

On mean-field stochastic maximum principle for near-optimal controls for Poisson jump diffusion with applications

Mokhtar Hafayed · Abdelmadjid Abba · Syed Abbas

Received: 4 September 2013 / Revised: 10 November 2013 / Accepted: 12 November 2013 / Published online: 3 December 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract In this paper, we study mean-field type stochastic control problems for systems described by mean-field stochastic differential equations with jump processes, in which the coefficients contains not only the state process but also its marginal distribution. Moreover, the cost functional is also 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 with non-negative derivatives, we prove that the near-maximum condition on the Hamiltonian function in integral form is a sufficient condition for near-optimality. Our result differs from the classical one in the sense that here the adjoint equation has a mean-field type, while the second-order adjoint equation remains the same as in the classical case. As an application, our results are applied to a mean-variance portfolio selection where explicit expression of the near-optimal portfolio selection strategy is obtained in the state feedback form involving both state process and its marginal distribution, via the solutions of Riccati ordinary differential equations.

Keywords Stochastic control · Controlled mean-field jump diffusion processes · Near-optimization · Necessary and sufficient conditions · Ekeland’s principle · McKean–Vlasov system · Time-inconsistent solution · Feedback control

Mathematics Subject Classification 60H10 · 93E20

1 Introduction

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 mean-field jump diffusion process

$$\begin{cases} dx^u(t) = f(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dt \\ \quad + \sigma(t, x^u(t), \mathbb{E}(x^u(t)), u(t))dW(t) \\ \quad + \int_{\Theta} g(t, x^u(t^-), u(t), \theta) N(d\theta, dt), \\ x^u(s) = \zeta, \end{cases} \quad (1)$$

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

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

M. Hafayed
Laboratory of Applied Mathematics, Biskra University,
Po Box 145, 07000 Biskra, Algeria
e-mail: hafayedmokhtar@yahoo.com

A. Abba (✉)
Departement of Mathematics, University of Biskra,
07000 Biskra, Algeria
e-mail: abdelmadjidabba@yahoo.com

S. Abbas
School of Basic Sciences, Indian Institute of Technology Mandi,
Mandi 175001, HP, India
e-mail: sabbas.iitk@gmail.com

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

$$J^{s,\zeta}(u(\cdot)) = \mathbb{E} \left[h(x^u(T), \mathbb{E}(x^u(T))) + \int_s^T \ell(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) dt \right]. \tag{2}$$

It worth mentioning that since the cost functional $J^{s,\zeta}$ 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. 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.

It is well-known that near-optimization is as sensible and important as optimization for both theory and applications. Since the recent work by Zhou [1], the concept of near-optimal controls was introduced for a class of stochastic control problems. Various kinds of near-optimal stochastic control problems have been investigated in [2–8]. In Hafayed et al. [2], the authors extended Zhou’s maximum principle of near-optimality [1] to singular stochastic control. The near-optimal stochastic control problem for systems governed by diffusions with jump processes, with application to finance has been investigated by Hafayed et al. [3]. The necessary and sufficient conditions of near-optimal mean-field singular stochastic control have been studied in Hafayed and Abbas [4]. Near-optimality necessary and sufficient conditions for singular control in jump diffusion processes has been investigated in Hafayed and Abbas [5]. The necessary and sufficient conditions for near-optimality for forward-backward stochastic differential equations with some applications have been studied in Huang et al. [6]. The near-optimal control problem for recursive stochastic problem has been studied in Hui et al. [7].

The stochastic optimal control problems for jump processes has been investigated by many authors, see for instance [9–17]. 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 and Li [9]. 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 [13, 18].

Mathematical mean-field problems play an important role in different fields of, economics, finance, physicals, chemistry and stochastic game theory. Many authors made contributions on mean-field systems and applications, see for instance [4, 10, 17, 19–25]. The existence and uniqueness result of mean-field backward stochastic differential equations (MF-BSDEs) as limit approach have been investigated in Buckdahn et al. [19]. The maximum principle for SDEs of mean-field type was introduced in [20]. Under some convexity assumptions, the mean-field type sufficient conditions for optimality have been established by Shi [21]. In Mayer-Brandis et al. [25] a stochastic maximum principle of optimality for systems governed by controlled Itô-Lévy process of mean-field type was proved by using Malliavin calculus. Various local maximum principles of optimality for mean-field stochastic control problem have been derived in [22, 23].

Our main goal in this paper 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 [26] 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 and Li [9]. 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 paper in the “Appendix”.

The rest of the paper 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 paper. In Sects. 3 and 4, we derive necessary and sufficient conditions for near-optimality respectively, which are our main results. An example of this kind of mean-field control problem is also given in the last section.

2 Problem formulation and preliminaries

Throughout this paper, we 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 $\tilde{N}(d\theta, dt)$ the random counting measure induced by η , defined on $\Theta \times \mathbb{R}_+$, where Θ is a fixed nonempty subset of \mathbb{R}^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

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

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 paper.

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 $\mathbf{1}_{\mathcal{B}}$ the indicator function of \mathcal{B} and $\bar{co}(\mathcal{B})$ the closure convex hull of \mathcal{B} and $Sgn(\cdot)$ 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 $\mathbb{L}_{\mathcal{F}}^2([s, T], \mathbb{R}^n)$ 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 $\Phi_x(t) = \frac{\partial \Phi}{\partial x}(t, x(t), \mathbb{E}(x(t)), u(t))$, $\Phi_{xx}(t) = \frac{\partial^2 \Phi}{\partial x^2}(t, x(t), \mathbb{E}(x(t)), u(t))$.

Basic Assumptions Throughout this paper we assume the following.

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

$$|\varphi(t, x, y, u) - \varphi(t, x', y', u)| + |\varphi_x(t, x, y, u) - \varphi_x(t, x', y', u)| \leq C [|x - x'| + |y - y'|]. \tag{3}$$

$$|\varphi(t, x, y, u)| \leq C (1 + |x| + |y|). \tag{4}$$

$$\sup_{\theta \in \Theta} |g(t, x, u, \theta) - g(t, x', u, \theta)| + \sup_{\theta \in \Theta} |g_x(t, x, u, \theta) - g_x(t, x', u, \theta)| \leq C |x - x'| \tag{5}$$

$$\sup_{\theta \in \Theta} |g(t, x, u, \theta)| \leq C (1 + |x|). \tag{6}$$

Assumption (H2) The function $h : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is twice continuously differentiable in (x, y) , and there exists a constant $C > 0$ such that

$$|h(x, y) - h(x', y')| + |h_x(x, y) - h_x(x', y')| \leq C [|x - x'| + |y - y'|]. \tag{7}$$

$$|h(x, y)| \leq C (1 + |x| + |y|). \tag{8}$$

Under the above assumptions, the SDE-(1) has a unique strong solution $x^u(t)$ which is given by

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

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

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

where $C(q)$ is a constant depending only on q and the functional $J^{s, \xi}$ 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 Tang and Li [9].

Definition 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*

$$\left\{ \begin{aligned} -d\Psi(t) &= \{f_x^*(t, x(t), \mathbb{E}(x(t), u(t))) \Psi(t) \\ &+ \mathbb{E} [f_y^*(t, x(t), \mathbb{E}(x(t), u(t))) \Psi(t)] \\ &+ \sigma_x^*(t, x(t), \mathbb{E}(x(t), u(t))) K(t) \\ &+ \mathbb{E} [\sigma_y^*(t, x(t), \mathbb{E}(x(t), u(t))) K(t)] \\ &+ \ell_x(t, x(t), \mathbb{E}(x(t), u(t))) + \mathbb{E} [\ell_y(t, x(t), \mathbb{E}(x(t), u(t)))] \\ &+ \int_{\Theta} g_x^*(t, x(t^-), u(t), \theta) \gamma_t(\theta) \mu(d\theta) \} dt \\ &- K(t) dW(t) - \int_{\Theta} \gamma_t(\theta) N(dt, d\theta) \\ \Psi(T) &= h_x(x(T), \mathbb{E}(x(T))) + \mathbb{E} [h_y(x(T), \mathbb{E}(x(T)))] \end{aligned} \right. \tag{9}$$

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

$$\left\{ \begin{aligned} -dQ(t) &= \{f_x^*(t, x(t), \mathbb{E}(x(t), u(t))) Q(t) \\ &+ Q_t f_x^*(t, x(t), \mathbb{E}(x(t), u(t))) \\ &+ \sigma_x^*(t, x(t), \mathbb{E}(x(t), u(t))) Q(t) \sigma_x^*(t, x(t), \mathbb{E}(x(t), u(t))) \\ &+ \sigma_x^*(t, x(t), \mathbb{E}(x(t), u(t))) R(t) \\ &+ R(t) \sigma_x(t, x(t), \mathbb{E}(x(t), u(t))) \\ &- \int_{\Theta} g_x^*(t, x(t^-), u(t), \theta) (\Gamma_t(\theta) + Q(t)) \\ &g_x(t, x(t^-), u(t), \theta) \mu(d\theta) - \int_{\Theta} \Gamma_t(\theta) g_x(t, x(t^-), u(t), \theta) \\ &+ g_x^*(t, x(t^-), u(t), \theta) \Gamma_t(\theta) \mu(d\theta) \\ &- H_{xx}(t, x(t), \mathbb{E}(x(t), u(t)), \Psi(t), K(t), \gamma_t(\theta)) \} dt \\ &- R(t) dW(t) - \int_{\Theta} \Gamma_t(\theta) N(dt, d\theta) \\ Q(T) &= h_{xx}(x(T), \mathbb{E}(x(T))), \end{aligned} \right. \tag{10}$$

As it is well known that under conditions **(H1)** and **(H2)** the first-order adjoint equation (7) admits one and only one \mathcal{F}_t -adapted solution pair $(\Psi(\cdot), K(\cdot), \gamma(\cdot)) \in \mathbb{L}_{\mathcal{F}}^2([s, T]; \mathbb{R}^n) \times \mathbb{L}_{\mathcal{F}}^2([s, T]; \mathbb{R}^{n \times d}) \times \mathbb{L}_{\mathcal{F}}^2([s, T]; \mathbb{R}^{n \times m})$. 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 (8) admits one and only one \mathcal{F}_t -adapted solution pair $(Q(\cdot), R(\cdot), \Gamma(\cdot)) \in \mathbb{L}_{\mathcal{F}}^2([s, T]; \mathbb{R}^{n \times n}) \times \mathbb{L}_{\mathcal{F}}^2([s, T]; (\mathbb{R}^{n \times n})^d) \times \mathbb{L}_{\mathcal{F}}^2([s, T]; (\mathbb{R}^{n \times n})^m)$. Moreover, since $f_x, f_y, \sigma_x, \sigma_y, \ell_x, \ell_y$ and h_x are bounded, by C by assumptions **(H1)** and **(H2)**, we have the following estimate

$$\mathbb{E} \left[\sup_{s \leq t \leq T} |\Psi(t)|^2 + \int_s^T |K(t)|^2 dt + \int_s^T \int_{\Theta} |\gamma_t(\theta)|^2 \mu(d\theta) dt + \sup_{s \leq t \leq T} |Q(t)|^2 + \int_s^T |R(t)|^2 dt + \int_s^T \int_{\Theta} |\Gamma_t(\theta)|^2 \mu(d\theta) dt \right] \leq C. \tag{11}$$

Definition 2.2 (*Usual Hamiltonian and \mathcal{H} -function*). We define the usual Hamiltonian associated with the mean-field stochastic control problem (3)–(4) as follows

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

where $(t, X, u) \in [s, T] \times \mathbb{R}^n \times \mathbb{A}$ 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

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

This shows that

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

where $\Psi(t), K(t), \gamma_t(\theta)$ and $Q(t)$ are determined by adjoint equations (9) and (10) corresponding to $(z(\cdot), v(\cdot))$.

