Discussiones Mathematicae
Differential Inclusions, Control and Optimization 35 (2015) 165–195
doi:10.7151/dmdico.1171

OPTIMAL CONTROL OF GENERAL MCKEAN-VLASOV STOCHASTIC EVOLUTION EQUATIONS ON HILBERT SPACES AND NECESSARY CONDITIONS OF OPTIMALITY

N.U. Ahmed

University of Ottawa, Canada
e-mail: ahmed@site.uottawa.ca

Abstract

In this paper we consider controlled McKean-Vlasov stochastic evolution equations on Hilbert spaces. We prove existence and uniqueness of solutions and regularity properties thereof. We use relaxed controls, adapted to a current of sub-sigma algebras generated by observable processes, and taking values from a Polish space. We introduce an appropriate topology based on weak star convergence. We prove continuous dependence of solutions on controls with respect to appropriate topologies. Theses results are then used to prove existence of optimal controls for Bolza problems. Then we develop the necessary conditions of optimality based on semi-martingale representation theory on Hilbert spaces. Next we show that the adjoint processes arising from the necessary conditions optimality can be constructed from the solution of certain BSDE.

Keywords: McKean-Vlasov stochastic differential equation, Hilbert spaces, relaxed controls, existence of optimal controls.

2010 Mathematics Subject Classification: 49J27, 60H15,93E20.

1. Introduction

It is well known that stochastic differential equations of Itô type generate linear diffusion. A more general class of stochastic systems is governed by McKean-Vlasov equations in which the coefficients are not only functions of the state but also of the probability measure induced by the state itself. This makes the corresponding diffusion nonlinear. A special case of McKean-Vlasov equation is the mean-field equation in which the coefficients depend not only on the state but also on its mean. This class of systems have been studied extensively in the literature [13, 14, 17,] after McKean introduced this model in [18]. Control problems involving this general model have been studied in [1, 2, 3, 4].

In recent years intensive research has been going on in the area of necessary conditions of optimality for stochastic systems governed by Itô differential equations defined on finite as well as infinite dimensional spaces along the line of the Pontryagin minimum principle [1, 2, 3, 4, 5, 6, 7, 9, 10, 16]. See also the extensive references given therein. Control of McKean-Vlasov type stochastic differential equations were studied in [1, 2, 3, 4, 19]. In [19] Shen and Siu consider maximum principle for jump-diffusion mean-field model on finite dimensional spaces giving some examples from finance. Here in this paper we wish to study the question of existence of optimal controls as well as present necessary conditions of optimality for the general class of McKean-Vlasov evolution equations on infinite dimensional Hilbert spaces.

Shen and Siu [19] presented maximum principle for a class of finite dimensional jump-diffusion stochastic differential equations. The cost functional is of Bolza type. In [16], Hu and Peng developed some fundamental results on the question of existence and uniqueness of a class of backward stochastic evolution equations (BSDE) on Hilbert spaces. In [2], we considered control of McKean-Vlasov equations and presented existence of optimal controls. In [4], we considered McKean-Vlasov equations on finite dimensional spaces and developed HJB equations. The author is not aware of any literature where the question of existence of optimal controls and necessary conditions of optimality for the general McKean-Vlasov stochastic evolution equations on infinite dimensional Hilbert spaces have been considered. This is what motivates us to consider optimal control of these equations on infinite dimensional spaces and develop necessary conditions of optimality thereof.

The paper is organized as follows. In Section 2, we present the mathematical model of the controlled system followed by some mathematical framework in Section 3. In Section 4, after basic assumptions are introduced, we prove the existence and uniqueness of mild solutions and their regularity properties. Existence of optimal control is proved in Section 5. In Section 6, we present the necessary conditions of optimality. For illustration of the abstract results, in Section 7 we consider some examples of linear quadratic regulator problems involving linear McKean-Vlasov dynamics.

2. System model

Let X and H denote a pair of real separable Hilbert spaces and $\{\Omega, \mathcal{F}, \mathcal{F}_t, t \in I, P\}$ a complete filtered probability space with $\mathcal{F}_t \subset \mathcal{F}$ a family of nondecreasing complete sub-sigma algebras of the sigma algebra $\mathcal{F}$ and $I \equiv [0, T]$, $T < \infty$. Let $W \equiv \{W(t), t \in I\}$, denote an H-Wiener process with covariance operator $\mathcal{R}$ in the sense that for any $h \in H$, $(W(t), h)$ is a real Brownian motion on I with variance $\mathbf{E}(W(t), h)^2 = t(\mathcal{R}h, h)$. If the operator $\mathcal{R} = I_H$, the identity operator in H, we say that W is a cylindrical Brownian motion or cylindrical Wiener process; and if $\mathcal{R}$ is nuclear we have the H-valued Wiener process. Since we are interested in controlled evolution equation we must now introduce the class of feasible controls. Let U be a compact Polish space and $\mathcal{M}(U)$ the space of Borel measures on the sigma algebra $\mathcal{B}(U)$ on U. Let $\mathcal{M}_1(U) \subset \mathcal{M}(U)$ denote the space of probability measures on U. Let $\mathcal{G}_t \subset \mathcal{F}_t$ denote another current of nondecreasing family of sub-sigma algebras of sigma-algebras $\mathcal{F}_t$ and let $L_\infty^\alpha(I, \mathcal{M}_1(U))$ denote the class of weak star measurable $\mathcal{G}_t$-adapted $\mathcal{M}_1(U)$ valued random processes. For any Banach space Z, let $C(U, Z)$ denote the Banach space of Z-valued continuous functions defined on U furnished with the standard sup-norm topology, that is, for any $\phi \in C(U, Z)$, its norm is given by $\| \phi \| \equiv \sup\{|\phi(\xi)|_Z, \xi \in U\}$. For any $\phi \in C(U, Z)$ and $u \in \mathcal{M}_1(U)$, the integral $\Phi(u) \equiv \int_U \phi(\xi) u(d\xi)$ is well defined as Bochner integral with values in Z.

Now we are prepared to introduce the system considered in this paper. It is governed by the following McKean-Vlasov controlled evolution equation on the Hilbert space X driven by the H-Brownian motion W and the control measure u:

$$\begin{aligned} &dx = Axdt + f(t, x, \mu, u)dt + \sigma(t, x, \mu, u)dW, \ x(0) = x_0, \\ &\text{and} \quad \mu(t) = \mathcal{P}(x(t)), \ t \in I \equiv [0, T]. \end{aligned} \tag{1}$$

where A is the infinitesimal generator a C_0-semigroup $S(t)$, $t \in I$, on X and f is a Borel measurable map from $I \times X \times \mathcal{M}_1(X) \times \mathcal{M}_1(U)$ to X and σ is also a Borel measurable map from $I \times X \times \mathcal{M}_1(X) \times \mathcal{M}_1(U)$ to $\mathcal{L}(H, X)$, the space of bounded linear operators from H to X, and x_0 is the initial state. We have denoted the probability law of any stochastic process $\{\zeta(t), t \geq 0\}$ by $\mathcal{P}(\zeta(t))$, $t \geq 0$.

The drift f and the diffusion σ are not only dependent on the current state $x(t)$ but also its probability law $\mu(t) \equiv \mathcal{P}(x(t))$, the measure induced by the X-valued random variable $x(t)$. We assume throughout the paper that both f and σ are given by

$$f(t, x, \mu, u) \equiv \int_U f(t, x, \mu, \xi) u(d\xi), \sigma(t, x, \mu, u) \equiv \int_U \sigma(t, x, \mu, \xi) u(d\xi)$$

for any $u \in \mathcal{M}_1(U)$. In case both X and H are finite dimensional, this class of models arise naturally in finance where the objective functional is of mean-variance type maximizing terminal wealth while minimizing variance. Also such models are known to arise in biological population process.

