République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
UNIVERSITÉ MOHAMED KHIDER, BISKRA
FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE
DÉPARTEMENT DE MATHÉMATIQUES.

THESE DE DOCTORAT EN SCIENCES
Option : **Mathématiques**

Par

Samira Boukaf

Titre

Sur un problem de contrôle optimal stochastique pour certain aspect des équations differentielles stochastiques de type mean-field et applications

---

Dr. Naceur Khelil, *MCA, Université de Biskra,* ______ **Président**
Dr. Mokhtar Hafayed, *MCA, Université de Biskra* ______ **Rapporteur**
Prof. Dahmen Achour, *Prof. Université de M'sila* ______ **Examinateur**
Dr. Saadi Khalil, *MCA, Université de Msila* ______ **Examinateur**
Dr. Boulakhras Gherbal , *MCA, Université de Biskra,* ______ **Examinateur**
Dr. Abdelmouman Tiaiba, *MCA. Université de M'sila* ______ **Examinateur**

2017

Contents

I dedicate this work to my family.

Acknowledgments

I would like to express my deepest gratitude to my advisor **Dr. Mokhtar Hafayed** not only because this work would have been not possible without his help, but above all because in this years he taught me with passion and patience the art of being a mathematician.

My sincere thanks to **Dr. Naceur Khelil** , Professor **Zouhir Mokhtari** , **Dr. Abdelmouman Tiaiba, Dr. Boulakhras Gherbal** and **Dr. Khalil Saadi** because they agreed to spend their times for reading and evaluating my thesis.
I thank all my colleagues of the Mathematics Department and Economics department in University Hamma lakhdar of El Oued, Algeria.

Résumé

Cette thèse de doctorat s'inscrit dans le cadre de l'analyse stochastique dont le thème central est: les conditions nécessaires et suffisantes sous forme du maximum stochastique de type champ moyen d'optimalité et de presque optimalité et ces applications. L'objectif de ce travail est d'étudier des problèmes d'optimisation stochastique. Il s'agira ensuite de faire le point sur les conditions nécessaires et suffisantes d'optimalité et de presque optimalité pour un system gouverné par des équations différentielles stochastiques de type champ moyen. Cette thèse s'articule autour de quatre chapitres:

Le chapitre 1 est essentiellement un rappel. La candidate présente quelques concepts et résultats qui lui permettent d'aborder son travail; tels que les processus stochastiques, l'espérance conditionnelle, les martingales, les formules d'Ito, les classes de contrôle stochastique, ... etc.

Dans le deuxième chapitre, on a établie et on a prouvé les conditions nécessaires et suffisantes de presque optimalité d'order 3b5{ 3bb} vérifiées par un contrôle optimal stochastique, pour un system différentiel gouverné par des équations différentielles stochastiques EDSs. Le domaine de contrôle stochastique est supposé convexe. La méthode utilisée est basée sur le lemme d'Ekeland. Les résultats obtenus dans le chapitre 2, sont tous nouveaux et font l'objet d'un premier article intitulé :

Boukaf Samira & Mokhtar Hafayed, & Ghebouli Messaoud: A study on optimal control problem with ε^{λ}-error bound for stochastic systems with application to linear quadratic problem, International Journal of Dynamics and Control, Springer DOI: 10.1007/s40435-015-0178-x (2015).

Dans le troisième chapitre, on a démontré le principe du maximum stochastique de presque optimalité, où le system est gouverné par des equations différentielles stochastiques progressive rétrogrades avec saut (FBSDEs). Ces resultats ont été appliqués pour résoudre un problème d'optimisation en finance. Ces resultats généralisent le principe du maximum de Zhou (SIAM. Control. Optim. (36)-3, 929-947 (1998)). Les résultats obtenus dans le chapitre 3 sont tous nouveaux et font l'objet d'un deuxième article intitulé:

Mokhtar Hafayed, & Abdelmadjid Abba & Samira Boukaf: On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications International Journal of Modelling, Identification and Control . 25 (1), 1-16, (2016).

De plus, et dans le chapitre 4, on a prouvé un principe du maximum stochastique de type de Pontryagin pour des systems gouvernés par FBSDEs avec saut. Ces resultats ont été établi avec M. Hafayed, et M. Tabet, sous le titre :

Mokhtar Hafayed, & Moufida Tabet & Samira Boukaf: Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem, Communication in Mathematics and Statistics, Springer, Doi: 10.1007/s40304-015-0054-1, Volume 3, Issue 2, pp 163-186 (2015) .

Dans le chapitre 5, on a aborde un probléme de contrôle singulier, où le problème est d'établir des conditions nécessaires et suffisantes d'optimalité pour un control singulier ou le system est gouverné par des équations différentielles stochastiques progressive rétrograde de type McKean-Vlasov. Dans ces cas, le domaine de contrôle admissible est supposé convexe. Les résultats obtenus dans le chapitre 5 sont tous nouveaux et font l'objet d'un article intitulé :

Mokhtar Hafayed, & Samira Boukaf & Yan Shi, & Shahlar Meherrem.: A McKean-Vlasov optimal mixed regular-singular control problem, for nonlinear stochastic systems with Poisson jump processes, Neurocomputing. Doi 10.1016/j.neucom.2015.11.082, Volume 182, 19, pages 133-144 (2016).

Abstract

This thesis is concerned with stochastic control of mean-field type. The central theme is the necessary and sufficient conditions in the form of the Pontryagin's stochastic maximum of the mean-field type for optimality and for near-optimality with some applications. Recently, the main purpose of this thesis is to derive a set of necessary as well as sufficient conditions of optimality and near optimality, where the system is governed by stochastic differential equations of the mean field type. This thesis is structured around four chapters:

Chapter 1 is essentially a reminder. we presents some concepts and results that allow us to prove our results, such as stochastic processes, conditional expectation, martingales, Ito formulas, class of stochastic control, etc. In the second chapter, we have proved the necessary and sufficient conditions of near-optimality of order 3b5 { 3bb} satisfied by an optimal stochastic control, where the system is governed by stochastic differential equations EDSs. The stochastic control domain is assumed to be convex. The method used is based on the Ekeland lemma. The results obtained in Chapter 2 are all new and are the subject of a first article entitled:

Boukaf Samira & Mokhtar Hafayed and Ghebouli Messaoud: A study on optimal control problem with ε^{λ}-error *bound for stochastic systems with application to linear quadratic problem, International Journal of Dynamics and Control, Springer DOI: 10.1007 / s40435-015-0178-x (2015).*

In the third chapter, we have proved the stochastic maximum principle of near-optimality, where the system is governed by forward backward stochastic differential equations (FBSDEs). These results have been applied to solve an optimization problem in finance. These results generalize the Zhou's maximum principle (SIAM, Control, Optim (36) -3, 929-947 (1998)). The results obtained in Chapter 3 are all new and are the subject of a second article entitled:

Mokhtar Hafayed, & Abdelmadjid Abba & Samira Boukaf: On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications. 25 (1), 1-16, (2016).

Moreover, and in Chapter 4, we have proved a Pontryagin type stochastic maximum principle for systems governed by FBSDEs with jumps. These results have been established with M. Hafayed, and M. Tabet, under the title

Mokhtar Hafayed, & Moufida Tabet & Samira Boukaf: Mean-field maximal for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem, Communication in Mathematics and Statistics, Springer, Doi: 10.1007 / s40304- 015-0054-1, Volume 3, Issue 2, pp 163-186 (2015).

In Chapter 5, we have addressed a singular control problem, where the problem is to establish a set of necessary and sufficient conditions of optimality for a singular control for a system governed by forward backward stochastic differential equations of McKean-Vlasov type. In these cases, the control domain is assumed to be convex. An application to finance is given to ullistrated our new results The results obtained in Chapter 5 are all new and are the subject of an article entitled:

Mokhtar Hafayed, & Samira Boukaf & Yan Shi, & Shahlar Meherrem: A McKean-Vlasov optimal mixed regular-singular control problem for nonlinear stochastic systems with Poisson jump processes, Neurocomputing. Doi 10.1016 / j.neucom.2015.11.082, Volume 182, 19, 133-144 (2016).

Symbols and Acronyms

- **a.e**. almost everywhere
- **a.s.** almost surely
- **càdlàg** continu à droite, limite à gauche
- **cf**. compare (abbreviation of Latin confer)
- **e.g**. for example (abbreviation of Latin exempli gratia)
- **i.e,**. that is (abbreviation of Latin id est)
- **HJB** *The Hamilton-Jacobi-Bellman equation*
- **SDE**: Stochastic differential equations.
- **BSDE:** Backward stochastic differential equation.
- **FBSDEs**: Forward-backward stochastic differential equations.
- **FBSDEJs**: Forward-Backward stochastic differential equations with jumps.
- **PDE**: Partial differential equation.
- **ODE:** Ordinary differential equation.
- $\mathbb{R}$: Real numbers.
- $\mathbb{R}_+$: Nonnegative real numbers.
- $\mathbb{N}$: Natural numbers.
- $\dfrac{\partial f}{\partial x}, f_x$: The derivatives with respect to x.
- $\mathbb{P}\otimes dt$: The product measure of $\mathbb{P}$ with the Lebesgue measure dt on $[0,T]$.
- $E(\cdot), E(\cdot \mid G)$ Expectation; conditional expectation
- $\sigma(A)$: $\sigma-$algebra generated by A.
- I_A : Indicator function of the set A.
- $\mathcal{F}^X$: The filtration generated by the process X.
- $W(\cdot), B(\cdot)$: Brownian motions
- $\mathcal{F}_t^B$ the natural filtration generated by the brownian motion $B(\cdot)$,
- $\mathrm{F}_1 \vee \mathrm{F}_2$ denotes the σ-field generated by $\mathrm{F}_1 \cup \mathrm{F}_2$.

Part I
Introduction

Optimal control theory can be described as the study of strategies to optimally influence a system x with dynamics evolving over time according to a differential equation. The influence on the system is modeled as a vector of parameters, u, called the control. It is allowed to take values in some set U, which is known as the action space. For a control to be optimal, it should minimize a cost functional (or maximize a reward functional), which depends on the whole trajectory of the system x and the control u over some time interval $[0, T]$. The infimum of the cost functional is known as the value function (as a function of the initial time and state). This minimization problem is infinite dimensional, since we are minimizing a functional over the space of functions $u(t), t \in [0, T]$. Optimal control theory essentially consists of different methods of reducing the problem to a less transparent, but more manageable problem.

1. Formulations of stochastic optimal control problems

In this section, we present two mathematical formulations (strong and weak formulations) of stochastic optimal control problems in the following two subsections, respectively.

1.1. Strong formulation

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ be a given filtered probability space satisfying the usual condition, on which an d-dimensional standard Brownian motion $W(\cdot)$ is defined, consider the following controlled stochastic differential equation:

$$\begin{cases} dx(t) &= f(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \\ x(0) &= x_0 \in \mathbb{R}^n, \end{cases} \tag{1}$$

where

$$f : [0, T] \times \mathbb{R}^n \times A \longrightarrow \mathbb{R}^n,$$
$$\sigma : [0, T] \times \mathbb{R}^n \times A \longrightarrow \mathbb{R}^{n\times d},$$

and $x(\cdot)$ is the variable of state.

The function $u(\cdot)$ is called the control representing the action of the decision-makers (controller). At any time instant the controller has some information (as specified by the information field $\{\mathcal{F}_t\}_{t\in[0,T]}$) of what has happened up to that moment, but not able to foretell what is going to happen afterwards due to the uncertainty of the system (as a consequence, for any t the controller cannot exercise his/her decision $u(t)$ before the time t really comes), This nonanticipative restriction in mathematical terms can be expressed as "$u(\cdot)$ is $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted".

The control $u(\cdot)$ is an element of the set

$$\mathcal{U}[0,T]=\{u:[0,T]\times\Omega\longrightarrow A \text{ such that } u(\cdot) \text{ is } \{\mathcal{F}_t\}_{t\in[0,T]}-\text{adapted}\}.$$

We introduce the cost functional as follows

$$J(u(\cdot)) \doteq E\left[\int_0^T l(t,x(t),u(t))dt+g(x(T))\right], \tag{2}$$

where

$$l:[0,T]\times\mathbb{R}^n\times A\longrightarrow\mathbb{R},$$
$$g:\mathbb{R}^n\longrightarrow\mathbb{R}.$$

Definition 1.1. Let $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$ be given satisfying the usual conditions and let $W(t)$ be a given d-dimensional standard $\{\mathcal{F}_t\}_{t\in[0,T]}$-Brownian motion. A control $u(\cdot)$ is called an admissible control, and $(x(\cdot),u(\cdot))$ an admissible pair, if

i) $u(\cdot)\in\mathcal{U}[0,T]$;

ii) $x(\cdot)$ is the unique solution of equation (25);

iii) $l(\cdot,x(\cdot),u(\cdot))\in\mathbb{L}^1_{\mathcal{F}}([0,T];\mathbb{R})$ and $g(x(T))\in\mathbb{L}^1_{\mathcal{F}_T}(\Omega;\mathbb{R})$.

The set of all admissible controls is denoted by $\mathcal{U}([0,T])$. Our stochastic optimal control problem under strong formulation can be stated as follows:
Problem 1.1 Minimize (26) over $\mathcal{U}([0,T])$. The goal is to find $u^*(\cdot)\in\mathcal{U}([0,T])$, such that

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}([0,T])} J(u(\cdot)). \tag{3}$$

For any $u^*(\cdot)\in\mathcal{U}^s([0,T])$ satisfying (27) is called an strong optimal control. The corresponding state process $x^*(\cdot)$ and the state control pair $(x^*(\cdot),u^*(\cdot))$ are called an strong optimal state process and an strong optimal pair, respectively.

1.2. Weak formulation

In stochastic control problems, there exists for the optimal control problem another formulation of a more mathematical aspect, it is the weak formulation of the stochastic optimal control problem. Unlike in the strong formulation the filtered probability space $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$ on which we define the Brownian motion $W(\cdot)$ are all fixed, but it is not the case in the weak formulation, where we consider them as a parts of the control.

Definition 1.2. A 6-tuple $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P},W(\cdot),u(\cdot)\right)$ is called weak-admissible control and $(x(\cdot),u(\cdot))$ an weak admissible pair, if

1. $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$ is a filtered probability space satisfying the usual conditions;

2. $W(\cdot)$ is an d-dimensional standard Brownian motion defined on $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$;

3. $u(\cdot)$ is an $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted process on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in U;

4. $x(\cdot)$ is the unique solution of equation (25),

5. $l(\cdot, x(\cdot), u(\cdot)) \in \mathbb{L}^1_{\mathcal{F}}([0,T];\mathbb{R})$ and $g(x(T)) \in \mathbb{L}^1_{\mathcal{F}}(\Omega;\mathbb{R})$.

The set of all weak admissible controls is denoted by $\mathcal{U}^w([0,T])$. Sometimes, might write $u(\cdot)) \in \mathcal{U}^w([0,T])$ instead of $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, W(\cdot), u(\cdot)\right) \in \mathcal{U}^w([0,T])$.
Our stochastic optimal control problem under weak formulation can be formulated as follows:

Problem 1.2. The objective is to minimize the cost functional given by equation (26) over the of admissible controls $\mathcal{U}^w([0,T])$. Namely, one seeks $\pi^*(\cdot) = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, W(\cdot), u(\cdot)\right) \in \mathcal{U}^w([0,T])$ such that

$$J(\pi^*(\cdot)) = \inf_{\pi(\cdot)\in\mathcal{U}^w([0,T])} J(\pi(\cdot)).$$

2. Methods to solving optimal control problem

Two major tools for studing optimal control are Bellman's dynamic programming method and Pontryagin's maximum principle.

2.1. The Dynamic Programming Principle.

We present an approach to solving optimal control problems, namely, the method of dynamic programming. Dynamic programming, originated by R. Bellman (*Bellman, R.: Dynamic programming, Princeton Univ. Press., (1957)*) is a mathematical technique for making a sequence of interrelated decisions, which can be applied to many optimization problems (including optimal control problems). The basic idea of this method applied to optimal controls is to consider a family of optimal control problems with different initial times and states, to establish relationships among these problems via the so-called Hamilton-Jacobi-Bellman equation (HJB, for short), which is a nonlinear first-order (in the deterministic case) or second-order (in the stochastic case) partial differential equation. If the HJB equation is solvable (either analytically or numerically), then one can obtain an optimal feedback control by taking the maximize/minimize of the Hamiltonian or generalized Hamiltonian involved in the HJB equation. This is the so-called verification technique. Note that this approach actually gives solutions to the whole family of problems (with different initial times and states).

The Bellman principle. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, satisfying the usual conditions, $T > 0$ a finite time, and W a d-dimensional Brownian motion defined on the filtered probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\in[0,T]}\right)$.
We consider the state stochastic differential equation

$$dx(s) = f(s, x(s), u(s))ds + \sigma(s, x(s), u(s))dW(s), \text{ s} \in [0,T] \quad (4)$$

The control $u = u(s)_{0\le s\le T}$ is a progressively measurable process valued in the control set U, a subset of $\mathbb{R}^k$, satisfies a square integrability condition. We denote by $\mathcal{U}([t,T])$ the set of control processes u.

Conditions. To ensure the existence of the solution to SDE-(**??**), the Borelian functions

$$\begin{aligned} f &: [0,T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^n \\ \sigma &: [0,T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^{n\times d} \end{aligned}$$

satisfy the following conditions:

$$|f(t,x,u) - f(t,y,u)| + |\sigma(t,x,u) - \sigma(t,y,u)| \leq C\,|x-y|\,,$$

$$|f(t,x,u)| + |\sigma(t,x,u)| \leq C\,[1+|x|]\,,$$

for some constant $C > 0$. We define the gain function as follows:

$$J(t,x,u) \;=\; E\left[\int_t^T l(s,x(s),u(s))ds + g(x\,(T))\right], \tag{5}$$

where

$$\begin{aligned} l &: [0,T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}, \\ g &: \mathbb{R}^n \longrightarrow \mathbb{R}, \end{aligned}$$

be given functions. We have to impose integrability conditions on f and g in order for the above expectation to be well-defined, e.g. a lower boundedness or quadratic growth condition. The objective is to maximize this gain function. We introduce the so-called value function:

$$V(t,x) \;=\; \sup_{u\in\mathcal{U}([t,T])} J(t,x,u), \tag{6}$$

where $x(t) = x$ is the initial state given at time t. For an initial state (t,x), we say that $u^* \in \mathcal{U}\,([t,T])$ is an optimal control if

$$V(t,x) = J(t,x,u^*).$$

Theorem 1.1. Let $(t,x) \in [0,T] \times \mathbb{R}^n$ be given. Then we have

$$V(t,x) = \sup_{u\in\mathcal{U}([t,T])} E\left[\int_t^{t+h} l(s,x(s),u(s))dt + V(t+h,x(t+h))\right], \text{ for } t \leq t+h \leq T. \tag{7}$$

Proof. The proof of the dynamic programming principle is technical and has been studied by different methods, we refer the reader to Yong and Zhou [69].

The Hamilton-Jacobi-Bellman equation The HJB equation is the infinitesimal version of the dynamic programming principle. It is formally derived by assuming that the value function is $C^{1,2}\,([0,T] \times \mathbb{R}^n)$, applying Itô's formula to $V(s, x^{t,x}(s))$ between $s = t$ and $s = t+h$, and then sending h to zero into (6). The classical HJB equation associated to the stochastic control problem (6) is

$$-V_t(t,x) - \sup_{u\in U}\left[\mathcal{L}^u V(t,x) + l(t,x,u)\right] \;=\; 0, \text{ on } [0,T] \times \mathbb{R}^n, \tag{8}$$

where $\mathcal{L}^u$ is the second-order infinitesimal generator associated to the diffusion x with control u

$$\mathcal{L}^u V = f(x,u).D_x V + \frac{1}{2} tr\left(\sigma(x,u)\sigma^\intercal(x,u) D_x^2 V\right).$$

This partial differential equation (PDE) is often written also as:

$$-V_t(t,x) - H(t,x,D_x V(t,x), D_x^2 V(t,x)) = 0, \ \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{9}$$

where for $(t,x,\Psi,Q) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$ ($\mathcal{S}_n$ is the set of symmetric $n \times n$ matrices) :

$$H(t,x,\Psi,Q) = \sup_{u \in U}\left[f(t,x,u).\Psi + \frac{1}{2} tr\left(\sigma\sigma^\intercal(t,x,u) Q\right) + l(t,x,u)\right]. \tag{10}$$

The function H is sometimes called Hamiltonian of the associated control problem, and the PDE (8) or (9) is the dynamic programming or HJB equation.
There is also an a priori terminal condition:

$$V(T,x) = g(x), \ \forall x \in \mathbb{R}^n,$$

which results from the very definition of the value function V.

The classical verification approach The classical verification approach consists in finding a smooth solution to the HJB equation, and to check that this candidate, under suitable sufficient conditions, coincides with the value function. This result is usually called a verification theorem and provides as a byproduct an optimal control. It relies mainly on Itô's formula. The assertions of a verification theorem may slightly vary from problem to problem, depending on the required sufficient technical conditions. These conditions should actually be adapted to the context of the considered problem. In the above context, a verification theorem is roughly stated as follows:

Theorem 1.2. Let W be a $C^{1,2}$ function on $[0,T] \times \mathbb{R}^n$ and continuous in T, with suitable growth condition. Suppose that for all $(t,x) \in [0,T] \times \mathbb{R}^n$, there exists $u^*(t,x)$ mesurable, valued in U such that W solves the HJB equation:

$$\begin{aligned} 0 &= -W_t(t,x) - \sup_{u \in U}\left[\mathcal{L}^u W(t,x) + l(t,x,u)\right] \\ &= -W_t(t,x) - \mathcal{L}^{u^*(t,x)} W(t,x) - l(t,x,u^*(t,x)), \ \text{on } [0,T] \times \mathbb{R}^n, \end{aligned}$$

together with the terminal condition $W(T,\cdot) = g$ on $\mathbb{R}^n$, and the stochastic differential equation:

$$dx(s) = f(s,x(s),u^*(s,x(s)))ds + \sigma(s,x(s),u^*(s,x(s)))dW(t),$$

admits a unique solution x^*, given an initial condition $x(t) = x$. Then, $W = V$ and $u^*(s,x^*)$ is an optimal control for $V(t,x)$.

A proof of this verification theorem can be found in book, by Yong & Zhou [69].

2.2. The Pontryagin's maximum principle

The pioneering works on the stochastic maximum principle were written by Kushner [37, 38]. Since then there have been a lot of works on this subject, among them, in particular, those by Bensoussan [6], Peng [40], and so on. The stochastic maximum principle gives some necessary conditions for optimality for a stochastic optimal control problem.
The original version of Pontryagin's maximum principle was first introduced for deterministic control problems in the 1960's by Pontryagin et al. (Pontryagin,L.S., Boltyanski,V.G., Gamkrelidze, R.V., Mischenko, E.F.) [1] as in classical calculus of variation. The basic idea is to perturbe an optimal control and to use some sort of Taylor expansion of the state trajectory around the optimal control, by sending the perturbation to zero, one obtains some inequality, and by duality.

The maximum principle. As an illustration, we present here how the maximum principle for a deterministic control problem is derived. In this setting, the state of the system is given by the ordinary differential equation (ODE) of the form

$$\begin{cases} dx(t) &= f(t,x(t),u(t))dt,\ t\in[0,T]\,, \\ x(0) &= x_0, \end{cases} \tag{11}$$

where

$$f:[0,T]\times\mathbb{R}\times\mathcal{A}\longrightarrow\mathbb{R},$$

and the action space $\mathcal{A}$ is some subset of $\mathbb{R}$. The objective is to minimize some cost function of the form:

$$J(u(\cdot)) \;=\; \textstyle\int_0^T l(t,x(t),u(t))+g(x(T)), \tag{12}$$

where

$$l:[0,T]\times\mathbb{R}\times\mathcal{A}\longrightarrow\mathbb{R},$$
$$g:\mathbb{R}\longrightarrow\mathbb{R}.$$

That is, the function l inflicts a running cost and the function g inflicts a terminal cost. We now assume that there exists a control $u^*(t)$ which is optimal, i.e.

$$J(u^*(\cdot)) \;=\; \inf_u J(u(\cdot)).$$

We denote by $x^*(t)$ the solution to (11) with the optimal control $u^*(t)$. We are going to derive necessary conditions for optimality, for this we make small perturbation of the optimal control. Therefore we introduce a so-called spike variation, i.e. a control which is equal to u^* except on some small time interval:

$$u^\varepsilon(t)=\begin{cases} v & \text{for } \tau-\varepsilon\le t\le\tau, \\ u^*(t) & \text{otherwise.}\end{cases} \tag{13}$$

---

[1] Pontryagin,L.S., Boltyanski,V.G., Gamkrelidze,R.V., Mischenko, E.F. Mathematical Theory of Optimal Processes, Wiley, New York, 1962.

We denote by $x^\varepsilon(t)$ the solution to (11) with the control $u^\varepsilon(t)$. We set that $x^*(t)$ and $x^\varepsilon(t)$ are equal up to $t = \tau - \varepsilon$ and that

$$\begin{aligned} x^\varepsilon(\tau) - x^*(\tau) &= (f(\tau, x^\varepsilon(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)))\varepsilon + o(\varepsilon) \\ &= (f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)))\varepsilon + o(\varepsilon), \end{aligned} \tag{14}$$

where the second equality holds since $x^\varepsilon(\tau) - x^*(\tau)$ is of order ε. We look at the Taylor expansion of the state with respect to ε. Let

$$z(t) = \frac{\partial}{\partial \varepsilon} x^\varepsilon(t) \mid_{\varepsilon=0},$$

i.e. the Taylor expansion of $x^\varepsilon(t)$ is

$$x^\varepsilon(t) = x^*(t) + z(t)\varepsilon + o(\varepsilon). \tag{15}$$

Then, by (14)

$$z(\tau) = f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)). \tag{16}$$

Moreover, we can derive the following differential equation for $z(t)$.

$$\begin{aligned} dz(t) &= \frac{\partial}{\partial \varepsilon} dx^\varepsilon(t) \mid_{\varepsilon=0} \\ &= \frac{\partial}{\partial \varepsilon} f(t, x^\varepsilon(t), u^\varepsilon(t))dt \mid_{\varepsilon=0} \\ &= f_x(t, x^\varepsilon(t), u^\varepsilon(t)) \frac{\partial}{\partial \varepsilon} x^\varepsilon(t)dt \mid_{\varepsilon=0} \\ &= f_x(t, x^*(t), u^*(t))z(t)dt, \end{aligned}$$

where f_x denotes the derivative of f with respect to x. If we for the moment assume that $l = 0$, the optimality of $u^*(t)$ leads to the inequality

$$\begin{aligned} 0 \le \frac{\partial}{\partial \varepsilon} J(u^\varepsilon)\bigg|_{\varepsilon=0} &= \frac{\partial}{\partial \varepsilon} g(x^\varepsilon(T)) \mid_{\varepsilon=0} \\ &= g_x(x^\varepsilon(T)) \frac{\partial}{\partial \varepsilon} x^\varepsilon(T) \mid_{\varepsilon=0} \\ &= g_x(x^*(T))\, z(T). \end{aligned}$$

We shall use duality to obtain a more explicit necessary condition from this. To this end we introduce the adjoint equation:

$$\begin{cases} d\Psi(t) = -f_x(t, x^*(t), u^*(t))\Psi(t)dt, t \in [0, T], \\ \Psi(T) = g_x(x^*(T)). \end{cases}$$

Then it follows that

$$d(\Psi(t)z(t)) = 0,$$

i.e. $\Psi(t)z(t)) =$ constant. By the terminal condition for the adjoint equation we have

$$\Psi(t)z(t) \quad = \quad g_x(x^*(T))z(T) \geq 0, \text{ for all } 0 \leq t \leq T.$$

In particular, by (16)

$$\Psi(\tau)\left(f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau))\right) \geq 0.$$

Since τ was chosen arbitrarily, this is equivalent to

$$\Psi(t)f(t, x^*(t), u^*(t)) \quad = \quad \inf_{v \in \mathcal{U}} \Psi(t)f(t, x^*(t), v), \text{ for all } 0 \leq t \leq T.$$

By repeating the calculations above for this two-dimensional system, one can derive the necessary condition

$$H(t, x^*(t), u^*(t), \Psi(t)) \quad = \quad \inf_{v} H(t, x^*(t), v, \Psi(t)) \text{ for all } 0 \leq t \leq T, \tag{17}$$

where H is the so-called Hamiltonian (sometimes defined with a minus sign which turns the minimum condition above into a maximum condition) :

$$H(x, u, \Psi) \quad = \quad l(x, u) + \Psi f(x, u),$$

and the adjoint equation is given by

$$\begin{cases} d\Psi(t) & = & -(l_x(t, x^*(t), u^*(t)) + f_x(t, x^*(t), u^*(t))\Psi(t))dt, \\ \Psi(T) & = & g_x(x^*(T)). \end{cases} \tag{18}$$

The minimum condition (17) together with the adjoint equation (18) specifies the Hamiltonian system for our control problem.

The stochastic maximum principle. Stochastic control is the extension of optimal control to problems where it is of importance to take into account some uncertainty in the system. One possibility is then to replace the differential equation by an SDE:

$$dx(t) \quad = \quad f(t, x(t), u(t))dt + \sigma(t, x(t))dW(t), t \in [0, T], \tag{19}$$

where f and σ are deterministic functions and the last term is an Itô integral with respect to a Brownian motion W defined on a probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$.
More generally, the diffusion coefficient σ may has an explicit dependence on the control: $t \in [0, T]$.

$$dx(t) \quad = \quad f(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \tag{20}$$

The cost function for the stochastic case is the expected value of the cost function (**??**), i.e. we want to minimize

$$J(u(\cdot)) = E\left[\int_0^T l(t, x(t), u(t)) + g(x(T))\right].$$

For the case (19) the adjoint equation is given by the following Backward SDE:

$$\begin{cases} -d\Psi(t) & = & \{f_x(t,x^*(t),u^*(t))\Psi(t) + \sigma_x(t,x^*(t))Q(t) \\ & & +(l_x(t,x^*(t),u^*(t))\}dt - Q(t)dW(t), \\ \Psi(T) & = & g_x(x^*(T)). \end{cases} \tag{21}$$

A solution to this backward SDE is a pair $(\Psi(t), Q(t))$ which fulfills (21). The Hamiltonian is

$$H(x,u,\Psi(t),Q(t)) \;\; = \;\; l(t,x,u) + \Psi(t)f(t,x,u) + Q(t)\sigma(t,x),$$

and the maximum principle reads for all $0 \leq t \leq T$,

$$H(t,x^*(t),u^*(t),\Psi(t),Q(t)) \;\; = \;\; \inf_{u\in\mathcal{U}} H(t,x^*(t),u,\Psi(t),Q(t)) \;\; \mathbb{P}-\text{a.s.} \tag{22}$$

Noting that there is also third case: if the state is given by (20) but the action space $\mathcal{A}$ is assumed to be convex, it is possible to derive the maximum principle in a local form. This is accomplished by using a convex perturbation of the control instead of a spike variation, see Bensoussan 1983 [6]. The necessary condition for optimality is then given by the following: for all $0 \leq t \leq T$

$$E\int_0^T H_u(t,x^*(t),u^*(t),\Psi^*(t),Q^*(t))\,(u-u^*(t)) \geq 0.$$

3. Some classes of stochastic controls

Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ be a complete filtred probability space.

1. Admissible control An admissible control is $\mathcal{F}_t$-adapted process $u(t)$ with values in a borelian $A \subset \mathbb{R}^n$

$$\mathcal{U} := \{u(\cdot) : [0,T]\times\Omega \to A : u(t) \text{ is } \mathcal{F}_t\text{-adapted}\}\,. \tag{23}$$

2. Optimal control The optimal control problem consists to minimize a cost functional $J(u)$ over the set of admissible control $\mathcal{U}$. We say that the control $u^*(\cdot)$ is an optimal control if

$$J(u^*(t)) \leq J(u(t)), \text{ for all } u(\cdot) \in \mathcal{U}.$$

3. Near-optimal control Let $\varepsilon > 0$, a control is a near-optimal control (or ε-optimal) if for all control $u(\cdot) \in \mathcal{U}$ we have

$$J(u^\varepsilon(t)) \leq J(u(t)) + \varepsilon. \tag{24}$$

See for some applications.

4. Singular control An admissible control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathcal{F}_t-$adapted processes, such that $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the left with right limits and $\xi(0_-) = 0$. Since $d\xi(t)$ may be singular with respect to Lebesgue measure dt, we call $\xi(\cdot)$ the singular part of the control and the process $u(\cdot)$ its absolutely continuous part.

5. Feedback control: We say that $u(\cdot)$ is a feedback control if $u(\cdot)$ depends on the state variable $X(\cdot)$. If $\mathcal{F}_t^X$ the natural filtration generated by the process X, then $u(\cdot)$ is a feedback control if $u(\cdot)$ is $\mathcal{F}_t^X-$adapted.

6. Impulsive control. Impulse control: Here one is allowed to reset the trajectory at stopping times (τ_i) from $X_{\tau_{i-}}$ (the value immediately before i) to a new (non-anticipative) value X_{τ_i}, resp., with an associated cost $L\left(X_{\tau_{i-}}, X_{\tau_i}\right)$. The aim of the controlled is to minimizes the cost functional:

$$\begin{aligned}
&E\int_0^T \exp\left[-\int_0^t C(X(s),u(s))ds\right] K(X(t),u(t)) \\
&+\sum_{\tau_i<T} \exp\left[-\int_0^{\tau_i} C(X(s),u(s))ds\right] g(X_\tau, X_{\tau_{i-}}) \\
&+\exp\left[-\int_0^{\tau_i} C(X(s),u(s))ds\right] h(X(T)).
\end{aligned}$$

7. Ergodic control Some stochastic systems may exhibit over a long period a stationary behavior characterized by an invariant measure. This measure, if it does exists, is obtained by the average of the states over a long time. An ergodic control problem consists in optimizing over the long term some criterion taking into account this invariant measure. (See Pham [47], Borkar [11]). The cost functional is given by

$$\lim \sup_{T\to+\infty} \frac{1}{T} E\int_0^T f(x(t),u(t))dt.$$

8. Robust control In the problems formulated above, the dynamics of the control system is assumed to be known and fixed. Robust control theory is a method to measure the performance changes of a control system with changing system parameters. This is of course important in engineering systems, and it has recently been used in finance in relation with the theory of risk measure. Indeed, it is proved that a coherent risk measure for an uncertain payoff $x(T)$ at time T is represented by :

$$\rho(-X(t)) = \sup_{Q\in S} E^Q(X(T)),$$

where S is a set of absolutly continuous probability measures with respect to the original probability P.

9. Partial observation control problem It is assumed so far that the controller completely observes the state system. In many real applications, he is only able to observe partially the state via other variables and there is noise in the observation system. For example in financial models, one may observe the asset price but not completely its rate of return and/or its volatility, and the portfolio investment is based only on the asset price information. We are facing a partial observation control problem. This may be formulated in a general form as follows : we have a controlled signal (unobserved) process governed by the following SDE:

$$dx(t) = f(t, x(t), y(t), u(t))\,dt + \sigma(t, x(t), y(t), u(t))\,dW(t),$$

and

$$dy\left(t\right)=g\left(t,x\left(t\right),y\left(t\right),u\left(t\right)\right)dt+h\left(t,x\left(t\right),y\left(t\right),u\left(t\right)\right)dB\left(t\right),$$

where $B\left(t\right)$ is another Brownian motion, eventually correlated with $W\left(t\right)$. The control $u(t)$ is adapted with respect to the filtration generated by the observation F_t^Y and the functional to optimize is :

$$J\left(u\left(\cdot\right)\right)=E\left[h\left(x\left(T\right),y(T)\right)+\int_0^T g\left(t,x\left(t\right),y(t),u\left(t\right)\right)dt\right].$$

10. Random horizon In classicla problem, the time horizon is fixed until a deterministic terminal time T. In some real applications, the time horizon may be random, the cost functional is given by the following:

$$J\left(u\left(\cdot\right)\right)=E\left[h\left(x\left(\tau\right)\right)+\int_0^\tau g\left(t,x\left(t\right),y(t),u\left(t\right)\right)dt\right],$$

where τ s a finite random time.

11. Relaxed control The idea is then to compactify the space of controls $\mathcal{U}$ by extending the definition of controls to include the space of probability measures on U. The set of relaxed controls $\mu_t\left(du\right)dt$, where μ_t is a probability measure, is the closure under weak* topology of the measures $\delta_{u(t)}(du)dt$ corresponding to usual, or strict, controls. This notion of relaxed control is introduced for deterministic optimal control problems in Young (*Young, L.C. Lectures on the calculus of variations and optimal control theory, W.B. Saunders Co., 1969.*) (See Borkar [11]).

A study on optimal control problem with $\varepsilon^{\lambda}-$error bound for stochastic systems with applications to linear quadratic problem

Part II

A study on optimal control problem with $\varepsilon^{\lambda}-$error bound for stochastic systems with applications to linear quadratic problem

Abstract. In this part, we study near-optimal stochastic control problem with $\varepsilon^{\lambda}-$error bound for systems governed by nonlinear controlled Itô stochastic differential equations (SDEs in short). The control is allowed to enter into both drift and diffusion coefficients and the control domain need be convex. The proof of our main result is based on Ekeland's variational principle and some approximation arguments on the state variable and adjoint process with respect to the control variable. Finally, as an example, the linear quadratic control problem is given to illustrate our theoretical results.

AMS Subject Classification: *93E20, 60H10.*

Keywords: Stochastic control with $\varepsilon^{\lambda}-$error bound, Weak maximum principle, Necessary and sufficient of conditions of near-optimality, Ekeland's variational principle, Convex perturbation.

4. Introduction

Stochastic near-optimization is as sensible and important as optimization for both theory and applications. In this work, we consider stochastic control problem with $\varepsilon^{\lambda}-$error bound for systems driven by non linear controlled SDEs of the form

$$\begin{cases} dx\left(t\right)=f\left(w,t,x\left(t\right),u\left(t\right)\right)dt+\sigma\left(w,t,x\left(t\right),u\left(t\right)\right)dW\left(t\right), \\ x(0)=\xi, \end{cases} \tag{25}$$

where $\left(W\left(t\right)\right)_{t\in[0,T]}$ is a standard $n-$dimensional Brownian motion defined on the filtered probability space $\left(\Omega,\mathcal{F},\left(\mathcal{F}_t\right)_{t\in[0,T]},P\right)$. The filtration $\mathcal{F}_t$ is a canonical filtration of $W\left(t\right)$ augmented by P-null sets. The initial condition ξ is an $\mathcal{F}_0$-measurable random variable. We associate to this state equation the following cost functional

$$J\left(u\left(\cdot\right)\right)=E\left[h\left(x\left(T\right)\right)+\int_0^T g\left(w,t,x\left(t\right),u\left(t\right)\right)dt\right], \tag{26}$$

and the value function is defined as

$$V = \inf_{u(\cdot)\in\mathcal{U}} \{J(u(\cdot))\}. \tag{27}$$

The maximum principle has been and remains an important tool in many situations in which optimal control plays a role. Near-optimization is as sensible and important as optimization for both theory and applications. The theory of stochastic near-optimization was introduced by Zhou [71]. Various kinds of near-optimal stochastic control problems have been investigated in [17, 18, 19, 20, 21, 28, 57, 70, 35]. The necessary and sufficient conditions of near-optimal mean-field singular stochastic control have been studied in Hafayed and Abbas [17]. The necessary and sufficient conditions for near-optimality for mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [18]. Near-optimality necessary and sufficient conditions for singular controls in jump diffusion processes have been investigated in Hafayed and Abbas [19]. In Hafayed, Veverka and Abbas [20], the authors extended Zhou's maximum principle of near-optimality [71] to singular stochastic control. The near-optimal stochastic control problem for jump diffusions has been investigated by Hafayed, Abbas and Veverka [21]. The near-optimality necessary and sufficient conditions for classical controlled FBSDEJs with applications to finance have been investigated in Hafayed, Veverka and Abbas [28]. Stochastic maximum principle of near-optimal control of fully coupled forward-backward stochastic differential equation has been investigated in Tang [57]. Near-optimal stochastic control problem for linear general controlled FBSDEs has been studied in Zhang, Huang and Li [70]. The near-optimal control problem for recursive stochastic problem has been studied in Hui, Huang, Li and Wang [35].

It is shown that the near-optimal controls in stochastic control problems, as the alternative to the exact optimal ones, are of great importance for both the theoretical analysis and practical application purposes due to its nice structure and broad-range availability as well as feasibility. The near-optimal controls in stochastic control problems are more available than the exact optimal ones, in the sense that the near-optimal controls always exist, while the exact optimal stochastic controls may not even exist in many situations. Moreover, since there are many near-optimal controls, it is possible to select among them appropriate ones that are easier for analysis and implementation. This justifies the use of near-optimal stochastic controls, which exist under minimal hypothesis and are sufficient in most practical cases.

Motivated by the arguments above and inspired by [71, 17, 18, 19, 21, 70], our purpose in this work is to derive a first-order necessary and sufficient conditions for any near-optimal stochastic control with $\varepsilon^{\lambda}-$error bound, where the diffusion coefficient can contain a control variable, and the control domain is necessarily convex. The proof of our main result is based on Ekeland's variational principle [14] and some approximation arguments on the state variable and adjoint process with respect to the control variable. As an applications, a linear quadratic control problem is discussed.

The rest of the chapter is organized as follows. In the second section we present the assumptions and the formulation of the problem. The necessary conditions for any near-optimal stochastic control is given in the third section. The sufficient conditions are given in the fourth section. An application to the linear quadratic control problem is given in the last section.

5. Assumptions and Preliminaries

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a fixed filtered probability space satisfying the usual conditions, in which a $n-$dimentional Brownian motion $W(t)$ is defined. We list some notations that will be used throughout this work. Any element $x \in \mathbb{R}^d$ will be identified to a column vector with i^{th} component, and the norm $|x| = \sum_{i=1}^{d} |x_i|$. We denote $\mathcal{A}^*$ the transpose of any vector or matrix $\mathcal{A}$. We denote $sgn(\cdot)$ the sign function. For a function Ψ, we denote by Ψ_x the gradient or Jacobian of a scalar function Ψ with respect to the variable x. We denote by $\mathbb{L}^2_{\mathcal{F}}([0,T], \mathbb{R}^n)$ the Hilbert space of $\mathcal{F}_t-$adapted processes $(x(t))$ such that $E\int_0^T |x(t)|^2 dt < +\infty$.
Throughout this work we assume the following.
Let $\sigma : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n\otimes\mathbb{R}^n$, $f : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, $g : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $h : \Omega\times\mathbb{R}^n\to\mathbb{R}$, are Borel measurable functions such that $\forall (w,t,x,y,u)\in\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m$.

Assumption (H1) f, σ, g and h are continuously differentiable with respect to x, u, dominated by $C(1+|x|)$, and their derivatives are bounded functions.

Assumption (H2) $\left|\frac{\partial\rho}{\partial u}(w,t,x,u) - \frac{\partial\rho}{\partial u}(w,t,y,v)\right| \leq C(|x-y|^\beta + |u-v|^\beta)$, for $\rho := f$, σ and $\beta\in(0,1)$.

Assumption (H3) The derivatives $\frac{\partial f}{\partial x}, \frac{\partial \sigma}{\partial x}, \frac{\partial g}{\partial x}, \frac{\partial g}{\partial u}$ are Lipschitz in x, u and h_x is Lipschitz in x.

Definition 1.2.1.. Let $T > 0$ be a fixed strictly positive real number and $\mathbb{U}$ be a nonempty compact convex subset of $\mathbb{R}^m$. An admissible control is defined as a function $u(\cdot) : [0,T]\times\Omega \longrightarrow \mathbb{U}$ which is $\mathcal{F}_t$-predictable, such that the SDE-(25) has a unique solution and write $u(\cdot)\in\mathcal{U}$. The set $\mathcal{U}$ is called the set of admissible controls.
From assumption (H1), the SDE-(25) has a unique strong solution given by

$$x(t) = \xi + \int_0^t f(w,s,x(s),u(s))\,ds + \int_0^t \sigma(w,s,x(s),u(s))\,dW(s).$$

The criteria to be minimized over the set of admissible controls given in (26) is well defined.
We introduce the adjoint equation for our control problem (25)-(26) as follows

$$\begin{cases} -dp(t) = [\frac{\partial f^*}{\partial x}(w,t,x(t),u(t))\,p(t) + \frac{\partial \sigma^*}{\partial x}(w,t,x(t),u(t))\,q(t) \\ \qquad + \frac{\partial g}{\partial x}(w,t,x(t),u(t))]dt - q(t)\,dW(t), \\ p(T) = h_x(x(T)), \end{cases} \tag{28}$$

and the Hamiltonian associated with our control problem (25)-(26) is given as

$$H(t,x,u,p(t),q(t)) = p(t)\,f(t,x,u) + q(t)\,\sigma(t,x,u) + g(t,x,u). \tag{29}$$

To simplify our notation, we suppress "w" in $f(w,t,x(t),u(t))$ and write $f(t,x(t),u(t))$ for $f(w,t,x(t),u(t))$. Similarly for the functions f, σ, g, h.
We aim at using Ekeland's variational principle [14] to establish necessary conditions of ε-optimality satisfied by a sequence of ε-optimal controls.

Lemma 1.2.1. (Ekeland's Lemma [14]) *Let* (E, d) *be a complete metric space and* $f : E \to \overline{\mathbb{R}}$ *be a lower semi-continuous and bounded from below. If for each* $\varepsilon > 0$*, there exists* $u^{\varepsilon} \in E$ *satisfies* $f(u^{\varepsilon}(\cdot)) \leq \inf_{u(\cdot)\in E}(f(u(\cdot))) + \varepsilon$*. Then for any* $\delta > 0$*, there exists* $u^{\delta}(\cdot) \in E$ *such that*
(1) $f\left(u^{\delta}(\cdot)\right) \leq f(u^{\varepsilon}(\cdot))$*.*
(2) $d\left(u^{\delta}(\cdot), u^{\varepsilon}(\cdot)\right) \leq \delta$*.*
(3) $f\left(u^{\delta}(\cdot)\right) \leq f(u(\cdot)) + \frac{\varepsilon}{\delta} d\left(u^{\delta}(\cdot), u(\cdot)\right)$*, for all* $u(\cdot) \in E$*.*
To apply Ekeland's variational principle to our problem, we must define a distance d on the space of admissible controls such that $(\mathcal{U}, d)$ becomes a complete metric space. For any $u(\cdot)$, $v(\cdot) \in \mathcal{U}$ we lay

$$d(u(\cdot), v(\cdot)) = \left[E \int_0^T |u(t) - v(t)|^2 dt\right]^{\frac{1}{2}}. \tag{30}$$

6. Stochastic maximum principle with $\varepsilon^{\lambda}-$error bound

Our goal in this section is to derive necessary conditions with $\varepsilon^{\lambda}-$error bound for SDEs with controlled diffusion coefficient, where the control domain is necessarily convex. We give the definition of ε-optimal control as given in [71].

Definition 1.3.1. For a given $\varepsilon > 0$ the admissible control $u^{\varepsilon}(\cdot)$ is ε-optimal if

$$|J(u^{\varepsilon}(\cdot)) - V| \leq \mathcal{Q}(\varepsilon),$$

where $\mathcal{Q}$ is a function of ε satisfying $\lim_{\varepsilon\to 0} \mathcal{Q}(\varepsilon) = 0$. The estimater $\mathcal{Q}(\varepsilon)$ is called an error bound. If $\mathcal{Q}(\varepsilon) = C\varepsilon^{\delta}$ for some $\delta > 0$ independent of the constant C, then $u^{\varepsilon}(\cdot)$ is called ε-optimal control with order ε^{δ}. If $\mathcal{Q}(\varepsilon) = \varepsilon$, the admissible control $u^{\varepsilon}(\cdot)$ called $\varepsilon-$optimal.

Now we are able to state and prove the Pontryagin's maximum principle of ε-optimality for our control problem, which is the main result in this section.

Theorem 1.3.1. *Assume that (H1), (H2) and (H3) hold. For any* $\lambda \in \left[0, \frac{1}{2}\right)$*, there exists a positive constant* $C = C(\lambda)$ *such that for each* $\varepsilon > 0$ *and any* ε*-optimal control* $u^{\varepsilon}(\cdot)$ *there exists a constant* $C > 0$ *such that for all* $u \in \mathbb{U}$

$$E \int_0^T \frac{\partial H}{\partial u}(t, x^{\varepsilon}(t), u^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t))(u(t) - u^{\varepsilon}(t))dt \geq -C\varepsilon^{\lambda},\ dt - a.e., \tag{31}$$

where $x^{\varepsilon}(\cdot)$ *denotes the solution of the state equation (25) and the pair* $(p^{\varepsilon}(\cdot), q^{\varepsilon}(\cdot))$ *is the solution of the adjoint equation (28) associated with* u^{ε}*.*
To prove the above Theorem, we need the following auxiliary results on the variation of the state and adjoint processes with respect to the control variable.