Before concluding this section, let us recall the definition of near-optimal controls as given in Zhou ([1], Definitions (2.1)–(2.2)), and Ekeland’s variational principle, which will be used in the sequel.

Definition 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) iff

$$|J^{s,\zeta}(u^\varepsilon(\cdot)) - V(s, \zeta)| \leq \mathcal{O}(\varepsilon), \tag{12}$$

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

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

Lemma 2.1 (Ekeland’s Variational Principle [26]) *Let (F, d_F) be a complete metric space and $f : F \rightarrow \overline{\mathbb{R}}$ be a lower semi-continuous function which is bounded from below. For a given $\varepsilon > 0$, suppose that $u^\varepsilon \in F$ satisfying $f(u^\varepsilon) \leq \inf_{u \in F} f(u) + \varepsilon$. Then for any $\delta > 0$, there exists $u^\delta \in F$ such that*

1. $f(u^\delta) \leq f(u^\varepsilon)$.
2. $d_F(u^\delta, u^\varepsilon) \leq \delta$.
3. $f(u^\delta) \leq f(u) + \frac{\varepsilon}{\delta}d_F(u, u^\delta)$, 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

$$\begin{aligned} d(u(\cdot), v(\cdot)) &= \mathbb{P} \otimes dt \{(w, t) \in \Omega \times [s, T] : u(w, t) \neq v(w, t)\}, \end{aligned} \tag{13}$$

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 ([27], 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} .

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 follows the general ideas as in [1, 9].

The following theorem constitutes the main contribution of this paper.

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 (7) and (8) respectively, corresponding to $u^\varepsilon(\cdot)$.

Theorem 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*

$$\begin{aligned} &\mathbb{E} \int_s^T \left\{ \frac{1}{2} [\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \right. \\ &\quad - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))]^* Q^\varepsilon(t) [\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ &\quad - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] + \Psi^\varepsilon(t) [f(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ &\quad - f(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] + K^\varepsilon(t) [\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ &\quad - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] + \int_{\Theta} \gamma^\varepsilon(t) g(t, x^\varepsilon(t), u, \theta) \\ &\quad - g(t, x^\varepsilon(t), u^\varepsilon(t), \theta) \mu(d\theta) + \frac{1}{2} \int_{\Theta} (g^*(t, x^\varepsilon(t), u, \theta) \\ &\quad - g^*(t, x^\varepsilon(t), u^\varepsilon(t), \theta)) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\ &\quad \times (g(t, x^\varepsilon(t), u, \theta) - g(t, x^\varepsilon(t), u^\varepsilon(t), \theta)) \mu(d\theta), \\ &\quad \left. + \ell(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) - \ell(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t)) \right\} dt \\ &\geq -C\varepsilon^\delta. \end{aligned} \tag{14}$$

Corollary 3.1 *Under the assumptions of Theorem 3.1 it holds that*

$$\begin{aligned} &\mathbb{E} \int_s^T \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t)) dt \\ &\geq \sup_{u(\cdot) \in \mathcal{U}} \mathbb{E} \int_s^T \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t)) dt - C\varepsilon^\delta. \end{aligned} \tag{15}$$

To prove Theorem 3.1 and Corollary 3.1, we need the following auxiliary results on the stability of the state and adjoint processes with respect to the control variable.

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 3.1 *If $x^u(t)$ and $x^v(t)$ be the solution of the state equation (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 $C = C(T, \alpha, \beta, \mu(\Theta))$ such that*

$$\mathbb{E} \left(\sup_{s \leq t \leq T} |x^u(t) - x^v(t)|^{2\beta} \right) \leq Cd^{\alpha\beta}(u(\cdot), v(\cdot)). \tag{16}$$

Proof Case 1. First, we assume that $\beta \geq 1$. Using Burkholder–Davis–Gundy inequality for the martingale part and Proposition 6.2 (see “Appendix”) we can compute, for any $r \geq s$:

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

where

$$\begin{aligned} I_1 & \leq C \mathbb{E} \int_s^r \left\{ |f(x^u(t), \mathbb{E}(x^u(t)), u(t)) \right. \\ & \quad - f(x^u(t), \mathbb{E}(x^u(t)), v(t))|^{2\beta} \\ & \quad + |\sigma(x^u(t), \mathbb{E}(x^u(t)), u(t)) - \sigma(x^u(t), \mathbb{E}(x^u(t)), v(t))|^{2\beta} \\ & \quad + \mu(\Theta) \sup_{\theta \in \Theta} |g(t, x^u(t), u(t), \theta) \\ & \quad \left. - g(t, x^v(t), v(t), \theta)|^{2\beta} \right\} \mathbf{1}_{\{u(t) \neq v(t)\}}(t) dt \end{aligned}$$

and

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

Now arguing as in ([1], 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

$$\begin{aligned} & \mathbb{E} \int_s^r |f(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & \quad - f(x^{u,\eta}(t), \mathbb{E}(x^u(t)), v(t))|^{2\beta} \mathbf{1}_{\{u(t) \neq v(t)\}}(t) dt \end{aligned}$$

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

by using definition of d and linear growth condition on f with respect to x and y , (Assumption 4) we obtain

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

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

$$\begin{aligned} & \mathbb{E} \int_s^r |\sigma(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & \quad - \sigma(t, x^u(t), \mathbb{E}(x^u(t)), v(t))|^{2\beta} \mathbf{1}_{\{u(t) \neq v(t)\}}(t) dt \\ & \leq Cd(u(\cdot), v(\cdot))^{\alpha\beta}. \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} \int_s^r \left(\sup_{\theta \in \Theta} |g(t, x^u(t), u, \theta) - g(t, x^v(t), v(t), \theta)| \right)^{2\beta} \\ & \quad \times \mathbf{1}_{\{u(t) \neq v(t)\}}(t) dt \leq Cd(u(\cdot), v(\cdot))^{\alpha\beta}. \end{aligned}$$

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

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

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

Hence (16) 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

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

This completes the proof of Lemma 3.1. \square

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 [1] to mean-field SDEs with jump processes.

Lemma 3.2 *For any $\alpha \in (0, 1)$ and $\beta \in (1, 2)$ satisfying $(1 + \alpha)\beta < 2$, there exist a positive constant $C = C(\alpha, \beta, \mu(\Theta))$ 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 (9)–(10), it holds that*

$$\mathbb{E} \int_s^T (|\Psi^u(t) - \Psi^v(t)|^\beta + |K^u(t) - K^v(t)|^\beta) dt + \mathbb{E} \int_s^T \int_\Theta |\gamma_t^u(\theta) - \gamma_t^v(\theta)|^\beta \mu(d\theta) dt \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}, \tag{17}$$

and

$$\mathbb{E} \int_s^T (|Q^u(t) - Q^v(t)|^\beta + |R^u(t) - R^v(t)|^\beta) dt + \mathbb{E} \int_s^T \int_\Theta |\Gamma_t^u(\theta) - \Gamma_t^v(\theta)|^\beta \mu(d\theta) dt \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \tag{18}$$

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

$$\left\{ \begin{aligned} -d\tilde{\Psi}(t) &= \left[f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \tilde{\Psi}(t) \right. \\ &+ \sigma_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \tilde{K}(t) \\ &+ \int_\Theta g_x^*(t, x^u(t), u, \theta) \tilde{\gamma}_t(\theta) \mu(d\theta) + \mathcal{L}(t) \left. \right] dt \\ &- \tilde{K}(t) dW(t) - \int_\Theta \tilde{\gamma}_t(\theta) N(d\theta, dt) \\ \tilde{\Psi}(T) &= h_x(x^u(T), \mathbb{E}(x^u(T))) - h_x(x^v(T), \mathbb{E}(x^v(T))) \\ &+ \mathbb{E}[h_y(x^u(T), \mathbb{E}(x^u(T))) - h_y(x^v(T), \mathbb{E}(x^v(T)))] \end{aligned} \right. \tag{19}$$

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

$$\begin{aligned} \mathcal{L}(t) &= [f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ &- f_x^*(t, x^v(t), \mathbb{E}(x^v(t)), v(t))] \Psi^v(t) \\ &+ [\sigma_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ &- \sigma_x^*(t, x^v(t), \mathbb{E}(x^v(t)), v(t))] K^v(t) \\ &+ [\ell_x(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ &- \ell_x(t, x^v(t), \mathbb{E}(x^v(t)), v(t))] \end{aligned}$$

$$\begin{aligned} &- \ell_x(t, x^v(t), \mathbb{E}(x^v(t)), v(t))] \\ &+ \mathbb{E}[f_y^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \Psi^u(t) \\ &- f_y^*(t, x^v(t), \mathbb{E}(x^v(t)), v(t)) \Psi^v(t)] \\ &+ \mathbb{E}[\sigma_y^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) K^u(t) \\ &- \sigma_y^*(t, x^v(t), \mathbb{E}(x^v(t)), v(t)) K^v(t)] \\ &+ \mathbb{E}[\ell_y(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ &- \ell_y(t, x^v(t), \mathbb{E}(x^v(t)), v(t))] \\ &+ \int_\Theta g_x^*(t, x^u(t-), u, \theta) \\ &- g_x^*(t, x^v(t-), v, \theta)) \gamma_t^v(\theta) \mu(d\theta). \end{aligned} \tag{20}$$

