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

Thèse présentée en vue de l'obtention du Diplôme de

Doctorat en Mathématiques

Option: **Probabilités**

Par

Moufida Tabet

Titre

Sur certain type du problème de contrôle optimal stochastique de type champ moyen et leur applications

Sous la direction de

Dr. Mokhtar Hafayed, MCA, Université de Biskra

Membres du Comité d'Examen

Dr. Zohir Mokhtari	MCA	Université de Biskra	**Président**
Dr. Mokhtar Hafayed	MCA	Université de Biskra	**Rapporteur**
Dr. Khalil Saadi	MCA	Université de M'sila	**Examinateur**
Dr. Badreddine Mansouri	MCA	Université de Biskra	**Examinateur**
Dr. Boulakhras Gherbal	MCA	Université de Biskra	**Examinateur**
Dr. Abdelmoumen Tiaiba	MCA	Université de M'sila	**Examinateur**

2016

Dédicace

Je dédie cette thèse

A mes très chers parents

qui veillent sans cesse sur moi avec leurs prières et leurs recommandations. Que Dieu le tout puissant les protège et leur réserve une longue et meilleure vie.

A mes très chers frères et sœurs.

A toute ma famille.

A mon époux et sa famille.

A mes chères amies.

Moufida.

REMERCIEMENTS

Mes remerciements vont premièrement à Dieu le tout puissant pour la volonté, la santé et la patience qu'il m'a donnée pour achever mon travail de recherche.

Je tiens à exprimer ma profonde gratitude, mes sincères et chaleureux remerciements à mon directeur de thèse, Monsieur le Dr. ***HAFAYED Mokhtar****, pour la confiance qu'il m'a accordée en acceptant d'encadrer ce travail doctoral, par ses multiples conseils. J'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de cette recherche.*

Je remercie vivement Monsieur le Dr. ***ZOHIR Mokhtari*** *qui m'a fait le grand honneur de présider le jury.*

Je voudrais également remercier tous les membres du jury, Monsieur le Dr. ***TIAIBA Abdelmoumen****, Monsieur le Dr.* ***SAADI Khalil****, Monsieur le Dr.* ***MANSOURI Badreddine*** *et Monsieur le Dr.* ***GHERBAL Boulakhras****, pour avoir accepté d'examiner et bien évaluer ma thèse.*

Je n'oublierai jamais l'aide, y compris financière, qui m'a été prodiguée par le Laboratoire de Mathématiques Appliquées (LMA) de Biskra.

Je ne peux pas oublier également de remercier ma famille et toute personne, de près ou de loin, qui ont contribué à la réalisation de ce travail de recherche.

Contents

Abstract

In this work, we are interesting in the necessary conditions of optimality of stochastic optimal control for systems governed by mean-field 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, mean-variance portfolio selection mixed with a recursive utility functional optimization problem is discussed.

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.

Résumé

Dans ce travail, nous intéressons aux conditions nécessaires d'optimalité en contrôle optimal stochastique pour des systèmes gouvernés par des équations différentielles stochastiques progressives rétrogrades de type champ-moyen avec sauts, où les coefficients dépendent de la loi marginale du processus de l'état par l'espérance de sa valeur. La variable de contrôle entre à la fois dans les coefficients de diffusion et de saut. De plus, la fonction du coût est aussi de type champ-moyen. Les conditions nécessaires d'optimalité pour ses systèmes seront établies sous la forme de principe du maximum par les techniques de perturbation convexe. Comme une application, la sélection de portefeuille moyenne-variance avec un problème d'optimisation fonctionnelle d'utilité récursive est discutée.

Mots clés: Equation différentielle stochastique progressive rétrograde de type champ-moyen avec sauts, contrôle optimal stochastique, principe du maximum de type champ-moyen, sélection du portefeuille moyenne-variance avec utilité récursive fonctionnelle.

Introduction

The two famous approach in solving control problems are the Bellman Dynamic Programming and Pontryagin's Maximum principles. The first method consists to find a solution of a stochastic partial differential equation (SPDE) which is not linear, verified by the value function. It is called Hamilton-Jacobi-Bellman (HJB) equation. We refer to [3] for more details about this method. The second method which will be the center of our interest in this work which consists to find an admissible control u^* that minimizes a cost functional subject to an stochastic differential equation on a finite time horizon. If u^* is some optimal control, we may ask how we can characterize it, in other words, what conditions must u^* necessarily satisfy? These conditions are called the stochastic maximum principle or the necessary conditions for optimality. The first version of the stochastic maximum principle was extensively established in the 1970's by Bismut [6], Kushner [28], Bensoussan [5] and Haussmann [18].

In this work, our main goal is to derive a maximum principle for optimal stochastic control of mean-field forward-backward stochastic differential equations with Poisson jump processes (FBSDEJs) of the form

$$\begin{cases} dx(t) &= f(t, x(t), E(x(t)), u(t))dt + \sigma(t, x(t), E(x(t)), u(t))dW(t) \\ &\quad + \int_\Theta c(t, x(t-), E(x(t-)), u(t), \theta)N(d\theta, dt), \\ dy(t) &= -\int_\Theta g(t, x(t), E(x(t)), y(t), E(y(t)), z(t), E(z(t)), r(t,\theta), u(t))\mu(d\theta)\, dt \\ &\quad + z(t)dW(t) + \int_\Theta r(t,\theta)N(d\theta, dt), \\ x(0) &= \zeta, y(T) = h(x(T), E(x(T))), \end{cases} \tag{1}$$

where f, σ, c, g et h are given maps and the initial condition ζ is an $\mathcal{F}_0-$measurable ran-

dom variable, $W = (W(t))_{t\in[0,T]}$ is a one-dimensional standard Brownian motion defined on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ satisfying the usual conditions and $N(d\theta, dt)$ is a Poisson martingale measure with local characteristics $\mu(d\theta)dt$.
The control variable $u(\cdot) = u(t)_{t\in[0,T]}$ is called admissible control, it is an $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted process and square-integrable with values in a nonempty convex subset $\mathcal{A}$ of $\mathbb{R}$. We denote by $\mathcal{U}([0,T])$ the set of all admissible controls.
Our stochastic optimal control problem is to minimize over the class of admissible control the following cost function:

$$\begin{aligned} J(u(\cdot)) &= E[\int_0^T \int_\Theta l(t, x(t), E(x(t)), y(t), E(y(t)), z(t), E(z(t)), \\ &\quad r(t,\theta), u(t))\mu(d\theta)\, dt + \phi(x(T), E(x(T))) + \varphi(y(0), E(y(0)))], \end{aligned} \tag{2}$$

where l, ϕ and φ is an appropriate functions. This cost functional is also of mean-field type, as the functions l, ϕ and φ depend on the marginal law of the state process through its expected value.
Any admissible control $u^*(\cdot) \in \mathcal{U}([0,T])$ satisfying

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

is called an optimal control.

The stochastic maximum principle of optimality for systems governed by forward-backward stochastic differential equations (FBSDEs) has been studied by many authors, Peng [41] firstly studied one kind of forward-backward stochastic control system which had the economic background and could be used to study the recursive utility problem in the mathematical finance. He obtained the maximum principle for this kind of control system with the control domain being convex. The difficulty to get the stochastic maximum principle for the control problems for systems governed by a forward and backward SDE for controlled diffusion and non convex control domain is how to use spike variation method for the variational equations with enough higher estimate order and use the duality technique for the adjoint equation. Xu [57] studied the non convex control domain case and obtained the corresponding maximum principle. But he assumed that the diffusion coefficient in the forward control system does not contain the control variable. Wu [55] first obtained stochastic maximum principle for fully coupled forward-backward stochastic control system in local form. Shi and Wu [50] extended this result to the global form while the diffusion coefficient doesn't contain the control variable. A good account and an extensive list of references on stochastic maximum principle for FBSDEs with applications can be found in Ma and Yong [33].

The mean-field stochastic differential equation was introduced by Kac [27] in 1956 as a stochastic model for the Vlasov kinetic equation of plasma and the study of this model was initiated by McKean [34] in 1966. Since then, many authors made contributions on mean-field stochastic problems and their applications. The mean-field games for large population multi-agent systems with Markov jump parameters have been investigated in Wang and Zhang [53]. Decentralized tracking-type games for large population multi-agent systems with mean-field coupling have been studied in Li and Zhang [30]. Discrete-time indefinite mean field linear-quadratic optimal control problem has been investigated in Ni et al. [36]. Discrete time mean-field stochastic linear-quadratic optimal control problems with applications have been derived by Elliott et al. [11]. In Buckdahn, Li, and Peng [9], 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 et al. [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 [47]. In Meyer-Brandis, Øksendal, and Zhou [35], a stochastic maximum principle of optimality for systems governed by controlled Itô-Levy process of mean-field type was proved using Malliavin calculus. Mean-field singular stochastic control problems have been investigated in Hafayed and Abbas [19]. More interestingly, mean-field type stochastic maximum principle for optimal singular control has been studied in Hafayed [20], in which convex perturbations used for both absolutely continuous and singular components. The maximum principle for optimal control of mean-field FBSDEJs with uncontrolled diffusion has been studied in Hafayed [21]. The necessary and sufficient conditions for near-optimality of mean-field jump diffusions with applications have been derived by Hafayed et al. [22]. Singular optimal control for mean-field forward-backward stochastic systems and applications to finance have been investigated in Hafayed [23]. Second-order necessary conditions for optimal control of mean-field jump diffusion have been obtained by Hafayed and Abbas [24]. Under partial information, mean-field type stochastic maximum principle for optimal control has been investigated in Wang, Zhang, and Zhang [54]. Under the condition that the control domain is convex, Andersson and Djehiche [1] and Li [29] 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 [58] and Shi [47]. The mean-field stochastic maximum principle for jump diffusions with applications has been investigated in Shen and Siu [45] 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 [46]. Mean-field optimal control for backward stochastic evolution equations in Hilbert spaces has been investigated in Xu and Wu [56].

The optimal control problems for stochastic systems described by Brownian motions and Poisson jumps have been investigated by many authors including [48], [49], [8], [26], [17], [37]. The necessary and sufficient conditions of optimality for FBSDEJs were obtained by Shi and Wu [48]. General maximum principle for fully coupled FBSDEJs has been obtained in Shi [49], where the author generalized Yong's maximum principle [59] to jump case.

This thesis is organized as follows:

- ▶ **Chapter 1:** This introductory chapter, we give a short introduction to stochastic control problems.

- ▶ **Chapter 2:** (Maximum principle for forward-backward stochastic control system with jumps and application to finance). In this chapter we present the maximum principle for systems governed by the forward-backward stochastic differential equations with jumps (FBSDEJs for short) in which the control domain is convex. This result was obtained by Shi and Wu [48].

- ▶ **Chapter 3:** (Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem). This chapter contains the main result of this thesis, 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 coeficients of the state equation depends 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. Our main goal is to derive necessary conditions for optimality satisfied by some optimal control which are also known as the stochastic maximum principle. The proof of our result is based on convex perturbation method. These necessary conditions are described in terms of two adjoint processes, corresponding to the mean-field forward and backward components with jumps and a maximum condition 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 travail, we only study the one-dimensional case.

Chapter 1

Introduction to stochastic Control Problems

Chapter 1

Introduction to stochastic control problems

This chapter will be organized as follows. In section 1, we give the optimal control theory. In section 2, we present strong and weak formulations of stochastic optimal control problems. Section 3 and 4 are concerned to the presentation of the two important methods which are dynamic programming and the maximum principle.

1.1 Optimal control theory

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. The two main methods are dynamic programming and the maximum principle.

1.2 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.

1.2.1 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 :

$$\begin{cases} dx(t) &= f(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dW(t), \\ x(0) &= x_0 \in \mathbb{R}^n, \end{cases} \tag{1.1}$$

where

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

and $x(\cdot)$ is the variable of state.
The function $u(\cdot)$ is called the control representing the action of the decision-makers (controllers). 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), which can be expressed in mathematical term as "$u(\cdot)$ is $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted".
The control $u(\cdot)$ is an element of the set:

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

We introduce the cost functional as follows:

$$J(u(\cdot)) \doteq E\left[\int_0^T l(t,x(t),u(t))dt + g(x(T))\right], \tag{1.2}$$

where

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

Definition 1.2.1 *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 s-admissible control, and $(x(\cdot),u(\cdot))$ an s-admissible pair, if*

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

ii) $x(\cdot)$ *is the unique solution of equation (1.1);*

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

The set of all s-admissible controls is denoted by $\mathcal{U}^s([0,T])$.
Our stochastic optimal control problem under strong formulation can be stated as follows:

Problem 1.2.1 *Minimize (1.2) over* $\mathcal{U}^s([0,T])$.

The goal is to find $u^*(\cdot) \in \mathcal{U}^s([0,T])$, such that

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}^s([0,T])} J(u(\cdot)). \tag{1.3}$$

For any $u^*(\cdot) \in \mathcal{U}^s([0,T])$ satisfying (1.3) is called an s-optimal control. The corresponding state process $x^*(\cdot)$ and the state-control pair $(x^*(\cdot),u^*(\cdot))$ are called an s-optimal state process and an s-optimal pair, respectively.

1.2.2 Weak formulation

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

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

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

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

iv) $x(\cdot)$ *is the unique solution of equation (1.1);*

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

The set of all w-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 1.2.2 *Minimize (1.2) over* $\mathcal{U}^w([0,T])$.

Namely, one seeks $\pi^*(\cdot) \in \mathcal{U}^w([0,T])$ such that

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

1.3 The Dynamic Programming Principle

In this section, we study an approach to solving optimal control problems, namely, the method of dynamic programming. Dynamic programming, originated by R. Bellman [3] in the early 1950's, 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).

1.3.1 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

$$dx(s) \quad = \quad f(s, x(s), u(s))ds + \sigma(s, x(s), u(s))dW(s). \tag{1.4}$$

The control $u = u(s)_{0\leq s\leq T}$ 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 $\mathcal{U}\left([t,T]\right)$ the set of control processes u.
To ensure the existence of a solution to SDE (1.4), the Borelian functions $f : [0,T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^{n\times d}$ 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.
We define the gain function as follows:

$$J(t,x,u) \quad = \quad E\left[\int_t^T l(s, x(s), u(s))ds + g(x\,(T))\right], \tag{1.5}$$

where

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

$$g : \mathbb{R}^n \longrightarrow \mathbb{R},$$

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.

Our objective is to maximize this gain function, we introduce the so-called value function:

$$V(t,x) \;=\; \sup_{u\in\mathcal{U}([t,T])} J(t,x,u), \tag{1.6}$$

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

$$V(t,x)=\sup_{u\in\mathcal{U}([t,T])} E\left[\int_t^{t+h} l(s,x(s),u(s))dt+V(t+h,x(t+h))\right],\ \textit{for } t\leq t+h\leq T. \tag{1.7}$$

The proof of the dynamic programming principle is technical and has been studied by different methods, we refer the reader to Lions [31], Fleming and Soner [15] and Yong and Zhou [60].

1.3.2 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 (1.7). The classical HJB equation associated to the stochastic control problem (1.6) is

$$-V_t(t,x)-\sup_{u\in U}\left[\mathcal{L}^u V(t,x)+l(t,x,u)\right] \;=\; 0,\ \text{on } [0,T]\times\mathbb{R}^n, \tag{1.8}$$

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

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

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

$$-V_t(t,x) - H(t,x,D_xV(t,x),D_x^2V(t,x)) \;=\; 0, \quad \forall (t,x) \in [0,T]\times\mathbb{R}^n, \tag{1.9}$$

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) :

$$H(t,x,\Psi,Q) \;=\; \sup_{u\in U}\left[f(t,x,u).\Psi + \frac{1}{2}tr\left(\sigma\sigma^\intercal\,(t,x,u)\,Q\right) + l(t,x,u)\right]. \tag{1.10}$$

The function H is sometimes called Hamiltonian of the associated control problem, and the PDE (1.8) or (1.9) 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.

1.3.3 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 1.3.2 *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:*

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

together with the terminal condition

$$W(T,\cdot) = g \text{ on } \mathbb{R}^n,$$

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 may be found in any textbook on stochastic control see. e.g. [14], [15], [60] or [42].

1.4 The Pontryagin stochastic maximum principle

Another classical approach for optimization and control problems is to derive necessary conditions satisfied by an optimal solution. The argument is to use an appropriate calculus of variations of the cost functional $J(u)$ with respect to the control variable in order to derive a necessary condition of optimality. The maximum principle initiated by Pontryagin, states that an optimal state trajectory must solve a Hamilton system together with a maximum condition of a function called a generalized Hamilton. In principle, solve a Hamilton should be easier than solving the original control problem.

The original version of Pontryagin's maximum principle was derived for deterministic problems in the 1950's and 1960's by Pontryagin and al. [43], as in classical calculus of variation. The basic idea is to perturb 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 is expressed in terms of an adjoint variable. More recent results for the study of optimal control in the deterministic cases were treated by Fleming [14] and [15], where the authors present the results of fundamental control theory. The first version of the stochastic maximum principle was extensively established in the 1970's by Bismut [6], Kushner [28], Bensoussan [5] and Haussmann [18]. However, at that time, the results were essentially obtained under the condition that there is no control on the diffusion coefficient. The first version of the stochastic maximum principle when the diffusion coefficient depends explicitly on the control variable and the control domain is not convex, was obtained by Peng [40], in which he studied the second order term in the Taylor expansion of the perturbation method arising from the Itô integral. He then obtained a maximum principle for control dependent diffusion, which involves in addition to the first-order adjoint process, a second-order adjoint process. The adjoint processes are described by what is called adjoint equation. In fact, the adjoint equation is in general a linear backward stochastic differential equation (BSDE) with a specified a random terminal condition on the state. Unlike a forward stochastic differential equation, the solution of a BSDE is a pair of adapted solutions. The linear BSDE was first proposed by Bismut [7] in 1973. Pardoux and Peng [39] got the uniqueness and existence theorem for the solutions of nonlinear BSDE driven by Brownian motion under Lipschitz condition in 1990. Now BSDE theory is playing a key role not only in dealing with stochastic optimal control problems, but also in mathematical finance, particularly in hedging and nonlinear pricing theory for imperfect market.

1.4.1 The maximum principle

As an illustration, we provide a sketch of how the maximum principle for a deterministic control problem is derived. In this setting, the state of the system is given by the differential equation

$$\begin{cases} dx(t) & = & f(t, x(t), u(t))dt, t \in [0, T], \\ x(0) & = & x_0, \end{cases} \tag{1.11}$$

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

$$J(u(\cdot)) = \int_0^T l(t, x(t), u(t)) + g(x(T)), \tag{1.12}$$

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.

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

We denote by $x^*(t)$ the solution to (1.11) 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:

$$u^{\varepsilon}(t) = \begin{cases} v & \text{for } \tau - \varepsilon \leq t \leq \tau, \\ u^*(t) & \text{otherwise.} \end{cases} \tag{1.13}$$

We denote by $x^{\varepsilon}(t)$ the solution to (1.11) with the control $u^{\varepsilon}(t)$. We set that $x^*(t)$ and $x^{\varepsilon}(t)$ are equal up to $t = \tau - \varepsilon$ and that

$$\begin{aligned} x^{\varepsilon}(\tau) - x^*(\tau) &= (f(\tau, x^{\varepsilon}(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)))\varepsilon + o(\varepsilon) \\ &= (f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)))\varepsilon + o(\varepsilon), \end{aligned} \tag{1.14}$$

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

$$x^{\varepsilon}(t) \quad = \quad x^*(t) + z(t)\varepsilon + o(\varepsilon). \tag{1.15}$$

Then, by (1.14)

$$z(\tau) = f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)). \tag{1.16}$$

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

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

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. $\Psi(t)z(t) = \text{constant}$. By the terminal condition for the adjoint equation we have

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

In particular, by (1.16)

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

Since τ was chosen arbitrarily, this is equivalent to

$$\Psi(t)f(t,x^*(t),u^*(t)) = \inf_v \Psi(t)f(t,x^*(t),v), \text{ for all } 0 \leq t \leq T.$$

This specifies a necessary condition for $u^*(t)$ to be optimal when $l = 0$.
To account for the running cost l one can construct an extra state

$$dx^0(t) = l(t,x(t),u(t))dt,$$

which allows us to write the cost function in terms of two terminal costs:

$$J(u(\cdot)) = x^0(T) + g(x(T)).$$

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

$$H(t,x^*(t),u^*(t),\Psi(t)) = \inf_v H(t,x^*(t),v,\Psi(t)) \text{ for all } 0 \leq t \leq T, \tag{1.17}$$

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

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

and the adjoint equation is given by

$$\begin{cases} d\Psi(t) = -(l_x(t,x^*(t),u^*(t)) + f_x(t,x^*(t),u^*(t))\Psi(t))dt, \\ \Psi(T) = g_x(x^*(T)). \end{cases} \tag{1.18}$$

The minimum condition (1.17) together with the adjoint equation (1.18) specifies the Hamiltonian system for our control problem.

