A mean-field maximum principle for optimal control of forward–backward stochastic differential equations with Poisson jump processes

Mokhtar Hafayed

Received: 12 June 2013 / Revised: 2 August 2013 / Accepted: 6 August 2013 / Published online: 10 September 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract We consider mean-field type stochastic optimal control problems for systems governed by special mean-field forward–backward 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. Necessary conditions of optimal control for these systems in the form of a maximum principle are established by means of spike variation techniques. Our result differs from the classical one in the sense that here the adjoint equation has a mean-field type. The control domain is not assumed to be convex.

Keywords Optimal stochastic control · Mean-field forward–backward stochastic differential equation with jump processes · Stochastic maximum principle · Spike variation techniques · Poisson martingale measure

Mathematics Subject Classification 93E20 · 60H10

1 Introduction

The stochastic maximum principle of optimality for systems governed by forward–backward stochastic differential equations (FBSDEs) has been studied by many authors, see e.g. [1–6]. The maximum principle for fully coupled forward–backward stochastic control system in global form has been investigated in Shi and Wu [1]. The necessary conditions for partially-observed optimal control of fully-coupled forward–backward stochastic system has been studied by Shi and Wu [2]. However, the necessary conditions of opti-

mality for FBSDEs in global form, with uncontrolled diffusions coefficient was derived by Xu [3]. In a recent paper by Yong [5], the author completely solved the problem of maximum principle of optimality for fully coupled FBSDEs. He considered an optimal control problem for general coupled FBSDEs with mixed initial-terminal conditions and derived the necessary conditions for optimality when the control variable appears in the diffusion coefficients of the forward-equation and the control domain is not necessarily convex. A good account and an extensive list of references on stochastic maximum principle for FBSDEs with applications can be found in Ma and Yong [6].

As is well known, the stochastic optimal control problems for jump processes have been investigated by many authors, see e.g. [7–14]. The necessary and sufficient conditions of optimality for forward–backward stochastic differential equations with jump processes (FBSDEJs) was obtained by Shi and Wu [7,8]. In an interesting paper, Shi [9] generalized Yong’s maximum principle for FBSDEs obtained in Yong [5] to jump case. He established the stochastic maximum principle for optimality for fully coupled FBSDEJs when the control variable appears both in the diffusion and jump coefficients and the control domain is not assumed to be convex.

The mean-field model was initially introduced by Kac [15] as a stochastic model for the Vlasov-kinetic equation of plasma and the study of which was initiated by McKean [16]. Since then, many authors made contributions on mean-field stochastic problems with some applications, see for instance, [17–28]. Recently, mean-field stochastic maximum principle of optimality was considered by many authors, see for instance [17,20,22–27]. In Lazry and Lions [23] the authors introduced a general mathematical modeling approach for high-dimensional systems of evolution equations corresponding to a large number of particles (or agents). They extended the field of such mean-field approaches also to

M. Hafayed (✉)
Laboratory of Applied Mathematics,
Biskra University, P.O. Box 145, 07000 Biskra, Algeria
e-mail: hafayedmokhtar@yahoo.com

problems in economics, finance and game theory. In Buckdahn et al. [19] a general notion of mean-field backward stochastic differential equations (BSDEs) associated with a mean-field stochastic differential equation (SDE) is obtained in a natural way as a limit of some high dimensional system of FBSDEs governed by a d -dimensional Brownian motion, and influenced by positions of a large number of other particles. In Buckdahn et al. [20] a general maximum principle was introduced for a class of stochastic control problems involving SDEs of mean-field type. However, the sufficient conditions of optimality for mean-field SDE have been established by Shi [17]. In Mayer-Brandis et al. [26] a stochastic maximum principle of optimality for systems governed by controlled Itô-Lévy process of mean-field type was proved by using Malliavin calculus. Under the conditions that the control domains are convex, Andersson and Djehiche [24] and Li [25] studied problems for two types of more general controlled SDEs and cost functionals, respectively. The linear-quadratic optimal control problem for mean-field SDEs has been studied by Yong [27] and in Shi [17]. The mean-field stochastic maximum principle for jump diffusion processes has been investigated in Hafayed and Abbas [22]. The macroscopic limit of stochastic partial differential equations of McKean-Vlasov type has been studied by Kotelenetz and Kurtz [21]. The existence result for the solution of mean-field FBSDEs has been studied by Carmona and Dularue [28].

Our goal in this paper is to establish a set of necessary conditions of optimality for mean-field FBSDEJs, where the coefficient depend not only on the state process but also its marginal law of the state process through its expected value. The cost functional is also of mean-field type. The mean-field problem under consideration is not simple extension from the mathematical point of view, but also provide interesting models in many applications such as mathematical finance, optimal control for mean-field systems. The proof of our main result is based on spike variation method. These necessary conditions are described in terms of two adjoint processes, corresponding to the mean-field forward and backward components with jumps and a maximum conditions on the Hamiltonian. To streamline the presentation of this paper, we only study the one dimensional case.

The plan of the rest of this paper is organized as follows. In Sect. 2 we formulate the mean-field stochastic control problem and describe the assumptions of the model. The last section is devoted to prove our stochastic maximum principle for mean-field FBSDEJs which is the main result of this paper.

2 Problem formulation and preliminaries

We consider in this paper stochastic optimal control problems of mean-field type of the following kind. Let $T > 0$

be a fixed time horizon and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ 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 $\mathbb{R}$ 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 $N(\cdot, \cdot)$ is Poisson martingale measure on $\mathcal{B}(\Theta) \times \mathcal{B}(\mathbb{R}_+)$ with local characteristics $\mu(d\theta) dt$. We assume that $(\mathcal{F}_t)_{t \in [0, T]}$ is $\mathbb{P}$ -augmentation of the natural filtration $(\mathcal{F}_t^{(W, N)})_{t \in [0, T]}$ defined as follows

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

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

In the present paper, we study stochastic optimal control for a system governed by mean-field forward-backward stochastic differential equations with jump processes (FBSDEJs) of the form

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

where f, σ, b, g, h are given maps and the initial condition ζ is an $\mathcal{F}_0$ -measurable random variable. The special mean-field FBSDEJs-(1), called McKean-Vlasov systems are obtained as a limit approach, by the mean-square limit, when $n \rightarrow +\infty$ of a system of interacting particles of the form:

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

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

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

$$J(u(\cdot)) = \mathbb{E}[\varphi(y(0), \mathbb{E}(y(0)))]. \quad (2)$$

This cost functional is also of mean-field type, as the function φ depends on the marginal law of the state process through its expected value, which makes the control problem time inconsistent in the sense that the dynamic programming principle for optimality does not hold.

An admissible control $u(\cdot)$ is an $\mathcal{F}_t$ -adapted and square-integrable process with values in a nonempty subset $\mathbb{A}$ of $\mathbb{R}$. We denote the set of all admissible controls by $\mathcal{U}([0, T])$. Any admissible control $u^*(\cdot) \in \mathcal{U}([0, T])$ satisfying

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

is called an optimal control.

The corresponding state process, solution of FBSDEJ-(1), is denoted by $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot, \cdot))$.

2.1 Notations

For convenience, we will use the following notations in this paper

1. In the sequel, $\mathbb{L}_{\mathcal{F}}^2([0, T]; \mathbb{R})$ denotes the Hilbert space of $\mathcal{F}_t$ -adapted processes $(x(t))_{t \in [0, T]}$ such that $\mathbb{E} \int_0^T |x(t)|^2 dt < +\infty$.
2. $\mathbb{M}_{\mathcal{F}}^2([0, T]; \mathbb{R})$ denotes the Hilbert space of $\mathcal{F}_t$ -predictable processes $(\psi(t, \theta))_{t \in [0, T]}$ defined on $[0, T] \times \Omega \times \Theta$ such that $\mathbb{E} \int_0^T \int_{\Theta} |\psi(t, \theta)|^2 \mu(\theta) dt < +\infty$.
3. We denote by $\mathbf{1}_{\mathcal{A}}$ the indicator function of $\mathcal{A}$ and $(\tilde{x}, \tilde{y}, \tilde{z})$ denotes the expected value of (x, y, z) .
4. For a differentiable function Φ we denote by $\Phi_{\rho}(t)$ its gradient with respect to the variable ρ .
5. For convenience, we will use the following

$$\begin{aligned} \delta f(t) &= f(t, x^*(t), \mathbb{E}(x^*(t)), u(t)) \\ &\quad - f(t, x^*(t), \mathbb{E}(x^*(t)), u^*(t)), \end{aligned}$$

and

$$\begin{aligned} \delta g(t, \theta) &= g(t, x^*(t), \mathbb{E}(x^*(t)), y^*(t), \mathbb{E}(y^*(t)), \\ &\quad z^*(t), \mathbb{E}(z^*(t)), r^*(t, \theta), u(t)) \end{aligned}$$

$$\begin{aligned} &-g(t, x^*(t), \mathbb{E}(x^*(t)), y^*(t), \mathbb{E}(y^*(t)), \\ &z^*(t), \mathbb{E}(z^*(t)), r^*(t, \theta), u^*(t)). \end{aligned}$$

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

2.2 Conditions

Throughout this paper, we also assume that the coefficients functions

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

satisfy the following standing assumptions

2.2.1 Condition (H1)

The functions $f, \sigma, b, g, h, \varphi$ are continuously differentiable with respect to $(x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r)$.

2.2.2 Condition (H2)

The derivatives of f, σ, b, h, φ with respect to $(x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r)$ are bounded and there is a constant $C > 0$ such that $\sup_{\theta \in \Theta} |g_x(t, \theta)| < C$ for $x = x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r$.

Under the assumptions (H1) and (H2), the FBSDEJ-(1) has an unique solution $(x(t), y(t), z(t), r(t, \cdot)) \in \mathbb{L}_{\mathcal{F}}^2([0, T]; \mathbb{R}) \times \mathbb{L}_{\mathcal{F}}^2([0, T]; \mathbb{R}) \times \mathbb{L}_{\mathcal{F}}^2([0, T]; \mathbb{R}) \times \mathbb{L}_{\mathcal{F}}^2([0, T]; \mathbb{R})$.

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

$$\begin{cases} d\Psi(t) = -\{f_x(t) \Psi(t) + \mathbb{E}[f_{\tilde{x}}(t) \Psi(t)] \\ \quad + \sigma_x(t) Q(t) + \mathbb{E}[\sigma_{\tilde{x}}(t) Q(t)] \\ \quad + \int_{\Theta} [g_x(t, \theta) K(t) + \mathbb{E}(g_{\tilde{x}}(t, \theta) K(t))] \\ \quad + c_x(t, \theta) R_t(\theta)] \mu(d\theta)\} dt \\ \quad + Q(t) dW(t) + \int_{\Theta} R_t(\theta) N(d\theta, dt), \\ \Psi(T) = -\{h_x(T) K(T) + \mathbb{E}[h_{\tilde{x}}(T) K(T)]\}, \\ dK(t) = \int_{\Theta} [g_y(t, \theta) K(t) + \mathbb{E}(g_{\tilde{y}}(t, \theta) K(t))] \mu(d\theta) dt \\ \quad + \int_{\Theta} [g_z(t, \theta) K(t) + \mathbb{E}(g_{\tilde{z}}(t, \theta) K(t))] \mu(d\theta) dW(t) \\ \quad - \int_{\Theta} g_r(t, \theta) K(t) N(d\theta, dt) \\ K(0) = -\{\varphi_y(y(0), \mathbb{E}(y(0))) \\ \quad + \mathbb{E}[\varphi_{\tilde{y}}(y(0), \mathbb{E}(y(0)))]\}, \end{cases} \quad (4)$$

where

$$\begin{aligned} f_{\rho}(t) &= \frac{\partial f}{\partial \rho}(t, x(t), \mathbb{E}(x(t)), u(t)), \quad \text{for } \rho = x, \tilde{x}, \\ \sigma_{\rho}(t) &= \frac{\partial \sigma}{\partial \rho}(t, x(t), \mathbb{E}(x(t))), \quad \text{for } \rho = x, \tilde{x}, \\ c_x(t, \theta) &= c_x(t, x, \theta), \\ h_{\rho}(t) &= \frac{\partial h}{\partial \rho}(x(t), \mathbb{E}(x(t))), \quad \text{for } \rho = x, \tilde{x}, \\ \varphi_{\rho}(t) &= \frac{\partial \varphi}{\partial \rho}(y(t), \mathbb{E}(y(t))), \quad \text{for } \rho = y, \tilde{y}, \\ g_{\rho}(t, \theta) &= \frac{\partial g}{\partial \rho}(t, x(t), \mathbb{E}(x(t)), y(t), \mathbb{E}(y(t)), z(t), \mathbb{E}(z(t)), \\ &\quad r(t, \theta), u(t)), \quad \text{for } \rho = x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r. \end{aligned}$$

Note that the first adjoint equation (backward) corresponding to the forward component $x(\cdot)$ turns out to be a linear mean-field backward SDE with jumps, and the second adjoint equation (forward) corresponding to the backward component $(y(\cdot), z(\cdot), r(\cdot, \cdot))$ turns out to be a linear mean-field (forward) SDE with jump processes.