Lemma 1.3.2. *Let* $x^u(t)$ *and* $x^v(t)$ *be the solution of the state equation (25) associated with* $u(\cdot)$ and $v(\cdot)$ *respectively. Then there exists a positive constant* C *such that, for* $\alpha > 0$ *:*

$$E\left[\sup_{0\leq t\leq T} |x^u(t) - x^v(t)|^{\alpha}\right] \leq C d^{\frac{\alpha}{2}}(u(\cdot), v(\cdot)).$$

Proof.
First we assume $\alpha \geq 2$. Using Hölder's and Burkholder-Davis-Gundy inequalities, we obtain

$$E\left[|x^u(t)-x^v(t)|^\alpha\right] \leq E\left|\int_0^t (f(s,x^u(s),u(s))-f(s,x^v(s),v(s)))ds \right.$$
$$\left. +\int_0^t (\sigma(s,x^u(s),u(s))-\sigma(s,x^v(s),v(s)))\,dW(s)\right|^\alpha$$
$$\leq CE\int_0^t |f(s,x^u(s),u(s))-f(s,x^v(s),v(s))|^\alpha\,ds$$
$$+CE\int_0^t |\sigma(s,x^u(s),u(s))-\sigma(s,x^v(s),v(s))|^\alpha\,ds$$

by adding and subtracting $f(s,x^v(s),u(s))$, $\sigma(s,x^v(s),u(s))$ and applying the Lipschitz continuity of the coefficients f and σ it holds that

$$E\left[|x^u(t)-x^v(t)|^\alpha\right] \leq CE\int_0^t |x^u(s)-x^v(s)|^\alpha\,ds + CE\int_0^t |u(s)-v(s)|^\alpha\,ds$$
$$\leq \quad CE\int_0^t |x^u(s)-x^v(s)|^\alpha\,ds + C\left[E\int_0^t |u(s)-v(s)|^2\,ds\right]^{\frac{\alpha}{2}},$$

using Gronwall's inequality, we get the desired inequality.
Now, we assume $0<\alpha<2$. Since $\frac{2}{\alpha}>1$, then by using Hölder's inequality and the above result, we have

$$E\left[|x^u(t)-x^v(t)|^\alpha\right] \leq \left[E\,|x^u(t)-x^v(t)|^2\right]^{\frac{\alpha}{2}} \leq Cd^{\frac{\alpha}{2}}(u(\cdot),v(\cdot)).$$

This completes the proof of Lemma 1.3.2. □

Lemma 1.3.3. *Let $(p^u(t),q^u(t))$ and $(p^v(t),q^v(t))$ be two adjoint processes corresponding to u and v respectively. Then we have the following estimate: for any $\alpha \geq 1$*

$$E\int_0^T (\,|p^u(t)-p^v(t)|^\alpha + |q^u(t)-q^v(t)|^\alpha)dt \leq Cd^\alpha(u(\cdot),v(\cdot)).$$

Proof. First we denote by $p(t)=(p^u(t)-p^v(t))$ and $q(t)=(q^u(t)-q^v(t))$, then $(\tilde{p}(t),\tilde{q}(t))$ satisfies the following backward stochastic differential equation:

$$\begin{cases} -d\tilde{p}(t) = [\frac{\partial f^*}{\partial x}(t,x^u(t),u(t))\,\tilde{p}(t) + \frac{\partial \sigma^*}{\partial x}(t,x^u(t),u(t))\,\tilde{q}(t) \\ \qquad + G(t)]dt - \tilde{q}(t)\,dW(t), \\ \tilde{p}(t) = h_x(x^u(T)) - h_x(x^v(T)), \end{cases}$$

where the process $G(t)$ is given by

$$G(t) = (\tfrac{\partial f}{\partial x}(t,x^u(t),u(t)) - \tfrac{\partial f}{\partial x}(t,x^v(t),v(t)))p^v(t)$$
$$+(\tfrac{\partial \sigma}{\partial x}(t,x^u(t),u(t)) - \tfrac{\partial \sigma}{\partial x}(t,x^v(t),v(t)))q^v(t)$$
$$+(\tfrac{\partial g}{\partial x}(t,x^u(t),u(t)) - \tfrac{\partial g}{\partial x}(t,x^v(t),v(t))).$$

Let η be the solution of the following linear SDE

$$\begin{cases} d\eta_t = [\frac{\partial f}{\partial x}\left(t, x^u\left(t\right), u\left(t\right)\right)\eta_t + \left|\tilde{p}\left(t\right)\right|^{\alpha-1} sgn(\tilde{p}\left(t\right))]dt \\ +[\frac{\partial \sigma}{\partial x}\left(t, x^u\left(t\right), u\left(t\right)\right)\eta_t + \left|\tilde{q}\left(t\right)\right|^{\alpha-1} sgn(\tilde{q}\left(t\right))]dW\left(t\right), \\ \eta_0 = 0, \end{cases} \tag{32}$$

where $sgn\left(y\right) \equiv \left(sgn(y_1), sgn(y_2), ..., sgn(y_n)\right)^*$ for any vector $y = \left(y_1, y_2, .., y_n\right)^*$. It is worth mentioning that since $\frac{\partial f}{\partial x}$ and $\frac{\partial \sigma}{\partial x}$ are bounded and the fact that

$$E\int_0^T \left\{\left|\left|\tilde{p}\left(t\right)\right|^{\alpha-1} sgn\left(\tilde{p}\left(t\right)\right)\right|^2 + \left|\left|\tilde{q}\left(t\right)\right|^{\alpha-1} sgn\left(\tilde{q}\left(t\right)\right)\right|^2\right\} dt < \infty,$$

then the SDE (32) has a unique strong solution. Let $\gamma \geq 2$ such that $\frac{1}{\gamma} + \frac{1}{\alpha} = 1$ then we get

$$\begin{aligned} & E\left\{\sup_{t\leq T}\left|\eta_t\right|^{\gamma}\right\} \leq CE\int_0^T \left\{\left|\tilde{p}\left(t\right)\right|^{\alpha\gamma-\gamma} + \left|\tilde{q}\left(t\right)\right|^{\alpha\gamma-\gamma}\right\} dt \\ & = CE\int_0^T \left\{\left|\tilde{p}\left(t\right)\right|^{\alpha} + \left|\tilde{q}\left(t\right)\right|^{\alpha}\right\} dt. \end{aligned} \tag{33}$$

Now applying Itô's formula to $p\left(t\right)\eta_t$ on $[0, T]$ and taking expectations, we obtain

$$E\left(\tilde{p}\left(t\right)\eta_T - \tilde{p}(0)\eta_0\right) = E\int_0^T -G\left(t\right)\eta_t dt + E\int_0^T (\left|\tilde{p}\left(t\right)\right|^{\alpha} + \left|\tilde{q}\left(t\right)\right|^{\alpha})dt,$$

using the fact that $\eta_0 = 0$ we can easily show that

$$\begin{aligned} E\int_0^T (\left|\tilde{p}\left(t\right)\right|^{\alpha} + \left|\tilde{q}\left(t\right)\right|^{\alpha})dt &= E\int_0^T G\left(t\right)\eta_t dt + E(\tilde{p}\left(T\right)\eta_T) \\ = \quad E\int_0^T G\left(t\right)\eta_t dt &+ E\left[\left(h_x\left(x^u\left(T\right)\right) - h_x\left(x^v\left(T\right)\right)\right)\eta_T\right], \end{aligned}$$

by applying Hölder's inequality to the right hand side, it holds that

$$\begin{aligned} & E\int_0^T (\left|\tilde{p}\left(t\right)\right|^{\alpha} + \left|\tilde{q}\left(t\right)\right|^{\alpha})dt \leq \left[E\int_0^T \left|G\left(t\right)\right|^{\alpha} dt\right]^{\frac{1}{\alpha}} \left[E\int_0^T \left|\eta_t\right|^{\gamma} dt\right]^{\frac{1}{\gamma}} \\ & + \left[E\left|h_x(x^u\left(T\right)) - h_x(x^v\left(T\right))\right|^{\alpha}\right]^{\frac{1}{\alpha}} \left[E\left|\eta_T\right|^{\gamma}\right]^{\frac{1}{\gamma}} \end{aligned}$$

using inequality (33), it holds that

$$\begin{aligned} & E\int_0^T (\left|\tilde{p}\left(t\right)\right|^{\alpha} + \left|\tilde{q}\left(t\right)\right|^{\alpha})dt \leq C\left[E\int_0^T (\left|\tilde{p}\left(t\right)\right|^{\alpha} + \left|\tilde{q}\left(t\right)\right|^{\alpha})dt\right]^{\frac{1}{\gamma}} \left\{\left[E\int_0^T \left|G\left(t\right)\right|^{\alpha} dt\right]^{\frac{1}{\alpha}}\right. \\ & \left. + C\left[E\left|h_x(x^u\left(T\right)) - h_x(x^v\left(T\right))\right|^{\alpha}\right]^{\frac{1}{\alpha}}\right\}, \end{aligned}$$

which implies that

$$\left[E\int_0^T(|\tilde{p}(t)|^\alpha+|\tilde{q}(t)|^\alpha)dt\right]^{1-\frac{1}{\gamma}}\leq C\left[E\int_0^T|G(t)|^\alpha dt\right]^{\frac{1}{\alpha}}$$
$$+C\left[E\left|h_x(x^u(T))-h_x(x^v(T))\right|^\alpha\right]^{\frac{1}{\alpha}},$$

thus

$$E\int_0^T(|\tilde{p}(t)|^\alpha+|\tilde{q}(t)|^\alpha)dt\leq CE\int_0^T|G(t)|^\alpha dt$$
$$+CE\left|h_x(x^u(T))-h_x(x^v(T))\right|^\alpha.$$

Since h_x is Lipschitz in x and due to Lemma 1.3.2, we have

$$E\left\{|h_x(x^u(T))-h_x(x^v(T))|^\alpha\right\}\leq Cd^\alpha(u(\cdot),v(\cdot)). \tag{34}$$

We proceed to estimate the first term on the right hand side, then we have

$$E\int_0^T|G(t)|^\alpha dt\leq E\int_0^T\left\{\left|\frac{\partial f}{\partial x}(t,x^u(t),u(t))-\frac{\partial f}{\partial x}(t,x^v(t),v(t))\right||p^v(t)|\right.$$
$$+\left|\frac{\partial \sigma}{\partial x}(t,x^v(t),u(t))-\frac{\partial \sigma}{\partial x}(t,x^v(t),v(t))\right||q^v(t)|+\left|\frac{\partial g}{\partial x}(t,x^u(t),u(t))\right.$$
$$\left.-\left.\frac{\partial g}{\partial x}(t,x^v(t),v(t))\right|\right\}^\alpha dt$$

$$\begin{aligned}
&\leq CE\int_0^T\left|\frac{\partial f}{\partial x}(t,x^u(t),u(t))-\frac{\partial f}{\partial x}(t,x^v(t),v(t))\right|^\alpha|p^v(t)|^\alpha dt\\
&\quad+CE\int_0^T\left|\frac{\partial \sigma}{\partial x}(t,x^u(t),u(t))-\frac{\partial \sigma}{\partial x}(t,x^v(t),v(t))\right|^\alpha|q^v(t)|^\alpha dt\\
&\quad+CE\int_0^T\left|\frac{\partial g}{\partial x}(t,x^u(t),u(t))-\frac{\partial g}{\partial x}(t,x^v(t),v(t))\right|^\alpha dt\\
&=\mathbb{I}_1+\mathbb{I}_2+\mathbb{I}_3.
\end{aligned}$$

Using the bounded of $p^v(t)$ and Hölder's inequality with $\frac{1}{2/(2-\alpha)}+\frac{1}{2/\alpha}=1$ we have

$$\begin{aligned}
\mathbb{I}_1&\leq C\left[E\int_0^T\left|\frac{\partial f}{\partial x}(t,x^u(t),u(t))-\frac{\partial f}{\partial x}(t,x^v(t),v(t))\right|^{\frac{2\alpha}{2-\alpha}}dt\right]^{1-\frac{\alpha}{2}}\\
&\quad\times\left[E\int_0^T|p^v(t)|^2dt\right]^{\frac{\alpha}{2}},
\end{aligned}$$

adding and subtracting $\frac{\partial f}{\partial x}(t, x^u, v)$, then by using the Lipschitz continuity on $\frac{\partial f}{\partial x}(t, x^u, v)$ in x and u (Assumption (H3)) and Lemma 1.3.2, we have

$$\mathbb{I}_1 \leq Cd^{\alpha}(u(\cdot), v(\cdot)).$$

Using similar argument developed above, we can prove $\mathbb{I}_2+\mathbb{I}_3 \leq Cd^{\alpha}(u(\cdot), v(\cdot))$. Then we conclude

$$E\int_0^T |G(t)|^{\alpha}\, dt \leq Cd^{\alpha}(u(\cdot), v(\cdot)). \tag{35}$$

Finally, combining (34) and (35), the proof of Lemma 1.3.3 is complete ☐

Lemma 1.3.4 *(Maximum principle for ε-optimality). For each $\varepsilon > 0$ there exists $\overline{u}^{\varepsilon}(\cdot) \in \mathcal{U}$* processes $\overline{p}^{\varepsilon}(t)$ and $\overline{q}^{\varepsilon}(t)$ *such that,* $\forall u(\cdot) \in \mathcal{U}$

$$E\int_0^T \frac{\partial H}{\partial u}(t, \overline{x}^{\varepsilon}, \overline{u}^{\varepsilon}, \overline{p}^{\varepsilon}(t), \overline{q}^{\varepsilon}(t))(u(t) - \overline{u}^{\varepsilon}(t))\, dt \geq -C\varepsilon^{\lambda},\ dt - a.e. \tag{36}$$

Proof. Applying Ekeland's variational principle with $\delta = \varepsilon^{1/2}$ there exists an admissible control $\overline{u}^{\varepsilon}$ such that
(i) $d(\overline{u}^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot)) \leq \varepsilon^{1/2}$,
(ii) $\overline{J}(\overline{u}^{\varepsilon}(\cdot)) \leq \overline{J}(u(\cdot))$, for any $u(\cdot) \in \mathcal{U}$ where

$$\bar{J}(u(\cdot)) := J(u(\cdot)) + \varepsilon^{1/2} d(\overline{u}^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot)). \tag{37}$$

Notice that $\overline{u}^{\varepsilon}(\cdot)$ which is ε-optimal for the initial cost J is optimal for the new cost $\overline{J}$ defined by (37).
Let us denote $u^{\varepsilon,\theta}(\cdot)$ a perturbed control given by $\overline{u}^{\varepsilon,\theta}(t) = \overline{u}^{\varepsilon}(t) + \theta(v(t) - \overline{u}^{\varepsilon}(t))$. By using the fact that
(i) $\overline{J}(\overline{u}^{\varepsilon}(\cdot)) \leq \overline{J}(u^{\varepsilon,\theta}(\cdot))$, (ii) $d(\overline{u}^{\varepsilon}(\cdot), u^{\varepsilon,\theta}(\cdot)) \leq C\theta$, we get

$$J(u^{\varepsilon,\theta}(\cdot)) - J(\overline{u}^{\varepsilon}(\cdot)) \geq -\varepsilon^{1/2} d(u^{\varepsilon}(\cdot), u^{\varepsilon,\theta}(\cdot)) \geq -C\varepsilon^{1/2}\theta. \tag{38}$$

Dividing (38) by θ and sending θ to zero we get

$$\frac{d}{d\theta}\left(J(u^{\varepsilon,\theta}(t))\right)\Big|_{\theta=0} \geq -C\varepsilon^{\frac{1}{2}} \geq -C\varepsilon^{\lambda}. \tag{39}$$

Arguing as in [6] for the left hand side of inequality (39), the desired result follows ☐

Proof of Theorem 1.3.1.
First, for each $\varepsilon > 0$ by using Lemma 1.3.4, there exists $\overline{u}^{\varepsilon}(\cdot)$ and $\mathcal{F}_t-$adapted processes $\overline{p}^{\varepsilon}(t)$ and $\overline{q}^{\varepsilon}(t)$ such that, $\forall u(\cdot) \in \mathcal{U}$:

$$E\int_0^T \frac{\partial H}{\partial u}(t, \overline{x}^{\varepsilon}, \overline{u}^{\varepsilon}, \overline{p}^{\varepsilon}(t), \overline{q}^{\varepsilon}(t))(u(t) - \overline{u}^{\varepsilon}(t))\, dt \geq -C\varepsilon^{\lambda},\ dt - a.e.$$

Now, to prove (31) it remains to estimate the following difference:

$$\begin{aligned}&E\int_0^T \frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)\left(u\left(t\right)-\overline{u}^{\varepsilon}\left(t\right)\right)dt\\ &-E\int_0^T \frac{\partial H}{\partial u}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt.\end{aligned}$$

First, by adding and subtracting $E\int_0^T \frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt$, we have

$$\begin{aligned}&E\int_0^T \frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)\left(u\left(t\right)-\overline{u}^{\varepsilon}(t)\right)dt\\ &-E\int_0^T \frac{\partial H}{\partial u}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt\\ \leq\ &E\int_0^T \frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)\left(u^{\varepsilon}\left(t\right)-\overline{u}^{\varepsilon}\left(t\right)\right)dt\\ &+E\int_0^T \left(\frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)-\frac{\partial H}{\partial u}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\right)\\ &\times\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt\\ =\ &\mathbb{I}_1+\mathbb{I}_2,\end{aligned}$$

by using Schwarz inequality and the bounded of $\frac{\partial H}{\partial u}$ in integral sense, we get

$$\begin{aligned}\mathbb{I}_1 &\leq E\int_0^T \left|\frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)\right|\left|\left(u^{\varepsilon}\left(t\right)-\overline{u}_t^{\varepsilon}\right)\right|dt\\ &\leq \left[E\left\{\int_0^T \left|\frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)\right|^2 dt\right\}\right]^{\frac{1}{2}}\left[E\left\{\int_0^T \left|\left(u^{\varepsilon}\left(t\right)-\overline{u}_t^{\varepsilon}\right)\right|^2 dt\right\}\right]^{\frac{1}{2}}\\ &\leq Cd\left(u^{\varepsilon}(\cdot),\overline{u}^{\varepsilon}(\cdot)\right)\leq C\varepsilon^{\frac{1}{2}}.\end{aligned}$$

Let us turn to the second term, it holds that

$$\begin{aligned}\mathbb{I}_2 &= E\int_0^T \left(\frac{\partial H}{\partial u}\left(t,\overline{x}^{\varepsilon},\overline{u}^{\varepsilon},\overline{p}^{\varepsilon}\left(t\right),\overline{q}^{\varepsilon}\left(t\right)\right)-\frac{\partial H}{\partial u}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\right)\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt\\ &= E\int_0^T \left[\overline{p}^{\varepsilon}(t)\frac{\partial f}{\partial u}\left(t,\overline{x}^{\varepsilon}(t),\overline{u}^{\varepsilon}(t)\right)-p^{\varepsilon}\left(t\right)\frac{\partial f}{\partial u}\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right)\right)\right]\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt\\ &\quad+E\int_0^T \left[\overline{q}^{\varepsilon}(t)\frac{\partial \sigma}{\partial u}\left(t,\overline{x}^{\varepsilon}\left(t\right),\overline{u}^{\varepsilon}(t)\right)-q^{\varepsilon}\left(t\right)\frac{\partial \sigma}{\partial u}\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right)\right)\right]\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt\\ &\quad+E\int_0^T \left[\frac{\partial g}{\partial u}\left(t,\overline{x}^{\varepsilon}\left(t\right),\overline{u}^{\varepsilon}(t)\right)-\frac{\partial g}{\partial u}\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right)\right)\right]\left(u\left(t\right)-u^{\varepsilon}\left(t\right)\right)dt\\ &= \mathbb{J}_1+\mathbb{J}_2+\mathbb{J}_3.\end{aligned}$$

We estimate the first term on the right hand side $\mathbb{J}_1$ by adding and subtracting $p^{\varepsilon}(t)\frac{\partial f}{\partial u}(t,\overline{x}^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))$ then we have

$$\mathbb{J}_1 \leq E\int_0^T |\overline{p}^{\varepsilon}(t)-p^{\varepsilon}(t)|\left|\frac{\partial f}{\partial u}(t,\overline{x}^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))(u(t)-u^{\varepsilon}(t))\right|dt$$
$$+E\int_0^T \left|\frac{\partial f}{\partial u}(t,\overline{x}^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))-\frac{\partial f}{\partial u}(t,x^{\varepsilon}(t),u^{\varepsilon}(t)))p^{\varepsilon}(t)(u(t)-u^{\varepsilon}(t))\right|dt.$$

First, by adding and subtracting $\frac{\partial f}{\partial u}(t,x^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))$ it holds that

$$\mathbb{J}_1 \leq E\int_0^T |\overline{p}^{\varepsilon}(t)-p^{\varepsilon}(t)|\left|\frac{\partial f}{\partial u}(t,\overline{x}_t^{\varepsilon},\overline{u}_t^{\varepsilon})(u(t)-u^{\varepsilon}(t))\right|dt$$
$$+E\int_0^T \left|\frac{\partial f}{\partial u}(t,\overline{x}^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))-\frac{\partial f}{\partial u}(t,x^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))\right||p^{\varepsilon}(t)(u(t)-u^{\varepsilon}(t))|dt$$
$$+E\int_0^T \left|\frac{\partial f}{\partial u}(t,x^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))-\frac{\partial f}{\partial u}(t,x^{\varepsilon}(t),u^{\varepsilon}(t))\right||p^{\varepsilon}(t)(u(t)-u^{\varepsilon}(t))|dt$$
$$=\mathbb{J}_1^1+\mathbb{J}_1^2+\mathbb{J}_1^3.$$

Using Hölder inequality, the bounded of $\frac{\partial f}{\partial u}$, Lemma 1.3.2 and integral properties of admissible controls, we obtain, for $\frac{1}{\gamma}+\frac{1}{\alpha}=1$,

$$\mathbb{J}_1^1 \leq \left[E\left\{\int_0^T \left|\frac{\partial f}{\partial u}(t,\overline{x}^{\varepsilon}(t),\overline{u}^{\varepsilon}(t))(u(t)-u^{\varepsilon}(t))\right|^{\gamma}dt\right\}\right]^{\frac{1}{\gamma}}\left[E\left\{\int_0^T |\overline{p}^{\varepsilon}(t)-p^{\varepsilon}(t)|^{\alpha}dt\right\}\right]^{\frac{1}{\alpha}}$$
$$\leq C\left(E\left\{\int_0^T |\overline{p}^{\varepsilon}(t)-p^{\varepsilon}(t)|^{\alpha}dt\right\}\right)^{\frac{1}{\alpha}}$$
$$\leq C\left(d^{\alpha}(\overline{u}^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))\right)^{\frac{1}{\alpha}} \leq C\varepsilon^{\frac{1}{2}}.$$

To estimate the second term $\mathbb{J}_1^2$ we use assumption (H2), then we have

$$\mathbb{J}_1^2 \leq CE\int_0^T |\overline{x}^{\varepsilon}(t)-x^{\varepsilon}(t)|^{\beta}|p^{\varepsilon}(t)(u(t)-u^{\varepsilon}(t))|dt,$$

using Hölder inequality, where $\frac{1}{\gamma}+\frac{1}{\alpha}=1$ then a simple computations gets

$$\begin{aligned}\mathbb{J}_1^2 &\leq C\left(E\int_0^T |\overline{x}^{\varepsilon}(t)-x^{\varepsilon}(t)|^{\alpha\beta}|p^{\varepsilon}(t)|^{\alpha}\right)^{\frac{1}{\alpha}}\left(E\int_0^T |(u(t)-u^{\varepsilon}(t))|^{\gamma}dt\right)^{\frac{1}{\gamma}}\\ &\leq C\left(E\int_0^T |\overline{x}^{\varepsilon}(t)-x^{\varepsilon}(t)|^{\alpha\beta}|p^{\varepsilon}(t)|^{\alpha}\right)^{\frac{1}{\alpha}},\end{aligned}$$

applying Hölder inequality for $\frac{1}{2/(2-\alpha)}+\frac{1}{2/\alpha}=1$ it holds that

$$\begin{aligned}\mathbb{J}_1^2 &\leq C\left[\left(E\int_0^T |\overline{x}^{\varepsilon}(t)-x^{\varepsilon}(t)|^{\frac{2\alpha\beta}{2-\alpha}}\right)^{\frac{2-\alpha}{\alpha}}\times\left(E\int_0^T |p^{\varepsilon}(t)|^{\alpha.\frac{2}{\alpha}}\right)^{\frac{\alpha}{2}}\right]^{\frac{1}{\alpha}}\\ &\leq C\left(d^{\frac{2\alpha\beta}{2-\alpha}}(u^{\varepsilon}(\cdot),\overline{u}^{\varepsilon}(\cdot))\right)^{\frac{2-\alpha}{2}\cdot\frac{1}{\alpha}} \leq C\varepsilon^{\lambda}.\end{aligned}$$

Next by applying assumption (H2) and Hölder inequality then we can proceed to estimate $\mathbb{J}_1^3$ as follows

$$\begin{aligned}\mathbb{J}_1^3 &\leq CE\int_0^T \left|\overline{u}^\varepsilon(t)-u^\varepsilon(t)\right|^\beta \left|p^\varepsilon(t)\right| \left|(u(t)-u^\varepsilon(t))\right| dt\\ &\leq C\left(E\int_0^T \left|\overline{u}^\varepsilon(t)-u^\varepsilon(t)\right|^{\alpha\beta} \left|p^\varepsilon(t)\right|^\alpha dt\right)^{\frac{1}{\alpha}} \left(E\int_0^T \left|(u(t)-u^\varepsilon(t))\right|^\gamma dt\right)^{\frac{1}{\gamma}}\\ &\leq C\left(\left(E\int_0^T \left|\overline{u}^\varepsilon(t)-u^\varepsilon(t)\right|^{\frac{2\alpha\beta}{2-\alpha}} dt\right)^{\frac{2-\alpha}{2}} \left(E\int_0^T \left|p^\varepsilon(t)\right|^{\alpha\frac{2}{\alpha}} dt\right)^{\frac{2}{\alpha}}\right)^{\frac{1}{\alpha}}\\ &\leq C\varepsilon^\beta.\end{aligned}$$

Using similar arguments developed above for $\mathbb{J}_2$ and $\mathbb{J}_3$, then a simple computations we can prove that $\mathbb{I}_1 \leq C\varepsilon^\lambda$. Applying similar method developed above for $\mathbb{I}_2$ and $\mathbb{I}_3$ we conclude

$$\begin{aligned}&E\int_0^T \tfrac{\partial H}{\partial u}\left(t,\overline{x}^\varepsilon,\overline{u}^\varepsilon,\overline{p}^\varepsilon(t),\overline{q}^\varepsilon(t)\right)\left(u(t)-\overline{u}^\varepsilon(t)\right)dt\\ &-E\int_0^T \tfrac{\partial H}{\partial u}\left(t,x^\varepsilon,u^\varepsilon,p^\varepsilon(t),q^\varepsilon(t)\right)\left(u(t)-u^\varepsilon(t)\right)dt \leq C\varepsilon^\lambda.\end{aligned} \tag{40}$$

Finally combining (36) and (40) the proof of Theorem 1.3.1 is complete. □

7. Sufficient conditions for ε-optimality

In this section, we will prove that under an additional hypothesis, the ε-maximum condition on the Hamiltonian function is a sufficient condition for ε-optimality.

Theorem 1.4.2. Assume that $H(t,\cdot,\cdot,p^\varepsilon(\cdot),q^\varepsilon(\cdot))$ is convex for $a.e.\ t\in[0,T]$, $P-a.s$, and h is convex. Let $(u^\varepsilon(\cdot),x^\varepsilon(\cdot))$ be a ε-optimal solution of the control problem (25)-(26) and $(p^\varepsilon(t),q^\varepsilon(t))$ be the solution of the adjoint equation associated with $u^\varepsilon(\cdot)$. If for some $\varepsilon>0$ and for any $u(\cdot)\in\mathcal{U}$:

$$E\int_0^T \frac{\partial H}{\partial u}\left(t,x^\varepsilon,u^\varepsilon,p^\varepsilon(t),q^\varepsilon(t)\right)\left(u(t)-u^\varepsilon(t)\right)dt \geq -C\varepsilon^\lambda, \tag{41}$$

then $u^\varepsilon(\cdot)$ is an ε-optimal control of order ε^λ, i.e.,

$$J\left(u^\varepsilon(\cdot)\right) \leq \inf_{v(\cdot)\in\mathcal{U}} J\left(v(\cdot)\right) + C\varepsilon^\lambda,$$

where C is a positive constant independent from ε.

Proof. Let $u^\varepsilon(\cdot)$ be an arbitrary element of $\mathcal{U}$ (candidate to be $\varepsilon-$optimal) and $x^\varepsilon(\cdot)$ is the corresponding trajectory. For any $v(\cdot)\in\mathcal{U}$ and its corresponding trajectory $x^v(\cdot)$, we have

$$\begin{aligned}&J\left(u^\varepsilon(\cdot)\right)-J\left(v(\cdot)\right)=E\int_0^T \left(g\left(t,x^\varepsilon(t),u^\varepsilon(t)\right)-g\left(t,x^v(t),v(t)\right)\right)dt\\ &+E\left[h\left(x^\varepsilon(T)\right)-h\left(x^v(T)\right)\right].\end{aligned}$$

Since h is convex, we have

$$\begin{aligned}
& J\left(u^{\varepsilon}(\cdot)\right)-J\left(v(\cdot)\right) \\
\leq\ & E\left[h_x\left(x^{\varepsilon}\left(T\right)\right)\left(x^{\varepsilon}\left(T\right)-x^{v}\left(T\right)\right)\right]+E\int_0^T\left(g\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right)\right)-g\left(t,x^{v}\left(t\right),v\left(t\right)\right)\right)dt,
\end{aligned}$$

replacing $h_x\left(x^{\varepsilon}(T)\right)$ with its value, see (28) we have

$$\begin{aligned}
& J\left(u^{\varepsilon}(\cdot)\right)-J\left(v(\cdot)\right)\leq E\left[p^{\varepsilon}\left(T\right)\left(x^{\varepsilon}\left(T\right)-x^{v}\left(T\right)\right)\right] \\
& +E\int_0^T\left(g\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right)\right)-g\left(t,x^{v}\left(t\right),v\left(t\right)\right)\right)dt.
\end{aligned} \tag{42}$$

On the other hand, by applying Itô's formula to $p^{\varepsilon}\left(T\right)\left(x^{\varepsilon}\left(T\right)-x^{v}\left(T\right)\right)$, and by taking expectation, we obtain

$$\begin{aligned}
& E\left[p^{\varepsilon}\left(T\right)\left(x^{\varepsilon}\left(T\right)-x^{v}\left(T\right)\right)\right] \\
& =E\int_0^T(H\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right),p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)-H\left(t,x^{v}\left(t\right),v,p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right))dt \\
& -E\int_0^T\frac{\partial H}{\partial x}\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right),p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(x^{\varepsilon}\left(t\right)-x^{v}\left(t\right)\right)dt \\
& -E\int_0^T\left(g\left(t,x^{\varepsilon}\left(t\right),u^{\varepsilon}\left(t\right)\right)-g\left(t,x^{v}\left(t\right),v\left(t\right)\right)\right)dt,
\end{aligned} \tag{43}$$

then by combining (42) and (43) we have

$$\begin{aligned}
& J\left(u^{\varepsilon}(\cdot)\right)-J\left(v(\cdot)\right)\leq E\int_0^T(H\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)-H\left(t,x^{v},v,p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right))dt \\
& -E\int_0^T\frac{\partial H}{\partial x}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(x^{\varepsilon}\left(t\right)-x^{v}\left(t\right)\right)dt.
\end{aligned} \tag{44}$$

Since H is convex in (x,u) we obtain

$$\begin{aligned}
& H\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)-H\left(t,x^{v},v,p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right) \\
& \leq\frac{\partial H}{\partial x}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(x^{\varepsilon}\left(t\right)-x^{v}\left(t\right)\right) \\
& +\frac{\partial H}{\partial u}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(u^{\varepsilon}\left(t\right)-v\left(t\right)\right),
\end{aligned}$$

then by using the necessary optimality conditions (41), it follows that

$$\begin{aligned}
& C\varepsilon^{\lambda}\geq H\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)-H\left(t,x^{v},v,p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right) \\
& -\frac{\partial H}{\partial x}\left(t,x^{\varepsilon},u^{\varepsilon},p^{\varepsilon}\left(t\right),q^{\varepsilon}\left(t\right)\right)\left(x^{\varepsilon}\left(t\right)-x^{v}\left(t\right)\right).
\end{aligned} \tag{45}$$

Finally combining (44) and (45) the desired result follows. □

8. Application: linear quadratic control problem

In this section, we consider a linear quadratic control problem as a particular case of our control problem. First, we restrict ourselves to the one dimensional case. We assume that $T = 1$ and the convex control domain be $U = [0,1]$, $f(t, x(t), u(t)) = -u(t)$, $\sigma(t, x(t), u(t)) = u(t)$, $g(t, x(t), u(t)) = \frac{1}{2}u^2(t)$ and $h(x(t)) = x(t)$.
Consider the following stochastic control problem

$$\begin{cases} dx(t) = -u(t)\,dt + u(t)\,dW(t), \\ x(0) = \frac{1}{2}, \end{cases} \tag{46}$$

and the cost functional being

$$J(u(\cdot)) = E\left\{x(1) + \int_0^1 \tfrac{1}{2}u^2(t)\,dt\right\}. \tag{47}$$

The Hamiltonian function gets the form

$$H(t, x, u, p(t), q(t)) = (q(t) - p(t))\,u + \frac{1}{2}u^2, \tag{48}$$

and the corresponding adjoint equation is given as follows

$$-dp(t) = q(t)\,dW(t),\ p(1) = 1. \tag{49}$$

It is clear that $(p(t), q(t)) = (1, 0)$ is the only unique adapted solution to (49). Moreover, the Hamiltonian function has the form

$$H(t, x, u, p(t), q(t)) = -u + \frac{1}{2}u^2. \tag{50}$$

If the admissible control $u^{\varepsilon}(\cdot)$ is $\varepsilon-$optimal in the sense that $J(u^{\varepsilon}(\cdot)) \leq \inf_{u(\cdot)\in\mathcal{U}} J(u(\cdot)) + \varepsilon$, then by applying Theorem 1.3.1, we obtain for any $u \in [0,1]$.

$$E\int_0^1 (u^{\varepsilon}(t) - 1)(u(t) - u^{\varepsilon}(t))\,dt \geq -C\varepsilon^{\lambda}. \tag{51}$$

For example, a simple computation shows that the admissible control $u^{\varepsilon}(t) = 1-\varepsilon$, satisfies the above inequality, where $\varepsilon > 0$ is sufficiently small. Conversely, for the sufficient part, let $u^{\varepsilon}(t) = 1 - \varepsilon$ which satisfy (51) candidate to be ε-optimal. Since H is convex in u and by using *Theorem 1.4.2* it follows that $u^{\varepsilon}(t)$ satisfies inequality (51), which means that $u^{\varepsilon}(\cdot)$ is ε-optimal for our control problem (46)-(47), and its corresponding trajectory is

$$x^{\varepsilon}(t) = \frac{1}{2} - (1-\varepsilon)t + (1-\varepsilon)W(t).$$

9. Concluding remarks and future research

In this chapter, necessary and sufficient conditions for near-optimal control with $\varepsilon^{\lambda}-$error bound for SDEs have been established. Linear quadratic control problem has been studied to illustrate our theoretical results. If we assume that $\varepsilon = 0$, our maximum principle (*Theorem 1.3.1*) reduces to maximum principle of optimality developed in Benssoussan [6].

An open questions are to establish necessary and sufficient conditions for near-optimality with $\varepsilon^{\lambda}-$error bound for SDEs with impulse control, Linear quadratic stochastic control with $\varepsilon^{\lambda}-$error bound for SDEs with impulse and SDEs with random jumps. We will work for this interesting issue in the future research.

On Zhou's maximum principle for near optimal control of mean-field forward backward stochastic systems with jumps and its applications

Part III
On Zhou's maximum principle for near optimal control of mean-field forward backward stochastic systems with jumps and its applications

Abstract. This chapter is concerned with stochastic maximum principle for near-optimal control of nonlinear controlled mean-field forward-backward stochastic systems driven by Bownian motions and random Poisson martingale measure (FBSDEJs in short) where the coefficients depend on the state of the solution process as well as on its marginal law through its expected value. Necessary conditions of near-optimality are derived where the control domain is non-convex. Under some additional hypotheses, we prove that the near-maximum condition on the Hamiltonian function in integral form is a sufficient condition for ε-optimality. Our result is derived by using spike variation method, Ekeland's variational principle and some estimates of the state and adjoint processes, along with Clarke's generalized gradient for nonsmooth data. This work extends the results obtained in (Zhou, X.Y.: SIAM J. Control Optim. **36**(3), 929–947, 1998) to a class of mean-field stochastic control problems involving mean-field FBSDEJs. As an application, mean-variance portfolio selection mixed with a recursive utility functional optimization problem is discussed to illustrate our theoretical results.

Keywords. Maximum principle. Stochastic near-optimal control. Mean-field forward-backward stochastic differential equations with jumps. Necessary and sufficient conditions of near-optimality. Ekeland's variational principle.

10. Introduction

The mean-field stochastic systems have attracted much attention because of their practical applications in many areas such as physics, chemistry, economics, finance and other areas of science and engineering. Discrete-time indefinite mean-field linear-quadratic optimal control problem has been investigated in Ni, Zhang and Li [41]. In a recent work, mean-field games for large population multiagent systems with Markov jump parameters have been investigated in Wang and Zhang [56]. Decentralized tracking-type games for large population multi-agent systems with mean-field coupling have been studied in Li and Zhang [42]. Mean-field stochastic control problems have been investigated by many authors, see for instance, [41, 62, 15, 17, 23, 24, 18, 25, 10, 5, 49, 26, 51, 52, 67, 66]. Mean-field type stochastic maximum principle for optimal control under partial information has been investigated in Wang, Zhang and Zhang [62]. Discrete time mean-field stochastic linear-quadratic optimal control problems with applications have been investigated in Elliott, Li and Ni [15]. Second order necessary and sufficient conditions of near-optimal singular control for mean-field SDE were established in Hafayed and Abbas [17]. Mean-field type stochastic maximum principle for optimal singular control has been studied in Hafayed [23], where convex perturbation was used for both absolutely continuous and singular components. The maximum principle for optimal control of mean-field FBSDEJs has been studied in Hafayed [24]. The necessary and sufficient conditions for near-optimality for mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [18]. Singular optimal control for mean-field forward-backward stochastic systems driven by Brownian motions has been investigated in Hafayed [25]. A general mean-field maximum principle was introduced in Buckdahn, Djehiche and Li [10]. Under the conditions that the control domains are convex, a various local maximum principle have been studied in [5, 49]. Second-order maximum principle for optimal stochastic control for mean-field jump diffusions was proved in Hafayed and Abbas [26]. Necessary and sufficient conditions for controlled jump diffusion with recent application in bicriteria mean-variance portfolio selection problem have been proved in Shen and Siu [51]. Recently, maximum principle for mean-field jump-diffusions stochastic delay differential equations and its applications to finance have been investigated in Yang, Meng and Shi [52]. A linear quadratic optimal control problem for mean-field stochastic differential equations has been studied in Yong [67]. Mean-field optimal control for backward stochastic evolution equations in Hilbert spaces have been investigated in Xu and Wu [66]. In Buckdahn, Djehiche, Li and Peng [8] a general notion of mean-field BSDE associated with a mean-field SDE is obtained in a natural way as a limit of some high dimensional system of FBSDEs governed by a $d-$dimensional Brownian motion, and influenced by positions of a large number of other particles.

Near-optimization is as sensible and important as optimization for both theory and applications. The theory of stochastic near-optimization was introduced by Zhou [71]. Various kinds of near-optimal stochastic control problems have been investigated in [17, 18, 19, 20, 21, 28, 57, 36, 70, 35]. The necessary and sufficient conditions of near-optimal mean-field singular stochastic control have been studied in Hafayed and Abbas [17]. The necessary and sufficient conditions for near-optimality for mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [18]. Near-optimality necessary and sufficient conditions for singular controls in jump diffusion processes have been investigated in Hafayed and Abbas [19]. The near-optimal stochastic control problem for jump diffusions has been investigated by Hafayed, Abbas and Veverka [21]. The near-optimality necessary and sufficient conditions for classical controlled FBSDEJs with applications to finance have been investigated in Hafayed, Veverka and Abbas [28]. Stochastic maximum principle of near-optimal control of fully coupled forward-backward stochastic differential equation has been investigated in Tang [57]. Near-optimal control problem for linear FBSDE have been studied in Huang, Li and Wang [36]. Near-optimal stochastic control problem for linear general controlled FBSDEs has been studied in Zhang, Huang and Li [70]. The near-optimal control problem for recursive stochastic problem has been studied in Hui, Huang, Li and Wang [35].

It is shown that the near-optimal controls in mean-field stochastic control problems, as the alternative to the exact optimal ones, are of great importance for both the theoretical analysis and practical application purposes due to its nice structure and broad-range availability as well as feasibility. The near-optimal controls in mean-field stochastic control problems are more available than the exact optimal ones, in the sense that the near-optimal controls always exist, while the exact optimal stochastic controls may not even exist in many situations. Moreover, since there are many near-optimal controls, it is possible to select among them appropriate ones that are easier for analysis and implementation. This justifies the use of near-optimal stochastic controls, which exist under minimal hypothesis and are sufficient in most practical cases.

Motivated by the arguments above and inspired by [71], our aim in this work is to establish a set of necessary conditions for near-optimality for systems governed by nonlinear controlled mean-field FBSDEJs. Moreover, we prove that under some additional assumptions and by applying Clarke's generalized gradient for nonsmooth functions, these necessary conditions are also sufficient for near-optimality. As an illustration, mean-variance portfolio selection problem: time-inconsistent solution is discussed.

The plan of the rest of the chaper is organized as follows. In Section 2, we present some typical notations and formulate the mean-field stochastic control problem considered in this work. In Section 3 we prove our main results. As an illustration, Time-inconsistent mean-variance portfolio selection problem is discussed in the last section.

11. Formulation of the problem and preliminaries

In the present work, we consider mean-field stochastic near-optimal control problem of the following kind. Let $T>0$ be a fixed time horizon and $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$ be a fixed filtered probability space equipped with a $\mathbb{P}-$completed right continuous filtration on which a $d-$dimensional Brownian motion $W=(W(t))_{t\in[0,T]}$ is defined. Let η be a homogeneous $(\mathcal{F}_t)$-Poisson point process independent of W. We denote by $\widetilde{N}(d\theta,dt)$ the random counting measure induced by η, defined on $\Theta\times\mathbb{R}_+$, where Θ is a fixed nonempty subset of $\mathbb{R}$ with its Borel $\sigma-$field $\mathcal{B}(\Theta)$. Further, let $\mu(d\theta)$ be the local characteristic measure of η, i.e. $\mu(d\theta)$ is a σ-finite measure on $(\Theta,\mathcal{B}(\Theta))$ with $\mu(\Theta)<+\infty$. We then define $N(d\theta,dt)=\widetilde{N}(d\theta,dt)-\mu(d\theta)\,dt$, where $N(\cdot,\cdot)$ is Poisson martingale measure on $\mathcal{B}(\Theta)\times\mathcal{B}(\mathbb{R}_+)$ with local characteristics $\mu(d\theta)\,dt$. We assume that $(\mathcal{F}_t)_{t\in[0,T]}$ is $\mathbb{P}-$augmentation of the natural filtration $(\mathcal{F}_t^{(W,N)})_{t\in[0,T]}$ defined as follows:

$$\mathcal{F}_t^{(W,N)}=\sigma\left\{W(s):0\leq s\leq t\right\}\vee\sigma\left\{\int_0^s\int_B N(d\theta,dr),0\leq s\leq t,\ B\in\mathcal{B}(\Theta)\right\}\vee\mathcal{G}_0,$$

where $\mathcal{G}_0$ denotes the totality of $\mathbb{P}-$null sets, and $\sigma_1\vee\sigma_2$ denotes the σ-field generated by $\sigma_1\cup\sigma_2$.

In the present work, we study stochastic near-optimal control problem for system described by mean-field forward-backward stochastic differential equations with Poisson jumps processes (mean-field

FBSDEJs) of the form:

$$\begin{cases} dx(t) = f\left(t, x(t), E(x(t)), u(t)\right) dt + \sigma\left(t, x(t), E(x(t))\right) dW(t) \\ \qquad + \int_{\Theta} c\left(t, x(t_{-}), \theta\right) N\left(d\theta, dt\right). \\ dy(t) = -\int_{\Theta} g(t, x(t), E(x(t)), y(t), E(y(t)), z(t), E(z(t)), r\left(t, \theta\right), u(t))\mu\left(d\theta\right) dt \\ \qquad + z(t) dW(t) + \int_{\Theta} r\left(t, \theta\right) N\left(d\theta, dt\right). \\ x(0) = \zeta,\ y(T) = h\left(x(T), E\left(x(T)\right)\right), \end{cases} \tag{52}$$

where $f, \sigma, b, g, h,$ are given maps and the initial condition ζ is an $\mathcal{F}_0$-measurable random variable. The mean-field FBSDEJs-(52), called McKean-Vlasov systems are obtained as a limit approach, by the mean-square limit, when $n \to +\infty$ of a system of interacting particles of the form:

$$\begin{cases} dx_n^j(t) = f(t, x_n^j(t), \frac{1}{n}\sum_{i=1}^n x_n^i(t), u(t))dt + \sigma(t, x_n^j(t), \frac{1}{n}\sum_{i=1}^n x_n^i(t))dW^j(t) \\ \qquad + \int_{\Theta} c(t, x_n^j(t_{-}), \theta) N^j\left(d\theta, dt\right). \\ dy_n^j(t) = -\int_{\Theta} g(t, x_n^j(t), \frac{1}{n}\sum_{i=1}^n x_n^i(t), y_n^j(t), \frac{1}{n}\sum_{i=1}^n y_n^i(t), z_n^j(t), \frac{1}{n}\sum_{i=1}^n z_n^i(t), \\ \qquad r(t, \theta), u(t))\, \mu\left(d\theta\right) dt + z_n^j(t) dW^j(t) + \int_{\Theta} r\left(t, \theta\right) N^j\left(d\theta, dt\right), \end{cases}$$

where $(W^j(\cdot) : j \geq 1)$ is a collection of independent Brownian motions and $(N^j(\cdot, \cdot) : j \geq 1)$ is a collection of independent Poisson martingale measure. Noting that mean-field FBSDEJs-(52) occur naturally in the probabilistic analysis of financial optimization problems and the optimal control of dynamics of the McKean-Vlasov type. Moreover, the above mathematical mean-field approaches play an important role in different fields of economics, finance, physics, chemistry and game theory.
The criteria to be minimized associated with the state equation (52) is defined by

$$J\left(\zeta, u(\cdot)\right) = E\left\{\phi\left(y(0), E\left(y(0)\right)\right)\right\}. \tag{53}$$

It's worth mentioning that since the cost functional J is possibly a nonlinear function of the expected value stands in contrast to the standard formulation of a control problem. This leads to a so-called time-inconsistent control problem where the Bellman Dynamic programming does not hold. The reason for this is that one cannot apply the law of iterated expectations on the cost functional.
This section sets out the notations and assumptions used in the sequel.

Notations. We use the following notations. In the sequel, $\mathbb{L}^2_{\mathcal{F}}\left([0, T]; \mathbb{R}^n\right)$ denotes the Hilbert space of $\mathcal{F}_t-$adapted processes $(x(t))_{t \in [0,T]}$ such that $E\int_0^T |x(t)|^2 dt < +\infty$. $\mathbb{M}^2_{\mathcal{F}}\left([0, T]; \mathbb{R}\right)$ denotes the Hilbert space of $\mathcal{F}_t-$ predictable processes $\left(\psi\left(\cdot, t, \theta\right)\right)_{t \in [0,T]}$ defined on $\Omega \times [0, T] \times \Theta$ such that $E\int_0^T \int_{\Theta} |\psi\left(w, t, \theta\right)|^2 \mu(\theta) dt < +\infty$. Any element $x \in \mathbb{R}^n$ will be identified to a column vector with its j^{th} component x_j and the norm $|x| = \sum_{j=1}^n |x_j|$. We denote A^* the transpose of any vector or matrix A. We denote by E the expectation with respect to $\mathbb{P}$. For a function $f \in \mathcal{C}^1$ we denote by f_x its gradient or Jacobian with respect to the variable x. We denote by $\mathbf{1}_{\mathcal{A}}$ the indicator function of $\mathcal{A}$, $conv\left(\mathcal{A}\right)$ the closure convex hull of $\mathcal{A}$ and $Sgn(\cdot)$ be the sign function. We denote by $(\widetilde{x}(t), \widetilde{y}(t), \widetilde{z}(t)) = (E(x(t)), E(y(t)), E(z(t)))$.

Definition 2.1.1 Let $T > 0$ be a fixed strictly positive real number and $\mathbb{U}$ be a nonempty subset of $\mathbb{R}^k$. An admissible control is defined as a function $u(\cdot) : [0,T] \times \Omega \longrightarrow \mathbb{U}$ which is $\mathcal{F}_t-$predictable, such that the mean–field FBSDEJs-(52) has a unique solution and write $u(\cdot) \in \mathcal{U}([0,T])$.
The value function is defined as

$$V(\zeta) = \inf \left\{ J\left(\zeta, u(\cdot)\right) : u(\cdot) \in \mathcal{U}([0,T]) \right\}. \tag{54}$$

Throughout this work, we also assume that the coefficients functions:

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

satisfy the following standard assumptions:
Assumption (H1) The functions f, σ, g, h, c, ϕ are continuous and continuously differentiable with respect to $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r)$, and there exists a constant $C > 0$ such that $|f(t, x, \widetilde{x}, u)| + |\sigma(t, x, \widetilde{x})| < C(1 + |x| + |\widetilde{x}|)$, $\sup_{\theta \in \Theta} |c(t, x, \theta)| < C(1 + |x|)$.
Assumption (H2) The derivatives of f, σ, c, h, ϕ with respect to $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r)$ are bounded and there is a constant $C > 0$ such that $\sup_{\theta \in \Theta} |g_{\varkappa}(t, \theta)| < C$ for $\varkappa = x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r$.
Assumption (H3) There is a constant $C > 0$ and $\gamma \in [0,1]$ such that

$$\begin{aligned}
&\left|f_x\left(t, x, \widetilde{x}, u\right) - f_x\left(t, x', \widetilde{x}', u\right)\right| + \left|f_{\widetilde{x}}\left(t, x, \widetilde{x}, u\right) - f_{\widetilde{x}}\left(t, x', \widetilde{x}', u\right)\right|\\
&+ \left|\sigma_x\left(t, x, \widetilde{x}\right) - \sigma_x\left(t, x', \widetilde{x}'\right)\right| + \left|\sigma_{\widetilde{x}}\left(t, x, \widetilde{x}\right) - \sigma_{\widetilde{x}}\left(t, x', \widetilde{x}'\right)\right| \leq C(|x - x'|^\gamma + |\widetilde{x} - \widetilde{x}'|^\gamma).\\
&\left|h_x\left(x, \widetilde{x}\right) - h_x\left(x', \widetilde{x}'\right)\right| + \left|h_{\widetilde{x}}\left(x, \widetilde{x}\right) - h_{\widetilde{x}}\left(x', \widetilde{x}'\right)\right| \leq C\left(|x - x'|^\gamma + |\widetilde{x} - \widetilde{x}'|^\gamma\right).
\end{aligned}$$

Further,

$$\begin{aligned}
&\left|g_{\varkappa}\left(t, x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, u, \theta\right) - g_{\varkappa}\left(t, x', \widetilde{x}', y', \widetilde{y}', z', \widetilde{z}', u, \theta\right)\right|\\
&\leq C(|x - x'|^\gamma + |\widetilde{x} - \widetilde{x}'|^\gamma + |y - y'|^\gamma + |\widetilde{y} - \widetilde{y}'|^\gamma + |z - z'|^\gamma + |\widetilde{z} - \widetilde{z}'|^\gamma),
\end{aligned}$$

where $\varkappa = x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}$.
Under the Assumptions (H1) and (H2) the mean–field FBSDEJs-(52) has a unique solution $(x(t), y(t), z(t), r(\cdot,\cdot)) \in \mathbb{L}^2_{\mathcal{F}}\left([0,T]; \mathbb{R}^n\right) \times \mathbb{L}^2_{\mathcal{F}}\left([0,T]; \mathbb{R}^m\right) \times \mathbb{L}^2_{\mathcal{F}}\left([0,T]; \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m)\right) \times \mathbb{M}^2_{\mathcal{F}}\left([0,T]; \mathbb{R}\right)$, (see Hafayed [24]).
Adjoint equations. For any $u(\cdot) \in \mathcal{U}([0,T])$ with its corresponding state trajectories $\left(x\left(\cdot\right), y\left(\cdot\right), z\left(\cdot\right), r(\cdot,\cdot)\right)$ we introduce the following adjoint equations, which differ from the classical ones in the sense that here the adjoint equation turns out to be a linear mean-field forward-backward

stochastic differential equations with jumps

$$\begin{cases} d\Psi(t) = -\{f_x(t)\,\Psi(t) + E\left[f_{\widetilde{x}}(t)\,\Psi(t)\right] + \sigma_x(t)\,Q(t) + E\left[\sigma_{\widetilde{x}}(t)\,Q(t)\right] \\ \qquad + \int_\Theta \left[g_x(t,\theta)K(t) + E\left(g_{\widetilde{x}}(t,\theta)K(t)\right) + c_x(t,\theta)\,R(t,\theta)\right]\mu(d\theta)\}dt \\ \qquad + Q(t)dW(t) + \int_\Theta R(t,\theta)\,N(d\theta,dt). \\ \Psi(T) = -\left\{h_x(T)\,K(T) + E\left[\left(h_{\widetilde{x}}(T)\right)K(T)\right]\right\}. \\ dK(t) = \int_\Theta \left[g_y(t,\theta)\,K(t) + E\left(g_{\widetilde{y}}(t,\theta)\,K(t)\right)\right]\mu(d\theta)\,dt \\ \qquad + \int_\Theta \left[g_z(t,\theta)K(t) + E\left(g_{\widetilde{z}}(t,\theta)K(t)\right)\right]\mu(d\theta)\,dW(t) - \int_\Theta g_r(t,\theta)K(t)N(d\theta,dt). \\ K(0) = -\phi_y\left(y(0),E\left(y(0)\right)\right) - E\left(\phi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)\right), \end{cases} \tag{55}$$

where
$f_\varkappa(t) = f_\varkappa(t,x(t),\widetilde{x}(t),u(t))$, $\sigma_\varkappa(t) = \sigma_\varkappa(t,x(t),\widetilde{x}(t))$, $h_\varkappa(t) = h_\varkappa(x(t),\widetilde{x}(t))$, for $\varkappa := x,\widetilde{x}$.
$\phi_\varkappa(t) = \phi_\varkappa(y(t),\widetilde{y}(t))$, for $\varkappa := y,\widetilde{y}$.
$g_\varkappa(t,\theta) = g_\varkappa(t,x(t),\widetilde{x}(t),y(t),\widetilde{y}(t),z(t),\widetilde{z}(t),u(t),\theta)$, for $\varkappa := x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r$.
Hamiltonian function. We define the Hamiltonian function

$$H:[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m\times\mathcal{L}(\mathbb{R}^d,\mathbb{R}^m)\times\mathcal{L}(\mathbb{R}^d,\mathbb{R}^m)\times\mathbb{U}\times\mathbb{R}^n\times\mathbb{R}^m\times\mathcal{L}\left(\mathbb{R}^d,\mathbb{R}^m\right)\to\mathbb{R}^n,$$

associated with the mean-field stochastic control problem (52)-(53) as follows

$$\begin{aligned} &H\left(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r(\cdot,\cdot),u,\Psi,Q,K,R(\cdot,\cdot)\right) \\ &= -\Psi(t)f\left(t,x,\widetilde{x},u\right) - Q(t)\sigma\left(t,x,\widetilde{x}\right) - \int_\Theta \left[K(t)g\left(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r(t,\theta),u\right)\right. \\ &\left. - R(t,\theta)\,c(t,x,\theta))\right]\mu(d\theta). \end{aligned} \tag{56}$$

If we denote by

$$H(t) = H(t,x(t),\widetilde{x}(t),y(t),\widetilde{y}(t),z(t),\widetilde{z}(t),r(t,\theta),u(t),\Psi(t),Q(t),K(t),R(t,\theta)),$$

the adjoint equation (55) can be rewritten as the following stochastic Hamiltonian system's type

$$\begin{cases} d\Psi(t) = \left\{H_x(t) + E\left[H_{\widetilde{x}}(t)\right]\right\}dt + Q(t)dW(t) + \int_\Theta R(t,\theta)\,N(d\theta,dt). \\ \Psi(T) = -\left[h_x(T) + E\left(h_{\widetilde{x}}(T)\right)\right]K(T). \\ -dK(t) = \left[H_y(t) + E\left(H_{\widetilde{y}}(t)\right)\right]dt + \left[H_z(t) + E\left(H_{\widetilde{z}}(t)\right)\right]dW(t)\int_\Theta H_r(t_-)\,N(d\theta,dt). \\ K(0) = -\left\{\varphi_y\left(y(0),E\left(y(0)\right)\right) + E\left[\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)\right]\right\}. \end{cases} \tag{57}$$

It is a well known fact that under Assumptions (H1) and (H2), the adjoint equation (55) admits a unique solution $(\Psi(t),Q(t),K(t),R(t,\cdot))$ such that

$$(\Psi(t),Q(t),K(t),R(t,\cdot))\in\mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R})\times\mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R})\times\mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R})\times\mathbb{M}^2_{\mathcal{F}}([0,T];\mathbb{R}).$$

Moreover, since the derivatives of $f_x, f_{\widetilde{x}}, \sigma_x, \sigma_{\widetilde{x}}, c_x, g_x,\ g_{\widetilde{x}}, g_y, g_{\widetilde{y}},\ g_z, g_{\widetilde{z}},\ g_r, h_x, h_{\widetilde{x}},\ \varphi_y$ and $\varphi_{\widetilde{y}}$ are bounded, we deduce from standard arguments that there exists a constant $C > 0$ such that

$$E\left\{\sup_{t\in[0,T]}|\Psi(t)|^2 + \sup_{t\in[0,T]}|K(t)|^2\,dt + \int_0^T |Q(t)|^2\,dt + \int_0^T\int_\Theta |R(t,\theta)|^2\,\mu(d\theta)\,dt\right\} < C. \tag{58}$$

Let us recall the definition of near-optimal control as given in (Zhou [71] Definition 2.1, and Definition 2.2) and Ekeland's variational principle which will be used in the sequel.