3. Mathematical framework

Let $\mathcal{B}(X)$ denote the Borel σ-algebra generated by closed (or open) subsets of the Hilbert space X and $\mathcal{M}_1(X)$ is the space of probability measures on $\mathcal{B}(X)$ carrying the usual topology of weak convergence. Let $C(X)$ denote the space of continuous functions on X. We use the notation $(\mu, \varphi) \equiv \mu(\varphi) \equiv \int_X \varphi(x)\mu(dx)$ whenever this integral makes sense. Throughout this paper we let γ denote the continuous function $\gamma(x) \equiv 1 + |x|, x \in X$, and introduce the Banach space

$$C_\rho(X) = \left\{ \varphi \in C(X) : ||\varphi||_{C_\rho(X)} \equiv \sup_{x\in X} \frac{|\varphi(x)|}{\gamma^2(x)} + \sup_{x\neq y} \frac{|\varphi(x) - \varphi(y)|}{|x-y|} < \infty \right\}.$$

For $p \geq 1$, let $\mathcal{M}^s_{\gamma^p}(X)$ denote the Banach space of signed measures m on X satisfying $||\mu||_{\gamma^p} \equiv \left(\int_X \gamma^p(x)|m|(dx)\right)^{1/p} < \infty$, where $|m| = m^+ + m^-$ denotes the total variation of the signed measure m, with $m = m^+ - m^-$ being the Jordan decomposition of m. Let $\mathcal{M}_{\gamma^2}(X) = \mathcal{M}^s_{\gamma^2}(X) \cap \mathcal{M}_1(X)$ denote the class of probability measures possessing second moments. We put on $\mathcal{M}_{\gamma^2}(X)$ a topology induced by the following metric:

$$\rho(\mu, \nu) = \sup\ \{(\mu - \nu)(\varphi) \equiv (\varphi, \mu - \nu) : \varphi \in C_\rho(X) \text{ and } ||\varphi||_{C_\rho(X)} \leq 1\}.$$

Then $(\mathcal{M}_{\gamma^2}(X), \rho) \equiv \mathcal{M}_{2,\rho}(X)$ forms a complete metric space. Note that this is a closed bounded subset of the closed unit ball in the linear metric space $\mathcal{M}^s_{2,\rho}(X) \equiv (\mathcal{M}^s_{\gamma^2}, \rho)$. Define $I \equiv [0, T]$ with $T < \infty$. We denote by $C(I, \mathcal{M}_{2,\rho}(X))$ the complete metric space of continuous functions from I to $\mathcal{M}_{2,\rho}(X)$ with the metric:

$$D(\mu, \nu) = \sup\{\rho(\mu(t), \nu(t)),\ t \in I\}$$

for any $\mu, \nu \in C(I, M_{2,\rho}(X))$. From now on all stochastic processes considered in this paper are assumed to be based on the complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ with $\mathcal{F}_T \subseteq \mathcal{F}$. For convenience of notation we denote the space $L_2((\Omega, \mathcal{F}, P), X)$ by $L_2(\Omega, X)$ and let $C(I, L_2(\Omega, X))$ denote the Banach space of continuous $\mathcal{F}$-measurable functions defined on I and taking values from $L_2(\Omega, X)$ satisfying the condition $\sup_{t\in I} E|x(t)|^2_X < \infty$. Let Λ_2 denote the closed subspace of $C(I; L_2(\Omega, X))$ consisting of continuous F_t-adapted (progressively measurable) X-valued random processes $x = \{x(t) : t \in I \equiv [0, T]\}$. Then, Λ_2 is a Banach space with respect to the norm topology given by $|x|_{\Lambda_2} = (\sup_{t\in I} E|x(t)|^2)^{1/2}$. We denote by $L_2^{\mathcal{F}_T}(\Omega, X)$ the space of $\mathcal{F}_T$ measurable X valued random variables having finite second moments. Similarly, we use $L_2^{\mathcal{F}}(I, X) \equiv L_2^{\mathcal{F}}(I \times \Omega, X)$ to denote the Banach space of $\mathcal{F}_t$-adapted X-valued norm-square integrable random processes defined on I. Let $\mathcal{L}_\mathcal{R}(H, X)$ denote the completion of the space of linear operators from H to X with respect to the inner product $< K, L >\equiv Tr(K\mathcal{R}L^*)$

and norm $|K|_{\mathcal{R}} \equiv \sqrt{Tr(K\mathcal{R}K^*)}$. Clearly this is a Hilbert space. In the sequel we also need the Hilbert space $L_2^{\mathcal{F}}(I, \mathcal{L}_{\mathcal{R}}(H,X))$ which consists of $\mathcal{F}_t$-adapted $\mathcal{L}_{\mathcal{R}}(H,X)$ valued random processes having finite square integrable norms in the sense that for any $K \in L_2^{\mathcal{F}}(I, \mathcal{L}_{\mathcal{R}}(H,X))$ we have $\mathbf{E}\int_I |K|_{\mathcal{R}}^2 dt < \infty$.

4. Basic assumptions and existence of solutions

Now we are prepared to introduce the basic assumptions. In order to study control problems involving the system (1) we must now state the basic properties of the drift and the diffusion operators $\{f, \sigma\}$ including the semigroup generator.

Basic Assumptions:

(A1): The operator A is the infinitesimal generator of a C_0-semigroup $S(t)$, $t \geq 0$, on the Hilbert space X satisfying

$$\sup\left\{\| S(t) \|_{\mathcal{L}(E)},\ t \in I\right\} \leq M < \infty.$$

(A2): The function $f : I \times X \times \mathcal{M}_1(X) \times U \longrightarrow X$ is measurable in the first argument and continuous with respect to the rest of the arguments. Further, there exists a constant $K \neq 0$ such that

$$|f(t,x,\mu,\xi)|_X^2 \leq K^2\{1 + |x|_X^2 + |\mu|_{\mathcal{M}_{\gamma^2}}^2\},\ \forall\ x, y \in X, \xi \in U$$
$$|f(t,x_1,\mu_1,\xi) - f(t,x_2,\mu_2,\xi)|_X^2 \leq K^2\{|x_1 - x_2|_X^2 + \rho^2(\mu_1,\mu_2)\},$$

for all $x_1, x_2 \in X, \mu_1, \mu_2 \in \mathcal{M}_{2,\rho}(X)$ uniformly with respect to $t \in I,\ \xi \in U$.

(A3): The incremental covariance of the Brownian motion W denoted by $\mathcal{R} \in \mathcal{L}_s^+(H)$ (symmetric, positive). The diffusion $\sigma : I \times X \times \mathcal{M}_1(X) \times U \longrightarrow \mathcal{L}(H,X)$ is measurable in the first argument and continuous with respect to the rest of the variables and there exists a constant $K_{\mathcal{R}} \neq 0$ such that

$$|\sigma(t,x,\mu,\xi)|_{\mathcal{R}}^2 \leq K_{\mathcal{R}}^2\{1 + |x|_X^2 + |\mu|_{\mathcal{M}_{\gamma^2}}^2\},\ \forall\ x, y \in X, \mu \in \mathcal{M}_{\gamma^2}$$
$$|\sigma(t,x_1,\mu_1,\xi) - \sigma(t,x_2,\mu_2,\xi)|_{\mathcal{R}}^2 \leq K_{\mathcal{R}}^2\{|x_1 - x_2|_X^2 + \rho^2(\mu_1,\mu_2)\}$$

for all $x_1, x_2 \in X$ and $\mu_1, \mu_2 \in M_{2,\rho}(X)$ uniformly with respect to $(t,\xi) \in I \times U$, where $|\sigma|_{\mathcal{R}}^2 = tr(\sigma\mathcal{R}\sigma^*)$.

For admissible controls, let $\mathcal{G}_t, t \geq 0$, denote a nondecreasing family of subsigma algebras of the current of sigma algebras $\mathcal{F}_t, t \geq 0$. Let U be a compact Polish space and $\mathcal{M}_1(U)$ the space of probability measures on U. For admissible controls, we choose the set $\mathcal{U}_{ad} \equiv L_\infty^{\mathcal{G}}(I, \mathcal{M}_1(U)) \subset L_\infty^{\mathcal{F}}(I, \mathcal{M}(U))$ which consist of $\mathcal{G}_t$-adapted $\mathcal{M}_1(U)$-valued random processes, endowed with the weak star topology. This is the class of relaxed controls. It follows from Alaoglu's theorem that $\mathcal{U}_{ad}$ is weak star compact. In contrast, let $\mathcal{U}_r$ denote the class of measurable functions on I with values in U, called regular controls. It is clear that the following embedding $\mathcal{U}_r \hookrightarrow \mathcal{U}_{ad}$, through the mapping $\mathcal{U}_r \ni u(\cdot) \longrightarrow \delta_{u(\cdot)} \in \mathcal{U}_{ad}$, is continuous. It follows from the well known Krien-Millman theorem that the closed convex hull of the extremals of any weak star compact set is weak star compact. The set of extremals of $\mathcal{U}_{ad}$ is given by $\mathcal{U}_r$, and hence $\mathcal{U}_{ad} = clco(\mathcal{U}_r)$. Thus any relaxed control from $\mathcal{U}_{ad}$ can be approximated as closely as required by regular controls from $\mathcal{U}_r$. There are several reasons for choosing relaxed controls. For example, it is well known from control theory of deterministic systems that there are examples (time optimal control) where optimal control does not exist in the regular class $\mathcal{U}_r$ but does so in the relaxed class $\mathcal{U}_{ad}$. For relaxed controls, the set U can be non-convex, discrete etc.