1.4.2 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:

$$dx(t) = f(t, x(t), u(t))dt + \sigma(t, x(t))dW(t), t \in [0, T], \tag{1.19}$$

where f and σ are deterministic functions and the last term is an Itô integral with respect to a Brownian motion $W(\cdot)$ 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 have an explicit dependence on the control:

$$dx(t) = f(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), t \in [0, T]. \tag{1.20}$$

The cost function for the stochastic case is the expected value of the cost function (1.12), 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 (1.19) the adjoint equation is given by

$$\begin{cases} d\Psi(t) = -(l_x(t, x^*(t), u^*(t)) + f_x(t, x^*(t), u^*(t))\Psi(t) \\ \qquad +\sigma_x(t, x^*(t))Q(t))dt + Q(t)dW(t), \\ \Psi(T) = g_x(x^*(T)). \end{cases} \tag{1.21}$$

A solution to this kind of backward SDE is a pair $(\Psi(t), Q(t))$ which fulfills (1.21).
The Hamiltonian is

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

and the maximum principle reads

$$H(t, x^*(t), u^*(t), \Psi(t), Q(t)) = \inf_u H(t, x^*(t), u, \Psi(t), Q(t)) \text{ for all } 0 \le t \le T, \mathbb{P} - \text{a.s.} \tag{1.22}$$

There is also third case: if the state is given by (1.20) but the action space $\mathcal{A}$ is 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 [4].

The necessary condition for optimality is then the following

$$H_u(t, x^*(t), u^*(t), \Psi^*(t), Q^*(t))\,(u - u^*(t)) \geq 0 \text{ for all } 0 \leq t \leq T, \mathbb{P} - \text{a.s.}$$

Chapter 2

Maximum principle for forward-backward stochastic control system with jumps and application to finance

Chapter 2

Maximum principle for forward-backward stochastic control system with jumps and application to finance

2.1 Introduction

In this chapter we will give the maximum principle for systems governed by the forward-backward stochastic differential equations with Jumps (FBSDEJ for short) in which the control domain is convex, the control variable is allowed to enter both diffusion and jump coefficients. This result was obtained by Shi and Wu [48].

This chapter is organized as follow. In section 2, we consider some assumptions to achieve the goal. In section 3, we obtain the maximum principle for one kind of forward-backward stochastic system with jumps. In section 4, we prove that under some additional conditions, the maximum principle is also sufficient. In section 5, we apply this result to study a mean-variance portfolio selection mixed with a recursive utility functional optimization problem and give the explicit expression of the optimal portfolio selection strategy.

2.2 Problem formulation and assumptions

Let $T>0$ be a fixed time horizon and $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$ 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 $\{\mathcal{F}_t\}_{t\in[0,T]}-$Poisson point process independent of W. We denote by $\tilde{N}(d\theta,dt)$ the random counting measure induced by η, defined on $\Theta\times\mathbb{R}^+$, where Θ is a fixed nonempty subset of $\mathbb{R}^l$ with its Borel σ-finite measure on $(\Theta,\mathcal{B}(\Theta))$ with $\mu(d\theta)<\infty$. We then define

$$N(d\theta,dt):=\tilde{N}(d\theta,dt)-\mu(d\theta),$$

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

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

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

Throughout this chapter, we use the notations

1. $\langle\cdot,\cdot\rangle$ the Euclidean scalar product.
2. $|\cdot|$ the Euclidean norm on $\mathbb{R}^n,\forall n\in\mathbb{N}$.
3. $\top$ appearing in the superscripts denotes the transpose of a matrix.

Let U be a nonempty convex subset of $\mathbb{R}^k$.

We define the admissible control set

$$\mathcal{U}([0,T]) = \left\{u(\cdot)\in L^2_{\mathcal{F}}\left([0,T];\mathbb{R}^k\right);u(t)\in U, a.e.\ \ t\in[0,T],\mathbb{P}-a.s.\right\}.$$

For any given admissible control $u(\cdot)\in\mathcal{U}([0,T])$ and initial state $x_0\in\mathbb{R}^n$, we consider the following forward-backward stochastic control system with jumps:

$$\begin{cases} dx(t) & = & f(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dW(t) \\ & & +\int_{\Theta} c(t,x(t-),u(t),\theta)N\left(d\theta,dt\right), \\ -dy(t) & = & \int_{\Theta} g(t,x(t),y(t),z(t),r(t,\theta),u(t))\mu\left(d\theta\right)dt \\ & & -z(t)dW(t) - \int_{\Theta} r(t,\theta)N\left(d\theta,dt\right), \\ x(0) & = & x_0, y(T) = h(x(T)), \end{cases} \tag{2.1}$$

where

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

are given maps.

The stochastic optimal control problem is to minimize over $\mathcal{U}([0,T])$ the cost functional of the form

$$J(u(\cdot)) \doteq E\left[\int_0^T\int_{\Theta} l(t,x(t),y(t),z(t),r(t,\theta),u(t))\mu\left(d\theta\right)dt + \phi(x(T)) + \varphi(y(0))\right], \tag{2.2}$$

where

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

are given maps.

An admissible control which solves this problem is called an optimal control and it is denoted by u^*, i.e.

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

Assumptions (H2.1)

Throughout this chapter, we assume that the above functions satisfy the following assumptions:

1. f, σ, c are global Lipschitz in (x, u) and g is global Lipschitz in (x, y, z, r, u) ;

2. $f, \sigma, c, g, l, \phi, \varphi$ are continuously differentiable in their variables including (x, y, z, r, u) ;

3. $f_x, f_u, \sigma_x, \sigma_u, g_x, g_y, g_z, g_r, g_u$ and $\int_\Theta |c_x(\cdot,\cdot,\theta)|^2 \mu(d\theta)$, $\int_\Theta |c_u(\cdot,\cdot,\theta)|^2 \mu(d\theta)$ are bounded;

4. l_x, l_y, l_z, l_r, l_u are bounded by $C(1+|x|+|y|+|z|+|r|+|u|)$, ϕ_x and φ_y are bounded by $C(1+|x|)$, $C(1+|y|)$ respectively;

5. $\forall x \in \mathbb{R}^n, h(x) \in L^2(\Omega, \mathcal{F}_T; \mathbb{R}^m)$ and for fixed $w \in \Omega, h(x)$ continuously differentiable in x, h_x is bounded;

6. For all $t \in [0,T]$, $f(t,0,0), g(t,0,0,0,0,0) \in L^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$, $\sigma(t,0,0) \in L^2_{\mathcal{F}}\left([0,T];\mathbb{R}^{n\times d}\right)$ and $c(t,0,0,\cdot) \in \mathcal{M}^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$.

Under assumption (**H2.1**) the forward-backward equation (2.1) admits a unique solution $(x(\cdot), y(\cdot), z(\cdot), r(\cdot,\cdot)) \in L^2_{\mathcal{F}}([0,T];\mathbb{R}^n) \times L^2_{\mathcal{F}}([0,T];\mathbb{R}^m) \times L^2_{\mathcal{F}}\left([0,T];\mathbb{R}^{m\times d}\right) \times \mathcal{M}^2_{\mathcal{F}}([0,T];\mathbb{R}^m)$.

2.3 Necessary conditions for optimal control of FBSDEJs

The goal in this section is to get the maximum principle and to find the necessary condition of optimality verified by some optimal control, for this we use the classic convex variation method. We assume the existence of an optimal control $u^*(\cdot)$ minimizing the cost J over $\mathcal{U}([0,T])$ and $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot,\cdot))$ denotes the optimal trajectory, that is, the solution of (2.1) corresponding to $u^*(\cdot)$.

Since U is convex, then for any $0 \leq \varepsilon \leq 1$, we define an admissible control $u^\varepsilon(\cdot)$ by the following perturbation

$$u^\varepsilon(\cdot) \doteq u^*(\cdot) + \varepsilon u(\cdot),$$

which is also a admissible control has value U.

We denote by $(x^\varepsilon(\cdot), y^\varepsilon(\cdot), z^\varepsilon(\cdot), r^\varepsilon(\cdot,\cdot))$ the trajectory corresponding to $u^\varepsilon(\cdot)$.

2.3.1 Variational equations and variational inequality

Let $(x_1^\varepsilon(\cdot), y_1^\varepsilon(\cdot), z_1^\varepsilon(\cdot), r_1^\varepsilon(\cdot,\cdot))$ be the solution of the following variational equation:

$$\begin{cases} dx_1^\varepsilon(t) &= [f_x(t)x_1^\varepsilon(t) + f_u(t)u(t)]\,dt + [\sigma_x(t)x_1^\varepsilon(t) + \sigma_u(t)u(t)]\,dW(t) \\ & \quad + \int_\Theta [c_x(t,\theta)x_1^\varepsilon(t-) + c_u(t,\theta)u(t)]\,N(d\theta, dt)\,, \\ -dy_1^\varepsilon(t) &= \int_\Theta [g_x(t,\theta)x_1^\varepsilon(t) + g_y(t,\theta)y_1^\varepsilon(t) + g_z(t,\theta)z_1^\varepsilon(t) \\ & \quad + g_r(t,\theta)r_1^\varepsilon(t,\theta) + g_u(t,\theta)u(t)]\mu(d\theta)\,dt \\ & \quad - z_1^\varepsilon(t)dW(t) - \int_\Theta r_1^\varepsilon(t,\theta)N(d\theta, dt)\,, \\ x_1^\varepsilon(0) &= 0, y_1^\varepsilon(T) = h_x(x^*(T))x_1^\varepsilon(T), \end{cases} \tag{2.3}$$

where and in the sequel

$$f(t) \equiv f(t, x^*(t), u^*(t)), \qquad \sigma(t) \equiv \sigma(t, x^*(t), u^*(t)),$$
$$c(t,\theta) \equiv c(t, x^*(t), u^*(t), \theta), \qquad l(t,\theta) \equiv l(t, x^*(t), y^*(t), z^*(t), r^*(t,\theta), u^*(t)),$$
$$g(t,\theta) \equiv g(t, x^*(t), y^*(t), z^*(t), r^*(t,\theta), u^*(t)).$$

and their derivatives.

Under (**H2.1**), there exists a unique solution

$$\begin{aligned} &(x_1^\varepsilon(\cdot), y_1^\varepsilon(\cdot), z_1^\varepsilon(\cdot), r_1^\varepsilon(\cdot,\cdot)) \\ &\in L^2_{\mathcal{F}}([0,T]\,;\mathbb{R}^n) \times L^2_{\mathcal{F}}([0,T]\,;\mathbb{R}^m) \times L^2_{\mathcal{F}}\left([0,T]\,;\mathbb{R}^{m\times d}\right) \times \mathcal{M}^2_{\mathcal{F}}([0,T]\,;\mathbb{R}^m)\,, \end{aligned}$$

satisfying (2.3).

The following lemma is needed to establish our result.

For all $t \in [0,T]\,, \varepsilon \geq 0$, we set

$$\begin{aligned} \hat{x}^\varepsilon(t) &\doteq \varepsilon^{-1}(x^\varepsilon(t) - x^*(t)) - x_1^\varepsilon(t), \\ \hat{y}^\varepsilon(t) &\doteq \varepsilon^{-1}(y^\varepsilon(t) - y^*(t)) - y_1^\varepsilon(t), \\ \hat{z}^\varepsilon(t) &\doteq \varepsilon^{-1}(z^\varepsilon(t) - z^*(t)) - z_1^\varepsilon(t), \\ \hat{r}^\varepsilon(t,\theta) &\doteq \varepsilon^{-1}(r^\varepsilon(t,\theta) - r^*(t,\theta)) - r_1^\varepsilon(t,\theta). \end{aligned} \tag{2.4}$$

Lemma 2.3.1 *Let assumption (**H2.1**) hold. Then*

$$\begin{aligned}
&\lim_{\varepsilon\longrightarrow 0}\sup_{0\leq t\leq T} E\left|\hat{x}^{\varepsilon}(t)\right|^{2}=0,\\
&\lim_{\varepsilon\longrightarrow 0}\sup_{0\leq t\leq T} E\left|\hat{y}^{\varepsilon}(t)\right|^{2}=0,\\
&\lim_{\varepsilon\longrightarrow 0} E\int_{0}^{T}\left|\hat{z}^{\varepsilon}(t)\right|^{2}dt=0,\\
&\lim_{\varepsilon\longrightarrow 0} E\int_{0}^{T}\int_{\Theta}\left|\hat{r}^{\varepsilon}(t,\theta)\right|^{2}\mu\left(d\theta\right)dt=0.
\end{aligned} \tag{2.5}$$

Proof After defining $x^{\varepsilon}(t), x^{*}(t)$ and $x_{1}^{\varepsilon}(t)$, we have

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

So we can well define $\hat{x}^{\varepsilon}(t)$ as

$$\hat{x}^{\varepsilon}(t) = \varepsilon^{-1}\left(x^{\varepsilon}(t)-x^{*}(t)\right)-x_{1}^{\varepsilon}(t).$$

Applying Itô's formula, we get

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

with $\hat{x}^{\rho}(0)=0$.

By replacing $x^{\varepsilon}(t)$ with $x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right)$ and $u^{\varepsilon}(t)$ with $u^*(t)+\varepsilon u(t)$, we find

$$\begin{aligned} d\hat{x}^{\varepsilon}(t) \;=\;& \varepsilon^{-1}\left[f(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\varepsilon u(t))-f(t)\right]dt \\ & -\left[f_x(t)x_1^{\varepsilon}(t)+f_u(t)u(t)\right]dt \\ & +\varepsilon^{-1}\left[\sigma(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\varepsilon u(t))-\sigma(t)\right]dW(t) \\ & -\left[\sigma_x(t)x_1^{\varepsilon}(t)+\sigma_u(t)u(t)\right]dW(t) \\ & +\varepsilon^{-1}\left[\int_{\Theta}c(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\varepsilon u(t),\theta)-c(t,\theta)\right]N\left(d\theta,dt\right) \\ & -\int_{\Theta}\left[c_x(t,\theta)x_1^{\varepsilon}(t-)+c_u(t,\theta)u(t)\right]N\left(d\theta,dt\right). \end{aligned} \tag{2.6}$$

By Taylor's expansion with a simple computation, we show that

$$\begin{aligned} & \varepsilon^{-1}\left[f(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\varepsilon u(t))-f(t)\right] \\ =\; & \int_0^1 f_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right)d\lambda \\ & +\int_0^1 f_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))u\left(t\right)d\lambda. \end{aligned} \tag{2.7}$$

$$\begin{aligned} & \varepsilon^{-1}\left[\sigma(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\varepsilon u(t))-\sigma(t)\right] \\ =\; & \int_0^1 \sigma_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right)d\lambda \\ & +\int_0^1 \sigma_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))u\left(t\right)d\lambda, \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} & \varepsilon^{-1}\left[c(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\varepsilon u(t),\theta)-c(t,\theta)\right] \\ =\; & \int_0^1 c_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t)),\theta)\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right)d\lambda \\ & +\int_0^1 c_u(t,x^*(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t)),\theta)u\left(t\right)d\lambda. \end{aligned} \tag{2.9}$$

We replace (2.7), (2.8) and (2.9) in (2.6), it becomes

$$\begin{aligned}
d\hat{x}^{\varepsilon}(t) &= \int_0^1 f_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))\hat{x}^{\varepsilon}(t)d\lambda dt \\
&+\int_0^1 \sigma_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))\hat{x}^{\varepsilon}(t)d\lambda dW(t) \\
&+\int_0^1 c_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t),\theta)\hat{x}^{\varepsilon}(t)d\lambda N\left(d\theta,dt\right) \\
&+\int_0^1 \left[f_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))-f_x(t)\right]x_1^{\varepsilon}(t)d\lambda dt \\
&+\int_0^1 \left[f_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))-f_u(t)\right]u(t)d\lambda dt \\
&+\int_0^1 \left[\sigma_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))-\sigma_x(t)\right]x_1^{\varepsilon}(t)d\lambda dW(t) \\
&+\int_0^1 \left[\sigma_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))-\sigma_u(t)\right]u(t)d\lambda dW(t) \\
&+\int_0^1 \left[c_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t),\theta)-c_x(t,\theta)\right]x_1^{\varepsilon}(t)d\lambda N\left(d\theta,dt\right) \\
&+\int_0^1 \left[c_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t),\theta)-c_u(t,\theta)\right]u(t)d\lambda N\left(d\theta,dt\right).
\end{aligned}$$

So, we obtain

$$\begin{cases}
d\hat{x}^{\varepsilon}(t) = \left[A^{\varepsilon}(t)\hat{x}^{\varepsilon}(t)+G^{1\varepsilon}(t)\right]dt+\left[B^{\varepsilon}(t)\hat{x}^{\varepsilon}(t)+G^{2\varepsilon}(t)\right]dW(t) \\
\qquad +\int_{\Theta}\left[C^{\varepsilon}(t-,\theta)\hat{x}^{\varepsilon}(t-)+G^{3\varepsilon}(t-,\theta)\right]N\left(d\theta,dt\right), \\
\hat{x}^{\varepsilon}(0) = 0,
\end{cases}$$

where

$$\begin{aligned}
A^{\varepsilon}(t) &\doteq \int_0^1 \left[f_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t))\right]d\lambda, \\
G^{1\varepsilon}(t) &\doteq \left[A^{\varepsilon}(t)-f_x(t)\right]x_1^{\varepsilon}(t)+\int_0^1\left[f_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),\right. \\
&\quad \left. u^*(t)+\lambda\varepsilon u(t))-f_u(t)\right]u\left(t\right)d\lambda, \\
B^{\varepsilon}(t) &\doteq \int_0^1 \left[\sigma_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),u^*(t)+\lambda\varepsilon u(t)\right]d\lambda,
\end{aligned}$$

$$
\begin{aligned}
G^{2\varepsilon}(t) &\doteq [B^\varepsilon(t)-\sigma_x(t)]\,x_1^\varepsilon(t)+\int_0^1[\sigma_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),\\
&\quad u^*(t)+\lambda\varepsilon u(t))-\sigma_u(t)]u\left(t\right)d\lambda,\\
C^\varepsilon(t,\theta) &\doteq \int_0^1 c_x[(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),u^*(t)+\lambda\varepsilon u(t),\theta)]d\lambda,\\
G^{3\varepsilon}(t,\theta) &\doteq [C^\varepsilon(t,\theta)-c_x(t,\theta)]\,x_1^\varepsilon(t)+\int_0^1[c_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),\\
&\quad u^*(t)+\lambda\varepsilon u(t),\theta)-c_u(t,\theta)]u\left(t\right)d\lambda.
\end{aligned}
$$