Definition 2.2.2. (*Near-optimal control of order ε^δ*) For a given $\varepsilon > 0$ the admissible control $u^\varepsilon(\cdot)$ is called near-optimal if

$$|J(\zeta, u^\varepsilon(\cdot)) - V(\zeta)| \leq \mathcal{O}(\varepsilon), \tag{59}$$

where $\mathcal{O}(\cdot)$ is a function of ε satisfying $\lim_{\varepsilon\to 0}\mathcal{O}(\varepsilon) = 0$. The estimator $\mathcal{O}(\varepsilon)$ is called an *error bound.*

1. If $\mathcal{O}(\varepsilon) = C\varepsilon^\delta$ for $\delta > 0$, where C and δ are independent of ε, then $u^\varepsilon(\cdot)$ is called near-optimal control of order ε^δ.
2. If $\mathcal{O}(\varepsilon) = \varepsilon$, the admissible control $u^\varepsilon(\cdot)$ called $\varepsilon-$optimal.

Lemma 2.2.1. (*Ekeland's Variational Principle*) Let (F, ρ) be a complete metric space and $f : F \to \overline{\mathbb{R}}$ be a lower semi-continuous function which is bounded from below. For a given $\varepsilon > 0$, suppose that there is $u^\varepsilon \in F$ satisfying $f(u^\varepsilon) \leq \inf_{u\in F} f(u) + \varepsilon$. Then for any $\lambda > 0$ there exists $u^\lambda \in F$ such that

1. $f\left(u^\lambda\right) \leq f(u^\varepsilon)$;
2. $\rho\left(u^\lambda, u^\varepsilon\right) \leq \lambda$;
3. $f\left(u^\lambda\right) \leq f(u) + \dfrac{\varepsilon}{\lambda}\rho\left(u, u^\lambda\right)$, *for all* $u \in F$.

To apply Ekeland's variational principle to our mean-field control problem, we must define a metric ρ on the space of admissible controls such that $(\mathcal{U}([0,T]), \rho)$ becomes a complete metric space. For any $u(\cdot)$, $v(\cdot) \in \mathcal{U}([0,T])$ we define

$$\rho(u(\cdot), v(\cdot)) = \mathbb{P}\otimes dt\,\{(\omega,t) \in \Omega\times[0,T] : u(\omega,t) \neq v(\omega,t)\}, \tag{60}$$

where $\mathbb{P}\otimes dt$ is the product measure of $\mathbb{P}$ with the Lebesgue measure dt on $[0,T]$.

Lemma 2.2.1. $(\mathcal{U}([0,T]), \rho)$ is a complete metric space.

2) The cost function $J(\cdot)$ is continuous from $\mathcal{U}([0,T])$ into $\mathbb{R}$.

See Yong and Zhou ([69], *Lemma 6.4,* pp. 146-147).

Proof. 1. See Yong and Zhou ([69], *Lemma 6.4,* pp. 146-147).

2. From (53), we have

$$\begin{aligned} |J(\zeta, u(\cdot)) - J(\zeta, v(\cdot))| &= |E[\phi(y^u(0), \widetilde{y}^u(0))] - E[\phi(y^v(0), \widetilde{y}^v(0))]| \\ &\leq E\,|\phi(y^u(0), \widetilde{y}^u(0)) - \phi(y^v(0), \widetilde{y}^v(0))|. \end{aligned}$$

Since $\phi(\cdot,\cdot)$ is continuously differentiable with respect to $y, \widetilde{y}$ with bounded derivatives, we get

$$|J(\zeta, u(\cdot)) - J(\zeta, v(\cdot))| \leq C\rho(u(\cdot), v(\cdot))^{\frac{1}{2}}.$$

Now, let $(u_n)_{n\geq 0}$ be a sequace of controls convergs to u in $(\mathcal{U}([0,T]), \rho)$,.then we have

$$|J(\zeta, u_n(\cdot)) - J(\zeta, u(\cdot))| \leq C\rho(u_n(\cdot), u(\cdot))^{\frac{1}{2}},$$

since $\rho(u_n(\cdot), u(\cdot)) \to 0$ as $n \to +\infty$, then $J(\zeta, u_n(\cdot))$ converges to $J(\zeta, u(\cdot))$ as $n \to +\infty$. □

12. Main results

12.1. Maximum principle of near-optimality for mean-field FBSDEJs

Our purpose in this section is to derive a set necessary conditions of near-optimality in the form of maximum principle for systems governed by nonlinear controlled mean–field FBSDE with Jumps. The proof of our main result is based on Ekeland's variational principle [14] and some estimates of the state and adjoint processes with respect to the control variable.
Now we are able to derive necessary conditions for a control to be near-optimal for systems governed by mean-field FBSDEJs, which is the main result of this work.
Let $(x^{\varepsilon}(\cdot), y^{\varepsilon}(\cdot), z^{\varepsilon}(\cdot), r^{\varepsilon}(\cdot))$ be the solution of state equation (52) and $(\Psi^{\varepsilon}(\cdot), Q^{\varepsilon}(\cdot), K^{\varepsilon}(\cdot), R^{\varepsilon}(\cdot))$ be the solution of the adjoint equation (55) corresponding to $u^{\varepsilon}(\cdot)$.

Theorem 2.3.1. *(Necessary Conditions of Near-optimality for mean–field FBSDEJs in integral form)* Let the assumptions (**H1**), (**H2**) and (**H3**) hold. Then for any $\delta \in \left[0, \frac{1}{3}\right[$, there exists a positive constant $C = C(\delta, T) > 0$ such that for any $\varepsilon > 0$ and any near-optimal control $u^{\varepsilon}(\cdot)$, it holds that $\forall u \in \mathbb{U}$:

$$\begin{aligned} &E\int_0^T [H(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u, \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) \\ &-H(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta))]dt \geq -C\varepsilon^{\delta}. \end{aligned} \tag{61}$$

where $(\Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta))) = (x^{\varepsilon}(t), E(x^{\varepsilon}(t)), y^{\varepsilon}(t), E(y^{\varepsilon}(t)), z^{\varepsilon}(t), E(z^{\varepsilon}(t)), r^{\varepsilon}(t,\theta))$.

Corollary 2.3.1. Under the hypotheses of *Theorem 3.1*, it holds that

$$\begin{aligned} &H(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) \\ &\geq \sup_{u(\cdot)\in\mathcal{U}([0,T]} H(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u(\cdot), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) - C\varepsilon^{\delta}. \\ &\quad \mathbb{P}-a.s.,\ a.e.\ t \in [0,T]. \end{aligned} \tag{62}$$

Remark 2.3.1 Note that Corollary 3.1 says that any $\varepsilon-$optimal control nearly maximizes the Hamiltonian functional with an error bound of order of $\varepsilon^{\frac{1}{3}}$ we believe, although we are not able to prove at this moment, that the error bound can be improved. To prove our mean-field maximum principle (*Theorem 3.1* and *Corollary 3.1)*, we need the following auxiliary results on the stability of the state and adjoint processes with respect to the control variable.
Our first Lemma below deals with the continuity of the state processes under distance ρ.

Lemma 2.3.1. (*Continuity Lemma*) If $(x^{u}(\cdot), y^{u}(\cdot), z^{u}(\cdot), r^{u}(\cdot,\cdot))$ and $(x^{v}(\cdot), y^{v}(\cdot), z^{v}(\cdot), r^{v}(\cdot,\cdot))$ be the solution of the state equation (52) associated respectively with $u(\cdot)$ and $v(\cdot)$. For any $\alpha \in]0,1[$ and $\beta \in]0,2]$ satisfying $\alpha\beta < 1$, there exists a positive constants $C = C(T, \alpha, \beta, \mu(\Theta))$ such that

$$E(\sup_{s\leq t\leq T} |x^{u}(t) - x^{v}(t)|^{\beta}) \leq C\rho(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \tag{63}$$

$$\begin{aligned} &\sup_{0\leq t\leq T} E(|y^{u}(t) - y^{v}(t)|^{\beta}) + E\int_t^T |z^{u}(s) - z^{v}(s)|^{\beta} ds \\ &+ \int_t^T \int_{\Theta} |r^{u}(s,\theta) - r^{v}(s,\theta)|^{\beta} \mu(d\theta)\, ds \Big\} \leq C\rho(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{64}$$

Proof.

Proof of estimate (63).

Case 1. First, we assume that $\beta \in [1,2]$. We can compute, for any $r \geq s$, with the helps of (Proposition A2, Appendix), we get

$$\begin{aligned}
& E\left[\sup_{s\leq t\leq r} |x^u(t) - x^v(t)|^\beta\right] \\
\leq\ & CE\int_s^r \{|f(t, x^u(t), E(x^u(t)), u(t)) - f(t, x^u(t), E(x^u(t)), v(t))|^\beta dt \\
& + |\sigma(t, x^u(t), E(x^u(t))) - \sigma(t, x^v(t), E(x^v(t)))|^\beta \\
& + \int_\Theta |c(t, x^u(t), \theta) - c(t, x^v(t), \theta)|^\beta \mu(d\theta)\Big\} \times \mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\, dt \\
& + CE\int_s^r \{|f(t, x^u(t), E(x^u(t)), v(t)) - f(t, x^v(t), E(x^v(t)), v(t))|^\beta \\
& + |\sigma(t, x^u(t), E(x^u(t))) - \sigma(t, x^v(t), E(x^v(t)))|^\beta\Big\} dt.
\end{aligned}$$

Setting $b = \frac{2}{\alpha\beta} > 1$ and $a > 1$ such that $\frac{1}{a} + \frac{1}{b} = 1$. By arguing as in Zhou [71], *Lemma 3.1*) and from *Cauchy-Schwartz inequality*, we get

$$\begin{aligned}
& E\int_s^r |f(t, x^u(t), E(x^u(t)), u(t)) - f(t, x^u(t), E(x^u(t)), v(t))|^\beta \mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\, dt \\
\leq\ & \left\{E\int_s^r |f(t, x^u(t), E(x^u(t)), u(t)) - f(t, x^u(t), E(x^u(t)), v(t))|^{\beta a} dt\right\}^{\frac{1}{a}} \\
& \times \left\{E\int_s^r \mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\, dt\right\}^{\frac{1}{b}},
\end{aligned}$$

by using definition of ρ and linear growth condition on f, we obtain

$$\begin{aligned}
& E\int_s^r |f(t, x^u(t), E(x^u(t)), u(t)) - f(t, x^u(t), E(x^u(t)), v(t))|^\beta \mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\, dt \\
& \leq C\left\{E\int_s^r (1 + |x^u(t)|^{2\beta a} + |E(x^u(t))|^{2\beta a}) dt\right\}^{\frac{1}{a}} \rho(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}} \\
& \leq C\rho(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}.
\end{aligned}$$

Similarly, the same inequality holds if f above is replaced by σ and c we get

$$\begin{aligned}
& E\int_0^r |\sigma(t, x^u(t), E(x^u(t))) - \sigma(t, x^v(t), E(x^v(t)))|^\beta \mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\, dt \\
& \leq C\rho(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \\
& E\int_0^r \int_\Theta |c(t, x^u(t_-), \theta) - c(t, x^v(t_-), \theta)|^\beta \mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\, \mu(d\theta)\, dt \\
& \leq C\rho(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}.
\end{aligned}$$

Therefore, by using Assumption (H1), we conclude that

$$E(\sup_{0\leq t\leq r}|x^u(t)-x^v(t)|^\beta)\leq CE\int_s^r\sup_{0\leq r\leq \tau}|x^u(t)-x^v(t)|^\beta d\tau+\rho\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}}.$$

Hence (63) follows immediately from *Definition 2.1* and *Gronwall's inequality*.
Case 2. Now we assume $\beta\in\,]0,1[$. By using the *Cauchy-Schwartz inequality* together with *Case 1.*, we get

$$\begin{aligned}E(\sup_{0\leq t\leq T}|x^u(t)-x^v(t)|^\beta)&\leq E(\sup_{0\leq t\leq T}|x^u(t)-x^v(t)|^2)^{\frac{\beta}{2}}\\&\leq\left[C\rho\left(u(\cdot),v(\cdot)\right)^\alpha\right]^{\frac{\beta}{2}}\leq C\rho\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}}.\end{aligned}$$

This completes the proof of inequality (63) □

Proof of estimate (64). Setting

$$(\Lambda^u(t,\theta),E(\Lambda^u(t,\theta))=(x^u(t),E\left(x^u(t)\right),y^u(t),E\left(y^u(t)\right),z^u(t),E\left(z^u(t)\right),r^u(t,\theta)).$$
$$(\Lambda^v(t,\theta),E(\Lambda^v(t,\theta)))=(x^v(t),E\left(x^v(t)\right),y^v(t),E\left(y^v(t)\right),z^v(t),E\left(z^v(t)\right),r^v(t,\theta)).$$

Case 1. First we assume $\beta=2$. From the backward component $(y(\cdot),z(\cdot))$, we get

$$\begin{aligned}&-\left(y^u(t)-y^v(t)\right)-\textstyle\int_t^T(z^u(s)-z^v(s))dW(s)-\int_t^T\int_\Theta\left(r^u(s,\theta)-r^v(s,\theta)\right)N(d\theta,ds)\\&=-\left[h(x^u(T),E(x^u(T)))-h(x^v(T),E(x^v(T)))\right]+\textstyle\int_t^T\int_\Theta\left[g\left(s,\Lambda^u(s,\theta),E(\Lambda^u(s,\theta)),u(s)\right)\right.\\&\left.-\,g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(s,\theta)),v(s)\right)\right]\mu\left(d\theta\right)ds.\end{aligned}$$

By squaring both sides of the above equation (see Hafayed [24]) and the fact that

$$\begin{aligned}&E\left\{\left(y^u(t)-y^v(t)\right)\textstyle\int_t^T\left(z^u(s)-z^v(s)\right)dW(s)\right\}=0.\\&E\left\{\left(y^u(t)-y^v(t)\right)\textstyle\int_t^T\int_\Theta\left(r^u(s,\theta)-r^v(s,\theta)\right)N(d\theta,ds)\right\}=0.\\&E\left\{\textstyle\int_t^T\left(z^u(s)-z^v(s)\right)dW(s)\times\int_t^T\int_\Theta\left(r^u(s,\theta)-r^v(s,\theta)\right)\mu\left(d\theta\right)ds\right\}=0.\end{aligned}$$

with the help of *Proposition A2*, we obtain

$$\begin{aligned}&E\left\{|y^u(t)-y^v(t)|^2\right\}+E\textstyle\int_t^T|z^u(s)-z^v(s)|^2\,ds+E\int_t^T\int_\Theta|r^u(s,\theta)-r^v(s,\theta)|^2\mu\left(d\theta\right)ds\\&\leq E\left\{|h(x^u(T),E(x^u(T)))-h(x^v(T),E(x^v(T)))|^2\right\}\\&+E\left\{\textstyle\int_t^T\int_\Theta[g\left(s,\Lambda^u(s,\theta),E(\Lambda^u(t,\theta)),u(s)\right)-g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(s,\theta)),v(s)\right)]\mu\left(d\theta\right)ds\right\}^2\\&\leq I_1+I_2.\end{aligned}\tag{65}$$

Let us estimate the first term I_1. Using Assumption (H1) then from inequality (63), we get

$$\begin{aligned} I_1 &= E\left\{\left|h(x^u(T),E(x^u(T))) - h(x^v(T),E(x^v(T)))\right|^2\right\} \\ &\leq CE\left\{\left|x^u(T)-x^v(T)\right|^2 + \left|E(x^u(T)) - E(x^v(T))\right|^2\right\} \\ &\leq CE\left\{\left|x^u(T)-x^v(T)\right|^2 + \left|E\left[x^u(T)) - x^v(T)\right]\right|^2\right\} \\ &\leq CE\left\{\left|x^u(T)-x^v(T)\right|^2 + E\left|x^u(T)) - x^v(T)\right|^2\right\} \\ &\leq CE\left\{\left|x^u(T)-x^v(T)\right|^2\right\} \leq C\rho\left(u(\cdot),v(\cdot)\right)^{\alpha}. \end{aligned} \tag{66}$$

Let us turn to estimate the second term I_2. By adding and subtracting $g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(s,\theta)),u(s)\right)$ from I_2 with the help of *Propositions A2*, we get

$$\begin{aligned} &I_2 = E\left\{\int_t^T\int_\Theta \left|g\left(s,\Lambda^u(s,\theta),E(\Lambda^u(t,\theta)),u(s)\right)\ -\ g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(t,\theta)),v(s)\right)\right| \right. \\ &\left.\times\ \mathbf{1}_{\{u(\omega,s)\neq v(\omega,s)\}}\left(s\right)\mu\left(d\theta\right)ds\right\}^2 \\ &\leq CE\left\{\int_t^T\int_\Theta \left|g\left(s,\Lambda^u(s,\theta),E(\Lambda^u(t,\theta)),u(s)\right)\ -\ g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(t,\theta)),u(s)\right)\right|\mu\left(d\theta\right)ds\right\}^2 \\ &+CE\left\{\int_t^T\int_\Theta \left|g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(s,\theta)),u(s)\right)\ -\ g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(t,\theta)),v(s)\right)\right|^2 \right. \\ &\left.\times\ \mathbf{1}_{\{u(\omega,s)\neq v(\omega,s)\}}\left(s\right)\mu\left(d\theta\right)ds\right\} \\ &= I_2^1 + I_2^2. \end{aligned} \tag{67}$$

Using Assumption (H1), we get

$$\begin{aligned} &I_2^1 = CE\left\{\int_t^T\int_\Theta \left|g\left(s,\Lambda^u(s,\theta),E(\Lambda^u(t,\theta)),u(s)\right)\ -g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(t,\theta)),u(s)\right)\right|\mu\left(d\theta\right)ds\right\}^2 \\ &\leq CE\int_t^T\left\{\left|x^u(s)-x^v(s)\right|^2 + \left|E\left(x^u(s)\right)-E\left(x^v(s)\right)\right|^2 + \left|y^u(s)-y^v(s)\right|^2ds + \left|E\left(y^u(s)\right)-E\left(y^v(s)\right)\right|^2 \right. \\ &\left. + \left|z^u(s)-z^v(s)\right|^2 + \left|E\left(z^u(s)\right)-E\left(z^v(s)\right)\right|^2 + \int_\Theta \left|r^u(s,\theta)-z^v(s,\theta)\right|^2\mu\left(d\theta\right)\right\}ds \\ &\leq CE\int_t^T\left|x^u(s)-x^v(s)\right|^2ds + CE\int_t^T\left|y^u(s)-y^v(s)\right|^2ds + C(T-t)E\left[\int_t^T\left|z^u(s)-z^v(s)\right|^2ds\right. \\ &\left. + \int_t^T\int_\Theta\left|r^u(s,\theta)-r^v(s,\theta)\right|^2\mu\left(d\theta\right)ds\right]. \end{aligned} \tag{68}$$

Now, taking $a=\frac{1}{1-\alpha}>1$ and $b=\frac{1}{\alpha}>1$ such that $\frac{1}{a}+\frac{1}{b}=1$, then from *Hölder's inequality* and the fact that g is bounded by $C(1+|x|+|\widetilde{x}|+|y|)$, (see Assumption (H2)) we can shows that

$$\begin{aligned} &I_2^2 = E\left\{\int_t^T\int_\Theta \left|g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(s,\theta)),u(s)\right)\ -\ g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(t,\theta)),v(s)\right)\right|^2 \right. \\ &\left.\times\ \mathbf{1}_{\{u(\omega,s)\neq v(\omega,s)\}}\left(s\right)\mu\left(d\theta\right)ds\right\} \\ &\leq CE\left\{\int_t^T\int_\Theta \left|g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(s,\theta)),u(s)\right)\ -\ g\left(s,\Lambda^v(s,\theta),E(\Lambda^v(t,\theta)),v(s)\right)\right|^{\frac{2}{1-\alpha}}\mu\left(d\theta\right)ds\right\}^{1-\alpha} \\ &\times\left\{E\int_t^T\mathbf{1}_{\{u(\omega,s)\neq v(\omega,s)\}}\left(s\right)ds\right\}^{\alpha}, \end{aligned}$$

which implies

$$
\begin{aligned}
&I_2^2 \le C\mu(\Theta)\left\{1+E\left[\sup_{t\le s\le T}|x^v(s)|^{\frac{2}{1-\alpha}}\right]+E\left[\sup_{t\le s\le T}|E(x^v(s))|^{\frac{2}{1-\alpha}}\right]\right.\\
&\left.+E\left[\sup_{t\le s\le T}|y^v(s)|^{\frac{2}{1-\alpha}}\right]\right\}\left\{E\int_t^T \mathbf{1}_{\{u(\omega,s)\neq v(\omega,s)\}}(s)\,ds\right\}^{\alpha}\\
&\le C_{\mu(\Theta)}\rho\left(u(\cdot),v(\cdot)\right)^{\alpha}.
\end{aligned} \tag{69}
$$

By combining (66)∼(69) together with (65), we get

$$
\begin{aligned}
&E\left|y^u(t)-y^v(t)\right|^2+E\int_t^T\left|z^u(s)-z^v(s)\right|^2ds+E\int_t^T\int_\Theta\left|r^u(s,\theta)-r^v(s,\theta)\right|^2\mu(d\theta)\,ds\\
&\le C\mu(\Theta)\,\rho\left(u(\cdot),v(\cdot)\right)^{\alpha}+C\mu(\Theta)\int_t^T E\left|y^u(s)-y^v(s)\right|^2ds\\
&+C\mu(\Theta)\,(T-t)E\int_t^T\left[\left|z^u(s)-z^v(s)\right|^2+\int_\Theta\left|r^u(s,\theta)-r^v(s,\theta)\right|^2\mu(d\theta)\right]ds.
\end{aligned}
$$

For every $\tau=T-t$ we obtain by choosing $\tau=\dfrac{1}{2C\mu(\Theta)}$ we shows that

$$
\begin{aligned}
&E\left|y^u(t)-y^v(t)\right|^2+\frac{1}{2}E\int_{T-\tau}^T\left|z^u(s)-z^v(s)\right|^2ds\\
&+\frac{1}{2}E\int_{T-\tau}^T\int_\Theta\left|r^u(s,\theta)-r^v(s,\theta)\right|^2\mu(d\theta)\,ds\\
&\le C\mu(\Theta)\,\rho\left(u(\cdot),v(\cdot)\right)^{\alpha}+C\mu(\Theta)\int_{T-\tau}^T E\left|y^u(s)-y^v(s)\right|^2ds.
\end{aligned}
$$

Using *Gronwall's inequality*, $t\in[T-\tau,T]$

$$
\begin{aligned}
&E\left|y^u(t)-y^v(t)\right|^2+\frac{1}{2}E\int_t^T\left|z^u(s)-z^v(s)\right|^2ds\\
&+\frac{1}{2}E\int_t^{T-\tau}\int_\Theta\left|r^u(s,\theta)-r^v(s,\theta)\right|^2\mu(d\theta)\,ds\\
&\le C\mu(\Theta)\,\rho\left(u(\cdot),v(\cdot)\right)^{\alpha},
\end{aligned}
$$

by similar argument, we obtain for $t\in[T-2\tau,T-\tau]$,

$$
\begin{aligned}
&E\left|y^u(t)-y^v(t)\right|^2+E\int_t^{T-\tau}\left|z^u(s)-z^v(s)\right|^2ds+E\int_t^{T-\tau}\int_\Theta\left|r^u(s,\theta)-r^v(s,\theta)\right|^2\mu(d\theta)\,ds\\
&\le C\mu(\Theta)\,\rho\left(u(\cdot),v(\cdot)\right)^{\alpha}.
\end{aligned}
$$

After a finite number of iterations, the desired result follows.

Case 2. First we assume $0 < \beta < 2$. Then by using *Hölder's inequality* and *Case 1*, we get

$$\sup_{0\leq t\leq T} E(\left|y^u(t) - y^v(t)\right|^\beta) + E\int_t^T \left|z^u(s) - z^v(s)\right|^\beta ds + E\int_t^T\int_\Theta \left|r^u(s,\theta) - r^v(s,\theta)\right|^\beta \mu\left(d\theta\right) ds$$
$$\leq C\mu\left(\Theta\right)\left\{\sup_{0\leq t\leq T} E\left|y^u(t) - y^v(t)\right|^2 + E\int_t^T \left|z^u(s) - z^v(s)\right|^2 ds\right.$$
$$\left.+E\int_t^T\int_\Theta \left|r^u(s,\theta) - r^v(s,\theta)\right|^2 \mu\left(d\theta\right) ds\right\}^{\frac{\beta}{2}} \leq C\left\{\rho\left(u(\cdot),v(\cdot)\right)^\alpha\right\}^{\frac{\beta}{2}}.$$

This completes the proof of (64) $\square$

Since the adjoint equations corresponding to our mean-field control problem (52)-(53) are given by forward-backward stochastic system of mean-type, the next result gives the β^{-th} moment continuity of the solutions to adjoint equations with respect to the metric ρ, This Lemma may be considered as an extension of *Lemma 3.3.2* in [71], to mean-field FBSDEs with jumps.

Lemma 2.3.2. For any $\alpha \in \left]0,1\right[$ and $\beta \in \left]0,2\right]$ satisfying $(1+\alpha)\beta < 2$, there exist a positive constant $C = C\left(\mu(\Theta),\alpha,\beta\right)$ such that for any $u(\cdot),\ v(\cdot) \in \mathcal{U}([0,T])$, along with the corresponding trajectories $(x^u(\cdot),y^u(\cdot),z^u(\cdot))$, $(x^v(\cdot),y^v(\cdot),z^v(\cdot))$ and the solutions $(\Psi^u(\cdot),Q^u(\cdot),K^u(\cdot),R^u(\cdot,\cdot))$ and $(\Psi^v(\cdot),Q^v(\cdot),K^v(\cdot),R^v(\cdot,\cdot))$ of the corresponding adjoint equations (55), it holds that

$$\begin{aligned} &E\int_0^T\left\{\left|\Psi^u(t) - \Psi^v(t)\right|^\beta + \left|Q^u(t) - Q^v(t)\right|^\beta\right. \\ &\left.+\int_\Theta \left|R^u(t,\theta) - R^v(t,\theta)\right|^\beta \mu\left(d\theta\right)\right\} dt \leq C\rho\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta\gamma}{2}}, \end{aligned} \tag{70}$$

and

$$E\int_0^T \left|K^u(t) - K^v(t)\right|^\beta dt \leq C\rho\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta\gamma}{2}}. \tag{71}$$

Proof. For each $t \in [0,T]$ we denote $\widetilde{\Psi}(t) = \Psi^u(t) - \Psi^v(t)$, $\widetilde{R}(t,\theta) = R^u(t,\theta) - R^v(t,\theta)$, $\widetilde{K}(t) = K^u(t) - K^v(t)$, and $\widetilde{Q}(t) = Q^u(t) - Q^v(t)$. First we proceed to prove inequality-(71).

Proof of estimate (71). Note that the process $\widetilde{K}(t)_{t\in[0,T]}$ satisfies the following mean-field SDE

$$\begin{cases} d\widetilde{K}(t) = \int_\Theta\left\{g_y^*(t,\theta)\widetilde{K}(t) + \mathbb{G}_y\left(t,\theta\right) + E[g_{\widetilde{y}}^*(t,\theta)\widetilde{K}(t) + \mathbb{G}_{\widetilde{y}}\left(t,\theta\right)]\right\}\mu\left(d\theta\right) dt \\ \quad + \int_\Theta\left\{g_z^*(t,\theta)\widetilde{K}(t) + \mathbb{G}_z\left(t,\theta\right) + E[g_{\widetilde{z}}^*(t,\theta)\widetilde{K}(t) + \mathbb{G}_{\widetilde{z}}\left(t,\theta\right)]\right\}\mu\left(d\theta\right) dW(t) \\ \quad + \int_\Theta\left[g_r^*(t,\theta)\widetilde{K}(t) + \mathbb{G}_r\left(t,\theta\right)\right] N\left(d\theta,dt\right). \\ \widetilde{K}(0) = -\left\{(\phi_y\left(y^u(0),E(y^u(0)\right) - \phi_y\left(y^v(0),E(y^v(0)\right)\right\} - E\left\{\phi_{\widetilde{y}}\left(y^u(0),E(y^u(0)\right)\right. \\ \quad \left. - \phi_{\widetilde{y}}\left(y^v(0),E(y^v(0)\right)\right\}, \end{cases} \tag{72}$$

where

$$\mathbb{G}_y\left(t,\theta\right) = [g_y\left(t,\Lambda^u(t,),E(\Lambda^u(t,\theta)),u(t)\right) - g_y\left(t,\Lambda^v(t,\theta),E(\Lambda^v(t,\theta)),v(t)\right)]K^v(t).$$
$$\mathbb{G}_{\widetilde{y}}\left(t,\theta\right) = [g_{\widetilde{y}}\left(t,\Lambda^u(t,\theta),E(\Lambda^u(t,\theta)),u(t)\right) - g_{\widetilde{y}}\left(t,\Lambda^v(t,\theta),E(\Lambda^v(t,\theta)),v(t)\right)]K^v(t).$$

$$\mathbb{G}_{z}(t,\theta)=[g_{z}(t,\Lambda^{u}(t,\theta),E(\Lambda^{u}(t,\theta)),u(t))\ -g_{z}(t,\Lambda^{v}(t,\theta),E(\Lambda^{v}(t,\theta)),v(t))]K^{v}(t),$$

$$\mathbb{G}_{\widetilde{z}}(t,\theta)=[g_{\widetilde{z}}(t,\Lambda^{u}(t,\theta),E(\Lambda^{u}(t,\theta)),u(t))-g_{\widetilde{z}}(t,\Lambda^{v}(t,\theta),E(\Lambda^{v}(t,\theta)),v(t))]K^{v}(t).$$

$$\mathbb{G}_{r}(t,\theta)=[g_{r}(t,\Lambda^{u}(t,\theta),E(\Lambda^{u}(t,\theta)),u(t))-g_{r}(t,\Lambda^{v}(t,\theta),E(\Lambda^{v}(t,\theta)),v(t))]K^{v}(t).$$

Since the derivatives $g_y, g_{\widetilde{y}}, g_z, g_{\widetilde{z}}, g_r, \phi_y, \phi_{\widetilde{y}}$, are bounded, the mean-field SDE-(72) admits one and only one $\mathcal{F}_t-$adapted solution given by: for each $t\in[0,T]$

$$\begin{aligned}\widetilde{K}(t) &= \{(\phi_y(y^u(0),E(y^u(0))-\phi_y(y^v(0),E(y^v(0))\}+E\{\phi_{\widetilde{y}}(y^u(0),E(y^u(0))-\phi_{\widetilde{y}}(y^v(0),E(y^v(0))\}\\ &+\int_0^T\int_\Theta\left\{g_y^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_y(s,\theta)+E[g_{\widetilde{y}}^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_{\widetilde{y}}(s,\theta)]\right\}\mu(d\theta)\,ds\\ &+\int_0^t\int_\Theta\left\{g_z^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_z(s,\theta)+E[g_{\widetilde{z}}^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_{\widetilde{z}}(s,\theta)]\right\}\mu(d\theta)\,dW(s)\\ &+\int_0^t\int_\Theta[g_r^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_r(s,\theta)]N(d\theta,ds)\,.\end{aligned} \tag{73}$$

Case 1. First, we assume that $\beta=2$. By squaring both sides of equation-(73), taking expectation and using the fact that $g_y, g_{\widetilde{y}}, g_z, g_{\widetilde{z}}$ and g_r are bounded, and Assumption (H3), we shows that

$$\begin{aligned}&E[\left|\widetilde{K}(t)\right|^2]\leq C\mu(\Theta)\left\{E(\left|y^u(0)-y^v(0)\right|^2)+E\int_0^T\left|\widetilde{K}(s)\right|^2ds+E\int_0^T\int_\Theta\left[\left|\mathbb{G}_y(s,\theta)\right|^2+\left|\mathbb{G}_z(s,\theta)\right|^2\right.\right.\\ &\left.\left.+\left|E(g_{\widetilde{y}}^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_{\widetilde{y}}(s,\theta)))\right|^2+\left|E(g_{\widetilde{z}}^*(s,\theta)\widetilde{K}(s)+\mathbb{G}_{\widetilde{z}}(s,\theta))\right|^2\right]\mu(d\theta)\,ds\right\}\\ &\leq C\mu(\Theta)\left\{E(\left|y^u(0)-y^v(0)\right|^2)+E\int_0^T\left|\widetilde{K}(s)\right|^2ds+E\int_0^T\int_\Theta[\left|\mathbb{G}_y(s,\theta)\right|^2+\left|\mathbb{G}_z(s,\theta)\right|^2+\left|\mathbb{G}_r(s,\theta)\right|^2\right.\\ &+E\left|\mathbb{G}_{\widetilde{y}}(s,\theta)\right|^2+E\left|\mathbb{G}_{\widetilde{z}}(s,\theta))\right|^2]\mu(d\theta)\,ds.\end{aligned}$$

We estimate the right hand side of the above inequality. By applying *Lemma 3.1* we can shows immediately that

$$E(\left|y^u(0)-y^v(0)\right|^2)\leq C\rho\left(u(\cdot),v(\cdot)\right)^\alpha. \tag{74}$$

By a simple computation, we shows that

$$\begin{aligned}&E\int_0^T\int_\Theta\left|\mathbb{G}_y(t,\theta)\right|^2\mu(d\theta)\,dt\\ &\leq C\mu(\Theta)E\int_0^T\int_\Theta\left|g_y(t,\Lambda^u(t,\theta),E(\Lambda^u(t,\theta)),u(t))-g_y(t,\Lambda^v(t,\theta),E(\Lambda^v(t,\theta)),u(t))\right|^2\left|K^v(t)\right|^2\mu(d\theta)\,dt\\ &+C\mu(\Theta)E\int_0^T\int_\Theta\left|g_y(t,\Lambda^v(t,\theta),E(\Lambda^v(t,\theta)),u(t))-g_y(t,\Lambda^v(t,\theta),E(\Lambda^u(t,\theta)),v(t))\right|^2\left|K^v(t)\right|^2\mu(d\theta)\,dt.\end{aligned}$$

Under Assumption (H3) and from definition of metric ρ, we get: for $\gamma\in[0,1]$

$$\begin{aligned}&E\int_0^T\int_\Theta\left|\mathbb{G}_y(t,\theta)\right|^2\mu(d\theta)\,dt\leq CE\int_0^T\int_\Theta\left[\left|x^u(t)-x^v(t)\right|^{2\gamma}+\left|y^u(t)-y^v(t)\right|^{2\gamma}+\left|z^u(t)-z^v(t)\right|^{2\gamma}\right.\\ &\left.+\left|E(y^u(t)-y^v(t))\right|^{2\gamma}+\left|E(y^u(t)-y^v(t))\right|^{2\gamma}+\left|E(y^u(t)-y^v(t))\right|^{2\gamma}\right]\left|K^v(t)\right|^2dt\\ &+C\mu(\Theta)E\int_0^T\mathbf{1}_{\{u(\omega,t)\neq v(\omega,t)\}}(t)\left|K^v(t)\right|^2dt,\end{aligned}$$

by using *Hölder's inequality* we get

$$E\int_0^T\int_\Theta\left|\mathbb{G}_y\left(t,\theta\right)\right|^2\mu\left(d\theta\right)dt\leq\left\{C\mu(\Theta)\left[E\int_0^T\left|x^u(t)-x^v(t)\right|^2\right]^\gamma+\left[E\int_0^T\left|y^u(t)-y^v(t)\right|^2\right]^\gamma\right.$$
$$+\left[E\int_0^T\left|z^u(t)-z^v(t)\right|^2\right]^\gamma+\left[E\int_0^T\left|E(x^u(t)-x^v(t))\right|^2\right]^\gamma+\left[E\int_0^T\left|E(y^u(t)-y^v(t))\right|^2\right]^\gamma$$
$$\left.+\left[E\int_0^T\left|E(z^u(t)-z^v(t))\right|^2\right]^\gamma\right\}\left[E\int_0^T\left|K^v(t)\right|^{2/(1-\gamma)}dt\right]^{1-\gamma}$$
$$+C\mu(\Theta)E\left[\int_0^T\left|K^v(t)\right|^{2/(1-\alpha\gamma)}dt\right]^{1-\alpha\gamma}\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}.$$

Using *Lemma 3.1,* and (58) we obtain

$$E\int_0^T\int_\Theta\left|\mathbb{G}_y\left(t,\theta\right)\right|^2\mu\left(d\theta\right)dt\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}. \tag{75}$$

Applying the same arguments developed above, we easily shows that $\gamma\in[0,1]$

$$\begin{array}{l} E\int_0^T\int_\Theta\left|\mathbb{G}_{\widetilde{y}}\left(t,\theta\right)\right|^2\mu\left(d\theta\right)dt\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}.\\ E\int_0^T\int_\Theta\left|\mathbb{G}_z\left(t,\theta\right)\right|^2\mu\left(d\theta\right)dt\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}.\\ E\int_0^T\int_\Theta\left|\mathbb{G}_{\widetilde{z}}\left(t,\theta\right)\right|^2\mu\left(d\theta\right)dt\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}.\\ E\int_0^T\int_\Theta\left|\mathbb{G}_r\left(t,\theta\right)\right|^2\mu\left(d\theta\right)dt\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}.\end{array} \tag{76}$$

Combining (74), (75) and (76) it follows that

$$E(\left|\widetilde{K}(t)\right|^2)\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\alpha\gamma}.$$

Case 2. Now assume that $\beta\in\,]1,2[$, then by using *Hölder's inequality* we get the inequality (71). □
Let us turn to prove inequality (70).
Proof of estimate (70)**.** Noting that the processes $\widetilde{\Psi}(t)=\Psi^u(t)-\Psi^v(t)$, $\widetilde{Q}(t)=Q^u(t)-Q^v(t)$ and $\widetilde{R}(t,\theta)=R^u(t,\theta)-R^v(t,\theta)$ satisfies the following mean-field BSDE with jumps processes

$$\begin{cases} -d\widetilde{\Psi}(t)=\left\{f_x^*\left(t,x^u(t),E(x^u(t)),u(t)\right)\widetilde{\Psi}(t)+\sigma_x^*\left(t,x^u(t),E(x^u(t))\right)\widetilde{Q}(t)\right.\\ +\int_\Theta\left\{g_x\left(t,\Lambda^u(t,\theta),E(\Lambda^u(t,\theta)),u(t)\right)\widetilde{K}(t)+c_x(t,\theta)\widetilde{R}(t,\theta)\right\}\mu\left(d\theta\right)\\ +E\left[f_{\widetilde{x}}^*\left(t,x^u(t),E(x^u(t)),u(t)\right)\widetilde{\Psi}(t)+\sigma_{\widetilde{x}}^*\left(t,x^u(t),E(x^u(t))\right)\widetilde{Q}(t)\right.\\ \left.+\int_\Theta g_{\widetilde{x}}\left(t,\Lambda^u(t,\theta),E(\Lambda^u(t,\theta)),u(t)\right)\widetilde{K}(t)\mu\left(d\theta\right)\right]+M(t)\}\,dt\\ -\widetilde{Q}(t)dW(t)+\int_\Theta\widetilde{R}\left(t,\theta\right)N\left(d\theta,dt\right).\\ \widetilde{\Psi}(T)=-h_x\left(x^u(T),E(x^u(T))\right)K^u(T)-h_x\left(x^v(T),E(x^v(T))\right)K^v(T)\\ -E\{h_{\widetilde{x}}\left(x^u(T),E(x^u(T))\right)K^u(T)-h_{\widetilde{x}}\left(x^v(T),E(x^v(T)\right)K^v(T)]\}, \end{cases} \tag{77}$$

where the process $M(\cdot)$ depend to marginal law of the state processes such that

$$
\begin{aligned}
&M(t) = [f_x^*(t, x^u(t), E(x^u(t)), u(t)) - f_x^*(t, x^v(t), E(x^v(t)), v(t))]\,\Psi^v(t) \\
&+ [\sigma_x^*(t, x^u(t), E(x^u(t))) - \sigma_x^*(t, x^v(t), E(x^v(t)))]\,Q^v(t) \\
&+ \int_\Theta [g_x^*(t, \Lambda^u(t,\theta), E(\Lambda^u(t,\theta)), u(t))) - g_x^* t, \Lambda^v(t,\theta), E(\Lambda^v(t,\theta)), v(t))] K^v(t))\mu(d\theta) \\
&+ E\{[f_{\widetilde{x}}^*(t, x^u(t), E(x^u(t)), u(t)) - f_{\widetilde{x}}^*(t, x^v(t), E(x^v(t)), v(t))]\,\Phi^v(t)\} \\
&+ E\{[\sigma_{\widetilde{x}}^*(t, x^u(t), E(x^u(t))) - \sigma_{\widetilde{x}}^*(t, x^v(t), E(x^v(t)))]\,Q^v(t)\} \\
&+ \int_\Theta E[g_{\widetilde{x}}^*(t, \Lambda^u(t,\theta), E(\Lambda^u(t,\theta)), u(t))) - g_{\widetilde{x}}^*(t, \Lambda^v(t,\theta), E(\Lambda^v(t,\theta)), v(t)))\,K^v(t)]\mu(d\theta).
\end{aligned} \tag{78}
$$

Let $(U(t))_{t\in[0,T]}$ be the solution of the following linear mean-field SDE

$$
\begin{cases}
dU(t) = \{f_x(t, x^u(t), E(x^u(t)), u(t))\,U(t) + f_{\widetilde{x}}(t, x^u(t), E(x^u(t)), u(t))\,E(U(t)) \\
+ \left|\widetilde{\Psi}(t)\right|^{\beta-1} Sgn(\widetilde{\Psi}(t))\Big\}\,dt + \{[\sigma_x(t, x^u(t), E(x^u(t)))\,U(t) \\
+ \sigma_{\widetilde{x}}(t, x^u(t), E(x^u(t)))\,E(U(t)) + \left|\widetilde{Q}(t)\right|^{\beta-1} Sgn(\widetilde{Q}(t))\Big\}\,dW(t) \\
+ \int_\Theta \left[c_x(t,\theta)\left|\widetilde{R}(t,\theta)\right|^{\beta-1} Sgn(\widetilde{R}(t,\theta))\right] N(d\theta, dt). \\
U(0) = 0,
\end{cases} \tag{79}
$$

where $Sgn\,(x) = (Sgn(x_1), Sgn(x_2), ..., Sgn(x_n))^*$ for any vector $x = (x_1, x_2, .., x_n)^*$.
Note that since $f_x, f_{\widetilde{x}}, \sigma_x\ f_{\widetilde{x}}, g_x\ g_{\widetilde{x}}$ are bounded with the helps of (*Proposition A2, Appendix*) and due to the fact that

$$
\begin{aligned}
&E\int_0^T \left\{\left|\left|\widetilde{\Psi}(t)\right|^{\beta-1} Sgn(\widetilde{\Psi}(t))\right|^2 + \left|\left|\widetilde{Q}(t)\right|^{\beta-1} Sgn(\widetilde{Q}(t))\right|^2\right\} \\
&+E\int_0^T \int_\Theta \left|\left|\widetilde{R}(t,\theta)\right|^{\beta-1} Sgn(\widetilde{R}(t,\theta))\right|^2 \mu(d\theta)\,dt < \infty,
\end{aligned}
$$

the SDE-(79) has a unique strong solution.
Let $q > 2$ such that $\frac{1}{q} + \frac{1}{\beta} = 1$, $\beta \in (1, 2)$ then according to (58), we get

$$
E(\sup_{t\in[0,T]} |U(t)|^q) \leq CE\int_0^T \left\{\left|\widetilde{\Psi}(t)\right|^{\beta q - q} + \left|\widetilde{Q}(t)\right|^{\beta q - q} + \int_\Theta \left|\widetilde{R}(t,\theta)\right|^{\beta q - q} \mu(d\theta)\right\} dt < \infty.
$$

By applying *Integration by parts formula* for jumps to $\widetilde{\Psi}(t)U(t)$, (see Lemma A1, *Appendix*) on $[0, T]$

and taking expectation, we get

$$
\begin{aligned}
&E\left[\widetilde{\Psi}(T)U(T)\right]+E\int_0^T\int_\Theta U(t)\{g_x\left(t,\Lambda^u(t,\theta),E(\Lambda^u(t,\theta)),u(t)\right)\widetilde{K}(t)\\
&+c_x(t,\theta)\widetilde{R}(t,\theta)\}\mu\left(d\theta\right)dt\\
&=E\int_0^T\left\{\widetilde{\Psi}(t)\left|\widetilde{\Psi}(t)\right|^{\beta-1}Sgn(\widetilde{\Psi}(t))+\widetilde{Q}(t)\left|\widetilde{Q}(t)\right|^{\beta-1}Sgn(\widetilde{Q}(t))\right.\\
&\left.+\int_\Theta\left|\left|\widetilde{R}(t,\theta)\right|^{\beta-1}Sgn(\widetilde{R}(t,\theta))\right|^2\mu\left(d\theta\right)\right\}dt-E\int_0^T M(t)U(t)dt.
\end{aligned}
$$

Since

$$
\begin{aligned}
&E\int_0^T\left\{\widetilde{\Psi}(t)\left|\widetilde{\Psi}(t)\right|^{\beta-1}Sgn(\widetilde{\Psi}(t))+\widetilde{Q}(t)\left|\widetilde{Q}(t)\right|^{\beta-1}Sgn(\widetilde{Q}(t))\right.\\
&\left.+\int_\Theta\widetilde{R}(t,\theta)\left|\widetilde{R}(t,\theta)\right|^{\beta-1}Sgn(\widetilde{R}(t,\theta))\mu\left(d\theta\right)\right\}dt\\
&=E\int_0^T\left[\left|\widetilde{\Psi}(t)\right|^\beta+\left|\widetilde{Q}(t)\right|^\beta+\int_\Theta\left|\widetilde{R}(t,\theta)\right|^\beta\mu\left(d\theta\right)\right]dt,
\end{aligned}
$$

we have

$$
\begin{aligned}
&E\int_0^T\left[\left|\widetilde{\Psi}(t)\right|^\beta+\left|\widetilde{Q}(t)\right|^\beta+\int_\Theta\left|\widetilde{R}(t,\theta)\right|^\beta\mu\left(d\theta\right)\right]dt\\
&\leq E\left\{\int_0^T M(t)U(t)dt+\widetilde{\Psi}(T)U(T)\right\}\\
&\leq E\left\{\int_0^T M(t)U(t)dt+\left[h_x\left(x^u(T),E(x^u(T))\right)K^u(T)\right.\right.\\
&-h_x\left(x^v(T),E(x^v(T))\right)K^v(T)]\,U(T)\\
&\left.+E[h_{\widetilde{x}}\left(x^u(T),E(x^u(T))\right)K^u(T)-h_{\widetilde{x}}\left(x^v(T),E(x^v(T)\right)K^v(T)]U(T)\right\}.
\end{aligned}
$$

By a simple computation we get

$$
\begin{aligned}
&E\textstyle\int_0^T\left[\left|\widetilde{\Psi}(t)\right|^\beta+\left|\widetilde{Q}(t)\right|^\beta+\int_\Theta\left|\widetilde{R}(t,\theta)\right|^\beta\mu\left(d\theta\right)\right]dt\\
&\leq CE\textstyle\int_0^T\left|M(t)\right|^\beta dt+CE\left\{\left|h_x(x^u(T),E(x^u(T)))K^u(T)\right.\right.\\
&-h_x(x^v(T),E(x^v(T)))K^v(T)|^\beta+|E\left\{h_{\widetilde{x}}(x^u(T),E(x^u(T)))K^u(T)\right.\\
&\left.\left.-\ h_{\widetilde{x}}(x^v(T),E(x^v(T)))K^v(T)\right\}|^\beta\right\}.
\end{aligned}
\tag{80}
$$

We proceed to estimate the right hand side of (80). From assumption (H3), *Lemma 3.1 and* (58), we easily see that

$$
\begin{aligned}
&E\left\{\left|h_x(x^u(T),E(x^u(T)))K^u(T)\ -\ h_x(x^v(T),E(x^v(T)))K^v(T)\right|^{\beta}\right.\\
&\left.+E\left|\left\{h_{\widetilde{x}}(x^u(T),E(x^u(T)))K^u(T)\ -\ h_{\widetilde{x}}(x^v(T),E(x^v(T)))K^v(T)\right\}\right|^{\beta}\right\}\\
&\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\frac{\alpha\beta\gamma}{2}}.
\end{aligned}
\tag{81}
$$

$$
E\int_0^T\left|M(t)\right|^{\beta}dt\leq C\rho\left(u\left(\cdot\right),v\left(\cdot\right)\right)^{\frac{\alpha\beta\gamma}{2}}.
\tag{82}
$$

Finally, the desired result (70) follows immediately by combining (80), (81) and (82). This completes the proof of *Lemma 3.2*. □

Lemma 2.3.3. For any $\varepsilon>0$ there exists near-optimal control $\overline{u}^{\varepsilon}(\cdot)\in\mathcal{U}([0,T])$ and an $\mathcal{F}_t-$adapted process $(\overline{\Psi}^{\varepsilon}(\cdot),\overline{Q}^{\varepsilon}(\cdot),\overline{K}^{\varepsilon}(\cdot),\overline{R}^{\varepsilon}(\cdot,\cdot))$ such that for all $u\in\mathbb{U}$

$$
\begin{aligned}
&E\left\{\int_0^T\overline{\Psi}^{\varepsilon}(t)\left[f\left(t,\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),u(t)\right)-f\left(t,\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right]\right.\\
&+\int_{\Theta}\overline{K}^{\varepsilon}(t)\left[g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u(t))\right.\\
&\left.\left.-g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))\right]\mu\left(d\theta\right)\right\}dt\geq-\varepsilon^{1/2},
\end{aligned}
\tag{83}
$$

where $(\overline{\Lambda}^{\varepsilon}(t),E(\overline{\Lambda}^{\varepsilon}(t))):=(\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),\overline{y}^{\varepsilon}(t),E(\overline{y}^{\varepsilon}(t)),\ \overline{z}^{\varepsilon}(t),E(\overline{z}^{\varepsilon}(t)))$ and $(\overline{x}^{\varepsilon}(\cdot),\overline{y}^{\varepsilon}(\cdot),\overline{z}^{\varepsilon}(\cdot),\overline{r}^{\varepsilon}(\cdot,\cdot))$ denotes the solution of mean-field FBSDEJs-(52) and $(\overline{\Phi}^{\varepsilon}(\cdot),\overline{Q}^{\varepsilon}(\cdot),\overline{K}^{\varepsilon}(\cdot),\overline{R}^{\varepsilon}(\cdot,\cdot))$ is the solution of the adjoint equation (55) corresponding to $\overline{u}^{\varepsilon}(\cdot)$.