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

$$\left\{ \begin{aligned} d\phi(t) &= [f_x(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \phi(t) \\ &+ |\tilde{\Psi}(t)|^{\beta-1} Sgn(\tilde{\Psi}(t))] dt \\ &+ [\sigma_x(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \phi(t) \\ &+ |\tilde{K}(t)|^{\beta-1} Sgn(\tilde{K}(t))] dW(t) \\ &+ [\int_\Theta g_x^*(t, x^u(t-), u, \theta) \phi(t) \\ &+ |\tilde{\gamma}_t(\theta)|^{\beta-1} Sgn(\tilde{\gamma}_t(\theta))] N(d\theta, dt), \phi(s) = 0, \end{aligned} \right. \tag{21}$$

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

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

$$\mathbb{E} \int_s^T \left\{ \left| |\tilde{\Psi}(t)|^{\beta-1} Sgn(\tilde{\Psi}(t)) \right|^2 + \left| |\tilde{K}(t)|^{\beta-1} Sgn(\tilde{K}(t)) \right|^2 \right\} dt + \mathbb{E} \int_s^T \int_\Theta \left| |\tilde{\gamma}_t(\theta)|^{\beta-1} Sgn(\tilde{\gamma}_t(\theta)) \right|^2 \mu(d\theta) dt < \infty, \tag{22}$$

then the SDE (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

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

Note that the right hand side term of the above inequality is bounded due to (9), then we get

$$\mathbb{E} \left(\sup_{s \leq t \leq T} |\phi(t)|^\eta \right) < \infty. \tag{23}$$

By applying Itô’s formula for jump processes (see Appendix Lemma 6.1) to $\tilde{\Psi}(t)\phi(t)$ on $[s, T]$ and taking expectation, we get

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

Since

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

and fact that

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

then according to (23) we deduce

$$\begin{aligned} & \mathbb{E} \int_s^T \left[|\tilde{\Psi}(t)|^\beta + |\tilde{K}(t)|^\beta + \int_{\Theta} |\tilde{\gamma}_t(\theta)|^\beta \mu(d\theta) \right] dt \\ & \leq C \mathbb{E} \int_s^T |\mathcal{L}(t)|^\beta dt \\ & \quad + C \mathbb{E} \left\{ |h_x(x^u(T), \mathbb{E}(x^u(T))) - h_x(x^v(T), \mathbb{E}(x^v(T)))|^\beta \right. \\ & \quad \left. + |\mathbb{E}(h_y(x^u(T), \mathbb{E}(x^u(T)))) - \mathbb{E}(h_y(x^v(T), \mathbb{E}(x^v(T))))|^\beta \right\}. \tag{24} \end{aligned}$$

We proceed to estimate the right hand side of (24). First noting that $\frac{\alpha\beta}{2} < 1 - \frac{\beta}{2} < 1$ then by using assumption (H2) and Lemma 3.1, we obtain

$$\begin{aligned} & \mathbb{E} |h_x(x^u(T), \mathbb{E}(x^u(T))) - h_x(x^v(T), \mathbb{E}(x^v(T)))|^\beta \\ & \leq C \mathbb{E} |x^u(T) - x^v(T)|^\beta \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \\ & \mathbb{E} |\mathbb{E}(h_y(x^u(T), \mathbb{E}(x^u(T)))) - \mathbb{E}(h_y(x^v(T), \mathbb{E}(x^v(T))))|^\beta \\ & \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \tag{25} \end{aligned}$$

Now, to prove inequality (17) it sufficient to estimate $\mathbb{E} \int_s^T |\mathcal{L}(t)|^\beta dt$. By repeatedly using Cauchy–Schwarz inequality and assumption (H2) we can estimate

$$\begin{aligned} & \mathbb{E} \int_s^T |f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & \quad - f_x^*(t, x^v(t), \mathbb{E}(x^v(t)), v(t))|^\beta |\Psi^v(t)|^\beta dt \\ & \leq C \mathbb{E} \int_s^T \left\{ |f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \right. \\ & \quad - f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), v(t))|^\beta |\Psi^v(t)|^\beta \\ & \quad \left. + |f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), v(t)) \right. \\ & \quad \left. - f_x^*(t, x^v(t), \mathbb{E}(x^v(t)), v(t))|^\beta |\Psi^v(t)|^\beta \right\} dt \\ & \leq C \mathbb{E} \int_s^T \left\{ \mathbf{1}_{\{u(t) \neq v(t)\}}(t) |\Psi^v(t)|^\beta \right. \\ & \quad \left. + [|x^u(t) - x^v(t)| + |\mathbb{E}(x^u(t)) - \mathbb{E}(x^v(t))|]^\beta |\Psi^v(t)|^\beta \right\} dt \\ & \leq C \left[\mathbb{E} \int_s^T |\Psi^v(t)|^2 dt \right]^{\frac{\beta}{2}} d(u(\cdot), v(\cdot))^{\frac{2-\beta}{2}} \\ & \quad + C \left[\mathbb{E} \int_s^T |\Psi^v(t)|^2 dt \right]^{\frac{\beta}{2}} \left[\mathbb{E} \int_s^T |x^u(t)x^v(t)|^{\frac{2\beta}{2-\beta}} dt \right]^{\frac{2-\beta}{2}}. \end{aligned}$$

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 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

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

we conclude that

$$\begin{aligned} & \mathbb{E} \int_s^T |f_x^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & - f_x^*(x^v(t), \mathbb{E}(x^v(t)), v(t))|^\beta |\Psi^v(t)|^\beta dt \\ & \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{26}$$

A similar argument shows that

$$\begin{aligned} & \mathbb{E} \int_s^T |\sigma_x(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & - \sigma_x(t, x^v(t), \mathbb{E}(x^v(t)), v(t))|^\beta |K^v(t)|^\beta dt \\ & \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}, \end{aligned} \tag{27}$$

and

$$\begin{aligned} & \mathbb{E} \int_s^T |\ell_x(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & - \ell_x(t, x^v(t), \mathbb{E}(x^{v,\xi}(t)), v(t))|^\beta dt \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{28}$$

Now, by using similar arguments developed above and (9) we get

$$\begin{aligned} & \mathbb{E} \int_s^T \left| \mathbb{E} \left\{ [f_y^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \right. \right. \\ & \quad \left. \left. - f_y^*(x^v(t), \mathbb{E}(x^v(t)), v(t))] \Psi^v(t) \right\} \right|^\beta dt \\ & \leq C \mathbb{E} \int_s^T \mathbb{E} f_y^* | (t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & \quad - f_y^*(x^v(t), \mathbb{E}(x^v(t)), v(t)) |^\beta \mathbb{E} [|\Psi^v(t)|]^\beta dt \\ & \leq C \mathbb{E} \int_s^T \mathbb{E} |f_y^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \\ & \quad - f_y^*(x^v(t), \mathbb{E}(x^v(t)), v(t))|^\beta dt \\ & \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{29}$$

A similar argument shows that

$$\begin{aligned} & \mathbb{E} \int_s^T \left| \mathbb{E} \{ [\sigma_y^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \right. \\ & \quad \left. - \sigma_y^*(x^v(t), \mathbb{E}(x^v(t)), v(t))] \Psi^v(t) \right\}^\beta dt \\ & \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}, \end{aligned} \tag{30}$$

$$\begin{aligned} & \mathbb{E} \int_s^T \left| \mathbb{E} \{ [f_y^*(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \right. \\ & \quad \left. - f_y^*(x^v(t), \mathbb{E}(x^v(t)), v(t))] \Psi^v(t) \right\}^\beta dt \\ & \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}, \end{aligned} \tag{31}$$

and

$$\begin{aligned} & \mathbb{E} \int_s^T \left| \mathbb{E} [\ell_y(t, x^u(t), \mathbb{E}(x^u(t)), u(t)) \right. \\ & \quad \left. - \ell_y(x^v(t), \mathbb{E}(x^v(t)), v(t))] \right|^\beta dt \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{32}$$

Next, by applying Cauchy–Schwarz inequality, we get

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

where

$$\begin{aligned} \mathbb{I}_1 &= \mathbb{E} \int_s^T \left| \int_{\Theta} (g_x^*(t, x^u(t_-), u(t), \theta) \right. \\ & \quad \left. - g_x^*(t, x^u(t_-), v(t), \theta)) \gamma_t^v(\theta) \mu(d\theta) \right|^\beta \mathbf{1}_{\{u(t) \neq v(t)\}}(t) dt, \end{aligned}$$

and

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

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 (11) we get

$$\begin{aligned} \mathbb{I}_1 &\leq C\mathbb{E} \left\{ \int_s^T \int_{\Theta} |\gamma_t^v(\theta)|^2 \mu(d\theta) \right\}^{\frac{\beta}{2}} \\ &\quad \times \left\{ \int_s^T \mathbf{1}_{\{u(t) \neq v(t)\}}(t) dt \right\}^{1-\frac{\beta}{2}} \\ &\leq C\mathbb{E} \left\{ \int_s^T \int_{\Theta} |\gamma_t^v(\theta)|^2 \mu(d\theta) \right\}^{\frac{\beta}{2}} d(u(\cdot), v(\cdot))^{1-\frac{\beta}{2}} \\ &\leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{33}$$

Further, since $\frac{\alpha\beta}{2-\beta} < 1$ we conclude from Lemma 3.1 and (11) that

$$\begin{aligned} \mathbb{I}_2 &\leq C\mathbb{E} \left(\int_s^T |x^u(t) - x^v(t)|^{\frac{2\beta}{2-\beta}} dt \right)^{1-\frac{\beta}{2}} \\ &\quad \mathbb{E} \left(\int_s^T \left| \int_{\Theta} \gamma_t^v(\theta) \mu(d\theta) \right|^2 dt \right)^{\frac{\beta}{2}} \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}, \end{aligned} \tag{34}$$

It follows from (33) and (34) that

$$\begin{aligned} \mathbb{E} \int_s^T \left| \int_{\Theta} (g_x^*(t, x^u(t), u(t), \theta) - g_x^*(t, x^v(t), v(t), \theta)) \right. \\ \left. \times \gamma_t^v(\theta) \mu(d\theta) \right|^{\beta} dt \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \end{aligned} \tag{35}$$

We conclude from (26)–(35) that

$$\mathbb{E} \int_s^T |\mathcal{L}(t)|^{\beta} dt \leq Cd(u(\cdot), v(\cdot))^{\frac{\alpha\beta}{2}}. \tag{36}$$

Finally, combining (24)–(25) and (36), the proof of (17) is complete. Similarly one can prove (19). This completes the proof of Lemma 3.2. \square

Now, let $(\bar{\Psi}^\varepsilon(\cdot), \bar{K}^\varepsilon(\cdot), \bar{\gamma}^\varepsilon(\cdot))$ and $(\bar{Q}^\varepsilon(\cdot), \bar{R}^\varepsilon(\cdot), \bar{\Gamma}^\varepsilon(\cdot))$ be the solution of adjoint equations (9)–(10) corresponding to $(\bar{x}^\varepsilon(\cdot), \mathbb{E}(\bar{x}^\varepsilon(\cdot)), \bar{u}^\varepsilon(\cdot))$.

