République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ DES SCIENCES EXACTES ET DES SCIENCES DE LA NATURE ET DE LA VIE

DÉPARTEMENT DE MATHÉMATIQUES

THESE DE DOCTORAT EN SCIENCES

Option: **Mathématiques**

Par

Abdelmadjid Abba

Titre

Sur le principe du maximum stochastique de presque optimalité et applications

---

Sous la direction de

Dr. Mokhtar Hafayed, MCA, Université de Biskra

Membres du Comité d'Examen

Dr. Naceur Khelil, *MCA, Université de Biskra,* ____ **Président**

Dr. Mokhtar Hafayed, *MCA, Université de Biskra* ____ **Rapporteur**

Prof. Dahmane Achour, *Prof. Université de M'sila* ____ **Examinateur**

Dr. Khalil Saadi, *MCA, Université de Msila* ____ **Examinateur**

Dr. Boulakhras Gherbal, *MCA, Université de Biskra* ____ **Examinateur**

Dr. Abdelmoumen Tiaiba, *MCA. Université de M'sila* ____ **Examinateur**

2016

I dedicate this work in memory of my mother Khadidja Bourdji

and my father Moussa.

To my family.

Acknowledgments

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

I would like to express my sincere thanks to Dr. Naceur Khelil because he agreed to spend his time for reading and evaluating my thesis.

I thank Professor Dahmen Achour (M'sila University) , Dr. Abdelmouman Tiaiba (M'sila University) and Dr. Khalil Saadi (M'sila University) because they agreed to spend their time for reading and evaluating my thesis and for their constructive corrections and valuable suggestions that improved the manuscript considerably.

I thank Dr. Boulakhras Gherbal because he played a fundamental role in my education and he was always ready to help me every time I asked.

I thank all my colleagues of the Mathematics Department, specially Dr. Badreddine Mansouri and Nacira Agram, and all my colleagues of Economics department of Biskra University.

Contents

SYMPOLS AND ACRONYMS

$a.e$	almost evrywhere
$a.s$	almost surly
$c\grave{a}dl\grave{a}g$	continu à droit, limit à gauche
$e.g$	for example
resp.	respectively
$\mathbb{R}$	real numbers
$\mathbb{R}_+$	nonnegative real numbers
$\sigma(A)$	$\sigma-$algebra generated by A.
$(\Omega, \mathcal{F})$	measurable space
$(\Omega, \mathcal{F}, \mathbb{P})$	probability space
$\mathbb{E}\left(\cdot\right)$	expectation
$\mathbb{E}\left(\cdot \mid \mathcal{G}\right)$	conditional expectation
$\mathcal{O}\left(\varepsilon\right)$	*error bound.*
$W(t)$	Brownian motion
$\mathbb{L}^2_{\mathcal{F}}([s,T], \mathbb{R}^n)$	the Hilbert space of $\mathcal{F}_t-$adapted processes $x(\cdot)$ such that $\mathbb{E}\int_s^T \left\|x(t)\right\|^2 dt < +\infty$
f_x	the gradient or Jacobian of a scalar function f with respect to the variable x.
f_{xx}	the Hessian of a scalar function f with respect to the variable x.
$\partial^{\circ}_x f$	the Clarke's generalized gradient of f with respect to x
$\mathcal{A}^*$	the transpose of any vector or matrix $\mathcal{A}$
$\langle x, y\rangle$	the scalar product of any two vectors x and y on $\mathbb{R}^d$
$\mathbf{I}_{\mathcal{B}}$	the indicator function of B
$\overline{co}\left(\mathcal{B}\right)$	the closure convex hull of $\mathcal{B}$
$Sgn(.)$	the sign function.
$L(\cdot) = \left(L(t)\right)_{t\in[0,T]}$	$\mathbb{R}$-valued Lévy process
$H(t) = \left(H^j(t)\right)_{j\geq 1}$	Teugels martingales

$\mathbb{P}\otimes dt$	the product measure of $\mathbb{P}$ with the Lebesgue measure dt
$l^2(\mathbb{R}^n)$	the space of $\mathbb{R}^n$-valued $(f_n)_{n\geq 1}$, $\left[\sum_{n=1}^{\infty}\|f_n\|_{\mathbb{R}^n}^2\right]^{\frac{1}{2}} < +\infty$.
$l_{\mathcal{F}}^2([0,T];\mathbb{R}^n)$	the Banach space of $\mathcal{F}_t$-adapted proc $\mathbb{E}\left(\int_0^T \|x(t)\|_{\mathbb{R}^n}^2 dt\right)^{\frac{1}{2}} < +\infty$.
$\mathbb{L}_{\mathcal{F}}^2([0,T];\mathbb{R}^n)$	the Banach space of $\mathcal{F}_t-$predictable proc $\mathbb{E}\left(\int_0^T \sum_{n=1}^{\infty}\|f_n\|_{\mathbb{R}^n}^2 dt\right)^{\frac{1}{2}} < +\infty$.
$\mathbb{S}_{\mathcal{F}}^2([0,T];\mathbb{R}^n)$	the Banach space of $\mathcal{F}_t-$adapted and càdlàg processes
	such that $\mathbb{E}(\sup\|x(t)\|^2)^{\frac{1}{2}} < +\infty$.
$\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{P},\mathbb{R}^n)$	the Banach space of $\mathbb{R}^n$-valued, square integrable r.v on $(\Omega,\mathcal{F},\mathbb{P})$.
$\mathbb{M}^{n\times m}(\mathbb{R})$	the space of $n\times m$ real matrices.
$\mathcal{F}_t^W$	the $\sigma-$algebra generated by $W(s)$ and $\sigma\{W(s): 0\leq s\leq t\}$.
$\mathcal{G}_0$	the totality of $\mathbb{P}-$null sets.
$\mathcal{F}_1\vee\mathcal{F}_2$	the σ-field generated by $\mathcal{F}_1\cup\mathcal{F}_2$.
$\Delta\xi(t)=\xi(t)-\xi(t_-)$.	the jumps of a singular control $\xi(\cdot)$ at any jumping time t.
ODE	ordinary differential equations
$SDEs$	stochastic differential equations
$BSDEs$	Backward stochastic differential equations
$FBSDEs$	Forward-Backward stochastic differential equations
$\mathcal{U}_{\mathcal{G}}^1\times\mathcal{U}_{\mathcal{G}}^2([0,T])$	the set of admissible controls.

INTRODUCTION

In this thesis, we study stochastic control problems, where the system is governed by stochastic differential equations of mean-field type. The main part of the thesis is divided in fort chapters.

In chapter 1., we collect some basis results of probability theory and stochastic analysis in particular, we recall some basic proprieties of conditional expectation ,class of controls , martingales... .

In chapter 2., we establish the necessary and sufficient conditions of near-optimality for systems governed by stochastic differential equations with of poison jumps mean-field type. The results have been proved by applying Ekeland's Lemma, spike variation method and some estimates of the state and adjoint processes. Under certain concavity conditions, we prove that the near-maximum condition on the Hamiltonian function in integral form is a sufficient condition for near-optimality. An example is presented to illustrate the theoretical results. These results generalize the maximum principle proved in Zhou (*SIAM. Control Optim. (36), 929-947, 1998* [45]) and Tang and Li (*SIAM. Control Optim. (32), 1147-1475, (1994)* [40]) to a class of stochastic control problems involving jump diffusion processes of mean-field type. We note that since the work by Zhou [45], the concept of near-optimal stochastic controls was introduced for a class of stochastic control problems involving classical stochastic differential equations (SDEs). A *near-optimal control of order* ε^{λ} *is an admissible control defined by*

For a given $\varepsilon > 0$ *the admissible control* $u^{\varepsilon}(\cdot)$ *is near-optimal with respect* (s,ζ) *if*

$$\left| J^{s,\zeta}\left(u^{\varepsilon}(\cdot)\right) - V\left(s,\zeta\right)\right| \leq O\left(\varepsilon\right),$$

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

- If $O(\varepsilon) = C\varepsilon^{\lambda}$ for some $\lambda > 0$ independent of the constant C then $u^{\varepsilon}(\cdot)$ is called near-optimal control of order ε^{λ}.

- If $O(\varepsilon) = C\varepsilon$, the admissible control $u^{\varepsilon}(\cdot)$ called $\varepsilon-$optimal.

In this chapter, we obtain a Zhou-type necessary conditions of near-optimality, where the system is described by nonlinear controlled jump diffusion processes of mean-field type of the form

$$\begin{cases} dx^u(t) = f(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t))dt + \sigma(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t))dW(t), \\ +\int_{\Theta} g\left(t, x^u(t^-), u(t), \theta\right) N(d\theta, dt),\ \ x^u(s) = \zeta, \end{cases}$$

and the cost functional has the form

$$J^{s,\zeta}\left(u(\cdot)\right) = \mathbb{E}\left[h(x^u(T), \mathbb{E}\left(x^u(T)\right)) + \int_s^T \ell(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t))dt\right].$$

The control domain is not need to be convex. (a general action space). The proof of our results follows the general ideas as in Zhou [45], Buckdahn et al., [5], and Tang et al., [40]. Finally, for the reader's convenience, we give some analysis results used in this chapter in the Appendix.

In chapter 3., In this chapter, we study partial information stochastic optimal control problem of mean-field type, where the system is governed by controlled stochastic differential equation driven by Teugels martingales associated with some Lévy process and an independent Brownian motion. We establish necessary and sufficient conditions of optimal control for these mean-field models in the form of maximum principle. The control domain is assumed to be convex. As an application, partial information linear quadratic control problem of mean-field type is discussed, where the optimal control is given in feedback form.

The system under consideration is governed by stochastic differential equations driven by Teugels martingales associated with some Lévy process and an independent Brownian motion of the form:

$$\begin{cases} dx^u(t) = f\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) dt + \sum_{j=1}^{d} \sigma^j \left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) dW^j(t), \\ + \sum_{j=1}^{\infty} g^j \left(t, x^u(t_-), \mathbb{E}(x^u(t_-)), u(t)\right) dH^j(t), \\ x^u(0) = x_0, \end{cases}$$

and the expected cost on the time interval $[0, T]$ has the form

$$J\left(u(\cdot)\right) := \mathbb{E}\Big\{ \int_0^T \ell(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dt + h\left(x^u(T), \mathbb{E}\left(x^u(T)\right)\right) \Big\},$$

where $W(\cdot)$ is a standard $d-$dimensional Brownian motion and $H(t) = \left(H^j(t)\right)_{j\geq 1}$ are pairwise strongly orthonormal Teugels martingales, associated with some Lévy process, having moments of all orders. The control $u(\cdot) = (u(t))_{t\geq 0}$ is required to be valued in some subset of $\mathbb{R}^k$ and adapted to a subfiltration $(\mathcal{G}_t)_{t\geq 0}$ of $(\mathcal{F}_t)_{t\geq 0}$. The maps f, σ, g, ℓ and h are an appropriate functions. In this chapter, we derive a partial information maximum principle for stochastic differential equations, with Lévy processes. Necessary and sufficient conditions of optimality have been established with an application to finance. Some discussions with remarks are given in the last of this chapter.

In chapter 4., we prove a necessary and sufficient conditions of optimality singular control for systems driven by stochastic differential equations with Teugels martingales associated with Lévy processes with applications to linear quadratic control problem of the form:

$$\begin{cases} dx^{u,\xi}(t) = f\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)\right) dt + \sum_{j=1}^{d} \sigma^j \left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)\right) dW^j(t), \\ + \sum_{j=1}^{\infty} g^j \left(t, x^{u,\xi}(t_-), \mathbb{E}(x^{u,\xi}(t_-)), u(t)\right) dH^j(t) + \mathcal{C}(t) d\xi(t), \\ x^{u,\xi}(0) = x_0, \end{cases}$$

and the cost functional has the form

$$\begin{aligned} J\left(u(\cdot), \xi(\cdot)\right) = \mathbb{E}\Big\{ & \int_0^T \ell(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)) dt \\ & + h(x^{u,\xi}(T), \mathbb{E}(x^{u,\xi}(T))) + \int_{[0,T]} \mathcal{M}(t) d\xi(t) \Big\}, \end{aligned}$$

where $W(\cdot)$ is a standard $d-$dimensional Brownian motion and $H(t) = \left(H^j(t)\right)_{j\geq 1}$ are pairwise strongly orthonormal Teugels martingales, associated with some Lévy processes, having moments of all orders, and $\xi(\cdot)$ is the singular part of the control, which is called intervention control. The continuous control $u(\cdot) = (u(t))_{t\geq 0}$ is required to be valued in some subset of $\mathbb{R}^k$ and adapted to a subfiltration $(\mathcal{G}_t)_{t\geq 0}$. In some finance models, the mean-field term $\mathbb{E}(x^{u,\xi}(t))$ represents an approximation to the weighted average $\frac{1}{n}\sum_{i=1}^{n} x_n^{u,\xi,i}(t)$ for large n, $\xi(t)$ representing the harvesting effort, while $\mathcal{C}(t)$ is a given harvesting efficiency coefficient.

As an illustration, linear quadratic control problem of mean-field type involving continuous-singular control is discussed, where the optimal control is given in feedback form. Note that in our mean-field control problem, there are two types of jumps for the state processes, the inaccessible ones which come from the Lévy martingale part and the predictable ones which come from the singular control part. Finally, some discussions with concluding remarks are given in the last of this chapter.

Chapter-I

Stochastic Control Problem

Chapter 1

Stochastic Control Problem

1.1 Stochastic Processes

Definition. (Filtration) A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family $(\mathcal{F}_t)_{t\in[0,T]}$ of $\sigma-$fields of $\mathcal{F} : \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0 \leq s \leq t$ T. $\mathcal{F}_t$ is interpreted as the information known at time t and increases as time elapses.

In this section we recall some results on stochastic processes.

Definition 1.1.1. Let I be a nonempty index set and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. A family $(X_t, t \in I)$ of random variables from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\mathbb{R}^n$ is called a stochastic process.For any $w \in \Omega$, the map $t \longmapsto X(w,t)$.is called a sample path.

In what follows, we set $I = [0,T]$, or $I = [0,\infty)$. We shall interchangeably use $(X_t, t \in I)$, X, X_t to denote a stochastic process.

For any given stochastic process $(X_t, t \in I)$, we can defined the following

$$F_{t_1}(x) \triangleq \mathbb{P}(X_{t_1} \leq x_1),$$

$$F_{t_1,,t_2}(x_1,x_2) \triangleq \mathbb{P}(X_{t_1} \leq x_1, X_{t_2} \leq x_2)$$

$$F_{t_1,,t_2...,t_n}(x_1,x_{2,...}x_n) \triangleq \mathbb{P}(X_{t_1} \leq x_1, X_{t_2} \leq x_2, ...X_{t_n} \leq x_n),$$

where $t_i \in I$, $x_i \in \mathbb{R}^n$, and $X_i \leq x_i$ stands for component twice inequalities, the functions define F are called the finite-dimensional distributions of the process X_t.

Definition 1.1.2.(stochastically equivalent) Two processes X_t and Y_t are said to be *stochastically equivalent* if

$$X_t = Y_t,\ \mathbb{P} - a.s,\ \forall t \in [0, T]\,.$$

In this case, one is called a modification of the other.

If X_t and Y_t are stochastically equivalent ,then for any $t \in [0, T]$ there exists a $\mathbb{P}$ -null set $N_t \in \mathcal{F}$ such that

$$X_t = Y_t, \quad \forall w \in \Omega \mid N_t.$$

Example Let $\Omega = [0, 1]\,, T \geq 1$, $\mathbb{P}$ the Lebesgue measure, $X\,(w, t) = 0$, and

$$Y_t\,(w) = \begin{cases} 0, & w \neq t, \\ 1, & w = t. \end{cases}$$

Then X_t and Y_t are said to be *stochastically equivalent*. But each sample path $X\,(., t)$ is continuous , and none of the sample paths $Y_t\,(., w)$ is continuous. In the present case, we actually have

$$\bigcup_{t \in [0,t]} N_t = [0, 1] = \Omega.$$

Definition 1.1.3. The process at $s \in [0, T]$ if for any $\varepsilon > 0$

$$\lim_{t \longrightarrow s} \mathbb{P}\,(w \in \Omega, |X_t\,(w) - X_s\,(w)| > \varepsilon) = 0.$$

Morover, X_t is said to be continuous if there exists a $\mathbb{P}$ -null set $N \in \mathcal{F}$ such that for any $w \in \Omega \mid N$,the sample path $X\,(\cdot, t)$ is continuous

Then X_t and Y_t are said to be stochastically equivalent. But each sample path $X\,(\cdot, t)$ is continuous ,and none of the sample paths $Y_t\,(\cdot, w)$ is continuous .

In the present case, we actually have

$$\bigcup_{t\in[0,t]} N_t = [0,1] = \Omega.$$

Definition 1.1.4. The process at $s \in [0,T]$ if for any $\varepsilon > 0$

$$\lim_{t\rightarrow s} \mathbb{P}\left(w \in \Omega, \left|X_t\left(w\right) - X_s\left(w\right)\right| > \varepsilon\right) = 0.$$

Moreover,X_t is said to be continuous if there exists a $\mathbb{P}$ -null set $N \in \mathcal{F}$ such that for any $w \in \Omega \mid N$, the sample path $X\left(\cdot, t\right)$ is continuous.

1.2 Lévy process

To model the sudden crashes in finance, it is natural to allow jumps in the model because this makes it more realistic. This models can be represented by Lévy processes which are used throughout this work. This term (Lévy process) honors the work of the French mathematician *Paul Lévy.*

Definition 1.2.1. A process $X = (X(t))_{t\geq 0} \subset \mathbb{R}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a Lévy process if it possesses the following properties:

(1) The paths of X are $\mathbb{P}$-almost surely right continuous with left limits.

(2) $\mathbb{P}(X(0) = 0) = 1$.

(3) Stationary increments, i.e., for $0 \leqslant s \leqslant t$, $X(t) - X(s)$ has the same distribution as $X(t-s)$

(4) Independent increments, i.e., for $0 \leqslant s \leqslant t$, $X(t) - X(s)$ is independent of $X(u)$, $u \leqslant s$.

Example. The known examples are the standard Brownian motion and the Poisson process.

Definition 1.2.2. A stochastic process $W = (W(t))_{t\geq 0}$ on $\mathbb{R}^n$ is a Brownian motion if it is a Lévy process and if

(1) For all $t > 0$, has a Gaussian distribution with mean 0 and covariance matrix tI_d.

(2) There is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that, for every $w \in \Omega_0$, $W(t, w)$ is continuous in t.

Definition 1.2.3. A stochastic process $N = (N(t))_{t \geq 0}$ on $\mathbb{R}$ such that

$$\mathbb{P}\left[N(t) = n\right] = \frac{(\lambda t)^n}{n!} e^{-\lambda t}; \quad n = 0, 1,$$

is a Poisson process with parameter $\lambda > 0$ if it is a Lévy process and for $t > 0$, $N(t)$ has a Poissson distribution with mean λt.

Remark 1.2.4. (1) Note that the properties of stationarity and independent increments imply that a Lévy process is a Markov process.

(2) Thanks to almost sure right continuity of paths, one may show in addition that Lévy processes are also strong Markov processes.

Any random variable can be characterized by its characteristic function. In the case of a Lévy process X, this characterization for all time t gives the *Lévy-Khintchine formula* and it is also called Lévy-Khintchine representation.

1.3 Stochastic integral with respect to Lévy process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space with the σ-algebra $(\mathcal{F}_t)_{t \geq 0}$ generated by the underline driven processes; Brownian motion $W(t)$ and an independent compensated Poisson random measure $\tilde{N}$, such that

$$\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt.$$

For any t, let $\tilde{N}(ds, dz)$, $z \in \mathbb{R}$, $s \leqslant t$, augmented for all the sets of $\mathbb{P}$-zero probability.

For any $\mathcal{F}_t-$adapted stochastic process $\theta = \theta(t, z)$, $t \geqslant 0$, such that

$$E\left[\int_0^T \int_{\mathbb{R}} \theta^2(t, z)\nu(dz)dt\right] < \infty, \text{ for some } T > 0,$$

we can see that the process

$$M_n(t) = \int_0^t \int_{|z|\geq \frac{1}{n}} \theta(s,z)\tilde{N}(ds,dz), \quad 0 \leqslant t \leqslant T,$$

is a martingale in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and its limit

$$M(t) = \lim_{n\to\infty} M_n(t) := \int_0^T \int_{|z|\geq \frac{1}{n}} \theta(s,z)\tilde{N}(ds,dz), \quad 0 \leqslant t \leqslant T,$$

in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is also a martingale. Moreover, we have the Itô isometry

$$E\left[\left(\int_0^T \int_{\mathbb{R}_0} \theta(s,z)\tilde{N}\left(ds,dz\right)\right)^2\right] = E\left[\left(\int_0^T \int_{\mathbb{U}} \theta^2(t,z)\nu(dz)dt\right)\right].$$

Such processes can be expressed as the sum of two independent parts, a continuous part and a part expressible as a compensated sum of independent jumps. That is the *Itô-Lévy* decomposition.

Theorem 1.3.1 (Itô-Lévy decomposition) The Itô-Lévy decomposition for a Lévy process X is given by

$$X(t) = \alpha t + \beta W(t) + \int_{|z|<1} z\tilde{N}(dt,dz) + \int_{|z|\geqslant 1} zN(dt,dz),$$

where $\alpha, \beta \in \mathbb{R}$, $\tilde{N}(dt,dz)$ is the compensated Poisson random measure of $X(.)$ and $B(t)$ is an independent Brownian motion with the jump measure $N(dt,dz)$.

We assume that

$$E\left[X^2(t)\right] < \infty, \quad t \geqslant 0,$$

then

$$\int_{|z|\geq 1} |z|^2 \nu(dz) < \infty.$$

We can represent as

$$X(t) = \alpha t + \beta B(t) + \int_{\mathbb{R}} z \tilde{N}(dt, dz),$$

where $X(t) = \alpha + \int_{|z|\geq 1} z\nu(dz)$. If $\beta = 0$, then a Lévy process is called a pure jump Lévy process.

Let us consider that the process $X(t)$ admits the stochastic integral representation as follows

$$X(t) = x + \int_0^t \alpha(s)ds + \int_0^t \beta(s)dW(s) + \int_0^t \int_{\mathbb{R}} \theta(s,z)\tilde{N}(ds, dz),$$

where $\alpha(t)$, $\beta(t)$, and $\theta(t,\cdot)$ are predictable processes such that, for all $t > 0, z \in \mathbb{R}$,

$$\int_0^t \left[|b(s)| + \sigma^2(s) + \int_{\mathbb{R}} \theta^2(s,z)\nu(dz) \right] ds < \infty \quad \mathbb{P} - a.s.$$

Under this assumption, the stochastic integrals are well-defined and local martingales. If we strengthened the condition to

$$E\left[\int_0^t \left[|b(s)| + \sigma^2(s) + \int_{\mathbb{R}} \theta^2(s,z)\nu(dz) \right] ds \right] < \infty,$$

for all $t > 0$, then the corresponding stochastic integrals are martingales.

We call such a process an Itô–Lévy process. In analogy with the Brownian motion case, we use the short-hand differential notation

$$\begin{cases} dX(t) = b(t)dt + \sigma(t)dB(t) + \int_{\mathbb{R}} \theta(t,z)\tilde{N}(dt, dz), \\ X(0) = x \in \mathbb{R}. \end{cases}$$

The Itô formula and related results

We now come to the important Itô formula for Itô-Lévy processes. Let $X(t)$ be a process given by 1.3.1

$$X(t) = \alpha(t) + \beta(t) B(t) + \int_{\mathbb{R}} \gamma(t,z) \tilde{N}(dt,dz), \tag{1.1}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is a C^2 function is the process $Y(t) := f(t, X(t))$ again an Itô-Lévy process and if so, how do we represent it in the form (1.1).

Let $X^c(t)$ be the continuous part of $X(t)$,i.e $X^c(t)$ is obtained by removing the jumps from $X(t)$.

$$\begin{aligned} dY(t) &= \frac{\partial f}{\partial t}(t; X(t))\, dt + \frac{\partial f}{\partial x}(t; X(t))\, dX^c(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t; X(t))\, \beta^2(t) \\ &+ \int_R \left\{ f\left(t, X(t^-)\right) + \gamma(t,z) - f\left(t, X(t^-)\right) \right\} \tilde{N}(dt,dz). \end{aligned}$$

It can be proved that our guess is correct. Since

$$dX^c(t) = \left(\alpha(t)dt - \int_{|F|<r} \gamma(t,z) v(dz) \right) + \beta(t) dB(t),$$

this given the following result;

Theorem 1.3.2 Let $X(t) \in \mathbb{R}$ is an Itô-Lévy process of the form

$$dX(t) = \alpha(t) + \beta(t) B(t) + \int_{\mathbb{R}} \gamma(t,z) \tilde{N}(dt,dz), \tag{1.2}$$

where

$$\tilde{N}(dt,dz) = \begin{cases} N(dt,dz) - v(dz)\, dt, & \text{if } |F| < r. \\ N(dt,dz) & \text{if } |F| \geq r, \end{cases}$$

for some $r \in [0, \infty]$. Let $f \in C^2(R^2)$ and define $Y(t) = f(t, X(t))$. Then $Y(t)$ is again an

Itô -Lévy process

$$dY(t) = \frac{\partial f}{\partial t}(t; X(t))\,dt + \frac{\partial f}{\partial x}(t; X(t))\,(\alpha(t)dt + \beta(t)dB(t)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t; X(t))\,\beta^2(t)$$
$$+ \int_{|F|<r} \left\{ f\left(t, X(t^-)\right) + \gamma(t,z) - f\left(t, X(t^-)\right) - \frac{\partial f}{\partial x}(t; X(t))\,\gamma(t,z) \right\} v(dz)$$
$$\int_{\mathbb{R}} \left\{ f\left(t, X(t^-)\right) + \gamma(t,z) - f\left(t, X(t^-)\right) \right\} \tilde{N}(dt, dz),$$

Remark 1.3.3. if $r = 0$ then $\tilde{N} = N$ every where. If $r = \infty$ then $\tilde{N} = N$ every where.

Theorem 1.3.3. (The multi-dimensional Itô formula).Let $X(t) \in \mathbb{R}^n$ be an Itô-Lévy process of the form

$$dX(t) = \alpha(t; w)\,dt + \sigma(t; X(t,w))\,dB(t) + \int_{\mathbb{R}^n} \gamma(t,z,w)\,\tilde{N}(dt,dz),$$

where $\alpha : [0,T] \times \Omega \to \mathbb{R}^n, \sigma : [0,T] \times \Omega \to \mathbb{R}^{n+m}$ and $\gamma : [0,T] \times \mathbb{R}^n \times \Omega\times \to \mathbb{R}^{n\times l}$ are adapted processes such that the integrals exist. Here $B(t)$ is an multidimensional Brownian motion and

$$\tilde{N}(dt,dz)^T = \left(\tilde{N}_1(dt,dz), ..., \tilde{N}_l(dt,dz)\right)$$
$$= \left(\tilde{N}_1(dt,dz) - I_{|z_1|<r} v_1(dz_1)\,dt, ..., \tilde{N}_l(dt,dz) - I_{|z_l|<r_l} v_l(dz_l)\,dt\right),$$

where $(N_j(\cdot,\cdot))$ are independent Poisson random measures with Lévy processes $(\eta_1, ..., \eta_l)$. Note that each column $\gamma^{(k)}$ of the $n \times l$ matrix $\gamma = (\gamma_{ij})$ depends on z only through the k^{th}coordinate z_k, i.e.,

$$\gamma^{(k)}(t,z,w) = \gamma^{(k)}((t, z_k, w));\ z = (z_1, ..., z_l) \in \mathbb{R}^l.$$

Thus the integral on the right of (1.2) is just a short hand matrix notation. When written out in detail component number i of $X(t)$ in (1.2), $X_i(t)$, gets the form

$$dX_i(t) = \alpha_i(t;w)\,dt + \sum_{j=1}^{m} \sigma_{ij}(t,w)\,dB_j(t) + \sum_{j=1}^{l} \int_{\mathbb{R}^n} \gamma_{ij}(t,z_j,w)\,\tilde{N}_j(dt,dz_j),$$

$$1 \leq i \leq n.$$

Theorem 1.3.4.(*The Itô-Lévy isometry*) Let $X(t) \in \mathbb{R}^n$ is be as in (1.2) but with $X(0)$ and $\alpha = 0$. Then

$$\begin{aligned} E\left[X^2(t)\right] &= E\left[\int_0^T \left\{\sum_{j=1}^{m} \sigma_{ij}^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{l} \int_{\mathbb{R}^n} \gamma_{ij}^2(t,z_j)\,v_j(dz_j)\right\} dt\right], \\ &= \sum_{i=1}^{n} E\left[\int_0^T \left\{\sum_{j=1}^{m} \sigma_{ij}^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{l} \int_{\mathbb{R}^n} \gamma_{ij}^2(t,z_j)\,v_j(dz_j)\right\} dt\right]. \end{aligned}$$

1.4 Some classes of stochastic control problems

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

(1) **Admissible control** An admissible control is a measurable and $\mathbb{F}$-adapted process $u(t)$ with values in a borelian $A \subset \mathbb{R}^n$. We denote by $\mathcal{U}$ the set of all admissible controls, such that

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

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

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

(3) **Near optimal control** Let $\varepsilon > 0$, a control is a near optimal control (or ε-optimal) if for all control $u \in U$ we have that

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

(4) **Feedback control** Let $u(\cdot)$ be an $\mathcal{F}$-adapted control and we denote by $\mathcal{F}_t^X$ the natural filtration generated by the process X. We say that $u(\cdot)$ is a feedback control if and only if $u(\cdot)$ depends on X.

(5) **Optimal stopping** In the formulation of such models, an admissible control stopping time is a pair $(u(\cdot), \tau)$ defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ along with an n-dimensional Brownian motion $W(\cdot)$, where $u(\cdot)$ is the contol satisfying the usual conditions and τ is an $(\mathcal{F}_t)_{t\geq 0}$-stopping time the optimal control stopping problem is to minize

$$J(u(\cdot), \tau) = E\left\{\int_0^\tau f(t, x(t), u(t))\, dt + h(x(\tau))\right\}.$$

$$\tau = \inf\{t \geq 0 : x(t) \in O\}, O \subseteq \mathbb{R}^n.$$

(6) **Singular control** Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, \mathbb{P})$ be a complete filtred probability space. An admissible control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathcal{F}_t-$adapted processes, such that $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the left with right limits and $\xi(0_-) = 0$. Moreover,

$$\mathbb{E}(\sup_{0\leq t\leq T} |u(t)|^2 + |\xi(T)|^2) < \infty.$$

Note that the jumps of a singular control $\xi(\cdot)$ at any jumping time t is denoted by

$$\Delta\xi(t) \triangleq \xi(t) - \xi(t_-).$$

Let us define the continuous part of the singular control by

$$\xi^{(c)}(t) \triangleq \xi(t) - \sum_{0\leq\tau_j\leq t} \Delta\xi(\tau_j),$$

i.e., the process obtained by removing the jumps of $\xi(t)$.

We denote $\mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2([0, T])$, the set of all admissible controls. Since $d\xi(t)$ may be singular with respect to Lebesgue measure dt, we call $\xi(\cdot)$ the singular part of the control and the process $u(\cdot)$ its absolutely continuous part.

(7) **Relaxed controls** Let $U \subset \mathbb{R}^d$. A relaxed control with values in U is a measure q over $[0,T] \times U$ such that the projection on $[0,T]$ is the Lebesgue measure. If there exists $v : [0,T] \to U$ such that

$$q\left(dt, dv\right) = \delta_{v(t)}\left(dv\right) dt,$$

q is identified with v_t and said to be a control process.

Noting that if q be a relaxed control with values in $\mathcal{U}$. Then, for all $t \in [0,T]$ there exists a probability measure q_t over $\mathcal{U}$ such that

$$q\left(dt, dv\right) = dtq_t\left(dv\right).$$

The proof is application of Fubini theorem.

Chapter-II

On stochastic Near-optimal Control Problems for Mean-field Jump Diffusion Processes

Chapter 2

On Stochastic Near-optimal Control Problems for Mean-field Jump Diffusion Processes

Abstract. In a recent work by Zhou [45], the concept of near-optimal stochastic controls was introduced for a class of stochastic control problems involving classical stochastic differential equations (SDEs in short). Necessary and sufficient conditions for near-optimal controls were derived. This work extends the results obtained by Zhou [45] to a class of stochastic control problems involving jump diffusion processes of mean-field type. We derive necessary as well as sufficient conditions of near-optimality for our model, using Ekeland's variational principle, spike variation method and some estimates of the state and adjoint processes. Under certain concavity conditions, we prove that the near-maximum condition on the Hamiltonian function in integral form is a sufficient condition for near-optimality. An example is presented to illustrate the theoretical results.

2.1 Introduction

In this work, we consider a stochastic control problem for systems driven by a nonlinear controlled jump diffusion processes of mean-field type, which is also called McKean-Vlasov equations, where the coefficients depend on the state of the solution process as well as of its expected value. More precisely, the system under consideration evolves according to the jump diffusion process

$$\begin{cases} dx^u(t) = f(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dt + \sigma(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dW(t) \\ \qquad\qquad + \int_\Theta g(t, x^u(t^-), u(t), \theta) N(d\theta, dt), \\ x^u(s) = \zeta, \end{cases} \tag{2.1}$$

for some functions f, σ, g. This mean-field jump diffusion processes are obtained as the mean-square limit, when $n \to +\infty$ of a system of interacting particles of the form

$$dx_n^{j,u}(t) = f(t, x_n^{j,u}(t), \frac{1}{n}\sum_{i=1}^{n} x_n^{i,u}(t), u(t))dt + \sigma(t, x_n^{j,u}(t), \frac{1}{n}\sum_{i=1}^{n} x_n^{i,u}(t), u(t))dW^j(t)$$
$$+ \int_\Theta g(t, x_n^{j,u}(t^-), u(t), \theta) N(d\theta, dt).$$

where $(W^j(\cdot) : j \geq 1)$ is a collection of independent Brownian motions. The expected cost to be near-minimized over the class of admissible controls is also of mean-field type, which has the form

$$J^{s,\zeta}(u(\cdot)) = \mathbb{E}\left[h(x^u(T), \mathbb{E}(x^u(T))) + \int_s^T \ell(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dt\right]. \tag{2.2}$$

The value function is defined as

$$V(s, \zeta) = \inf_{u(\cdot)\in\mathcal{U}} J^{s,\zeta}(u(\cdot)),$$

where the initial time s and the initial state ζ of the system are fixed.

The optimal control theory has been developed since early 1960s, when Pontryagin et al., [35] published their work on the maximum principle and Bellman [7] put forward the dynamic programming method. The pioneering works on the stochastic maximum principle was written by Kushner ([29, 30]). Since then there have been a lot of works on this subject, among them, in particular, see [2, 3, 80, 32, 27, 36, 109] and the references therein.

It is well-known that near-optimization is as sensible and important as optimization for both theory and applications. The Modern near-optimal control theory has been well developed when Zhou published their works on necessary and sufficient conditions for any near-optimal controls for both deterministic and stochastic controls see ([42, 43, 45]). The near-optimal deterministic control problems have been investigated in ([42, 43, 44, 14, 12, 25, 34]. The necessary conditions for some near-optimal controls have been established by Ekeland [12], The necessary and sufficient conditions for any near-optimal deterministic controls are investigated in Zhou [42]. Dynamic programming and viscosity solutions approach for near-optimal deterministiccontrols have been studied in [43]. In Pan et al., [34] the authors extended the results obtained by Zhou [42] to a class of optimal control problems involving Volterra integral equations.

It is well documented (e.g. Zhou (1998) [45]) that the near-optimal stochastic controls, as the alternative to the exact optimal controls, are of great importance for both the theoretical analysis and practical application purposes due to its nice structure and broad-range availability, feasibility as well as flexibility. In this recent work, Zhou [45] established the second-order necessary as well as sufficient conditions for near-optimal stochastic controls for classical controlled diffusion, where the coefficients were assumed to be twice continuously differentiable and the control domain not necessarily convex. In Hafayed et al., [17], the authors extended Zhou's maximum principle of near-optimality to singular stochastic controls. The near-optimal control problems for systems described by SDEs with jumps have been studied in Hafayed et al., [16]. The second-order maximum principle of near-optimality for jump diffusions was obtained in [11]. The near-optimal stochastic control problem for Forward backward SDEs has been investigated in Huang et al., [21] and Bahlali et al. [20]. The near-optimal control problem for recursive stochastic problem has been studied in Hui el al., [19].

The stochastic optimal control problems for jump processes has been investigated by many authors, see for instance, ([9, 13, 33, 37, 62, 39, 40]. The general case, where the control domain is not necessarily convex and the diffusion coefficient depends explicitly on the control variable, was derived via spike variation method by Tang et al., [40], extending the Peng stochastic maximum principle of optimality [36]. These conditions are described in terms of two adjoint processes, which are linear classical backward SDEs. A good account and an extensive list of references on stochastic optimal control for jump processes can be founded in Øksendal et al., [33], and Shi [38].

The SDE of mean-field type was suggested by Kac [15] in 1956 as a stochastic model for the Vlasov-kinetic equation of plasma and the study of which was initiated by McKean [24] in 1966. Since then, many authors made contributions on SDEs of mean-field type and applications, see for instance, ([1, 8, 41, 15, 6, 5, 60, 26]). Mean- field stochastic maximum principle of optimality was considered by many authors, see for instance ([6, 5, 18, 60, 26, 64]). In Buckdahn et al., [5] the authors obtained mean-field backward stochastic differential equations. The general maximum principle of optimality for mean-field control problem has been investigated in Buckdahn et al., [5], where the authors obtained a stochastic maximum principle differs from the classical one in the sense that the first-order adjoint equation turns out to be a linear mean-field backward SDE, while the second-order adjoint equation remains the same as in Peng's stochastic maximum principle [36]. The stochastic maximum principle of optimality for mean-field jump diffusion processes has been studied by Hafayed et al, [18]. The local maximum principle of optimality for mean-field stochastic control problem has been derived by Li [60]. The linear-quadratic optimal control problem for mean-field SDEs has been studied by Yong [64]. In Mayer-Brandis et al., [26] a maximum principle of optimality for SDEs of mean-field type was proved by using Malliavin calculus. An extensive list of references on mean-field control problems can be founded in Yong [64].

Our main goal in this work is to establish necessary as well as sufficient conditions of near-optimality for mean-field jump diffusion processes, in which the coefficients depend on the state of the solution process as well as of its expected value. Moreover, the cost functional is also of mean-field type. The proof of our main result is based on some stability results with respect to the control variable of the state process and adjoint processes, along with Ekeland's variational principle [12] and spike variation method. This near-optimality necessary and sufficient conditions differs from the classical one in the sense that here the first-order adjoint equation turns out to be a linear mean-field backward stochastic differential equation, while the second-order adjoint equation remains the same as in stochastic maximum principle for jump diffusions developed in Tang et al., [40]. The control domain under consideration is not necessarily convex. It is shown that stochastic optimal control may fail to exist even in simple cases, while near-optimal controls always exist. This justifies the use of near-optimal stochastic controls, which exist under minimal conditions and are sufficient in most practical cases. Moreover, since there are many near-optimal controls, it is possible to select among them appropriate ones that are easier for analysis and implementation. Finally, for the reader's convenience we give some analysis results used in this work in the Appendix.

The rest of the work is organized as follows. Section 2 begins with a general formulation of a Mean-field control problem with jump processes and give the notations and assumptions used throughout the work. In Sections 3 and 4, we derive necessary and sufficient conditions for near-optimality respectively, which are our main results. An example of this kind of control problem is also given in the last section.