To prove the existence of solution of the stochastic evolution equation (1) we need the following.

Lemma 4.1. *Consider the system* (1) *and suppose the assumptions* (A1)–(A3) *hold. Further, suppose that* $W \equiv \{W(t), t \geq 0\}$ *is an* H*-Brownian motion with incremental covariance* (*operator*) $\mathcal{R} \in \mathcal{L}_1^+(H)$. *Then, for every* $\mathcal{F}_0$ *measurable* X *valued random variable* $x_0 \in L_2^{\mathcal{F}_0}(\Omega, X)$, *and control* $u \in \mathcal{U}_{ad}$, *and* $\nu \in C(I, \mathcal{M}_{2,\rho}(X))$, *the stochastic evolution equation given by*

$$dx = Axdt + f(t, x, \nu, u)dt + \sigma(t, x, \nu, u)dW, \; x(0) = x_0, \; t \in I \equiv [0, T], \tag{2}$$

has a unique mild solution $x = x_\nu \in \Lambda_2$ *in the sense that it satisfies the following stochastic integral equation:*

$$\begin{aligned} x_\nu(t) \equiv S(t)x_0 &+ \int_0^t S(t-\tau) f(\tau, x_\nu(\tau), \nu(\tau), u_\tau) d\tau \\ &+ \int_0^t S(t-\tau)\sigma(\tau, x_\nu(\tau), \nu(\tau), u_\tau) dW(\tau) \;\; t \in I. \end{aligned} \tag{3}$$

Further the solution has a continuous modification.

Proof. First we show that for every given $\nu \in C(I, \mathcal{M}_{2,\rho}(X))$, the solution of the integral equation (3), if one exists, has an a-priori bound. Clearly, for the given $\nu \in C(I, \mathcal{M}_{2,\rho}(X))$, there exists a finite positive number b such that $\| \nu \|_{C(I,M_{2,\rho})(X)} \equiv \sup\{\| \nu(t) \|_{\gamma^2}, t \in I\} \leq b$. Then using equation (3) and computing the expected value of the square of the norm of $x_\nu(t)$ one can easily obtain the following inequality,

$$\mathbf{E}|x_\nu(t)|_X^2 \leq C_1 + C_2 \int_0^t \mathbf{E}|x_\nu(s)|_X^2 ds, \tag{4}$$

where

$$C_1 \equiv 4M^2\{\mathbf{E}|x_0|_X^2 + (TK^2 + K_R^2)\int_0^T (1 + |\nu(s)|_{\gamma^2}^2)ds\}$$
$$C_2 \equiv 4M^2(TK^2 + K_R^2).$$

Hence, it follows from Gronwall inequality applied to (4) that

$$\sup\{\mathbf{E}|x_\nu(t)|_X^2, t \in I\} \leq C_1 \exp\{C_2 T\}. \tag{5}$$

Next we show that under the assumptions (A1)–(A3), the integral equation has a unique solution $x_\nu \in \Lambda_2$. For the fixed ν, define the operator F_ν by

$$\begin{aligned}(F_\nu x)(t) &\equiv S(t)x_0 + \int_0^t S(t-\tau)f(\tau, x(\tau), \nu(\tau), u_\tau)d\tau \\ &+ \int_0^t S(t-\tau)\sigma(\tau, x(\tau), \nu(\tau), u_\tau)dW(\tau) \;\; t \in I.\end{aligned} \tag{6}$$

It is clear from the a priori bound proved above that $F_\nu : \Lambda_2 \longrightarrow \Lambda_2$. We prove that it has unique fixed point in Λ_2. For any pair of $x, y \in \Lambda_2$, it follows from the Lipschitz property of f and σ (see (A2)-(A3)) that

$$\sup_{0\leq s\leq t} \mathbf{E}|(F_\nu x)(s) - (F_\nu y)(s)|_X^2 \leq \alpha(t) \sup_{0\leq s\leq t}\{\mathbf{E}|x(s) - y(s)|_X^2\} \tag{7}$$

where

$$\alpha(t) = 2M^2\{K^2t^2 + K_{\mathcal{R}}^2 t\},\;\; t \in I.$$

For $s, t \in I$, $s < t$, let $\Lambda_2[s,t]$ denote the restriction of the Banach space Λ_2 over the interval $[s,t] \subset I$. Clearly, it follows from the inequality (7) that

$$\| F_\nu x - F_\nu y \|_{\Lambda_2[0,t]} \leq \sqrt{\alpha}(t) \| x - y \|_{\Lambda_2[0,t]},\;\; t \in I. \tag{8}$$

Since α is a continuous and monotone increasing function of $t \in I$, with $\alpha(0) = 0$, there exists $t_1 \in I \equiv (0,T]$ such that $\alpha(t_1) < 1$. Thus it follows from the expression (8) that F_ν is a contraction on $\Lambda_2[0,t_1]$ and therefore by Banach fixed point theorem it has a unique fixed point $x^1 \in \Lambda_2[0,t_1]$. Further, it follows from the well known factorization technique [12] that x^1 has a continuous modification which we continue to denote by x^1. Clearly $x^1(t_1)$ is $\mathcal{F}_{t_1}$ measurable and it belongs to $L_2(\Omega, X)$. Using this $x^1(t_1)$ as the initial condition, we consider the integral operator F_ν over the interval $[t_1, T]$ giving

$$\begin{aligned}(F_\nu x)(t) &\equiv S(t-t_1)x_1(t_1) + \int_{t_1}^t S(t-\tau)f(\tau, x(\tau), \nu(\tau), u_\tau)d\tau \\ &+ \int_{t_1}^t S(t-\tau)\sigma(\tau, x(\tau), \nu(\tau), u_\tau)dW(\tau) \;\; t \in [t_1, T].\end{aligned} \tag{9}$$

Again it follows from the property of the function α that there exists $t_2 \in (t_1, T]$ such that $\alpha(t_2 - t_1) < 1$ and therefore the operator F_ν restricted to the Banach space $\Lambda_2[t_1, t_2]$ is a contraction and hence by the Banach fixed point theorem, it has unique fixed point $x^2 \in \Lambda_2[t_1, t_2]$ having continuous modification. We continue this process starting with $x(t_2) \equiv x^2(t_2)$ for the remaining interval $[t_2, T]$. Since I is a compact interval, it can be covered by the union of a finite number of compact subintervals $\{[t_i, t_{i+1}]\}_{i=0}^{n-1}$, with $t_0 = 0$ and $t_n = T$. Then the solution of the integral equation (3) is given by the concatenation of the processes $\{x^1, x^2, \ldots, x^n\}$ defined on the intervals $\{I_i, i = 1, 2 \ldots, n\}$ proving that $x_\nu \in \Lambda_2$ is a unique fixed point of the operator F_ν. This proves that the integral equation (3) has a unique solution and hence the evolution equation (2) has a unique mild solution. This completes the proof. ■

Now we are prepared to consider the question of existence of solution of the McKean-Vlasov evolution equation (1). By a solution of this equation, we mean the solution of the following integral equation

$$\begin{aligned} x(t) &= S(t)x_0 + \int_0^t S(t-s)f(s, x(s), \mu(s), u_s)ds \\ &\quad + \int_0^t S(t-s)\sigma(s, x(s), \mu(s), u_s)dW(s), \ t \in I, \\ &\text{with} \ \ \mu(t) = \mathcal{P}(x(t)), \ t \in I. \end{aligned} \tag{10}$$

Theorem 4.2. *Consider the system* (1) *and suppose the assumptions of Lemma* 4.1 *hold. Then the system* (1) *has a unique mild solution $x \in \Lambda_2$ satisfying the integral equation* (10) *with probability law $\mu \in C(I, \mathcal{M}_{2,\rho}(X))$ such that $\mathcal{P}(x(t)) = \mu(t)$ for all $t \in I$.*