Applying Itô's formula to $|\hat{x}^\varepsilon(t)|^2$, noting that assumption (**H2.1**), we have

$$
\begin{aligned}
E\left|\hat{x}^\varepsilon(t)\right|^2 &= E\int_0^T[\langle 2\hat{x}^\varepsilon(t),A^\varepsilon(s)\hat{x}^\varepsilon(t)+G^{1\varepsilon}(t)\rangle]\,dt+E\int_0^T|B^\varepsilon(t)\hat{x}^\varepsilon(s)+G^{2\varepsilon}(t)|^2]dt\\
&\quad+\int_0^T\int_\Theta\left|C^\varepsilon(t,\theta)\hat{x}^\varepsilon(t)+G^{3\varepsilon}(t,\theta)\right|^2\mu(d\theta)dt\\
&\leq CE\int_0^T\left|\hat{x}^\varepsilon(t)\right|^2dt+o(\rho).
\end{aligned}
$$

Then we can to get the first convergence result of (2.5) from Gronwall's inequality. Afterwards, we will prove the last three results, first we have:

$$
\begin{cases}
-dy^\varepsilon(t) = \displaystyle\int_\Theta g(t,x^\varepsilon(t),y^\varepsilon(t),z^\varepsilon(t),r^\varepsilon(t,\theta),u^\varepsilon(t))\mu\left(d\theta\right)dt-z^\varepsilon(t)dW(t)\\
\qquad\qquad\displaystyle-\int_\Theta r^\varepsilon(t,\theta)N\left(d\theta,dt\right),\\
-dy^*(t) = \displaystyle\int_\Theta g(t,x^*(t),y^*(t),z^*(t),r^*(t,\theta),u(t))\mu\left(d\theta\right)dt-z^*(t)dW(t)\\
\qquad\qquad\displaystyle-\int_\Theta r^*(t,\theta)N\left(d\theta,dt\right),\\
-dy_1^\varepsilon(t) = \displaystyle\int_\Theta[g_x(t,\theta)x_1^\varepsilon(t)+g_y(t,\theta)y_1^\varepsilon(t)+g_z(t,\theta)z_1^\varepsilon(t)+g_r(t,\theta)r_1^\varepsilon(t,\theta)\\
\qquad\qquad\displaystyle+g_u(t,\theta)u(t)]\mu\left(d\theta\right)dt-z_1^\varepsilon(t)dW(t)-\int_\Theta r_1^\varepsilon(t,\theta)N\left(d\theta,dt\right).
\end{cases}
$$

So we can define $\hat{y}^\varepsilon(t)$ as

$$
\hat{y}^\varepsilon(t) = \varepsilon^{-1}\left(y^\varepsilon(t)-y^*(t)\right)-y_1^\varepsilon(t),
$$

and from Itô's formula, we get

$$\begin{aligned}
-d\hat{y}^{\varepsilon}(t) &= \varepsilon^{-1}\int_{\Theta}\left[g(t,x^{\varepsilon}(t),y^{\varepsilon}(t),z^{\varepsilon}(t),r^{\varepsilon}(t,\theta),u^{\varepsilon}(t))-g(t,\theta)\right]\mu\left(d\theta\right)dt \\
&\quad -\varepsilon^{-1}\left[z^{\varepsilon}(t)-z^{*}(t)\right]dW(t)-\varepsilon^{-1}\int_{\Theta}\left[r^{\varepsilon}(t,\theta)-r^{*}(t,\theta)\right]N\left(d\theta,dt\right) \\
&\quad -\int_{\Theta}[g_x(t,\theta)x_1^{\varepsilon}(t)+g_y(t,\theta)y_1^{\varepsilon}(t)+g_z(t,\theta)z_1^{\varepsilon}(t)+g_r(t,\theta)r_1^{\varepsilon}(t,\theta) \\
&\quad +g_u(t,\theta)u(t)]\mu\left(d\theta\right)dt+z_1^{\varepsilon}(t)dW(t)+\int_{\Theta}r_1^{\varepsilon}(t,\theta)N\left(d\theta,dt\right).
\end{aligned}$$

Using the fact that

$$\begin{aligned}
x^{\varepsilon}(t) &= x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right), \\
y^{\varepsilon}(t) &= y^{*}(t)+\varepsilon\left(\hat{y}^{\varepsilon}(t)+y_1^{\varepsilon}(t)\right), \\
z^{\varepsilon}(t) &= z^{*}(t)+\varepsilon\left(\hat{z}^{\varepsilon}(t)+z_1^{\varepsilon}(t)\right) \\
r^{\varepsilon}(t,\theta) &= r^{*}(t,\theta)+\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta)+r_1^{\varepsilon}(t,\theta)\right) \\
u^{\varepsilon}(t) &= u^{*}(t)+\varepsilon u(t),
\end{aligned}$$

we get

$$\begin{aligned}
-d\hat{y}^{\varepsilon}(t) &= \varepsilon^{-1}\int_{\Theta}[g(t,x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_1^{\varepsilon}(t)\right),y^{*}(t)+\varepsilon\left(\hat{y}^{\varepsilon}(t)+y_1^{\varepsilon}(t)\right), \\
&\quad z^{*}(t)+\varepsilon\left(\hat{z}^{\varepsilon}(t)+z_1^{\varepsilon}(t)\right),r^{*}(t,\theta)+\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta)+r_1^{\varepsilon}(t,\theta)\right),u^{*}(t)+\varepsilon u(t)) \\
&\quad -g(t,\theta)]\mu\left(d\theta\right)dt-\varepsilon^{-1}\left[z^{\varepsilon}(t)-z^{*}(t)\right]dW(t)-\varepsilon^{-1}\int_{\Theta}\left[r^{\varepsilon}(t,\theta)-r^{*}(t,\theta)\right]N\left(d\theta,dt\right) \\
&\quad -\int_{\Theta}[g_x(t,\theta)x_1^{\varepsilon}(t)+g_y(t,\theta)y_1^{\varepsilon}(t)+g_z(t,\theta)z_1^{\varepsilon}(t)+g_r(t,\theta)r_1^{\varepsilon}(t,\theta) \\
&\quad +g_u(t,\theta)u(t)]\mu\left(d\theta\right)dt+z_1^{\varepsilon}(t)dW(t)+\int_{\Theta}r_1^{\varepsilon}(t,\theta)N\left(d\theta,dt\right).
\end{aligned} \tag{2.10}$$

By Taylor's expansion with a simple computation, we show that

$$\begin{aligned}
&\varepsilon^{-1}[g(t,x^*(t)+\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\varepsilon\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)+\varepsilon\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),\\
&r^*(t,\theta)+\varepsilon\left(\hat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\varepsilon u(t))-g(t,\theta)]\\
&=\int_0^1[g_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\lambda\varepsilon\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)+\lambda\varepsilon\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),\\
&r^*(t,\theta)+\lambda\varepsilon\left(\hat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\lambda\varepsilon u(t))\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right)d\lambda\\
&+\int_0^1 g_y(t,x^*(t)+\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\lambda\varepsilon\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)+\lambda\varepsilon\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),\\
&r^*(t,\theta)+\lambda\varepsilon\left(\hat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\lambda\varepsilon u(t))\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right)d\lambda\\
&+\int_0^1 g_z(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\lambda\varepsilon\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)+\lambda\varepsilon\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),\\
&r^*(t,\theta)+\lambda\varepsilon\left(\hat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\lambda\varepsilon u(t))\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right)d\lambda\\
&+\int_0^1 g_r(t,x^*(t)+\lambda\varepsilon\left(\tilde{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\lambda\varepsilon\left(\tilde{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)+\lambda\varepsilon\left(\tilde{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),\\
&r^*(t,\theta)+\lambda\varepsilon\left(\tilde{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\lambda\varepsilon u(t))\left(\tilde{r}^\varepsilon(t)+r_1^\varepsilon(t)\right)d\lambda\\
&+\int_0^1 g_u(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\lambda\varepsilon\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)+\lambda\varepsilon\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),\\
&r^*(t,\theta)+\lambda\varepsilon\left(\hat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\lambda\varepsilon u(t))u(t)d\lambda.
\end{aligned}\tag{2.11}$$

by substituting (2.11) in (2.10), we obtain

$$\begin{cases}
-d\hat{y}^\varepsilon(t) &= \displaystyle\int_\Theta[D^\varepsilon(t,\theta)\hat{x}^\varepsilon(t)+I^\varepsilon(t,\theta)\hat{y}^\varepsilon(t)+F^\varepsilon(t,\theta)\hat{z}^\varepsilon(t)+\Lambda^\varepsilon(t,\theta)\hat{r}^\varepsilon(t,\theta)\\
&\quad +G^{4\varepsilon}(t,\theta)]\mu(d\theta)dt-\hat{z}^\varepsilon(t)dW(t)-\displaystyle\int_\Theta\hat{r}^\varepsilon(t,\theta)N\left(d\theta,dt\right),\\
\hat{y}^\varepsilon(T) &= \varepsilon^{-1}\left[\left(h\left(x^\varepsilon(T)\right)-h(x^*(T)\right)\right]-h_x\left(x^*(T)\right)x_1^\varepsilon(T),
\end{cases}$$

where

$$\begin{aligned}
D^\varepsilon(t,\theta)\doteq\int_0^1&[g_x(t,x^*(t)+\lambda\varepsilon\left(\hat{x}^\varepsilon(t)+x_1^\varepsilon(t)\right),y^*(t)+\lambda\varepsilon\left(\hat{y}^\varepsilon(t)+y_1^\varepsilon(t)\right),z^*(t)\\
&+\lambda\varepsilon\left(\hat{z}^\varepsilon(t)+z_1^\varepsilon(t)\right),r^*(t,\theta)+\lambda\varepsilon\left(\hat{r}^\varepsilon(t,\theta)+r_1^\varepsilon(t,\theta)\right),u^*(t)+\lambda\varepsilon u(t))d\lambda.
\end{aligned}$$

$$I^{\varepsilon}(t,\theta) \doteq \int_0^1 g_y(t, x^*(t) + \lambda\varepsilon\left(\hat{x}^{\varepsilon}(t) + x_1^{\varepsilon}(t)\right), y^*(t) + \lambda\varepsilon\left(\hat{y}^{\varepsilon}(t) + y_1^{\varepsilon}(t)\right), z^*(t)$$
$$+ \lambda\varepsilon\left(\hat{z}^{\varepsilon}(t) + z_1^{\varepsilon}(t)\right), r^*(t,\theta) + \lambda\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta) + r_1^{\varepsilon}(t,\theta)\right), u^*(t) + \lambda\varepsilon u(t))d\lambda.$$

$$F^{\varepsilon}(t,\theta) \doteq \int_0^1 g_z(t, x^*(t) + \lambda\varepsilon\left(\hat{x}^{\varepsilon}(t) + x_1^{\varepsilon}(t)\right), y^*(t) + \lambda\varepsilon\left(\hat{y}^{\varepsilon}(t) + y_1^{\varepsilon}(t)\right), z^*(t)$$
$$+ \lambda\varepsilon\left(\hat{z}^{\varepsilon}(t) + z_1^{\varepsilon}(t)\right), r^*(t,\theta) + \lambda\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta) + r_1^{\varepsilon}(t,\theta)\right), u^*(t) + \lambda\varepsilon u(t))d\lambda,$$

$$\Lambda^{\varepsilon}(t,\theta) \doteq \int_0^1 g_r(t, x^*(t) + \lambda\varepsilon\left(\hat{x}^{\varepsilon}(t) + x_1^{\varepsilon}(t)\right), y^*(t) + \lambda\varepsilon\left(\hat{y}^{\varepsilon}(t) + y_1^{\varepsilon}(t)\right), z^*(t)$$
$$+ \lambda\varepsilon\left(\hat{z}^{\varepsilon}(t) + z_1^{\varepsilon}(t)\right), r^*(t,\theta) + \lambda\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta) + r_1^{\varepsilon}(t,\theta)\right), u^*(t) + \lambda\varepsilon u(t))d\lambda.$$

$$\begin{aligned} G^{4\varepsilon}(t,\theta) \doteq\ & \left[D^{\varepsilon}(t,\theta) - g_x(t,\theta)\right] x_1^{\varepsilon}(t) + \left[I^{\varepsilon}(t,\theta) - g_y(t,\theta)\right] y_1^{\varepsilon}(t) \\ & + \left[F_1^{\varepsilon}(t,\theta) - g_z(t,\theta)\right] z_1^{\varepsilon}(t) + \left[\Lambda^{\varepsilon}(t,\theta) - g_r(t,\theta)\right] r_1^{\varepsilon}(t,\theta) \\ & + \int_0^1 [g_u(t, x^*(t) + \lambda\varepsilon\left(\hat{x}^{\varepsilon}(t) + x_1^{\varepsilon}(t)\right), y^*(t) + \lambda\varepsilon\left(\hat{y}^{\varepsilon}(t) + y_1^{\varepsilon}(t)\right), \\ & \quad z^*(t) + \lambda\varepsilon\left(\hat{z}^{\varepsilon}(t) + z_1^{\varepsilon}(t)\right), r^*(t,\theta) + \lambda\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta) + r_1^{\varepsilon}(t,\theta)\right), u^*(t) \\ & + \lambda\varepsilon u(t)) - g_u(t,\theta)]u(t)d\lambda. \end{aligned}$$

Applying Itô's formula $\left|\hat{y}^{\varepsilon}(t)\right|^2$, noting assumption (**H2.1**), we have

$$\begin{aligned} & E\left[\left|\hat{y}^{\varepsilon}(t)\right|^2\right] + E\int_t^T \left|\hat{z}^{\varepsilon}(s)\right|^2 ds + E\int_t^T \int_{\Theta} \left|\hat{r}^{\varepsilon}(s,\theta)\right|^2 \mu(d\theta)ds \\ = \ & E\int_t^T \int_{\Theta} \langle 2\hat{y}^{\varepsilon}(s), D^{\varepsilon}(s,\theta)\hat{x}^{\varepsilon}(s) + I^{\varepsilon}(s,\theta)\hat{y}^{\varepsilon}(s) + F^{\varepsilon}(s,\theta)\hat{z}^{\varepsilon}(s) \\ & +\Lambda^{\varepsilon}(s,\theta)\hat{r}^{\varepsilon}(s,\theta) + G^{4\varepsilon}(s,\theta)\rangle\, \mu(d\theta)ds \\ & +E\left[\varepsilon^{-1}\left[\left(h\left(x^{\varepsilon}(T)\right) - h(x^*(T))\right] - h_x\left(x^*(T)\right)x_1^{\varepsilon}(T)\right]^2 \\ \leq \ & CE\int_t^T \left|\hat{y}^{\varepsilon}(s)\right|^2 ds + \frac{1}{2}\mathbb{E}\int_t^T \left|\hat{z}^{\varepsilon}(s)\right|^2 ds + \frac{1}{2}\mathbb{E}\int_t^T \int_{\Theta} \left|\hat{r}^{\varepsilon}(s,\theta)\right|^2 \mu(d\theta)ds + o(\rho). \end{aligned}$$

By Gronwall's inequality again, we can get the last three convergence results of (2.5). ■

Variational inequality

Since $u^*(\cdot)$ is an optimal control, then

$$\varepsilon^{-1}\left[J\left(u^{\varepsilon}(\cdot)\right)-J\left(u^*(\cdot)\right)\right]\geq 0. \tag{2.12}$$

From this and Lemma (2.3.1), we have the following.

Lemma 2.3.2 *Let assumption (**H2.1**) hold. Then the following Variational inequality holds:*

$$\begin{aligned} o(\rho) \leq\ & E\int_0^T\int_\Theta [l_x(t,\theta)x_1^{\varepsilon}(t)+l_y(t,\theta)y_1^{\varepsilon}(t)+l_z(t,\theta)z_1^{\varepsilon}(t)+l_r(t,\theta)r_1^{\varepsilon}(t,\theta) \\ & +l_u(t,\theta)u(t)\mu(d\theta)]dt+E\left[\phi_x(x^*(T))x_1^{\varepsilon}(T)\right]+E\left[\varphi_y(y^*(0))y_1^{\varepsilon}(0)\right]. \end{aligned} \tag{2.13}$$

Proof From (2.12), we have

$$\begin{aligned} & \varepsilon^{-1}\left[J\left(u^{\varepsilon}(\cdot)\right)-J\left(u^*(\cdot)\right)\right] \\ =\ & \varepsilon^{-1}E\int_0^T\int_\Theta l(t,x^{\varepsilon}(t),y^{\varepsilon}(t),z^{\varepsilon}(t),r^{\varepsilon}(t,\theta),u^{\varepsilon}(t)) \\ & -l(t,x^*(t),y^*(t),z^*(t),r^*(t,\theta),u^*(t))]\mu(d\theta)\,dt \\ & +\varepsilon^{-1}E\left[\phi(x^{\varepsilon}(T))-\phi(x^*(T))\right]+\varepsilon^{-1}E\left[\varphi(y^{\varepsilon}(0))-\varphi(y^*(0))\right]\geq 0. \end{aligned}$$

By applying Taylor's expansion and the first result of (2.5) respectively, we get

$$\begin{aligned} \varepsilon^{-1}E\left[\phi(x^{\varepsilon}(T))-\phi(x^*(T))\right] &= \varepsilon^{-1}E\int_0^1\phi_x(x^*(T)+\lambda\left(x^{\varepsilon}(T)-x^*(T)\right))\left(x^{\varepsilon}(T)-x^*(T)\right)d\lambda \\ &\longrightarrow E\left[\phi_x(x^*(T)x_1^{\varepsilon}(T)\right]. \end{aligned}$$

Similarly, we have

$$\begin{aligned} \varepsilon^{-1}E\left[\varphi(y^{\varepsilon}(0))-\varphi(y^*(0))\right] &= \varepsilon^{-1}E\int_0^1\varphi_y(y^*(0)+\lambda\left(y^{\varepsilon}(0)-y^*(0)\right))\left(y^{\varepsilon}(0)-y^*(0)\right)d\lambda \\ &\longrightarrow E\left[\varphi_y(y^*(0)y_1^{\varepsilon}(0)\right], \end{aligned}$$

and

$$\varepsilon^{-1}E\int_0^T\int_\Theta[l(t,x^\varepsilon(t),y^\varepsilon(t),z^\varepsilon(t),r^\varepsilon(t,\theta),u^\varepsilon(t))$$
$$-l(t,x^*(t),y^*(t),z^*(t),r^*(t,\theta),u^*(t))]\mu\,(d\theta)\,dt$$
$$\longrightarrow E\int_0^T\int_\Theta[l_x(t,\theta)x_1^\varepsilon(t)+l_y(t,\theta)y_1^\varepsilon(t)+l_z(t,\theta)z_1^\varepsilon(t)$$
$$+l_r(t,\theta)r_1^\varepsilon(t,\theta)+l_u(t,\theta)u(t)]\mu\,(d\theta)\,dt.$$

Thus (2.13) follows. ■

2.3.2 Adjoint equations and adjoint processes

We introduce the following adjoint equations:

$$\begin{cases} -d\Psi(t) &= [f_x^\intercal(t)\Psi(t)-\int_\Theta g_x^\intercal(t,\theta)K(t)\mu(d\theta)+\sigma_x^\intercal(t)Q(t) \\ &\quad +\int_\Theta\left(c_x^\intercal(t,\theta)R(t,\theta)+l_x^\intercal(t,\theta)\right)\mu(d\theta)]dt \\ &\quad -Q(t)dW(t)-\int_\Theta R(t,\theta)N(d\theta,dt), \\ \Psi(T) &= -h_x^\intercal(x^*(T))K(T)+\phi_x(x^*(T)), \\ dK(t) &= \int_\Theta\left[g_y^\intercal(t,\theta)K(t)-l_y^\intercal(t,\theta)\right]\mu(d\theta)dt \\ &\quad +\int_\Theta\left[g_z^\intercal(t,\theta)K^*(t)-l_z^\intercal(t,\theta)\right]\mu(d\theta)dW\,(t) \\ &\quad +\int_\Theta\left[g_r^\intercal(t-,\theta)K^*(t-)-l_r^\intercal(t-,\theta)\right]N(d\theta,dt), \\ K(0) &= -\varphi_y(y^*(0)). \end{cases} \tag{2.14}$$

The same as (2.3), under assumption (**H2.1**), there exists a unique quartet $(\Psi(t),Q(t),K(t),R(t,\cdot))\in L^2_{\mathcal{F}}([0,T];\mathbb{R}^n)\times L^2_{\mathcal{F}}\left([0,T];\mathbb{R}^{n\times d}\right)\times L^2_{\mathcal{F}}([0,T];\mathbb{R}^m)\times\mathcal{M}^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ satisfying (2.14).