2.2 Problem formulation and preliminaries

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a fixed filtered probability space equipped with a $\mathbb{P}-$completed right continuous filtration on which a $d-$dimensional Brownian motion $W = (W(t))_{t\in[0,T]}$ is defined. Let η be a homogeneous $(\mathcal{F}_t)$-Poisson point process independent of W. We denote by $\widetilde{N}(d\theta, dt)$ the random counting measure induced by η, defined on $\Theta \times \mathbb{R}_+$, where Θ is a fixed nonempty subset of $\mathbb{R}^k$ with its Borel σ-field $\mathcal{B}(\Theta)$. Further, let $\mu(d\theta)$ be the local characteristic measure of η, i.e. $\mu(d\theta)$ is a σ-finite measure on $(\Theta, \mathcal{B}(\Theta))$ with $\mu(\Theta) < +\infty$. We then define

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

where N 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

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

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

Basic Notations. We list some notations that will be used throughout this work.

1. Any element $x \in \mathbb{R}^d$ will be identified to a column vector with i^{th} component, and the norm $|x| = \sum_{i=1}^d |x_i|$.

2. The scalar product of any two vectors x and y on $\mathbb{R}^d$ is denoted by $\langle x, y\rangle$.

3. We denote $\mathcal{A}^*$ the transpose of any vector or matrix $\mathcal{A}$.

4. For a set $\mathcal{B}$, we denote by $\mathbf{I}_{\mathcal{B}}$ the indicator function of B and $\overline{co}(\mathcal{B})$ the closure convex hull of $\mathcal{B}$ and $Sgn(.)$ the sign function.

5. For a function Φ, we denote by Φ_x (*resp.* Φ_{xx}) the gradient or Jacobian (*resp.* the Hessian) of a scalar function Φ with respect to the variable x. We denote $\partial_x^\circ \Phi$ the Clarke's generalized gradient of Φ with respect to x.

6. We denote by $\mathbb{L}^2_{\mathcal{F}}([s,T], \mathbb{R}^n)$ the Hilbert space of $\mathcal{F}_t-$adapted processes $x(\cdot)$ such that $\mathbb{E}\int_s^T |x(t)|^2 dt < +\infty$.

7. For convenience, we will use $\Phi_x(t) = \dfrac{\partial \Phi}{\partial x}(t, x(t), \mathbb{E}(x(t)), u(t))$, and $\Phi_{xx}(t) = \frac{\partial^2 \Phi}{\partial x^2}(t, x(t), \mathbb{E}(x(t)), u(t))$.

Basic Assumptions. Throughout this work we assume the following.

Assumption (H1). The functions $f : [s,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \to \mathbb{R}^n$, $\sigma : [s,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \to \mathcal{M}_{n\times d}(\mathbb{R})$ and $\ell : [s,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \to \mathbb{R}$ are measurable in (t,x,y,u) and twice continuously differentiable in (x,y), $g : [s,T] \times \mathbb{R}^n \times \mathbb{A} \times \Theta \to \mathbb{R}^{n\times m}$ is twice continuously differentiable in x, and there exists a constant $C>0$ such that, for $\varphi = f, \sigma, \ell$:

$$\begin{aligned} &|\varphi(t,x,y,u) - \varphi(t,x',y',u)| + |\varphi_x(t,x,y,u) - \varphi_x(t,x',y',u)| \\ &\leq C\left[|x-x'| + |y-y'|\right]. \end{aligned} \tag{2.3}$$

$$|\varphi(t,x,y,u)| \leq C\left(1 + |x| + |y|\right). \tag{2.4}$$

$$\begin{aligned} &\sup\nolimits_{\theta\in\Theta} |g(t,x,u,\theta) - g(t,x',u,\theta)| + \sup\nolimits_{\theta\in\Theta} |g_x(t,x,u,\theta) - g_x(t,x',u,\theta)| \\ &\leq C|x-x'| \end{aligned} \tag{2.5}$$

$$\sup_{\theta\in\Theta} |g(t,x,u,\theta)| \leq C\left(1+|x|\right). \tag{2.6}$$

Assumption (H2). The function $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable in

(x, y), and there exists a constant $C > 0$ such that

$$|h(x,y) - h(x',y'))| + |h_x(x,y) - h_x(x',y'))| \leq C\left[|x - x'| + |y - y'|\right]. \tag{2.7}$$

$$|h(x,y)| \leq C\left(1 + |x| + |y|\right). \tag{2.8}$$

Under the above assumptions, the SDE-(2.1) has a unique strong solution $x^u(t)$ which is given by

$$\begin{aligned} x^u(t) = \zeta &+ \int_s^t f\left(r, x^u(r), \mathbb{E}(x^u(r)), u(r)\right) dr + \int_s^t \sigma\left(r, x^u(r), \mathbb{E}(x^u(r)), u(r)\right) dW(r) \\ &+ \int_s^t \int_\Theta g\left(t, x^u(r^-), u(r), \theta\right) N\left(d\theta, dr\right), \end{aligned}$$

and by standard arguments it is easy to show that for any $q > 0$, it holds that

$$\mathbb{E}(\sup_{t\in[s,T]} |x^u(t)|^q) < C\left(q\right),$$

where $C\left(q\right)$ is a constant depending only on q and the functional $J^{s,\zeta}$ is well defined.

We introduce the adjoint equations as follows. The first-order adjoint equation turns out to be a linear mean-field backward SDE, while the second-order adjoint equation remains the same as in Peng [36], see also Zhou [45].

Definition 2.2.1. (*Adjoint equation for mean-field jump diffusion processes*) For any $u(\cdot) \in \mathcal{U}$ and the corresponding state trajectory $x(\cdot)$, we define the first-order adjoint process $(\Psi(\cdot), K(\cdot), \gamma(\cdot))$ and the second-order adjoint process $(Q(\cdot), R(\cdot), \Gamma(\cdot))$ as the ones satisfying the following equations:

(1) First-order adjoint equation: linear Backward SDE of mean-field type with jump processes

$$
\begin{cases}
-d\Psi(t) = \left\{ f_x^* \left(t, x(t), \mathbb{E}(x(t), u(t)\right) \Psi(t) + \mathbb{E}\left[f_y^* \left(t, x(t), \mathbb{E}(x(t), u(t)\right) \Psi(t) \right] \right. \\
\quad + \sigma_x^* \left(t, x(t), \mathbb{E}(x(t), u(t)\right) K(t) + \mathbb{E}\left[\sigma_y^* \left(t, x(t), \mathbb{E}(x(t), u(t)\right) K(t) \right] \\
\quad + \ell_x \left(t, x(t), \mathbb{E}(x(t), u(t)\right) + \mathbb{E}\left[\ell_y \left(t, x(t), \mathbb{E}(x(t), u(t)\right) \right] \\
\quad \left. + \int_\Theta g_x^* \left(t, x(t^-), u(t), \theta\right) \gamma_t(\theta) \mu(d\theta) \right\} dt \\
\quad - K(t) dW(t) - \int_\Theta \gamma_t(\theta) N(dt, d\theta) \\
\Psi(T) = h_x \left(x(T), \mathbb{E}(x(T)\right) + \mathbb{E}\left[h_y \left(x(T), \mathbb{E}(x(T)\right) \right].
\end{cases}
\tag{2.9}
$$

(2) Second-order adjoint equation: classical linear Backward SDE with jump processes

$$
\begin{cases}
-dQ(t) = \left\{ f_x^* \left(t, x(t), \mathbb{E}(x(t)), u(t)\right) Q(t) + Q_t f_x^* \left(t, x(t), \mathbb{E}(x(t), u(t)\right) \right. \\
\quad + \sigma_x^* \left(t, x(t), \mathbb{E}(x(t)), u(t)\right) Q(t) \sigma_x^* \left(t, x(t), \mathbb{E}(x(t)), u(t)\right) \\
\quad + \sigma_x^* \left(t, x(t), \mathbb{E}(x(t)), u(t)\right) R(t) + R(t) \sigma_x \left(t, x(t), \mathbb{E}(x(t)), u(t)\right) \\
\quad - \int_\Theta g_x^* \left(t, x(t^-), u(t), \theta\right) \left(\Gamma_t(\theta) + Q(t)\right) g_x \left(t, x(t^-), u(t), \theta\right) \mu(d\theta) \\
\quad - \int_\Theta \Gamma_t(\theta) g_x \left(t, x(t^-), u(t), \theta\right) + g_x^* \left(t, x(t^-), u(t), \theta\right) \Gamma_t(\theta) \mu(d\theta) \\
\quad \left. - H_{xx}(t, x(t), E(x(t)), u(t), \Psi(t), K(t), \gamma_t(\theta)) \right\} dt - R(t) dW(t) \\
\quad - \int_\Theta \Gamma_t(\theta) N(dt, d\theta) \\
Q(T) = h_{xx} \left(x(T), \mathbb{E}(x(T))\right),
\end{cases}
\tag{2.10}
$$

As it is well known that under conditions (**H1**) and (**H2**) the first-order adjoint equation (2.7) admits one and only one $\mathcal{F}_t$−adapted solution pair $(\Psi(\cdot), K(\cdot), \gamma(\cdot)) \in \mathbb{L}^2_{\mathcal{F}}([s,T]; \mathbb{R}^n) \times \mathbb{L}^2_{\mathcal{F}}\left([s,T]; \mathbb{R}^{n\times d}\right) \times \mathbb{L}^2_{\mathcal{F}}([s,T]; \mathbb{R}^{n\times m})$. This equation reduces to the standard one, when the coefficients do not explicitly depend on the expected value (or the marginal law) of the underlying diffusion process. Also the second-order adjoint equation (2.8) admits one and only one $\mathcal{F}_t$−adapted solution pair $(Q(\cdot), R(\cdot), \Gamma(\cdot)) \in \mathbb{L}^2_{\mathcal{F}}([s,T]; \mathbb{R}^{n\times n}) \times \mathbb{L}^2_{\mathcal{F}}\left([s,T]; (\mathbb{R}^{n\times n})^d\right) \times$

$\mathbb{L}^2_{\mathcal{F}}\left([s,T];(\mathbb{R}^{n\times n})^m\right)$. Moreover, since $f_x, f_y, \sigma_x, \sigma_y, \ell_x, \ell_x$ and h_x are bounded, by C by assumptions (**H1**) and (**H2**), we have the following estimate

$$\begin{aligned}\mathbb{E}\Big[&\sup\nolimits_{s\leq t\leq T}|\Psi(t)|^2+\textstyle\int_s^T|K(t)|^2dt+\int_s^T\int_\Theta|\gamma_t(\theta)|^2\mu(d\theta)dt\\&+\sup\nolimits_{s\leq t\leq T}|Q(t)|^2+\textstyle\int_s^T|R(t)|^2dt++\int_s^T\int_\Theta|\Gamma_t(\theta)|^2\mu(d\theta)dt\Big]\leq C.\end{aligned}\tag{2.11}$$

Definition 2.2.2. (Usual Hamiltonian and $\mathcal{H}$-function). We define the usual Hamiltonian associated with the mean-field stochastic control problem (2.3)-(2.4) as follows

$$\begin{aligned}H\left(t,X,\mathbb{E}\left(X\right),u,p,q,\varphi\right):=&-pf\left(t,X,\mathbb{E}\left(X\right),u\right)-q\sigma\left(t,X,\mathbb{E}\left(X\right),u\right)\\&-\int_\Theta\varphi g\left(t,x(t^-),u(t),\theta\right)\mu(d\theta)\\&-\ell\left(t,X,\mathbb{E}\left(X\right),u\right),\end{aligned}$$

where $(t,X,u)\in[s,T]\times R^n\times\mathbb{A}$ and X is a random variable such that $X\in\mathbb{L}^1([s,T];\mathbb{R}^n)$. Furthermore, we define the $\mathcal{H}$-function corresponding to a given admissible pair $(z(\cdot),v(\cdot))$ as follows

$$\begin{aligned}\mathcal{H}^{(z(\cdot),v(\cdot))}(t,x,u)=H&\left(t,x,\mathbb{E}\left(x\right),u,\Psi(t),K(t)-Q(t)\sigma\left(t,z(t),\mathbb{E}\left(z(t)\right),v(t)\right),\right.\\&\left.\gamma_t(\theta)-\left(Q(t)+\gamma_t(\theta)\right)g\left(t,z(t^-),v(t),\theta\right)\right)\\&-\frac{1}{2}\sigma^*\left(t,x,\mathbb{E}(x),u\right)Q(t)\sigma\left(t,x,\mathbb{E}(x),u\right),\\&-\frac{1}{2}\int_\Theta g^*\left(t,x,u,\theta\right)\left(Q(t)+\gamma_t(\theta)\right)g\left(t,x,u,\theta\right)\mu(d\theta).\end{aligned}$$

This shows that

$$
\begin{aligned}
\mathcal{H}^{(z(.),v(\cdot))}(t,x,u) = {} & H\left(t,x,\mathbb{E}\left(x\right),u,\Psi(t),K(t),\gamma_t(\theta)\right) \\
& + \sigma^*\left(t,x,\mathbb{E}\left(x\right),u\right)Q(t)\sigma\left(t,z(t),\mathbb{E}\left(z(t)\right),v(t)\right) \\
& - \frac{1}{2}\sigma^*\left(t,x,\mathbb{E}(x),u\right)Q(t)\sigma\left(t,x,\mathbb{E}(x),u\right) \\
& + \int_\Theta g^*\left(t,x,u,\theta\right)\left(Q(t)+\gamma_t(\theta)\right)g\left(t,z(t^-),v(t),\theta\right)\mu(d\theta) \\
& - \frac{1}{2}\int_\Theta g^*\left(t,x,u,\theta\right)\left(Q(t)+\gamma_t(\theta)\right)g\left(t,x,u,\theta\right)\mu(d\theta),
\end{aligned}
$$

where $\Psi(t), K(t), \gamma_t(\theta)$ and $Q(t)$ are determined by adjoint equations (2.9) and (2.10) corresponding to $\left(z\left(\cdot\right),v(\cdot)\right)$.

Before concluding this section, let us recall the definition of near-optimal controls as given in Zhou [[45], *Definitions (2.1)-(2.2)*], and Ekeland's variational principle, which will be used in the sequel.

Definition 2.2.3. *(Near-optimal control of order ε^λ.) For a given $\varepsilon > 0$ the admissible control $u^\varepsilon(\cdot)$ is near-optimal with respect (s,ζ) if*

$$
\left|J^{s,\zeta}\left(u^\varepsilon(\cdot)\right) - V\left(s,\zeta\right)\right| \leq \mathcal{O}\left(\varepsilon\right), \tag{2.12}
$$

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

1. If $\mathcal{O}\left(\varepsilon\right) = C\varepsilon^\lambda$ for some $\lambda > 0$ independent of the constant C then $u^\varepsilon(\cdot)$ is called near-optimal control of order ε^λ.

2. If $\mathcal{O}\left(\varepsilon\right) = C\varepsilon$, the admissible control $u^\varepsilon(\cdot)$ called $\varepsilon-$optimal.

Lemma 2.2.1. (*Ekeland's Variational Principle [12]*) *Let $(F,\, d_F)$ be a complete metric space and $f : F \to \overline{\mathbb{R}}$ be a lower semi-continuous function which is bounded from below. For* *a given* $\varepsilon > 0$*, suppose that* $u^{\varepsilon} \in F$ *satisfying*

$$f\left(u^{\varepsilon}\right) \leq \inf_{u \in F}\left(f(u)\right) + \varepsilon.$$

Then for any $\delta > 0$*, there exists* $u^{\delta} \in F$ *such that*

1. $f\left(u^{\delta}\right) \leq f\left(u^{\varepsilon}\right).$

2. $d_F\left(u^{\delta}, u^{\varepsilon}\right) \leq \delta.$

3. $f\left(u^{\delta}\right) \leq f\left(u\right) + \dfrac{\varepsilon}{\delta} d_F\left(u, u^{\delta}\right)$*, for all* $u \in F$.

Now, in order to apply Ekeland's principle to our Mean-field control problem, we have to endow the set of admissible controls $\mathcal{U}$ with an appropriate metric. We define a distance function d on the space of admissible controls $\mathcal{U}$ such that $(\mathcal{U}, d)$ becomes a complete metric space. For any $u(\cdot)$ and $v(\cdot) \in \mathcal{U}$ we set

$$d\left(u(\cdot), v(\cdot)\right) = \mathbb{P}\otimes dt\left\{(w,t) \in \Omega \times [s,T] : u\left(w,t\right) \neq v\left(w,t\right)\right\}, \tag{2.13}$$

where $\mathbb{P}\otimes dt$ is the product measure of $\mathbb{P}$ with the Lebesgue measure dt on $[s,T]$. Moreover, it has been shown in the book by Yong and Zhou ([109], 146-147) that

1. $(\mathcal{U}, d)$ is a complete metric space

2. The cost function $J^{s,\zeta}$ is continuous from $\mathcal{U}$ into $\mathbb{R}$.

2.3 Necessary conditions of near-optimality for mean-field jump diffusion processes

In this section, we obtain a Zhou-type necessary conditions of near-optimality, where the system is described by nonlinear controlled jump diffusion processes of mean-field type. The control domain is not need to be convex. (a general action space).

The proof of our theorem follows the general ideas as in Zhou [45], Buckdahn et al., [5], and Tang et al., [40].

The following theorem constitutes the main contribution of this work.

Let $(\Psi^{\varepsilon}(\cdot), K^{\varepsilon}(\cdot), \gamma^{\varepsilon}(\cdot))$ and $(Q^{\varepsilon}(\cdot), R^{\varepsilon}(\cdot), \Gamma^{\varepsilon}(\cdot))$ be the solution of adjoint equations (2.7) and (2.8) respectively, corresponding to $u^{\varepsilon}(\cdot)$.

Theorem 2.3.1. *(Mean-field stochastic maximum principle for any near-optimal control). For any $\delta \in [0, \frac{1}{3})$, and any near-optimal control $u^{\varepsilon}(\cdot)$ there exists a positive constant $C = C(\delta, \mu(\Theta))$ such that for each $\varepsilon > 0$* it holds that

$$\begin{aligned}
&\mathbb{E}\int_s^T \Big\{ \tfrac{1}{2}\left(\sigma\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u\right) - \sigma\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t)\right)\right)^* Q^{\varepsilon}(t) \\
&\quad \times \left(\sigma\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u\right) - \sigma\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t)\right)\right) \\
&\quad + \Psi^{\varepsilon}(t)\left(f\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u\right) - f\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t)\right)\right) \\
&\quad + K^{\varepsilon}(t)\left(\sigma\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u\right) - \sigma\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t)\right)\right) \\
&\quad + \int_{\Theta} \gamma^{\varepsilon}(t) g\left(t, x^{\varepsilon}(t), u, \theta\right) - g\left(t, x^{\varepsilon}(t), u^{\varepsilon}(t), \theta\right)\mu(d\theta) \\
&\quad + \tfrac{1}{2}\int_{\Theta}\left(g^*\left(t, x^{\varepsilon}(t), u, \theta\right) - g^*\left(t, x^{\varepsilon}(t), u^{\varepsilon}(t), \theta\right)\right)\left(Q^{\varepsilon}(t) + \gamma_t^{\varepsilon}(\theta)\right) \\
&\quad \times \left(g\left(t, x^{\varepsilon}(t), u, \theta\right) - g\left(t, x^{\varepsilon}(t), u^{\varepsilon}(t), \theta\right)\right)\mu(d\theta), \\
&\quad + \left(\ell\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u\right) - \ell\left(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t)\right)\right)\Big\}\, dt \geq -C\varepsilon^{\delta},
\end{aligned} \tag{2.14}$$

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

$$\begin{aligned}
&\mathbb{E}\int_s^T \mathcal{H}^{(x^{\varepsilon}(.), u^{\varepsilon}(.))}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t)) dt \\
&\geq \sup_{u(\cdot)\in\mathcal{U}} \mathbb{E}\int_s^T \mathcal{H}^{(x^{\varepsilon}(.), u^{\varepsilon}(.))}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u(t)) dt - C\varepsilon^{\delta}.
\end{aligned} \tag{2.15}$$

To prove *Theorem 2.3.1* and *Corollary 2.3.1*, we need the following auxiliary results on the stability of the state and adjoint processes with respect to the control variable.

In what follows, C represents a generic constant, which can be different from line to line.

Our first Lemma below deals with the continuity of the state processes under distance d.

Lemma 2.3.1. *If $x^u(t)$ and $x^v(t)$ be the solution of the state equation (2.1) associated respectively with $u(t)$* and *$v(t)$. For any $\alpha \in (0,1)$* and *$\beta \geq 0$ satisfying $\alpha\beta < 1$, there exists a positive constants $C = C\left(T, \alpha, \beta, \mu(\Theta)\right)$ such that*

$$\mathbb{E}(\sup_{s\leq t\leq T} \left|x^u(t) - x^v(t)\right|^{2\beta}) \leq Cd^{\alpha\beta}\left(u(\cdot), v(\cdot)\right). \tag{2.16}$$

Proof. We consider the following two cases:

Case 1. First, we assume that $\beta \geq 1$. Using Burkholder-Davis-Gundy inequality for the martingale part and Propositions A2 (see Appendix) we can compute, for any $r \geq s$:

$$\begin{aligned}
&\mathbb{E}(\sup_{s\leq t\leq r} \left|x^u(t) - x^v(t)\right|^{2\beta}) \\
&\leq C\mathbb{E}(\int_s^r \Big\{\left|f\left(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t)\right) - f(x^v(t), \mathbb{E}\left(x^v(t)\right), v(t))\right|^{2\beta} \\
&+ \int_s^r \left|\sigma\left(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t)\right) - \sigma(x^v(t), \mathbb{E}\left(x^v(t)\right), v(t)\right|^{2\beta} \\
&+ \int_\Theta \left|g\left(t, x^u(t), u, \theta\right) - g\left(t, x^v(t), v(t), \theta\right)\right|^{2\beta} \mu(d\theta)\Big\}\, dt \\
&\leq I_1 + I_2,
\end{aligned}$$

where

$$\begin{aligned}
I_1 \leq\ & C\mathbb{E}(\int_s^r \Big\{\left|f\left(x^u(t), \mathbb{E}\left(x^u(t)\right), u(t)\right) - f\left(x^u(t), \mathbb{E}\left(x^u(t)\right), v(t)\right)\right|^{2\beta} \\
&+ \int_s^r \left|\sigma\left(x^u(t), \mathbb{E}\left(x^u(t)\right), u(t)\right) - \sigma\left(x^u(t), \mathbb{E}\left(x^u(t)\right), v(t)\right)\right|^{2\beta} \\
&+ \mu(\Theta) \sup_{\theta\in\Theta} \left|g\left(t, x^u(t), u(t), \theta\right) - g\left(t, x^v(t), v(t), \theta\right)\right|^{2\beta}\Big\} \mathbf{I}_{\{u(t)\neq v(t)\}}\left(t\right) dt
\end{aligned}$$

and

$$I_2 \leq C\mathbb{E}(\int_s^r \Big\{ \left| f\left(x^u(t), \mathbb{E}\left(x^u(t)\right), v(t)\right) - f\left(x^v(t), \mathbb{E}\left(x^v(t)\right), v(t)\right)\right|^{2\beta}$$
$$+ \int_s^r \left| \sigma\left(x^u(t), \mathbb{E}\left(x^u(t)\right), v(t)\right) - \sigma\left(x^v(t), \mathbb{E}(x^v(t)), v\left(t\right)\right)\right|^{2\beta}$$
$$+ \mu(\Theta)(\sup_{\theta\in\Theta} \left| g\left(t, x^u(t), v(t), \theta\right) - g\left(t, x^v(t), v(t), \theta\right) \right|)^{2\beta} \Big\}$$

Now arguing as in ([45], *Lemma 3.1*) taking $b = \frac{1}{\alpha\beta} > 1$ and $a > 1$ such that $\frac{1}{a} + \frac{1}{b} = 1$, and applying Cauchy-Schwarz inequality, we get

$$\mathbb{E} \int_s^r \left| f\left(t, x^{u,\eta}(t), \mathbb{E}\left(x^{u,\eta}(t)\right), u(t)\right) - f\left(x^{u,\eta}(t), \mathbb{E}\left(x^{u,\eta}(t)\right), v(t)\right)\right|^{2\beta} \mathbf{I}_{\{u(t)\neq v(t)\}}\left(t\right) dt$$
$$\leq \left\{ \mathbb{E} \int_s^r \left| f\left(t, x^{u,\eta}(t), \mathbb{E}\left(x^{u,\eta}(t)\right), u(t)\right) - f\left(x^{u,\eta}(t), \mathbb{E}\left(x^{u,\eta}(t)\right), v(t)\right)\right|^{2\beta a} dt \right\}^{\frac{1}{a}}$$
$$\times \left\{ \mathbb{E} \int_s^r \mathbf{I}_{\{u(t)\neq v(t)\}}\left(t\right) dt \right\}^{\frac{1}{b}},$$

by using definition of d and linear growth condition on f with respect to x and y, (assumption 2.4) we obtain

$$\mathbb{E} \int_s^r \left| f\left(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t)\right) - f\left(t, x^u(t), \mathbb{E}\left(x^u(t)\right), v(t)\right)\right|^{2\beta} \mathbf{I}_{\{u(t)\neq v(t)\}}\left(t\right) dt$$
$$\leq C \left\{ \mathbb{E} \int_s^r \left(1 + \left|x^u(t)\right|^{2\beta a} + \left|\mathbb{E}\left(x^u(t)\right)\right|^{2\beta a}\right) dt \right\}^{\frac{1}{a}} d\left(u(.), v(.)\right)^{\alpha\beta} \leq Cd\left(u(.), v(.)\right)^{\alpha\beta}.$$

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

$$\mathbb{E} \int_s^r \left| \sigma\left(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t)\right) - \sigma\left(t, x^u(t), \mathbb{E}\left(x^u(t)\right), v(t)\right)\right|^{2\beta} \mathbf{I}_{\{u(t)\neq v(t)\}}\left(t\right) dt$$
$$\leq Cd\left(u(.), v(.)\right)^{\alpha\beta}.$$

and

$$\mathbb{E}\int_s^r \left(\sup_{\theta\in\Theta}\left|g\left(t,x^u(t),u,\theta\right)-g\left(t,x^v(t),v(t),\theta\right)\right|\right)^{2\beta}\mathbf{I}_{\{u(t)\neq v(t)\}}\left(t\right)dt\leq Cd\left(u(.),v(.)\right)^{\alpha\beta}.$$

This implied that $I_1\leq Cd\left(u(.),v(.)\right)^{\alpha\beta}$.

Since the coefficients f,σ and g are Lipschitz with respected to x and y (assumption (**H1**)) we conclude that

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

Hence (2.17) follows immediately from Gronwall's inequality.

Case 2. Now we assume $0\leq\beta<1$. Since $\frac{2}{\alpha}>1$ then the Cauchy-Schwarz inequality yields

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

This completes the proof of *Lemma 3.1*. □

The next result gives the β^{-th} moment continuity of the solutions to adjoint equations with respect to the metric d. This Lemma is an extension of *Lemma 3.2* in Zhou [45] to mean-field SDEs with jump processes.

Lemma 2.3.2. *For any $\alpha\in(0,1)$ and $\beta\in(1,2)$ satisfying $(1+\alpha)\beta<2$, there exist a positive constant $C=C\left(\alpha,\beta,\mu(\Theta)\right)$* such that for any $u(\cdot),\ v(\cdot)\in\mathcal{U}$, along with the corresponding trajectories $x^u(\cdot)$, $x^v(\cdot)$ and the solutions $(\Psi^u(\cdot),K^u(\cdot),\gamma^u(\cdot),Q^u(\cdot),R^u(\cdot),\Gamma^u(\cdot))$ and $(\Psi^v(\cdot),K^v(\cdot),\gamma^v(\cdot),Q^v(\cdot),R^v(\cdot),\Gamma^v(\cdot))$ *of the corresponding adjoint equations (2.9)-(2.10), it holds that*

$$\begin{aligned}&\mathbb{E}\int_s^T\left(\left|\Psi^u(t)-\Psi^v(t)\right|^{\beta}+\left|K^u(t)-K^v(t)\right|^{\beta}\right)dt\\&+\mathbb{E}\int_s^T\int_\Theta\left|\gamma_t^u(\theta)-\gamma_t^v(\theta)\right|^{\beta}\mu(d\theta)dt\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}},\end{aligned}\tag{2.17}$$

and

$$
\begin{aligned}
&\mathbb{E}\int_s^T \left(|Q^u(t) - Q^v(t)|^\beta + |R^u(t) - R^v(t)|^\beta \right) dt \\
&+\mathbb{E}\int_s^T \int_\Theta |\Gamma_t^u(\theta) - \Gamma_t^v(\theta)|^\beta \mu(d\theta)dt \leq Cd\left(u(\cdot), v(\cdot)\right)^{\frac{\alpha\beta}{2}} .
\end{aligned}
\tag{2.18}
$$

Proof. Note that $\widetilde{\Psi}(t) = \Psi^u(t) - \Psi^v(t)$, $\widetilde{K}(t) = K^u(t) - K^v(t)$ and $\widetilde{\gamma}_t(\theta) = \gamma_t^u(\theta) - \gamma_t^v(\theta)$ satisfied the following BSDEs:

$$
\begin{cases}
-d\widetilde{\Psi}(t) = \Big[f_x^*\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) \widetilde{\Psi}(t) + \sigma_x^*\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) \widetilde{K}(t) \\
\qquad\qquad + \int_\Theta g_x^*\left(t, x^u(t), u, \theta\right) \widetilde{\gamma}_t(\theta)\mu(d\theta) + \mathcal{L}(t) \Big] \, dt \\
\qquad\qquad - \widetilde{K}(t)dW(t) - \int_\Theta \widetilde{\gamma}_t(\theta) N(d\theta, dt) \\
\widetilde{\Psi}(T) = h_x\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_x\left(x^v(T), \mathbb{E}(x^v(T))\right) \\
\qquad\qquad + \mathbb{E}[h_y\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_y\left(x^v(T), \mathbb{E}(x^v(T))\right)].
\end{cases}
\tag{2.19}
$$

where the process $\mathcal{L}(t)$ is given by

$$
\begin{aligned}
\mathcal{L}(t) = {} & \left[f_x^*\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) - f_x^*\left(t, x^v(t), \mathbb{E}(x^v(t)), v(t)\right) \right] \Psi^v(t) \\
& + \left[\sigma_x^*\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) - \sigma_x^*\left(t, x^v(t), \mathbb{E}(x^v(t)), v(t)\right) \right] K^v(t) \\
& + \left(\ell_x\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) - \ell_x\left(t, x^v(t), \mathbb{E}(x^v(t)), v(t)\right) \right) \\
& + \mathbb{E}\left\{ f_y^*\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) \Psi^u(t) - f_y^*\left(t, x^v(t), \mathbb{E}(x^v(t)), v(t)\right) \Psi^v(t) \right\} \\
& + \mathbb{E}\left\{ \left[\sigma_y^*\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) K^u(t) - \sigma_y^*\left(t, x^v(t), \mathbb{E}(x^v(t)), v(t)\right) K^v(t) \right] \right\} \\
& + \mathbb{E}\left(\ell_y\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) - \ell_y\left(t, x^v(t), \mathbb{E}(x^v(t)), v(t)\right) \right) \\
& + \int_\Theta (g_x^*\left(t, x^u(t_-), u, \theta\right) - g_x^*\left(t, x^v(t_-), v, \theta\right)) \gamma_t^v(\theta)\mu(d\theta).
\end{aligned}
\tag{2.20}
$$

Let $\phi(\cdot)$ be the solution of the following linear SDE

$$\begin{cases} d\phi(t) = \left[f_x\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) \phi(t) + \left|\widetilde{\Psi}(t)\right|^{\beta-1} Sgn(\widetilde{\Psi}(t)) \right] dt \\ \qquad + \left[\sigma_x\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) \phi(t) + \left|\widetilde{K}(t)\right|^{\beta-1} Sgn(\widetilde{K}(t)) \right] dW(t) \\ \qquad + \left[\int_\Theta g_x^*\left(t, x^u(t_-), u, \theta\right) \phi(t) + \left|\widetilde{\gamma}_t(\theta)\right|^{\beta-1} Sgn(\widetilde{\gamma}_t(\theta)) \right] N(d\theta, dt) \\ \phi(s) = 0, \end{cases} \tag{2.21}$$

where $Sgn\,(a) \equiv (Sgn(a_1), Sgn(a_2), ..., Sgn(a_n))^*$ for any vector $a = (a_1, a_2, .., a_n)^*$.

It is worth mentioning that since f_x σ_x and g_x are bounded and the fact that

$$\begin{cases} \mathbb{E} \int_s^T \left\{ \left| \left|\widetilde{\Psi}(t)\right|^{\beta-1} Sgn(\widetilde{\Psi}(t)) \right|^2 + \left| \left|\widetilde{K}(t)\right|^{\beta-1} Sgn(\widetilde{K}(t)) \right|^2 \right\} dt \\ \qquad + \mathbb{E} \int_s^T \int_\Theta \left| \left|\widetilde{\gamma}_t(\theta)\right|^{\beta-1} Sgn(\widetilde{\gamma}_t(\theta)) \right|^2 \mu(d\theta) dt < \infty, \end{cases} \tag{2.22}$$

then the SDE (2.21) has a unique strong solution.

Let $\eta \geq 2$ such that $\frac{1}{\eta} + \frac{1}{\beta} = 1$, $\beta \in (1, 2)$ then we get

$$\begin{aligned} \mathbb{E}(\sup_{s \leq t \leq T} |\phi(t)|^\eta) &\leq C\mathbb{E} \int_s^T \left\{ \left|\widetilde{\Psi}(t)\right|^{\beta\eta-\eta} + \left|\widetilde{K}(t)\right|^{\beta\eta-\eta} \right\} dt \\ &\quad + \mathbb{E} \int_s^T \int_\Theta \left|\widetilde{\gamma}_t(\theta)\right|^{\beta\eta-\eta} \mu(d\theta) dt \\ &\leq C\mathbb{E} \int_s^T \left\{ \left|\widetilde{\Psi}(t)\right|^\beta + \left|\widetilde{K}(t)\right|^\beta + \int_\Theta \left|\widetilde{\gamma}_t(\theta)\right|^\beta \mu(d\theta) \right\} dt \end{aligned}$$

Note that the right hand side term of the above inequality is bounded due to (2.9), then we get

$$\mathbb{E}(\sup_{s \leq t \leq T} |\phi(t)|^\eta) < \infty. \tag{2.23}$$

By applying Itô's formula for jump processes (see *Appendix Lemma A1*) to $\widetilde{\Psi}(t)\phi(t)$ on $[s, T]$

and taking expectation, we get

$$\begin{aligned}
&\mathbb{E}\int_s^T \left\{ \widetilde{\Psi}(t) \left|\widetilde{\Psi}(t)\right|^{\beta-1} Sgn(\widetilde{\Psi}(t)) + \widetilde{K}(t) \left|\widetilde{K}(t)\right|^{\beta-1} Sgn(\widetilde{K}(t)) \right. \\
&\left. + \int_\Theta \widetilde{\gamma}_t(\theta) \left|\widetilde{\gamma}_t(\theta)\right|^{\beta-1} Sgn(\widetilde{\gamma}_t(\theta))\mu(d\theta) \right\} dt \\
&= \mathbb{E}\left\{ \int_s^T \mathcal{L}(t)\phi(t)dt + \widetilde{\Psi}(T)\phi(T) \right\} \\
&= \mathbb{E}\int_s^T \mathcal{L}(t)\phi(t)dt + \mathbb{E}\left\{ (h_x\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_x\left(x^v(T), \mathbb{E}(x^v(T))\right))\phi(T) \right\} \\
&+ \mathbb{E}[h_y\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_y\left(x^v(T), \mathbb{E}(x^v(T)\right)]\mathbb{E}\left(\phi(T)\right).
\end{aligned}$$

Since

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

and fact that

$$\begin{aligned}
&\mathbb{E}\left\{ \int_s^T \mathcal{L}(t)\phi(t)dt + [(h_x\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_x\left(x^v(T), \mathbb{E}(x^v(T))\right)) \right. \\
&\left. + \mathbb{E}(h_y\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_y\left(x^v(T), \mathbb{E}(x^v(T))\right))]\left(\phi(T)\right) \right\} \\
&\leq \left[\mathbb{E}\int_s^T \left|\mathcal{L}(t)\right|^\beta dt \right]^{\frac{1}{\beta}} \left[\mathbb{E}\int_s^T \left|\phi(t)\right|^\eta dt \right]^{\frac{1}{\eta}} \\
&+ [\mathbb{E}\left|(h_x\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_x\left(x^v(T), \mathbb{E}(x^v(T))\right))\right. \\
&+ \left.\mathbb{E}(h_y\left(x^u(T), \mathbb{E}(x^u(T))\right) - h_y\left(x^v(T), \mathbb{E}(x^v(T))\right))\right|^\eta]^{\frac{1}{\beta}} \left[\mathbb{E}\left|\phi(T)\right|^\eta\right]^{\frac{1}{\eta}}
\end{aligned}$$

then according to (2.23) we deduce

$$\begin{aligned}&\mathbb{E}\int_s^T\left[\left|\widetilde{\Psi}(t)\right|^\beta+\left|\widetilde{K}(t)\right|^\beta+\int_\Theta\left|\widetilde{\gamma}_t(\theta)\right|^\beta\mu(d\theta)\right]dt\leq C\mathbb{E}\int_s^T\left|\mathcal{L}(t)\right|^\beta dt\\&\quad+C\mathbb{E}\left\{\left|h_x(x^u(T),\mathbb{E}(x^u(T)))-h_x(x^v(T),\mathbb{E}(x^v(T)))\right|^\beta\right.\\&\quad\left.+\left|\mathbb{E}(h_y(x^u(T),\mathbb{E}(x^u(T))))-\mathbb{E}(h_y(x^v(T),\mathbb{E}(x^v(T))))\right|^\beta\right\}.\end{aligned}\tag{2.24}$$

We proceed to estimate the right hand side of (2.24). First noting that $\frac{\alpha\beta}{2}<1-\frac{\beta}{2}<1$ then by using assumption (**H2**) and *Lemma 2.3.1*, we obtain

$$\begin{aligned}&\mathbb{E}\left|h_x(x^u(T),\mathbb{E}(x^u(T)))-h_x(x^v(T),\mathbb{E}(x^v(T)))\right|^\beta\\&\qquad\leq C\mathbb{E}\left|x^u(T)-x^v(T)\right|^\beta\leq Cd(u(\cdot),v(\cdot))^{\frac{\alpha\beta}{2}}.\\&\mathbb{E}\left|\mathbb{E}(h_y(x^u(T),\mathbb{E}(x^u(T))))-\mathbb{E}(h_y(x^v(T),\mathbb{E}(x^v(T))))\right|^\beta\\&\qquad\leq Cd(u(\cdot),v(\cdot))^{\frac{\alpha\beta}{2}}.\end{aligned}\tag{2.25}$$