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

$$\begin{aligned} &\mathbb{E} \int_s^T \left\{ \frac{1}{2} (\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t)))^* \right. \\ &\quad \times \bar{Q}^\varepsilon(t) (\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))) \\ &\quad + \bar{\Psi}^\varepsilon(t) (f(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - f(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))) \\ &\quad + \bar{K}^\varepsilon(t) (\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))) \\ &\quad + \int_{\Theta} \bar{\gamma}^\varepsilon(t) g(t, \bar{x}^\varepsilon(t-), u, \theta) - g(t, \bar{x}^\varepsilon(t-), \bar{u}^\varepsilon(t), \theta) \mu(d\theta) \\ &\quad + \frac{1}{2} \int_{\Theta} (g^*(t, \bar{x}^\varepsilon(t-), u, \theta) - g^*(t, \bar{x}^\varepsilon(t-), \bar{u}^\varepsilon(t), \theta)) \\ &\quad \left. (\bar{Q}^\varepsilon(t) + \bar{\gamma}_t^\varepsilon(\theta)) (g(t, \bar{x}^\varepsilon(t-), u, \theta) - g(t, \bar{x}^\varepsilon(t-), \bar{u}^\varepsilon(t), \theta)) \mu(d\theta), \right. \\ &\quad \left. + (\ell(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - \ell(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))) \right\} dt \\ &\geq -\varepsilon^{\frac{1}{3}}, \end{aligned} \tag{37}$$

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

$$d(u^\varepsilon(\cdot), \bar{u}^\varepsilon(\cdot)) \leq \varepsilon^{\frac{2}{3}}, \tag{38}$$

and

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

Notice that $u^\varepsilon(\cdot)$ which is near-optimal for the initial cost $J^{s,\zeta}$ defined in (2) is an optimal control for the new cost $J^{s,\zeta,\varepsilon}$ given by

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

Therefore we have

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

Next, we use the spike variation techniques for $\bar{u}^\varepsilon(\cdot)$ 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 $\bar{u}^\varepsilon(\cdot)$:

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

where u is an arbitrary element of \mathbb{A} be fixed. By using the fact that $J^{s,\zeta,\varepsilon}(\bar{u}^\varepsilon(\cdot)) \leq J^{s,\zeta,\varepsilon}(\bar{u}^{\varepsilon,\hbar}(\cdot))$, and $d(\bar{u}^{\varepsilon,\hbar}(\cdot), \bar{u}^\varepsilon(\cdot)) = d(\bar{u}^{\varepsilon,\hbar}(\cdot), \bar{u}^\varepsilon(\cdot)) \leq \hbar$, we get

$$\begin{aligned} J^{s,\zeta}(\bar{u}^{\varepsilon,\hbar}(\cdot)) - J^{s,\zeta}(\bar{u}^\varepsilon(\cdot)) &\geq -\varepsilon^{1/3} d(\bar{u}^\varepsilon(\cdot), \bar{u}^{\varepsilon,\hbar}(\cdot)) \\ &\geq -\varepsilon^{1/3} \hbar. \end{aligned} \tag{39}$$

Arguing as in Hafayed and Abbas ([17], Theorem 3.1), the left-hand side of (39) is equal to

$$\begin{aligned} & \mathbb{E} \int_{\mathcal{E}_h} \left\{ \frac{1}{2} [\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \right. \\ & - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))]^* \times \bar{Q}^\varepsilon(t) \\ & [\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))] \\ & + \bar{\Psi}^\varepsilon(t)[f(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \\ & - f(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))] \\ & + \bar{K}^\varepsilon(t)[\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \\ & - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))] \\ & + \int_{\Theta} \bar{\gamma}^\varepsilon(t) g(t, \bar{x}^\varepsilon(t_-), u, \theta) \\ & - g(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta) \mu(d\theta) \\ & + \frac{1}{2} \int_{\Theta} [g^*(t, \bar{x}^\varepsilon(t_-), u, \theta) - g^*(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta)] \\ & \times (\bar{Q}^\varepsilon(t) + \bar{\gamma}_t^\varepsilon(\theta)) [g(t, \bar{x}^\varepsilon(t_-), u, \theta) \\ & - g(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta)] \mu(d\theta) + [\ell(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \\ & - \ell(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))] \Big\} dt + \tau(\hbar), \end{aligned} \tag{40}$$

where $\tau(\hbar) \rightarrow 0$ as $\hbar \rightarrow 0$. Finally, replacing (40) in (39), then dividing inequality (39) by \hbar and sending \hbar to zero, the near-maximum condition (37) follows. \square

Proof of Theorem 3.1 First, we are about to derive an estimate for the term similar to the left side of inequality (34) and (35) with all the $(\bar{x}^\varepsilon(\cdot), \mathbb{E}(\bar{x}^\varepsilon(\cdot)), \bar{u}^\varepsilon(\cdot))$ etc. replaced by $(x^\varepsilon(\cdot), \mathbb{E}(x^\varepsilon(\cdot)), u^\varepsilon(\cdot))$ etc.,

Now, to prove (14) it remains to estimate the following differences

$$\begin{aligned} S_1(\varepsilon) = & \mathbb{E} \int_s^T \left[\bar{K}^\varepsilon(t) (\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \right. \\ & - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))) \\ & - K^\varepsilon(t) (\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ & \left. - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))) \right] dt, \end{aligned} \tag{41}$$

$$\begin{aligned} S_2(\varepsilon) = & \mathbb{E} \int_s^T \left\{ \frac{1}{2} (\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \right. \\ & - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t)))^* \bar{Q}^\varepsilon(t) \\ & \times (\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \\ & - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))) \\ & - \frac{1}{2} (\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ & - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t)))^* \\ & \left. \times Q^\varepsilon(t) [\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \right. \end{aligned}$$

$$\begin{aligned} & - \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] \\ & + \bar{\Psi}^\varepsilon(t) [f(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \\ & - f(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))] \\ & - \Psi^\varepsilon(t) [f(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ & - f(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] \\ & + [\ell(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \\ & - \ell(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))] \\ & - [\ell(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\ & - \ell(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] \Big\} dt. \end{aligned} \tag{42}$$

and

$$\begin{aligned} S_3(\varepsilon) = & \mathbb{E} \int_s^T \int_{\Theta} [\bar{\gamma}_t^\varepsilon(\theta) (g(t, \bar{x}^\varepsilon(t_-), u, \theta) \\ & - g(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta)) - \gamma_t^\varepsilon(\theta) (g(t, x^\varepsilon(t_-), u, \theta) \\ & - g(t, x^\varepsilon(t_-), u^\varepsilon(t), \theta))] \mu(d\theta) dt, \end{aligned} \tag{43}$$

Then we have

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

We estimate the first term on the right-hand side $\mathbb{I}_1(\varepsilon) = \mathbb{E} \int_s^T [\bar{K}^\varepsilon(t) - K^\varepsilon(t)] [\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))]$. 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 3.2 and note (4) we obtain

$$\begin{aligned} \mathbb{I}_1(\varepsilon) \leq & \left[\mathbb{E} \int_s^T |\bar{K}^\varepsilon(t) - K^\varepsilon(t)|^\beta dt \right]^{\frac{1}{\beta}} \\ & \times \left[\mathbb{E} \int_s^T |\sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), u) \right. \\ & \left. - \sigma(t, \bar{x}^\varepsilon(t), \mathbb{E}(\bar{x}^\varepsilon(t)), \bar{u}^\varepsilon(t))|^q dt \right]^{\frac{1}{q}} \end{aligned}$$

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

We estimate now the second term $\mathbb{I}_2(\varepsilon)$. Then by applying Cauchy–Schwarz inequality, note (9), assumption (H1), and Lemma 3.1, we get

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

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

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

then by using Cauchy–Schwarz inequality, we have

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

We proceed as in $I_2(\varepsilon)$ to estimate the second term in the right of above inequality, then by applying Cauchy–Schwarz inequality, Assumption (H1) and (9) we obtain

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

thus, we have proved that

$$S_1(\varepsilon) = \mathbb{I}_1(\varepsilon) + \mathbb{I}_2(\varepsilon) + \mathbb{I}_3(\varepsilon) \leq C\varepsilon^\delta. \tag{44}$$

By using similar arguments developed above, we can prove that

$$S_2(\varepsilon) \leq C\varepsilon^\delta. \tag{45}$$

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

$$\begin{aligned} S_3(\varepsilon) &\leq \mathbb{E} \int_s^T \int_{\Theta} (\bar{\gamma}_t^\varepsilon(\theta) - \gamma_t^\varepsilon(\theta)) \\ &\quad \times (g(t, \bar{x}^\varepsilon(t_-), u, \theta) - g(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta)) \mu(d\theta) dt \\ &\quad + \mathbb{E} \int_s^T \int_{\Theta} [\gamma_t^\varepsilon(\theta) [g(t, \bar{x}^\varepsilon(t_-), u, \theta) \\ &\quad - g(t, x^\varepsilon(t_-), u)] \mu(d\theta) dt \\ &\quad + \mathbb{E} \int_s^T \int_{\Theta} \gamma_t^\varepsilon(\theta) (g(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta) \\ &\quad - g(t, x^\varepsilon(t_-), u^\varepsilon(t), \theta)) \mu(d\theta) dt, \\ &= \mathbb{J}_1(\varepsilon) + \mathbb{J}_2(\varepsilon) + \mathbb{J}_3(\varepsilon). \end{aligned}$$