Proof. By applying Ekeland's variational principle (*Lemma 2.1*) with $\lambda=\varepsilon^{2/3}$, there exists an admissible control $\overline{u}^{\varepsilon}(\cdot)$ such that

$$
\rho\left(\overline{u}^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot)\right)\leq\varepsilon^{2/3},
\tag{84}
$$

and $J^{\varepsilon}\left(\zeta,\overline{u}^{\varepsilon}(\cdot)\right)\leq J^{\varepsilon}\left(\zeta,u(\cdot)\right)$, for any $u(\cdot)\in\mathcal{U}([0,T])$ where

$$
J^{\varepsilon}\left(\zeta,u(\cdot)\right):=J\left(\zeta,u(\cdot)\right)+\varepsilon^{1/3}\rho\left(\overline{u}^{\varepsilon}(\cdot),u(\cdot)\right).
\tag{85}
$$

Notice that $\overline{u}^{\varepsilon}(\cdot)$ which is near-optimal for the initial cost J is optimal for the new cost J^{ε} defined by (85).

Next, we use the spike variation techniques for $\overline{u}^{\varepsilon}(\cdot)$ to derive the variational inequality as follows. For $0<\hbar<T$, we choose a Borel subset $\mathcal{B}_{\hbar}\subset[0,T]$ such that $\mu(\mathcal{B}_{\hbar})=\hbar$, where $\mu(\mathcal{B}_{\hbar})$ denote the *Lebesgue measure* of the subset $\mathcal{B}_{\hbar}$, and we consider the control process which is the spike variation of $\overline{u}^{\varepsilon}(\cdot)$, i.e., $t\in[0,T]$

$$
\overline{u}^{\varepsilon,\hbar}(t)=\begin{cases}u:t\in\mathcal{B}_{\hbar}.\\ \overline{u}^{\varepsilon}(t):t\in[0,T]\setminus\mathcal{B}_{\hbar}.\end{cases}
$$

By using the fact that $J^{\varepsilon}\left(\zeta,\overline{u}^{\varepsilon}(\cdot)\right)\leq J^{\varepsilon}(\zeta,\overline{u}^{\varepsilon,\hbar}(\cdot))$ and $\rho(\overline{u}^{\varepsilon}(\cdot),\overline{u}^{\varepsilon,\hbar}(\cdot))\leq\hbar$, we obtain

$$
J(\zeta,\overline{u}^{\varepsilon,\hbar}(\cdot))-J(\zeta,\overline{u}^{\varepsilon}(\cdot))\geq-\varepsilon^{1/2}\rho(\overline{u}^{\varepsilon}(\cdot),\overline{u}^{\varepsilon,\hbar}(\cdot))\geq-\varepsilon^{1/3}\hbar.
\tag{86}
$$

Arguing as in Hafayed ([24], Theorem 3.1), the left-hand side of inequality (86) is equal to

$$E\int_0^T\left\{\overline{\Psi}^{\varepsilon}(t)[f(t,\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),u)-f(t,\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t))]\right.$$
$$\left.+\int_{\Theta}\overline{K}^{\varepsilon}(t)\,[g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u)-g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),\overline{u}^{\varepsilon}(t))]\mu(d\theta)\right\}\mathbf{1}_{\mathcal{B}_{\hbar}}(t)dt$$
$$+o(\hbar).$$

Finally, dividing (86) by $\hbar$ and sending $\hbar$ to zero, the desired result follows □

Proof of Theorem 2.3.1. To prove (61) it remains to estimate the following differences

$$\begin{aligned}\Gamma_1(\varepsilon)=E\int_0^T[\overline{\Phi}^{\varepsilon}(t)\{f(t,\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),u))-f(t,\overline{x}^{\varepsilon}(t),E(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t))\}\\-\Psi^{\varepsilon}(t)\{f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t)),u)-f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t)),u^{\varepsilon}(t)\}]\,dt,\end{aligned}\tag{87}$$

$$\begin{aligned}\Gamma_2(\varepsilon)=E\int_0^T\int_{\Theta}\{\overline{K}^{\varepsilon}(t)[g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u)-g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),\overline{u}^{\varepsilon}(t))]\\-K^{\varepsilon}(t)[g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u)-g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t)]\}\,\mu(d\theta)\,dt.\end{aligned}\tag{88}$$

Estimate (88). First, by adding and subtracting $E\int_0^T\int_{\Theta}K^{\varepsilon}(t)g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),\overline{u}^{\varepsilon}(t))\mu(d\theta)\,dt$ from $\Gamma_2(\varepsilon)$ we get

$$\begin{aligned}&\Gamma_2(\varepsilon)=E\int_0^T\int_{\Theta}(\overline{K}^{\varepsilon}(t)-K^{\varepsilon}(t))[g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u(t)))\\&-g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),\overline{u}^{\varepsilon}(t))]\mu(d\theta)\,dt+E\int_0^T\int_{\Theta}K^{\varepsilon}(t)[g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u)\\&-g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u)]\mu(d\theta)\,dt-E\int_0^T\int_{\Theta}K^{\varepsilon}(t)[g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),\overline{u}^{\varepsilon}(t))\\&-g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))]\mu(d\theta)\,dt\\&=\mathbb{I}_1(\varepsilon)+\mathbb{I}_2(\varepsilon)+\mathbb{I}_3(\varepsilon).\end{aligned}$$

We estimate the first term on the right-hand side $\mathbb{I}_1(\varepsilon)$. For any $\delta\in[0,\frac{1}{3}[$, let $\alpha\gamma=3\delta\in[0,1[$. Let β be a fixed real number such that $1<\beta<2$ so that $(1+\alpha)\beta<2$. Taking $q>2$ such that $\frac{1}{\beta}+\frac{1}{q}=1$ then by using Hölder's inequality, Lemma 3.2 and note (84) and the fact that $\mu(\Theta)<\infty$, we obtain

$$\begin{aligned}&\mathbb{I}_1(\varepsilon)\leq\left[E\int_0^T\left|\overline{K}^{\varepsilon}(t)-K^{\varepsilon}(t)\right|^{\beta}dt\right]^{\frac{1}{\beta}}\left[E\int_0^T\int_{\Theta}\left|g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),u)\right.\right.\\&\left.\left.-\;g(t,\overline{\Lambda}^{\varepsilon}(t,\theta),E(\overline{\Lambda}^{\varepsilon}(t,\theta)),\overline{u}^{\varepsilon}(t))\right|^{q}\mu(d\theta)\,dt\right]^{\frac{1}{q}}\\&\leq C\left[\mu(\Theta)\right]^{\frac{1}{q}}\left[\rho(\overline{u}^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))^{\frac{\alpha\beta\gamma}{2}}\right]^{\frac{1}{\beta}}\left[E\int_0^T(1+|\overline{x}^{\varepsilon}(t)|^{q}+|E(\overline{x}^{\varepsilon}(t))|^{q}+|\overline{y}^{\varepsilon}(t)|^{q}+|E(\overline{y}^{\varepsilon}(t))|^{q})dt\right]^{\frac{1}{q}}\\&\leq C(\varepsilon^{\frac{2}{3}})^{\frac{\alpha\beta\gamma}{2}\frac{1}{\beta}}=C\varepsilon^{\frac{\alpha\gamma}{3}}=C\varepsilon^{\delta}.\end{aligned}$$

Let us turn to the second term $\mathbb{I}_2(\varepsilon)$. By applying *Cauchy-Schwartz inequality*, note (58), Assumption

(H1), and Lemma 3.1, we get

$$
\begin{aligned}
&\mathbb{I}_2(\varepsilon) = E\int_0^T\int_\Theta K^\varepsilon(t)[g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u) - g(t,\Lambda^\varepsilon(t,\theta),E(\Lambda^\varepsilon(t,\theta)),u)]\mu\,(d\theta)\,dt\\
&\leq \left[E\int_0^T |K^\varepsilon(t)|^2\,dt\right]^{\frac{1}{2}}\left[E\int_0^T\int_\Theta \left|g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u)\right.\right.\\
&-\left.\left. g(t,\Lambda^\varepsilon(t,\theta),E(\Lambda^\varepsilon(t,\theta)),u)\right|^2\mu\,(d\theta)\,dt\right]^{\frac{1}{2}}\\
&\leq C\left[\mu\,(\Theta)\right]^{\frac{1}{2}}\left\{E\int_0^T\left[|\overline{x}^\varepsilon(t)-x^\varepsilon(t)|^2+|E(\overline{x}^\varepsilon(t)-x^\varepsilon(t))|^2\right.\right.\\
&\left.\left.+|\overline{y}^\varepsilon(t)-y^\varepsilon(t)|^2+|E(\overline{y}^\varepsilon(t)-y^\varepsilon(t))|^2+|\overline{z}^\varepsilon(t)-z^\varepsilon(t)|^2+|E(\overline{z}^\varepsilon(t)-z^\varepsilon(t))|^2\right\}^{\frac{1}{2}}\right.\\
&\leq C\left[\mu\,(\Theta)\right]^{\frac{1}{2}}\left[\rho(\overline{u}^\varepsilon(\cdot),u^\varepsilon(\cdot))^{\alpha\gamma}\right]^{\frac{1}{2}} \leq C(\varepsilon^{\frac{2}{3}})^{\alpha\gamma\frac{1}{2}} = C\varepsilon^{\frac{\alpha\gamma}{3}} = C\varepsilon^\delta.
\end{aligned}
$$

Now, let us estimate the third term $\mathbb{I}_3(\varepsilon)$. By adding and subtracting $g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u^\varepsilon(t))$ from $\mathbb{I}_3(\varepsilon)$, we have

$$
\begin{aligned}
&\mathbb{I}_3(\varepsilon)\\
&= -E\int_0^T\int_\Theta K^\varepsilon(t)[g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),\overline{u}^\varepsilon(t)) - g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u^\varepsilon(t))]\mu\,(d\theta)\,dt\\
&- E\int_0^T\int_\Theta K^\varepsilon(t)[g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u^\varepsilon(t)) - g(t,\Lambda^\varepsilon(t),E(\Lambda^\varepsilon(t,\theta)),u^\varepsilon(t))]\mu\,(d\theta)\,dt,
\end{aligned}
$$

then by using Cauchy-Schwatrz inequality, we have

$$
\begin{aligned}
&\mathbb{I}_3(\varepsilon) \leq \left[E\int_0^T |K^\varepsilon(t)|^2\,dt\right]^{\frac{1}{2}}\left\{E\int_0^T\int_\Theta g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),\overline{u}^\varepsilon(t))\right.\\
&\left.-g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u^\varepsilon(t))\right|^2\mathbf{1}_{\{\overline{u}^\varepsilon(w,t)\neq u^\varepsilon(w,t\}}(t)\,\mu\,(d\theta)\,dt\Big\}^{\frac{1}{2}}\\
&+E\int_0^T\int_\Theta |K^\varepsilon(t)|\left|[g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u^\varepsilon(t)) - g\left(t,\Lambda^\varepsilon(t,\theta),E(\Lambda^\varepsilon(t,\theta)),u^\varepsilon(t)\right)]\right|\mu\,(d\theta)\,dt.
\end{aligned}
$$

We proceed as in $\mathbb{I}_2^\varepsilon$ to estimate the second term in the right of above inequality, then by applying *Cauchy-Schwartz inequality,* Assumption (H1), Lemma 3.1 and note (58) we get

$$
\begin{aligned}
&\mathbb{I}_3(\varepsilon) \leq \left[E\int_0^T |K^\varepsilon(t)|^2\,dt\right]^{\frac{1}{2}}\left\{\left[E\int_0^T\left|\int_\Theta g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),\overline{u}^\varepsilon(t))\right.\right.\right.\\
&-\left.\left.\left. g(t,\overline{\Lambda}^\varepsilon(t,\theta),E(\overline{\Lambda}^\varepsilon(t,\theta)),u^\varepsilon(t))\mu\,(d\theta)\right|^4 dt\right]^{\frac{1}{2}}\left[E\int_0^T \mathbf{1}_{\{\overline{u}^\varepsilon(w,t)\neq u^\varepsilon(w,t)\}}(t)\,dt\right]^{\frac{1}{2}}\right\}^{\frac{1}{2}} + C\varepsilon^\delta\\
&\leq C\left[\rho(\overline{u}^\varepsilon(\cdot),u^\varepsilon(\cdot))^{\frac{1}{2}}\right]^{\frac{1}{2}} + C\varepsilon^\delta \leq C\varepsilon^\delta,
\end{aligned}
$$

this implies that

$$
\Gamma_2(\varepsilon) = \mathbb{I}_1(\varepsilon) + \mathbb{I}_2(\varepsilon) + \mathbb{I}_3(\varepsilon) \leq C\varepsilon^\delta. \tag{89}
$$

Estimate (87). Using similar arguments developed above, we can prove that

$$
\Gamma_1(\varepsilon) \leq C\varepsilon^\delta. \tag{90}
$$

By combining (89), (90) and *Lemma 2.3.3,* we get

$$E\left\{\int_0^T \Psi^{\varepsilon}(t)\left[f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u)-f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u^{\varepsilon}(t))\right]\right.$$
$$\left.+\int_{\Theta} K^{\varepsilon}(t)\left[g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u)-g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))\right]\mu(d\theta)\right\}dt$$
$$\geq -C\varepsilon^{\delta}.$$

Finally, the desired result (61) follows immediately from the fact that

$$\Psi^{\varepsilon}(t)\left[f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u)-f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u^{\varepsilon}(t))\right]$$
$$+\int_{\Theta} K^{\varepsilon}(t)\left[g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u)-g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))\right]\mu(d\theta)$$
$$=H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u,\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))$$
$$-H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta)).$$

This completes the proof of Theorem 3.1. □

Proof of Corollary 2.3.1. Let $u\in\mathbb{U}$ be a deterministic element and $\mathcal{G}$ be an arbitrary element of $\sigma-$algebra $\mathcal{F}_t$, and by setting

$$v(t)=u\mathbf{1}_{\mathcal{G}}+u^{\varepsilon}(t)\mathbf{1}_{\Omega-\mathcal{G}}.$$

It is obvious that $v(\cdot)\in\mathcal{U}([0,T])$ is an admissible control. Applying (61) with $v(\cdot)$, we get

$$E[\mathbf{1}_{\mathcal{G}}H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))]$$
$$\geq E[\mathbf{1}_{\mathcal{G}}H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u,\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))]-C\varepsilon^{\delta},$$
$$\forall u\in\mathbb{U},\ \mathcal{G}\in\mathcal{F}_t,$$

which implies that $\forall u\in\mathbb{U}$.

$$E[H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))\,|\mathcal{F}_t]$$
$$\geq E[H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u,\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))\,|\mathcal{F}_t]-C\varepsilon^{\delta}.$$

Noting that the above quantity inside the conditional expectation is $\mathcal{F}_t-$measurable, then we get $\forall u\in\mathbb{U}:$

$$\begin{gathered}H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))\\ \geq H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u,\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))\\ -C\varepsilon^{\delta}.\ \mathbb{P}-a.s.,\ a.e.\ t\in[0,T].\end{gathered}\tag{91}$$

Moreover, in the spike variations technique for the perturbed control $\overline{u}^{\varepsilon,\hbar}(\cdot)$ the point $u\in\mathbb{U}$ may be replaced by any admissible control $u(\cdot)\in\mathcal{U}([0,T])$ and the subsequent argument still goes through. So the inequality in the estimate (91) holds for any $u(\cdot)\in\mathcal{U}([0,T])$. This completes the proof of (62). □

12.2. Sufficient conditions for near-optimality of mean-field FBSDEJs

The sufficient condition of near-optimality is of significant importance in the stochastic maximum principle for computing optimal controls. It says that if an admissible control satisfies the near-maximum condition on the Hamiltonian then the control is indeed near-optimal for the stochastic control problem. In this section, we will prove that under an additional hypotheses with non negative derivatives with respect to $\widetilde{x}, \widetilde{y}, \widetilde{z}$, the near-maximality condition on the Hamiltonian function is a sufficient condition for near-optimality. This is the second main result of this work.

Assumption (H4). We assume

$$|f(t,x,\widetilde{x},u) - f(t,x,\widetilde{x},v)| + |f_u(t,x,\widetilde{x},u) - f_u(t,x,\widetilde{x},v)| \leq C\,|u-v|\,. \tag{92}$$

$$\begin{aligned} & |g(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},u) - g(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},v)| + |g_u(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},u) - g_u(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},v)| \\ \leq \; & C\,|u-v|\,. \end{aligned} \tag{93}$$

$$\begin{aligned} & H\left(t,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta)\right) \text{ is concave with respect to } (x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r,u) \\ & \qquad a.e.t \in [0,T]\,,\ \mathbb{P}-a.s. \end{aligned} \tag{94}$$

$$h\left(\cdot,\cdot\right) \text{ concave, } \phi\left(\cdot,\cdot\right) \text{ convex with respect to } (x,\widetilde{x})\,. \tag{95}$$

Duality relations. Our Lemma below deals with the duality relations between $\Psi^{\varepsilon}(T)$, $x^{*}(T) - x^{\varepsilon}(T)$ and $K^{\varepsilon}(T)$, $y^{*}(T) - y^{\varepsilon}(T)$. This Lemma is very important for the proof of *Theorem 3.2.1*

Lemma 2.4.1 Let $(x^{*}(\cdot), y^{*}(\cdot), z^{*}(\cdot), r^{*}(\cdot,\cdot))$ be the solution of state equation (52) corresponding to any admissible control $u^{*}(\cdot)$. We have

$$\begin{aligned} & E\left[\Psi^{\varepsilon}(T)\left(x^{*}(T)-x^{\varepsilon}(T)\right)\right] = E\int_0^T \Psi^{\varepsilon}(t)\left[f(t,x^{*}(t),E(x^{*}(t),u^{*}(t))\right. \\ & -\left.f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u^{\varepsilon}(t))\right]dt \\ & +E\int_0^T H_x^{\varepsilon}(t)\left(x^{*}(t)-x^{\varepsilon}(t)\right)dt + E\int_0^T E[H_{\widetilde{x}}^{\varepsilon}(t)]\left(E(x^{*}(t))-E(x^{\varepsilon}(t))\right)dt \\ & +E\int_0^T Q^{\varepsilon}(t)\left[\sigma(t,x^{*}(t),E(x^{*}(t)) - \sigma(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t))\right]dt \\ & +E\int_0^T\int_{\Theta} R^{\varepsilon}(t,\theta)\left[c(t,x^{*}(t),\theta)-c(t,x^{\varepsilon}(t),\theta)\right]\mu\left(d\theta\right)dt, \end{aligned} \tag{96}$$

similarly

$$\begin{aligned} & E\left[K^{\varepsilon}(T)\left(y^{*}(T)-y^{\varepsilon}(T)\right)\right] = -E\left(\phi_y\left(y(0),E\left(y(0)\right)\right)\left(y^{\varepsilon}(0)-y^{*}(0)\right)\right) \\ & -E\left(\phi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)\right)\left(E\left(y^{\varepsilon}(0)\right)-E\left(y^{*}(0)\right)\right) \\ & +E\int_0^T\int_{\Theta} K^{\varepsilon}(t)\left\{g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))\right. \\ & -\left.g(t,\Lambda^{*}(t,\theta),E(\Lambda^{*}(t,\theta)),u^{*}(t))\right\}\mu\left(d\theta\right)dt + E\int_0^T H_y^{\varepsilon}(t)\left(y^{*}(t)-y^{\varepsilon}(t)\right)dt \\ & +E\int_0^T E(H_{\widetilde{y}}^{\varepsilon}(t))\left(E(y^{*}(t))-E(y^{\varepsilon}(t))\right))dt + E\int_0^T H_z^{\varepsilon}(t)\left(z^{*}(t)-z^{\varepsilon}(t)\right)dt \\ & +E\int_0^T E(H_{\widetilde{z}}^{\varepsilon}(t))\left(E(z^{*}(t))-E(z^{\varepsilon}(t))\right)dt + E\int_0^T\int_{\Theta} H_r^{\varepsilon}(t)\left[r^{*}(t,\theta)-r^{\varepsilon}(t,\theta)\right]\mu\left(d\theta\right)dt, \end{aligned} \tag{97}$$

and

$$
\begin{aligned}
&E\left[\Psi^{\varepsilon}(T)\left(x^{*}(T)-x^{\varepsilon}(T)\right)\right]+E\left[K^{\varepsilon}(T)\left(y^{*}(T)-y^{\varepsilon}(T)\right)\right]\\
&+E\left(\phi_{y}\left(y(0),E\left(y(0)\right)\right)\left(y^{\varepsilon}(0)-y^{*}(0)\right)\right)+E[\phi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)]\left(E\left(y^{\varepsilon}(0)\right)-E\left(y^{*}(0)\right)\right)\\
&=E\int_{0}^{T}\Psi^{\varepsilon}(t)(f^{*}(t,x^{*}(t),E(x^{*}(t),u^{*}(t))-f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u^{\varepsilon}(t)))dt\\
&+E\int_{0}^{T}Q^{\varepsilon}(t)[\sigma(t,x^{*}(t),E(x^{*}(t))-\sigma(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t))]dt\\
&+E\int_{0}^{T}\int_{\Theta}K^{\varepsilon}(t)[g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))-g(t,\Lambda^{*}(t,\theta),E(\Lambda^{*}(t,\theta)),u^{*}(t)))]\mu\left(d\theta\right)dt\\
&+E\int_{0}^{T}\int_{\Theta}R^{\varepsilon}(t,\theta)\left[c(t,x^{*}(t),\theta)-c(t,x^{\varepsilon}(t),\theta)\right]\mu\left(d\theta\right)dt+E\int_{0}^{T}H_{x}^{\varepsilon}(t)\left(x^{*}(t)-x^{\varepsilon}(t)\right)dt\\
&+E\int_{0}^{T}E\left[H_{\widetilde{x}}^{\varepsilon}(t)\right]\left(E(x^{*}(t))-E(x^{\varepsilon}(t))\right)dt+E\int_{0}^{T}H_{y}^{\varepsilon}(t)\left(y^{*}(t)-y^{\varepsilon}(t)\right)dt\\
&+E\int_{0}^{T}E(H_{\widetilde{y}}^{\varepsilon}(t))\left(E(y^{*}(t))-E(y^{\varepsilon}(t)))\right)dt+E\int_{0}^{T}H_{z}^{\varepsilon}(t)\left(z^{*}(t)-z^{\varepsilon}(t)\right)dt\\
&+E\int_{0}^{T}E(H_{\widetilde{z}}^{\varepsilon}(t))\left(E(z^{*}(t))-E(z^{\varepsilon}(t))\right)dt+E\int_{0}^{T}\int_{\Theta}H_{r}^{\varepsilon}(t)\left[r^{*}(t,\theta)-r^{\varepsilon}(t,\theta)\right]\mu\left(d\theta\right)dt.
\end{aligned}
\tag{98}
$$

Proof. First, by a simple computation, we get

$$
\begin{aligned}
&d\left(x^{*}(t)-x^{\varepsilon}(t)\right)\\
&=[f(t,x^{*}(t),E(x^{*}(t),u^{*}(t))-f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t),u^{\varepsilon}(t))]dt\\
&+(\sigma(t,x^{*}(t),E(x^{*}(t))-\sigma(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t)))\,dW(t)\\
&+\int_{\Theta}\left[c(t,x^{*}(t_{-}),\theta)-c(t,x^{\varepsilon}(t_{-}),\theta)\right]N(d\theta,dt)
\end{aligned}
\tag{99}
$$

$$
\begin{aligned}
&d\left(y^{*}(t)-y^{\varepsilon}(t)\right)\\
&=\int_{\Theta}[g(t,\Lambda^{*}(t,\theta),E(\Lambda^{*}(t,\theta)),u^{*}(t))-g(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t))]\mu\left(d\theta\right)dt\\
&+\left(z^{*}(t)-z^{\varepsilon}(t)\right)dW(t)+\int_{\Theta}\left[r^{*}(t,\theta)-r^{\varepsilon}(t,\theta)\right]N(d\theta,dt).
\end{aligned}
\tag{100}
$$

By applying integration by parts formula for jumps to $\Psi^{\varepsilon}(T)\left(x^{*}(T)-x^{\varepsilon}(T)\right)$, and since $x^{\varepsilon}(0)-x^{*}(0)=0$, (see *Lemma A1, Appendix*) we get

$$
\begin{aligned}
&E\left\{\Psi^{\varepsilon}(T)\left(x^{*}(T)-x^{\varepsilon}(T)\right)\right\}\\
&=E\int_{0}^{T}\Psi^{\varepsilon}(t)d\left(x^{*}(t)-x^{\varepsilon}(t)\right)+E\int_{0}^{T}\left(x^{*}(t)-x^{\varepsilon}(t)\right)d\Psi^{\varepsilon}(t)\\
&+E\int_{0}^{T}Q^{\varepsilon}(t)[\sigma(t,x^{*}(t),E(x^{*}(t))-\sigma(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t)))]dt\\
&+E\int_{0}^{T}\int_{\Theta}R^{\varepsilon}(t,\theta)\left[c(t,x^{*}(t),\theta)-c(t,x^{\varepsilon}(t),\theta)\right]\mu\left(d\theta\right)dt\\
&=I_{1}(\varepsilon)+I_{2}(\varepsilon)+I_{3}(\varepsilon)+I_{4}(\varepsilon).
\end{aligned}
\tag{101}
$$

From (99), we obtain

$$
\begin{aligned}
&I_{1}(\varepsilon)=E\int_{0}^{T}\Psi^{\varepsilon}(t)d\left(x^{*}(t)-x^{\varepsilon}(t)\right)\\
&=E\int_{0}^{T}\Psi^{\varepsilon}(t)[f(t,x^{*}(t),E(x^{*}(t),u^{*}(t))-f(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t)),u^{\varepsilon}(t))]dt,
\end{aligned}
\tag{102}
$$

similarly, by applying (57), we get

$$\begin{aligned}I_2(\varepsilon) &= E\int_0^T (x^*(t)-x^\varepsilon(t))\,d\Psi^\varepsilon(t)\\ &= E\int_0^T (x^*(t)-x^\varepsilon(t))\,[H_x^\varepsilon(t)+E(H_{\widetilde{x}}^\varepsilon(t))]dt\\ &= E\int_0^T H_x^\varepsilon(t)\,(x^*(t)-x^\varepsilon(t))\,dt+\int_0^T E(H_{\widetilde{x}}^\varepsilon(t))\,(E(x^*(t))-E(x^\varepsilon(t)))dt.\end{aligned} \tag{103}$$

By standard arguments, we obtain

$$I_3(\varepsilon)=E\int_0^T Q^\varepsilon(t)[\sigma(t,x^*(t),E(x^*(t))-\sigma(t,x^\varepsilon(t),E(x^\varepsilon(t))]dt, \tag{104}$$

and

$$I_4(\varepsilon)=E\int_0^T\int_\Theta R^\varepsilon(t,\theta)\,[c(t,x^*(t),\theta)-c(t,x^\varepsilon(t),\theta)]\,\mu\,(d\theta)\,dt. \tag{105}$$

The duality relation (96) follows from combining (102)∼(105) together with (101).
Let us turn to the second duality relation (97). By applying *integration by parts formula* to $K^\varepsilon(t)\,[y^\varepsilon(t)-y^*(t)]$, we get

$$\begin{aligned}&E\,(K^\varepsilon(T)\,(y^\varepsilon(T)-y^*(T)))=E\,\{K^\varepsilon(0)\,(y^\varepsilon(0)-y^*(0))\}\\ &+E\int_0^T K^\varepsilon(t)d\,(y^*(t)-y^\varepsilon(t))+E\int_0^T (y^*(t)-y^\varepsilon(t))\,dK^\varepsilon(t)\\ &+E\int_0^T (z^*(t)-z^\varepsilon(t))\,[H_z^\varepsilon(t)+E(H_{\widetilde{z}}^\varepsilon(t))]dt+E\int_0^T\int_\Theta H_r^\varepsilon(t)\,[r^*(t,\theta)-r^\varepsilon(t,\theta)]\,\mu\,(d\theta)\,dt\\ &=I_1(\varepsilon)+I_2(\varepsilon)+I_3(\varepsilon)+I_4(\varepsilon)+I_5(\varepsilon).\end{aligned} \tag{106}$$

Let us turn to the first term $I_2(\varepsilon)$. From (100) we get

$$\begin{aligned}&I_2(\varepsilon)=E\int_0^T K^\varepsilon(t)d\,(y^*(t)-y^\varepsilon(t))\\ &=E\int_0^T K^\varepsilon(t)[g(t,\Lambda^\varepsilon(t,\theta),E(\Lambda^\varepsilon(t,\theta)),u^\varepsilon(t))-g(t,\Lambda^*(t,\theta),E(\Lambda^*(t,\theta)),u^*(t))]dt,\end{aligned} \tag{107}$$

from (57), we obtain

$$\begin{aligned}I_3(\varepsilon)&=E\int_0^T (y^*(t)-y^\varepsilon(t))\,dK^*(t)\\ &=E\int_0^T (y^*(t)-y^\varepsilon(t))\,(H_y^\varepsilon(t)+E(H_{\widetilde{y}}^\varepsilon(t)))dt\\ &=E\int_0^T H_y^\varepsilon(t)\,(y^*(t)-y^\varepsilon(t))\,dt+E\int_0^T E(H_{\widetilde{y}}^\varepsilon(t))\,(E(y^*(t))-E(y^\varepsilon(t))))dt\end{aligned} \tag{108}$$

and

$$\begin{aligned}I_4(\varepsilon)&=E\int_0^T (z^*(t)-z^\varepsilon(t))\,[H_z^\varepsilon(t)+E(H_{\widetilde{z}}^\varepsilon(t))]\,dt\\ &=E\int_0^T H_z^\varepsilon(t)\,(z^*(t)-z^\varepsilon(t))\,dt+E\int_0^T E(H_{\widetilde{z}}^\varepsilon(t))\,(E(z^*(t))-E(z^\varepsilon(t)))\,dt.\end{aligned} \tag{109}$$

and

$$I_5(\varepsilon)=E\int_0^T\int_\Theta H_r^\varepsilon(t)\,[r^*(t,\theta)-r^\varepsilon(t,\theta)]\,\mu\,(d\theta)\,dt. \tag{110}$$

From (55) and since

$$I_1(\varepsilon) = E\left\{K^*(0)\left(y^\varepsilon(0) - y^*(0)\right)\right\}$$
$$= -E\{\left[\phi_y\left(y(0), E\left(y(0)\right)\right) + E(\phi_{\widetilde{y}}(y(0), E\left(y(0)\right)))\right]\left(y^\varepsilon(0) - y^*(0)\right)\}$$
$$= -E\left[\phi_y\left(y(0), E\left(y(0)\right)\right)\left(y^\varepsilon(0) - y^*(0)\right)\right] - E(\phi_{\widetilde{y}}\left(y(0), E\left(y(0)\right)\right))[E\left(y^\varepsilon(0)\right) - E(y^*(0))], \tag{111}$$

the duality relation (97) follows immediately by combining (107)∼(111) together with (106). Finally inequality (98) follows from combining (96) and (97) □

Now we are able to state and prove the sufficient conditions for near-optimality for our mean-field control problem, which is the second main result of this work.

Theorem 2.4.1 (*Sufficient Near-optimality Maximum Principle*) Let Assumptions (H4) hold. Let $u^\varepsilon(\cdot)$ be some admissible control, $(x^\varepsilon(\cdot), y^\varepsilon(\cdot), z^\varepsilon(\cdot), r^\varepsilon(\cdot,\cdot))$ and $(\Psi^\varepsilon(\cdot), Q^\varepsilon(\cdot), K^\varepsilon(\cdot), R^\varepsilon(\cdot,\cdot))$ be the solution to (52) and (55) respectively associated with $u^\varepsilon(\cdot)$. If some $\varepsilon > 0$ and for any $u(\cdot) \in \mathcal{U}([0,T])$ the following near-maximality relation holds

$$\begin{aligned} &E\int_0^T H(t, \Lambda^\varepsilon(t,\theta), E(\Lambda^\varepsilon(t,\theta)), u^\varepsilon(t), \Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta))dt \\ &\geq \max_{u(\cdot)\in\mathcal{U}([0,T])} E\int_0^T H(t, \Lambda^\varepsilon(t,\theta), E(\Lambda^\varepsilon(t,\theta)), u(t), \Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta))dt - \varepsilon, \end{aligned} \tag{112}$$

then $u^\varepsilon(\cdot)$ is a near-optimal control of order $\varepsilon^{\frac{1}{2}}$, i.e,

$$J\left(\zeta, u^\varepsilon(\cdot)\right) \leq \inf_{u(\cdot)\in\mathcal{U}([0,T])} J\left(\zeta, u(\cdot)\right) + C\varepsilon^{\frac{1}{2}}, \tag{113}$$

where C is a positive constant independent from ε.

Proof. The key step in the proof of our result is to show that $H_u(t, \Lambda^\varepsilon(t,\theta), E(\Lambda^\varepsilon(t,\theta)), u^\varepsilon(t), \Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta))$ is very small and estimate it in terms of ε. We first fix an $\varepsilon > 0$ and define a new metric $\widehat{\rho}$ on $\mathcal{U}([0,T])$, by setting

$$\widehat{\rho}(u(\cdot), v(\cdot)) = E\int_0^T |u(t) - v(t)|\,\varsigma_t^\varepsilon dt,$$

where ς_t^ε is defined by

$$\varsigma_t^\varepsilon = 1 + |\Psi^\varepsilon(t)| + |K^\varepsilon(t)| + |Q^\varepsilon(t)| + |R^\varepsilon(t,\theta)| \geq 1.$$

Obviously $\widehat{\rho}$ is a metric and it is a complete metric as a weighted $\mathbb{L}^1-$norm. Now, we define a functional $\mathcal{L}$ on $\mathcal{U}([0,T])$ by

$$\mathcal{L}\left(u(\cdot)\right) = E\int_0^T H(t, \Lambda^\varepsilon(t,\theta), E(\Lambda^\varepsilon(t,\theta)), u(t), \Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta))dt. \tag{114}$$

By using assumption (92) then a simple computation shows that

$$\begin{aligned} &\left|\mathcal{L}\left(u(\cdot)\right) - \mathcal{L}\left(v(\cdot)\right)\right| \\ &\leq E\int_0^T |H(t, \Lambda^\varepsilon(t,\theta), E(\Lambda^\varepsilon(t,\theta)), u(t), \Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta)) \\ &-H(t, \Lambda^\varepsilon(t,\theta), E(\Lambda^\varepsilon(t,\theta)), v(t), \Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta)))|\,dt \\ &\leq CE\int_0^T |u(t) - v(t)|\,\varsigma_t^\varepsilon dt \leq C\widehat{\rho}(u(\cdot), v(\cdot)), \end{aligned}$$

which implies that $\mathcal{L}$ given by (114) is continuous on $\mathcal{U}([0,T])$ with respect to $\widehat{\rho}$. Now by using (52) and Ekeland's variational principle**,** there exists a $\overline{u}^{\varepsilon}(\cdot) \in \mathcal{U}([0,T])$ such that

$$\widehat{\rho}(\overline{u}^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot)) \leq \varepsilon^{\frac{1}{2}}, \tag{115}$$

and

$$\begin{aligned} &E\int_0^T \mathcal{H}(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), \overline{u}^{\varepsilon}(t))dt \\ &= \max_{u(\cdot)\in\mathcal{U}([0,T])} E\int_0^T \mathcal{H}(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), u(t))dt, \end{aligned} \tag{116}$$

where

$$\begin{aligned} &\mathcal{H}(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), u(t)) \\ &= H(t, \Lambda(t,\theta), E(\Lambda(t,\theta)), u(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) \\ &- \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon} \left| u(t) - \overline{u}^{\varepsilon}(t) \right|. \end{aligned}$$

The maximum condition (116) implies a pointwise maximum condition namely, for $a.e.\ t \in [0,T]$ and $\mathbb{P}-a.s$,

$$\begin{aligned} &\mathcal{H}(t, \Lambda^{\varepsilon}(t), E(\Lambda^{\varepsilon}(t)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), \overline{u}^{\varepsilon}(t)) \\ &= \max_{u\in\mathbb{U}} \mathcal{H}(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), u), \end{aligned}$$

then by using Clarke's generalized gradient, see [17, Proposition A.1, (3)] we have

$$0 \in \partial_u^{\circ} \mathcal{H}(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), \overline{u}^{\varepsilon}(t)). \tag{117}$$

Since the absolute value function $u \longmapsto |u - \overline{u}^{\varepsilon}(t)|$ is not differentiable in $\overline{u}^{\varepsilon}(t)$ (locally Lipschitz), then the Clarke's generalized gradient, see [17, Proposition A.1, Example] shows that

$$\partial_u^{\circ}(\varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon} \left| u - \overline{u}^{\varepsilon}(t) \right|) = conv\{-\varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon},\ \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon}\} = [-\varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon},\ \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon}]. \tag{118}$$

By using (118) and fact that the Clarke's generalized gradient of the sum of two functions is contained in the sum of the Clarke's generalized gradient of the two functions, (*Proposition A1, (4)*) we get

$$\begin{aligned} &\partial_u^{\circ} \mathcal{H}(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta), \overline{u}^{\varepsilon}(t)) \\ &\subset H_u(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \overline{u}^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) + \left[-\varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon},\ \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon}\right]. \end{aligned}$$

Since H is differentiable in u by Assumption (H4), the differential inclusion (117) implies that there is $e^{\varepsilon}(t)$ such that $e^{\varepsilon}(t) \in \left[-\varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon},\ \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon}\right]$, satisfies

$$H_u(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), \overline{u}^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) + e^{\varepsilon}(t) = 0. \tag{119}$$

Consequently, by noting Assumption (H4) and the fact that $|e^{\varepsilon}(t)| \leq \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon}$ we have

$$
\begin{aligned}
&|H_u(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t)\ , R^{\varepsilon}(t,\theta))| \\
&\leq |H_u(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) \\
&- H_u(t, \Lambda^{\varepsilon}(t), E(\Lambda^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t),\ \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta))| \\
&+ |H_u(t, \Lambda^{\varepsilon}(t), E(\Lambda^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta))| \\
&\leq C\varsigma_t^{\varepsilon} |u^{\varepsilon}(t) - \overline{u}^{\varepsilon}(t)| + \varepsilon^{\frac{1}{2}} \varsigma_t^{\varepsilon}.
\end{aligned}
$$

Now, since $H(t, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta))$ is concave with respect to $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$, we obtain

$$
\begin{aligned}
&H(t, \Lambda(t,\theta), E(\Lambda(t,\theta)), u(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) \\
&- H(t, \Lambda^{\varepsilon}(t), E(\Lambda^{\varepsilon}(t)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta)) \\
&\leq H_x^{\varepsilon}(t)(x(t) - x^{\varepsilon}(t)) + E(H_{\widetilde{x}}^{\varepsilon}(t))(E(x(t) - x^{\varepsilon}(t))) + H_y^{\varepsilon}(t)(y(t) - y^{\varepsilon}(t)) \\
&+ E(H_{\widetilde{y}}^{\varepsilon}(t))(E(y(t) - y^{\varepsilon}(t))) + H_z^{\varepsilon}(t)(z(t) - z^{\varepsilon}(t))) + E(H_{\widetilde{z}}^{\varepsilon}(t))(E(z(t) - z^{\varepsilon}(t))) \\
&+ \int_{\Theta} H_r^{\varepsilon}(t)(r(t,\theta) - r^{\varepsilon}(t,\theta)))\mu(d\theta) + H_u^{\varepsilon}(t)(u(t) - u^{\varepsilon}(t)).
\end{aligned}
\tag{120}
$$

Integrating this inequality with respect to t and taking expectations, we obtain

$$
\begin{aligned}
&E\int_0^T \{H(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t),\ R^{\varepsilon}(t,\theta)) \\
&- H(t, \Lambda^{\varepsilon}(t,\theta), E(\Lambda^{\varepsilon}(t,\theta)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), Q^{\varepsilon}(t), K^{\varepsilon}(t), R^{\varepsilon}(t,\theta))\}dt \\
&\leq E\int_0^T \{H_x^{\varepsilon}(t)(x(t) - x^{\varepsilon}(t))) + E(H_{\widetilde{x}}^{\varepsilon}(t))(E(x(t) - x^{\varepsilon}(t))) \\
&+ H_y^{\varepsilon}(t)(y(t) - y^{\varepsilon}(t))) + E(H_{\widetilde{y}}^{\varepsilon}(t))(E(y(t) - y^{\varepsilon}(t))) \\
&+ H_z^{\varepsilon}(t)(z(t) - z^{\varepsilon}(t))) + E(H_{\widetilde{z}}^{\varepsilon}(t))(E(z(t) - z^{\varepsilon}(t))) \\
&+ \int_{\Theta} H_r^{\varepsilon}(t)(r(t,\theta) - r^{\varepsilon}(t,\theta)))\mu(d\theta)\}dt + C\varepsilon^{\frac{1}{2}},
\end{aligned}
\tag{121}
$$

by applying Lemma 4.1 (98), the concavity of $h(\cdot,\cdot)$, and the convexity of $\phi(\cdot,\cdot)$ we have

$$
J(\zeta, u(\cdot)) \geq J(\zeta, u^{\varepsilon}(\cdot)) - C\varepsilon^{\frac{1}{2}}.
$$

Since $u(\cdot)$ is an arbitrary admissible control, we get

$$
J(\zeta, u^{\varepsilon}(\cdot)) \leq \inf_{u(\cdot)\in\mathcal{U}([0,T])} J(\zeta, u(\cdot)) + C\varepsilon^{\frac{1}{2}}.
$$

This completes the proof of (113) □

The following corollary gives a sufficient condition for an admissible control $u^{\varepsilon}(\cdot)$ to be $\varepsilon-$optimal for our mean-field control problem (52)-(53).

Corollary 2.4.1 (Sufficient condition for $\varepsilon-$optimality). Under the assumptions of *Theorem 3.2.1* a sufficient condition for an admissible control $u^{\varepsilon}(\cdot)$ to be $\varepsilon-$optimal for our mean-field control problem (52)-(53) is

$$E\int_0^T H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))dt+(\frac{\varepsilon}{C})^2$$
$$\geq \sup_{u(\cdot)\in\mathcal{U}([0,T])} E\int_0^T H(t,\Lambda^{\varepsilon}(t,\theta),E(\Lambda^{\varepsilon}(t,\theta)),u(t),\Psi^{\varepsilon}(t),Q^{\varepsilon}(t),K^{\varepsilon}(t),R^{\varepsilon}(t,\theta))dt.$$

13. Applications: Time-inconsistent mean-variance portfolio selection problem combined with a recursive utility functional maximization

It is well known that mean-variance portfolio selection problem introduced by Markowitz [43] is a time-inconsistent optimal control problem in the sense that it does not satisfy Bellman's optimality principle and therefore the usual dynamic programming approach fails. In this section, we will apply our maximum principle of near-optimality to study mean-variance portfolio selection problem mixed with a recursive utility functional optimization in a financial market. In this section we will apply our maximum principle to study a perturbed mean-variance portfolio selection problem mixed with a recursive utility functional optimization in a financial market and we will derive the explicit expression for the optimal (and any near-optimal) portfolio selection strategy.

Suppose that we are given a mathematical market consisting of two investment possibilities:
The first asset is a risk-free security whose price $P_0(t)$ evolves according to the ordinary differential equation (ODE):

$$dP_0\left(t\right)=P_0\left(t\right)\rho(t)dt,\ t\in\left[0,T\right],\ \ P_0\left(0\right)>0, \tag{122}$$

where $\rho(\cdot):[0,T]\rightarrow\mathbb{R}_+$ is a locally bounded deterministic function.
A risky security (e.g. a stock) where the price $P_1\left(t\right)$ at time t is given by

$$\begin{cases} dP_1\left(t\right)=P_1\left(t_-\right)\left[\varsigma(t)dt+\sigma(t)dW(t)+\int_\Theta\xi_t\left(\theta\right)N\left(d\theta,dt\right)\right]. \\ P_1\left(0\right)>0, \end{cases} \tag{123}$$

where $\varsigma(\cdot),\sigma(\cdot)\ :\ [0,T]\ \rightarrow\ \mathbb{R}$ are bounded deterministic functions such that $\varsigma(t),\sigma(t)\ \neq\ 0$ and $\varsigma(t)>\rho(t),\ \forall t\in[0,T]$.

In order to ensure that $P_1\left(t\right)>0$ for all $t\in[0,T]$ we assume that: $\xi_t\left(\theta\right)>-1$ for $\mu-$almost all $\theta\in\Theta$ and all $t\in[0,T]$ and $\int_\Theta\xi_t^2\left(\theta\right)\mu(d\theta)$ is bounded.
A portfolio is an $(\mathcal{F}_t)-$predictable process $\left(e_0\left(t\right),e_1\left(t\right)\right)$ giving the number of units of the risk free and the risky security held at time t. Let $\nu(t)=e_1\left(t\right)P_1\left(t\right)$ denote the amount invested in the risky security. We call the control process $\nu(\cdot)$ a portfolio strategy.
Let $x^{\nu}(0)=\zeta>0$ be an initial wealth. By combining (122) and (123) we introduce the wealth dynamics:

$$\begin{cases} dx^{\nu}(t)=\left[\rho(t)x^{\nu}(t)+(\varsigma(t)-\rho(t))\nu(t)\right]dt \\ +\sigma(t)\nu(t)dW(t)+\int_\Theta\xi_t\left(\theta\right)\nu(t)N\left(d\theta,dt\right),\ \ x^{\nu}(0)=\zeta. \end{cases} \tag{124}$$

Let $\mathbb{U}$ be a compact convex subset of $\mathbb{R}$. We denote $\mathcal{U}([0,T])$ the set of admissible $(\mathcal{F}_t)-$predictable portfolio strategies $\nu(\cdot)$ valued in $\mathbb{U}$.
Under the conditions that

$$E((x^{\nu}(T)) = a, \tag{125}$$

where a is a given real number. This problem has been formulated and solved using both the Bellman's dynamic programming and using the maximum principle, see ([16, 28]).

In this section, without condition (125) the objective is to use our near-optimal maximum principle to study the mean-variance portfolio selection problem mixed with a recursive utility functional maximization. We consider a small investor endowed with an initial wealth $x^{\nu}(0) > 0$ who chooses at each time t his or her portfolio strategy $\nu(t)$. The investor wants to choose a portfolio strategy $\nu^{\varepsilon}(\cdot) \in \mathcal{U}([0,T])$ which near-maximizes the expected utility functional.
We assume that we originally have a family of optimization problems parameterized by some parameter $\varepsilon > 0$ representing the complexity of the cost functional

$$J(\zeta, \nu(\cdot)) = \frac{\gamma}{2} Var(x^{\nu}(T)) + E\left\{\int_0^T \varepsilon\varphi(\nu(t))dt - x^{\nu}(T)\right\} + y^{\nu}(0), \tag{126}$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is a nonlinear, convex and continuously differentiable function independent of ε. Further, we define the wealth process $(x(t))$ and the recursive utility process $y(t)$ corresponding to $\nu(\cdot) \in \mathcal{U}([0,T])$ as the solutions to the following FBSDEJs:

$$\begin{cases} dx(t) = [\rho(t)x(t) + (\varsigma(t) - \rho(t))\nu(t)]\,dt + \sigma(t)\nu(t)dW(t) + \int_\Theta \xi_t(\theta)\,\nu(t)N(d\theta, dt). \\ -dy(t) = [\rho(t)x(t) + (\varsigma(t) - \rho(t))\nu(t) - cy(t)]\,dt - z(t)dW(t) - \int_\Theta r(t,\theta)\,\nu(t)N(d\theta, dt). \\ x(0) = \zeta, \quad y(T) = x(T). \end{cases} \tag{127}$$

By setting $\varepsilon = 0$ in (126) leads to

$$\begin{aligned} J_0(\zeta, \nu(\cdot)) &= \frac{\gamma}{2} Var(x^{\nu}(T)) - E(x^{\nu}(T)) + y^{\nu}(0). \\ &= E\left[\frac{\gamma}{2} x^{\nu}(T)^2 - x^{\nu}(T)\right] - \frac{\gamma}{2}\left[E(x^{\nu}(T))\right]^2 + y^{\nu}(0). \end{aligned} \tag{128}$$

According to the maximum condition ((61), *Theorem 3.1 with* $\varepsilon = 0$*)*, and since $\nu^*(\cdot)$ is optimal we immediately get

$$(\varsigma(t) - \rho(t))\,(\Psi^*(t) + K^*(t)) + \sigma(t)Q^*(t) + \int_\Theta \xi_t(\theta)\,R^*(t,\theta)\,\mu(d\theta) = 0. \tag{129}$$

The adjoint equation (55) being

$$\begin{cases} d\Psi^*(t) = -\rho(t)\,(K^*(t) + \Psi^*(t))\,dt + Q^*(t)dW(t) + \int_\Theta R^*(t,\theta)\,N(d\theta, dt). \\ \Psi^*(T) = \gamma\,(x^*(T) + E(x^*(T))) - 1 - K^*(T). \\ dK^*(t) = -cK^*(t)dt, \ K^*(0) = -1, \ t \in [0,T]. \end{cases} \tag{130}$$

In order to solve the above equation (130) and find the expression of optimal portfolio strategy $\nu^*(\cdot)$ we conjecture a process $\Psi^*(t)$ of the form:

$$\Psi^*(t) = \phi_1(t)x^*(t) + \phi_2(t)E\left(x^*(t)\right) + \phi_3(t), \tag{131}$$

where $\phi_1(\cdot), \phi_2(\cdot)$ and $\phi_3(\cdot)$ are deterministic differentiable functions. (see [28, 17, 23, 49, 5, 69] for other models of conjecture). From last equation in (130), which is a simple ODE, we get

$$K^*(t) = -\exp\left(-ct\right). \tag{132}$$

From (127), we get

$$d(E(x^*(t)) = (\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\nu^*(t)))dt.$$

Applying Itô's formula to (131) (*see Lemma A1, Appendix*) in virtue of (127), we have

$$\begin{aligned} d\Psi^*(t) &= \phi_1(t)\left\{\left[\rho(t)x^*(t) + (\varsigma(t) - \rho(t))\nu^*(t)\right]dt + \sigma(t)\nu^*(t)dW(t)\right. \\ &\quad \left. + \int_\Theta \xi_t\left(\theta\right)\nu^*(t)N\left(d\theta, dt\right)\right\} + x^*(t)\phi_1'(t)dt \\ &\quad + \phi_2(t)\left[\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\nu^*(t))\right]dt + E\left(x^*(t)\right)\phi_2'(t)dt + \phi_3'(t)dt, \end{aligned}$$

which implies that

$$\begin{cases} d\Psi^*(t) = \left\{\phi_1(t)\left[\rho(t)x^*(t) + (\varsigma(t) - \rho(t))\,\nu^*(t)\right]\right. \\ +x^*(t)\phi_1'(t) + \phi_2(t)\rho(t)E(x^*(t)) + \phi_2(t)(\varsigma(t) - \rho(t))E(\nu^*(t)) \\ \left. + \phi_2'(t)E\left(x^*(t)\right) + \phi_3'(t)\right\}dt + \phi_1(t)\sigma(t)\nu^*(t)dW(t) \\ + \int_\Theta \phi_1(t)\xi_t\left(\theta\right)\nu^*(t)N\left(d\theta, dt\right). \\ \Psi^*(T) = \phi_1(T)x^*(T) + \phi_2(T)E\left(x^*(T)\right) + \phi_3(T), \end{cases} \tag{133}$$

where $\phi_1'(t)$, $\phi_2'(t)$, and $\phi_3'(t)$ denotes the derivatives with respect to t.
Next, comparing (133) with (130), we get

$$\begin{aligned} &-\rho(t)\left(K^*(t) + \Psi^*(t)\right) \\ &= \phi_1(t)\left[\rho(t)x^*(t) + (\varsigma(t) - \rho(t))\nu^*(t)\right] + x^*(t)\phi_1'(t) \\ &+\phi_2(t)\left[\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\nu^*(t))\right] + \phi_2'(t)E\left(x^*(t)\right) + \phi_3'(t), \end{aligned} \tag{134}$$

$$Q^*(t) = \phi_1(t)\sigma(t)\nu^*(t), \tag{135}$$

$$R^*(t, \theta) = \phi_1(t)\xi_t\left(\theta\right)\nu^*(t). \tag{136}$$

By looking at the terminal condition of $\Psi^*(t)$, in (133), it is reasonable to get

$$\phi_1(T) = \gamma,\ \phi_2(T) = -\gamma,\ \phi_3(T) = -1 - K^*(T). \tag{137}$$

Combining (134) and (131) we deduce that $\phi_1(\cdot), \phi_2(\cdot)$ and $\phi_3(\cdot)$ satisfying the following ODEs*:*

$$\begin{cases} \phi_1'(t) = -2\rho(t)\phi_1(t), \quad \phi_1(T) = \gamma. \\ \phi_2'(t) = -2\rho(t)\phi_2(t), \quad \phi_2(T) = -\gamma. \\ \phi_3'(t) + \rho(t)\phi_3(t) = \rho(t)\exp\{-ct\}. \\ \phi_3(T) = \exp\{-cT\} - 1. \end{cases} \tag{138}$$

By solving the first two ordinary differential equations in (138) we obtain

$$\phi_1(t) = -\phi_2(t) = \gamma \exp\left\{2\int_t^T \rho(s)ds\right\}. \tag{139}$$

Using *Integrating factor method* for the third equation in (138), we get

$$\phi_3(t) = a(t)^{-1}\left[\exp(-cT) - 1 - \int_t^T a(s)\rho(s)\exp\{-cs\}\,ds\right], \tag{140}$$

where the integrating factor is

$$a(t) = \exp\left\{\int_t^T \rho(s)ds\right\}, \; a(T) = 1.$$

Combining (129), (132), (135) and (136) we get the explicit optimal portfolio section strategy in the state feedback form involving both $x^*(\cdot)$ and $E(x^*(\cdot)$:

$$\begin{aligned} &\nu^*(t, x^*(t), E(x^*(t))) \\ &= (\rho(t) - \varsigma(t))\left[\phi_1(t)\left(\sigma^2(t) + \textstyle\int_\Theta \xi_t^2(\theta)\,\mu(d\theta)\right)\right]^{-1} \\ &\times\left[\phi_1(t)\left(x^*(t) - E(x^*(t))\right) + \phi_3(t) - \exp\{-ct\}\right], \end{aligned} \tag{141}$$

and

$$E(\nu^*(t, x^*(t), E(x^*(t)))) = \frac{(\rho(t) - \varsigma(t))\left[\phi_3(t) - \exp\{-ct\}\right]}{\phi_1(t)\left(\sigma^2(t) + \int_\Theta \xi_t^2(\theta)\,\mu(d\theta)\right)}. \tag{142}$$

See [**?**, **?**] for other class of control problems in state feedback form.
However, the Hamiltonian H^ε for the problem (127)-(126) can be rewritten in the form

$$\begin{aligned} &H^\varepsilon(t, x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, \nu, \Psi, Q, K, R) \\ &= [\rho(t)x(t) + (\varsigma(t) - \rho(t))\nu(t)](K(t) - \Psi(t)) \\ &+\sigma(t)Q(t)\nu(t) + \nu(t)\textstyle\int_\Theta \xi_t(\theta)\,R(t,\theta)\,\mu(d\theta) - \varepsilon\varphi(\nu(t)) \\ &= H(t, x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, \nu, \Psi, Q, K, R) - \varepsilon\varphi(\nu(t)), \end{aligned} \tag{143}$$

for all $(x, y, z, r, \nu, \Psi, Q, K, R)$. Therefore, if $(x^*(t), y^*(t), z^*(t), r^*(\cdot))$ denotes the optimal trajectory to the (unperturbed) control problem (127)-(128) we can express the difference of Hamiltonian at different control points but at this fixed optimal trajectory in the following way

$$\begin{aligned}
&H^{\varepsilon}\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&-H^{\varepsilon}\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu^{*}, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&= H\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&-H\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu^{*}, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&-\varepsilon(\varphi\left(\nu(t)\right)-\varphi\left(\nu^{*}(t)\right)).
\end{aligned} \tag{144}$$

Using the fact that the function $\varphi\left(\cdot\right)$ is continuously differentiable and $\mathbb{U}$ is compact convex subset in $\mathbb{R}$ it follows that

$$-\varepsilon\big[\varphi\left(\nu(t)\right)-\varphi\left(\nu^{*}(t)\right)\big] \leq \varepsilon\big|\varphi'\left(\nu(t)\right)\big|\left|\nu(t)-\nu^{*}(t)\right| \leq C\varepsilon.$$

Now, employing the above fact, taking $max_{\nu\in\mathbb{U}}$ in (144) and using the optimality of $\nu^{*}(\cdot)$ we get

$$\begin{aligned}
&\max_{\nu\in\mathbb{U}} H^{\varepsilon}\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&- H^{\varepsilon}\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu^{*}(t), \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&\leq \max_{\nu\in\mathbb{U}} H\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&- H\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, \nu^{*}(t), \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&+ \varepsilon \max_{\nu\in\mathbb{U}} \left\{\big|\varphi'\left(\nu(t)\right)\big|\left|\nu(t)-\nu^{*}(t)\right|\right\} \\
&\leq H\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu^{*}(t), \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&- H\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu^{*}(t), \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) + C\varepsilon \\
&= C\varepsilon,
\end{aligned}$$

which implies the near-maximality property of $\nu^{*}(\cdot)$

$$\begin{aligned}
&H^{\varepsilon}\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu^{*}(t), \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) \\
&\geq \max_{\nu\in\mathbb{U}} H^{\varepsilon}\left(t, x^{*}, \widetilde{x}^{*}, y^{*}, \widetilde{y}^{*}, z^{*}, \widetilde{z}^{*}, r^{*}, \nu, \Psi^{*}, Q^{*}, K^{*}, R^{*}\right) - C\varepsilon.
\end{aligned}$$

Finally, since the function $\varphi(\cdot)$ is convex, the Hamiltonian H^{ε} is concave. By using sufficient maximum principle (*Theorem 2*), the portfolio strategy $\nu^{*}(\cdot)$ is indeed a near-optimal for the problem (127)-(126). □

Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem

Part IV
Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem

Absract. In this chapter, we study mean-field type optimal stochastic control problem for systems governed by mean-field controlled forward-backward stochastic differential equations with jump processes, in which the coefficients depend on the marginal law of the state process through its expected value. The control variable is allowed to enter both diffusion and jump coefficients. Moreover, the cost functional is also of mean-field type. Necessary conditions for optimal control for these systems in the form of maximum principle are established by means of convex perturbation techniques. As an application, time-inconsistent mean-variance portfolio selection mixed with a recursive utility functional optimization problem is discussed to illustrate the theoretical results.
Keywords: Mean-field forward-backward stochastic differential equation with jumps; Optimal stochastic control; Mean-field maximum principle; Mean-variance portfolio selection with recursive utility functional; Time-inconsistent control problem.