Further, we define the Hamiltonian function

$$H : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R},$$

associated with the stochastic control problem (1)–(2) as follows

$$\begin{aligned} H(t, x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r(\cdot), u, \Psi, Q, K, R(\cdot)) \\ = \Psi(t)f(t, x, \tilde{x}, u) + Q(t)\sigma(t, x, \tilde{x}) \\ + \int_{\Theta} [K(t)g(t, x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r(t, \theta), u) \\ + R(t, \theta)c(t, x, \theta)] \mu(d\theta). \end{aligned} \quad (5)$$

If we denote by

$$H(t) = H(t, x(t), \tilde{x}(t), y(t), \tilde{y}(t), z(t), \tilde{z}(t), r(t, \cdot), u(t), \Psi(t), Q(t), K(t), R_t(\cdot)),$$

then the adjoint equation (4) can be rewritten as the following stochastic Hamiltonian system's type

$$\begin{cases} d\Psi(t) = -\{H_x(t) + \mathbb{E}[H_{\tilde{x}}(t)]\} dt \\ \quad + Q(t)dW(t) + \int_{\Theta} R_t(\theta) N(d\theta, dt), \\ \Psi(T) = -[h_x(T) + \mathbb{E}(h_{\tilde{x}}(T))] K(T) \\ -dK(t) = [H_y(t) + \mathbb{E}(H_{\tilde{y}}(t))] dt \\ \quad + [H_z(t) + \mathbb{E}(H_{\tilde{z}}(t))] dW(t) \\ \quad + \int_{\Theta} H_r(t, \theta) N(d\theta, dt) \\ K(0) = -\{\varphi_y(y(0), \mathbb{E}(y(0))) \\ \quad + \mathbb{E}[\varphi_{\tilde{y}}(y(0), \mathbb{E}(y(0)))]\}. \end{cases} \quad (6)$$

It is a well known fact that under assumptions (H1) and (H2), the adjoint equation (4) admits a unique solution

$(\Psi(t), Q(t), K(t), R(t, \cdot))$ such that

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

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

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} |\Psi(t)|^2 + \sup_{t \in [0, T]} |K(t)|^2 dt \right. \\ \left. + \int_0^T |Q(t)|^2 dt + \int_0^T \int_{\Theta} |R_t(\theta)|^2 \mu(d\theta) dt \right] < C. \end{aligned} \quad (7)$$

3 Mean-field type maximum principle for FBSDEJs

In this section, we establish a set of necessary conditions of Pontryagin's type for a stochastic control to be optimal where the system evolves according to nonlinear controlled mean-field FBSDEJs. Spike variation techniques are applied to prove our mean-field stochastic maximum principle.

The following theorem constitutes the main contribution of this paper.

Let $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot, \cdot))$ be the trajectory of the mean-field FBSDEJ-(1) corresponding to the optimal control $u^*(\cdot)$, and $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot), R^*(\cdot))$ be the solution of adjoint equation (4) corresponding to $u^*(\cdot)$.

Theorem 3.1 *Let Assumptions (H1) and (H2) hold. If $(u^*(\cdot), x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot, \cdot))$ is an optimal solution of the mean-field control problem (1)–(2). Then the maximum principle holds, that is*

$$\begin{aligned} H(t, \lambda^*(t, \theta), u^*(t), \Lambda^*(t, \theta)) \\ = \min_{u \in \mathbb{A}} H(t, \lambda^*(t, \theta), u, \Lambda^*(t, \theta)) \\ \mathbb{P} - a.s., \quad a.e. t \in [0, T], \end{aligned} \quad (8)$$

where

$$\begin{aligned} \lambda^*(t, \theta) &= (x^*(t), \mathbb{E}(x^*(t)), y^*(t), \mathbb{E}(y^*(t)), z^*(t), \\ &\quad \mathbb{E}(z^*(t)), r^*(t, \theta)) \\ \Lambda^*(t, \theta) &= (\Psi^*(t), Q^*(t), K^*(t), R^*(t, \theta)). \end{aligned}$$

To prove Theorem 3.1 we need some preliminary results given in the following Lemmas. We derive the variational inequality (8) in several steps, from the fact that

$$J(u^{\varepsilon}(\cdot)) \geq J(u^*(\cdot)), \quad (9)$$

where $u^{\varepsilon}(\cdot)$ is the so called spike variation of optimal control $u^*(\cdot)$ defined as follows.

For $\varepsilon > 0$, we choose a Borel measurable set $\mathcal{A}_\varepsilon \subset [0, T]$ such that $\nu(\mathcal{A}_\varepsilon) = \varepsilon$, where $\nu(\mathcal{A}_\varepsilon)$ denote the Lebesgue measure of the subset $\mathcal{A}_\varepsilon$, and we consider the control process which is the *spike variation* of $u^*(\cdot)$

$$u^\varepsilon(t) = \begin{cases} u : t \in \mathcal{A}_\varepsilon, \\ u^*(t) : t \in \mathcal{A}_\varepsilon^c, \end{cases}$$

where $\varepsilon > 0$ is sufficiently small and u is an arbitrary element $\mathcal{F}_t$ -measurable random variable with values in $\mathbb{A}$, such that $\sup_{w \in \Omega} |u(w)| < \infty$, which we consider as fixed from now on.

Let $\lambda^\varepsilon(\cdot) = (x^\varepsilon(\cdot), y^\varepsilon(\cdot), z^\varepsilon(\cdot), r^\varepsilon(\cdot, \cdot))$ be the solution of FBSDEJs-(1) and $\Lambda^\varepsilon(\cdot) = (\Psi^\varepsilon(\cdot), Q^\varepsilon(\cdot), K^\varepsilon(\cdot), R^\varepsilon(\cdot))$ be the solution of the adjoint equation (4) corresponding to $u^\varepsilon(\cdot)$.

3.1 Variational equations

Now, we introduce the following variational equations. For simplicity of notation, we will still use $\delta\Phi(t)$ where $u(t) = u^\varepsilon(t)$, for $\Phi = f, g$, etc. and $f_x(t) = \frac{\partial f}{\partial x}(t, x^*, \mathbb{E}(x^*), u^*)$, etc. Let $(x_1^\varepsilon(\cdot), y_1^\varepsilon(\cdot), z_1^\varepsilon(\cdot), r_1^\varepsilon(\cdot, \cdot))$ be the solution of the following mean-field forward backward stochastic system

$$\begin{cases} dx_1^\varepsilon(t) = \{f_x(t)x_1^\varepsilon(t) + f_{\bar{x}}(t)\mathbb{E}(x_1^\varepsilon(t)) \\ \quad + \delta f(t)\mathbf{I}_{\mathcal{A}_\varepsilon}(t)\} dt \\ \quad + \{\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\bar{x}}(t)\mathbb{E}(x_1^\varepsilon(t))\} dW(t) \\ \quad + \int_{\Theta} c_x(t, \theta)x_1^\varepsilon(t)N(d\theta, dt), \\ x_1^\varepsilon(0) = 0, \\ dy_1^\varepsilon(t) = \int_{\Theta} \{g_x(t, \theta)x_1^\varepsilon(t) + g_{\bar{x}}(t, \theta)\mathbb{E}(x_1^\varepsilon(t)) \\ \quad + g_y(t, \theta)y_1^\varepsilon(t) + g_{\bar{y}}(t, \theta)\mathbb{E}(y_1^\varepsilon(t)) \\ \quad + g_z(t, \theta)z_1^\varepsilon(t) + g_{\bar{z}}(t, \theta)\mathbb{E}(z_1^\varepsilon(t)) \\ \quad + g_r(t, \theta)r_1^\varepsilon(t, \theta) + \delta g(t, \theta)\mathbf{I}_{\mathcal{A}_\varepsilon}(t)\} \mu(d\theta)dt \\ \quad + z_1^\varepsilon(t)dW(t) - \int_{\Theta} r_1^\varepsilon(t, \theta)N(d\theta, dt) \\ y_1^\varepsilon(T) = -[h_x(T) + \mathbb{E}(h_{\bar{x}}(T))]\mathbf{I}_{\mathcal{A}_\varepsilon}(T). \end{cases} \quad (10)$$

Noting that the variational equations (10) is also a FBSDEJs of mean-field type.

3.2 Duality

Our first Lemma below deals with the *duality relations* between $\Psi^*(t)$, $x_1^\varepsilon(t)$ and $K^*(t)$, $y_1^\varepsilon(t)$.

Lemma 3.1 *We have*

$$\begin{aligned} \mathbb{E}(\Psi^*(T)x_1^\varepsilon(T)) &= \mathbb{E} \int_0^T \Psi^*(t)\delta f(t)\mathbf{I}_{\mathcal{A}_\varepsilon}(t)dt \\ &\quad - \mathbb{E} \int_0^T \int_{\Theta} x_1^\varepsilon(t) [\mathbb{E}(g_{\bar{x}}(t, \theta)K^*(t)) \\ &\quad + g_x(t, \theta)K^*(t)] \mu(d\theta)dt, \end{aligned} \quad (11)$$

$$\begin{aligned} \mathbb{E}(K^*(T)y_1^\varepsilon(T)) &= -\mathbb{E}\{\varphi_y(y(0), \mathbb{E}(y(0))) \\ &\quad + \mathbb{E}(\varphi_{\bar{y}}(y(0), \mathbb{E}(y(0))))y_1^\varepsilon(0)\} \\ &\quad + \mathbb{E} \int_0^T \int_{\Theta} \{K^*(t)g_x(t, \theta)x_1^\varepsilon(t) \\ &\quad + K^*(t)g_{\bar{x}}(t, \theta)\mathbb{E}(x_1^\varepsilon(t)) \\ &\quad + K^*(t)\delta g(t, \theta)\mathbf{I}_{\mathcal{A}_\varepsilon}(t)\} \mu(d\theta)dt, \end{aligned} \quad (12)$$

and

$$\begin{aligned} \mathbb{E}\{[\varphi_y(y(0), \mathbb{E}(y(0))) + \mathbb{E}(\varphi_{\bar{y}}(y(0), \mathbb{E}(y(0))))]y_1^\varepsilon(0)\} \\ = \mathbb{E} \int_0^T \left[\Psi^*(t)\delta f(t) + \int_{\Theta} K^*(t)\delta g(t, \theta)\mu(d\theta)\mathbf{I}_{\mathcal{A}_\varepsilon}(t) \right] dt. \end{aligned} \quad (13)$$