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 $\frac{1}{\beta} + \frac{1}{q} = 1$. By Hölder’s inequality, Lemma 3.2 and (5) we obtain

$$\begin{aligned} \mathbb{J}_1(\varepsilon) &= \mathbb{E} \int_s^T \int_{\Theta} (\bar{\gamma}_t^\varepsilon(\theta) - \gamma_t^\varepsilon(\theta)) \\ &\quad \times (g(t, \bar{x}^\varepsilon(t_-), u, \theta) - g(t, \bar{x}^\varepsilon(t_-), \bar{u}^\varepsilon(t), \theta)) \mu(d\theta) dt \\ &\leq \left[\mathbb{E} \int_s^T \int_{\Theta} |\bar{\gamma}_t^\varepsilon(\theta) - \gamma_t^\varepsilon(\theta)|^\beta \mu(d\theta) dt \right]^{\frac{1}{\beta}} \\ &\quad \times \mathbb{E} \int_s^T ((\sup_{\theta \in \Theta} |g(t, \bar{x}^\varepsilon(t), u, \theta)| \end{aligned}$$

$$\begin{aligned} & - g(t, \bar{x}^\varepsilon(t), \bar{u}^\varepsilon(t), \theta)|^q dt \Big\}^{\frac{1}{q}} \mu(\Theta)^{\frac{1}{q}} \\ & \leq C \left[d(\bar{u}^\varepsilon(\cdot), u^\varepsilon(\cdot))^{\frac{\alpha\beta}{2}} \right]^{\frac{1}{\beta}} \\ & \quad \times \left[\mathbb{E} \int_s^T (1 + |\bar{x}^\varepsilon(t)|^q + |\mathbb{E}(\bar{x}^\varepsilon(t))|^q) dt \right]^{\frac{1}{q}} \\ & \leq C(\varepsilon^{\frac{2}{3}})^{\frac{\alpha\beta}{2} \cdot \frac{1}{\beta}} = C\varepsilon^{\frac{\alpha}{3}}. \end{aligned}$$

Applying assumption **(H3)**, Cauchy–Schwarz inequality, Lemma 3.2, note (10) and the fact that $\mu(\Theta) < \infty$ we get

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

by using (38) we get $d(\bar{u}^\varepsilon(\cdot), u^\varepsilon(\cdot))^\alpha \leq (\varepsilon^{\frac{2}{3}})^\alpha$, it holds that

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

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

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

by applying Cauchy–Schwarz inequality, Lemma 3.2 and (11) it follows that

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

Thus, we have proved that

$$S_3(\varepsilon) = \mathbb{J}_1(\varepsilon) + \mathbb{J}_2(\varepsilon) + \mathbb{J}_3(\varepsilon) \leq C\varepsilon^\delta. \tag{46}$$

The desired result (14) follows immediately by combining (44), (45), (46) and (34). This completes the proof of Theorem 3.1. \square

Proof of Corollary 3.1 In the spike variations technique for the perturbed control $\bar{u}^{\varepsilon, \theta}(\cdot)$ in (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 (15) holds for any $u(\cdot) \in \mathcal{U}$ and the subsequent argument still goes through. So the inequalities in the estimate (15) holds for any $u(\cdot) \in \mathcal{U}$. \square

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

We will show 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

$$\begin{aligned} & |\psi(t, x, y, u) - \psi(t, x, y, u')| \\ & \quad + |\psi_u(t, x, y, u) - \psi_u(t, x, y, u')| \leq C |u - u'|, \\ & \quad \sup_{\theta \in \Theta} |g(t, x, u, \theta) - g(t, x, u', \theta)| \\ & \quad + \sup_{\theta \in \Theta} |g_u(t, x, u, \theta) - g_u(t, x, u', \theta)| \leq C |u - u'|. \end{aligned} \tag{47}$$

$$h(\cdot, \cdot) \text{ convex with respect to } (x, y). \tag{48}$$

$$H(t, \cdot, \cdot, \cdot, \Psi^\varepsilon(\cdot), K^\varepsilon(\cdot), \gamma^\varepsilon(\cdot)) \text{ is concave with respect to } (x, y, u), \text{ for a.e. } t \in [0, T], \mathbb{P} - a.s. \tag{49}$$

$$\text{The derivatives } f_y, \sigma_y, h_y, \ell_y \text{ are non-negative.} \tag{50}$$

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 paper.

Let $u^\varepsilon(\cdot)$ be an admissible control and $(\Psi^\varepsilon(\cdot), K^\varepsilon(\cdot), \gamma^\varepsilon(\cdot)), (Q^\varepsilon(\cdot), R^\varepsilon(\cdot), \Gamma^\varepsilon(\cdot))$ be the solution of the adjoint equations (9)–(10) corresponding to $u^\varepsilon(\cdot)$.

Theorem 4.1 *Sufficient conditions for near-optimality of order $\varepsilon^{\frac{1}{2}}$. Let conditions (47)–(49) holds. If for some $\varepsilon > 0$ and for any $u(\cdot) \in \mathcal{U}$:*

$$\begin{aligned} & \mathbb{E} \int_s^T \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t)) dt + \varepsilon \\ & \geq \sup_{u(\cdot) \in \mathcal{U}} \mathbb{E} \int_s^T \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t)) dt, \end{aligned} \tag{51}$$

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

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

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

Corollary 4.1 (Sufficient Conditions for ε -optimality) *Under the assumptions of Theorem 4.1 a sufficient condition for an admissible control $u^\varepsilon(\cdot)$ to be ε -optimal for our mean-field control problem (1)–(2) is*

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

Proof of Theorem 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 $v(\cdot) \in \mathcal{U}$:

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

where

$$\begin{aligned} \mathfrak{F}^\varepsilon(t) &= 1 + |\Psi^\varepsilon(t)| + |K^\varepsilon(t)| \\ &+ 2 |Q^\varepsilon(t)| [1 + |x^\varepsilon(t)| + |\mathbb{E}(x^\varepsilon(t))|] \\ &+ 2 \left[|Q^\varepsilon(t)| + \left| \int_{\Theta} \gamma_t^\varepsilon(\theta) \mu(d\theta) \right| \right] [1 + |x^\varepsilon(t)|]. \end{aligned}$$

Obviously \widehat{d} is a metric on \mathcal{U} satisfied $\mathfrak{F}^\varepsilon(t) > 1$, 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 (47) then a simple computation shows that

$$\begin{aligned} & |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. \\ & \quad \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)) \\ & \quad - H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))| dt \\ & \quad + \mathbb{E} \int_s^T |\sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t)) \\ & \quad - \sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v(t))| |Q^\varepsilon(t)| \\ & \quad \times |\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))| dt \\ & \quad + \frac{1}{2} \mathbb{E} \int_s^T |\sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t)) Q(t) \\ & \quad \times \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t)) \\ & \quad - \sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v(t)) Q^\varepsilon(t) \\ & \quad \times \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v(t))| dt \\ & \quad + \mathbb{E} \int_s^T \int_{\Theta} |g^*(t, x^\varepsilon(t), u(t), \theta) - g^*(t, x^\varepsilon(t), v(t), \theta)| \\ & \quad \times |(Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) g(t, x^\varepsilon(t), u^\varepsilon(t), \theta)| \mu(d\theta) dt \\ & \quad + \frac{1}{2} \mathbb{E} \int_s^T \int_{\Theta} |g^*(t, x^\varepsilon(t), u(t), \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\ & \quad \times g(t, x^\varepsilon(t), u(t), \theta) \\ & \quad - g^*(t, x^\varepsilon(t), v(t), \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\ & \quad \times g(t, x^\varepsilon(t), v(t), \theta)| \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 and assumption (H3)

$$\begin{aligned} \mathcal{I}_1^\varepsilon &= \mathbb{E} \int_s^T |H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u, \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ & \quad - H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v, \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))| dt \\ & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \left[|\Psi^\varepsilon(t)| + |K^\varepsilon(t)| \right. \end{aligned}$$

$$\begin{aligned}
 & + \left| \int_{\Theta} \gamma_t^\varepsilon(\theta) \mu(d\theta) \right| dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \mathbb{F}^\varepsilon(t) dt. \tag{52}
 \end{aligned}$$

Since σ is linear growth with respect to x and y then by using assumption (47) we get

$$\begin{aligned}
 \mathcal{I}_2^\varepsilon & = \mathbb{E} \int_s^T |\sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) \\
 & \quad - \sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v)| \\
 & \quad \times |Q^\varepsilon(t) \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))| dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| |Q^\varepsilon(t)| [1 + |x^\varepsilon(t)| + |\mathbb{E}(x^\varepsilon(t))|] dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \mathbb{F}^\varepsilon(t) dt. \tag{53}
 \end{aligned}$$

Similarly, since g is linear growth with respect to x then by assumptions (47) we can prove that

$$\begin{aligned}
 \mathcal{I}_4^\varepsilon & = \mathbb{E} \int_s^T \int_{\Theta} |g^*(t, x^\varepsilon(t), u, \theta) - g^*(t, x^\varepsilon(t), v, \theta)| \\
 & \quad \times |(Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) g(t, x^\varepsilon(t), u^\varepsilon(t), \theta)| \mu(d\theta) dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \left[|Q^\varepsilon(t)| + \left| \int_{\Theta} \gamma_t^\varepsilon(\theta) \mu(d\theta) \right| \right] \\
 & \quad \times [1 + |x^\varepsilon(t)|] dt \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \mathbb{F}^\varepsilon(t) dt. \tag{54}
 \end{aligned}$$

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

$$\begin{aligned}
 \mathcal{I}_3^\varepsilon & = \frac{1}{2} \mathbb{E} \int_s^T |\sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) Q^\varepsilon(t) \\
 & \quad \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u) - \sigma^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v) \\
 & \quad \times Q^\varepsilon(t) \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), v)| dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \frac{1}{2} |Q^\varepsilon(t)| [1 + |x^\varepsilon(t)| + |\mathbb{E}(x^\varepsilon(t))|] dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \mathbb{F}^\varepsilon(t) dt, \tag{55}
 \end{aligned}$$

and

$$\begin{aligned}
 \mathcal{I}_5^\varepsilon & = \frac{1}{2} \mathbb{E} \int_s^T \int_{\Theta} |g^*(t, x^\varepsilon(t), u, \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\
 & \quad \times g(t, x^\varepsilon(t), u, \theta) - g^*(t, x^\varepsilon(t), v, \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\
 & \quad \times g(t, x^\varepsilon(t), v, \theta)| \mu(d\theta) dt, \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \frac{1}{2} |Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)| [1 + |x^\varepsilon(t)|] dt \\
 & \leq C \mathbb{E} \int_s^T |u(t) - v(t)| \mathbb{F}^\varepsilon(t) dt, \tag{56}
 \end{aligned}$$