Proof. For any given $\nu \in C(I, \mathcal{M}_{2,\rho}(X))$, consider the evolution equation (2). By Lemma (4.1), we know that it has a unique mild solution $x_\nu \in \Lambda_2$ having continuous modification. Define the operator $\Phi : C(I, \mathcal{M}_{2,\rho}(X)) \longrightarrow C(I, \mathcal{M}_{2,\rho}(X))$ taking values

$$\Phi(\nu)(t) \equiv \mathcal{P}(x_\nu(t)), \ t \in I.$$

It is clear that if the operator Φ has a fixed point in $C(I, \mathcal{M}_{2,\rho}(X))$, that is $\Phi(\mu) = \mu$, then equation (1) has a unique mild solution and conversely, if equation (1) has a mild solution $x \in \Lambda_2$, then $\mathcal{P}(x(t)) = \mu(t)$, $t \in I$, and μ is the fixed point of the operator Φ. Thus it suffices to prove that Φ has a unique fixed point $C(I, \mathcal{M}_{2,\rho}(X))$. For any fixed but arbitrary $\mathcal{F}_0$-measurable initial condition $x_0 \in L_2(\Omega, X)$ and control $u \in \mathcal{U}_{ad}$, consider the evolution equation (2) corresponding to $\nu = \lambda$ and $\nu = \vartheta$ separately where $\lambda, \vartheta \in C(I, \mathcal{M}_{2,\rho}(X))$. By Lemma 4.1, equation (2) has unique mild solutions $x_\lambda, x_\vartheta \in \Lambda_2$ corresponding to λ and ϑ respectively. Clearly, these are solutions of the following integral equations

$$\begin{aligned} x_\lambda(t) &\equiv S(t)x_0 + \int_0^t S(t-\tau)f(\tau, x_\lambda(\tau), \lambda(\tau), u_\tau)d\tau \\ &+ \int_0^t S(t-\tau)\sigma(\tau, x_\lambda(\tau), \lambda(\tau), u_\tau)dW(\tau) \ \ t \in I. \end{aligned} \tag{11}$$

$$\begin{aligned} x_\vartheta(t) &\equiv S(t)x_0 + \int_0^t S(t-\tau)f(\tau, x_\vartheta(\tau), \vartheta(\tau), u_\tau)d\tau \\ &+ \int_0^t S(t-\tau)\sigma(\tau, x_\vartheta(\tau), \vartheta(\tau), u_\tau)dW(\tau) \ \ t \in I. \end{aligned} \tag{12}$$

Subtracting equation (12) from equation (11) and following similar steps as in the proof of Lemma 4.1, the reader can easily verify that

$$\begin{aligned} &\sup_{0\le t\le \tau} \mathbf{E}|x_\lambda(t) - x_\vartheta(t)|_X^2 \\ &\le \alpha(\tau)\Big\{\sup_{0\le t\le \tau} \mathbf{E}|x_\lambda(t) - x_\vartheta(t)|_X^2 + \sup_{0\le t\le \tau} \rho^2(\lambda(t).\vartheta(t))\Big\}. \end{aligned} \tag{13}$$

Using the inequality (13) and choosing $\tau = t_1 \in (0, T]$, sufficiently small, so that $\alpha(t_1) < (1/3)$, we arrive at the following inequality

$$\sup_{0\le t\le t_1} \mathbf{E}|x_\lambda(t) - x_\vartheta(t)|_X^2 \le (1/2) \sup_{0\le t\le t_1} \rho^2(\lambda(t), \vartheta(t)). \tag{14}$$

Recall that by definition of the operator Φ, $(\Phi\lambda)(t) = \mathcal{P}(x_\lambda(t))$ and $(\Phi\vartheta)(t) = \mathcal{P}(x_\vartheta(t))$ for $t \in I$. Then computing the distance between the measures $(\Phi\lambda)(t)$ and $(\Phi\vartheta)(t)$, it follows from the definition of the metric ρ that

$$\begin{aligned} &\rho((\Phi\lambda)(t), (\Phi\vartheta)(t)) \\ &= \sup\big\{< \varphi, (\Phi\lambda)(t) - (\Phi\vartheta)(t) >: \varphi \in C_\rho, \| \varphi \|_{C_\rho} \le 1\big\} \\ &= \sup\big\{\mathbf{E}[\varphi(x_\lambda(t)) - \varphi(x_\vartheta(t))] : \varphi \in C_\rho, \| \varphi \|_{C_\rho} \le 1\big\} \\ &\le E|x_\lambda(t) - x_\vartheta(t)|_X. \end{aligned} \tag{15}$$

Clearly, it follows from the above inequality that

$$\sup_{0\le t\le t_1} \rho^2((\Phi\lambda)(t), (\Phi\vartheta)(t)) \le \sup_{0\le t\le t_1} \mathbf{E}|x_\lambda(t) - x_\vartheta(t)|_X^2. \tag{16}$$

Therefore, it follows from the inequalities (14) and (16) that

$$\sup_{0\le t\le t_1} \rho^2((\Phi\lambda)(t), (\Phi\vartheta)(t)) \le (1/2) \sup_{0\le t\le t_1} \rho^2(\lambda(t), \vartheta(t)).$$

from which we arrive at the following inequality

$$\sup_{0\leq t\leq t_1} \rho((\Phi\lambda)(t),(\Phi\vartheta)(t)) \leq (1/\sqrt{2}) \sup_{0\leq t\leq t_1} \rho(\lambda(t),\vartheta(t)). \tag{17}$$

This shows that Φ is a contraction on the restriction $C([0,t_1],\mathcal{M}_{2,\rho}(X))$ of the metric space $C([0,T],\mathcal{M}_{2,\rho}(X))$ and hence by Banach fixed point theorem, it has a unique fixed point, say, $\mu^1 \in C([0,t_1],\mathcal{M}_{2,\rho}(X))$, that is, $(\Phi\mu^1)(t) = \mu^1(t)$, $t \in [0,t_1]$. Next, choosing $t_2 \in (t_1,T]$ such that $\alpha(t_2 - t_1) \leq (1/3)$ and carrying out similar analysis, we arrive at the following inequality,

$$\sup_{t_1\leq t\leq t_2} \rho((\Phi\lambda)(t),(\Phi\vartheta)(t)) \leq (1/\sqrt{2}) \sup_{t_1\leq t\leq t_2} \rho(\lambda(t),\vartheta(t)). \tag{18}$$

Thus Φ, restricted to the metric space $C([t_1,t_2],\mathcal{M}_{2,\rho}(X))$, is again a contraction and hence it has a unique fixed point $\mu^2 \in C([t_1,t_2],\mathcal{M}_{2,\rho}(X))$ giving $\Phi\mu^2 = \mu^2$ for $t \in [t_1,t_2]$ with $\mu^2(t_1) = \mu^1(t_1)$. Continuing this process we can exhaust the interval in a finite number steps and obtain a finite sequence of measure valued functions $\{\mu^i \in C([t_{i-1},t_i],\mathcal{M}_{2,\rho}(X)),\ i = 1,2,\ldots,n\}$ with $t_0 = 0, t_n = T$. Again, by concatenation of these measure valued functions, we obtain μ which coincides with μ^i on the interval $[t_{i-1},t_i]$ for $i \in \{1,2,\ldots,n\}$ satisfying $(\Phi\mu)(t) = \mu(t)$, $t \in I$ proving that Φ has a unique fixed point in $C(I,\mathcal{M}_{2,\rho}(X))$. Hence the McKean-Vlasov evolution equation (1) has unique mild solution $x \in \Lambda_2$ with probability law $\mu \in C(I,\mathcal{M}_{2,\rho}(X))$. This completes the proof. ■

Corollary 4.3. *Suppose the assumptions of Theorem 4.2 hold with the admissible controls $\mathcal{U}_{ad} \equiv L^{\mathcal{G}}_{\infty}(I,\mathcal{M}_1(U))$. Then the solution set $\Xi \equiv \{x(u), u \in \mathcal{U}_{ad}\}$ is a bounded subset of Λ_2 and the corresponding set of measure valued functions lies in a bounded subset of $C(I,\mathcal{M}_{2,\rho}(X))$.*

Proof. We present a brief outline. Let $x(u) \in \Lambda_2$ denote the solution of the integral equation (10) corresponding to any control $u \in \mathcal{U}_{ad}$ and let $\mu^u \in C(I,\mathcal{M}_{2,\rho}(X))$ denote the associated measure valued function. It follows from the first part of the assumptions (A2)–(A3) that, for any given $x \in X$ and $\mu \in \mathcal{M}_{\gamma^2}(X)$, both f and σ are uniformly bounded with respect to controls. Hence, using the integral equation (10), and the fact that

$$\begin{aligned} |\mu^u(s)|^2_{\mathcal{M}_{\gamma^2}} &\equiv \int_X \gamma^2(x)\mu^u(s)(dx) \\ &= \int_X (1+|x|_X)^2\mu^u(s)(dx) \leq 2(1+E|x^u(s)|^2_X), \end{aligned} \tag{19}$$

it is easy to verify that, for every $u \in \mathcal{U}_{ad}$, we have

$$\mathbf{E}|x(u)(t)|^2_E \leq C_1(T) + C_2(T)\int_0^t \mathbf{E}|x(u)(s)|^2_E ds,\ t \in I, \tag{20}$$

where

$$C_1(T) \equiv 4M^2\{\mathbf{E}|x_0|_E^2 + 3K^2T^2 + 3K_R^2T\} \text{and } C_2(T) \equiv 12M^2(K^2T + K_R^2).$$

The constants C_1 and C_2 are independent of control. Thus the first conclusion follows from Gronwall inequality applied to the expression (20) and the second conclusion follows from the first and the inequality (19). This completes the proof. ∎

Remark 4.4. In Theorem 4.1, we assumed that $\{f, \sigma\}$ satisfy uniform Lipschitz condition. In fact this uniform Lipschitz condition is not essential. By using stopping time arguments this can be relaxed to local Lipschitz condition.