Now, to prove inequality (2.17) it sufficient to estimate $\mathbb{E}\int_s^T\left|\mathcal{L}(t)\right|^\beta dt$. By repeatedly using Cauchy-Schwarz inequality and assumption (**H2**) we can estimate

$$\begin{aligned}&\mathbb{E}\int_s^T\left|f_x^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_x^*\left(t,x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right|^\beta\left|\Psi^v(t)\right|^\beta dt\\&\leq C\mathbb{E}\int_s^T\left\{\left|f_x^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_x^*\left(t,x^u(t),\mathbb{E}(x^u(t)),v(t)\right)\right|^\beta\left|\Psi^v(t)\right|^\beta\right.\\&\left.+\left|f_x^*\left(t,x^u(t),\mathbb{E}(x^u(t)),v(t)\right)-f_x^*\left(t,x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right|^\beta\left|\Psi^v(t)\right|^\beta\right\}dt\\&\leq C\mathbb{E}\int_s^T\left\{\mathbf{I}_{\{u(t)\neq v(t)\}}(t)\left|\Psi^v(t)\right|^\beta\right.\\&\left.+\left[\left|x^u(t)-x^v(t)\right|+\left|\mathbb{E}(x^u(t))-\mathbb{E}(x^v(t))\right|\right]^\beta\left|\Psi^v(t)\right|^\beta\right\}dt\\&\leq C\left[\mathbb{E}\int_s^T\left|\Psi^v(t)\right|^2dt\right]^{\frac{\beta}{2}}d(u(.),v(.))^{\frac{2-\beta}{2}}\\&+C\left[\mathbb{E}\int_s^T\left|\Psi^v(t)\right|^2dt\right]^{\frac{\beta}{2}}\left[\mathbb{E}\int_s^T\left|x^u(t)-x^v(t)\right|^{\frac{2\beta}{2-\beta}}dt\right]^{\frac{2-\beta}{2}}.\end{aligned}$$

By using the fact that $d(u(\cdot),v(\cdot))\leq 1$ and $\frac{\alpha\beta}{2}<1-\frac{\beta}{2}$, the first term of the right side of the above inequality is dominated by $d(u(\cdot),v(\cdot))^{\frac{\alpha\beta}{2}}$. Since $\frac{\alpha\beta}{2-\beta}<1$ and we have from *Lemma 2.3.1* that

$$\mathbb{E}\int_s^T |x^u(t)-x^v(t)|^{\frac{2\beta}{2-\beta}}\,dt\leq d(u(\cdot),v(\cdot))^{\frac{\alpha\beta}{2-\beta}},$$

then we have

$$\begin{aligned}&C\left[\mathbb{E}\int_s^T |\Psi^v(t)|^2\,dt\right]^{\frac{\beta}{2}} d(u(\cdot),v(\cdot))^{\frac{2-\beta}{2}}\\&+\left[\mathbb{E}\int_s^T |\Psi^v(t)|^2\,dt\right]^{\frac{\beta}{2}}\left[\mathbb{E}\int_s^T |x^u(t)-x^v(t)|^{\frac{2\beta}{2-\beta}}\,dt\right]^{\frac{2-\beta}{2}}\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}},\end{aligned}$$

we conclude that

$$\begin{aligned}&\mathbb{E}\int_s^T \left|f_x^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_x^*\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right|^{\beta}|\Psi^v(t)|^{\beta}\,dt\\&\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}}.\end{aligned}\tag{2.26}$$

A similar argument shows that

$$\begin{aligned}&\mathbb{E}\int_s^T \left|\sigma_x\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-\sigma_x\left(t,x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right|^{\beta}|K^v(t)|^{\beta}\,dt\\&\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}},\end{aligned}\tag{2.27}$$

and

$$\begin{aligned}&\mathbb{E}\int_s^T \left|\ell_x\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-\ell_x\left(t,x^v(t),\mathbb{E}(x^{v,\xi}(t)),v(t)\right)\right|^{\beta}dt\\&\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}},\end{aligned}\tag{2.28}$$

Now, by using similar arguments developed above and (2.9) we get

$$\begin{aligned}
&\mathbb{E}\int_s^T \left|\mathbb{E}\left\{\left[f_y^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_y^*\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right]\right.\right.\\
&\qquad \left.\left.\times\ \Psi^v(t)\right\}\right|^\beta dt\\
&\leq C\mathbb{E}\int_s^T \mathbb{E}\left|f_y^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_y^*\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right|^\beta \\
&\qquad \times\mathbb{E}\left[\left|\Psi^v(t)\right|\right]^\beta dt\\
&\leq C\mathbb{E}\int_s^T \mathbb{E}\left|f_y^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_y^*\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right|^\beta dt\\
&\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}}.
\end{aligned} \tag{2.29}$$

A similar argument shows that

$$\begin{aligned}
\mathbb{E}\int_s^T \left|\mathbb{E}\left\{\left[\sigma_y^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-\sigma_y^*\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right]\Psi^v(t)\right\}\right|^\beta dt&\\
\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}},&
\end{aligned} \tag{2.30}$$

$$\begin{aligned}
\mathbb{E}\int_s^T \left|\mathbb{E}\left\{\left[f_y^*\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-f_y^*\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right]\Psi^v(t)\right\}\right|^\beta dt&\\
\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}},&
\end{aligned} \tag{2.31}$$

and

$$\begin{aligned}
\mathbb{E}\int_s^T \left|\mathbb{E}\left\{\ell_y\left(t,x^u(t),\mathbb{E}(x^u(t)),u(t)\right)-\ell_y\left(x^v(t),\mathbb{E}(x^v(t)),v(t)\right)\right\}\right|^\beta dt&\\
\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}}.&
\end{aligned} \tag{2.32}$$

Next , by applying Cauchy-Schwarz inequality, we get

$$\begin{aligned}
&\mathbb{E}\int_s^T \left| \int_\Theta (g_x^*(t,x^u(t_-),u(t),\theta) - g_x^*(t,x^v(t_-),v(t),\theta))\gamma_t^v(\theta)\mu(d\theta) \right|^\beta dt \\
&= \mathbb{E}\int_s^T \left| \int_\Theta (g_x^*(t,x^u(t_-),u(t),\theta) - g_x^*(t,x^u(t_-),v(t),\theta))\gamma_t^v(\theta)\mu(d\theta) \right|^\beta dt \\
&+ \mathbb{E}\int_s^T \left| \int_\Theta (g_x^*(t,x^u(t_-),v(t),\theta) - g_x^*(t,x^v(t_-),v(t),\theta))\gamma_t^v(\theta)\mu(d\theta) \right|^\beta dt \\
&\leq \mathbb{I}_1 + \mathbb{I}_2,
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{I}_1 = \mathbb{E}\int_s^T & \left| \int_\Theta (g_x^*(t,x^u(t_-),u(t),\theta) - g_x^*(t,x^u(t_-),v(t),\theta))\gamma_t^v(\theta)\mu(d\theta) \right|^\beta \\
&\times \mathbf{I}_{\{u(t)\neq v(t)\}}(t)dt,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{I}_2 = \mathbb{E}\int_s^T & \left(\sup_{\theta\in\Theta} |(g_x^*(t,x^u(t_-),u(t),\theta) - g_x^*(t,x^u(t_-),v(t),\theta))| \right)^\beta \\
&\times \left(\left| \int_\Theta \gamma_t^v(\theta)\mu(d\theta) \right|^\beta \right) dt,
\end{aligned}$$

by using the fact that g_x is bounded, $d(u(\cdot),v(\cdot)) \leq 1$ and $\frac{\alpha\beta}{2} < 1 - \frac{\beta}{2}$, then due to (2.11) we get

$$\begin{aligned}
\mathbb{I}_1 &\leq C\mathbb{E}\left\{ \int_s^T \int_\Theta |\gamma_t^v(\theta)|^2 \mu(d\theta) \right\}^{\frac{\beta}{2}} \times \left\{ \int_s^T \mathbf{I}_{\{u(t)\neq v(t)\}}(t)dt \right\}^{1-\frac{\beta}{2}} \\
&\leq C\mathbb{E}\left\{ \int_s^T \int_\Theta |\gamma_t^v(\theta)|^2 \mu(d\theta) \right\}^{\frac{\beta}{2}} d\left(u(\cdot),v(\cdot)\right)^{1-\frac{\beta}{2}} \\
&\leq Cd\left(u(\cdot),v(\cdot)\right)^{\frac{\alpha\beta}{2}}.
\end{aligned} \tag{2.33}$$

Further, since $\frac{\alpha\beta}{2-\beta} < 1$ we conclude from *Lemma 2.3.1* and (2.11) that

$$\begin{aligned} \mathbb{I}_2 &\leq C\mathbb{E}\left(\int_s^T |x^u(t) - x^v(t)|^{\frac{2\beta}{2-\beta}} dt\right)^{1-\frac{\beta}{2}} \mathbb{E}\left(\int_s^T \left|\int_\Theta \gamma_t^v(\theta)\mu(d\theta)\right|^2 dt\right)^{\frac{\beta}{2}} \\ &\leq Cd\left(u(\cdot), v(\cdot)\right)^{\frac{\alpha\beta}{2}}, \end{aligned} \tag{2.34}$$

It follows from (2.33) and (2.34) that

$$\begin{aligned} &\mathbb{E}\int_s^T \left|\int_\Theta (g_x^*\left(t, x^u(t_-), u(t), \theta\right) - g_x^*\left(t, x^v(t_-), v(t), \theta\right))\gamma_t^v(\theta)\mu(d\theta)\right|^\beta dt \\ &\qquad\qquad \leq Cd\left(u(\cdot), v(\cdot)\right)^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{2.35}$$

We conclude from (2.26)$\sim$(2.35) that

$$\mathbb{E}\int_s^T \left|\mathcal{L}(t)\right|^\beta dt \leq Cd\left(u(\cdot), v(\cdot)\right)^{\frac{\alpha\beta}{2}}. \tag{2.36}$$

Finally, combining (2.24)-(2.25) and (2.36), the proof of (2.17) is complete. Similarly one can prove (2.19). This completes the proof of *Lemma 2.3.2*. $\square$

Now, let $(\overline{\Psi}^\varepsilon(\cdot), \overline{K}^\varepsilon(\cdot), \overline{\gamma}^\varepsilon(\cdot))$ and $(\overline{Q}^\varepsilon(\cdot), \overline{R}^\varepsilon(\cdot), \overline{\Gamma}^\varepsilon(\cdot))$ be the solution of adjoint equations (2.9)-(2.10) corresponding to $(\overline{x}^\varepsilon(\cdot), \mathbb{E}\left(\overline{x}^\varepsilon(\cdot)\right), \overline{u}^\varepsilon(\cdot))$.

Lemma 3.3. *For any $\varepsilon > 0$, there exists near-optimal control* $\overline{u}^{\varepsilon}(\cdot)$ such that for any $u \in \mathbb{A}$:

$$\begin{aligned}
&\mathbb{E}\int_s^T \Big\{\tfrac{1}{2}\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)^{*}\overline{Q}^{\varepsilon}(t)\\
&\quad\times\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)\\
&\quad+\overline{\Psi}^{\varepsilon}(t)\left(f\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-f\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)\\
&\quad+\overline{K}^{\varepsilon}(t)\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)\\
&\quad+\int_{\Theta}\overline{\gamma}^{\varepsilon}(t)g\left(t,\overline{x}^{\varepsilon}(t_{-}),u,\theta\right)-g\left(t,\overline{x}^{\varepsilon}(t_{-}),\overline{u}^{\varepsilon}(t),\theta\right)\mu(d\theta)\\
&\quad+\tfrac{1}{2}\int_{\Theta}(g^{*}\left(t,\overline{x}^{\varepsilon}(t_{-}),u,\theta\right)-g^{*}\left(t,\overline{x}^{\varepsilon}(t_{-}),\overline{u}^{\varepsilon}(t),\theta\right))(\overline{Q}^{\varepsilon}(t)+\overline{\gamma}_t^{\varepsilon}(\theta))\\
&\quad\times(g\left(t,\overline{x}^{\varepsilon}(t_{-}),u,\theta\right)-g\left(t,\overline{x}^{\varepsilon}(t_{-}),\overline{u}^{\varepsilon}(t),\theta\right))\mu(d\theta),\\
&\quad+\left(\ell\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\ell\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)\Big\}\,dt\geq-\varepsilon^{\frac{1}{3}},
\end{aligned} \tag{2.37}$$

Proof. By using Ekeland's variational principle with $\lambda = \varepsilon^{\frac{2}{3}}$, there is an admissible control $\overline{u}^{\varepsilon}(\cdot)$ such that for any $u(\cdot) \in \mathcal{U}$:

$$d\left(u^{\varepsilon}(\cdot),\overline{u}^{\varepsilon}(\cdot)\right)\leq\varepsilon^{\frac{2}{3}}, \tag{2.38}$$

and

$$J^{s,\zeta}\left(u^{\varepsilon}(\cdot)\right)\leq J^{s,\zeta}\left(u^{\varepsilon}(\cdot)\right)+\varepsilon^{\frac{1}{3}}d\left(u(\cdot),\overline{u}^{\varepsilon}(\cdot)\right).$$

Notice that $u^{\varepsilon}(\cdot)$ which is near-optimal for the initial cost $J^{s,\zeta}$ defined in (2.2) is an optimal control for the new cost $J^{s,\zeta,\varepsilon}$ given by

$$J^{s,\zeta,\varepsilon}\left(u(\cdot)\right)=J^{s,\zeta}\left(u(\cdot)\right)+\varepsilon^{\frac{1}{3}}d\left(u(\cdot),\overline{u}^{\varepsilon}(\cdot)\right).$$

Therefore we have

$$J^{s,\zeta,\varepsilon}\left(\overline{u}^{\varepsilon}(\cdot)\right)\leq J^{s,\zeta,\varepsilon}\left(u(\cdot)\right)\ \text{ for any } u(\cdot)\in\mathcal{U}.$$

Next, we use the spike variation techniques for $\overline{u}^{\varepsilon}(\cdot)$ to derive the variational inequality as follows. For $\hbar > 0$, we choose a Borel subset $\mathcal{E}_\hbar \subset [s, T]$ such that $|\mathcal{E}_\hbar| = \hbar$, and we consider the control process which is the spike variation of $\overline{u}^\varepsilon(\cdot)$:

$$\overline{u}^{\varepsilon,\hbar}(t) = \begin{cases} u : t \in \mathcal{E}_\hbar, \\ \overline{u}^\varepsilon(t) : t \in [s, T] \mid \mathcal{E}_\hbar, \end{cases}$$

where u is an arbitrary element of $\mathbb{A}$ be fixed. By using the fact that

1. $J^{s,\zeta,\varepsilon}\left(\overline{u}^\varepsilon(\cdot)\right) \leq J^{s,\zeta,\varepsilon}(\overline{u}^{\varepsilon,\hbar}(\cdot))$,

2. $d(\overline{u}^{\varepsilon,\hbar}(\cdot), \overline{u}^\varepsilon(\cdot)) = d(\overline{u}^{\varepsilon,\hbar}(\cdot), \overline{u}^\varepsilon(\cdot)) \leq \hbar$, we get

$$J^{s,\zeta}(\overline{u}^{\varepsilon,\hbar}(\cdot)) - J^{s,\zeta}(\overline{u}^\varepsilon(\cdot)) \geq -\varepsilon^{1/3} d(\overline{u}^\varepsilon(\cdot), \overline{u}^{\varepsilon,\hbar}(\cdot)) \geq -\varepsilon^{1/3}\hbar. \tag{2.39}$$

Arguing as in Hafayed et al., ([18], *Theorem 3.1*), the left-hand side of (2.39) is equal to

$$\begin{aligned}
&\mathbb{E}\int_{\mathcal{E}_\hbar} \left\{\tfrac{1}{2}\left(\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - \sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right)\right)^* \overline{Q}^\varepsilon(t)\right. \\
&\quad \times \left(\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - \sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right)\right) \\
&\quad + \overline{\Psi}^\varepsilon(t)\left(f\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - f\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right)\right) \\
&\quad + \overline{K}^\varepsilon(t)\left(\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - \sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right)\right) \\
&\quad + \int_\Theta \overline{\gamma}^\varepsilon(t) g\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right)\mu(d\theta) \\
&\quad + \tfrac{1}{2}\int_\Theta (g^*\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g^*\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right))\left(\overline{Q}^\varepsilon(t) + \overline{\gamma}_t^\varepsilon(\theta)\right) \\
&\quad \times (g\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right))\mu(d\theta), \\
&\quad \left. + \left(\ell\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - \ell\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right)\right)\right\} dt + \tau(\hbar),
\end{aligned} \tag{2.40}$$

where $\tau(\hbar) \longrightarrow 0$ as $\hbar \longrightarrow 0$. Finally, replacing (2.40) in (2.39), then dividing inequality (2.39) by $\hbar$ and sending $\hbar$ to zero, the near-maximum condition (2.37) follows. □

Proof of Theorem 2.3.1. First, we are about to derive an estimate for the term similar to the left side of inequality (2.34) and (2.35) with all the $(\overline{x}^{\varepsilon}(\cdot),\mathbb{E}(\overline{x}^{\varepsilon}(\cdot)),\overline{u}^{\varepsilon}(\cdot))$ etc. replaced by $(x^{\varepsilon}(\cdot),\mathbb{E}(x^{\varepsilon}(\cdot)),\ u^{\varepsilon}(\cdot))$ etc.

Now, to prove (2.14) it remains to estimate the following differences

$$\begin{aligned} S_1(\varepsilon) = \mathbb{E}\int_s^T &\left[\overline{K}^{\varepsilon}(t)\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)\right. \\ &\left.-\ K^{\varepsilon}(t)\left(\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u\right)-\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)\right)\right]dt, \end{aligned} \tag{2.41}$$

$$\begin{aligned} S_2(\varepsilon) = \mathbb{E}\int_s^T &\left\{\tfrac{1}{2}\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right)^*\overline{Q}^{\varepsilon}(t)\right. \\ &\times\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right) \\ &-\tfrac{1}{2}\left(\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u\right)-\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)\right)^*Q^{\varepsilon}(t) \\ &\times\left(\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u\right)-\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)\right) \\ &+\overline{\Psi}^{\varepsilon}(t)\left[f\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-f\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right] \\ &-\Psi^{\varepsilon}(t)\left[f\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u\right)-f\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)\right] \\ &+\left[\ell\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\ell\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right] \\ &\left.-\left[\ell\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u\right)-\ell\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)\right]\right\}dt. \end{aligned} \tag{2.42}$$

and

$$\begin{aligned} S_3(\varepsilon) = \mathbb{E}\int_s^T\int_\Theta &\left[\overline{\gamma}_t^{\varepsilon}(\theta)\left(g\left(t,\overline{x}^{\varepsilon}(t_-),u,\theta\right)-g\left(t,\overline{x}^{\varepsilon}(t_-),\overline{u}^{\varepsilon}(t)\right)\right)\right. \\ &\left.-\ \gamma_t^{\varepsilon}(\theta)\left(g\left(t,x^{\varepsilon}(t_-),u,\theta\right)-g\left(t,x^{\varepsilon}(t_-),u^{\varepsilon}(t),\theta\right)\right)\right]\mu(d\theta)dt, \end{aligned} \tag{2.43}$$

Then we have

$$\begin{aligned} S_1(\varepsilon) &= \mathbb{E}\int_s^T\left[\overline{K}^{\varepsilon}(t)-K^{\varepsilon}(t)\right]\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)\right) \\ &\quad+\mathbb{E}\int_s^T K^{\varepsilon}(t)\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),u\right)-\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u\right)\right)dt \\ &\quad-\mathbb{E}\int_s^T K^{\varepsilon}(t)\left(\sigma\left(t,\overline{x}^{\varepsilon}(t),\mathbb{E}(\overline{x}^{\varepsilon}(t)),\overline{u}^{\varepsilon}(t)\right)-\sigma\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)\right)dt \\ &\quad=\mathbb{I}_1\left(\varepsilon\right)+\mathbb{I}_2\left(\varepsilon\right)+\mathbb{I}_3\left(\varepsilon\right), \end{aligned}$$

We estimate the first term on the right-hand side $\mathbb{I}_1(\varepsilon)$. For any $\delta \in [0, \frac{1}{3})$ so that $\alpha = 3\delta \in [0, 1)$. Now, let β be a fixed real number such that $1 < \beta < 2$ so that $(1+\alpha)\beta < 2$. Taking $q > 2$ such that $\frac{1}{\beta} + \frac{1}{q} = 1$ then by using Hôlder's inequality, *Lemma 2.3.2* and note (2.4) we obtain

$$\begin{aligned}
\mathbb{I}_1(\varepsilon) &\leq \left[\mathbb{E}\int_s^T \left|\overline{K}^\varepsilon(t) - K^\varepsilon(t)\right|^\beta dt\right]^{\frac{1}{\beta}} \\
&\quad \times \left[\mathbb{E}\int_s^T \left|\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - \sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right)\right|^q dt\right]^{\frac{1}{q}} \\
&\leq C\left[d(\overline{u}^\varepsilon(\cdot), u^\varepsilon(\cdot))^{\frac{\alpha\beta}{2}}\right]^{\frac{1}{\beta}}\left[\mathbb{E}\int_s^T (1 + |\overline{x}^\varepsilon(t)|^q + |\mathbb{E}(\overline{x}^\varepsilon(t))|^q)dt\right]^{\frac{1}{q}} \leq C\left[\varepsilon^{\frac{2}{3}}\right]^{\frac{\alpha\beta}{2}\cdot\frac{1}{\beta}} = C\varepsilon^\delta.
\end{aligned}$$

We estimate now the second term $\mathbb{I}_2(\varepsilon)$. Then by applying Cauchy-Schwarz inequality, note (2.9), assumption (**H1**), and *Lemma 2.3.1*, we get

$$\begin{aligned}
\mathbb{I}_2(\varepsilon) &\leq \left[\mathbb{E}\int_s^T |K^\varepsilon(t)|^2 dt\right]^{\frac{1}{2}}\left[\mathbb{E}\int_s^T \left|\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u\right) - \sigma\left(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u\right)\right|^2 dt\right]^{\frac{1}{2}} \\
&\leq C\left[\mathbb{E}\int_s^T (|\overline{x}^\varepsilon(t) - x^\varepsilon(t)|^2 + |\mathbb{E}(\overline{x}^\varepsilon(t)) - \mathbb{E}(x^\varepsilon(t))|^2)dt\right]^{\frac{1}{2}} \\
&\leq C\left[d(\overline{u}^\varepsilon(\cdot), u^\varepsilon(\cdot))^\alpha\right]^{\frac{1}{2}} \leq C\left[\varepsilon^{\frac{2}{3}}\right]^{\alpha\cdot\frac{1}{2}} = C\varepsilon^{\frac{\alpha}{3}} = C\varepsilon^\delta.
\end{aligned}$$

Now, let us turn to estimate the third term $\mathbb{I}_3(\varepsilon)$. By adding and subtracting $\sigma(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u^\varepsilon(t))$ then we have

$$\begin{aligned}
\mathbb{I}_3(\varepsilon) = &-\mathbb{E}\int_s^T K^\varepsilon(t)[\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), \overline{u}^\varepsilon(t)\right) - \sigma(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u^\varepsilon(t))]dt \\
&- \mathbb{E}\int_s^T K^\varepsilon(t)\sigma\left(t, \overline{x}^\varepsilon(t), \mathbb{E}(\overline{x}^\varepsilon(t)), u^\varepsilon(t)\right) - \sigma\left(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t)\right))dt,
\end{aligned}$$

then by using Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\mathbb{I}_3(\varepsilon) &\leq \left[\mathbb{E}\int_s^T |K^\varepsilon(t)|^2 \, dt\right]^{\frac{1}{2}} \\
&\times \left[\mathbb{E}\int_s^T |\sigma(t,\overline{x}^\varepsilon(t),\mathbb{E}(\overline{x}^\varepsilon(t)),\overline{u}^\varepsilon(t)) - \sigma(t,\overline{x}^\varepsilon(t),\mathbb{E}(\overline{x}^\varepsilon(t)),u^\varepsilon(t))|^2\right. \\
&\times \left.\mathbf{I}_{\{\overline{u}^\varepsilon(\cdot)\neq u^\varepsilon(\cdot)\}}(t)\, dt\right]^{\frac{1}{2}} \\
&+ \mathbb{E}\int_s^T |K^\varepsilon(t)|\, |[\sigma(t,\overline{x}^\varepsilon(t),\mathbb{E}(\overline{x}^\varepsilon(t)),u^\varepsilon(t)) - \sigma(t,x^\varepsilon(t),\mathbb{E}(x^\varepsilon(t)),u^\varepsilon(t))]|\, dt,
\end{aligned}$$

We proceed as in $I_2(\varepsilon)$ to estimate the second term in the right of above inequality, then by applying Cauchy-Schwartz inequality, Assumption (**H1**) and (2.9) we obtain

$$\begin{aligned}
\mathbb{I}_3(\varepsilon) &\leq \left[\mathbb{E}\int_s^T |K^\varepsilon(t)|^2 \, dt\right]^{\frac{1}{2}} \\
&\times \left\{\left[\mathbb{E}\int_s^T |\sigma(t,\overline{x}^\varepsilon(t),\mathbb{E}(\overline{x}^\varepsilon(t)),\overline{u}^\varepsilon(t)) - \sigma(t,\overline{x}^\varepsilon(t),\mathbb{E}(\overline{x}^\varepsilon(t)),u^\varepsilon(t))|^4 \, dt\right]^{\frac{1}{2}}\right. \\
&\times \left.\left[\mathbb{E}\int_s^T \mathbf{I}_{\{\overline{u}^\varepsilon(\cdot)\neq u^\varepsilon(\cdot)\}}(t)\, dt\right]^{\frac{1}{2}}\right\}^{\frac{1}{2}} + C\varepsilon^\delta, \\
&\leq C\left[d(\overline{u}^\varepsilon(\cdot),u^\varepsilon(\cdot))^{\frac{1}{2}}\right]^{\frac{1}{2}} + C\varepsilon^\delta \\
&\leq C\varepsilon^\delta,
\end{aligned}$$

thus, we have proved that

$$S_1(\varepsilon) = \mathbb{I}_1(\varepsilon) + \mathbb{I}_2(\varepsilon) + \mathbb{I}_3(\varepsilon) \leq C\varepsilon^\delta. \tag{2.44}$$

By using similar arguments developed above, we can prove that

$$S_2(\varepsilon) \leq C\varepsilon^\delta. \tag{2.45}$$

Now, let us turn to estimate the third term $S_3(\varepsilon)$. By applying the Cauchy-Schwarz inequality, we get

$$S_3(\varepsilon) \leq \mathbb{E}\int_s^T \int_\Theta \left(\overline{\gamma}_t^\varepsilon(\theta) - \gamma_t^\varepsilon(\theta)\right)\left(g\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right)\right)\mu(d\theta)dt$$
$$+\mathbb{E}\int_s^T \int_\Theta \left[\gamma_t^\varepsilon(\theta)\left(g\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g\left(t, x^\varepsilon(t_-), u\right)\right)\mu(d\theta)dt,\right.$$
$$+\mathbb{E}\int_s^T \int_\Theta \gamma_t^\varepsilon(\theta)\left(g\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right) - g\left(t, x^\varepsilon(t_-), u^\varepsilon(t), \theta\right)\right)\mu(d\theta)dt,$$
$$= \mathbb{J}_1(\varepsilon) + \mathbb{J}_2(\varepsilon) + \mathbb{J}_3(\varepsilon).$$

For any $\delta \in [0, \frac{1}{3})$ so that $\alpha = 3\delta \in [0, 1)$. Now, let β be a fixed real number such that $\beta \in (1, 2)$ so that $(1 + \alpha)\beta < 2$. Taking $q > 2$ such that $\frac{1}{\beta} + \frac{1}{q} = 1$ then by using Hôlder's inequality, *Lemma 2.3.2* and note (2.5) we obtain

$$\mathbb{J}_1(\varepsilon) \leq \left[\mathbb{E}\int_s^T \int_\Theta \left|\overline{\gamma}_t^\varepsilon(\theta) - \gamma_t^\varepsilon(\theta\right|^\beta \mu(d\theta)dt\right]^{\frac{1}{\beta}}$$
$$\times \mathbb{E}\left\{\int_s^T (\sup_{\theta\in\Theta} \left|g\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right)\right|)^q dt\right\}^{\frac{1}{q}} \times \mu(\Theta)^{\frac{1}{q}}$$
$$\leq C\left[d(\overline{u}^\varepsilon(\cdot), u^\varepsilon(\cdot))^{\frac{\alpha\beta}{2}}\right]^{\frac{1}{\beta}} \left[\mathbb{E}\int_s^T (1 + |\overline{x}^\varepsilon(t)|^q + |\mathbb{E}(\overline{x}^\varepsilon(t))|^q)dt\right]^{\frac{1}{q}}$$
$$\leq C\left[\varepsilon^{\frac{2}{3}}\right]^{\frac{\alpha\beta}{2}\cdot\frac{1}{\beta}} = C\varepsilon^{\frac{\alpha}{3}}.$$

Applying assumption **(H3),** Cauchy-Schwarz inequality, *Lemma 2.3.2*, note (2.10) and the fact that $\mu(\Theta) < \infty$ we get

$$\mathbb{J}_2(\varepsilon) \leq \left[\mathbb{E}\int_s^T \int_\Theta \left|\gamma_t^\varepsilon(\theta\right|^2 \mu(d\theta)dt\right]^{\frac{1}{2}} [\mu(\Theta)]^{\frac{1}{2}}$$
$$\times \mathbb{E}\left\{\int_s^T (\sup_{\theta\in\Theta} \left|g\left(t, \overline{x}^\varepsilon(t_-), u, \theta\right) - g\left(t, \overline{x}^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta\right)\right|)^2 dt\right\}^{\frac{1}{2}}$$
$$\leq C\mathbb{E}\left\{\int_s^T \left|\overline{x}^\varepsilon(t) - x^\varepsilon(t)\right|^2 dt\right\}^{\frac{1}{2}}$$
$$\leq C\left[d(\overline{u}^\varepsilon(\cdot), u^\varepsilon(\cdot))^\alpha\right]^{\frac{1}{2}}$$

by using (2.38) we get $d(\overline{u}^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot))^{\alpha} \leq \left(\varepsilon^{\frac{2}{3}}\right)^{\alpha}$, it holds that

$$\mathbb{J}_2(\varepsilon) \leq C\left[\varepsilon^{\frac{2\alpha}{3}}\right]^{\frac{1}{2}} = C\varepsilon^{\frac{\alpha}{3}} = C\varepsilon^{\delta}.$$

We proceed to estimate $\mathbb{J}_3(\varepsilon)$. By adding and subtracting $g(t, \overline{x}^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta)$ and Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
\mathbb{J}_3(\varepsilon) = {} & \mathbb{E}\int_s^T\int_{\Theta}\gamma_t^{\varepsilon}(\theta)(g(t, \overline{x}^{\varepsilon}(t_-), \overline{u}^{\varepsilon}(t), \theta) - g(t, \overline{x}^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta)\,\mathbf{I}_{\{\overline{u}^{\varepsilon}(\cdot)\neq u^{\varepsilon}(\cdot)\}}(t)\,\mu(d\theta)dt \\
& + \mathbb{E}\int_s^T\int_{\Theta}\gamma_t^{\varepsilon}(\theta)(g(t, \overline{x}^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta) - g(t, x^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta))\mu(d\theta)dt \\
& \leq \mathbb{E}\left\{\int_s^T\int_{\Theta}\left|\gamma_t^{\varepsilon}(\theta)\right|^2\mu(d\theta)dt\right\}^{\frac{1}{2}} \times [\mu(\Theta)]^{\frac{1}{2}} \\
& \times \mathbb{E}\left\{\int_s^T(\sup_{\theta\in\Theta}|g(t, \overline{x}^{\varepsilon}(t_-), \overline{u}^{\varepsilon}(t), \theta) - g(t, \overline{x}^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta)|)^2\mathbf{I}_{\{\overline{u}^{\varepsilon}(\cdot)\neq u^{\varepsilon}(\cdot)\}}(t)\,dt\right\}^{\frac{1}{2}} \\
& + \mathbb{E}\left\{\int_s^T\int_{\Theta}\left|\gamma_t^{\varepsilon}(\theta)\right|^2\mu(d\theta)dt\right\}^{\frac{1}{2}} \times \mathbb{E}\left\{\int_s^T\left|\overline{x}^{\varepsilon}(t) - x^{\varepsilon}(t)\right|^2dt\right\}^{\frac{1}{2}},
\end{aligned}$$

by applying Cauchy-Schwarz inequality, *Lemma 2.3.2* and (2.11) it follows that

$$\begin{aligned}
\mathbb{J}_3(\varepsilon) \leq {} & \mathbb{E}\left\{\int_s^T(1 + |\overline{x}^{\varepsilon}(t)|^4)dt\right\}^{\frac{1}{2}} d(\overline{u}^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot))^{\frac{1}{2}} \\
& + C\mathbb{E}\left\{\int_s^T\left|\overline{x}^{\varepsilon}(t) - x^{\varepsilon}(t)\right|^2dt\right\}^{\frac{1}{2}} \leq C\varepsilon^{\delta}.
\end{aligned}$$

Thus, we have proved that

$$S_3(\varepsilon) = \mathbb{J}_1(\varepsilon) + \mathbb{J}_2(\varepsilon) + \mathbb{J}_3(\varepsilon) \leq C\varepsilon^{\delta}. \tag{2.46}$$

The desired result (2.14) follows immediately from combining (2.44), (2.45), (2.46) and (2.34). This completes the proof of *Theorem 2.3.1.* $\square$

Proof of Corollary 2.3.1. In the spike variations technique for the perturbed control

$\overline{u}^{\varepsilon,\theta}(\cdot)$ in (2.37) the point $u \in \mathbb{A}$ may be replaced by any admissible control $u(\cdot) \in \mathcal{U}$, and the subsequent argument still goes through. So the inequality in the estimate (2.15) holds for any $u(\cdot) \in \mathcal{U}$ and the subsequent argument still goes through. So the inequalities in the estimate (2.15) holds for any $u(\cdot) \in \mathcal{U}$. $\square$

2.4 Sufficient conditions of near-optimality for mean-field jump diffusion processes

We will shows in this section, that under certain concavity conditions on the Hamiltonian H and some convexity conditions on the function $h(\cdot,\cdot)$, the ε-maximum condition on the Hamiltonian function $\mathcal{H}$ in the integral form is sufficient for near-optimality. We assume:

Assumption (H3) ψ is differentiable in u for $\psi =: f, \sigma, \ell, g$ and there is a constant $C > 0$ such that

$$\begin{aligned} &|\psi(t,x,y,u) - \psi(t,x,y,u')| + |\psi_u(t,x,y,u) - \psi_u(t,x,y,u')| \\ &\qquad \leq C\,|u-u'|\,, \\ &\sup\nolimits_{\theta\in\Theta} |g(t,x,u,\theta) - g(t,x,u',\theta)| + \sup\nolimits_{\theta\in\Theta} |g_u(t,x,u,\theta) - g_u(t,x,u',\theta)| \\ &\qquad \leq C\,|u-u'|\,, \end{aligned} \tag{2.47}$$

$$h(x,y) - h(x',y') \geq (h_x(x',y') + h_y(x',y'))(x-x'), \tag{2.48}$$

and

$$\begin{aligned} &H(t,x,\mathbb{E}(x),u,\Psi,K,R) - H(t,x',\mathbb{E}(x'),u',\Psi,K,R) \\ &\leq (H_x(t,x',\mathbb{E}(x'),u',\Psi,K,R) + H_y(t,x',\mathbb{E}(x'),u',\Psi,K,R))\,(x-x') \\ &+H_u(t,x',\mathbb{E}(x'),u',\Psi,K,R)(u-u'), \quad a.e.,\ t\in[s,T]\,,\ \mathbb{P}-a.s. \end{aligned} \tag{2.49}$$

Now we are able to state and prove the sufficient conditions for near-optimality for systems governed by mean-field SDEs with jump processes, which is the second main result of this work.

Let $u^{\varepsilon}(\cdot)$ be an admissible control and $\left(\Psi^{\varepsilon}(\cdot), K^{\varepsilon}(\cdot), \gamma^{\varepsilon}(\cdot)\right)$, $\left(Q^{\varepsilon}(\cdot), R^{\varepsilon}(\cdot), \Gamma^{\varepsilon}(\cdot)\right)$ be the solution of the adjoint equations (2.9)-(2.10) corresponding to $u^{\varepsilon}(\cdot)$.

Theorem 2.4.1. (Sufficient conditions for near-optimality of order $\varepsilon^{\frac{1}{2}}$). *Let conditions (2.47)∼(2.49) holds. If for some $\varepsilon > 0$ and for any $u(\cdot) \in \mathcal{U}$:*

$$\begin{aligned} &\mathbb{E}\int_s^T \mathcal{H}^{(x^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t))dt + \varepsilon \\ &\geq \sup_{u(\cdot)\in\mathcal{U}} \mathbb{E}\int_s^T \mathcal{H}^{(x^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u(t))dt, \end{aligned} \tag{2.50}$$

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

$$J^{s,\zeta}\left(u^{\varepsilon}(\cdot)\right) \leq \inf_{u(\cdot)\in\mathcal{U}} J^{s,\zeta}\left(u(\cdot)\right) + C\varepsilon^{\frac{1}{2}},$$

where $C > 0$ is a positive constant independent of ε.

Corollary 2.4.1. (Sufficient Conditions for $\varepsilon-$optimality) Under the assumptions of *Theorem 2.4.1* a sufficient condition for an admissible control $u^{\varepsilon}(\cdot)$ to be $\varepsilon-$optimal for our mean-field control problem (2.1)-(2.2) is

$$\begin{aligned} &\mathbb{E}\left\{\int_s^T \mathcal{H}^{(x^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t))dt\right\} + \left(\frac{\varepsilon}{C}\right)^2 \\ &\geq \sup_{u(\cdot)\in\mathcal{U}} \mathbb{E}\left\{\int_s^T \mathcal{H}^{(x^{\varepsilon}(.),u^{\varepsilon}(.))}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u(t))dt\right\}. \end{aligned}$$

Proof of Theorem 2.4.1. The key step in the proof is to show that $H_u(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t), \Psi^{\varepsilon}(t), K^{\varepsilon}(t), \gamma_t^{\varepsilon}(\theta))$ is very small and estimate it in terms of ε. We first fix an $\varepsilon > 0$ and define a new metric $\widehat{d}$ on $\mathcal{U}$, by setting: for any $u(\cdot)$ and $v(\cdot) \in \mathcal{U}$:

$$\widehat{d}(u(\cdot), v(\cdot)) = \mathbb{E}\int_s^T \left|u(t) - v(t)\right| \pounds^{\varepsilon}(t)dt,$$

where

$$\mathcal{L}^{\varepsilon}(t) = 1 + |\Psi^{\varepsilon}(t)| + |K^{\varepsilon}(t)| + 2\,|Q^{\varepsilon}(t)|\,[1 + |x^{\varepsilon}(t)| + |\mathbb{E}(x^{\varepsilon}(t))|]$$
$$+ 2\left[|Q^{\varepsilon}(t)| + \left|\int_{\Theta} \gamma_t^{\varepsilon}(\theta)\mu(d\theta)\right|\right][1 + |x^{\varepsilon}(t)|]$$

Obviously $\widehat{d}$ is a metric on $\mathcal{U}$ satisfied $\mathcal{L}^{\varepsilon}(t) > 1$, and it is a complete metric as a weighted $\mathbb{L}^1$-norm.