By combining (52)–(56) we conclude that

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

which implies that g is continuous on \mathcal{U} with respect to \widehat{d} . Now by using (51) and Ekeland’s Variational Principle (Lemma 2.1), there exists $\bar{u}^\varepsilon(\cdot) \in \mathcal{U}$ such that

$$\widehat{d}(\bar{u}^\varepsilon(\cdot), u^\varepsilon(\cdot)) \leq \sqrt{\varepsilon}, \tag{57}$$

and

$$\begin{aligned}
 & \mathbb{E} \int_s^T \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) dt \\
 & = \max_{u(\cdot) \in \mathcal{U}} \mathbb{E} \int_s^T \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u(t)) dt, \tag{58}
 \end{aligned}$$

where

$$\begin{aligned}
 \widetilde{\mathcal{H}}(t, x, y, u) & = \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x, y, u) - \sqrt{\varepsilon} |u - \bar{u}^\varepsilon(t)| \mathbb{F}^\varepsilon(t). \tag{59}
 \end{aligned}$$

The maximum condition (58) implies a pointwise maximum condition namely, for \mathbb{P} –a.s. and a.e., $t \in [s, T]$

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

Using [Item 3, Proposition 6.1], then we have

$$0 \in \partial_u^\circ \widetilde{\mathcal{H}}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)). \tag{60}$$

Since the function $u : \cdot \mapsto |u - \bar{u}^\varepsilon(t)|$ is locally Lipschitz but not differentiable in $\bar{u}^\varepsilon(t)$, then Clarke’s generalized gradient (see Proposition 6.1, Example, “Appendix”) shows that

$$\begin{aligned}
 \partial_u^\circ (\sqrt{\varepsilon} |u - \bar{u}^\varepsilon(t)| \mathbb{F}^\varepsilon(t)) & = \overline{co} \{ -\mathbb{F}^\varepsilon(t) \sqrt{\varepsilon}, \mathbb{F}^\varepsilon(t) \sqrt{\varepsilon} \} \\
 & = [-\mathbb{F}^\varepsilon(t) \sqrt{\varepsilon}, \mathbb{F}^\varepsilon(t) \sqrt{\varepsilon}]. \tag{61}
 \end{aligned}$$

By using (61) 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 6.1] we get

$$\begin{aligned} &\partial_u^\circ \tilde{H}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) \\ &\subset \partial_u^\circ \mathcal{H}^{(x^\varepsilon(\cdot), u^\varepsilon(\cdot))}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) \\ &\quad + [-\sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t), \sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t)]. \end{aligned}$$

By applying assumption (47), the Hamiltonian H is differentiable in u , then [Item 4, Proposition 6.1] shows that

$$\begin{aligned} &\partial_u^\circ \tilde{H}(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) \\ &\subset \left\{ \begin{aligned} &H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &+ \{ \sigma_u^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) Q^\varepsilon(t) \\ &\times (\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))) \\ &- \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) \} \\ &+ \int_{\Theta} g_u^*(t, x^\varepsilon(t-), \bar{u}^\varepsilon(t), \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\ &\times (g(t, x^\varepsilon(t), u^\varepsilon(t), \theta) - g(t, x^\varepsilon(t), \bar{u}^\varepsilon(t), \theta)) \mu(d\theta) \} \\ &+ [-\sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t), \sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t)]. \end{aligned} \right. \end{aligned}$$

Next, the differential inclusion (60) implies that there is

$$\tau^\varepsilon(t) \in [-\sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t), \sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t)],$$

such that

$$\begin{aligned} &H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &+ \sigma_u^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) Q^\varepsilon(t) \\ &\times (\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))) \\ &- \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) \\ &+ \int_{\Theta} g_u^*(t, x^\varepsilon(t), \bar{u}^\varepsilon(t), \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \\ &\times (g(t, x^\varepsilon(t), u^\varepsilon(t), \theta) - g(t, x^\varepsilon(t), \bar{u}^\varepsilon(t), \theta)) \mu(d\theta) \} \\ &+ \tau^\varepsilon(t) = 0. \end{aligned} \tag{62}$$

By using assumption (47) we can prove that

$$\begin{aligned} &|H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &- H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))| \\ &\leq C |u^\varepsilon(t) - \bar{u}^\varepsilon(t)| \mathfrak{F}^\varepsilon(t), \end{aligned} \tag{63}$$

hence from (62) and (63), assumption (47) and the fact that $|\tau^\varepsilon(t)| \leq \sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t)$ we get

$$\begin{aligned} &|H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))| \\ &\leq C |u^\varepsilon(t) - \bar{u}^\varepsilon(t)| \mathfrak{F}^\varepsilon(t) \\ &+ |\sigma_u^*(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t)) Q^\varepsilon(t) \\ &\times (\sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))) \end{aligned}$$

$$\begin{aligned} &- \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), \bar{u}^\varepsilon(t))| \\ &+ \left| \int_{\Theta} g_u^*(t, x^\varepsilon(t-), \bar{u}^\varepsilon(t), \theta) (Q^\varepsilon(t) + \gamma_t^\varepsilon(\theta)) \right. \\ &\times (g(t, x^\varepsilon(t-), u^\varepsilon(t), \theta) - g(t, x^\varepsilon(t-), \bar{u}^\varepsilon(t), \theta)) \mu(d\theta) \left. \right| \\ &+ |\tau^\varepsilon(t)| \\ &\leq C |u^\varepsilon(t) - \bar{u}^\varepsilon(t)| \mathfrak{F}^\varepsilon(t) + |\tau^\varepsilon(t)| \\ &\leq C |u^\varepsilon(t) - \bar{u}^\varepsilon(t)| \mathfrak{F}^\varepsilon(t) + \sqrt{\varepsilon}\mathfrak{F}^\varepsilon(t), \end{aligned} \tag{64}$$

Now, using (49), we obtain for any $u(\cdot) \in \mathcal{U}$

$$\begin{aligned} &H(t, x(t), \mathbb{E}(x(t)), u(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &- H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\leq H_x(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (x(t) - x^\varepsilon(t)) \\ &+ H_y(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (x(t) - x^\varepsilon(t)) \\ &+ H_u(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (u(t) - u^\varepsilon(t)). \end{aligned} \tag{65}$$

Integrating this inequality with respect to t and taking expectations we obtain from (52) and (64)

$$\begin{aligned} &\mathbb{E} \int_s^T [H(t, x(t), \mathbb{E}(x(t)), u(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &- H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta))] dt \\ &\leq \mathbb{E} \int_s^T H_x(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (x(t) - x^\varepsilon(t)) dt \\ &+ \mathbb{E} \int_s^T H_y(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (x(t) - x^\varepsilon(t)) dt \\ &+ C(\widehat{d}(u^\varepsilon(\cdot), \bar{u}^\varepsilon(\cdot)) + \varepsilon^{\frac{1}{2}}) \\ &\leq \mathbb{E} \int_s^T H_x(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (x(t) - x^\varepsilon(t)) dt \\ &+ \mathbb{E} \int_s^T H_y(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &\times (x(t) - x^\varepsilon(t)) dt + C\varepsilon^{\frac{1}{2}}. \end{aligned} \tag{66}$$

On the other hand, by using (48) we get

$$\begin{aligned} &h(x(T), \mathbb{E}(x(T))) - h(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T))) \geq \\ &[h_x(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T))) + h_y(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T)))] \\ &\times [x(T) - x^\varepsilon(T)]. \end{aligned}$$

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

$$\mathbb{E} \{h(x(T), \mathbb{E}(x(T))) - h(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T)))\} \geq \mathbb{E} \{\Psi^\varepsilon(T)(x(T) - x^\varepsilon(T))\} \tag{67}$$

By integration by parts formula for jumps process $\Psi^\varepsilon(t)(x(t) - x^\varepsilon(t))$ (see Lemma 6.1) we get

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

with the help of (1), and (9) we obtain

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

then from (49) and (66) we get

$$\begin{aligned} \mathbb{E} \{ \Psi^\varepsilon(T)(x(T) - x^\varepsilon(T)) \} &\geq \mathbb{E} \int_s^T \left\{ H(t, x(t), \mathbb{E}(x(t)), u(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \right. \\ &- H(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t), \Psi^\varepsilon(t), K^\varepsilon(t), \gamma_t^\varepsilon(\theta)) \\ &+ \Psi^\varepsilon(t)[f(t, x(t), \mathbb{E}(x(t)), u(t)) \\ &- f(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] \\ &+ K^\varepsilon(t)[\sigma(t, x(t), \mathbb{E}(x(t)), u(t)) \end{aligned}$$

$$\begin{aligned} &- \sigma(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t))] \\ &+ \int_{\Theta} \gamma_t^\varepsilon(\theta) [g(t, x(t), u(t), \theta) \\ &- g(t, x^\varepsilon(t), u^\varepsilon(t), \theta)] \mu(d\theta) \left. \right\} dt - C\varepsilon^{\frac{1}{2}} \\ &= \mathbb{E} \int_s^T [\ell(t, x^\varepsilon(t), \mathbb{E}(x^\varepsilon(t)), u^\varepsilon(t)) \\ &- \ell(t, x(t), \mathbb{E}(x(t)), u(t))] dt - C\varepsilon^{\frac{1}{2}}. \tag{68} \end{aligned}$$

Combining (67) and (68) we get

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

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

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

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

5 Application to finance: penalized mean-variance portfolio selection

In this section, we will apply our necessary and sufficient conditions of near-optimality to study a penalized mean-variance portfolio selection and we derive the explicit expression of the optimal portfolio selection strategy. Our method inspired from Zhou ([1], Example 6.1).

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

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

$$\begin{cases} dP_0(t) = P_0(t) \rho(t) dt, & t \in [0, T] \\ P_0(0) > 0, \end{cases} \tag{69}$$

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

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

$$\begin{cases} P_1(t) = P_1(t) \zeta(t) dt + \sigma_t dW(t) P_1(t) \\ \quad + P_1(t) \int_{\Theta} \xi_t(\theta) N(d\theta, dt), \\ P_1(0) > 0, \end{cases} \tag{70}$$

where $\zeta(t)$, σ_t and $\xi_t(\theta)$ are bounded deterministic functions such that $\zeta(t) \neq 0$, $\sigma_t \neq 0$ and $\zeta(t) > \rho(t)$. 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

$$x^\pi(t) = \pi_0(t)P_0(t) + \pi_1(t)P_1(t). \tag{71}$$

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

$$x^\pi(t) = x^\pi(0) + \int_0^t \pi_0(r)dP_0(r) + \int_0^t \pi_1(r)dP_1(r). \tag{72}$$

We denote by

$$v(t) = \pi_1(t)P(t), \tag{73}$$

the amount invested in the risky security. Now, by combining (71) and (72) together with (73) we introduce the wealth dynamics as follows

$$\begin{cases} dx^v(t) = [\rho(t)x^v(t) + (\zeta(t) - \rho(t))v(t)] dt \\ \quad + \sigma_t v(t)dW(t) + \int_{\Theta} \xi_{t-}(\theta) v(t)N(d\theta, dt), \\ x^v(0) = \zeta, \end{cases} \tag{74}$$

where $\zeta \in \mathbb{R}$. If the corresponding wealth process $x^v(\cdot)$ given by SDE-(74) is square integrable, the control variable $v(\cdot)$ is called tame. We denote \mathcal{U} the set of admissible portfolio valued in $\mathbb{A} = \mathbb{R}$.

