]$, $i_0 \in \mathcal{M}$, and $(x,y,z),(x',y',z') \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p}$,

$$\begin{aligned}
&\big|f(t,x,y,z,i_0) - f(t,x,y',z',i_0)\big| \leqslant K\big(|y-y'| + \|z-z'\|\big),\\
&\big|g(t,x,y,z,i_0) - g(t,x',y,z',i_0)\big| \leqslant K\big(|x-x'| + \|z-z'\|\big),\\
&\big|h(x,i_0) - h(x',i_0)\big| \leqslant K|x-x'|,\\
&\big\|\sigma(t,x,y,i_0) - \sigma(t,x',y',i_0)\big\|^2 \leqslant K^2\big(|x-x'|^2 + |y-y'|^2\big).
\end{aligned}$$

(A2) For all $t \in [0,T]$, $(x,y,z,i_0) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M}$, and $(x',y') \in \mathbb{R}^p \times \mathbb{R}^q$,

$$\begin{aligned}
&\big\langle x-x', f(t,x,y,z,i_0) - f(t,x',y,z,i_0)\big\rangle \leqslant K|x-x'|^2,\\
&\big\langle y-y', g(t,x,y,z,i_0) - g(t,x,y',z,i_0)\big\rangle \leqslant K|y-y'|^2.
\end{aligned}$$

(A3) For all $t \in [0,T]$ and $(x,y,z,i_0) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M}$,

$$\begin{aligned}
&\big|f(t,x,y,z,i_0)\big| \leqslant L\big(1 + |x| + |y| + \|z\|\big),\\
&\big|g(t,x,y,z,i_0)\big| \leqslant L\big(1 + |x| + |y| + \|z\|\big),\\
&\big|h(x,i_0)\big| \leqslant L\big(1 + |x|\big),\\
&\big\|\sigma(t,x,y,i_0)\big\| \leqslant L\big(1 + |x| + |y|\big).
\end{aligned}$$

(A4) For all $t \in [0,T], x \in \mathbb{R}^p, y \in \mathbb{R}^q, z \in \mathbb{R}^{q\times p}$, and $i_0 \in \mathcal{M}$, the functions $u \mapsto f(t,u,y,z,i_0)$ and $v \mapsto g(t,x,v,z,i_0)$ are continuous.

Theorem 2.1. *Assume that Assumption (A) holds. Then there exists a constant $C_1 = C_1(K) > 0$, only depending on K, such that for every $T \leqslant C_1$, the equation* (2.0.1) *admits a unique solution.*

Proof. Consider the mapping

$$\begin{aligned}
\Gamma: \quad & \mathcal{S}^2(0,T;\mathbb{R}^p) \times \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p}) \times \mathcal{M}^2(0,T;\mathbb{R}^q)\\
&\to \mathcal{S}^2(0,T;\mathbb{R}^p) \times \mathcal{S}^2(0,T;\mathbb{R}^q) \times \mathcal{L}^2(0,T;\mathbb{R}^{q\times p}) \times \mathcal{M}^2(0,T;\mathbb{R}^q),
\end{aligned}$$

$$(X_t, Y_t, Z_t, \Lambda_t) \mapsto (\bar{X}_t, \bar{Y}_t, \bar{Z}_t, \bar{\Lambda}_t),$$

where $(\bar{X}_t, \bar{Y}_t, \bar{Z}_t, \bar{\Lambda}_t)$ is defined as follows

$$\bar{X}_t = \xi + \int_0^t f(s, \bar{X}_s, Y_s, Z_s, \alpha_s)ds + \int_0^t \sigma(s, \bar{X}_s, Y_s, \alpha_s)dW_s, \quad t \in [0, T], \tag{2.1.1}$$

and

$$\bar{Y}_t = h(\bar{X}_T, \alpha_T) + \int_t^T g(s, \bar{X}_s, \bar{Y}_s, \bar{Z}_s, \alpha_s)ds - \int_t^T \bar{Z}_s dW_s - \int_t^T \bar{\Lambda}_s \bullet dM_s, \quad t \in [0, T]. \tag{2.1.2}$$

Note that $\bar{X}_t$ is defined as the solution of (2.1.1) which is a forward SDE, whereas $(\bar{Y}_t, \bar{Z}_s, \bar{\Lambda}_s)$ are defined as the solution of the BSDE (2.1.2). The solutions to BSDEs with regime switching are investigated in [33] with respect to the the σ-field generated by the Brownian and Markov chain only. However, in our case with $\sigma\{\xi\}$ is included in $\mathcal{F}_0$, the martingale representation theorem (see [21, Theorem B.4.6]) is still valid. That is, every square integrable $\{\mathcal{F}_t\}$-martingale can be represented as a stochastic integral with respect to W_t and M_t. As a consequence, the existence and uniqueness of solutions to BSDEs given in [33, Theorem 5.15] can be extended to our case. Therefore, (2.1.2) has a unique solution and Γ is well-defined.

Next, we will show that there exists a constant $C_1 = C_1(K) > 0$, only depending on K, such that for every $T \leqslant C_1$, Γ is a contraction. Without any loss of generality we can assume that $T \leqslant 1$. Let $(S_t, U_t, V_t, \Upsilon_t) \in \mathcal{S}^2(0, T; \mathbb{R}^p) \times \mathcal{S}^2(0, T; \mathbb{R}^q) \times \mathcal{L}^2(0, T; \mathbb{R}^{q \times p}) \times \mathcal{M}^2(0, T; \mathbb{R}^q)$ and denote $(\bar{S}_t, \bar{U}_t, \bar{V}_t, \bar{\Upsilon}_t) = \Gamma(S_t, U_t, V_t, \Upsilon_t)$. In view of Assumption (A) and the Itô's formula for $|\bar{X} - \bar{S}|^2$, there exists a constant γ_K, only depending on K, such that

$$
\begin{aligned}
\left|\bar{X}_t-\bar{S}_t\right|^2 = {}& 2\int_0^t \langle \bar{X}_s-\bar{S}_s, f(s,\bar{X}_s,Y_s,Z_s,\alpha_s)-f(s,\bar{S}_s,U_s,V_s,\alpha_s)\rangle ds \\
&+\ 2\int_0^t \langle \bar{X}_s-\bar{S}_s, (\sigma(s,\bar{X}_s,Y_s,\alpha_s)-\sigma(s,\bar{S}_s,U_s,\alpha_s))dW_s\rangle \\
&+\int_0^t \left|\sigma(s,\bar{X}_s,Y_s,\alpha_s)-\sigma(s,\bar{S}_s,U_s,\alpha_s)\right|^2 ds \\
= {}& 2\int_0^t \langle \bar{X}_s-\bar{S}_s, f(s,\bar{X}_s,Y_s,Z_s,\alpha_s)-f(s,\bar{X}_s,U_s,V_s,\alpha_s)\rangle ds \\
&+\ 2\int_0^t \langle \bar{X}_s-\bar{S}_s, f(s,\bar{X}_s,U_s,V_s,\alpha_s)-f(s,\bar{S}_s,U_s,V_s,\alpha_s)\rangle ds \\
&+\ 2\int_0^t \langle \bar{X}_s-\bar{S}_s, (\sigma(s,\bar{X}_s,Y_s,\alpha_s)-\sigma(s,\bar{S}_s,U_s,\alpha_s))dW_s\rangle \\
&+\int_0^t \left|\sigma(s,\bar{X}_s,Y_s,\alpha_s)-\sigma(s,\bar{S}_s,U_s,\alpha_s)\right|^2 ds \\
\leqslant {}& 2K\int_0^t \left|\bar{X}_s-\bar{S}_s\right|\Big(\left|Y_s-U_s\right|+\left\|Z_s-V_s\right\|\Big)ds \\
&+\ 2K\int_0^t \left|\bar{X}_s-\bar{S}_s\right|^2 ds \\
&+\ 2\int_0^t \langle \bar{X}_s-\bar{S}_s, (\sigma(s,\bar{X}_s,Y_s,\alpha_s)-\sigma(s,\bar{S}_s,U_s,\alpha_s))dW_s\rangle \\
&+\ K^2\int_0^t \Big(\left|\bar{X}_s-\bar{S}_s\right|^2+\left|Y_s-U_s\right|^2\Big)ds
\end{aligned}
$$

Hence, there exists a constant γ_K, only depending on K, such that

$$
\begin{aligned}
&\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2 \\
&\leqslant\gamma_K\Bigg[\mathbb{E}\int_0^T \left|\bar{X}_s-\bar{S}_s\right|\Big(\left|\bar{X}_s-\bar{S}_s\right|+\left|Y_s-U_s\right|+\left\|Z_s-V_s\right\|\Big)ds \\
&\quad+\mathbb{E}\int_0^T \Big(\left|\bar{X}_s-\bar{S}_s\right|^2+\left|Y_s-U_s\right|^2\Big)ds\Bigg] \\
&\quad+2\mathbb{E}\Bigg[\sup_{[0,T]}\left|\int_0^t \Big\langle \bar{X}_s-\bar{S}_s, \Big(\sigma\big(s,\bar{X}_s,Y_s,\alpha_s\big)-\sigma\big(s,\bar{S}_s,U_s,\alpha_s\big)\Big)dW_s\Big\rangle\right|\Bigg].
\end{aligned}
$$

In what follows, we can modify γ_K if necessary, so this constant may vary from place to place. By Burkholder-Davis-Gundy's inequality and Young's inequality,

$$\begin{aligned}
&\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2\\
\leqslant&\gamma_K\bigg\{\mathbb{E}\int_0^T\left|\bar{X}_s-\bar{S}_s\right|\Big(\left|\bar{X}_s-\bar{S}_s\right|+\left|Y_s-U_s\right|+\left\|Z_s-V_s\right\|\Big)ds\\
&+\mathbb{E}\int_0^T\Big(\left|\bar{X}_s-\bar{S}_s\right|^2+\left|Y_s-U_s\right|^2\Big)ds\\
&+\mathbb{E}\bigg[\bigg|\int_0^T\Big\langle\bar{X}_s-\bar{S}_s,\Big(\sigma\big(s,\bar{X}_s,Y_s,\alpha_s\big)-\sigma\big(s,\bar{S}_s,U_s,\alpha_s\big)\Big)dW_s\Big\rangle\bigg|^2\bigg]^{1/2}\bigg\}\\
\leqslant&\gamma_K\bigg\{\mathbb{E}\int_0^T\left|\bar{X}_s-\bar{S}_s\right|\Big(\left|\bar{X}_s-\bar{S}_s\right|+\left|Y_s-U_s\right|+\left\|Z_s-V_s\right\|\Big)ds\\
&+\mathbb{E}\int_0^T\Big(\left|\bar{X}_s-\bar{S}_s\right|^2+\left|Y_s-U_s\right|^2\Big)ds\\
&+\mathbb{E}\bigg[\int_0^T\Big|\bar{X}_s-\bar{S}_s\Big|^2\Big\|\sigma\big(s,\bar{X}_s,Y_s,\alpha_s\big)-\sigma\big(s,\bar{S}_s,U_s,\alpha_s\big)\Big\|^2ds\bigg]^{1/2}\bigg\}\\
\leqslant&\gamma_K\bigg\{\mathbb{E}\int_0^T\left|\bar{X}_s-\bar{S}_s\right|\Big(\left|\bar{X}_s-\bar{S}_s\right|+\left|Y_s-U_s\right|+\left\|Z_s-V_s\right\|\Big)ds\\
&+\mathbb{E}\int_0^T\Big(\left|\bar{X}_s-\bar{S}_s\right|^2+\left|Y_s-U_s\right|^2\Big)ds\\
&+\mathbb{E}\bigg[\int_0^T\left|\bar{X}_s-\bar{S}_s\right|^2\Big(\left|\bar{X}_s-\bar{S}_s\right|^2+\left|Y_s-U_s\right|^2\Big)ds\bigg]^{1/2}\bigg\}.
\end{aligned}$$

By Cauchy-Schwarz's inequality we arrive at

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2\leqslant\gamma_K\sqrt{T}\bigg(\mathbb{E}\sup_{0\leqslant s\leqslant T}\left|\bar{X}_s-\bar{S}_s\right|^2+\mathbb{E}\sup_{0\leqslant s\leqslant T}\left|Y_s-U_s\right|^2+\mathbb{E}\int_0^T\left\|Z_s-V_s\right\|^2ds\bigg).$$

As a consequence,

$$\big(1-\gamma_K\sqrt{T}\big)\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2\leqslant\gamma_K\sqrt{T}\bigg(\mathbb{E}\sup_{0\leqslant s\leqslant T}\left|Y_s-U_s\right|^2+\mathbb{E}\int_0^T\left\|Z_s-V_s\right\|^2ds\bigg).\quad(2.1.3)$$

Next, by Itô's formula for non-continuous semimartingales,

$$
\begin{aligned}
&\left|\bar{Y}_t-\bar{U}_t\right|^2+\int_t^T\left\|\bar{Z}_s-\bar{V}_s\right\|^2 ds+\int_t^T\left|\bar{\Lambda}_s-\bar{\Upsilon}_s\right|^2 \bullet d[M]_s\\
&=\left|h(\bar{X}_T,\alpha_T)-h(\bar{S}_T,\alpha_T)\right|^2+2\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s, g\big(s,\bar{X}_s,\bar{Y}_s,\bar{Z}_s,\alpha_s\big)-g\big(s,\bar{S}_s,\bar{U}_s,\bar{V}_s,\alpha_s\big)\Big\rangle ds\\
&\quad-2\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s,\big(\bar{Z}_s-\bar{V}_s\big)dW_s\Big\rangle-2\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s,\bar{\Lambda}_s-\bar{\Upsilon}_s\Big\rangle \bullet dM_s. \qquad (2.1.4)
\end{aligned}
$$

Note that

$$
\mathbb{E}\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s,\big(\bar{Z}_s-\bar{V}_s\big)dW_s\Big\rangle=\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s,\bar{\Lambda}_s-\bar{\Upsilon}_s\Big\rangle \bullet dM_s=0,
$$

$\forall\ i_0, j_0 \in \mathcal{M}$, by using Assumption (A) and Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
&\mathbb{E}\int_t^T\left\|\bar{Z}_s-\bar{V}_s\right\|^2 ds+\mathbb{E}\int_t^T\left|\bar{\Lambda}_s-\bar{\Upsilon}_s\right|^2 \bullet d[M]_s\\
&=\mathbb{E}\left|h(\bar{X}_T,\alpha_T)-h(\bar{S}_T,\alpha_T)\right|^2+2\mathbb{E}\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s, g\big(s,\bar{X}_s,\bar{Y}_s,\bar{Z}_s,\alpha_s\big)-g\big(s,\bar{S}_s,\bar{Y}_s,\bar{Z}_s,\alpha_s\big)\Big\rangle ds\\
&\quad+2\mathbb{E}\int_t^T\Big\langle \bar{Y}_s-\bar{U}_s, g\big(s,\bar{S}_s,\bar{Y}_s,\bar{Z}_s,\alpha_s\big)-g\big(s,\bar{S}_s,\bar{U}_s,\bar{V}_s,\alpha_s\big)\Big\rangle ds\\
&\leqslant\gamma_K\left[\mathbb{E}\left|\bar{X}_T-\bar{S}_T\right|^2+\mathbb{E}\int_t^T\left|\bar{Y}_s-\bar{U}_s\right|\Big(\left|\bar{X}_s-\bar{S}_s\right|+\left|\bar{Y}_s-\bar{U}_s\right|+\left\|\bar{Z}_s-\bar{V}_s\right\|\Big)ds\right]\\
&\leqslant\gamma_K\Big((1+T)\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2+T\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{Y}_t-\bar{U}_t\right|^2\Big)+\frac{1}{2}\mathbb{E}\int_t^T\left\|\bar{Z}_s-\bar{V}_s\right\|^2 ds.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\mathbb{E}\int_t^T\left\|\bar{Z}_s-\bar{V}_s\right\|^2 ds+\mathbb{E}\int_t^T\left|\bar{\Lambda}_s-\bar{\Upsilon}_s\right|^2 \bullet d[M]_s\\
&\leqslant\gamma_K\Big((1+T)\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2+T\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{Y}_t-\bar{U}_t\right|^2\Big). \qquad (2.1.5)
\end{aligned}
$$

By simple estimate, we can prove the following inequalities

$$\mathbb{E}\Big[\int_0^T \big|\bar{Y}_s-\bar{U}_s\big|^2\big\|\bar{Z}_s-\bar{V}_s\big\|^2 ds\Big]^{1/2}\leqslant\mathbb{E}\Big[\frac{1}{4\gamma_K}\sup_{0\leqslant s\leqslant T}\big|\bar{Y}_s-\bar{U}_s\big|^2+4\gamma_K\int_0^T\big\|\bar{Z}_s-\bar{V}_s\big\|^2 ds\Big],$$

$$\mathbb{E}\Big[\int_0^T \big|\bar{Y}_s-\bar{U}_s\big|^2\big|\bar{\Lambda}_s-\bar{\Upsilon}_s\big|^2\bullet d\big[M\big]_s\Big]^{1/2}$$
$$\leqslant\mathbb{E}\Big[\frac{1}{4\gamma_K}\sup_{0\leqslant s\leqslant T}\big|\bar{Y}_s-\bar{U}_s\big|^2+4\gamma_K\int_0^T\big|\bar{\Lambda}_s-\bar{\Upsilon}_s\big|^2\bullet d\big[M\big]_s\Big].$$

Using these inequalities and Burkholder-Davis-Gundy's inequality and (2.1.5) for (2.1.4) we get

$$\begin{aligned}
&\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{Y}_t-\bar{U}_t\big|^2\\
\leqslant&\gamma_K\Big\{\mathbb{E}\big|\bar{X}_T-\bar{S}_T\big|^2+\mathbb{E}\int_0^T\big|\bar{Y}_s-\bar{U}_s\big|\Big(\big|\bar{X}_s-\bar{S}_s\big|+\big|\bar{Y}_s-\bar{U}_s\big|+\big\|\bar{Z}_s-\bar{V}_s\big\|\Big)ds\\
&\quad+\mathbb{E}\Big[\int_0^T\big|\bar{Y}_s-\bar{U}_s\big|^2\big\|\bar{Z}_s-\bar{V}_s\big\|^2ds\Big]^{1/2}+\mathbb{E}\Big[\int_0^T\big|\bar{Y}_s-\bar{U}_s\big|^2\big|\bar{\Lambda}_s-\bar{\Upsilon}_s\big|^2\bullet d\big[M\big]_s\Big]^{1/2}\Big\}\\
\leqslant&\gamma_K\Big((1+T)\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{X}_t-\bar{S}_t\big|^2+T\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{Y}_t-\bar{U}_t\big|^2\Big)+\frac{1}{2}\mathbb{E}\int_t^T\big\|\bar{Z}_s-\bar{V}_s\big\|^2ds\\
&+\mathbb{E}\Big[\frac{1}{4\gamma_K}\sup_{0\leqslant s\leqslant T}\big|\bar{Y}_s-\bar{U}_s\big|^2+4\gamma_K\int_0^T\big\|\bar{Z}_s-\bar{V}_s\big\|^2ds\Big]\\
&+\mathbb{E}\Big[\frac{1}{4\gamma_K}\sup_{0\leqslant s\leqslant T}\big|\bar{Y}_s-\bar{U}_s\big|^2+4\gamma_K\int_0^T\big|\bar{\Lambda}_s-\bar{\Upsilon}_s\big|^2\bullet d\big[M\big]_s\Big].\\
\leqslant&\gamma_K\Big((1+T)\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{X}_t-\bar{S}_t\big|^2+T\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{Y}_t-\bar{U}_t\big|^2\Big)
\end{aligned}$$

This implies

$$(1-\gamma_K T)\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{Y}_t-\bar{U}_t\big|^2\leqslant\gamma_K(1+T)\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|\bar{X}_t-\bar{S}_t\big|^2. \tag{2.1.6}$$

Combining (2.1.3), (2.1.5) and (2.1.6) we obtain

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2\leqslant\frac{\gamma_K\sqrt{T}}{1-\gamma_K\sqrt{T}}\left(\mathbb{E}\sup_{0\leqslant s\leqslant T}\left|Y_s-U_s\right|^2+\mathbb{E}\int_0^T\left\|Z_s-V_s\right\|^2ds\right),$$

$$\mathbb{E}\int_t^T\left\|\bar{Z}_s-\bar{V}_s\right\|^2ds+\mathbb{E}\int_t^T\left|\bar{\Lambda}_s-\bar{\Upsilon}_s\right|^2\bullet d[M]_s \leqslant\gamma_K(1+T)\left(\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2+\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{Y}_t-\bar{U}_t\right|^2\right).$$

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{Y}_t-\bar{U}_t\right|^2\leqslant\frac{\gamma_K(1+T)}{1-\gamma_K T}\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|\bar{X}_t-\bar{S}_t\right|^2.$$

It is easily seen that there exists a constant $C_1=C_1(K)>0$ only depending on K such that for $T\leqslant C_1$, the mapping Γ is contractive from $\mathcal{S}^2(0,T;\mathbb{R}^p)\times\mathcal{S}^2(0,T;\mathbb{R}^q)\times\mathcal{L}^2(0,T;\mathbb{R}^{q\times p})\times\mathcal{M}^2(0,T;\mathbb{R}^q)$ to itself. By the contraction mapping theorem, there exists a unique $\{\mathcal{F}_t\}$-progressively measurable solution to (2.0.1). $\square$

Proposition 2.2. *Under Assumption (A) there exists a constant $C_2=C_2(K)\in(0,C_1]$ only depending on K such that for every $T\leqslant C_2$, for every quadruplet of functions $(\tilde{f},\tilde{g},\tilde{h},\tilde{\sigma})$ satisfying Assumption (A) with the same constants K and L as (f,g,h,σ), for every $A\in\mathcal{F}_0$, and for all $\mathcal{F}_0$-measurable random vectors ξ and $\tilde{\xi}$ with finite second moment, we have the following estimate*

$$\begin{aligned}&\mathbb{E}\left(\mathbb{1}_A\sup_{0\leqslant s\leqslant T}\left|X_s-\tilde{X}_s\right|^2\right)+\mathbb{E}\left(\mathbb{1}_A\sup_{0\leqslant s\leqslant T}\left|Y_s-\tilde{Y}_s\right|^2\right)+\mathbb{E}\int_0^T\mathbb{1}_A\left\|Z_s-\tilde{Z}_s\right\|^2ds\\&\qquad+\mathbb{E}\int_0^T\mathbb{1}_A\left|\Lambda_s-\tilde{\Lambda}_s\right|^2\bullet d\big[M\big]_s\\&\leqslant\gamma_K\Bigg\{\mathbb{E}\left(\mathbb{1}_A\left|\xi-\tilde{\xi}\right|^2\right)+\mathbb{E}\left(\mathbb{1}_A\left|h-\tilde{h}\right|^2(X_T)\right)+\mathbb{E}\int_0^T\mathbb{1}_A\left\|\sigma-\tilde{\sigma}\right\|^2(s,X_s,Y_s,\alpha_s)ds\\&\qquad+\mathbb{E}\left[\left(\int_0^T\mathbb{1}_A\left|f-\tilde{f}\right|(s,X_s,Y_s,Z_s,\alpha_s)ds\right)^2+\left(\int_0^T\mathbb{1}_A\left|g-\tilde{g}\right|(s,X_s,Y_s,Z_s,\alpha_s)ds\right)^2\right]\Bigg\}\end{aligned}\tag{2.1.7}$$

where γ_K is a constant only depending on K, the processes $\big(X_s,Y_s,Z_s,\Lambda_s\big)$ and $\big(\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\tilde{\Lambda}_s\big)$, $0\leqslant s\leqslant T$, are respectively the solutions to the problems associated to the coefficients (f,g,h,σ) and $(\tilde{f},\tilde{g},\tilde{h},\tilde{\sigma})$ and to the initial conditions $(0,\xi)$ and $(0,\tilde{\xi})$.

Proof. According to Itô's formula and Burkholder-Davis-Gundy inequality, we have

$$
\begin{aligned}
&\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|\tilde{X}_t - X_t\big|^2\Big)\\
\leqslant&\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi}-\xi\big|^2\Big) + \mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{\sigma}(s,\tilde{X}_s,\tilde{Y}_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big\|^2 ds\\
&\quad + 2\mathbb{E}\sup_{0\leqslant t\leqslant T}\bigg(\int_0^t \mathbb{1}_A\Big\langle \tilde{X}_s - X_s, \tilde{f}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s) - f(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds\bigg)\\
&\quad + 2\mathbb{E}\sup_{0\leqslant t\leqslant T}\bigg(\int_0^t \mathbb{1}_A\Big\langle \tilde{X}_s - X_s, \big(\tilde{\sigma}(s,\tilde{X}_s,\tilde{Y}_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big)dW_s\Big\rangle\bigg)\\
\leqslant&\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi}-\xi\big|^2\Big) + \mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{\sigma}(s,\tilde{X}_s,\tilde{Y}_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big\|^2 ds\\
&\quad + 2\mathbb{E}\sup_{0\leqslant t\leqslant T}\bigg(\int_0^t \mathbb{1}_A\Big\langle \tilde{X}_s - X_s, \tilde{f}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s) - f(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds\bigg)\\
&\quad + \gamma\mathbb{E}\bigg(\int_0^t \mathbb{1}_A\big|\tilde{X}_s - X_s\big|^2\big\|\tilde{\sigma}(s,\tilde{X}_s,\tilde{Y}_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big\|^2 ds\bigg)^{1/2}.
\end{aligned}
$$

By using simple estimates for the last two terms in the above inequalities and modifying γ we arrive at

$$
\begin{aligned}
&\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|\tilde{X}_t - X_t\big|^2\Big)\\
\leqslant&\gamma\bigg[\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi}-\xi\big|^2\Big) + \mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{\sigma}(s,\tilde{X}_s,\tilde{Y}_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big\|^2 ds\\
&\quad + \mathbb{E}\sup_{0\leqslant t\leqslant T}\bigg(\int_0^t \mathbb{1}_A\Big\langle \tilde{X}_s - X_s, \tilde{f}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s) - \tilde{f}(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds\bigg)\\
&\quad + \mathbb{E}\bigg(\int_0^t \mathbb{1}_A\big|\tilde{f}(s,X_s,Y_s,Z_s,\alpha_s) - f(s,X_s,Y_s,Z_s,\alpha_s)\big|ds\bigg)^2\bigg].
\end{aligned}
$$

Therefore, Assumption (A) implies that there exists a constant γ_K only depending on K such that

$$
\begin{aligned}
&\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T} \big|\tilde{X}_t - X_t\big|^2\Big)\\
&\leqslant \gamma_K\Bigg[\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi}-\xi\big|^2\Big) + \mathbb{E}\int_0^T \mathbb{1}_A\Big(\big|\tilde{X}_s - X_s\big|^2 + \big|\tilde{Y}_s - Y_s\big|^2\Big)ds\\
&\qquad + \mathbb{E}\int_0^T \mathbb{1}_A\big|\tilde{X}_s - X_s\big|\big\|\tilde{Z}_s - Z_s\big\|ds + \mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{\sigma}(s,X_s,Y_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big\|^2 ds\\
&\qquad + \mathbb{E}\Big(\int_0^T \mathbb{1}_A\big|\tilde{f}(s,X_s,Y_s,Z_s,\alpha_s) - f(s,X_s,Y_s,Z_s,\alpha_s)\big|ds\Big)^2\Bigg]. \qquad (2.1.8)\\
&\leqslant \gamma_K\Bigg[\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi}-\xi\big|^2\Big) + \mathbb{E}\int_0^T \mathbb{1}_A\Big(\big|\tilde{X}_s - X_s\big|^2 + \big|\tilde{Y}_s - Y_s\big|^2 + \big\|\tilde{Z}_s - Z_s\big\|^2\Big)ds\\
&\qquad + \mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{\sigma}(s,X_s,Y_s,\alpha_s) - \sigma(s,X_s,Y_s,\alpha_s)\big\|^2 ds\\
&\qquad + \mathbb{E}\Big(\int_0^T \mathbb{1}_A\big|\tilde{f}(s,X_s,Y_s,Z_s,\alpha_s) - f(s,X_s,Y_s,Z_s,\alpha_s)\big|ds\Big)^2\Bigg].
\end{aligned}
$$

Next, similar to (2.1.4), by Itô's formula for non-continuous semimartingales, for $0\leqslant t\leqslant T$

$$
\begin{aligned}
&\big|Y_t - \tilde{Y}_t\big|^2 + \int_t^T \big\|\tilde{Z}_s - Z_s\big\|^2 ds + \int_t^T \big|\tilde{\Lambda}_s - \Lambda_s\big|^2 \bullet d\big[M\big]_s\\
&= \big|\tilde{h}(\tilde{X}_T,\alpha_T) - h(X_T,\alpha_T)\big|^2 + 2\int_t^T \Big\langle \tilde{Y}_s - Y_s, \tilde{g}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s) - g(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds\\
&\quad - 2\int_t^T \Big\langle \tilde{Y}_s - Y_s, \big(\tilde{Z}_s - Z_s\big)dW_s\Big\rangle - 2\int_t^T \Big\langle \tilde{Y}_s - Y_s, \tilde{\Lambda}_s - \Lambda_s\Big\rangle \bullet dM_s.
\end{aligned}
$$

Hence, for any $A \in \mathcal{F}_0$ and $0\leqslant t\leqslant T$,

$$
\begin{aligned}
&\mathbb{E}\Big(\mathbb{1}_A\big|Y_t - \tilde{Y}_t\big|^2\Big) + \mathbb{E}\int_t^T \mathbb{1}_A\big\|\tilde{Z}_s - Z_s\big\|^2 ds + \mathbb{E}\int_t^T \mathbb{1}_A\big|\tilde{\Lambda}_s - \Lambda_s\big|^2 \bullet d\big[M\big]_s\\
&= \mathbb{E}\Big(\mathbb{1}_A\big|\tilde{h}(\tilde{X}_T,\alpha_T) - h(X_T,\alpha_T)\big|^2\Big)\\
&\quad + 2\mathbb{E}\int_t^T \mathbb{1}_A\Big\langle \tilde{Y}_s - Y_s, \tilde{g}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s) - g(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds.
\end{aligned}
$$

In addition, by using Burkholder-Davis-Gundy inequality we get

$$\begin{aligned}\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|Y_t-\tilde{Y}_t\big|^2\Big)\leqslant&\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{h}(\tilde{X}_T,\alpha_T)-h(X_T,\alpha_T)\big|^2\Big)\\&+2\mathbb{E}\sup_{0\leqslant t\leqslant T}\int_t^T \mathbb{1}_A\Big\langle \tilde{Y}_s-Y_s,\tilde{g}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s)-g(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds\\&+\gamma\mathbb{E}\Big(\int_0^T \mathbb{1}_A\big|\tilde{Y}_s-Y_s\big|^2\big\|\tilde{Z}_s-Z_s\big\|^2ds\Big)^{1/2}+\gamma\mathbb{E}\Big(\int_0^T \mathbb{1}_A\big|\tilde{Y}_s-Y_s\big|^2\big|\tilde{\Lambda}_s-\Lambda_s\big|^2\bullet d\big[M\big]_s\Big)^{1/2}\end{aligned}$$

which, together with the above equation, yields

$$\begin{aligned}&\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|Y_t-\tilde{Y}_t\big|^2\Big)+\mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{Z}_s-Z_s\big\|^2ds+\mathbb{E}\int_0^T \mathbb{1}_A\big|\tilde{\Lambda}_s-\Lambda_s\big|^2\bullet d\big[M\big]_s\\&\leqslant\gamma\Bigg[\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{h}(\tilde{X}_T,\alpha_T)-h(X_T,\alpha_T)\big|^2\Big)\\&\qquad+\mathbb{E}\sup_{0\leqslant t\leqslant T}\int_t^T \mathbb{1}_A\Big\langle \tilde{Y}_s-Y_s,\tilde{g}(s,\tilde{X}_s,\tilde{Y}_s,\tilde{Z}_s,\alpha_s)-\tilde{g}(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds\\&\qquad+\mathbb{E}\Big(\int_0^T \mathbb{1}_A\Big|\tilde{g}(s,X_s,Y_s,Z_s,\alpha_s)-g(s,X_s,Y_s,Z_s,\alpha_s)\Big|ds\Big)^2\Bigg].\end{aligned}\tag{2.1.9}$$

Combining (2.1.8) and (2.1.9) and using Assumption (A) lead to

$$\begin{aligned}&\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|X_t-\tilde{X}_t\big|^2\Big)+\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|Y_t-\tilde{Y}_t\big|^2\Big)+\mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{Z}_s-Z_s\big\|^2ds\\&\qquad\qquad+\mathbb{E}\int_0^T \mathbb{1}_A\big|\tilde{\Lambda}_s-\Lambda_s\big|^2\bullet d\big[M\big]_s\\&\leqslant\gamma_K\Bigg[\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi}-\xi\big|^2\Big)+\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{h}(\tilde{X}_T,\alpha_T)-h(X_T,\alpha_T)\big|^2\Big)\\&\qquad+\mathbb{E}\int_0^T \mathbb{1}_A\Big\|\tilde{\sigma}(s,X_s,Y_s,\alpha_s)-\sigma(s,X_s,Y_s,\alpha_s)\Big\|^2ds\\&\qquad+\mathbb{E}\Big(\int_0^T \mathbb{1}_A\Big|\tilde{f}(s,X_s,Y_s,Z_s,\alpha_s)-f(s,X_s,Y_s,Z_s,\alpha_s)\Big|ds\Big)^2\\&\qquad+\mathbb{E}\Big(\int_0^T \mathbb{1}_A\Big|\tilde{g}(s,X_s,Y_s,Z_s,\alpha_s)-g(s,X_s,Y_s,Z_s,\alpha_s)\Big|ds\Big)^2\\&\qquad+\mathbb{E}\int_0^T \mathbb{1}_A\big|\tilde{X}_s-X_s\big|^2ds+\mathbb{E}\int_0^T \mathbb{1}_A\big|\tilde{Y}_s-Y_s\big|^2ds+\mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{Z}_s-Z_s\big\|^2ds\Bigg]\end{aligned}\tag{2.1.10}$$

for some constant γ_K only depending on K. This implies that there exists a constant $C_2 = C_2(K)$ only depending on K such that for any $T \leqslant C_2$,

$$
\begin{aligned}
&\mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|X_t - \tilde{X}_t\big|^2\Big) + \mathbb{E}\Big(\mathbb{1}_A \sup_{0\leqslant t\leqslant T}\big|Y_t - \tilde{Y}_t\big|^2\Big) + \mathbb{E}\int_0^T \mathbb{1}_A\big\|\tilde{Z}_s - Z_s\big\|^2 ds \\
&\qquad\qquad + \mathbb{E}\int_0^T \mathbb{1}_A\big|\tilde{\Lambda}_s - \Lambda_s\big|^2 \bullet d\big[M\big]_s \\
&\leqslant \gamma_K\Bigg[\mathbb{E}\Big(\mathbb{1}_A\big|\tilde{\xi} - \xi\big|^2\Big) + \mathbb{E}\Big(\mathbb{1}_A\big|\tilde{h}(\tilde{X}_T, \alpha_T) - h(X_T, \alpha_T)\big|^2\Big) \qquad (2.1.11)\\
&\quad + \mathbb{E}\int_0^T \mathbb{1}_A\Big\|\tilde{\sigma}(s, X_s, Y_s, \alpha_s) - \sigma(s, X_s, Y_s, \alpha_s)\Big\|^2 ds \\
&\quad + \mathbb{E}\bigg(\int_0^T \mathbb{1}_A\Big|\tilde{f}(s, X_s, Y_s, Z_s, \alpha_s) - f(s, X_s, Y_s, Z_s, \alpha_s)\Big|ds\bigg)^2 \\
&\quad + \mathbb{E}\bigg(\int_0^T \mathbb{1}_A\Big|\tilde{g}(s, X_s, Y_s, Z_s, \alpha_s) - g(s, X_s, Y_s, Z_s, \alpha_s)\Big|ds\bigg)^2. \qquad (2.1.12)
\end{aligned}
$$

This completes the proof. □

Corollary 2.3. *Assume that Assumption (A) holds then for any $T \leqslant C_2$ and $t \in [0, T]$ and for any $\mathcal{F}_t$-measurable random vector ξ with finite second moment, we define the process $\big(X_s^{t,\xi,\alpha_t}, Y_s^{t,\xi,\alpha_t}, Z_s^{t,\xi,\alpha_t}, \Lambda_s^{t,\xi,\alpha_t}\big)$, $t \leqslant s \leqslant T$ as the unique solution of the problem*

$$
\begin{cases}
X_s &= \xi + \int_t^s f(r, X_r, Y_r, Z_r, \alpha_r)dr + \int_t^s \sigma(r, X_r, Y_r, \alpha_r)dW_r, \\
Y_s &= h(X_T, \alpha_T) + \int_s^T g(r, X_r, Y_r, Z_r, \alpha_r)dr - \int_s^T Z_r dW_r - \int_s^T \Lambda_r \bullet dM_r, \quad s \in [t, T].
\end{cases}
\tag{2.1.13}
$$

extended to the whole interval $[0, T]$ if $\xi = x$ a.s. $x \in \mathbb{R}^p$ by putting

$$
X_s^{t,x,\alpha_t} = x, \quad Y_s^{t,x,\alpha_t} = Y_t^{t,x,\alpha_t}, \quad Z_s^{t,x,\alpha_t} = 0, \quad \Lambda_s^{i_0,j_0,t,x,\alpha_t} = 0, \quad \textit{for } 0 \leqslant s \leqslant t,\ i_0, j_0 \in \mathcal{M}.
$$

Then the following assertions hold.