Define a functional g on $\mathcal{U}$ as follows

$$g\,(u(\cdot)) = \mathbb{E}\int_s^T \mathcal{H}^{(x^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))}\,(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u(t))\,dt.$$

By using assumption (**??**) then a simple computation shows that

$$|g\,(u(\cdot)) - g\,(v(\cdot))| = \mathbb{E}\int_s^T \left\{\mathcal{H}^{(x^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))}\,(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u(t))\right.$$
$$\left.-\mathcal{H}^{(x^{\varepsilon}(\cdot),u^{\varepsilon}(\cdot))}\,(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), v(t))\right\}dt.$$
$$\leq \mathbb{E}\int_s^T |H\,(t, x^{\varepsilon}(t), \mathbb{E}\,(x^{\varepsilon}(t))\,, u(t), \Psi^{\varepsilon}(t), K^{\varepsilon}(t), \gamma_t^{\varepsilon}(\theta))$$
$$-\; H\,(t, x^{\varepsilon}(t), \mathbb{E}\,(x^{\varepsilon}(t))\,, v(t), \Psi^{\varepsilon}(t), K^{\varepsilon}(t), \gamma_t^{\varepsilon}(\theta))|\,dt$$
$$+\,\mathbb{E}\int_s^T |\sigma^*\,(t, x^{\varepsilon}(t), \mathbb{E}\,(x^{\varepsilon}(t))\,, u(t)) - \sigma^*\,(t, x^{\varepsilon}(t), \mathbb{E}\,(x^{\varepsilon}(t))\,, v(t))|$$
$$\times\,|Q^{\varepsilon}(t)|\,|\sigma\,(t, x^{\varepsilon}(t), \mathbb{E}\,(x^{\varepsilon}(t))\,, u^{\varepsilon}(t))|\,dt$$

$$
\begin{aligned}
&+\frac{1}{2}\mathbb{E}\int_s^T \left|\sigma^*\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u(t)\right)Q(t)\sigma\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u(t)\right)\right.\\
&-\left.\sigma^*\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),v(t)\right)Q^\varepsilon(t)\sigma\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),v(t)\right)\right|dt\\
&+\mathbb{E}\int_s^T\int_\Theta \left|g^*\left(t,x^\varepsilon(t),u(t),\theta\right)-g^*\left(t,x^\varepsilon(t),v(t),\theta\right)\right|\\
&\times\left|\left(Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right)g\left(t,x^\varepsilon(t_-),u^\varepsilon(t),\theta\right)\right|\mu(d\theta)dt\\
&+\frac{1}{2}\mathbb{E}\int_s^T\int_\Theta \left|g^*\left(t,x^\varepsilon(t),u(t),\theta\right)\left(Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right)g\left(t,x^\varepsilon(t),u(t),\theta\right)\right.\\
&-\left.g^*\left(t,x^\varepsilon(t),v(t),\theta\right)\left(Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right)g\left(t,x^\varepsilon(t),v(t),\theta\right)\right|\mu(d\theta)dt,\\
&=\mathcal{I}_1^\varepsilon+\mathcal{I}_2^\varepsilon+\mathcal{I}_3^\varepsilon+\mathcal{I}_4^\varepsilon+\mathcal{I}_5^\varepsilon
\end{aligned}
$$

Now, by using *definition 2.2.2* and assumption **(H3)**

$$
\begin{aligned}
\mathcal{I}_1^\varepsilon &= \mathbb{E}\int_s^T \left|H\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u,\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta)\right)\right.\\
&\quad-\left.H\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),v,\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta)\right)\right|dt\\
&\le C\mathbb{E}\int_s^T |u(t)-v(t)|\left(|\Psi^\varepsilon(t)|+|K^\varepsilon(t)|+\left|\int_\Theta \gamma_t^\varepsilon(\theta)\mu(d\theta)\right|\right)dt\\
&\le C\mathbb{E}\int_s^T |u(t)-v(t)|\,\pounds^\varepsilon(t)dt
\end{aligned}
\tag{2.51}
$$

Since σ is linear growth with respect to x and y then by using assumption (2.47) we get

$$
\begin{aligned}
\mathcal{I}_2^\varepsilon &= \mathbb{E}\int_s^T \left|\sigma^*\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u\right)-\sigma^*\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),v\right)\right|\\
&\quad\times\left|Q^\varepsilon(t)\sigma\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u^\varepsilon(t)\right)\right|dt\\
&\le C\mathbb{E}\int_s^T |u(t)-v(t)|\,|Q^\varepsilon(t)|\left[1+|x^\varepsilon(t)|+\left|\mathbb{E}\left(x^\varepsilon(t)\right)\right|\right]dt\\
&\le C\mathbb{E}\int_s^T |u(t)-v(t)|\,\pounds^\varepsilon(t)dt.
\end{aligned}
\tag{2.52}
$$

Similarly, since g is linear growth with respect to x then by assumptions (2.47) we can prove that

$$\begin{aligned}\mathcal{I}_4^{\varepsilon} &= \mathbb{E}\int_s^T\int_\Theta \left|g^*\left(t,x^\varepsilon(t),u,\theta\right)-g^*\left(t,x^\varepsilon(t),v,\theta\right)\right| \\ &\quad\times\left|\left(Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right)g\left(t,x^\varepsilon(t_-),u^\varepsilon(t),\theta\right)\right|\mu(d\theta)dt \\ &\leq C\mathbb{E}\int_s^T\left|u(t)-v(t)\right|\left[\left|Q^\varepsilon(t)\right|+\left|\int_\Theta\gamma_t^\varepsilon(\theta)\mu(d\theta)\right|\right]\left[1+\left|x^\varepsilon(t)\right|\right]dt \\ &\leq C\mathbb{E}\int_s^T\left|u(t)-v(t)\right|\mathcal{L}^\varepsilon(t)dt.\end{aligned} \tag{2.53}$$

Next, since σ is linear growth with respect to x and y then we deduce that

$$\begin{aligned}\mathcal{I}_3^{\varepsilon} &= \tfrac{1}{2}\mathbb{E}\int_s^T\left|\sigma^*\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u\right)Q^\varepsilon(t)\sigma\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),u\right)\right. \\ &\quad\left.-\,\sigma^*\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),v\right)Q^\varepsilon(t)\sigma\left(t,x^\varepsilon(t),\mathbb{E}\left(x^\varepsilon(t)\right),v\right)\right|dt \\ &\leq C\mathbb{E}\int_s^T\left|u(t)-v(t)\right|\tfrac{1}{2}\left|Q^\varepsilon(t)\right|\left[1+\left|x^\varepsilon(t)\right|+\left|\mathbb{E}\left(x^\varepsilon(t)\right)\right|\right]dt \\ &\leq C\mathbb{E}\int_s^T\left|u(t)-v(t)\right|\mathcal{L}^\varepsilon(t)dt,\end{aligned} \tag{2.54}$$

and

$$\begin{aligned}\mathcal{I}_5^{\varepsilon} &= +\tfrac{1}{2}\mathbb{E}\int_s^T\int_\Theta\left|g^*\left(t,x^\varepsilon(t),u,\theta\right)\left(Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right)g\left(t,x^\varepsilon(t),u,\theta\right)\right. \\ &\quad\left.-\,g^*\left(t,x^\varepsilon(t),v,\theta\right)\left(Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right)g\left(t,x^\varepsilon(t),v,\theta\right)\right|\mu(d\theta)dt, \\ &\leq C\mathbb{E}\int_s^T\left|u(t)-v(t)\right|\tfrac{1}{2}\left|Q^\varepsilon(t)+\gamma_t^\varepsilon(\theta)\right|\left[1+\left|x^\varepsilon(t)\right|\right]dt \\ &\leq C\mathbb{E}\int_s^T\left|u(t)-v(t)\right|\mathcal{L}^\varepsilon(t)dt,\end{aligned} \tag{2.55}$$

By combining (2.51)$\sim$(2.55) we conclude that

$$\left|g\left(u(\cdot)\right)-g\left(v(\cdot)\right)\right|\leq C\widehat{d}\left(u(\cdot),v(\cdot)\right),$$

which implies that g is continuous on $\mathcal{U}$ with respect to $\widehat{d}$. Now by using (2.50) and Ekeland's Variational Principle (*Lemma 2.2.1*), there exists $\overline{u}^\varepsilon(\cdot)\in\mathcal{U}$ such that

$$\widehat{d}(\overline{u}^\varepsilon(\cdot),u^\varepsilon(\cdot))\leq\sqrt{\varepsilon}, \tag{2.56}$$

and

$$\mathbb{E}\int_s^T \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t))dt = \max_{u(\cdot)\in\mathcal{U}} \mathbb{E}\int_s^T \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t))dt, \tag{2.57}$$

where

$$\widetilde{\mathcal{H}}(t, x, y, u) = \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x, y, u) - \sqrt{\varepsilon}\left|u - \overline{u}^\varepsilon(t)\right| \mathcal{L}^\varepsilon(t). \tag{2.58}$$

The maximum condition (2.57) implies a pointwise maximum condition namely, for $\mathbb{P}-a.s$, and $a.e.$, $t \in [s, T]$

$$\widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t)) = \max_{u\in\mathbb{A}} \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u).$$

Using [Item 3, Proposition A1], then we have

$$0 \in \partial_u^\circ \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t)). \tag{2.59}$$

Since the function $u :\longmapsto |u - \overline{u}^\varepsilon(t)|$ is locally Lipschitz but not differentiable in $\overline{u}^\varepsilon(t)$, then Clarke's generalized gradient (see Proposition A1, Example, Appendix) shows that

$$\begin{aligned}\partial_u^\circ\left(\sqrt{\varepsilon}\left|u - \overline{u}^\varepsilon(t)\right| \mathcal{L}^\varepsilon(t)\right) &= \overline{co}\left\{-\mathcal{L}^\varepsilon(t)\sqrt{\varepsilon},\ \mathcal{L}^\varepsilon(t)\sqrt{\varepsilon}\right\} \\ &= \left[-\mathcal{L}^\varepsilon(t)\sqrt{\varepsilon},\ \mathcal{L}^\varepsilon(t)\sqrt{\varepsilon}\right].\end{aligned} \tag{2.60}$$

By using (2.60) and fact that the Clarke's generalized gradient of the sum of two functions is contained in the sum of the Clarke's generalized gradient of the two functions, ([Item 5, Proposition A1] we get

$$\begin{aligned}\partial_u^\circ \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t)) \subset\ & \partial_u^\circ \mathcal{H}^{(x^\varepsilon(.), u^\varepsilon(.))}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t)) \\ &+ \left[-\sqrt{\varepsilon}\mathcal{L}^\varepsilon(t),\ \sqrt{\varepsilon}\mathcal{L}^\varepsilon(t)\right].\end{aligned}$$

By applying assumption (2.47), the Hamiltonian H is differentiable in u, then [Item 4, Proposition A1] shows that

$$
\begin{aligned}
&\partial_u^{\circ}\widetilde{\mathcal{H}}(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t)) \\
&\qquad \subset \{H_u(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t), \Psi^{\varepsilon}(t), K^{\varepsilon}(t), \gamma_t^{\varepsilon}(\theta)) \\
&\qquad + \{\sigma_u^*(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t)) Q^{\varepsilon}(t) \\
&\qquad \times (\sigma(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t))) - \sigma(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t)))\} \\
&\qquad + \int_{\Theta} g_u^*(t, x^{\varepsilon}(t_-), \overline{u}^{\varepsilon}(t), \theta)(Q^{\varepsilon}(t) + \gamma_t^{\varepsilon}(\theta)) \\
&\qquad \times (g(t, x^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta) - g(t, x^{\varepsilon}(t_-), \overline{u}^{\varepsilon}(t), \theta))\mu(d\theta)\} \\
&\qquad + [-\sqrt{\varepsilon}\pounds^{\varepsilon}(t), \sqrt{\varepsilon}\pounds^{\varepsilon}(t)].
\end{aligned}
$$

Next, the differential inclusion (2.59) implies that there is

$$
\tau^{\varepsilon}(t) \in \left[-\sqrt{\varepsilon}\pounds^{\varepsilon}(t),\ \sqrt{\varepsilon}\pounds^{\varepsilon}(t)\right],
$$

such that

$$
\begin{aligned}
&H_u(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t), \Psi^{\varepsilon}(t), K^{\varepsilon}(t), \gamma_t^{\varepsilon}(\theta)) \\
&+\sigma_u^*(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t)) Q^{\varepsilon}(t) \\
&\times(\sigma(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), u^{\varepsilon}(t))) - \sigma(t, x^{\varepsilon}(t), \mathbb{E}(x^{\varepsilon}(t)), \overline{u}^{\varepsilon}(t))) \\
&+ \int_{\Theta} g_u^*(t, x^{\varepsilon}(t_-), \overline{u}^{\varepsilon}(t), \theta)(Q^{\varepsilon}(t) + \gamma_t^{\varepsilon}(\theta)) \\
&\times(g(t, x^{\varepsilon}(t_-), u^{\varepsilon}(t), \theta) - g(t, x^{\varepsilon}(t_-), \overline{u}^{\varepsilon}(t), \theta))\mu(d\theta)\} + \tau^{\varepsilon}(t) = 0.
\end{aligned} \tag{2.61}
$$

By using assumption (2.47) we can prove that

$$
\begin{aligned}
&|H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\
&-H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))| \\
&\leq C\,|u^\varepsilon(t) - \overline{u}^\varepsilon(t)|\, \pounds^\varepsilon(t),
\end{aligned} \tag{2.62}
$$

hence from (2.61) and (2.62), assumption (2.47) and the fact that $|\tau^\varepsilon(t)| \leq \sqrt{\varepsilon}\pounds^\varepsilon(t)$ we get

$$
\begin{aligned}
&|H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))| \\
&\leq C\,|u^\varepsilon(t) - \overline{u}^\varepsilon(t)|\, \pounds^\varepsilon(t) + |\sigma_u^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t)) Q^\varepsilon(t) \\
&\times\, (\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))) - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \overline{u}^\varepsilon(t)))| \\
&\quad + \left|\int_\Theta g_u^*\,(t, x^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta)\,(Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta))\right. \\
&\quad \times(g\,(t, x^\varepsilon(t_-), u^\varepsilon(t), \theta) - g\,(t, x^\varepsilon(t_-), \overline{u}^\varepsilon(t), \theta))\mu(d\theta)| + |\tau^\varepsilon(t)| \\
&\leq C\,|u^\varepsilon(t) - \overline{u}^\varepsilon(t)|\, \pounds^\varepsilon(t) + |\tau^\varepsilon(t)| \\
&\leq C\,|u^\varepsilon(t) - \overline{u}^\varepsilon(t)|\, \pounds^\varepsilon(t) + \sqrt{\varepsilon}\pounds^\varepsilon(t).
\end{aligned} \tag{2.63}
$$

Now, using (2.49), we obtain for any $u(\cdot) \in \mathcal{U}$

$$
\begin{aligned}
&H(t, x(t), \mathbb{E}(x(t)), u(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\
&-H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\
&\leq H_x(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))(x(t) - x^\varepsilon(t)) \\
&+ H_y(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))(x(t) - x^\varepsilon(t)) \\
&+ H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))(u(t) - u^\varepsilon(t)).
\end{aligned} \tag{2.64}
$$

Integrating this inequality with respect to t and taking expectations we obtain from (2.51) and (2.63)

$$\begin{aligned}
&\mathbb{E}\int_s^T [H(t,x(t),\mathbb{E}(x(t)),u(t),\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta)) \\
&-H(t,x^\varepsilon(t),\mathbb{E}(x^\varepsilon(t)),u^\varepsilon(t),\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta))]\,dt \\
&\leq \mathbb{E}\int_s^T H_x(t,x^\varepsilon(t),\mathbb{E}(x^\varepsilon(t)),u^\varepsilon(t),\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta))(x(t)-x^\varepsilon(t))dt \\
&+\mathbb{E}\int_s^T H_y(t,x^\varepsilon(t),\mathbb{E}(x^\varepsilon(t)),u^\varepsilon(t),\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta))(x(t)-x^\varepsilon(t))dt \\
&+C(\widehat{d}(u^\varepsilon(\cdot),\overline{u}^\varepsilon(\cdot))+\varepsilon^{\frac{1}{2}}) \\
&\leq \mathbb{E}\int_s^T H_x(t,x^\varepsilon(t),\mathbb{E}(x^\varepsilon(t)),u^\varepsilon(t),\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta))(x(t)-x^\varepsilon(t))dt \\
&+\mathbb{E}\int_s^T H_y(t,x^\varepsilon(t),\mathbb{E}(x^\varepsilon(t)),u^\varepsilon(t),\Psi^\varepsilon(t),K^\varepsilon(t),\gamma_t^\varepsilon(\theta))(x(t)-x^\varepsilon(t))dt \\
&+C\varepsilon^{\frac{1}{2}}.
\end{aligned} \tag{2.65}$$

On the other hand, by using (2.48) we get

$$h\left(x(T),\mathbb{E}\left(x(T)\right)\right)-h(x^\varepsilon(T),\mathbb{E}\left(x^\varepsilon(T)\right))\geq$$
$$\left[h_x(x^\varepsilon(T),\mathbb{E}(x^\varepsilon(T)))+h_y(x^\varepsilon(T),\mathbb{E}(x^\varepsilon(T)))\right](x(T)-x^\varepsilon(T))$$

Noticing that since $\Psi^\varepsilon(T)=h_x(x^\varepsilon(T),\mathbb{E}(x^\varepsilon(T)))+\mathbb{E}\left(h_y(x^\varepsilon(T),\mathbb{E}(x^\varepsilon(T)))\right)$ then we have

$$\mathbb{E}\left\{h\left(x(T),\mathbb{E}\left(x(T)\right)\right)-h(x^\varepsilon(T),\mathbb{E}\left(x^\varepsilon(T)\right))\right\}\geq\mathbb{E}\left\{\Psi^\varepsilon(T)(x(T)-x^\varepsilon(T))\right\}. \tag{2.66}$$

By integration by parts formula for jumps process $\Psi^\varepsilon(t)(x(t)-x^\varepsilon(t))$ (see Lemma A1) we get

$$
\begin{aligned}
&\mathbb{E}\left[\Psi^{\varepsilon}(T)(x(T)-x^{\varepsilon}(T))\right] \\
&=\mathbb{E}\int_s^T \Psi^{\varepsilon}(t)d(x(t)-x^{\varepsilon}(t))+\mathbb{E}\int_s^T (x(t)-x^{\varepsilon}(t))d\Psi^{\varepsilon}(t) \\
&+\mathbb{E}\int_s^T K^{\varepsilon}(t)\left(\sigma(t,x(t),\mathbb{E}(x(t)),u(t))-\sigma(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t))\right)dt \\
&+\mathbb{E}\int_s^T\int_{\Theta}\gamma_t^{\varepsilon}(\theta)\left(g\left(t,x(t),u(t),\theta\right)-g\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t),\theta\right)\right)\mu(d\theta)dt,
\end{aligned}
$$

with the help of (2.1), and (2.9) we obtain

$$
\begin{aligned}
&\mathbb{E}\left\{\Psi^{\varepsilon}(T)(x(T)-x^{\varepsilon}(T))\right\} \\
&=\mathbb{E}\int_s^T \{[H_x(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),K^{\varepsilon}(t),\gamma_t^{\varepsilon}(\theta)) \\
&+\mathbb{E}(H_y(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),K^{\varepsilon}(t),\gamma_t^{\varepsilon}(\theta)))](x(t)-x^{\varepsilon}(t)) \\
&+\Psi^{\varepsilon}(t)\left[f(t,x(t),\mathbb{E}(x(t)),u(t))-f(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t))\right] \\
&+K^{\varepsilon}(t)\left[\sigma(t,x(t),\mathbb{E}(x(t)),u(t))-\sigma(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t))\right] \\
&+\left.\int_{\Theta}\gamma_t^{\varepsilon}(\theta)\left(g\left(t,x(t_-),u(t),\theta\right)-g\left(t,x^{\varepsilon}(t_-),u^{\varepsilon}(t),\theta\right)\right)\mu(d\theta)\right\}dt,
\end{aligned}
$$

then from (2.49) and (2.65) we get

$$
\begin{aligned}
&\mathbb{E}\left\{\Psi^{\varepsilon}(T)(x(T)-x^{\varepsilon}(T))\right\} \\
&\geq \mathbb{E}\textstyle\int_s^T\{H(t,x(t),\mathbb{E}(x(t)),u(t),\Psi^{\varepsilon}(t),K^{\varepsilon}(t),\gamma_t^{\varepsilon}(\theta)) \\
&-H(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t),\Psi^{\varepsilon}(t),K^{\varepsilon}(t),\gamma_t^{\varepsilon}(\theta)) \\
&+\Psi^{\varepsilon}(t)\left[f(t,x(t),\mathbb{E}(x(t)),u(t))-f(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t))\right] \\
&+K^{\varepsilon}(t)\left[\sigma(t,x(t),\mathbb{E}(x(t)),u(t))-\sigma(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t))\right] \\
&+\textstyle\int_{\Theta}\gamma_t^{\varepsilon}(\theta)\left(g\left(t,x(t_-),u(t),\theta\right)-g\left(t,x^{\varepsilon}(t_-),u^{\varepsilon}(t),\theta\right)\right)\mu(d\theta)\}\,dt-C\varepsilon^{\frac{1}{2}} \\
&=\mathbb{E}\textstyle\int_s^T\left(\ell\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)-\ell\left(t,x(t),\mathbb{E}(x(t)),u(t)\right)\right)dt-C\varepsilon^{\frac{1}{2}}.
\end{aligned}
\tag{2.67}
$$

Combining (2.66) and (2.67) we get

$$
\begin{aligned}
&\mathbb{E}\left\{h\left(x(T),\mathbb{E}\left(x(T)\right)\right)-h(x^{\varepsilon}(T),\mathbb{E}\left(x^{\varepsilon}(T)\right))\right\}\\
&\geq \mathbb{E}\int_s^T\left(\ell\left(t,x^{\varepsilon}(t),\mathbb{E}(x^{\varepsilon}(t)),u^{\varepsilon}(t)\right)-\ell\left(t,x(t),\mathbb{E}(x(t)),u(t)\right)\right)dt-C\varepsilon^{\frac{1}{2}},
\end{aligned}
$$

then by using definition of $J^{s,\zeta}$ we conclude

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

Finally, since $u(\cdot)$ is arbitrary element of $\mathcal{U}$, the desired result follows. □

2.5 Application to finance: Parameterized mean-variance portfolio selection

In this section, we will apply our necessary and sufficient conditions of near-optimality to study a parameterized mean-variance portfolio selection and we derive the explicit expression of the optimal portfolio selection strategy.

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

The first asset is a bond whose price $P_0(t)$ evolves according to the ordinary differential equation

(1) Risk-free security: (e.g., a bond), where the price $P_0(t)$ at time t is given by the following equation:

$$
\begin{cases}
dP_0\left(t\right)=P_0\left(t\right)\rho(t)dt,\ t\in[0,T]\\
P_0\left(0\right)>0,
\end{cases}
\tag{2.68}
$$

where $\rho(\cdot)$ is a bounded deterministic function.

(2) Risky security (e.g. a stock), where the price $P_1(t)$ at time t is given by

$$P_1(t) = P_1(t_-)\left[\varsigma(t)dt + \sigma_t dW(t) + \int_\Theta \xi_t(\theta) N(d\theta, dt)\right],\ P_1(0) > 0, \tag{2.69}$$

where $\varsigma(t)$, σ_t and $\xi_t(\theta)$ are bounded deterministic functions such that $\varsigma(t) \neq 0, \sigma_t \neq 0$ and $\varsigma(t) > \rho(t)$. and as above $N(d\theta, dt)$ is a compensated random measure.

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

(1) $\xi_t(\theta) > -1$ for any $\theta \in \Theta$.

(2) The function $t \rightarrow \int_\Theta \xi_t^2(\theta) \mu(d\theta)$ is a locally bounded

Portfolio and wealth dynamics: A portfolio is a predictable process $\pi(t) = (\pi_0(t), \pi_1(t))$ giving the number of units held at time t of the bond and the stock**.** The corresponding wealth process $x^\pi(t)$, $t \geq 0$ is then given by

$$x^\pi(t) = \pi_0(t)P_0(t) + \pi_1(t)P_1(t). \tag{2.70}$$

The portfolio $\pi(\cdot)$ is called *Self-financing* if

$$x^\pi(t) = x^\pi(0) + \int_0^t \pi_0(t)dP_0(t) + \int_0^t \pi_1(t_-)dP_1(t). \tag{2.71}$$

We denote by

$$v(t) = \pi_1(t)P(t), \tag{2.72}$$

the amount invested in the risky security. Now, by combining (2.70) and (2.71) together with

(2.72) we introduce the wealth dynamics as follows

$$\begin{cases} dx^v(t) = [\rho(t)x^v(t) + (\varsigma(t) - \rho(t))v(t)]\,dt + \sigma_t v(t)dW(t) \\ \qquad\qquad + \int_\Theta \xi_{t_-}(\theta)\, v(t)N(d\theta, dt)\,, \\ x^v(0) = \zeta, \end{cases} \tag{2.73}$$

where $\zeta \in \mathbb{R}$. If the corresponding wealth process $x^v(\cdot)$ given by SDE-(2.73) is square integrable, the control variable $v(\cdot)$ is called tame. We denote $\mathcal{U}$ the set of admissible portfolio valued in $\mathbb{A} = \mathbb{R}$.

Parameterized mean-variance portfolio selection. We assume that we have a family of optimization problem parameterized by ε, where ε is a small parameter $\varepsilon > 0$ may be represent the complexity of the cost functional

$$J^{\zeta,\varepsilon}(v(\cdot)) = \mathbb{E}\left\{\left(x^v(T) - \mathbb{E}(x^v(T)) - \frac{\varepsilon}{2})\right)^2 + \int_0^T \frac{\varepsilon^2}{4}L(v(t))dt\right\}, \tag{2.74}$$

subject to $x^v(T)$ solution of SDE-(2.73) at time T given by

$$\begin{aligned} x^v(T) = \zeta + &\int_0^T [\rho(t)x^v(t) + (\varsigma(t) - \rho(t))v(t)]\,dt + \int_0^T \sigma_t v(t)dW(t) \\ &+ \int_0^T \int_\Theta \xi_{t_-}(\theta)\, v(t)N(d\theta, dt)\,, \end{aligned}$$

where $L(\cdot)$ is a nonlinear, convex and bounded function, satisfying assumption (2.47) and independent of ε.

Our objective is to find an admissible portfolio $v^*(\cdot)$ which minimizes the cost function (2.74) of mean-field type (i.e., with $\ell \equiv \frac{\varepsilon^2}{4}L(v(t)), s = 0$, $h\left(x(t), \mathbb{E}(x(t))\right) = \left(x(t) - \mathbb{E}(x(t)) - \frac{\varepsilon}{2}\right)^2$). Explicit solution of problem (2.73)-(2.74), called $\mathcal{P}_\varepsilon$, may be a difficult problem. The idea is to show that we can easily get a near-optimal control (in feedback form) analytically based on the optimal control of the simpler problem, called $\mathcal{P}_0$ which is obtained by setting $\varepsilon = 0$

in (2.74), then we get

$$J_0^{\zeta}(v(\cdot)) = \mathbb{E}\left\{(x^v(T) - \mathbb{E}(x^v(T)))^2\right\}, \tag{2.75}$$

We study the optimal control problem where the state is governed by SDE-(2.73) with a new cost function (2.75). In a second step, we solve the control problem (2.73)-(2.75), and obtain an optimal solution explicitly. Finally, we solve the control problem $\mathcal{P}_\varepsilon$ of near-optimally.

Problem $\mathcal{P}_0$: (*optimal solution of mean-field stochastic control problem (2.73)-(2.75)).* By a standard argument, problem $\mathcal{P}_0$ can be solved as follows.

Since $f(t, x(t), \mathbb{E}(x(t), v(t)) = \rho(t)x(t) + (\varsigma(t) - \rho(t))v(t)$, $\sigma(t, x(t), \mathbb{E}(x(t), v(t)) = \sigma_t v(t)$, $g(t, x(t_-), v(t), \theta) = v(t)\xi_{t_-}(\theta)$, then the Hamiltonian H gets the form

$$\begin{aligned} H(t, x, \mathbb{E}(x), v(t), \Psi(t), K(t), \gamma_t(\theta)) &= -\Psi(t)\left[\rho(t)x(t) + (\varsigma(t) - \rho(t))v(t)\right] - K(t)\sigma_t v(t) \\ &\quad - v(t)\int_\Theta \gamma_t(\theta)\,\xi_t(\theta)\,\mu(d\theta) \\ &= -\Psi(t)\rho(t)x(t) - v(t)\left[\Psi(t)(\varsigma(t) - \rho(t))\right. \\ &\quad \left. + K(t)\sigma_t + \int_\Theta \gamma_t(\theta)\,\xi_t(\theta)\,\mu(d\theta)\right]. \end{aligned}$$

Consequently, since this is a linear expression of $v(\cdot)$ then it is clear that the supremum is attained at $v^*(t)$ satisfying

$$\Psi^*(t)(\varsigma(t) + \rho(t)) + K^*(t)\sigma_t + \int_\Theta \gamma_t^*(\theta)\,\xi_t(\theta)\,\mu(d\theta) = 0. \tag{2.76}$$

Since $h_x(x(T), \mathbb{E}(x(T)) = 2(x(T) - \mathbb{E}(x(T))$, $h_y(x(T), \mathbb{E}(x(T)) = -2(x(T) - \mathbb{E}(x(T))$ then a simple computation shows that the first-order adjoint equation (2.9) associated with $v^*(t)$ gets the form

$$\begin{cases} d\Psi^*(t) = -\rho(t)\Psi^*(t)dt + K^*(t)dW(t) + \int_\Theta \gamma_t^*(\theta)N(dt, d\theta) \\ \Psi^*(T) = 2(x^*(T) - \mathbb{E}(x^*(T)). \end{cases} \tag{2.77}$$

In order to solve the above equation (2.77) and to find the expression of $v^*(t)$ we conjecture a process $\Psi^*(t)$ of the form

$$\Psi^*(t) = \Phi_1(t)x^*(t) + \Phi_2(t)\mathbb{E}\left(x^*(t)\right) + \Phi_3(t), \tag{2.78}$$

where $\Phi_1(\cdot), \Phi_2(\cdot)$ and $\Phi_3(\cdot)$ are deterministic differentiable functions. (see Shi et al., [62] and Framstad et al, [13], Ma et al, [23] and Li [60] for other models of conjecture).

Applying Itô's formula to (2.78), in virtue of SDE-(2.73), we get

$$\begin{aligned} d\Psi^*(t) &= \Phi_1(t)\left\{\left[\rho(t)x^*(t) + (\varsigma(t)-\rho(t))v^*(t)\right]dt + \sigma_t v^*(t)dW(t)\right.\\ &\qquad \left. + \int_\Theta v^*(t)\xi_{t_-}(\theta)\,N(d\theta,dt)\right\} + x^*(t)\dot{\Phi}_1(t)dt \\ &\qquad + \Phi_2(t)\left[\rho(t)\mathbb{E}(x^*(t)) + (\varsigma(t)-\rho(t))v^*(t)\right]dt \\ &\qquad + \mathbb{E}\left(x^*(t)\right)\dot{\Phi}_2(t)dt + \dot{\Phi}_3(t)dt \\ &= \Big\{\Phi_1(t)\left[\rho(t)x^*(t) + (\varsigma(t)-\rho(t))v^*(t)\right] + x^*(t)\dot{\Phi}_1(t) \\ &\qquad + \Phi_2(t)\left[\rho(t)\mathbb{E}(x^*(t)) + (\varsigma(t)-\rho(t))v^*(t)\right] \\ &\qquad + \dot{\Phi}_2(t)\mathbb{E}\left(x^*(t)\right) + \dot{\Phi}_3(t)\Big\}\,dt \\ &\qquad + \Phi_1(t)\sigma_t v^*(t)dW(t) \\ &\qquad + \int_\Theta \Phi_1(t)v^*(t)\xi_{t_-}(\theta)\,N(d\theta,dt)\,, \\ \Psi^*(T) &= \Phi_1(T)x^*(T) + \Phi_2(T)\mathbb{E}\left(x^*(T)\right) + \Phi_3(T), \end{aligned} \tag{2.79}$$

Next, comparing (2.79) with (2.77), we get

$$\begin{aligned} -\rho(t)\Psi^*(t) &= \Phi_1(t)\left[\rho(t)x^*(t) + (\varsigma(t)-\rho(t))v^*(t)\right] + x^*(t)\dot{\Phi}_1(t) \\ &\qquad + \Phi_2(t)\left[\rho(t)\mathbb{E}(x^*(t)) + (\varsigma(t)-\rho(t))v^*(t)\right] \\ &\qquad + \dot{\Phi}_2(t)\mathbb{E}\left(x^*(t)\right) + \dot{\Phi}_3(t), \end{aligned} \tag{2.80}$$

$$K^*(t) = \Phi_1(t)\sigma_t v^*(t), \tag{2.81}$$

$$\gamma_t^*(\theta) = \Phi_1(t)v^*(t)\xi_t(\theta), \tag{2.82}$$

and

$$\Phi_1(T) = 2, \Phi_2(T) = -2, \Phi_3(T) = 0. \tag{2.83}$$

Combining (2.81) and (2.83) together with (2.76) we get

$$v^*(t) = \frac{-(\varsigma(t) - \rho(t))\Psi^*(t)}{\Phi_1(t)\left[\sigma_t^2 + \int_\Theta \xi_t^2(\theta)\,\mu(d\theta)\right]}. \tag{2.84}$$

We denote

$$A(t) = \sigma_t^2 + \int_\Theta \xi_t^2(\theta)\,\mu(d\theta), \tag{2.85}$$

by using (2.76) together with (2.84) and (2.85) then we can get

$$\begin{aligned}
&\Phi_3(t) = 0 \text{ for } t \in [0, T], \\
&v^*(t) = (\rho(t) - \varsigma(t))\,(A(t))^{-1}\,\frac{(\Phi_1(t)x^*(t) + \Phi_2(t)\mathbb{E}\,(x^*(t)))}{\Phi_1(t)}. \\
&\quad = \left\{(\rho(t) - \varsigma(t))\,(A(t))^{-1}\right\} x^*(t) \\
&\quad + \left\{(\rho(t) - \varsigma(t))\,(A(t))^{-1}\,\tfrac{\Phi_2(t)}{\Phi_1(t)}\right\} \mathbb{E}\,(x^*(t))
\end{aligned} \tag{2.86}$$

Now combining (2.80) with (2.78) we deduce

$$\begin{aligned}
v^*(t)\,(\Phi_1(t) + \Phi_2(t))\,(\rho(t) - \varsigma(t)) &= \left[2\rho(t)\Phi_1(t) + \dot{\Phi}_1(t)\right] x^*(t) \\
&+ \left[2\rho(t)\Phi_2(t) + \dot{\Phi}_2(t)\right] \mathbb{E}(x^*(t))
\end{aligned} \tag{2.87}$$

By comparing the terms containing $x^*(t)$ and $\mathbb{E}\,(x^*(t))$, we obtain from (2.86) with (2.87) the two ordinary differential equations (ODEs in short):

$$\begin{aligned}&\left[(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}-2\rho(t)\right]\Phi_1(t)+(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}\Phi_2(t)=\dot{\Phi}_1(t).\\&\left[(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}-2\rho(t)\right]\Phi_2(t)+(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}\tfrac{\Phi_2^2(t)}{\Phi_1(t)}=\dot{\Phi}_2(t),\end{aligned}\tag{2.88}$$

a simple computation from (2.88) we obtain

$$\dot{\Phi}_1(t)\Phi_2(t)=\dot{\Phi}_2(t)\Phi_1(t),\tag{2.89}$$

which is equivalent to $|\Phi_1(t)|=c_0\,|\Phi_2(t)|$ where c_0 is a positive constant. Since $\Phi_1(T)=2, \Phi_2(T)=-2$, (see (2.83)) we deduce $c_0=1$, then we get

$$|\Phi_1(t)|=|\Phi_2(t)|\,,\tag{2.90}$$

Let us turn to calculate explicitly $\Phi_1(t)$ and $\Phi_2(t)$. From (2.90) we have

$$\frac{\Phi_2(t)}{\Phi_1(t)}=Sgn(\Phi_1(t)\Phi_2(t)),$$

then by dividing the first ODE in (2.88) by $\Phi_1(t)$ and the second ODE by $\Phi_2(t)$ we get

$$\begin{aligned}&\left[(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}-2\rho(t)\right]+(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}Sgn(\Phi_1(t)\Phi_2(t))=\tfrac{\dot{\Phi}_1(t)}{\Phi_1(t)},\\&\left[(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}-2\rho(t)\right]+(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}Sgn(\Phi_1(t)\Phi_2(t))=\tfrac{\dot{\Phi}_2(t)}{\Phi_2(t)},\end{aligned}$$

from (2.83) then a simple computations shows that for any $t\in[0,T]$

$$\begin{aligned}|\Phi_1(t)|=2\exp\Big\{-\textstyle\int_t^T&\left[(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}-2\rho(t)\right]\\&+(\rho(t)-\varsigma(t))^2\left(A(t)\right)^{-1}Sgn(\Phi_1(t)\Phi_2(t))dt\Big\}.\end{aligned}\tag{2.91}$$

With this choice of $\Phi_1(t)$ and $\Phi_2(t)$, we conclude that $v^*(t)$ is given by

$$\begin{aligned} v^*(t) &= \left[(\rho(t) - \varsigma(t))\,(A(t))^{-1}\right] x^*(t) \\ &\quad + \left[(\rho(t) - \varsigma(t))\,(A(t))^{-1}\, Sgn(\Phi_1(t)\Phi_2(t))\right] \mathbb{E}\,(x^*(t)) \end{aligned} \tag{2.92}$$

and the adjoint processes

$$\begin{aligned} \Psi^*(t) &= \Phi_1(t)x^*(t) + \Phi_2(t)\mathbb{E}\,(x^*(t))\,, \\ K^*(t) &= \Phi_1(t)\sigma_t v^*(t), \\ \gamma_t^*(\theta) &= \Phi_1(t)\xi_t\,(\theta)\, v^*(t), \end{aligned}$$

satisfying the adjoint equation (2.9). Moreover, with this choice of $v^*(t)$, the maximum condition (2.14) of *Theorem 2.3.1* holds. Since $h\,(x(t), \mathbb{E}x(t)) = (x(t) - \mathbb{E}x(t))^2$ is convex and $H\,(\cdot,\cdot,\cdot,\Psi(t), K(t), \gamma_t(\theta))$ is concave, we can assert that our admissible portfolio $v^*(t)$ is optimal and the sufficient conditions in *Theorem 2.4.1* are satisfied where $v^*(t)$ achieves the maximum. Finally, we give the explicit optimal portfolio in the state feedback form in the following theorem.

Theorem 2.5.1. The optimal solution of our mean-field stochastic control problem $\mathcal{P}_0$ is given in the state feedback form by

$$\begin{aligned} v^*(t, x^*(t), \mathbb{E}\,(x^*(t))) &= \left[(\rho(t) - \varsigma(t))\,(A(t))^{-1}\right] x^*(t) \\ &\quad + \left[(\rho(t) - \varsigma(t))\,(A(t))^{-1}\, Sgn(\Phi_1(t)\Phi_2(t))\right] \mathbb{E}\,(x^*(t))\,, \end{aligned} \tag{2.93}$$

where $A(t)$, $\Phi_1(t)$ and $\Phi_2(t))$ are given by (2.85), (2.91) and (2.90) respectively.