Proof This Lemma is very important for the proof of our result. So that we give its detailed proof. By applying integration by parts formula for jump processes (see Lemma A1) to $\Psi^*(t)x_1^\varepsilon(t)$, we get

$$\begin{aligned} \mathbb{E}(\Psi^*(T)x_1^\varepsilon(T)) &= \mathbb{E} \int_0^T \Psi^*(t)dx_1^\varepsilon(t) + \mathbb{E} \int_0^T x_1^\varepsilon(t)d\Psi^*(t) \\ &\quad + \mathbb{E} \int_0^T Q^*(t) [\sigma_x(t)x_1^\varepsilon(t) + \sigma_{\bar{x}}(t)\mathbb{E}(x_1^\varepsilon(t))] dt \\ &\quad + \mathbb{E} \int_0^T \int_{\Theta} c_x(t, \theta)x_1^\varepsilon(t)R_t(\theta)\mu(d\theta)dt \\ &= I_1^\varepsilon + I_2^\varepsilon + I_3^\varepsilon + I_4^\varepsilon. \end{aligned} \quad (14)$$

A simple computation shows that

$$\begin{aligned} I_1^\varepsilon &= \mathbb{E} \int_0^T \Psi^*(t)dx_1^\varepsilon(t) \\ &= \mathbb{E} \int_0^T \{ \Psi^*(t)f_x(t)x_1^\varepsilon(t) + \Psi^*(t)f_{\bar{x}}(t)\mathbb{E}(x_1^\varepsilon(t)) \\ &\quad + \Psi^*(t)\delta f(t)\mathbf{I}_{\mathcal{A}_\varepsilon}(t) \} dt, \end{aligned} \quad (15)$$

and

$$\begin{aligned} I_2^\varepsilon &= \mathbb{E} \int_0^T x_1^\varepsilon(t)d\Psi^*(t) \\ &= -\mathbb{E} \int_0^T \{ x_1^\varepsilon(t)f_x(t)\Psi^*(t) + x_1^\varepsilon(t)\mathbb{E}(f_{\bar{x}}(t)\Psi^*(t)) \\ &\quad + x_1^\varepsilon(t)\sigma_x(t)Q^*(t) + x_1^\varepsilon(t)\mathbb{E}(\sigma_{\bar{x}}(t)Q^*(t)) \} dt \end{aligned}$$

$$+ \int_{\Theta} (x_1^\varepsilon(t) g_x(t, \theta) K^*(t) + x_1^\varepsilon(t) \mathbb{E}(g_{\tilde{x}}(t, \theta) K^*(t)) + c_x(t, \theta) R_t(\theta)) \mu(d\theta) dt. \quad (16)$$

By standard arguments we get

$$\begin{aligned} I_3^\varepsilon &= \mathbb{E} \int_0^T Q^*(t) \sigma_x(t) x_1^\varepsilon(t) dt \\ &\quad + \mathbb{E} \int_0^T Q^*(t) \sigma_{\tilde{x}}(t) \mathbb{E}(x_1^\varepsilon(t)) dt \\ I_4^\varepsilon &= \mathbb{E} \int_0^T \int_{\Theta} c_x(t, \theta) x_1^\varepsilon(t) R_t(\theta) \mu(d\theta) dt. \end{aligned} \quad (17)$$

The duality relation (11) follows immediately from combining (15)–(17) and (14).

Let us turn to second duality relation (12). By applying integration by parts formula for jump process (Lemma A1) to $K^*(t) y_1^\varepsilon(t)$ we get

$$\begin{aligned} \mathbb{E}(K^*(T) y_1^\varepsilon(T)) &= \mathbb{E}(K^*(0) y_1^\varepsilon(0)) \\ &\quad + \mathbb{E} \int_0^T K^*(t) dy_1^\varepsilon(t) + \mathbb{E} \int_0^T y_1^\varepsilon(t) dK^*(t) \\ &\quad - \mathbb{E} \int_0^T z_1^\varepsilon(t) [g_z(t) K^*(t) + \mathbb{E}(g_{\tilde{z}}(t) K^*(t))] dt \\ &\quad - \mathbb{E} \int_0^T \int_{\Theta} r_1^\varepsilon(t, \theta) g_r(t, \theta) K(t) \mu(d\theta) dt. \\ &= I_1^\varepsilon + I_2^\varepsilon + I_3^\varepsilon + I_4^\varepsilon + I_5^\varepsilon. \end{aligned} \quad (18)$$

From (10) we obtain

$$\begin{aligned} I_2^\varepsilon &= \mathbb{E} \int_0^T K^*(t) dy_1^\varepsilon(t) \\ &= \mathbb{E} \int_0^T \int_{\Theta} \{K^*(t) g_x(t, \theta) x_1^\varepsilon(t) \\ &\quad + K^*(t) g_{\tilde{x}}(t, \theta) \mathbb{E}(x_1^\varepsilon(t)) \\ &\quad + K^*(t) g_y(t, \theta) y_1^\varepsilon(t) + K^*(t) g_{\tilde{y}}(t, \theta) \mathbb{E}(y_1^\varepsilon(t)) \\ &\quad + K^*(t) g_z(t, \theta) z_1^\varepsilon(t) + K^*(t) g_{\tilde{z}}(t, \theta) \mathbb{E}(z_1^\varepsilon(t)) \\ &\quad + K^*(t) g_r(t, \theta) r_1^\varepsilon(t, \theta) \\ &\quad + K^*(t) \delta g(t, \theta) \mathbf{I}_{A_\varepsilon}(t)\} \mu(d\theta) dt, \end{aligned} \quad (19)$$

from (4), we obtain

$$\begin{aligned} I_3^\varepsilon &= \mathbb{E} \int_0^T y_1^\varepsilon(t) dK^*(t) \\ &= -\mathbb{E} \int_0^T \int_{\Theta} \{y_1^\varepsilon(t) g_y(t, \theta) K^*(t) \\ &\quad + y_1^\varepsilon(t) \mathbb{E}(g_{\tilde{y}}(t, \theta) K^*(t))\} \mu(d\theta) dt, \end{aligned} \quad (20)$$

and

$$\begin{aligned} I_4^\varepsilon &= -\mathbb{E} \int_0^T \int_{\Theta} [z_1^\varepsilon(t) g_z(t, \theta) K^*(t) \\ &\quad + z_1^\varepsilon(t) \mathbb{E}(g_{\tilde{z}}(t, \theta) K^*(t))] \mu(d\theta) dt \\ I_5^\varepsilon &= -\mathbb{E} \int_0^T \int_{\Theta} r_1^\varepsilon(t, \theta) g_r(t, \theta) K^*(t) \mu(d\theta) dt. \end{aligned} \quad (21)$$

Since

$$\begin{aligned} I_1^\varepsilon &= \mathbb{E}(K^*(0) y_1^\varepsilon(0)) \\ &= -\mathbb{E}\{\varphi_y(y(0), \mathbb{E}(y(0))) \\ &\quad + \mathbb{E}(\varphi_{\tilde{y}}(y(0), \mathbb{E}(y(0)))) y_1^\varepsilon(0)\}, \end{aligned}$$

the duality relation (12) follows immediately by combining (19)–(21) and (18).

Let us turn to (13). From (4) and (11) we deduce

$$\mathbb{E}(\Psi^*(T) x_1^\varepsilon(T)) + \mathbb{E}(K^*(T) y_1^\varepsilon(T)) = 0.$$

Combining (11) and (12), then (13) follows immediately. $\square$

To this end we give the following estimations.

3.3 Some prior estimates

Lemma 3.2 Under Assumptions (H1) and (H2), the following estimations holds: for any $k \geq 1$

$$\begin{aligned} \mathbb{E}(\sup_{0 \leq t \leq T} |x_1^\varepsilon(t)|^{2k}) &\leq C_k \varepsilon^k, \\ \mathbb{E}(\sup_{0 \leq t \leq T} |y_1^\varepsilon(t)|^2) \\ &\quad + \mathbb{E} \int_0^T \left[|z_1^\varepsilon(s)|^2 + \int_{\Theta} |r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) \right] ds \leq C \varepsilon^2, \end{aligned} \quad (22)$$

$$\begin{aligned} \sup_{0 \leq t \leq T} |\mathbb{E}(x_1^\varepsilon(t))|^2 &\leq C \varepsilon^2, \\ \sup_{0 \leq t \leq T} |\mathbb{E}(y_1^\varepsilon(t))|^2 &\quad + \int_t^T |\mathbb{E}(z_1^\varepsilon(s))|^2 ds \\ &\quad + \int_0^T \int_{\Theta} |\mathbb{E}(r_1^\varepsilon(s, \theta))|^2 \mu(d\theta) ds \leq C \tilde{\rho}(\varepsilon). \end{aligned} \quad (23)$$

$$\begin{aligned}
& \mathbb{E}(\sup_{0 \leq t \leq T} |x^\varepsilon(t) - x^*(t)|^{2k}) \leq C_k \varepsilon^k, \\
& \mathbb{E}(\sup_{0 \leq t \leq T} |y^\varepsilon(t) - y^*(t)|^2) \leq C \varepsilon^2, \\
& \mathbb{E} \int_t^T |z^\varepsilon(s) - z^*(s)|^2 ds \leq C \varepsilon^2, \\
& \mathbb{E} \int_0^T \int_\Theta |r^\varepsilon(s, \theta) - r^*(s, \theta)|^2 \mu(d\theta) ds \leq C \varepsilon^2, \quad (24)
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E}(\sup_{0 \leq t \leq T} |x^\varepsilon(t) - x^*(t) - x_1^\varepsilon(t)|^{2k}) \leq C_k \varepsilon^{2k}, \\
& \mathbb{E}(\sup_{0 \leq t \leq T} |y^\varepsilon(t) - y^*(t) - y_1^\varepsilon(t)|^{2k}) \leq C_k \varepsilon^{2k}, \\
& \mathbb{E} \int_t^T |z^\varepsilon(s) - z^*(s) - z_1^\varepsilon(s)|^2 ds \leq C \varepsilon^2, \\
& \mathbb{E} \int_t^T \int_\Theta |r^\varepsilon(s, \theta) - r^*(s, \theta) - r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \leq C \varepsilon^2, \quad (25)
\end{aligned}$$

where C_k is a positive constant depend to k and $\tilde{\rho}(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Noting that estimate (24) follows from standard arguments (see Shi [7], Lemma 2.1). Estimate (23) will be derived as a consequence of the following Lemma.

Lemma 3.3 For any progressively measurable processes $(\Phi(s))_{s \in [0, T]}$, $(\psi(s))_{s \in [t, T]}$ and $(\phi(s, \theta))_{(s, \theta) \in [t, T] \times \Theta}$ for any $p \geq 1$, there exists a positive constants C and C_p such that

$$\begin{aligned}
& \mathbb{E}(\sup_{0 \leq t \leq T} |\Phi(t)|^p) \leq C_p, \\
& \mathbb{E} \int_t^T |\psi(s)|^2 ds \leq C, \\
& \mathbb{E} \int_t^T \int_\Theta |\phi(s, \theta)|^2 \mu(d\theta) ds \leq C, \quad (26)
\end{aligned}$$

then we have for $\varepsilon > 0$

$$|\mathbb{E}(x_1^\varepsilon(T) \Phi(T))|^2 + \int_0^T |\mathbb{E}(x_1^\varepsilon(t) \Phi(t))|^2 dt \leq C \tilde{\rho}(\varepsilon), \quad (27)$$

and

$$\begin{aligned}
& |\mathbb{E}(y_1^\varepsilon(T) \Phi(T))|^2 + \int_0^T |\mathbb{E}(y_1^\varepsilon(s) \Phi(s))|^2 ds \leq C \tilde{\rho}(\varepsilon), \\
& |\mathbb{E}(z_1^\varepsilon(T) \psi(T))|^2 + \int_\Theta |\mathbb{E}(r_1^\varepsilon(T, \theta) \phi(T, \theta))|^2 \mu(d\theta)
\end{aligned}$$

$$\begin{aligned}
& + \int_t^T \left\{ |\mathbb{E}(z_1^\varepsilon(s) \psi(s))|^2 \right. \\
& \left. + \int_\Theta |\mathbb{E}(r_1^\varepsilon(s, \theta) \phi(s, \theta))|^2 \mu(d\theta) \right\} ds \\
& \leq C_{(T, \mu(\Theta))} \tilde{\rho}(\varepsilon), \quad (28)
\end{aligned}$$

where $\tilde{\rho}(\varepsilon) = (\varepsilon + \varepsilon^2) \rightarrow 0, \varepsilon \rightarrow 0$.