(i) There exists a constant $\gamma_{K,L}$ only depending on K and L such that for all $(t,x) \in [0,T] \times \mathbb{R}^p$,

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|X_s^{t,x,\alpha_t}\right|^2 + \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|Y_s^{t,x,\alpha_t}\right|^2 + \mathbb{E}\int_0^T \left\|Z_s^{t,x,\alpha_t}\right\|^2 ds + \mathbb{E}\int_0^T \left|\Lambda_s^{t,x,\alpha_t}\right|^2 \bullet d\big[M\big]_s$$
$$\leqslant \gamma_{K,L}\big(1+|x|^2\big). \tag{2.1.14}$$

(ii) There exist a constant γ_K only depending on K and a constant $\gamma_{K,L}$ only depending on K, L such that for all $t, t' \in [0,T]$ and $x, x' \in \mathbb{R}^p$,

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}\left|X_s^{t',x',\alpha_t} - X_s^{t,x,\alpha_t}\right|^2 + \mathbb{E}\sup_{0\leqslant t\leqslant T}\left|Y_s^{t',x',\alpha_t} - Y_s^{t,x,\alpha_t}\right|^2 + \mathbb{E}\int_0^T \left\|Z_s^{t',x',\alpha_t} - Z_s^{t,x,\alpha_t}\right\|^2 ds$$
$$+ \mathbb{E}\int_0^T \left|\Lambda_s^{t',x',\alpha_t} - \Lambda_s^{t,x,\alpha_t}\right|^2 \bullet d\big[M\big]_s$$
$$\leqslant \gamma_K|x-x'|^2 + \gamma_{K,L}\big(1+|x|^2\big)|t-t'|. \tag{2.1.15}$$

Proof. Let $0\leqslant t\leqslant T\leqslant C_2$. We observe that the quadruplets of functions $\big(\mathbb{1}_{[t,T]}f, \mathbb{1}_{[t,T]}g, \mathbb{1}_{[t,T]}\sigma, h\big)$ and $\big(0,0,0,0\big)$ satisfy Assumption (A). Moreover,

$\big(X_s^{t,x,\alpha_t}, Y_s^{t,x,\alpha_t}, Z_s^{t,x,\alpha_t}, \Lambda_s^{t,x,\alpha_t}\big)$, $0\leqslant s\leqslant T$ is the unique solution of the FBSDE

$$\begin{cases} X_s &= x + \int_t^s \mathbb{1}_{[t,T]}(r) f(r, X_r, Y_r, Z_r, \alpha_r)dr + \int_t^s \mathbb{1}_{[t,T]}(r)\sigma(r, X_r, Y_r, \alpha_r)dW_r, \\ Y_s &= h(X_T, \alpha_T) + \int_s^T \mathbb{1}_{[t,T]}(r) g(r, X_r, Y_r, Z_r, \alpha_r)dr - \int_s^T Z_r dW_r - \int_s^T \Lambda_r \bullet dM_r, \\ s &\in [0,T]. \end{cases}$$

Therefore, as a direct consequence of Proposition 2.2 we obtain

$$
\begin{aligned}
&\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|X_s^{t,x,\alpha_t}\big|^2+\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|Y_s^{t,x,\alpha_t}\big|^2+\mathbb{E}\int_0^T\big\|Z_s^{t,x,\alpha_t}\big\|^2ds+\mathbb{E}\int_0^T\big|\Lambda_s^{t,x,\alpha_t}\big|^2\bullet d\big[M\big]_s\\
\leqslant&\gamma_K\Bigg[|x|^2+|h(0)|^2+\mathbb{E}\bigg(\int_0^T\big|f(s,0,0,0,\alpha_s)\big|ds\bigg)^2+\mathbb{E}\bigg(\int_0^T\big|g(s,0,0,0,\alpha_s)\big|ds\bigg)^2\\
&\quad+\mathbb{E}\int_0^T\Big\|\sigma(s,0,0,\alpha_s)\Big\|^2ds\Bigg]\\
\leqslant&\gamma_{K,L}\big(1+|x|^2\big).
\end{aligned}
$$

Note that we have used Assumption (A) in the last inequality.

Next, let $t,t'\in[0,T]$ and $x,x'\in\mathbb{R}^p$. Again, as $\big(\mathbb{1}_{[t,T]}f,\mathbb{1}_{[t,T]}g,\mathbb{1}_{[t,T]}\sigma,h\big)$ and $\big(\mathbb{1}_{[t',T]}f,\mathbb{1}_{[t',T]}g,\mathbb{1}_{[t',T]}\sigma,h\big)$ both satisfy Assumption (A), a similar argument to that in the above proves that

$$
\begin{aligned}
&\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|X_s^{t',x',\alpha_t}-X_s^{t,x,\alpha_t}\big|^2+\mathbb{E}\sup_{0\leqslant t\leqslant T}\big|Y_s^{t',x',\alpha_t}-Y_s^{t,x,\alpha_t}\big|^2+\mathbb{E}\int_0^T\big\|Z_s^{t',x',\alpha_t}-Z_s^{t,x,\alpha_t}\big\|^2ds\\
&\quad+\mathbb{E}\int_0^T\big|\Lambda_s^{t',x',\alpha_t}-\Lambda_s^{t,x,\alpha_t}\big|^2\bullet d\big[M\big]_s\\
\leqslant&\gamma_K\Bigg[|x'-x|^2+\mathbb{E}\int_{t'\wedge t}^{t'\vee t}\Big\|\sigma(s,X_s^{t,x,\alpha_t},Y_s^{t,x,\alpha_t},\alpha_s)\Big\|^2ds\\
&\qquad+\mathbb{E}\bigg(\int_{t'\wedge t}^{t'\vee t}\big|f(s,X_s^{t,x,\alpha_t},Y_s^{t,x,\alpha_t},Z_s^{t,x,\alpha_t},\alpha_s)\big|ds\bigg)^2\\
&\qquad\quad+\mathbb{E}\bigg(\int_{t'\wedge t}^{t'\vee t}\big|g(s,X_s^{t,x,\alpha_t},Y_s^{t,x,\alpha_t},Z_s^{t,x,\alpha_t},\alpha_s)\big|ds\bigg)^2\\
\leqslant&\gamma_K|x'-x|^2+\gamma_{K,L}\big(1+|x|^2\big)|t'-t|.
\end{aligned}
$$

We have used Assumption (A) (with the note that $Z_s^{t,x,\alpha_t}=0$ if $s\leqslant t$ and that we can assume $t\leqslant t'$ as t and t' play an equal role) and the estimates in part (i) in the last inequality. This completes the proof. □

Proposition 2.4. *Suppose that Assumption (A) is in force, then for any $T\leqslant C_2$, the mapping*

$$
\theta:[0,T]\times\mathbb{R}^p\times\mathcal{M}\to\mathbb{R}^q,\qquad (t,x,\iota)\mapsto Y_t^{t,x,\iota},
$$

satisfies for any $t, t' \in [0,T]$, $x, x' \in \mathbb{R}^p$, *and* $\iota \in \mathcal{M}$,

$$\big|\theta(t,x,\iota)\big|^2 \leqslant \gamma_{K,L}\big(1+|x|^2\big), \tag{2.1.16}$$

$$\big|\theta(t',x',\iota)-\theta(t,x,\iota)\big|^2 \leqslant \gamma_K|x'-x|^2+\gamma_{K,L}\big(1+|x|^2\big)|t'-t|, \tag{2.1.17}$$

and for every $0 \leqslant t \leqslant T$, *for every* $\mathcal{F}_t$*-measurable random vector* ξ *with finite second moment, there exists a* $\mathbb{P}$*-null set* $N_t^{t,\xi,\alpha_t} \in \mathcal{F}_0$ *such that*

$$Y_s^{t,\xi,\alpha_t}(\omega) = \theta\big(s, X_s^{t,\xi,\alpha_t}(\omega), \alpha_s(\omega)\big), \quad \forall s \in [t,T],\ \omega \notin N_t^{t,\xi,\alpha_t}. \tag{2.1.18}$$

Proof. Note that for any $0 \leqslant t \leqslant T$ and $x \in \mathbb{R}^p$, Y_t^{t,x,α_t} is $\sigma\{\alpha_t\}$-measurable, so it is a function of α_t. That is, $Y_t^{t,x,\alpha_t} = \eta^{t,x}(\alpha_t)$ for some function $\eta^{t,x} : \mathcal{M} \to \mathbb{R}^q$. This implies that $\theta(t,x,\iota) = \eta^{t,x}(\iota)$ is well-defined and that $\theta(t,x,\alpha_t) = Y_t^{t,x,\alpha_t}$. It is easily seen that (2.1.16) and (2.1.17) respectively follows from (2.1.14) and (2.1.15). It remains to prove (2.1.18). To this end, let ξ be a $\mathcal{F}_t$-measurable random vector such that $\mathbb{E}|\xi|^2 < \infty$. In view of Proposition 2.2, for any $\epsilon > 0$ we have

$$\mathbb{E}\Big(\mathbb{1}_{\{|\xi-x|<\epsilon\}}\big|Y_s^{t,\xi,\alpha_t}-Y_s^{t,x,\alpha_t}\big|^2\Big) \leqslant \gamma_K \mathbb{E}\Big(\mathbb{1}_{\{|\xi-x|<\epsilon\}}|\xi-x|^2\Big).$$

Thus, the Lipschitz property (2.1.17) implies

$$\begin{aligned}
&\mathbb{E}\Big(\mathbb{1}_{\{|\xi-x|<\epsilon\}}\big|\theta(t,\xi,\alpha_t)-Y_t^{t,\xi,\alpha_t}\big|^2\Big)\\
&\leqslant 2\Big[\gamma_K\mathbb{E}\Big(\mathbb{1}_{\{|\xi-x|<\epsilon\}}|\xi-x|^2\Big)+\mathbb{E}\Big(\mathbb{1}_{\{|\xi-x|<\epsilon\}}\big|\theta(t,\xi,\alpha_t)-\theta(t,x,\alpha_t)\big|^2\Big)\Big]\\
&\leqslant 4\gamma_K\mathbb{E}\Big(\mathbb{1}_{\{|\xi-x|<\epsilon\}}|\xi-x|^2\Big).
\end{aligned}$$

As a consequence, for any positive integer N,

$$\sum_{k\in\mathbb{Z}^p}\mathbb{E}\Big(\mathbb{1}_{\{|\xi-k/N|_\infty<1/N\}}\big|\theta(t,\xi,\alpha_t)-Y_t^{t,\xi,\alpha_t}\big|^2\Big) \leqslant \frac{4\gamma_K}{N^2}\sum_{k\in\mathbb{Z}^p}\mathbb{E}\mathbb{1}_{\{|\xi-k/N|_\infty<1/N\}},$$

where $|\cdot|_\infty$ denotes the sup norm on $\mathbb{R}^p$. This gives

$$\mathbb{E}\big|\theta(t,\xi,\alpha_t)-Y_t^{t,\xi,\alpha_t}\big|^2 \leqslant \frac{2^{p+2}\gamma_K}{N^2} \qquad \text{for all positive integer } N$$

which means

$$\theta(t,\xi,\alpha_t) = Y_t^{t,\xi,\alpha_t} \quad \text{a.s.} \tag{2.1.19}$$

Moreover, for $t \leqslant s \leqslant T$, $\left(X_u^{t,\xi,\alpha_t}, Y_u^{t,\xi,\alpha_t}, Z_u^{t,\xi,\alpha_t}, \Lambda_u^{t,\xi,\alpha_t}\right)_{s \leqslant u \leqslant T}$ is the solution to the FBSDE

$$\begin{cases} X_u &= X_s^{t,\xi,\alpha_t} + \int_s^u f(r, X_r, Y_r, Z_r, \alpha_r)dr + \int_s^u \sigma(r, X_r, Y_r, \alpha_r)dW_r, \\ Y_u &= h(X_T, \alpha_T) + \int_u^T g(r, X_r, Y_r, Z_r, \alpha_r)dr - \int_u^T Z_r dW_r - \int_u^T \Lambda_r \bullet dM_r. \end{cases}$$

Hence, (2.1.19) shows that

$$Y_u^{t,\xi,\alpha_t} = \theta(u, X_u^{t,\xi,\alpha_t}, \alpha_u) \quad \text{a.s.}$$

Since θ and the trajectories of X_s^{t,ξ,α_t} and X_s^{t,ξ,α_t} are all continuous, we have almost surely for all $u \in [t,T]$,

$$Y_u^{t,\xi,\alpha_t} = \theta(u, X_u^{t,\xi,\alpha_t}, \alpha_u).$$

□

Keep using the notations of Corollary 2.3, we have the following consequence on the dependence of the solutions of the FBSDE on the coefficients.

Corollary 2.5. *Assume that the Assumption (A) hold and $T \leqslant C_2$. Let $(f_n, g_n, h_n, \sigma_n)_{n \geqslant 1}$ be a sequence of functions satisfying Assumption (A) with the same constants K, L as (f, g, h, σ) such that for almost all $t \in [0,T]$ and all $(x, y, z, i_0) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q \times p} \times \mathcal{M}$,*

$$\big(f_n, g_n, h_n, \sigma_n\big)(x, y, z, i_0) \to \big(f, g, h, \sigma\big)(x, y, z, i_0) \quad \text{as } n \to \infty.$$

Let $\left(X_t^{n,0,\xi,\alpha_0}, Y_t^{n,0,\xi,\alpha_0}, Z_t^{n,0,\xi,\alpha_0}, \Lambda_t^{n,0,\xi,\alpha_0}\right)$, $0 \leqslant t \leqslant T$, be the solution of the problem

$$\begin{cases} X_t &= \xi + \int_0^t f_n(s, X_s, Y_s, Z_s, \alpha_s)ds + \int_0^t \sigma_n(s, X_s, Y_s, \alpha_s)dW_s, \\ Y_t &= h_n(X_T, \alpha_T) + \int_t^T g_n(s, X_s, Y_s, Z_s, \alpha_s)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \\ t &\in [0,T]. \end{cases}$$

Then, as $n \to \infty$,

$$\begin{aligned}&\mathbb{E}\sup_{0\leqslant s\leqslant T}\left|X_s^{n,0,\xi,\alpha_0}-X_s^{0,\xi,\alpha_0}\right|^2+\mathbb{E}\sup_{0\leqslant s\leqslant T}\left|Y_s^{n,0,\xi,\alpha_0}-Y_s^{0,\xi,\alpha_0}\right|^2\\&\quad+\mathbb{E}\int_0^T\left\|Z_s^{n,0,\xi,\alpha_0}-Z_s^{0,\xi,\alpha_0}\right\|^2ds+\mathbb{E}\int_0^T\left|\Lambda_s^{n,0,\xi,\alpha_0}-\Lambda_s^{0,\xi,\alpha_0}\right|^2\bullet d\big[M\big]_s\to 0.\end{aligned}\tag{2.1.20}$$

As a consequence, as $n \to \infty$, $\theta_n \to \theta$ uniformly on every compact set of $[0,T]\times\mathbb{R}^p\times\mathcal{M}$.

2 Existence and Uniqueness of Global Solution in Non-Degenerate Diffusion Coefficient Case

Assumption (B) We say that the functions f, g, h, and σ satisfy Assumption (B) if for some constants K and L and there exist constants k, λ such that

(B1) For all $t \in [0,T]$, $i_0 \in \mathcal{M}$, and $(x,y),(x',y') \in \mathbb{R}^p\times\mathbb{R}^q$,

$$\begin{aligned}&\big|h(x,i_0)-h(x',i_0)\big|\leqslant k|x-x'|,\\&\big\|\sigma(t,x,y,i_0)-\sigma(t,x',y',i_0)\big\|^2\leqslant k^2\big(|x-x'|^2+|y-y'|^2\big).\end{aligned}$$

(B2) For all $t \in [0,T]$, $(x,y,z,i_0) \in \mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}^{q\times p}\times\mathcal{M}$, and $(x',y') \in \mathbb{R}^p\times\mathbb{R}^q$,

$$\begin{aligned}&\big|f(t,x,y,z,i_0)\big|\leqslant L\big(1+|y|+\|z\|\big),\\&\big|g(t,x,y,z,i_0)\big|\leqslant L\big(1+|y|+\|z\|\big),\\&\big|h(x,i_0)\big|\leqslant L,\\&\big\|\sigma(t,x,y,i_0)\big\|\leqslant L\big(1+|y|\big).\end{aligned}$$

(B3) For all $(t,x,y,i_0) \in [0,T]\times\mathbb{R}^p\times\mathbb{R}^q\times\mathcal{M}$, and $\zeta \in \mathbb{R}^p$,

$$\big\langle\zeta, a(t,x,y,i_0)\zeta\big\rangle\geqslant\lambda|\zeta|^2,$$

where the function a is defined on $[0,T]\times\mathbb{R}^p\times\mathbb{R}^q\times\mathcal{M}$ as follows

$$a(t,x,y,i_0)=\sigma(t,x,y,i_0)\sigma^\top(t,x,y,i_0)\quad\text{for all }(t,x,y,i_0)\in[0,T]\times\mathbb{R}^p\times\mathbb{R}^q\times\mathcal{M}.$$

(B4) For each $i_0 \in \mathcal{M}$, the function $\sigma(\cdot,\cdot,\cdot,i_0)$ is continuous on $[0,T]\times\mathbb{R}^p\times\mathbb{R}^q$.

Lemma 2.6. *Assume that for each* $i_0 \in \mathcal{M}$*, the functions* $\tilde{f}(\cdot,\cdot,\cdot,\cdot,i_0)$*,* $\tilde{g}(\cdot,\cdot,\cdot,\cdot,i_0)$*,* $\tilde{h}(\cdot,i_0)$*, and* $\tilde{\sigma}(\cdot,\cdot,\cdot,i_0)$ *satisfy Assumption (B). In addition, assume that they are all bounded* C^∞ *functions with bounded derivatives of all orders. Then the following system of PDE*

$$\begin{cases} \dfrac{\partial\tilde{\theta}_k}{\partial t}(t,x,i_0) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{p} \tilde{a}_{ij}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big)\dfrac{\partial^2\tilde{\theta}_k}{\partial x_i \partial x_j}(t,x,i_0) \\ +\displaystyle\sum_{i=1}^{p} \tilde{f}_i\Big(t,x,\tilde{\theta}(t,x,i_0),\nabla_x\tilde{\theta}(t,x,i_0)\tilde{\sigma}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big),i_0\Big)\dfrac{\partial\tilde{\theta}_k}{\partial x_i}(t,x,i_0) \\ +\tilde{g}_k\Big(t,x,\tilde{\theta}(t,x,i_0),\nabla_x\tilde{\theta}(t,x,i_0)\tilde{\sigma}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big),i_0\Big) + \displaystyle\sum_{j_0\in\mathcal{M}} q_{i_0j_0}\tilde{\theta}_k(t,x,j_0) = 0, \\ \tilde{\theta}(T,x,i_0) = \tilde{h}(x,i_0), \qquad \forall\, t\in[0,T], x\in\mathbb{R}^p, k=1,2,\ldots,q \end{cases} \tag{2.2.1}$$

admits a unique bounded solution $\tilde{\theta}(\cdot,\cdot,i_0) \in C^{1,2}\big([0,T]\times\mathbb{R}^p,\mathbb{R}^q\big)$ *satisfying*

$$\frac{\partial\tilde{\theta}}{\partial t}, \frac{\partial\tilde{\theta}}{\partial x_i}, \frac{\partial^2\tilde{\theta}}{\partial x_i\partial x_j} \text{ are bounded on } \mathbb{R}^p \text{ for any } i,j=1,2,\ldots,p. \tag{2.2.2}$$

Furthermore, there exists a constant $\tilde{K}$ *only depending on the constants* $K, L, T, k, \lambda, p, q$ *such that*

$$\sup_{(t,x,i_0)\in[0,T]\times\mathbb{R}^p\times\mathcal{M}} \big|\tilde{\theta}(t,x,i_0)\big| \leqslant \tilde{K}, \tag{2.2.3}$$

$$\sup_{(t,x,i_0)\in[0,T]\times\mathbb{R}^p\times\mathcal{M}} \big|\nabla_x\tilde{\theta}(t,x,i_0)\big| \leqslant \tilde{K}, \tag{2.2.4}$$

$$\mathbb{E}\big|\tilde{\theta}(t,x,\alpha_t) - \tilde{\theta}(t',x,\alpha_{t'})\big| \leqslant \tilde{K}|t'-t|^{1/2}, \quad \forall\, t,t'\in[0,T], x\in\mathbb{R}^p. \tag{2.2.5}$$

In addition, for every $t\in[0,T]$ *and* $\mathcal{F}_t$*-measurable random vector* ξ *with finite second moment, the SDE*

$$\begin{aligned} \tilde{X}_s = \xi &+ \int_t^s \tilde{f}\Big(r,\tilde{X}_r,\tilde{\theta}(r,\tilde{X}_r,\alpha_r),\nabla_x\tilde{\theta}(r,\tilde{X}_r,\alpha_r)\tilde{\sigma}\big(r,\tilde{X}_r,\tilde{\theta}(r,\tilde{X}_r,\alpha_r),\alpha_r\big),\alpha_r\Big)dr \\ &+ \int_t^s \tilde{\sigma}\big(r,\tilde{X}_r,\tilde{\theta}(r,\tilde{X}_r,\alpha_r),\alpha_r\big)dW_r, \quad s\in[t,T] \end{aligned} \tag{2.2.6}$$

admits a unique solution, denoted by $\tilde{X}_s^{t,\xi,\alpha_t}$*,* $s \in [t,T]$*, and the process* $\big(\tilde{X}_s^{t,\xi,\alpha_t}, \tilde{Y}_s^{t,\xi,\alpha_t}, \tilde{Z}_s^{t,\xi,\alpha_t}, \tilde{\Lambda}_s^{i_0,j_0,t,\xi,\alpha_t}\big)$ *given by*

$$\tilde{Y}_s^{t,\xi,\alpha_t} = \tilde{\theta}\big(s, \tilde{X}_s^{t,\xi,\alpha_t}, \alpha_s\big), \quad \tilde{Z}_s^{t,\xi,\alpha_t} = \nabla_x \tilde{\theta}\big(s, \tilde{X}_s^{t,\xi,\alpha_t}, \alpha_s\big)\tilde{\sigma}\big(s, \tilde{X}_s^{t,\xi,\alpha_t}, \tilde{Y}_s^{t,\xi,\alpha_t}, \alpha_s\big), \tag{2.2.7}$$

and

$$\tilde{\Lambda}_s^{i_0,j_0,t,\xi,\alpha_t} = \tilde{\theta}\big(s, \tilde{X}_s^{t,\xi,\alpha_t}, j_0\big) - \tilde{\theta}\big(s, \tilde{X}_s^{t,\xi,\alpha_t}, i_0\big), \quad i_0, j_0 \in \mathcal{M}, s \in [t,T]$$

satisfies the FBSDE associate to $\big(\tilde{f}, \tilde{g}, \tilde{\sigma}, \tilde{h}\big)$ *and to the initial condition* (t,ξ)*.*

Proof. For each $i_0 \in \mathcal{M}$ and for any matrix $\boldsymbol{\vartheta}$ that contains qm_0 rows, we denote by $\boldsymbol{\vartheta}_{i_0}$ the submatrix of $\boldsymbol{\vartheta}$ that contains rows $(i_0-1)q+1, \ldots, (i_0-1)q+q$; that is,

$$\boldsymbol{\vartheta} = \big(\boldsymbol{\vartheta}_1^\top, \boldsymbol{\vartheta}_2^\top, \ldots, \boldsymbol{\vartheta}_{m_0}^\top\big)^\top.$$

Let I_q be the identity matrix of size q and $\otimes$ be the tensor product. For any $\boldsymbol{\theta} \in \mathbb{R}^{qm_0 \times 1}$ and $\boldsymbol{\vartheta} \in \mathbb{R}^{qm_0 \times p}$, put $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{m_0}) \in \mathbb{R}^{q \times m_0}$ and

$$\begin{aligned}
\mathbf{h}(x) &= \operatorname{vec}\big(h(x,1), \ldots, h(x,m_0)\big),\\
\mathbf{a}_{ij}(t,x,\boldsymbol{\theta}) &= \operatorname{diag}\Big(a_{ij}(t,x,\boldsymbol{\theta}_1,1), \ldots, a_{ij}(t,x,\boldsymbol{\theta}_{m_0},m_0)\Big) \otimes I_q,\\
\mathbf{f}_i(t,x,\boldsymbol{\theta},\boldsymbol{\vartheta}) &= \operatorname{diag}\Big(f_i\big(t,x,\boldsymbol{\theta}_1,\boldsymbol{\vartheta}_1\sigma(t,x,\boldsymbol{\theta}_1,1),1\big), \ldots, f_i\big(t,x,\boldsymbol{\theta}_{m_0},\boldsymbol{\vartheta}_{m_0}\sigma(t,x,\boldsymbol{\theta}_{m_0},m_0),m_0\big)\Big) \otimes I_q,\\
\bar{\mathbf{g}}(t,x,\boldsymbol{\theta},\boldsymbol{\vartheta}) &= \begin{pmatrix} g_1\big(t,x,\boldsymbol{\theta}_1,\boldsymbol{\vartheta}_1\sigma(t,x,\boldsymbol{\theta}_1,1),1\big) & \ldots & g_1\big(t,x,\boldsymbol{\theta}_{m_0},\boldsymbol{\vartheta}_{m_0}\sigma(t,x,\boldsymbol{\theta}_{m_0},m_0),m_0\big) \\ \vdots & \ddots & \vdots \\ g_q\big(t,x,\boldsymbol{\theta}_1,\boldsymbol{\vartheta}_1\sigma(t,x,\boldsymbol{\theta}_1,1),1\big) & \ldots & g_q\big(t,x,\boldsymbol{\theta}_{m_0},\boldsymbol{\vartheta}_{m_0}\sigma(t,x,\boldsymbol{\theta}_{m_0},m_0),m_0\big) \end{pmatrix} + \tilde{\boldsymbol{\theta}} Q^\top,\\
\mathbf{g}(t,x,\boldsymbol{\theta},\boldsymbol{\vartheta}) &= \operatorname{vec}\Big(\bar{\mathbf{g}}(t,x,\boldsymbol{\theta},\boldsymbol{\vartheta})\Big).
\end{aligned}$$

Moreover, we denote $\boldsymbol{\theta}(t,x) = \operatorname{vec}\big(\tilde{\boldsymbol{\theta}}(t,x)\big)$, where

$$\tilde{\boldsymbol{\theta}}(t,x) = \begin{pmatrix} \tilde{\theta}_1(t,x,1) & \cdots & \tilde{\theta}_1(t,x,m_0) \\ \vdots & \ddots & \vdots \\ \tilde{\theta}_q(t,x,1) & \cdots & \tilde{\theta}_q(t,x,m_0) \end{pmatrix}.$$

Then (2.2.1) becomes

$$\begin{cases} \dfrac{\partial \boldsymbol{\theta}}{\partial t}(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{p} \mathbf{a}_{ij}(t,x,\boldsymbol{\theta})\dfrac{\partial^2 \boldsymbol{\theta}}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{p} \mathbf{f}_i(t,x,\boldsymbol{\theta},\nabla_x \boldsymbol{\theta})\dfrac{\partial \boldsymbol{\theta}}{\partial x_i}(t,x) + \mathbf{g}(t,x,\boldsymbol{\theta},\nabla_x\boldsymbol{\theta}) = 0, \\ \boldsymbol{\theta}(T,x) = \mathbf{h}(x). \end{cases} \tag{2.2.8}$$

In view of [34, Proposition 3.3] and [31, Theorem VII.7.1], under Assumption (B) the system (2.2.8) admits a unique bounded classical solution $\boldsymbol{\theta} \in C^{1,2}([0,T]\times\mathbb{R}^p, \mathbb{R}^{qm_0})$. Moreover, the solution $\boldsymbol{\theta}$ has bounded partial derivatives $\frac{\partial \boldsymbol{\theta}}{\partial t}(t,x)$, $\frac{\partial \boldsymbol{\theta}}{\partial x_i}(t,x)$, and $\frac{\partial^2 \boldsymbol{\theta}}{\partial x_i \partial x_j}(t,x)$, $1 \leqslant i,j \leqslant p$, on $[0,T]\times\mathbb{R}^p$. This implies (2.2.2).

Next, to prove the remaining part we first define for each $(t,x,i_0) \in [0,T]\times\mathbb{R}^p\times\mathcal{M}$

$$\tilde{F}(t,x,i_0) = \tilde{f}\Big(t,x,\tilde{\theta}(t,x,i_0), \nabla_x\tilde{\theta}(t,x,i_0)\tilde{\sigma}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big),i_0\Big),$$

$$\tilde{\Sigma}(t,x,i_0) = \tilde{\sigma}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big).$$

Then for each $t \in [0,T]$ and a $\mathcal{F}_t$-measurable $\mathbb{R}^d$-valued random vector ξ with bounded second moment the equation

$$\tilde{X}_s^{t,\xi,\alpha_t} = \xi + \int_t^s \tilde{F}(r,\tilde{X}_r^{t,\xi,\alpha_t},\alpha_r)dr + \int_t^s \tilde{\Sigma}(r,\tilde{X}_r^{t,\xi,\alpha_t},\alpha_r)dW_r$$

posses a unique solution. For $s \in [t,T]$ denote

$$\tilde{Y}_s^{t,\xi,\alpha_t} = \tilde{\theta}\big(s,\tilde{X}_s^{t,\xi,\alpha_t},\alpha_s\big), \quad \tilde{Z}_s^{t,\xi,\alpha_t} = \nabla_x\tilde{\theta}\big(s,\tilde{X}_s^{t,\xi,\alpha_t},\alpha_s\big)\tilde{\Sigma}(s,\tilde{X}_s^{t,\xi,\alpha_t},\alpha_s),$$

and

$$\tilde{\Lambda}_s^{i_0,j_0,t,\xi,\alpha_t} = \tilde{\theta}\big(s,\tilde{X}_s^{t,\xi,\alpha_t},j_0\big) - \tilde{\theta}\big(s,\tilde{X}_s^{t,\xi,\alpha_t},i_0\big).$$

Using Itô formula for $\tilde{\theta}\big(s, \tilde{X}_s^{t,\xi,\alpha_t}, \alpha_s\big)$ and then applying (2.2.1) we obtain

$$\tilde{Y}_s^{t,\xi,\alpha_t} = \tilde{h}(\tilde{X}_T^{t,\xi,\alpha_t}, \alpha_T) + \int_s^T \tilde{g}\big(r, \tilde{X}_r^{t,\xi,\alpha_t}, \tilde{Y}_r^{t,\xi,\alpha_t}, \tilde{Z}_r^{t,\xi,\alpha_t}, \alpha_r\big)dr - \int_s^T \tilde{Z}_r^{t,\xi,\alpha_t} dW_r - \int_s^T \tilde{\Lambda}_r^{t,\xi,\alpha_t} \bullet dM_r$$

for each $s \in [t,T]$. This implies that $\big(\tilde{X}^{t,\xi,\alpha_t}, \tilde{Y}^{t,\xi,\alpha_t}, \tilde{Z}^{t,\xi,\alpha_t}, \tilde{\Lambda}^{t,\xi,\alpha_t}\big)$ is a solution of the FBSDE associated to the coefficients $\tilde{f}, \tilde{g}, \tilde{h}$, and $\tilde{\sigma}$ and to the initial conditions t, ξ, α_t.