Problem $\mathcal{P}_{\varepsilon}$: The Hamiltonian function $\mathcal{H}$ for the problem $\mathcal{P}$ is

$$\begin{aligned}\mathcal{H}^{(z(.),v(\cdot))}(t,x,u) = &-\Psi(t)\rho(t)x(t) - u(t)\left\{\Psi(t)(\varsigma(t)-\rho(t))\right.\\ &\left.+K(t)\sigma_t + \int_{\Theta}\gamma_t(\theta)\,\xi_t(\theta)\,\mu(d\theta)\right\}\\ &+\sigma_t^2 v(t)u(t)Q(t) - \frac{1}{2}\sigma_t^2 u^2(t)Q(t)\\ &+u(t)v(t)\int_{\Theta}(\xi_t(\theta))^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta),\\ &-\frac{1}{2}v(t)\int_{\Theta}(\xi_t(\theta))^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta),\end{aligned}$$

where $Q^*(\cdot)$ is given by second-order adjoint equation

$$\begin{cases} dQ^*(t) = -2\rho(t)Q^*(t)dt + R^*(t)dW(t) + \int_{\Theta}\Gamma_t^*(\theta)N(d\theta,dt) \\ Q^*(T) = 2. \end{cases}$$

By uniqueness of the solution of the above classical backward SDE it is easy to show that

$$(Q^*(t),R^*(t),\Gamma_t^*(\theta)) = (2\exp\left\{2\int_t^T\rho(r)dr\right\},0,0),$$

then we get

$$\begin{aligned}\mathcal{H}^{x^*(.),v^*(\cdot)}(t,x,v) = &-\Psi(t)\rho(t)x(t) - v(t)\left\{\Psi^*(t)(\varsigma(t)-\rho(t))\right.\\ &\left.+K^*(t)\sigma_t + \textstyle\int_{\Theta}\gamma_t^*(\theta)\,\xi_t(\theta)\,\mu(d\theta)\right\}\\ &+\sigma_t^2 v^*(t)v(t)Q^*(t) - \tfrac{1}{2}\sigma_t^2 v^2(t)Q^*(t)\\ &+v(t)v^*(t)\textstyle\int_{\Theta}(\xi_t(\theta))^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta)\\ &-\tfrac{1}{2}v^2(t)\textstyle\int_{\Theta}(\xi_t(\theta))^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta).\end{aligned} \tag{2.94}$$

Since $v^*(\cdot)$ is optimal, by stochastic maximum principle, it necessary that $v^*(\cdot)$ maximizes the $\mathcal{H}$-function $a.s.$ namely,

$$\Psi^{*}(t)(\varsigma(t)-\rho(t))+K^{*}(t)\sigma_{t}+\int_{\Theta}\gamma_{t}^{*}(\theta)\,\xi_{t}(\theta)\,\mu(d\theta)=0 \qquad \mathbb{P}-a.s,\ a.e.\ t. \tag{2.95}$$

The Hamiltonian $\mathcal{H}_{\varepsilon}$ for the problem $\mathcal{P}_{\varepsilon}$ is

$$\begin{aligned}\mathcal{H}_{\varepsilon}^{(x^{*}(.),v^{*}(\cdot))}(t,x,v)=&-\Psi(t)\rho(t)x(t)-v(t)\left\{\Psi^{*}(t)(\varsigma(t)-\rho(t))\right.\\&\left.+K^{*}(t)\sigma_{t}+\int_{\Theta}\gamma_{t}^{*}(\theta)\,\xi_{t}(\theta)\,\mu(d\theta)\right\}\\&+\sigma_{t}^{2}v^{*}(t)v(t)Q^{*}(t)-\tfrac{1}{2}\sigma_{t}^{2}v^{2}(t)Q^{*}(t)\\&+v(t)v^{*}(t)\int_{\Theta}(\xi_{t}(\theta))^{2}\,(Q^{*}(t)+\gamma_{t}^{*}(\theta))\,\mu(d\theta)\\&-\tfrac{1}{2}v^{2}(t)\int_{\Theta}(\xi_{t}(\theta))^{2}\,(Q^{*}(t)+\gamma_{t}^{*}(\theta))\,\mu(d\theta)\\&-\tfrac{\varepsilon^{2}}{4}L(v(t)).\end{aligned} \tag{2.96}$$

The above function is maximized at $v^{\varepsilon}(t)$ which satisfies

$$\begin{aligned}&\Psi^{*}(t)(\varsigma(t)-\rho(t))+K^{*}(t)\sigma_{t}+\int_{\Theta}\gamma_{t}^{*}(\theta)\,\xi_{t}(\theta)\,\mu(d\theta)+\sigma_{t}^{2}v^{*}(t)Q^{*}(t)\\&-\sigma_{t}^{2}v^{\varepsilon}(t)Q^{*}(t)+v^{*}(t)\int_{\Theta}(\xi_{t}(\theta))^{2}\,(Q^{*}(t)+\gamma_{t}^{*}(\theta))\,\mu(d\theta)\\&-v^{\varepsilon}(t)\int_{\Theta}(\xi_{t}(\theta))^{2}\,(Q^{*}(t)+\gamma_{t}^{*}(\theta))\,\mu(d\theta)-\tfrac{\varepsilon^{2}}{4}\dot{L}(v^{\varepsilon}(t))=0,\\&\qquad\qquad\qquad\qquad\qquad\qquad\mathbb{P}-a.s,\ a.e.\ t.\end{aligned}$$

by applying (2.95) we have

$$\begin{aligned}(v^{*}(t)-v^{\varepsilon}(t))\left[\sigma_{t}^{2}Q^{*}(t)+\int_{\Theta}(\xi_{t}(\theta))^{2}\,(Q^{*}(t)+\gamma_{t}^{*}(\theta))\,\mu(d\theta)\right]\\-\frac{\varepsilon^{2}}{4}\dot{L}(v^{\varepsilon}(t))=0.\end{aligned} \tag{2.97}$$

Combining (2.96)-(2.95) then we can shows that

$$\begin{aligned}
&\max_{v(\cdot)\in\mathcal{U}} \mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x(t),v(t)) - \mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x(t),v^*(t)) \\
&= \mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x(t),v^{\varepsilon}(t)) - \mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x(t),v^*(t)) \\
&= \sigma_t^2 v^*(t)v^{\varepsilon}(t)Q^*(t) - \frac{1}{2}\sigma_t^2 \left(v^{\varepsilon}(t)\right)^2 Q^*(t) - \frac{\varepsilon^2}{4}L(v^{\varepsilon}(t)). \\
&+ \left(v^{\varepsilon}(t)v^*(t) - \frac{1}{2}\left(v^{\varepsilon}(t)\right)^2\right)\int_{\Theta}\left(\xi_{t_-}(\theta)\right)^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta) \\
&- \left\{\frac{1}{2}\sigma_t^2\left(v^*(t)\right)^2 Q^*(t) - \frac{\varepsilon^2}{4}L(v^*(t)) \right. \\
&\left. + \frac{1}{2}\left(v^*(t)\right)^2\int_{\Theta}\left(\xi_t(\theta)\right)^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta)\right\}
\end{aligned}$$

$$\begin{aligned}
&= \sigma_t^2 Q^*(t)\left[v^{\varepsilon}(t)v^*(t) - \frac{1}{2}\left(v^{\varepsilon}(t)\right)^2 - \frac{1}{2}\left(v^*(t)\right)^2\right] \\
&+ \left(v^{\varepsilon}(t)v^*(t) - \frac{1}{2}\left(v^{\varepsilon}(t)\right)^2 - \frac{1}{2}\left(v^*(t)\right)^2\right)\int_{\Theta}\left(\xi_t(\theta)\right)^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta) \\
&- \frac{\varepsilon^2}{4}\left(L(v^{\varepsilon}(t)) - L(v^*(t))\right),
\end{aligned}$$

since

$$v^{\varepsilon}(t)v^*(t) - \frac{1}{2}\left(v^{\varepsilon}(t)\right)^2 - \frac{1}{2}\left(v^*(t)\right)^2 = -\frac{1}{2}\left(v^*(t) - v^{\varepsilon}(t)\right)^2,$$

then by simple computation we get

$$\begin{aligned}
&\max_{v(\cdot)\in\mathcal{U}} \mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x,v(t)) - \mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x,v^*(t)) \\
&= -\frac{1}{2}\left(v^*(t) - v^{\varepsilon}(t)\right)^2\left[\sigma_t^2 Q^*(t) + \int_{\Theta}\left(\xi_{t_-}(\theta)\right)^2\left(Q^*(t)+\gamma_t^*(\theta)\right)\mu(d\theta)\right] \\
&- \frac{\varepsilon^2}{4}\left(L(v^{\varepsilon}(t)) - L(v^*(t))\right)
\end{aligned}$$

using (2.97), (2.47), and the fact that $L(\cdot)$ is convex and bounded we obtain

$$\begin{aligned}&\max_{v(\cdot)\in\mathcal{U}}\mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x,v(t))-\mathcal{H}_{\varepsilon}^{(x^*(.),v^*(\cdot))}(t,x,v^*(t))\\&=-\frac{\varepsilon^2}{8}\left(v^*(t)-v^{\varepsilon}(t)\right)\dot{L}(v^{\varepsilon}(t))+\frac{\varepsilon^2}{4}\left(L(v^*(t))-L(v^{\varepsilon}(t))\right)\leq C\varepsilon^2.\end{aligned}$$

Moreover, by using (2.95) the hamiltonian H_ε of problem $\mathcal{P}_\varepsilon$ is

$$\begin{aligned}H_{\varepsilon}\left(t,x,\mathbb{E}\left(x\right),v(t),\Psi(t),K(t),\gamma_t(\theta)\right)&=-\Psi(t)\rho(t)x(t)-v(t)\left\{\Psi(t)(\varsigma(t)-\rho(t))\right.\\&\left.+K(t)\sigma_t+\int_{\Theta}\gamma_t\left(\theta\right)\xi_t\left(\theta\right)\mu(d\theta)\right\}-\frac{\varepsilon^2}{4}L(v(t))\\&=-\Psi(t)\rho(t)x(t)-\frac{\varepsilon^2}{4}L(v(t)).\end{aligned}$$

Since $L(\cdot)$ is convex then the Hamiltonian $H_\varepsilon\left(t,\cdot,\cdot,\cdot,\Psi(t),K(t),\gamma_t(\theta)\right)$ is concave. By applying *Theorem 2.4.1*, this proves that, the control $v^*(t)$ given by (2.93) is indeed a near-optimal for stochastic control problem $\mathcal{P}_\varepsilon$.

2.6 Concluding remarks.

In this chapter, necessary and sufficient conditions of near-optimal stochastic control for systems governed by mean-field jump diffusion processes of mean-field type is proved. The control variable is allowed to enter both diffusion and jump coefficients and also the diffusion coefficients depend on the state of the solution process as well as of its expected value. Moreover, the cost functional is also of mean-field type. Our result is applied to financial optimization problem, where explicit expression of the optimal (and near-optimal) portfolio is obtained in the state feedback form. If we assume that $\varepsilon=0$ *Theorem 2.3.1* reduces to stochastic maximum principle of optimality developed in Hafayed et al., ([18], *Theorem 3.1*).

Moreover, if we assume that $\varepsilon=0$ and when the coefficients f, σ of the underlying jump diffusion processes and the cost functional do not explicitly depend on the expected value, *Theorem 2.3.1* reduces to necessary conditions of optimality developed in Tang et al., ([40], *Theorem 2.1*) and *Theorem 2.4.1* reduces to sufficient conditions of optimality developed in Framstad el al., ([13] *Theorem 2.1*).

Chapter-III

On Mean-field Partial Information Maximum Principle of Optimal Control for Stochastic Systems with Lévy Processes

Chapter 3

On Mean-field Partial Information Maximum Principle of Optimal Control for Stochastic Systems with Lévy Processes

Abstract. In this work, we study mean-field type partial information stochastic optimal control problem, where the system is governed by controlled stochastic differential equation driven by Teugels martingales associated with some Lévy process and an independent Brownian motion. We prove necessary and sufficient conditions of optimal control for these mean-field models in the form of maximum principle. The control domain is assumed to be convex. As an application, partial information linear quadratic control problem of mean-field type is discussed, where the optimal control is given in feedback form.

3.1 Introduction

We consider a mean-field stochastic control problem under partial information, where the controlled mean-field stochastic differential equation (SDEs) driven by Teugels martingales and an independent Brownian motion of the form

$$\begin{cases} dx^u(t) = f\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) dt \\ \qquad\qquad + \sum_{j=1}^{d} \sigma^j\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t)\right) dW^j(t) \\ \qquad\qquad + \sum_{j=1}^{\infty} g^j\left(t, x^u(t_-), \mathbb{E}(x^u(t_-)), u(t)\right) dH^j(t) \\ x^u(0) = x_0, \end{cases} \tag{3.1}$$

where f, σ and g are given maps and the initial condition x_0 is an $\mathcal{F}_0-$measurable random variable. The mean-field SDEs-(3.1) which is also called McKean-Vlasov systems are obtained as a limit approach, by the mean-square limit, as n goes to infinity of a system of interacting particles of the form:

$$\begin{aligned} dx_n^{u,j}(t) &= f(t, x_n^{u,j}(t), \frac{1}{n}\sum_{i=1}^{n} x_n^{u,i}(t), u(t))dt \\ &\quad + \sum_{k=1}^{d} \sigma^k(t, x_n^{u,j}(t), \frac{1}{n}\sum_{i=1}^{n} x_n^{u,i}(t), u(t))dW^{k,j}(t) \\ &\quad + \sum_{k=1}^{\infty} g^k(t, x_n^{u,j}(t_-), \frac{1}{n}\sum_{j=1}^{n} x_n^{u,j}(t_-), u(t))dH^{k,j}(t), \end{aligned}$$

where $W(\cdot)$ is a standard $d-$dimensional Brownian motion and $H(t) = \left(H^j(t)\right)_{j\geq 1}$ are pairwise strongly orthonormal Teugels martingales, associated with some Lévy process, having moments of all orders. The control $u(\cdot) = (u(t))_{t\geq 0}$ is required to be valued in some subset of $\mathbb{R}^k$ and adapted to a subfiltration $(\mathcal{G}_t)_{t\geq 0}$ of $(\mathcal{F}_t)_{t\geq 0}$. These Teugels martingales are the natural martingales, which generate the Hilbert space of square integrable martingales, with respect to the natural filtration of a Lévy process having moments of all orders.

The main new purpose here is the formulation of the partial information stochastic control in mean-field system with Lévy processes, which requires special attention. Noting that mean-field SDE associated with Lévy processes (3.1) under partial information occur naturally in the probabilistic analysis of financial optimization problems (incomplete financial market). 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 on the time interval $[0, T]$ is defined by

$$J(u(\cdot)) := \mathbb{E}\Big\{ \int_0^T \ell(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dt + h(x^u(T), \mathbb{E}(x^u(T))) \Big\}, \tag{3.2}$$

where ℓ and h are an appropriate functions. This cost functional is also of mean-field type, as the functions ℓ and h depend on the marginal law of the state process through its expected value. It worth mentioning that since the cost functional J is possibly a nonlinear function of the expected value stands in contrast to the standard formulation of a control problem. This leads to a so called *time-inconsistent control problem* where the Bellman dynamic programming does not hold. The reason for this is that one cannot apply the law of iterated expectations on the cost functional. This is a type of a control problem which, it seems, has not been studied before. An admissible control $u^*(\cdot)$ is called optimal iff it satisfies

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

The corresponding state processes, solution of mean-field system (3.1) is denoted by $x^*(\cdot) = x^{u^*}(\cdot)$.

Partial information or incomplete information, means that the information available to the controller is possibly less than the whole information. That is, any admissible control is adapted to a subfiltration $(\mathcal{G}_t)_t$ of $(\mathcal{F}_t)_t$ $t \geq 0$. This kind of problem, which has potential applications in mathematical finance and mathematical economics, arises naturally, because it may fail to obtain an admissible control with full information in real world applications.

To the best to our knowledge, the stochastic optimal control problems related to Teugel's martingales have been investigated by many authors. For example, [90, 91, 98, 99, 100, 101]. In Meng and Tang [90] the authors studied the general stochastic optimal control problem for the stochastic systems driven by Teugel's martingales and an independent multi-dimensional Brownian motion and recently, they prove the corresponding stochastic maximum principle. Optimal control problem for a backward stochastic control systems associated with Lévy processes under partial information has been investigated in Meng, Zhang and Tang [91]. The stochastic linear-quadratic problem with Lévy processes was studied by Mitsui and Tabata [98] and Tang and Wu [99]. Optimal control of BSDEs and FBSDEs driven by Teugels martingales has been studied in Tang and Zhang [100]. Stochastic maximum principle for SDEs with jumps under partial information was proved in Baghery and Øksendal [102].

Under complete information, the mean-field stochastic model was introduced by Kac [15] as a stochastic system for the Vlasov-kinetic equation of plasma and the study of which was initiated by McKean model [24]. Since then, many authors made contributions on mean-field stochastic control and applications, see for instance, [81, 82, 103, 104, 83, 58, 5, 6, 1, 105, 59, 60, 106, 107, 108, 96]. Second order necessary and sufficient conditions of near-optimal singular control for mean-field SDE have been established in Hafayed and Abbas [81]. More interestingly, mean-field type stochastic maximum principle for optimal singular control has been studied in Hafayed [82], in which convex perturbations used for both absolutly continuous and singular components. The maximum principle for optimal control of mean-field FBSDEJs has been studied in Hafayed [103]. The necessary and sufficient conditions for near-optimality for mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [104]. Singular optimal control for mean-field forward-backward stochastic systems and applications to finance has been investigated in Hafayed [83]. However, sufficient conditions of optimality for mean-field SDE with application have been investigated in Shi [58]. In Buckdahn, Djehiche, Li and Peng [5] a general notion of mean-field BSDE associated with a mean-field SDE is obtained in a natural way as a limit of some high dimensional system of FBSDEs governed by a $d-$dimensional Brownian motion, and influenced by positions of a large number of other particles. General maximum principle was introduced for a class of stochastic control problems involving SDEs of mean-field type in Buckdahn, Djehiche and Li [6]. Optimal control of nonlinear mean-field diffusion on Hilbert space was investigated in Ahmed [1]. In Lazry and Lions [105] the authors introduced a general mathematical modeling approach for high-dimensional systems of evolution equations corresponding to a large number of particles (or agents). Under the conditions that the control domains are convex, a various local maximum principle have been studied in [59, 60]. Second-order maximum principle for optimal stochastic control for mean-field jump diffusions was proved in Hafayed and Abbas [106]. Necessary and sufficient conditions for controlled jump diffusion with recent application in bicriteria mean-variance portfolio selection problem have been proved in Shen and Siu [107]. Recently, maximum principle for mean-field jump-diffusions stochastic delay differential equations and its applicationt to finance have been investigated in Yang, Meng and Shi [108]. A linear quadratic optimal control problem for mean-field stochastic differential equations has been studied in Yong [96]. Under partial information, mean-field type stochastic maximum principle for optimal control has been investigated in Wang, Zhang and Zhang [51].

Our main goal in this work is to establish a partial information necessary and sufficient conditions for optimal stochastic control of systems governed by mean-field SDEs associated with Lévy processes, where the coefficient of the system and the performance functional depend not only on the state process but also its marginal law of the state process through its expected value. The partial information mean-field control problem under consideration is not simple extension from the mathematical point of view, but also provide interesting models in many applications such as mathematical finance. An application is given to illustrate the theoretical results. Our result could be seen as an extension of necessary and sufficient conditions of stochastic systems associated with Lévy processes proved in Meng and Tang [90] to the mean-field models under partial information.

The rest of this work is structured as follows. The assumptions, notations and some basic definitions are given in Section 2. Sections 3 and 4 are devoted to prove our main results. As an illustration, time inconsistent linear quadratic mean-field problem is discussed in the section 5. Finally, Section 6 concludes the work and outlines some possible future developments.

3.2 Assumptions and Statement of the Control Problem

In this chapter, we study stochastic optimal control problems of mean-field type SDEs associated with Lévy processes of the following kind. Let $T > 0$ be a fixed time horizon and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a fixed filtered probability space equipped with a $\mathbb{P}-$completed right continuous filtration on which a $d-$dimensional Brownian motion $W = (W(t))_{t\in[0,T]}$ is defined. Let $L(\cdot) = (L(t))_{t\in[0,T]}$ be a $\mathbb{R}$-valued Lévy process, independent of the Brownian motion $W(\cdot)$, of the form $L(t) = bt + \lambda(t)$, where $\lambda(t)$ is a pure jump process. Assume that the Lévy measure $\mu(d\theta)$ corresponding to the Lévy process $\lambda(t)$ satisfies

1. $\int_{\mathbb{R}} (1 \wedge \theta^2)\mu(d\theta) < \infty.$
2. There exist $\gamma > 0$ such that for every $\delta > 0 : \int_{]-\delta,\delta[} \exp(\gamma |\theta|)\, \mu(d\theta) < \infty.$

We assume that $(\mathcal{F}_t)_{t\in[0,T]}$ is $\mathbb{P}-$augmentation of the natural filtration $(\mathcal{F}_t^{(W,L)})_{t\in[0,T]}$ defined as follows:

$$\mathcal{F}_t^{(W,L)} := \mathcal{F}_t^W \vee \sigma\{L(s) : 0 \leq s \leq t\} \vee \mathcal{G}_0,$$

where $\mathcal{F}_t^W := \sigma\{W(s) : 0 \leq s \leq t\}$, $\mathcal{G}_0$ denotes the totality of $\mathbb{P}-$null sets, and $\mathcal{F}_1 \vee \mathcal{F}_2$ denotes the σ-field generated by $\mathcal{F}_1 \cup \mathcal{F}_2$.

We denote $\mathcal{U}_{\mathcal{G}}([0,T])$ the set of all admissible controls.

Throughout this work, the power jump processes is defined by

$$\begin{cases} L_{(k)}(t) = \sum_{0<\tau\leq t} \left(\Delta L(\tau)\right)_k : k > 1 \\ L_{(1)}(t) = L(t), \end{cases}$$

where $\Delta L(\tau) := L(\tau) - L(\tau_-)$. Moreover, we define the continuous part of the control by

$$L_{(k)}{}^{(c)}(t) := L_{(k)}(t) - \sum_{0<\tau\leq t} \left(\Delta L(\tau)\right)_k : k > 1,$$

i.e., the process obtained by removing the jumps of $L(t)$. If we define

$$N_{(k)}(t) := L_{(k)}(t) - \mathbb{E}\left\{L_{(k)}(t)\right\} : k \geq 1,$$

then the family of Teugels martingales $\left(H_j(\cdot)\right)_{j\geq 1}$ is defined by $H_j(t) := \sum_{1<k\leq j} \alpha_{jk} N_k(t)$ where the coefficients α_{jk} associated with the orthonormalization of the polynomials $\{1, x, x^2, ...\}$ with respect to the measure $m(dx) = x^2\mu(dx)$. The jumps of $x^u(t)$ caused by the Lévy martingals $\Delta_L x^u(t)$ is defined by

$$\Delta_L x^u(t) := g\left(t, x^u(t_-), \mathbb{E}(x^u(t_-)), u(t)\right) \Delta L(t).$$

For convenience, we will use the following notation in this work.

1. l^2: the Hilbert space of real-valued sequences $x = (x_n)_{n\geq 0}$ such that $\|x\| := \left[\sum_{n=1}^{\infty} x_n\right]^2 < +\infty$, and $l^2\left(\mathbb{R}^n\right)$: the space of $\mathbb{R}^n$-valued $\left(f_n\right)_{n\geq 1}$ such that $\left[\sum_{n=1}^{\infty} \|f_n\|_{\mathbb{R}^n}^2\right]^{\frac{1}{2}} < +\infty$.
2. $l_{\mathcal{F}}^2\left([0,T]\,;\mathbb{R}^n\right)$ denotes the Banach space of $\mathcal{F}_t$-adapted processes such that $\mathbb{E}\left(\int_0^T |x(t)|_{\mathbb{R}^n}^2\, dt\right)^{\frac{1}{2}} < +\infty$.
3. $\mathbb{L}_{\mathcal{F}}^2\left([0,T]\,;\mathbb{R}^n\right)$ denotes the Banach space of $\mathcal{F}_t-$predictable processes such that $\mathbb{E}\left(\int_0^T \sum_{n=1}^{\infty} \|f_n\|_{\mathbb{R}^n}^2\, dt\right)^{\frac{1}{2}} < +\infty$.
4. $\mathbb{S}_{\mathcal{F}}^2\left([0,T]\,;\mathbb{R}^n\right)$ denotes the Banach space of $\mathcal{F}_t-$adapted and cadlag processes such that $\mathbb{E}(\sup|x(t)|^2)^{\frac{1}{2}} < +\infty$.

5. $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^n)$ the Banach space of $\mathbb{R}^n$-valued, square integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

6. $\mathbb{M}^{n\times m}(\mathbb{R})$ denotes the space of $n \times m$ real matrices.

7. For a differentiable function Φ we denote by $\Phi_x(t)$ its gradient with respect to the variable x.

8. $1_{[t,t+r]}(\cdot)$ denotes the indicator function on the set $[t, t+r]$.

In this work, we assume

$$\begin{aligned}
&f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \to \mathbb{R}^n, \\
&\sigma : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \to \mathbb{M}^{n\times d}(\mathbb{R}), \\
&g : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \to l^2(\mathbb{R}^n), \\
&h : \mathbb{R}^n \to \mathbb{R}.
\end{aligned}$$

Conditions (A1) The functions f, σ, ℓ, g and h are continuously differentiable in their variables including $(x, \widetilde{x}, u)$. The maps f, σ, g are progressively measurable processes such that $f(\cdot, 0, 0, 0)$ and $g(\cdot, 0, 0, 0) \in \mathbb{L}^2_{\mathcal{F}}([0,T]; \mathbb{R}^n)$, and $\sigma(\cdot, 0, 0, 0) \in \mathbb{M}^2_{\mathcal{F}}([0,T]; \mathbb{R}^n)$.

Conditions (A2) The derivatives of f, σ and g with respect to their variables including $(x, \widetilde{x}, u)$ are bounded. Further the map ℓ are dominated by $C(1 + |x| + |u|)$ and its derivatives are dominated by $C(1+|x|^2 + |u|^2)$. The map h is dominated by $C(1 + |x|)$ and its derivatives with respect to $(x, \widetilde{x})$ are dominated by $C(1 + |x|^2)$.

Thanks to Lemma 2.1 in Meng and Tang [90], and under conditions (A1) and (A2), the

SDE-(4.1) has an unique solution $x^u(\cdot) \in \mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ such that

$$\begin{aligned} x^u(t) = x_0 &+ \int_0^t f(s, x^u(s), \mathbb{E}(x^u(s)), u(s))\, ds \\ &+ \int_0^t \sum_{j=1}^d \sigma^j(s, x^u(s), \mathbb{E}(x^u(s)), u(s)) dW^j(s) \\ &+ \int_0^t \sum_{j=1}^\infty g^j(s, x^u(s), \mathbb{E}(x^u(s)), u(s))\, dH^j(s). \end{aligned}$$

Mean-field Adjoint Equations. We introduce the new adjoint equations involved in the stochastic maximum principle for our mean-field control problem (3.1)-(3.2). For simplicity of notation, we will still use $f_x(t) := \dfrac{\partial f}{\partial x}(t, x^u(\cdot), \mathbb{E}(x^u(\cdot)), u(\cdot))$, So for any admissible control $u(\cdot) \in \mathcal{U}_{\mathcal{G}}([0,T])$ and the corresponding state trajectory $x(\cdot) = x^u(\cdot)$, we consider the following adjoint equations of mean-field type

$$\begin{cases} -d\Psi(t) = [f_x(t)\Psi(t) + \mathbb{E}[f_{\widetilde{x}}(t)\Psi(t)] + \sum_{j=1}^d (\sigma_x^j(t) Q^j(t) + \mathbb{E}[\sigma_{\widetilde{x}}^j(t) Q^j(t)]) \\ \qquad + \sum_{j=1}^\infty (g_x^j(t) K^j(t) + \mathbb{E}[g_{\widetilde{x}}^j(t) K^j(t)]) + \ell_x(t) + \mathbb{E}(\ell_x(t))]dt \\ \qquad - \sum_{j=1}^d Q^j(t) dW(t) - \sum_{j=1}^\infty K^j(t) dH^j(t) \\ \Psi(T) = h_x(x(T), \mathbb{E}(x(T))) + \mathbb{E}[h_{\widetilde{x}}(x(T), \mathbb{E}(x(T)))]. \end{cases} \tag{3.3}$$

We define the Hamiltonian function

$$\mathcal{H} : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A} \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times l^2(\mathbb{R}^n) \to \mathbb{R}^n,$$

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

$$\begin{aligned} \mathcal{H}(t, x, \widetilde{x}, u, \Psi(\cdot), Q(\cdot), K(\cdot)) := \Psi(t) f(t, x, \widetilde{x}, u) + \textstyle\sum_{j=1}^d Q^j(t)\sigma(t, x, \widetilde{x}, u) \\ + \textstyle\sum_{j=1}^\infty K^j(t) g^j(t, x, \widetilde{x}, u) + \ell(t, x, \widetilde{x}, u). \end{aligned} \tag{3.4}$$

If we denote by $\mathcal{H}(t) := \mathcal{H}(t, x(t), \widetilde{x}(t), u(t), \Psi(t), Q(t), K(t))$, then the adjoint equation (3.3) can be rewritten as the following stochastic Hamiltonian system's type

$$\begin{cases} -d\Psi(t) = \{\mathcal{H}_x(t) + \mathbb{E}[\mathcal{H}_{\widetilde{x}}(t)]\} dt - \sum_{j=1}^{d} Q^j(t) dW(t) - \sum_{j=1}^{\infty} K^j(t) dH^j(t) \\ \Psi(T) = h_x(x(T), \mathbb{E}(x(T))) + \mathbb{E}[h_{\widetilde{x}}(x(T), \mathbb{E}(x(T)))]. \end{cases} \tag{3.5}$$

It is a well known fact that under assumptions (A1) and (A2), the adjoint equations (3.3) or (3.5) admits a unique solution $(\Psi(t), Q(t), K(t))$ such that $(\Psi(t), Q(t), K(t)) \in \mathbb{S}^2_{\mathcal{F}}([0,T]; \mathbb{R}^n) \times \mathbb{L}^2_{\mathcal{F}}([0,T]; \mathbb{R}^{n\times d}) \times l^2_{\mathcal{F}}([0,T]; \mathbb{R}^n)$. Moreover, since the derivatives of f, σ, g, h with respect to $(x, \widetilde{x})$ are bounded, we deduce from standard arguments that there exists a constant $C > 0$ such that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\Psi(t)|^2 + \sup_{t\in[0,T]} |K(t)|^2 + \int_0^T |Q(t)|^2 dt\right] < C. \tag{3.6}$$

3.3 Partial Information Necessary Conditions for Optimal Control of Mean-field SDEs with Lévy Processes

In this section, we establish a set of necessary conditions for a stochastic control to be optimal where the system evolves according to nonlinear controlled mean-field SDEs associated with Lévy processes. In addition to the assumptions in Section 2 we assume the following

Conditions (A3)

1. For all t, r such that $0 \leq t \leq t + r \leq T$, all $i : 1, 2, ..., k$ and all bounded $\mathcal{G}_t-$measurable $\alpha = \alpha(w)$, the control $\beta(t) = (0, ..., 0, \beta_i(t), 0, ..., 0) \in \mathbb{A} \subset \mathbb{R}^k$, with $\beta_i(s) = \alpha_i 1_{[t,t+r]}(s)$, $s \in [0,T]$ belong to $\mathcal{U}_{\mathcal{G}}([0,T])$.

2. For all $u(\cdot), \beta \in \mathcal{U}_{\mathcal{G}}([0,T])$, with β bounded, there exist $\delta > 0$ such that $u + y\beta \in \mathcal{U}_{\mathcal{G}}([0,T])$ for all $y \in [-\delta, \delta]$.

For a given $u(\cdot), \beta \in \mathcal{U}_{\mathcal{G}}([0,T])$ bounded, we define the process $Z(\cdot)$ by

$$Z(t) = Z^{u,\beta}(t) := \frac{d}{dy}(x^{u,\beta})(t).$$

Note that $Z(t)$ satisfies the following mean-field linear stochastic differential equation driven by both Brownian motion and Teugels martingales

$$\begin{cases} dZ(t) = [f_x(t)Z(t) + f_{\widetilde{x}}(t)\mathbb{E}(Z(t)) + f_u(t)\beta(t)]dt \\ \qquad + \sum_{j=1}^{d}[\sigma_x^j(t)Z(t) + \sigma_{\widetilde{x}}^j(t)\mathbb{E}\left(Z(t)\right) + \sigma_u^j(t)\beta(t)]dW^j(t) \\ \qquad + \sum_{j=1}^{\infty}[g_x^j(t)Z(t) + g_{\widetilde{x}}^j(t)\mathbb{E}\left(Z(t)\right) + g_u^j(t)\beta(t)]dH^j(t) \\ Z(0) = 0. \end{cases}$$

The following theorem constitutes the main contribution of this work.

Let $u^*(\cdot)$ be a local minimum for the cost J over $\mathcal{U}_{\mathcal{G}}\left([0,T]\right)$ in the sense that for all bounded β, there exist $\delta > 0$ such that $(u^* + y\beta) \in \mathcal{U}_{\mathcal{G}}\left([0,T]\right)$ for all $y \in [-\delta, \delta]$ and $\varphi(y) = J(u^* + y\beta)$ is maximal at $y = 0$:

$$\varphi'(y) = \frac{d}{dy}J(u^* + y\beta) = 0. \tag{3.7}$$

Let $x^*(\cdot)$ be the solution of the mean-field SDEs-(3.1) corresponding to $u^*(\cdot)$.

Theorem 3.3.1. (Partial information necessary condition for optimality in integral form). *Let conditions* (A1), (A2) *and* (A3) *hold. Then there exists a unique triplet of adapted process* $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot))$ *solution of adjoint equation* (3.3) *such that* $u^*(\cdot)$ *is a stationary point for* $\mathbb{E}\left[\mathcal{H} \mid \mathcal{G}_t\right]$ *in the sense that for almost all* $t \in [0,T]$ *we have*

$$\mathbb{E}\left[\mathcal{H}_u(t, x^*(t_-), \mathbb{E}(x^*(t_-)), u^*\left(t\right), \Psi^*(t), Q^*(t), K^*(t)) \mid \mathcal{G}_t\right] = 0, \ a.e., \ t \in [0,T]\,. \tag{3.8}$$

Proof. From (3.7) we have

$$
\begin{aligned}
0 &= \frac{d}{dy} J(u^* + y\beta) \qquad (3.9)\\
&= \mathbb{E}\Big\{ \int_0^T \ell_x(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t)) Z^*(t) dt + \mathbb{E}(\ell_{\widetilde{x}}(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t))) Z^*(t)\\
&+ \int_0^T \ell_u(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t)) \beta(t) dt\\
&+ h_x\left(x^*(T), \mathbb{E}\left(x^*(T)\right)\right) Z^*(T) + \mathbb{E}(h_{\widetilde{x}}\left(x^*(T), \mathbb{E}\left(x^*(T)\right)\right) Z^*(T) \Big\}.
\end{aligned}
$$

By applying Itô's formula to $\Psi^*(t)Z^*(t)$ and take expectation we get

$$
\begin{aligned}
\mathbb{E}(\Psi^*(T)Z^*(T)) &= \mathbb{E}\int_0^T \Psi^*(t) dZ^*(t) + \mathbb{E}\int_0^T Z^*(t) d\Psi^*(t)\\
&+ \mathbb{E}\int_0^T \sum_{j=1}^d Q^{j*}(t)[\sigma_x^j(t) Z^*(t) + \sigma_{\widetilde{x}}^j(t)\mathbb{E}\left(Z^*(t)\right) + \sigma_u^j(t)\beta(t)] dt\\
&+ \mathbb{E}\int_0^T \sum_{j=1}^d K^{j*}(t)[g_x^j(t) Z^*(t) + g_{\widetilde{x}}^j(t)\mathbb{E}\left(Z^*(t)\right) + g_u^j(t)\beta(t)] dt\\
&= \mathbb{I}_1 + \mathbb{I}_2 + \mathbb{I}_3 + \mathbb{I}_4,
\end{aligned} \qquad (3.10)
$$

where

$$
\begin{aligned}
\mathbb{I}_1 &= \mathbb{E}\int_0^T \Psi^*(t) dZ^*(t)\\
&= \mathbb{E}\int_0^T \Psi^*(t)\left[f_x\left(t\right) Z^*(t) + f_{\widetilde{x}}\left(t\right)\mathbb{E}(Z^*(t)) + f_u(t)\beta(t)\right] dt \qquad (3.11)\\
&= \mathbb{E}\int_0^T \Psi^*(t) f_x\left(t\right) Z^*(t) + \mathbb{E}\int_0^T \Psi^*(t) f_{\widetilde{x}}\left(t\right)\mathbb{E}(Z^*(t)) + \mathbb{E}\int_0^T \Psi^*(t) f_u(t)\beta(t)) dt.
\end{aligned}
$$

By simple computations we get

$$
\begin{aligned}
\mathbb{I}_2 &= \mathbb{E}\int_0^T Z^*(t)d\Psi^*(t)\\
&= -\mathbb{E}\int_0^T Z^*(t)\Big\{f_x(t)\,\Psi^*(t)+\mathbb{E}\left(f_{\widetilde{x}}(t)\,\Psi^*(t)\right)+\textstyle\sum_{j=1}^d(\sigma_x^j(t)\,Q^{j*}(t)+\mathbb{E}(\sigma_{\widetilde{x}}^j(t)\,Q^{j*}(t)))\\
&\quad+\textstyle\sum_{j=1}^\infty(g_x^j(t)\,K^{j*}(t)+\mathbb{E}[g_{\widetilde{x}}^j(t)\,K^{j*}(t)])+\ell_x(t)+\mathbb{E}\left(\ell_{\widetilde{x}}(t)\right)\Big\}\,dt\\
&= -\mathbb{E}\int_0^T Z^*(t)f_x(t)\,\Psi^*(t)dt-\mathbb{E}\int_0^T Z^*(t)\mathbb{E}\left(f_{\widetilde{x}}(t)\,\Psi^*(t)\right)dt\\
&\quad-\mathbb{E}\int_0^T\textstyle\sum_{j=1}^d Z^*(t)(\sigma_x^j(t)\,Q^{j*}(t)+\mathbb{E}(\sigma_{\widetilde{x}}^j(t)\,Q^{j*}(t)))\\
&\quad-\mathbb{E}\int_0^T\textstyle\sum_{j=1}^\infty Z^*(t)(g_x^j(t)\,K^{j*}(t)+\mathbb{E}[g_{\widetilde{x}}^j(t)\,K^{j*}(t)])\\
&\quad-\mathbb{E}\int_0^T Z^*(t)\ell_x(t)\,dt-\mathbb{E}\int_0^T Z^*(t)\mathbb{E}\left(\ell_{\widetilde{x}}(t)\right)dt.
\end{aligned}
\tag{3.12}
$$

$$
\begin{aligned}
\mathbb{I}_3 &= \mathbb{E}\int_0^T\textstyle\sum_{j=1}^d Q^{j*}(t)[\sigma_x^j(t)Z^*(t)+\sigma_{\widetilde{x}}^j(t)\mathbb{E}\left(Z^*(t)\right)+\sigma_u^j(t)\beta(t)]dt\\
&= \mathbb{E}\int_0^T\textstyle\sum_{j=1}^d Q^{j*}(t)[\sigma_x^j(t)Z^*(t)dt+\mathbb{E}\int_0^T\sum_{j=1}^d Q^{j*}(t)\sigma_{\widetilde{x}}^j(t)\mathbb{E}\left(Z^*(t)\right)dt\\
&\quad+\mathbb{E}\int_0^T\textstyle\sum_{j=1}^d Q^{j*}(t)\sigma_u^j(t)\beta(t)dt,
\end{aligned}
\tag{3.13}
$$

and

$$
\begin{aligned}
\mathbb{I}_4 &= \mathbb{E}\int_0^T\textstyle\sum_{j=1}^d K^{j*}(t)[g_x^j(t)Z^*(t)+g_{\widetilde{x}}^j(t)\mathbb{E}\left(Z^*(t)\right)+g_u^j(t)\beta(t)]dt\\
&= \mathbb{E}\int_0^T\textstyle\sum_{j=1}^d K^{j*}(t)g_x^j(t)Z^*(t)dt+\mathbb{E}\int_0^T\sum_{j=1}^d K^{j*}(t)g_{\widetilde{x}}^j(t)\mathbb{E}\left(Z^*(t)\right)dt\\
&\quad+\mathbb{E}\int_0^T\textstyle\sum_{j=1}^d K^{j*}(t)g_u^j(t)\beta(t)]dt.
\end{aligned}
\tag{3.14}
$$

Combining (3.10), (3.11), (3.12), (3.13), (3.14) and the fact that

$$
\Psi^*(T)=h_x\left(x^*(T),\mathbb{E}(x^*(T))+\mathbb{E}\left[h_{\widetilde{x}}\left(x^*(T),\mathbb{E}(x^*(T))\right)\right]\ \text{and}\ Z^*(0)=0,
$$

we get

$$\begin{aligned}
&\mathbb{E}\left\{\left[h_x\left(x(T),\mathbb{E}(x(T)\right)+\mathbb{E}\left(h_{\widetilde{x}}\left(x(T),\mathbb{E}(x(T)\right)\right)\right]Z^*(T)\right\} \\
&=\mathbb{E}\int_0^T\Big\{\Psi^*(t)f_u(t)\beta(t)dt+\textstyle\sum_{j=1}^d Q^{j*}(t)\sigma_u^j\left(t\right)\beta(t)dt+\sum_{j=1}^d K^{j*}(t)g_u^j(t)\beta(t) \\
&\quad-\ell_x\left(t\right)Z^*(t)-\mathbb{E}\left(\ell_{\widetilde{x}}\left(t\right)\right)Z^*(t)\Big\}dt.
\end{aligned} \tag{3.15}$$

Combining (3.9) and (3.15) we obtain

$$\begin{aligned}
&\mathbb{E}\int_0^T\Big\{\Psi^*(t)f_u(t)\beta(t)dt+\sum_{j=1}^d Q^{j*}(t)\sigma_u^j\left(t\right)\beta(t)dt+\sum_{j=1}^d K^{j*}(t)g_u^j(t)\beta(t) \\
&+\ell_u(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))\beta(t)\Big\}dt=0,
\end{aligned}$$

which implied that

$$\mathbb{E}\int_0^T\mathcal{H}_u\left(t,x^*(t_-),\widetilde{x}^*(t_-),u^*(t),\Psi^*(t),Q^*(t),K^*(t)\right)\beta(t)dt=0. \tag{3.16}$$

Fix $t\in[0,T]$ and apply the above to $\beta=(0,...,\beta_i,...,0)$ where $\beta_i(s)=\alpha_i 1_{[t,t+r]}(s)$, $s\in[0,T]$, $t+r\leq T$ and $\alpha_i=\alpha_i(w)$ is bounded, $\mathcal{G}_t-$measurable. Then from (3.16) we get

$$\mathbb{E}\int_t^{t+r}\frac{\partial}{\partial u_i}\mathcal{H}\left(s,x^*(s_-),\widetilde{x}^*(s_-),u^*(s),\Psi^*(s),Q^*(s),K^*(s)\right)\alpha_i(w)ds=0.$$

Differentiating with respect to r at $r=0$ gives

$$\mathbb{E}\left[\frac{\partial}{\partial u_i}\mathcal{H}\left(s,x^*(s_-),\widetilde{x}^*(s_-),u^*(s),\Psi^*(s),Q^*(s),K^*(s)\right)\alpha_i\right]=0. \tag{3.17}$$

Since (3.17) holds for all bounded $\mathcal{G}_t-$measurable α_i, we have

$$\mathbb{E}\left[\mathcal{H}_u\left(t,x^*(t_-),\widetilde{x}^*(t_-),u^*(t),\Psi^*(t),Q^*(t),K^*(t)\right)\mid\mathcal{G}_t\right]=0.\qquad\mathbb{P}-a.s.$$

This completes the proof of Theorem 3.3.1 □

3.4 Partial Information Sufficient Conditions for Optimal Control of Mean-field SDEs with Lévy Processes

The sufficient condition of optimality is of significant importance in the stochastic maximum principle for computing optimal controls. In this section, we will prove that under some additional hypotheses on the Hamiltonian function is a sufficient condition for optimality.