Proof of estimate (27) First we set for $t \in [0, T]$: $\eta(t) = \exp\{M(t)\}$, where

$$\begin{aligned}
M(t) = & - \int_0^t \left[f_x(s) - \frac{1}{2} |\sigma_x(s)|^2 - \frac{1}{2} \int_\Theta (c_x(s, \theta))^2 \mu(d\theta) \right] ds \\
& - \int_0^t \sigma_x(s) dW(s) - \int_0^t \int_\Theta c_x(s, \theta) N(d\theta, ds),
\end{aligned}$$

and we denote by $\rho(t) = \eta(t)^{-1} = \exp\{-M(t)\}$.

By using Itô formula for $\exp\{M(t)\}$, we get

$$\begin{aligned}
d(\exp\{M(t)\}) &= \exp\{M(t)\} dM(t) \\
&+ \frac{1}{2} \exp\{M(t)\} d\langle M(t); M(t) \rangle,
\end{aligned}$$

this shows that

$$\begin{aligned}
d\eta(t) &= d(\exp\{M(t)\}) \\
&= -\eta(t) \left\{ \left[f_x(t) - (\sigma_x(t))^2 - \int_\Theta (c_x(t, \theta))^2 \mu(d\theta) \right] dt \right. \\
&\quad \left. + \sigma_x(t) dW(t) + \int_\Theta c_x(t, \theta) N(d\theta, dt) \right\}. \quad (29)
\end{aligned}$$

By using Integration by parts formula for jumps processes $\eta(t) x_1^\varepsilon(t)$ we have

$$\begin{aligned}
d(\eta(t) x_1^\varepsilon(t)) &= \eta(t) dx_1^\varepsilon(t) + x_1^\varepsilon(t) d\eta(t) \\
&+ d\langle \eta(t), x_1^\varepsilon(t) \rangle, \\
&= I_1^\varepsilon + I_2^\varepsilon + I_3^\varepsilon.
\end{aligned}$$

From (10) we get

$$\begin{aligned}
I_1^\varepsilon &= \eta(t) dx_1^\varepsilon(t) \\
&= \eta(t) \left\{ \left[f_x(t) x_1^\varepsilon(t) + f_{\bar{x}}(t) \mathbb{E}(x_1^\varepsilon(t)) \right. \right. \\
&\quad \left. + \delta f(t) \mathbf{I}_{\mathcal{A}_\varepsilon}(t) \right] dt \\
&\quad \left. + [\sigma_x(t) x_1^\varepsilon(t) + \sigma_{\bar{x}}(t) \mathbb{E}(x_1^\varepsilon(t))] dW(t) \right. \\
&\quad \left. + \int_\Theta c_x(t, \theta) x_1^\varepsilon(t) N(d\theta, dt) \right\}.
\end{aligned}$$

By using (29) we get

$$\begin{aligned} I_2^\varepsilon &= x_1^\varepsilon(t) d\eta(t) \\ &= -\eta(t) f_x(t) x_1^\varepsilon(t) dt - \eta(t) \sigma_x(t) x_1^\varepsilon(t) dW(t) \\ &\quad - \eta(t) x_1^\varepsilon(t) \int_{\Theta} c_x(t-, \theta) N(d\theta, dt) \\ &\quad + \eta(t) x_1^\varepsilon(t) (\sigma_x(t))^2 dt \\ &\quad + \eta(t) \int_{\Theta} (c_x(t-, \theta))^2 x_1^\varepsilon(t) \mu(d\theta) dt, \end{aligned}$$

and by a simple computation we get

$$\begin{aligned} I_3^\varepsilon &= d\langle \eta(t), x_1^\varepsilon(t) \rangle \\ &= -\eta(t) \sigma_x(t) [\sigma_x(t) x_1^\varepsilon(t) + \sigma_x^-(t) \mathbb{E}(x_1^\varepsilon(t))] dt \\ &\quad - \int_{\Theta} \eta(t) (c_x(t, \theta))^2 x_1^\varepsilon(t) \mu(d\theta) dt. \end{aligned}$$

Consequently, from the above equations we deduce that

$$\begin{aligned} d(\eta(t) x_1^\varepsilon(t)) &= I_1^\varepsilon + I_2^\varepsilon + I_3^\varepsilon \\ &= \eta(t) \{ (f_x(t) \mathbb{E}(x_1^\varepsilon(t)) + \delta f(t) \mathbf{I}_{\mathcal{A}_\varepsilon}(t) \\ &\quad - \sigma_x(t) \sigma_x^-(t) \mathbb{E}(x_1^\varepsilon(t))) dt \\ &\quad + \sigma_x^-(t) \mathbb{E}(x_1^\varepsilon(t)) dW(t) \}, \end{aligned}$$

by integrating the above equation and the fact $\rho(t) = \eta(t)^{-1}$ we obtain

$$\begin{aligned} x_1^\varepsilon(t) &= \rho(t) \int_0^t \eta(s) \times \{ [f_x(s) \mathbb{E}(x_1^\varepsilon(s)) + \delta f(s) \mathbf{I}_{\mathcal{A}_\varepsilon}(s)] \\ &\quad - \sigma_x(s) \sigma_x^-(s) \mathbb{E}(x_1^\varepsilon(s)) \} ds \\ &\quad + \rho(t) \int_0^t \eta(s) \sigma_x^-(s) \mathbb{E}(x_1^\varepsilon(s)) dW(s). \end{aligned} \quad (30)$$

From (30) and the fact that $f_x, \sigma_x, c_x(\cdot, \theta)$ are bounded, then by using Proposition A1, we get

for all $p > 1$ there exists a positive constant $C = C_{(T, p, \mu(\Theta))}$ such that

$$\begin{aligned} &\mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \int_{\Theta} c_x(s-, \theta) N(d\theta, ds) \right|^p \right] \\ &\leq C_{(T, p, \mu(\Theta))} \mathbb{E} \left[\int_0^T \int_{\Theta} |c_x(s, \theta)|^p \mu(d\theta) ds \right], \end{aligned}$$

which shows that for all $p > 1$

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} (|\eta(t)|^p + |\rho(t)|^p) \right] \leq C_{(T, p, \mu(\Theta))}. \quad (31)$$

Moreover, it follows from (26) that for all $p > 1$

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |\Phi(t) \rho(t)|^p \right] \leq C_{(T, p, \mu(\Theta))}. \quad (32)$$

Next, since $\mathcal{F}_t = (\mathcal{F}_t^{(W, N)})_{t \in [s, T]}$ then by applying *Martingale Representation Theorem for Jump Processes* (see Lemma A2), there exists a unique process $\gamma_t(\cdot) \in \mathbb{L}_{\mathcal{F}}^2([0, t])$ and unique $\xi_t(\cdot, \theta) \in \mathbb{M}_{\mathcal{F}}^2([0, t])$ such that $\forall t \in [0, T]$:

$$\begin{aligned} \Phi(t) \rho(t) &= \mathbb{E}(\Phi(t) \rho(t)) + \int_0^t \gamma_t(s) dW(s) \\ &\quad + \int_0^t \int_{\Theta} \xi_t(s, \theta) N(d\theta, ds). \quad \mathbb{P} - a.s. \end{aligned} \quad (33)$$

Noting that, for every $p > 1$, with the help of (30) it follows from the Burkholder–Davis–Gundy inequality and Proposition A1 that there exists a constant $C_{(T, p, \mu(\Theta))}$ such that for $p > 1$,

$$\begin{aligned} &\mathbb{E} \left[\left(\int_0^t |\gamma_t(s)|^2 ds \right)^{\frac{p}{2}} \right] \\ &\quad + \mathbb{E} \left[\left(\int_0^t \int_{\Theta} |\xi_t(s, \theta)|^2 \mu(d\theta) ds \right)^{\frac{p}{2}} \right] \\ &\leq C_p \mathbb{E} \left[\sup_{0 \leq \tau \leq t} \left| \int_0^\tau \gamma_t(s) dW(s) \right|^p \right] \\ &\quad + C_{(T, p, \mu(\Theta))} \mathbb{E} \left[\sup_{0 \leq \tau \leq t} \left| \int_0^\tau \int_{\Theta} \xi_t(s-, \theta) N(d\theta, ds) \right|^p \right] \\ &\leq C_p \left(1 + \frac{1}{p-1} \right)^p \mathbb{E} \left[\left| \int_0^t \gamma_t(s) dW(s) \right|^p \right] \\ &\quad + C_{(T, p, \mu(\Theta))} \mathbb{E} \left[\left| \int_0^t \int_{\Theta} \xi_t(s-, \theta) N(d\theta, ds) \right|^p \right] \\ &\leq C_{(T, p, \mu(\Theta))} \mathbb{E} [|\Phi(t) \rho(t) - \mathbb{E}(\Phi(t) \rho(t))|^p] \\ &\leq C_{(T, p, \mu(\Theta))} \{ \mathbb{E}(|\Phi(t) \rho(t)|^p) + |\mathbb{E}(\Phi(t) \rho(t))|^p \} \\ &\leq C_{(T, p, \mu(\Theta))} \mathbb{E} [|\Phi(t) \rho(t)|^p] \\ &\leq C_{(T, p, \mu(\Theta))} \mathbb{E} [\sup_{0 \leq t \leq T} |\Phi(t) \rho(t)|^p] \leq C_{(T, p, \mu(\Theta))} \end{aligned}$$

This implies that

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[\left(\int_0^t |\gamma_t(r)|^2 dr \right)^{\frac{p}{2}} \right] \leq C_{(T, p, \mu(\Theta))}, \quad (34)$$

and

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[\left(\int_0^t \int_{\Theta} |\xi_t(s, \theta)|^2 \mu(d\theta) ds \right)^{\frac{p}{2}} \right] \leq C_{(T, p, \mu(\Theta))}. \quad (35)$$

Now, we consider

$$\Phi(t) x_1^\varepsilon(t) = \mathcal{J}_1^\varepsilon(t) + \mathcal{J}_2^\varepsilon(t), \quad t \in [0, T], \quad (36)$$

where

$$\begin{aligned} \mathcal{J}_1^\varepsilon(t) &= \Phi(t) \rho(t) \\ &\quad \times \int_0^t \eta(s) \{ [f_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) + \delta f(t) \mathbf{I}_{\mathcal{A}_\varepsilon}(s)] \\ &\quad - \sigma_x(s) \sigma_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) \} ds, \end{aligned}$$

and

$$\mathcal{J}_2^\varepsilon(t) = \Phi(t) \rho(t) \int_0^t \eta(s) \sigma_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) dW(s).$$