5. Existence of optimal control

For the proof of existence of optimal controls we use lower semicontinuity and compactness arguments. For this we prove the continuity of the map $u \longrightarrow x$, that is, the control to solution map. Since continuity is critically dependent on the topology, we must mention the topologies used for the control space and the solution space. For the solution space we have already the norm topology on Λ_2 (see Section 3). So we consider an admissible topology for the control space. In a recent paper [1], we introduced a topology on the control space which is weaker than the one we introduce here. The reason for this shift is to remove the compactness assumption on the semigroup $S(t), t \geq 0$, used in [1]. Let U be a compact Polish space and $C(U)$ the Banach space of continuous functions with the usual sup-norm topology. Let $\mathcal{M}(U)$ denote the space of finite Borel measures on U (more precisely on $\mathcal{B}(U)$ the class of Borel subsets of U). Equipped with the norm topology induced by the total variation, this is a Banach space. It is well known that $\mathcal{M}(U)$ is the topological dual of $C(U)$ and hence for any continuous linear functional $\ell \in (C(U))^*$, there exists a unique $u \in \mathcal{M}(U)$ such that

$$\ell(\varphi) = \int_U \varphi(\xi)u(d\xi).$$

Since U is a compact Polish space, the space $C(U)$ with the usual sup-norm topology is a separable Banach space. We are interested in partially observed relaxed controls. Let $\mathcal{G}_t, t \geq 0$, denote a nondecreasing family of complete sub-sigma algebras of the current of sigma algebras $\mathcal{F}_t$, $t \geq 0$. Let λ denote the Lebesgue measure on I and P the probability measure on Ω and $\lambda \times P$ the product measure on $I \times \Omega$. Let $\mathcal{P}_r$ denote the sigma algebra generated by $\mathcal{G}_t$-predictable subsets of the set $I \times \Omega$ and μ the restriction of the product measure $\lambda \times P$ onto $\mathcal{P}_r$. We assume that $(I \times \Omega, \mathcal{P}_r, \mu)$ is a complete separable measure space. Let $L_1(\mu, C(U))$ denote the Lebegue-Bochner space. Since $\mathcal{M}(U)$ does not satisfy RNP (Radon-Nikodym property), it follows from the theory of lifting that its topological dual is given by $L^{\alpha}_{\infty}(\mu, \mathcal{M}(U))$ which consists of weak star μ-measurable essentially bounded random processes with values in $\mathcal{M}(U)$. In other words

$$(L_1(\mu, C(U)))^* \cong L^{\alpha}_{\infty}(\mu, \mathcal{M}(U)).$$

Thus for any continuous linear functional $\ell \in L_1(\mu, C(U)))^*$ there exists a unique $u \in L^{\alpha}_{\infty}(\mu, \mathcal{M}(U))$ such that

$$\ell(\varphi) = \int_{U\times I\times\Omega} \varphi(t,\omega,\xi) u_{t,\omega}(d\xi) d\mu \equiv \int_{I\times\Omega} u_{t,\omega}(\varphi_{t,\omega}) d\mu.$$

For the set of admissible controls our natural choice is the set $\mathcal{U} \equiv L^{\alpha}_{\infty}(\mu, \mathcal{M}_1(U)) \subset L^{\alpha}_{\infty}(\mu, \mathcal{M}(U))$. Since the measure space $(I \times \Omega, \mathcal{P}_r, \mu)$ is complete separable, the Banach space $L_1(\mu, C(U))$ is separable and hence, it follows from [see Dunford & Schwartz [15], Theorem V.5.1, p. 426] that the set $L^{\alpha}_{\infty}(\mu, \mathcal{M}_1(U))$ is metrizable with the metric δ given by,

$$\delta(u,v) = \sum_{n=1}^{\infty} (1/2^n) \frac{|\int_{I\times\Omega}\{u_{t,\omega}(g^n_{t,\omega}) - v_{t,\omega}(g^n_{t,\omega}\} d\mu|}{1 + |\int_{I\times\Omega}\{u_{t,\omega}(g^n_{t,\omega}) - v_{t,\omega}(g^n_{t,\omega}\} d\mu|},$$

where the set $\{g^n\}$ is dense in $L_1(\mu, C(U))$. With respect to this metric topology, $(\mathcal{U}, \delta) \equiv \mathcal{U}_\delta$ is a compact metric space. This topology is rather weak. We use slightly stronger topology. For the space $\mathcal{U}$ we introduce the following metric topology. Let $D \equiv \{g_n\}$ be a dense subset of $L_2(\mu, C(U))$ and define the function $d : \mathcal{U} \times \mathcal{U} \longrightarrow [0, 1]$ by

$$d(u,v) \equiv \sum_{n=1}^{\infty} (1/2^n) \min\left\{1, \left(\int_{I\times\Omega} |u_{t,\omega}(g_n) - v_{t,\omega}(g_n)|^2 d\mu\right)^{1/2}\right\}$$

for $u, v \in \mathcal{U}$ where $u_{t,\omega}(g) \equiv \int_U g_{t,\omega}(\xi) u_{t,\omega}(d\xi)$. The reader can easily verify that d defines a metric on $\mathcal{U}$. We denote this metric space by $\mathcal{U}_d$ and show that it is complete. Let $\{u^k\} \subset \mathcal{U}_d$ be a Cauchy sequence. Then it follows from the expression for d that, for each $g \in L_2(\mu, C(U))$, $\{u^k(g)\}$ is a Cauchy sequence in $L_2(\mu) \equiv L_2(I \times \Omega, \mathcal{P}_r, \mu)$ and therefore it has a unique limit, say $h_g \in L_2(\mu)$. On the other hand, it follows from Alaoglu's theorem that $L^{\alpha}_{\infty}(\mu, \mathcal{M}_1(U))$ is weak star compact and since this space is Hausdorf, there exists a unique $u^o \in L^{\alpha}_{\infty}(\mu, \mathcal{M}_1(U))$ such that for every $g \in L_1(\mu, C(U))$ we have

$$\lim_{k\to\infty} \int_{I\times\Omega} u^k(g) d\mu \longrightarrow \int_{I\times\Omega} u^o(g) d\mu.$$

Equivalently, $u^k(g) \xrightarrow{w} u^o(g)$ in $L_1(\mu)$ for every $g \in L_1(\mu, C(U))$. Since $(I \times \Omega, \mathcal{P}_r, \mu)$ is a finite measure space it is clear that $L_2(\mu, C(U)) \subset L_1(\mu, C(U))$ and therefore $u^k(g) \xrightarrow{w} u^o(g)$ in $L_2(\mu)$ for every $g \in L_2(\mu, C(U))$. Clearly, $h_g - u^o(g) \in L_2(\mu)$ and it follows from Hahn-Banach theorem that there exists an $e \in L_2(\mu)$ with $\| e \|_{L_2(\mu)} = 1$ such that $\| h_g - u^o(g) \|_{L_2(\mu)} = (e, h_g - u^o(g))$. The reader can easily verify from this that $h_g = u^o(g)$ for every $g \in L_2(\mu, C(U))$. Hence we conclude that

$$d(u^k, u^o) \longrightarrow 0, \quad \text{as} \quad k \to \infty,$$

and therefore $\mathcal{U}_d$ is a complete metric space. Let V be a closed and totally bounded subset of $\mathcal{U}_d$. Then it follows from well known Borel-Lebesgue theorem that $V_d \equiv (V, d)$ is a compact metric space. For admissible controls, we choose $\mathcal{U}_{ad} \equiv (V, d) \equiv V_d$.

To prove the existence of optimal controls we use the following result on continuous dependence of solutions on controls.

Theorem 5.1. *Consider the control system* (1) *with the admissible controls* $\mathcal{U}_{ad} = V_d$. *Suppose the assumptions of Theorem* 4.2 *hold. Then, the control to solution map* $u \longrightarrow x$ *is continuous with respect to the metric topology* d *on* $\mathcal{U}_{ad}$ *and the norm topology on* Λ_2.