14. Introduction

In this work, we consider stochastic optimal control for systems governed by nonlinear mean-field controlled forward-backward stochastic differential equations with Poisson jump processes (FBSDEJs) of the form

$$\begin{cases} dx(t) = f\left(t, x(t), E(x(t)), u(t)\right) dt + \sigma\left(t, x(t), E(x(t)), u(t)\right) dW(t) \\ \qquad\qquad + \int_\Theta c(t, x(t_-), E(x(t_-)), u(t), \theta) N\left(d\theta, dt\right), \\ dy(t) = -\int_\Theta g(t, x(t), E(x(t)), y(t), E(y(t)), z(t), E(z(t)), r\left(t, \theta\right), \\ \qquad\qquad u(t))\mu\left(d\theta\right) dt + z(t) dW(t) + \int_\Theta r\left(t, \theta\right) N\left(d\theta, dt\right), \\ x(0) = \zeta, y(T) = h\left(x(T), E\left(x(T)\right)\right), \end{cases} \tag{145}$$

where f, σ, c, g, h are given maps and the initial condition ζ is an $\mathcal{F}_0$-measurable random variable. The mean-field FBSDEJs-(145) called McKean-Vlasov systems are obtained as the mean square limit of an interacting particle system of the form

$$\begin{cases} dx_n^j(t) = f(t, x_n^j(t), \frac{1}{n}\sum_{i=1}^n x_n^i(t), u(t))dt \\ \qquad\qquad +\sigma(t, x_n^j(t), \frac{1}{n}\sum_{i=1}^n x_n^i(t), u(t))dW^j(t) \\ \qquad\qquad + \int_\Theta c(t, x_n^j(t_-), \frac{1}{n}\sum_{i=1}^n x_n^i(t_-), u(t), \theta) N^j(d\theta, dt), \\ dy_n^j(t) = -\int_\Theta g(t, x_n^j(t), \frac{1}{n}\sum_{i=1}^n x_n^i(t), y_n^j(t), \frac{1}{n}\sum_{i=1}^n y_n^i(t), z_n^j(t), \\ \qquad\qquad \frac{1}{n}\sum_{i=1}^n z_n^i(t), r(t,\theta), u(t))\mu(d\theta)\, dt \\ \qquad\qquad +z_n^j(t)dW^j(t) + \int_\Theta r(t,\theta) N^j(d\theta, dt), \end{cases}$$

where $(W^j(\cdot) : j \geq 1)$ is a collection of independent Brownian motions and $(N^j(\cdot,\cdot) : j \geq 1)$ is a collection of independent Poisson martingale measure. Noting that mean-field FBSDEJs-(145) occur naturally in the probabilistic analysis of financial optimization problems and the optimal control of dynamics of the McKean-Vlasov type. Moreover, the above mathematical mean-field approaches play an important role in different fields of economics, finance, physics, chemistry and game theory.
The expected cost to be minimized over the class of admissible control has the form

$$\begin{aligned} J(u(\cdot)) = E\Big[\int_0^T \int_\Theta \ell(t, x(t), E(x(t)), y(t), E(y(t)), z(t), E(z(t)), \\ r(t,\theta), u(t))\, \mu(d\theta)\, dt + \phi(x(T), E(x(T))) + \varphi(y(0), E(y(0)))\Big]. \end{aligned} \tag{146}$$

where ℓ, ϕ, φ is an appropriate functions. This cost functional is also of mean-field type, as the functions ℓ, ϕ, φ depend on the marginal law of the state process through its expected value. It worth mentioning that since the cost functional J is possibly a nonlinear function of the expected value stands in contrast to the standard formulation of a control problem. This leads to a so-called time-inconsistent control problem where the Bellman dynamic programming does not hold. The reason for this is that one cannot apply the law of iterated expectations on the cost functional.
An admissible control $u(\cdot)$ is an $\mathcal{F}_t-$adapted and square-integrable process with values in a nonempty convex subset $\mathcal{A}$ of $\Re$. We denote by $\mathcal{U}([0,T])$ the set of all admissible controls. Any admissible control $u^*(\cdot) \in \mathcal{U}([0,T])$ satisfying

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}([0,T])} J(u(\cdot)), \tag{147}$$

is called an optimal control.
The mean-field stochastic differential equation was introduced by Kac [39] as a stochastic model for the Vlasov-kinetic equation of plasma and the study of this model was initiated by McKean [45]. Since then, many authors made contributions on mean-field stochastic problems and their applications, see for instance [56, 42, 46, 15, 8, 10, 55, 48, 17, 23, 24, 18, 25, **?**, 62, 5, 49, 51, 52, 66]. In a recent work, mean-field games for large population multiagent systems with Markov jump parameters have been investigated in Wang and Zhang [56]. Decentralized tracking-type games for large population multi-agent systems with mean-field coupling have been studied in Li and Zhang [42].

Discrete-time indefinite mean-field linear-quadratic optimal control problem has been investigated in Ni, Zhang and Li [46]. Discrete time mean-field stochastic linear-quadratic optimal control problems with applications have been derived by in Elliott, Li and Ni [15]. In Buckdahn, Li and Peng [8] a general notion of mean-field BSDE associated with a mean-field SDE was obtained in a natural way as a limit of some high dimensional system of FBSDEs governed by a $d-$dimensional Brownian motion, and influenced by positions of a large number of other particles. In Buckdahn, Djehiche and Li [10], a general maximum principle was introduced for a class of stochastic control problems involving SDEs of mean-field type. However, sufficient conditions of optimality for mean-field SDE have been established by Shi [55]. In Meyer-Brandis, ∅ksendal and Zhou [48] a stochastic maximum principle of optimality for systems governed by controlled Itô-Levy process of mean-field type was proved by using Malliavin calculus. Mean-field singular stochastic control problems have been investigated in Hafayed and Abbas [17]. More interestingly, mean-field type stochastic maximum principle for optimal singular control has been studied in Hafayed [23], in which convex perturbations used for both absolutly continuous and singular components. The maximum principle for optimal control of mean-field FBSDEJs with uncontrolled diffusion has been studied in Hafayed [24]. The necessary and sufficient conditions for near-optimality of mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [18]. Singular optimal control for mean-field forward-backward stochastic systems and applications to finance have been investigated in Hafayed [25]. Second-order necessary conditions for optimal control of mean-field jump diffusion have been obtained by Hafayed and Abbas [**?**]. Under partial information, mean-field type stochastic maximum principle for optimal control has been investigated in Wang, Zhang and Zhang [62]. Under the condition that the control domain is convex, Andersson and Djehiche [5] and Li [49] investigated problems for two types of more general controlled SDEs and cost functionals, respectively. The linear-quadratic optimal control problem for mean-field SDEs has been studied by Yong [67] and Shi [55]. The mean-field stochastic maximum principle for jump diffusions with applications has been investigated in Shen and Siu [51] Recently, maximum principle for mean-field jump-diffusions stochastic delay differential equations and its applications to finance have been derived by Yang, Meng and Shi [52]. Mean-field optimal control for backward stochastic evolution equations in Hilbert spaces has been investigated in Xu and Wu [66].

The optimal control problems for stochastic systems described by Brownian motions and Poisson jumps have been investigated by many authors including [53, 54, 9, 21, 61, 16]. The necessary and sufficient conditions of optimality for FBSDEJs were obtained by Shi and Wu [53]. General maximum principle for fully coupled FBSDEJs has been obtained in Shi [54], where the author generalized Yong's maximum principle [68] to jump case.

In this work, our main goal is to derive a maximum principle for optimal stochastic control of mean-field FBSDEJs, where the coefficient depend not only on the state process but also its marginal law of the state process through its expected value. The cost functional is also of mean-field type. The mean-field problem under consideration is not simple extension from the mathematical point of view, but also provide interesting models in many applications such as mathematical finance; (mean-variance portfolio selection problems), optimal control for mean-field systems. The proof of our result is based on convex perturbation method. These necessary conditions are described in terms two adjoint processes, corresponding to the mean-field forward and backward components with jumps and a maximum conditions on the Hamiltonian. In the end, as an application to finance; a mean-variance portfolio selection mixed with a recursive utility optimization problem is given, where explicit expression of the optimal portfolio selection strategy is obtained in feedback form involving both state process and its marginal distribution, via the solutions of Riccati ordinary differential equations. To streamline the presentation of this work, we only study the one dimensional case.

The rest of this work is structured as follows. In Section 2 we formulate the mean-field stochastic control problem and describe the assumptions of the model. Section 3 is devoted to prove our mean-field stochastic maximum principle. As an illustration, using these results, a mean-variance portfolio selection mixed problem with recursive utility (time-inconsistent solution) is discussed in the last section.

15. Problem statement and preliminaries

We consider stochastic optimal control problem of mean-field type of the following kind. Let $T > 0$ be a fixed time horizon and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a fixed filtered probability space equipped with a $P-$completed right continuous filtration on which a $one-$dimensional Brownian motion $W = (W(t))_{t\in[0,T]}$ is defined. Let η be a homogeneous $(\mathcal{F}_t)$-Poisson point process independent of W. We denote by $\widetilde{N}(d\theta, dt)$ the random counting measure induced by η, defined on $\Theta \times \Re_+$, where Θ is a fixed nonempty subset of $\Re$ with its Borel $\sigma-$field $\mathcal{B}(\Theta)$. Further, let $\mu(d\theta)$ be the local characteristic measure of η, i.e. $\mu(d\theta)$ is a σ-finite measure on $(\Theta, \mathcal{B}(\Theta))$ with $\mu(\Theta) < +\infty$. We then define $N(d\theta, dt) := \widetilde{N}(d\theta, dt) - \mu(d\theta)\,dt$, where $N(\cdot,\cdot)$ is Poisson martingale measure on $\mathcal{B}(\Theta) \times \mathcal{B}(\Re_+)$ with local characteristics $\mu(d\theta)\,dt$. We assume that $(\mathcal{F}_t)_{t\in[0,T]}$ is $P-$augmentation of the natural filtration $(\mathcal{F}_t^{(W,N)})_{t\in[0,T]}$ defined as follows

$$\mathcal{F}_t^{(W,N)} := \sigma\left(W(s) : s \in [0,t]\right) \vee \sigma\Big(\int_0^s \int_B N(d\theta, dr) : s \in [0,t], B \in \mathcal{B}(\Theta)\Big) \vee \mathcal{G}_0,$$

where $\mathcal{G}_0$ denotes the totality of $P-$null sets, and $\sigma_1 \vee \sigma_2$ denotes the σ-field generated by $\sigma_1 \cup \sigma_2$.

In the sequel, $L^2_{\mathcal{F}}([0,T];\Re)$ denotes the Hilbert space of $\mathcal{F}_t-$adapted processes $(X(t))_{t\in[0,T]}$ such that $E\int_0^T |X(t)|^2\,dt < +\infty$ and $\mathcal{M}^2_{\mathcal{F}}([0,T];\Re)$ denotes the Hilbert space of $\mathcal{F}_t-$ predictable processes $(\psi(t,\theta))_{t\in[0,T]}$ defined on $[0,T]\times\Theta$ such that $E\int_0^T\int_\Theta |\psi(t,\theta)|^2\,\mu(\theta)dt < +\infty$. In what follows, C represents a generic constants, which can be different from line to line. For simplicity of notation, we still use $f_x(t) = \frac{\partial f}{\partial x}(t, x^*(\cdot), E(x^*(\cdot)), u^*(\cdot))$, *etc.*

Throughout this work, we also assume that the functions $f, \sigma : [0,T]\times\Re\times\Re\times\mathcal{A}\to\Re$, $c : [0,T]\times\Re\times\mathcal{A}\times\Theta\to\Re$, $g, \ell : [0,T]\times\Re\times\Re\times\Re\times\Re\times\Re\times\Re\times\Re\times\mathcal{A}\to\Re$ and $h, \phi, \varphi : \Re\times\Re\to\Re$ satisfy the following standing assumptions

Assumption (H1) 1. The functions f, σ and c are global Lipschitz in $(x, \widetilde{x}, u)$ and g is global Lipschitz in $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$.

2. The functions $f, \sigma, \ell, c, g, h, \phi, \varphi$ are continuously differentiable in their variables including $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$.

Assumption (H2) 1. The derivatives of f, σ, g, ϕ with respect to their variables including $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$ are bounded, and

$$\int_{\Theta}(|c_x(t,x,\widetilde{x},u,\theta)|^2 + |c_{\widetilde{x}}(t,x,\widetilde{x},u,\theta)|^2 + |c_u(t,x,\widetilde{x},u,\theta)|^2)\mu(d\theta) < +\infty.$$

2. The derivatives b_ρ are bounded by $C(1+|x|+|\widetilde{x}|+|y|+|\widetilde{y}|+|z|+|\widetilde{y}|+|r|+|u|)$ for $\rho = x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u$ and $b = f, \sigma, g, c, \ell$. Moreover, $\varphi_y, \varphi_{\widetilde{y}}$ are bounded by $C(1+|y|+|\widetilde{y}|)$ and $h_x, h_{\widetilde{x}}$ are bounded by $C(1+|x|+|\widetilde{x}|)$.

3. For all $t \in [0,T]$, $f(t,0,0,0), g(t,0,0,0,0,0,0,0,0,0) \in L^2_{\mathcal{F}}([0,T];\Re)$, $\sigma(t,0,\ 0,\ 0) \in L^2_{\mathcal{F}}([0,T];\Re\times\Re)$ and $c(t,0,0,0,\cdot) \in \mathcal{M}^2_{\mathcal{F}}([0,T];\Re)$.

Under the assumptions (H1) and (H2), the FBSDEJ-(145) has a unique solution $(x(t), y(t), z(t), r(t,\cdot)) \in L^2_{\mathcal{F}}([0,T];\Re) \times L^2_{\mathcal{F}}([0,T];\Re) \times L^2_{\mathcal{F}}([0,T];\Re) \times L^2_{\mathcal{F}}([0,T];\Re)$. (See [51] Theorem 3.1, for mean-field BSDE with jumps)

For any $u(\cdot) \in \mathcal{U}([0,T])$ with its corresponding state trajectories $(x(\cdot), y(\cdot),\ z(\cdot), r(\cdot,\cdot))$ we introduce the following adjoint equations

$$\begin{cases} d\Psi(t) = -\{f_x(t)\Psi(t) + E(f_{\widetilde{x}}(t)\Psi(t)) + \sigma_x(t)Q(t) + E(\sigma_{\widetilde{x}}(t)Q(t)) \\ \quad + \int_\Theta [g_x(t,\theta)K(t) + E(g_{\widetilde{x}}(t,\theta)K(t)) + c_x(t,\theta)R(t,\theta) \\ \quad + E(c_{\widetilde{x}}(t,\theta)R(t,\theta)) + \ell_x(t,\theta) + E(\ell_{\widetilde{x}}(t,\theta))]\mu(d\theta)\}dt \\ \quad + Q(t)dW(t) + \int_\Theta R_t(\theta)N(d\theta,dt), \\ \Psi(T) = -[h_x(x(T),E(x(T)))K(T) + E(h_{\widetilde{x}}(x(T),E(x(T)))K(T))] \\ \quad + \phi_x(x(T),E(x(T))) + E(\phi_{\widetilde{x}}(x(T),E(x(T)))), \\ dK(t) = \int_\Theta [g_y(t,\theta)K(t) + E(g_{\widetilde{y}}(t,\theta)K(t)) - \ell_y(t,\theta) - E(\ell_{\widetilde{y}}(t,\theta))]\mu(d\theta)dt \\ \quad + \int_\Theta [g_z(t,\theta)K(t) + E(g_{\widetilde{z}}(t,\theta)K(t)) - \ell_z(t,\theta) - E(\ell_{\widetilde{z}}(t,\theta))]\mu(d\theta)dW(t) \\ \quad + \int_\Theta (g_r(t,\theta)K(t) - \ell_r(t,\theta))N(d\theta,dt), \\ K(0) = -(\varphi_y(0) + E(\varphi_{\widetilde{y}}(0))). \end{cases} \tag{148}$$

Note that the first adjoint equation (backward) corresponding to the forward component turns out to be a linear mean-field backward SDE with jumps, and the second adjoint equation (forward) corresponding to the backward component turns out to be a linear mean-field (forward) SDE with jump processes. Further, we define the Hamiltonian function

$$H : [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathcal{A}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R},$$

associated with the stochastic control problem (145)-(146) as follows

$$
\begin{aligned}
H\left(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r,u,\Psi,Q,K,R\right) &:= \Psi(t)f\left(t,x,\widetilde{x},u\right)+Q(t)\sigma\left(t,x,\widetilde{x},u\right) \\
&\quad -\textstyle\int_{\Theta}\left[K(t)g\left(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r,u\right)+R\left(t,\theta\right)c\left(t,x,\widetilde{x},u,\theta\right)\right) \\
&\quad +\ell\left(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r,u\right)]\,\mu\left(d\theta\right).
\end{aligned}
\tag{149}
$$

If we denote by

$$
H(t):=H(t,x(t),\widetilde{x}(t),y(t),\widetilde{y}(t),z(t),\widetilde{z}(t),r(t,\cdot),u(t),\Psi(t),Q(t),K(t),R(t,\cdot)),
$$

then the adjoint equation (148) can be rewritten as the following stochastic Hamiltonian system's type

$$
\begin{cases}
-d\Psi(t)=(H_{x}\left(t\right)+E(H_{\widetilde{x}}\left(t\right)))dt-Q(t)dW(t)-\int_{\Theta}R\left(t,\theta\right)N\left(d\theta,dt\right), \\
\Psi(T)=-[h_{x}\left(x(T),E(x(T))\right)K(T)+E(h_{\widetilde{x}}(x(T),E(x_{T}(t)))K(T))] \\
\qquad +\phi_{x}(x(T),E(x(T)))+E(\phi_{\widetilde{x}}(x(T),E(x(T)))). \\
-dK(t)=(H_{y}\left(t\right)+E(H_{\widetilde{y}}\left(t\right)))dt+(H_{z}\left(t\right)+E\left(H_{\widetilde{z}}\left(t\right)\right))dW(t) \\
\qquad +\int_{\Theta}H_{r}\left(t,\theta\right)N\left(d\theta,dt\right) \\
K(0)=-(\varphi_{y}\left(0\right)+E(\varphi_{\widetilde{y}}\left(0\right))).
\end{cases}
\tag{150}
$$

Thanks to Lemma 3.1 in Shen and Siu [51], under assumptions (H1), (H2), the adjoint equations (148) admits a unique solution $(\Psi(t),Q(t),K(t),R(t,\cdot))$ such that

$$
\begin{aligned}
&(\Psi(t),Q(t),K(t),R(t,\cdot)) \\
\in\;& L_{\mathcal{F}}^{2}([0,T]\,;\Re)\times L_{\mathcal{F}}^{2}([0,T]\,;\Re)\times L_{\mathcal{F}}^{2}([0,T]\,;\Re)\times\mathcal{M}_{\mathcal{F}}^{2}([0,T]\,;\Re).
\end{aligned}
$$

Moreover, since the derivatives of $f,\sigma,c,g,h,\varphi,\phi$ with respect to $x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r$ are bounded, we deduce from standard arguments that there exists a constant $C>0$ such that

$$
\begin{aligned}
&E\Big\{\sup_{t\in[0,T]}\left|\Psi(t)\right|^{2}+\sup_{t\in[0,T]}\left|K(t)\right|^{2}+\textstyle\int_{0}^{T}\left|Q(t)\right|^{2}dt \\
&+\textstyle\int_{0}^{T}\int_{\Theta}\left|R\left(t,\theta\right)\right|^{2}\mu\left(d\theta\right)dt\Big\}<C.
\end{aligned}
\tag{151}
$$

16. Mean-field type necessary conditions for optimal control of FBSDEJs

In this section, we establish a set of necessary conditions of Pontraygin's type for a stochastic control to be optimal where the system evolves according to nonlinear controlled mean-field FBSDEJs. Convex perturbation techniques are applied to prove our mean-field stochastic maximum principle. The following theorem constitutes the main contribution of this work.

Let $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot,\cdot))$ be the trajectory of the mean-field FBSDEJ-(145) corresponding to the optimal control $u^*(\cdot)$, and $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot), R^*(\cdot,\cdot))$ be the solution of adjoint equation (148) corresponding to $u^*(\cdot)$.

Theorem 3.3.1. *(Maximum principle for mean-field FBSDEJs). Let Assumptions (H1) and (H2) hold. If $(u^*(\cdot), x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot,\cdot))$ is an optimal solution of the mean-field control problem (145)-(146). Then the maximum principle holds, that is $\forall u \in \mathcal{A}$*

$$H_u(t, \lambda^*(t,\theta), u^*, \Lambda^*(t,\theta))(u - u^*(t)) \geq 0,\ P-a.s., a.e.,\ t \in [0,T]\,, \tag{152}$$

where $\lambda^*(t,\theta) = (x^*(t), E(x^*(t)), y^*(t), E(y^*(t)), z^*(t), E(z^*(t)), r^*(t,\theta))$ and
$\Lambda^*(t,\theta) = (\Psi^*(t), Q^*(t), K^*(t), R^*(t,\theta))$.

We derive the variational inequality (152) in several steps, from the fact that

$$J\left(u^\varepsilon(\cdot)\right) \geq J\left(u^*(\cdot)\right), \tag{153}$$

Since the control domain $\mathcal{A}$ is convex and for any given admissible control $u(\cdot) \in \mathcal{U}([0,T])$ the following perturbed control process

$$u^\varepsilon(t) = u^*(t) + \varepsilon\left(u(t) - u^*(t)\right),$$

is also an element of $\mathcal{U}([0,T])$.
Let $\lambda^\varepsilon(t,\theta) = (x^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), r^\varepsilon(t,\theta))$ be the solution of state equation (145) and $\Lambda^\varepsilon(t,\theta) = (\Psi^\varepsilon(t), Q^\varepsilon(t), K^\varepsilon(t), R^\varepsilon(t,\theta))$ be the solution of the adjoint equation (148) corresponding to perturbed control $u^\varepsilon(\cdot)$.

Variational equations. We introduce the following variational equations which have a mean-field type. Let $(x_1^\varepsilon(\cdot), y_1^\varepsilon(\cdot), z_1^\varepsilon(\cdot), r_1^\varepsilon(\cdot,\cdot))$ be the solution of the following forward-backward stochastic system described by Brownian motions and Poisson jumps of mean-field type

$$\begin{cases} dx_1^\varepsilon(t) = \left\{f_x(t)x_1^\varepsilon(t) + f_{\widetilde{x}}(t)E(x_1^\varepsilon(t)) + f_u(t)u(t)\right\}dt \\ \ + \left\{\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + \sigma_u(t)u(t)\right\}dW(t) \\ \ + \int_\Theta [c_x(t,\theta)x_1^\varepsilon(t) + c_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t)) + c_u(t,\theta)u(t)]N(d\theta,dt), \\ x_1^\varepsilon(0) = 0, \\ dy_1^\varepsilon(t) = -\int_\Theta\{g_x(t,\theta)x_1^\varepsilon(t) + g_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t)) + g_y(t,\theta)y_1^\varepsilon(t) \\ +g_{\widetilde{y}}(t,\theta)E(y_1^\varepsilon(t)) + g_z(t,\theta)z_1^\varepsilon(t) + g_{\widetilde{z}}(t,\theta)E(z_1^\varepsilon(t)) + g_r(t,\theta)r_1^\varepsilon(t,\theta) \\ +g_u(t,\theta)u(t)\}\mu(d\theta)dt + z_1^\varepsilon(t)dW(t) + \int_\Theta r_1^\varepsilon(t,\theta)N(d\theta,dt), \\ y_1^\varepsilon(T) = -\left[h_x(T) + E(h_{\widetilde{x}}(T))\right]x_1^\varepsilon(T). \end{cases} \tag{154}$$

Duality relations. Our first Lemma below deals with the duality relations between $\Psi^*(t)$, $x_1^\varepsilon(t)$ and $K^*(t)$, $y_1^\varepsilon(t)$. This Lemma is very important for the proof of Theorem 3.1.

Lemma 3.3.1. We have

$$\begin{aligned}
&E\left(\Psi^*(T)x_1^\varepsilon(T)\right)=E\int_0^T[\Psi^*(t)f_u(t)u(t)+Q^*(t)\sigma_u(t)u(t)\\
&+\int_\Theta R^*(t,\theta)c_u(t,\theta)u(t)\mu(d\theta)]dt-E\int_0^T\int_\Theta\{x_1^\varepsilon(t)g_x(t,\theta)K(t)\\
&+x_1^\varepsilon(t)E(g_{\widetilde{x}}(t,\theta)K(t))+x_1^\varepsilon(t)\ell_x(t,\theta)+x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta))\}\,\mu(d\theta)dt,
\end{aligned}\tag{155}$$

similarly, we get

$$\begin{aligned}
&E\left(K^*(T)y_1^\varepsilon(T)\right)\\
&=-E\left\{[\varphi_y(y(0),E(y(0)))+E(\varphi_{\widetilde{y}}(y(0),E(y(0))))]y_1^\varepsilon(0)\right\}\\
&+E\int_0^T\int_\Theta\{K^*(t)g_x(t,\theta)x_1^\varepsilon(t)+K^*(t)g_{\widetilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right)\\
&-K^*(t)g_u(t,\theta)u(t)-y_1^\varepsilon(t)\ell_y(t,\theta)-y_1^\varepsilon(t)E(\ell_{\widetilde{y}}(t,\theta))\\
&-z_1^\varepsilon(t)\ell_z(t,\theta)-z_1^\varepsilon(t)E(\ell_{\widetilde{z}}(t,\theta))-r_1^\varepsilon(t,\theta)\ell_r(t,\theta)\}\,\mu(d\theta)dt,
\end{aligned}\tag{156}$$

and

$$\begin{aligned}
&E\left\{[\phi_x(x(T),E(x(T)))+E(\phi_{\widetilde{x}}(x(T),E(x(T))))]\,x_1^\varepsilon(T)\right\}\\
&+E\left\{[\varphi_y\left(y(0),E(y(0))\right)+E(\varphi_{\widetilde{y}}(y(0),E(y(0))))]y_1^\varepsilon(0)\right\}\\
&=E\int_0^T\int_\Theta\{x_1^\varepsilon(t)\ell_x(t,\theta)+x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta))-y_1^\varepsilon(t)\ell_y(t,\theta)\\
&-y_1^\varepsilon(t)E(\ell_{\widetilde{y}}(t,\theta))-z_1^\varepsilon(t)\ell_z(t,\theta)-z_1^\varepsilon(t)E(\ell_{\widetilde{z}}(t,\theta))\\
&-r_1^\varepsilon(t,\theta)\ell_r(t,\theta)-\ell_u(t,\theta)u(t)\}\,\mu(d\theta)dt+E\int_0^T H_u(t)u(t)dt.
\end{aligned}\tag{157}$$

Proof. By applying integration by parts formula for jump processes (see Lemma A1) to $\Psi^*(t)x_1^\varepsilon(t)$, we get

$$\begin{aligned}
&E\left(\Psi^*(T)x_1^\varepsilon(T)\right)=E\int_0^T\Psi^*(t)dx_1^\varepsilon(t)+E\int_0^T x_1^\varepsilon(t)d\Psi^*(t)\\
&+E\int_0^T Q^*(t)\left[\sigma_x(t)x_1^\varepsilon(t)+\sigma_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right)+\sigma_u(t)u(t)\right]dt\\
&+E\int_0^T\int_\Theta\left[c_x(t,\theta)x_1^\varepsilon(t)+c_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t))+c_u(t,\theta)u(t)\right]R(t,\theta)\mu(d\theta)dt\\
&=I_1^\varepsilon+I_2^\varepsilon+I_3^\varepsilon+I_4^\varepsilon.
\end{aligned}\tag{158}$$

A simple computation shows that

$$\begin{aligned}
&I_1^\varepsilon=E\int_0^T\Psi^*(t)dx_1^\varepsilon(t)\\
&=E\int_0^T\left\{\Psi^*(t)f_x(t)x_1^\varepsilon(t)+\Psi^*(t)f_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right)+\Psi^*(t)f_u\left(t\right)u(t)\right\}dt,
\end{aligned}\tag{159}$$

and

$$
\begin{aligned}
I_2^\varepsilon &= E\int_0^T x_1^\varepsilon(t)d\Psi^*(t)\\
&= -E\int_0^T \{x_1^\varepsilon(t)f_x\,(t)\,\Psi^*(t)+x_1^\varepsilon(t)E\,(f_{\widetilde{x}}(t)\Psi^*(t))\\
&+x_1^\varepsilon(t)\sigma_x\,(t)\,Q^*(t)+x_1^\varepsilon(t)E\,(\sigma_{\widetilde{x}}(t)Q^*(t))\\
&+\int_\Theta [x_1^\varepsilon(t)g_x\,(t,\theta)\,K^*(t)\,+x_1^\varepsilon(t)E\,(g_{\widetilde{x}}(t,\theta)K^*(t))\\
&+x_1^\varepsilon(t)c_x\,(t,\theta)\,R\,(t,\theta)+x_1^\varepsilon(t)E(c_{\widetilde{x}}\,(t,\theta)\,R\,(t,\theta))\\
&+\,x_1^\varepsilon(t)\ell_x(t,\theta)\,+\,x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta))]\,\mu\,(d\theta)\}\,dt.
\end{aligned}
\tag{160}
$$

From (158), we get

$$
\begin{aligned}
I_3^\varepsilon &= E\int_0^T Q^*(t)\,[\sigma_x(t)x_1^\varepsilon(t)+\sigma_{\widetilde{x}}(t)E\,(x_1^\varepsilon(t))+\sigma_u(t)u(t)]\,dt\\
&= E\int_0^T Q^*(t)\sigma_x(t)x_1^\varepsilon(t)dt+E\int_0^T Q^*(t)\sigma_{\widetilde{x}}(t)E\,(x_1^\varepsilon(t))\,dt\\
&+E\int_0^T Q^*(t)\sigma_u(t)u(t)dt\\
I_4^\varepsilon &= E\int_0^T \int_\Theta [c_x(t,\theta)x_1^\varepsilon(t)+c_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t))\\
&+c_u(t,\theta)u(t)]R(t,\theta)\mu(d\theta)dt\\
&= E\int_0^T \int_\Theta c_x(t,\theta)x_1^\varepsilon(t)R(t,\theta)\mu(d\theta)dt\\
&+E\int_0^T \int_\Theta c_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t))R(t,\theta)\mu(d\theta)dt\\
&+E\int_0^T \int_\Theta c_u(t,\theta)u(t)R(t,\theta)\mu(d\theta)dt.
\end{aligned}
\tag{161}
$$

The duality relation (155) follows immediately from combining (159)∼(161) and (158).

Let us turn to second duality relation (156). By applying integration by parts formula for jump process (Lemma A1) to $K^*(t)y_1^\varepsilon(t)$, we get

$$
\begin{aligned}
&E\,(K^*(T)y_1^\varepsilon(T))\\
&= E\,(K^*(0)y_1^\varepsilon(0))+E\int_0^T K^*(t)dy_1^\varepsilon(t)+E\int_0^T y_1^\varepsilon(t)dK^*(t)\\
&+E\int_0^T \int_\Theta z_1^\varepsilon(t)[g_z(t,\theta)K^*(t)+E\,(g_{\widetilde{z}}(t,\theta)K^*(t))\\
&-\ell_z(t,\theta)-E\,(\ell_{\widetilde{z}}\,(t,\theta))]\mu\,(d\theta)\,dt\\
&+E\int_0^T \int_\Theta [r_1^\varepsilon(t,\theta)(g_r(t,\theta)K^*(t)-\ell_r(t,\theta))]\,\mu(d\theta)dt.\\
&= I_1^\varepsilon+I_2^\varepsilon+I_3^\varepsilon+I_4^\varepsilon+I_5^\varepsilon.
\end{aligned}
\tag{162}
$$

From (155), we obtain

$$
\begin{aligned}
I_2^{\varepsilon} &= E\int_0^T K^*(t)dy_1^{\varepsilon}(t) \\
&= -E\int_0^T\int_{\Theta}\{K^*(t)g_x(t,\theta)x_1^{\varepsilon}(t)+K^*(t)g_{\widetilde{x}}(t,\theta)E\left(x_1^{\varepsilon}(t)\right) \\
&+K^*(t)g_y(t,\theta)y_1^{\varepsilon}(t)+K^*(t)g_{\widetilde{y}}(t,\theta)E\left(y_1^{\varepsilon}(t)\right)+K^*(t)g_z(t,\theta)z_1^{\varepsilon}(t) \\
&+K^*(t)g_{\widetilde{z}}(t,\theta)E\left(z_1^{\varepsilon}(t)\right)+K^*(t)g_r(t,\theta)r_1^{\varepsilon}(t,\theta) \\
&+K^*(t)g_u(t,\theta)u(t)\}\,\mu(d\theta)dt,
\end{aligned}
\tag{163}
$$

from (148), we obtain

$$
\begin{aligned}
I_3^{\varepsilon} &= E\int_0^T y_1^{\varepsilon}(t)dK^*(t) \\
&= E\int_0^T\int_{\Theta}\{y_1^{\varepsilon}(t)g_y\left(t,\theta\right)K^*(t)+y_1^{\varepsilon}(t)E(g_{\widetilde{y}}(t,\theta)K^*(t)) \\
&-y_1^{\varepsilon}(t)\ell_y(t,\theta)-y_1^{\varepsilon}(t)E(\ell_{\widetilde{y}}\left(t,\theta\right))\}\mu(d\theta)dt,
\end{aligned}
\tag{164}
$$

and

$$
\begin{aligned}
I_4^{\varepsilon} &= E\int_0^T\int_{\Theta}\left[z_1^{\varepsilon}(t)g_z(t,\theta)K^*(t)+z_1^{\varepsilon}(t)E\left(g_{\widetilde{z}}(t,\theta)K^*(t)\right)\right. \\
&\qquad \left.-\,z_1^{\varepsilon}(t)\ell_z(t,\theta)-z_1^{\varepsilon}(t)E\left(\ell_{\widetilde{z}}\left(t,\theta\right)\right)\right]\mu(d\theta)dt, \\
I_5^{\varepsilon} &= E\int_0^T\int_{\Theta}\left\{r_1^{\varepsilon}(t,\theta)g_r(t,\theta)K^*(t)-r_1^{\varepsilon}(t,\theta)\ell_r(t,\theta)\right\}\mu(d\theta)dt.
\end{aligned}
\tag{165}
$$

Since

$$
\begin{aligned}
I_1^{\varepsilon} &= E\left(K^*(0)y_1^{\varepsilon}(0)\right) \\
&= -E\left\{[\varphi_y(y(0),E(y(0)))+E(\varphi_{\widetilde{y}}(y(0),E(y(0))))]y_1^{\varepsilon}(0)\right\},
\end{aligned}
$$

the duality relation (156) follows immediately by combining (163)∼(165) and (162). Let us turn to (157). Combining (155) and (156) we get

$$
\begin{aligned}
&E\left(\Psi^*(T)x_1^{\varepsilon}(T)\right)+E\left(K^*(T)y_1^{\varepsilon}(T)\right) \\
&= -E[\varphi_y(y(0),E(y(0)))+E(\varphi_{\widetilde{y}}(y(0),E(y(0)))]y_1^{\varepsilon}(0) \\
&+E\int_0^T\int_{\Theta}\{x_1^{\varepsilon}(t)\ell_x(t,\theta)+x_1^{\varepsilon}(t)E(\ell_{\widetilde{x}}(t,\theta))-\ell_y(t,\theta)-E(\ell_{\widetilde{y}}(t,\theta)) \\
&-\ell_u(t,\theta)u(t)-\ell_z(t,\theta)-E(\ell_{\widetilde{z}}(t,\theta))-r_1^{\varepsilon}(t,\theta)\ell_r(t,\theta)\}\,\mu(d\theta)dt \\
&+E\int_0^T H_u(t)u(t)dt.
\end{aligned}
$$

From (150) and (154), we get

$$
\begin{aligned}
&E\left(\Psi^*(T)x_1^{\varepsilon}(T)\right)+E\left(K^*(T)y_1^{\varepsilon}(T)\right) \\
&= [\phi_x(x(T),E(x(T)))+E(\phi_{\widetilde{x}}(x(T),E(x(T)))]\,x_1^{\varepsilon}(T).
\end{aligned}
$$

Using (149), we obtain

$$E\int_0^T \{\Psi(t)f_u(t)u(t)+Q(t)\sigma_u(t)u(t)$$
$$+\int_\Theta[-K(t)g_u(t)u(t)+R(t,\theta)c_u(t,\theta)u(t))$$
$$+\ \ell_u(t,\theta)u(t)]\mu(d\theta)\}dt=E\int_0^T H_u(t)u(t)dt,$$

which implies that

$$E\{[\phi_x(x(T),E(x(T)))+E(\phi_{\widetilde{x}}(x(T),E(x(T))))]x_1^\varepsilon(T)\}$$
$$+E\{[\varphi_y(y(0),E(y(0)))+E(\varphi_{\widetilde{y}}(y(0),E(y(0))))]y_1^\varepsilon(0)\}$$
$$=E\int_0^T\int_\Theta\{x_1^\varepsilon(t)\ell_x(t,\theta)+x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta))$$
$$-y_1^\varepsilon(t)\ell_y(t,\theta)-y_1^\varepsilon(t)E(\ell_{\widetilde{y}}(t,\theta))-z_1^\varepsilon(t)\ell_z(t,\theta)-z_1^\varepsilon(t)E(\ell_{\widetilde{z}}(t,\theta))$$
$$-\ r_1^\varepsilon(t,\theta)\ell_r(t,\theta)-\ell_u(t,\theta)u(t)\}\mu(d\theta)dt+E\int_0^T H_u(t)u(t)dt.$$

This completes the proof of (157). □

The second Lemma present the estimates of the perturbed state process $(x^\varepsilon(\cdot),y^\varepsilon(\cdot),z^\varepsilon(\cdot),r^\varepsilon(\cdot,\cdot))$.

Lemma 3.3.2. Under assumptions (H1) and (H2), the following estimations holds

$$\begin{aligned}&E(\sup_{0\leq t\leq T}|x_1^\varepsilon(t)|^2)\to 0,\ as\ \varepsilon\to 0,\\&E(\sup_{0\leq t\leq T}|y_1^\varepsilon(t)|^2)+E\int_0^T[|z_1^\varepsilon(s)|^2\\&+\int_\Theta|r_1^\varepsilon(s,\theta)|^2\mu(d\theta)]ds\to 0,\ as\ \varepsilon\to 0,\end{aligned}\tag{166}$$

$$\begin{aligned}&\sup_{0\leq t\leq T}|E(x_1^\varepsilon(t))|^2\to 0,\ as\ \varepsilon\to 0,\\&\sup_{0\leq t\leq T}|E(y_1^\varepsilon(t))|^2+\int_t^T|E(z_1^\varepsilon(s))|^2ds\\&+\int_0^T\int_\Theta|E(r_1^\varepsilon(s,\theta))|^2\mu(d\theta)ds\to 0,\ as\ \varepsilon\to 0,.\end{aligned}\tag{167}$$

$$\begin{aligned}&E(\sup_{0\leq t\leq T}|x^\varepsilon(t)-x^*(t)|^2)\to 0,\ as\ \varepsilon\to 0,\\&E(\sup_{0\leq t\leq T}|y^\varepsilon(t)-y^*(t)|^2)+E(\int_0^T|z^\varepsilon(t)-z^*(t)|^2)dt\\&+E\int_0^T\int_\Theta|r^\varepsilon(t,\theta)-r^*(t,\theta)|^2\mu(d\theta)dt\to 0,\ as\ \varepsilon\to 0,\end{aligned}\tag{168}$$

and

$$\begin{aligned}&E(\sup_{0\leq t\leq T}\left|\tfrac{1}{\varepsilon}[x^\varepsilon(t)-x^*(t)]-x_1^\varepsilon(t)\right|^2)\to 0,\ as\ \varepsilon\to 0,\\&E(\sup_{0\leq t\leq T}\left|\tfrac{1}{\varepsilon}[y^\varepsilon(t)-y^*(t)]-y_1^\varepsilon(t)\right|^2)\to 0,\ as\ \varepsilon\to 0,\\&E\int_0^T\left|\tfrac{1}{\varepsilon}[z^\varepsilon(s)-z^*(s)]-z_1^\varepsilon(s)\right|^2ds\to 0,\ as\ \varepsilon\to 0,\\&E\int_0^T\int_\Theta\left|\tfrac{1}{\varepsilon}[r^\varepsilon(s,\theta)-r^*(s,\theta)]-r_1^\varepsilon(s,\theta)\right|^2\mu(d\theta)ds\to 0,\ as\ \varepsilon\to 0.\end{aligned}\tag{169}$$

Let us also point out that the above estimates (166)-(168) can be proved by using similar arguments developed in (Lemma 4.2, Lemma 4.3 [51]) and (Lemma 2.1, [53]). So we omit their proofs.