We define the Hamiltonian function

$$H:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^m\times U\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times\mathbb{R}^m\times\mathbb{R}^n\longrightarrow\mathbb{R},$$

as follows:

$$\begin{aligned} & H(t,x,y,z,r(\cdot),u,\Psi,Q,K,R(\cdot)) \\ \doteq\ & \langle \Psi, f(t,x,u)\rangle + \langle Q,\sigma(t,x,u)\rangle - \int_\Theta [\langle K, g(t,x,y,z,r(\theta),u)\rangle \\ & -l(t,x,y,z,r(\theta),u) - \langle R(\theta), c(t,x,u,\theta)\rangle]\mu(d\theta). \end{aligned} \tag{2.15}$$

Denote $H(t) = H(t,x^*(t),y^*(t),z^*(t),r^*(t,\cdot),u^*(t),\Psi(t),Q(t),K(t),R(t,\cdot))$ and its derivatives, then adjoint equations (2.14) can be rewritten as the following stochastic Hamiltonian system's type:

$$\begin{cases} -d\Psi(t) &= H_x(t)dt - Q(t)dW(t) - \int_\Theta R(t,\theta)N(d\theta,dt), \\ \Psi(T) &= -h_x^\intercal(x^*(T))K(T) + \phi_x(x^*(T)), \\ dK(t) &= -H_y(t)dt - H_z(t)dW(t) - \int_\Theta H_r(t-,\theta)N(d\theta,dt), \\ K(0) &= -\varphi_y(y^*(0)). \end{cases} \tag{2.16}$$

The main result in this chapter is the following.

2.3.3 Necessary conditions of optimality

In this part, our objective is to derive optimality necessary conditions satisfied by some optimal control, known as the stochastic maximum which is given by the following theorem.

Theorem 2.3.1 (Stochastic maximum principle) *Assume (**H2.1**) hold. Let $u^*(\cdot)$ be an optimal control for our stochastic optimal control problem, and $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot,\cdot))$ be the corresponding optimal trajectory. Then we have*

$$\langle H_u(t), u - u^*(t)\rangle \ \geq\ 0, \quad \forall u \in U, a.e.t \in [0,T], \mathbb{P}-a.s. \tag{2.17}$$

Proof Applying Itô's formula to $\langle \Psi(t), x_1^\varepsilon(t)\rangle$, we obtain

$$\begin{aligned}
E\left[\langle \Psi(T), x_1^\varepsilon(T)\rangle\right] &= E\int_0^T \langle \Psi(t), dx_1^\varepsilon(t)\rangle + E\int_0^T \langle x_1^\varepsilon(t), d\Psi(t)\rangle \\
&\quad + E\int_0^T \langle Q(t), \sigma_x(t)x_1^\varepsilon(t) + \sigma_u(t)u(t)\rangle\, dt \\
&\quad + E\int_0^T\int_\Theta \langle R(t,\theta), c_x(t)x_1^\varepsilon(t-) + c_u(t,\theta)u(t)\rangle\, \mu(d\theta)dt.
\end{aligned} \tag{2.18}$$

A simple computation shows that

$$E\int_0^T \langle \Psi(t), dx_1^\varepsilon(t)\rangle = E\int_0^T \langle \Psi(t), f_x(t)x_1^\varepsilon(t) + f_u(t)u(t)\rangle\, dt, \tag{2.19}$$

and

$$\begin{aligned}
E\int_0^T \langle x_1^\varepsilon(t), d\Psi(t)\rangle &= -E\int_0^T [\Big\langle x_1^\varepsilon(t), f_x^\intercal(t)\Psi(t) - \int_\Theta g_x^\intercal(t,\theta)K(t)\mu(d\theta) + \sigma_x^\intercal(t)Q(t) \\
&\quad + \int_\Theta (c_x^\intercal(t,\theta)R(t,\theta) + l_x^\intercal(t,\theta))\mu(d\theta)\Big\rangle]dt.
\end{aligned} \tag{2.20}$$

We replace (2.19), (2.20) in (2.18) and using the fact that

$$\Psi(T) = -h_x^\intercal(x^*(T))K(T) + \phi_x(x^*(T)),$$

we get

$$\begin{aligned}
& E\left[\langle -h_x^\intercal(x^*(T))K(T) + \phi_x(x^*(T))], x_1^\varepsilon(T)\rangle\right] \\
= \; & E\int_0^T\int_\Theta \langle x_1^\varepsilon(t), g_x^\intercal(t,\theta)K(t) - l_x^\intercal(t,\theta))\rangle\, \mu(d\theta)dt \\
& + E\int_0^T \langle \Psi(t), f_u(t)u(t)\rangle\, dt + E\int_0^T \langle Q(t), \sigma_u(t)u(t)\rangle\, dt \\
& + E\int_0^T\int_\Theta \langle R(t,\theta), c_u(t,\theta)u(t)\rangle\, \mu(d\theta)dt.
\end{aligned} \tag{2.21}$$

Applying Ito's formula also to $\langle K(t), y_1^\varepsilon(t)\rangle$, we obtain

$$\begin{aligned}
& E\left[\langle K(T), y_1^\varepsilon(T)\rangle\right] - E\left[\langle K(0), y_1^\varepsilon(0)\rangle\right] \\
= \; & E\int_0^T \langle K(t), dy_1^\varepsilon(t)\rangle + E\int_0^T \langle y_1^\varepsilon(t), dK(t)\rangle \\
& + E\int_0^T\int_\Theta \langle z_1^\varepsilon(t), g_z^\intercal(t,\theta)K(t) - l_z^\intercal(t,\theta)\rangle\,\mu\left(d\theta\right)dt \\
& + E\int_0^T\int_\Theta \langle r_1^\varepsilon(t,\theta), g_r^\intercal(t-,\theta)K(t-) - l_r^\intercal(t-,\theta)\rangle\,\mu(d\theta)dt,
\end{aligned} \tag{2.22}$$

where

$$\begin{aligned}
E\int_0^T \langle K(t), dy_1^\varepsilon(t)\rangle \;=\; & -E\int_0^T\int_\Theta \langle K(t), g_x(t,\theta)x_1^\varepsilon(t) + g_y(t,\theta)y_1^\varepsilon(t) + g_z(t,\theta)z_1^\varepsilon(t) \\
& + g_r(t,\theta)r_1^\varepsilon(t,\theta) + g_u(t,\theta)u(t)\rangle\,\mu\left(d\theta\right)dt
\end{aligned} \tag{2.23}$$

and

$$E\int_0^T \langle y_1^\varepsilon(t), dK(t)\rangle \;=\; E\int_0^T\int_\Theta \left\langle y_1^\varepsilon(t), g_y^\intercal(t,\theta)K(t) - l_y^\intercal(t,\theta)\right\rangle\mu(d\theta)dt. \tag{2.24}$$

We replace (2.23), (2.24) in (2.22) and use the fact that $K(0) = -\varphi_y(y^*(0))$, $y_1^\varepsilon(T) = h_x(x^*(T))x_1^\varepsilon(T)$, we get

$$\begin{aligned}
& E\left[\langle K(T), h_x(x^*(T))x_1^\varepsilon(T)\rangle\right] + E\left[\langle \varphi_y(y^*(0)), y_1^\varepsilon(0)\rangle\right] \\
= \; & -E\int_0^T\int_\Theta \left\langle y_1^\varepsilon(t), l_y^\intercal(t,\theta)\right\rangle\mu(d\theta)dt \\
& -E\int_0^T\int_\Theta \langle K(t), g_x(t,\theta)x_1^\varepsilon(t) + g_u(t,\theta)u(t)\rangle\,\mu\left(d\theta\right)dt \\
& -E\int_0^T\int_\Theta \langle z_1^\varepsilon(t), l_z^\intercal(t,\theta)\rangle\,\mu(d\theta)dt \\
& -E\int_0^T\int_\Theta \langle r_1^\varepsilon(t,\theta), l_r^\intercal(t-,\theta)\rangle\,\mu(d\theta)dt.
\end{aligned} \tag{2.25}$$

Combining (2.21) and (2.25), we get

$$
\begin{aligned}
& E\left[\phi_x(x^*(T)x_1^\varepsilon(T)\right] + E\left[\varphi_y(y^*(0)y_1^\varepsilon(0)\right] \\
= \; & -E\int_0^T\int_\Theta [l_x(t,\theta)x_1^\varepsilon(t) + l_y(t,\theta)y_1^\varepsilon(t) + l_z(t,\theta)z_1^\varepsilon(t) + l_r(t,\theta)r_1^\varepsilon(t,\theta) \\
& + l_u(t,\theta)u(t)]\mu\left(d\theta\right)dt + E\int_0^T \left\langle H_u(t), u(t)\right\rangle dt.
\end{aligned}
$$

This together with the variational inequality (2.13) implies, for $u(\cdot) \in \mathcal{U}\left([0,T]\right)$,

$$
E\int_0^T \left\langle H_u(t), u(t)\right\rangle dt \quad \geq \quad 0.
$$

So (2.17) holds. ■

2.4 Sufficient conditions for optimal control of FBSDEJs

In this section, we give sufficient conditions of optimality with the same notations used in the previous section. Therefore, we add some hypothesis:

Assumptions (H2.2)

i) ϕ is convex in x,

ii) φ is convex in y,

iii) H is convex in $(x, y, z, r(\cdot), u)$.

Then we have the following result.

Theorem 2.4.1 (Suffisient Conditions of Optimality) *Assume (**H2.1**) and (**H2.2**) holds. Let $u^*(\cdot)$ be an admissible control and $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot,\cdot))$ be the corresponding trajectory with $y^*(T) = M_T x^*(T)$, $M_T \in \mathbb{R}^{m\times n}$. Let $(\Psi(\cdot), Q(\cdot), K(\cdot), R(\cdot,\cdot))$ be the solution of the adjoint equations (2.14). Then $u^*(\cdot)$ is an optimal control if it satisfies (2.17).*

Proof Let $u(\cdot)$ be an arbitrary admissible control and $(x(\cdot), y(\cdot), z(\cdot), r(\cdot,\cdot))$ be the corresponding trajectory. We consider

$$\begin{aligned} J(u^*(\cdot)) - J(u(\cdot)) &= E\int_0^T \int_\Theta [l(t, x^*(t), y^*(t), z^*(t), r^*(t,\theta), u^*(t)) \\ &\quad -l(t, x(t), y(t), z(t), r(t,\theta), u(t))]\mu(d\theta)\, dt \\ &\quad +E\left[\phi(x^*(T) - \phi(x(T))\right] + E\left[\varphi(y^*(0) - \varphi(y(0))\right]. \end{aligned} \tag{2.26}$$

We first note that, by the convexity of ϕ and Itô's formula to $(x^*(t) - x(t))^\intercal \Psi(t)$, we get

$$\begin{aligned} & E\left[\phi(x^*(T) - \phi(x(T))\right] \\ \leq\ & E\left[(x^*(T) - x(T))^\intercal \phi_x(x^*(T))\right] \\ =\ & E\left[(x^*(T) - x(T))^\intercal \Psi(T)\right] + \mathbb{E}\left[(x^*(T) - x(T))^\intercal M_T^\intercal K(T)\right] \\ =\ & E\int_0^T [(x^*(t) - x(t))^\intercal (-f_x^\intercal(t)\Psi(t) + \int_\Theta g_x^\intercal(t,\theta)K(t)\mu(d\theta) - \sigma_x^\intercal(t)Q(t) \\ & -\int_\Theta \left(c_x^\intercal(t,\theta)R(t,\theta) - l_x^\intercal(t,\theta)\right)\mu(d\theta)) + \langle \Psi(t), f(t) - f(t, x(t), u(t))\rangle \\ & \langle Q(t), \sigma(t) - \sigma(t, x(t), u(t))\rangle + \int_\Theta \langle R(t,\theta), c(t,\theta) - c(t, x(t), u(t), \theta)\rangle\, \mu(d\theta)]dt \\ & +E\left[(x^*(T) - x(T))^\intercal M_T^\intercal K(T)\right]. \end{aligned}$$

And similarly, by the convexity of φ and Itô's formula to $(y^*(t) - y(t))^\intercal K(t)$, it becomes

$$\begin{aligned} & E\left[\varphi(y^*(0) - \varphi(y(0))\right] \\ \leq\ & E\left[(y^*(0) - y(0))^\intercal \varphi_y(y(0))\right] \\ =\ & -E\left[(y^*(0) - y(0))^\intercal K(0)\right] \\ =\ & -E\left[(x^*(T) - x(T))^\intercal M_T^\intercal K(T)\right] \\ & +E\int_0^T \int_\Theta [(y^*(t) - y(t))^\intercal \left(g_y^\intercal(t,\theta)K(t) - l_y^\intercal(t,\theta)\right) \\ & +\left[(z^*(t) - z(t))^\intercal \left(g_z^\intercal(t,\theta)K(t) - l_z^\intercal(t,\theta)\right)\right] \\ & +\left[(r^*(t,\theta) - r(t,\theta))^\intercal \left(g_r^\intercal(t,\theta)K(t) - l_r^\intercal(t,\theta)\right)\right] \\ & -\langle K(t), g(t,\theta) - g(t, x(t), y(t), z(t), r(t,\theta), u(t))\rangle]\mu(d\theta)dt. \end{aligned}$$

By the definition (2.15) of H, we have

$$
\begin{aligned}
& E\int_0^T\int_\Theta [l(t,x^*(t),y^*(t),z^*(t),r^*(t,\theta),u^*(t))-l(t,x(t),y(t),z(t),r(t,\theta),u(t))]\mu(d\theta)dt \\
= \; & E\int_0^T [H(t)-H(t,x(t),y(t),z(t),r(t,\cdot),u(t),\Psi(t),Q(t),K(t),R(t,\cdot))]dt \\
& +E\int_0^T\int_\Theta [-\langle \Psi(t), f(t)-f(t,x(t),u(t))\rangle - \langle Q(t),\sigma(t)-\sigma(t,x(t),u(t))\rangle \\
& +\langle K(t), g(t,\theta)-g(t,x(t),y(t),z(t),r(t,\theta),u(t))\rangle \\
& -\langle R(t,\theta), c(t,\theta)-c(t,x(t),u(t),\theta)\rangle \, \mu(d\theta)]dt.
\end{aligned}
$$

Adding the above (in)equalities up, from (2.26), we can get

$$
\begin{aligned}
& J(u^*(\cdot))-J(u(\cdot)) \\
\leq \; & E\int_0^T [H(t)-H(t,x(t),y(t),z(t),r(t,\cdot),u(t),\Psi(t),Q(t),K(t),R(t,\cdot)) \\
& -\langle H_x(t), x^*(t)-x(t)\rangle - \langle H_y(t), y^*(t)-y(t)\rangle - \langle H_z(t), z^*(t)-z(t)\rangle \\
& -\langle H_r(t), r^*(t,\cdot)-r(t,\cdot)\rangle]dt,
\end{aligned}
\tag{2.27}
$$

Since H is convex in $(x,y,z,r(\cdot),u)$.

$$
\begin{aligned}
& H(t)-H(t,x(t),y(t),z(t),r(t,\cdot),u(t),\Psi(t),K(t),Q(t),R(t,\cdot)) \\
\leq \; & \langle H_x(t), x^*(t)-x(t)\rangle + \langle H_y(t), y^*(t)-y(t)\rangle + \langle H_z(t), z^*(t)-z(t)\rangle \\
& +\langle H_r(t), r^*(t,\theta)-r(t,\theta)\rangle + \langle H_u(t), u^*(t)-u(t)\rangle .
\end{aligned}
\tag{2.28}
$$

Combing (2.27) and (2.28), we get

$$
J(u^*(\cdot))-J(u(\cdot)) \;\leq\; E\int_0^T \langle H_u(t), u^*(t)-u(t)\rangle \, dt. \tag{2.29}
$$

Then from the maximum condition (2.17), we deduce that $J(u^*(\cdot)) \leq J(u(\cdot))$ for all $u(\cdot)\in U$, which proves that $u^*(\cdot)$ is optimal. ■

2.5 Applications to finance

In this section, we will apply the maximum principle (2.3.1) to study a mean-variance portfolio selection mixed with a recursive utility functional optimization problem. Suppose we have two kinds of securities in the market for possible investment choice:

1. *A risk-free security (Bond price)*, where the price $P_0(t)$ at time t is given by

$$dP_0(t) = \rho(t) P_0(t)dt, P_0(0) > 0, \tag{2.30}$$

   here $\rho(t)$ is a bounded deterministic function;

2. *A risk security (Stock price)*, where the price $P_1(t)$ at time t is given by

$$dP_1(t) = P_1(t-)\left[\zeta(t)\,dt + \sigma(t)\,dW(t) + \int_\Theta \xi(t,\theta)N(d\theta,dt)\right], \; P_1(0) > 0, \tag{2.31}$$

   here $\zeta(t), \sigma(t) \neq 0$, are bounded deterministic functions and $\zeta(t) > \rho(t)$. To ensure that $P_1(t) > 0$ for all t, we assume that $\xi(t,\theta) > -1$, $\forall\theta \in \Theta$ and in addition we assume that $\int_\Theta \xi^2(t)\mu(d\theta)$ is abounded function.

Let $u(t) \doteq e_1(t)P_1(t)$ denote the amount invested in the risky security which we call portfolio strategy. Let be $x(0) = x_0 \geq 0$ the initial wealth. By combining (2.30) and (2.31), we introduce the wealth process $x(\cdot)$ and the recursive utility process $y(\cdot)$ as the solution of the following FBSDEJs:

$$\begin{cases} dx(t) = [\rho(t)x(t) + (\zeta(t) - \rho(t))u(t)]\,dt + \sigma(t)u(t)dW(t) + \int_\Theta \xi(t,\theta)u(t-)N(d\theta,dt), \\ -dy(t) = [\rho(t)x(t) + (\zeta(t) - \rho(t))u(t) - \beta y(t)]\,dt - z(t)dW(t) - \int_\Theta r(t,\theta)N(d\theta,dt), \\ x(0) = x_0, y(T) = x(T). \end{cases} \tag{2.32}$$

We denote by $\mathcal{U}([0,T])$ the set of admissible portfolios valued in $U = \mathbb{R}$.
The cost functional, to be minimized, is given by

$$J(u(\cdot)) \doteq E[\frac{1}{2}(x(T) - a)^2] - y(0), \tag{2.33}$$

where some $a \in \mathbb{R}$ is given.
The optimization problem can be written as

$$J(u^*(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}([0,T])} J(u(\cdot)). \tag{2.34}$$

In this case the adjoint equations (2.14) reduce to

$$\begin{cases} -d\Psi(t) &= \rho(t) \left[\Psi(t) - K(t)\right] dt - Q(t) dW(t) - \displaystyle\int_\Theta R(t,\theta) N(d\theta, dt), \\ dK(t) &= -\beta K(t) dt, \\ \Psi(T) &= x(T) - a - K(T), K(0) = 1. \end{cases} \tag{2.35}$$

Let $u^*(\cdot)$ be an optimal portfolio strategy and $x^*(\cdot), y^*(\cdot)$ be the corresponding wealth process, recursive utility process, respectively, with corresponding solution $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot), R^*(\cdot,\cdot))$ of the adjoint equations (2.35).
Hamiltonian function (2.15) reduce to

$$\begin{aligned} & H(t, x^*(t), y^*(t), z^*(t), r^*(t,\theta), u, \Psi^*(t), Q^*(t), K^*(t), R^*(t,\theta)) \\ = & - \left[\rho(t) x^*(t) + (\zeta(t) - \rho(t))\, u\right] \left[K^*(t) - \Psi^*(t)\right] dt + \sigma(t) Q^*(t) u \\ & + \beta K^*(t) y^*(t) + \int_\Theta \xi(t,\theta) R^*(t,\theta)\, u \mu(d\theta). \end{aligned} \tag{2.36}$$

Since this is a linear expression of u, by the maximum condition ((2.17), Theorem (2.3.1)), we have

$$- (\zeta(t) - \rho(t)) \left[K^*(t) - \Psi^*(t)\right] dt + \sigma(t) K^*(t) + \int_\Theta \xi(t,\theta) R^*(t,\theta)\, \mu(d\theta) = 0. \tag{2.37}$$

In order to find the expression of $u^*(t)$, we conjecture a process $\Psi^*(t)$ with form

$$\Psi^*(t) = A(t) x^*(t) + B(t), \tag{2.38}$$

where $A(t), B(t)$ are deterministic differentiable functions.