Proof. Let $\{u^n, u^o\} \in \mathcal{U}_{ad}$ be a sequence and suppose $u^n \xrightarrow{d} u^o$. Let $\{x^n, x^o\} \in \Lambda_2$, with $x^n(0) = x^o(0) = x_0$, denote the solutions of the integral equation (11) corresponding to the controls $\{u^n, u^o\}$ respectively and let $\{\mu^n, \mu^o\} \in C(I, \mathcal{M}_{2,\rho}(X))$ denote the corresponding measure valued functions with $\mu^n(0) = \mu^o(0) = \mathcal{P}(x_0)$. We show that $x^n \xrightarrow{s} x^o$ in Λ_2 and $\mu^n \xrightarrow{s} \mu^o$ in $C(I, \mathcal{M}_{2,\rho}(X))$. Using the integral equation (10) corresponding to controls $\{u^n\}$ and u^o respectively we have

$$\begin{aligned} x^n(t) &= S(t)x_0 + \int_0^t S(t-s) f(s, x^n(s), \mu^n(s), u^n_s) ds \\ &\quad + \int_0^t S(t-s)\sigma(s, x^n(s), \mu^n(s), u^n_s) dW(s), \ t \in I, \end{aligned} \tag{21}$$

$$\begin{aligned} x^o(t) &= S(t)x_0 + \int_0^t S(t-s) f(s, x^o(s), \mu^o(s), u^o_s) ds \\ &\quad + \int_0^t S(t-s)\sigma(s, x^o(s), \mu^o(s), u^o_s) dW(s), \ t \in I. \end{aligned} \tag{22}$$

Subtracting equation (22) from (21) and rearranging terms suitably we arrive at the following expression,

$$x^n(t) - x^o(t) = \int_0^t S(t-s)\big[f(s, x^n(s), \mu^n(s), u_s^n) - f(s, x^o(s), \mu^o(s), u_s^n)\big]ds$$

$$+ \int_0^t S(t-s)\big[\sigma(s, x^n(s), \mu^n(s), u_s^n) - \sigma(s, x^o(s), \mu^o(s), u_s^n)\big]dW(s) \tag{23}$$

$$+ e_1^n(t) + e_2^n(t),\ t \in I,$$

where the processes $\{e_1^n, e_2^n\}$ are given by

$$e_1^n(t) = \int_0^t S(t-s)\big[f(s, x^o(s), \mu^o(s), u_s^n) - f(s, x^o(s), \mu^o(s), u_s^o)\big]ds \tag{24}$$

$$e_2^n(t) = \int_0^t S(t-s)\big[\sigma(s, x^o(s), \mu^o(s), u_s^n) - \sigma(s, x^o(s), \mu^o(s), u_s^o)\big]dW(s). \tag{25}$$

Using the assumptions (A2)–(A3) and computing the expected value of the square of the X-norm, it follows from the expression (23) that

$$\begin{aligned} &\mathbf{E}|x^n(t) - x^o(t)|_X^2 \\ &\leq\ 2^3 M^2(K^2 t + K_R^2) \int_0^t \big\{\mathbf{E}|x^n(s) - x^o(s)|_X^2 + \rho^2(\mu^n(s), \mu^o(s))\big\}ds \\ &+\ 2^3\big(\mathbf{E}|e_1^n(t)|_X^2 + \mathbf{E}|e_2^n(t)|_X^2\big),\ t \in I. \end{aligned} \tag{26}$$

¿From the definition of the metric ρ, the reader can easily verify that

$$\rho^2(\mu^n(s), \mu^o(s)) \leq \mathbf{E}|x^n(s) - x^o(s)|_X^2,\ \forall\ s \in I. \tag{27}$$

Then using this inequality in (26) we obtain

$$\begin{aligned} \mathbf{E}|x^n(t) - x^o(t)|_X^2 \leq 2^4 M^2(K^2 t + K_R^2) \int_0^t \big\{\mathbf{E}|x^n(s) - x^o(s)|_X^2\big\}ds \\ + 2^3\big(\mathbf{E}|e_1^n(t)|_X^2 + \mathbf{E}|e_2^n(t)|_X^2\big),\ t \in I. \end{aligned} \tag{28}$$

For each $n \in N$, define the function η_n as follows

$$\eta_n(t) \equiv 2^3\big(\mathbf{E}|e_1^n(t)|_X^2 + \mathbf{E}|e_2^n(t)|_X^2\big),\ t \in I, \tag{29}$$

and a function C given by $C(t) = 2^4 M^2(K^2 t + K_R^2)$, $t \in I$. Then, by virtue of Gronwall inequality, it follows from the inequality (28) that

$$\mathbf{E}|x^n(t) - x^o(t)|_X^2 \leq \eta_n(t) + C(t) \int_0^t \exp\left\{\int_\theta^t C(s)ds\right\}\eta_n(\theta)d\theta. \tag{30}$$

Considering the processes e_1^n and e_2^n, it is easy to verify that, for each $t \in I$, we have

$$(31)\quad \mathbf{E}|e_1^n(t)|_X^2 \le M^2 t\mathbf{E}\int_0^t |f(s,x^o(s),\mu^o(s),u_s^n) - f(s,x^o(s),\mu^o(s),u_s^o)|_X^2 ds$$

$$(32)\quad \mathbf{E}|e_2^n(t)|_X^2 \le M^2 \mathbf{E}\int_0^t \parallel (\sigma(s,x^o(s),\mu^o(s),u_s^n) - \sigma(s,x^o(s),\mu^o(s),u_s^o)) \parallel_{\mathcal{R}}^2 ds$$

where we have used the notation $\parallel \sigma \parallel_{\mathcal{R}}^2 = Tr(\sigma\mathcal{R}\sigma^*)$. Using the elementary properties of conditional expectations and the fact that $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \ge 0$, it follows from the above inequalities that

$$\sup_{t\in I} \mathbf{E}|e_1^n(t)|_X^2 \le (M^2T)\mathbf{E}\int_0^T |f(s,x^o(s),\mu^o(s),u_s^n) - f(s,x^o(s),\mu^o(s),u_s^o)|_X^2 ds$$

$$(33)\qquad = (M^2T)\int_{I\times\Omega} \mathbf{E}\big\{|f(s,x^o(s),\mu^o(s),u_s^n) - f(s,x^o(s),\mu^o(s),u_s^o)|_X^2|\mathcal{G}_s\big\}\, d\mu$$

and

$$\sup_{t\in I} \mathbf{E}|e_2^n(t)|_X^2 \le M^2\mathbf{E}\int_0^T \parallel (\sigma(s,x^o(s),\mu^o(s),u_s^n) - \sigma(s,x^o(s),\mu^o(s),u_s^o)) \parallel_{\mathcal{R}}^2 ds$$

$$(34)\qquad = M^2\int_{I\times\Omega} \mathbf{E}\big\{\parallel (\sigma(s,x^o(s),\mu^o(s),u_s^n) - \sigma(s,x^o(s),\mu^o(s),u_s^o)) \parallel_{\mathcal{R}}^2 |\mathcal{G}_s\big\} d\mu$$

It follows from the assumptions (A1)–(A3), particularly the growth properties, that along the process $\{x^o, \mu^o\}$ the integrands in the expressions (33) and (34) belong to $L_1(\mu, C(U))$ and by Corollary 4.3 they are dominated by integrable functions (processes). Thus, by Lebesgue dominated convergence theorem, as $u^n \xrightarrow{d} u^o$, the integrals on the righthand side of the above expressions converge to zero. Consequently, it follows from (29) that the function $\eta_n(t) \to 0$ uniformly on I as $n \to \infty$. Using this fact in the inequality (30) we conclude that

$$(35)\qquad \lim_{n\to\infty} \sup\{\mathbf{E}|x^n(t) - x^o(t)|_X^2,\ t \in I\} = 0,$$

and hence $x^n \xrightarrow{s} x^o$ in Λ_2. This proves the continuity as stated in the theorem. ■

Remark 5.2. As a corollary of the above theorem, we observe that as $u^n \xrightarrow{d} u^o$, the probability measure valued process $\mu^n \xrightarrow{s} \mu^o$ in $C(I, \mathcal{M}_{2,\rho}(X))$. This follows readily from Theorem 5.1 and the inequality (27).

Now we are prepared to consider the question of existence of optimal control. The objective functional (cost) is given by

$$(36)\qquad J(u) \equiv \mathbf{E}\bigg\{\int_0^T \ell(t,x(t),\mu(t),u_t)dt + \Phi(x(T),\mu(T))\bigg\},$$

where x is the mild solution of the McKean-Vlasov evolution (1) corresponding to control $u \in \mathcal{U}_{ad}$. Our objective is to find a control that minimizes this functional.