Let $\gamma > 0$ be an arbitrary fixed number. Note that M_t is a purely discontinuous and square integrable martingale. Using generalized Itô formula [48, Theorem V.18] for $\big\{e^{\gamma s}\big|\tilde{Y}_s^{t,x,\alpha_t}\big|^2\big\}_{t \leqslant s \leqslant T}$ we arrive at

$$\begin{aligned}
&e^{\gamma T}\big|\tilde{Y}_T^{t,x,\alpha_t}\big|^2 \\
&= e^{\gamma s}\big|\tilde{Y}_s^{t,x,\alpha_t}\big|^2 + \gamma \int_s^T e^{\gamma r}\big|\tilde{Y}_r^{t,x,\alpha_t}\big|^2 dr - 2\int_s^T e^{\gamma r}\Big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{g}\big(r, \tilde{X}_r^{t,x,\alpha_t}, \tilde{Y}_r^{t,x,\alpha_t}, \tilde{Z}_r^{t,x,\alpha_t}, \alpha_r\big)\Big\rangle dr \\
&\quad + 2\int_s^T e^{\gamma r}\big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{Z}_r^{t,x,\alpha_t} dW_r\big\rangle + 2\int_s^T e^{\gamma r}\big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{\Lambda}_r^{t,x,\alpha_t}\big\rangle dM_r + \int_s^T e^{\gamma r}\big\|\tilde{Z}_r^{t,x,\alpha_t}\big\|^2 dr \\
&\quad + \int_s^T e^{\gamma r}|\tilde{\Lambda}_r^{t,x,\alpha_t}|^2 \bullet d[M]_r
\end{aligned}$$

for all $0 \leqslant t \leqslant s \leqslant T$ and $x \in \mathbb{R}^p$, which, together with Assumption (B), yields

$$\begin{aligned}
&e^{\gamma s}\big|\tilde{Y}_s^{t,x,\alpha_t}\big|^2 + \int_s^T e^{\gamma r}\big\|\tilde{Z}_r^{t,x,\alpha_t}\big\|^2 dr + \sum_{s<r\leqslant T} e^{\gamma r}\big|\tilde{Y}_r^{t,x,\alpha_t} - \tilde{Y}_{r-}^{t,x,\alpha_t}\big|^2 \\
&\leqslant e^{\gamma T}\big|\tilde{Y}_T^{t,x,\alpha_t}\big|^2 + \int_s^T e^{\gamma r}\Big[2L\Big(1 + \big|\tilde{Y}_r^{t,x,\alpha_t}\big| + \big\|\tilde{Z}_r^{t,x,\alpha_t}\big\|\Big)\big|\tilde{Y}_r^{t,x,\alpha_t}\big| - \gamma\big|\tilde{Y}_r^{t,x,\alpha_t}\big|^2\Big] dr \\
&\quad - 2\int_s^T e^{\gamma r}\big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{Z}_r^{t,x,\alpha_t} dW_r\big\rangle - 2\int_s^T e^{\gamma r}\big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{\Lambda}_r^{t,x,\alpha_t}\big\rangle dM_r \\
&\leqslant e^{\gamma T}\big|\tilde{Y}_T^{t,x,\alpha_t}\big|^2 + \int_s^T e^{\gamma r}\Big[L + \big(3L + 2L^2 - \gamma\big)\big|\tilde{Y}_r^{t,x,\alpha_t}\big|^2 + \frac{1}{2}\big\|\tilde{Z}_r^{t,x,\alpha_t}\big\|^2\Big] dr \\
&\quad - 2\int_s^T e^{\gamma r}\big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{Z}_r^{t,x,\alpha_t} dW_r\big\rangle - 2\int_s^T e^{\gamma r}\big\langle \tilde{Y}_r^{t,x,\alpha_t}, \tilde{\Lambda}_r^{t,x,\alpha_t}\big\rangle dM_r.
\end{aligned}$$

By choosing $\gamma = 3L + 2L^2$, taking the conditional expectations with $\mathcal{F}_t$, and using Assumption (B) we get

$$\mathbb{E}\Big(e^{\gamma s}\big|\tilde{Y}_s^{t,x,\alpha_t}\big|^2\Big|\mathcal{F}_t\Big) \leqslant e^{\gamma T}L^2 + L\int_s^T \mathbb{E}\Big(e^{\gamma r}\Big|\mathcal{F}_t\Big) dr = C, \quad t \leqslant s \leqslant T,$$

where $C = C(L,T)$ is a constant depending on L and T. Therefore, there exists a constant C such that $\left|\tilde{\theta}(t,x,\alpha_t)\right| \leqslant C$ for all $(t,x) \in [0,T] \times \mathbb{R}^p$. Since the Markov chain (α_t) is irreducible, it follows that

$$\left|\tilde{\theta}(t,x,i_0)\right| \leqslant C \quad \forall (t,x,i_0) \in [0,T] \times \mathbb{R}^p \times \mathcal{M}.$$

Next, to estimate $\left|\nabla_x \tilde{\theta}(t,x,i_0)\right|$ we apply [31, Theorem VII.6.8] for the solution $\boldsymbol{\theta}$ of (2.2.8) to the cylinders $[0,T] \times \{x \in \mathbb{R}^p, |x| \leqslant n\}$ and $[0,T] \times \{x \in \mathbb{R}^p, |x| \leqslant n+1\}$. It follows that $\sup_{\{t \in [0,T], |x| \leqslant n\}} \left|\nabla_x \boldsymbol{\theta}(t,x)\right|^2$ is bounded by a constant depending on the constants $C(L,T), K, L, k, \lambda, p, q$ and the distance between $\{x \in \mathbb{R}^p, |x| \leqslant n\}$ and $\partial\{x \in \mathbb{R}^p, |x| \leqslant n+1\}$ which is 1. As a consequence, there exists a constant $C(K,L,T,k,\lambda,p,q)$ such that

$$\sup_{(t,x,i_0) \in [0,T] \times \mathbb{R}^p \times \mathcal{M}} \left|\nabla_x \tilde{\theta}(t,x,i_0)\right| \leqslant C(K,L,T,k,\lambda,p,q).$$

Finally, in order to prove the remaining inequality (2.2.5) take t and t' such that $0 \leqslant t \leqslant t' \leqslant T$. According to Corollary 2.3, Assumption (B), (2.2.3), and (2.2.4), there exists a constant $\tilde{C} = \tilde{C}(K,L,T,k,\lambda,p,q)$ such that

$$\mathbb{E}\left|\tilde{Y}_{t'}^{t,x,\alpha_t} - \tilde{Y}_t^{t,x,\alpha_t}\right|^2 \leqslant \tilde{C}(t'-t), \quad \mathbb{E}\left|\tilde{X}_{t'}^{t,x,\alpha_t} - \tilde{X}_t^{t,x,\alpha_t}\right|^2 \leqslant \tilde{C}(t'-t).$$

Since $\tilde{Y}_t^{t,x,\alpha_t} = \tilde{\theta}(t,x,\alpha_t)$ and $\tilde{Y}_{t'}^{t,x,\alpha_t} = \tilde{\theta}(t', \tilde{X}_{t'}^{t,x,\alpha_t}, \alpha_{t'})$, by modifying $\tilde{C}$ if necessary and Proposition 2.4, we obtain

$$\begin{aligned}\mathbb{E}\left|\tilde{\theta}(t,x,\alpha_t) - \tilde{\theta}(t',x,\alpha_{t'})\right|^2 &\leqslant 2\Big(\mathbb{E}\left|\tilde{\theta}(t,x,\alpha_t) - \tilde{Y}_{t'}^{t,x,\alpha_t}\right|^2 + \mathbb{E}\left|\tilde{\theta}(t', \tilde{X}_{t'}^{t,x,\alpha_t}, \alpha_{t'}) - \tilde{\theta}(t',x,\alpha_{t'})\right|^2\Big) \\ &\leqslant \tilde{C}\Big[(t'-t) + \mathbb{E}\left|\tilde{X}_{t'}^{t,x,\alpha_t} - x\right|^2\Big] = \tilde{C}(t'-t).\end{aligned}$$

This completes the proof. □

Remark 2.7. (i) In view of [20, Proposition 2.2], under Assumption (B) there exists a sequence of C^∞ functions $(f_n, g_n, h_n, \sigma_n)_{n \geqslant 1}$ satisfying for every n Assumption (B) with respect to the constants $K + 4L$, k, $2L$, and $\lambda/2$ such that for almost all $t \in [0,T]$ and all

$(x, y, z, i_0) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^{q\times p} \times \mathcal{M}$,

$$(f_n, g_n, h_n, \sigma_n)(t, x, y, z, i_0) \to (f, g, h, \sigma)(t, x, y, z, i_0) \quad \text{as } n \to \infty.$$

Denote $a_n = \sigma_n\sigma_n^\top$ then as a consequence of Lemma 2.6, the following system of PDE

$$\begin{cases} \dfrac{\partial(\theta_n)_k}{\partial t}(t,x,i_0) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{p}(a_n)_{ij}\big(t,x,\theta_n(t,x,i_0),i_0\big)\dfrac{\partial^2(\theta_n)_k}{\partial x_i\partial x_j}(t,x,i_0) \\ +\displaystyle\sum_{i=1}^{p}(f_n)_i\Big(t,x,\theta_n(t,x,i_0),\nabla_x\theta_n(t,x,i_0)\sigma_n\big(t,x,\theta_n(t,x,i_0),i_0\big),i_0\Big)\dfrac{\partial(\theta_n)_k}{\partial x_i}(t,x,i_0) \\ +(g_n)_k\Big(t,x,\theta_n(t,x,i_0),\nabla_x\theta_n(t,x,i_0)\sigma_n\big(t,x,\theta_n(t,x,i_0),i_0\big),i_0\Big) + \displaystyle\sum_{j_0\in\mathcal{M}} q_{i_0j_0}(\theta_n)_k(t,x,j_0) = 0, \\ \theta_n(T,x,i_0) = h_n(x,i_0), \qquad \forall\, t\in[0,T], x\in\mathbb{R}^p, k=1,2,\ldots,q \end{cases} \tag{2.2.9}$$

admits a unique bounded solution $\theta_n(\cdot,\cdot,i_0) \in C^{1,2}\big([0,T]\times\mathbb{R}^p,\mathbb{R}^q\big)$ satisfying

$$\frac{\partial\theta_n}{\partial t}, \frac{\partial\theta_n}{\partial x_i}, \frac{\partial^2\theta_n}{\partial x_i\partial x_j} \text{ are bounded on } \mathbb{R}^p \text{ for any } i, j = 1, 2, \ldots, p.$$

Furthermore, there exists a constant $\tilde{K}$ only depending on the constants $K, L, T, k, \lambda, p, q$ such that

$$\sup_{(t,x,i_0)\in[0,T]\times\mathbb{R}^p\times\mathcal{M}} \big|\theta_n(t,x,i_0)\big| \leqslant \tilde{K}, \tag{2.2.10}$$

$$\sup_{(t,x,i_0)\in[0,T]\times\mathbb{R}^p\times\mathcal{M}} \big|\nabla_x\theta_n(t,x,i_0)\big| \leqslant \tilde{K}, \tag{2.2.11}$$

$$\mathbb{E}\big|\theta_n(t,x,\alpha(t)) - \theta_n(t',x,\alpha(t'))\big| \leqslant \tilde{K}|t'-t|^{1/2}, \quad \forall\, t,t'\in[0,T], x\in\mathbb{R}^p. \tag{2.2.12}$$

(ii) Next, for simplicity, we denote the following constants

$$K^* = \max\{k, K+4L, \tilde{K}\} \quad \text{and} \quad T^* = C_2(K^*), \tag{2.2.13}$$

where $\tilde{K} = \tilde{K}(K, L, T, \lambda, p, q)$ is the constant given in Lemma 2.6 and $C_2(K^*)$ is the constant given in Proposition 2.2. Let N be the integer satisfying $(N-1)T^* \leqslant T < NT^*$ and put $t_0 = 0$, $t_i = T - (N-i)T^*$ for each $i = 1, 2, \ldots, N$. Then according to Theorem 2.1, for each $n \geqslant 1$,

$t \in [t_i, t_{i+1})$ for some $0 \leqslant i \leqslant N-1$, and a $\mathcal{F}_t$-measurable random vector ξ with bounded second moment, the following equation

$$\begin{cases} X_s &= \xi + \int_t^s f_n(r, X_r, Y_r, Z_r, \alpha_r)dr + \int_t^s \sigma_n(r, X_r, Y_r, \alpha_r)dW_r, \\ Y_s &= \theta_n(t_{i+1}, X_{t_{i+1}}, \alpha_{t_{i+1}}) + \int_s^{t_{i+1}} g_n(r, X_r, Y_r, Z_r, \alpha_r)dr - \int_s^{t_{i+1}} Z_r dW_r - \int_s^{t_{i+1}} \Lambda_r \bullet dM_r, \end{cases}$$

where $s \in [t, t_{i+1}]$, admits a unique solution denoted by $\left(X_s^{n,i,t,\xi,\alpha_t}, Y_s^{n,i,t,\xi,\alpha_t}, Z_s^{n,i,t,\xi,\alpha_t}, \Lambda_s^{n,i,t,\xi,\alpha_t}\right)$. In view of Lemma 2.6, this solution also satisfies the following equations

$$\begin{aligned} Y_s^{n,i,t,\xi,\alpha_t} &= \theta_n\big(s, X_s^{n,i,t,\xi,\alpha_t}, \alpha_s\big), \\ Z_s^{n,i,t,\xi,\alpha_t} &= \nabla_x \theta_n\big(s, X_s^{n,i,t,\xi,\alpha_t}, \alpha_s\big)\sigma_n\big(s, X_s^{n,i,t,\xi,\alpha_t}, Y_s^{n,i,t,\xi,\alpha_t}, \alpha_s\big), \end{aligned}$$

and

$$\Lambda_s^{n,i,i_0,j_0,t,\xi,\alpha_t} = \theta_n\big(s, X_s^{n,i,t,\xi,\alpha_t}, j_0\big) - \theta_n\big(s, X_s^{n,i,t,\xi,\alpha_t}, i_0\big), \quad i_0, j_0 \in \mathcal{M}, s \in [t, t_{i+1}].$$

As a consequence, Assumption (B) together with the inequalities (2.2.10) and (2.2.11) imply that

$$\big\|Z_s^{n,i,t,\xi,\alpha_t}\big\|, \big|\Lambda_s^{n,i,i_0,j_0,t,\xi,\alpha_t}\big| \leqslant \tilde{K}', \quad \forall\ i_0, j_0 \in \mathcal{M}, s \in [t, t_{i+1}]$$

for some constants $\tilde{K}'$ only depending on the constants $K, L, T, k, \lambda, p, q$.

Proposition 2.8. *Under Assumption (B) and the notations in Remark* 2.7*, there exists a mapping* $\theta : [0,T] \times \mathbb{R}^p \times \mathcal{M} \to \mathbb{R}^p$ *such that, as* $n \to \infty$*,* $\theta_n(t, x, \alpha_t)$ *converges uniformly in* $\mathcal{L}^2(\mathbb{R}^p)$ *to* $\theta(t, x, \alpha_t)$ *on every compact subset of* $[0,T] \times \mathbb{R}^p$*. In addition, the mapping* θ *satisfies the following properties:*

$$\begin{gathered} \theta(T, x, i_0) = h(x, i_0), \quad \forall\ (t, x, i_0) \in [0,T] \times \mathbb{R}^p \times \mathcal{M}, \\ \big|\theta(t, x, i_0)\big| \leqslant C, \quad \forall\ (t, x, i_0) \in [0,T] \times \mathbb{R}^p \times \mathcal{M}, \\ \mathbb{E}\big|\theta(t, x, \alpha_t) - \theta(t', x', \alpha_{t'})\big| \leqslant \tilde{K}\big(|t-t'|^{1/2} + |x-x'|\big), \quad \forall\ (t, x), (t', x') \in [0,T] \times \mathbb{R}^p. \end{gathered}$$

Furthermore, for each $n \geqslant 1$*,* $t \in [t_i, t_{i+1})$ *for some* $0 \leqslant i \leqslant N-1$*, and a* $\mathcal{F}_t$*-measurable random vector* ξ *with bounded second moment, the following equation*

$$\begin{cases} X_s &= \xi + \int_t^s f(r, X_r, Y_r, Z_r, \alpha_r)dr + \int_t^s \sigma(r, X_r, Y_r, \alpha_r)dW_r, \\ Y_s &= \theta(t_{i+1}, X_{t_{i+1}}, \alpha_{t_{i+1}}) + \int_s^{t_{i+1}} g(r, X_r, Y_r, Z_r, \alpha_r)dr - \int_s^{t_{i+1}} Z_r dW_r - \int_s^{t_{i+1}} \Lambda_r \bullet dM_r, \end{cases}$$

where $s \in [t, t_{i+1}]$*, admits a unique solution denoted by* $\big(X_s^{i,t,\xi,\alpha_t}, Y_s^{i,t,\xi,\alpha_t}, Z_s^{i,t,\xi,\alpha_t}, \Lambda_s^{i,t,\xi,\alpha_t}\big)$ *which satisfies*

$$\mathbb{P}\Big(Y_s^{i,t,\xi,\alpha_t} = \theta\big(s, X_s^{i,t,\xi,\alpha_t}, \alpha_s\big) \text{ for all } s \in [t, t_{i+1}]\Big) = 1,$$
$$\mathbb{P} \otimes \mu\Big\{(\omega, s) \in \Omega \times [t, t_{i+1}],\ \big|X_s^{i,t,\xi,\alpha_t}\big| > \tilde{K}'\Big\} = 0.$$

Proof. Let T^*, N, and t_k, $0 \leqslant k \leqslant N$, be defined as in previous Remark. We will define by induction the mapping θ on the intervals $[t_{k-1}, t_k)$ with k running downward from N to 1.

First, for $k = N$, in virtue of Theorem 2.1, for any $t \in [t_{N-1}, T)$ and $\mathcal{F}_t$-measurable random vector ξ with bounded second moment, the FBSDE

$$\begin{cases} X_s &= \xi + \int_t^s f(r, X_r, Y_r, Z_r, \alpha_r)dr + \int_t^s \sigma(r, X_r, Y_r, \alpha_r)dW_r, \\ Y_s &= h(X_T, \alpha_T) + \int_s^T g(r, X_r, Y_r, Z_r, \alpha_r)dr - \int_s^T Z_r dW_r - \int_s^T \Lambda_r \bullet dM_r, \quad s \in [t, T]. \end{cases}$$

has a unique solution. Let us denote this solution by $\big(X_s^{N-1,t,\xi,\alpha_t}, Y_s^{N-1,t,\xi,\alpha_t}, Z_s^{N-1,t,\xi,\alpha_t}, \Lambda_s^{N-1,t,\xi,\alpha_t}\big)_{t \leqslant s \leqslant T}$. Define

$$\theta : [t_{N-1}, T] \times \mathbb{R}^p \times \mathcal{M} \to \mathbb{R}^q, \qquad (t, x, \alpha_t) \mapsto \theta(t, x, \alpha_t) = Y_t^{N-1,t,x,\alpha_t}.$$

According to Corollary 2.3, it follows that $Y_t^{N-1,t,\xi,\alpha_t} = \theta(t, \xi, \alpha_t)$ for all $t \leqslant s \leqslant T$ with probability 1. $\square$

3 Related PDEs: Weak sense

Put

$$\tilde{U}(t,x,i_0) = \nabla_x \tilde{\theta}(t,x,i_0)\tilde{\sigma}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big),$$
$$\tilde{W}(t,x,i_0,j_0) = \tilde{\theta}(t,x,j_0) - \tilde{\theta}(t,x,i_0),$$

and

$$\tilde{\mathcal{L}}\tilde{\theta}_k(t,x,i_0) = \frac{1}{2}\sum_{i,j=1}^{p} \tilde{a}_{ij}\big(t,x,\tilde{\theta}(t,x,i_0),i_0\big)\frac{\partial^2\tilde{\theta}_k}{\partial x_i\partial x_j}(t,x,i_0) + \sum_{i=1}^{p}\tilde{f}_i\Big(t,x,\tilde{\theta}(t,x,i_0),\tilde{U}(t,x,i_0),i_0\Big)\frac{\partial\tilde{\theta}_k}{x_i}(t,x,i_0)$$

Then (2.2.1) becomes

$$\begin{cases} \frac{\partial\tilde{\theta}_k}{\partial t}(t,x,i_0) + \tilde{\mathcal{L}}\tilde{\theta}_k(t,x,i_0) + \tilde{g}_k\Big(t,x,\tilde{\theta}(t,x,i_0),\tilde{U}(t,x,i_0),i_0\Big) + \sum_{j0\in\mathcal{M}} q_{i_0j_0}\tilde{\theta}_k(t,x,j_0) = 0, \\ \tilde{\theta}(T,x,i_0) = \tilde{h}(x,i_0), \qquad \forall\, t\in[0,T], x\in\mathbb{R}^p, k=1,2,\dots,q. \end{cases} \tag{2.3.1}$$

Put

$$U(t,x,i_0) = \nabla_x\theta(t,x,i_0)\sigma\big(t,x,\theta(t,x,i_0),i_0\big),$$

and

$$W(t,x,i_0,j_0) = \theta(t,x,j_0) - \theta(t,x,i_0).$$

We define $\mathcal{H}$ the set of functions $\theta(s,x,i)$ such that $(\theta,U)\in\mathcal{L}^2\left([0,T]\times\mathbb{R}^p;\mathbb{R}^q\right)\otimes\mathcal{L}^2\left([0,T]\times\mathbb{R}^p;\mathbb{R}^{q\times p}\right)$ for each $i\in\mathcal{M}$ with the norm

$$||\theta||_{\mathcal{H}} = \left(\int_0^T\int_{\mathbb{R}^p}\left(|\theta(s,x,i)|^2 + \|U(s,x,i)\|^2 + \sum_{j\in\mathcal{M}} q_{ij}|W(s,x,i,j)|^2\right)dxds\right)^{1/2}.$$

Definition 2.9. *We say that θ is a weak solution of the PDE* (2.2.1) *if θ satisfies*

$$\begin{aligned}
&\int_t^T \int_{\mathbb{R}^p} \theta_k(s,x,i_0) D_s\varphi_k(s,x)dxds + \int_{\mathbb{R}^p} \theta_k(t,x,i_0)\varphi_k(t,x)dx - \int_{\mathbb{R}^p} h_k(x,i_0)\varphi_k(T,x)dx \\
&+ \frac{1}{2}\int_t^T \int_{\mathbb{R}^p} \sum_{i,j=1}^p a_{ij}\big(s,x,\theta(s,x,i_0),i\big) D_i\theta_k(s,x,i_0) D_j\varphi_k(s,x)dxds \\
&+ \int_t^T \int_{\mathbb{R}^p} \theta_k(s,x,i_0) \sum_{i=1}^p D_i((f_i - A_i)\varphi_k(s,x))dxds \\
&- \int_t^T \int_{\mathbb{R}^p} g_k(t,x,\theta(s,x,i_0),U(s,x,i_0),i_0)\varphi_k(s,x)dxds \\
&- \int_t^T \int_{\mathbb{R}^p} \sum_{j_0\in\mathcal{M}} q_{i_0j_0}\theta_k(t,x,j_0)\varphi_k(s,x)dxds = 0,
\end{aligned} \tag{2.3.2}$$

where $A_i(s,x,i_0) = \frac{1}{2}\sum_{j=1}^p D_j(a_{ij}(s,x,\theta(s,x,i_0),i_0))$.

Definition 2.10. *We say that $\boldsymbol{\theta}$ is a weak solution of the PDE* (2.2.1) *if $\boldsymbol{\theta}$ satisfies*

$$\begin{aligned}
&\int_t^T \int_{\mathbb{R}^p} \boldsymbol{\theta}(s,x) D_s\varphi(s,x)dxds + \int_{\mathbb{R}^p} \boldsymbol{\theta}(t,x)\varphi(t,x)dx - \int_{\mathbb{R}^p} \mathbf{h}(x)\varphi(T,x)dx \\
&\quad + \int_t^T \int_{\mathbb{R}^p} \frac{1}{2}\sum_{i,j=1}^p \mathbf{a}_{ij}(s,x,\boldsymbol{\theta}) D_i\boldsymbol{\theta}(s,x) D_j\varphi(s,x)dxds \\
&\quad + \int_t^T \int_{\mathbb{R}^p} \boldsymbol{\theta}(s,x) \sum_{i=1}^p D_i\big((\mathbf{f}_i - A_i)\varphi(s,x)\big)dsdx \\
&= \int_t^T \int_{\mathbb{R}^p} \mathbf{g}(s,x,\boldsymbol{\theta},\nabla_x\boldsymbol{\theta})\varphi(s,x)dxds.
\end{aligned} \tag{2.3.3}$$

where $A_i = \frac{1}{2}\sum_{i=1}^p D_j(\mathbf{a}_{ij}(s,x,\boldsymbol{\theta}))$.

To proceed further, we need to use the following lemma (see [36, Lemma 2.10]).

Lemma 2.11 (Generalized equivalence of norm principle)**.** *We take $\rho(x) := e^{F(x)}$ as the weight function, where $F : \mathbb{R}^d \to \mathbb{R}$ is a continuous function. Moreover, we assume that there exists a constant $R > 0$ such that for $|x| > R, F \in C^2_{l,b}\left(\mathbb{R}^d;\mathbb{R}\right)$ and $\sup_{x\in\mathbb{R}^d}|F'(x)x| < +\infty$. For instance, we can take $\rho(x) = (1+|x|)^q$, with $q \in \mathbb{R}$ or $\rho(x) = e^{\frac{\alpha}{1+|x|}}$ with $\alpha \in \mathbb{R}$. If $\varphi\rho^{-1} \in L^1\left(\mathbb{R}^d\right)$. Then there exist two constants $c > 0$ and $C > 0$ such that*

$$c\int_{\mathbb{R}^d} |\varphi(x)|\rho^{-1}(x)dx \leqslant \mathbb{E}\left[\int_{\mathbb{R}^d} \left|\varphi\left(X_s^{t,x,i}\right)\right|\rho^{-1}(x)dx\right] \leqslant C\int_{\mathbb{R}^d} |\varphi(x)|\rho^{-1}(x)dx$$

Moreover if $\Psi : \Omega \times [t,T] \times \mathbb{R}^d \to \mathbb{R}, \Psi(s,\cdot)$ *is* F_s^α *measurable for* $s \in [t,T]$ *and* $\Psi\rho^{-1} \in L^1\left(\Omega \times [0,T] \times \mathbb{R}^d\right)$*, then there exist two constants* $c > 0$ *and* $C > 0$ *such that*

$$c\mathbb{E}\int_t^T \int_{\mathbb{R}^d} |\Psi(s,x)|\rho(x)^{-1}dx \leqslant \mathbb{E}\int_t^T \int_{\mathbb{R}^d} \left|\Psi\left(s, X_s^{t,x,i}\right)\right| \rho(x)^{-1}dx$$
$$\leqslant C\mathbb{E}\int_t^T \int_{\mathbb{R}^d} |\Psi(s,x)|\rho(x)^{-1}dx$$

The constants c *and* C *depend on* T, ρ*, the bounds of* σ *and the bounds of the first (resp. first and second) derivatives of* b *(resp. of* σ *).*

Theorem 2.12. *Let* $\left(X_s^{t,\xi,\alpha_t}, Y_s^{t,\xi,\alpha_t}, Z_s^{t,\xi,\alpha_t}, \Lambda_s^{t,\xi,\alpha_t}\right)$*,* $t \leqslant s \leqslant T$ *be the unique solution of the problem* (2.1.13)*. Then* $\theta(t,x,i) = Y_t^{t,x,\alpha_t}$ *is a weak solution of* (2.2.8) *with* $\theta(T,x,i) = h(x)$*.*

Proof. Let $\tilde{f}^m$ (resp. $\tilde{g}^m, \theta^m, \tilde{\sigma}^m$) be smooth functions which approximate $\tilde{f}$ (resp. $\tilde{g}, \theta, \tilde{\sigma}$) and satisfy Assumption (A) , and $\left(X_{s,m}^{t,x,\imath}, Y_{s,m}^{t,x,\imath}, Z_{s,m}^{t,x,\imath}, \Lambda_{s,m}^{t,x,\imath}\right)$, $t \leqslant s \leqslant T$ be the unique solution of the following equations

$$\begin{cases} X_{s,m} &= x + \int_t^s \tilde{f}^m(r, X_{r,m}, Y_{r,m}, Z_{r,m}, \alpha_r)dr + \int_t^s \tilde{\sigma}^m(r, X_{r,m}, Y_{r,m}, Z_{r,m}, \alpha_r)dW_r, \\ Y_{s,m} &= \theta^m(X_T, \alpha_T) + \int_s^T \tilde{g}^m(r, X_{r,m}, Y_{r,m}, Z_{r,m}, \alpha_r)dr - \int_s^T Z_{r,m}dW_r - \int_s^T \Lambda_{r,m} \bullet dM_r, \\ s &\in [t,T]. \end{cases}$$

Put $\theta(t,\cdot,i_0) = Y_t^{t,\cdot,i_0}$ and $\theta^m(t,\cdot,i_0) = Y_{t,m}^{t,\cdot,i_0}$. From Lemma 2.6, we know that $\theta^m(t,x,i_0)$ is the unique solution of the following partial differential equation

$$\frac{\partial \theta_k^m}{\partial t}(t,x,i_0) + \tilde{\mathcal{L}}^m \theta_k^m(t,x,i_0) + \tilde{g}_k^m\Big(t,x,\theta^m(t,x,i_0), \tilde{U}^m(t,x,i_0), i_0\Big) + \sum_{j_0 \in \mathcal{M}} q_{i_0 j_0}\theta_k^m(t,x,j_0) = 0,$$
$$\theta(T,x,i_0) = \tilde{h}(x,i_0), \qquad \forall\, t \in [0,T], x \in \mathbb{R}^p, k = 1,2,\ldots,q, \tag{2.3.4}$$

where

$$\tilde{U}^m(t,x,i_0) = \nabla_x \theta^m(t,x,i_0)\tilde{\sigma}^m\big(t,x,\theta^m(t,x,i_0),i_0\big),$$
$$\tilde{W}^m(t,x,i_0,j_0) = \theta^m(t,x,j_0) - \theta^m(t,x,i_0),$$

and

$$\tilde{\mathcal{L}}^m\theta_k^m(t,x,i_0) = \frac{1}{2}\sum_{i,j=1}^{p} \tilde{a}_{ij}^m\big(t,x,\theta^m(t,x,i_0),i_0\big)\frac{\partial^2\theta_k^m}{\partial x_i\partial x_j}(t,x,i_0)$$
$$+\sum_{i=1}^{p}\tilde{f}_i^m\Big(t,x,\theta^m(t,x,i_0),\tilde{U}^m(t,x,i_0),i_0\Big),i_0\Big)\frac{\partial\theta_k^m}{x_i}(t,x,i_0)$$

Thus

$$\begin{aligned}
&\int_t^T\int_{\mathbb{R}^p}\theta_k^m(s,x,i_0)D_s\varphi_k(s,x)dxds+\int_{\mathbb{R}^p}\theta_k^m(T,x,i_0)\varphi_k(t,x)dx-\int_{\mathbb{R}^p}\theta_k^m(x,i_0)\varphi_k(T,x)dx\\
&+\frac{1}{2}\int_t^T\int_{\mathbb{R}^p}\sum_{i,j=1}^{p}\tilde{a}_{ij}^m\big(s,x,\theta^m(s,x,i_0),i\big)D_i\theta_k^m(s,x,i_0)D_j\varphi_k(s,x)dxds\\
&+\int_t^T\int_{\mathbb{R}^p}\theta_k^m(s,x,i_0)\sum_{i=1}^{p}D_i((\tilde{f}_i^m-\tilde{A}_i^m)\varphi_k(s,x))dxds\\
&-\int_t^T\int_{\mathbb{R}^p}\tilde{g}_k^m(t,x,\theta^m(s,x,i_0),\tilde{U}^m(s,x,i_0),i_0),i_0)\varphi_k(s,x)dxds\\
&-\int_t^T\int_{\mathbb{R}^p}\sum_{j_0\in\mathcal{M}}q_{i_0j_0}\theta_k^m(t,x,j_0)\varphi_k(s,x)dxds=0,
\end{aligned}\tag{2.3.5}$$

where $\tilde{A}_i^m(s,x,i_0)=\frac{1}{2}\sum_{j=1}^{p}D_j(a_{ij}^m(s,x,\theta^m(s,x,i_0),i_0))$.

On the other hand, let ρ be a weight function as mentioned in Lemma 2.11. Then

$$\begin{aligned}
&||\theta^{m_1}(t,x,i)-\theta^{m_2}(t,x,i)||_{\mathcal{H}}\\
&=\mathbb{E}\Big[\int_t^T\int_{\mathbb{R}^d}\Big(\big|\theta^{m_1}\big(s,X_s^{t,x,i},i\big)-\theta^{m_2}\big(s,X_s^{t,x,i},i\big)\big|^2+\Big\|\tilde{U}^{m_1}(s,X_s^{t,x,i},i)-\tilde{U}^{m_2}(s,X_s^{t,x,i},i)\Big\|^2\\
&\quad+\sum_{j\in I}\Big|\tilde{W}^{m_1}(s,X_s^{t,x,i},i,j)-\tilde{W}^{m_2}(s,X_s^{t,x,i},i,j)\Big|^2q_{ij}\Big)\rho^{-1}(x)dxds\Big]\\
&=\mathbb{E}\Big[\int_t^T\int_{\mathbb{R}^d}\Big(\Big|\tilde{Y}_{s,m_1}^{t,x,i}-\tilde{Y}_{s,m_2}^{t,x,i}\Big|^2+\Big\|\tilde{Z}_{s,m_1}^{t,x,i}-\tilde{Z}_{s,m_2}^{t,x,i}\Big\|^2+\sum_{j\in\mathcal{M}}\Big|\tilde{\Lambda}_{s,m_1}^{i,j,t,x,i}-\tilde{\Lambda}_{s,m_2}^{i,j,t,x,i}\Big|^2q_{ij}\Big)\rho^{-1}(x)dxds\Big].
\end{aligned}$$

In view of Assumption (B), standard calculations (see Corollary 2.5) show that

$$\mathbb{E}\Big[\int_t^T\int_{\mathbb{R}^d}\Big(\Big|\tilde{Y}_{s,m}^{t,x,i}-\tilde{Y}_{s}^{t,x,i}\Big|^2+\Big\|\tilde{Z}_{s,m}^{t,x,i}-\tilde{Z}_{s}^{t,x,i}\Big\|^2+\sum_{j\in\mathcal{M}}\Big|\tilde{\Lambda}_{s,m}^{i,j,t,x,i}-\tilde{\Lambda}_{s}^{i,j,t,x,i}\Big|^2q_{ij}\Big)\rho^{-1}(x)dxds\Big]\to 0$$

as $m\to\infty$. So $||\theta^{m_1}(t,x,i)-\theta^{m_2}(t,x,i)||_{\mathcal{H}}\to 0$ as $m_1,m_2\to\infty$, which means that $\theta^m(t,x,i)$ is a Cauchy sequence in $\mathcal{H}$. Therefore, there exists $\theta\in\mathcal{H}$ such that $\theta^m\to\theta$ in $\mathcal{H}$, which implies that $(\theta^m, \nabla_x \theta^m \tilde{\sigma}^m) \to (\theta, \nabla_x \theta \sigma)$. This result together with passing the limit as $m \to \infty$ in (2.3.5) can verify that θ is a weak solution of (2.2.8). $\square$

CHAPTER 3
FBSDEs with Regime Switching Under Monotoncity Conditions

In this chapter, we first work on backward stochastic differential equations with Markovian switching and derive useful estimates for the solutions. Then, we focus on forward-backward stochastic differential equations and provide sufficient conditions for the existence and uniqueness of the solutions.