Proof of (169). We set:

$$\begin{aligned}
\widehat{x}^{\varepsilon}(t) &= \tfrac{1}{\varepsilon}\left[x^{\varepsilon}(t)-x^{*}(t)\right]-x_{1}^{\varepsilon}(t),\\
\widehat{y}^{\varepsilon}(t) &= \tfrac{1}{\varepsilon}\left[y^{\varepsilon}(t)-y^{*}(t)\right]-y_{1}^{\varepsilon}(t),\\
\widehat{z}^{\varepsilon}(t) &= \tfrac{1}{\varepsilon}\left[z^{\varepsilon}(t)-z^{*}(t)\right]-z_{1}^{\varepsilon}(t),\\
\widehat{r}^{\varepsilon}(t,\theta) &= \tfrac{1}{\varepsilon}\left[r^{\varepsilon}(t,\theta)-r^{*}(t,\theta)\right]-r_{1}^{\varepsilon}(t,\theta),
\end{aligned} \tag{170}$$

and

$$\begin{aligned}
&f(t)=f(t,x^{*}(t),E(x^{*}(t)),u^{*}(t)),\sigma(t)=\sigma(t,x^{*}(t),E(x^{*}(t)),u^{*}(t)),\\
&c(t,\theta)=c(t,x^{*}(t),E(x^{*}(t)),u^{*}(t),\theta)\\
&g(t,\theta)=g(x^{*}(t),E(x^{*}(t)),y^{*}(t),E(y^{*}(t)),z^{*}(t),E(z^{*}(t)),r^{*}(t,\theta),u^{*}(t)).
\end{aligned}$$

From equation (145) we have

$$\begin{aligned}
&d\widehat{x}^{\varepsilon}(t)=\tfrac{1}{\varepsilon}\left[dx^{\varepsilon}(t)-dx^{*}(t)\right]-dx_{1}^{\varepsilon}(t)\\
&=\tfrac{1}{\varepsilon}[f(t,x^{*}(t)+\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)),E(x^{*}(t)+\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)),u^{\varepsilon}(t))-f(t)]dt\\
&-\left[f_{x}(t)x_{1}^{\varepsilon}(t)+f_{\widetilde{x}}(t)E\left(x_{1}^{\varepsilon}(t)\right)+f_{u}(t)u(t))\right]dt\\
&+\tfrac{1}{\varepsilon}[\sigma(t,x^{*}(t)+\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)),E(x^{*}(t)+\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t))),u^{\varepsilon}(t))\\
&-\sigma(t)]dW(t)-\left[\sigma_{x}(t)x_{1}^{\varepsilon}(t)+\sigma_{\widetilde{x}}(t)E\left(x_{1}^{\varepsilon}(t)\right)+\sigma_{u}(t)u(t)\right]dW(t)\\
&+\textstyle\int_{\Theta}[c(t,x^{*}(t)+\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)),E(x^{*}(t)+\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t))),u^{\varepsilon}(t),\theta)\\
&-c(t,\theta)]N(d\theta,dt)-\textstyle\int_{\Theta}[c_{x}(t,\theta)x_{1}^{\varepsilon}(t)+c_{\widetilde{x}}(t,\theta)E(x_{1}^{\varepsilon}(t))+c_{u}(t,\theta)u(t)]N(d\theta,dt).
\end{aligned} \tag{171}$$

We denote

$$\begin{aligned}
x^{\lambda,\varepsilon}(t)&=x^{*}(t)+\lambda\varepsilon(\widehat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)),\\
y^{\lambda,\varepsilon}(t)&=y^{*}(t)+\lambda\varepsilon(\widehat{y}^{\varepsilon}(t)+y_{1}^{\varepsilon}(t)),\\
z^{\lambda,\varepsilon}(t)&=z^{*}(t)+\lambda\varepsilon(\widehat{z}^{\varepsilon}(t)+z_{1}^{\varepsilon}(t)),\\
r^{\lambda,\varepsilon}(t,\theta)&=r^{*}(t,\theta)+\lambda\varepsilon(\widehat{r}^{\varepsilon}(t,\theta)+r_{1}^{\varepsilon}(t,\theta)),\\
u^{\lambda,\varepsilon}(t)&=u^{*}(t)+\lambda\varepsilon u(t).
\end{aligned} \tag{172}$$

By Taylor's expansion with a simple computations we show that

$$\widehat{x}^{\varepsilon}(t)=\frac{1}{\varepsilon}\left[x^{\varepsilon}(t)-x^{*}(t)\right]-x_{1}^{\varepsilon}(t)=\widetilde{I}_{1}(\varepsilon)+\widetilde{I}_{2}(\varepsilon)+\widetilde{I}_{3}(\varepsilon), \tag{173}$$

where

$$\begin{aligned}
\widetilde{I}_{1}(\varepsilon)&=\textstyle\int_{0}^{t}\int_{0}^{1}f_{x}(s,x^{\lambda,\varepsilon}(s),E(x^{\lambda,\varepsilon}(s)),u^{\lambda,\varepsilon}(s))(\widehat{x}^{\varepsilon}(s)+x_{1}^{\varepsilon}(s))d\lambda ds\\
&+\textstyle\int_{0}^{t}\int_{0}^{1}f_{\widetilde{x}}(s,x^{\lambda,\varepsilon}(s),E(x^{\lambda,\varepsilon}(s)),u^{\lambda,\varepsilon}(s))E(\widehat{x}^{\varepsilon}(s)+x_{1}^{\varepsilon}(s))d\lambda ds\\
&+\textstyle\int_{0}^{t}\int_{0}^{1}\left[f_{x}(s,x^{\lambda,\varepsilon}(s),E(x^{\lambda,\varepsilon}(s)),u^{\lambda,\varepsilon}(s))-f_{x}(s)\right]x_{1}^{\varepsilon}(s)d\lambda ds\\
&+\textstyle\int_{0}^{t}\int_{0}^{1}\left[f_{\widetilde{x}}(s,x^{\lambda,\varepsilon}(s),E(x^{\lambda,\varepsilon}(s)),u^{\lambda,\varepsilon}(s))-f_{\widetilde{x}}(s)\right]E(x_{1}^{\varepsilon}(s))d\lambda ds\\
&+\textstyle\int_{0}^{t}\int_{0}^{1}\left[f_{u}(s,x^{\lambda,\varepsilon}(s),E(x^{\lambda,\varepsilon}(s)),u^{\lambda,\varepsilon}(s))-f_{u}(s)\right]u(s)d\lambda ds,
\end{aligned} \tag{174}$$

$$
\begin{aligned}
\widetilde{I}_2(\varepsilon) &= \int_0^t \int_0^1 \sigma_x(s, x^{\lambda,\varepsilon}(s), E(x^{\lambda,\varepsilon}(s)), u^{\lambda,\varepsilon}(s))[\widehat{x}^\varepsilon(s) + x_1^\varepsilon(s)]d\lambda ds \\
&+ \int_0^t \int_0^1 \sigma_{\widetilde{x}}(s, x^{\lambda,\varepsilon}(s), E(x^{\lambda,\varepsilon}(s)), u^{\lambda,\varepsilon}(s))E[\widehat{x}^\varepsilon(s) + x_1^\varepsilon(s)]d\lambda ds \\
&+ \int_0^t \int_0^1 [\sigma_x(s, x^{\lambda,\varepsilon}(s), E(x^{\lambda,\varepsilon}(s)), u^{\lambda,\varepsilon}(s)) - \sigma_x(s)]x_1^\varepsilon(s)d\lambda ds \\
&+ \int_0^t \int_0^1 [\sigma_{\widetilde{x}}(s, x^{\lambda,\varepsilon}(s), E(x^{\lambda,\varepsilon}(s)), u^{\lambda,\varepsilon}(s)) - \sigma_{\widetilde{x}}(s)]E(x_1^\varepsilon(s))d\lambda ds \\
&+ \int_0^t \int_0^1 \left[\sigma_u(s, x^{\lambda,\varepsilon}(s), E(x^{\lambda,\varepsilon}(s)), u^{\lambda,\varepsilon}(s)) - \sigma_u(s)\right] u(s)d\lambda ds,
\end{aligned} \tag{175}
$$

and

$$
\begin{aligned}
&\widetilde{I}_3(\varepsilon) \\
&= \int_0^t \int_\Theta \int_0^1 c_x(s, x^{\lambda,\varepsilon}(s_-), E(x^{\lambda,\varepsilon}(s_-)), u^{\lambda,\varepsilon}(s), \theta)[\widehat{x}^\varepsilon(s) + x_1^\varepsilon(s)]d\lambda N(d\theta, ds) \\
&+ \int_0^t \int_\Theta \int_0^1 c_{\widetilde{x}}(s, x^{\lambda,\varepsilon}(s_-), E(x^{\lambda,\varepsilon}(s_-)), u^{\lambda,\varepsilon}(s), \theta)E[\widehat{x}^\varepsilon(s) + x_1^\varepsilon(s)]d\lambda N(d\theta, ds). \\
&+ \int_0^t \int_\Theta \int_0^1 \left[c_x(s, x^{\lambda,\varepsilon}(s_-), E(x^{\lambda,\varepsilon}(s_-)), u^{\lambda,\varepsilon}(s), \theta) - c_x(s, \theta)\right] x_1^\varepsilon(s)d\lambda N(d\theta, ds) \\
&+ \int_0^t \int_\Theta \int_0^1 [c_{\widetilde{x}}(s, x^{\lambda,\varepsilon}(s_-), E(x^{\lambda,\varepsilon}(s_-)), u^{\lambda,\varepsilon}(s), \theta) - c_{\widetilde{x}}(s, \theta)]E(x_1^\varepsilon(s))d\lambda N(d\theta, ds) \\
&+ \int_0^t \int_\Theta \int_0^1 [c_u(s, x^{\lambda,\varepsilon}(s), E(x^{\lambda,\varepsilon}(s)), u^{\lambda,\varepsilon}(s), \theta) - c_u(s, \theta)]u(s)d\lambda N(d\theta, ds),
\end{aligned} \tag{176}
$$

we proceed as in Anderson and Djehiche [5], pp 7-8, we get

$$
\begin{aligned}
&E(\sup_{0\leq t\leq T} \left|\widetilde{I}_1(\varepsilon)\right|^2) \to 0, \quad as \;\; \varepsilon \to 0, \\
&E(\sup_{0\leq t\leq T} \left|\widetilde{I}_2(\varepsilon)\right|^2) \to 0, \quad as \;\; \varepsilon \to 0,
\end{aligned} \tag{177}
$$

Applying similar estimations for the third term with the help of *Proposition 3.2* (in Appendix Bouchard and Elie [9]) we have

$$
E(\sup_{0\leq t\leq T} \left|\widetilde{I}_3(\varepsilon)\right|^2) \to 0, \quad as, \;\; \varepsilon \to 0. \tag{178}
$$

From (177) and (178) we obtain

$$
E(\sup_{0\leq t\leq T} \left|\frac{1}{\varepsilon}\left[x^\varepsilon(t) - x^*(t)\right] - x_1^\varepsilon(t)\right|^2) \to 0, \quad as \;\; \varepsilon \to 0. \tag{179}
$$

We proceed to estimate the last terms in (169). First, from (170) and since $\widehat{y}^\varepsilon(t) = \frac{1}{\varepsilon}\left[y^\varepsilon(t) - y^*(t)\right] - y_1^\varepsilon(t)$ we get

$$d\widehat{y}^\varepsilon(t)=-\frac{1}{\varepsilon}\int_\Theta[g(t,x^*(t)+\varepsilon(\widehat{x}^\varepsilon(t)+x_1^\varepsilon(t)),E(x^*(t)+\varepsilon(\widehat{x}^\varepsilon(t)+x_1^\varepsilon(t))),$$
$$y^*(t)+\varepsilon(\widehat{y}^\varepsilon(t)+y_1^\varepsilon(t)),E(y^*(t)+\varepsilon(\widehat{y}^\varepsilon(t)+y_1^\varepsilon(t))),z^*(t)+\varepsilon(\widehat{z}^\varepsilon(t)+z_1^\varepsilon(t)),$$
$$E(z^*(t)+\varepsilon(\widehat{z}^\varepsilon(t)+z_1^\varepsilon(t))),r^*(t,\theta)+\varepsilon(\widehat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)),u^\varepsilon(t))-g(t,\theta)]\mu(d\theta)dt$$
$$-\int_\Theta[g_x(t,\theta)x_1^\varepsilon(t)+g_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t))+g_y(t,\theta)y_1^\varepsilon(t)+g_{\widetilde{y}}(t,\theta)E(y_1^\varepsilon(t))$$
$$+g_z(t,\theta)z_1^\varepsilon(t)+g_{\widetilde{z}}(t,\theta)E(z_1^\varepsilon(t))+g_r(t,\theta)r_1^\varepsilon(t,\theta)+g_u(t,\theta)u(t)]\mu(d\theta)dt$$
$$+\widehat{z}^\varepsilon(t)dW(t)+\int_\Theta\widehat{r}^\varepsilon(t,\theta)N(d\theta,dt),$$

and

$$\widehat{y}^\varepsilon(T)=\frac{1}{\varepsilon}[h(x^\varepsilon(T),E(x^\varepsilon(T)))-h(x(T),E(x(T)))]$$
$$+[h_x(x(T),E(x(T)))+h_{\widetilde{x}}(x(T),E(x(T)))]x_1^\varepsilon(T).$$

Applying Taylor's expansion we get

$$-d\widehat{y}^\varepsilon(t)=\int_\Theta\int_0^1 g_x(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),$$
$$r^{\lambda,\varepsilon}(t,\theta),u^{\lambda,\varepsilon}(t))\times(\widehat{x}^\varepsilon(t)+x_1^\varepsilon(t))d\lambda\mu(d\theta)dt$$
$$+\int_\Theta\int_0^1 g_{\widetilde{x}}(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))\times E(\widehat{x}^\varepsilon(t)+x_1^\varepsilon(t))d\lambda\mu(d\theta)dt$$
$$+\int_\Theta\int_0^1[g_x(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))-g_x(t,\theta)]x_1^\varepsilon(t)d\lambda\mu(d\theta)dt$$

$$+\int_\Theta\int_0^1[g_{\widetilde{x}}(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))-g_{\widetilde{x}}(t,\theta)]E(x_1^\varepsilon(t))d\lambda\mu(d\theta)dt$$
$$+\int_\Theta\int_0^1[g_u(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))-g_u(t,\theta)]u(t)d\lambda\mu(d\theta)dt$$
$$+\int_\Theta\int_0^1 g_y(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))\times(\widehat{y}^\varepsilon(t)+y_1^\varepsilon(t))d\lambda\mu(d\theta)dt$$

$$+\int_\Theta\int_0^1 g_{\widetilde{y}}(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))\times E(\widehat{y}^\varepsilon(t)+y_1^\varepsilon(t))d\lambda\mu(d\theta)dt$$
$$+\int_\Theta\int_0^1[g_y(t,x^{\lambda,\varepsilon}(t),E(x^{\lambda,\varepsilon}(t)),y^{\lambda,\varepsilon}(t),E(y^{\lambda,\varepsilon}(t)),z^{\lambda,\varepsilon}(t),E(z^{\lambda,\varepsilon}(t)),r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t))-g_y(t,\theta)]y_1^\varepsilon(t)d\lambda\mu(d\theta)dt$$

$$+\int_\Theta \int_0^1 [g_{\widetilde{y}}(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) - g_{\widetilde{y}}(t,\theta)] \times E(y_1^\varepsilon(t)) d\lambda \mu(d\theta) dt$$
$$+\int_\Theta \int_0^1 g_z(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) \times (\widehat{z}^\varepsilon(t) + z_1^\varepsilon(t)) d\lambda \mu(d\theta) dt$$
$$+\int_\Theta \int_0^1 g_{\widetilde{z}}(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) \times E(\widehat{z}^\varepsilon(t) + z_1^\varepsilon(t)) d\lambda \mu(d\theta) dt$$
$$+\int_\Theta \int_0^1 [g_z(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) - g_z(t,\theta)] \times z_1^\varepsilon(t) d\lambda \mu(d\theta) dt$$

$$+\int_\Theta \int_0^1 [g_{\widetilde{z}}(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) - g_{\widetilde{z}}(t,\theta)] E(z_1^\varepsilon(t)) d\lambda \mu(d\theta) dt$$
$$+\int_\Theta \int_0^1 g_r(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) \times (\widehat{r}^\varepsilon(t,\theta) + r_1^\varepsilon(t,\theta)) d\lambda \mu(d\theta) dt$$
$$+\int_\Theta \int_0^1 [g_r(t, x^{\lambda,\varepsilon}(t), E(x^{\lambda,\varepsilon}(t)), y^{\lambda,\varepsilon}(t), E(y^{\lambda,\varepsilon}(t)), z^{\lambda,\varepsilon}(t), E(z^{\lambda,\varepsilon}(t)), r^{\lambda,\varepsilon}(t,\theta),$$
$$u^{\lambda,\varepsilon}(t)) - g_r(t,\theta)] r_1^\varepsilon(t,\theta) d\lambda \mu(d\theta) dt$$
$$-\widehat{z}^\varepsilon(t) dW(t) - \int_\Theta \widehat{r}^\varepsilon(t,\theta) N(d\theta, dt),$$

finally, by using similar arguments developed in [53], pp 222-224, the desired result follows. This completes the proof of (169) □

Lemma 3.3.3. Let assumptions (H1) and (H2) hold. The following variational inequality holds

$$E \int_0^T \int_\Theta [\ell_x(t,\theta) x_1^\varepsilon(t) + \ell_{\widetilde{x}}(t,\theta) E(x_1^\varepsilon(t)) + \ell_y(t,\theta) y_1^\varepsilon(t) + \ell_{\widetilde{y}}(t,\theta) E(y_1^\varepsilon(t))$$
$$+\ell_z(t,\theta) z_1^\varepsilon(t) + \ell_{\widetilde{z}}(t,\theta) E(z_1^\varepsilon(t)) + \ell_r(t,\theta) r_1^\varepsilon(t,\theta) + \ell_u(t,\theta) u(t)] \mu(d\theta) dt$$
$$+E[\phi_x(T) x_1^\varepsilon(T) + \phi_{\widetilde{x}}(T) E(x_1^\varepsilon(T))] + E[\varphi_y(0) y_1^\varepsilon(0) + \varphi_{\widetilde{y}}(0) E(y_1^\varepsilon(0))] \geq o(\varepsilon).$$

Proof. From (153) we have

$$\begin{aligned} &J(u^\varepsilon(\cdot)) - J(u^*(\cdot)) \\ &= E\Big\{ \int_0^T \int_\Theta [\ell(t, x^\varepsilon(t), E(x^\varepsilon(t)), y^\varepsilon(t), E(y^\varepsilon(t)), z^\varepsilon(t), E(z^\varepsilon(t)), r^\varepsilon(t,\theta), u^\varepsilon(t)) \\ &-\ell(t, x^*(t), E(x^*(t)), y^*(t), E(y^*(t)), z^*(t), E(z^*(t)), r^*(t,\theta), u^*(t)] \mu(d\theta) dt \\ &+ [\phi(x^\varepsilon(T), E(x^\varepsilon(T))) - \phi(x^*(T), E(x^*(T)))] \\ &+ [\varphi(x^\varepsilon(0), E(x^\varepsilon(0))) - \varphi(y^*(0), E(y^*(0)))] \Big\} \geq 0. \end{aligned} \tag{180}$$

By applying Taylor's expansion and Lemma 3.3.2 we have

$$\begin{aligned}
&\tfrac{1}{\varepsilon}E[\phi(x^{\varepsilon}(T),\widetilde{x}^{\varepsilon}(T))-\phi(x^{*}(T),\widetilde{x}^{*}(T))]\\
&=\tfrac{1}{\varepsilon}E\left\{\int_0^1 \phi_x(x^{*}(T)+\lambda(x^{\varepsilon}(T)-x^{*}(T)),\widetilde{x}^{*}(T)\right.\\
&+\lambda(\widetilde{x}^{\varepsilon}(T)-\widetilde{x}^{*}(T)))d\lambda(x^{\varepsilon}(T)-x^{*}(T))\\
&+\int_0^1 \phi_{\widetilde{x}}(x^{*}(T)+\lambda(x^{\varepsilon}(T)-x^{*}(T)),\widetilde{x}^{*}(T)\\
&+\lambda(\widetilde{x}^{\varepsilon}(T)-\widetilde{x}^{*}(T)))d\lambda(\widetilde{x}^{\varepsilon}(T)-\widetilde{x}^{*}(T)))\}+o\left(\varepsilon\right).
\end{aligned}$$

From estimate (169), we get

$$\begin{aligned}
&\tfrac{1}{\varepsilon}E[\phi(x^{\varepsilon}(T),\widetilde{x}^{\varepsilon}(T))-\phi(x^{*}(T),\widetilde{x}^{*}(T))]\\
&\to E[\phi_x(x^{*}(T),E(x^{*}(T)))x_1^{\varepsilon}(T)+\phi_{\widetilde{x}}(x^{*}(T),E(x^{*}(T)))E(x_1^{\varepsilon}(T))]\\
&=E\left[\phi_x(T)x_1^{\varepsilon}(T)+\phi_{\widetilde{x}}(T)E(x_1^{\varepsilon}(T))\right],\quad as\ \varepsilon\to 0.
\end{aligned}\tag{181}$$

Similarly, we have

$$\begin{aligned}
&\tfrac{1}{\varepsilon}E[\varphi(y^{\varepsilon}(0),\widetilde{y}^{\varepsilon}(0))-\varphi(y^{*}(0),\widetilde{y}^{*}(0))]\\
&\to E[\varphi_y(y^{*}(0),\widetilde{y}^{*}(0))y_1^{\varepsilon}(0)+\varphi_{\widetilde{y}}(y^{*}(0),\widetilde{y}^{*}(0))E\left(y_1^{\varepsilon}(0)\right)]\\
&=E\left[\varphi_y(0)y_1^{\varepsilon}(0)+\varphi_{\widetilde{y}}(0)E\left(y_1^{\varepsilon}(0)\right)\right],\quad as\ \varepsilon\to 0.
\end{aligned}\tag{182}$$

and

$$\begin{aligned}
&\tfrac{1}{\varepsilon}E\int_0^T\int_{\Theta}[\ell(t,x^{\varepsilon}(t),E(x^{\varepsilon}(t)),y^{\varepsilon}(t),E(y^{\varepsilon}(t)),z^{\varepsilon}(t),E(z^{\varepsilon}(t)),r^{\varepsilon}\left(t,\theta\right),u^{\varepsilon}(t))\\
&-\ell(t,x^{*}(t),E(x^{*}(t)),y^{*}(t),E(y^{*}(t)),z^{*}(t),E(z^{*}(t)),r^{*}\left(t,\theta\right),u^{*}(t))]\mu\left(d\theta\right)dt\\
&\to E\int_0^T\int_{\Theta}[\ell_x(t,\theta)x_1^{\varepsilon}(t)+\ell_{\widetilde{x}}(t,\theta)E(x_1^{\varepsilon}(t))+\ell_y(t,\theta)y_1^{\varepsilon}(t)+\ell_{\widetilde{y}}(t,\theta)E(y_1^{\varepsilon}(t))\\
&+\ell_z(t,\theta)z_1^{\varepsilon}(t)+\ell_{\widetilde{z}}(t,\theta)E(z_1^{\varepsilon}(t))+\ell_r(t,\theta)r_1^{\varepsilon}(t,\theta)+\ell_u(t,\theta)u(t)]\mu(d\theta)dt,\\
&as\ \ \varepsilon\to 0.
\end{aligned}\tag{183}$$

The desired result follows by combining (180)∼(183). This completes the proof of Lemma 3.3.3. □
Proof of Theorem 3.3.1. The desired result follows immediately by combining (157) in Lemma 3.3.2 and Lemma 3.3.3. □

17. Application: mean-variance portfolio selection problem mixed with a recursive utility functional, time-inconsistent solution

The mean-variance portfolio selection theory, which was first proposed in Markowitz [58] is a millstone in mathematical finance and has laid down the foundation of modern finance theory. By using sufficient maximum principle, the authors in [16] gave the expression for the optimal portfolio selection in a jump diffusion market with time consistent solutions. The near-optimal consumption-investment problem has been discussed in Hafayed, Abbas and Veverka [21]. The continuous time mean-variance portfolio selection problem has been studied in Zhou and Li [72]. The mean-variance portfolio selection problem where the state driven by SDE (without jump terms) has been studied in [5]. Optimal dividend, harvesting rate and optimal portfolio for systems governed by jump diffusion processes have been investigated in [48]. Mean-variance portfolio selection problem mixed with a recursive utility functional has been studied by Shi and Wu [53], under the condition that

$$E(x^{\pi}(T)) = c,$$

where c is a given real positive number.

In this section we will apply our mean-field stochastic maximum principle of optimality to study a mean-variance portfolio selection problem mixed with a recursive utility functional *time-inconsistent solutions* in a financial market and we will derive the explicit expression for the optimal portfolio selection strategy. This optimal control is represented by a state *feedback* form involving both $x(\cdot)$ and $E(x(\cdot))$.

Suppose that we are given a mathematical market consisting of two investment possibilities:

1. *Risk-free security (Bond price)***.** The first asset is a risk-free security whose price $P_0(t)$ evolves according to the ordinary differential equation

$$\begin{cases} dP_0(t) = \rho(t) P_0(t)\, dt, \quad t \in [0,T], \\ P_0(0) > 0, \end{cases} \tag{184}$$

where $\rho(\cdot) : [0,T] \to \Re_+$ is a locally bounded and continuous deterministic function.

2. *Risk-security (Stock price).* A risky security (e.g. a stock), where the price $P_1(t)$ at time t is given by

$$\begin{cases} dP_1(t) = P_1(t_-) \left[\varsigma(t)dt + G(t)dW(t) + \int_\Theta \xi(t,\theta) N(d\theta, dt) \right], \\ P_1(0) > 0, \quad t \in [0,T]. \end{cases} \tag{185}$$

Assumptions. In order to ensure that $P_1(t) > 0$ for all $t \in [0,T]$ we assume

1. The functions $\varsigma(\cdot) : [0,T] \to \Re$, $G(\cdot) : [0,T] \to \Re$ are bounded deterministic such that

$$\varsigma(t), G(t) \neq 0, \quad \varsigma(t) > \rho(t), \forall t \in [0,T].$$

2. $\xi(t,\theta) > -1$ for $\mu-$almost all $\theta \in \Theta$ and all $t \in [0,T]$,
3. $\int_\Theta \xi^2(t,\theta)\, \mu(d\theta)$ is bounded.

Portfolio strategy, the price dynamic with recursive utility process. A portfolio is a $\mathcal{F}_t-$predictable process $e(t) = (e_1(t), e_2(t))$ giving the number of units of the risk-free and the risky security held at time t. Let $\pi(t) = e_1(t) P_0(t)$ denote the amount invested in the risky security. We call the control process $\pi(\cdot)$ a portfolio strategy.

Let $x^{\pi}(0) = \zeta > 0$ be an initial wealth. By combining (184) and (185) we introduce the wealth process $x^{\pi}(\cdot)$ and the recursive utility process $y^{\pi}(\cdot)$ corresponding to $\pi(\cdot) \in \mathcal{U}([0,T])$ as solution of the following FBSDEJs

$$\begin{cases} dx^{\pi}(t) = [\rho(t)x^{\pi}(t) + (\varsigma(t) - \rho(t))\pi(t)]\, dt \\ \quad +G(t)\pi(t)dW(t) + \int_{\Theta} \xi\,(t,\theta)\,\pi(t)N\,(d\theta, dt)\,, \\ -dy^{\pi}(t) = [\rho(t)x^{\pi}(t) + (\varsigma(t) - \rho(t))\pi(t) - \alpha y^{\pi}(t)]\, dt \\ \quad -z^{\pi}(t)dW(t) - \int_{\Theta} r^{\pi}\,(t,\theta)\, N\,(d\theta, dt)\,, \\ x^{\pi}(0) = \zeta, \quad y^{\pi}(T) = x^{\pi}(T). \end{cases} \tag{186}$$

Mean-variance portfolio selection problem mixed with a recursive utility functional: In this section, the objective is to apply our maximum principle to study the mean-variance portfolio selection problem mixed with a recursive utility functional maximization.
The cost functional, to be minimized, is given by

$$J\,(\pi\,(\cdot)) = \frac{\gamma}{2} Var(x^{\pi}(T)) - E(x^{\pi}(T)) - y^{\pi}(0). \tag{187}$$

By a simple computation, we can shows that

$$J\,(\pi\,(\cdot)) = E[\frac{\gamma}{2} x^{\pi}(T)^2 - x^{\pi}(T)] - \frac{\gamma}{2}\,[E(x^{\pi}(T))]^2 - y^{\pi}(0), \tag{188}$$

where the wealth process $x^{\pi}(\cdot)$ and the recursive utility process $y^{\pi}(\cdot)$ corresponding to $\pi\,(\cdot) \in \mathcal{U}\,([0,T])$ is given by FBSDEJ-(186). We note that the cost functional (188) becomes a time-inconsistent control problem. Let $\mathcal{A}$ be a compact convex subset of $\Re$. We denote $\mathcal{U}\,([0,T])$ the set of admissible $\mathcal{F}_t-$predictable portfolio strategies $\pi\,(\cdot)$ valued in $\mathcal{A}$. The optimal solution is denoted by $(x^*(\cdot), \pi^*(\cdot))$. The Hamiltonian functional (149) gets the form

$$\begin{aligned} &H\,(t, x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, \pi, \Psi, Q, K, R) \\ &= [\rho(t)x(t) + (\varsigma(t) - \rho(t))\pi(t)]\,(\Psi(t) + K(t)) \\ &+G(t)\pi(t)Q(t) - \alpha K(t)y(t) + \int_{\Theta} \xi\,(t,\theta)\,\pi(t)R\,(t,\theta)\,\mu\,(d\theta)\,. \end{aligned}$$

According to the maximum condition ((152), *Theorem 3.1),* and since $\pi^*(\cdot)$ is optimal we immediately get

$$\begin{aligned} &(\varsigma(t) - \rho(t))\,(\Psi^*(t) + K^*(t)) + G(t)Q^*(t) \\ &+ \int_{\Theta} \xi\,(t,\theta)\, R^*\,(t,\theta)\,\mu\,(d\theta) = 0. \end{aligned} \tag{189}$$

The adjoint equation (148) being

$$\begin{cases} d\Psi^*(t) = -\rho(t)\,(K^*(t) + \Psi^*(t))\,dt + Q^*(t)dW(t) \\ \quad + \int_{\Theta} R^*\,(t,\theta)\, N\,(d\theta, dt)\,. \\ \Psi^*(T) = \gamma\,(x^*(T) + E(x^*(T))) - 1 - K^*(T), \\ dK^*(t) = -\alpha K^*(t)dt, \quad K^*(0) = 1, \quad t \in [0,T]\,. \end{cases} \tag{190}$$

In order to solve the above equation (190) and to find the expression of optimal portfolio strategy $\pi^*(\cdot)$ we conjecture a process $\Psi^*(t)$ of the form:

$$\Psi^*(t) = A_1(t)x^*(t) + A_2(t)E\left(x^*(t)\right) + A_3(t), \tag{191}$$

where $A_1(\cdot)$, $A_2(\cdot)$ and $A_3(\cdot)$ are deterministic differentiable functions. (see Shi and Wu [53], Shi [55], Framstad, Øksendal and Sulem [16], Li [49], Yong [50], for other models of conjecture). From last equation in (190), which is a simple ordinary differential equation (ODE in short), we get immediately

$$K^*(t) = \exp\left(-\alpha t\right). \tag{192}$$

Noting that from (186), we get

$$d(E(x^*(t)) = \left\{\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\pi^*(t))\right\} dt.$$

Applying Itô's formula to (191) (see Lemma A1, Appendix) in virtue of SDE-(186), we get

$$\begin{aligned}
&d\Psi^*(t) = A_1(t)\left\{\left[\rho(t)x^*(t) + (\varsigma(t) - \rho(t))\pi^*(t)\right] dt\right.\\
&\left.+G(t)\pi^*(t)dW(t) + \int_\Theta \xi\left(t_-, \theta\right)\pi^*(t)N\left(d\theta, dt\right)\right\}\\
&+x^*(t)A_1'(t)dt + A_2(t)\left[\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\pi^*(t))\right] dt\\
&+E\left(x^*(t)\right)A_2'(t)dt + A_3'(t)dt,
\end{aligned}$$

which implies that

$$\begin{cases}
d\Psi^*(t) = \left\{A_1(t)\left[\rho(t)x^*(t) + (\varsigma(t) - \rho(t))\pi^*(t)\right] + x^*(t)A_1'(t)\right.\\
\qquad +A_2(t)\left[\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\pi^*(t))\right]\\
\qquad \left.+ A_2'(t)E\left(x^*(t)\right) + A_3'(t)\right\} dt + A_1(t)G(t)\pi^*(t)dW(t)\\
\qquad + \int_\Theta A_1(t)\xi\left(t_-, \theta\right)\pi^*(t)N\left(d\theta, dt\right),\\
\Psi^*(T) = A_1(T)x^*(T) + A_2(T)E\left(x^*(T)\right) + A_3(T),
\end{cases} \tag{193}$$

where $A_1'(t)$, $A_2'(t)$, and $A_3'(t)$ denotes the derivatives with respect to t.
Next, comparing (193) with (190), we get

$$\begin{aligned}
&-\rho(t)\left(K^*(t) + \Psi^*(t)\right)\\
&= A_1(t)\left[\rho(t)x^*(t) + (\varsigma(t) - \rho(t))\pi^*(t)\right] + x^*(t)A_1'(t)\\
&\quad +A_2(t)\left[\rho(t)E(x^*(t)) + (\varsigma(t) - \rho(t))E(\pi^*(t))\right]\\
&\quad +A_2'(t)E\left(x^*(t)\right) + A_3'(t),
\end{aligned} \tag{194}$$

$$Q^*(t) = A_1(t)G(t)\pi^*(t), \tag{195}$$

$$R^*(t, \theta) = A_1(t)\xi\left(t, \theta\right)\pi^*(t). \tag{196}$$

By looking at the terminal condition of $\Psi^*(t)$, in (193), it is reasonable to get

$$A_1(T) = \gamma, \ \ A_2(T) = -\gamma, \ \ A_3(T) = -1 - K^*(T). \tag{197}$$

Combining (194) and (191) we deduce that $A_1(\cdot)$, $A_2(\cdot)$ and $A_3(\cdot)$ satisfying the following ODEs

$$\begin{cases} A_1'(t) = -2\rho(t)A_1(t), \quad A_1(T) = \gamma, \\ A_2'(t) = -2\rho(t)A_2(t), \quad A_2(T) = -\gamma, \\ A_3'(t) + \rho(t)A_3(t) = \rho(t)\exp\{-\alpha t\}, \ A_3(T) = -\exp\{-\alpha T\} - 1. \end{cases} \tag{198}$$

By solving the first two ordinary differential equations in (198) we obtain

$$A_1(t) = -A_2(t) = \gamma \exp\left\{ 2\int_t^T \rho(s)ds \right\}. \tag{199}$$

Using integrating factor method for the third equation in (198), we get

$$A_3(t) = -\chi(t)^{-1}\left[\exp(-\alpha T) + 1 + \int_t^T \chi(s)\rho(s)\exp\{-\alpha s\}\, ds \right], \tag{200}$$

where the integrating factor is $\chi(t) = \exp\left\{ \int_t^T \rho(s)ds \right\}, \ \ \chi(T) = 1.$
Combining (189), (192), (195) and (196) and denoting

$$\Gamma(t) = A_1(t)\left(G^2(t) + \int_\Theta \xi^2(t,\theta)\,\mu(d\theta) \right), \tag{201}$$

we get

$$\pi^*(t) = \Gamma(t)^{-1}(\rho(t) - \varsigma(t))\left[A_1(t)\left(x^*(t) - E(x^*(t)) \right) + A_3(t) - \exp(-\alpha t) \right], \tag{202}$$

and

$$E(\pi^*(t)) = \Gamma(t)^{-1}(\rho(t) - \varsigma(t))\left[A_3(t) - \exp\{-\alpha t\} \right]. \tag{203}$$

Finally, we give the explicit optimal portfolio selection strategy in the state feedback form involving both $x^*(\cdot)$ and $E(x^*(\cdot))$.

Theorem 3.4.1 The optimal portfolio strategy $\pi^*(t)$ of our mean-variance portfolio selection problem (186)-(188) is given in feedback form by

$$\begin{aligned} &\pi^*(t, x^*(t), E(x^*(t)) \\ &= \Gamma(t)^{-1}(\rho(t) - \varsigma(t))\left[A_1(t)\left(x^*(t) - E(x^*(t)) \right) + A_3(t) - \exp\{-\alpha t\} \right], \end{aligned}$$

and

$$E(\pi^*(t, x^*(t), E(x^*(t))) = \Gamma(t)^{-1}(\rho(t) - \varsigma(t))\left[A_3(t) - \exp\{-\alpha t\} \right],$$

where $A_1(t)$, $A_3(t)$ and $\Gamma(t)$ are given by (199), (200), (201) respectively.

Conclusions and future research.

In this chapter, we have discussed the necessary conditions for optimal stochastic control of mean-field forward-backward stochastic differential equations with Poisson jumps (FBSDEJs). Time-inconsistent mean-variance portfolio selection mixed with recursive utility functional optimization problem has been studied to illustrate our theoretical results.

We would like to indicate that the general maximum principle for fully coupled mean-field FBSDEJs is not addressed, and we will work for this interesting issue in the future research.

Appendix

The following result gives special case of the Itô formula for mean-field jump diffusions.

Lemma A1. (*Integration by parts formula for mean-field jump diffusions.*) Suppose that the processes $x_1(t)$ and $x_2(t)$ are given by: for $i = 1, 2$, $t \in [0, T]$

$$\begin{cases} dx_i(t) = f\left(t, x_i(t), E(x_i(t)), u(t)\right) dt + \sigma\left(t, x_i(t), E(x_i(t)), u(t)\right) dW(t) \\ \qquad + \int_\Theta g\left(t, x_i(t_-), E(x_i(t_-)), u(t), \theta\right) N\left(d\theta, dt\right), \\ x_i(0) = 0. \end{cases}$$

Then we get

$$\begin{aligned} & E\left(x_1(T)x_2(T)\right) = E\left[\int_0^T x_1(t)dx_2(t) + \int_0^T x_2(t)dx_1(t)\right] \\ & + E\int_0^T \sigma\left(t, x_1(t), E(x_1(t)), u(t)\right)\sigma\left(t, x_2(t), E(x_2(t)), u(t)\right) dt \\ & + E\int_0^T \int_\Theta g\left(t, x_1(t), E(x_1(t)), u(t), \theta\right) g\left(t, x_2(t), E(x_2(t)), u(t), \theta\right) \mu(d\theta)dt. \end{aligned}$$

Applying a similar method as in [16, Lemma 2.1], for the proof of the above Lemma.

A McKean-Vlasov optimal mixed regular-singular control problem for nonlinear stochastic systems with Poisson jump processes

Part V

A McKean-Vlasov optimal mixed regular-singular control problem for nonlinear stochastic systems with Poisson jump processes

Abstract. In this chapter, we develop the necessary conditions of optimality for a new class of mixed regular-singular control problem for nonlinear forward-backward stochastic systems with Poisson jump processes of McKean-Vlasov type. The coefficients of the system and the performance functional depend not only on the state process but also its marginal law of the state process through its expected value. The control variable has two components, the first being absolutely continuous and the second singular control. Our optimality conditions for this McKean-Vlasov's systems are established by means of convex perturbation techniques for both continuous and singular parts. In our class of McKean-Vlasov control problem, there are two types of jumps for the state processes, the inaccessible ones which come from the Poission martingale part and the predictable ones which come from the singular control part.

Keywords McKean-Vlasov systems, Empirical measures, Probability distribution, Mixed regular-singular control, Poisson jump processes. Necessary and sufficient conditions for optimal control. Mean-field Forward bakcward stochastic systems.

18. Introduction

The stochastic control problems of McKean-Vlasov type have attracted much attention because of their practical applications in many areas such as physics, chemistry, economics, finance and other areas of science and engineering. Historically, the stochastic differential equation of McKean-Vlasov type was suggested by Kac [39] in 1956 as a stochastic model for the Vlasov-kinetic equation of plasma and the study of which was initiated by McKean [45] in 1966. Since then, many authors have made contributions on stochastic differential systems of McKean-Vlasov type and applications, see, for instance, [1, 62, 15, 10, 27, 17, 23, 24, 18, 25, 29, 5, 49, 26, 51, 52, 67]. Optimal control problems for nonlinear diffusions governed by McKean-Vlasov equations on Hilbert space have been investigated by Ahmed [1]. McKean-Vlasov type stochastic maximum principle for optimal control under partial information has been investigated in Wang, Zhang and Zhang [62]. Discrete time mean-field stochastic linear-quadratic optimal control problems with applications have been investigated in Elliott, Li and Ni [15]. Stochastic maximum principle for mean-field stochastic systems governed by Lévy processes, associated with Teugels martingales measures have been investigated by Hafayed, Abbas and Abba [27]. Second order necessary and sufficient conditions of near-optimal singular control for mean-field stochastic differential equation (SDE) have been established in Hafayed and Abbas [17]. Mean-field type stochastic maximum principle for optimal singular control has been studied in Hafayed [23], in which convex perturbations have been used for both absolutely continuous and singular components. The maximum principle for optimal control of mean-field forward-backward stochastic differential equations (FBSDEs) with Poisson jump process has been studied in Hafayed [24]. The necessary and sufficient conditions for near-optimality for mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [18]. Singular optimal control for mean-field forward-backward stochastic systems driven by Brownian motions has been investigated in Hafayed [25]. Necessary and sufficient optimality conditions for mean-field forward-backward stochastic differential equations with jumps (FBSDEJs) have been established by Hafayed, Tabet and Boukaf [29]. General mean-field maximum principle was introduced in Buckdahn, Djehiche and Li [10]. Under the conditions that the control domains are convex, a various local maximum principle have been studied in [5, 49]. Second-order maximum principle for optimal stochastic control for mean-field jump diffusions was proved in Hafayed and Abbas [26]. Necessary and sufficient conditions for controlled jump diffusion with recent application in bicriteria mean-variance portfolio selection problem have been proved in Shen and Siu [51]. Recently, maximum principle for mean-field jump-diffusions stochastic delay differential equations and its applications to finance have been investigated in Yang, Meng and Shi [52]. A linear quadratic optimal control problem for mean-field stochastic differential equations has been studied in Yong [67]. In Buckdahn, Djehiche, Li and Peng [8] a general notion of mean-field backward stochastic differential equation (BSDE) associated with mean-field SDE is obtained in a natural way as a limit of some high dimensional system of FBSDEs governed by a $d-$dimensional Brownian motion, and influenced by positions of a large number of other particles. Mean-field games for large population multiagent systems with Markov jump parameters have been investigated in Wang and Zhang [56].

Stochastic maximum principle for optimal continuous control for classical FBSDEs has been investigated by many authors, see e.g. [24, 68, 9, 61]. The near-optimal stochastic control problem for jump diffusions has been investigated by Hafayed, Abbas and Veverka [21]. The near-optimality necessary and sufficient conditions for controlled FBSDEJs with applications to finance have been investigated in Hafayed, Veverka and Abbas [28]. A survey on Markovian jump systems has been investigated by Shi and Li [59]. The stochastic finite-time state estimation for discrete time-delay neural networks with Markovian jumps have been studied in Shi, Zhang and Agarwal [60].

The stochastic singular control problems have received considerable research attention in recent years due to wide applicability in a number of different areas, see for instance [2, 12, 20, 63, 22, 3, 64, 65, 4, 44]. In most classical cases, the optimal singular control problem was investigated through dynamic programming principle. The first version of maximum principle for singular stochastic control problems was obtained by Cadenillas and Haussmann [12]. In Dufour and Miller [13], the authors derived stochastic maximum principle where the singular part has a linear form. For this type of singular control problem, the reader may consult the works by Haussmann and Suo [34] and the list of references therein. The necessary and sufficient conditions for near-optimal singular control was obtained by Hafayed Abbas and Veverka [20]. Stochastic maximum principle for optimal control problems of forward backward systems involving impulse controls has been studied in Wu and Zhang [63]. The necessary and sufficient conditions of near-optimality for singular control for jump diffusion processes have been investigated in Hafayed and Abbas [22]. Necessary and sufficient conditions for near-optimal mixed singular jump control have been proved in Hafayed and Abbas [22]. A good account on stochastic optimal control for jump diffusions and mixed singular stochastic control in Poisson jump problems with applications in finance can be found in Alvarez and Rakkolainen [3] and recently in ∅ksendal and Sulem [64, 65]. A combined singular stochastic control and optimal stopping in the jump-diffusion model was studied in An [4]. Some cases of mixed singular-jump control problems when the payoff functional does not depend explicitly on the control have been investigated in Menaldi and Rebin [44].

Our main goal in this work is to study a new class of mixed regular-singular optimal stochastic control of systems governed by McKean-Vlasov FBSDEJs, where the coefficients of the system and the performance functional depend not only on the state process but also its marginal law of the state process through its expected value. Necessary and sufficient conditions for the optimal regular-singular control are established for McKean-Vlasov FBSDEJs. The McKean-Vlasov mixed control problem under consideration is not simple extension from the mathematical point of view, but also provides an interesting models in many applications such as mathematical finance, where the singular components of the control means the interventions. The convexity of the control state space allows to use an argument of convex perturbation for both continuous and singular parts of our control process in order to deduce the stochastic maximum principle. In order to illustrate the study motivation and application background of this optimal intelligent control strategy, we present an example of a utility maximization problem.

Example 4.1.1. Suppose that we are given a mathematical market consisting of two investment possibilities:
(i) Bond: The first asset is a risk-free security whose price $S_0(t)$ evolves according to the ordinary differential equation

$$\mathrm{d}S_0\left(t\right)=S_0\left(t\right)\rho(t)\mathrm{d}t,\ t\in\left[0,T\right],\ S_0\left(0\right)>0, \tag{204}$$

where $\rho\left(\cdot\right):\left[0,T\right]\to\mathbb{R}_+$ is a locally bounded continuous deterministic function.
(ii) Stock: a risky security where the price $S_1\left(t\right)$ at time t is given by

$$\mathrm{d}S_1\left(t\right)=\varsigma(t)S_1\left(t\right)\mathrm{d}t+\sigma(t)S_1\left(t\right)\mathrm{d}W(t)+\int_{\Theta}A\left(t,\theta\right)N\left(\mathrm{d}\theta,\mathrm{d}t\right),\ S_1\left(0\right)>0, \tag{205}$$

In order to ensure that $S_1\left(t\right)>0$ for all $t\in\left[0,T\right]$ we assume
(1) The functions $\varsigma(\cdot):\left[0,T\right]\to\mathbb{R}$, $\sigma(\cdot):\left[0,T\right]\to\mathbb{R}$ are bounded continuous deterministic maps such that

$$\varsigma(t),\ \sigma(t)\neq 0 \text{ and } \varsigma(t)-\rho(t)>0,\ \forall t\in\left[0,T\right].$$

(2) $A\left(t,\theta\right)>-1$ for any $\theta\in\Theta$ and any $t\in\left[0,T\right]$.
(3) $\int_{\Theta}A^2\left(t,\theta\right)m(\mathrm{d}\theta)$ is bounded.

Let $x^{u,\xi}(0) = a > 0$ be an initial wealth and $G \geq 0$. By combining (204) and (205), we introduce the wealth dynamics

$$\begin{cases} \mathrm{d}x^{u,\xi}(t) = \left[\rho(t)x^{u,\xi}(t) + (\varsigma(t) - \rho(t))u(t)\right] \mathrm{d}t + \sigma(t)u(t)\mathrm{d}W(t) \\ \qquad\qquad + \int_{\Theta} A\,(t,\theta)\, u(t) N\,(\mathrm{d}\theta, \mathrm{d}t) - G\mathrm{d}\xi(t), \\ -\mathrm{d}y^{u,\xi}(t) = \left[\rho(t)x^{u,\xi}(t) + (\varsigma(t) - \rho(t))u(t) - \alpha y^{u,\xi}(t)\right] \mathrm{d}t - z^{u,\xi}(t)\mathrm{d}W(t) \\ \qquad\qquad - \int_{\Theta} r^{u,\xi}\,(t,\theta)\, N\,(\mathrm{d}\theta, \mathrm{d}t) + \beta\mathrm{d}\xi(t), \\ x^{u,\xi}(0) = a, \quad y^{u,\xi}(T) = x^{u,\xi}(T). \end{cases} \tag{206}$$

More precisely, for any admissible control $(u(\cdot), \xi(\cdot))$ the utility functional is given by

$$J\,(u(\cdot), \xi(\cdot)) = \frac{\delta}{2} Var(x^{u,\xi}(T)) - E(x^{u,\xi}(T)) + y^{u,\xi}(0) + E \int_{[0,T]} M(t)\mathrm{d}\xi(t). \tag{207}$$

By a simple computation, we shows that

$$J\,(u(\cdot), \xi(\cdot)) = E\left[\frac{\delta}{2} x^{u,\xi}(T)^2 - x^{u,\xi}(T)\right] - \frac{\delta}{2}\left[E(x^{u,\xi}(T))\right]^2 + y^{u,\xi}(0) + E \int_{[0,T]} M(t)\mathrm{d}\xi(t). \tag{208}$$

This is a time-inconsistent optimal control problem in the sense that it does not satisfy Bellman's optimality principle and therefore the usual dynamic programming approach fails.

The rest of this work is structured as follows. The assumptions, notations and some basic definitions are given in Section 2. Sections 3 and 4 are devoted to prove our main results. As an illustration, mean-variance portfolio selection problem: time-inconsistent solution is discussed in the last section.

19. Assumptions and statement of the mixed control problem

In this work, we study mixed stochastic optimal control problems of McKean-Vlasov type of the following kind. Let $T > 0$ be a fixed time horizon and $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a fixed filtered probability space equipped with a $\mathbb{P}-$complete right continuous filtration on which a $one-$dimensional Brownian motion $W = (W(t))_{t\in[0,T]}$ is defined. Let η be a homogeneous $\mathbb{F}_t$-Poisson point process independent of W. We denote by $\widetilde{N}(\mathrm{d}\theta, \mathrm{d}t)$ the random counting measure induced by η, defined on $\Theta \times \mathbb{R}_+$, where Θ is a fixed nonempty subset of $\mathbb{R}$ with its Borel $\sigma-$field $\mathcal{B}\,(\Theta)$. Further, let $m\,(\mathrm{d}\theta)$ be the local characteristic measure of η, i.e. $m\,(\mathrm{d}\theta)$ is a σ-finite measure on $(\Theta, \mathcal{B}\,(\Theta))$ with $m\,(\Theta) < +\infty$. We then define $N(\mathrm{d}\theta, \mathrm{d}t) := \widetilde{N}(\mathrm{d}\theta, \mathrm{d}t) - m\,(\mathrm{d}\theta)\,\mathrm{d}t$, where $N\,(\cdot, \cdot)$ is Poisson martingale measure on $\mathcal{B}\,(\Theta) \times \mathcal{B}\,(\mathbb{R}_+)$ with local characteristics $m\,(\mathrm{d}\theta)\,\mathrm{d}t$. We assume that $(\mathbb{F}_t)_{t\in[0,T]}$ is $\mathbb{P}-$augmentation of the natural filtration $(\mathbb{F}_t^{(W,N)})_{t\in[0,T]}$ defined as follows

$$\mathbb{F}_t^{(W,N)} := \mathbb{F}_t^W \vee \sigma\left\{\int_0^s \int_A N(\mathrm{d}\theta, \mathrm{d}r) : 0 \leq s \leq t,\ A \in \mathcal{B}\,(\Theta)\right\} \vee \mathbb{G}_0,$$

where $\mathbb{F}_t^W := \sigma\{W(s) : 0 \leq s \leq t\}$, $\mathbb{G}_0$ denotes the totality of $\mathbb{P}-$null sets, and $\mathbb{F}_1 \vee \mathbb{F}_2$ denotes the σ-field generated by $\mathbb{F}_1 \cup \mathbb{F}_2$.

We consider the following controlled nonlinear McKean-Vlasov coupled forward-backward stochastic differential equations which are governed both by Brownian motions and an independent Poisson random measure of the form:

$$\begin{cases} \mathrm{d}x^{u,\xi}(t) = f(t, x^{u,\xi}(t), \mu^{x,u,\xi}(t), u(t))\mathrm{d}t + \sigma(t, x^{u,\xi}(t), \mu^{x,u,\xi}(t), u(t))\mathrm{d}W(t) \\ \qquad + \mathcal{C}(t)\mathrm{d}\xi(t) + \int_\Theta \gamma(t, x^{u,\xi}(t_-), \mu^{x,u,\xi}(t_-), u(t), \theta) N(\mathrm{d}\theta, \mathrm{d}t), \ t \in [0,T], \\ \mathrm{d}y^{u,\xi}(t) = g(t, x^{u,\xi}(t), \mu^{x,u,\xi}(t), y^{u,\xi}(t), \mu^{y,u,\xi}(t), z^{u,\xi}(t), \mu^{z,u,\xi}(t), u(t))\mathrm{d}t \\ \qquad + z^{u,\xi}(t)\mathrm{d}W(t) + \mathcal{D}(t)\mathrm{d}\xi(t) + \int_\Theta r^{u,\xi}(t,\theta) N(\mathrm{d}\theta, \mathrm{d}t), \\ x^{u,\xi}(0) = a, \ y^{u,\xi}(T) = h(x^{u,\xi}(T), \mu^{x,u,\xi}(T)), \\ \mu^{x,u,\xi}(t) : \text{Probability distribution of } x^{u,\xi}(t), \\ \mu^{y,u,\xi}(t) : \text{Probability distribution of } y^{u,\xi}(t), \\ \mu^{z,u,\xi}(t) : \text{Probability distribution of } z^{u,\xi}(t), \end{cases} \tag{209}$$

where $f, \sigma, \gamma, g, h, \mathcal{C}$ and $\mathcal{D}$ are measurable given maps and the initial condition a is an $\mathbb{F}_0-$measurable random variable. The main new purpose here is the introduction of the combined singular control in McKean-Vlasov forward-backward stochastic system with random Poisson jumps. In particular this control might be discontinuous and it is necessary to distinguish between the jumps coming from the jump Poisson measure in the McKean-Vlasov forward-backward dynamics (209) and those from the interventions of controls. Noting that McKean-Vlasov FBSDEJs-(209) occur naturally in the probabilistic analysis of financial optimization problems and the optimal control of dynamics of the McKean-Vlasov type. Moreover, the above mathematical McKean-Vlasov approaches play an important role in different fields of economics, finance, physics, chemistry and game theory. For example, one may think of a biological, chemical or physical, interacting particle system in which each particle moves in the space according to the dynamics described by McKean-Vlasov FBSDEJs-(209) with $(\mu^{x,u,\xi}(t), \mu^{y,u,\xi}(t), \mu^{z,u,\xi}(t))$ being replaced by the empirical measure:

$$\mu_N^{x,u,\xi}(t) \triangleq \frac{1}{N}\sum_{j=1}^N \delta_{x_j^{u,\xi}(t)}, \ \mu_N^{y,u,\xi}(t) \triangleq \frac{1}{N}\sum_{j=1}^N \delta_{y_j^{u,\xi}(t)}, \text{ and } \mu_N^{z,u,\xi}(t) \triangleq \frac{1}{N}\sum_{j=1}^N \delta_{z_j^{u,\xi}(t)},$$

of $N-$particles $(x_1^{u,\xi}(t), y_1^{u,\xi}(t), z_1^{u,\xi}(t)), ..., (x_N^{u,\xi}(t), y_N^{u,\xi}(t), z_N^{u,\xi}(t))$ at time t. According to McKean-Vlasov theory, (see, McKean [45], Ahmed [1]), under proper conditions, the empirical measure-valued processes $(\mu_N^{x,u,\xi}(t), \mu_N^{y,u,\xi}(t), \mu_N^{z,u,\xi}(t))$ converges in probability as N goes to infinity to a deterministic measure-valued function $(\mu^{x,u,\xi}(t), \mu^{y,u,\xi}(t), \mu^{z,u,\xi}(t))$ which corresponds to the probability distribution of the processes determined by the McKean-Vlasov FBSDEJs-(209).

The cost functional on the time interval $[0,T]$ is defined by

$$\begin{aligned} & J(u(\cdot), \xi(\cdot)) \\ \triangleq \ & E\Big\{ \int_0^T \int_\Theta \ell\left(t, x^{u,\xi}(t), \mu^{x,u,\xi}(t), y^{u,\xi}(t), \mu^{y,u,\xi}(t), z^{u,\xi}(t), \mu^{z,u,\xi}(t), r^{u,\xi}(t,\theta), u(t)\right) m(\mathrm{d}\theta)\,\mathrm{d}t \\ & + \phi(x^{u,\xi}(T), \mu^{x,u,\xi}(T)) + \varphi(y^{u,\xi}(0), \mu^{y,u,\xi}(0)) + \int_{[0,T]} M(t)\mathrm{d}\xi(t) \Big\}, \end{aligned} \tag{210}$$

where ℓ, ϕ, φ and M are an appropriate functions. This cost functional is also of McKean-Vlasov type, as the functions ℓ, ϕ, φ depend on the marginal law of the state process through its expected value. It is worth mentioning that since the cost functional J is possibly a nonlinear function of the expected value stands in contrast to the standard formulation of a control problem. This leads to so called time-inconsistent control problem where the Bellman dynamic programming does not hold. The reason for this is that one cannot apply the law of iterated expectations on the cost functional. Noting that in most cases, the classical singular control problem (without McKean-Vlasov terms) was studied through dynamic programming principle. This is a type of a control problem which, it seems, has not been studied before.

An admissible control $(u^*(\cdot), \xi^*(\cdot))$ is called optimal if it satisfies

$$J\left(u^*(\cdot), \xi^*(\cdot)\right) \triangleq \inf_{(u(\cdot),\xi(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2([0,T])} J\left(u(\cdot), \xi(\cdot)\right). \tag{211}$$

The corresponding state processes, solution of Eq-(209), is denoted by $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot,\cdot)) = (x^{u^*,\xi^*}(\cdot), y^{u^*,\xi^*}(\cdot), z^{u^*,\xi^*}(\cdot), r^{u^*,\xi^*}(\cdot,\cdot))$.