Conditions (A4) . We assume

1. $\mathcal{H}\left(t,\cdot,\cdot,\cdot,\Psi^{*}(t),Q^{*}(t),K^{*}(t)\right)$ is convex with respect to $(x,\widetilde{x},u)$ for $a.e.t\in[0,T]\,,\mathbb{P}-a.s.$
2. $h(\cdot,\cdot)$ is convex with respect to $(x,\widetilde{x})$.

Theorem 3.4.1. *Let conditions* (A1), (A2), (A3) and (A4) *hold. Then* $u^{*}(\cdot)$ *is a partial information optimal control, i.e.,*

$$J(u^{*}(\cdot))=\inf_{u(\cdot)\in\mathcal{U}_{\mathcal{G}}([0,T])}J\left(u(\cdot)\right). \tag{3.18}$$

if satisfies (3.8).

To prove Theorem 3.4.1, we need the following auxiliary result, which deals with the duality relations between $\Psi^{*}(t)$, $[x^{u}(t)-x^{*}(t)]$. This Lemma is very important for proving our sufficient maximum principle. We denote by $\mathcal{H}_{x}^{*}(t):=\mathcal{H}_{x}(t,x^{*}(t),\mathbb{E}(x^{*}(t)),u^{*}(t),\Psi^{*}(t),Q^{*}(t),K^{*}(t))$ etc,.

Lemma 3.4.1 *Let* $x^{u}(\cdot)$ *be the solution of state mean-field* SDE-(3.1) *corresponding to any* *admissible control* $u(\cdot)$. *We have*

$$\mathbb{E}\left[\Psi^*(T)\left(x^u(T)-x^*(T)\right)\right]=\mathbb{E}\int_0^T\Psi^*(t)\left[f(t,x^u(t),\mathbb{E}(x^u(t)),u(t))\,-\,f(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))\right]dt$$
$$+\,\mathbb{E}\int_0^T\mathcal{H}_x^*(t)\left(x^u(t)-x^*(t)\right)dt+\mathbb{E}\int_0^T\mathbb{E}[\mathcal{H}_{\widetilde{x}}^*(t)]\left(\mathbb{E}(x^u(t))-\mathbb{E}(x^*(t))\right)dt$$
$$+\,\mathbb{E}\int_0^T\sum_{j=1}^d Q^{*,j}(t)\left[\sigma^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t))\,-\,\sigma^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))\right]dt$$
$$+\,\mathbb{E}\int_0^T\sum_{j=1}^d K^{j*}(t)[g^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-g^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))]dt. \tag{3.19}$$

Proof. First, by simple computations, we get

$$\begin{aligned}d\left(x^u(t)-x^*(t)\right)&=\left[f(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-f(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))\right]dt\\&+\sum_{j=1}^d\left[\sigma^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-\sigma^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t)\right]dW^j(t)\\&+\sum_{j=1}^\infty\left[g^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-g^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t)\right]dH^j(t)\,.\end{aligned} \tag{3.20}$$

By applying integration by parts formula to $\Psi^*(t)\left(x^u(t)-x^*(t)\right)$ and the fact that $x^u(0)-x^*(0)=0$, we get

$$\begin{aligned}&\mathbb{E}\left\{\Psi^*(T)\left(x^u(T)-x^*(T)\right)\right\}=\mathbb{E}\int_0^T\Psi^*(t)d\left(x^u(t)-x^*(t)\right)+\mathbb{E}\int_0^T\left(x^u(t)-x^*(t)\right)d\Psi^*(t)\\&+\mathbb{E}\int_0^T\sum_{j=1}^d Q^{*,j}(t)[\sigma^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-\sigma^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))]dt\\&+\mathbb{E}\int_0^T\sum_{j=1}^d K^{*,j}(t)[g^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-g^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))]dt\\&=I_1+I_2+I_3+I_4.\end{aligned} \tag{3.21}$$

From (3.20), we obtain

$$\begin{aligned}I_1&=\mathbb{E}\int_0^T\Psi^*(t)d\left(x^u(t)-x^*(t)\right)\\&=\mathbb{E}\int_0^T\Psi^*(t)[f(t,x^u(t),\mathbb{E}(x^u(t)),u(t))-f(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))]dt,\end{aligned} \tag{3.22}$$

similarly, by applying (3.5), we get

$$\begin{aligned} I_2 &= \mathbb{E}\int_0^T (x^u(t)-x^*(t))\, d\Psi^*(t) = \mathbb{E}\int_0^T (x^u(t)-x^*(t))\, [\mathcal{H}_x^*(t)+\mathbb{E}(\mathcal{H}_{\widetilde{x}}^*(t))]dt \\ &= \mathbb{E}\int_0^T \mathcal{H}_x^*(t)\,(x^u(t)-x^*(t))\,dt + \mathbb{E}\int_0^T \mathbb{E}(\mathcal{H}_{\widetilde{x}}(t))\,(\mathbb{E}(x^u(t))-\mathbb{E}(x^*(t)))dt. \end{aligned} \tag{3.23}$$

By standard arguments, we obtain

$$I_3 = \mathbb{E}\int_0^T \sum_{j=1}^d Q^*(t)[\sigma^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t)) - \sigma^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))]dt, \tag{3.24}$$

and

$$I_4 = \mathbb{E}\int_0^T \sum_{j=1}^d K^*(t)[g^j(t,x^u(t),\mathbb{E}(x^u(t)),u(t)) - g^j(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t))]dt. \tag{3.25}$$

The duality relation (3.19) follows from combining (3.22), (3.23), (3.24) and (3.25) together with (3.21). □

Proof of Theorem 3.4.1. Let $x^*(\cdot)$ be the solution of the state equation (3.1) and $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot))$ be the solution of the adjoint equation (3.3), corresponding to $u^*(\cdot) \in \mathcal{U}_{\mathcal{G}}([0,T])$ (condidate to be optimal). For any $u(\cdot) \in \mathcal{U}_{\mathcal{G}}([0,T])$ and from (3.2) we get

$$\begin{aligned} J(u^*(\cdot)) - J(u(\cdot)) &= \mathbb{E}\,[h(x^*(T),\mathbb{E}(x^*(t)) - h(x^u(T),\mathbb{E}(x^u(T))] \\ &+ \mathbb{E}\int_0^T [\ell(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t) - \ell(t,x^u(t),\mathbb{E}(x^u(t)),u(t)]\,dt. \end{aligned}$$

From the convexity of $h(\cdot,\cdot)$ we get

$$\begin{aligned} J(u^*(\cdot)) - J(u(\cdot)) &\le \mathbb{E}\,[(h_x(x^*(T),\mathbb{E}(x^*(T)) + \mathbb{E}(h_{\widetilde{x}}(x^*(T),\mathbb{E}(x^*(T)))))\,(x^*(T)-x^u(T))] \\ &+ \mathbb{E}\int_0^T [\ell(t,x^*(t),\mathbb{E}(x^*(t)),u^*(t)) - \ell(t,x^u(t),\mathbb{E}(x^u(t)),u(t))]\,dt. \end{aligned} \tag{3.26}$$

Since $\Psi^*(T) = h_x(x^*(T), \mathbb{E}\left(x^*(T)\right) + \mathbb{E}(h_{\widetilde{x}}(x^*(T), \mathbb{E}\left(x^*(T)\right)$, we get

$$\begin{aligned}
J\left(u^*(\cdot)\right) - J\left(u(\cdot)\right) &\leq \mathbb{E}\left[\Psi^*(T)\left(x^*(T) - x^u(T)\right)\right] \\
&\quad + \mathbb{E}\int_0^T \left[\ell(t, x^*(t), \mathbb{E}\left(x^*(t)\right), u^*(t)) - \ell(t, x^u(t), \mathbb{E}\left(x^u(t)\right), u(t))\right] dt.
\end{aligned}$$

By applying Lemma 3.4.1, we have

$$\begin{aligned}
J\left(u^*(\cdot)\right) - J\left(u(\cdot)\right) &\leq \mathbb{E}\int_0^T (\mathcal{H}\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right) \\
&- \mathcal{H}\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t), \Psi^*(t), Q^*(t), K^*(t)\right))dt \\
&- \mathbb{E}\int_0^T (x^*(t) - x^u(t))\left[\mathcal{H}_x\left(t, x^*(t), \mathbb{E}\left(x^*(t)\right), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\right. \\
&+ \left.\mathbb{E}\left(\mathcal{H}_{\widetilde{x}}\left(t, x^*(t), \mathbb{E}\left(x^*(t)\right), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\right)\right] dt.
\end{aligned} \tag{3.27}$$

By the convexity of $\mathcal{H}\left(t, \cdot, \cdot, \cdot, \Psi^*(t), Q^*(t), K^*(t)\right)$ (Conditions (A4), (2)) it hold that

$$\begin{aligned}
&\mathcal{H}\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^u(t), Q^u(t), K^u(t)\right) \\
&-\mathcal{H}\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t), \Psi^*(t), Q^*(t), K^*(t)\right) \\
&\leq \mathcal{H}_x\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\left(x^*(t) - x^u(t)\right) \\
&+\mathbb{E}\left(\mathcal{H}_{\widetilde{x}}\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\right)\left(x^*(t) - x^u(t)\right) \\
&+\mathcal{H}_u\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\left(u^*(t) - u(t)\right).
\end{aligned} \tag{3.28}$$

Since $\mathbb{E}\left[\mathcal{H}_u\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right) \mid \mathcal{G}_t\right]$, $u(t)$ and $u^*(t)$ are $\mathcal{G}_t-$measurable we get

$$\begin{aligned}
&\mathbb{E}\left[\mathcal{H}_u\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right) \mid \mathcal{G}_t\right]\left(u(t) - u^*(t)\right) \\
&= \mathbb{E}\left[\mathcal{H}_u\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\left(u(t) - u^*(t)\right) \mid \mathcal{G}_t\right],
\end{aligned} \tag{3.29}$$

from (3.8), (3.28) and (3.29) we obtain

$$\begin{aligned}
&\mathcal{H}\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right) \\
&-\mathcal{H}\left(t, x^u(t), \mathbb{E}(x^u(t)), u(t), \Psi^*(t), Q^*(t), K^*(t)\right)dt \\
&-\mathbb{E}\int_0^T \left[\mathcal{H}_x\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\right. \\
&\left.+\,\mathbb{E}\left(\mathcal{H}_{\widetilde{x}}\left(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right)\right)\right](x^*(t)-x^u(t))dt \leq 0,
\end{aligned} \tag{3.30}$$

by combining (3.27) and (3.30) we get

$$J(u(\cdot)) - J(u^*(\cdot)) \geq 0.$$

Finally, since $u(\cdot)$ is an arbitrary element of $\mathcal{U}_{\mathcal{G}}\left([0,T]\right)$ the desired result (3.18) follows. This completes the proof of Theorem 3.4.1. □

3.5 Application: Partial Information Mean-field Linear Quadratic Control Problem

In this section, partial information optimal stochastic linear quadratic control problem of mean-field type is considered. We give a mean-field partial information counterpart for the example studied in Meng and Tang [90]. The optimal control is represented by a state *feedback* form involving both $x(\cdot)$ and $\mathbb{E}(x(\cdot))$, via the solutions of Riccati ordinary differential equations. Then mean-field SDE (3.1), but now with linear coefficients, writes as follows

$$\begin{cases}
dx(t) = (Ax(t) + \widetilde{A}\mathbb{E}\left(x(t)\right) + Bu(t))dt + \sum_{j=1}^{d}(C^j x(t) + \widetilde{C}^j\mathbb{E}\left(x(t)\right) + D^j u(t))dW^j(t), \\
\qquad\qquad + \sum_{j=1}^{\infty}\left[\varrho^j x(t) + \widetilde{\varrho}^j\mathbb{E}\left(x(t)\right) + F^j u(t)\right]dH^j(t), \\
x(0) = x_0,
\end{cases} \tag{3.31}$$

where $A, \widetilde{A}, B, C, \widetilde{C}, D, \varrho, \widetilde{\varrho}$ and F are constants and $u(\cdot) \in \mathcal{U}_{\mathcal{G}}([0,T])$. The cost where R, N and Π are positive constants. Noting that the admissible controls $u = (u(t))$ are adapted to a subfiltration $(\mathcal{G}_t) : t \geq 0$. For a given control $u(\cdot)$, then due to (3.4) the Hamiltonian functional $\mathcal{H}$ corresponding to control problem (3.31)-(**??**) gets the form:

$$\begin{aligned} \mathcal{H}(t, x, \widetilde{x}, u, \Psi(\cdot), Q(\cdot), K(\cdot)) = \Psi(t)(Ax(t) + \widetilde{A}\mathbb{E}(x(t)) + Bu(t)) \\ + \textstyle\sum_{j=1}^{d} Q^j(t)(C^j x(t) + \widetilde{C}^j \mathbb{E}(x(t)) + D^j u(t)) \\ + \textstyle\sum_{j=1}^{\infty} K^j(t)(\varrho^j x(t) + \widetilde{\varrho}^j \mathbb{E}(x(t)) + F^j u(t)) \\ + \tfrac{1}{2}(Rx(t)^2 + Nu(t)^2), \end{aligned} \tag{3.32}$$

and due to (3.5) the corresponding adjoint equation gets the form

$$\begin{cases} d\Psi(t) = -[A\Psi(t) + \widetilde{A}\mathbb{E}(\Psi(t)) + \sum_{j=1}^{d}(C^j Q^j(t) + \widetilde{C}^j \mathbb{E}(Q^j(t))) \\ \qquad + \sum_{j=1}^{\infty}(\varrho^j K^j(t) + \widetilde{\varrho}\mathbb{E}[K^j(t)]) + Rx(t)]dt \\ \qquad + \sum_{j=1}^{d} Q^j(t)dW^j(t) + \sum_{j=1}^{\infty} K^j(t)dH^j(t) \\ \Psi(T) = \Pi x(T). \end{cases} \tag{3.33}$$

Let $u^*(\cdot)$ be a local optimal control of the partial information problem. For example, $\mathcal{G}_t$ could be the δ-delayed information defined by

$$\mathcal{G}_t = \mathcal{F}_{(t-\delta)^+} : t \geq 0,$$

where δ is a given constant delay. Then by applying Theorem 3.4.1 and the fact that

$$\mathcal{H}_u(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t), \Psi^*(t), Q^*(t)) = B\Psi^*(t) + DQ^*(t) + FK^*(t) + Nu^*(t),$$

we deduce that the optimal control is given by

$$\mathbb{E}[B\Psi^*(t) + DQ^*(t) + FK^*(t) + Nu^*(t) \mid \mathcal{G}_t] = 0, \; t \in [0,T]. \tag{3.34}$$

Since $u^*(t)$ is adapted to $\mathcal{G}_t$ we get

$$u^*(t) = -\frac{1}{N}\left\{B\mathbb{E}[\Psi^*(t) \mid \mathcal{G}_t] + D\mathbb{E}[Q^*(t) \mid \mathcal{G}_t] + F\mathbb{E}[K^*(t) \mid \mathcal{G}_t]\right\}. \tag{3.35}$$

In order to solve explicitly the above equation (3.35), we conjecture the adjoint process $\Psi^*(\cdot)$ as follows

$$\Psi^*(t) = \Phi_1(t)\,x^*(t) + \Phi_2(t)\,\mathbb{E}\left(x^*(t)\right) + \Phi_3(t), \tag{3.36}$$

where $\Phi_1(\cdot)$, $\Phi_2(\cdot)$ and $\Phi_3(\cdot)$ are deterministic differentiable functions. See Hafayed [82, 83], Li [60], and Anderson, Djehiche [59] for other models of conjecture.

Applying Itô's formula to (3.36) we get

$$\begin{aligned} d\Psi^*(t) &= d(\Phi_1(t)\,x^*(t)) + d(\Phi_2(t)\,\mathbb{E}\left(x^*(t)\right)) + d\Phi_3(t) \\ &= \Phi_1(t)\,dx^*(t) + x^*(t)\Phi_1'(t)\,dt + \Phi_2(t)\,d\mathbb{E}\left(x^*(t)\right) + \mathbb{E}\left(x^*(t)\right)\Phi_2'(t)\,dt \\ &+ \Phi_3'(t)\,dt. \end{aligned}$$

Since $d\mathbb{E}\left(x^*(t)\right) = [(A+\widetilde{A})\mathbb{E}\left(x^*(t)\right) + B\mathbb{E}\left(u^*(t)\right)]dt$, we get

$$\begin{aligned} d\Psi^*(t) = &\Big\{\Phi_1(t)\left[Ax^*(t) + \widetilde{A}\mathbb{E}\left(x^*(t)\right) + Bu^*(t)\right)] \\ &+ \Phi_2(t)\left[(A+\widetilde{A})\mathbb{E}\left(x^*(t)\right) + B\mathbb{E}\left(u^*(t)\right)\right] \\ &+ x^*(t)\Phi_1'(t) + \mathbb{E}\left(x^*(t)\right)\Phi_2'(t) + \Phi_3'(t)\}\,dt \\ &+ \textstyle\sum_{j=1}^{d}[C^j x^*(t) + \widetilde{C}^j\mathbb{E}\left(x^*(t)\right) + D^j u^*(t)]\Phi_1(t)\,dW^j(t), \\ &+ \textstyle\sum_{j=1}^{\infty}\left[\varrho^j x^*(t) + \widetilde{\varrho}^j\mathbb{E}\left(x^*(t)\right) + F^j u^*(t)\right]\Phi_1(t)\,dH^j(t), \\ \Psi^*(T) = &\ \Phi_1(T)\,x^*(T) + \Phi_2(T)\,\mathbb{E}\left(x^*(T)\right) + \Phi_3(T). \end{aligned} \tag{3.37}$$

From (3.33) and (3.37) we have $\Phi_3(t) \equiv 0$, $\forall t \in [0,T]$, and

$$\begin{cases} \Phi_1(t)\left[Ax^*(t)+\widetilde{A}\mathbb{E}(x^*(t))+Bu^*(t))\right]+\Phi_2(t)\left[(A+\widetilde{A})\mathbb{E}(x^*(t))+B\mathbb{E}(u^*(t))\right] \\ +x^*(t)\Phi_1'(t)+\mathbb{E}(x^*(t))\Phi_2'(t) \\ =-[A\Psi^*(t)+\widetilde{A}\mathbb{E}(\Psi^*(t))+(CQ^*(t)+\widetilde{C}\mathbb{E}(Q^*(t))) \\ +(\varrho K^*(t)+\widetilde{\varrho}\mathbb{E}[K^*(t)])+Rx^*(t)], \end{cases} \tag{3.38}$$

$$Q^*(t)=[Cx^*(t)+\widetilde{C}\mathbb{E}(x^*(t))+Du^*(t)]\Phi_1(t), \tag{3.39}$$

$$K^*(t)=[\varrho x^*(t)+\widetilde{\varrho}\mathbb{E}(x^*(t))+Fu^*(t)]\Phi_1(t). \tag{3.40}$$

By comparing the coefficient of $x^*(t)$ and $\mathbb{E}(x^*(t))$ in equation (3.38) and last equation in (3.37) (terminal condition) we immediately deduce that $\Phi_1(\cdot)$, $\Phi_2(\cdot)$ are given by the following ordinary differential equations (ODEs in short)

$$\begin{cases} \Phi_1'(t)+(2A+C^2+\varrho^2)\Phi_1(t)+R=0,\ \ \Phi_1(T)=\Pi, \\ \Phi_2'(t)+2(A+\widetilde{A})\Phi_2(t)+(2\widetilde{A}+\widetilde{C}^2+\widetilde{\varrho}^2+2(C\widetilde{C}+\varrho\widetilde{\varrho}))\Phi_1(t)=0 \\ \Phi_2(T)=0. \end{cases} \tag{3.41}$$

By solving the ODEs (3.41) we obtain

$$\Phi_1(t)=-R\left(2A+C^2+\varrho^2\right)^{-1}+\left[\Pi+R\left(2A+C^2+\varrho^2\right)^{-1}\right]\exp\left\{(2A+C^2+\varrho^2)(T-t)\right\}$$
$$\Phi_2(t)=(2\widetilde{A}+\widetilde{C}^2+\widetilde{\varrho}^2+2(C\widetilde{C}+\varrho\widetilde{\varrho}))\exp\left\{-2(A+\widetilde{A})t\right\}\int_t^T\Phi_1(s)\exp\left\{2(A+\widetilde{A})s\right\}ds.$$

Finally, by combining Theorem 3.3.1 and Theorem 3.4.1 we give the explicit optimal control in feedback form involving both $x^*(t)$ and $\mathbb{E}(x^*(t))$.

Theorem 3.5.1 *The optimal control* $u^*(\cdot) \in \mathcal{U}_{\mathcal{G}}([0,T])$ *for the mean-field linear quadratic* *control problem* (3.31)-(**??**) *is given in feedback form by*

$$u^*(t, x^*(t), \mathbb{E}(x^*(t))) = -\frac{1}{N}\left\{B\mathbb{E}[\Psi^*(t) \mid \mathcal{G}_t] + D\mathbb{E}[Q^*(t) \mid \mathcal{G}_t] + F\mathbb{E}[K^*(t) \mid \mathcal{G}_t]\right\}.$$

3.6 Conclusions

In this chaptre, under partial information, optimal control problem for mean-field stochastic differential equations driven by Lévy process has been discussed. Necessary and sufficient conditions of optimal control are established. As an illustration, using these results, linear quadratic control problem (time-inconsistent solution) has been studied. Apparently, there are many problems left unsolved. To mention a few, necessary and sufficient conditions for mean-field nonlinear controlled forward-backward stochastic systems governed by Teugels martingales associated with some Lévy process.

Chapter-IV

On partial-information optimal singular control problem for mean-field stochastic differential equations driven by Teugels martingales measures

Chapter 4

On optimal singular control for mean-field SDEs driven by Teugels martingales measures under partial information

Abstract. This work is concerned with partial-information mixed optimal stochastic continuous-singular control problem for mean-field stochastic differential equation driven by Teugels martingales associated with some Lévy processes and an independent Brownian motion. The control variable has two components; the first being absolutely continuous, and the second singular. Partial-information necessary and sufficient conditions of optimal continuous-singular control for these mean-field models are investigated. The control domain is assumed to be convex. As an illustration, this work studies a partial-information linear quadratic control problem of mean-field type involving continuous-singular control.

4.1 Introduction

Stochastic control problems related to Lévy processes and Teugels martingales are an important and challenging class of problems in control theory. These appear in various fields like mathematical finance, problem of optimal consumption, etc. A number of results have been obtained for these types of problems, see Meng, Tang [90]; Meng, Zhang and Tang [91]; Mitsui andTabata [98]; Tang and Zhang [100]; Tang and Wu [99], and references therein. Under partial-information, the necessary and sufficient optimality conditions for stochastic differential equations (SDEs), driven by Teugels martingales and an independent multi-dimensional Brownian motion have been proved by using convex perturbation, see Meng and Tang [90]. Partial-information optimal control problems for backward stochastic differential equations (BSDEs), and for forward-backward stochastic differential equations (FBSDEs) associated with Lévy processes have been investigated in Meng, Zhang and Tang [91]; Tang and Zhang [100]. The stochastic linear-quadratic problems with Lévy processes have been studied by Mitsui and Tabata [98] and Tang and Wu [99].

Mean-field stochastic control theory has been an active area of research and a useful tool in many applications, particularly in biology, game theory, economics and finances. A general mean-field maximum principle for SDEs was obtained by using spike variational method, see Buckdahn, Djehiche and Li [6]. A mean-field type stochastic maximum principle for Risk-Sensitive control has been proved by Djehiche, Tembine and Tempone [79]. For decentralized tracking-type games for large population multi-agent systems with mean-field coupling, we refer to Li and Zhang [50], and for discrete time mean-field stochastic linear-quadratic optimal control problems with applications, we refer to Elliott, Li and Ni [80]. Under complete information, second order necessary and sufficient conditions of near-optimal singular control for mean-field SDE have been established in Hafayed and Abbas [81]. Mean-field type stochastic maximum principle for optimal singular control has been studied in Hafayed [82], in which convex perturbations used for both absolutely continuous and singular components. The maximum principle for optimal control of mean-field FBSDEJs has been studied in Hafayed [103]. The necessary and sufficient conditions for near-optimality for mean-field jump diffusions with applications have been derived by Hafayed, Abba and Abbas [104]. Singular optimal control for mean-field forward-backward stochastic systems with applications to finance has been investigated in Hafayed [83]. Second-order maximum principle for optimal stochastic control for mean-field jump diffusions was proved in Hafayed and Abbas [106]. For singular mean-field control games with applications to optimal harvesting and investment problems, we refer to Hu, Øksendal and Sulem [88], and for mean-field games for large population multiagent systems with Markov jump parameters, we refer to Wang and Zhang [49]. Various forms of necessary and sufficient optimality conditions, for systems of SDEs with jumps with their applications have been studied in Shen, Meng and Shi [108]; Shen and Siu [107]; Meng and Yang [93]. Special attention has been paid to applying the maximum principle to mean-field linear quadratic control problems, see Ni, Zhang and Li [94]; Yong [96] and the references therein.

Partial-information or incomplete information means that the information available to the controller is possibly less than the whole information. That is, any admissible control is adapted to a subfiltration $(\mathcal{G}_t)_t$ of $(\mathcal{F}_t)_t$ $t \geq 0$. This kind of problem, which has potential applications in mathematical finance and mathematical economics, arises naturally, because it may fail to obtain an admissible control with full information in real world applications. Under partial-information, mean-field type stochastic maximum principle for optimal control has been investigated by Wang, Zhang and Zhang [51]. Stochastic maximum principles for partially observed mean-field stochastic systems with application has been investigated by Wang, Wu and Zhang [95].

The singular stochastic control problems have received considerable research attention in recent years due to wide applicability in a number of different areas, see Alvarez [77]; Cadenillas and Haussmann [9]; Dufour and Miller [78]; Hafayed and Abbas [81]; Hafayed [82]; Haussmann and Suo [84], and the list of references therein. In most classical cases, the optimal singular control problem was investigated through dynamic programming principle.

The first version of stochastic maximum principle for singular control was obtained, where the coefficient of SDEs are random, see Cadenillas and Haussmann [9]. In Dufour and Miller [78], the authors derived stochastic maximum principle where the singular part has a linear form. The maximum principle for mixed regular-singular stochastic control of FBSDEs have been proved by using the approach of relaxed controls, where the set of regular controls is not necessarily convex and the regular control enters the diffusion coefficient, see Zhang [97].

The mixed continuous-singular control problems in stochastic systems with jumps have been studied by only a few researchers. A maximum principle for singular stochastic control problems and optimal stopping with partial-information of Itô–Lévy processes have been studied by using Malliavin calculus, see ∅ksendal and Sulem [89]. For some cases of mixed singular-jump control problems when the payoff functional does not depend explicitly on the control, see Menaldi and Rebin [76]. Necessary and sufficient conditions for near-optimal mixed singular jump control have been proved by using Ekeland's variational principle, see Hafayed and Abbas [73].

Our main goal in this work is to derive partial-information necessary and sufficient conditions of optimal stochastic continuous-singular control in the form of a stochastic maximum principle, where the system is governed by mean-field controlled SDE, driven by Teugels martingales associated with some Lévy processes and an independent Brownian motion. The coefficients of the system and the cost functional depend not only on the state process but also on its marginal law of the state process through its expected value. The mean-field mixed continuous-singular control problem under consideration is not a simple extension from the mathematical point of view. It also provide interesting models in many applications such as mathematical finance, where the singular components of the control means the interventions. As an illustration, linear quadratic control problem of mean-field type involving continuous-singular control is discussed, where the optimal control is given in feedback form. Note that in our mean-field control problem, there are two types of jumps for the state processes, the inaccessible ones which come from the Lévy martingale part and the predictable ones which come from the singular control part.

The rest of the work is structured as follows. Section 2 begins with general formulation of the mean-field mixed control problem and gives the notations, assumptions and some basic definitions used throughout the work. In Sections 3 and 4, respectively, we derive necessary and sufficient conditions for optimality. An application is discussed in section 5. Finally, some discussions with concluding remarks are given in the last section.

4.2 Formulation of the problem

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a fixed filtered probability space equipped with a $\mathbb{P}-$completed right continuous filtration on which a $d-$dimensional Brownian motion $W = (W(t))_t : t \in [0,T]$ is defined. Let $(\mathcal{G}_t)_t$ a subfiltration of $(\mathcal{F}_t)_t$. For example $\mathcal{G}_t$ could be the δ-delayed information defined by $\mathcal{G}_t = \mathcal{F}_{(t-\delta)^+} : t \geq 0$. Let $L(\cdot) = (L(t))_{t\in[0,T]}$ be a $\mathbb{R}$-valued Lévy process, independent of the Brownian motion $W(\cdot)$, and of the form $L(t) = bt + \lambda(t)$, where $\lambda(\cdot)$ is a pure jump process. Assume that the Lévy measure $\mu(d\theta)$ corresponding to the Lévy process $\lambda(\cdot)$ satisfies $\int_{\mathbb{R}}(1 \wedge \theta^2)\mu(d\theta) < \infty$, and for every $\delta > 0$: there exist $\gamma > 0$ such that

$$\int_{]-\delta,\delta[} \exp\left(\gamma \left|\theta\right|\right) \mu(d\theta) < \infty.$$

Let $\mathcal{F}_t$ be $\mathbb{P}-$augmentation of the natural filtration $\mathcal{F}_t^{(W,L)}$ defined as follows: for $t \in [0,T]$

$$\mathcal{F}_t^{(W,L)} \triangleq \mathcal{F}_t^W \vee \sigma\left\{L(s) : 0 \leq s \leq t\right\} \vee \mathcal{F}_0,$$

where $\mathcal{F}_t^W = \sigma\left\{W(s) : 0 \leq s \leq t\right\}$, $\mathcal{F}_0$ denotes the totality of $\mathbb{P}-$null sets, and $\mathcal{F}_1 \vee \mathcal{F}_2$ denotes the σ-field generated by $\mathcal{F}_1 \cup \mathcal{F}_2$. Since the purpose of this work is to study optimal singular-continuous stochastic control for mean-field systems, we give here the precise definition of the singular part of an admissible control.

Consider the following sets: $\mathbb{A}_1$ is a nonempty convex subset of $\mathbb{R}^k$ and $\mathbb{A}_2 \triangleq ([0,\infty[)^m$.

Definition 4.1.1. An admissible control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathcal{G}_t-$adapted processes, such that $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the left with right limits and $\xi(0_-) = 0$. Moreover,

$$\mathbb{E}(\sup_{0\leq t\leq T} |u(t)|^2 + |\xi(T)|^2) < \infty.$$

Note that the jumps of a singular control $\xi(\cdot)$ at any jumping time t is denoted by

$$\Delta\xi(t) \triangleq \xi(t) - \xi(t_-).$$

Let us define the continuous part of the singular control by

$$\xi^{(c)}(t) \triangleq \xi(t) - \sum_{0\leq\tau_j\leq t} \Delta\xi(\tau_j),$$

i.e., the process obtained by removing the jumps of $\xi(t)$.

We denote $\mathcal{U}^1_{\mathcal{G}}\times\mathcal{U}^2_{\mathcal{G}}\left([0,T]\right)$, the set of all admissible controls. Since $d\xi(t)$ may be singular with respect to Lebesgue measure dt, we call $\xi(\cdot)$ the singular part of the control and the process $u(\cdot)$ its absolutely continuous part.

Remark 4.1.1. (*Jumps caused by the singular control and by the Lévy martingales*)

Throughout this work, we distinguish between the jumps caused by the singular control $\xi(\cdot)$ at any jumping time t defined by

$$\Delta_\xi x^{u,\xi}(t) \triangleq \mathcal{C}(t)\Delta\xi(t) = \mathcal{C}(t)(\xi(t) - \xi(t_-)),$$

and the jumps of $x^{u,\xi}(t)$ caused by the Lévy martingales, where $\Delta L(t) = L(t) - L(t_-)$ and the power jump processes is defined by

$$\begin{cases} L_{(k)}(t) \triangleq \sum_{0<\tau\leq t} (\Delta L(\tau))_k : k > 1 \\ L_{(1)}(t) \triangleq L(t). \end{cases}$$

Moreover, we define the continuous part of the control by

$$L_{(k)}{}^{(c)}(t) \triangleq L_{(k)}(t) - \sum_{0<\tau\leq t} (\Delta L(\tau))_k : k > 1,$$

i.e., the process obtained by removing the jumps of $L(t)$. If we define

$$N_{(k)}(t) \triangleq L_{(k)}(t) - \mathbb{E}\left\{L_{(k)}(t)\right\} : k \geq 1,$$

then the family of Teugels martingales $(H_j(\cdot))_{j\geq 1}$ is defined by $H_j(t) = \sum_{1<k\leq j} \alpha_{jk} N_k(t)$, where the coefficients α_{jk} associated with the orthonormalization of the polynomials $\{1, x, x^2, ...\}$ with respect to the measure $m(dx) = x^2\mu(dx)$. The Teugels martingales $(H_j(\cdot))_{j\geq 1}$ are pathwise strongly orthogonal and their predictable quadratic variation processes are given by $\langle H_i(t), H_j(t)\rangle = \delta_{ij}t$. The jumps of $x^{u,\xi}(t)$ caused by the Lévy martingales $\Delta_L x^{u,\xi}(t)$ is defined by

$$\Delta_L x^{u,\xi}(t) \triangleq g(t, x^{u,\xi}(t_-), \mathbb{E}(x^{u,\xi}(t_-)), u(t))\Delta L(t).$$

The general jump of the state processes $x^{u,\xi}(\cdot)$ at any jumping time t is given by

$$\Delta x^{u,\xi}(t) \triangleq x^{u,\xi}(t) - x^{u,\xi}(t_-) \triangleq \Delta_\xi x^{u,\xi}(t) + \Delta_{_L} x^{u,\xi}(t).$$

In this chapter, we consider the following mean-field continuous-singular stochastic control problem under partial-information, where the system is governed by a mean-field SDEs driven by Teugels martingales, associated with some Lévy processes and an independent Brownian motion, whose state equation is related to a kind of McKean-Vlasov equation of the type

$$\begin{cases} dx^{u,\xi}(t) = f\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)\right) dt \\ + \sum_{j=1}^{d} \sigma^j \left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)\right) dW^j(t) \\ + \sum_{j=1}^{\infty} g^j \left(t, x^{u,\xi}(t_-), \mathbb{E}(x^{u,\xi}(t_-)), u(t)\right) dH^j(t) + \mathcal{C}(t) d\xi(t), \\ x^{u,\xi}(0) = x_0, \end{cases} \tag{4.1}$$

where f, σ, g and $\mathcal{C}$ are given maps and the initial condition x_0 is an $\mathcal{F}_0-$measurable random variable. The mean-field SDEs-(4.1) may be obtained as a limit approach, by the mean-square limit, as n goes to infinity of a system of interacting particles of the form:

$$\begin{aligned} dx_n^{u,\xi,j}(t) &= f(t, x_n^{u,\xi,j}(t), m_n^{u,\xi}(t), u(t))dt \\ &\quad + \sum_{k=1}^{d} \sigma^k(t, x_n^{u,\xi,j}(t), m_n^{u,\xi}(t), u(t))dW^{k,j}(t) \\ &\quad + \sum_{k=1}^{\infty} g^k(t, x_n^{u,\xi,j}(t_-), m_n^{u,\xi}(t_-), u(t))dH^{k,j}(t) \\ &\quad + \mathcal{C}(t)d\xi(t), \end{aligned}$$

where $m_n^{u,\xi}(t) = \frac{1}{n}\sum_{i=1}^{n} x_n^{u,\xi,i}(t)$, $W(\cdot)$ is a standard $d-$dimensional Brownian motion and $H(t) = (H^j(t))_{j\geq 1}$ are pairwise strongly orthonormal Teugels martingales, associated with some Lévy processes, having moments of all orders, and $\xi(\cdot)$ is the singular part of the control, which is called intervention control. The continuous control $u(\cdot) = (u(t))_{t\geq 0}$ is required to be valued in some subset of $\mathbb{R}^k$ and adapted to a subfiltration $(\mathcal{G}_t)_{t\geq 0}$. In some finance models, the mean-field term $\mathbb{E}(x^{u,\xi}(t))$ represents an approximation to the weighted average $\frac{1}{n}\sum_{i=1}^{n} x_n^{u,\xi,i}(t)$ for large n, $\xi(t)$ representing the harvesting effort, while $\mathcal{C}(t)$ is a given harvesting efficiency coefficient.