1 Estimate of BSDEs with Markovian Switching

Lemma 3.1. *Let $F_i : [0,T] \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathcal{M} \times \Omega \to \mathbb{R}^d$, $i = 1,2$, satisfy*

$$\big|F_i(t,y_1,z_1,i_0) - F_i(t,y_2,z_2,i_0)\big| \leqslant C\big(|y_1 - y_2| + \|z_1 - z_2\|\big), \quad \mathbb{P}\text{-}a.s.$$

for any $t \in [0,T]$, $y_1, y_2 \in \mathbb{R}^d$, $z_1, z_2 \in \mathbb{R}^{d\times d}$, and $i_0 \in \mathcal{M}$ (with the constant C being deterministic). In addition, $F_i(\cdot,0,0,i_0) \in \mathcal{L}^2(0,T;\mathbb{R}^d)$ and $F_i(t,y,z,i_0)$ is $\mathcal{F}$-adapted for any $i = 1,2$, $y \in \mathbb{R}^d$, $z \in \mathbb{R}^{d\times d}$, and $i_0 \in \mathcal{M}$. Let $\xi_i \in \mathcal{L}^2_{\mathcal{F}_T}(\mathbb{R}^d)$ and (Y_i, Z_i, Λ_i), $i = 1,2$, be the solutions of the BSDEs

$$Y_i(t) = \xi_i - \int_t^T F_i\big(s, Y_i(s), Z_i(s), \alpha(s)\big)ds - \int_t^T Z_i(s)dW_s - \int_t^T \Lambda_i(s) \bullet dM_s, \quad t \in [0,T].$$

Then

$$\mathbb{E}\bigg(\sup_{0\leqslant t\leqslant T}\big|Y_1(t) - Y_2(t)\big|^2 + \int_0^T \big\|Z_1(t) - Z_2(t)\big\|^2 dt + \int_0^T \big|\Lambda_1(t) - \Lambda_2(t)\big|^2 \bullet d[M]_t\bigg)$$
$$\leqslant C\bigg(\mathbb{E}|\xi_1 - \xi_2|^2 + \mathbb{E}\int_0^T \big|F_1\big(t, Y_1(s), Z_1(s), \alpha(s)\big) - F_2\big(t, Y_1(s), Z_1(s), \alpha(s)\big)\big|^2 ds\bigg).$$

Proof. Denote $\widehat{Y}(\cdot) = Y_1(\cdot) - Y_2(\cdot)$, $\widehat{Z}(\cdot) = Z_1(\cdot) - Z_2(\cdot)$, $\widehat{\xi} = \xi_1 - \xi_2$, and

$$\widehat{F}(\cdot) = F_1\big(\cdot, Y_1(\cdot), Z_1(\cdot), \alpha(\cdot)\big) - F_2\big(\cdot, Y_1(\cdot), Z_1(\cdot), \alpha(\cdot)\big).$$

Using Itô formula for $|\widehat{Y}(\cdot)|^2$ we have

$$\begin{aligned}
&|\widehat{Y}(t)|^2 + \int_t^T \|\widehat{Z}(s)\|^2 ds + \int_t^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s \\
&= |\widehat{\xi}|^2 - 2\int_t^T \Big\langle \widehat{Y}(s), F_1\big(s, Y_1(s), Z_1(s), \alpha(s)\big) - F_2\big(s, Y_2(s), Z_2(s), \alpha(s)\big)\Big\rangle ds \\
&\qquad - 2\int_t^T \Big\langle \widehat{Y}(s), \widehat{Z}(s) dW_s\Big\rangle - 2\int_t^T \Big\langle \widehat{Y}(s), \widehat{\Lambda}(s) \bullet dM_s\Big\rangle \\
&\leqslant |\widehat{\xi}|^2 + 2\int_t^T \Big[|\widehat{Y}(s)||\widehat{F}(s)| + C|\widehat{Y}(s)|\Big(|\widehat{Y}(s)| + \|\widehat{Z}(s)\|\Big)\Big] ds \\
&\qquad - 2\int_t^T \Big\langle \widehat{Y}(s), \widehat{Z}(s) dW_s\Big\rangle - 2\int_t^T \Big\langle \widehat{Y}(s), \widehat{\Lambda}(s) \bullet dM_s\Big\rangle,
\end{aligned}$$

which together with the Cauchy-Schwarz inequality implies

$$\begin{aligned}
&|\widehat{Y}(t)|^2 + \frac{1}{2}\int_t^T \|\widehat{Z}(s)\|^2 ds + \int_t^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s \\
&\leqslant |\widehat{\xi}|^2 + \int_t^T \Big[\big(1 + 2C + 2C^2\big)|\widehat{Y}(s)|^2 + |\widehat{F}(s)|^2\Big] ds \\
&\qquad - 2\int_t^T \Big\langle \widehat{Y}(s), \widehat{Z}(s) dW_s\Big\rangle - 2\int_t^T \Big\langle \widehat{Y}(s), \widehat{\Lambda}(s) \bullet dM_s\Big\rangle.
\end{aligned} \tag{3.1.1}$$

Taking the expectations in above inequality, we obtain

$$\begin{aligned}
\mathbb{E}|\widehat{Y}(t)|^2 &\leqslant \mathbb{E}\left(|\widehat{Y}(t)|^2 + \frac{1}{2}\int_t^T \|\widehat{Z}(s)\|^2 ds + \int_t^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s\right) \\
&\leqslant \mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds\Big) + \big(1 + 2C + 2C^2\big)\int_t^T \mathbb{E}|\widehat{Y}(s)|^2 ds.
\end{aligned} \tag{3.1.2}$$

By the Gronwall inequality,

$$\mathbb{E}|\widehat{Y}(t)|^2 \leqslant K_1 \mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds\Big), \qquad \forall t \in [0, T]. \tag{3.1.3}$$

Using this inequality and (3.1.2) (with $t = 0$) yields

$$\mathbb{E}\left(\int_0^T \|\widehat{Z}(s)\|^2 ds + \int_0^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s\right) \leqslant K_2 \mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds\Big). \tag{3.1.4}$$

Next, using $\int_t^T = \int_0^T - \int_0^t$ for stochastic integrals in (3.1.1), we obtain

$$
\begin{aligned}
&|\widehat{Y}(t)|^2 + \frac{1}{2}\int_t^T \|\widehat{Z}(s)\|^2 ds + \int_t^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s \\
\leqslant &|\widehat{\xi}|^2 + \int_t^T \Big[\big(1 + 2C + 2C^2\big)|\widehat{Y}(s)|^2 + |\widehat{F}(s)|^2\big)\Big] ds \\
&\quad - 2\int_0^T \langle \widehat{Y}(s), \widehat{Z}(s)dW_s\rangle - 2\int_0^T \langle \widehat{Y}(s), \widehat{\Lambda}(s) \bullet dM_s\rangle \\
&\qquad + 2\int_0^t \langle \widehat{Y}(s), \widehat{Z}(s)dW_s\rangle + 2\int_0^t \langle \widehat{Y}(s), \widehat{\Lambda}(s) \bullet dM_s\rangle.
\end{aligned}
$$

Subsequently, according to the Burkholder-Davis-Gundy inequality, the Hölder inequality, and the Cauchy-Schwartz inequality,

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} |\widehat{Y}(t)|^2\Big] \\
\leqslant &\mathbb{E}\left[|\widehat{\xi}|^2 + \big(1 + 2C + 2C^2\big)\int_0^T |\widehat{Y}(s)|^2 ds + \int_0^T |\widehat{F}(s)|^2 ds\right] \\
&\quad + 4\mathbb{E}\Big[\sup_{t\in[0,T]} \Big|\int_0^t \langle \widehat{Y}(s), \widehat{Z}(s)dW_s\rangle\Big|\Big] + 4\mathbb{E}\Big[\sup_{t\in[0,T]} \Big|\int_0^t \langle \widehat{Y}(s), \widehat{\Lambda}(s) \bullet dM_s\rangle\Big|\Big] \\
\leqslant &K\mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds\Big) + \frac{1}{4}\mathbb{E}\Big[\sup_{t\in[0,T]} |\widehat{Y}(t)|^2\Big] + K^2\mathbb{E}\int_0^T \|\widehat{Z}(s)\|^2 ds \\
&\quad + \frac{1}{4}\mathbb{E}\Big[\sup_{t\in[0,T]} |\widehat{Y}(t)|^2\Big] + K^2\mathbb{E}\int_0^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{t\in[0,T]} |\widehat{Y}(t)|^2\Big] \leqslant &C\mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds + \int_0^T \|\widehat{Z}(s)\|^2 ds + \int_0^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s\Big) \\
\leqslant &C\mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds\Big).
\end{aligned}
\tag{3.1.5}
$$

Note that we have used (3.1.4) in the last inequality. Again, combining (3.1.4) with (3.1.5) lead to

$$
\mathbb{E}\Big(\sup_{t\in[0,T]} |\widehat{Y}(t)|^2 + \int_0^T \|\widehat{Z}(s)\|^2 ds + \int_0^T |\widehat{\Lambda}(s)|^2 \bullet d[M]_s\Big) \leqslant C\mathbb{E}\Big(|\widehat{\xi}|^2 + \int_0^T |\widehat{F}(s)|^2 ds\Big).
$$

This completes the proof. □

2 FBSDEs with Markovian Switching

In this section we will develop the theory of FBSDEs with Markovian switching using a different approach. More precisely, we focus on the forward backward equations without the presence of mean-field terms and will use the continuation method and monotonicity condition to examine the well-posedness of the underlying FBSDEs. We will provide several different conditions that guarantee the existence and uniqueness of solutions of such FBSDEs with Markovian switching. The obtained results here are vital to further study these systems in a much more general settings, where the conditional mean-fields are coupled in the systems. For this section we will be working with equation

$$\begin{cases} X_t = \xi + \int_0^t f_0(s, X_s, Y_s, Z_s, \alpha_s)ds + \int_0^t \sigma_0(s, X_s, Y_s, Z_s, \alpha_s)dW_s, \\ Y_t = h_0(X_T, \alpha_T) - \int_t^T g_0(s, X_s, Y_s, Z_s, \alpha_s)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \quad t \in [0, T]. \end{cases} \tag{3.2.1}$$

Throughout this section, we assume the following assumption.

Assumption ($\mathbf{C}_0$).

($C_0$1) For each fixed $i_0 \in \mathcal{M}$, the functions $f_0(\cdot,\cdot,\cdot,\cdot,\cdot,i_0)$, $g_0(\cdot,\cdot,\cdot,\cdot,\cdot,i_0)$, and $\sigma_0(\cdot,\cdot,\cdot,\cdot,\cdot,i_0)$ are $\mathcal{F}$-progressively measurable and Lipschitz in (x, y, z) uniformly in (t, ω). That is, for $\varphi_0 = f_0$, g_0, or σ_0, there exists a (deterministic) constant C_θ such that for any $t \in [0, T]$, $i_0 \in \mathcal{M}$, and $\theta_1 = (x_1, y_1, z_1), \theta_2 = (x_2, y_2, z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$ we have

$$\big|\varphi_0\big(t, \theta_1, i_0\big) - \varphi_0\big(t, \theta_2, i_0\big)\big| \leqslant C_\theta \big\|\theta_1 - \theta_2\big\|, \quad \mathbb{P}\text{-a.s.}$$

where $\|\theta_1 - \theta_2\| = |x_1 - x_2| + |y_1 - y_2| + \|z_1 - z_2\|$.

($C_0$2) For each fixed $(x, i_0) \in \mathbb{R}^d \times \mathcal{M}$, the function $h_0(\cdot, x, i_0)$ belongs to $\mathcal{L}^2(\mathbb{R}^d)$. In addition, $h_0(\omega, \cdot, i_0)$ is Lipschitz uniformly in (ω, i_0). That is, for any $i_0 \in \mathcal{M}$, and $x_1, x_2 \in \mathbb{R}^d$, there exist a (deterministic) constant c such that

$$\big|h_0\big(x_1, i_0\big) - h_0\big(x_2, i_0\big)\big| \leqslant c\big|x_1 - x_2\big|, \quad \mathbb{P}\text{-a.s.}$$

For $t \in [0,T]$, $i_0 \in \mathcal{M}$, and $\theta_1 = (x_1, y_1, z_1), \theta_2 = (x_2, y_2, z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$, denote

$$\begin{aligned}\Psi_0\big(t,\theta_1,\theta_2,i_0\big) &= \big\langle f_0\big(t,\theta_1,i_0\big) - f_0\big(t,\theta_2,i_0\big), y_1 - y_2\big\rangle + \big\langle g_0\big(t,\theta_1,i_0\big) - g_0\big(t,\theta_2,i_0\big), x_1 - x_2\big\rangle \\ &\quad + \big[\sigma_0\big(t,\theta_1,i_0\big) - \sigma_0\big(t,\theta_2,i_0\big), z_1 - z_2\big]. \end{aligned} \tag{3.2.2}$$

Recall that $\langle x, y\rangle$ is the dot product of x and y and $[A,B] = \sum_{i=1}^{d}\langle A_i, B_i\rangle$, where A_i and B_i, $i = 1, 2, \ldots, d$, are the i-th column of $d \times d$ matrices A and B. To obtain the existence and uniqueness of solution of the forward backward system (3.2.1), we first make the following assumption.

Assumption $(\mathbf{H}_0)$.

($\mathrm{H}_0$1) There exists a positive constant K_h such that for any $x_1, x_2 \in \mathbb{R}^d$ and $i_0 \in \mathcal{M}$,

$$\big\langle h_0\big(x_1,i_0\big) - h_0\big(x_2,i_0\big), x_1 - x_2\big\rangle \geqslant K_h\big|x_1 - x_2\big|^2, \quad \mathbb{P}\text{-a.s.}$$

($\mathrm{H}_0$2) There exists a positive constant K_Ψ such that for any $t \in [0,T]$, $\theta_1 = (x_1, y_1, z_1), \theta_2 = (x_2, y_2, z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$, and $i_0 \in \mathcal{M}$,

$$\Psi_0\big(t,\theta_1,\theta_2,i_0\big) \leqslant -K_\Psi\Big(\big|x_1 - x_2\big|^2 + \big|y_1 - y_2\big|^2 + \big\|z_1 - z_2\big\|^2\Big), \quad \mathbb{P}\text{-a.s.}$$

We are now in a position to state the following theorem.

Theorem 3.2. *Under assumptions* $(\mathbf{C}_0)$ *and* $(\mathbf{H}_0)$, *there exists a unique quadruple of processes* (X, Y, Z, Λ) *which solves the system of FBSDEs with Markovian switching (3.2.1).*

In order to prove Theorem 3.2, we present some preliminary results on FBSDEs with Markov switching. First, we consider linear FBSDEs in the following lemma.

Lemma 3.3. *Suppose that* $\big(\bar f_0(\cdot), \bar\sigma_0(\cdot), \bar g_0(\cdot)\big) \in \mathcal{L}^2(0,T;\mathbb{R}^d) \times \mathcal{L}^2(0,T;\mathbb{R}^{d\times d}) \times \mathcal{L}^2(0,T;\mathbb{R}^d)$ *and* $\bar h_0 \in \mathcal{L}^2(\mathbb{R}^d)$, *then the following linear FBSDE*

$$\begin{cases} X_t = x + \displaystyle\int_0^t \big(-Y_s + \bar f_0(s)\big)ds + \int_0^t \big(-Z_s + \bar\sigma_0(s)\big)dW_s, \\ Y_t = \big(X_T + \bar h_0\big) - \displaystyle\int_t^T \big(-X_s + \bar g_0(s)\big)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s \end{cases} \tag{3.2.3}$$

has a unique adapted solution (X, Y, Z, Λ).

Proof. We consider the following BSDE:

$$\bar{Y}_t = \bar{h}_0 - \int_t^T \big(\bar{Y}_s + \bar{g}_0(s) - \bar{f}_0(s)\big)ds - \int_t^T \big(2\bar{Z}_s - \bar{\sigma}_0(s)\big)dW_s - \int_t^T \bar{\Lambda}_s \bullet dM_s.$$

According to [8, Theorem 3.4], the above equation has a unique adapted solution $(\bar{Y}, \bar{Z}, \bar{\Lambda})$. Then we solve the following forward equation:

$$X_t = x + \int_0^t \big(- X_s - \bar{Y}_s + \bar{f}_0(s)\big)ds + \int_0^t \big(- \bar{Z}_s + \bar{\sigma}_0(s)\big)dW_s$$

and set $Y = X + \bar{Y}$, $Z = \bar{Z}$, and $\Lambda = \bar{\Lambda}$. It is easily seen that (X, Y, Z, Λ) is a solution of equation (3.2.3). Hence, the existence is proved. Finally, the proof of uniqueness is similar to that of Theorem 3.2. $\square$

Next, for $\gamma \in \mathbb{R}$ define

$$\begin{aligned} f_0^\gamma(t, x, y, z, i_0) &= \gamma f_0(t, x, y, z, i_0) + (1 - \gamma)(-y), \\ \sigma_0^\gamma(t, x, y, z, i_0) &= \gamma \sigma_0(t, x, y, z, i_0) + (1 - \gamma)(-z), \\ g_0^\gamma(t, x, y, z, i_0) &= \gamma g_0(t, x, y, z, i_0) + (1 - \gamma)(-x), \\ h_0^\gamma(x, i_0) &= \gamma h_0(x, i_0) + (1 - \gamma)x, \end{aligned} \tag{3.2.4}$$

and consider the following system of equations

$$\begin{cases} X_t = x + \displaystyle\int_0^t \Big(f_0^\gamma(s, \Theta_s, \alpha_s) + \bar{f}_0(s)\Big)ds + \int_0^t \Big(\sigma_0^\gamma(s, \Theta_s, \alpha_s) + \bar{\sigma}_0(s)\Big)dW_s, \\ Y_t = \big(h_0^\gamma(X_T, \alpha_T) + \bar{h}_0\big) - \displaystyle\int_t^T \Big(g_0^\gamma(s, \Theta_s, \alpha_s) + \bar{g}_0(s)\Big)ds - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \end{cases} \tag{3.2.5}$$

where $\Theta = (X, Y, Z)$. To proceed, we present the following lemma.

Lemma 3.4. *Assume that Assumptions* $(\mathbf{C}_0)$ *and* $(\mathbf{H}_0)$ *are in force. In addition, assume that for a given* $\gamma_0 \in [0, 1)$ *and for any* $\big(\bar{f}_0(\cdot), \bar{\sigma}_0(\cdot), \bar{g}_0(\cdot), \bar{h}_0\big) \in \mathcal{L}^2(0, T; \mathbb{R}^d) \times \mathcal{L}^2(0, T; \mathbb{R}^{d\times d}) \times \mathcal{L}^2(0, T; \mathbb{R}^d) \times \mathcal{L}^2(\mathbb{R}^d)$, *the system* (3.2.5) *has an adapted solution. Then there exists a constant* $\delta_0 \in (0, 1)$ *depending only on* C_θ, c, K_h, K_Ψ *and* T, *such that for any* $\gamma \in [\gamma_0, \gamma_0 + \delta_0]$

and for any $\big(\bar{f}_0(\cdot),\bar{\sigma}_0(\cdot),\bar{g}_0(\cdot),\bar{h}_0\big)\in\mathcal{L}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{L}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(\mathbb{R}^d)$, *system* (3.2.5) *has a solution* $(X,Y,Z,\Lambda)\in\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{S}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{M}^2(0,T;\mathbb{R}^d)$.

Proof. In view of (3.2.4),

$$\begin{aligned}
f_0^{\gamma_0+\delta}(t,x,y,z,i_0) &= f_0^{\gamma_0}(t,x,y,z,i_0)+\delta\big(y+f_0(t,x,y,z,i_0)\big),\\
\sigma_0^{\gamma_0+\delta}(t,x,y,z,i_0) &= \sigma_0^{\gamma_0}(t,x,y,z,i_0)+\delta\big(z+\sigma_0(t,x,y,z,i_0)\big),\\
g_0^{\gamma_0+\delta}(t,x,y,z,i_0) &= g_0^{\gamma_0}(t,x,y,z,i_0)+\delta\big(x+g_0(t,x,y,z,i_0)\big),\\
h_0^{\gamma_0+\delta}(x,i_0) &= h_0^{\gamma_0}(x,i_0)+\delta\big(-x+h_0(x,i_0)\big).
\end{aligned}$$

Put $(X^0,Y^0,Z^0,\Lambda^0)=(0,0,0,0)$ and $\Theta^n=(X^n,Y^n,Z^n)$ for $n=0,1,2,\ldots$. According to the assumption, the following recursive systems always have unique solutions.

$$\begin{cases}
X_t^{n+1} = x+\displaystyle\int_0^t\Big[f_0^{\gamma_0}(s,\Theta_s^{n+1},\alpha_s)+\delta\big(Y_s^n+f_0(s,\Theta_s^n,\alpha_s)\big)+\bar{f}_0(s)\Big]ds\\
\qquad\qquad +\displaystyle\int_0^t\Big[\sigma_0^{\gamma_0}(s,\Theta_s^{n+1},\alpha_s)+\delta\big(Z_s^n+\sigma_0(s,\Theta_s^n,\alpha)\big)+\bar{\sigma}_0(s)\Big]dW_s,\\
Y_t^{n+1} = \Big[h_0^{\gamma_0}(X_T^{n+1},\alpha_T)+\delta\big(-X_T^n+h_0(X_T^n,\alpha_T)\big)+\bar{h}_0\Big]\\
\qquad\qquad -\displaystyle\int_t^T\Big[g_0^{\gamma_0}(s,\Theta_s^{n+1},\alpha_s)+\delta\big(X_s^n+g_0(s,\Theta_s^n,\alpha_s)\big)+\bar{g}_0(s)\Big]ds\\
\qquad\qquad -\displaystyle\int_t^T Z_s^{n+1}dW_s-\int_t^T\Lambda_s^{n+1}\bullet dM_s.
\end{cases}\tag{3.2.6}$$

Denote $\hat{\Theta}_t^{n+1}=\Theta_t^{n+1}-\Theta_t^n$. In addition, put

$$\hat{X}_t^{n+1}=X_t^{n+1}-X_t^n,\ \hat{Y}_t^{n+1}=Y_t^{n+1}-Y_t^n,\ \hat{Z}_t^{n+1}=Z_t^{n+1}-Z_t^n,\ \hat{\Lambda}_t^{n+1}=\Lambda_t^{n+1}-\Lambda_t^n.$$

$$\hat{h}_0^n(t):=h_0(X_t^n,\alpha_t)-h_0(X_t^{n-1},\alpha_t),\quad \hat{h}_0^{n,\gamma}(t):=h_0^\gamma(X_t^n,\alpha_t)-h_0^\gamma(X_t^{n-1},\alpha_t),$$

and, for $\varphi=f_0,g_0,\sigma_0$,

$$\hat{\varphi}^{n+1}(t)=\varphi(t,\Theta_t^{n+1},\alpha_t)-\varphi(t,\Theta_t^n,\alpha_t),\quad \hat{\varphi}^{n+1,\gamma}(t)=\varphi^\gamma(t,\Theta_t^{n+1},\alpha_t)-\varphi^\gamma(t,\Theta_t^n,\alpha_t).$$

Applying Itô's formula to $\langle \hat{X}_t^{n+1}, \hat{Y}_t^{n+1}\rangle$ and taking the expectation we obtain

$$\begin{aligned}&\mathbb{E}\Big\langle \hat{X}_T^{n+1}, \hat{h}_0^{n+1,\gamma_0}(T)\Big\rangle = \delta\mathbb{E}\Big\langle \hat{X}_T^{n+1}, \hat{X}_T^{n} - \hat{h}_0^{n}(T)\Big\rangle \\ &\quad + \mathbb{E}\int_0^T \Big(\Big\langle \hat{X}_s^{n+1}, \hat{g}_0^{n+1,\gamma_0}(s)\Big\rangle + \Big\langle \hat{Y}_s^{n+1}, \hat{f}_0^{n+1,\gamma_0}(s)\Big\rangle + \big[\hat{Z}_s^{n+1}, \hat{\sigma}_0^{n+1,\gamma_0}(s)\big]\Big)ds \\ &\quad\quad + \delta\mathbb{E}\int_0^T \Big(\Big\langle \hat{X}_s^{n+1}, \hat{X}_s^{n} + \hat{g}_0^{n}(s)\Big\rangle + \Big\langle \hat{Y}_s^{n+1}, \hat{Y}_s^{n} + \hat{f}_0^{n}(s)\Big\rangle + \big[\hat{Z}_s^{n+1}, \hat{Z}_s^{n} + \hat{\sigma}_0^{n}(s)\big]\Big)ds.\end{aligned}$$

From Assumptions $(\mathbf{C}_0)$ and $(\mathbf{H}_0)$, it follows that

$$\mathbb{E}\left(|\hat{X}_T^{n+1}|^2 + \int_0^T \|\hat{\Theta}_s^{n+1}\|^2 ds\right) \leqslant \frac{\delta(1+C)}{K}\mathbb{E}\left(|\hat{X}_T^{n+1}||\hat{X}_T^{n}| + \int_0^T \|\hat{\Theta}_s^{n+1}\|\|\hat{\Theta}_s^{n}\| ds\right),$$

where $K = \min(1, K_h, K_\Psi)$ and $C = \max(c, C_\theta)$. Young's inequality implies

$$\mathbb{E}\left(|\hat{X}_T^{n+1}|^2 + \int_0^T \|\hat{\Theta}_s^{n+1}\|^2 ds\right) \leqslant \left(\frac{\delta(1+C)}{K}\right)^2 \mathbb{E}\left(|\hat{X}_T^{n}|^2 + \int_0^T \|\hat{\Theta}_s^{n}\|^2 ds\right).$$

Recall that $\forall n \geqslant 1$,

$$\hat{X}_T^{n} = \int_0^T \Big[\hat{f}_0^{n,\gamma_0}(s) + \delta\big(\hat{Y}_s^{n-1} + \hat{f}_0^{n-1}(s)\big)\Big] ds + \int_0^T \Big[\hat{\sigma}_0^{n,\gamma_0}(s) + \delta\big(\hat{Z}_s^{n-1} + \hat{\sigma}_0^{n-1}(s)\big)\Big] dW_s.$$

We can derive that there exists a constant $c_1 > 0$ that depends only on C and T such that

$$\mathbb{E}|\hat{X}_T^{n}|^2 \leqslant c_1\mathbb{E}\left(\int_0^T \|\hat{\Theta}_s^{n}\|^2 ds + \int_0^T \|\hat{\Theta}_s^{n-1}\|^2 ds\right), \quad \forall n \geqslant 1.$$

Hence, there exists a constant $c_2 > 0$ that depends only on C, K, and T such that

$$\mathbb{E}\int_0^T \|\hat{\Theta}_s^{n+1}\|^2 ds \leqslant c_2\delta^2\mathbb{E}\left(\int_0^T \|\hat{\Theta}_s^{n}\|^2 ds + \int_0^T \|\hat{\Theta}_s^{n-1}\|^2 ds\right), \quad \forall n \geqslant 1.$$

So, there exists a $\delta_0 \in (0,1)$ that depends only on C, K, and T such that when $0 < \delta \leqslant \delta_0$,

$$\mathbb{E}\int_0^T \|\hat{\Theta}_s^{n+1}\|^2 ds \leqslant \frac{1}{4}\mathbb{E}\int_0^T \|\hat{\Theta}_s^{n}\|^2 ds + \frac{1}{8}\mathbb{E}\int_0^T \|\hat{\Theta}_s^{n-1}\|^2 ds, \quad \forall n \geqslant 1.$$

In view of [25, Lemma 4.1], there exists a constant $\hat{c} > 0$ such that

$$\mathbb{E}\int_0^T \|\hat{\Theta}_s^{n}\|^2 ds \leqslant \hat{c}\left(\frac{1}{2}\right)^n, \quad \forall n \geqslant 0.$$

This implies that $(X_T^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{L}^2(\mathbb{R}^d)$ and $(X^n)_{n\geqslant 0}, (Y^n)_{n\geqslant 0}$ and $(Z^n)_{n\geqslant 0}$ are Cauchy sequences in $\mathcal{L}^2(0,T;\mathbb{R}^d)$, $\mathcal{L}^2(0,T;\mathbb{R}^d)$, $\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})$, respectively. Moreover, similar to (4.2.13) in the proof of Theorem 4.1, we have that $(\Lambda^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{M}^2(0,T;\mathbb{R}^d)$. We denote their limits by (X,Y,Z,Λ), respectively. Passing to the limit in equations (3.2.6), when $0<\delta\leqslant\delta_0$, (X,Y,Z,Λ) solves equations (3.2.5) for $\gamma=\gamma_0+\delta$. By standard estimates we can show that $X,Y\in\mathcal{S}^2(0,T;\mathbb{R}^d)$. $\square$

Now we are ready to prove Theorem 3.2.

Proof of Theorem 3.2.

First, we prove the uniqueness of the solution. Suppose that (X,Y,Z,Λ) and (X',Y',Z',Λ') are two solutions of (3.2.1). Taking Itô's formula and the expectations we get

$$
\begin{aligned}
&\mathbb{E}\big\langle X'_T-X_T, h_0(X'_T,\alpha_T)-h_0(X_T,\alpha_T)\big\rangle\\
&=\mathbb{E}\int_0^T\Big(\big\langle X'_s-X_s, g_0(s,X'_s,Y'_s,Z'_s,\alpha_s)-g_0(s,X_s,Y_s,Z_s,\alpha_s)\big\rangle\\
&\qquad\qquad+\big\langle Y'_s-Y_s, f_0(s,X'_s,Y'_s,Z'_s,\alpha_s)-f_0(s,X_s,Y_s,Z_s,\alpha_s)\big\rangle\\
&\qquad\qquad\qquad+\big[Z'_s-Z_s,\sigma_0(s,X'_s,Y'_s,Z'_s,\alpha_s)-\sigma_0(s,X_s,Y_s,Z_s,\alpha_s)\big]\Big)ds.
\end{aligned}
$$

By virtue of Assumption $(\mathbf{H}_0)$, we obtain

$$
K_h\mathbb{E}\Big[|X'_T-X_T|^2\Big]+K_\Psi\mathbb{E}\int_0^T\Big(|X'_s-X_s|^2+|Y'_s-Y_s|^2+\|Z'_s-Z_s\|^2\Big)ds\leqslant 0,
$$

which implies $X'_T=X_T$, $X'=X$, $Y'=Y$, and $Z'=Z$. Moreover, according to Lemma 3.1, assumption ($C_0$2), and the fact that $X'_T=X_T$, we get

$$
\mathbb{E}\int_0^T\big|\Lambda'_s-\Lambda_s\big|^2\bullet d[M]_s\leqslant C\mathbb{E}\big|h_0(X'_T,\alpha_T)-h_0(X_T,\alpha_T)\big|^2\leqslant C\mathbb{E}\big|X'_T-X_T\big|^2=0.
$$

This yields $\Lambda'=\Lambda$ in $\mathcal{M}^2(0,T;\mathbb{R}^d)$.

Next, we prove the existence of solution. From Lemma 3.3, when $\gamma=0$, for any $\big(\bar f_0(\cdot),\bar\sigma_0(\cdot),\bar g_0(\cdot)\big)\in\mathcal{L}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{L}^2(0,T;\mathbb{R}^d)$ and $\bar h_0\in\mathcal{L}^2(\mathbb{R}^d)$ the forward backward system (3.2.5) has an adapted solution. According to Lemma 3.4, for any

$\big(\bar{f}_0(\cdot), \bar{\sigma}_0(\cdot), \bar{g}_0(\cdot)\big) \in \mathcal{L}^2(0,T;\mathbb{R}^d) \times \mathcal{L}^2(0,T;\mathbb{R}^{d\times d}) \times \mathcal{L}^2(0,T;\mathbb{R}^d)$ and $\bar{h}_0 \in \mathcal{L}^2(\mathbb{R}^d)$, we can solve the system (3.2.5) successively for the case $\gamma \in [0,\delta_0]$, $[\delta_0, 2\delta_0]$, ... It turns out that, when $\gamma = 1$, for any $\big(\bar{f}_0(\cdot), \bar{\sigma}_0(\cdot), \bar{g}_0(\cdot)\big) \in \mathcal{L}^2(0,T;\mathbb{R}^d) \times \mathcal{L}^2(0,T;\mathbb{R}^{d\times d}) \times \mathcal{L}^2(0,T;\mathbb{R}^d)$ and $\bar{h}_0 \in \mathcal{L}^2(\mathbb{R}^d)$, the adapted solution of (3.2.5) exists, then we deduce immediately that the adapted solution of (3.2.1) exists. □

In what follows, we study how to weaken assumption $(\mathbf{H}_0)$ to get a solution of the FBSDE with Markovian switching (3.2.1). In fact, we can establish the existence and uniqueness of a solution to equation (3.2.1) if one of the following assumptions $(\mathbf{I}_0)$ and $(\mathbf{J}_0)$ below is satisfied.

Assumption $\mathbf{I}_0$.