For convenience, we will use the following notation in this work. $\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$ denotes the Hilbert space of $\mathbb{F}_t-$adapted processes such that $E\left(\int_0^T |x(t)|^2 \mathrm{d}t\right) < +\infty$ and $\mathbb{M}^2_{\mathbb{F}}([0,T];\mathbb{R})$ denotes the Hilbert space of $\mathbb{F}_t-$ predictable processes $(\psi(t,\theta))_{t\in[0,T]}$ defined on $[0,T]\times\Theta$ such that $E\int_0^T\int_\Theta |\psi(t,\theta)|^2 m(\theta)\mathrm{d}t < +\infty$. For a differentiable function Φ we denote by $\nabla_x\Phi(t)$ its gradient with respect to the variable x. To simplify our notation, we suppress "w" in $f(t,w,x,\mu^{x,u,\xi},u)$ and write $f(t,x,\mu^{x,u,\xi},u)$ for $f(t,w,x,\mu^{x,u,\xi},u)$ *etc*. Since the purpose of this work is to study optimal combined stochastic control for McKean-Vlasov systems, we give here the precise definition of the singular part of an admissible control.

Definition 4.2.1. *An admissible control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{U}_1\times\mathbb{U}_2-$valued, $\mathbb{F}^W_t-$adapted processes, such that:*
1. $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the left with right limits and $\xi(0_-)=0$.
2. $E\left[\sup_{t\in[0,T]}|u(t)|^2 + |\xi(T)|^2\right] < \infty$.
Notice that the jumps of a singular control $\xi(\cdot)$ at any jumping time t_j denote by $\mathbf{\Delta}\xi(t_j) \triangleq \xi(t_j) - \xi(t_{j-})$ and we define the continuous part of the singular control by

$$\xi^{(c)}(t) \triangleq \xi(t) - \sum_{0\leq t_j\leq t}\mathbf{\Delta}\xi(t_j),$$

i.e., the process obtained by removing the jumps of $\xi(t)$.
We denote $\mathcal{U}_1\times\mathcal{U}_2([0,T])$ the set of all admissible controls. Since $\mathrm{d}\xi(t)$ may be singular with respect to Lebesgue measure $\mathrm{d}t$, we call $\xi(\cdot)$ the singular part of the control and the process $u(\cdot)$ its absolutely continuous part.
Throughout this work, we distinguish between the jumps caused by the singular control $\xi(\cdot)$ and the jumps caused by the random Poisson measure at any jumping time t.

Definition 4.2.2. *We define the jumps of $x^{u,\xi}(t)$ and $y^{u,\xi}(t)$ caused by the singular control $\xi(\cdot)$* by

$$\begin{aligned}\mathbf{\Delta}_\xi x^{u,\xi}(t) &\triangleq \mathcal{C}(t)\mathbf{\Delta}\xi(t) = \mathcal{C}(t)(\xi(t)-\xi(t_-)),\\ \mathbf{\Delta}_\xi y^{u,\xi}(t) &\triangleq \mathcal{D}(t)\mathbf{\Delta}\xi(t) = \mathcal{D}(t)(\xi(t)-\xi(t_-)),\end{aligned}$$

and we define the jumps of $x^{u,\xi}(t)$ *and* $y^{u,\xi}(t)$ *caused by the Poisson measure of* $\widetilde{N}(\theta,t)$ *by*

$$\begin{aligned}
\mathbf{\Delta}_{N} x^{u,\xi}(t) &\triangleq \int_{\Theta} \gamma\left(t, x^{u,\xi}(t_{-}), \mu^{x,u,\xi}(t_{-}), u(t_{-}), \theta\right) \widetilde{N}\left(\mathrm{d}\theta, \{t\}\right) \\
&\triangleq \begin{cases} \gamma\left(t, x^{u,\xi}(t_{-}), \mu^{x,u,\xi}(t_{-}), u(t_{-}), \theta\right) : \text{if } \eta \text{ has a jump of size } \theta \text{ at } t. \\ 0 : \text{otherwise} \end{cases} \\
\mathbf{\Delta}_{N} y^{u,\xi}(t) &\triangleq \int_{\Theta} r^{u,\xi}\left(t,\theta\right) \widetilde{N}\left(\mathrm{d}\theta, \{t\}\right), \\
&\triangleq \begin{cases} r^{u,\xi}\left(t,\theta\right) : \text{if } \eta \text{ has a jump of size } \theta \text{ at } t. \\ 0 : \text{otherwise,} \end{cases}
\end{aligned}$$

where $\widetilde{N}\left(\mathrm{d}\theta, \{t\}\right)$ *means the jump in the Poisson random measure, occurring at time* t

Definition 4.2.3. *The general jump of the state processes* $x^{u,\xi}(\cdot), y^{u,\xi}(\cdot)$ *at any jumping time* t *is given by* (see Figure 1.)

$$\begin{aligned}
\mathbf{\Delta} x^{u,\xi}(t) \triangleq x^{u,\xi}(t) - x^{u,\xi}(t_{-}) = \mathbf{\Delta}_{\xi} x^{u,\xi}(t) + \mathbf{\Delta}_{N} x^{u,\xi}(t). \\
\mathbf{\Delta} y^{u,\xi}(t) \triangleq y^{u,\xi}(t) - y^{u,\xi}(t_{-}) = \mathbf{\Delta}_{\xi} y^{u,\xi}(t) + \mathbf{\Delta}_{N} y^{u,\xi}(t).
\end{aligned}$$

In this work, we also assume that the coefficients $f, \sigma, g, \ell, \gamma, h, \varphi, \phi, \mathcal{C}, \mathcal{D}$ and M satisfy the following standing assumptions:

Condition (H1) The functions $f, \sigma, \ell, \gamma, g, h, \phi, \varphi$ are continuously differentiable in their variables including $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$.

Condition (H2) *(i)* The derivatives of $f, \sigma, g, \phi, \gamma$ with respect to their variables including $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$ are bounded, and $\int_{\Theta}(\left|\nabla_{x}\gamma\left(t,x,\widetilde{x},u,\theta\right)\right|^{2} + \left|\nabla_{\widetilde{x}}\gamma\left(t,x,\widetilde{x},u,\theta\right)\right|^{2} + \left|\nabla_{u}\gamma\left(t,x,\widetilde{x},u,\theta\right)\right|^{2})m\left(\mathrm{d}\theta\right) < +\infty$.

(ii) The derivatives of $f, \sigma, g, \gamma, \ell$ with respect to $(x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r, u)$ are dominated by $C(1 + |x| + |\widetilde{x}| + |y|$
$+ |\widetilde{y}| + |z| + |\widetilde{y}| + |r| + |u|)$. Moreover, $\varphi_{y}, \varphi_{\widetilde{y}}$ are bounded by $C\left(1 + |y| + |\widetilde{y}|\right)$ and $h_{x}, h_{\widetilde{x}}$ are bounded by $C\left(1 + |x| + |\widetilde{x}|\right)$.

(iii) For all $t \in [0,T]$, $f(t,0,0,0), g(t,0,0,0,0,0,0,0,0) \in \mathbb{L}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\right)$, $\sigma(t,0,0,0) \in \mathbb{L}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\times\mathbb{R}\right)$, and $\gamma(t,0,0,0,\cdot) \in \mathbb{M}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\right)$.

Conditions (H3) The functions $\mathcal{C} : [0,T] \to \mathbb{R}$, $\mathcal{D} : [0,T] \to \mathbb{R}$ and $M : [0,T] \to \mathbb{R}^{+}$ are continuous and bounded.

Under conditions (H1)∼(H3), the FBSDEJ-(209) has an unique solution $\left(x^{u,\xi}\left(\cdot\right), y^{u,\xi}\left(\cdot\right), z^{u,\xi}\left(\cdot\right), r^{u,\xi}(\cdot,\cdot)\right)$
$\in \mathbb{L}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\right) \times \mathbb{L}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\right) \times \mathbb{L}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\right) \times \mathbb{L}^{2}_{\mathbb{F}}\left([0,T];\mathbb{R}\right)$ such that

$$\begin{aligned}
x^{u,\xi}(t) = a &+ \int_{0}^{t} f\left(s, x^{u,\xi}(s), \mu^{x,u,\xi}(s), u(s)\right) \mathrm{d}s + \int_{0}^{t} \sigma(s, x^{u,\xi}(s), \mu^{x,u,\xi}(s)), u(s)) \mathrm{d}W(s) \\
&+ \int_{0}^{t}\int_{\Theta} \gamma\left(s, x^{u,\xi}(s_{-}), \mu^{x,u,\xi}(s_{-}), u(s), \theta\right) N\left(\mathrm{d}\theta, \mathrm{d}s\right) + \int_{[0,t]} \mathcal{C}(s)\mathrm{d}\xi(s),
\end{aligned}$$

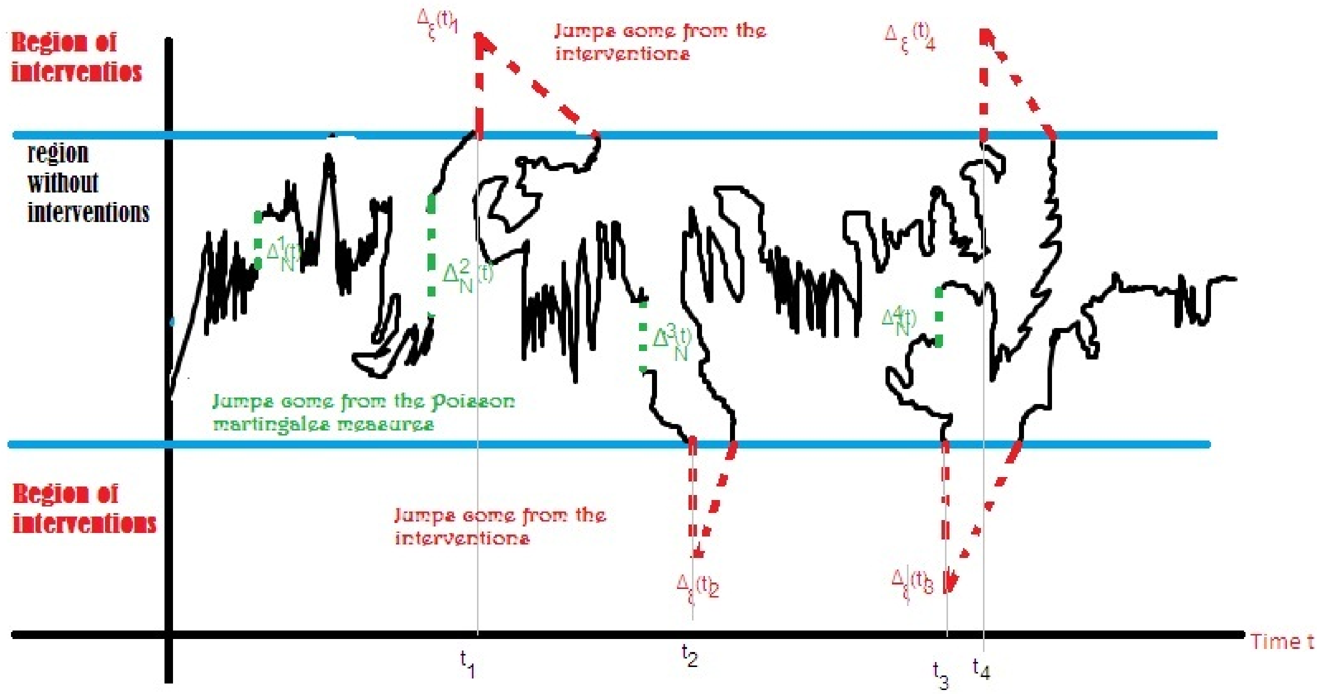

Figure 1. Region and type of Jumps

and for $t \in [0,T]$

$$y^{u,\xi}(t) = y^{u,\xi}(T) - \int_t^T \int_\Theta g(s, x^{u,\xi}(s), \mu^{x,u,\xi}(s), y^{u,\xi}(s), \mu^{y,u,\xi}(s), z^{u,\xi}(s)$$

$$, \mu^{z,u,\xi}(s), r^{u,\xi}(t,\theta), u(s)) m(\mathrm{d}\theta)\,\mathrm{d}s + \int_t^T z^{u,\xi}(s)\mathrm{d}W(s) + \int_t^T \int_\Theta r^{u,\xi}(s,\theta) N(\mathrm{d}\theta, \mathrm{d}s)$$

$$+ \int_{[t,T]} \mathcal{D}(s)\mathrm{d}\xi(s).$$

Since the coefficients $\mathcal{C}$ and $\mathcal{D}$ are continuous and bounded, the existence and uniqueness can be proved similar to ([51], Lemma 3.1 and Theorem 3.1).

Adjoint equations. We introduce the new adjoint equations involved in the stochastic maximum principle for our mixed singular-jump McKean-Vlasov control problem (209)-(210). For simplicity of notation, we will still use $\nabla_x f(t) \triangleq \dfrac{\partial f}{\partial x}(t, x^{u,\xi}(\cdot), \mu^{x,u,\xi}(\cdot), u(\cdot))$, and $\nabla_x g(t,\theta) \triangleq \dfrac{\partial g}{\partial x}(t, x(t), \mu^{x,u,\xi}(t), y(t), \mu^{y,u,\xi}(t), z(t), \mu^{z,u,\xi}(\cdot), \; r(t,\theta), u(t))$ *etc.* So for any admissible control $(u(\cdot),\xi(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2([0,T])$ and the corresponding state trajectory $(x^{u,\xi}(\cdot), y^{u,\xi}(\cdot), z^{u,\xi}(\cdot), r^{u,\xi}(\cdot,\cdot)) \triangleq (x(\cdot), y(\cdot), z(\cdot), r(\cdot,\cdot))$, we consider the following adjoint equations of McKean-Vlasov type, which are independent to singular control. In what follows, we simply write $\mu^{x,u,\xi}(t) = E(x^{u,\xi}(t))$ *etc.*

$$
\begin{cases}
\mathrm{d}\Phi^u(t) = -\{\nabla_x f(t)\,\Phi^u(t) + E(\nabla_{\widetilde{x}} f(t)\,\Phi^u(t)) + \nabla_x\sigma(t)\,Q^u(t) + E[\nabla_{\widetilde{x}}\sigma(t)\,Q^u(t)] \\
\qquad + \int_\Theta [\nabla_x g(t,\theta)K^u(t) + E(\nabla_{\widetilde{x}} g(t,\theta)K^u(t)) + \nabla_x\gamma(t,\theta)\,R^u(t,\theta) \\
\qquad + E(\nabla_{\widetilde{x}}\gamma(t,\theta)R^u(t,\theta)) + \nabla_x\ell(t,\theta) + E(\nabla_{\widetilde{x}}\ell(t,\theta))]\,m(\mathrm{d}\theta)\}\mathrm{d}t \\
\qquad + Q^u(t)\mathrm{d}W(t) + \int_\Theta R^u(\theta,t)\,N(\mathrm{d}\theta,\mathrm{d}t), \\
\Phi^u(T) = -\{\nabla_x h(T)\,K^u(T) + E[(\nabla_{\widetilde{x}} h(T))\,K^u(T)]\} + \nabla_x\phi(T) + E(\nabla_{\widetilde{x}}\phi(T)). \\
-\mathrm{d}K^u(t) = \int_\Theta [\nabla_y g(t,\theta)\,K^u(t) + E(\nabla_{\widetilde{y}} g(t,\theta)\,K^u(t)) + \nabla_y\ell(t,\theta) + E(\nabla_{\widetilde{y}}\ell(t,\theta))]m(\mathrm{d}\theta)\,\mathrm{d}t \\
\qquad + \int_\Theta [\nabla_z g(t,\theta)K^u(t) + E(\nabla_{\widetilde{z}} g(t,\theta)K^u(t)) + \nabla_z\ell(t,\theta) + E(\nabla_{\widetilde{z}}\ell(t,\theta))]m(\mathrm{d}\theta)\,\mathrm{d}W(t) \\
\qquad - \int_\Theta [\nabla_r g(t,\theta)K^u(t) + \nabla_r\ell(t,\theta)]\,N(\mathrm{d}\theta,\mathrm{d}t) \\
K^u(0) = -\{\nabla_y\varphi(y(0), E(y(0))) + E[\nabla_{\widetilde{y}}\varphi(y(0), E(y(0)))]\}.
\end{cases}
\tag{212}
$$

Note that the first adjoint equation (backward) corresponding to the forward component turns out to be a linear McKean-Vlasov backward SDE with jumps, and the second adjoint equation (forward) corresponding to the backward component turns out to be a linear McKean-Vlasov forward SDE with jumps.
We define the Hamiltonian function $H : [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{U}_1\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, associated with the singular stochastic control problem (209)-(210) as follows

$$
\begin{aligned}
&H(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},u,r(\cdot),\Phi(\cdot),Q(\cdot),K(\cdot),R(\cdot,\cdot)) = -\Phi^u(t)f(t,x,\widetilde{x},u) - Q^u(t)\sigma(t,x,\widetilde{x},u) \\
&- \int_\Theta [K^u(t)g(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r(\cdot),u) + R^u(t,\theta)\,\gamma(t,x,\widetilde{x},u,\theta)) + \ell(t,x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r(\cdot),u)]\,m(\mathrm{d}\theta).
\end{aligned}
\tag{213}
$$

If we denote by $H(t) := H(t,x(t),\widetilde{x}(t),y(t),\widetilde{y}(t),z(t),\widetilde{z}(t),r(t,\cdot),u(t),\Phi(t),Q(t),K(t),R(t,\cdot))$, then the adjoint equation (212) can be rewritten as the following stochastic Hamiltonian system:

$$
\begin{cases}
\mathrm{d}\Phi^u(t) = \{H_x(t) + E[H_{\widetilde{x}}(t)]\}\,\mathrm{d}t + Q(t)\mathrm{d}W(t) + \int_\Theta R(t,\theta)\,N(\mathrm{d}\theta,\mathrm{d}t), \\
\Phi^u(T) = -\{h_x(T)\,K^u(T) + E[(h_{\widetilde{x}}(T))\,K^u(T)]\} + \phi_x(T) + E(\phi_{\widetilde{x}}(T)). \\
-\mathrm{d}K^u(t) = [H_y(t) + E(H_{\widetilde{y}}(t))]\,\mathrm{d}t + [H_z(t) + E(H_{\widetilde{z}}(t))]\,\mathrm{d}W(t) \\
\qquad - \int_\Theta [g_r(t,\theta)K^u(t) + \ell_r(t,\theta)]\,N(\mathrm{d}\theta,\mathrm{d}t) \\
K^u(0) = -\{\varphi_y(y(0), E(y(0))) + E[\varphi_{\widetilde{y}}(y(0), E(y(0)))]\}.
\end{cases}
\tag{214}
$$

It is a well known fact that under assumptions (H1) and (H2), the adjoint equations (212) or (214) admits a unique solution $(\Phi^u(t), Q^u(t), K^u(t), R^u(t,\cdot)) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{M}^2_{\mathbb{F}}([0,T];\mathbb{R})$. Moreover, since the derivatives of $f,\sigma,\gamma,g,h,\varphi,\phi$ with respect to $(x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r)$ are bounded, we deduce from standard arguments that there exists a constant $C>0$

such that

$$E\left[\sup_{t\in[0,T]}\left|\Phi^{u}(t)\right|^{2}+\sup_{t\in[0,T]}\left|K^{u}(t)\right|^{2}+\int_{0}^{T}\left|Q^{u}(t)\right|^{2}\mathrm{d}t+\int_{0}^{T}\int_{\Theta}\left|R^{u}\left(t,\theta\right)\right|^{2}m\left(\mathrm{d}\theta\right)\mathrm{d}t\right]<C. \quad (215)$$

20. Necessary conditions for optimal mixed continuous-singular control of McKean-Vlasov FBSDEJs

In this section, we establish a set of necessary conditions for a stochastic singular control to be optimal where the system evolves according to controlled McKean-Vlasov FBSDEJs. Convex perturbation techniques for both continuous and singular parts are applied to derive our McKean-Vlasov stochastic maximum principle.

The following theorem constitutes the main contribution of this work.

Let $(x^*(\cdot),y^*(\cdot),z^*(\cdot),r^*(\cdot,\cdot))$ be the solution of the McKean-Vlasov FBSDEJs-(209) and $(\Phi^*(\cdot),Q^*(\cdot),K^*(\cdot),R^*(\cdot,\cdot))$ be the solution of adjoint equation (212) corresponding to the optimal singular control $(u^*(\cdot),\xi^*(\cdot))$.

Theorem 4.3.1. (Necessary condition for optimal mixed control in Integral form). *Let Conditions (H1), (H2) and (H3) hold. If* $(u^*(\cdot),\xi^*(\cdot))$ and $(x^*(\cdot),y^*(\cdot),z^*(\cdot),r^*(\cdot,\cdot))$ *is an optimal solution of the McKean-Vlasov singular control problem (209)-(210). Then the maximum principle holds, that is for all* $(u,\xi)\in\mathbb{U}_1\times\mathbb{U}_2$:

$$\begin{gathered}0\leq E\int_0^T\nabla_u H(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*,\Lambda^*(t,\theta))(u-u^*(t))\mathrm{d}t\\ +E\int_{[0,T]}(M(t)+\mathcal{C}(t)\Phi^*(t)+\mathcal{D}(t)K^*(t))\mathrm{d}\left(\xi-\xi^*\right)(t),\\ a.e.,\ t\in[0,T]\,,\end{gathered} \tag{216}$$

where $(\lambda^*(t,\theta),E(\lambda^*(t,\theta)))\triangleq(x^*(t),E(x^*(t)),y^*(t),E(y^*(t)),z^*(t),E(z^*(t)),r^*(t,\theta))$ *and* $\Lambda^*(t,\theta)\triangleq(\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta))$.

To prove *Theorem 3.1* we need some preliminary results given in the following Lemmas.

We derive the variational inequality (216) in several steps, from the fact that

$$J\left(u^*(\cdot),\xi^*(\cdot)\right)\leq J\left(u^{\varepsilon}(\cdot),\xi^{\varepsilon}(\cdot)\right), \tag{217}$$

where $(u^{\varepsilon}(\cdot),\xi^{\varepsilon}(\cdot))$ is the so called convex perturbation of optimal control $(u^*(\cdot),\xi^*(\cdot))$ defined as follows

$$u^{\varepsilon}(t)=u^*(t)+\varepsilon\left(u(t)-u^*(t)\right)\text{ and }\xi^{\varepsilon}(t)=\xi^*(t)+\varepsilon\left(\xi(t)-\xi^*(t)\right),$$

where $\varepsilon\in[0,1]$ is sufficiently small, $(u(\cdot),\xi(\cdot))$ is an arbitrary element of $\mathbb{F}_t-$measurable random variable with values in $\mathbb{U}_1\times\mathbb{U}_2$ which we consider as fixed from now on.

We emphasize that the convexity of $\mathbb{U}_1\times\mathbb{U}_2$ has the consequence that the perturbed control $(u^{\varepsilon}(\cdot),\xi^{\varepsilon}(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2\left([0,T]\right)$ where

$$(u^{\varepsilon}(t),\xi^{\varepsilon}(t))=(u^*(t),\xi^*(t))+\varepsilon\left[(u(t),\xi(t))-(u^*(t),\xi^*(t))\right].$$

Variational equations. Now, we introduce the following new variational equations which have a McKean-Vlasov type. Let $(x_1^\varepsilon(\cdot), y_1^\varepsilon(\cdot), z_1^\varepsilon(\cdot), r_1^\varepsilon(\cdot,\cdot))$ be the solution of the following forward-backward stochastic system described by Brownian motions and Poisson jumps of McKean-Vlasov type

$$\begin{cases} \mathrm{d}x_1^\varepsilon(t) = \{f_x(t)x_1^\varepsilon(t) + f_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + f_u(t)u(t))\}\,\mathrm{d}t \\ \qquad + \{\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + \sigma_u(t)u(t)\}\,\mathrm{d}W(t) \\ \qquad + \int_\Theta [\gamma_x(t,\theta)x_1^\varepsilon(t) + \gamma_{\widetilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right) + \gamma_u(t,\theta)u(t)]N(\mathrm{d}\theta,\mathrm{d}t) \\ \qquad + \mathcal{C}(t)\mathrm{d}\xi(t), \quad x_1^\varepsilon(0) = 0, \\ \mathrm{d}y_1^\varepsilon(t) = \int_\Theta \{g_x(t,\theta)x_1^\varepsilon(t) + g_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t)) + g_y(t,\theta)y_1^\varepsilon(t) + g_{\widetilde{y}}(t,\theta)E(y_1^\varepsilon(t)) \\ \qquad + g_z(t,\theta)z_1^\varepsilon(t) + g_{\widetilde{z}}(t,\theta)E(z_1^\varepsilon(t)) + g_r(t,\theta)r_1^\varepsilon(t,\theta) + g_u(t,\theta)u(t)\}\,m(\mathrm{d}\theta)\mathrm{d}t \\ \qquad + z_1^\varepsilon(t)\mathrm{d}W(t) - \int_\Theta r_1^\varepsilon(t,\theta)N(\mathrm{d}\theta,\mathrm{d}t) + \mathcal{D}(t)\mathrm{d}\xi(t) \\ y_1^\varepsilon(T) = [h_x(T) + E\left(h_{\widetilde{x}}(T)\right))]\,x_1^\varepsilon(T). \end{cases} \tag{218}$$

Duality relations. Our first Lemma below deals with the duality relations between $\Phi^*(t)$, $x_1^\varepsilon(t)$ and $K^*(t)$, $y_1^\varepsilon(t)$. This Lemma is very important for the proof of our main result.

Lemma 4.3.1. We have

$$\begin{aligned} &E\left(\Phi^*(T)x_1^\varepsilon(T)\right) \\ &= E\int_0^T \left[\Phi^*(t)f_u(t)u(t) + Q^*(t)\sigma_u(t)u(t) + \int_\Theta R^*(t,\theta)\gamma_u(t,\theta)u(t)m(\mathrm{d}\theta)\right]\mathrm{d}t \\ &\quad - E\int_0^T\int_\Theta \{x_1^\varepsilon(t)g_x(t,\theta)K^*(t) + x_1^\varepsilon(t)E(g_{\widetilde{x}}(t,\theta)K^*(t)) + x_1^\varepsilon(t)\ell_x(t,\theta) \\ &\quad + x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta))\}\,m(\mathrm{d}\theta)\mathrm{d}t + E\int_{[0,T]}\Phi^*(t)\mathcal{C}(t)\mathrm{d}\xi(t), \end{aligned} \tag{219}$$

similarly, we get

$$\begin{aligned} &E\left(K^*(T)y_1^\varepsilon(T)\right) = -E\left\{[\varphi_y(0) + E\left(\varphi_{\widetilde{y}}(0)\right)]\,y_1^\varepsilon(0)\right\} \\ &+E\int_0^T\int_\Theta \{K^*(t)g_x(t,\theta)x_1^\varepsilon(t) + K^*(t)g_{\widetilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right) \\ &+K^*(t)g_u(t,\theta)u(t)) - y_1^\varepsilon(t)\ell_y(t,\theta) - y_1^\varepsilon(t)E(\ell_{\widetilde{y}}(t,\theta)) \\ &-z_1^\varepsilon(t)\ell_z(t,\theta) - z_1^\varepsilon(t)E(\ell_{\widetilde{z}}(t,\theta)) \\ &-r_1^\varepsilon(t,\theta)\ell_r(t,\theta)\}\,m(\mathrm{d}\theta)\mathrm{d}t + E\int_{[0,T]}K^*(t)\mathcal{D}(t)\mathrm{d}\xi(t), \end{aligned} \tag{220}$$

and

$$\begin{aligned} &E\left\{[\phi_x(T) + E(\phi_{\widetilde{x}}(T))]\,x_1^\varepsilon(T)\right\} + E\left\{[\varphi_y(0) + E\left(\varphi_{\widetilde{y}}(0)\right)]\,y_1^\varepsilon(0)\right\} \\ &= -E\int_0^T\int_\Theta \{x_1^\varepsilon(t)\ell_x(t,\theta) + x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta)) + y_1^\varepsilon(t)\ell_y(t,\theta) + y_1^\varepsilon(t)E(\ell_{\widetilde{y}}(t,\theta)) \\ &\quad + z_1^\varepsilon(t)\ell_z(t,\theta) + z_1^\varepsilon(t)E(\ell_{\widetilde{z}}(t,\theta)) + r_1^\varepsilon(t,\theta)\ell_r(t,\theta) + \ell_u(t,\theta)u(t)\}\,m(\mathrm{d}\theta)\mathrm{d}t \\ &\quad + E\int_0^T H_u(t)u(t)\mathrm{d}t + E\int_{[0,T]}\left[\Phi^*(t)\mathcal{C}(t) + K^*(t)\mathcal{D}(t)\right]\mathrm{d}\xi(t). \end{aligned} \tag{221}$$

Proof.

Proof of duality relation (219). By applying integration by parts formula for Poisson jump processes to $\Phi^*(t)x_1^\varepsilon(t)$, we get

$$\begin{aligned}
&E\left(\Phi^*(T)x_1^\varepsilon(T)\right) \\
&= E\int_0^T \Phi^*(t)\mathrm{d}x_1^\varepsilon(t) + E\int_0^T x_1^\varepsilon(t)\mathrm{d}\Phi^*(t) \\
&\quad + E\int_0^T Q^*(t)\left[\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + \sigma_u(t)u(t)\right]\mathrm{d}t \\
&\quad + E\int_0^T \int_\Theta \left[\gamma_x(t,\theta)x_1^\varepsilon(t) + \gamma_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t)) + \gamma_u(t,\theta)u(t)\right]R(t,\theta)m(\mathrm{d}\theta)\mathrm{d}t \\
&\quad = I_1(\varepsilon) + I_2(\varepsilon) + I_3(\varepsilon) + I_4(\varepsilon).
\end{aligned} \tag{222}$$

A simple computation shows that

$$\begin{aligned}
&I_1(\varepsilon) = E\int_0^T \Phi^*(t)\mathrm{d}x_1^\varepsilon(t) = E\int_0^T \left\{\Phi^*(t)f_x(t)x_1^\varepsilon(t) + \Phi^*(t)f_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right)\right. \\
&\left. + \Phi^*(t)f_u\left(t\right)u(t)\right\}\mathrm{d}t + E\int_{[0,T]} \Phi^*(t)\mathcal{C}(t)\mathrm{d}\xi(t),
\end{aligned} \tag{223}$$

and

$$\begin{aligned}
&I_2(\varepsilon) = E\int_0^T x_1^\varepsilon(t)\mathrm{d}\Phi^*(t) \\
&\quad = -E\int_0^T \left\{x_1^\varepsilon(t)f_x\left(t\right)\Phi^*(t) + x_1^\varepsilon(t)E\left(f_{\widetilde{x}}(t)\Phi^*(t)\right)\right. \\
&+ x_1^\varepsilon(t)\sigma_x\left(t\right)Q^*(t) + x_1^\varepsilon(t)E\left(\sigma_{\widetilde{x}}(t)Q^*(t)\right) \\
&\quad + \int_\Theta \left[x_1^\varepsilon(t)g_x\left(t,\theta\right)K^*(t) + x_1^\varepsilon(t)E\left(g_{\widetilde{x}}(t,\theta)K^*(t)\right)\right. \\
&+ x_1^\varepsilon(t)\gamma_x\left(t,\theta\right)R\left(t,\theta\right) + x_1^\varepsilon(t)E(\gamma_{\widetilde{x}}\left(t,\theta\right)R\left(t,\theta\right)) \\
&\quad \left.\left. + x_1^\varepsilon(t)\ell_x(t,\theta) + x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta))m\left(\mathrm{d}\theta\right)\right]\right\}\mathrm{d}t.
\end{aligned} \tag{224}$$

By standard arguments, we get

$$I_3(\varepsilon) = E\int_0^T Q^*(t)\sigma_x(t)x_1^\varepsilon(t)\mathrm{d}t + E\int_0^T Q^*(t)\sigma_{\widetilde{x}}(t)E\left(x_1^\varepsilon(t)\right)\mathrm{d}t + E\int_0^T Q^*(t)\sigma_u(t)u(t)\mathrm{d}t, \tag{225}$$

and

$$\begin{aligned}
&I_4(\varepsilon) = E\int_0^T \int_\Theta \gamma_x(t,\theta)x_1^\varepsilon(t)R(t,\theta)m(\mathrm{d}\theta)\mathrm{d}t + E\int_0^T \int_\Theta \gamma_{\widetilde{x}}(t,\theta)E(x_1^\varepsilon(t))R(t,\theta)m(\mathrm{d}\theta)\mathrm{d}t \\
&\quad + E\int_0^T \int_\Theta \gamma_u(t,\theta)u(t)R(t,\theta)m(\mathrm{d}\theta)\mathrm{d}t.
\end{aligned} \tag{226}$$

The duality relation (219) follows immediately from combining (222)∼(226).

Proof of duality relation (220). Let us turn to second duality relation (220). By applying integration by parts formula for Poisson jump process to $K^*(t)y_1^\varepsilon(t)$, we get

$$\begin{aligned}
&E\left(K^*(T)y_1^\varepsilon(T)\right)\\
&= E\left(K^*(0)y_1^\varepsilon(0)\right) + E\int_0^T K^*(t)\mathrm{d}y_1^\varepsilon(t) + E\int_0^T y_1^\varepsilon(t)\mathrm{d}K^*(t)\\
&- E\int_0^T\int_\Theta z_1^\varepsilon(t)[g_z(t,\theta)K^*(t) + E\left(g_{\widetilde{z}}(t,\theta)K^*(t)\right)\\
&+ \ell_z(t,\theta) + E\left(\ell_{\widetilde{z}}\left(t,\theta\right)\right)]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\
&- E\int_0^T\int_\Theta \left[r_1^\varepsilon(t,\theta)(g_r(t,\theta)K^*(t) + \ell_r(t,\theta))\right]m(\mathrm{d}\theta)\mathrm{d}t.\\
&= \tilde{I}_1(\varepsilon) + \tilde{I}_2(\varepsilon) + \tilde{I}_3(\varepsilon) + \tilde{I}_4(\varepsilon) + \tilde{I}_5(\varepsilon).
\end{aligned} \tag{227}$$

From (219) we obtain

$$\begin{aligned}
&\tilde{I}_2(\varepsilon) = E\int_0^T K^*(t)\mathrm{d}y_1^\varepsilon(t)\\
&= E\int_0^T\int_\Theta \{K^*(t)g_x(t,\theta)x_1^\varepsilon(t) + K^*(t)g_{\widetilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right)\\
&+ K^*(t)g_y(t,\theta)y_1^\varepsilon(t)\\
&+ K^*(t)g_{\widetilde{y}}(t,\theta)E\left(y_1^\varepsilon(t)\right) + K^*(t)g_z(t,\theta)z_1^\varepsilon(t) + K^*(t)g_{\widetilde{z}}(t,\theta)E\left(z_1^\varepsilon(t)\right)\\
&+ K^*(t)g_r(t,\theta)r_1^\varepsilon(t,\theta) + K^*(t)g_u(t,\theta)u(t)\}\, m(\mathrm{d}\theta)\mathrm{d}t\\
&+E\int_{[0,T]} K^*(t)\mathcal{D}(t)\mathrm{d}\xi(t),
\end{aligned} \tag{228}$$

from (212) we obtain

$$\begin{aligned}
&\tilde{I}_3(\varepsilon) = E\int_0^T y_1^\varepsilon(t)\mathrm{d}K^*(t)\\
&= -E\int_0^T\int_\Theta y_1^\varepsilon(t)g_y\left(t,\theta\right)K^*(t) + y_1^\varepsilon(t)E\left(g_{\widetilde{y}}(t,\theta)K^*(t)\right)\\
&+ y_1^\varepsilon(t)\ell_y(t,\theta) + y_1^\varepsilon(t)E\left(\ell_{\widetilde{y}}\left(t,\theta\right)\right)\}\, m(\mathrm{d}\theta)\mathrm{d}t,
\end{aligned} \tag{229}$$

and

$$\begin{aligned}
\tilde{I}_4(\varepsilon) &= -E\int_0^T\int_\Theta \left[z_1^\varepsilon(t)g_z(t,\theta)K^*(t) + z_1^\varepsilon(t)E\left(g_{\widetilde{z}}(t,\theta)K^*(t)\right)\right.\\
&\left.+ z_1^\varepsilon(t)\ \ell_z(t,\theta) + z_1^\varepsilon(t)E\left(\ell_{\widetilde{z}}\left(t,\theta\right)\right)\right]m(\mathrm{d}\theta)\mathrm{d}t\\
\tilde{I}_5(\varepsilon) &= -E\int_0^T\int_\Theta [r_1^\varepsilon(t,\theta)g_r(t,\theta)K^*(t) + r_1^\varepsilon(t,\theta)\ell_r(t,\theta)]m(\mathrm{d}\theta)\mathrm{d}t.
\end{aligned} \tag{230}$$

Since $\tilde{I}_1(\varepsilon) = E\left(K^*(0)y_1^\varepsilon(0)\right) = -E\left\{\left[\varphi_y\left(0\right) + E\left(\varphi_{\widetilde{y}}\left(0\right)\right)\right]y_1^\varepsilon(0)\right\}$, the duality relation (220) follows immediately by combining (227)∼(230).

Proof of duality relation (221)**.** Combining (219) and (220) we get

$$\begin{aligned}
&E\left(\Phi^*(T)x_1^\varepsilon(T)\right) + E\left(K^*(T)y_1^\varepsilon(T)\right)\\
&= -E\left\{\left[\varphi_y\left(0\right) + E\left(\varphi_{\widetilde{y}}\left(0\right)\right)\right]y_1^\varepsilon(0)\right\} - E\int_0^T\int_\Theta \{x_1^\varepsilon(t)\ell_x(t,\theta) + x_1^\varepsilon(t)E(\ell_{\widetilde{x}}(t,\theta)) + y_1^\varepsilon(t)\ell_y(t,\theta)\\
&+ y_1^\varepsilon(t)E(\ell_{\widetilde{y}}(t,\theta)) + \ell_u(t,\theta)u(t) + z_1^\varepsilon(t)\ell_z(t,\theta) + z_1^\varepsilon(t)E(\ell_{\widetilde{z}}(t,\theta)) + r_1^\varepsilon(t,\theta)\ell_r(t,\theta)\}\, m(\mathrm{d}\theta)\mathrm{d}t\\
&+ E\int_0^T H_u(t)u(t)\mathrm{d}t + E\int_{[0,T]}\left[\Phi^*(t)\mathcal{C}(t) + K^*(t)\mathcal{D}(t)\right]\mathrm{d}\xi(t).
\end{aligned}$$

From (214) and (218), we get

$$E\left(\Phi^{*}(T)x_1^{\varepsilon}(T)\right)+E\left(K^{*}(T)y_1^{\varepsilon}(T)\right)=E\left\{\left[\phi_x(T)+E(\phi_{\widetilde{x}}(T))\right]x_1^{\varepsilon}(T)\right\},$$

which implies that

$$\begin{aligned}
&E\left\{\left[\phi_x(T)+E(\phi_{\widetilde{x}}(T))\right]x_1^{\varepsilon}(T)\right\}+E\left\{\left[\varphi_y(0)+E\left(\varphi_{\widetilde{y}}(0)\right)\right]y_1^{\varepsilon}(0)\right\}\\
&=-E\int_0^T\int_{\Theta}\{x_1^{\varepsilon}(t)\ell_x(t,\theta)+x_1^{\varepsilon}(t)E(\ell_{\widetilde{x}}(t,\theta))+y_1^{\varepsilon}(t)\ell_y(t,\theta)+y_1^{\varepsilon}(t)E(\ell_{\widetilde{y}}(t,\theta))\\
&\quad+z_1^{\varepsilon}(t)\ell_z(t,\theta)+z_1^{\varepsilon}(t)E(\ell_{\widetilde{z}}(t,\theta))+r_1^{\varepsilon}(t,\theta)\ell_r(t,\theta)+\ell_u(t,\theta)u(t)\}\,m(\mathrm{d}\theta)\mathrm{d}t\\
&\quad+E\int_0^T H_u(t)u(t)\mathrm{d}t+E\int_{[0,T]}\left[\Phi^{*}(t)\mathcal{C}(t)+K^{*}(t)\mathcal{D}(t)\right]\mathrm{d}\xi(t).
\end{aligned}$$

This completes the proof of (221). □

To this end we give the following estimations.

Lemma 4.3.2. Under Assumptions (H1) and (H3), the following estimations holds

$$\begin{aligned}
&E\big(\sup_{0\le t\le T}\left|x_1^{\varepsilon}(t)\right|^2\big)\to 0,\text{ as }\varepsilon\to 0,\\
&E\big(\sup_{0\le t\le T}\left|y_1^{\varepsilon}(t)\right|^2\big)+E\textstyle\int_0^T[\left|z_1^{\varepsilon}(s)\right|^2+\int_{\Theta}\left|r_1^{\varepsilon}(s,\theta)\right|^2m(\mathrm{d}\theta)]\mathrm{d}s\to 0,\text{ as }\varepsilon\to 0,
\end{aligned}\tag{231}$$

$$\begin{aligned}
&\sup_{0\le t\le T}\left|E\left(x_1^{\varepsilon}(t)\right)\right|^2\to 0,\text{ as }\varepsilon\to 0,\\
&\sup_{0\le t\le T}\left|E\left(y_1^{\varepsilon}(t)\right)\right|^2+\textstyle\int_t^T\left|E\left(z_1^{\varepsilon}(s)\right)\right|^2\mathrm{d}s+\int_0^T\int_{\Theta}\left|E\left(r_1^{\varepsilon}(s,\theta)\right)\right|^2m(\mathrm{d}\theta)\mathrm{d}s\to 0,\text{ as }\varepsilon\to 0,
\end{aligned}\tag{232}$$

$$\begin{aligned}
&E\big(\sup_{0\le t\le T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^2\big)\to 0,\text{ as }\varepsilon\to 0,\\
&E\big(\sup_{0\le t\le T}\left|y^{\varepsilon}(t)-y^{*}(t)\right|^2\big)+E\textstyle\int_0^T\left|z^{\varepsilon}(t)-z^{*}(t)\right|^2\mathrm{d}t\\
&+E\textstyle\int_0^T\int_{\Theta}\left|r^{\varepsilon}(t,\theta)-r^{*}(t,\theta)\right|^2m(\mathrm{d}\theta)\mathrm{d}t\to 0,\text{ as }\varepsilon\to 0,
\end{aligned}\tag{233}$$

and

$$\begin{aligned}
&E\Big(\sup_{0\le t\le T}\left|\frac{1}{\varepsilon}\left[x^{\varepsilon}(t)-x^{*}(t)\right]-x_1^{\varepsilon}(t)\right|^2\Big)\to 0,\text{ as }\varepsilon\to 0.\\
&E\Big(\sup_{0\le t\le T}\left|\frac{1}{\varepsilon}\left[y^{\varepsilon}(t)-y^{*}(t)\right]-y_1^{\varepsilon}(t)\right|^2\Big)\to 0,\text{ as }\varepsilon\to 0.\\
&E\int_0^T\left|\frac{1}{\varepsilon}\left[z^{\varepsilon}(s)-z^{*}(s)\right]-z_1^{\varepsilon}(s)\right|^2\mathrm{d}s\to 0,\text{ as }\varepsilon\to 0.\\
&E\int_0^T\int_{\Theta}\left|\frac{1}{\varepsilon}\left[r^{\varepsilon}(s,\theta)-r^{*}(s,\theta)\right]-r_1^{\varepsilon}(s,\theta)\right|^2m(\mathrm{d}\theta)\mathrm{d}s\to 0,\text{ as }\varepsilon\to 0.
\end{aligned}\tag{234}$$

Let us also point out that the above estimates can be proved by using similar arguments developed in (Lemma 3.2 [24]), so we omit its proofs.

Lemma 4.3.3. Let Assumptions (H1) and (H3) hold. The following variational inequality holds

$$
\begin{aligned}
\tau(\varepsilon) \leq E\int_0^T \int_\Theta & \left[\ell_x(t,\theta)x_1^\varepsilon(t) + \ell_{\widetilde{x}}(t,\theta)E\left[x_1^\varepsilon(t)\right] + \ell_y(t,\theta)y_1^\varepsilon(t) + \ell_{\widetilde{y}}(t,\theta)E\left[y_1^\varepsilon(t)\right]\right. \\
& \left. + \ell_z(t,\theta)z_1^\varepsilon(t) + \ell_{\widetilde{z}}(t,\theta)E\left[z_1^\varepsilon(t)\right] + \ell_r(t,\theta)r_1^\varepsilon(t,\theta) + \ell_u(t,\theta)u(t)\right] m(\mathrm{d}\theta)\mathrm{d}t \\
& + E\left[\left(\phi_x(T)x_1^\varepsilon(T) + \phi_{\widetilde{x}}(T)E\left(x_1^\varepsilon(T)\right)\right)\right] + E\left[\left(\varphi_y(0)y_1^\varepsilon(0) + \varphi_{\widetilde{y}}(0)E\left(y_1^\varepsilon(0)\right)\right)\right] \\
& + E\int_{[0,T]} M(t)\mathrm{d}(\xi - \xi^*)(t).
\end{aligned}
$$

Proof. From (217) we have

$$
\begin{aligned}
& J\left(u(\cdot),\xi(\cdot)\right) - J\left(u^*(\cdot),\xi^*(\cdot)\right) \\
& = E\Big\{ \textstyle\int_0^T \int_\Theta [\ell(t, x^{u,\xi}(t), E(x^{u,\xi}(t)), y^{u,\xi}(t), E(y^{u,\xi}(t)), z^{u,\xi}(t), E(z^{u,\xi}(t)), r^{u,\xi}\left(t,\theta\right), u(t)) \\
& -\ell(t, x^*(t), E(x^*(t)), y^*(t), E(y^*(t)), z^*(t), E(z^*(t)), r^*\left(t,\theta\right), u^*(t)] m\left(\mathrm{d}\theta\right)\mathrm{d}t \\
& + \left[\phi\left(x(T), E(x(T))\right) - \phi\left(x^*(T), E(x^*(T))\right)\right] + \left[\varphi\left(y(0), E\left(y(0)\right)\right) - \varphi\left(y^*(0), E\left(y^*(0)\right)\right)\right] \\
& + \textstyle\int_{[0,T]} M(t)\mathrm{d}\left(\xi(t) - \xi^*(t)\right) \Big\} \geq 0,
\end{aligned} \tag{235}
$$

by applying Tylor's formula, we have

$$
\begin{aligned}
& \frac{1}{\varepsilon}E\left\{(\phi(x^\varepsilon(T),\widetilde{x}^\varepsilon(T)) - \phi(x^*(T),\widetilde{x}^*(T))\right\} \\
& = \frac{1}{\varepsilon}E\left\{\int_0^1 \phi_x(x^*(T) + \lambda\left[x^\varepsilon(T) - x^*(T)\right], \widetilde{x}^*(T) + \lambda\left[\widetilde{x}^\varepsilon(T) - \widetilde{x}^*(T)\right])\mathrm{d}\lambda\left(x^\varepsilon(T) - x^*(T)\right)\right. \\
& \left. + \int_0^1 \phi_{\widetilde{x}}(x^*(T) + \lambda\left[x^\varepsilon(T) - x^*(T)\right], \widetilde{x}^*(T) + \lambda\left[\widetilde{x}^\varepsilon(T) - \widetilde{x}^*(T)\right])\mathrm{d}\lambda\left(\widetilde{x}^\varepsilon(T) - \widetilde{x}^*(T)\right)\right\} + \tau(\varepsilon).
\end{aligned}
$$

From estimate (234), we get

$$
\begin{aligned}
& \frac{1}{\varepsilon}E\left\{(\phi(x^\varepsilon(T),\widetilde{x}^\varepsilon(T)) - \phi(x^*(T),\widetilde{x}^*(T))\right\} \\
& \to E[\left(\phi_x(x^*(T),\widetilde{x}^*(T))x_1^\varepsilon(T) + \phi_{\widetilde{x}}(x^*(T),\widetilde{x}^*(T))E\left(x_1^\varepsilon(T)\right)\right)] \\
& = E\left[\left(\phi_x(T)x_1^\varepsilon(T) + \phi_{\widetilde{x}}(T)E\left(x_1^\varepsilon(T)\right)\right)\right], \text{ as } \varepsilon \to 0.
\end{aligned} \tag{236}
$$

Similarly, we obtain

$$
\begin{aligned}
& \frac{1}{\varepsilon}E\left\{(\varphi(y^\varepsilon(0),\widetilde{y}^\varepsilon(0)) - \varphi(y^*(0),\widetilde{y}^*(0))\right\} \\
& \longrightarrow E\left\{(\varphi_y(y^*(0),\widetilde{y}^*(0))y_1^\varepsilon(0) + \varphi_{\widetilde{y}}(y^*(0),\widetilde{y}^*(0))E\left(y_1^\varepsilon(0)\right))\right\} \\
& = E\left[\left(\varphi_y(0)y_1^\varepsilon(0) + \varphi_{\widetilde{y}}(0)E\left(y_1^\varepsilon(0)\right)\right)\right], \text{ as } \varepsilon \to 0.
\end{aligned} \tag{237}
$$

and

$$\begin{aligned}
&\frac{1}{\varepsilon}E\int_0^T\int_\Theta[\ell(t,x^\varepsilon(t),E(x^\varepsilon(t)),y^\varepsilon(t),E(y^\varepsilon(t)),z^\varepsilon(t),E(z^\varepsilon(t)),r^\varepsilon(t,\theta),u^\varepsilon(t))\\
&-\ell(t,x^*(t),E(x^*(t)),y^*(t),E(y^*(t)),z^*(t),E(z^*(t)),r^*(t,\theta),u^*(t))]m(\mathrm{d}\theta)\,\mathrm{d}t\\
&\to E\int_0^T\int_\Theta[\ell_x(t,\theta)x_1^\varepsilon(t)+\ell_{\widetilde{x}}(t,\theta)E[x_1^\varepsilon(t)]+\ell_y(t,\theta)y_1^\varepsilon(t)+\ell_{\widetilde{y}}(t,\theta)E[y_1^\varepsilon(t)]\\
&+\ell_z(t,\theta)z_1^\varepsilon(t)+\ell_{\widetilde{z}}(t,\theta)E[z_1^\varepsilon(t)]+\ell_r(t,\theta)r_1^\varepsilon(t,\theta)+\ell_u(t,\theta)u(t)]\,m(\mathrm{d}\theta)\mathrm{d}t,\ \text{as } \varepsilon\to 0.
\end{aligned}\tag{238}$$

The desired result follows by combining (235)∼(238). This complete the proof of Lemma 3.3 □

Proof of Theorem 3.1. The desired result follows immediately from ((221) Lemma 3.2) and Lemma 3.3. □

21. Sufficient conditions for optimal mixed control of McKean-Vlasov FBSDEJs

The sufficient condition of optimality is of significant importance in the stochastic maximum principle for computing optimal controls. In this section, we will prove that under some additional hypotheses, the maximality condition on the Hamiltonian function is a sufficient condition for optimality.

Conditions (H4). We assume:

(i) The functional $H(t,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\Phi^*(t),Q^*(t),K^*(t),R^*(\cdot,\cdot))$ is convex with respect to $(x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r,u)$ for $a.e.t\in[0,T]$, $\mathbb{P}-a.s.$

(ii) The maps $\phi(\cdot,\cdot)$, $\varphi(\cdot,\cdot)$ are convex with respect to $(x,\widetilde{x})$ and $h(\cdot,\cdot)$ is concave with respect to $(x,\widetilde{x})$.

Now we are able to state and prove the sufficient conditions for optimality for our control problem (209)−(210), which is the main result of this work. Let $(u^*(\cdot),\xi^*(\cdot))$ be a given admissible control, $(x^*(\cdot),y^*(\cdot),z^*(\cdot),r^*(\cdot,\cdot))$ and $(\Phi^*(\cdot),Q^*(\cdot),K^*(\cdot),R^*(\cdot,\cdot))$ be the solution to (209) and (212) respectively, associated with $(u^*(\cdot),\xi^*(\cdot))$.

Theorem 4.4.1. Let conditions (H1)-(H4) hold. If for any $(u(\cdot),\xi(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2([0,T])$ the following maximality relation holds

$$\begin{aligned}
&E\int_0^T H_u(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*,\Lambda^*(t,\theta))(u-u^*(t))\mathrm{d}t\geq 0,\\
&E\int_{[0,T]}(M(t)+\mathcal{C}(t)\Phi^*(t)+\mathcal{D}(t)K^*(t))\mathrm{d}(\xi-\xi^*)(t)\geq 0,\\
&a.e.,\ t\in[0,T],
\end{aligned}\tag{239}$$

then we have

$$\inf_{(u(\cdot),\xi(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2([0,T])}J(u(\cdot),\xi(\cdot))=J(u^*(\cdot),\xi^*(\cdot)).\tag{240}$$

i.e., the regular-singular admissible control $(u^*(\cdot),\xi^*(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2([0,T])$ is an optimal control.