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

$$J^{\zeta, \varepsilon}(v(\cdot)) = \mathbb{E} \left(x^v(T) - \mathbb{E}(x^v(T)) - \frac{\varepsilon}{2} \right)^2 + \int_0^T \frac{\varepsilon^2}{4} L(v(t))dt, \tag{75}$$

subject to $x^v(T)$ solution of SDE-(74) at time T given by

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

where $L(\cdot)$ is a nonlinear, convex and bounded function, satisfying assumption (47) and independent of ε .

Inspired from (Zhou [1], example 6.1), our objective is to find an admissible portfolio $v^*(\cdot)$ which minimizes the cost function (75) of mean-field type (i.e., with $\ell \equiv \frac{\varepsilon^2}{4} L(v(t)), s = 0, h(x(t), \mathbb{E}(x(t))) = (x(t) - \mathbb{E}(x(t)) - \frac{\varepsilon}{2})^2$). Explicit solution of problem (74)–(75), called \mathcal{P}_ε , 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 (75), then we get $J_0^\zeta(v(\cdot)) = \mathbb{E} \left\{ (x^v(T) - \mathbb{E}(x^v(T)))^2 \right\}$, (76)

We study the optimal control problem where the state is governed by SDE-(74) with a new cost function (76). In a second step, we solve the control problem (74)–(76), and obtain an optimal solution explicitly. Finally, inspired by Zhou ([1], Example 6.1), we solve the control problem \mathcal{P}_ε of near-optimally.

Problem \mathcal{P}_0 : (Optimal solution of mean-field stochastic control problem (74)–(76)). By a standard argument, problem \mathcal{P}_0 can be solved as follows.

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

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

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

$$\begin{aligned} \Psi^*(t)(\zeta(t) + \rho(t)) + K^*(t)\sigma_t \\ + \int_{\Theta} \gamma_t^*(\theta) \xi_t(\theta) \mu(d\theta) = 0. \end{aligned} \tag{77}$$

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 (9) associated with $v^*(t)$ gets the form

$$\begin{cases} d\Psi^*(t) = -\rho(t)\Psi^*(t)dt + K^*(t)dW(t) \\ \quad + \int_{\Theta} \gamma_t^*(\theta)N(dt, d\theta) \\ \Psi^*(T) = 2(x^*(T) - \mathbb{E}(x^*(T))). \end{cases} \tag{78}$$

In order to solve the above Eq. (78) and to find the expression of $v^*(t)$ we conjecture a process $\Psi^*(t)$ of the form

$$\Psi^*(t) = \Phi_1(t)x^*(t) + \Phi_2(t)\mathbb{E}(x^*(t)) + \Phi_3(t), \tag{79}$$

where $\Phi_1(\cdot)$, $\Phi_2(\cdot)$ and $\Phi_3(\cdot)$ are deterministic differentiable functions. (see [4, 12, 15, 22] for other models of conjecture). Applying Itô's formula to (79), in virtue of SDE-(74), we get

$$\begin{aligned}
 d\Psi^*(t) = & \Phi_1(t) \left\{ \left[\rho(t)x^*(t) + (\zeta(t) - \rho(t))v^*(t) \right] dt \right. \\
 & \left. + \sigma_t v^*(t) dW(t) + \int_{\Theta} v^*(t) \xi_{t-}(\theta) N(d\theta, dt) \right\} \\
 & + x^*(t) \dot{\Phi}_1(t) dt + \Phi_2(t) [\rho(t) \mathbb{E}(x^*(t)) \\
 & + (\zeta(t) - \rho(t))v^*(t)] dt + \mathbb{E}(x^*(t)) \dot{\Phi}_2(t) dt + \dot{\Phi}_3(t) dt \\
 = & \left\{ \Phi_1(t) [\rho(t)x^*(t) + (\zeta(t) - \rho(t))v^*(t)] + x^*(t) \dot{\Phi}_1(t) \right. \\
 & + \Phi_2(t) [\rho(t) \mathbb{E}(x^*(t)) + (\zeta(t) - \rho(t))v^*(t)] \\
 & + \dot{\Phi}_2(t) \mathbb{E}(x^*(t)) + \dot{\Phi}_3(t) \left. \right\} dt \\
 & + \Phi_1(t) \sigma_t v^*(t) dW(t) + \int_{\Theta} \Phi_1(t) v^*(t) \xi_{t-}(\theta) N(d\theta, dt), \\
 \Psi^*(T) = & \Phi_1(T)x^*(T) + \Phi_2(T) \mathbb{E}(x^*(T)) + \Phi_3(T). \tag{80}
 \end{aligned}$$

Next, comparing (80) with (78), we get

$$\begin{aligned}
 -\rho(t)\Psi^*(t) = & \Phi_1(t) [\rho(t)x^*(t) + (\zeta(t) - \rho(t))v^*(t)] \\
 & + x^*(t) \dot{\Phi}_1(t) \\
 & + \Phi_2(t) [\rho(t) \mathbb{E}(x^*(t)) + (\zeta(t) \\
 & - \rho(t))v^*(t)] \\
 & + \dot{\Phi}_2(t) \mathbb{E}(x^*(t)) + \dot{\Phi}_3(t), \tag{81}
 \end{aligned}$$

$$K^*(t) = \Phi_1(t) \sigma_t v^*(t), \tag{82}$$

$$\gamma_t^*(\theta) = \Phi_1(t) v^*(t) \xi_t(\theta), \tag{83}$$

and

$$\Phi_1(T) = 2, \Phi_2(T) = -2, \Phi_3(T) = 0. \tag{84}$$

Combining (82) and (84) together with (77) we get

$$v^*(t) = \frac{-(\zeta(t) - \rho(t))\Psi^*(t)}{\Phi_1(t) [\sigma_t^2 + \int_{\Theta} \xi_t^2(\theta) \mu(d\theta)]}. \tag{85}$$

We denote

$$A(t) = \sigma_t^2 + \int_{\Theta} \xi_t^2(\theta) \mu(d\theta), \tag{86}$$

by using (77) together with (85) and (86) then we can get

$$\begin{aligned}
 \Phi_3(t) = & 0 \text{ for } t \in [0, T], v^*(t) = (\rho(t) - \zeta(t)) (A(t))^{-1} \\
 & \frac{(\Phi_1(t)x^*(t) + \Phi_2(t) \mathbb{E}(x^*(t)))}{\Phi_1(t)}. \\
 = & \left\{ (\rho(t) - \zeta(t)) (A(t))^{-1} \right\} x^*(t) \\
 & + \left\{ (\rho(t) - \zeta(t)) (A(t))^{-1} \frac{\Phi_2(t)}{\Phi_1(t)} \right\} \mathbb{E}(x^*(t)). \tag{87}
 \end{aligned}$$

Now combining (81) with (79) we deduce

$$\begin{aligned}
 v^*(t) (\Phi_1(t) + \Phi_2(t)) (\rho(t) - \zeta(t)) \\
 = & [2\rho(t)\Phi_1(t) + \dot{\Phi}_1(t)] x^*(t) \\
 & + [2\rho(t)\Phi_2(t) + \dot{\Phi}_2(t)] \mathbb{E}(x^*(t)). \tag{88}
 \end{aligned}$$

By comparing the terms containing $x^*(t)$ and $\mathbb{E}(x^*(t))$, we obtain from (87) with (88) the two ordinary differential equations (ODEs in short):

$$\begin{aligned}
 [(\rho(t) - \zeta(t))^2 (A(t))^{-1} - 2\rho(t)] \Phi_1(t) \\
 + (\rho(t) - \zeta(t))^2 (A(t))^{-1} \Phi_2(t) = & \dot{\Phi}_1(t). \\
 [(\rho(t) - \zeta(t))^2 (A(t))^{-1} - 2\rho(t)] \Phi_2(t) \\
 + (\rho(t) - \zeta(t))^2 (A(t))^{-1} \frac{\Phi_2^2(t)}{\Phi_1(t)} = & \dot{\Phi}_2(t), \tag{89}
 \end{aligned}$$

a simple computation from (89) we obtain

$$\dot{\Phi}_1(t) \Phi_2(t) = \dot{\Phi}_2(t) \Phi_1(t), \tag{90}$$

Since $\Phi_1(T) = 2, \Phi_2(T) = -2$, (see (84)) we deduce

$$\Phi_1(t) = -\Phi_2(t), \tag{91}$$

Let us turn to calculate explicitly $\Phi_1(t)$ and $\Phi_2(t)$. By dividing the first ODE in (89) by $\Phi_1(t)$ and the second ODE by $\Phi_2(t)$ we get

$$\begin{aligned}
 \dot{\Phi}_1(t) = -2\rho(t)\Phi_1(t), \Phi_1(T) = & 2, \\
 \dot{\Phi}_2(t) = -2\rho(t)\Phi_2(t), \Phi_2(T) = & -2.
 \end{aligned}$$

We now try to solve the above ODEs (See the book by Boyce and DiPrima [28], Chapter §2). By simple computations shows that for any $t \in [0, T]$

$$\begin{cases} \Phi_1(t) = 2 \exp \left[\int_t^T \rho(s) ds \right] \\ \Phi_2(t) = -2 \exp \left[\int_t^T \rho(s) ds \right]. \end{cases} \tag{92}$$

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

$$\begin{aligned}
 v^*(t) = & [(\rho(t) - \zeta(t)) (A(t))^{-1}] x^*(t) \\
 & - [(\rho(t) - \zeta(t)) (A(t))^{-1}] \mathbb{E}(x^*(t)), \tag{93}
 \end{aligned}$$

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 (9). Moreover, with this choice of $v^*(t)$, the maximum condition (14) of Theorem 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 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 5.1 *The optimal solution of our mean-field stochastic control problem \mathcal{P}_0 is given in the state feedback form by*

$$v^*(t, x^*(t), \mathbb{E}(x^*(t))) = \left[(\rho(t) - \zeta(t)) (A(t))^{-1} \right] x^*(t) - \left[(\rho(t) - \zeta(t)) (A(t))^{-1} \right] \mathbb{E}(x^*(t)), \tag{94}$$

where $A(t)$ is given by (86).