($\mathrm{I}_0$1) There exists a positive constant K_h such that for any $x_1, x_2 \in \mathbb{R}^d$ and $i_0 \in \mathcal{M}$,

$$\big\langle h_0(x_1,i_0) - h_0(x_2,i_0), x_1 - x_2 \big\rangle \geqslant K_h |x_1 - x_2|^2, \quad \mathbb{P}\text{-a.s.}$$

($\mathrm{I}_0$2) There exists a positive constant K_Ψ such that for any $t \in [0,T]$, $\theta_1 = (x_1,y_1,z_1), \theta_2 = (x_2,y_2,z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$ and $i_0 \in \mathcal{M}$,

$$\Psi_0(t,\theta_1,\theta_2,i_0) \leqslant -K_\Psi |x_1 - x_2|^2, \quad \mathbb{P}\text{-a.s.}$$

Assumption $\mathbf{J}_0$.

($\mathrm{J}_0$1) For any $x_1, x_2 \in \mathbb{R}^d$ and $i_0 \in \mathcal{M}$, we have

$$\big\langle h_0(x_1,i_0) - h_0(x_2,i_0), x_1 - x_2 \big\rangle \geqslant 0, \quad \mathbb{P}\text{-a.s.}$$

($\mathrm{J}_0$2) There exists a positive constant K_Ψ such that for any $t \in [0,T]$, $\theta_1 = (x_1,y_1,z_1), \theta_2 = (x_2,y_2,z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$ and $i_0 \in \mathcal{M}$, we have

$$\Psi_0(t,\theta_1,\theta_2,i_0) \leqslant -K_\Psi \Big(|y_1 - y_2|^2 + \|z_1 - z_2\|^2\Big), \quad \mathbb{P}\text{-a.s.}$$

It is easily seen that both assumptions ($\mathrm{I}_0 2$) and ($\mathrm{J}_0 2$) are weaker than assumption ($\mathrm{H}_0 2$). Hence, the following result is an improvement of Theorem 3.2.

Theorem 3.5. *Let Assumption* $(\mathbf{C}_0)$ *and either Assumption* $(\mathbf{I}_0)$ *or Assumption* $(\mathbf{J}_0)$ *hold. Then there exists a unique process* $(X,Y,Z,\Lambda) \in \mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{M}^2(0,T;\mathbb{R}^d)$ *which is the solution of the FBSDEs with Markovian switching (3.2.1).*

The proof of Theorem 3.5 is divided into two parts. First, Theorem 3.5 is proved under Assumptions $(\mathbf{C}_0)$ and $(\mathbf{I}_0)$. Then, it is proved under Assumptions $(\mathbf{C}_0)$ and $(\mathbf{J}_0)$.

***Proof of Theorem* 3.5 *under Assumptions* $(\mathbf{C}_0)$ *and* $(\mathbf{I}_0)$.**

(1) Existence of a solution: Let $\delta > 0$ and consider the sequence $(X^n, Y^n, Z^n, \Lambda^n)_{n\geqslant 0}$ of processes defined recursively as follows: $(X^0, Y^0, Z^0, \Lambda^0) = (0,0,0,0)$ and $(X^{n+1}, Y^{n+1}, Z^{n+1}, \Lambda^{n+1}) \in \mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{M}^2(0,T;\mathbb{R}^d)$ satisfies

$$\begin{cases} X_t^{n+1} = \xi + \displaystyle\int_0^t \Big(f_0\big(s, X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1}, \alpha_s\big) - \delta Y_s^{n+1} + \delta Y_s^n\Big) ds \\ \qquad\qquad + \displaystyle\int_0^t \Big(\sigma_0\big(s, X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1}, \alpha_s\big) - \delta Z_s^{n+1} + \delta Z_s^n\Big) dW_s, \\ Y_t^{n+1} = h_0(X_T^{n+1}, \alpha_T) - \displaystyle\int_t^T g_0\big(s, X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1}, \alpha_s\big) ds \\ \qquad\qquad - \int_t^T Z_s^{n+1} dW_s - \int_t^T \Lambda_s^{n+1} \bullet dM_s. \end{cases} \tag{3.2.7}$$

In view of Theorem 3.2, these recursive FBSDEs have unique solutions. For $n \geqslant 0$ and $t \in [0,T]$, denote

$$\hat{X}_t^{n+1} := X_t^{n+1} - X_t^n,\ \hat{Y}_t^{n+1} := Y_t^{n+1} - Y_t^n,\ \hat{Z}_t^{n+1} := Z_t^{n+1} - Z_t^n,\ \hat{\Lambda}_t^{n+1} := \Lambda_t^{n+1} - \Lambda_t^n,$$

$$\hat{h}_0^n(t) := h_0(X_t^n, \alpha_t) - h_0(X_t^{n-1}, \alpha_t).$$

In addition, put $\Theta_s^{n+1} = (X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1})$, $\hat{\Theta}_s^{n+1} = (\hat{X}_s^{n+1}, \hat{Y}_s^{n+1}, \hat{Z}_s^{n+1})$, and for $\varphi = f_0, g_0, \sigma_0$,

$$\hat{\varphi}^{n+1}(t) := \varphi(t, \Theta_t^{n+1}, \alpha_t) - \varphi(t, \Theta_t^n, \alpha_t).$$

By taking the expectations after using Itô's formula for $\langle \hat{X}_t^{n+1}, \hat{Y}_t^{n+1}\rangle$, it is clear that

$$\begin{aligned}&\mathbb{E}\Big\langle \hat{X}_T^{n+1}, \hat{h}_0^{n+1}(T)\Big\rangle + \delta\mathbb{E}\int_0^T \Big(|\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\Big)ds\\ &= \delta\mathbb{E}\int_0^T \Big(\Big\langle \hat{Y}_s^{n+1}, \hat{Y}_s^n\Big\rangle + \big[\hat{Z}_s^{n+1}, \hat{Z}_s^n\big]\Big)ds\\ &\qquad + \mathbb{E}\int_0^T \Big(\Big\langle \hat{Y}_s^{n+1}, \hat{f}_0^{n+1}(s)\Big\rangle + \Big\langle \hat{X}_s^{n+1}, \hat{g}_0^{n+1}(s)\Big\rangle + \big[\hat{Z}_s^{n+1}, \hat{\sigma}_0^{n+1}(s)\big]\Big)\, ds.\end{aligned}$$

According to Assumption ($\mathbf{I}_0$) and Young's inequality, we obtain

$$K_h\mathbb{E}|\hat{X}_T^{n+1}|^2 + K_\Psi\mathbb{E}\int_0^T |\hat{X}_s^{n+1}|^2 ds + \frac{\delta}{2}\mathbb{E}\int_0^T \Big(|\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\Big)ds \leqslant \frac{\delta}{2}\mathbb{E}\int_0^T \Big(|\hat{Y}_s^n|^2 + \|\hat{Z}_s^n\|^2\Big)ds. \tag{3.2.8}$$

Again, using Itô's formula for $|\hat{Y}^n|^2$ and then taking the expectation yield

$$\mathbb{E}|\hat{Y}_t^n|^2 + \mathbb{E}\int_t^T \|\hat{Z}_s^n\|^2 ds + \mathbb{E}\int_t^T |\hat{\Lambda}_s^n|^2 \bullet d[M]_s = \mathbb{E}|\hat{h}_0^n(T)|^2 - 2\mathbb{E}\int_t^T \Big\langle \hat{Y}_s^n, \hat{g}_0^n(s)\Big\rangle ds.$$

In view of Assumption ($\mathrm{C}_0$1) and Young's inequality, for $t \leqslant T$, we have

$$\begin{aligned}&\mathbb{E}|\hat{Y}_t^n|^2 + \frac{1}{2}\mathbb{E}\int_t^T \|\hat{Z}_s^n\|^2 ds + \mathbb{E}\int_t^T |\hat{\Lambda}_s^n|^2 \bullet d[M]_s\\ &\quad\leqslant \mathbb{E}|\hat{h}_0^n(T)|^2 + C\mathbb{E}\int_0^T |\hat{X}_s^n|^2 ds + C\mathbb{E}\int_t^T |\hat{Y}_s^n|^2 ds.\end{aligned}$$

Then, applying Gronwall's inequality and using ($\mathrm{C}_0$2), we arrive at

$$\mathbb{E}|\hat{Y}_t^n|^2 \leqslant C\mathbb{E}\Big(|\hat{h}_0^n(T)|^2 + \int_0^T |\hat{X}_s^n|^2 ds\Big) \leqslant C\mathbb{E}\Big(|\hat{X}_T^n|^2 + \int_0^T |\hat{X}_s^n|^2 ds\Big).$$

Similarly to (3.1.3)-(3.1.4), we obtain

$$\mathbb{E}\Big(\int_0^T |\hat{Y}_s^n|^2 ds + \int_0^T \|\hat{Z}_s^n\|^2 ds + \int_0^T |\hat{\Lambda}_s^n|^2 \bullet d[M]_s\Big) \leqslant C\mathbb{E}\Big(|\hat{X}_T^n|^2 + \int_0^T |\hat{X}_s^n|^2 ds\Big). \tag{3.2.9}$$

Combining (3.2.9) and (3.2.8) lead to

$$\mathbb{E}\Big(|\hat{X}_T^{n+1}|^2 + \int_0^T |\hat{X}_s^{n+1}|^2 ds\Big) \leqslant \frac{\delta C}{2\min(K_h, K_\Psi)}\mathbb{E}\Big(|\hat{X}_T^n|^2 + \int_0^T |\hat{X}_s^n|^2 ds\Big).$$

If we take $\delta = \frac{\min(K_h, K_\Psi)}{C}$, then for any $n \geqslant 1$,

$$\mathbb{E}\Big(\big|\hat{X}_T^{n+1}\big|^2 + \int_0^T |\hat{X}_s^{n+1}|^2 ds\Big) \leqslant \frac{1}{2^n}\mathbb{E}\Big(\big|\hat{X}_T^1\big|^2 + \int_0^T |\hat{X}_s^1|^2 ds\Big)$$

and from (3.2.9), we also get

$$\mathbb{E}\Big(\int_0^T \big|\hat{Y}_s^n\big|^2 ds + \int_0^T \big\|\hat{Z}_s^n\big\|^2 ds + \int_0^T \big|\hat{\Lambda}_s^n\big|^2 \bullet d[M]_s\Big) \leqslant \frac{C}{2^{n-1}}\mathbb{E}\Big(\big|\hat{X}_T^1\big|^2 + \int_0^T |\hat{X}_s^1|^2 ds\Big).$$

It follows that $(X_T^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{L}^2(\mathbb{R}^d)$, $(X^n)_{n\geqslant 0}$, $(Y^n)_{n\geqslant 0}$ and $(Z^n)_{n\geqslant 0}$ are Cauchy sequences in $\mathcal{L}^2(0,T;\mathbb{R}^d)$, $\mathcal{L}^2(0,T;\mathbb{R}^d)$, and $\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})$, respectively, and $(\Lambda^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{M}^2(0,T;\mathbb{R}^d)$. Let X, Y, Z and Λ respectively be the limits of $(X^n)_{n\geqslant 0}$, $(Y^n)_{n\geqslant 0}$, $(Z^n)_{n\geqslant 0}$ and $(\Lambda^n)_{n\geqslant 0}$ in the corresponding spaces. Then, passing the limit in equations (3.2.7) yields that (X, Y, Z, Λ) is a solution of the FBSDEs with Markovian switching (3.2.1).

(2) Uniqueness of the solution: Let (X, Y, Z, Λ) and (X', Y', Z', Λ') be two solutions of equations (3.2.1). Taking the expectations together with using Itô's formula for $\langle X' - X, Y' - Y\rangle$ yields

$$\begin{aligned}
&\mathbb{E}\big\langle X_T' - X_T, h_0(X_T', \alpha_T) - h_0(X_T, \alpha_T)\big\rangle \\
&= \mathbb{E}\int_0^T \Big\{\big\langle X_s' - X_s, g_0(s, X_s', Y_s', Z_s', \alpha_s) - g_0(s, X_s, Y_s, Z_s, \alpha_s)\big\rangle \\
&\quad + \big\langle Y_s' - Y_s, f_0(s, X_s', Y_s', Z_s', \alpha_s) - f_0(s, X_s, Y_s, Z_s, \alpha_s)\big\rangle \\
&\quad + \big[Z_s' - Z_s, \sigma_0(s, X_s', Y_s', Z_s', \alpha_s) - \sigma_0(s, X_s, Y_s, Z_s, \alpha_s)\big]\Big\} ds.
\end{aligned}$$

It follows from both assumptions of $(\mathbf{I}_0)$ that

$$K_h \mathbb{E}\,|X_T' - X_T|^2 + K_\Psi \mathbb{E}\int_0^T |X_s' - X_s|^2 ds \leqslant 0$$

which implies $X'_T = X_T$ and $X' = X$. Now we take the expectations in Itô's formula for $|Y' - Y|^2$,

$$\mathbb{E}|Y'_t - Y_t|^2 + \mathbb{E}\int_t^T \|Z'_s - Z_s\|^2 ds + \mathbb{E}\int_t^T |\Lambda'_s - \Lambda_s|^2 \bullet d[M]_s$$
$$= \mathbb{E}|h_0(X'_T, \alpha_T) - h_0(X_T, \alpha_T)|^2 - 2\mathbb{E}\int_t^T \Big\langle Y'_s - Y_s, g_0(s, X'_s, Y'_s, Z'_s, \alpha_s) - g_0(s, X_s, Y_s, Z_s, \alpha_s)\Big\rangle ds. \tag{3.2.10}$$

Note that by both assumptions of $(\mathbf{C}_0)$, Young's inequality and $X'_T = X_T$ and $X' = X$ imply

$$\begin{aligned}
&\mathbb{E}|Y'_t - Y_t|^2 + \frac{1}{2}\mathbb{E}\int_t^T \|Z'_s - Z_s\|^2 ds + \mathbb{E}\int_t^T |\Lambda'_s - \Lambda_s|^2 \bullet d[M]_s \\
&\leqslant \mathbb{E}|h_0(X'_T, \alpha_T) - h_0(X_T, \alpha_T)|^2 + \mathbb{E}\int_t^T |X'_s - X_s|^2 ds + \mathbb{E}\int_t^T |Y'_s - Y_s|^2 ds \qquad (3.2.11)\\
&\leqslant c\mathbb{E}|X'_T - X_T|^2 + \mathbb{E}\int_t^T |X'_s - X_s|^2 ds + \mathbb{E}\int_t^T |Y'_s - Y_s|^2 ds = \mathbb{E}\int_t^T |Y'_s - Y_s|^2 ds.
\end{aligned}$$

By Gronwall's inequality, we get

$$\mathbb{E}|Y'_t - Y_t|^2 \leqslant 0,$$

yielding $Y' = Y$ and, by (3.2.11), $Z' = Z$ in $\mathcal{L}^2(0, T; \mathbb{R}^{d\times d})$ and $\Lambda' = \Lambda$ in $\mathcal{M}^2(0, T; \mathbb{R}^d)$. □

Proof of Theorem* 3.5 *under Assumptions* $(\mathbf{C}_0)$ *and* $(\mathbf{J}_0)$*.

(1) Existence of a solution: Let $\delta > 0$ and consider the sequence $(X^n, Y^n, Z^n, \Lambda^n)_{n\geqslant 0}$ of processes defined recursively as follows: $(X^0, Y^0, Z^0, \Lambda^0) = (0, 0, 0, 0)$ and $(X^{n+1}, Y^{n+1}, Z^{n+1}, \Lambda^{n+1}) \in \mathcal{S}^2(0, T; \mathbb{R}^d) \times \mathcal{S}^2(0, T; \mathbb{R}^d) \times \mathcal{L}^2(0, T; \mathbb{R}^{d\times d}) \times \mathcal{M}^2(0, T; \mathbb{R}^d)$ satisfies

$$\left\{\begin{aligned}
X_t^{n+1} &= \xi + \int_0^t f_0(s, X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1}, \alpha_s) ds + \int_0^t \sigma_0(s, X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1}, \alpha_s) dW_s, \\
Y_t^{n+1} &= \Big(h_0(X_T^{n+1}, \alpha_T) + \delta X_T^{n+1} - \delta X_T^n\Big) \\
&\qquad - \int_t^T \Big(g_0(s, X_s^{n+1}, Y_s^{n+1}, Z_s^{n+1}, \alpha_s) - \delta X_s^{n+1} + \delta X_s^n\Big) ds \\
&\qquad - \int_t^T Z_s^{n+1} dW_s - \int_t^T \Lambda_s^{n+1} \bullet dM_s, \qquad 0 \leqslant t \leqslant T.
\end{aligned}\right. \tag{3.2.12}$$

By Theorem 3.2, the recursive FBSDEs have unique solutions. Define for $n \geqslant 1$, $\hat{X}^n$, $\hat{Y}^n$, $\hat{Z}^n$, $\hat{\Lambda}^n$, $\hat{h}_0^n$, $\hat{f}_0^n$, $\hat{g}_0^n$, and $\hat{\sigma}_0^n$ as in the proof of Theorem 3.5. Now, for $n \geqslant 0$ and $t \leqslant T$, taking the expectations in Itô's formula, we have

$$\mathbb{E}\Big(\big\langle \hat{X}_T^{n+1}, \hat{h}_0^{n+1}(T)\big\rangle + \delta|\hat{X}_T^{n+1}|^2 - \big\langle \delta\hat{X}_T^{n+1}, \hat{X}_T^n\big\rangle\Big) + \delta\mathbb{E}\int_0^T |\hat{X}_s^{n+1}|^2 ds$$
$$- \mathbb{E}\int_0^T \Big(\big\langle \hat{Y}_s^{n+1}, \hat{f}_0^{n+1}(s)\big\rangle + \big\langle \hat{X}_s^{n+1}, \hat{g}_0^{n+1}(s)\big\rangle + \big[\hat{Z}_s^{n+1}, \hat{\sigma}_0^{n+1}(s)\big]\Big)\, ds = \delta\mathbb{E}\int_0^T \big\langle \hat{X}_s^{n+1}, \hat{X}_s^n\big\rangle ds.$$

By both assumptions of $(\mathbf{J}_0)$ and Young's inequality, we arrive at

$$\frac{\delta}{2}\mathbb{E}|\hat{X}_T^{n+1}|^2 + \frac{\delta}{2}\mathbb{E}\int_0^T |\hat{X}_s^{n+1}|^2 ds + K_\Psi\mathbb{E}\int_0^T \big(|\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\big)ds \leqslant \frac{\delta}{2}\mathbb{E}\left(\big|\hat{X}_T^n\big|^2 + \int_0^T \big|\hat{X}_s^n\big|^2 ds\right). \tag{3.2.13}$$

Now, we show that $\mathbb{E}\big[\sup_{s\leqslant T}|\hat{X}_s^n|^2\big] \leqslant C\mathbb{E}\int_0^T \big(|\hat{Y}_s^n|^2 + \|\hat{Z}_s^n\|^2\big)ds$. Since $\hat{X}^n$ is a continuous semimartingale and f_0 and σ_0 are Lipschitz functions in (x, y, z) uniformly in (t, ω) then for $t \leqslant T$,

$$\begin{aligned}
\mathbb{E}\Big[\sup_{s\leqslant t}|\hat{X}_s^n|^2\Big] &\leqslant C\mathbb{E}\left[\left(\int_0^t |\hat{f}_0^n(s)|ds\right)^2 + \int_0^t \|\hat{\sigma}_0^n(s)\|^2 ds\right] \\
&\leqslant C\mathbb{E}\left[\int_0^t |\hat{f}_0^n(s)|^2 ds + \int_0^t \|\hat{\sigma}_0^n(s)\|^2 ds\right] \\
&\leqslant C\mathbb{E}\int_0^t \big(|\hat{X}_s^n|^2 + |\hat{Y}_s^n|^2 + \|\hat{Z}_s^n\|^2\big)ds.
\end{aligned}$$

Thus, for all $t \leqslant T$,

$$\mathbb{E}\Big[\sup_{s\leqslant t}|\hat{X}_s^n|^2\Big] \leqslant C\left\{\int_0^t \mathbb{E}\Big[\sup_{u\leqslant s}|\hat{X}_u^n|^2\Big]ds + \mathbb{E}\int_0^T \big(|\hat{Y}_s^n|^2 + \|\hat{Z}_s^n\|^2\big)ds\right\}.$$

Using Gronwall's inequality, we obtain the desired inequality

$$\mathbb{E}\Big[\sup_{s\leqslant T}|\hat{X}_s^n|^2\Big] \leqslant C\mathbb{E}\int_0^T \big(|\hat{Y}_s^n|^2 + \|\hat{Z}_s^n\|^2\big)ds. \tag{3.2.14}$$

Combining (3.2.14) and (3.2.13), it follows that

$$K_\Psi\mathbb{E}\int_0^T \big(|\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\big)ds \leqslant \delta C\mathbb{E}\int_0^T \big(|\hat{Y}_s^n|^2 + \|\hat{Z}_s^n\|^2\big)ds.$$

Taking δ small enough then the sequences $(Y^n)_{n\geqslant 0}$ and $(Z^n)_{n\geqslant 0}$ are Cauchy sequences in $\mathcal{L}^2(0,T;\mathbb{R}^d)$ and $\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})$, respectively. As in the proof of previous part, $(X_T^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{L}^2(\mathbb{R}^d)$ and $(X^n)_{n\geqslant 0}$, is a Cauchy sequence in $\mathcal{L}^2(0,T;\mathbb{R}^d)$; which implies $(\Lambda^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{M}^2(0,T;\mathbb{R}^d)$. Let X, Y, Z and Λ respectively be the limit of $(X^n)_{n\geqslant 0}$, $(Y^n)_{n\geqslant 0}$, $(Z^n)_{n\geqslant 0}$ and $(\Lambda^n)_{n\geqslant 0}$ in the corresponding space and then pass the limits in equations (3.2.12) we can show that (X,Y,Z,Λ) is a solution of the system (3.2.1) of FBSDEs with Markovian switching.

(2) Uniqueness of the solution: Let (X,Y,Z,Λ) and (X',Y',Z',Λ') be two solutions of equations (3.2.1). Taking the expectations in Itô's formula for $\langle X'-X, Y'-Y\rangle$ we get

$$\begin{aligned}
&\mathbb{E}\Big\langle X'_T - X_T, h_0(X'_T,\alpha_T) - h_0(X_T,\alpha_T)\Big\rangle \\
&= \mathbb{E}\int_0^T \Big(\Big\langle X'_s - X_s, g_0(s,X'_s,Y'_s,Z'_s,\alpha_s) - g_0(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle \\
&\quad + \Big\langle Y'_s - Y_s, f_0(s,X'_s,Y'_s,Z'_s,\alpha_s) - f_0(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle \\
&\quad + \big[Z'_s - Z_s, \sigma_0(s,X'_s,Y'_s,Z'_s,\alpha_s) - \sigma_0(s,X_s,Y_s,Z_s,\alpha_s)\big]\Big)ds.
\end{aligned}$$

By both assumptions of $(\mathbf{J}_0)$, we obtain

$$K_\Psi \mathbb{E}\int_0^T \big(|Y'_s - Y_s|^2 + \|Z'_s - Z_s\|^2\big)ds \leqslant 0$$

which implies $Y' = Y$ and $Z' = Z$.

Next, using Itô's formula for $|X'-X|^2$ and then taking expectations lead to

$$\begin{aligned}
\mathbb{E}|X'_t - X_t|^2 &= 2\mathbb{E}\int_0^t \big\langle X'_s - X_s, f_0(s,X'_s,Y'_s,Z'_s,\alpha_s) - f_0(s,X_s,Y_s,Z_s,\alpha_s)\big\rangle ds \\
&\quad + \mathbb{E}\int_0^t \|\sigma_0(s,X'_s,Y'_s,Z'_s,\alpha_s) - \sigma_0(s,X_s,Y_s,Z_s,\alpha_s)\|^2 ds.
\end{aligned}$$

Since f_0 and σ_0 are Lipschitz functions in (x,y,z), $Y'=Y$ and $Z'=Z$ give

$$\mathbb{E}|X'_t - X_t|^2 \leqslant 3\mathbb{E}\int_0^t |X'_s - X_s|^2 ds.$$

Then, by Gronwall's inequality,

$$\mathbb{E}|X'_t - X_t|^2 \leqslant 0,$$

yielding $X' = X$ and $X'_T = X_T$. Moreover, taking the expectations in Itô's formula for $|Y' - Y|^2$ we get (3.2.10). Since $X'_T = X_T$, $Y' = Y$, $Z' = Z$, it follows from assumption ($C_0$2) that

$$\mathbb{E}\int_t^T |\Lambda'_s - \Lambda_s|^2 \bullet d[M]_s = \mathbb{E}|h_0(X'_T, \alpha_T) - h_0(X_T, \alpha_T)|^2 \leqslant c\mathbb{E}|X'_T - X_T|^2 = 0.$$

Thus, $\Lambda' = \Lambda$ in $\mathcal{M}^2(0, T; \mathbb{R}^d)$. $\square$

CHAPTER 4
Conditional McKean-Vlasov Forward Backward Stochastic Differential Equations with Regime Switching

1 Conditional McKean-Vlasov FBSDEs with Regime Switching

Motivated by control and game problems for large-scale systems under random environments, in this section we will consider forward-backward stochastic systems with mean-field and regime-switching where the mean-field terms are represented by random measures conditioned on the history of the Markov chain. Conditions for existence and uniqueness of solutions are obtained, which can be considered as generalizations of the results in the previous section.

Let $W_2(\cdot,\cdot)$ be the 2-Wasserstein distance, defined in (1.7.3), on $\mathcal{P}_2(\mathbb{R}^d)$ defined by Note that

$$W_2(\mu,\nu) = \inf\Big\{\big(\mathbb{E}\big|\xi-\zeta\big|^2\big)^{1/2} : \xi,\zeta \in \mathcal{L}^2(\mathbb{R}^d), \mathbb{P}_\xi = \mu, \mathbb{P}_\zeta = \nu\Big\},$$

which implies that

$$W_2(\mathbb{P}_\xi,\mathbb{P}_\zeta) \leqslant \big(\mathbb{E}\big|\xi-\zeta\big|^2\big)^{1/2}.$$

In addition, it follows from [17, Lemma 7.2] that for any sub-σ-field $\mathcal{G}$ of $\mathcal{F}$ and $r \geqslant 2$, we have

$$W_2\big(\mathbb{P}_{(\xi|\mathcal{G})},\mathbb{P}_{(\zeta|\mathcal{G})}\big) \leqslant \big[\mathbb{E}\big(\big|\xi-\zeta\big|^r\big|\mathcal{G}\big)\big]^{1/r}, \quad \mathbb{P}\text{-a.s.} \tag{4.1.1}$$

Let

$$f:[0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}\times\mathcal{P}(\mathbb{R}^{2d})\times\mathcal{M}\to\mathbb{R}^d,$$
$$g:[0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}\times\mathcal{P}(\mathbb{R}^{2d})\times\mathcal{M}\to\mathbb{R}^d,$$
$$\sigma:[0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}\times\mathcal{P}(\mathbb{R}^{2d})\times\mathcal{M}\to\mathbb{R}^{d\times d},$$
$$h:\Omega\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)\times\mathcal{M}\to\mathbb{R}^d$$

be measurable functions with respect to the Borel σ-fields.

For this section we will be working with the following conditional McKean-Vlasov FB-SDE

$$\begin{cases} X_t=\xi+\int_0^t f(s,X_s,Y_s,Z_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})},\alpha_s)ds+\int_0^t\sigma(s,X_s,Y_s,Z_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})},\alpha_s)dW_s,\\ Y_t=h(X_T,\mathbb{P}_{(X_T|\mathcal{F}^\alpha_T)},\alpha_T)-\int_t^T g(s,X_s,Y_s,Z_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})},\alpha_s)ds-\int_t^T Z_s dW_s-\int_t^T\Lambda_s\bullet dM_s,\\ t\in[0,T]. \end{cases} \tag{4.1.2}$$

A quadruplet of measurable process (X_t,Y_t,Z_t,Λ_t) is called a solution of above equation if $(X_t,Y_t,\ Z_t,\ \Lambda_t)\in\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{M}^2(0,T;\mathbb{R}^d)$ and satisfies (4.1.2).

Assumption (C).

(C1) For each fixed $i_0\in\mathcal{M}$, the functions $f(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,i_0)$, $g(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,i_0)$, and $\sigma(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,i_0)$ are $\mathcal{F}$-progressively measurable and Lipschitz in (x,y,z,μ) uniformly in (t,ω). That is for $\varphi=f,\ g$, or σ, there exists (deterministic) constants C_θ and C_ν such that for any $t\in[0,T]$, $i_0\in\mathcal{M}$, $\theta_1=(x_1,y_1,z_1),\theta_2=(x_2,y_2,z_2)\in\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}$, and $\nu_1,\nu_2\in\mathcal{P}(\mathbb{R}^{2d})$, we have

$$\left|\varphi(t,\theta_1,\nu_1,i_0)-\varphi(t,\theta_2,\nu_2,i_0)\right|\leqslant C_\theta\|\theta_1-\theta_2\|+C_\nu W_2(\nu_1,\nu_2),\quad \mathbb{P}\text{-a.s.}$$

(C2) For each fixed $(x,\mu,i_0)\in\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)\times\mathcal{M}$, the function $h(\cdot,x,\mu,i_0)$ belongs to $\mathcal{L}^2(\mathbb{R}^d)$. In addition, $h(\omega,\cdot,\cdot,i_0)$ is Lipschitz uniformly in (ω,i_0). That is, for any $i_0\in\mathcal{M}$,

$x_1, x_2 \in \mathbb{R}^d$, and $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$ there exist (deterministic) constants c and C_μ such that

$$\big|h(x_1,\mu_1,i_0) - h(x_2,\mu_2,i_0)\big| \leqslant c|x_1 - x_2| + C_\mu W_2(\mu_1,\mu_2), \quad \mathbb{P}\text{-a.s.}$$

For $t \in [0,T]$, $i_0 \in \mathcal{M}$, $\theta_1 = (x_1,y_1,z_1), \theta_2 = (x_2,y_2,z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$, and $\nu \in \mathcal{P}(\mathbb{R}^{2d})$ we denote

$$\begin{aligned}\Psi\big(t,\theta_1,\theta_2,\nu,i_0\big) = &\big\langle f\big(t,\theta_1,\nu,i_0\big) - f\big(t,\theta_2,\nu,i_0\big), y_1 - y_2\big\rangle \\ &+ \big\langle g\big(t,\theta_1,\nu,i_0\big) - g\big(t,\theta_2,\nu,i_0\big), x_1 - x_2\big\rangle \\ &+ \big[\sigma\big(t,\theta_1,\nu,i_0\big) - \sigma\big(t,\theta_2,\nu,i_0\big), z_1 - z_2\big]. \end{aligned} \tag{4.1.3}$$

Similar to Assumption $(\mathbf{H}_0)$, we consider the following assumption:

Assumption (H).

(H1) There exists a positive constant K_h such that for any $x_1, x_2 \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d)$, and $i_0 \in \mathcal{M}$, we have

$$\big\langle h\big(x_1,\mu,i_0\big) - h\big(x_2,\mu,i_0\big), x_1 - x_2\big\rangle \geqslant K_h\big|x_1 - x_2\big|^2, \quad \mathbb{P}\text{-a.s.}$$

(H2) There exists a positive constant K_Ψ such that for any $t \in [0,T]$, $\theta_1 = (x_1,y_1,z_1), \theta_2 = (x_2,y_2,z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$, $\nu \in \mathcal{P}(\mathbb{R}^{2d})$, and $i_0 \in \mathcal{M}$, we have

$$\Psi\big(t,\theta_1,\theta_2,\nu,i_0\big) \leqslant -K_\Psi\Big(\big|x_1 - x_2\big|^2 + \big|y_1 - y_2\big|^2 + \big\|z_1 - z_2\big\|^2\Big), \quad \mathbb{P}\text{-a.s.}$$

The following theorem can be viewed as and extension of Theorem 3.2. To make the presentation more transparent, its proof is aggregated in the next subsection.

Theorem 4.1. *Let assumptions* **(C)** *and* **(H)** *hold. If the constant*

$$C_\nu, C_\mu < \min\Big\{\big(\sqrt{3} - 1\big)K_h, K_\Psi/\sqrt{3}\Big\} \tag{4.1.4}$$

then there exists a unique process (X, Y, Z, Λ) *that solves the system (4.1.2) of the conditional mean-field FBSDE with Markovian switching.*

Next, similar to Theorem 3.5, we shall study how to weaken Assumption **(H)** in Theorem 4.1. Different from the setting in Theorem 3.5, the appearance of the mean-field terms in (4.1.2) makes the problem more complicated. To simplify the calculations, we assume that σ does not depend on $\mathbb{P}_{(X_s,Y_s|\mathcal{F}^{\alpha}_{s-})}$. Subsequently, the conditional mean-field FBSDE with Markovian switching (4.1.2) becomes:

$$\begin{cases} X_t = \xi + \int_0^t f(s, X_s, Y_s, Z_s, \mathbb{P}_{(X_s,Y_s|\mathcal{F}^{\alpha}_{s-})}, \alpha_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s, \alpha_s)dW_s, \\ Y_t = h(X_T, \mathbb{P}_{(X_T|\mathcal{F}^{\alpha}_T)}, \alpha_T) - \int_t^T g(s, X_s, Y_s, Z_s, \mathbb{P}_{(X_s,Y_s|\mathcal{F}^{\alpha}_{s-})}, \alpha_s)ds \\ \qquad - \int_t^T Z_s dW_s - \int_t^T \Lambda_s \bullet dM_s, \\ t \in [0,T]. \end{cases} \tag{4.1.5}$$

We make the following assumption.

Assumption L.

(L1) For any $x_1, x_2 \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d)$, and $i_0 \in \mathcal{M}$ we have

$$\left\langle h\big(x_1, \mu, i_0\big) - h\big(x_2, \mu, i_0\big), x_1 - x_2 \right\rangle \geqslant K_h \big|x_1 - x_2\big|^2, \quad \mathbb{P}\text{-a.s.}$$

(L2) There exists a positive constant K_Ψ such that for any $t \in [0,T]$, $\theta_1 = (x_1, y_1, z_1), \theta_2 = (x_2, y_2, z_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$, $\nu \in \mathcal{P}(\mathbb{R}^{2d})$, and $i_0 \in \mathcal{M}$, we have

$$\Psi\big(t, \theta_1, \theta_2, \nu, i_0\big) \leqslant -K_\Psi\Big(\big|x_1 - x_2\big|^2 + \big|y_1 - y_2\big|^2\Big), \quad \mathbb{P}\text{-a.s.}$$

Notice that although Assumption **(L)** is weaker than Assumption **(H)**, it is slightly different from Assumption **($\mathbf{I}_0$)** and Assumption **($\mathbf{J}_0$)** in the previous section. Following similar steps as in the proof of Theorem 4.1, we have the following result. Its proof is provided in the next subsection.