The expected cost on the time interval $[0, T]$ is defined by

$$J\left(u(\cdot), \xi(\cdot)\right) \triangleq \mathbb{E}\Big\{ \int_0^T \ell(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t))dt \\ + h(x^{u,\xi}(T), \mathbb{E}(x^{u,\xi}(T))) + \int_{[0,T]} \mathcal{M}(t)d\xi(t) \Big\}, \tag{4.2}$$

where ℓ, h and $\mathcal{M}$ are an appropriate functions. This cost functional is also of mean-field type, as the functions ℓ and h depend on the marginal law of the state process through its expected value. It worth mentioning that since the cost functional J is possibly a nonlinear function of the expected value stands in contrast to the standard formulation of a control problem. This leads to so called time-inconsistent control problem where the Bellman dynamic programming does not hold. The reason for this is that one cannot apply the law of iterated expectations on the cost functional. Noting that the partial-information mixed control problem (4.1)-(4.2) occur naturally in the probabilistic analysis of financial optimization problems. Moreover, the above mathematical mean-field approaches play an important role in different fields of game theory, economics and finance, where the objective of the controller is to choose a couple $(u^*(\cdot), \xi^*(\cdot))$ of adapted processes, in order to minimize the performance functional.

Problem. Find an admissible control $(u^*(\cdot), \xi^*(\cdot)) \in \mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2\left([0, T]\right)$ such that

$$J\left(u^*(\cdot), \xi^*(\cdot)\right) = \inf_{(u(\cdot),\xi(\cdot)) \in \mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2([0,T])} J\left(u(\cdot), \xi(\cdot)\right). \tag{4.3}$$

The admissible control $(u^*(\cdot), \xi^*(\cdot))$ satisfying (4.3) is called an optimal control. The corresponding optimal state process, solution of mean-field system (4.1) is denoted by $x^*(\cdot) = x^{u^*,\xi^*}(\cdot)$.

Notations. We will use the following notations in this work.

- 1. The set $\mathbb{R}^n$ denotes the $n-$dimensional Euclidean space, l^2 denotes the Hilbert space of real-valued sequences $x = (x_n)_{n\geq 0}$ such that

$$\|x\|_{l^2} \triangleq \left[\sum_{n=1}^{\infty} x_n\right]^2 < \infty.$$

- 2. $l^2(\mathbb{R}^n)$ denotes the space of $\mathbb{R}^n$-valued $(f_n)_{n\geq 1}$ such that

$$\left[\sum_{n=1}^{\infty} \|f_n\|_{\mathbb{R}^n}^2\right]^{\frac{1}{2}} < \infty.$$

- 3. $l^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ denotes the Banach space of $\mathcal{F}_t-$adapted processes such that

$$\mathbb{E}\left(\int_0^T |x(t)|_{\mathbb{R}^n}^2\, dt\right)^{\frac{1}{2}} < \infty.$$

- 4. $\mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ denotes the Banach space of $\mathcal{F}_t-$predictable processes such that

$$\mathbb{E}\left(\int_0^T \sum_{n=1}^{\infty} \|f_n\|_{\mathbb{R}^n}^2\, dt\right)^{\frac{1}{2}} < \infty.$$

- 5. $S^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ denotes the Banach space of $\mathcal{F}_t-$adapted and *cadlag* processes such that

$$\mathbb{E}(\sup |x(t)|^2)^{\frac{1}{2}} < \infty.$$

- 6. $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^n)$ denotes the Banach space of $\mathbb{R}^n$-valued, square integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

- 7. $\mathbb{M}^{n\times m}(\mathbb{R})$ denotes the space of $n \times m$ real matrices.

- 8. For a differentiable function f, we denote by $f_x(t)$ its gradient with respect to the variable x.

- 9. We denote by I_A the indicator function of A and by $\mathbb{E}(\cdot \mid \cdot)$ the conditional expectation.

Assumptions. Throughout this work, we assume the following:

$$\begin{aligned} f &: [0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{A}_1\rightarrow\mathbb{R}^n, \\ \sigma &: [0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{A}_1\rightarrow\mathbb{M}^{n\times m}(\mathbb{R}), \\ g &: [0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{A}_1\rightarrow l^2\left(\mathbb{R}^n\right), \\ \ell &: [0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{A}_1\rightarrow\mathbb{R}, \\ h &: \mathbb{R}^n\times\mathbb{R}^n\rightarrow\mathbb{R}, \\ \mathcal{C} &: [0,T]\rightarrow\mathbb{M}^{n\times m}(\mathbb{R}), \\ \mathcal{M} &: [0,T]\rightarrow\left([0,\infty)\right)^m. \end{aligned}$$

Assumption (C1) The functions f,σ,ℓ,g and h are continuously differentiable in their variables including (x,y,u). The maps f,σ,g are progressively measurable processes such that $f(\cdot,0,0,0)$, $g(\cdot,0,0,0)\in\mathbb{L}^2_{\mathcal{F}}\left([0,T];\mathbb{R}^n\right)$ and $\sigma(\cdot,0,0,0)\in\mathbb{M}^2_{\mathcal{F}}\left([0,T];\mathbb{R}^n\right)$.

Assumption (C2) The derivatives of f,σ and g with respect to their variables including (x,y,u) are bounded. Further the map ℓ are dominated by $C(1+x^2+y^2+u^2)$ and its derivatives with respect to (x,y,u) are dominated by $C\left(1+|x|+|y|+|u|\right)$. The map h is dominated by $C\left(1+x^2+y^2\right)$ and its derivatives with respect to (x,y) are dominated by $C(1+|x|+|y|)$.

Assumption (C3) The functions $\mathcal{C}$ and $\mathcal{M}$ are $\mathcal{F}^W_t-$adapted, continuous and bounded.

From *Lemma 1* in Meng and Tang [90], and under assumptions (H1)$-$(H3), the mean-field

SDE-(4.1) has a unique solution $x^{u,\xi}(\cdot) \in \mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ such that

$$\begin{aligned} x^{u,\xi}(t) &= x_0 + \int_0^t f(s, x^{u,\xi}(s), \mathbb{E}(x^{u,\xi}(s)), u(s))ds \\ &+ \int_0^t \sum_{j=1}^d \sigma^j(s, x^{u,\xi}(s), \mathbb{E}(x^{u,\xi}(s)), u(s)))dW^j(s) \\ &+ \int_0^t \sum_{j=1}^\infty g^j(s, x^{u,\xi}(s), \mathbb{E}(x^{u,\xi}(s)), u(s))dH^j(s) \\ &+ \int_{[0,t]} \mathcal{C}(s)d\xi(s). \end{aligned}$$

Adjoint equation. We introduce the adjoint equations involved in the stochastic maximum principle for our mean-field mixed continuous-singular control problem (4.1)-(4.2). For simplicity of notation, we will still use $f_x(t) = \frac{\partial f}{\partial x}(t, x^{u,\xi}(\cdot), \mathbb{E}(x^{u,\xi}(\cdot)), u(\cdot))$, etc. So for any admissible control $(u(\cdot), \xi(\cdot)) \in \mathcal{U}^1_{\mathcal{G}} \times \mathcal{U}^2_{\mathcal{G}}([0,T])$ and the corresponding state trajectory $x(\cdot) = x^{u,\xi}(\cdot)$, we consider the following adjoint equations of mean-field type, which are independent to singular control

$$\begin{cases} -d\Psi(t) = [f_x(t)\Psi(t) + \mathbb{E}(f_y(t)\Psi(t)) + \sum_{j=1}^d[\sigma^j_x(t)Q^j(t) + \mathbb{E}(\sigma^j_y(t)Q^j(t))] \\ \qquad + \sum_{j=1}^\infty(g^j_x(t)K^j(t) + \mathbb{E}(g^j_y(t)K^j(t))) + \ell_x(t) + \mathbb{E}(\ell_x(t))]dt \\ \qquad - \sum_{j=1}^d Q^j(t)dW(t) - \sum_{j=1}^\infty K^j(t)dH^j(t) \\ \Psi(T) = h_x(x(T), \mathbb{E}(x(T))) + \mathbb{E}(h_y(x(T), \mathbb{E}(x(T)))). \end{cases} \tag{4.4}$$

Hamiltonian function. We define the Hamiltonian function $\mathcal{H} : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{A}_1 \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times l^2(\mathbb{R}^n) \to \mathbb{R}^n$, associated with the mean-field stochastic control problem (4.1)-(4.2) as follows:

$$\begin{aligned} \mathcal{H}(t,x,y,u,\Psi(\cdot),Q(\cdot),K(\cdot)) &\triangleq \Psi(t)f(t,x,y,u) + \textstyle\sum_{j=1}^d Q^j(t)\sigma(t,x,y,u) \\ &+ \textstyle\sum_{j=1}^\infty K^j(t)g^j(t,x,y,u) + \ell(t,x,y,u). \end{aligned} \tag{4.5}$$

If we denote by $\mathcal{H}(t)=\mathcal{H}(t,x(t),y(t),u(t),\Psi(t),Q(t),K(t))$, then the adjoint equation (4.4) can be rewritten in terms of the derivative of the Hamiltonian as

$$\begin{cases} -d\Psi(t)=\left\{\mathcal{H}_x\left(t\right)+\mathbb{E}\left[\mathcal{H}_y\left(t\right)\right]\right\}dt \\ \qquad\qquad -\sum_{j=1}^{d}Q^j(t)dW(t)-\sum_{j=1}^{\infty}K^j(t)dH^j(t), \\ \Psi(T)=h_x\left(x(T),\mathbb{E}(x(T))\right)+\mathbb{E}[h_y\left(x(T),\mathbb{E}(x(T))\right)]. \end{cases} \tag{4.6}$$

Since the derivatives of f,σ,g,h and ℓ with respect to (x,y) are bounded, by assumptions (C1)-(C2), the mean-field BSDEs-(4.4) and (4.6) admits a unique solution $(\Psi(\cdot),Q(\cdot),K(\cdot))\in\mathbb{S}^2_{\mathcal{F}}([0,T];\mathbb{R}^n)\ \times\mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R}^{n\times d})\ \times l^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$, such that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\Psi(t)\right|^2+\sup_{t\in[0,T]}\left|K(t)\right|^2+\int_0^T\left|Q(t)\right|^2dt\right]<C.$$

4.3 Necessary conditions for optimal continuous-singular control for mean-field SDEs driven by Teugels martingales

In this section, inspired by Meng and Tang [90], we establish partial-information mean-field type necessary conditions for optimal stochastic continuous-singular control, where the system evolves according to controlled mean–field SDEs-(4.1), driven by Teugels martingales associated with some Lévy processes and an independent Brownian motion. In addition to the assumptions in Section 2, we now assume the following:

Assumptions (C4)

(1) For all t, r such that $0\leq t\leq t+r\leq T$, all $i=1,...,k$ and all bounded $\mathcal{G}_t-$measurable $\alpha=\alpha(w)$, the control $\beta(t)=(0,...,0,\beta_i(t),0,...,0)\in\mathbb{A}_1$, with $\beta_i(s)=\alpha_iI_{[t,t+r]}(s)$, $s\in[0,T]$ belong to $\mathcal{U}^1_{\mathcal{G}}\left([0,T]\right)$.

(2) For all $u(\cdot),\beta\left(\cdot\right)\in\mathcal{U}^1_{\mathcal{G}}([0,T])$ with $\beta\left(\cdot\right)$ bounded, there exist $\delta_1>0$ such that $u(\cdot)+\theta\beta\in$

$\mathcal{U}_{\mathcal{G}}^{1}([0,T])$ for all $\theta \in [0,\delta_1]$.

(3) For $\xi(\cdot) \in \mathcal{U}_{\mathcal{G}}^{2}([0,T])$, we let $\mathcal{V}(\xi)$ denote the set of $\mathcal{G}_t$-adapted processes ζ of finite variation such that there exist $\delta_2 = \delta_2(\xi) > 0$ such that $\xi(\cdot) + \theta\zeta \in \mathcal{U}_{\mathcal{G}}^{2}([0,T])$ for all $\theta \in [0,\delta_2]$.

Now, let $\delta = \min(\delta_1,\delta_2)$, for all $\theta \in [0,\delta]$ and for a given $u(\cdot),\beta \in \mathcal{U}_{\mathcal{G}}^{1}([0,T])$ and $\xi(\cdot) \in \mathcal{U}_{\mathcal{G}}^{2}([0,T])$ with β bounded, $\zeta \in \mathcal{V}(\xi)$, we define the process $Z^{u,\xi}(\cdot)$ by

$$Z^{u,\xi}(t) = Z^{u(\cdot),\xi(\cdot),\beta,\zeta}(t) \triangleq \frac{d}{d\theta}(x^{u^*+\theta\beta,\xi^*+\theta\zeta}(t)) \mid_{\theta=0}. \tag{4.7}$$

Note that the process $Z^{u,\xi}(\cdot)$ satisfies the following mean-field linear SDEs driven by both Brownian motion and Teugels martingales:

$$\begin{cases} dZ^{u,\xi}(t) = [f_x(t)Z^{u,\xi}(t) + f_y(t)\mathbb{E}(Z^{u,\xi}(t)) + f_u(t)\beta(t)]dt + \sum_{j=1}^{d}[\sigma_x^j(t)Z^{u,\xi}(t) \\ \quad + \sigma_y^j(t)\mathbb{E}\left(Z^{u,\xi}(t)\right) + \sigma_u^j(t)\beta(t)]dW^j(t) + \sum_{j=1}^{\infty}[g_x^j(t)Z^{u,\xi}(t) + g_y^j(t)\mathbb{E}\left(Z^{u,\xi}(t)\right) \\ \quad + g_u^j(t)\beta(t)]dH^j(t) + \mathcal{C}(t)d\zeta(t), \\ Z^{u,\xi}(0) = 0, \end{cases}$$

The main result of this section is stated in the following theorem.

Let $(u^*(\cdot),\xi^*(\cdot))$ be a local minimum for the cost functional J over $\mathcal{U}_{\mathcal{G}}^{1} \times \mathcal{U}_{\mathcal{G}}^{2}([0,T])$ in the sense that for all bounded β and all $\zeta \in \mathcal{V}(\xi^*)$, there exist $\delta = \min(\delta_1,\delta_2) > 0$ such that $(u^*(\cdot)+\theta\beta, \xi^*(\cdot)+\theta\zeta) \in \mathcal{U}_{\mathcal{G}}^{1} \times \mathcal{U}_{\mathcal{G}}^{2}([0,T])$ for all $\theta \in [0,\delta]$ and a function φ defined by

$$\varphi(\theta) \triangleq J(u^*(\cdot)+\theta\beta, \xi^*(\cdot)+\theta\zeta),$$

is minimal at $\theta = 0$. Then it follows that

$$\frac{d}{d\theta}\varphi(\theta) \mid_{\theta=0} = \frac{d}{d\theta}J(u^*(\cdot)+\theta\beta, \xi^*(\cdot)+\theta\zeta) \mid_{\theta=0} = 0. \tag{4.8}$$

Let $x^*(\cdot)$ be the solution of the mean-field SDEs-(4.1) corresponding to $(u^*(\cdot),\xi^*(\cdot))$.

Theorem 4.3.1. (Partial-information mean-field necessary conditions). Let assumptions (C1)-(C4) hold. Then there exists a unique triplet of adapted process $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot))$ solution of adjoint equation (4.4) corresponding to $(u^*(\cdot), \xi^*(\cdot))$, such that $(u^*(\cdot), \xi^*(\cdot))$ is a critical point in the sense that for almost all $t \in [0,T]$, we have

$$\begin{aligned}
&\mathbb{E}[\mathcal{H}_u(t, x^*(t_-), \mathbb{E}(x^*(t_-)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)) \mid \mathcal{G}_t] \\
&+\mathbb{E}\left[\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t))\, d\xi^*(t) \mid \mathcal{G}_t\right] \\
&= 0, \quad a.e., \ t \in [0,T].
\end{aligned} \tag{4.9}$$

Proof: Let $Z^*(\cdot) = Z^{u^*,\xi^*}(\cdot)$. From (4.8) and (4.2), we have

$$\begin{aligned}
0 &= \frac{d}{d\theta} J(u^*(\cdot) + \theta\beta(\cdot), \xi^*(\cdot) + \theta\zeta(\cdot)) \mid_{\theta=0} \\
&= \mathbb{E}\Big\{ \int_0^T [\ell_x(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t)) Z^*(t) \\
&+\ell_y(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t)))\mathbb{E}(Z^*(t))]dt \\
&+ \int_0^T \ell_u(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t))\beta(t) dt \\
&+ \int_{[0,T]} \mathcal{M}(t) d\zeta(t) + h_x\left(x^*(T), \mathbb{E}\left(x^*(T)\right)\right) Z^*(T) \\
&+ h_y\left(x^*(T), \mathbb{E}\left(x^*(T)\right)\right) \mathbb{E}(Z^*(T)) \Big\}.
\end{aligned} \tag{4.10}$$

By applying Itô's formula to $\Psi^*(t)Z^*(t)$ and take expectation, we get

$$\begin{aligned}
&\mathbb{E}(\Psi^*(T)Z^*(T)) = \mathbb{E}\int_0^T \Psi^*(t) dZ^*(t) + \mathbb{E}\int_0^T Z^*(t) d\Psi^*(t) \\
&+\mathbb{E}\int_0^T \sum_{j=1}^d Q^{*j}(t)[\sigma_x^j(t) Z^*(t) + \sigma_y^j(t)\mathbb{E}\left(Z^*(t)\right) + \sigma_u^j(t)\beta(t)]dt \\
&+\mathbb{E}\int_0^T \sum_{j=1}^\infty K^{*j}(t)[g_x^j(t) Z^*(t) + g_y^j(t)\mathbb{E}\left(Z^*(t)\right) + g_u^j(t)\beta(t)]dt \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{4.11}$$

From (4.8), we obtain

$$\begin{aligned} I_1 &= \mathbb{E}\int_0^T \Psi^*(t)dZ^*(t) \\ &= \mathbb{E}\int_0^T \Psi^*(t)[f_x(t)\, Z^*(t) + f_y(t)\,\mathbb{E}(Z^*(t)) + f_u(t)\beta(t)]dt + \mathbb{E}\int_0^T \Psi^*(t)\mathcal{C}(t)d\zeta(t) \\ &= \mathbb{E}\int_0^T \Psi^*(t)f_x(t)\, Z^*(t)dt + \mathbb{E}\int_0^T \Psi^*(t)f_y(t)\,\mathbb{E}(Z^*(t))dt \\ &+ \mathbb{E}\int_0^T \Psi^*(t)f_u(t)\beta(t))dt + \mathbb{E}\int_{[0,T]} \Psi^*(t)\mathcal{C}(t)d\zeta(t). \end{aligned} \tag{4.12}$$

By applying (4.4) and (4.8), we have

$$\begin{aligned} I_2 &= \mathbb{E}\int_0^T Z^*(t)d\Psi^*(t) = -\mathbb{E}\int_0^T Z^*(t)\left\{f_x(t)\,\Psi^*(t) + \mathbb{E}\left(f_y(t)\,\Psi^*(t)\right)\right. \\ &+ \textstyle\sum_{j=1}^d (\sigma_x^j(t)\, Q^{j*}(t) + \mathbb{E}(\sigma_y^j(t)\, Q^{j*}(t))) + \sum_{j=1}^\infty (g_x^j(t)\, K^{j*}(t) + \mathbb{E}[g_y^j(t)\, K^{j*}(t)]) \\ &+ \left.\ell_x(t) + \mathbb{E}\left(\ell_y(t)\right)\right\} dt, \end{aligned} \tag{4.13}$$

$$\begin{aligned} I_3 &= \mathbb{E}\int_0^T \textstyle\sum_{j=1}^d Q^{*j}(t)[\sigma_x^j(t)Z^*(t) + \sigma_y^j(t)\mathbb{E}\left(Z^*(t)\right) + \sigma_u^j(t)\beta(t)]dt \\ &= \mathbb{E}\int_0^T \textstyle\sum_{j=1}^d Q^{*j}(t)\sigma_x^j(t)Z^*(t)dt + \mathbb{E}\int_0^T \sum_{j=1}^d Q^{*j}(t)\sigma_y^j(t)\mathbb{E}\left(Z^*(t)\right)dt \\ &+ \mathbb{E}\int_0^T \textstyle\sum_{j=1}^d Q^{*j}(t)\sigma_u^j(t)\beta(t)dt, \end{aligned} \tag{4.14}$$

and it follows easily by the same arguments that

$$\begin{aligned} I_4 &= \mathbb{E}\int_0^T \textstyle\sum_{j=1}^\infty K^{*j}(t)[g_x^j(t)Z^*(t) + g_y^j(t)\mathbb{E}\left(Z^*(t)\right) + g_u^j(t)\beta(t)]dt \\ &= \mathbb{E}\int_0^T \textstyle\sum_{j=1}^\infty K^{*j}(t)g_x^j(t)Z^*(t)dt + \mathbb{E}\int_0^T \sum_{j=1}^d K^{*j}(t)g_y^j(t)\mathbb{E}\left(Z^*(t)\right)dt \\ &+ \mathbb{E}\int_0^T \textstyle\sum_{j=1}^\infty K^{*j}(t)g_u^j(t)\beta(t)dt. \end{aligned} \tag{4.15}$$

By combining (4.11)−(4.15), together with (4.4) and the fact that $Z^{*}(0)=0$, we get

$$\begin{aligned}
&\mathbb{E}\{[h_x(x(T),\mathbb{E}(x(T))\\
&+\mathbb{E}(h_y(x(T),\mathbb{E}(x(T)))]Z^{*}(T)\}\\
&=\mathbb{E}\int_0^T\Big\{\Psi^{*}(t)f_u(t)\beta(t)dt\\
&+\sum_{j=1}^{d}Q^{j*}(t)\sigma_u^j(t)\beta(t)dt+\sum_{j=1}^{\infty}K^{j*}(t)g_u^j(t)\beta(t)\\
&-\ell_x(t)Z^{*}(t)-\mathbb{E}(\ell_y(t))Z^{*}(t)\Big\}dt\\
&+\mathbb{E}\int_{[0,T]}\Psi^{*}(t)\mathcal{C}(t)d\zeta(t).
\end{aligned}\tag{4.16}$$

From (4.10) and (4.16), we obtain

$$\begin{aligned}
&\mathbb{E}\int_0^T\Big\{\Psi^{*}(t)f_u(t)\beta(t)dt+\sum_{j=1}^{d}Q^{j*}(t)\sigma_u^j(t)\beta(t)dt\\
&+\sum_{j=1}^{d}K^{*j}(t)g_u^j(t)\beta(t)+\ell_u(t,x^{*}(t),\mathbb{E}(x^{*}(t)),u^{*}(t))\beta(t)\Big\}dt\\
&+\mathbb{E}\int_{[0,T]}[\mathcal{M}(t)+\mathcal{C}(t)\Psi^{*}(t)]d\zeta(t)=0.
\end{aligned}$$

Since

$$\begin{aligned}
&\mathcal{H}_u(t,x,y,u,\Psi(\cdot),Q(\cdot),K(\cdot))\\
&=\Psi(t)f_u(t,x,y,u)+\sum_{j=1}^{d}Q^j(t)\sigma_u(t,x,y,u)\\
&+\sum_{j=1}^{\infty}K^j(t)g_u^j(t,x,y,u)+\ell_u(t,x,y,u),
\end{aligned}$$

we get

$$\begin{aligned}
&\mathbb{E}\textstyle\int_0^T\mathcal{H}_u(t,x^{*}(t_-),y^{*}(t_-),u^{*}(t),\Psi^{*}(t),Q^{*}(t),K^{*}(t))\beta(t)dt\\
&+\mathbb{E}\textstyle\int_{[0,T]}[\mathcal{M}(t)+\mathcal{C}(t)\Psi^{*}(t)]d\zeta(t)=0.
\end{aligned}\tag{4.17}$$

We fix $0 \leq t \leq T$ and apply the above to $\beta = (0, ..., \beta_i, ..., 0)$, where $\beta_i(s) = \alpha_i I_{[t,t+r]}(s)$, $s \in [0, T]$, $t + r \leq T$ and $\alpha_i = \alpha_i(w)$ is bounded, $\mathcal{G}_t-$measurable. Then from (4.17), we get

$$\begin{aligned}&\mathbb{E}\int_t^{t+r} \frac{\partial}{\partial u_i}\mathcal{H}(s, x^*(s_-), y^*(s_-), u^*(s), \Psi^*(s), Q^*(s), K^*(s))\alpha_i(w)ds \\ &+\mathbb{E}\int_{[0,T]} [\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)]\, d\zeta(t) = 0.\end{aligned}$$

Now, differentiating the above equation with respect to r at $r = 0$, we obtain

$$\begin{aligned}&\mathbb{E}\left[\frac{\partial}{\partial u_i}\mathcal{H}(s, x^*(s_-), y^*(s_-), u^*(s), \Psi^*(s),\ Q^*(s), K^*(s))\alpha_i\right] \\ &+\mathbb{E}\int_{[0,T]} [\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)]\, d\zeta(t) = 0.\end{aligned} \tag{4.18}$$

Since (4.18) holds for all bounded $\mathcal{G}_t-$measurable α_i, and for all $\zeta \in \mathcal{V}(\xi^*)$, it is easy to show that

$$\begin{aligned}&\mathbb{E}\left[\mathcal{H}_u(t, x^*(t_-), y^*(t_-), u^*(t), \Psi^*(t), Q^*(t),\ K^*(t)) \mid \mathcal{G}_t\right] \\ &+\mathbb{E}\left[\int_{[0,T]} [\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)]\, d\xi^*(t) \mid \mathcal{G}_t\right] = 0, \quad \mathbb{P} - a.s,\end{aligned}$$

which completes the proof of *Theorem 4.3.1.* □

4.4 Sufficient conditions for optimal continuous-singular control for mean-field SDEs driven by Teugels martingales

The purpose of this section is to derive partial-information mean-field type sufficient conditions for optimal stochastic continuous-singular control, where the system evolves according to controlled mean–field SDEs-(4.1) driven by Teugels martingales associated with some Lévy processes and an independent Brownian motion. We prove that under some additional conditions, the maximality condition on the Hamiltonian function is a sufficient condition for optimality.

Assumptions (C5). We assume

1. The functional $\mathcal{H}(t,\cdot,\cdot,\cdot,\Psi^u(t),Q^u(t),K^u(t))$ is convex with respect to (x,y,u) for $a.e.t \in [0,T]$, $\mathbb{P}-a.s.$

2. The function $h(\cdot,\cdot)$ is convex with respect to (x,y).

Let $(\Psi^u(\cdot),Q^u(\cdot),K^u(\cdot))$ solution of adjoint equation (4.4) corresponding to $(u(\cdot),\xi(\cdot))$.

Now, we are able to state and prove the partial-information sufficient conditions of optimal continuous-singular mean-field control problem, which is the second main result of this work.

Theorem 4.4.2. (Partial-information mean-field sufficient conditions) Let assumptions (C1)-(C5) hold. Suppose that an admissible continuous-singular control $(u(\cdot),\xi(\cdot)) \in \mathcal{U}^1_{\mathcal{G}} \times \mathcal{U}^2_{\mathcal{G}}([0,T])$ satisfies

$$\begin{aligned}&\mathbb{E}\left[\mathcal{H}_u(t,x^{u,\xi}(t_-),\mathbb{E}(x^{u,\xi}(t_-)),u,\Psi^u(t),Q^u(t),\ K^u(t)) \mid \mathcal{G}_t\right]\\&+\mathbb{E}\left[\int_{[0,T]}\left(\mathcal{M}(t)+\mathcal{C}(t)\Psi^u(t)\right)d\xi(t)\mid \mathcal{G}_t\right]=0,\ a.e.,\ t\in[0,T].\end{aligned} \tag{4.19}$$

Then $(u(\cdot),\xi(\cdot))$ is a partial-information optimal control, i.e.,

$$J\left(u(\cdot),\xi(\cdot)\right)=\inf_{(v(\cdot),\eta(\cdot))\in\mathcal{U}^1_{\mathcal{G}}\times\mathcal{U}^2_{\mathcal{G}}([0,T])} J\left(v(\cdot),\eta(\cdot)\right). \tag{4.20}$$

The following Lemma gives the duality relations between $\Psi^u(t)$ and $(x^{v,\eta}(t)-x^{u,\xi}(t))$. It plays a key role in proving the sufficient optimality conditions (*Theorem 4.4.2.*)

Lemma 4.4.1. Let $x^{u,\xi}(\cdot)$ and $x^{v,\eta}(\cdot)$ be the solutions of the state equation (4.1) corresponding respectively to $(u(\cdot),\xi(\cdot))$ and $(v(\cdot),\eta(\cdot))$. Then they satisfy

$$\begin{aligned}
&\mathbb{E}\left[\Psi^u(T)\left(x^{v,\eta}(T)-x^{u,\xi}(T)\right)\right]\\
&=\mathbb{E}\int_0^T\Psi^u(t)\left[f(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t))\,-\,f(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))\right]dt\\
&+\mathbb{E}\int_0^T(\mathcal{H}_x(t)+\mathbb{E}(\mathcal{H}_y(t)))\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right)dt\\
&+\mathbb{E}\int_0^T\sum_{j=1}^d Q^{u,,j}(t)\left[\sigma^j(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t))\,-\,\sigma^j(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))\right]dt\\
&+\mathbb{E}\int_0^T\sum_{j=1}^\infty K^{u,j}(t)[g^j(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t))-g^j(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))]dt.\\
&+\mathbb{E}\int_{[0,T]}\mathcal{C}(t)\Psi^u(t)d(\eta-\xi)(t).
\end{aligned}\tag{4.21}$$

Remark 4.4.2. From *Lemme 4.4.1* and using the fact that

$$\begin{aligned}
&\mathcal{H}\left(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t),\Psi^u(t),Q^u(t),K^u(t)\right)\\
&=\Psi^u(t)f\left(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t)\right)\\
&+\sum_{j=1}^d Q^j(t)\sigma\left(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t)\right)\\
&+\sum_{j=1}^\infty K^j(t)g^j\left(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t)\right)\\
&+\ell\left(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t)\right),
\end{aligned}$$

we obtain

$$\begin{aligned}
&\mathbb{E}\left[\Psi^u(T)\left(x^{v,\eta}(T)-x^{u,\xi}(T)\right)\right]\\
&=\mathcal{H}\left(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t),\Psi^u(t),Q^u(t),K^u(t)\right)\\
&-\mathcal{H}\left(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t),\Psi^u(t),Q^u(t),K^u(t)\right)\\
&+\ell(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))\\
&-\ell(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t))\\
&+\mathbb{E}\int_0^T(\mathcal{H}_x(t)+\mathbb{E}(\mathcal{H}_y(t)))\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right)dt\\
&+\mathbb{E}\int_{[0,T]}\mathcal{C}(t)\Psi^u(t)d(\eta-\xi)(t).
\end{aligned}\tag{4.22}$$

Proof: First, from (4.1) and by simple computations, we get

$$\begin{aligned}
& d\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right)=[f(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t)) \\
& -f(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))]dt+\textstyle\sum_{j=1}^{d}[\sigma^j(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t)) \\
& -\sigma^j(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t)]dW^j(t)+\textstyle\sum_{j=1}^{\infty}[g^j(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t)) \\
& -g^j(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t)]dH^j(t)+\mathcal{C}(t)d(\eta-\xi)(t).
\end{aligned} \tag{4.23}$$

By applying integration by parts formula to $\Psi^u(t)\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right)$, we obtain

$$\begin{aligned}
& \mathbb{E}[\Psi^u(T)\left(x^{v,\eta}(T)-x^{u,\xi}(T)\right)] \\
& =\mathbb{E}\textstyle\int_0^T\Psi^u(t)d\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right)+\mathbb{E}\int_0^T\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right)d\Psi^u(t) \\
& +\mathbb{E}\textstyle\int_0^T\sum_{j=1}^{d}Q^{u,j}(t)[\sigma^j(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t)) \\
& -\sigma^j(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))]dt+\mathbb{E}\textstyle\int_0^T\sum_{j=1}^{\infty}K^{u,j}(t)[g^j(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t)) \\
& -g^j(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))]dt \\
& =I_1+I_2+I_3+I_4.
\end{aligned} \tag{4.24}$$

From (4.23), we obtain

$$\begin{aligned}
I_1 & =\mathbb{E}\textstyle\int_0^T\Psi^u(t)d\left(x^{v,\eta}(t)-x^{u,\xi}(t)\right) \\
& =\mathbb{E}\textstyle\int_0^T\Psi^u(t)[f(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t)) \\
& \quad -f(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t))]dt \\
& +\mathbb{E}\textstyle\int_{[0,T]}\mathcal{C}(t)\Psi^u(t)d(\eta-\xi)(t).
\end{aligned} \tag{4.25}$$

Similarly, from (4.6), we get

$$\begin{aligned}
I_2 &= \mathbb{E}\int_0^T \left(x^{v,\eta}(t) - x^{u,\xi}(t)\right) d\Psi^u(t) \\
&= \mathbb{E}\int_0^T \left(x^{v,\eta}(t) - x^{u,\xi}(t)\right) [\mathcal{H}_x(t) + \mathbb{E}(\mathcal{H}_y(t))]dt \\
&= \mathbb{E}\int_0^T \mathcal{H}_x(t) \left(x^{v,\eta}(t) - x^{u,\xi}(t)\right) dt \\
&+\mathbb{E}\int_0^T \mathbb{E}(\mathcal{H}_y(t))(x^{v,\eta}(t)) - x^{u,\xi}(t))dt.
\end{aligned} \tag{4.26}$$

By standard arguments, we obtain

$$\begin{aligned}
I_3 &= \mathbb{E}\int_0^T \textstyle\sum_{j=1}^d Q^u(t)[\sigma^j(t, x^{v,\eta}(t), \mathbb{E}(x^{v,\eta}(t)), v(t)) \\
&-\sigma^j(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t))]dt,
\end{aligned} \tag{4.27}$$

and a similar argument shows that

$$\begin{aligned}
I_4 &= \mathbb{E}\int_0^T \textstyle\sum_{j=1}^\infty K^u(t)[g^j(t, x^{v,\eta}(t), \mathbb{E}(x^{v,\eta}(t)), v(t)) \\
&-g^j(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t))]dt.
\end{aligned} \tag{4.28}$$

Finally, the duality relation (4.21) follows by combining (4.25)-(4.28) together with (4.24). This completes the proof of *Lemma 4.4.1.* □

Proof of Theorem 4.4.2.: Let $x^{u,\xi}(\cdot)$ be the solution of the state equation (4.1) and $(\Psi^u(\cdot), Q^u(\cdot), K^u(\cdot))$ be the solution of the adjoint equation (4.4) corresponding to $(u(\cdot), \xi(\cdot)) \in \mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2([0,T])$. For any $(v(\cdot), \eta(\cdot)) \in \mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2([0,T])$ and from (4.2), we obtain

$$\begin{aligned}
&J(u(\cdot), \xi(\cdot)) - J(v(\cdot), \eta(\cdot)) \\
&= \mathbb{E}(h(x^{u,\xi}(T), \mathbb{E}\left(x^{u,\xi}(T)\right) - h(x^{v,\eta}(T), \mathbb{E}\left(x^{v,\eta}(T)\right)) \\
&+\mathbb{E}\int_0^T [\ell(t, x^{u,\xi}(t), \mathbb{E}\left(x^{u,\xi}(t)\right), u(t) - \ell(t, x^{v,\eta}(t), \mathbb{E}\left(x^{v,\eta}(t)\right), v(t)]dt \\
&+\mathbb{E}\int_0^T \mathcal{M}(t)d(\xi - \eta)(t).
\end{aligned}$$

By the convexity condition on h (see assumptions (C5)), we get

$$\begin{aligned} &J(u(\cdot),\xi(\cdot)) - J(v(\cdot),\eta(\cdot)) \leq \mathbb{E}[(h_x(x^{u,\xi}(T),\mathbb{E}\left(x^{u,\xi}(T)\right) \\ &+\mathbb{E}(h_y(x^{u,\xi}(T),\mathbb{E}\left(x^{u,\xi}(T)\right)))\left(x^{u,\xi}(T) - x^{v,\eta}(T)\right)] \\ &+\mathbb{E}\int_0^T [\ell(t,x^{u,\xi}(t),\mathbb{E}\left(x^{u,\xi}(t)\right),u(t)) - \ell(t,x^{v,\eta}(t),\mathbb{E}\left(x^{v,\eta}(t)\right),v(t))]dt \\ &+\mathbb{E}\textstyle\int_0^T \mathcal{M}(t)d(\xi-\eta)(t). \end{aligned} \tag{4.29}$$

We observe that, from the adjoint equation (4.4), we get

$$\begin{aligned} &J(u(\cdot),\xi(\cdot)) - J(v(\cdot),\eta(\cdot)) \\ &\leq \mathbb{E}\left[\Psi^u(T)\left(x^{u,\xi}(T) - x^{v,\eta}(T)\right)\right] + \mathbb{E}\textstyle\int_0^T[\ell(t,x^{u,\xi}(t),\mathbb{E}\left(x^{u,\xi}(t)\right),u(t)) \\ &-\ell(t,x^{v,\eta}(t),\mathbb{E}\left(x^{v,\eta}(t)\right),v(t))]dt + \mathbb{E}\textstyle\int_0^T \mathcal{M}(t)d(\xi-\eta)(t). \end{aligned}$$

By applying *Lemma 4.4.1,* we have

$$\begin{aligned} &J(u(\cdot),\xi(\cdot)) - J(v(\cdot),\eta(\cdot)) \\ &\leq \mathbb{E}\textstyle\int_0^T[\mathcal{H}(t,x^{u,\xi}(t),\mathbb{E}(x^{u,\xi}(t)),u(t),\Psi^u(t),Q^u(t),K^u(t)) \\ &\quad - \mathcal{H}(t,x^{v,\eta}(t),\mathbb{E}(x^{v,\eta}(t)),v(t),\Psi^u(t),Q^u(t),K^u(t))]dt \\ &-\mathbb{E}\textstyle\int_0^T \mathcal{H}_x(t,x^{u,\xi}(t),\mathbb{E}\left(x^{u,\xi}(t)\right),u(t),\Psi^u(t),Q^u(t),K^u(t))(x^{u,\xi}(t) - x^{v,\eta}(t)) \\ &-\mathbb{E}\textstyle\int_0^T \mathbb{E}[\mathcal{H}_y(t,x^{u,\xi}(t),\mathbb{E}\left(x^{u,\xi}(t)\right),u(t),\Psi^u(t),Q^u(t),K^u(t))](x^{u,\xi}(t) - x^{v,\eta}(t)) \\ &+\mathbb{E}\textstyle\int_{[0,T]}(\Psi^u(t)\mathcal{C}(t) + \mathcal{M}(t))d(\xi-\eta)(t). \end{aligned} \tag{4.30}$$

By the convexity of the functional $\mathcal{H}\left(t,\cdot,\cdot,\cdot,\Psi^u(t),Q^u(t),K^u(t)\right),$ (see assumption (H5)) in the sense of Clarke's generalized gradient, the following holds

$$\begin{aligned}
&\mathbb{E}\int_0^T [\mathcal{H}(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)) \\
&-\mathcal{H}(t, x^{v,\eta}(t), \mathbb{E}(x^{v,\eta}(t)), v(t), \Psi^u(t), Q^u(t), K^u(t))]dt \\
&\leq \mathbb{E}\int_0^T \mathcal{H}_x(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t))(x^{u,\xi}(t) - x^{v,\eta}(t))dt \\
&+\mathbb{E}\int_0^T \mathbb{E}(\mathcal{H}_y(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)))(x^{u,\xi}(t) - x^{v,\eta}(t))dt \\
&+\mathbb{E}\int_0^T \mathcal{H}_u(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t))(u(t) - v(t))dt.
\end{aligned} \tag{4.31}$$

Since the conditional expectation $\mathbb{E}[\mathcal{H}_u(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)) \mid \mathcal{G}_t]$, $v(t)$ and $u(t)$ are $\mathcal{G}_t-$measurable, we have

$$\begin{aligned}
&\mathbb{E}[\mathcal{H}_u(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)) \mid \mathcal{G}_t](v(t) - u(t)) \\
&= \mathbb{E}[\mathcal{H}_u\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)\right)(v(t) - u(t)) \mid \mathcal{G}_t].
\end{aligned} \tag{4.32}$$

Using condition (4.19), (4.31) and (4.32), we obtain

$$\begin{aligned}
&\mathcal{H}\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)\right) \\
&-\mathcal{H}\left(t, x^{v,\eta}(t), \mathbb{E}(x^{v,\eta}(t)), v(t), \Psi^u(t), Q^u(t), K^u(t)\right))dt \\
&-\mathbb{E}\int_0^T \left[\mathcal{H}_x(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t),\ Q^u(t), K^u(t))\right. \\
&+\mathbb{E}(\mathcal{H}_y(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), \\
&\left.K^u(t)))\right](x^{u,\xi}(t) - x^{v,\eta}(t))dt \leq 0.
\end{aligned} \tag{4.33}$$

From (4.30) and (4.33), then for any continuous-singular control $(v(\cdot), \eta(\cdot)) \in \mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2([0, T])$, we obtain

$$J\left(u(\cdot), \xi(\cdot)\right) - J(v(\cdot), \eta(\cdot)) \leq 0.$$

Finally, we observe that since $(v(\cdot), \eta(\cdot))$ is an arbitrary admissible control of $\mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2\left([0, T]\right)$, the desired result (4.20) follows. This completes the proof of *Theorem 4.4.2.* □

4.5 Application: continuous-singular mean-field linear quadratic control problem with Teugels martingales

As an application, under partial-information, we study optimal continuous-singular stochastic linear quadratic control problem for linear mean-field SDEs driven by Teugels martingales associated with some Lévy processes and an independent Brownian motion. The optimal control $(u^*(t), \xi^*(t))$ is obtained in feedback form involving both $x^{u,\xi}(\cdot)$ and its marginal law through its expected value $\mathbb{E}(x^{u,\xi}(\cdot))$, via the solutions of Riccati ordinary differential equations (ODEs). Let $\mathcal{G}_t$ be a given subfiltration of $\mathcal{F}_t$, $t \geq 0$,. For example, $\mathcal{G}_t$ could be the δ-delayed information defined by $\mathcal{G}_t = \mathcal{F}_{(t-\delta)^+} : t \geq 0$, where δ is a given constant delay. The cost functional to be minimized, over the set of admissible controls $\mathcal{U}^1_{\mathcal{G}} \times \mathcal{U}^2_{\mathcal{G}}([0,T])$, has the quadratic form

$$\begin{aligned} J\left(u(\cdot), \xi(\cdot)\right) = {} & \frac{1}{2}\mathbb{E}\int_0^T (Rx^{u,\xi}(t)^2 + Nu(t)^2)dt + \frac{1}{2}\pi\mathbb{E}\left(x^{u,\xi}(T)^2\right) \\ & + \int_{[0,T]} \mathcal{M}(t)d\xi(t), \end{aligned} \tag{4.34}$$

where $(u(\cdot), \xi(\cdot)) \in \mathcal{U}^1_{\mathcal{G}} \times \mathcal{U}^2_{\mathcal{G}}([0,T])$ is adapted to a subfiltration $\mathcal{G}_t$, and R, N and π are positive constants, subject to $x^{u,\xi}(t)$ is the solution of the following linear mean-field SDE:

$$\begin{cases} dx^{u,\xi}(t) = (Ax^{u,\xi}(t) + \widetilde{A}\mathbb{E}\left(x^{u,\xi}(t)\right) + Bu(t))dt \\ + \sum_{j=1}^d (C^j x^{u,\xi}(t) + \widetilde{C}^j\mathbb{E}\left(x^{u,\xi}(t)\right) + D^j u(t))dW^j(t), \\ + \sum_{j=1}^\infty (\phi^j x^{u,\xi}(t) + \widetilde{\phi}^j\mathbb{E}\left(x^{u,\xi}(t)\right) + F^j u(t))dH^j(t), \\ +\mathcal{C}(t)d\xi(t),\ x^{u,\xi}(0) = x_0, \end{cases} \tag{4.35}$$

where A, $\widetilde{A}$, B, C^j, $\widetilde{C}^j$, D^j, ϕ^j, $\widetilde{\phi}^j$ and F^j are constants.