We estimate the first term in the right-hand side of (36). Since $f_{\tilde{x}}$, σ_x and $\sigma_{\tilde{x}}$ are bounded we get

$$\begin{aligned} |\mathbb{E}(\mathcal{J}_1^\varepsilon(t))| &= \left| \mathbb{E} \left\{ \Phi(t) \rho(t) \int_0^t \eta(s) [(f_{\tilde{x}}(s) - \sigma_x(s) \sigma_{\tilde{x}}(s)) \mathbb{E}(x_1^\varepsilon(s)) \right. \right. \\ &\quad \left. \left. + \delta f(t) \mathbf{I}_{\mathcal{A}_\varepsilon}(s)] ds \right\} \right| \\ &\leq C(\mu(\Theta)) \mathbb{E} \left\{ \sup_{t \in [0, T]} |\Phi(t) \rho(t)| \sup_{t \in [0, T]} |\eta(t)| \right. \\ &\quad \left. \times \left[\int_0^t |\mathbb{E}(x_1^\varepsilon(s))| ds + \varepsilon \right] \right\}, \end{aligned}$$

applying Cauchy–Schwarz inequality, then from (31) and (32) (with $p = 2$), we get

$$\begin{aligned} |\mathbb{E}(\mathcal{J}_1^\varepsilon(t))| &\leq C(\mu(\Theta)) \left[\mathbb{E} \left(\sup_{t \in [0, T]} |\Phi(t) \rho(t)|^2 \right) \right]^{\frac{1}{2}} \\ &\quad \times \left[\mathbb{E} \left(\sup_{t \in [0, T]} |\eta(t)|^2 \right) \right]^{\frac{1}{2}} \left[\int_0^t |\mathbb{E}(x_1^\varepsilon(s))| ds + \varepsilon \right] \\ &\leq C(T, \mu(\Theta)) \left[\int_0^t |\mathbb{E}(x_1^\varepsilon(s))| ds + \varepsilon \right], \end{aligned}$$

by applying Cauchy–Schwarz inequality and the fact that $(\alpha + \beta)^2 \leq 2\alpha^2 + 2\beta^2$ we can show that

$$\begin{aligned} |\mathbb{E}(\mathcal{J}_1^\varepsilon(t))|^2 &\leq C(T, \mu(\Theta)) \left[\left(\int_0^t |\mathbb{E}(x_1^\varepsilon(s))| ds \right)^2 + \varepsilon^2 \right] \\ &\leq C(T, \mu(\Theta)) \left[\int_0^t |\mathbb{E}(x_1^\varepsilon(s))|^2 ds + \varepsilon^2 \right]. \quad (37) \end{aligned}$$

Next, we proceed to estimate the second term $\mathcal{J}_2^\varepsilon(t)$. With the help of (33) and the Itô Isometry we have

$$\begin{aligned} \mathbb{E}(\mathcal{J}_2^\varepsilon(t)) &= \mathbb{E} \left\{ \Phi(t) \rho(t) \int_0^t \eta(s) \sigma_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) dW(s) \right\} \\ &= \mathbb{E} \left\{ \left[\mathbb{E}(\Phi(t) \rho(t)) + \int_0^t \gamma_t(s) dW(s) \right. \right. \\ &\quad \left. \left. + \int_0^t \int_{\Theta} \xi_t(s-, \theta) N(d\theta, ds) \right] \right. \\ &\quad \left. \times \int_0^t \eta(s) \sigma_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) dW(s) \right\} \\ &= \mathbb{E} \left[\int_0^t \gamma_t(s) \eta(s) \sigma_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) ds \right]. \end{aligned}$$

Applying (31)–(34) we obtain

$$\begin{aligned} \mathbb{E}(\mathcal{J}_2^\varepsilon(t)) &\leq \left| \mathbb{E} \int_0^t \gamma_t(s) \eta(s) \sigma_{\tilde{x}}(s) \mathbb{E}(x_1^\varepsilon(s)) ds \right|^2 \\ &\leq C \int_0^t |\mathbb{E}(x_1^\varepsilon(s))|^2 ds. \quad (38) \end{aligned}$$

By combining (35)–(38) and the fact that

$$|\mathbb{E}(\Phi(t) x_1^\varepsilon(t))|^2 \leq 2 |\mathbb{E}(\mathcal{J}_1^\varepsilon(t))|^2 + 2 |\mathbb{E}(\mathcal{J}_2^\varepsilon(t))|^2, \quad t \in [0, T], \quad (39)$$

we conclude

$$\begin{aligned} |\mathbb{E}(\Phi(t) x_1^\varepsilon(t))|^2 &\leq C(T, \mu(\Theta)) \left[\varepsilon^2 + \varepsilon + \int_0^t |\mathbb{E}(x_1^\varepsilon(s))|^2 ds \right], \quad (40) \end{aligned}$$

integrating the above inequality, we get

$$\begin{aligned} \int_0^t |\mathbb{E}(\Phi(s) x_1^\varepsilon(s))|^2 ds &\leq C(T, \mu(\Theta)) \left[\varepsilon^2 + \varepsilon + \int_0^t \int_0^r |\mathbb{E}(x_1^\varepsilon(s))|^2 ds dr \right]. \quad (41) \end{aligned}$$

Now, taking $\Phi(t) = 1$ in (41) and from Gronwall's Lemma we have

$$\int_0^t |\mathbb{E}(x_1^\varepsilon(s))|^2 ds \leq C(T, \mu(\Theta)) (\varepsilon^2 + \varepsilon). \quad (42)$$

Consequently, from (41) it holds that

$$\int_0^t |\mathbb{E}(\Phi(s) x_1^\varepsilon(s))|^2 ds \leq C_{(T, \mu(\Theta))} (\varepsilon^2 + \varepsilon). \quad (43)$$

By combining (40), and (42) we conclude

$$|\mathbb{E}(\Phi(T) x_1^\varepsilon(T))|^2 \leq C_{(T, \mu(\Theta))} (\varepsilon^2 + \varepsilon). \quad (44)$$

Finally, by setting $\tilde{\rho}(\varepsilon) = (\varepsilon + \varepsilon^2) \rightarrow 0, \varepsilon \rightarrow 0$, then the desired result (27) follows immediately from (43) and (44) $\square$

Proof of estimate (28) Applying similar argument developed above we can show that

$$\begin{aligned} \int_0^t |\mathbb{E}(y_1^\varepsilon(s) \Phi(s))|^2 ds &\leq C_{(T, \mu(\Theta))} \tilde{\rho}(\varepsilon) \quad \text{and} \\ |\mathbb{E}(\Phi(T) y_1^\varepsilon(T))|^2 &\leq C_{(T, \mu(\Theta))} \tilde{\rho}(\varepsilon), \\ \int_t^T \left\{ |\mathbb{E}(z_1^\varepsilon(s) \psi(s))|^2 + \int_{\Theta} |\mathbb{E}(r_1^\varepsilon(s, \theta) \phi(s, \theta))|^2 \mu(d\theta) \right\} ds \\ &\leq C_{(T, \mu(\Theta))} \tilde{\rho}(\varepsilon), \end{aligned}$$

and for $t \in [T - \tau, T]$.

$$\begin{aligned} |\mathbb{E}(z_1^\varepsilon(T) \psi(T))|^2 + \int_{\Theta} |\mathbb{E}(r_1^\varepsilon(T, \theta) \phi(T, \theta))|^2 \mu(d\theta) \\ \leq C_{(T, \mu(\Theta))} \tilde{\rho}(\varepsilon). \end{aligned}$$

Finally, by using similar iterative method (see Xu [3] and also proof of (22)) then after finite number of iterations the desired result (28) follows. This completes the proof of Lemma 3.3. $\square$

Proof of estimate (22) The first estimate can be proved by standard arguments. By squaring both sides of

$$\begin{aligned} -y_1^\varepsilon(t) - \int_t^T z_1^\varepsilon(s) dW(s) + \int_t^T \int_{\Theta} r_1^\varepsilon(s, \theta) N(d\theta, ds) \\ = -[h_x(T) + \mathbb{E}(h_{\tilde{x}}(T))] x_1^\varepsilon(T) \\ + \int_t^T \int_{\Theta} \{g_x(s, \theta) x_1^\varepsilon(s) + g_{\tilde{x}}(s, \theta) \mathbb{E}(x_1^\varepsilon(s)) \\ + g_y(s, \theta) y_1^\varepsilon(s) + g_{\tilde{y}}(s, \theta) \mathbb{E}(y_1^\varepsilon(s)) \\ + g_z(s, \theta) z_1^\varepsilon(s) + g_{\tilde{z}}(s, \theta) \mathbb{E}(z_1^\varepsilon(s)) \\ + g_r(t, \theta) r_1^\varepsilon(t, \theta) + \delta g(t, \theta) \mathbf{I}_{A_\varepsilon}(t)\} \mu(d\theta) ds, \end{aligned}$$

and the fact that

$$\begin{aligned} \mathbb{E} \left[y_1^\varepsilon(t) \int_t^T z_1^\varepsilon(s) dW(s) \right] &= 0, \\ \mathbb{E} \left[y_1^\varepsilon(t) \int_t^T \int_{\Theta} r_1^\varepsilon(s, \theta) N(d\theta, ds) \right] &= 0, \\ \mathbb{E} \left[\int_t^T z_1^\varepsilon(s) dW(s) \int_t^T \int_{\Theta} r_1^\varepsilon(s, \theta) N(d\theta, ds) \right] &= 0, \end{aligned}$$

we get

$$\begin{aligned} \mathbb{E} |y_1^\varepsilon(t)|^2 + \mathbb{E} \int_t^T |z_1^\varepsilon(s)|^2 ds \\ + \mathbb{E} \int_t^T \int_{\Theta} |r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \\ = \mathbb{E} \{ -[h_x(T) + \mathbb{E}(h_{\tilde{x}}(T))] x_1^\varepsilon(T) \\ + \int_t^T \int_{\Theta} \{g_x(s, \theta) x_1^\varepsilon(s) + g_{\tilde{x}}(s, \theta) \mathbb{E}(x_1^\varepsilon(s)) \\ + g_y(s, \theta) y_1^\varepsilon(s) + g_{\tilde{y}}(s, \theta) \mathbb{E}(y_1^\varepsilon(s)) + g_z(s, \theta) z_1^\varepsilon(s) \\ + g_{\tilde{z}}(s, \theta) \mathbb{E}(z_1^\varepsilon(s)) + g_r(t, \theta) r_1^\varepsilon(t, \theta) \\ + \delta g(s, \theta) \mathbf{I}_{A_\varepsilon}(s)\} \mu(d\theta) ds \}^2, \end{aligned}$$

from the bounded of $h_x, h_{\tilde{x}}, \sup_{\theta \in \Theta} (g_x(t, \theta))$ for $x = x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r$ we obtain

$$\begin{aligned} \mathbb{E} |y_1^\varepsilon(t)|^2 + \mathbb{E} \int_t^T |z_1^\varepsilon(s)|^2 ds + \mathbb{E} \int_t^T \int_{\Theta} |r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \\ \leq C \mathbb{E} |x_1^\varepsilon(T)|^2 + C_{(T, \mu(\Theta))} \mathbb{E} \int_t^T |x_1^\varepsilon(s)|^2 ds \\ + C_{(T, \mu(\Theta))} \mathbb{E} \int_t^T |y_1^\varepsilon(s)|^2 ds \\ + C_{(T, \mu(\Theta))} (T - t) \mathbb{E} \int_t^T |z_1^\varepsilon(s)|^2 ds \\ + C(T - t) \mathbb{E} \int_t^T \int_{\Theta} |r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \\ + \mathbb{E} \int_t^T |\mathbb{E}(x_1^\varepsilon(s))|^2 ds + \mathbb{E} \int_t^T |\mathbb{E}(y_1^\varepsilon(s))|^2 ds \\ + \mathbb{E} \int_t^T |\mathbb{E}(z_1^\varepsilon(s))|^2 ds + C \mathbb{E} \left(\int_t^T \int_{\Theta} \delta g(s, \theta) \mathbf{I}_{A_\varepsilon}(s) \mu(d\theta) ds \right)^2, \end{aligned}$$

for every τ such that $\tau = T - t$, and by choosing $\tau = \frac{1}{2C(T, \mu(\Theta))}$, in virtue of estimate (23) we get

$$\begin{aligned} \mathbb{E} |y_1^\varepsilon(t)|^2 + \frac{1}{2} \mathbb{E} \int_t^T |z_1^\varepsilon(s)|^2 ds \\ + \frac{1}{2} \mathbb{E} \int_t^T \int_{\Theta} |r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \leq C\varepsilon^2, \\ \text{for } t \in [T - \tau, T]. \end{aligned}$$