Theorem 4.2. *Let Assumptions* **(C)** *and* **(L)** *hold. If the constant*

$$C_\nu, C_\mu < \min\Big\{2\big(\sqrt{2} - 1\big)K_h, K_\Psi/\sqrt{2}\Big\} \tag{4.1.6}$$

then there exists a unique process (X,Y,Z,Λ), which solves the system (4.1.5) of the conditional mean-field FBSDE with Markovian switching.

2 Proofs of Main Theorems

2.1 Proof of Theorem 4.1

In this section we will prove the existence and uniqueness of solutions to conditional mean-field FBSDEs with Markovian switching (4.1.2) under Assumption (**C**) and (**H**). The proof is divided into two parts. First, we prove the existence of solutions.

(1) Existence of a solution: Let δ be a fixed positive number. We recursively define the sequence of processes $(X^n,Y^n,Z^n,\Lambda^n)_{n\geqslant 0}$ in $\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{M}^2(0,T;\mathbb{R}^d)$ as follows: $(X^0,Y^0,Z^0,\Lambda^0)\equiv(0,0,0,0)$ and, for $n\geqslant 0$, $(X^{n+1},Y^{n+1},Z^{n+1},\Lambda^{n+1})\in\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{S}^2(0,T;\mathbb{R}^d)\times\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})\times\mathcal{M}^2(0,T;\mathbb{R}^d)$ satisfies the FBSDE

$$\begin{cases} X_t^{n+1} = \xi + \displaystyle\int_0^t \Big(f(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1},\nu_s^n,\alpha_s) - \delta Y_s^{n+1} + \delta Y_s^n\Big)ds \\ \qquad\qquad + \displaystyle\int_0^t \Big(\sigma(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1},\nu_s^n,\alpha_s) - \delta Z_s^{n+1} + \delta Z_s^n\Big)dW_s, \\ Y_t^{n+1} = h(X_T^{n+1},\mu_T^n,\alpha_T) - \displaystyle\int_t^T g(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1},\nu_s^n,\alpha_s)ds \\ \qquad\qquad - \displaystyle\int_t^T Z_s^{n+1}dW_s - \int_t^T \Lambda_s^{n+1}\bullet dM_s, \qquad 0\leqslant t\leqslant T, \end{cases} \tag{4.2.1}$$

where $\nu_s^n := \mathbb{P}_{(X_s^n,Y_s^n|\mathcal{F}_{s-}^\alpha)}$ and $\mu_T^n := \mathbb{P}_{(X_T^n|\mathcal{F}_T^\alpha)}$. In view of Theorem 3.2, this FBSDE has a unique solution. For $n\geqslant 0$ and $t\in[0,T]$, denote

$$\hat{X}_t^{n+1} := X_t^{n+1} - X_t^n,\ \hat{Y}_t^{n+1} := Y_t^{n+1} - Y_t^n,\ \hat{Z}_t^{n+1} := Z_t^{n+1} - Z_t^n,\ \hat{\Lambda}_t^{n+1} := \Lambda_t^{n+1} - \Lambda_t^n, \tag{4.2.2}$$

$$\hat{h}^n(t) := h(X_t^n,\mu_t^{n-1},\alpha_t) - h(X_t^{n-1},\mu_t^{n-2},\alpha_t)$$

and for $\Theta^n = (X^n,Y^n,Z^n)$ and $\varphi = f,g,\sigma$,

$$\begin{aligned} \hat{\varphi}^{n+1}(t) &:= \varphi(t,\Theta_t^{n+1},\nu_t^n,\alpha_t) - \varphi(t,\Theta_t^n,\nu_t^{n-1},\alpha_t), \\ \bar{\varphi}^n(t) &:= \varphi(t,\Theta_t^n,\nu_t^n,\alpha_t) - \varphi(t,\Theta_t^n,\nu_t^{n-1},\alpha_t). \end{aligned} \tag{4.2.3}$$

It follows from (4.1.1) that

$$W_2^2(\nu_t^n,\nu_t^{n-1}) \leqslant \mathbb{E}\big(|\hat{X}_t^n|^2+|\hat{Y}_t^n|^2\big|\mathcal{F}_{t-}^{\alpha}\big). \tag{4.2.4}$$

According to Itô's formula, we obtain

$$\begin{aligned}
&\big\langle \hat{X}_T^{n+1},\hat{Y}_T^{n+1}\big\rangle-\big\langle \hat{X}_0^{n+1},\hat{Y}_0^{n+1}\big\rangle\\
&=\int_0^T\big\langle \hat{Y}_s^{n+1},\hat{f}^{n+1}(s)-\delta\hat{Y}_s^{n+1}+\delta\hat{Y}_s^n\big\rangle ds+\int_0^T\big\langle \hat{Y}_s^{n+1},\big(\hat{\sigma}^{n+1}(s)-\delta\hat{Z}_s^{n+1}+\delta\hat{Z}_s^n\big)dW_s\big\rangle\\
&\quad+\int_0^T\big\langle \hat{X}_s^{n+1},\hat{g}^{n+1}(s)\big\rangle ds+\int_0^T\big\langle \hat{X}_s^{n+1},\hat{Z}_s^{n+1}dW_s\big\rangle+\int_0^T\big\langle \hat{X}_s^{n+1},\hat{\Lambda}_s^{n+1}\bullet dM_s\big\rangle\\
&\quad+\int_0^T\big[\hat{Z}_s^{n+1},\hat{\sigma}^{n+1}(s)-\delta\hat{Z}_s^{n+1}+\delta\hat{Z}_s^n\big]ds.
\end{aligned} \tag{4.2.5}$$

Applying standard estimates of BSDEs and the Burkholder-Davis-Gundy inequality, it is easy to see that the stochastic integrals in (4.2.5) are true martingales. After taking expectations, we obtain

$$\begin{aligned}
&\mathbb{E}\big\langle \hat{X}_T^{n+1},h(X_T^{n+1},\mu_T^n,\alpha_T)-h(X_T^n,\mu_T^{n-1},\alpha_T)\big\rangle+\delta\mathbb{E}\int_0^T\big(|\hat{Y}_s^{n+1}|^2+\|\hat{Z}_s^{n+1}\|^2\big)ds\\
&=\delta\mathbb{E}\int_0^T\Big(\big\langle \hat{Y}_s^{n+1},\hat{Y}_s^n\big\rangle+\big[\hat{Z}_s^{n+1},\hat{Z}_s^n\big]\Big)ds\\
&\qquad+\mathbb{E}\int_0^T\Big(\big\langle \hat{Y}_s^{n+1},\hat{f}^{n+1}(s)\big\rangle+\big\langle \hat{X}_s^{n+1},\hat{g}^{n+1}(s)\big\rangle+\big[\hat{Z}_s^{n+1},\hat{\sigma}^{n+1}(s)\big]\Big)ds.
\end{aligned} \tag{4.2.6}$$

Using the Lipschitz continuity of h (assumption (C2)), Young's inequality, (4.1.1) and (H1), we have

$$
\begin{aligned}
&\mathbb{E}\Big\langle \hat{X}_T^{n+1}, h(X_T^{n+1},\mu_T^n,\alpha_T) - h(X_T^n,\mu_T^{n-1},\alpha_T)\Big\rangle \\
&= \mathbb{E}\Big\langle \hat{X}_T^{n+1}, h(X_T^{n+1},\mu_T^n,\alpha_T) - h(X_T^n,\mu_T^n,\alpha_T)\Big\rangle + \mathbb{E}\Big\langle \hat{X}_T^{n+1}, h(X_T^n,\mu_T^n,\alpha_T) - h(X_T^n,\mu_T^{n-1},\alpha_T)\Big\rangle \\
&\geqslant K_h\mathbb{E}|\hat{X}_T^{n+1}|^2 - C_\mu\mathbb{E}\Big[\big|\hat{X}_T^{n+1}\big|\ W_2(\mu_T^n,\mu_T^{n-1})\Big] \\
&\geqslant K_h\mathbb{E}|\hat{X}_T^{n+1}|^2 - \frac{C_\mu\epsilon}{2}\mathbb{E}|\hat{X}_T^{n+1}|^2 - \frac{C_\mu}{2\epsilon}\mathbb{E}W_2^2(\mu_T^n,\mu_T^{n-1}) \\
&\geqslant \left(K_h - \frac{C_\mu\epsilon}{2}\right)\mathbb{E}|\hat{X}_T^{n+1}|^2 - \frac{C_\mu}{2\epsilon}\mathbb{E}\Big[\mathbb{E}\big(|\hat{X}_T^n|^2|\mathcal{F}_T^\alpha\big)\Big] \\
&= \left(K_h - \frac{C_\mu\epsilon}{2}\right)\mathbb{E}|\hat{X}_T^{n+1}|^2 - \frac{C_\mu}{2\epsilon}\mathbb{E}|\hat{X}_T^n|^2, \quad \forall\, \epsilon > 0. \qquad (4.2.7)
\end{aligned}
$$

Again, by the Lipschitz continuity of $f,\ h,\ \sigma$, Young's inequality, (4.1.1) and (H2), we also have

$$
\begin{aligned}
&\Big\langle \hat{X}_t^{n+1}, \hat{g}^{n+1}(t)\Big\rangle + \Big\langle \hat{Y}_t^{n+1}, \hat{f}^{n+1}(t)\Big\rangle + \big[\hat{Z}_t^{n+1}, \hat{\sigma}^{n+1}(t)\big] \\
&= \Psi\big(t,\Theta_t^{n+1},\Theta_t^n,\nu_t^n,\alpha_t\big) + \Big\langle \hat{X}_t^{n+1}, \bar{g}^n(t)\Big\rangle + \Big\langle \hat{Y}_t^{n+1}, \bar{f}^n(t)\Big\rangle + \big[\hat{Z}_t^{n+1}, \bar{\sigma}^n(t)\big] \\
&\leqslant -K_\Psi\Big(|\hat{X}_t^{n+1}|^2 + |\hat{Y}_t^{n+1}|^2 + \|\hat{Z}_t^{n+1}\|^2\Big) + |\hat{X}_t^{n+1}||\bar{g}^n(t)| + |\hat{Y}_t^{n+1}||\bar{f}^n(t)| + \|\hat{Z}_t^{n+1}\|\|\bar{\sigma}^n(t)\| \\
&\leqslant -K_\Psi\Big(|\hat{X}_t^{n+1}|^2 + |\hat{Y}_t^{n+1}|^2 + \|\hat{Z}_t^{n+1}\|^2\Big) + C_\nu W_2\big(\nu_t^n,\nu_t^{n-1}\big)\Big(|\hat{X}_t^{n+1}| + |\hat{Y}_t^{n+1}| + \|\hat{Z}_t^{n+1}\|\Big) \\
&\leqslant \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\Big(|\hat{X}_t^{n+1}|^2 + |\hat{Y}_t^{n+1}|^2 + \|\hat{Z}_t^{n+1}\|^2\Big) + \frac{3\gamma C_\nu}{2}W_2^2\big(\nu_t^n,\nu_t^{n-1}\big), \quad \forall\, t\in[0,T], \gamma>0.
\end{aligned}
$$

Now, it follows from (4.2.4) that

$$
\begin{aligned}
&\mathbb{E}\int_0^T \Big(\Big\langle \hat{X}_s^{n+1}, \hat{g}^{n+1}(s)\Big\rangle + \Big\langle \hat{Y}_s^{n+1}, \hat{f}^{n+1}(s)\Big\rangle + \big[\hat{Z}_s^{n+1}, \hat{\sigma}^{n+1}(s)\big]\Big)ds \\
&\leqslant \mathbb{E}\int_0^T \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\Big(\big|\hat{X}_s^{n+1}\big|^2 + \big|\hat{Y}_s^{n+1}\big|^2 + \|\hat{Z}_s^{n+1}\|^2\Big)ds + \frac{3\gamma C_\nu}{2}\int_0^T \mathbb{E}\big(\big|\hat{X}_s^n\big|^2 + \big|\hat{Y}_s^n\big|^2\big)ds. \qquad (4.2.8)
\end{aligned}
$$

On the other hand, in view of Young's inequality, we also have for any $\rho > 0$,

$$
\mathbb{E}\int_0^T \Big(\Big\langle \hat{Y}_s^{n+1}, \hat{Y}_s^n\Big\rangle + \big[\hat{Z}_s^{n+1}, \hat{Z}_s^n\big]\Big)ds \leqslant \frac{1}{2}\mathbb{E}\int_0^T \Big(\rho|\hat{Y}_s^{n+1}|^2 + \rho\|\hat{Z}_s^{n+1}\|^2 + \frac{1}{\rho}|\hat{Y}_s^n|^2 + \frac{1}{\rho}\|\hat{Z}_s^n\|^2\Big)ds. \qquad (4.2.9)
$$

Combining (4.2.7), (4.2.8), and (4.2.9), to (4.2.6), yields

$$\begin{aligned}
&\Big(K_h - \frac{C_\mu \epsilon}{2}\Big)\mathbb{E}\big|\hat{X}_T^{n+1}\big|^2 - \frac{C_\mu}{2\epsilon}\mathbb{E}|\hat{X}_T^n|^2 + \delta\mathbb{E}\int_0^T \Big(|\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\Big)ds \\
&\qquad - \mathbb{E}\int_0^T \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\Big(|\hat{X}_s^{n+1}|^2 + |\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\Big)ds \\
&\leqslant \frac{3\gamma C_\nu}{2}\int_0^T \mathbb{E}\left(\big|\hat{X}_s^n\big|^2 + \big|\hat{Y}_s^n\big|^2\right)ds + \delta\mathbb{E}\int_0^T \Big(\frac{\rho}{2}|\hat{Y}_s^{n+1}|^2 + \frac{\rho}{2}\|\hat{Z}_s^{n+1}\|^2 + \frac{1}{2\rho}|\hat{Y}_s^n|^2 + \frac{1}{2\rho}\|\hat{Z}_s^n\|^2\Big)ds.
\end{aligned}$$

Rearranging the terms, we obtain

$$\begin{aligned}
&\Big(K_h - \frac{C_\mu\epsilon}{2}\Big)\mathbb{E}|\hat{X}_T^{n+1}|^2 + \Big(K_\Psi - \frac{C_\nu}{2\gamma}\Big)\mathbb{E}\int_0^T |\hat{X}_s^{n+1}|^2 ds \\
&\qquad + \Big(\delta\big(1 - \frac{\rho}{2}\big) + K_\Psi - \frac{C_\nu}{2\gamma}\Big)\mathbb{E}\int_0^T \Big(|\hat{Y}_s^{n+1}|^2 + \|\hat{Z}_s^{n+1}\|^2\Big)ds \\
&\leqslant \frac{C_\mu}{2\epsilon}\mathbb{E}|\hat{X}_T^n|^2 + \mathbb{E}\int_0^T \left[\frac{3\gamma C_\nu}{2}|\hat{X}_s^n|^2 + \Big(\frac{3\gamma C_\nu}{2} + \frac{\delta}{2\rho}\Big)|\hat{Y}_s^n|^2 + \frac{\delta}{2\rho}\|\hat{Z}_s^n\|^2\right] ds.
\end{aligned} \tag{4.2.10}$$

Denote $\|\Theta\|^2 = |X|^2 + |Y|^2 + \|Z\|^2$ and

$$\begin{aligned}
\lambda(\epsilon,\delta,\gamma,\rho) &\triangleq \min\left\{K_h - \frac{C_\mu\epsilon}{2}, K_\Psi - \frac{C_\nu}{2\gamma}, \delta\left(1 - \frac{\rho}{2}\right) + K_\Psi - \frac{C_\nu}{2\gamma}\right\}, \\
\theta(\epsilon,\delta,\gamma,\rho) &\triangleq \max\left\{\frac{C_\mu}{2\epsilon}, \frac{3\gamma C_\nu}{2} + \frac{\delta}{2\rho}\right\},
\end{aligned}$$

we obtain

$$\mathbb{E}\left(|\hat{X}_T^{n+1}|^2 + \int_0^T \|\hat{\Theta}_s^{n+1}\|^2 ds\right) \leqslant \frac{\theta(\epsilon,\delta,\gamma,\rho)}{\lambda(\epsilon,\delta,\gamma,\rho)}\mathbb{E}\left(|\hat{X}_T^n|^2 + \int_0^T \|\hat{\Theta}_s^n\|^2 ds\right). \tag{4.2.11}$$

To proceed, we temporarily assume there exist constants ϵ, δ, γ, and ρ such that

$$\lambda(\epsilon,\delta,\gamma,\rho) > \theta(\epsilon,\delta,\gamma,\rho). \tag{4.2.12}$$

Then the inequality (4.2.11) becomes a contraction, which subsequently implies that $(X_T^n)_{n\geqslant 0}$ is a Cauchy sequence in $L^2(\Omega,\mathbb{P})$ and $(X^n)_{n\geqslant 0}, (Y^n)_{n\geqslant 0}$ and $(Z^n)_{n\geqslant 0}$ are Cauchy sequences in $L^2([0,T]\times\Omega, dt\otimes d\mathbb{P})$. Taking the expectation of Itô's formula, we have

$$\mathbb{E}|\hat{Y}_t^n|^2 + \mathbb{E}\int_t^T \|\hat{Z}_s^n\|^2 ds + \mathbb{E}\int_t^T |\hat{\Lambda}_s^n|^2 \bullet d[M]_s = \mathbb{E}|\hat{h}^n(T)|^2 - 2\mathbb{E}\int_t^T \big\langle \hat{Y}_s^n, \hat{g}^n(s)\big\rangle ds \to 0 \tag{4.2.13}$$

since $(X^n_T)_{n\geqslant 0}$, $(X^n)_{n\geqslant 0}$, $(Y^n)_{n\geqslant 0}$ and $(Z^n)_{n\geqslant 0}$ are Cauchy. Thus, $(\Lambda^n)_{n\geqslant 0}$ is also a Cauchy sequence. This yields

$$\mathbb{E}\left[\sup_{s\leqslant T}\left(|X^n_s - X^m_s|^2 + |Y^n_s - Y^m_s|^2\right)\right] \to 0 \text{ as } n,\ m\to\infty.$$

By Banach fixed point theorem, there exist $\mathbb{F}$-adapted càdlàg processes X and Y, an $\mathcal{F}$-progressively measurable process Z, and a collection of $\mathcal{F}$-progressively measurable functions Λ such that

$$\mathbb{E}\left[\sup_{s\leqslant T}\left(|X^n_s - X_s|^2 + |Y^n_s - Y_s|^2\right) + \int_0^T \|Z^n_s - Z_s\|^2 ds + \int_0^T \left|\Lambda^n_s - \Lambda_s\right|^2 \bullet d[M]_s\right] \to 0 \text{ as } n\to\infty$$

and

$$\mathbb{E}\left[\sup_{s\leqslant T}\left(|X_s|^2 + |Y_s|^2\right) + \int_0^T \|Z_s\|^2 ds + \int_0^T \left|\Lambda_s\right|^2 \bullet d[M]_s\right] < \infty.$$

Moreover, taking the limits in equation (4.2.1) we obtain that (X, Y, Z, Λ) is a solution of (4.1.2).

To complete the proof, we are to show that such γ, ϵ, δ and ρ exist when the condition (4.1.4) is satisfied. In fact, to make the contraction meaningful, we assume $K_h - \frac{C_\mu \epsilon}{2}$, $K_\Psi - \frac{C_\nu}{2\gamma}$, and $1-\frac{\rho}{2}$ are positive. Since $(1-\frac{\rho}{2}) \leqslant \frac{1}{2\rho}$ with equality obtained if and only if $\rho = 1$, we will take $\rho = 1$ and set

$$\theta^*(\epsilon,\gamma) = \lim_{\delta\to 0}\theta(\epsilon,\delta,\gamma,1) = \max\left\{\frac{C_\mu}{2\epsilon}, \frac{3\gamma C_\nu}{2}\right\}$$

and

$$\lambda^*(\epsilon,\gamma) = \lim_{\delta\to 0}\lambda(\epsilon,\delta,\gamma,1) = \min\left\{K_h - \frac{C_\mu\epsilon}{2}, K_\Psi - \frac{C_\nu}{2\gamma}\right\}.$$

If we have $\lambda^*(\epsilon,\gamma) > \theta^*(\epsilon,\gamma)$ for some ϵ and γ, then there exists δ small enough such (4.2.12) is satisfied. Note that $\lambda^*(\epsilon,\gamma) > \theta^*(\epsilon,\gamma)$, equivalent to having the following inequalities:

$$K_h > \max\left\{\frac{C_\mu}{2\epsilon} + \frac{C_\mu\epsilon}{2}, \frac{3\gamma C_\nu}{2} + \frac{C_\mu\epsilon}{2}\right\}, \quad K_\Psi > \max\left\{\frac{C_\mu}{2\epsilon} + \frac{C_\nu}{2\gamma}, \frac{3\gamma C_\nu}{2} + \frac{C_\nu}{2\gamma}\right\}. \tag{4.2.14}$$

To minimize $\frac{C_\mu\epsilon}{2} + \frac{C_\mu}{2\epsilon}$ and $\frac{C_\nu}{2\gamma} + \frac{3\gamma C_\nu}{2}$ we take $\epsilon = 1$ and $\gamma = \frac{1}{\sqrt{3}}$. Let $\eta_1, \eta_2 > 0$ such that $C_\nu, C_\mu < \min\{\eta_1 K_h, \eta_2 K_\Psi\}$. Without loss of the generality, we can assume $\eta_1 K_h \leqslant \eta_2 K_\Psi$.

Then (4.2.14) holds if following inequalities hold

$$K_h > C_\mu, \quad K_h > \frac{\eta_1 K_h}{2}(\sqrt{3}+1), \quad K_\Psi > \frac{\eta_1 K_h}{2}(\sqrt{3}+1), \quad K_\Psi > \sqrt{3}C_\nu. \tag{4.2.15}$$

From the second inequality, we obtain $\eta_1 < \sqrt{3}-1$. The third inequality in (4.2.15) holds if the next inequality

$$\frac{\eta_1 K_h}{2}(\sqrt{3}+1) < \frac{\eta_1 K_h}{\eta_2}$$

holds. This implies $\eta_2 < \sqrt{3}-1$. From the fourth inequality in (4.2.15) we have $C_\nu < \frac{\sqrt{3}}{3}K_\Psi$. Note that $\frac{\sqrt{3}}{3} < \sqrt{3}-1$. Combining these we obtain the sufficient condition $C_\nu, C_\mu < \min\left\{(\sqrt{3}-1)K_h, \frac{\sqrt{3}}{3}K_\Psi\right\}$ for which $\lambda(\epsilon,\delta,\alpha,\rho) > \theta(\epsilon,\delta,\alpha,\rho)$ when $\epsilon = \rho = 1$, $\gamma = \frac{\sqrt{3}}{3}$ and $\delta > 0$ small enough.

(2) Uniqueness of the solution: Let (X', Y', Z', Λ') be another solution to (4.1.2). Let $\Theta = (X, Y, Z)$ and $\Theta' = (X', Y', Z')$. Applying Itô's formula to the product $\langle X' - X, Y' - Y\rangle$ and taking expectation, we obtain

$$\mathbb{E}\Big\langle X'_T - X_T, Y'_T - Y_T\Big\rangle = \Gamma_T, \tag{4.2.16}$$

where

$$\begin{aligned}\Gamma_T \triangleq \mathbb{E}\int_0^T \Bigg(&\Big\langle X'_s - X_s, g\big(s, \Theta'_s, \mathbb{P}_{(X'_s, Y'_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big) - g\big(s, \Theta_s, \mathbb{P}_{(X_s, Y_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big)\Big\rangle \\ &+ \Big\langle Y'_s - Y_s, f\big(s, \Theta'_s, \mathbb{P}_{(X'_s, Y'_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big) - f\big(s, \Theta_s, \mathbb{P}_{(X_s, Y_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big)\Big\rangle \\ &+ \big[\sigma\big(s, \Theta'_s, \mathbb{P}_{(X'_s, Y'_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big) - \sigma\big(s, \Theta_s, \mathbb{P}_{(X_s, Y_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big), Z'_s - Z_s\big]\Bigg) ds.\end{aligned}$$

Using Assumption (C), Assumption (H1), (4.1.1), and Hölder's inequality, we can obtain the following lower bound for Γ_T.

$$
\begin{aligned}
\Gamma_T &= \mathbb{E}\Big\langle X'_T - X_T, h\big(X'_T, \mathbb{P}_{(X'_T|\mathcal{F}^\alpha_T)}, \alpha_T\big) - h\big(X_T, \mathbb{P}_{(X_T|\mathcal{F}^\alpha_T)}, \alpha_T\big)\Big\rangle \\
&\geqslant K_h \mathbb{E}|X'_T - X_T|^2 - C_\mu \mathbb{E}\Big(|X'_T - X_T| \; W_2\big(\mathbb{P}_{(X'_T|\mathcal{F}^\alpha_T)}, \mathbb{P}_{(X_T|\mathcal{F}^\alpha_T)}\big)\Big) \\
&\geqslant K_h \mathbb{E}|X'_T - X_T|^2 - C_\mu \mathbb{E}\left\{|X'_T - X_T|\Big[\mathbb{E}\big(|X'_T - X_T|^2\big|\mathcal{F}^\alpha_T\big)\Big]^{\frac{1}{2}}\right\} \\
&= K_h \mathbb{E}|X'_T - X_T|^2 - C_\mu \mathbb{E}\left\{\mathbb{E}\big(|X'_T - X_T|\big|\mathcal{F}^\alpha_T\big)\Big[\mathbb{E}\big(|X'_T - X_T|^2\big|\mathcal{F}^\alpha_T\big)\Big]^{\frac{1}{2}}\right\} \\
&\geqslant (K_h - C_\mu)\mathbb{E}|X'_T - X_T|^2.
\end{aligned}
\tag{4.2.17}
$$

On the other hand, we can also have an upper bound for Γ_T as follow. First, by the triangle inequality,

$$
\begin{aligned}
\Gamma_T \leqslant \mathbb{E}\int_0^T \Big[&\Psi\big(s, \Theta_s, \Theta'_s, \mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})}, \alpha_s\big) \\
&+ C_\nu W_2\big(\mathbb{P}_{(X'_s,Y'_s|\mathcal{F}^\alpha_{s-})}, \mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})}\big)\Big(|X'_s - X_s| + |Y'_s - Y_s| + \|Z'_s - Z_s\|\Big)\Big] ds.
\end{aligned}
$$

Then, using the estimate (4.1.1),

$$
W_2\big(\mathbb{P}_{(X'_s,Y'_s|\mathcal{F}^\alpha_{s-})}, \mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})}\big) \leqslant \sqrt{\mathbb{E}\big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2\big|\mathcal{F}^\alpha_{s-}\big)} \tag{4.2.18}
$$

together with (H2) and the Cauchy-Schwarz inequality three times yields

$$
\begin{aligned}
\Gamma_T \leqslant \mathbb{E}\int_0^T \Big[&-K_\Psi\Big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2 + \|Z'_s - Z_s\|^2\Big) \\
&+ C_\nu\sqrt{\mathbb{E}\big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2\big|\mathcal{F}^\alpha_{s-}\big)}\Big(|X'_s - X_s| + |Y'_s - Y_s| + \|Z'_s - Z_s\|\Big)\Big] ds \\
\leqslant \mathbb{E}\int_0^T \Big[&-K_\Psi\Big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2 + \|Z'_s - Z_s\|^2\Big) \\
&+ \frac{C_\nu}{2}\Big(\frac{1}{\gamma} + 3\gamma\Big)|X'_s - X_s|^2 + \frac{C_\nu}{2}\Big(\frac{1}{\gamma} + 3\gamma\Big)|Y'_s - Y_s|^2 + \frac{C_\nu}{2\gamma}\|Z'_s - Z_s\|^2\Big] ds.
\end{aligned}
\tag{4.2.19}
$$

Combining (4.2.17) and (4.2.19), we obtain

$$0 \leqslant (C_\mu - K_h)\mathbb{E}|X'_T - X_T|^2 + \mathbb{E}\int_0^T \Big\{\Big[\frac{C_\nu}{2}\Big(\frac{1}{\gamma}+3\gamma\Big) - K_\Psi\Big]|X'_s - X_s|^2$$
$$+\Big[\frac{C_\nu}{2}\Big(\frac{1}{\gamma}+3\gamma\Big) - K_\Psi\Big]|Y'_s - Y_s|^2 + \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\|Z'_s - Z_s\|^2\Big\}ds.$$

Noting that if $C_\mu, C_\nu < \min\Big\{\big(\sqrt{3}-1\big)K_h, K_\Psi/\sqrt{3}\Big\}$, with $\gamma = \frac{1}{\sqrt{3}}$, then all the coefficients of the right-hand side of the above inequality are negative. This implies that $X'_T = X_T$ $\mathbb{P}$-a.s. and for all $0\leqslant s\leqslant T$, $X'_s = X_s$, $Y'_s = Y_s$, and $Z'_s = Z_s$ $\mathbb{P}$-a.s. Next, we take the expectation of Itô's formula for $|Y' - Y|^2$,

$$\mathbb{E}|Y'_0 - Y_0|^2 + \mathbb{E}\int_0^T \|Z'_s - Z_s\|^2 ds + \mathbb{E}\int_0^T |\Lambda'_s - \Lambda_s|^2 \bullet d[M]_s$$
$$= \mathbb{E}|h(X'_T,\alpha_T) - h(X_T,\alpha_T)|^2 - 2\mathbb{E}\int_0^T \Big\langle Y'_s - Y_s, g(s,X'_s,Y'_s,Z'_s,\alpha_s) - g(s,X_s,Y_s,Z_s,\alpha_s)\Big\rangle ds.$$

Since $X'_T = X_T$ $\mathbb{P}$-a.s. and, for all $0\leqslant s\leqslant T$, $X'_s = X_s$, $Y'_s = Y_s$, and $Z'_s = Z_s$ $\mathbb{P}$-a.s., we get

$$\mathbb{E}\int_0^T |\Lambda'_s - \Lambda_s|^2 \bullet d[M]_s = 0.$$

Thus, $\Lambda'_s = \Lambda_s$ in $\mathcal{M}^2(0,T;\mathbb{R}^d)$, yielding that the solution of (4.1.2) is unique. □

2.2 Proof of Theorem 4.2.

(1) Existence of a solution: We will follow the same approach as in Theorem 4.1. Let $\delta > 0$ and consider the sequence $(X^n, Y^n, Z^n, \Lambda^n)_{n\geqslant 0}$ defined recursively as follows: For $n = 0$, put $(X^0, Y^0, Z^0, \Lambda^0) = (0,0,0,0)$. For $n \geqslant 0$, let $(X^{n+1}, Y^{n+1}, Z^{n+1}, \Lambda^{n+1})$ be the stochastic process in $\mathcal{S}^2(0,T;\mathbb{R}^d) \times \mathcal{S}^2(0,T;\mathbb{R}^d) \times \mathcal{L}^2(0,T;\mathbb{R}^{d\times d}) \times \mathcal{M}^2(0,T;\mathbb{R}^d)$ which is the unique solution of the following FBSDE

$$\begin{cases} X_t^{n+1} &= \xi + \displaystyle\int_0^t \Big(f(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1},\nu_s^n,\alpha_s) - \delta Y_s^{n+1} + \delta Y_s^n\Big)ds \\ & \quad + \displaystyle\int_0^t \sigma(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1},\alpha_s)dW_s, \\ Y_t^{n+1} &= h(X_T^{n+1},\mu_T^n,\alpha_T) - \displaystyle\int_t^T g(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1},\nu_s^n,\alpha_s)ds \\ & \quad - \displaystyle\int_t^T Z_s^{n+1}dW_s - \int_t^T \Lambda_s^{n+1}\bullet dM_s, \quad 0\leqslant t\leqslant T, \end{cases} \tag{4.2.20}$$

where $\nu_s^n := \mathbb{P}_{(X_s^n, Y_s^n | \mathcal{F}_{s-}^\alpha)}$ and $\mu_T^n := \mathbb{P}_{(X_T^n | \mathcal{F}_T^\alpha)}$. By virtue of Theorem 3.2 (under Assumptions $(\mathbf{C}_0)$ and $(\mathbf{H}_0)$), we know that this recursive FBSDE has a unique solution. For $n \geqslant 1$, $t \in [0, T]$, we define $\hat{X}_t^{n+1}$, $\hat{Y}_t^{n+1}$, $\hat{Z}_t^{n+1}$, $\hat{\Lambda}_t^{n+1}$, $\hat{h}^n(t)$, $\hat{\varphi}^{n+1}(t)$, $\bar{\varphi}^n(t)$ as in (4.2.2) and (4.2.3). As the diffusion coefficient σ is assumed to be independent of the conditional measure, (4.2.3) implies $\bar{\sigma}^n(t) \equiv 0$.