Applying Itô's formula to (2.38), in virtue of (2.32), we can get

$$\begin{aligned} d\Psi^*(t) &= A(t)\left\{\left[\rho(t)x^*(t) + \left(\zeta(t)-\rho(t)\right)u^*(t)\right]dt + \sigma(t)u^*(t)dW(t) \right. \\ &\quad \left. + \int_\Theta \xi(t,\theta)u^*(t-)\,N(d\theta,dt)\right\} + x^*(t)A'(t)dt + B'(t)dt, \\ &= \left[A(t)\rho(t)x^*(t) + A(t)\left(\zeta(t)-\rho(t)\right)u^*(t) + x^*(t)A'(t) + B'(t)\right]dt \\ &\quad + A(t)\sigma(t)u^*(t)dW(t) + \int_\Theta A(t)\xi(t,\theta)u^*(t-)\,N(d\theta,dt), \end{aligned} \tag{2.39}$$

where $A'(t)$ and $B'(t)$ denote the derivatives with respect to t.
Comparing (2.39) with the BSDEJ the (2.35), we get (noting that $K^*(t) = e^{-\beta t}$)

$$\begin{aligned} & A(t)\rho(t)x^*(t) + A(t)\left(\zeta(t)-\rho(t)\right)u^*(t) + x^*(t)A'(t) + B'(t) \\ = & -\rho(t)\left(A(t)x^*(t) + B(t)\right) + \rho(t)e^{-\beta t}, \end{aligned} \tag{2.40}$$

$$Q^*(t) = A(t)\sigma(t)u^*(t), \tag{2.41}$$

$$R^*(t,\theta) = A(t)\xi(t,\theta)u^*(t). \tag{2.42}$$

Substituting (2.41), (2.42) in (2.37) and denoting

$$\Lambda(t) \doteq \sigma^2(t) + \int_\Theta \xi^2(t,\theta)\mu(d\theta), \tag{2.43}$$

we can get

$$u^*(t) = \frac{\left(\rho(t)-\zeta(t)\right)\left(A(t)x^*(t) + B(t) - e^{-\beta t}\right)}{A(t)\Lambda(t)}. \tag{2.44}$$

On the other hand, (2.40) gives

$$u^*(t) = \frac{\left(2A(t)\rho(t) + A'(t)\right)x^*(t) + \rho(t)B(t) + B'(t) - \rho(t)e^{-\beta t}}{A(t)\left(\rho(t)-\zeta(t)\right)}. \tag{2.45}$$

Combining (2.44) and (2.45) (noting the terminal condition in (2.35)), we get

$$\begin{cases} A'(t) = \left[\dfrac{\left(\rho(t)-\zeta(t)\right)^2}{\Lambda(t)} - 2\rho(t)\right]A(t), \\ A(T) = 1, \end{cases}$$

and

$$\begin{cases} B'(t) &= \left[\frac{(\rho(t)-\zeta(t))^2}{\Lambda(t)} - \rho(t)\right] B(t) - e^{-\beta t}\left[\frac{(\rho(t)-\zeta(t))^2}{\Lambda(t)} - \rho(t)\right], \\ B(T) &= -a-1. \end{cases}$$

The solutions of these equations are

$$\begin{cases} A(t) &= \exp\left\{-\int_t^T [\frac{(\rho(s)-\zeta(s))^2}{\Lambda(s)} - 2\rho(s)]ds\right\}, \\ B(t) &= \exp\left\{-\int_t^T [\frac{(\rho(s)-\zeta(s))^2}{\Lambda_s} - \rho(s)]ds\right\} \\ & \left\{\int_t^T e^{-\beta s}\left[\frac{(\rho(s)-\zeta(s))^2}{\Lambda(s)} - \rho(s)\right]\exp\left\{\int_s^T [\frac{(\rho(r)-\zeta(r))^2}{\Lambda(r)} - \rho(r)]dr\right\}ds - a - e^{-\beta T}\right\}. \end{cases} \tag{2.46}$$

With this choice of $A(t)$ and $B(t)$ the process

$$\Psi^*(t) = A(t)x^*(t) + B(t), K^*(t) = e^{-\beta t}, Q^*(t) = A(t)\sigma(t)u^*(t), R^*(t,\theta) = A(t)\xi(t,\theta)u^*(t),$$

satisfying the adjoint equation (2.35) with $u^*(t)$ given by (2.44). Moreover, with this choice of $u^*(t)$, the maximum condition (2.17) of Theorem (2.3.1)) holds.
Finally, we give the explicit optimal portfolio section strategy in the state feedback form.

Theorem 2.5.1 *The optimal solution $u^*(t)$ of our mean-variance portfolio selection mixed with a recursive utility optimization problem (2.34), when the wealth dynamics obeys (2.32), is given in the state feedback form by*

$$u^*(t,x^*) = \frac{(\rho(t)-\zeta(t))\left(A(t)x^*(t) + B(t) - \rho(t)e^{-\beta t}\right)}{A(t)\Lambda(t)}.$$

where $\Lambda(t)$, $A(t)$ and $B(t)$ are given by (2.43) and (2.46) respectively.

Chapter 3

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

Chapter 3

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

3.1 Introduction

In this chapter, by means of convex variation methods and duality techniques, we will give the necessary conditions of optimality satisfied by an optimal control in the form of maximum principle. This maximum principle differs from the classical one, where the adjoint equation is a linear forward-backward stochastic differential equation with Poisson jump processes, since here the adjoint equation turns out to be a linear mean-field forward-backward stochastic differential equation with Poisson jump processes.

We consider stochastic optimal control for systems governed by nonlinear mean-field controlled forward-backward stochastic differential equation with Poisson jump processes (FB-

SDEJs) of the form

$$\begin{cases} dx(t) &= f(t,x(t),E(x(t)),u(t))dt + \sigma(t,x(t),E(x(t)),u(t))dW(t) \\ & \quad + \int_\Theta c(t,x(t-),E(x(t-)),u(t),\theta)N(d\theta,dt), \\ dy(t) &= -\int_\Theta g(t,x(t),E(x(t)),y(t),E(y(t)),z(t),E(z(t)),r(t,\theta),u(t))\mu(d\theta)\,dt \\ & \quad + z(t)dW(t) + \int_\Theta r(t,\theta)N(d\theta,dt), \\ x(0) &= \zeta, y(T) = h(x(T),E(x(T))), \end{cases} \tag{3.1}$$

where f, σ, c, g et h are given maps and the initial condition ζ is an $\mathcal{F}_0-$measurable random variable. The mean-field FBSDEJs-(3.1) called McKean-Vlasov systems are obtained as the mean square limit of an interacting particle system of the form

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

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 (3.1) 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

$$\begin{aligned} J(u(\cdot)) &= E[\int_0^T \int_\Theta l(t, x(t), E(x(t)), y(t), E(y(t)), z(t), E(z(t)), \\ &\quad r(t,\theta), u(t))\mu(d\theta)\, dt + \phi(x(T), E(x(T))) + \varphi(y(0), E(y(0)))], \end{aligned} \tag{3.2}$$

where l, ϕ and φ is an appropriate functions. This cost functional is also of mean-field type, as the functions l, ϕ and 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 the 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 $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted and square-integrable process with values in a nonempty convex subset $\mathcal{A}$ of $\mathbb{R}$. We denote by $\mathcal{U}([0,T])$ the set of all admissible controls.
Any admissible control $u^*(\cdot) \in \mathcal{U}([0,T])$ satisfying

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}([0,T])} J(u(\cdot)), \tag{3.3}$$

is called an optimal control.
This chapter is organized as follows. In the second section, we formulate the mean-field stochastic control problem and describe the assumptions of the model. The third section 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 is discussed in the fourth section.

3.2 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 $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ be a fixed filtered probability space equipped with a $\mathbb{P}-$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

$\{\mathcal{F}_t\}_{t\in[0,T]}-$ Poisson point process independent of W. We denote by $\tilde{N}(d\theta, dt)$ 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 σ-finite measure on $(\Theta, \mathcal{B}(\Theta))$ with $\mu(d\theta) < \infty$.
We then define

$$N(d\theta, dt) := \tilde{N}(d\theta, dt) - \mu(d\theta),$$

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

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

where $\mathcal{G}_0$ denotes the totality of $\mathbb{P}-$null sets and $\sigma_1 \vee \sigma_2$ denotes the σ-field generated by $\sigma_1 \cup \sigma_2$.
Throughout this paper, we also assume that the functions

$$\begin{aligned}
&f, \sigma : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathcal{A} \longrightarrow \mathbb{R},\\
&c : [0,T] \times \mathbb{R} \times \mathcal{A} \times \Theta \longrightarrow \mathbb{R},\\
&g : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{A} \longrightarrow \mathbb{R},\\
&h : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R},\\
&l : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{A} \longrightarrow \mathbb{R},\\
&\phi, \varphi : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R},
\end{aligned}$$

satisfy the following standing assumptions:

Assumptions (H1):

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

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

Assumptions (H2):

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

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

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

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

Under the assumptions **(H1)** and **(H2)**, the FBSDEJs (3.1) has a unique solution $(x(t), y(t), z(t), r(t, \cdot)) \in L_{\mathcal{F}}^2([0, T] ; \mathbb{R}) \times L_{\mathcal{F}}^2([0, T] ; \mathbb{R}) \times L_{\mathcal{F}}^2([0, T] ; \mathbb{R}) \times \mathcal{M}_{\mathcal{F}}^2([0, T] ; \mathbb{R})$. (See [45] 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:

$$\begin{cases}
d\Psi(t) = -\{f_x(t)\Psi(t)+E(f_{\tilde{x}}(t)\Psi(t))+\sigma_x(t)Q(t)+E(\sigma_{\tilde{x}}(t)Q(t)) \\
\quad +\int_{\Theta}[-g_x(t,\theta)K(t)-E(g_{\tilde{x}}(t,\theta)K(t))+c_x(t,\theta)R(t,\theta) \\
\quad +E(c_{\tilde{x}}(t,\theta)R(t,\theta))+l_x(t,\theta)+E(l_{\tilde{x}}(t,\theta))]\mu(d\theta)\}dt \\
\quad +Q(t)dW(t)+\int_{\Theta}R(t,\theta)N(d\theta,dt), \\
\Psi(T) = -[h_x(x(T),E(x(T))K(T)+E(h_{\tilde{x}}(x(T),E(x(T)))K(T))] \\
\quad +\phi_x(x(T),E(x(T)))+E(\phi_{\tilde{x}}(x(T),E(x(T)))), \\
dK(t) = \int_{\Theta}[g_y(t,\theta)K(t)+E(g_{\tilde{y}}(t,\theta)K(t))-l_y(t,\theta)-E(l_{\tilde{y}}(t,\theta))]\mu(d\theta)dt \\
\quad +\int_{\Theta}[g_z(t,\theta)K(t)+E(g_{\tilde{z}}(t,\theta)K(t))-l_z(t,\theta)-E(l_{\tilde{z}}(t,\theta))]\mu(d\theta)dW(t) \\
\quad +\int_{\Theta}[g_r(t,\theta)K(t)-l_r(t,\theta)]N(d\theta,dt), \\
K(0) = -(\varphi_y(y(0))+E(\varphi_{\tilde{y}}(y(0)))).
\end{cases} \tag{3.4}$$

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

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

associated with the stochastic control problem (3.1)-(3.2) as follows

$$\begin{aligned} & H\left(t,x,\tilde{x},y,\tilde{y},z,\tilde{z},r,u,\Psi,Q,K,R\right) \\ := \; & \Psi(t) f\left(t,x,\tilde{x},u\right) + Q(t)\sigma\left(t,x,\tilde{x},u\right) \\ & + \int_{\Theta} [-K(t) g\left(t,x,\tilde{x},y,\tilde{y},z,\tilde{z},r,u\right) + R(t,\theta) c\left(t,x,\tilde{x},u,\theta\right) \\ & + l\left(t,x,\tilde{x},y,\tilde{y},z,\tilde{z},r,u\right)]\mu(d\theta). \end{aligned} \tag{3.5}$$

If we denote by

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

then the adjoint equation (3.4) can be rewritten as the following stochastic Hamiltonian system's type

$$\begin{cases} -d\Psi(t) &= \left(H_x(t) + E\left(H_{\tilde{x}}(t)\right)\right) dt - Q(t) dW(t) - \displaystyle\int_{\Theta} R(t,\theta) N(d\theta, dt), \\ \Psi\left(T\right) &= -[h_x(x(T), E(x(T)))K(T) + E\left(h_{\tilde{x}}(x(T), E\left(x(T)\right))\right)K(T)] \\ & \quad + \phi_x(x\left(T\right), E\left(x\left(T\right)\right)) + E\left(\phi_{\tilde{x}}(x\left(T\right), E\left(x\left(T\right)\right))\right), \\ -dK(t) &= \left(H_y(t) + E\left(H_y(t)\right)\right) dt + \left(H_z(t) + E\left(H_{\tilde{z}}(t)\right)\right) dW(t) \\ & \quad + \displaystyle\int_{\Theta} H_r(t,\theta) N(d\theta, dt), \\ K(0) &= -\left(\varphi_y(y(0)) + E\left(\varphi_{\tilde{y}}(y(0))\right)\right). \end{cases} \tag{3.6}$$

Thanks to Lemma 3.1 in Shen and Siu [45], under assumptions (**H1**), (**H2**), the adjoint equations (3.4) admit a unique solution $(\Psi(t), Q(t), K(t), R(t,\cdot))$ such that

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

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

$$E\left[\sup_{t\in[0,T]} |\Psi(t)|^2 + \sup_{t\in[0,T]} |K(t)|^2 + \int_0^T |Q(t)|^2 dt + \int_0^T \int_\Theta |R(t,\theta)|^2 \mu(d\theta)dt\right] < C. \quad (3.7)$$

3.3 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 and duality 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-(3.1) corresponding to the optimal control $u^*(\cdot)$ and $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot), R^*(\cdot,\cdot))$ be the solution of adjoint equation (3.4) 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 (3.1)-(3.2). Then the maximum principle holds, that is, $\forall u \in \mathcal{A}$*

$$H_u(t, \lambda^*(t,\theta), u^*, \Lambda^*(t,\theta))(u - u^*(t)) \quad \geq \quad 0, \quad \mathbb{P}-a.s., a.e., t \in [0,T], \quad (3.8)$$

where

$$\lambda^*(t,\theta) \quad = \quad (x^*(t), E(x^*(t)), y^*(t), E(y^*(t)), z^*(t), E(z^*(t)), r^*(t,\theta)),$$

and

$$\Lambda^{*}(t,\theta) = (\Psi^{*}(t), Q^{*}(t), K^{*}(t), R^{*}(t,\theta)).$$

We derive the variational inequality (3.8) in several steps, from the fact that

$$J(u^{\varepsilon}(\cdot)) \geq J(u^{*}(\cdot). \tag{3.9}$$

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

$$u^{\varepsilon}(t) = u^{*}(t) + \varepsilon(u(t) - u^{*}(t)),$$

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 (3.1) 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 (3.4) corresponding to perturbed control $u^{\varepsilon}(\cdot)$.

Variational equations:

We introduce the following variational equations which have a mean-field type.

For simplicity of notation, we still use $f_x(t) = \dfrac{\partial f}{\partial x}(t, x^{*}(\cdot), E(x^{*}(\cdot)), u^{*}(\cdot))$, etc.

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

$$\begin{cases} dx_1^{\varepsilon}(t) = [f_x(t)x_1^{\varepsilon}(t) + f_{\tilde{x}}(t)E(x_1^{\varepsilon}(t)) + f_u(t)u(t)]\,dt \\ \quad + [\sigma_x(t)x_1^{\varepsilon}(t) + \sigma_{\tilde{x}}(t)E(x_1^{\varepsilon}(t)) + \sigma_u(t)u(t)]\,dW(t) \\ \quad + \int_{\Theta}(c_x(t,\theta)x_1^{\varepsilon}(t) + c_{\tilde{x}}(t,\theta)E(x_1^{\varepsilon}(t)) + c_u(t,\theta)u(t))\,N(d\theta,dt), \\ x_1^{\varepsilon}(0) = 0, \\ dy_1^{\varepsilon}(t) = -\int_{\Theta}[g_x(t,\theta)x_1^{\varepsilon}(t) + g_{\tilde{x}}(t,\theta)E(x_1^{\varepsilon}(t)) + g_y(t,\theta)y_1^{\varepsilon}(t) \\ \quad + g_{\tilde{y}}(t,\theta)E(y_1^{\varepsilon}(t)) + g_z(t,\theta)z_1^{\varepsilon}(t) + g_{\tilde{z}}(t,\theta)E(z_1^{\varepsilon}(t)) + g_r(t,\theta)r_1^{\varepsilon}(t,\theta) \\ \quad + g_u(t,\theta)u(t)]\mu(d\theta)dt + z_1^{\varepsilon}(t)dW(t) + \int_{\Theta} r_1^{\varepsilon}(t,\theta)N(d\theta,dt), \\ y_1^{\varepsilon}(T) = -[h_x(T) + E(h_{\tilde{x}}(T))]\,x_1^{\varepsilon}(T). \end{cases} \tag{3.10}$$

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 (3.3.1).