For a given continuous-singular control $(u(\cdot), \xi(\cdot) \in \mathcal{U}^1_{\mathcal{G}} \times \mathcal{U}^2_{\mathcal{G}}([0,T])$, then from (4.5) the Hamiltonian functional $\mathcal{H}$ corresponding to the partial-information control problem (4.34)-

(4.35) gets the form

$$
\begin{aligned}
&\mathcal{H}\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t), \Psi^u(t), Q^u(t), K^u(t)\right) \\
&= \Psi^u(t)(Ax^{u,\xi}(t) + \widetilde{A}\mathbb{E}\left(x^{u,\xi}(t)\right) + Bu(t)) \\
&+ \textstyle\sum_{j=1}^{d} Q^{u,j}(t)(C^j x^{u,\xi}(t) + \widetilde{C}^j\mathbb{E}\left(x^{u,\xi}(t)\right) + D^j u(t)) \\
&+ \textstyle\sum_{j=1}^{\infty} K^{u,j}(t)(\phi^j x^{u,\xi}(t) + \widetilde{\phi}^j\mathbb{E}\left(x^{u,\xi}(t)\right) + F^j u(t)) \\
&+ \tfrac{1}{2}(Rx^{u,\xi}(t)^2 + Nu(t)^2).
\end{aligned}
\tag{4.36}
$$

From (4.4), the corresponding adjoint equation gets the form

$$
\begin{cases}
d\Psi^u(t) = -[A\Psi^u(t) + \widetilde{A}\mathbb{E}(\Psi^u(t)) \\
+ \sum_{j=1}^{d}(C^j Q^{u,j}(t) + \widetilde{C}^j\mathbb{E}(Q^{u,j}(t))) \\
+ \sum_{j=1}^{\infty}(\phi^j K^{u,j}(t) + \widetilde{\phi}\mathbb{E}[K^{u,j}(t)]) + Rx^{u,\xi}(t)]dt \\
+ \sum_{j=1}^{d} Q^{u,j}(t)dW^j(t) + \sum_{j=1}^{\infty} K^{u,j}(t)dH^j(t), \\
\Psi^u(T) = \pi x^{u,\xi}(T).
\end{cases}
\tag{4.37}
$$

Let $(u^*(t), \xi^*(t)$ be a local optimal control of the partial-information control problem (4.34)-(4.35). Then by applying *Theorem 4.4.2* and the fact that

$$
\begin{aligned}
&\mathcal{H}_u\left(t, x^*(t), \mathbb{E}\left(x^*(t)\right), u^*(t), \Psi^*(t), Q^*(t), K^*(t)\right) \\
&= B\Psi^*(t) + DQ^*(t) + FK^*(t) + Nu^*(t),
\end{aligned}
$$

we deduce that the optimal control is given by

$$
\begin{aligned}
&\mathbb{E}\left[B\Psi^*(t) + DQ^*(t) + FK^*(t) + Nu^*(t) \mid \mathcal{G}_t\right] = 0, \\
&t \in [0, T].
\end{aligned}
\tag{4.38}
$$

Since $u^*(t)$ is adapted to $\mathcal{G}_t$, we get

$$u^*(t) = -N^{-1}\{B\mathbb{E}[\Psi^*(t) \mid \mathcal{G}_t] + D\mathbb{E}[Q^*(t) \mid \mathcal{G}_t] \\ +F\mathbb{E}[K^*(t) \mid \mathcal{G}_t]\}. \tag{4.39}$$

Now, to solve explicitly the above equation (4.39), we assume that the adjoint process $\Psi^*(\cdot)$ has the following form:

$$\Psi^*(t) = U_1(t)\, x^*(t) + U_2(t)\, \mathbb{E}(x^*(t)) + U_3(t), \tag{4.40}$$

where $U_1(\cdot)$, $U_2(\cdot)$ and $U_3(\cdot)$ are deterministic differentiable functions. Applying Itô's formula to (4.40), we get

$$\begin{aligned} d\Psi^*(t) &= d(U_1(t)\, x^*(t)) + d(U_2(t)\, \mathbb{E}(x^*(t))) + dU_3(t) \\ &= U_1(t)\, dx^*(t) + x^*(t)U_1'(t)\, dt + U_2(t)\, d\mathbb{E}(x^*(t)) + \mathbb{E}(x^*(t))\, U_2'(t)\, dt \\ &\quad + U_3'(t)\, dt. \end{aligned}$$

By simple computation and the fact that $d[\mathbb{E}(x^*(t))] = [(A + \widetilde{A})\mathbb{E}(x^*(t)) + B\mathbb{E}(u^*(t))]dt$, we obtain

$$\begin{aligned} d\Psi^*(t) &= \Big\{U_1(t)\, [Ax^*(t) + \widetilde{A}\mathbb{E}(x^*(t)) + Bu^*(t))] \\ &+U_2(t)\, [(A + \widetilde{A})\mathbb{E}(x^*(t)) + B\mathbb{E}(u^*(t))] \\ &+ x^*(t)U_1'(t) + \mathbb{E}(x^*(t))\, U_2'(t) + U_3'(t)\}\, dt \\ &+ \textstyle\sum_{j=1}^{d}[C^j x^*(t) + \widetilde{C}^j\mathbb{E}(x^*(t)) + D^j u^*(t)]U_1(t)\, dW^j(t), \\ &+ \textstyle\sum_{j=1}^{\infty}\Big[\phi^j x^*(t) + \widetilde{\phi}^j\mathbb{E}(x^*(t)) + F^j u^*(t)\Big]\, U_1(t)\, dH^j(t), \\ \Psi^*(T) &= U_1(T)\, x^*(T) + U_2(T)\, \mathbb{E}(x^*(T)) + U_3(T). \end{aligned} \tag{4.41}$$

Now, from (4.37) and (4.41), we can easily prove that $U_3(t) \equiv 0$, $\forall t \in [0,T]$,

$$\begin{cases} U_1(t)\left[Ax^*(t) + \widetilde{A}\mathbb{E}(x^*(t)) + Bu^*(t)\right] \\ +U_2(t)\left[(A+\widetilde{A})\mathbb{E}(x^*(t)) + B\mathbb{E}(u^*(t))\right] \\ +x^*(t)U_1'(t) + \mathbb{E}(x^*(t))U_2'(t) \\ = -[A\Psi^*(t) + \widetilde{A}\mathbb{E}(\Psi^*(t)) + (CQ^*(t) \\ +\widetilde{C}\mathbb{E}(Q^*(t))) + (\phi K^*(t) + \widetilde{\phi}\mathbb{E}[K^*(t)]) + Rx^*(t)]. \end{cases} \tag{4.42}$$

A similar argument shows that

$$Q^*(t) = [Cx^*(t) + \widetilde{C}\mathbb{E}(x^*(t)) + Du^*(t)]U_1(t), \tag{4.43}$$

and

$$K^*(t) = \left[\phi x^*(t) + \widetilde{\phi}\mathbb{E}(x^*(t)) + Fu^*(t)\right]U_1(t). \tag{4.44}$$

By comparing the coefficients of $x^*(t)$ and $\mathbb{E}(x^*(t))$ in equation (4.42) and last equation in (4.41), we immediately deduce that $U_1(\cdot)$, $U_2(\cdot)$ are given by the following ODEs:

$$\begin{cases} U_1'(t) + (2A + C^2 + \phi^2)U_1(t) + R = 0, \quad U_1(T) = \pi, \\ U_2'(t) + 2(A+\widetilde{A})U_2(t) + (2\widetilde{A} + \widetilde{C}^2 + \widetilde{\phi}^2 + 2(C\widetilde{C} + \phi\widetilde{\phi}))U_1(t) = 0, \\ U_2(T) = 0. \end{cases} \tag{4.45}$$

By solving the ODEs-(4.45), we obtain

$$U_1(t) = [\pi + R\left(2A + C^2 + \phi^2\right)^{-1}]\exp\left\{(2A + C^2 + \phi^2)(T-t)\right\} - R\left(2A + C^2 + \phi^2\right)^{-1}$$
$$U_2(t) = (2\widetilde{A} + \widetilde{C}^2 + \widetilde{\phi}^2 + 2(C\widetilde{C} + \phi\widetilde{\phi}))\exp[-2(A+\widetilde{A})t]\int_t^T U_1(s)\exp[2(A+\widetilde{A})s]ds.$$

Now, from *Theorem 4.4.2*, the singular part $\xi^*(\cdot)$ satisfies: for any $\xi(\cdot) \in \mathcal{U}^2_{\mathcal{G}}([0,T])$

$$\mathbb{E}\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)) d\left(\xi - \xi^*\right)(t) \geq 0. \tag{4.46}$$

If we define

$$\mathfrak{D} \triangleq \left\{(w,t) \in \Omega \times [0,T] : \mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t) \geq 0\right\},$$

and let $\xi(\cdot) \in \mathcal{U}^2_{\mathcal{G}}([0,T])$ such that

$$d\xi(t) = \begin{cases} 0 \text{ if } \mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t) \geq 0 \\ \\ d\xi^*(t) \text{ otherwise,} \end{cases} \tag{4.47}$$

then by a simple computations, it is easy to see that

$$\begin{aligned} 0 &\leq \mathbb{E}\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)) d\left(\xi - \xi^*\right)(t) \\ &= \mathbb{E}\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)) I_{\mathsf{D}}\, d\left(-\xi^*\right)(t) \\ &= -\mathbb{E}\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)) I_{\mathsf{D}}\, d\xi^*(t), \end{aligned}$$

which implies that $\xi^*(t)$ satisfies for any $t \in [0,T]$

$$\mathbb{E}\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t)) I_{\mathsf{D}}\, d\xi^*(t) = 0. \tag{4.48}$$

Finally, from (4.47) and (4.48) we can easy shows that the optimal singular control $\xi^*(\cdot)$ has the form

$$\xi^*(t) = \xi(t) + \int_0^t I_{\mathsf{D}^c}(s,w) ds = \xi(t) + \int_0^t I_{\{(w,s)\in\Omega\times[0,T]:\mathcal{M}(s)+\mathcal{C}(s)\Psi^*(s)<0\}}(s,w) ds. \tag{4.49}$$

Finally, we give the explicit optimal continuous-singular control in feedback form involving both the state process and its expected value.

Theorem 4.5.3. The optimal continuous-singular control $(u^*(\cdot), \xi^*(\cdot)) \in \mathcal{U}_{\mathcal{G}}^1 \times \mathcal{U}_{\mathcal{G}}^2([0,T])$ of the partial-information mean-field linear quadratic control problem (4.34)-(4.35) is given in feedback form by

$$u^*(t) = u^*(t, x^*(t), \mathbb{E}(x^*(t))) = -N^{-1}\mathbb{E}[B\Psi^*(t) + DQ^*(t) + FK^*(t) \mid \mathcal{G}_t].$$

$$\xi^*(t) = \xi(t) + \int_0^t I_{\{(w,s)\in\Omega\times[0,T]:\mathcal{M}(s)+\mathcal{C}(s)\Psi^*(s)<0\}}(s,w)ds, \quad t \in [0,T].$$

4.6 Some discussion and concluding remarks

In this last chaptre, under partial-information, necessary and sufficient conditions for optimal continuous-singular control for mean-field SDEs driven by Teugels martingales associated with some Lévy processes and an independent Brownian motion have been established. A partial-information linear quadratic control problem has been studied to illustrate our theoretical results. In our mean-field control problem, there are two types of jumps for the state processes, the predictable ones which come from the discrete interventions of singular control and the inaccessible ones which come from the Teugels martingale measure.

- 1. In *Theorem 4.3.1*, equation (4.9) is equivalent to

$$\mathbb{E}[\mathcal{H}_u(t, x^*(t_-), \mathbb{E}(x^*(t_-)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)) \mid \mathcal{G}_t] \\ +\mathbb{E}\left[\int_{[0,T]} (\mathcal{M}(t) + \mathcal{C}(t)\Psi^*(t))\, d\xi^*(t) \mid \mathcal{G}_t\right] = 0, \ a.e., \ t \in [0,T].$$

- 2. If $\xi^*(t) = \sum_{j\geq 1} \xi_j I_{[\tau_j,T]}(t)$, $\tau_j \in [0,T]$, then (4.9)-*Theorem 4.3.1* is equivalent to

$$\mathbb{E}[\mathcal{H}_u(t, x^*(t_-), \mathbb{E}(x^*(t_-)), u^*(t), \Psi^*(t), Q^*(t), K^*(t)) \mid \mathcal{G}_t] \\ + \sum_{\tau_j \leq T} \mathbb{E}\{[\mathcal{M}(\tau_j) + \mathcal{C}(\tau_j)\Psi^*(\tau_j)] \mid \mathcal{G}(\tau_j)\} = 0, \quad a.e., \ \tau_j \in [0,T].$$

- 3. If $\mathcal{G}_t = \mathcal{F}_t$ and $g = 0$, our maximum principle (*Theorem 4.3.1*) coincides with the stochastic maximum principle (*Theorem 1*) developed in Hafayed [82].

- 4. If $\mathcal{G}_t = \mathcal{F}_t$ and $\mathcal{C}(t) = \mathcal{M}(t) = 0$ and without mean-field terms, our maximum principle (*Theorem 4.3.1*) coincides with the stochastic maximum principle developed in Meng and Tang [90].

- 5. Apparently, there are many problems left unsolved. To mention a few, necessary and sufficient conditions for optimality for mean-field general controlled SDEs driven by Teugels martingales and an independent Brownian motion of the form

$$\begin{cases} dx^{u,\xi}(t) = f\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)\right) dt \\ +\sum_{j=1}^{d} \sigma^j\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t)\right) dW^j(t) \\ +\sum_{j=1}^{\infty} g^j\left(t, x^{u,\xi}(t_-), \mathbb{E}(x^{u,\xi}(t_-)), u(t)\right) dH^j(t) \\ +\mathcal{C}\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t))\right) d\xi(t), \\ x^{u,\xi}(0) = x_0, \end{cases} \tag{4.50}$$

and the expected cost has the general form

$$\begin{aligned} J\left(u(\cdot), \xi(\cdot)\right) = \mathbb{E}\Big\{ \int_0^T \ell(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t)), u(t))dt \\ +h\left(x^{u,\xi}(T), \mathbb{E}\left(x^{u,\xi}(T)\right)\right) \\ + \int_{[0,T]} \mathcal{M}\left(t, x^{u,\xi}(t), \mathbb{E}(x^{u,\xi}(t))\right) d\xi(t)\Big\}. \end{aligned} \tag{4.51}$$

where the coefficients of the singular parts $\mathcal{C}$ and $\mathcal{M}$ depend on the state of the solution process as well as of its expected value. Moreover, the second-order maximum principle for the problem (4.50)-(4.51), where the control domain is not assumed to be convex is still an open problem. It is worthwhile pointing out that we can derive these results by using the singular version of the Hamiltonian and the adjoint processes should be defined within singular control. Such topics will be studied in our future works.

References

Bibliography

[1] Ahmed, N.U.: Nonlinear diffusion governed by McKean-Vlasov equation on Hilbert space and optimal control, SIAM J. Control Optim. (**46**), 356-378 (2007).

[2] Bensoussan, A.: *Lectures on stochastic control.* In: Lecture Notes in Mathematics, Vol. (**972**)1–62, Springer, Berlin (1981).

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

[4] Bouchard, B., Elie, R.: *Discrete time approximation of decoupled Forward-Backward SDE with jumps*, Stochastic Process. Appl. (**118**)1 53-75 (2008).

[5] Buckdahn, R., Li, J., Peng, S. *Mean-field backward stochastic differential equations and related partial differential equations.* Stochastic Processes and their Applications, (**119**), 3133–3154 (2009).

[6] Buckdahn, R., Djehiche, B., and Li, J.: *A General Stochastic Maximum Principle for SDEs of Mean-field Type.* Appl. Math. Optim. (**64**) 197–216 (2011).

[7] Bellman, R.: *Dynamic programming,* Princeton Univ. Press., (1957).

[8] Borkar V. S., Kumar, K.S.: *Mean-field limit in portfolio optimization*, Stoch. Anal. Appl.,**(28)**, 884–906 (2010).

[9] A. Cadenillas, *A stochastic maximum principle for system with jumps, with applications to finance,* Systems and control letters (**47**) 433-444 (2002).

[10] Clarke, F.H.: *Optimization and Nonsmooth Analysis,* Wiley, New York, (1983).

[11] Chighoub, F., Mezerdi, B.: *Near optimality conditions in stochastic control of jump diffusion processes.* Syst. Cont Let.(**60)**, 907–916 (2011).

[12] Ekeland, I.: *Non convex mninmisations,* Bull. Amer. Math. Soc., NS (**1**) 324-353 (1979).

[13] 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).

[14] Gabasov, R. Kirillova M., Mordukhovich, B.S.: *The* $\varepsilon-$*maximum principle for sub-optimal controls,* Soviet Math Dokh (**27**), 95-99 (1983).

[15] Kac, M.: *Foundations of kinetic theory, Proc. 3-rd Berkeley Sympos.* Math. Statist. Prob. (**3**), 171-197 (1956).

[16] Hafayed, M., Veverka, P., Abbas, S.: *On Maximum Principle of near-optimality for diffusions with Jumps, with Application to Consumption-Investment Problem,* Differential Equations and Dynamical Systems, Springer-Verlag, (**20)-**2, 111-125 (2012).

[17] Hafayed, M., Abbas, S., Veverka, P.: *On necessary and sufficient conditions for near-optimal singular stochastic controls.* Optimization Letters, Springer-Verlag, DOI 10.1007/s11590-012-0484-6 (2012).

[18] Hafayed, M., Abbas, S.: *A General Maximum Principle for Stochastic Differential Equations of Mean-field type with Jump Processes.* Technical report (2013).

[19] Hui, E., Huang, J., Li, X., and Wang. G.,: Near-optimal control for stochastic recursive problems, Syst. Cont. Letters (**60**), 161-168 (2011).

[20] Bahlali K., Khelfallah N, Mezerdi B., Necessary and sufficient conditions for near-optimality in stochastic control of FBSDEs, Systems & Control Letters **58,** 857 864 (2009)

[21] Huang, J., Li, X., Wang, G.: Near-optimal control problems for linear forward-backward stochastic systems, Automatica, (**46**)2, 397-404 (2010).

[22] Jeanblanc-Picqué, M., Pontier, M.: *Optimal portfolio for small investor in a market model with discontinuous prices,* Appl. Math. Optim. (**22**), 287-310 (1990).

[23] Ma, J. Yong J.: *Forward-Backward Stochastic Differential Equations and Their Applications.* Lecture Notes in Mathematics, Springer (1999).

[24] McKean, H.P.: *A class of Markov processes associated with nonlinear parabolic equations,* Proc. Natl. Acad. Sci. USA, (**56**), 1907-1911 (1966).

[25] Mordukhovich, B.S.: *Approximation Methods in Problems of Optimization and Control,* Nauka, Moscow (1988).

[26] Meyer-Brandis, T., Øksendal, B., Zhou, X.Y.: *A Mean-field stochastic maximum principle via Malliavin calculus.* A special issue for Mark Davis' Festschrift, to appear in Stochastics.

[27] Cadenillas, A., Karatzas, I.: *The stochastic maximum principle for linear, convex optimal control with random coefficients.* SIAM J. Control Optim. (**33**)-2, 590–624 (1995).

[28] Elliott, R.J.: *The optimal control of diffusions.* Appl. Math. Optim. (**22**), 229–240 (1990).

[29] Kushner, H.J.: *On the stochastic maximum principle: fixed time of control.* J. Math. Anal. Appl. (**11**), 78–92 (1965).

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

[31] Mezerdi B. Necessary conditions for optimality for a diusion with a non-smooth drift. Stochastics, **24**, 305-326 (1988)

[32] Haussmann, U.G.: *A Stochastic Maximum Principle for Optimal Control of Diffusions. Longman.* Scientific and Technical, Essex (1986).

[33] ∅ksendal, B., Sulem. A.: *Applied Stochastic Control of Jump Diffusions*, 2^{nd} ed., Springer-Verlag, Berlin, (2007).

[34] Pan., L. P., Teo, K.L.: *Near-Optimal Controls of Class of Volterra Integral Systems,* J. Optim Theory Appli., (**101**)-2, 355-373 (1999).

[35] Pontryagin, L.S., Boltanskii, V.G., Gamkrelidze, R.V.: *The mathematical theory of optimal processes.* Inter-scien. N.Y., (1962).

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

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

[38] Shi, J., Wu, Z.: *The Maximum principle for fully coupled Forward-backward stochastic control system*, Acta Automatica sinica, (**32**)-2, 161-169 (2006).

[39] Shi, J., and Wu, Z.: *A Stochastic Maximum Principle for Optimal Control of Jump Diffusions and Application to Finance.* Chinese Journal of Applied Probability and statistics, (**27**)-2, (2011).

[40] Tang, S. J., Li, X.J.: *Necessary conditions for optimal control of stochastic systems with random jumps.* SIAM J. Control Optim. (**32**)-5, 1447–1475 (1994).

[41] Veretennikov, A.Y.: *On ergodic measures for Mean-field stochastic equations, From stochastic calculus to mathematical finance*, Springer Berlin, 623-633 (2006).

[42] Zhou, X.Y.: *Deterministic near-optimal controls. Part I: Necessary and sufficient conditions for near optimality,* J. Optim. Theory Appl. (**85**), 473-488 (1995).

[43] Zhou, X.Y.: *Deterministic near-optimal controls. Part II: Dynamic programming and viscosity solution approach,* Math. Oper. Rese., (**21**), 655-674 (1996).

[44] Zhou, X.Y., Sethi S.: *A Sufficient Condition for Near-Optimal Stochastic Controls and its Application to Manufacturing Systems* , Appl. Math. Optim. (**29)-**1, 67-92, (1994).

[45] Zhou, X.Y.: *Stochastic near-optimal controls: Necessary and sufficient conditions for near-optimality,* SIAM. Control. Optim. (**36**)-3, 929-947 (1998).

[46] Hafayed M, and Abbas S.: Stochastic near-optimal singular controls for jump diffusions: necessary and sufficient conditions, Journal of Dynamical and Control Systems, **19**(4), 503-517 (2013).

[47] Zhang L, Huang J, Li X.: Necessary condition for near optimal control of linear forward-backward stochastic differential equations. International Journal of Control, Doi 10.1080/00207179 pp. 1-29, (2015).

[48] Ni. Y.H, Zhang J.F., and Li X., (2014) Indefinite mean-field stochastic linear-quadratic optimal control, IEEE Transactions on automatic control, Doi: 10.1109/TAC.2014.2385253

[49] Wang B.C. and Zhang J.F.: (2012).Mean-field games for larg-population multiagent systems with Markov jump parameters, SIAM J. Control. Vol 50 No. 4, pp. 2308–2334.

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

[51] Wang G, Zhang C, and Zhang W.: (2014). Stochastic maximum principle for mean-field type optimal control under partial information, IEEE Transactions on automatic control, Vol 59 No 2, pp. 522-528.

[52] Xu R, and Wu T.: (2014) Mean-field backward stochastic evolution equations in Hilbert spaces and optimal control for BSPDEs, Abstract and Applied Analysis Vol 2014, Article ID 839467, 15 pages.

[53] Hafayed M, Veverka P and Abbas A.: (2014) On Near-optimal Necessary and Sufficient Conditions for Forward-backward Stochastic Systems with Jumps, with Applications to Finance. Applications of Mathematics, Vol 59 No.4, pp. 407-440.

[54] Tang M.: (2014) Stochastic maximum principle of near-optimal control of fully coupled forward-backward stochastic differential equation, Abstract and Applied Analysis Volume 2014, Article ID 361259, 12 pages

[55] Zhang L, Huang J, Li X.: (2015) Necessary condition for near optimal control of linear forward-backward stochastic differential equations. International Journal of Control, Doi 10.1080/00207179 pp. 1-29.

[56] Markowitz H.: (1952) Portfolio selection. J. Finance Vol. **7**, pp. 77-91

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

[58] Shi J.: Sufficient conditions of optimality for mean-field stochastic control problems. 12^{-th} International Conference on Control, Automation, Robotics and Vision Guangzhou, China, 5-7^{th} December, ICARCV 2012, 747-752 (2012)

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

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

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

[62] 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)

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

[64] 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)

[65] Markowitz, H.: Portfolio selection, J. of Finance, **7**: 77-91 (1952)

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

[67] M. Hafayed, S. Abbas, A. Abba : On mean-field partial information maximum principle of optimal control for stochastic systems with Lévy processes, J. Optim Theory Appl, 10.1007/s10957-015-0762-4, (2015).

[68] M. Hafayed, M. Tabet., S. Boukaf: Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem, Commun. Math. Stat (3) (2015) 163–186.

[69] P. Shi and F. Li, A survey on Markovian jump systems: modeling and design, Int. J. of Control, Automation and Systems, 13(1) (2015) 1-16.

[70] P. Shi, Y. Zhang and R. Agarwal, Stochastic finite-time state estimation for discrete time-delay neural networks with Markovian jumps, Neurocomputing, 151, (2015) 168-174.

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

[72] Z. Wu, and F. Zhang: Stochastic maximum principle for optimal control problems of forward-backward systems involving impulse controls, IEEE Transactions on Automatic Control, 56(6) (2011) 1401-1406.

[73] M. Hafayed, S. Abbas: Stochastic near-optimal singular controls for jump diffusions: Necessary and sufficient conditions, Journal of Dynamical and Control Systems, 19(4) (2013) 503-517.

[74] L. Alvarez and T.A. Rakkolainen: On singular stochastic control and optimal stopping of spectrally negative jump diffusions, Stochastics An International Journal of Probability and Stochastics Processes 81(1) (2009) 55-78.

[75] T.T K. An: Combined optimal stopping and singular stochastic control, Stochastic Analysis and Applications, 28 (2010) 401-414.

[76] J. Menaldi and M. Robin: On singular stochastic control problems for diffusions with jumps, IEEE Transactions on Automatic Control 29(11), (1984) 991-1004.

[77] Alvarez, L.H.R. (2001) Singular stochastic control linear diffusion and optimal stopping: A class of solvable problems, SIAM J. Control Optim. 39, 1697-1710.

[78] Dufour F., and Miller, B. (2006) Maximum principle for singular stochastic control problem. SIAM J. Control Optim., 45(2), 668-698.

[79] Djehiche, B., Tembine, H. and Tempone, R., (2015) A stochastic maximum principle for Risk-Sensitive mean-field-type control, IEEE Transactions on Automatic Control, Doi 10.1109/TAC.2015.2406973.

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

[81] Hafayed, M., Abbas, S. (2014) On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality, J. Optim Theory Appl, 160(3), 778-808.

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

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

[84] Haussmann, U.G., and Suo, W., (1995) Singular optimal control I, II, SIAM J. Control Optim., 33(3), 916-936, 937-959.

[85] Haussmann U.G.: Extremal controls for completely observable diffusions, in Lecture Notes in Control and Information Sciences, 42, Springer-Verlag, Berlin, New York, 1982, pp. 149-160.

[86] Haussmann U.G.: Optimal control of partially observed diffusions via the separation principle, in Lecture Notes in Control and Information Sciences, 43, Springer-Verlag, Berlin, New York, (1982) 302-311.

[87] Haussman U.G.: A Stochastic Maximum Principle for Optimal Control of Diffusions, Longman Scientific and Technical, Essex, U.K., 1986.

[88] Hu, Y., ∅ksendal, B., and Sulem A. (2014) Singular mean-field control games with applications to optimal harvesting and investment problems, Technical report, arXiv:1406.1863v1.

[89] ∅ksendal, B., and Sulem A. (2012) Singular stochastic control and optimal stopping with partial information of Itô-Lévy processes, SIAM J Control Optim, 50(4), 2254-2287.

[90] Meng, Q.X., and Tang, M.N. (2009) Necessary and sufficient conditions for optimal control of stochastic systems associated with Lévy processes Sci. China Ser F-Inf Sci. 52(11), 1982–1992.

[91] Meng, Q.X., Zhang F., and Tang M.N. (2012) Maximum principle for backward stochastic systems associated with Lévy processes under partial information, Proceedings of the 31 st Chinese control conference, July 25-27, Hefei, China.

[92] Meng, Q.X: (2014) General linear quadratic optimal stochastic control problem driven by a brownian motion and a poisson random martingale measure with random coefficients, Stochastic Analysis and Applications, 32(1) 88-109.

[93] Meng, Q.X and Shen Y. (2015) Optimal control of mean-field jump-diffusion systems with delay: A stochastic maximum principle approach, Journal of Computational and Applied Mathematics, 279, 13-30.

[94] Ni, Y.H., Zhang, J.F., and Li, X. (2015) Indefinite Mean-field stochastic linear-quadratic optimal control, IEEE Transactions on Automatic Control, DOI: 10.1109/TAC.2014.2385253.

[95] Wang, G., Wu, Z., and Zhang, C..(2014) Maximum Principles for Partially Observed Mean-Field Stochastic Systems with Application to Financial Engineering, Proceedings of the 33rd Chinese Control Conference July 28-30.

[96] Yong, J. (2013) A Linear-Quadratic Optimal Control Problem for Mean-Field Stochastic Differential Equations. SIAM J. Control Optim. 51(4), 2809-2838.

[97] Zhang, F. (2013) Stochastic maximum principle for mixed regular-singular control problems of forward-backward systems, J Syst Sci Complex 26, 886–901.

[98] Mitsui, K., and Tabata, M.: A stochastic linear quadratic problem with Lévy, process and its application to finance, Stoch. Proc. Appl. **118**, 120-152 (2008)

[99] Tang, H. and Wu, Z.: Stochastic differential equations and stochastic linear quadratic optiaml control problem with Lévy processes. J. Syst. Sci. Complex, **22**, 122-136 (2009)

[100] Tang, M., Zhang, Q.: Optimal variational principle for backward stochastic control systems associated with Lévy processes, arXiv: 1010.4744v1, (2010)

[101] Nualart, D., Schoutens, W.: BSDE's and Feynman–Kac formula for Lévy process with application in finance, Bernoulli **7,** 761–776 (2001)

[102] Baghery, F., ∅ksendal, B.: A maximum principal for stochastic control with partial information, Stoch. Anal. Appl. **25**, 493–514 (2007)

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

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

[105] Lasry, J.M., Lions, P.L.: Mean field games. Japan Jour. Math. **2**, 229–260 (2007)

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

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

[108] 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)

[109] Yong J., and Zhou XY.: (1999) Stochastic Controls. Hamiltonian Systems and HJB Equations. Springer-Verlag. New York.

Appendix

The following result gives the definition and some basic properties of the generalized gradient.

Definition A1. Let F be a convex set in $\mathbb{R}^n$ and let $f : F \to \mathbb{R}$ be a locally Lipschitz function. The generalized gradient of f at $\widehat{x} \in F$, denoted by $\partial_x^\circ f(\widehat{x})$, is a set defined by

$$\partial_x^\circ f(\widehat{x}) = \{\xi \in \mathbb{R}^n : \langle \xi, \upsilon \rangle \leq f^\circ(\widehat{x}, \upsilon), \text{ for any } \upsilon \in \mathbb{R}^n\},$$

where $f^\circ(\widehat{x}, \upsilon) = \limsup_{y \to \widehat{x}, t \to 0} \frac{1}{t} (f(y + t\upsilon) - f(y))$.

Proposition A1. If $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz at $x \in \mathbb{R}^n$, then the following statements holds

1. $\partial_x^\circ f(x)$ is nonempty, compact and convex set in $\mathbb{R}^n$.
2. $\partial_x^\circ(-f)(x) = -\partial_x^\circ(f)(x)$.
3. $\partial_x^\circ f(x) \ni 0$ if f attains a local minimum or maximum at x.
4. If f is continuously differentiable at x, then $\partial_x^\circ f(x) = \{f'(x)\}$.
5. If f, $g : \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz functions at $x \in \mathbb{R}^d$, then

$$\partial_x^\circ(f + g)(x) \subset \partial_x^\circ f(x) + \partial_x^\circ g(x).$$

For the detailed proof of the above Proposition see Clarke [10] or the book by Yong et al., ([109] *Lemma 2.3*).

As a simple example of the generalized gradient, we consider the absolute value function $f : x \mapsto |x-a|$ which is continuously differentiable everywhere except at $x = a$. Since $f'(x) = 1$ for $x > a$ and $f'(x) = -1$ for $x < a$, then a simple calculation shows that the generalized gradient of f at $x = a$ is given by $\partial_x^\circ f(a) = \overline{co}\{-1, 1\} = [-1, 1]$.

The following result gives special case of the Itô formula for jump diffusions.

Lemma A1. (*Integration by parts formula for jumps processes*) Suppose that the processes $x_1(t)$ and $x_2(t)$ are given by: for $j = 1, 2$, $t \in [s, T]$:

$$\begin{cases} dx_j(t) = f(t, x_j(t), u(t))\, dt + \sigma(t, x_j(t), u(t))\, dW(t) \\ \qquad\qquad + \int_\Theta g(t, x_j(t^-), u(t), \theta)\, N(d\theta, dt), \\ x_j(s) = 0. \end{cases}$$

Then we get

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

See Framstad et al., ([13], *Lemma 2.1*) for the detailed proof of the above Lemma.

Proposition A2. Let $\mathcal{G}$ be the predictable $\sigma-$field on $\Omega \times [s, T]$, and f be a $\mathcal{G} \times \mathcal{B}(\Theta)-$measurable function such that

$$\mathbb{E}\int_s^T \int_\Theta |f(r, \theta)|^2\, \mu(d\theta)dr < \infty,$$

then for all $\beta \geq 2$ there exists a positive constant $C = C(T, \beta, \mu(\Theta))$ such that

$$\mathbb{E}\left[\sup_{0 \leq t \leq T}\left|\int_s^t \int_\Theta f(r, \theta)\, N(d\theta, dr)\right|^\beta\right] < C\mathbb{E}\left[\int_s^T \int_\Theta |f(r, \theta)|^\beta\, \mu(d\theta)dr\right].$$

Proof. See Bouchard et al., ([4], *Appendix*). □