Using Itô's formula and then taking the expectations, we obtain

$$
\begin{aligned}
&\mathbb{E}\Big\langle \hat{X}_T^{n+1}, h(X_T^{n+1}, \mu_T^n, \alpha_T) - h(X_T^n, \mu_T^{n-1}, \alpha_T)\Big\rangle + \delta\mathbb{E}\int_0^T |\hat{Y}_s^{n+1}|^2 ds \\
&= \delta\mathbb{E}\int_0^T \big\langle \hat{Y}_s^{n+1}, \hat{Y}_s^n \big\rangle ds + \mathbb{E}\int_0^T \Big(\big\langle \hat{Y}_s^{n+1}, \hat{f}^{n+1}(s) \big\rangle + \big\langle \hat{X}_s^{n+1}, \hat{g}^{n+1}(s) \big\rangle + \big[\hat{Z}_s^{n+1}, \hat{\sigma}^{n+1}(s)\big]\Big) ds.
\end{aligned}
\tag{4.2.21}
$$

Similar to (4.2.7), for any $\epsilon > 0$, we have

$$
\mathbb{E}\Big\langle \hat{X}_T^{n+1}, h(X_T^{n+1}, \mu_T^n, \alpha_T) - h(X_T^n, \mu_T^{n-1}, \alpha_T)\Big\rangle \geqslant \Big(K_h - \frac{C_\mu \epsilon}{2}\Big)\mathbb{E}|\hat{X}_T^{n+1}|^2 - \frac{C_\mu}{2\epsilon}\mathbb{E}|\hat{X}_T^n|^2.
\tag{4.2.22}
$$

By the Lipschitz continuity of f, h, σ (assumption (C1)), Young's inequality, and using $\bar{\sigma}^n(t) \equiv 0$ and applying assumption (L2) instead of (H2), for every $0 \leqslant t \leqslant T$ and any $\gamma > 0$, we have

$$
\begin{aligned}
&\big\langle \hat{X}_t^{n+1}, \hat{g}^{n+1}(t) \big\rangle + \big\langle \hat{Y}_t^{n+1}, \hat{f}^{n+1}(t) \big\rangle + \big[\hat{Z}_t^{n+1}, \hat{\sigma}^{n+1}(t)\big] \\
&= \Psi\big(t, \Theta_t^{n+1}, \Theta_t^n, \nu_t^n, \alpha_t\big) + \big\langle \hat{X}_t^{n+1}, \bar{g}^n(t) \big\rangle + \big\langle \hat{Y}_t^{n+1}, \bar{f}^n(t) \big\rangle \\
&\leqslant -K_\Psi\Big(\big|\hat{X}_t^{n+1}\big|^2 + \big|\hat{Y}_t^{n+1}\big|^2\Big) + \big|\hat{X}_t^{n+1}\big|\big|\bar{g}^n(t)\big| + \big|\hat{Y}_t^{n+1}\big|\big|\bar{f}^n(t)\big| \\
&\leqslant -K_\Psi\Big(\big|\hat{X}_t^{n+1}\big|^2 + \big|\hat{Y}_t^{n+1}\big|^2\Big) + C_\nu W_2\big(\nu_t^n, \nu_t^{n-1}\big)\Big(\big|\hat{X}_t^{n+1}\big| + \big|\hat{Y}_t^{n+1}\big|\Big) \\
&\leqslant \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\Big(\big|\hat{X}_t^{n+1}\big|^2 + \big|\hat{Y}_t^{n+1}\big|^2\Big) + \gamma C_\nu W_2^2\big(\nu_t^n, \nu_t^{n-1}\big).
\end{aligned}
$$

Using (4.2.4) leads to the following inequality which is similar to (4.2.8)

$$
\begin{aligned}
&\mathbb{E}\int_0^T \Big(\big\langle \hat{X}_s^{n+1}, \hat{g}^{n+1}(s) \big\rangle + \big\langle \hat{Y}_s^{n+1}, \hat{f}^{n+1}(s) \big\rangle + \big[\hat{Z}_s^{n+1}, \hat{\sigma}^{n+1}(s)\big]\Big) ds \\
&\leqslant \mathbb{E}\int_0^T \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\Big(\big|\hat{X}_s^{n+1}\big|^2 + \big|\hat{Y}_s^{n+1}\big|^2\Big) ds + \gamma C_\nu \mathbb{E}\int_0^T \Big(\big|\hat{X}_s^n\big|^2 + \big|\hat{Y}_s^n\big|^2\Big) ds.
\end{aligned}
\tag{4.2.23}
$$

On the other hand, in view of Young's inequality, we also have

$$\delta\mathbb{E}\int_0^T \big\langle \hat{Y}_s^{n+1}, \hat{Y}_s^n \big\rangle ds \leqslant \frac{\delta}{2}\mathbb{E}\int_0^T \left(\big|\hat{Y}_s^{n+1}\big|^2 + \big|\hat{Y}_s^n\big|^2 \right) ds. \tag{4.2.24}$$

Applying now (4.2.22), (4.2.23), and (4.2.24), to (4.2.21), yields

$$\begin{aligned}
&\Big(K_h - \frac{C_\mu \epsilon}{2}\Big)\mathbb{E}\big|\hat{X}_T^{n+1}\big|^2 - \frac{C_\mu}{2\epsilon}\mathbb{E}|\hat{X}_T^n|^2 + \delta\mathbb{E}\int_0^T \big|\hat{Y}_s^{n+1}\big|^2 ds \\
&\leqslant \Big(\frac{C_\nu}{2\gamma} - K_\Psi\Big)\mathbb{E}\int_0^T \Big(\big|\hat{X}_s^{n+1}\big|^2 + \big|\hat{Y}_s^{n+1}\big|^2\Big) ds \\
&\quad + \gamma C_\nu \mathbb{E}\int_0^T \Big(\big|\hat{X}_s^n\big|^2 + \big|\hat{Y}_s^n\big|^2\Big) ds + \frac{\delta}{2}\mathbb{E}\int_0^T \Big(\big|\hat{Y}_s^{n+1}\big|^2 + \big|\hat{Y}_s^n\big|^2\Big) ds
\end{aligned}$$

Rearranging terms, we obtain

$$\begin{aligned}
&\Big(K_h - \frac{C_\mu \epsilon}{2}\Big)\mathbb{E}\big|\hat{X}_T^{n+1}\big|^2 + \Big(K_\Psi - \frac{C_\nu}{2\gamma}\Big)\mathbb{E}\int_0^T \big|\hat{X}_s^{n+1}\big|^2 ds + \Big(\frac{\delta}{2} + K_\Psi - \frac{C_\nu}{2\gamma}\Big)\mathbb{E}\int_0^T \big|\hat{Y}_s^{n+1}\big|^2 ds \\
&\leqslant \frac{C_\mu}{2\epsilon}\mathbb{E}\big|\hat{X}_T^n\big|^2 + \gamma C_\nu \mathbb{E}\int_0^T \big|\hat{X}_s^n\big|^2 ds + \Big(\gamma C_\nu + \frac{\delta}{2}\Big)\mathbb{E}\int_0^T \big|\hat{Y}_s^n\big|^2 ds.
\end{aligned} \tag{4.2.25}$$

Define

$$\lambda(\epsilon,\delta,\gamma) \triangleq \min\left\{K_h - \frac{C_\mu \epsilon}{2},\ K_\Psi - \frac{C_\nu}{2\gamma}\right\}, \quad \theta(\epsilon,\delta,\gamma) \triangleq \max\left\{\frac{C_\mu}{2\epsilon},\ \gamma C_\nu + \frac{\delta}{2}\right\}.$$

Then (4.2.25) implies

$$\mathbb{E}\left[|\hat{X}_T^{n+1}|^2 + \int_0^T \Big(|\hat{X}_s^{n+1}|^2 + |\hat{Y}_s^{n+1}|^2\Big) ds\right] \leqslant \frac{\theta(\epsilon,\delta,\gamma)}{\lambda(\epsilon,\delta,\gamma)}\mathbb{E}\left[|\hat{X}_T^n|^2 + \int_0^T \Big(|\hat{X}_s^{n+1}|^2 + |\hat{Y}_s^{n+1}|^2\Big) ds\right]. \tag{4.2.26}$$

Now, we assume temporarily that if there exist ϵ, δ, and γ such that

$$\lambda(\epsilon,\delta,\gamma) > \theta(\epsilon,\delta,\gamma). \tag{4.2.27}$$

Then the inequality (4.2.26) becomes a contraction, which implies that $(X_T^n)_{n\geqslant 0}$ is a Cauchy sequence in $\mathcal{L}^2(\Omega,\mathbb{P})$ and $(X^n)_{n\geqslant 0}$ and $(Y^n)_{n\geqslant 0}$ are Cauchy sequences in $\mathcal{L}^2([0,T]\times\Omega, dt\otimes d\mathbb{P})$. As a consequence, we can show that (4.1.5) has a solution. Indeed, taking the expectation of Itô's formula, using (C1) and (4.2.4), we have

$$
\begin{aligned}
&\mathbb{E}|\hat{Y}_t^n|^2 + \mathbb{E}\int_t^T \|\hat{Z}_s^n\|^2 ds + \mathbb{E}\int_t^T |\hat{\Lambda}_s^n|^2 \bullet d[M]_s \\
&= \mathbb{E}|\hat{h}^n(T)|^2 - 2\mathbb{E}\int_t^T \big\langle \hat{Y}_s^n, \hat{g}^n(s)\big\rangle ds \\
&\leqslant \mathbb{E}|\hat{h}^n(T)|^2 + 2\mathbb{E}\int_t^T |\hat{Y}_s^n|\Big[C_\theta\big(|\hat{X}_s^n| + |\hat{Y}_s^n| + \|\hat{Z}_s^n\|\big) + C_\nu\sqrt{\mathbb{E}\big(|\hat{X}_s^{n-1}|^2 + |\hat{Y}_s^{n-1}|^2\big|\mathcal{F}_{s-}^\alpha\big)}\Big]ds \\
&\leqslant \mathbb{E}|\hat{h}^n(T)|^2 + \mathbb{E}\int_t^T \bigg[\Big(2C_\theta^2 + 3C_\theta + C_\nu\Big)|\hat{Y}_s^n|^2 + C_\theta|\hat{X}_s^n|^2 + C_\nu|\hat{X}_s^{n-1}|^2 + C_\nu|\hat{Y}_s^{n-1}|^2 + \frac{1}{2}\|\hat{Z}_s^n\|^2\bigg]ds.
\end{aligned}
$$

Subsequently,

$$
\begin{aligned}
&\mathbb{E}|\hat{Y}_0^n|^2 + \frac{1}{2}\mathbb{E}\int_0^T \|\hat{Z}_s^n\|^2 ds + \mathbb{E}\int_0^T |\hat{\Lambda}_s^n|^2 \bullet d[M]_s \\
&\leqslant \mathbb{E}|\hat{h}^n(T)|^2 + \mathbb{E}\int_0^T \bigg[\Big(2C_\theta^2 + 3C_\theta + C_\nu\Big)|\hat{Y}_s^n|^2 + C_\theta|\hat{X}_s^n|^2 + C_\nu|\hat{X}_s^{n-1}|^2 + C_\nu|\hat{Y}_s^{n-1}|^2\bigg]ds \longrightarrow 0
\end{aligned}
$$

since $(X_T^n)_{n\geqslant 0}$, $(X^n)_{n\geqslant 0}$, and $(Y^n)_{n\geqslant 0}$ are Cauchy sequences. Thus, $(Z^n)_{n\geqslant 0}$ and $(\Lambda^n)_{n\geqslant 0}$ are also Cauchy sequences. As a result,

$$
\mathbb{E}\left[\sup_{s\leqslant T}\big(|X_s^n - X_s^m|^2 + |Y_s^n - Y_s^m|^2\big)\right] \to 0 \text{ as } n,\ m \to \infty.
$$

By the Banach fixed point theorem, there exist $\mathbb{F}$-adapted càdlàg processes X and Y, an $\mathcal{F}$-progressively measurable process Z and a collection of $\mathcal{F}$-progressively measurable functions Λ such that

$$
\mathbb{E}\left[\sup_{s\leqslant T}\big(|X_s^n - X_s|^2 + |Y_s^n - Y_s|^2\big) + \int_0^T \|Z_s^n - Z_s\|^2 ds + \int_0^T \big|\Lambda_s^n - \Lambda_s\big|^2 \bullet d[M]_s\right] \to 0 \text{ as } n \to \infty
$$

and

$$
\mathbb{E}\left[\sup_{s\leqslant T}\big(|X_s|^2 + |Y_s|^2\big) + \int_0^T \|Z_s\|^2 ds + \int_0^T \big|\Lambda_s\big|^2 \bullet d[M]_s\right] < \infty.
$$

Taking the limits in equation (4.2.20), we obtain that (X, Y, Z, Λ) is a solution of (4.1.5).

Next, we are to show that such γ, ϵ, and δ exist when the condition (4.1.6) is satisfied. In fact, to make the contraction meaningful, we assume $K_h - \frac{C_\mu \epsilon}{2}$ and $K_\Psi - \frac{C_\nu}{2\gamma}$ are positive.

Since $\delta > 0$ can be chosen small enough, denote

$$\theta^*(\epsilon,\gamma) = \lim_{\delta\to 0}\theta(\epsilon,\delta,\gamma) = \max\left\{\frac{C_\mu}{2\epsilon}, \gamma C_\nu\right\},$$
$$\lambda^*(\epsilon,\gamma) = \lim_{\delta\to 0}\lambda(\epsilon,\delta,\gamma) = \min\left\{K_h - \frac{C_\mu\epsilon}{2}, K_\Psi - \frac{C_\nu}{2\gamma}\right\}.$$

If we have $\lambda^*(\epsilon,\gamma) > \theta^*(\epsilon,\gamma)$ for some ϵ and γ, then there exists δ small enough such (4.2.27) is satisfied. Note that $\lambda^*(\epsilon,\gamma) > \theta^*(\epsilon,\gamma)$, equivalent to having the following inequalities:

$$K_h > \frac{C_\mu}{2}\Big(\epsilon + \frac{1}{\epsilon}\Big), \quad K_h > \frac{C_\mu\epsilon}{2} + \gamma C_\nu, \quad K_\Psi > \frac{C_\nu}{2\gamma} + \frac{C_\mu}{2\epsilon}, \quad K_\Psi > C_\nu\Big(\gamma + \frac{1}{2\gamma}\Big). \tag{4.2.28}$$

Similar to the proof in Theorem 4.1, we choose $\epsilon = 1$ and $\gamma = \frac{\sqrt{2}}{2}$. Let $\eta_1, \eta_2 > 0$ such that $C_\nu, C_\mu < \min\big\{\eta_1 K_h, \eta_2 K_\Psi\big\}$. Without loss of the generality, we assume that $\eta_1 K_h \leqslant \eta_2 K_\Psi$ (the result in the other direction turns out to be the same). Then (4.2.28) holds if the following system of inequalities hold

$$K_h > C_\mu, \quad K_h > \frac{\eta_1 K_h}{2}(\sqrt{2}+1), \quad K_\Psi > \frac{\eta_1 K_h}{2}(\sqrt{2}+1), \quad K_\Psi > \sqrt{2}C_\nu. \tag{4.2.29}$$

From the second inequality, we obtain $\eta_1 < 2(\sqrt{2}-1)$. The third inequality in (4.2.29) holds if the next inequality

$$\frac{\eta_1 K_h}{2}(\sqrt{2}+1) < \frac{\eta_1 K_h}{\eta_2}$$

holds. This implies $\eta_2 < 2(\sqrt{2}-1)$. From the fourth inequality in (4.2.29) we have $C_\nu < \frac{\sqrt{2}}{2}K_\Psi$. Note that $\frac{\sqrt{2}}{2} < 2(\sqrt{2}-1)$. Combining these we obtain the sufficient condition $C_\nu, C_\mu < \min\{2(\sqrt{2}-1)K_h, \frac{\sqrt{2}}{2}K_\Psi\}$ for which $\lambda(\epsilon,\delta,\alpha) > \theta(\epsilon,\delta,\alpha)$ when $\epsilon = 1$, $\gamma = \frac{\sqrt{2}}{2}$ and $\delta > 0$ small enough.

(2) Uniqueness of the solution: Let (X', Y', Z', Λ') be another solution to (4.1.5). Let $\Theta = (X, Y, Z)$ and $\Theta' = (X', Y', Z')$. Similar to the proof of Theorem 4.1, we take the expectation of Itô's formula to the product $\langle X' - X, Y' - Y\rangle$ and obtain

$$\mathbb{E}\Big\langle X'_T - X_T, Y'_T - Y_T\Big\rangle = \Gamma_T \tag{4.2.30}$$

where

$$\Gamma_T \triangleq \mathbb{E}\int_0^T \bigg(\Big\langle X'_s - X_s, g(s,\Theta'_s,\mathbb{P}_{(X'_s,Y'_s|\mathcal{F}^\alpha_{s-})},\alpha_s) - g(s,\Theta_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})},\alpha_s)\Big\rangle$$
$$+\Big\langle Y'_s - Y_s, f(s,\Theta'_s,\mathbb{P}_{(X'_s,Y'_s|\mathcal{F}^\alpha_{s-})},\alpha_s) - f(s,\Theta_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})},\alpha_s)\Big\rangle \tag{4.2.31}$$
$$+\Big[\sigma(s,\Theta'_s,\alpha_s) - \sigma(s,\Theta_s,\alpha_s), Z'_s - Z_s\Big]\bigg)ds.$$

Note that Assumption (L1) is the same as Assumption (H1), so (4.2.17) still holds true. That is,

$$\Gamma_T \geqslant (K_h - C_\mu)\mathbb{E}|X'_T - X_T|^2. \tag{4.2.32}$$

On the other hand, in view of Assumption (L2) and (4.2.18), we have

$$\begin{aligned}
\Gamma_T &\leqslant \mathbb{E}\int_0^T \Big[\Psi\big(s,\Theta_s,\Theta'_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})},\alpha_s\big) \\
&\quad + C_\nu W_2\big(\mathbb{P}_{(X'_s,Y'_s|\mathcal{F}^\alpha_{s-})},\mathbb{P}_{(X_s,Y_s|\mathcal{F}^\alpha_{s-})}\big)\Big(|X'_s - X_s| + |Y'_s - Y_s|\Big)\Big]ds \\
&\leqslant \mathbb{E}\int_0^T \Big[-K_\Psi\big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2\big) \\
&\quad + C_\nu\sqrt{\mathbb{E}\big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2\big|\mathcal{F}^\alpha_{s-}\big)}\Big(|X'_s - X_s| + |Y'_s - Y_s|\Big)\Big]ds \\
&\leqslant \Big[\frac{C_\nu}{2}\Big(2\gamma + \frac{1}{\gamma}\Big) - K_\Psi\Big]\mathbb{E}\int_0^T \Big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2\Big)ds.
\end{aligned} \tag{4.2.33}$$

Combining (4.2.32) and (4.2.33), we obtain

$$0 \leqslant (C_\mu - K_h)\mathbb{E}|X'_T - X_T|^2 + \Big[\frac{C_\nu}{2}\Big(2\gamma + \frac{1}{\gamma}\Big) - K_\Psi\Big]\mathbb{E}\int_0^T \Big(|X'_s - X_s|^2 + |Y'_s - Y_s|^2\Big)ds.$$

Noting now that $C_\mu, C_\nu < \min\Big\{2\big(\sqrt{2}-1\big)K_h, K_\Psi/\sqrt{2}\Big\}$, with $\gamma = \frac{1}{\sqrt{2}}$, all the coefficients of the right hand side of the above inequality are negative, which implies that, $X'_T = X_T$ $\mathbb{P}$-a.s. and for all $0 \leqslant s \leqslant T$, $X'_s = X_s$ and $Y'_s = Y_s$, $\mathbb{P}$-a.s.

Now we take the expectation of Itô's formula for $|Y'-Y|^2$,

$$
\begin{aligned}
&\mathbb{E}|Y_0'-Y_0|^2+\mathbb{E}\int_0^T\|Z_s'-Z_s\|^2ds+\mathbb{E}\int_0^T|\Lambda_s'-\Lambda_s|^2\bullet d[M]_s\\
&=\mathbb{E}|h(X_T',\alpha_T)-h(X_T,\alpha_T)|^2\\
&\quad-2\mathbb{E}\int_0^T\Big\langle Y_s'-Y_s, g(s,\Theta_s',\mathbb{P}_{(X_s',Y_s'|\mathcal{F}_{s-}^{\alpha})},\alpha_s)-g(s,\Theta_s,\mathbb{P}_{(X_s,Y_s|\mathcal{F}_{s-}^{\alpha})},\alpha_s)\Big\rangle ds.
\end{aligned}
$$

Since $X_T'=X_T$ $\mathbb{P}$-a.s. and for all $0\leqslant s\leqslant T$, $X_s'=X_s$ and $Y_s'=Y_s$ $\mathbb{P}$-a.s., we get

$$
\mathbb{E}\int_0^T\|Z_s'-Z_s\|^2ds+\mathbb{E}\int_0^T|\Lambda_s'-\Lambda_s|^2\bullet d[M]_s=0.
$$

Thus, $Z'=Z$ (in $\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})$) and $\Lambda'=\Lambda$ (in $\mathcal{M}^2(0,T;\mathbb{R}^d)$), yielding that the solution of (4.1.2) is unique. $\square$

Note that with a little effort, we can show that Theorem 4.1 still hold true if $\mathbb{P}_{(X_s,Y_s|\mathcal{F}_{s-}^{\alpha})}$ is replaced by $\mathbb{P}_{(X_s,Y_s,Z_s|\mathcal{F}_{s-}^{\alpha})}$ in (4.1.2). Nevertheless, with the present approach, we cannot include Λ_s to these probability measures. If Λ_s is included, $\nu_s^n=\mathbb{P}_{(X_s^n,Y_s^n,Z_s^n,\Lambda_s^n|\mathcal{F}_s^{\alpha})}$ and by virtue of (4.1.1), to estimate the term involving $W_2^2(\nu_s^n,\nu_s^{n-1})$, one needs some estimate of the expected value $\mathbb{E}\int_0^T|\hat{\Lambda}_s^n|^2ds$ which is not possible because the quadratic variations of the martingale associate with the Markov chain are random.

CHAPTER 5
Application in Conditional Mean-Field Nonzero-sum Game

In this chapter, we consider a nonzero-sum game problem with N players in which the dynamics and cost functionals of each player depend on conditional mean-field terms and a regime-switching process. Let N, d_i, $1\leqslant i\leqslant N$, be positive integers. For each i, $1\leqslant i\leqslant N$, let $\mathcal{U}^i = \mathcal{L}(0,T;\mathbb{R}^{d_i})$ be the set of admissible controls of the player i and denote $\mathcal{U} = \mathcal{U}^1 \times \mathcal{U}^2 \times \ldots \times \mathcal{U}^N$. The dynamics of the system is given by the following conditional mean-field SDE

$$\begin{aligned} X_t = x_0 + \int_0^t \Big[A(s,\alpha_s)X_s + \bar{A}(s,\alpha_s)\mathbb{E}\big(X_s\big|\mathcal{F}^\alpha_{s-}\big) + \sum_{i=1}^N B^i(s,\alpha_s)u^i_s + f(s,\alpha_s)\Big]ds \\ + \int_0^t \Big[\sigma(s,\alpha_s)X_s + g(s,\alpha_s)\Big]dW_s \end{aligned} \tag{5.0.1}$$

where for each $i_0 \in \mathcal{M}$ and $1\leqslant i\leqslant N$, $u^i \in \mathcal{U}^i$, $A(\cdot,i_0)$, $\bar{A}(\cdot,i_0)$, $B^i(\cdot,i_0)$, $f(\cdot,i_0)$, and $g(\cdot,i_0)$ are bounded continuous functions taking values in $\mathbb{R}^{d\times d}$, $\mathbb{R}^{d\times d}$, $\mathbb{R}^{d\times d_i}$, $\mathbb{R}^d$, and $\mathbb{R}^d$, respectively. In addition, $\sigma(\cdot,i_0)$ is a continuous function taking values in $\mathbb{R}^{d\times d}$.

Note that the mean-field SDE (5.0.1) is obtained as the mean-square limit as $n\to\infty$ of a system of interacting players of the form

$$\begin{aligned} X^{k,n}_t = x_0 + \int_0^t \Big[A(s,\alpha_s)X^{k,n}_s + \frac{1}{n}\bar{A}(s,\alpha_s)\mathbb{E}\sum_{j=1}^n X^{j,n}_s + \sum_{i=1}^N B^i(s,\alpha_s)u^{i,k,n}_s + f(s,\alpha_s)\Big]ds \\ + \int_0^t \Big[\sigma(s,\alpha_s)X^{k,n}_s + g(s,\alpha_s)\Big]dW^k_s, \qquad 1\leqslant k\leqslant n, \end{aligned}$$

where $(W^k, k\geqslant 1)$ is a collection of independent standard Brownian motions. Due to the symmetry of the dynamics, we assume that the social planners apply the same control policies for all players in the feedback forms

$$u_t^{i,k,n} = \phi^i\left(t, X_t^{k,n}, \frac{1}{n}\sum_{j=1}^{n} X_t^{j,n}, \alpha_t\right).$$

In (5.0.1), the conditional expectation $\mathbb{E}(X_s|\mathcal{F}_s^\alpha)$ appears instead of the expectation $\mathbb{E}(X_s)$ because of the effect of the common switching process α_t to all the players.

For simplicity, through out this section we assume that W_s is a scalar Brownian motion. The case with multidimensional Brownian motion W_s can be treated in a similar way. Given a control $\mathbf{u} = (u^1, u^2, \ldots, u^N) \in \mathcal{U}$, the cost functional of the player i is given by

$$\begin{aligned} J_i(\mathbf{u}) = \frac{1}{2}\mathbb{E}\Bigg\{ \int_0^T \Big[X_s^\intercal M^i(s,\alpha_s) X_s + \mathbb{E}\big(X_s^\intercal \big|\mathcal{F}_{s-}^\alpha\big)\, \overline{M}^i(s,\alpha_s) \mathbb{E}\big(X_s \big|\mathcal{F}_{s-}^\alpha\big) + (u_s^i)^\intercal N^i(s,\alpha_s) u_s^i \Big] ds \\ + X_T^\intercal R^i(\alpha_T) X_T + \mathbb{E}\big(X_T^\intercal \big|\mathcal{F}_T^\alpha\big)\, \bar{R}^i(\alpha_T) \mathbb{E}\big(X_T \big|\mathcal{F}_T^\alpha\big) \Bigg\}, \end{aligned} \tag{5.0.2}$$

where for each $i_0 \in \mathcal{M}$ and $1 \leqslant i \leqslant N$, $M^i(\cdot, i_0)$ and $\overline{M}^i(\cdot, i_0)$ are bounded continuous symmetric non-negative matrices with values in $\mathbb{R}^{d\times d}$, $N^i(\cdot, i_0)$ and its inverse $\big(N^i(\cdot, i_0)\big)^{-1}$ are bounded continuous symmetric positive matrices with values in $\mathbb{R}^{d_i \times d_i}$, and $R^i(i_0)$ and $\bar{R}^i(i_0)$ are symmetric non-negative matrices with values in $\mathbb{R}^{d\times d}$.

An admissible control $\mathbf{u}^* = (u^{*,1}, u^{*,2}, \ldots, u^{*,N}) \in \mathcal{U}$ is called a Nash equilibrium point if for any $1 \leqslant i \leqslant N$,

$$J_i(\mathbf{u}^*) \leqslant J_i(\mathbf{u}^{*,-i}, u^i), \quad \forall u^i \in \mathcal{U}^i$$

where $(\mathbf{u}^{*,-i}, u^i) = (u^{*,1}, u^{*,2}, \ldots, u^{*,i-1}, u^i, u^{*,i+1}, \ldots, u^{*,N})$. We are interested in finding a Nash equilibrium point for the aforementioned game problem.

For each $(t,x,\bar{x},u^1,u^2,\dots,u^N,p^i,q^i,i_0)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\times\ldots\times\mathbb{R}^{d_N}\times\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{M}$, $1\leqslant i\leqslant N$, denote the Hamiltonian associated with the player i as follow

$$\begin{aligned}H_i\Big(t,x,\bar{x},u^1,u^2,\dots,u^N,p^i,q^i,i_0\Big) &= (p^i)^\intercal\Big[A(t,i_0)x+\bar{A}(t,i_0)\bar{x}+\sum_{j=1}^N B^j(t,i_0)u^j+f(t,i_0)\Big]\\&\quad+\frac{1}{2}\Big[x^\intercal M^i(t,i_0)x+\bar{x}^\intercal\overline{M}^i(t,i_0)\bar{x}+(u^i)^\intercal N^i(t,i_0)u^i\Big]\\&\quad+\Big(\sigma^\intercal(t,i_0)x+g(t,i_0)\Big)^\intercal q^i.\end{aligned}$$

In addition, define the function $\hat{u}^i$ by

$$\hat{u}^i(t,p^i)=-\big(N^i(t,\alpha_t)\big)^{-1}\big(B^i(t,\alpha_t)\big)^\intercal p^i,\quad 0\leqslant t\leqslant T.$$

It is easy to check that the functions $\hat{u}^i$, $1\leqslant i\leqslant N$, satisfy

$$\begin{aligned}&H_i\Big(t,x,\bar{x},\hat{u}^1(t,p^1),\hat{u}^2(t,p^2),\dots,\hat{u}^N(t,p^N),p^i,q^i,i_0\Big)\\&\leqslant H_i\Big(t,x,\bar{x},\hat{u}^1(t,p^1),\hat{u}^2(t,p^2),\dots,\hat{u}^{i-1}(t,p^{i-1}),u^i,\hat{u}^{i+1}(t,p^{i+1}),\dots,\hat{u}^N(t,p^N),p^i,q^i,i_0\Big)\end{aligned}$$

for any $u^i\in\mathbb{R}^{d_i}$, $p^i,q^i\in\mathbb{R}^d$, $1\leqslant i\leqslant N$. We have the following proposition.

Proposition 5.1. *The process* $\big(X_t,(p_t^1,q_t^1),\dots,(p_t^N,q_t^N),\lambda_t^1,\dots,\lambda_t^N\big)$ *solves the following conditional mean-field FBSDE with regime switching*

$$\begin{cases}X_t=x_0+\displaystyle\int_0^t\Big[A(s,\alpha_s)X_s+\bar{A}(s,\alpha_s)\mathbb{E}\big(X_s\big|\mathcal{F}_{s-}^\alpha\big)+\sum_{i=1}^N B^i(s,\alpha_s)\hat{u}^i(s,p_s^i)+f(s,\alpha_s)\Big]ds\\ \qquad\quad+\displaystyle\int_0^t\Big[\sigma(s,\alpha_s)X_s+g(s,\alpha_s)\Big]dW_s,\\ p_t^i=\Big[R^i(\alpha_T)X_T+\bar{R}^i(\alpha_T)\mathbb{E}\big(X_T\big|\mathcal{F}_T^\alpha\big)\Big]+\displaystyle\int_t^T\Big[A(s,\alpha_s)^\intercal p_s^i+\mathbb{E}\big(\bar{A}(s,\alpha_s)^\intercal p_s^i\big|\mathcal{F}_{s-}^\alpha\big)\\ \qquad\quad+M^i(s,\alpha_s)X_s+\overline{M}^i(s,\alpha_s)\mathbb{E}\big(X_s\big|\mathcal{F}_{s-}^\alpha\big)+\sigma(s,\alpha_s)^\intercal q_s^i\Big]ds-\displaystyle\int_t^T q_s^i dW_s-\int_t^T\lambda_s^i\bullet dM_s,\\ i=1,\dots,N\end{cases}\tag{5.0.3}$$

if and only if the admissible control $\hat{\mathbf{u}} = (\hat{u}^1, \hat{u}^2, \dots, \hat{u}^N) = \big(\hat{u}^1(t, p_t^1), \hat{u}^2(t, p_t^2), \dots, \hat{u}^N(t, p_t^N)\big)$ *is a Nash equilibrium point of the conditional mean-field nonzero-sum quadratic stochastic differential game.*

Proof. First, we shall show that the condition is sufficient. Suppose that $\big(X_t, (p_t^1, q_t^1), \dots, (p_t^N, q_t^N), \lambda_t^1, \dots, \lambda_t^N\big)$ is a solution of (5.0.3). Fix i, $1 \leqslant i \leqslant N$. Let $u^i \in \mathcal{U}^i$ and denote $\mathbf{u}^i = (\hat{\mathbf{u}}^{-i}, u^i) = \big(\hat{u}^1, \dots, \hat{u}^{i-1}, u^i, \hat{u}^{i+1}, \dots, \hat{u}^N\big)$. Let X_t^i be the state dynamics corresponding to the control $\mathbf{u}^i$. For simplicity, denote $\hat{u}^i(s) = \hat{u}^i(s, p_s^i)$, $\bar{X}_s = \mathbb{E}\big(X_s \big| \mathcal{F}_{s-}^\alpha\big)$, $\bar{X}_s^i = \mathbb{E}\big(X_s^i \big| \mathcal{F}_{s-}^\alpha\big)$, and $\bar{p}_s^i = \mathbb{E}\big(p_s^i \big| \mathcal{F}_{s-}^\alpha\big)$ for $0 \leqslant s \leqslant T$. As a consequence, $\bar{X}_s^\intercal = \mathbb{E}\big(X_s^\intercal \big| \mathcal{F}_{s-}^\alpha\big)$ and $(\bar{X}_s^i)^\intercal = \mathbb{E}\big[(X_s^i)^\intercal \big| \mathcal{F}_{s-}^\alpha\big]$. It suffices to prove that $J_i(\mathbf{u}^i) \geqslant J_i(\hat{\mathbf{u}})$.