Problem \mathcal{P}_ε : The Hamiltonian function \mathcal{H} for the problem \mathcal{P} is

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

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

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

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

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

then we get

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

$$\begin{aligned} &+ \sigma_t^2 v^*(t)v(t)Q^*(t) - \frac{1}{2}\sigma_t^2 v^2(t)Q^*(t) \\ &+ v(t)v^*(t) \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta) \\ &- \frac{1}{2}v^2(t) \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta). \end{aligned} \tag{95}$$

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

$$\begin{aligned} &\Psi^*(t)(\zeta(t) - \rho(t)) + K^*(t)\sigma_t \\ &+ \int_{\Theta} \gamma_t^*(\theta) \xi_t(\theta) \mu(d\theta) = 0. \\ &\mathbb{P} - a.s, a.e.t. \end{aligned} \tag{96}$$

The Hamiltonian \mathcal{H}_ε for the problem \mathcal{P}_ε is

$$\begin{aligned} \mathcal{H}_\varepsilon^{(x^*(\cdot), v^*(\cdot))}(t, x, v) &= -\Psi(t)\rho(t)x(t) - v(t) \left\{ \Psi^*(t)(\zeta(t) - \rho(t)) \right. \\ &\quad \left. + K^*(t)\sigma_t + \int_{\Theta} \gamma_t^*(\theta) \xi_t(\theta) \mu(d\theta) \right\} \\ &\quad + \sigma_t^2 v^*(t)v(t)Q^*(t) - \frac{1}{2}\sigma_t^2 v^2(t)Q^*(t) \\ &\quad + v(t)v^*(t) \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta) \\ &\quad - \frac{1}{2}v^2(t) \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta) \\ &\quad - \frac{\varepsilon^2}{4}L(v(t)). \end{aligned} \tag{97}$$

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

$$\begin{aligned} &\Psi^*(t)(\zeta(t) - \rho(t)) + K^*(t)\sigma_t \\ &+ \int_{\Theta} \gamma_t^*(\theta) \xi_t(\theta) \mu(d\theta) + \sigma_t^2 v^*(t)Q^*(t) \\ &- \sigma_t^2 v^\varepsilon(t)Q^*(t) \\ &+ v^*(t) \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta) \\ &- v^\varepsilon(t) \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta) \\ &- \frac{\varepsilon^2}{4}\dot{L}(v^\varepsilon(t)) = 0, \\ &\mathbb{P} - a.s, a.e.t. \end{aligned}$$

by applying (96) we have

$$\left[\sigma_t^2 Q^*(t) + \int_{\Theta} (\xi_t(\theta))^2 (Q^*(t) + \gamma_t^*(\theta)) \mu(d\theta) \right] \times (v^*(t) - v^\varepsilon(t)) - \frac{\varepsilon^2}{4} \dot{L}(v^\varepsilon(t)) = 0. \tag{98}$$

Combining (97)–(96) then we can shows that

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

since

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

then by simple computation we get

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

using (98), (47), and the fact that $L(\cdot)$ is convex and bounded we obtain

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

Moreover, by using (96) the hamiltonian H_ε of problem \mathcal{P}_ε is

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

Since $L(\cdot)$ is convex then the Hamiltonian $H_\varepsilon(t, \cdot, \cdot, \cdot, \cdot, \Psi(t), K(t), \gamma_t(\theta))$ is concave. By applying Theorem 4.1, this proves that, the control $v^*(t)$ given by (94) is indeed a near-optimal for stochastic control problem \mathcal{P}_ε .

Concluding remarks In this paper, necessary and sufficient conditions of near-optimal stochastic control for systems governed by mean-field jump diffusion processes are 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 3.1 reduces to stochastic maximum principle of optimality developed in Hafayed and Abbas ([17], 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 3.1 reduces to necessary conditions of optimality developed in Tang and Li ([9], Theorem 2.1) and Theorem 4.1 reduces to sufficient conditions of optimality developed in Framstad et al. ([12] Theorem 2.1).

Acknowledgments The authors would like to thank the editor and anonymous referees for their constructive corrections and valuable suggestions that improved the manuscript. The first author was partially supported by Algerian PNR project grant 08-u07-857, ATRST-ANDRU 2011-2013.

6 Appendix

The following result gives the definition and some basic properties of the Clarke’s generalized gradient.

Definition 6.1 Let F be a convex set in \mathbb{R}^n and let $f : F \rightarrow \mathbb{R}$ be a locally Lipschitz function. The generalized gradient of f at $\hat{x} \in F$, denoted by $\partial_x^\circ f(\hat{x})$, is a set defined by

$$\partial_x^\circ f(\hat{x}) = \{ \xi \in \mathbb{R}^n : \langle \xi, v \rangle \leq f^\circ(\hat{x}, v), \text{ for any } v \in \mathbb{R}^n \},$$

where $f^\circ(\hat{x}, v) = \limsup_{y \rightarrow \hat{x}, t \rightarrow 0} \frac{1}{t} (f(y + tv) - f(y))$.

Proposition 6.1 If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 [29] or the book by Yong and Zhou ([27] 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 $\partial_x^\circ f(a) = \overline{\text{co}}\{-1, 1\} = [-1, 1]$.

The following result gives special case of the Itô formula for jump diffusions.

Lemma 6.1 (Integration by parts formula for jumps processes) Suppose that the processes $x_1(t)$ and $x_2(t)$ are given by: for $j = 1, 2, t \in [s, T]$:

$$\begin{cases} dx_j(t) = f(t, x_j(t), u(t)) dt + \sigma(t, x_j(t), u(t)) dW(t) \\ + \int_{\Theta} g(t, x_j(t^-), u(t), \theta) N(d\theta, dt), x_j(s) = 0. \end{cases}$$

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. ([12], Lemma 2.1) for the detailed proof of the above Lemma.

Proposition 6.2 Let \mathcal{G} be the predictable σ -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

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_s^t \int_{\Theta} f(r, \theta) N(d\theta, dr) \right|^\beta \right] \\ & \leq C \mathbb{E} \left[\int_s^T \int_{\Theta} |f(r, \theta)|^\beta \mu(d\theta) dr \right]. \end{aligned}$$

See Bouchard and Elie ([30], Appendix).

References

1. Zhou XY (1998) Stochastic near-optimal controls: necessary and sufficient conditions for near-optimality. *SIAM J Control Optim* 36(3):929–947
2. Hafayed M, Abbas S, Veverka P (2013) On necessary and sufficient conditions for near-optimal singular stochastic controls. *Optim Lett* 7(5):949–966
3. Hafayed M, Veverka P, Abbas S (2012) On maximum principle of near-optimality for diffusions with jumps, with application to consumption-investment problem. *Differ Equ Dyn Syst* 20(2):111–125
4. Hafayed M, Abbas S (2013) On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality. *J Optim Theory Appl*. doi:10.1007/s10957-013-0361-1
5. Hafayed M, Abbas S (2013) Stochastic near-optimal singular controls for jump diffusions: necessary and sufficient conditions. *J Dyn Control Syst*. doi:10.1007/s10883-013-9191-6
6. Huang J, Li X, Wang G (2010) Near-optimal control problems for linear forward-backward stochastic systems. *Automatica* 46(2):397–404
7. Hui E, Huang J, Li X, Wang G (2011) Near-optimal control for stochastic recursive problems. *Syst Control Lett* 60:161–168
8. Chighoub F, Mezerdi B (2011) Optimality conditions in stochastic control of jump diffusion processes. *Syst Control Lett* 60:907–916
9. Tang SJ, Li XJ (1994) Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J Control Optim* 32(5):1447–1475
10. Hafayed M (2013) A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes. *Int J Dyn Control*. doi:10.1007/s40435-013-0027-8
11. Cadenillas A (2002) A stochastic maximum principle for system with jumps, with applications to finance. *Syst Control Lett* 47:433–444
12. Framstad NC, Øksendal B, Sulem A (2004) Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *J Optim Theory Appl* 121:77–98
13. Øksendal B, Sulem A (2007) Applied stochastic control of jump diffusions, 2nd edn. Springer, Berlin
14. Rishel R (1975) A minimum principle for controlled jump processes. *Lecture notes in economics and mathematical systems*, vol 107. Springer, Berlin, pp 493–508

15. Shi J, Wu Z (2010) Maximum principle for Forward-backward stochastic control system with random jumps and application to finance. *J Syst Sci Complex* 23:219–231
16. Shi J, Wu Z (2011) A stochastic maximum principle for optimal control of jump diffusions and application to finance. *Chin J Appl Prob Stat* 27(2)
17. Hafayed M, Abbas S (2013) A general maximum principle for stochastic differential equations of mean-field type with jump processes. Technical report. arXiv:1301.7327v4
18. Shi J, Wu Z (2006) The Maximum principle for fully coupled Forward-backward stochastic control system. *Acta Autom Sinica* 32(2):161–169
19. Buckdahn R, Djehiche B, Li J, Peng S (2009) Mean-field backward stochastic differential equations: a limit approach. *Ann Prob* 37(4):1524–1565
20. Buckdahn R, Djehiche B, Li J (2011) A general stochastic maximum principle for SDEs of mean-field type. *Appl Math Optim* 64:197–216
21. Shi J (2012) Sufficient conditions of optimality for mean-field stochastic control problems. In: 12th international conference on control, automation, robotics and vision Guangzhou, China, 5–7th December
22. Li J (2012) Stochastic maximum principle in the Mean-field controls. *Automatica* 48:366–373
23. Andersson D, Djehiche B (2011) A maximum principle for SDEs of mean-field type. *Appl Math Optim* 63:341–356
24. Shen Y, Siu TK (2013) The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem. *Nonlinear Anal* 86:58–73
25. Meyer-Brandis T, Øksendal B, Zhou, XY (2012) A mean-field stochastic maximum principle via Malliavin calculus. *Stoch Int J Prob Stoch Proc* 84(5–6):643–666
26. Ekeland I (1974) On the variational principle. *J Math Anal Appl* 47:324–353
27. Yong J, Zhou XY (1999) Stochastic controls, Hamiltonian systems and HJB equations. Springer, New York
28. Boyce WE, DiPrima RC (2000) Elementary differential equations and boundary value problems, 7th edn. Wiley, New York
29. Clarke FH (1983) Optimization and nonsmooth analysis. Wiley, New York
30. Bouchard B, Elie R (2008) Discrete time approximation of decoupled Forward-Backward SDE with jumps. *Stoch Process Appl* 118(1):53–75