Similarly we can show that for $t \in [T - 2\tau, T - \tau]$

$$\begin{aligned} \mathbb{E} |y_1^\varepsilon(t)|^2 + \frac{1}{2} \mathbb{E} \int_t^{T-\tau} |z_1^\varepsilon(s)|^2 ds \\ + \frac{1}{2} \mathbb{E} \int_t^{T-\tau} \int_{\Theta} |r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \leq C\varepsilon^2. \end{aligned}$$

Finally, after a finite number of iterations, estimate (22) is proved. $\square$

Proof of estimate (23) By applying (10) we get

$$\begin{aligned} \mathbb{E}(x_1^\varepsilon(t)) = \int_0^t \{ \mathbb{E}[f_x(s)x_1^\varepsilon(s)] + \mathbb{E}(f_{\tilde{x}}(s)) \mathbb{E}(x_1^\varepsilon(s)) \\ + \mathbb{E}(\delta f(s)\mathbf{I}_{\mathcal{A}_\varepsilon}(s)) \} ds, \end{aligned}$$

then we have

$$\begin{aligned} |\mathbb{E}(x_1^\varepsilon(t))|^2 \leq 2 \left| \int_0^t \mathbb{E}[f_x(s)x_1^\varepsilon(s)] ds \right|^2 \\ + 2 \left| \int_0^t \mathbb{E}((f_{\tilde{x}}(s)) \mathbb{E}(x_1^\varepsilon(s)) + \mathbb{E}(\delta f(s)\mathbf{I}_{\mathcal{A}_\varepsilon}(s))) ds \right|^2, \end{aligned} \quad (45)$$

by setting $\Phi(t) = f_x(t)$ in (41), then by Cauchy–Schwarz inequality and fact that $t \leq T$, we get

$$\begin{aligned} \left| \int_0^t \mathbb{E}[f_x(s)x_1^\varepsilon(s)] \right|^2 \leq T \int_0^t |\mathbb{E}[f_x(s)x_1^\varepsilon(s)]|^2 ds \\ \leq C\varepsilon^2, \end{aligned}$$

thus, in view of assumption (H1), we obtain

$$\begin{aligned} |\mathbb{E}(x_1^\varepsilon(t))| \leq [C\varepsilon^2]^{\frac{1}{2}} + C \int_0^t (\varepsilon + \mathbb{E}|x_1^\varepsilon(s)|) ds \\ = C\varepsilon + C \int_0^t (\varepsilon + \mathbb{E}|x_1^\varepsilon(s)|) ds. \end{aligned}$$

Finally the first estimate in (23) follows immediately by applying Gronwall's Lemma.

Let us turn to second estimate in (23). From (10) we have

$$\begin{aligned} y_1^\varepsilon(t) \\ = y_1^\varepsilon(T) - \int_t^T \int_{\Theta} \{ g_x(s, \theta)x_1^\varepsilon(s) + g_{\tilde{x}}(s, \theta)\mathbb{E}(x_1^\varepsilon(s)) \\ + g_y(s, \theta)y_1^\varepsilon(s) + g_{\tilde{y}}(s, \theta)\mathbb{E}(y_1^\varepsilon(s)) + g_z(s, \theta)z_1^\varepsilon(s) \\ + g_{\tilde{z}}(s, \theta)\mathbb{E}(z_1^\varepsilon(s)) + g_r(s, \theta)r_1^\varepsilon(s, \theta) \\ + \delta g(s, \theta)\mathbf{I}_{\mathcal{A}_\varepsilon}(s) \} \mu(d\theta) ds - \int_t^T z_1^\varepsilon(s) dW(s) \\ + \int_t^T \int_{\Theta} r_1^\varepsilon(s, \theta) N(d\theta, ds), \end{aligned}$$

by taking expectation and squaring both side of the above equation we get

$$\begin{aligned} |\mathbb{E}(y_1^\varepsilon(t))|^2 \\ \leq C \left| \int_t^T \int_{\Theta} \{ \mathbb{E}[g_x(s, \theta)x_1^\varepsilon(s)] + \mathbb{E}[g_y(s, \theta)y_1^\varepsilon(s)] \right. \\ + \mathbb{E}[g_z(s, \theta)z_1^\varepsilon(s)] + \mathbb{E}[g_r(s, \theta)r_1^\varepsilon(s, \theta)] \} \mu(d\theta) ds \Big|^2 \\ + C \left| \int_t^T (\mathbb{E}(g_{\tilde{x}}(s, \theta)) \mathbb{E}(x_1^\varepsilon(s)) + \mathbb{E}(g_{\tilde{y}}(s, \theta)) \mathbb{E}(y_1^\varepsilon(s)) \right. \\ + \mathbb{E}(g_{\tilde{z}}(s, \theta)) \mathbb{E}(z_1^\varepsilon(s)) + \mathbb{E}[\delta g(s, \theta)\mathbf{I}_{\mathcal{A}_\varepsilon}(s)]) ds \Big|^2. \end{aligned} \quad (46)$$

Arguing as in (41) and by setting $\psi(t) = g_x(t, \theta)$, with the help of Cauchy–Schwarz inequality we get

$$\begin{aligned} \left| \int_t^T \int_{\Theta} \mathbb{E}[g_x(s, \theta)x_1^\varepsilon(s)] \mu(d\theta) ds \right|^2 \\ \leq (T-t)C \int_t^T \int_{\Theta} |\mathbb{E}[g_x(s, \theta)x_1^\varepsilon(s)]|^2 \mu(d\theta) ds, \end{aligned}$$

for every τ such that $\tau = T - t$, we get by choosing $\tau = \frac{1}{C}$, for $t \in [T - \tau, T]$.

$$\left| \int_t^T \int_{\Theta} \mathbb{E}[g_x(s, \theta)x_1^\varepsilon(s)] \mu(d\theta) ds \right|^2 \leq \varepsilon^2.$$

Similarly for the terms in the right hand side of inequality (46), then by setting $\psi(t) = g_z(t, \theta)$, $\phi(t, \theta) = g_r(t, \theta)$ in (28), with the fact that $\sup_{\theta \in \Theta} g_\varkappa(s, \theta)$, for $\varkappa := x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r$ are bounded and $\int_0^T \mathbf{I}_{\mathcal{A}_\varepsilon}(s) ds = \varepsilon$ we obtain

$$\begin{aligned}
& |\mathbb{E}(y_1^\varepsilon(t))|^2 \\
& \leq C\varepsilon + C_{\mu(\Theta)} \int_t^T \left\{ \varepsilon + |\mathbb{E}(y_1^\varepsilon(s))|^2 + |\mathbb{E}(z_1^\varepsilon(s))|^2 \right. \\
& \quad \left. + \int_{\Theta} |\mathbb{E}(r_1^\varepsilon(s, \theta))|^2 \mu(d\theta) \right\} ds. \quad (47)
\end{aligned}$$

Applying similar arguments developed in (42)–(43) and taking $\psi(t) = \phi(t, \theta) = 1$ we can get

$$\begin{aligned}
& \int_t^T |\mathbb{E}(z_1^\varepsilon(s))|^2 ds \leq C_{\mu(\Theta)} \tilde{\rho}(\varepsilon). \\
& \int_t^T \int_{\Theta} |\mathbb{E}(r_1^\varepsilon(s, \theta))|^2 \mu(d\theta) ds \leq C_{\mu(\Theta)} \tilde{\rho}(\varepsilon), \quad (48)
\end{aligned}$$

combining (47)–(48) we get

$$\begin{aligned}
|\mathbb{E}(y_1^\varepsilon(t))|^2 & \leq C\varepsilon + (T-t)C_{\mu(\Theta)}\tilde{\rho}(\varepsilon) \\
& \quad + C_{\mu(\Theta)} \int_t^T [\varepsilon + |\mathbb{E}(y_1^\varepsilon(s))|^2] ds,
\end{aligned}$$

for every τ such that $\tau = T - t$, we get by choosing $\tau = \frac{1}{C_{\mu(\Theta)}}$,

$$\begin{aligned}
|\mathbb{E}(y_1^\varepsilon(t))|^2 & \leq C\varepsilon + \tilde{\rho}(\varepsilon) \\
& \quad + C_{\mu(\Theta)} \int_t^T [\varepsilon + |\mathbb{E}(y_1^\varepsilon(s))|^2] ds, \quad t \in [T - \tau, T], \quad (49)
\end{aligned}$$

by applying Gronwall's Lemma we get

$$|\mathbb{E}(y_1^\varepsilon(t))|^2 \leq \tilde{\rho}(\varepsilon), \quad t \in [T - \tau, T]. \quad (50)$$

Finally, after a finite number of iterations, estimate (23) follows by combining (48) and (50). $\square$

Proof of estimate (25) We set $t \in [0, T]$, $\theta \in \Theta$, $\alpha \in [0, 1]$

$$\begin{aligned}
A(t, \theta, \alpha) &= (x^*(t) + \alpha x_1^\varepsilon(t), \mathbb{E}(x^*(t) + \alpha x_1^\varepsilon(t)), \\
& \quad y^*(t) + \alpha y_1^\varepsilon(t), \mathbb{E}(y^*(t) + \alpha y_1^\varepsilon(t)), \\
& \quad z^*(t) + \alpha z_1^\varepsilon(t), \mathbb{E}(z^*(t) + \alpha z_1^\varepsilon(t)), r^*(t, \theta)).
\end{aligned}$$

and

$$\begin{aligned}
& G^\varepsilon(t, \theta) \\
& = x_1^\varepsilon(t) \int_0^1 \{g_x(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_x(t, \theta) \\
& \quad + \mathbb{E}[g_x(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_{\tilde{x}}(t, \theta)]\} d\alpha \\
& \quad + y_1^\varepsilon(t) \int_0^1 \{g_y(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_y(t, \theta) \\
& \quad + \mathbb{E}[g_{\tilde{y}}(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_{\tilde{y}}(t, \theta)]\} d\alpha \\
& \quad + z_1^\varepsilon(t) \int_0^1 \{g_z(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_z(t, \theta) \\
& \quad + \mathbb{E}[g_{\tilde{z}}(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_{\tilde{z}}(t, \theta)]\} d\alpha \\
& \quad + r_1^\varepsilon(t, \theta) \int_0^1 \{g_r(t, A(t, \theta, \alpha), u^\varepsilon(t)) - g_r(t, \theta)\} d\alpha.
\end{aligned}$$