To proceed, we observe that for any symmetric non-negative $n \times n$ matrix S and $v^1, v^2 \in \mathbb{R}^n$ we have

$$(v^1)^\intercal S v^1 - (v^2)^\intercal S v^2 = (v^1 - v^2)^\intercal S (v^1 - v^2) + 2(v^1 - v^2)^\intercal S v^2 \geqslant 2(v^1 - v^2)^\intercal S v^2.$$

Note that $M^i(s, \alpha_s), \overline{M}^i(s, \alpha_s), N^i(s, \alpha_s), R^i(\alpha_T)$ and $\bar{R}^i(\alpha_T)$ are all symmetric non-negative matrices. Hence, using the definition of $J_i(\cdot)$ and then above inequality yields

$$\begin{aligned}
&J_i(\mathbf{u}^i) - J_i(\hat{\mathbf{u}}) = J_i\big(\hat{u}_1, \dots, \hat{u}_{i-1}, u^i, \hat{u}_{i+1}, \dots, \hat{u}_N\big) - J_i(\hat{u}) \\
&= \frac{1}{2}\mathbb{E}\Bigg\{ \int_0^T \Bigg[(X_s^i)^\intercal M^i(s, \alpha_s) X_s^i - X_s^\intercal M^i(s, \alpha_s) X_s + (\bar{X}_s^i)^\intercal \overline{M}^i(s, \alpha_s) \bar{X}_s^i - \bar{X}_s^\intercal \overline{M}^i(s, \alpha_s) \bar{X}_s \\
&\qquad + (u_s^i)^\intercal N^i(s, \alpha_s) u_s^i - (\hat{u}_s^i)^\intercal N^i(s, \alpha_s) \hat{u}_s^i \Bigg] ds \\
&\qquad\qquad + (X_T^i)^\intercal R^i(\alpha_T) X_T^i - X_T^\intercal R^i(\alpha_T) X_T + (\bar{X}_T^i)^\intercal \bar{R}^i(\alpha_T) \bar{X}_T^i - \bar{X}_T^\intercal \bar{R}^i(\alpha_T) \bar{X}_T \Bigg\} \\
&\geqslant \mathbb{E}\Bigg\{ \int_0^T \Bigg[(X_s^i - X_s)^\intercal M^i(s, \alpha_s) X_s + (\bar{X}_s^i - \bar{X}_s)^\intercal \overline{M}^i(s, \alpha_s) \bar{X}_s + (u_s^i - \hat{u}_s^i)^\intercal N^i(s, \alpha_s) \hat{u}_s^i \Bigg] ds \\
&\qquad + (X_T^i - X_T)^\intercal R^i(\alpha_T) X_T + (\bar{X}_T^i - \bar{X}_T)^\intercal \bar{R}^i(\alpha_T) \bar{X}_T \Bigg\}. \qquad (5.0.4)
\end{aligned}$$

Next, we show that the rightmost of (5.0.4) is equal to 0. Note that

$$X_t^i - X_t = \int_0^t \Big[A(s,\alpha_s)\big(X_s^i - X_s\big) + \bar{A}(s,\alpha_s)(\bar{X}_s^i - \bar{X}_s\big) + B^i(s,\alpha_s)\big(u_s^i - \hat{u}^i(s,p_s^i)\big)\Big]ds + \int_0^t \sigma(s,\alpha_s)\big(X_s^i - X_s\big)dW_s$$

and $p_T^i = \Big[R^i(\alpha_T)X_T + \bar{R}^i(\alpha_T)\mathbb{E}\big(X_T\big|\mathcal{F}_T^\alpha\big)\Big]$. Hence, by Itô formula and backward equation in (5.0.3) we have

$$\begin{aligned}
&\big(X_T^i - X_T\big)^\intercal p_T^i \\
&= -\int_0^T \big(X_s^i - X_s\big)^\intercal\Big[A(s,\alpha_s)^\intercal p_s^i + \bar{A}(s,\alpha_s)^\intercal \bar{p}_s^i + M^i(s,\alpha_s)X_s + \overline{M}^i(s,\alpha_s)\bar{X}_s + \sigma(s,\alpha_s)^\intercal q_s^i\Big]ds \\
&\quad + \int_0^T \Big[\big(X_s^i - X_s\big)^\intercal A(s,\alpha_s)^\intercal + (\bar{X}_s^i - \bar{X}_s\big)^\intercal \bar{A}(s,\alpha_s)^\intercal + \big(u_s^i - \hat{u}^i(s,p_s^i)\big)^\intercal B^i(s,\alpha_s)^\intercal\Big]p_s^i ds \\
&\quad + \int_0^T \big(X_s^i - X_s\big)^\intercal \sigma(s,\alpha_s)^\intercal q_s^i ds \\
&\quad + \int_0^t \big(X_s^i - X_s\big)^\intercal q_s^i dW_s + \int_0^t \big(X_s^i - X_s\big)^\intercal \sigma(s,\alpha_s)^\intercal p_s^i dW_s + \int_0^t \big(X_s^i - X_s\big)^\intercal \lambda_s^i \bullet dM_s.
\end{aligned} \tag{5.0.5}$$

Since

$$\begin{aligned}
\mathbb{E}\Big[(\bar{X}_s^i - \bar{X}_s)^\intercal \bar{A}(s,\alpha_s)^\intercal p_s^i\Big] &= \mathbb{E}\Big\{\mathbb{E}\Big[(\bar{X}_s^i - \bar{X}_s)^\intercal \bar{A}(s,\alpha_s)^\intercal p_s^i\Big|\mathcal{F}_{s-}^\alpha\Big]\Big\} \\
&= \mathbb{E}\Big[(\bar{X}_s^i - \bar{X}_s\big)^\intercal \bar{A}(s,\alpha_s)^\intercal \bar{p}_s^i\Big] \\
&= \mathbb{E}\Big[(X_s^i - X_s\big)^\intercal \bar{A}(s,\alpha_s)^\intercal \bar{p}_s^i\Big],
\end{aligned}$$

simplifying the right hand side of (5.0.5) and taking the expectations its both sides we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\big(X_T^i - X_T\big)^\intercal p_T^i\Big] = \mathbb{E}\Big\{\big(X_T^i - X_T\big)^\intercal\Big[R^i(\alpha_T)X_T + \bar{R}^i(\alpha_T)\bar{X}_T\Big]\Big\} \\
&= \mathbb{E}\int_0^T \Big\{ -\big(X_s^i - X_s\big)^\intercal\Big[M^i(s,\alpha_s)X_s + \overline{M}^i(s,\alpha_s)\bar{X}_s\Big] + \big(u_s^i - \hat{u}_s^i\big)^\intercal B^i(s,\alpha_s)^\intercal p_s^i\Big\}ds.
\end{aligned}$$

The equation $\hat{u}^i_s = -\big(N^i(s,\alpha_s)\big)^{-1} B^i(s,\alpha_s)^\intercal p^i_s$ then implies

$$\begin{aligned}&\mathbb{E}\Big\{\big(X^i_T - X_T\big)^\intercal\Big[R^i(\alpha_T)X_T + \bar{R}^i(\alpha_T)\bar{X}_T\Big]\Big\}\\&= -\mathbb{E}\int_0^T\Big\{\big(X^i_s - X_s\big)^\intercal\Big[M^i(s,\alpha_s)X_s + \overline{M}^i(s,\alpha_s)\bar{X}_s\Big] + \big(u^i_s - \hat{u}^i_s\big)^\intercal N^i(s,\alpha_s)\hat{u}^i_s\Big\}ds,\end{aligned}$$

which subsequently proves that the rightmost hand side of (5.0.4) equals 0. Therefore, it follows from (5.0.4) that $J_i(\mathbf{u}^i) - J_i(\hat{\mathbf{u}})\geqslant 0$.

To complete the proof, we show that the condition is necessary. Suppose that $\hat{\mathbf{u}} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N)$ is a Nash equilibrium point of the game. Denote the corresponding state trajectory by $\hat{X}$. Clearly, if we fix the control $\hat{u}^j$ for $j \neq i$, $1\leqslant i, j\leqslant N$, then $\hat{u}^i$ is the optimal control for the player i and the corresponding optimal trajectory is $\hat{X}$. Since the control problem for the player i is of conditional mean-field type with Markovian switching, we can apply the maximum principle in [8] to get the necessary condition for optimality. The adjoint equation associate to the control problem of player i is

$$\begin{aligned}p^i_t = \Big[R^i(\alpha_T)X_T + \bar{R}^i(\alpha_T)\bar{X}_T\Big] + \int_t^T\Big[&A(s,\alpha_s)^\intercal p^i_s + \bar{A}(s,\alpha_s)^\intercal\mathbb{E}\big(p^i_s\big|\mathcal{F}^\alpha_{s-}\big)\\&+ M^i(s,\alpha_s)X_s + \overline{M}^i(s,\alpha_s)\bar{X}_s + \sigma(s,\alpha_s)^\intercal q^i_s\Big]ds - \int_t^T q^i_s dW_s - \int_t^T \lambda^i_s \bullet dM_s,\end{aligned}$$

which always admits a unique solution (see [8, Theorem 3.4]). According the [8, Theorem 3.7], for any vector $v^i \in \mathbb{R}^{d_i}$ and $t \in [0,T]$,

$$\frac{d}{du^i}H_i\Big(t, X_t, \bar{X}_t, \hat{u}^1_t, \ldots, \hat{u}^{i-1}_t, \hat{u}^i_t, \hat{u}^{i+1}_t, \ldots, \hat{u}^N(t), p^i_t, q^i_t, \alpha_{t-}\Big)\big(v^i - \hat{u}^i_t\big)\geqslant 0, \qquad \mathbb{P}-a.s.$$

or, equivalently,

$$\Big[B^i(t,\alpha_t)^\intercal p^i_t + N^i(t,\alpha_t)\hat{u}^i_t\Big]\big(v^i - \hat{u}^i_t\big)\geqslant 0, \qquad \mathbb{P}-a.s.$$

Since the inequality holds for any $v^i \in \mathbb{R}^{d_i}$, we must have $B^i(t,\alpha_t)^\intercal p^i_t + N^i(t,\alpha_t)\hat{u}^i_t = 0$. As a consequence, $\hat{u}^i_t = -\big(N^i(t,\alpha_t)\big)^{-1}B^i(t,\alpha_t)^\intercal p^i_t$. Plug this value of $\hat{u}^i_t$ in the above adjoint equation we derive that $\big(X_t,(p_t^1,q_t^1),\ldots,(p_t^N,q_t^N),\lambda_t^1,\ldots,\lambda_t^N\big)$ is a solution of the FBSDEs (5.0.3). This completes the proof. $\square$

Next, we present conditions on the coefficients such that a Nash equilibrium point of the differential game exists. To this end, we first need the following assumptions.

Assumption (K)

(K1) For $1\leqslant i\leqslant N$, the matrices $B^i(t,i_0)\equiv B^i$ and $N^i(t,i_0)\equiv N^i$ are independent of t and i_0. Denote

$$K^i = B^i(N^i)^{-1}(B^i)^\intercal.$$

(K2) There exist constants $\beta_1,\beta_2>0$ such that for any $x\in\mathbb{R}^d$ and $0\leqslant t\leqslant T$,

$$x^\intercal\Big[\sum_{i=1}^N K^iR^i(i_0)\Big]x\geqslant\beta_1|x|^2,\quad x^\intercal\Big[\sum_{i=1}^N K^iM^i(t,i_0)\Big]x\geqslant\beta_2|x|^2.$$

(K3) For $1\leqslant i\leqslant N$ and $0\leqslant t\leqslant T$, $\mathbb{P}$-a.s.

$$K^iA^\intercal(t,i_0)=A^\intercal(t,i_0)K^i,\quad K^i\bar{A}^\intercal(t,i_0)=\bar{A}^\intercal(t,i_0)K^i,\quad K^i\sigma^\intercal(t,i_0)=\sigma^\intercal(t,i_0)K^i.$$

For $(t,i_0)\in[0,T]\times\mathcal{M}$, denote

$$M(t,i_0)=\sum_{i=1}^N K^iM^i(t,i_0),\quad R(i_0)=\sum_{i=1}^N K^iR^i(i_0),$$

and, similarly, $\bar{M}(t,i_0)=\sum_{i=1}^N K^i\bar{M}^i(t,i_0)$, $\bar{R}(i_0)=\sum_{i=1}^N K^i\bar{R}^i(i_0)$.

Let us consider the following conditional mean-field FBSDEs with regime switching:

$$\begin{aligned}
X_t &= x_0+\int_0^t\Big[A(s,\alpha_s)X_s+\bar{A}(s,\alpha_s)\bar{X}_s-Y_s+f(s,\alpha_s)\Big]ds+\int_0^t\Big[\sigma(s,\alpha_s)X_s+g(s,\alpha_s)\Big]dW_s,\\
Y_t &= \big(R(\alpha_T)X_T+\bar{R}(\alpha_T)\bar{X}_T\big)-\int_t^T Z_sdW_s-\int_t^T\Lambda_s\bullet dM_s \qquad (5.0.6)\\
&\quad+\int_t^T\Big[A(s,\alpha_s)^\intercal Y_s+\bar{A}(s,\alpha_s)^\intercal\bar{Y}_s+M(s,\alpha_{s-})X_s+\bar{M}(s,\alpha_{s-})\bar{X}_s+\sigma(s,\alpha_s)^\intercal Z_s\Big]ds,
\end{aligned}$$

where $X,Y\in\mathcal{S}^2(0,T;\mathbb{R}^d)$, $Z\in\mathcal{L}^2(0,T;\mathbb{R}^{d\times d})$, $\Lambda\in\mathcal{M}^2(0,T;\mathbb{R}^d)$.

Note that if $\big(X_t,\,(p_t^1,q_t^1),\,\ldots,\,(p_t^N,q_t^N),\,\lambda_t^1,\ldots,\lambda_t^N\big)$ is a solution of (5.0.3) and Assumption **(K)** holds, then applying Itô's formula we obtain

$$K^i p_t^i = \Big[K^i R^i(\alpha_T)X_T + K^i\bar{R}^i(\alpha_T)\mathbb{E}\big(X_T\big|\mathcal{F}^{\alpha}_{T-}\big)\Big] + \int_t^T \Big[A(s,\alpha_s)^\intercal K^i p_s^i + \mathbb{E}(\bar{A}(s,\alpha_s)^\intercal K^i p_s^i|\mathcal{F}^{\alpha}_{s-}) \\ + K^i M^i(s,\alpha_s)X_s + K^i\bar{M}^i(s,\alpha_s)\bar{X}_s + \sigma(s,\alpha_s)^\intercal K^i q_s^i\Big]ds - \int_t^T K^i q_s^i dW_s - \int_t^T K^i\lambda_s^i \bullet dM_s. \tag{5.0.7}$$

By taking the sum of (5.0.7) where $i = 1,2,\ldots,N$, we easily see that the process $\big(X_t, Y_t = \sum_{i=1}^N K^i p_t^i,\, Z_t = \sum_{i=1}^N K^i q_t^i,\, \Lambda_t = \sum_{i=1}^N K^i\lambda_t^i\big)$ is a solution of FBSDEs (5.0.6).

Now we are in a position to show that the coefficients in (5.0.6) satisfy all conditions in Section 3. For any t, x, y, z, i_0, and $\nu \in \mathcal{P}(\mathbb{R}^{2d}), \mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\begin{aligned} f(t,x,y,z,\nu,i_0) &= A(t,i_0)x + \bar{A}(t,i_0)\int_{\mathbb{R}^{d+d}} \zeta_1\nu(d\zeta_1,d\zeta_2) - y + f(t,i_0), \\ g(t,x,y,z,\nu,i_0) &= -A(t,i_0)^\intercal y - \bar{A}(t,i_0)^\intercal \int_{\mathbb{R}^{d+d}} \zeta_2\nu(d\zeta_1,d\zeta_2) \\ &\qquad - M(t,i_0)x - \bar{M}(t,i_0)\int_{\mathbb{R}^{d+d}} \zeta_1\nu(d\zeta_1,d\zeta_2) - \sigma(t,i_0)^\intercal z, \\ \sigma(t,x,y,z,\nu,i_0) &= \sigma(t,i_0)x + g(t,i_0), \\ h(x,\mu,i_0) &= R(i_0)x + \bar{R}(i_0)\int_{\mathbb{R}^d} \zeta\mu(d\zeta). \end{aligned}$$

Because of the boundedness of $\bar{A}(s,i_0)$, f is uniformly Lipschitz with respect to ν. The linearity implies that f satisfies assumption (C1). Similarly, we obtain that g and σ satisfy assumption (C1) and h satisfy assumption (C2). More precisely, it can be shown that in the present setting, we can take $C_\theta = \max_{t,i_0}\{1, \|A(t,i_0)\|, \|M(t,i_0)\|, \|\sigma(t,i_0)\|\}$ and $C_\nu = \sqrt{2}\max_{t,i_0}\{\|\bar{A}(t,i_0)\|, \|\bar{M}(t,i_0)\|\}$ in (C1) and $c = \max_{i_0}\|R(i_0)\|$ and $C_\mu = \max_{i_0}\|\bar{R}(i_0)\|$ in (C2). Due to the linearity, it is trivial to verify the constants C_θ and c. For C_ν, it suffices to verify for $\varphi = g$ in assumption (C1). If $\nu_1 = \mathbb{P}_{(\Upsilon_{1,1},\Upsilon_{1,2})}$ and $\nu_2 = \mathbb{P}_{(\Upsilon_{2,1},\Upsilon_{2,2})}$, then

$$\begin{aligned}
&\big|g(t,x,y,z,\nu_1,i_0) - g(t,x,y,z,\nu_2,i_0)\big| \\
&= \Big|\bar{A}(t,i_0)^\intercal(\mathbb{E}\Upsilon_{1,2} - \mathbb{E}\Upsilon_{2,2}) - \bar{M}(t,i_0)^\intercal(\mathbb{E}\Upsilon_{1,1} - \mathbb{E}\Upsilon_{2,1})\Big| \\
&\leqslant \sqrt{2}\max\Big\{\|\bar{A}(t,i_0)\|, \|\bar{M}(t,i_0)\|\Big\}\Big(\mathbb{E}\big|\Upsilon_{1,2} - \Upsilon_{2,2}\big|^2 + \mathbb{E}\big|\Upsilon_{1,1} - \Upsilon_{2,1}\big|^2\Big)^{1/2} \\
&= \sqrt{2}\max\Big\{\|\bar{A}(t,i_0)\|, \|\bar{M}(t,i_0)\|\Big\}\Big(\mathbb{E}\big|\Upsilon_1 - \Upsilon_2\big|^2\Big)^{1/2}.
\end{aligned}$$

Hence,

$$\big|g(t,x,y,z,\nu_1,i_0) - g(t,x,y,z,\nu_2,i_0)\big| \leqslant \sqrt{2}\max\Big\{\|\bar{A}(t,i_0)\|, \|\bar{M}(t,i_0)\|\Big\}W_2(\nu_1,\nu_2).$$

By a similar way, we can verify $C_\mu \leqslant \max_{i_0} \|\bar{R}(i_0)\|$ in (C2).

In addition, in the present setting, the operator Ψ defined in (4.1.3), related to (5.0.6), becomes

$$\Psi\big(t,\theta,\theta',\nu,i_0\big) = -|y-y'|^2 - M(t,i_0)|x-x'|^2.$$

In Proposition 5.2 below, we show that if Assumption **(K)** holds then Ψ and h satisfy Assumption **(L)** with the constants $K_\Psi = \min\{1,\beta_2\}$ and $K_h = \beta_1$.

Proposition 5.2. *Assume that Assumption* **(K)** *holds and*

$$\begin{aligned}
&(1) \quad \|\bar{A}(t,i_0)\|, \|\bar{M}(t,i_0)\| < \min\Big\{(2-\sqrt{2})\beta_1, \frac{1}{2}\beta_2\Big\} \\
&(2) \quad \Big\|\bar{R}(i_0)\Big\| < \min\Big\{2(\sqrt{2}-1)\beta_1, \frac{\sqrt{2}}{2}\beta_2\Big\}.
\end{aligned} \tag{5.0.8}$$

Then the conditional mean-field FBSDEs (5.0.3) has a unique solution $\big(X_t, (p_t^1, q_t^1), \ldots, (p_t^N, q_t^N), \lambda_t^1, \ldots, \lambda_t^N\big)$, *where* $X, p^i \in \mathcal{S}^2(0,T;\mathbb{R}^d)$, $q^i \in \mathcal{L}^2(0,T;\mathbb{R}^d)$, *and* $\lambda^i \in \mathcal{M}^2(0,T;\mathbb{R}^d)$ *for all* $i = 1,\ldots,N$.

Proof. Because Assumption **(K)** holds,

$$\Psi\big(t,\theta,\theta',\nu,i_0\big) = -|y-y'|^2 - M(t,i_0)|x-x'|^2 \leqslant -|y-y'|^2 - \beta_2|x-x'|^2. \tag{5.0.9}$$

This implies that $K_\Psi = \min\{1, \beta_2\}$. In addition, for any $x, x' \in \mathbb{R}^d$ and $\mu \in \mathcal{P}(\mathbb{R}^{2d})$,

$$\big\langle h(x, \mu, i_0) - h(x', \mu, i_0), x - x' \big\rangle = (x - x')^\intercal R(i_0)(x - x') \geqslant \beta_1 |x - x'|^2, \tag{5.0.10}$$

which implies that $K_h = \beta_1$. As a consequence, (5.0.8) and the inequalities

$$C_\nu \leqslant \sqrt{2} \max_{t\in[0,T], i_0 \in \mathcal{M}} \{\|\bar{A}(t, i_0)\|, \|\bar{M}(t, i_0)\|\}, \quad C_\mu \leqslant \max_{i_0 \in \mathcal{M}} \|\bar{R}(i_0)\|$$

imply that Assumptions **(C)** and **(L)** are satisfied. Hence, we can apply Theorem 4.2 to derive the existence of a unique process (X, Y, Z, Λ) that solves the conditional mean-field FBSDE with regime switching (5.0.6).

Next, according to [8, Theorem 3.4], for $i = 1, \ldots, N$, there exists $(p^i, q^i, \lambda^i) \in \mathcal{S}^2(0, T; \mathbb{R}^d) \times \mathcal{L}^2(0, T; \mathbb{R}^d) \times \mathcal{M}^2(0, T; \mathbb{R}^d)$ unique solution of the following BSDE:

$$\begin{aligned} p_t^i = \Big(R^i(\alpha_T)X_T + \bar{R}^i(\alpha_T)\bar{X}_T\Big) + \int_t^T \Big[A(s, \alpha_s)^\intercal p_s^i + \bar{A}(s, \alpha_s)^\intercal \mathbb{E}\big(p_s^i \big| \mathcal{F}_{s-}^\alpha\big) \\ + M^i(s, \alpha_s)X_s + \sigma(s, \alpha_s)^\intercal q_s^i\Big] ds - \int_t^T q_s^i dW_s - \int_t^T \lambda_s^i \bullet dM_s, \quad t \in [0, T]. \end{aligned} \tag{5.0.11}$$

Hence, the processes $\Big(X, Y = \sum_{i=1}^N K^i p^i, Z = \sum_{i=1}^N K^i q^i, \Lambda = \sum_{i=1}^N K^i \lambda^i\Big)$ is a solution of (5.0.6). Because the solution of (5.0.11) is unique, then $Y = \sum_{i=1}^N K^i p^i$, $Z = \sum_{i=1}^N K^i q^i$ and $\Lambda = \sum_{i=1}^N K^i \lambda^i$. Substitute Y, Z and Λ to (5.0.6) we obtain that $\big(X_t, (p_t^1, q_t^1), \ldots, (p_t^N, q_t^N), \lambda_t^1, \ldots, \lambda_t^N\big)$ is a solution of (5.0.3). This completes the proof. □

Combining Proposition 5.1 and Proposition 5.2, the next result follows.

Theorem 5.3. *Assume Assumption* **(K)** *holds and*

$$(1) \quad \|\bar{A}(t, i_0)\|, \|\bar{M}(t, i_0)\| < \min\Big\{(2 - \sqrt{2})\beta_1, \frac{1}{2}\beta_2\Big\}$$

$$(2) \quad \Big\|\bar{R}(i_0)\Big\| < \min\Big\{2(\sqrt{2} - 1)\beta_1, \frac{\sqrt{2}}{2}\beta_2\Big\}.$$

Then the admissible control $\hat{\mathbf{u}} = (\hat{u}^1, \hat{u}^2, \ldots, \hat{u}^N)$, *where* $\hat{u}_t^i = -\big(N^i\big)^{-1} B^i(\alpha_{t-})^\intercal p_t^i$, $0 \leqslant t \leqslant T$, $1 \leqslant i \leqslant N$, *and* $\big(X_t, (p_t^1, q_t^1), \ldots, (p_t^N, q_t^N), \lambda_t^1, \ldots, \lambda_t^N\big)$ *is the solution of FBSDE (5.1), is a*

Nash equilibrium point of the conditional mean-field nonzero-sum quadratic stochastic differential game.

CHAPTER 6
Conclusion and Future Work

In this work we derive useful estimates for the solutions of the backward stochastic differential equations with Markovian switching and forward-backward stochastic differential equations with Markovian switching.

We also provide sufficient conditions for the existence and uniqueness of the solutions of the FBSDEs with regime-switching and FBSDEs with mean-field and regime-switching. For the FBSDEs with regime-switching we use two different approaches. The first approach is obtained in two steps. The first step, find existence and uniqueness of a solution over a small enough time duration. The second step, by using the connection with a system of PDEs and the local result, we can deduce the existence and uniqueness of a solution (under a non-degeneracy assumption) over an arbitrarily prescribed time duration. The second approach is used for FBSDEs with regime-switching and FBSDEs with mean-field and regime-switching, it concentrates on developing the continuation method and monotonicity conditions to examine the well-posedness of the systems.

Then we consider a nonzero-sum game problem with N players in which the dynamics and cost functionals of each player depend on conditional mean-field terms and a regime-switching process, presenting conditions on the coefficients such that a Nash equilibrium point of the differential game exists and the relationship of the existence of the Nash equilibrium point and the solution of the conditional mean-field FBSDE with regime switching.

Achievements:

We have extended and developed the theory of FBSDE for a very general class of equations that can capture the sudden jumps in dynamics as well as describe the limit of weakly interaction systems with both initial and terminal conditions. Several conditions for the well-posedness of these systems were given using different approaches. Finally, the results are used to solve a nonzero-sum game problem.

Future Works:

This work has opened new venues for the studies of conditional mean-field systems with regime switching and both initial and terminal conditions and their applications in modelling, control, and game problems. To further these researches there are several interesting and important problems for our on-going projects. For instance, while this thesis managed to generalize the conditions required to guarantee the existence and uniqueness of solutions to FBSDEs that admit conditional mean-field and Markovian-switching dynamics, there are practical needs consider some more general models which can also admit jumps or common noise. Also, while $\mathcal{L}^2$ solutions are widely studied and have had a lot of applications, there are also needs to generalize the results to $\mathcal{L}^p$ solutions. From application point of view, delay systems are very important in the real life. Hence, we also would like to address the issue of modeling delays in the system together with the impact of a Markovian switching or jump. Finally, we have a long plan to study numerical approximation for systems of FBSDEs. These directions will have immediate applications to solve control problems numerically in very general settings.

REFERENCE LIST

[1] Carmona R. and Delarue F. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer Nature, 2018.

[2] Chung, K.L. *Markov Chains*. Springer-Verlag, New York, 1967.

[3] Donnelly C. Sufficient stochastic maximum principle in a regime-switching diffusion model. *Appl. Math. Optim.* **64**, no. 2 (2011) 155–169.

[4] Framstad N.C., Oksendal B. and Sulem A. Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *J. Optim. Theory Appl.* **121**, no. 1 (2004) 77–98.

[5] Guo X., Pham H., and Wei X. Itô's formula for flow of measures on semimartingales. *arXiv preprint* arXiv:2010.05288 (2020).

[6] Hao T. and Li J. Mean-field SDEs with jumps and nonlocal integral-PDEs. *Nonlinear Differential Equations and Applications NoDEA* **23**, no. 2 (2016) 1-51.

[7] Hafayed M., Meherrem S., Eren S. and Guçoglu D.H. On optimal singular control problem for general Mckean-Vlasov differential equations: necessary and sufficient optimality conditions. *Optimal Control Appl. Methods*. **39**, no. 3 (2018) 1202–1219.

[8] Nguyen S.L., Nguyen D.T. and Yin G. A stochastic maximum principle for switching diffusions using conditional mean-fields with applications to control problems. *ESAIM: Control, Optimisation and Calculus of Variations* **26** (2020) pp.69.

[9] Nguyen S.L., Yin G. and Hoang T.A. On laws of large numbers for systems with mean-field interactions and Markovian switching. *Stochastic Processes and their Applications* **130**, no. 1 (2020) 262-296.

[10] Shao J., Tian T., and Wang S. Conditional McKean-Vlasov SDEs with jumps and Markovian regime-switching: Wellposedness, propagation of chaos, averaging principle. *preprint* (2022).

[11] F. Antonelli, Backward-forward stochastic differential equations. *Ann. Appl. Probab.* **3** (1993), 777–793.

[12] A. Bensoussan, J. Frehse, and P. Yam, *Mean Field Games and Mean Field Type Control Theory*. Springer, New York, 2013.

[13] A. Bensoussan, S.C.P. Yam, and Z. Zhang, Well-posedness of mean-field type forward-backward stochastic differential equations. *Stochastic Process. Appl.* **125** (2015) 3327–3354.

[14] R. Buckdahn, J. Li, and J. Ma, A mean-field stochastic control problem with partial observations. *Ann. Appl. Probab.* **27** (2017), 3201–3245.

[15] R. Carmona, Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications, *SIAM*, (2016).

[16] R. Carmona and F. Delarue, Mean field forward-backward stochastic differential equations. *Electron. Commun. Probab.* **18** (2013), no. 68, 15 pp.

[17] R. Carmona and X. Zhu, A probabilistic approach to mean field games with major and minor players. *Ann. Appl. Probab.* **26** (2016), 1535–1580.

[18] J.F. Chassagneux, H. Chotai, M. Muûls, *A Forward Backward SDEs Approach to Pricing in Carbon Markets*. Springer, Cham, 2017.

[19] Y. Chen, B. Djehiche,and S. Hamadène, Mean-field backward-forward stochastic differential equations and nonzero sum stochastic differential games. *Stoch. Dyn.* **21** (2021), no. 6, Paper No. 2150036, 32 pp.

[20] F. Delarue, On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Process. Appl.* **99** (2002), 209–286.

[21] C. Donnelly, Convex duality in constrained mean-variance portfolio optimization under a regime-switching model. Ph.D. Thesis (2008).

[22] C. Donnelly and A.J. Heunis, Quadratic risk minimization in a regime-switching model with portfolio constraints. *SIAM J. Control Optim.* **50** (2012), 2431–2461.

[23] D. Duffie and L. Epstein, Stochastic differential utility, *Econometrica* **60** (1992), 353–394.

[24] S. Hamadène, Backward-forward SDE's and stochastic differential games. *Stochastic Process. Appl.* **77** (1998), 1–15.

[25] Y. Hu and S. Peng, Solution of forward-backward stochastic differential equations. *Probab. Theory Related Fields* **103** (1995), 273–283

[26] M. Huang, P.E. Caines, and R.P. Malhamé, Individual and mass behavior in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. *Proc. 42nd IEEE CDC.* (2003), 98–103.

[27] M. Huang, R.P. Malhamé, and P.E. Caines, Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006), 221–252.

[28] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, *North-Holland Publishing Company* (1989).

[29] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, *Math. Finance* **7** (1997), 1–71.

[30] V.N. Kolokoltsov and M. Troeva, On the mean field games with common noise and the McKean-Vlasov SPDEs. *Stoch. Anal. Appl.* **37** (2019), 522–549.

[31] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Uralceva, *Linear and quasilinear equations of parapolic type*, Amer. Math. Soc., R. I. Providence, 1968.

[32] J.M. Lasry and P.L. Lions, Jeux à champ moyen. I. Le cas stationnaire. *C.R. Math. Acad. Sci. Paris* **343** (2006), 619–625.

[33] Y. Li and H. Zheng, Weak necessary and sufficient stochastic maximum principle for Markovian regime-switching diffusion models. *Appl. Math. Optim.* **71** (2015), 39–77.

[34] J. Ma, P. Protter, and J. Yong, Solving forward-backward stochastic differential equations explicitly – a four step scheme. *Probab. Theory Related Fields* **98** (1994), 339–359.

[35] J. Ma, Z. Wu, D. Zhang, and J. Zhang, On well-posedness of forward-backward SDEs – a unified approach. *Ann. Appl. Probab.* **25** (2015), 2168–2214.

[36] Ma, Ning; Wu, Zhen; Zhao, Huaizhong, *Backward stochastic differential equations with Markov chains and associated PDEs*, Journal of Differential Equations, **302**, (2021), 854–894.

[37] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*. Lecture Notes in Mathematics, 1702. Springer-Verlag, Berlin, 1999.

[38] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*. Imperial College Press, London, 2006.

[39] S.L. Nguyen and M. Huang, Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players, *SIAM J. Control Optim.* **50** (2012), 2907–2937.

[40] S.L. Nguyen, G. Yin, and D.T. Nguyen, A general stochastic maximum principle for mean-field controls with regime switching. *Appl. Math. Optim.* **84** (2021), 3255–3294.

[41] M. Nourian and P.E. Caines, ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM J. Control Optim.* **51** (2013), 3302–3331.

[42] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, Berlin, 2003.

[43] E. Pardoux and S. Tang, Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Related Fields* **114** (1999), 123–150.

[44] S. Peng and Z. Wu, Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.* **37** (1999), 825–843.

[45] H. Pham, *Continuous-Time Stochastic Control and Optimization with Financial Applications.* Springer-Verlag, Berlin, 2009

[46] H. Pham and X. Wei, Dynamic programming for optimal control of stochastic McKean–Vlasov dynamics. *SIAM J. Control Optim.* **55** (2017), 1069–1101.

[47] E. Platen and N. Bruti-Liberati, *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. Springer Berlin, Heidelberg, 2010

[48] Protter, P.E. *Stochastic Integration and Differential Equations*. Second edition. Stochastic Modelling and Applied Probability, **21**. Springer, Berlin, Heidelberg, 2005.

[49] Rong, S. *Theory of Stochastic Differential Equations with Jumps and Applications: Mathematical and Analytical Techniques with Applications to Engineering*. Springer, Science+Business Media, 2005.

[50] N. Şen and P.E. Caines, Nonlinear filtering theory for McKean-Vlasov type stochastic differential equations. *SIAM J. Control Optim.* **54** (2016), 153–174.

[51] R. Sh. Lipster and A. N. Shiryayev *Theory of Martingales* KLUWER ACADEMIC PUBLISHERS, 1989

[52] J. Shi, G. Wang , and J. Xiong, Leader-follower stochastic differential game with asymmetric information and applications. *Automatica* **63** (2016), 60–73.

[53] A.S., Sznitman, Topics in propagation of chaos. *Ecole d'Eté de Probabilités de Saint-Flour XIX* (1989) 165-–251

[54] Wu, Z, Fully coupled FBSDE with Brownian motion and Poisson process in stopping time duration. *J. Aust. Math. Soc.***74** (2003), 249-–266.

[55] Wu, Z. and Yu, Z, Probabilistic interpretation for a system of quasilinear parabolic partial differential equation combined with algebra equations. *Stochas- tic Process. Appl.***124** (2014), 3921-–3947.

[56] F. Xi and G. Yin, Asymptotic properties of a mean-field model with a continuous-state-dependent switching process. *J. Appl. Probab.* **46** (2009), 221–243.

[57] G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Two-Time-Scale Approach.* 2nd Ed., Springer, New York, 2013.

[58] G. Yin, G. Zhao, and F. Xi, Mean-field models involving continuous-state-dependent random switching: nonnegativity constraints, moment bounds, and two-time-scale limits. *Taiwanese J. Math.* **15** (2011), 1783–1805.

[59] G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications.* Springer, New York, 2010.

[60] J. Yong, Finding adapted solutions of forward–backward stochastic differential equations: Method of continuation. *Probab. Theory Related Fields* **107** (1997), 537–572.

[61] J. Yong, A linear-quadratic optimal control problem for mean-field stochastic differential equations. *SIAM J. Control Optim.* **51** (2013), 2809–2838.

[62] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations.* Springer-Verlag, New York, 1999

[63] Yu, Z., Equivalent cost functionals and stochastic linear quadratic optimal control problems. *ESAIM Control Optim. Calc. Var.***19** (2013), 78-–90.

[64] J. Zhang, The wellposedness of FBSDEs. *Discrete Contin. Dyn. Syst. Ser. B* **6** (2006), 927–940.

[65] L. Zhang and X. Li, Optimal control of Markov regime-switching stochastic recursive utilities. *arXiv preprint* (2019) `arXiv:1904.08811`.

Markovian switching systems: conditional McKean-Vlasov backward and forward-backward equations and their applications

Esteban J. Rolón Gutiérrez

esteban.rolon@upr.edu

Department of Mathematics

Chair: Son L. Nguyen Ph.D.

Degree: Doctor of Philosophy in Mathematics

Graduation Date: January 2024