Theorem 5.3. *Consider the system* (1) *with the cost functional* (36) *and admissible controls $\mathcal{U}_{ad} = V_d$. Suppose ℓ and Φ are Borel measurable in all the arguments, and lower semicontinuous in (x, μ) on $X \times \mathcal{M}_{2,\rho}(X)$ and continuous on U satisfying the following properties:*

(C1): *There exist an $\alpha_1 \in L_1^+(I)$ and a nonnegative number $\alpha_2 < \infty$ so that*

$$|\ell(t, x, \mu, \xi)| \leq \alpha_1(t) + \alpha_2\{1 + |x|_X^2 + |\mu|_{\mathcal{M}_{\gamma^2}}^2\} \;\; \forall \;\; (t, x, \mu, \xi) \in I \times X \times \mathcal{M}_{\gamma^2} \times U.$$

(C2): *There exists a nonnegative constant $\beta < \infty$ so that*

$$|\Phi(x, \mu)| \leq \beta\{1 + |x|_X^2 + |\mu|_{\mathcal{M}_{\gamma^2}}^2\} \;\; \forall \;\; (x, \mu) \in X \times \mathcal{M}_{\gamma^2}.$$

Then there exists an optimal control u^o minimizing the functional (36).

Proof. Since $\mathcal{U}_{ad}$ is compact in the metric topology d, it suffices to verify that the functional $u \longrightarrow J(u)$ is lower semicontinuous in this topology. Let $\{u^n\} \in \mathcal{U}_{ad}$ and suppose $u^n \xrightarrow{d} u^o$. Let $\{x^n, \mu^n\}$ denote the mild solutions of equation (1) corresponding to the sequence of controls $\{u^n\}$, and $\{x^o, \mu^o\}$ the solution corresponding to control u^o. Then it follows from Theorem 5.1 that, as $u^n \xrightarrow{d} u^o$, $x^n \xrightarrow{s} x^o$ in Λ_2 and the corresponding sequence of measures $\mu^n \xrightarrow{s} \mu^o$ in $C(I, \mathcal{M}_{2,\rho}(X))$. Further, recall that $\{x^n, x^o\}$ have continuous (modifications) versions. Thus, it follows from lower semicontinuity of ℓ and Φ that, along a subsequence if necessary,

$$\ell(t, x^o(t), \mu^o(t), u_t^o) \leq \underline{\lim}\, \ell(t, x^n(t), \mu^n(t), u_t^n) \tag{37}$$

$$\Phi(x^o(T), \mu^o(T)) \leq \underline{\lim}\, \Phi(x^n(T), \mu^n(T)) \tag{38}$$

μ a.e. in $I \times \Omega$. Since the norm topology of Λ_2 is stronger than the norm topology of $L_2^{\mathcal{F}}(I, X)$ it is easy to verify that the subsequence referred to above can be chosen independently of $(t, \omega) \in I \times \Omega$. Clearly, it follows from (37) and (38) that

$$\begin{aligned} &\mathbf{E}\left\{ \int_I \ell(t, x^o(t), \mu^o(t), u_t^o)dt + \Phi(x^o(T), \mu^o(T)) \right\} \\ &\leq \mathbf{E}\left\{ \int_0^T \underline{\lim}\, \ell(t, x^n(t), \mu^n(t), u_t^n)dt + \underline{\lim}\, \Phi(x^n(T), \mu^n(T)) \right\}. \end{aligned} \tag{39}$$

Since ℓ and Φ satisfy the assumptions (C1) and (C2), it follows from generalized Fatou's Lemma that

$$\begin{aligned}&\mathbf{E}\left\{\int_0^T \underline{\lim}\,\ell(t,x^n(t),\mu^n(t),u_t^n)dt+\underline{\lim}\,\Phi(x^n(T),\mu^n(T))\right\}\\ &\leq \underline{\lim}\,\mathbf{E}\left\{\int_0^T \ell(t,x^n(t),\mu^n(t),u_t^n)dt+\Phi(x^n(T),\mu^n(T))\right\}.\end{aligned} \tag{40}$$

Thus it follows from the definition of the cost functional J and the inequalities (39) and (40) that

$$J(u^o)\leq \underline{\lim}\,J(u^n) \tag{41}$$

proving that J is lower semicontinuous on $\mathcal{U}_{ad}$ in the metric topology d. Since $\mathcal{U}_{ad}$ is compact in this metric topology, J attains its minimum at some point $u^*\in\mathcal{U}_{ad}$. This completes the proof. ∎

Remark 5.4. The metric topology d on the space of admissible controls can be replaced by the natural weak star topology on $L^{\alpha}_{\infty}(\mu,\mathcal{M}_1(U))$ provided the semigroup generated by the unbounded operator A is compact [1].

6. Necessary conditions of optimality

Given that optimal control exists, we can proceed to develop the necessary conditions of optimality which can be used to determine the optimal policy. To develop the necessary conditions one requires more regularity properties for the drift and the diffusion operators including the cost integrands. For this reason we introduce the following additional assumptions:

(A4): The drift $f=f(t,x,\mu,u)$ and the diffusion operator $\sigma=\sigma(t,x,\mu,u)$ are Borel measurable in all the arguments and once continuously Fréchet differentiable in their second and third argument, and the Fréchet derivatives are uniformly bounded on $I\times X\times\mathcal{M}_{\gamma^2}(X)\times U$ and measurable in the uniform operator topology.

(A5): The cost integrands $\ell=\ell(t,x,\mu,u)$ and $\Phi=\Phi(x,\mu)$ are Borel measurable in all the variables and once continuously Gâteaux differentiable with respect to the arguments $x,\mu\in X\times\mathcal{M}_{\gamma^2}(X)$, and there exist constants $C_1,C_2>0$ so that their Gâteaux derivatives satisfy the following growth conditions:

$|\ell_x(t,x,\mu,\xi)|_X\leq C_1(1+|x|_X+|\mu|_{\mathcal{M}_{\gamma^2}(X)})\ \forall\ (t,x,\mu,\xi)\in I\times X\times\mathcal{M}_{\gamma^2}(X)\times U;$

$|\ell_\mu(t,x,\mu,u)|_{C_\rho(X)}\leq C_1(1+|x|_X+|\mu|_{\mathcal{M}_{\gamma^2}(X)})\ \forall\ (t,x,\mu,\xi)\in I\times X\times\mathcal{M}_{\gamma^2}(X)\times U$

$|\Phi_x(x,\mu)|_X\leq C_2(1+|x|_E+|\mu|_{\mathcal{M}_{\gamma^2}(X)})\ \forall\ (x,\mu)\in X\times\mathcal{M}_{\gamma^2}(X)$

$|\Phi_\mu(x,\mu)|_{C_\rho(X)}\leq C_2(1+|x|_X+|\mu|_{\mathcal{M}_{\gamma^2}(X)})\ \forall(x,\mu)\in X\times\mathcal{M}_{\gamma^2}(X).$

In the sequel we will be required to use the properties of semimartingales. Let $\mathcal{SM}_2^c(I,X)$ denote the space of continuous, norm-square integrable $\mathcal{F}_t$ semi-martingales with values in the Hilbert space X starting from zero. Every such semimartingale has the following integral representation. For each $M \in \mathcal{SM}_2^c(I,X) \subset L_2^{\mathcal{F}}(I,X)$, there exists a unique pair of intensity $(\phi,Q) \in L_2^{\mathcal{F}}(I,X)\times L_2^{\mathcal{F}}(I,\mathcal{L}_{\mathcal{R}}(H,X))$ such that

$$M_t = \int_0^t \phi(s)ds + \int_0^t Q(s)dW(s),\ t \in I.$$

For $M^1, M^2 \in \mathcal{SM}_2^c(I,X)$, with the intensities (ϕ_1,Q_1) and (ϕ_2,Q_2) respectively, one introduces the scalar product

$$(M_1,M_2)_{\mathcal{SM}_2^c(I,X)} \equiv \mathbf{E}\left\{\int_0^T (\phi_1(s),\phi_2(s))_X ds + \int_0^T Tr(Q_1(s)\mathcal{R}Q_2^*(s))ds\right\}.$$

Completion of $\mathcal{SM}_2^c(I,X)$ with respect to the above inner product turns it into a Hilbert space which we continue to denote by the same symbol. The associated norm is given by $\|\cdot\|_{\mathcal{SM}_2^c(I,X)}$ where

$$\| M \|^2_{\mathcal{SM}_2^c(I,X)} = \mathbf{E}\left\{\int_0^T |\phi(s)|_X^2 ds + \int_0^T Tr(Q(s)\mathcal{R}Q^*(s))ds\right\}.$$

Now we return to the control problem. To develop the necessary conditions of optimality we need the so-called variational equation. This equation characterizes the Gâteaux differential of the solution of the state equation (1) with respect to controls $u \in \mathcal{U}_{ad}$. We present this in the following lemma.