By a simple computation we can show that

$$\begin{aligned}
& \int_t^T \int_{\Theta} g(s, x^*(s) + x_1^\varepsilon(s), \\
& \quad \mathbb{E}(x^*(s) + x_1^\varepsilon(s)), y^*(s) + y_1^\varepsilon(s), \\
& \quad \mathbb{E}(y^*(s) + y_1^\varepsilon(s)), z^*(s) + z_1^\varepsilon(s), \\
& \quad \mathbb{E}(z^*(s) + z_1^\varepsilon(s)), \\
& \quad r^*(s, \theta) + r_1^\varepsilon(s, \theta), u^\varepsilon(s)) \mu(d\theta) ds \\
& \quad + \int_t^T (z^*(s) + z_1^\varepsilon(s)) dW(s) \\
& \quad + \int_t^T \int_{\Theta} (r^*(s, \theta) + r_1^\varepsilon(s, \theta)) dN(d\theta, ds) \\
& = h(x^*(T), \mathbb{E}(x^*(T))) \\
& \quad + [h_x(x^*(T), \mathbb{E}(x^*(T))) \\
& \quad + \mathbb{E}(h_{\tilde{x}}(x^*(T), \mathbb{E}(x^*(T))))] x_1^\varepsilon(T) \\
& \quad - y^*(t) - y_1^\varepsilon(t) + \int_t^T \int_{\Theta} G^\varepsilon(t, \theta) \mu(d\theta) ds.
\end{aligned}$$

Since

$$\begin{aligned}
& -(y^\varepsilon(t) - y^*(t) - y_1^\varepsilon(t)) \\
& \quad - \int_t^T (z^\varepsilon(s) - z^*(s) + z_1^\varepsilon(s)) dW(s) \\
& \quad - \int_t^T \int_{\Theta} (r^\varepsilon(s, \theta) - r^*(s, \theta) - r_1^\varepsilon(s, \theta)) N(d\theta, ds)
\end{aligned}$$

$$\begin{aligned}
&= -[h(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T))) - h(x^*(T), \mathbb{E}(x^*(T)))] \\
&\quad + [h_x(x^*(T), \mathbb{E}(x^*(T))) \\
&\quad + \mathbb{E}(h_{\tilde{x}}(x^*(T), \mathbb{E}(x^*(T))))] x_1^\varepsilon(T) \\
&\quad + \int_t^T \int_{\Theta} \{g(s, x^\varepsilon(s), \mathbb{E}(x^\varepsilon(s)), y^\varepsilon(s), \\
&\quad \mathbb{E}(y^\varepsilon(s)), z^\varepsilon(s), \mathbb{E}(z^\varepsilon(s)), u^\varepsilon(s)) \\
&\quad - g(s, x^*(s) + x_1^\varepsilon(s), \mathbb{E}(x^*(s) + x_1^\varepsilon(s)), \\
&\quad y^*(s) + y_1^\varepsilon(s), \mathbb{E}(y^*(s) + y_1^\varepsilon(s)), z^*(s) + z_1^\varepsilon(s), \\
&\quad \mathbb{E}(z^*(s) + z_1^\varepsilon(s)), \\
&\quad r^*(s, \theta) + r_1^\varepsilon(s, \theta), u^\varepsilon(s))\} \mu(d\theta) ds \\
&\quad + \int_t^T \int_{\Theta} G^\varepsilon(s, \theta) \mu(d\theta) ds,
\end{aligned}$$

then by squaring both sides of the above equation, taking expectation and the fact that

$$\begin{aligned}
&\mathbb{E} \left\{ (y^\varepsilon(t) - y^*(t) - y_1^\varepsilon(t)) \right. \\
&\quad \times \left. \int_t^T (z^\varepsilon(s) - z^*(s) + z_1^\varepsilon(s)) dW(s) \right\} = 0, \\
&\mathbb{E} \left\{ (y^\varepsilon(t) - y^*(t) - y_1^\varepsilon(t)) \right. \\
&\quad \times \left. \int_{\Theta} (r^\varepsilon(s, \theta) - r^*(s, \theta) - r_1^\varepsilon(s, \theta)) N(d\theta, ds) \right\} = 0,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E} \left\{ \int_t^T (z^\varepsilon(s) - z^*(s) + z_1^\varepsilon(s)) dW(s) \right. \\
&\quad \times \left. \int_{\Theta} (r^\varepsilon(s, \theta) - r^*(s, \theta) - r_1^\varepsilon(s, \theta)) N(d\theta, ds) \right\} = 0,
\end{aligned}$$

we get

$$\begin{aligned}
&\mathbb{E} |y^\varepsilon(t) - y^*(t) - y_1^\varepsilon(t)|^2 \\
&\quad + \mathbb{E} \int_t^T |z^\varepsilon(s) - z^*(s) + z_1^\varepsilon(s)|^2 ds \\
&\quad + \mathbb{E} \int_t^T \int_{\Theta} |r^\varepsilon(s, \theta) - r^*(s, \theta) + r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds, \\
&= \mathbb{E} \left\{ -[h(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T))) \right.
\end{aligned}$$

$$\begin{aligned}
&\quad - h(x^*(T) + x_1^\varepsilon(t), \mathbb{E}(x^*(T)) + x_1^\varepsilon(t))] \\
&\quad - x_1^\varepsilon(T) \int_0^1 \{ [h_x(x^*(T) + \alpha x_1^\varepsilon(T), \mathbb{E}(x^*(T) + \alpha x_1^\varepsilon(T))) \\
&\quad - h_x(x^*(T), \mathbb{E}(x^*(T)))] \mathbb{E}[h_{\tilde{x}}(x^*(T) + \alpha x_1^\varepsilon(T), \\
&\quad \mathbb{E}(x^*(T) + \alpha x_1^\varepsilon(T)) - h_{\tilde{x}}(x^*(T), \mathbb{E}(x^*(T)))] \} d\alpha \\
&\quad + \int_t^T \int_{\Theta} [g(s, x^\varepsilon(s), \mathbb{E}(x^\varepsilon(s)), y^\varepsilon(s), \\
&\quad \mathbb{E}(y^\varepsilon(s)), z^\varepsilon(s), \mathbb{E}(z^\varepsilon(s)), u^\varepsilon(s)) \\
&\quad - g(s, x^*(s) + x_1^\varepsilon(s), \mathbb{E}(x^*(s) + x_1^\varepsilon(s)), \\
&\quad y^*(s) + y_1^\varepsilon(s), \mathbb{E}(y^*(s) + y_1^\varepsilon(s)), z^*(s) + z_1^\varepsilon(s), \\
&\quad \mathbb{E}(z^*(s) + z_1^\varepsilon(s)), r^*(s, \theta) + r_1^\varepsilon(s, \theta), u^\varepsilon(s))] \mu(d\theta) ds \\
&\quad + \left. \int_t^T \int_{\Theta} G^\varepsilon(s, \theta) \mu(d\theta) ds \right\}^2,
\end{aligned}$$

by using (22)–(23) and the fact that

$$\begin{aligned}
&\mathbb{E} \int_t^T \int_{\Theta} G^\varepsilon(s, \theta) \mu(d\theta) ds \\
&= \mathbb{E} \int_t^T \int_{\Theta} \int_0^1 x_1^\varepsilon(s) \{ g_x(s, A(s, \theta, \alpha), u^\varepsilon(s)) - g_x(s, \theta) \\
&\quad + \mathbb{E}[g_x(s, A(s, \theta, \alpha), u^\varepsilon(s)) - g_{\tilde{x}}(s, \theta)] \} d\alpha \mu(d\theta) ds \\
&\quad + \mathbb{E} \int_t^T \int_{\Theta} \int_0^1 y_1^\varepsilon(s) \{ g_y(s, A(s, \theta, \alpha), u^\varepsilon(s)) - g_y(s, \theta) \\
&\quad + \mathbb{E}[g_{\tilde{y}}(s, A(s, \theta, \alpha), u^\varepsilon(s)) - g_{\tilde{y}}(s, \theta)] \} d\alpha \mu(d\theta) ds \\
&\quad + \mathbb{E} \int_t^T \int_{\Theta} \int_0^1 z_1^\varepsilon(s) \{ g_z(s, A(s, \theta, \alpha), u^\varepsilon(s)) - g_z(s, \theta) \\
&\quad + \mathbb{E}[g_{\tilde{z}}(s, A(s, \theta, \alpha), u^\varepsilon(s)) - g_{\tilde{z}}(s, \theta)] \} d\alpha \mu(d\theta) ds \\
&\quad + \mathbb{E} \int_t^T \int_{\Theta} \int_0^1 r_1^\varepsilon(s, \theta) \{ g_r(s, A(s, \theta, \alpha), u^\varepsilon(s)) \\
&\quad - g_r(s, \theta) \} d\alpha \mu(d\theta) ds,
\end{aligned}$$

we can show that

$$\begin{aligned}
&\mathbb{E} [h(x^\varepsilon(T), \mathbb{E}(x^\varepsilon(T))) - h(x^*(T) \\
&\quad + x_1^\varepsilon(T), \mathbb{E}(x^*(T)) + x_1^\varepsilon(T))]^2 = o(\varepsilon), \\
&\sup_{t \in [0, T]} \mathbb{E} \left[\int_t^T \int_{\Theta} G^\varepsilon(s, \theta) \mu(d\theta) ds \right]^2 = o(\varepsilon^2).
\end{aligned}$$

Thus

$$\mathbb{E} |y^\varepsilon(t) - y^*(t) - y_1^\varepsilon(t)|^2 \leq C\varepsilon^2,$$

and for $t \in [T - \tau, T]$

$$\mathbb{E} \int_t^T |z^\varepsilon(s) - z^*(s) + z_1^\varepsilon(s)|^2 ds \leq C\varepsilon^2,$$

$$\mathbb{E} \int_t^T \int_{\Theta} |r^\varepsilon(s, \theta) - r^*(s, \theta) + r_1^\varepsilon(s, \theta)|^2 \mu(d\theta) ds \leq C\varepsilon^2.$$

Finally, by applying similar iterative method developed in proof of (22) then the desired result (25) follows. $\square$

3.4 Maximum principle

Lemma 3.4 *Let Assumptions (H1) and (H2) hold. The following variational inequality holds*

$$o(\varepsilon) \leq \mathbb{E} \left\{ [\varphi_y(y^*(0), \mathbb{E}(y^*(0))) + \mathbb{E}(\varphi_{\tilde{y}}(y^*(0), \mathbb{E}(y^*(0))))] y_1^\varepsilon(0) \right\}.$$

Proof From Lemma 3.2 we have

$$\mathbb{E} \left\{ \varphi(y^\varepsilon(0), \mathbb{E}(y^\varepsilon(0))) - \varphi(y^*(0), \mathbb{E}(y^*(0))) + y_1^\varepsilon(0) \right\} = o(\varepsilon).$$

Therefore

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

the desired result (25) follows. $\square$

Proof of Theorem 3.1 From (13) and Lemma 3.4, we obtain

$$o(\varepsilon) \leq \mathbb{E} \left\{ [\varphi_y(y^*(0), \mathbb{E}(y^*(0))) + \mathbb{E}(\varphi_{\tilde{y}}(y^*(0), \mathbb{E}(y^*(0))))] y_1^\varepsilon(0) \right\} \\ = \mathbb{E} \int_0^T \left\{ \Psi^*(t) \delta f(t) + \int_{\Theta} K^*(t) \delta g(t, \theta) \mu(d\theta) \right\} \\ \times \mathbf{I}_{\mathcal{A}_\varepsilon}(t) dt,$$

since

$$\Psi^*(t) \delta f(t) + \int_{\Theta} K^*(t) \delta g(t, \theta) \mu(d\theta) = \delta H(t),$$

we obtain

$$\mathbb{E} \int_0^T \delta H(t) \mathbf{I}_{\mathcal{A}_\varepsilon}(t) dt \geq o(\varepsilon).$$

Finally, we deduce

$$\mathbb{E} \int_0^T H(t, \lambda^*(t, \theta), u^*(t), \Lambda^*(t, \theta)) dt \\ = \min_{u \in \mathbb{A}} \mathbb{E} \int_0^T H(t, \lambda^*(t, \theta), u, \Lambda^*(t, \theta)) dt.$$

The proof of Theorem 3.1 is complete. $\square$

3.5 General cost functional of mean-field type

Noting that when the expected cost has the general mean-field form

$$J(u(\cdot)) = \mathbb{E} \left[\int_0^T \int_{\Theta} \ell(t, x(t), \mathbb{E}(x(t)), y(t), \mathbb{E}(y(t)), z(t), \mathbb{E}(z(t)), r(t, \theta), u(t)) \mu(d\theta) dt, \right. \\ \left. + \phi(x(T), \mathbb{E}(x(T))) + \varphi(y(0), \mathbb{E}(y(0))) \right], \quad (51)$$

where ϕ, ℓ are some appropriate functions, the mean-field maximum principle for the control problem (1)–(51) is given as follows. We assume

3.5.1 Condition (H3)

The functions $\ell : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{A} \rightarrow \mathbb{R}$, and $\phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ are continuously differentiable with respect to $(x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r)$, and the derivatives of ℓ, ϕ with respect to $(x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r)$ are bounded.