Lemma 3.3.1 *We have*

$$\begin{aligned}
& E\left(\Psi^*(T)x_1^\varepsilon(T)\right) \\
= & E\int_0^T [\Psi^*(t)f_u(t)u(t) + Q^*(t)\sigma_u(t)u(t) \\
& +\int_\Theta R^*(t,\theta)c_u(t,\theta)u(t)\mu(d\theta)]dt \\
& +E\int_0^T\int_\Theta \{x_1^\varepsilon(t)g_x(t,\theta)K^*(t) \\
& +x_1^\varepsilon(t)E\left(g_{\tilde{x}}(t,\theta)K^*(t)\right) - x_1^\varepsilon(t)l_x(t,\theta) \\
& -x_1^\varepsilon(t)E\left(l_{\tilde{x}}(t,\theta)\right)\}\mu(d\theta)dt,
\end{aligned} \tag{3.11}$$

similarly, we get

$$\begin{aligned}
& E\left(K^*(T)y_1^\varepsilon(T)\right) \\
= & -E\{[\varphi_y(y^*(0),E\left(y^*(0)\right)) + E\left(\varphi_{\tilde{y}}(y^*(0),E\left(y^*(0)\right))\right)]y_1^\varepsilon(0)\} \\
& -E\int_0^T\int_\Theta \{K^*(t)g_x(t,\theta)x_1^\varepsilon(t) + K^*(t)g_{\tilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right) \\
& +K^*(t)g_u(t,\theta)u(t) + y_1^\varepsilon(t)l_y(t,\theta) + y_1^\varepsilon(t)E\left(l_{\tilde{y}}(t,\theta)\right) \\
& +z_1^\varepsilon(t)l_z(t,\theta) + z_1^\varepsilon(t)E\left(l_{\tilde{z}}(t,\theta)\right) + r_1^\varepsilon(t,\theta)l_r(t,\theta)\}\mu(d\theta)dt,
\end{aligned} \tag{3.12}$$

and

$$\begin{aligned}
& E\left\{[\phi_x\left(x^*\left(T\right),E\left(x^*\left(T\right)\right)\right) + E\left(\phi_{\tilde{x}}\left(x^*\left(T\right),E\left(x^*\left(T\right)\right)\right)\right)]x_1^\varepsilon(T)\right\} \\
& +E\{[\varphi_y(y^*(0),E\left(y^*(0)\right)) + E\left(\varphi_{\tilde{y}}(y^*(0),E\left(y^*(0)\right))\right)]y_1^\varepsilon(0)\} \\
= & -E\int_0^T\int_\Theta x_1^\varepsilon(t)l_x(t,\theta) + x_1^\varepsilon(t)E\left(l_{\tilde{x}}(t,\theta)\right) + y_1^\varepsilon(t)l_y(t,\theta) + y_1^\varepsilon(t)E\left(l_{\tilde{y}}(t,\theta)\right) \\
& +z_1^\varepsilon(t)l_z(t,\theta) + z_1^\varepsilon(t)E\left(l_{\tilde{z}}(t,\theta)\right) + r_1^\varepsilon(t)l_r(t,\theta) + l_u(t,\theta)u(t)\}\mu(d\theta)dt \\
& +E\int_0^T H_u(t)u(t)dt.
\end{aligned} \tag{3.13}$$

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

$$\begin{aligned}E\left(\Psi^*(T)x_1^\varepsilon(T)\right) &= E\int_0^T \Psi^*(t)dx_1^\varepsilon(t) + E\int_0^T x_1^\varepsilon(t)d\Psi^*(t)\\ &+E\int_0^T Q^*(t)\left[\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\tilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + \sigma_u(t)u(t)\right]dt\\ &+E\int_0^T\int_\Theta R^*\left(t,\theta\right)\left[c_x(t,\theta)x_1^\varepsilon(t) + c_{\tilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right) + c_u(t)u(t)\right]\mu(d\theta)dt\\ &= I_1^\varepsilon + I_2^\varepsilon + I_3^\varepsilon + I_4^\varepsilon,\end{aligned} \tag{3.14}$$

A simple computation shows that

$$\begin{aligned}I_1^\varepsilon &= E\int_0^T \Psi^*(t)dx_1^\varepsilon(t)\\ &= E\int_0^T \Psi^*(t)\left[f_x(t)x_1^\varepsilon(t) + f_{\tilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + f_u(t)u(t)\right]dt\\ &= E\int_0^T \{\Psi^*(t)f_x(t)x_1^\varepsilon(t) + \Psi^*(t)f_{\tilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + \Psi^*(t)f_u(t)u(t)\}dt,\end{aligned} \tag{3.15}$$

and

$$\begin{aligned}I_2^\varepsilon &= E\int_0^T x_1^\varepsilon(t)d\Psi^*(t)\\ &= -E\int_0^T \{x_1^\varepsilon(t)f_x(t)\Psi^*(t) + x_1^\varepsilon(t)E\left(f_{\tilde{x}}(t)\Psi^*(t)\right) + x_1^\varepsilon(t)\sigma_x(t)Q^*(t)\\ &\quad +x_1^\varepsilon(t)E\left(\sigma_{\tilde{x}}(t)Q^*(t)\right) + \int_\Theta(-x_1^\varepsilon(t)g_x(t,\theta)K^*(t)\\ &\quad -x_1^\varepsilon(t)E\left(g_{\tilde{x}}(t,\theta)K^*(t)\right) + x_1^\varepsilon(t)c_x(t,\theta)R^*(t,\theta)\\ &\quad +x_1^\varepsilon(t)E\left(c_{\tilde{x}}(t,\theta)R^*(t,\theta)\right) + x_1^\varepsilon(t)l_x(t,\theta) + x_1^\varepsilon(t)E\left(l_{\tilde{x}}(t,\theta)\right))\mu(d\theta)\}dt.\end{aligned} \tag{3.16}$$

From (3.14), we get

$$\begin{aligned}I_3^\varepsilon &= E\int_0^T Q^*(t)\left[\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\tilde{x}}(t)E\left(x_1^\varepsilon(t)\right) + \sigma_u(t)u(t)\right]\\ &= E\int_0^T Q^*(t)\sigma_x(t)x_1^\varepsilon(t)dt + E\int_0^T Q^*(t)\sigma_{\tilde{x}}(t)E\left(x_1^\varepsilon(t)\right)dt + E\int_0^T Q^*(t)\sigma_u(t)u(t)dt,\end{aligned}$$

$$\begin{aligned} I_4^{\varepsilon} &= E\int_0^T \int_\Theta R^*(t,\theta)\left[c_x(t,\theta)x_1^{\varepsilon}(t) + c_{\tilde{x}}(t,\theta)E\left(x_1^{\varepsilon}(t)\right)\right. \\ &\quad \left. + c_u(t)u(t)\mu(d\theta)\right]dt \\ &= E\int_0^T \int_\Theta R^*(t,\theta)\, c_x(t,\theta)x_1^{\varepsilon}(t)\mu(d\theta)dt \\ &\quad + E\int_0^T \int_\Theta R^*(t,\theta)\, c_{\tilde{x}}(t,\theta)E\left(x_1^{\varepsilon}(t)\right)\mu(d\theta)dt \\ &\quad + E\int_0^T \int_\Theta R^*(t,\theta)\, c_u(t)u(t)\mu(d\theta)dt. \end{aligned} \tag{3.17}$$

The duality relation (3.11) follows immediately from combining (3.15)-(3.17) and (3.14). Let us turn to second duality relation (3.12). By applying integration by parts formula for jump process to $K^*(t)y_1^{\varepsilon}(t)$, we get

$$\begin{aligned} E\left(K^*(T)y_1^{\varepsilon}(T)\right) &= E\left(K^*(0)y_1^{\varepsilon}(0)\right) \\ &\quad + E\int_0^T K^*(t)dy_1^{\varepsilon}(t) + E\int_0^T y_1^{\varepsilon}(t)dK^*(t) \\ &\quad + E\int_0^T \int_\Theta z_1^{\varepsilon}(t)[g_z(t,\theta)K^*(t) + E\left(g_{\tilde{z}}(t,\theta)K^*(t)\right) \\ &\quad - l_z(t,\theta) - E\left(l_{\tilde{z}}(t,\theta)\right)]\mu(d\theta)dt \\ &\quad + E\int_0^T \int_\Theta r_1^{\varepsilon}(t,\theta)\left[g_r(t,\theta)K^*(t) - l_r(t,\theta)\right]\mu(d\theta)dt \\ &= I_1^{\varepsilon} + I_2^{\varepsilon} + I_3^{\varepsilon} + I_4^{\varepsilon} + I_5^{\varepsilon}. \end{aligned} \tag{3.18}$$

From (3.11), we get

$$\begin{aligned} I_2^{\varepsilon} &= E\int_0^T K^*(t)dy_1^{\varepsilon}(t) \\ &= -E\int_0^T \int_\Theta \{K^*(t)g_x(t,\theta)x_1^{\varepsilon}(t) + K^*(t)g_{\tilde{x}}(t,\theta)E\left(x_1^{\varepsilon}(t)\right) + K^*(t)g_y(t,\theta)y_1^{\varepsilon}(t) \\ &\quad + K^*(t)g_{\tilde{y}}(t,\theta)E\left(y_1^{\varepsilon}(t)\right) + K^*(t)g_z(t,\theta)z_1^{\varepsilon}(t) + K^*(t)g_{\tilde{z}}(t,\theta)E\left(z_1^{\varepsilon}(t)\right) \\ &\quad + K^*(t)g_r(t,\theta)r_1^{\varepsilon}(t,\theta) + K^*(t)g_u(t,\theta)u(t)\}\mu(d\theta)dt, \end{aligned} \tag{3.19}$$

from (3.4), we obtain

$$
\begin{aligned}
I_3^\varepsilon &= E\int_0^T y_1^\varepsilon(t)dK^*(t) \\
&= E\int_0^T\int_\Theta \{y_1^\varepsilon(t)g_y(t,\theta)K^*(t)+y_1^\varepsilon(t)E\left(g_{\tilde{y}}(t,\theta)K^*(t)\right) \\
&\quad -y_1^\varepsilon(t)l_y(t)-y_1^\varepsilon(t)E\left(l_{\tilde{y}}(t)\right)\}\mu(d\theta)dt,
\end{aligned}
\tag{3.20}
$$

and

$$
\begin{aligned}
I_4^\varepsilon &= E\int_0^T\int_\Theta \{z_1^\varepsilon(t)g_z(t,\theta)K^*(t)+z_1^\varepsilon(t)E\left(g_{\tilde{z}}(t,\theta)K^*(t)\right) \\
&\quad -z_1^\varepsilon(t)l_z(t,\theta)-z_1^\varepsilon(t)E\left(l_{\tilde{z}}(t,\theta)\right)\}\mu(d\theta)dt, \\
I_5^\varepsilon &= E\int_0^T\int_\Theta r_1^\varepsilon(t,\theta)\left[g_r(t,\theta)K^*(t)-l_r(t,\theta)\right]\mu(d\theta)dt.
\end{aligned}
\tag{3.21}
$$

Since

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

the duality relation (3.12) follows immediately by combining (3.19)-(3.21) and (3.18). Let us turn to (3.13). Combining (3.11) and (3.12) we get

$$
\begin{aligned}
& E\left(\Psi^*(T)x_1^\varepsilon(T)\right)+E\left(K^*(T)y_1^\varepsilon(T)\right) \\
= & -E\{[\varphi_y(y^*(0),E\left(y^*(0)\right))+E\left(\varphi_{\tilde{y}}(y^*(0),E\left(y^*(0)\right))\right)]y_1^\varepsilon(0)\} \\
& -E\int_0^T\int_\Theta \{x_1^\varepsilon(t)l_x(t,\theta)+x_1^\varepsilon(t)E\left(l_{\tilde{x}}(t,\theta)\right)+y_1^\varepsilon(t)l_y(t,\theta)+y_1^\varepsilon(t)E\left(l_{\tilde{y}}(t,\theta)\right) \\
& +z_1^\varepsilon(t)l_z(t,\theta)+z_1^\varepsilon(t)E\left(l_{\tilde{z}}(t,\theta)\right)+r_1^\varepsilon(t)l_r(t,\theta)+l_u(t,\theta)u(t)\}\mu(d\theta)dt \\
& +E\int_0^T H_u(t)u(t)dt.
\end{aligned}
$$

From (3.6) and (3.10), we get

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

Using (3.5), we obtain

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

which implies that

$$
\begin{aligned}
& E\left\{\left[\phi_x\left(x^*(T), E\left(x^*(T)\right)\right) + E\left(\phi_{\tilde{x}}\left(x^*(T), E\left(x^*(T)\right)\right)\right)\right] x_1^\varepsilon(t)\right\} \\
& +E\{[\varphi_y(y^*(0), E\left(y^*(0)\right)) + E\left(\varphi_{\tilde{y}}(y^*(0), E\left(y^*(0)\right))\right)] y_1^\varepsilon(0)\} \\
= \; & -E\int_0^T\int_\Theta x_1^\varepsilon(t) l_x(t,\theta) + x_1^\varepsilon(t) E\left(l_{\tilde{x}}(t,\theta)\right) + y_1^\varepsilon(t) l_y(t,\theta) + y_1^\varepsilon(t) E\left(l_{\tilde{y}}(t,\theta)\right) \\
& +z_1^\varepsilon(t) l_z(t,\theta) + z_1^\varepsilon(t) E\left(l_{\tilde{z}}(t,\theta)\right) + r_1^\varepsilon(t) l_r(t,\theta) + l_u(t,\theta)u(t)\}\mu(d\theta)dt \\
& +E\int_0^T H_u(t)u(t)dt.
\end{aligned}
$$

This completes the proof of (3.13). ■

The second Lemma presents the estimates of the perturbed state process $\left(x_1^\varepsilon(\cdot), y_1^\varepsilon(\cdot), z_1^\varepsilon(\cdot), r_1^\varepsilon(\cdot,\cdot)\right)$.

Lemma 3.3.2 *Under assumptions **(H1)** and **(H2)**, the following estimations hold*

$$
\begin{aligned}
& E\left(\sup_{0\le t\le T} \mid x_1^\varepsilon(t) \mid^2\right) \longrightarrow 0, \;\; as\ \varepsilon \longrightarrow 0, \\
& E\left(\sup_{0\le t\le T} \mid y_1^\varepsilon(t) \mid^2\right) + E\int_0^T [|z_1^\varepsilon(s)|^2 \\
& +\int_\Theta \left|r_1^\varepsilon(s,\theta)\right|^2 \mu(d\theta)]ds \longrightarrow 0, \;\; as\ \varepsilon \longrightarrow 0,
\end{aligned}
\tag{3.22}
$$

$$
\begin{aligned}
& \sup_{0\le t\le T} \left|E\left(x_1^\varepsilon(t)\right)\right|^2 \longrightarrow 0, \;\; as\ \varepsilon \longrightarrow 0, \\
& \sup_{0\le t\le T} \left|E\left(y_1^\varepsilon(t)\right)\right|^2 + \int_t^T \left|E\left(z_1^\varepsilon(s)\right)\right|^2 ds \\
& +\int_0^T\int_\Theta \left|E\left(r_1^\varepsilon(s,\theta)\right)\right|^2 \mu(d\theta)]ds \longrightarrow 0, \;\; as\ \varepsilon \longrightarrow 0,
\end{aligned}
\tag{3.23}
$$

$$
\begin{aligned}
&E\left(\sup_{0\leq t\leq T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^{2}\right)\longrightarrow 0,\ \ as\ \varepsilon\longrightarrow 0,\\
&E\left(\sup_{0\leq t\leq T}\left|y(t)-y^{*}(t)\right|^{2}\right)+E\int_{0}^{T}\left|z^{\varepsilon}(t)-z^{*}(t)\right|^{2}dt\\
&+E\int_{0}^{T}\int_{\Theta}\left|r^{\varepsilon}(t,\theta)-r^{*}(t,\theta)\right|^{2}\mu(d\theta)dt\longrightarrow 0,\ \ as\ \varepsilon\longrightarrow 0,
\end{aligned}
\tag{3.24}
$$

and

$$
\begin{aligned}
&E\left(\sup_{0\leq t\leq T}\left|\frac{1}{\varepsilon}\left[x^{\varepsilon}(t)-x^{*}(t)\right]-x_{1}^{\varepsilon}(t)\right|^{2}\right)\longrightarrow 0,\ \ as\ \varepsilon\longrightarrow 0,\\
&E\left(\sup_{0\leq t\leq T}\left|\frac{1}{\varepsilon}\left[y^{\varepsilon}(t)-y^{*}(t)\right]-y_{1}^{\varepsilon}(t)\right|^{2}\right)\longrightarrow 0,\ \ as\ \varepsilon\longrightarrow 0,\\
&E\int_{0}^{T}\left|\frac{1}{\varepsilon}\left[z^{\varepsilon}(s)-z^{*}(s)\right]-z_{1}^{\varepsilon}(s)\right|^{2}ds\longrightarrow 0,\ \ as\ \varepsilon\longrightarrow 0,\\
&+E\int_{0}^{T}\int_{\Theta}\left|\frac{1}{\varepsilon}\left[r^{\varepsilon}(s,\theta)-r^{*}(s,\theta)\right]-r_{1}^{\varepsilon}(s,\theta)\right|^{2}\mu(d\theta)ds\longrightarrow 0,\ \ as\ \varepsilon\longrightarrow 0.
\end{aligned}
\tag{3.25}
$$

Let us also point out that the above estimates (3.22)-(3.24) can be proved using similar arguments developed in ([[45], Lemma 4.2 and Lemma 4.3]) and ([[48], Lemma 2.1]). So we omit their proofs.

Proof We set

$$
\begin{aligned}
\hat{x}^{\varepsilon}(t) &= \frac{1}{\varepsilon}\left[x^{\varepsilon}(t)-x^{*}(t)\right]-x_{1}^{\varepsilon}(t),\\
\hat{y}^{\varepsilon}(t) &= \frac{1}{\varepsilon}\left[y^{\varepsilon}(t)-y^{*}(t)\right]-y_{1}^{\varepsilon}(t),\\
\hat{z}^{\varepsilon}(t) &= \frac{1}{\varepsilon}\left[z^{\varepsilon}(t)-z^{*}(t)\right]-z_{1}^{\varepsilon}(t),\\
\hat{r}^{\varepsilon}(t,\theta) &= \frac{1}{\varepsilon}\left[r^{\varepsilon}(t,\theta)-r^{*}(t,\theta)\right]-r_{1}^{\varepsilon}(t,\theta),
\end{aligned}
\tag{3.26}
$$

and

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

From (3.1) and since

$$x^{\varepsilon}(t) = x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right)$$

we have

$$\begin{aligned}
d\hat{x}^{\varepsilon}(t) &= \frac{1}{\varepsilon}\left[dx^{\varepsilon}(t)-dx^{*}(t)\right]-dx_{1}^{\varepsilon}(t)\\
&= \frac{1}{\varepsilon}\left[f(t,x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right),E\left(x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right)\right),u^{\varepsilon}(t))-f(t)\right]dt\\
&\quad-\left[f_{x}(t)x_{1}^{\varepsilon}(t)+f_{\tilde{x}}(t)E\left(x_{1}^{\varepsilon}(t)\right)+f_{u}(t)u(t)\right]dt\\
&\quad+\frac{1}{\varepsilon}[\sigma(t,x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right),E\left(x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right)\right),u^{\varepsilon}(t))\\
&\quad-\sigma(t)]dW(t)-\left[\sigma_{x}(t)x_{1}^{\varepsilon}(t)+\sigma_{\tilde{x}}(t)E\left(x_{1}^{\varepsilon}(t)\right)+\sigma_{u}(t)u(t)\right]dW\left(t\right)\\
&\quad+\frac{1}{\varepsilon}\int_{\Theta}[c(t,x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right),E\left(x^{*}(t)+\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right)\right),u^{\varepsilon}(t),\theta)\\
&\quad-c(t,\theta)]N(d\theta,dt)-\int_{\Theta}[c_{x}(t,\theta)x_{1}^{\varepsilon}(t)+c_{\tilde{x}}(t,\theta)E\left(x_{1}^{\varepsilon}(t)\right)+c_{u}(t,\theta)u(t)]N(d\theta,dt).
\end{aligned} \tag{3.27}$$

We denote

$$\begin{aligned}
x^{\lambda,\varepsilon}(t) &= x^{*}(t)+\lambda\varepsilon\left(\hat{x}^{\varepsilon}(t)+x_{1}^{\varepsilon}(t)\right),\\
y^{\lambda,\varepsilon}(t) &= y^{*}(t)+\lambda\varepsilon\left(\hat{y}^{\varepsilon}(t)+y_{1}^{\varepsilon}(t)\right),\\
z^{\lambda,\varepsilon}(t) &= z^{*}(t)+\lambda\varepsilon\left(\hat{z}^{\varepsilon}(t)+z_{1}^{\varepsilon}(t)\right),\\
r^{\lambda,\varepsilon}(t) &= r^{*}(t,\theta)+\lambda\varepsilon\left(\hat{r}^{\varepsilon}(t,\theta)+r_{1}^{\varepsilon}(t,\theta)\right),\\
u^{\lambda,\varepsilon}(t) &= u^{*}(t)+\lambda\varepsilon u(t).
\end{aligned} \tag{3.28}$$