To prove Theorem 4.1, we need the following auxiliary result, which deals with the duality relations between $\Phi^*(t)$, $[x(t)-x^*(t)]$ and $K^*(t)$, $[y(t)-y^*(t)]$. This Lemma is very important for proving our sufficient maximum principle.

Lemma 4.4.1. Let $(x(\cdot), y(\cdot), z(\cdot), r(\cdot,\cdot))$ be the solution of state McKean-Vlasov FBSDEJs-(209) corresponding to any admissible control $(u(\cdot), \xi(\cdot))$. We have

$$\begin{aligned}
&E\left[\Phi^*(T)\left(x(T)-x^*(T)\right)\right] = E\int_0^T \Phi^*(t)\left[f(t,x(t),E(x(t)),u(t)) - f(t,x^*(t),E(x^*(t)),u^*(t))\right]\mathrm{d}t \\
&\quad + E\int_0^T H_x^*(t)\left(x(t)-x^*(t)\right)\mathrm{d}t + E\int_0^T E[H_{\widetilde{x}}^*(t)]\left(E(x(t))-E(x^*(t))\right)\mathrm{d}t \\
&\quad + E\int_0^T Q^*(t)\left[\sigma(t,x(t),E(x(t)),u(t)) - \sigma(t,x^*(t),E(x^*(t)),u^*(t))\right]\mathrm{d}t \\
&\quad + E\int_0^T\int_\Theta R^*(t,\theta)\left[\gamma\left(t,x(t),E(x(t),u(t),\theta\right) - \gamma\left(t,x^*(t),E(x^*(t),u^*(t),\theta\right)\right]m(\mathrm{d}\theta)\mathrm{d}t \\
&\quad + E\int_{[0,T]}\Phi^*(t)\mathcal{C}(t)\mathrm{d}(\xi(t)-\xi^*(t)).
\end{aligned} \tag{241}$$

Similarly

$$\begin{aligned}
&E\left[K^*(T)\left(y(T)-y^*(T)\right)\right] = -E\left(\varphi_y\left(y(0),E\left(y(0)\right)\right)\left(y^*(0)-y(0)\right)\right) \\
&\quad - E\left(\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)\right)\left(E\left(y^*(0)\right)-E\left(y(0)\right)\right) \\
&\quad + E\int_0^T\int_\Theta K^*(t)\left\{g(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t)) - g(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))\right\}m\left(\mathrm{d}\theta\right)\mathrm{d}t \\
&\quad + E\int_0^T H_y(t)\left(y(t)-y^*(t)\right)\mathrm{d}t + E\int_0^T E(H_{\widetilde{y}}(t))\left(E(y(t))-E(y^*(t))\right))\mathrm{d}t \\
&\quad + E\int_0^T H_z^*(t)\left(z(t)-z^*(t)\right)\mathrm{d}t + E\int_0^T E(H_{\widetilde{z}}^*(t))\left(E(z(t))-E(z^*(t))\right)\mathrm{d}t \\
&\quad + E\int_0^T\int_\Theta H_r^*(t)\left[r(t,\theta)-r^*(t,\theta)\right]m\left(\mathrm{d}\theta\right)\mathrm{d}t + E\int_{[0,T]}K^*(t)\mathcal{D}(t)\mathrm{d}(\xi(t)-\xi^*(t)),
\end{aligned} \tag{242}$$

and

$$\begin{aligned}
&E\left[\Phi^*(T)\left(x(T)-x^*(T)\right)\right] + E\left[K^*(T)\left(y(T)-y^*(T)\right)\right] \\
&+E\left(\varphi_y\left(y(0),E\left(y(0)\right)\right)\left(y^*(0)-y(0)\right)\right) + E[\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)]\left(E\left(y^*(0)\right)-E\left(y(0)\right)\right) \\
&= E\int_0^T \Phi^*(t)(f(t,x(t),E(x(t)),u(t)-f(t,x^*(t),E(x^*(t)),u^*(t)))\mathrm{d}t \\
&+E\int_0^T Q^*(t)[\sigma(t,x(t),E(x(t)),u(t))-\sigma(t,x^*(t),E(x^*(t)),u^*(t))]\mathrm{d}t \\
&+E\int_0^T\int_\Theta K^*(t)[g(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t)))-g(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t \\
&+E\int_0^T H_x^*(t)\left(x(t)-x^*(t)\right)\mathrm{d}t + E\int_0^T E\left[H_{\widetilde{x}}^*(t)\right]\left(E(x(t))-E(x^*(t))\right)\mathrm{d}t \\
&+E\int_0^T H_y^*(t)\left(y(t)-y^*(t)\right)\mathrm{d}t + E\int_0^T E(H_{\widetilde{y}}^*(t))\left(E(y(t))-E(y^*(t))\right))\mathrm{d}t \\
&+E\int_0^T H_z^*(t)\left(z(t)-z^*(t)\right)\mathrm{d}t + E\int_0^T E(H_{\widetilde{z}}^*(t))\left(E(z(t))-E(z^*(t))\right)\mathrm{d}t \\
&+E\int_0^T\int_\Theta H_r^*(t)\left[r(t,\theta)-r^*(t,\theta)\right]m\left(\mathrm{d}\theta\right)\mathrm{d}t + E\int_{[0,T]}\left[\Phi^*(t)\mathcal{C}(t)+K^*(t)\mathcal{D}(t)\right]\mathrm{d}(\xi(t)-\xi^*(t)),
\end{aligned} \tag{243}$$

Proof. First, by simple computations, we get

$$\begin{aligned}
d\left(x(t)-x^*(t)\right) &= \left[f(t,x(t),E(x(t)),u(t)) - f(t,x^*(t),E(x^*(t)),u^*(t))\right]\mathrm{d}t \\
&\quad + \left[\sigma(t,x(t),E(x(t)),u(t)) - \sigma(t,x^*(t),E(x^*(t)),u^*(t)\right]\mathrm{d}W(t) \\
&\quad + \left[\int_\Theta(\gamma(t,x(t),E(x(t)),u(t),\theta) - \gamma(t,x^*(t),E(x^*(t)),u^*(t),\theta))\right]N(\mathrm{d}\theta,\mathrm{d}t) \\
&\quad + \mathcal{C}(t)\mathrm{d}(\xi(t)-\xi^*(t)),
\end{aligned} \tag{244}$$

$$
\begin{aligned}
d\left(y(t)-y^{*}(t)\right) &= \textstyle\int_{\Theta}[g(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t))-g(t,\lambda^{*}(t,\theta),E(\lambda^{*}(t,\theta)),u^{*}(t))]\mathrm{d}t \\
&\quad +\left(z(t)-z^{*}(t)\right)\mathrm{d}W(t)+\textstyle\int_{\Theta}\left[r(t,\theta)-r^{*}(t,\theta)\right]N(\mathrm{d}\theta,\mathrm{d}t) \\
&\quad +\mathcal{D}(t)\mathrm{d}(\xi(t)-\xi^{*}(t)).
\end{aligned} \tag{245}
$$

By applying integration by parts formula to $\Phi^{*}(t)\left(x(t)-x^{*}(t)\right)$ and the fact that $x(0)-x^{*}(0)=0$, we get

$$
\begin{aligned}
&E\left\{\Phi^{*}(T)\left(x(T)-x^{*}(T)\right)\right\}=E\textstyle\int_{0}^{T}\Phi^{*}(t)\mathrm{d}\left(x(t)-x^{*}(t)\right)+E\int_{0}^{T}\left(x(t)-x^{*}(t)\right)\mathrm{d}\Phi^{*}(t) \\
&\quad +E\textstyle\int_{0}^{T}Q^{*}(t)[\sigma(t,x(t),E(x(t)),u(t))-\sigma(t,x^{*}(t),E(x^{*}(t)),u^{*}(t))]\mathrm{d}t \\
&\quad +E\textstyle\int_{0}^{T}\int_{\Theta}R^{*}(t,\theta)\left[\gamma(t,x(t),E(x(t)),u(t),\theta)-\gamma(t,x^{*}(t),E(x^{*}(t)),u^{*}(t),\theta)\right]m\left(\mathrm{d}\theta\right)\mathrm{d}t \\
&\quad =I_{1}+I_{2}+I_{3}+I_{4}.
\end{aligned} \tag{246}
$$

From (244), we obtain

$$
\begin{aligned}
I_{1} &= E\textstyle\int_{0}^{T}\Phi^{*}(t)\mathrm{d}\left(x(t)-x^{*}(t)\right) \\
&= E\textstyle\int_{0}^{T}\Phi^{*}(t)[f(t,x(t),E(x(t)),u(t))-f(t,x^{*}(t),E(x^{*}(t)),u^{*}(t))]\mathrm{d}t \\
&\quad +E\textstyle\int_{0}^{T}\Phi^{*}(t)\mathcal{C}(t)\mathrm{d}(\xi(t)-\xi^{*}(t)),
\end{aligned} \tag{247}
$$

similarly, by applying (214), we get

$$
\begin{aligned}
I_{2} &= E\textstyle\int_{0}^{T}\left(x(t)-x^{*}(t)\right)\mathrm{d}\Phi^{*}(t) \\
&= E\textstyle\int_{0}^{T}\left(x(t)-x^{*}(t)\right)[H_{x}^{*}(t)+E(H_{\widetilde{x}}^{*}(t))]\mathrm{d}t \\
&= E\textstyle\int_{0}^{T}H_{x}^{*}(t)\left(x(t)-x^{*}(t)\right)\mathrm{d}t+\int_{0}^{T}E(H_{\widetilde{x}}^{*}(t))\left(E(x(t))-E(x^{*}(t))\right)\mathrm{d}t.
\end{aligned} \tag{248}
$$

By standard arguments, we obtain

$$
I_{3}=E\int_{0}^{T}Q^{*}(t)[\sigma(t,x(t),E(x(t)),u(t))-\sigma(t,x^{*}(t),E(x^{*}(t)),u^{*}(t))]\mathrm{d}t, \tag{249}
$$

and

$$
I_{4}=E\int_{0}^{T}\int_{\Theta}R^{*}(t,\theta)\left[\gamma(t,x(t),E(x(t),u(t),\theta)-\gamma(t,x^{*}(t),E(x^{*}(t),u^{*}(t),\theta)\right]m\left(\mathrm{d}\theta\right)\mathrm{d}t. \tag{250}
$$

The duality relation (241) follows from combining (247)∼(250) together with (246).
Let us turn to second duality relation (242). By applying *integration by parts formula* to $K^{*}(t)\left[y^{*}(t)-y(t)\right]$, we get

$$
\begin{aligned}
&E\left(K^{*}(T)\left(y^{*}(T)-y(T)\right)\right)=E\left\{K^{*}(0)\left(y^{*}(0)-y(0)\right)\right\} \\
&\quad +E\textstyle\int_{0}^{T}K^{*}(t)d\left(y(t)-y^{*}(t)\right)+E\int_{0}^{T}\left(y(t)-y^{*}(t)\right)dK^{*}(t) \\
&\quad +E\textstyle\int_{0}^{T}\left(z(t)-z^{*}(t)\right)[H_{z}^{*}(t)+E(H_{\widetilde{z}}^{*}(t))]\mathrm{d}t \\
&\quad +E\textstyle\int_{0}^{T}\int_{\Theta}H_{r}^{*}(t)\left[r(t,\theta)-r^{*}(t,\theta)\right]m\left(\mathrm{d}\theta\right)\mathrm{d}t \\
&\quad =I_{1}+I_{2}+I_{3}+I_{4}+I_{5}.
\end{aligned} \tag{251}
$$

Let us turn to the first term I_2. From (245) we get

$$
\begin{aligned}
I_2 &= E\int_0^T K^*(t)\mathrm{d}\left(y(t)-y^*(t)\right)\\
&= E\int_0^T\int_\Theta K^*(t)[g(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t))-g(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\
&+ E\int_{[0,T]} K^*(t)\mathcal{D}(t)\mathrm{d}(\xi(t)-\xi^*(t)),
\end{aligned}
\tag{252}
$$

from (214), we obtain

$$
\begin{aligned}
I_3 &= E\int_0^T \left(y(t)-y^*(t)\right)\mathrm{d}K^*(t) = E\int_0^T \left(y(t)-y^*(t)\right)\left(H_y^*(t)+E(H_{\widetilde{y}}^*(t))\right)\mathrm{d}t\\
&= E\int_0^T H_y^*(t)\left(y(t)-y^*(t)\right)\mathrm{d}t + E\int_0^T E(H_{\widetilde{y}}^*(t))\left(E(y(t))-E(y^*(t))\right)\mathrm{d}t.
\end{aligned}
\tag{253}
$$

$$
\begin{aligned}
I_4 &= E\int_0^T \left(z(t)-z^*(t)\right)\left[H_z^*(t)+E(H_{\widetilde{z}}^*(t))\right]\mathrm{d}t\\
&= E\int_0^T H_z^*(t)\left(z(t)-z^*(t)\right)\mathrm{d}t + E\int_0^T E(H_{\widetilde{z}}^*(t))\left(E(z(t))-E(z^*(t))\right)\mathrm{d}t,
\end{aligned}
\tag{254}
$$

and

$$
I_5 = E\int_0^T\int_\Theta H_r^*(t)\left[r(t,\theta)-r^*(t,\theta)\right]m\left(\mathrm{d}\theta\right)\mathrm{d}t.
\tag{255}
$$

From (212) and the fact that

$$
\begin{aligned}
I_1 &= E\left\{K^*(0)\left(y^*(0)-y(0)\right)\right\}\\
&= -E\Big\{\left[\varphi_y\left(y(0),E\left(y(0)\right)\right)+E(\varphi_{\widetilde{y}}(y(0),E\left(y(0)\right)))\right]\left(y^*(0)-y(0)\right)\Big\}\\
&= -E\left[\varphi_y\left(y(0),E\left(y(0)\right)\right)\left(y^*(0)-y(0)\right)\right]-E(\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right))[E\left(y^*(0)\right)-E(y(0))],
\end{aligned}
\tag{256}
$$

the duality relation (242) follows immediately by combining (252)∼(256) together with (251). Finally, inequality (243) follows from combining (241) and (242). □

Proof of Theorem 4.4.1. Let $(x(\cdot),y(\cdot),z(\cdot),r\left(\cdot,\cdot\right))$ be the solution of the state equation (209) and $\left(\Phi\left(\cdot\right),Q\left(\cdot\right),\ K\left(\cdot\right),\ R\left(\cdot,\cdot\right)\right)$ be the solution of the adjoint equation (212), corresponding to $(u(\cdot),\xi(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2\left([0,T]\right)$.
By concavity of $H\left(t,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\cdot,\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta)\right)$ with respect to $(x,\widetilde{x},y,\widetilde{y},z,\widetilde{z},r,u)$, we obtain

$$
\begin{aligned}
&H(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta))\\
&-H(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta))\\
&\geq H_x^*(t)(x(t)-x^*(t))+E(H_{\widetilde{x}}^*(t))(E(x(t)-x^*(t)))+H_y^*(t)(y(t)-y^*(t))\\
&+E(H_{\widetilde{y}}^*(t))(E(y(t)-y^*(t)))+H_z^*(t)(z(t)-z^*(t)))+E(H_{\widetilde{z}}^*(t))(E(z(t)-z^*(t)))\\
&+\int_\Theta H_r^*(t)(r(t,\theta)-r^*(t,\theta)))m\left(\mathrm{d}\theta\right)+H_u^*(t)(u(t)-u^*(t)).
\end{aligned}
$$

Integrating this inequality with respect to t and taking expectations, with the help of (239) we get

$$\begin{aligned}
&E\int_0^T(H(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta))\\
&-H(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta)))\mathrm{d}t\\
&\geq E\int_0^T\Big\{H_x^*(t)(x(t)-x^*(t))+E(H_{\widetilde{x}}^*(t))(E(x(t)-x^*(t)))+H_y^*(t)(y(t)-y^*(t))\\
&+E(H_{\widetilde{y}}^*(t))(E(y(t)-y^*(t)))+H_z^*(t)(z(t)-z^*(t)))+E(H_{\widetilde{z}}^*(t))(E(z(t)-z^*(t)))\\
&+\int_\Theta H_r^*(t)(r(t,\theta)-r^*(t,\theta)))m\,(\mathrm{d}\theta)\Big\}\mathrm{d}t.
\end{aligned}\tag{257}$$

Now, using (243) in Lemma 4.3.1 and definition of the Hamiltonian (213) and the fact that

$$\begin{aligned}
&E\int_0^T[H(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta))\\
&-H(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta))]\mathrm{d}t\\
&=E\int_0^T\Phi^*(t)(f(t,x(t),E(x(t)),u(t)-f(t,x^*(t),E(x^*(t)),u^*(t)))\mathrm{d}t\\
&+E\int_0^T Q^*(t)[\sigma(t,x(t),E(x(t)),u(t))-\sigma(t,x^*(t),E(x^*(t)),u^*(t))]\mathrm{d}t\\
&+E\int_0^T\int_\Theta K^*(t)[g(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t)))-g(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\,(\mathrm{d}\theta)\,\mathrm{d}t\\
&+E\int_0^T\int_\Theta R^*(t,\theta)[\gamma(t,x(t),E(x(t)),u(t),\theta))-\gamma(t,x^*(t),E(x^*(t)),u^*(t),\theta))]m\,(\mathrm{d}\theta)\,\mathrm{d}t\\
&+E\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t)))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\,(\mathrm{d}\theta)\,\mathrm{d}t,
\end{aligned}$$

we get

$$\begin{aligned}
&E\left[\Phi^*(T)\left(x(T)-x^*(T)\right)\right]+E\left[K^*(T)\left(y(T)-y^*(T)\right)\right]\\
&+E\left(\varphi_y\left(y(0),E\left(y(0)\right)\right)\left(y^*(0)-y(0)\right)\right)+E[\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)]\left(E\left(y^*(0)\right)-E\left(y(0)\right)\right)\\
&+E\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta),u(t)))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\,(\mathrm{d}\theta)\,\mathrm{d}t\\
&-E\int_{[0,T]}\left[\Phi^*(t)\mathcal{C}(t)+K^*(t)\mathcal{D}(t)\right]\mathrm{d}(\xi(t)-\xi^*(t))\\
&=E\int_0^T[H(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta)))\\
&-H(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t),\Phi^*(t),Q^*(t),K^*(t),R^*(t,\theta)))]\mathrm{d}t\\
&+E\int_0^T H_x^*(t)\left(x(t)-x^*(t)\right)\mathrm{d}t+E\int_0^T E\left[H_{\widetilde{x}}^*(t)\right]\left(E(x(t))-E(x^*(t))\right)\mathrm{d}t\\
&+E\int_0^T H_y^*(t)\left(y(t)-y^*(t)\right)\mathrm{d}t+E\int_0^T E(H_{\widetilde{y}}^*(t))\left(E(y(t))-E(y^*(t)))\right)\mathrm{d}t\\
&+E\int_0^T H_z^*(t)\left(z(t)-z^*(t)\right)\mathrm{d}t+E\int_0^T E(H_{\widetilde{z}}^*(t))\left(E(z(t))-E(z^*(t))\right)\mathrm{d}t\\
&+E\int_0^T\int_\Theta H_r^*(t)\left(r(t,\theta)-r^*(t,\theta)\right)m\,(\mathrm{d}\theta)\,\mathrm{d}t,
\end{aligned}\tag{258}$$

combining (257), (258) we get

$$\begin{aligned}&E\left[\Phi^*(T)\left(x(T)-x^*(T)\right)\right]+E\left[K^*(T)\left(y(T)-y^*(T)\right)\right]\\&+E\left(\varphi_y\left(y(0),E\left(y(0)\right)\right)\left(y^*(0)-y(0)\right)\right)+E[\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)]\left(E\left(y^*(0)\right)-E\left(y(0)\right)\right)\\&+E\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t)))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\&-E\int_{[0,T]}\left[\Phi^*(t)\mathcal{C}(t)+K^*(t)\mathcal{D}(t)\right]d(\xi(t)-\xi^*(t))\geq 0,\end{aligned}$$

by the fact that $E\int_{[0,T]}\left[M(t)+\Phi^*(t)\mathcal{C}(t)+K^*(t)\mathcal{D}(t)\right]d(\xi(t)-\xi^*(t))\geq 0$ (see (239))

$$\begin{aligned}&E\left[\Phi^*(T)\left(x(T)-x^*(T)\right)\right]+E\left[K^*(T)\left(y(T)-y^*(T)\right)\right]+E\left(\varphi_y\left(y(0),E\left(y(0)\right)\right)\left(y(0)-y^*(0)\right)\right)\\&+E[\varphi_{\widetilde{y}}\left(y(0),E\left(y(0)\right)\right)]\left(E\left(y(0)\right)-E\left(y^*(0)\right)\right)\\&+E\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t)))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\&+E\int_{[0,T]}M(t)\mathrm{d}(\xi(t)-\xi^*(t))\geq 0.\end{aligned}$$

Since $\Phi(T) = -\left\{h_x\left(T\right)K(T)+E\left[\left(h_{\widetilde{x}}\left(T\right)\right)K(T)\right]\right\} + \phi_x(T) + E(\phi_{\widetilde{x}}(T))$ and $y(T) = h\left(x(T),E\left(x(T)\right)\right)$ we get

$$\begin{aligned}&-E\left[\left\{h_x\left(T\right)K^*(T)+E\left[\left(h_{\widetilde{x}}\left(T\right)\right)K^*(T)\right]\right\}\left(x(T)-x^*(T)\right)\right]\\&+E[\left(\phi_x(T)+E(\phi_{\widetilde{x}}(T))\right)\left(x(T)-x^*(T)\right)]\\&+E\left[K^*(T)\left(h\left(x(T),E\left(x(T)\right)\right)-h\left(x^*(T),E\left(x^*(T)\right)\right)\right)\right]\\&+E\left(\varphi_y\left(y^*(0),E\left(y^*(0)\right)\right)\left(y(0)-y^*(0)\right)\right)+E[\varphi_{\widetilde{y}}\left(y^*(0),E\left(y^*(0)\right)\right)]\left(E\left(y(0)\right)-E\left(y^*(0)\right)\right)\\&+E\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\&+E\int_{[0,T]}M(t)\mathrm{d}(\xi(t)-\xi^*(t))\geq 0,\end{aligned} \tag{259}$$

from the concavity of $h(\cdot,\cdot)$, the convexity of $\phi(\cdot,\cdot)$ and $\varphi(\cdot,\cdot)$, we get

$$\begin{aligned}&0\leq E\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\&+E\left[\phi\left(x(T),E(x(T))\right)-\phi\left(x^*(T),E(x^*(T))\right)\right]+E\left[\varphi\left(y(0),E\left(y(0)\right)\right)-\varphi\left(y^*(0),E\left(y^*(0)\right)\right)\right]\\&+E\int_{[0,T]}M(t)\mathrm{d}\left(\xi(t)-\xi^*(t)\right).\end{aligned} \tag{260}$$

Combining (259) and (260) and the fact that

$$\begin{aligned}&J\left(u(\cdot),\xi(\cdot)\right)-J\left(u^*(\cdot),\xi^*(\cdot)\right)\\&\quad=E\Big\{\int_0^T\int_\Theta[\ell(t,\lambda(t,\theta),E(\lambda(t,\theta)),u(t))-\ell(t,\lambda^*(t,\theta),E(\lambda^*(t,\theta)),u^*(t))]m\left(\mathrm{d}\theta\right)\mathrm{d}t\\&\quad+\left[\phi\left(x(T),E(x(T))\right)-\phi\left(x^*(T),E(x^*(T))\right)\right]+\left[\varphi\left(y(0),E\left(y(0)\right)\right)-\varphi\left(y^*(0),E\left(y^*(0)\right)\right)\right]\\&\quad+\int_{[0,T]}M(t)\mathrm{d}\left(\xi(t)-\xi^*(t)\right)\Big\},\end{aligned}$$

we get

$$J\left(u(\cdot), \xi(\cdot)\right) \geq J\left(u^*(\cdot), \xi^*(\cdot)\right).$$

Finally, since $(u(\cdot), \xi(\cdot))$ is an arbitrary admissible control of $\mathcal{U}_1 \times \mathcal{U}_2([0,T])$, we have

$$J\left(u^*(\cdot), \xi^*(\cdot)\right) = \inf_{(u(\cdot),\xi(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2([0,T])} J\left(u(\cdot), \xi(\cdot)\right),$$

the desired result follows. □

22. Application: mean-variance portfolio selection problem with interventions control

It is well known that mean-variance portfolio selection problem introduced by Markowitz [43] is a time-inconsistent control problem, in the sense that it does not satisfy Bellman's optimality principle and therefore the usual dynamic programming approach fails. In this section, we will apply our time-inconsistent maximum principle of optimality to study mean-variance portfolio selection problem mixed with a recursive utility functional optimization in a financial market involving singular control.

We first come back to Example 1 and solve the optimal regular-singular control problem (206)-(208). We note that the cost functional (208) becomes a time-inconsistent control problem. Let $\mathbb{U}_1\times\mathbb{U}_2$ be a compact convex subset of $\mathbb{R}\times\mathbb{R}$. We denote $\mathcal{U}_1 \times \mathcal{U}_2([0,T])$ the set of admissible $\mathbb{F}_t-$predictable portfolio strategies $\left(u\left(\cdot\right), \xi(\cdot)\right)$ valued in $\mathbb{U}_1\times\mathbb{U}_2$.
Now, we start our attempt to solve our mean-variance portfolio selection problem mixed with a recursive utility functional, time-inconsistent solutions, (206)-(208) by writing down the Hamiltonian and the adjoint processes for this system. The Hamiltonian functional (213) gets the form

$$\begin{aligned} &H\left(t, x, \widetilde{x}, y, \widetilde{y}, z, \widetilde{z}, r(\cdot), u, \Phi, Q, K, R(\cdot,\cdot)\right) \\ &\quad= \left[\rho(t)x(t) + (\varsigma(t) - \rho(t))u(t)\right]\left(\Phi(t) + K(t)\right) \\ &\quad+ \sigma(t)u(t)Q(t) - \alpha K(t)y(t) + \textstyle\int_\Theta A\left(t,\theta\right)u(t)R\left(t,\theta\right)m\left(\mathrm{d}\theta\right). \end{aligned}$$

According to the maximum condition ((216), *Theorem 3.1),* and since $(u^*(\cdot), \xi^*(\cdot))$ is optimal we immediately get

$$\left(\varsigma(t) - \rho(t)\right)\left(\Phi^*(t) + K^*(t)\right) + \sigma(t)Q^*(t) + \int_\Theta A\left(t,\theta\right)R^*\left(t,\theta\right)m\left(\mathrm{d}\theta\right) = 0. \tag{261}$$

The adjoint equation (212) being

$$\begin{cases} \mathrm{d}\Phi^*(t) = -\rho(t)\left(K^*(t) + \Phi^*(t)\right)\mathrm{d}t + Q^*(t)\mathrm{d}W(t) + \int_\Theta R^*\left(t,\theta\right)N\left(\mathrm{d}\theta, \mathrm{d}t\right). \\ \Phi^*(T) = \delta\left(x^*(T) + E(x^*(T))\right) - 1 - K^*(T), \\ \mathrm{d}K^*(t) = -\alpha K^*(t)\mathrm{d}t,\ K^*(0) = -1,\ t\in[0,T]. \end{cases} \tag{262}$$

In order to solve the above equation (262) and to find the expression of optimal portfolio strategy $(u^*(\cdot),\xi^*(\cdot))$ we conjecture a process $\Phi^*(t)$ of the form

$$\Phi^*(t)=\psi_1(t)x^*(t)+\psi_2(t)E\left(x^*(t)\right)+\psi_3(t), \tag{263}$$

where $\psi_1(\cdot),\psi_2(\cdot)$ and $\psi_3(\cdot)$ are deterministic differentiable functions. (See Hafayed and Abbas [17], Hafayed [23], Li [49], Anderson and Djehiche [5], for other models of conjecture). From last equation in (262), which is a simple *ordinary differential equation (*ODE in short), we get immediately

$$K^*(t)=-\exp\left(-\alpha t\right). \tag{264}$$

Noting that from (206), we get

$$\mathrm{d}(E(x^*(t))=\left\{\rho(t)E(x^*(t))+(\varsigma(t)-\rho(t))E(u^*(t))\right\}\mathrm{d}t.$$

Applying Itô's formula to (263) (see Hafayed [24], Lemma 3.5, Appendix) in virtue of SDE-(206), we get

$$\begin{aligned}\mathrm{d}\Phi^*(t)&=\psi_1(t)\left\{\left[\rho(t)x^*(t)+(\varsigma(t)-\rho(t))u^*(t)\right]\mathrm{d}t+\sigma(t)u^*(t)\mathrm{d}W(t)\right.\\&\quad\left.+\int_\Theta A\left(t_-,\theta\right)u^*(t)N\left(\mathrm{d}\theta,\mathrm{d}t\right)\right\}+x^*(t)\psi_1'(t)\mathrm{d}t\\&\quad+\psi_2(t)\left[\rho(t)E(x^*(t))+(\varsigma(t)-\rho(t))E(u^*(t))\right]\mathrm{d}t+E\left(x^*(t)\right)\psi_2'(t)\mathrm{d}t+\psi_3'(t)\mathrm{d}t,\end{aligned}$$

which implies that

$$\begin{cases}\mathrm{d}\Phi^*(t)=\left\{\psi_1(t)\left[\rho(t)x^*(t)+(\varsigma(t)-\rho(t))\,u^*(t)\right]+x^*(t)\psi_1'(t)\right.\\\qquad\qquad+\psi_2(t)\left[\rho(t)E(x^*(t))+(\varsigma(t)-\rho(t))E(u^*(t))\right]\\\qquad\qquad\left.+\psi_2'(t)E\left(x^*(t)\right)+\psi_3'(t)\right\}\mathrm{d}t+\psi_1(t)\sigma(t)u^*(t)\mathrm{d}W(t)\\\qquad\qquad+\int_\Theta\psi_1(t)A\left(t_-,\theta\right)u^*(t)N\left(\mathrm{d}\theta,\mathrm{d}t\right),\\\Phi^*(T)=\psi_1(T)x^*(T)+\psi_2(T)E\left(x^*(T)\right)+\psi_3(T),\end{cases} \tag{265}$$

where $\psi_1'(t)$, $\psi_2'(t)$, and $\psi_3'(t)$ denotes the derivatives with respect to t.
Next, comparing (265) with (262), we get

$$\begin{aligned}-\rho(t)\left(K^*(t)+\Phi^*(t)\right)&=\psi_1(t)\left[\rho(t)x^*(t)+(\varsigma(t)-\rho(t))u^*(t)\right]+x^*(t)\psi_1'(t)\\&+\psi_2(t)\left[\rho(t)E(x^*(t))+(\varsigma(t)-\rho(t))E(u^*(t))\right]+\psi_2'(t)E\left(x^*(t)\right)+\psi_3'(t),\end{aligned} \tag{266}$$

$$Q^*(t)=\psi_1(t)\sigma(t)u^*(t), \tag{267}$$

$$R^*(t,\theta)=\psi_1(t)A\left(t,\theta\right)u^*(t). \tag{268}$$

By looking at the terminal condition of $\Phi^*(t)$, in (265), it is reasonable to get

$$\psi_1(T)=\delta,\ \psi_2(T)=-\delta,\ \psi_3(T)=-1-K^*(T). \tag{269}$$

Combining (266) and (263) we deduce that $\psi_1(\cdot), \psi_2(\cdot)$ and $\psi_3(\cdot)$ satisfying the following ordinary differential equation

$$\begin{cases} \psi_1'(t) = -2\rho(t)\psi_1(t), \quad \psi_1(T) = \delta, \\ \psi_2'(t) = -2\rho(t)\psi_2(t), \quad \psi_2(T) = -\delta, \\ \psi_3'(t) + \rho(t)\psi_3(t) = \rho(t) \exp\{-\alpha t\}, \\ \psi_3(T) = \exp\{-\alpha T\} - 1. \end{cases} \tag{270}$$

By solving the first two equations in (270) we obtain

$$\psi_1(t) = -\psi_2(t) = \delta \exp\left\{2\int_t^T \rho(s)\mathrm{d}s\right\}. \tag{271}$$

Using Integrating factor method for the third equation in (270), we get

$$\psi_3(t) = \chi(t)^{-1}\left[\exp(-\alpha T) - 1 - \int_t^T \chi(s)\rho(s)\exp\{-\alpha s\}\,\mathrm{d}s\right], \tag{272}$$

where the integrating factor is $\chi(t) = \exp(\int_t^T \rho(s)\mathrm{d}s)$, $\chi(T) = 1$. Combining (261), (264), (267) and (268) we get

$$u^*(t) = (\rho(t) - \varsigma(t))\frac{\psi_1(t)\left(x^*(t) - E(x^*(t))\right) + \psi_3(t) - \exp\{-\alpha t\}}{\psi_1(t)\left(\sigma^2(t) + \int_\Theta A^2(t,\theta)\,m(\mathrm{d}\theta)\right)}, \tag{273}$$

and

$$E(u^*(t)) = \frac{(\rho(t) - \varsigma(t))\left[\psi_3(t) - \exp\{-\alpha t\}\right]}{\psi_1(t)\left(\sigma^2(t) + \int_\Theta A^2(t,\theta)\,m(\mathrm{d}\theta)\right)}. \tag{274}$$

Let $\xi^*(\cdot)$ satisfies the maximum condition (239), we get: for any $\eta(\cdot) \in \mathcal{U}_2([0,T])$

$$E\int_{[0,T]} (M(t) + G\Phi^*(t) + \beta K^*(t))\mathrm{d}(\eta - \xi^*)(t) \geq 0,$$

where $(\Phi^*(t), K^*(t))$ is the adjoint processes corresponding to optimal control $u^*(\cdot)$.

Now, we define a set

$$B = \{(w,t) \in \Omega \times [0,T] : M(t) + G\Phi^*(t) + \beta K^*(t) > 0\}, \tag{275}$$

and let $\eta(\cdot) \in \mathcal{U}_2([0,T])$ such that

$$\mathrm{d}\eta(t) = \begin{cases} 0 : \text{if } M(t) + G\Phi^*(t) + \beta K^*(t) > 0, \\ \mathrm{d}\xi^*(t) : \text{if } M(t) + G\Phi^*(t) + \beta K^*(t) \leq 0. \end{cases} \tag{276}$$

By a simple computations it is easy to get

$$\begin{aligned} 0 &\leq E\int_{[0,T]} (M(t) + G\Phi^*(t) + \beta K^*(t))\mathrm{d}(\eta(t) - \xi^*(t)) \\ &= E\int_{[0,T]} (M(t) + G\Phi^*(t) + \beta K^*(t))\boldsymbol{I}_B(t,w)\mathrm{d}(-\xi^*)(t) \\ &= -E\int_{[0,T]} (M(t) + G\Phi^*(t) + \beta K^*(t))\boldsymbol{I}_B(t,w)\mathrm{d}\xi^*(t), \end{aligned}$$

this implies that $\xi^*(\cdot)$ satisfies for any $t \in [0,T]$:

$$E \int_{[0,T]} (M(t) + G\Phi^*(t) + \beta K^*(t)) \boldsymbol{I}_B(t,w) \mathrm{d}\xi^*(t) = 0.$$

Finally, from (275) and (276) we can easy shows that the optimal singular control $\xi^*(\cdot)$ has the form

$$\xi^*(t) = \eta(t) + \int_0^t \boldsymbol{I}_{\{(w,s)\in\Omega\times[0,T]:\, M(s)+G\Phi^*(s)+\beta K^*(s)\leq 0\}}(s,w)\mathrm{d}s.$$

Finally, we give the explicit optimal portfolio section strategy in feedback form involving both $x^*(\cdot)$ and $E(x^*(\cdot))$.

Theorem 4.5.1 The optimal portfolio strategies $(u^*(\cdot), \xi^*(\cdot))$ of our mean-variance portfolio selection mixed with a recursive utility optimization problem involving singular control (206)-(208) is given in the state *feedback* form by

$$\begin{aligned} u^*(t, x^*(t), E(x^*(t))) &= (\rho(t) - \varsigma(t)) \frac{\psi_1(t)\left(x^*(t) - E(x^*(t))\right) + \psi_3(t) - \exp\{-\alpha t\}}{\psi_1(t)\left(\sigma^2(t) + \int_\Theta A^2(t,\theta)\, m(\mathrm{d}\theta)\right)}, \\ E(u^*(t, x^*(t), E(x^*(t)))) &= (\rho(t) - \varsigma(t)) \frac{\psi_3(t) - \exp\{-\alpha t\}}{\psi_1(t)\left(\sigma^2(t) + \int_\Theta A^2(t,\theta)\, m(\mathrm{d}\theta)\right)}, \\ \xi^*(t) &= \eta(t) + \int_0^t \boldsymbol{I}_{\{(w,s)\in\Omega\times[0,T]:\, M(s)+G\Phi^*(s)+\beta K^*(s)\leq 0\}}(s,w)\mathrm{d}s, \end{aligned}$$

where $\psi_1(\cdot)$ and $\psi_3(\cdot)$ are given by (271), (272) respectively.

Conclusion. In this work, mixed stochastic optimal control problem for McKean-Vlasov FBSDEJs has been formulated and discussed. Necessary and sufficient conditions of optimal control for systems governed by McKean-Vlasov FBSDEJs are proved by means of convex perturbation techniques for both continuous and singular parts. In our combined singular McKean-Vlasov control problem (209)-(210), there are two types of jumps for the state processes $(x^{u,\xi}(\cdot), y^{u,\xi}(\cdot), z^{u,\xi}(\cdot), r^{u,\xi}(\cdot,\cdot))$, the inaccessible ones which come from the Poisson martingale measure $N(\cdot,\cdot)$ and the predictable ones which come from the singular control part $\xi(\cdot)$. As an illustration, using these results, mean-variance portfolio selection problem: time-inconsistent solution has been discussed. An open question is to derive a general maximum principle for optimal singular control of McKean-Vlasov fully coupled forward-backward stochastic differential equations with random Poisson jumps.

Part VI
Appendix

The following result gives a case of the Itô formula for jump diffusions of mean-field type.

Lemma A1. Suppose that the processes $x_1(t)$ and $x_2(t)$ are given by: for $j = 1, 2$, $t \in [s, T]$:

$$\begin{aligned} dx_j(t) &= f\left(t, x_j(t), E(x_j(t)), u(t)\right) dt + \sigma\left(t, E(x_j(t)), u(t)\right) dW(t) \\ &\quad + \int_{\Theta} g\left(t, x_j(t^-), E(x_j(t)), u(t), \theta\right) N\left(d\theta, dt\right), \\ x_j(s) &= 0, \end{aligned}$$

then we get

$$\begin{aligned} E\left(x_1(T)x_2(T)\right) &= E\left[\int_s^T x_1(t)dx_2(t) + \int_s^T x_2(t)dx_1(t)\right] \\ &\quad +E\int_s^T \sigma^*\left(t, x_1(t), E(x_1(t)), u(t)\right)\sigma\left(t, x_2(t), E(x_2(t)), u(t)\right) dt \\ &\quad +E\int_s^T \int_{\Theta} g^*\left(t, x_1(t), E(x_1(t)), u(t), \theta\right) g\left(t, x_2(t), E(x_2(t)), u(t), \theta\right)\mu(d\theta)dt. \end{aligned}$$

Applying a similar method as in [16, Lemma 2.1], for the proof of the above Lemma.

Proposition A2. [9, Appendix] Let $\mathcal{G}$ be the predictable $\sigma-$field on $\Omega \times [0, T]$, and f be a $\mathcal{G} \times \mathcal{B}\left(\Theta\right)-$measurable function such that

$$E\left\{\int_0^T \int_{\Theta} \left|f\left(w, s, \theta\right)\right|^2 \mu\left(d\theta\right) ds\right\} < +\infty.$$

Then for all $p \geq 2$ there exists a positive constant $C = C\left(p, T, \mu\left(\Theta\right)\right)$ such that

$$E\left\{\sup_{t\in[0,T]}\left|\int_0^t \int_{\Theta} f\left(w, s, \theta\right) N\left(ds, d\theta\right)\right|^p\right\} \leq CE\left\{\int_0^T \int_{\Theta} \left|f\left(w, s, \theta\right)\right|^p \mu\left(d\theta\right) ds\right\}.$$

Part VII
References

References

[1] Ahmed NU.: Nonlinear diffusion governed by McKean-Vlasov equation on Hilbert space and optimal control. SIAM J. Control Optim. 46, (2007) 356–378.

[2] Alvarez L.: Singular stochastic control linear diffusion and optimal stopping: A class of solvable problems, SIAM J. Control Optim., 39 (2001) 1697-1710.

[3] Alvarez L. and T.A. Rakkolainen: On singular stochastic control and optimal stopping of spectrally negative jump diffusions, Stochastics An International Journal of Probability and Stochastics Processes 81(1) (2009) 55-78.

[4] An T.T K.: Combined optimal stopping and singular stochastic control, Stochastic Analysis and Applications, 28 (2010) 401-414.

[5] Andersson D, Djehiche B.: A maximum principle for SDEs of mean-field type, Appl Math Optim, **63**, 341-356 (2011)

[6] Bensoussan A, Lectures on stochastic control. In Lect. Notes in Math. 972, Springer-Verlag, 1-62 (1983).

[7] Bellman, R.: Dynamic programming, Princeton Univ. Press., (1957)

[8] Buckdahn R, Li J, Peng S.: (2009) Mean-field backward stochastic differential equations and related partial differential equations. Stochastic Processes and their Applications, Vol 119, pp. 3133-3154.

[9] Bouchard B., Elie R.: (2008) Discrete time approximation of decoupled forward-backward SDE with jumps, Stoch. Proc. Appl. Vol 118 No 1 pp. 53-75.

[10] Buckdahn R, Djehiche B, Li, J.: (2011) A general stochastic maximum principle for SDEs of mean-field type. Appl. Math. Optim. Vol 64, pp. 197-216.

[11] Borkar V.: Controlled diffusion processes, Probability Surveys, **2** (2005) 213-244.

[12] Cadenillas A.,and Haussmann U.: The stochastic maximum principle for singular control problem, Stochastics, Stochastics rep., 49, N 3-4, (1994) 211-237.

[13] Dufour F, and B. Miller: Maximum principle for singular stochastic control problem. SIAM J. Control Optim., 45(2), (2006) 668-698.

[14] Ekeland I.: On the variational principle, J. Math. Anal. Appl. , (**47**) 424-353, (1974).

[15] Elliott R.J., Li X. and Ni. Y.H.: (2013) Discrete time mean-field stochastic linear-quadratic optimal control problems. Automatica, Vol 49, No 11, pp. 3222-3233.

[16] Framstad N. C., ∅ksendal B., and Sulem A: (2004) Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance, J. Optim. Theory Appl., Vol 121 pp.77-98.

[17] Hafayed M, Abbas S.: On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality, J. Optim Theory Appl, **160**(3), 778-808 (2014).

[18] Hafayed M., Abba A and Abbas S.: On mean-field stochastic maximum principle for near-optimal controls for Poisson jump diffusion with applications, Int. J. Dynam. Control, **2**, 262–284 (2014).

[19] Hafayed M, and Abbas S.: Stochastic near-optimal singular controls for jump diffusions: necessary and sufficient conditions, Journal of Dynamical and Control Systems, **19**(4), 503-517 (2013).

[20] Hafayed M., Abbas S., and Veverka P.: On necessary and sufficient conditions for near-optimal singular stochastic controls. Optim. Lett., (**7**)5, 949-966, (2013).

[21] Hafayed M., Veverka P., and Abbas S.: On maximum principle of near-optimality for diffusions with jumps, with application to Consumption-investment problem, Differ. Equ. Dyn. Syst. , **20**(2), 111-125 (2012).

[22] M. Hafayed, S. Abbas: Stochastic near-optimal singular controls for jump diffusions: Necessary and sufficient conditions, Journal of Dynamical and Control Systems, 19(4) (2013) 503-517.

[23] Hafayed M.: (2013) A mean-field necessary and sufficient conditions for optimal singular stochastic control, Commun. Math. Stat. Vol 1, No 4, pp. 417–435.

[24] Hafayed M.: (2013). A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes, Int. J. Dynam. Control, Vol 1 No 4, pp. 300-315.

[25] Hafayed M.: (2014) Singular mean-field optimal control for forward-backward stochastic systems and applications to finance, Int. J. Dynam. Control, Vol 2 No 4, pp. 542–554

[26] Hafayed M, Abbas S.: (2013) A general maximum principle for stochastic differential equations of mean-field type with jump processes. Technical report, arXiv: 1301.7327v4.

[27] Hafayed M, Abbas, S., Abba A. : On mean-field partial information maximum principle of optimal control for stochastic systems with Lévy processes, J. Optim Theory Appl, 10.1007/s10957-015-0762-4, (2015).

[28] Hafayed M, Veverka P and Abbas A.: (2014) On Near-optimal Necessary and Sufficient Conditions for Forward-backward Stochastic Systems with Jumps, with Applications to Finance. Applications of Mathematics, Vol 59 No.4, pp. 407-440.

[29] Hafayed, M., Tabet. M., Boukaf S.: Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem, Commun. Math. Stat (3) (2015) 163–186.

[30] Hafayed, M. Ghebouli M., Boukaf S., Shi Y.: Partial information optimal control of mean-field forward-backward stochastic system driven by Teugels martingales with applications (2016) DOI 10.1016/j.neucom. 2016.03.002. Neurocomputing Vol **200** pages 11–21 (2016).

[31] Hafayed M., Boukaf M., Shi Y. Meherrem S.: A McKean-Vlasov optimal mixed regular-singular control problem, for nonlinear stochastic systems with Poisson jump processes (2016) Neurocomputing. Doi 10.1016/j.neucom.2015.11.082, Volume **182**, 19, pages 133-144 (2016).

[32] Hafayed, M.Abba A, Abbas S: On partial-information optimal singular control problem for mean-field stochastic differential equations driven byTeugels martingales measures, ID: 1079648 DOI:10.1080/00207179.2015.1079648 Internat. J. Control **89** (2016), no. 2, 397–410.

[33] Hafayed, M, Abba A Boukaf S: On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications,"International Journal of Modelling, Identification and Control".25 (1), 1-16, (2016).

[34] Haussman UG. and W. Suo: Singular optimal control I, II, SIAM J. Control Optim., 33(3), (1995) 916-936, 937-959.

[35] Hui E, Huang J, Li X, and Wang. G.: Near-optimal control for stochastic recursive problems, Syst. Cont. Letters, (**60**), 161-168 (2011).

[36] Huang J., Li X., Wang G.: (2010) Near-optimal control problems for linear forward-backward stochastic systems, Automatica. , Vol. 46 pp. 397-404.

[37] Kushner, H.J.: Optimal stochastic control, IRE Trans. Auto. Control, AC-7 (1962), 120-122.

[38] Kushner, H.J.:. On the stochastic maximum principle: Fixed time of control, J. Math. Anal. Appl., 11 (1965), 78-92.

[39] Kac M.: Foundations of kinetic theory, Proc. 3-rd Berkeley Sympos. Math. Statist. Prob. **3**: 171-197 (1956)

[40] Peng, S.: A general stochastic maximum principle for optimal control problems. SIAM J. Control Optim. 28(4), 966–979 (1990)

[41] Ni. Y.H, Zhang J.F., and Li X., (2014) Indefinite mean-field stochastic linear-quadratic optimal control, IEEE Transactions on automatic control, Doi: 10.1109/TAC.2014.2385253

[42] Li T., and Zhang J.F.: (2013) Adaptive mean field games for large population coupled ARX Systems with unknown coupling strength, Dyn Games Appl, Vol 3: pp. 489–507.

[43] Markowitz H.: (1952) Portfolio selection. J. Finance Vol. **7**, pp. 77-91.

[44] Menaldi J. and M. Robin: On singular stochastic control problems for diffusions with jumps, IEEE Transactions on Automatic Control 29(11), (1984) 991-1004.

[45] McKean H P.: A class of markov processes associated with nonlinear parabolic equations, Proc. Natl. Acad. Sci. USA, **56**, 1907-1911 (1966)

[46] Ni. Y.H., Zhang J.F. and Li X.: Indefinite mean-field stochastic linear-quadratic optimal Control, IEEE Trans. Autom. Control, DOI: 10.1109/TAC.2014.2385253 (2014)

[47] Pham H. On some recent aspects of stochastic control and their applications, Probability Surveys Vol. 2 (2005) 506-549.

[48] Meyer-Brandis T, $\emptyset$ksendal B, Zhou X.Y. A mean-field stochastic maximum principle via malliavin calculus. Stochastics. **84**, 643-666 (2012)

[49] Li, J.: Stochastic maximum principle in the mean-field controls. Automatica, **48**, 366-373 (2012)

[50] Yong J.: A linear-quadratic optimal control problem for mean-field stochastic differential equations. SIAM J. Control Optim., **54**(4), 2809-2838 (2013)

[51] Shen Y, Siu T K. The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem, Nonlinear Analysis, **86**, 58-73 (2013)

[52] Shen Y, Meng Q, Shi P.: Maximum principle for mean-field jump-diffusions to stochastic delay differential equations and its applicationt to finance, Automatica **50**, 1565-1579 (2014)

[53] Shi J, Wu Z.: Maximum principle for Forward-backward stochastic control system with random jumps and application to finance, J. Syst. Sci. Complex, **23**: 219-231 (2010)

[54] Shi J.: Necessary conditions for optimal control of forward-backward stochastic systems with random jumps. Int. J. Stoch. Anal. Article ID 258674, 50 pages (2012)

[55] Shi J.: Sufficient conditions of optimality for mean-field stochastic control problems. 12^{-th} International Conference on Control, Automation, Robotics & Vision Guangzhou, China, 5-7^{th} December, ICARCV 2012, 747-752 (2012)

[56] Wang B.C. and Zhang J.F.: (2012).Mean-field games for larg-population multiagent systems with Markov jump parameters, SIAM J. Control. Vol 50 No. 4, pp. 2308–2334.

[57] Tang M.: (2014) Stochastic maximum principle of near-optimal control of fully coupled forward-backward stochastic differential equation, Abstract and Applied Analysis Volume 2014, Article ID 361259, 12 pages

[58] Markowitz, H.: Portfolio selection, J. of Finance, **7**: 77-91 (1952)

[59] Shi P. and F. Li, A survey on Markovian jump systems: modeling and design, Int. J. of Control, Automation and Systems, 13(1) (2015) 1-16.

[60] Shi, P Y. Zhang and R. Agarwal, Stochastic finite-time state estimation for discrete time-delay neural networks with Markovian jumps, Neurocomputing, 151, (2015) 168-174.

[61] $\emptyset$ksendal, B and A. Sulem: Maximum principles for optimal control of forward-backward stochastic differential equations with jumps, SIAM J. Control Optim. 48 (2009), 2845-2976.

[62] Wang G, Zhang C, and Zhang W.: (2014). Stochastic maximum principle for mean-field type optimal control under partial information, IEEE Transactions on automatic control, Vol 59 No 2, pp. 522-528.

[63] Wu, Z and F. Zhang: Stochastic maximum principle for optimal control problems of forward-backward systems involving impulse controls, IEEE Transactions on Automatic Control, 56(6) (2011) 1401-1406.

[64] $\emptyset$ksendal B and A. Sulem: Applied stochastic control of jump diffusions, 2^{nd} ed., Springer-Verlag, Berlin (2007).

[65] $\emptyset$ksendal B, A. Sulem: Singular stochastic control and optimal stopping with partial information of Ito-Levy processes, SIAM J Control Optim, 50(4) (2012) 2254-2287.

[66] Xu R, and Wu T.: (2014) Mean-field backward stochastic evolution equations in Hilbert spaces and optimal control for BSPDEs, Abstract and Applied Analysis Vol 2014, Article ID 839467, 15 pages.

[67] Yong J.: A linear-quadratic optimal control problem for mean-field stochastic differential equations. SIAM J. Control Optim. Vol 51 No 4, 2809-2838. (2013)

[68] Yong J.: Optimality variational principle for controlled forward-backward stochastic differential equations with mixed intial-terminal conditions. SIAM J. Control Optim. **48**(6), 4119-4156 (2010)

[69] Yong J., and Zhou XY.: (1999) Stochastic Controls. Hamiltonian Systems and HJB Equations. Springer-Verlag. New York.

[70] Zhang L, Huang J, Li X.: Necessary condition for near optimal control of linear forward-backward stochastic differential equations. International Journal of Control, 1-29, (2015).

[71] Zhou, X.Y.: Stochastic near-optimal controls: Necessary and sufficient conditions for near-optimality, SIAM. Control. Optim. (**36**)-3, 929-947 (1998).

[72] Zhou X Y., Li D.: Continuous time mean-variance portfolio selection: a stochastic LQ framework. Appl. Math. Optim. **42**, 19-33 (2000)