Let $(\Psi(\cdot), Q(\cdot), K(\cdot), R(\cdot))$ be solution of the adjoint equation

$$\begin{cases} d\Psi(t) = -\{f_x(t) \Psi(t) + \mathbb{E}[f_{\tilde{x}}(t) \Psi(t)] \\ \quad + \sigma_x(t) Q(t) + \mathbb{E}[\sigma_{\tilde{x}}(t) Q(t)] \\ \quad + \int_{\Theta} [g_x(t, \theta) K(t) + \mathbb{E}(g_{\tilde{x}}(t, \theta) K(t)) \\ \quad + c_x(t, \theta) R(t, \theta)] \mu(d\theta) dt \\ \quad + Q(t) dW(t) + \int_{\Theta} R_t(\theta) N(d\theta, dt), \\ \Psi(T) = -\{h_x(T) K(T) + \mathbb{E}[(h_{\tilde{x}}(T)) K(T)] \\ \quad + \phi_x(T) + \mathbb{E}(\phi_{\tilde{x}}(T))\}, \\ dK(t) = \int_{\Theta} [g_y(t, \theta) K(t) \\ \quad + \mathbb{E}(g_{\tilde{y}}(t, \theta) K(t))] \mu(d\theta) dt \\ \quad + \int_{\Theta} [g_z(t, \theta) K(t) + \mathbb{E}(g_{\tilde{z}}(t, \theta) K(t))] \mu(d\theta) dW(t) \\ \quad - \int_{\Theta} g_r(t, \theta) K(t) N(d\theta, dt) \\ K(0) = -\{\varphi_y(y(0), \mathbb{E}(y(0))) \\ \quad + \mathbb{E}[\varphi_{\tilde{y}}(y(0), \mathbb{E}(y(0)))]\}, \end{cases} \quad (52)$$

and the hamiltonian function $\mathcal{H}$ is given by

$$\mathcal{H}(t, x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r(\cdot), u, \Psi, Q, K, R(\cdot)) \\ = \Psi(t) f(t, x, \tilde{x}, u) + Q(t) \sigma(t, x, \tilde{x}) \\ + \int_{\Theta} [K(t) g(t, x, \tilde{x}, y, \tilde{y}, z, \tilde{z}, r(t, \theta), u) \\ + R(t, \theta) c(t, x, \theta)) - \ell(t, \lambda_t(\theta), u)] \mu(d\theta). \quad (53)$$

Corollary 3.1 *Let Assumptions (H1)–(H3) hold. Let $(x^*(\cdot), y^*(\cdot), z^*(\cdot), r^*(\cdot, \cdot))$ be the trajectory of the mean-field FBSDEJ-(1) corresponding to the optimal control $u^*(\cdot)$, and $(\Psi^*(\cdot), Q^*(\cdot), K^*(\cdot), R^*(\cdot))$ be the solution of adjoint equation (52) corresponding to $u^*(\cdot)$. Then the maximum principle holds, that is*

$$\begin{aligned}
& \mathcal{H}(t, \lambda^*(t, \theta), u^*(t), \Lambda^*(t, \theta)) \\
&= \min_{u \in \mathbb{A}} \mathcal{H}(t, \lambda^*(t, \theta), u, \Lambda^*(t, \theta)) \\
& \mathbb{P}\text{-a.s., a.e. } t \in [0, T].
\end{aligned} \tag{54}$$

The proof is obtained by applying similar arguments developed in Theorem 3.1, where

$$\begin{aligned}
\delta \mathcal{H}(t) &= \Psi^*(t) \delta f(t) \\
&+ \int_{\Theta} [K^*(t) \delta g(t, \theta) - \delta \ell(t, \theta)] \mu(d\theta),
\end{aligned}$$

and

$$\begin{aligned}
\delta \ell(t, \theta) &= \ell(t, x^*(t), \mathbb{E}(x^*(t)), y^*(t), \mathbb{E}(y^*(t)), \\
& \quad z^*(t), \mathbb{E}(z^*(t)), r^*(t, \theta), u(t)) \\
& \quad - \ell(t, x^*(t), \mathbb{E}(x^*(t)), y^*(t), \mathbb{E}(y^*(t)), \\
& \quad z^*(t), \mathbb{E}(z^*(t)), r^*(t, \theta), u^*(t)).
\end{aligned}$$

3.6 Some discussions

In this paper, we have established a set of necessary conditions for optimal stochastic control in the form of a stochastic maximum principle for systems governed by mean-field FBSDEJs, where the coefficients depend not only on the state process but also its marginal distribution of the state process through its expected value. The cost functional is also of mean field type. When the coefficients of the underlying mean-field FBSDEJs and the cost functional not explicitly depend on the marginal law, our result (Theorem 3.1), reduces as a particular case, to stochastic maximum principle of optimality, proved in Shi [9]. Apparently, there are many problems left unsolved. To mention a few, sufficient conditions, one may consider the controlled diffusion and jump coefficients, and more interestingly, the general case following the idea of Shi [9], where the system will evolve according to the following mean-field FBSDEJs

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

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

Appendix

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

Lemma 3.5 (Integration by parts formula for jumps processes) *Suppose that the processes $x_1(t)$ and $x_2(t)$ are given by: for $i = 1, 2, t \in [0, T]$:*

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

Then we get

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

See Framstad et al. ([12], Lemma 2.1) for the detailed proof of the above Lemma. $\square$

Proposition 3.1 *Let $\mathcal{G}$ be the predictable σ -field on $\Omega \times [0, T]$, and f be a $\mathcal{G} \times \mathcal{B}(\Theta)$ -measurable function such that*

$$\mathbb{E} \int_0^T \int_{\Theta} |f(s, \theta)|^2 \mu(d\theta)ds < \infty,$$

then for all $p \geq 2$ there exists a positive constant $C = C(T, p, \mu(\Theta))$ such that

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| \int_0^t \int_{\Theta} f(s, \theta)N(d\theta, ds) \right|^p \right] \\
& \leq C \mathbb{E} \left[\int_0^T \int_{\Theta} |f(s, \theta)|^p \mu(d\theta)ds \right].
\end{aligned}$$

See Bouchard and Elie ([10], Appendix). $\square$

Lemma 3.6 (Martingale representation theorem for jump diffusion processes) *Let $\mathcal{G}$ be a finite-dimensional space and let $M(t)$ be an $\mathcal{G}$ -valued $\mathcal{F}_t$ -adapted square-integrable martingale. Then there exist $q(\cdot) \in \mathbb{L}_{\mathcal{F}}^2([s, T], \mathcal{G})$ and $\zeta(\cdot, \cdot) \in \mathbb{M}_{\mathcal{F}}^2([s, T], \mathcal{G})$ such that*

$$M(t) = M(s) + \int_s^t q(r)dW(r) + \int_s^t \int_{\Theta} \zeta(r, \theta)N(d\theta, dr).$$

The proof can be found in Tang and Li ([29] Lemma 2.3). $\square$

References

1. Shi J, Wu Z (2006) The maximum principle for fully coupled forward–backward stochastic control system. *Acta Autom Sin* 32(2):161–169
2. Shi J, Wu Z (2010) Maximum principle for partially-observed optimal control of fully-coupled forward–backward stochastic systems. *J Optim Theory Appl* 145(3):543–578
3. Xu W (1995) Stochastic maximum principle for optimal control problem of forward and backward. *J Aust Math Soc Ser B* 37:172–185
4. Peng S, Wu Z (1999) Fully coupled forward backward stochastic differential equations and application to optimal control. *SIAM J Control Optim* 37(3):825–843
5. Yong J (2010) Optimality variational principle for controlled forward–backward stochastic differential equations with mixed initial-terminal conditions. *SIAM J Control Optim* 48(6):4119–4156
6. Ma J, Yong J (1999) Forward–backward stochastic differential equations and their applications. *Lecture notes in mathematics*, vol 1702. Springer, Berlin
7. 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
8. Shi J, Wu Z (2007) Maximum principle for fully coupled Forward–backward stochastic control system with random jumps. In: *Proceedings of the 26th Chinese Control Conference*, Zhangjiajie, Hunan, China, pp 375–380
9. Shi J (2012) Necessary conditions for optimal control of forward–backward stochastic systems with random jumps. *J Stoch Anal* 258674:50. doi:[10.1155/2012/258674](https://doi.org/10.1155/2012/258674)
10. Bouchard B, Elie R (2008) Discrete time approximation of decoupled forward–backward SDE with jumps. *Stoch Proc Appl* 118(1):53–75
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. 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
14. Øksendal B, Sulem A (2007) *Applied stochastic control of jump diffusions*, 2nd edn. Springer, Berlin
15. Kac M (1956) Foundations of kinetic theory. In: *Proceedings of 3rd Berkeley Symposium on mathematical statistics and probability*, vol 3, pp 171–197
16. McKean HP (1966) A class of Markov processes associated with nonlinear parabolic equations. *Proc Natl Acad Sci USA* 56:1907–1911
17. Shi J (2012) Sufficient conditions of optimality for mean-field stochastic control problems. 12th International Conference on control, automation, robotics & vision, (ICARCV 2012), Guangzhou, China, 5–7 Dec 2012
18. Ahmed NU (2007) Nonlinear diffusion governed by McKean-Vlasov equation on Hilbert space and optimal control. *SIAM J Control Optim* 46:356–378
19. Buckdahn R, Li J, Peng S (2009) Mean-field backward stochastic differential equations and related partial differential equations. *Stoch Process Appl* 119:3133–3154
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. Kotelenez PM, Kurtz TG (2010) Macroscopic limit for stochastic partial differential equations of McKean-Vlasov type. *Probab Theory Relat Fields* 146:189–222
22. Hafayed M, Abbas S (2013) A general maximum principle for stochastic differential equations of mean-field type with jump processes. Technical report
23. Lasry JM, Lions PL (2007) Mean field games. *Japan J Math* 2:229–260
24. Andersson D, Djehiche B (2011) A maximum principle for SDEs of mean-field type. *Appl Math Optim* 63:341–356
25. Li J (2012) Stochastic maximum principle in the mean-field controls. *Automatica* 48:366–373
26. Meyer-Brandis T, Øksendal B, Zhou XY (2012) A mean-field stochastic maximum principle via Malliavin calculus. *Stochastics* 84(5–6):643–666
27. Yong J (2011) A linear-quadratic optimal control problem for mean-field stochastic differential equations. Technical report
28. Carmona R, Dularue F (2012) Mean-field forward–backward stochastic differential equations, Technical report
29. Tang SL, Li XJ (1994) Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J Control Optim* 32:1447–1475