By Taylor's expansion with a simple computation, we show that

$$\begin{aligned}
\hat{x}^{\varepsilon}(t) &= \frac{1}{\varepsilon}\left[x^{\varepsilon}(t)-x^{*}(t)\right]-x_{1}^{\varepsilon}(t)\\
&= \tilde{I}_{1}\left(\varepsilon\right)+\tilde{I}_{2}\left(\varepsilon\right)+\tilde{I}_{3}\left(\varepsilon\right),
\end{aligned} \tag{3.29}$$

where

$$\begin{aligned}
\tilde{I}_1(\varepsilon) &= \int_0^t\int_0^1 f_x(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))\hat{x}^{\varepsilon}(s)d\lambda ds \\
&+\int_0^t\int_0^1 f_{\tilde{x}}(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))E\left(\hat{x}^{\varepsilon}(s)\right)d\lambda ds \\
&+\int_0^t\int_0^1 \left[f_x(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))-f_x(s)\right]x_1^{\varepsilon}(s)d\lambda ds \\
&+\int_0^t\int_0^1 \left[f_{\tilde{x}}(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))-f_{\tilde{x}}(s)\right]E\left(x_1^{\varepsilon}(s)\right)d\lambda ds \\
&+\int_0^t\int_0^1 \left[f_u(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))-f_u(s)\right]u(s)d\lambda ds,
\end{aligned} \tag{3.30}$$

$$\begin{aligned}
\tilde{I}_2(\varepsilon) &= \int_0^t\int_0^1 \sigma_x(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))\hat{x}^{\varepsilon}(s)d\lambda dW(s) \\
&+\int_0^t\int_0^1 \sigma_{\tilde{x}}(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))E\left(\hat{x}^{\varepsilon}(s)\right)d\lambda dW(s) \\
&+\int_0^t\int_0^1 \left[\sigma_x(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))-\sigma_x(s)\right]x_1^{\varepsilon}(s)d\lambda dW(s) \\
&+\int_0^t\int_0^1 \left[\sigma_{\tilde{x}}(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))-\sigma_{\tilde{x}}(s)\right]E\left(x_1^{\varepsilon}(s)\right)d\lambda dW(s) \\
&+\int_0^t\int_0^1 \left[\sigma_u(s,x^{\lambda,\varepsilon}(s),E\left(x^{\lambda,\varepsilon}(s)\right),u^{\lambda,\varepsilon}(s))-\sigma_u(s)\right]u(s)d\lambda dW(s),
\end{aligned} \tag{3.31}$$

and

$$\begin{aligned}
&\tilde{I}_3(\varepsilon) \\
&= \int_0^t\int_\Theta\int_0^1 c_x(s,x^{\lambda,\varepsilon}(s-),E\left(x^{\lambda,\varepsilon}(s-)\right),u^{\lambda,\varepsilon}(s),\theta)\hat{x}^{\varepsilon}(s)d\lambda N(d\theta,ds) \\
&+\int_0^t\int_\Theta\int_0^1 c_{\tilde{x}}(s,x^{\lambda,\varepsilon}(s-),E\left(x^{\lambda,\varepsilon}(s-)\right),u^{\lambda,\varepsilon}(s),\theta)E\left(\hat{x}^{\varepsilon}(s)\right)d\lambda N(d\theta,ds) \\
&+\int_0^t\int_\Theta\int_0^1 \left[c_x(s,x^{\lambda,\varepsilon}(s-),E\left(x^{\lambda,\varepsilon}(s-)\right),u^{\lambda,\varepsilon}(s),\theta)-c_x(s,\theta)\right]x_1^{\varepsilon}(s)d\lambda N(d\theta,ds) \\
&+\int_0^t\int_\Theta\int_0^1 \left[c_{\tilde{x}}(s,x^{\lambda,\varepsilon}(s-),E\left(x^{\lambda,\varepsilon}(s-)\right),u^{\lambda,\varepsilon}(s),\theta)-c_{\tilde{x}}(s,\theta)\right]E\left(x_1^{\varepsilon}(s)\right)d\lambda N(d\theta,ds) \\
&+\int_0^t\int_\Theta\int_0^1 \left[c_u(s,x^{\lambda,\varepsilon}(s-),E\left(x^{\lambda,\varepsilon}(s-)\right),u^{\lambda,\varepsilon}(s),\theta)-c_u(s,\theta)\right]u(s)d\lambda N(d\theta,ds),
\end{aligned} \tag{3.32}$$

we proceed as in Anderson and Djehiche [1], we get

$$E\left(\sup_{0\leq t\leq T}\left|\tilde{I}_1(\varepsilon)\right|^2\right)\longrightarrow 0,\ \text{as}\ \varepsilon\longrightarrow 0,$$
$$E\left(\sup_{0\leq t\leq T}\left|\tilde{I}_2(\varepsilon)\right|^2\right)\longrightarrow 0,\ \text{as}\ \varepsilon\longrightarrow 0. \tag{3.33}$$

Applying similar estimations for the third term with the help of Proposition 3.2 (in Appendix Bouchard and Elie [8]), we have

$$E\left(\sup_{0\leq t\leq T}\left|\tilde{I}_3(\varepsilon)\right|^2\right)\longrightarrow 0,\ \text{as}\ \varepsilon\longrightarrow 0. \tag{3.34}$$

From (3.33) and (3.34), we obtain

$$E\left(\sup_{0\leq t\leq T}\left|\frac{1}{\varepsilon}\left[x^\varepsilon(t)-x^*(t)\right]-x_1^\varepsilon(t)\right|^2\right)\longrightarrow 0,\ \text{as}\ \varepsilon\longrightarrow 0. \tag{3.35}$$

We proceed to estimate the last terms in (3.25). First, from (3.26) and since

$$\hat{y}^\varepsilon(t) = \frac{1}{\varepsilon}\left[y^\varepsilon(t)-y^*(t)\right]-y_1^\varepsilon(t),$$

we get

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

and

$$\begin{aligned}
\hat{y}^\varepsilon(T) &= \frac{1}{\varepsilon}\left[h\left(x^\varepsilon(T),E\left(x^\varepsilon(T)\right)\right)-h\left(x^*(T),E\left(x^*(T)\right)\right)\right]\\
&\quad+\left[h_x\left(x^*(T),E\left(x^*(T)\right)\right)-h_{\tilde{x}}\left(x^*(T),E\left(x^*(T)\right)\right)\right]x_1^\varepsilon(T).
\end{aligned}$$

Applying Taylor's expansion, we get

$$
\begin{aligned}
-d\hat{y}^{\varepsilon}(t) \;=\;& \int_{\Theta}\int_0^1 g_x(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right) \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times \hat{x}^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 g_{\tilde{x}}(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right) \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times E\left(\hat{x}^{\varepsilon}(t)\right) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 [g_x(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right) \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_x(t,\theta)] x_1^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 [g_{\tilde{x}}(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right) \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_{\tilde{x}}(t,\theta)] E\left(x_1^{\varepsilon}(t)\right) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 g_u(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right) \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_u(t,\theta)] u(t) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 g_y(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right), \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times \hat{y}^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 g_{\tilde{y}}(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right), \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times E\left(\hat{y}^{\varepsilon}(t)\right) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 [g_y(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right) \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_y(t,\theta)] y_1^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt \\
& + \int_{\Theta}\int_0^1 [g_{\tilde{y}}(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right), \\
& r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_{\tilde{y}}(t,\theta)] E\left(y_1^{\varepsilon}(t)\right) d\lambda \mu\left(d\theta\right) dt
\end{aligned}
$$

$$+\int_{\Theta}\int_{0}^{1} g_z(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right),$$

$$r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times \hat{z}^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt$$

$$+\int_{\Theta}\int_{0}^{1} g_{\tilde{z}}(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right),$$

$$r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times E\left(\hat{z}^{\varepsilon}(t)\right) d\lambda \mu\left(d\theta\right) dt$$

$$+\int_{\Theta}\int_{0}^{1} [g_z(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right),$$

$$r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_z(t,\theta)] z_1^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt$$

$$+\int_{\Theta}\int_{0}^{1} [g_{\tilde{z}}(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right)$$

$$r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_{\tilde{z}}(t,\theta)] E\left(z_1^{\varepsilon}(t)\right) d\lambda \mu\left(d\theta\right) dt$$

$$+\int_{\Theta}\int_{0}^{1} g_r(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right)$$

$$r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) \times \hat{r}^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt$$

$$+\int_{\Theta}\int_{0}^{1} [g_r(t, x^{\lambda,\varepsilon}(t), E\left(x^{\lambda,\varepsilon}(t)\right), y^{\lambda,\varepsilon}(t), E\left(y^{\lambda,\varepsilon}(t)\right), z^{\lambda,\varepsilon}(t), E\left(z^{\lambda,\varepsilon}(t)\right)$$

$$r^{\lambda,\varepsilon}(t,\theta), u^{\lambda,\varepsilon}(t)) - g_r(t,\theta)] r_1^{\varepsilon}(t) d\lambda \mu\left(d\theta\right) dt - \hat{z}^{\varepsilon}(t) dW(t) - \int_{\Theta} \hat{r}^{\varepsilon}(t) N\left(d\theta d, dt\right),$$

finally, using similar arguments developed in [48], the desired result follows. This completes the proof of (3.25). ■

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

$$E\int_0^T \int_{\Theta} [l_x(t,\theta) x_1^{\varepsilon}(t) + l_{\tilde{x}}(t,\theta) E\left(x_1^{\varepsilon}(t)\right) + l_y(t,\theta) y_1^{\varepsilon}(t) + l_{\tilde{y}}(t,\theta) E\left(y_1^{\varepsilon}(t)\right)$$

$$l_z(t,\theta) z_1^{\varepsilon}(t) + l_{\tilde{z}}(t,\theta) E\left(z_1^{\varepsilon}(t)\right) + l_r(t,\theta) r_1^{\varepsilon}(t,\theta) + l_u(t,\theta) u(t)] \mu\left(d\theta\right) dt$$

$$+E\left[\phi_x\left(T\right) x_1^{\varepsilon}(T) + \phi_{\tilde{x}}\left(T\right) E\left(x_1^{\varepsilon}(T)\right)\right] + E\left[\varphi_y(0) y_1^{\varepsilon}(0) + \varphi_{\tilde{y}}(0) E\left(y_1^{\varepsilon}(0)\right)\right] \geq o(\varepsilon).$$

Proof From (3.9), we have

$$\begin{aligned}
&J(u^{\varepsilon}(\cdot)) - J(u^{*}(\cdot)) \\
= \; & E\{\int_0^T \int_\Theta [l(t, x^{\varepsilon}(t), E\left(x^{\varepsilon}(t)\right), y^{\varepsilon}(t), E\left(y^{\varepsilon}(t)\right), z^{\varepsilon}(t), E\left(z^{\varepsilon}(t)\right), r^{\varepsilon}(t,\theta), u^{\varepsilon}(t)) \\
& - l(t, x^{*}(t), E\left(x^{*}(t)\right), y^{*}(t), E\left(y^{*}(t)\right), z^{*}(t), E\left(z^{*}(t)\right), r^{*}(t,\theta), u^{*}(t))]\mu\left(d\theta\right) dt \\
& + [\phi(x^{\varepsilon}(T), E\left(x^{\varepsilon}(T)\right)) - \phi(x^{*}(T), E\left(x^{*}(T)\right))] \\
& + [\varphi(y^{\varepsilon}(0), E\left(y^{\varepsilon}(0)\right)) - \varphi(y^{*}(0), E\left(y^{*}(0)\right))]\} \geq 0.
\end{aligned} \tag{3.36}$$

By applying Taylor's expansion and Lemma (3.3.2), we have

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

From estimate (3.25), we get

$$\begin{aligned}
& \frac{1}{\varepsilon} E\left[\phi(x^{\varepsilon}(T), \tilde{x}^{\varepsilon}(T)) - \phi(x^{*}(T), \tilde{x}^{*}(T))\right] \\
& \longrightarrow E\left[\phi_x(x^{*}(T), E\left(x^{*}(T)\right)) x_1^{\varepsilon}(T), \phi_{\tilde{x}}(x^{*}(T), E\left(x^{*}(T)\right)) E\left(x_1^{\varepsilon}(T)\right)\right] \\
& = E\left[\phi_x\left(T\right) x_1^{\varepsilon}(T) + \phi_{\tilde{x}}\left(T\right) E\left(x_1^{\varepsilon}(T)\right)\right], \text{ as } \varepsilon \longrightarrow 0.
\end{aligned} \tag{3.37}$$

Similarly, we have

$$\begin{aligned}
& \frac{1}{\varepsilon} E\left[\varphi(y^{\varepsilon}(0), \tilde{y}^{\varepsilon}(0)) - \varphi(y^{*}(0), \tilde{y}^{*}(0))\right] \\
& \longrightarrow E\left[\varphi_y(y^{*}(0), E((y^{*}(0))) y_1^{\varepsilon}(0), \varphi_{\tilde{y}}(y^{*}(0), E((y^{*}(0))) E\left(y_1^{\varepsilon}(0)\right)\right] \\
& = E\left[\varphi_y\left(0\right) y_1^{\varepsilon}(0) + \varphi_{\tilde{y}}\left(0\right) E\left(y_1^{\varepsilon}(0)\right)\right], \text{ as } \varepsilon \longrightarrow 0,
\end{aligned} \tag{3.38}$$

and

$$\begin{aligned}
& E\int_0^T\int_\Theta [l(t,x^\varepsilon(t),E\left(x^\varepsilon(t)\right),y^\varepsilon(t),E\left(y^\varepsilon(t)\right),z^\varepsilon(t),E\left(z^\varepsilon(t)\right),r^\varepsilon(t,\theta),u^\varepsilon(t)) \\
& -l(t,x^*(t),E\left(x^*(t)\right),y^*(t),E\left(y^*(t)\right),z^*(t),E\left(z^*(t)\right),r^*(t,\theta),u^*(t))]\mu\left(d\theta\right)dt \\
& \longrightarrow E\int_0^T\int_\Theta [l_x(t,\theta)x_1^\varepsilon(t)+l_{\tilde{x}}(t,\theta)E\left(x_1^\varepsilon(t)\right)+l_y(t,\theta)y_1^\varepsilon(t)+l_{\tilde{y}}(t,\theta)E\left(y_1^\varepsilon(t)\right) \\
& +l_z(t,\theta)z_1^\varepsilon(t)+l_{\tilde{z}}(t,\theta)E\left(z_1^\varepsilon(t)\right)+l_r(t,\theta)r_1^\varepsilon(t,\theta)+l_u(t,\theta)u(t)]\mu\left(d\theta\right)dt, \\
& \text{as } \varepsilon\longrightarrow 0.
\end{aligned} \tag{3.39}$$

The desired result follows by combining (3.36)∼(3.39), This completes the proof of Lemma (3.3.3). ∎

Proof of Theorem (3.3.1): The desired result follows immediately by combining (3.13) in Lemma (3.3.2) and Lemma (3.3.3).

3.4 Application: Mean-variance portfolio selection problem mixed with a recursive utility functional

The mean-variance portfolio selection theory, which was first proposed in Markowitz [32] is a milestone in mathematical finance and has laid down the foundation of modern finance theory. Using sufficient maximum principle, the authors in [17] gave the expression for the optimal portfolio selection in a jump diffusion market. The near-optimal consumption-investment problem has been discussed in Hafayed, Abbas, and Veverka [26]. The continuous time mean-variance portfolio selection problem has been studied in Zhou and Li [61]. The mean-variance portfolio selection problem where the state driven by SDE (without jump terms) has been studied in [1]. Optimal dividend, harvesting rate, and optimal portfolio for systems governed by jump diffusion processes have been investigated in [35]. Mean-variance portfolio selection problem mixed with a recursive utility functional has been studied by Shi and Wu [48], under the condition that

$$E\left(x^\pi(T)\right)=c,$$

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 optimization problem 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

$$\begin{cases} dP_0(t) = \rho(t)P_0(t)dt, t \in [0,T] \\ P_0(0) > 0, \end{cases} \tag{3.40}$$

   where $\rho(\cdot) : [0,T] \longrightarrow \mathbb{R}^+$ is a locally bounded and continuous deterministic function.

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

$$\begin{cases} dP_1(t) = P_1(t-)\left[\zeta(t)dt + G(t)dW(t) + \displaystyle\int_\Theta \xi\left(t,\theta\right) N\left(d\theta, dt\right)\right], \\ P_1(0) > 0, t \in \left[0,T\right]. \end{cases} \tag{3.41}$$

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

1. The functions $\zeta(\cdot) : [0,T] \longrightarrow \mathbb{R}, G(\cdot) : [0,T] \longrightarrow \mathbb{R}$ are bounded deterministic such taht

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

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

3. $\displaystyle\int_\Theta \xi^2(t,\theta)\mu\left(d\theta\right)$ 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 $\pi(t) = e_2(t)P_1(t)$ denote the amount invested in the risky security. We call the control process $\pi(\cdot)$ a portfolio strategy.

Let $x^\pi(0) = \xi$ be an initial wealth. By combining (3.40) and (3.41), 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

$$\begin{cases} dx^\pi(t) = & [\rho(t)x^\pi(t) + (\zeta(t) - \rho(t))\pi(t)]\,dt \\ & +G(t)\pi(t)dW(t) + \int_\Theta \xi(t,\theta)\,N(d\theta,dt), \\ -dy^\pi(t) = & [\rho(t)x^\pi(t) + (\zeta(t) - \rho(t))\pi(t) - \alpha y^\pi(t)]\,dt \\ & -z^\pi(t)dW(t) - \int_\Theta r^\pi(t,\theta)\,N(d\theta,dt), \\ x^\pi(0) = & \xi, y^\pi(T) = x^\pi(T). \end{cases} \tag{3.42}$$

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

$$J(\pi(\cdot)) = \frac{\gamma}{2}Var(x^\pi(T)) - E(x^\pi(T)) - y^\pi(0). \tag{3.43}$$

By a simple computation, we can show that

$$J(\pi(\cdot)) = E\left[\frac{\gamma}{2}x^\pi(T)^2 - x(T)\right] - \frac{\gamma}{2}\left[E\left(x^\pi(T)\right)\right]^2 - y^\pi(0), \tag{3.44}$$

where the wealth process $x^\pi(\cdot)$ and the recursive utility process $y^\pi(\cdot)$ corresponding $\pi(\cdot) \in \mathcal{U}([0,T])$ are given by FBSDEJ-(3.42). We note that the cost functional (3.44) becomes a time-inconsistent control problem. Let $\mathcal{A}$ be a compact convex subset of $\mathbb{R}$. 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 (3.5) gets the form

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

According to the maximum condition ((3.8), Theorem (3.3.1)), and since $\pi^{*}(\cdot)$ is optimal we immediately get,

$$
\begin{aligned}
&[(\zeta(t)-\rho(t))]\,(\Psi^{*}(t)-K^{*}(t))+G(t)Q^{*}(t)\\
&+\int_{\Theta}\xi\,(t,\theta)\,R^{*}(t,\theta)\mu\,(d\theta)=0.
\end{aligned}
\tag{3.45}
$$

The adjoint equation (3.4) being

$$
\begin{cases}
d\Psi^{*}(t) &= -\rho(t)\,(\Psi^{*}(t)-K^{*}(t))\,dt+Q^{*}(t)dW(t)\\
&\quad +\displaystyle\int_{\Theta}R^{*}\,(t,\theta)\,N\,(d\theta,dt)\,,\\
\Psi^{*}(T) &= \gamma(x^{*}(T)-E(x^{*}(T))-1-K^{*}(T)\\
dK^{*}(t) &= -\alpha K^{*}(t)dt,K^{*}(0)=1,t\in[0,T]\,.
\end{cases}
\tag{3.46}
$$

In order to solve the above equation (3.46) and to find the expression of optimal portfolio strategy $\pi^{*}(\cdot)$, we conjecture a process $\Psi^{*}(t)$ of the form:

$$
\Psi^{*}(t) = A_1(t)x^{*}(t)+A_2(t)E\,(x^{*}(t))+A_3(t),
\tag{3.47}
$$

where $A_1(\cdot),A_2(\cdot)$ and $A_3(\cdot)$ deterministic differentiable functions.(See Shi and Wu [48], Shi [47], Framstad, Øksendal and Sulem [17]).

From last equation in (3.46), which is a simple ordinary differential equation (ODE in short), we get immediately

$$
K^{*}(t) = \exp(-\alpha t).
\tag{3.48}
$$

Noting that from (3.42), we get

$$
d\,(E\,(x^{*}(t))) = \{\rho(t)E\,(x^{*}(t))+(\zeta(t)-\rho(t))\,E\,(\pi^{*}(t))\}dt.
$$

Applying Itô's formula to (3.47) in virtue of SDE-(3.42), we get

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

which implies that

$$\begin{cases} d\Psi^*(t) &= \{A_1(t)[\rho(t)x^*(t) + (\zeta(t) - \rho(t))\,\pi^*(t)] + x^*(t)A_1^{'}(t) \\ & \quad +A_2(t)[\rho(t)E\,(x^*(t)) + (\zeta(t) - \rho(t))\,E\,(\pi^*(t))] \\ & \quad +A_2^{'}(t)E\,(x^*(t)) + A_3^{'}(t)\}dt \\ & \quad +A_1(t)G(t)\pi^*(t)dW(t) + \int_\Theta A_1(t)\xi\,(t-,\theta)\,\pi^*(t)N\,(d\theta,dt) \\ \Psi^*(T) &= A_1(T)x^*(T) + A_2(t)E(x^*(T)) + A_3(T), \end{cases} \tag{3.49}$$

where $A_1^{'}(t), A_2^{'}(t)$ and $A_3^{'}(t)$ denote the derivatives with respect to t.
Next, comparing (3.49) with (3.46), we get

$$\begin{aligned} -\rho(t)\,(\Psi^*(t) - K^*(t)) &= A_1(t)[\rho(t)x^*(t) + (\zeta(t) - \rho(t))\,\pi^*(t)] + x^*(t)A_1^{'}(t) \\ & \quad +A_2(t)[\rho(t)E\,(x^*(t)) + (\zeta(t) - \rho(t))\,E\,(\pi^*(t))] \\ & \quad +A_2^{'}(t)E\,(x^*(t)) + A_3^{'}(t), \end{aligned} \tag{3.50}$$

$$Q^*(t) = A_1(t)G(t)\pi^*(t), \tag{3.51}$$

$$R^*(t) = A_1(t)\xi\,(t-,\theta)\,\pi^*(t). \tag{3.52}$$

By looking at the terminal condition of $\Psi^*(t)$, in (3.49), it is reasonable to get

$$A_1(T) = \gamma, A_2(T) = -\gamma, A_3(T) = -1 - K^*(T) \tag{3.53}$$

Combining (3.50) and (3.47), we deduce that $A_1^{'}(\cdot), A_2^{'}(\cdot)$ and $A_3^{'}(\cdot)$ satisfying the following

ODEs:

$$\begin{cases} A_1'(t) = -2\rho(t)A_1(t), & A_1(T) = \gamma, \\ A_2'(t) = -2\rho(t)A_2(t), & A_2(T) = -\gamma, \\ A_3'(t) + \rho(t)A_3(t) = \rho(t)\exp(-\alpha t), & A_3(T) = -\exp(-\alpha T) - 1. \end{cases} \tag{3.54}$$

By solving the first two ordinary differential equations in (3.54), we obtain

$$A_1(t) = -A_2(t) = \gamma \exp\left\{2\int_t^T \rho(s)ds\right\}. \tag{3.55}$$

Using integrating factor method for the third equation in (3.54), we get

$$A_3(t) = -\chi(t)^{-1}\left[\exp(-\alpha T) + 1 + \int_t^T \chi(s)\rho(s)\exp(-\alpha s)ds\right], \tag{3.56}$$

where the integrating factor is $\chi(t) = \displaystyle\int_t^T \rho(s)ds, \chi(T) = 1$.
Combining (3.45), (3.48), (3.51), and (3.52) and denoting

$$\Gamma(t) = A_1(t)\left(G^2(t) + \int_\Theta \xi^2(t,\theta)\mu(d\theta)\right) \tag{3.57}$$

we get

$$\pi^*(t) = \Gamma(t)^{-1}(\rho(t) - \zeta(t))\left[A_1(t)\left(x^*(t) - E\left(x^*(t)\right)\right) + A_3(t) - \exp(-\alpha t)\right], \tag{3.58}$$

and

$$E\left(\pi^*(t)\right) = \Gamma(t)^{-1}(\rho(t) - \zeta(t))\left[A_3(t) - \exp(-\alpha t)\right] \tag{3.59}$$

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 (3.42)-(3.44) is given in feedback form by*

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

and

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

where $A_1(t)$, $A_3(t)$ and $\Gamma(t)$ are given by (3.55), (3.56) and (3.57) respectively.

Conclusion

In this work, we have discussed the necessary conditions for optimal stochastic control of mean-field forward-backward stochastic differential equations with Poisson jumps (FBSDEJs). The cost functional is also of mean field type. 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.

Bibliography

[1] Andersson, D., Djehiche, B.: A maximum principle for SDEs of mean-field type. Appl. Math. Optim. 63, 341-356 (2011).

[2] Arkin, V.I. and Saksonov, M.T.: Necessary optimality conditions for stochastic differential equations, Soviet. Math. Dokl., Vol. 20, 1-5 (1979).

[3] Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, (1957).

[4] Bensoussan, A., Lectures on stochastic control, Nonlinear Filtering and Stochastic Control (S.K. Mitter, A. Moro, eds.). Lecture Notes in Math. 972, Springer-Verlag (1982).

[5] Bensoussan, A.: Stochastic maximum principle for distributed parameter system, J. of the Franklin Inst., 315, 387-406 (1983).

[6] Bismut, J. M.: An introductory approach to duality in optimal stochastic control. SIAM Rev. 20(1), 62-78 (1978).

[7] Bismut, J. M.: Conjugate convex functions in optimal stochastic control. J Math Anal Appl, 44: 384–404 (1973).

[8] Bouchard, B., Elie, R.: Discrete time approximation of decoupled forward–backward SDE with jumps. Stoch. Proc. Appl. 118(1), 53-75 (2008).

[9] Buckdahn, R., Li, J., Peng, S.: Mean-field backward stochastic differential equations and related partial differential equations. Stoch. Proc. Appl 119, 3133–3154 (2009).

[10] Buckdahn, R., Djehiche, B., Li, J.: A general stochastic maximum principle for SDEs of mean-field type. Appl. Math. Optim. 64, 197–216 (2011).

[11] Elliott, R.J., Li, X., Ni, Y.H.: Discrete time mean-field stochastic linear-quadratic optimal control problems. Automatica 49(11), 3222–3233 (2013).

[12] El Karoui, N., Peng, S., Quenez, M.C.: A dynamic maximum principle for the optimization of recursive utilities under constraints, Ann. Appl. Prob., 11, 664-693 (2001).

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

[14] Fleming W.H. and Rishel R.W.: Deterministic and stochastic optimal control, Springer Verlag. New York (1975).

[15] Fleming W.H. and Soner H.M.: Controlled Markov processes and viscosity solution. Springer Verlag. New York (1993).

[16] Framstad, N.C., ∅ksendal, B., Sulem, A.: Optimal Consumption and Portfolio in a Jump Diffusion Market with Proportional Transaction Costs, Journal of Mathematical Economics, 35, 233-257 (2001).

[17] Framstad, N.C., ∅ksendal, B., Sulem, A.: Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. J. Optim. Theory Appl. 121, 77–98 (2004).

[18] Haussmann, U. G.: A Stochastic maximum principle for optimal control of diffusions. Pitman Research Notes in Math. Series 151 (1986).

[19] Hafayed, M., Abbas, S.: On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality. J. Optim. Theory Appl. 160, 778–808 (2014).

[20] Hafayed, M.: A mean-field necessary and sufficient conditions for optimal singular stochastic control. Commun. Math. Stat. 1(4), 417–435 (2014).

[21] Hafayed, M.: A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes. Int. J. Dyn. Control 1(4), 300–315 (2013).

[22] Hafayed, M., Abba, A., Abbas, S.: On mean-field stochastic maximum principle for near-optimal controls for Poisson jump diffusion with applications. Int. J. Dyn. Control 2, 262–284 (2014).

[23] Hafayed,M.: Singular mean-field optimal control for forward-backward stochastic systems and applications to finance. Int. J. Dyn. Control 2(4), 542–554 (2014).

[24] Hafayed, M., Abbas, S.: A general maximum principle for stochastic differential equations of mean-field type with jump processes, arXiv: 1301.7327v4, (2013).

[25] 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(2), 163-186 (2015)

[26] Hafayed, M., Veverka, P., Abbas, S.: On maximum principle of near-optimality for diffusions with jumps, with application to consumption-investment problem. Diff. Equ. Dyn. Syst. 20(2), 111-125 (2012).

[27] Kac, M.: Foundations of kinetic theory. In: Berkeley Symposium on Mathematical Statistics and Probability, vol. 3, 171–197 (1956).

[28] Kushner, H.J.: Necessary conditions for continuous parameter stochastic optimization problems, SIAM J. Control, 10, 550-565 (1972).

[29] Li, J.: Stochastic maximum principle in the mean-field controls. Automatica 48, 366–373 (2012).

[30] Li, T., Zhang, J.F.: Adaptive mean field games for large population coupled ARX systems with unknown coupling strength. Dyn. Games Appl. 3, 489–507 (2013).

[31] Lions, P. L.: Optimal control of diusion processes and Hamilton-Jacobi-Bellman equations. 1: The dynamic programming principle and applications; 2: Viscosity solutions and uniqueness. Communications in Partial Differential Equations 8: 1101-1174, 1229-1276, (1983).

[32] Markowitz, H.: Portfolio selection. J. Fin. 7, 77–91 (1952).

[33] Ma, J., Yong, J.: Forward-backward stochastic differential equations and their applications. Lecture notes in mathematics, vol 1702. Springer, Berlin (1999).

[34] McKean, H.P.: A class ofMarkov processes associated with nonlinear parabolic equations. Proc. Natl. Acad. Sci. USA 56, 1907–1911 (1966).

[35] Meyer-Brandis, T., Øksendal, B., Zhou,X.Y.:A mean-field stochasticmaximumprinciple viamalliavin calculus. Stochastics 84, 643–666 (2012).

[36] Ni, Y.H., Zhang, J.F., Li, X.: Indefinite mean-field stochastic linear-quadratic optimal control. IEEE Trans. Autom. Control (2014). doi:10.1109/TAC.2014.2385253.

[37] Øksendal, B., Sulem, A.: Maximum principles for optimal control of forward-backward stochastic differential equations with jumps. SIAM J. Control Optim. 48(5), 2845-2976 (2009).

[38] Øksendal, B. and Sulem, A.: Applied Stochastic Control of Jump Diffusions. Springer, 2nd edition, (2007).

[39] Pardoux, E. Peng, S.: Adapted solution of a backward stochasticdiferential equation. Syst Control Lett, 14: 55–61 (1990).

[40] Peng, S.: A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim. 28(4), 966-979 (1990).

[41] Peng, S.: Backward stochastic differential equation and application to optimal control, Applied Mathematics and Optimization, 27, 125-144 (1993).

[42] Pham H.: Optimisation et contrôle stochastique appliqués à la finance, to appear, Series Mathématiques et Applications, Springer Verlag (2005).

[43] Pontrygain, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., Mischenko, E. F.: The mathematical theory of optimal control processes, John Wiley, New York, (1962).

[44] Rishel, R.: A minimum principle for controlled jump processes, Lecture Notes in Economics and Mathematical Systems, Vol. 107, Springer-Verlag, Berlin, Heidelberg, New York, 493-508 (1975).

[45] Shen, Y., Siu, T.K.: The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem. Nonlinear Anal. 86, 58–73 (2013).

[46] 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).

[47] Shi, J.: Sufficient conditions of optimality for mean-field stochastic control problems. In: 12th International Conference on Control, Automation, Robotics & Vision Guangzhou, 5-7 Dec, ICARCV 2012, 747–752 (2012).

[48] 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).

[49] Shi, J.: Necessary conditions for optimal control of forward-backward stochastic systems with random jumps. Int. J. Stoch. Anal. Article ID 258674, 50 (2012).

[50] Shi, J. T. and Wu, Z.: The maximum principle for fully coupled forward-backward stochastic control system, Acta Automatica Sinica, 32(2): 161–169 (2006).

[51] Situ. R.: A maximum principle for optimal controls of stochastic systems with random jumps, in Proc. National Conference on Control Theory and Its Applications, Qingdao, Shandong, People's Republic of China, October (1991).

[52] Tang S. J. and Li X. J., Necessary conditions for optimal control of stochastic systems with random jumps, SIAM J. on Control and Optimization, 32(5): 1447–1475.(1994).

[53] Wang, B.C., Zhang, J.F.:Mean-field games for large-populationmultiagent systems with Markov jump parameters. SIAM J. Control 50(4), 2308–2334 (2012).

[54] Wang, G., Zhang, C., Zhang, W.: Stochastic maximum principle for mean-field type optimal control under partial information. IEEE Trans. Autom. Control 59(2), 522–528 (2014).

[55] Wu, Z.: Maximum principle for optimal control problem of fully coupled forward-backward systems. Syst. Sci. Math. Sci. 11(3), 249-259 (1998).

[56] Xu, R., Wu, T.: Mean-field backward stochastic evolution equations in Hilbert spaces and optimal control for BSPDEs. Abstr. Appl. Anal. 2014, Article ID 839467, 15 (2014).

[57] Xu, W. S.: Stochastic maximum principle for optimal control problem of forward and backward system. Jor. Aus. Math. Socie., B37, 172-185 (1995).

[58] Yong, J.: A linear-quadratic optimal control problem for mean-field stochastic differential equations. SIAM J. Control Optim. 54(4), 2809-2838 (2013).

[59] 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).

[60] Yong.J and Zhou X.Y., Stochastic controls, Springer (1999).

[61] Zhou, X.Y., Li, D.: Continuous time mean-variance portfolio selection: a stochastic LQ framework. Appl. Math. Optim. 42, 19-33 (2000).

Appendix A: Abbreviations and Notations

The following notation is frequently used in this thesis

$\mathbb{P}-a.s.$	Almost surely for the probability measure $\mathbb{P}$.
$a.e.$	Almost everywhere.
e.g	for example
u^{ε}	perturbed control.
u^{*}	Optimal control.
W	Brownian motion.
$N\left(\cdot,\cdot\right)$	Poisson martingale measure.
$\mathcal{G}_0$	The totality of $\mathbb{P}-$ null sets.
$\mathbb{R}^n$	n-dimensional real Euclidean space.
$\mathbb{R}^{n\times d}$	The set of all $(n\times d)$ real matrixes.
$(\Omega,\mathcal{F},\mathbb{P})$	Probability space.
$\{\mathcal{F}_t\}_{t\in[0,T]}$	Filtration.
$(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P})$	Filtered probability space.
$L^2\left(\Omega,\mathcal{F}_T;\mathcal{H}\right)$	The space of all $\mathcal{H}-$ valued squared integrable $\mathcal{F}_T-$ measurable random variables.
$L^2_{\mathcal{F}}\left([0,T];\mathcal{H}\right)$	The Hilbert space of all $\mathcal{H}-$ valued $\{\mathcal{F}_t\}_{t\in[0,T]}-$ adapted processes $(X(t))_{t\in[0,T]}$ such that $E\int_0^T \left\vert X\left(t\right)\right\vert^2 dt<\infty$.

$\mathcal{M}^2_{\mathcal{F}}\left([0,T],\mathcal{H}\right)$	The sHilbert pace of all $\mathcal{H}-$ valued $\{\mathcal{F}_t\}_{t\in[0,T]}-$ predictable processes $(\psi(t,\theta))_{t\in[0,T]}$ defined on $[0,T]\times\Theta$ such that $E\int_0^T\int_\Theta\mid\psi(t,\theta)\mid^2\mu(d\theta)\,dt<\infty.$
$\mathcal{U}^s\left([0,T]\right)$	The set of (stochastic) strong admissible controls.
$\mathcal{U}^w\left([0,T]\right)$	The set of (stochastic) weak admissible controls.
$E(\cdot)$	The expectation with respect to $\mathbb{P}$.
SDE	Stochastic differential equation.
SDEJ	Stochastic differential equation with Poisson jump processes.
ODE	Ordinary differential equation.
BSDE	Backward stochastic differential equations.
FBSDEJs	Forward-backward stochastic differential equations with Poisson jump processes.
SPDE	stochastic partial differential equation.

Appendix B

Proposition

Let f be real function defined and derivable in the interval I, so the following properties are equivalent :

i) f is convex on I.

ii) $\forall\,(x,y)\in I, x<y,\;\; f(y)\geq f(x)+f_x(x)(y-x)$.

Gronwall Lemma

Let f and g be non negative integrable functions and c a nonnegative constant. if

$$f(t)\leq c+\int_0^t g(s)f(s)ds \qquad \text{for } t\geq 0,$$

then

$$f(t)\leq c\exp\left(\int_0^t g(s)ds\right) \qquad \text{for } t\geq 0.$$

Lemma: (Integration by parts formula for mean-field jump diffusions)

Suppose that the processes $x_1(t)$ and $x_2(t)$ are given by for $i=1,2, t\in[0,T]$

$$\begin{cases} dx_i(t) &= f(t,x_i(t),E(x_i(t)),u(t))dt+\sigma(t,x_i(t),E(x_i(t)),u(t))dW(t) \\ & \quad +\int_\Theta g(t,x_i(t-),E(x_i(t-)),u(t))N\left(d\theta,dt\right), \\ x_i(0) &= 0. \end{cases}$$

Then we get

$$
\begin{aligned}
& E\left[x_1(t)x_2(t)\right] \\
= \;& E\int_0^T x_1(t)dx_2(t) + E\int_0^T x_2(t)dx_1(t) \\
& + E\int_0^T \sigma(t, x_1(t), E(x_1(t)), u(t))\sigma(t, x_2(t), E(x_2(t)), u(t))dt \\
& + E\int_0^T \int_\Theta g(t, x_1(t), E(x_1(t)), u(t), \theta)g(t, x_2(t), E(x_2(t)), u(t), \theta)\mu(d\theta)dt.
\end{aligned}
$$

See Framstad et al., ([17] Lemma 2.1), for the proof of the above Lemma.

Integrating factor method

The integrating factor method for solving partial differential equations may be used to solve linear, first order differential equations of the form:

$$\frac{dy}{dt} + a(t)y = b(t)$$

where $a(t)$ and $b(t)$ are functions of t. and in some cases may be constants.
We will say that an equation written in the above way is written in the standard form. The method for solving linear, first order differential equations using the integrating factor method may be broken down into the following steps:

1. Write the differential equation in the standard form: $\dfrac{dy}{dt} + a(t)y = b(t)$.

2. Determine the integrating factor,

$$\text{Integrating Factor} = \exp\left(\int a(t)dt\right).$$

3. Multiply the equation in standard form by the integrating factor,

$$\exp\left(\int a(t)dt\right)\left(\frac{dy}{dt} + a(t)y\right) = b(t)\exp\left(\int a(t)dt\right).$$

4. Using the product and chain rule of differentiation, write the left hand side of the equation in the following way:

$$\exp\left(\int a(t)dt\right)\left(\frac{dy}{dt} + a(t)y\right) = \frac{d}{dt}\left(\exp\left(\int a(t)dt\right)y\right).$$

So,

$$\frac{d}{dt}\left(\exp\left(\int a(t)dt\right)y\right) = b(t)\exp\left(\int a(t)dt\right).$$

5. Integrate both sides of the new equation:

$$\exp\left(\int a(t)dt\right)y = \int b(t)\exp\left(\int a(t)dt\right)dt + C.$$

6. Divide by the integrating factor to get the solution:

$$y = \exp\left(-\int a(t)dt\right)\left[\int b(t)\exp\left(\int a(t)dt\right)dt + C\right].$$