Lemma 6.1. *Suppose the assumptions* (A1)–(A4) *including those of Theorem* 5.3 *hold, and let* $\{x^o,\mu^o,u^o\}$ *be the optimal state-control process with* $\mu^o(t) = \mathcal{P}(x^o(t))$, $t \in I$. *Then, for any* $u \in \mathcal{U}_{ad}$, *there exists a unique pair* $(z,\nu) \in \Lambda_2 \times C(I,\mathcal{M}^s_{\gamma^2}(X))$ *which is the mild solution of the following variational equation*

$$\begin{aligned} dz &= Azdt + f_x(t,x^o(t),\mu^o(t),u_t^o)zdt + f_\mu(t,x^o(t),\mu^o(t),u_t^o)\nu(t)dt \\ &+ \sigma_x(t,x^o(t),\mu^o(t),u_t^o;z(t))dW(t) + \sigma_\mu(t,x^o(t),\mu^o(t),u_t^o;\nu(t))dW(t) \\ &+ d\mathbf{\Lambda}^{u-u^o}(t), \quad z(0)=0, \quad t \in I, \end{aligned} \tag{42}$$

where $\mathbf{\Lambda} \in \mathcal{SM}_2^c(I,X)$ *is the semi-martingale given by*

$$d\mathbf{\Lambda}^{u-u^o}(t) = f(t,x^o(t),\mu^o(t),u_t-u_t^o)dt + \sigma(t,x^o(t),\mu^o(t),u_t-u_t^o)dW(t), \tag{43}$$

starting from $\mathbf{\Lambda}^{u-u^o}(0)=0$. *The solution* $\{z,\nu\}$ *is the strong limit of* $(1/\varepsilon)(x^\varepsilon - x^o)$ *and* $(1/\varepsilon)(\mu^\varepsilon - \mu^o)$ *in* Λ_2 *and* $C(I,\mathcal{M}^s_{\gamma^2}(X))$ *respectively where* $\{x^\varepsilon, x^o\} \in \Lambda_2$

and $\{\mu^\varepsilon, \mu^o\} \in C(I, \mathcal{M}_{2,\rho}(X))$ *are the solutions of the integral equation* (10) *corresponding to the controls* $\{u^\varepsilon, u^o\} \in \mathcal{U}_{ad}$ *respectively. Further, for any fixed* $u \in \mathcal{U}_{ad}$, $\mathbf{\Lambda}^{u-u^o} \longrightarrow z$ *is a continuous linear map from* $\mathcal{SM}_2^c(I,X)$ *to* Λ_2.

Proof. Let $u^o \in \mathcal{U}_{ad}$ denote the optimal control and $u \in \mathcal{U}_{ad}$ any other control. Since the admissible set $\mathcal{U}_{ad}$ consists of relaxed controls, it is evident that $u^\varepsilon \equiv u^o + \varepsilon(u - u^o) \in \mathcal{U}_{ad}$ for all $\varepsilon \in (0,1)$. Let $\{x^\varepsilon, \mu^\varepsilon\}$ and $\{x^o, \mu^o\}$ denote the solutions of the integral equation (10) corresponding to the controls $\{u^\varepsilon, u^o\}$ respectively. Define

$$z^\varepsilon(t) \equiv (1/\varepsilon)(x^\varepsilon(t) - x^o(t)), \quad \nu^\varepsilon(t) \equiv (1/\varepsilon)(\mu^\varepsilon(t) - \mu^o(t)), \; t \in I.$$

Considering the integral equation (10) corresponding to the controls u^ε and u^o respectively and subtracting one from the other and dividing by ε, we obtain

$$\begin{aligned} z^\varepsilon(t) &= \int_0^t S(t-s)(1/\varepsilon)[f(s, x^\varepsilon(s), \mu^\varepsilon(s), u^o) - f(s, x^o(s), \mu^o(s), u_s^o)]ds \\ &+ \int_0^t S(t-s)(1/\varepsilon)[\sigma(s, x^\varepsilon(s), \mu^\varepsilon(s), u^o) - \sigma(s, x^o(s), \mu^o(s), u_s^o)]dW(s) \\ &+ \int_0^t S(t-s) f(s, x^\varepsilon(s), \mu^\varepsilon(s), u_s - u_s^o)ds \\ &+ \int_0^t S(t-s)\sigma(s, x^\varepsilon(s), \mu^\varepsilon(s), u_s - u_s^o)dW(s), \; t \in I. \end{aligned} \tag{44}$$

Using the Lipschitz properties (A2) and (A3) and computing the expected value of the norm square it follows from (44) that

$$\begin{aligned} \mathbf{E}|z^\varepsilon(t)|_X^2 &\leq 16M^2(K^2 t + K_R^2)\int_0^t \mathbf{E}|z^\varepsilon(s)|_X^2 ds \\ &+ 16M^2(K^2 t + K_R^2)\int_0^t (1/\varepsilon^2)\rho^2(\mu^\varepsilon(s), \mu^o(s))ds \\ &+ \mathbf{E}|V_\varepsilon^{u-u^o}(t)|_X^2, \; t \in I, \end{aligned} \tag{45}$$

where the process $V_\varepsilon^{u-u^o}$ is given by the convolution integral of the semimartingale $\mathbf{\Lambda}_\varepsilon^{u-u^o} \in \mathcal{SM}_2^c(I,X)$ with respect to the semigroup $S(t), t \geq 0$, that is,

$$\begin{aligned} V_\varepsilon^{u-u^o}(t) &= \int_0^t S(t-s)d\mathbf{\Lambda}_\varepsilon^{u-u^o}(s) = \int_0^t S(t-s) f(s, x^\varepsilon(s), \mu^\varepsilon(s), u_s - u_s^o)ds \\ &+ \int_0^t S(t-s)\sigma(s, x^\varepsilon(s), \mu^\varepsilon(s), u_s - u_s^o)dW(s), \; t \in I. \end{aligned} \tag{46}$$

Clearly, the semimartingale $\mathbf{\Lambda}_\varepsilon^{u-u^o}$ is given by the Itô differential,

$$d\Lambda_\varepsilon^{u-u^o}(t) = f(t, x^\varepsilon(t), \mu^\varepsilon(t), u_t - u_t^o)dt + \sigma(t, x^\varepsilon(t), \mu^\varepsilon(t), u_t - u_t^o)dW(t), \; t \in I,$$

with $\mathbf{\Lambda}_\varepsilon^{u-u^o}(0) = 0$. Using the definition of the metric ρ the reader can easily verify that

$$(1/\varepsilon^2)\rho^2(\mu^\varepsilon(t), \mu^o(t)) \leq \mathbf{E}|z^\varepsilon(t)|_X^2,\ t \in I. \tag{47}$$

Considering the process $V_\varepsilon^{u-u^o}$, it follows from the growth properties in the assumptions (A2)–(A3), that

$$\sup_{t\in I} \mathbf{E}|V_\varepsilon^{u-u^o}(t)|_X^2 \leq 2M^2(K^2T+K_R^2)\int_0^T \{1+\mathbf{E}|x^\varepsilon(s)|_X^2+|\mu^\varepsilon(s)|^2_{\mathcal{M}_{\gamma_2}(X)}\}ds \tag{48}$$

for all $\varepsilon \in (0,1)$. By virtue of Corollary 4.3, it follows from the above inequality that there exists a finite positive number $\delta(T)$, dependent on T, such that

$$\sup\{\mathbf{E}|V_\varepsilon^{u-u^o}(t)|_X^2, t \in I, \varepsilon \in (0,1)\} \leq\ \delta(T),\ \forall\ u \in \mathcal{U}_{ad}. \tag{49}$$

Defining $\beta(T) \equiv 32M^2(K^2T + K_R^2)$ and using the estimates (47) and (49) in the inequality (45) we arrive at the following inequality,

$$\mathbf{E}|z^\varepsilon(t)|_X^2 \leq \beta(T)\int_0^t \mathbf{E}|z^\varepsilon(s)|_X^2 ds + \delta(T). \tag{50}$$

Hence it follows from Gronwall inequality that

$$\sup_{0\leq\varepsilon\leq 1}\ \sup_{t\in I} \mathbf{E}|z^\varepsilon(t)|_X^2 \leq \delta(T)\exp\{T\beta(T)\} < \infty,$$

and consequently it follows from (47) that

$$\sup\{|\nu^\varepsilon(t)|^2_{\mathcal{M}_{\gamma^2}}, t \in I, 0 \leq \varepsilon \leq 1\} \leq \delta(T)\exp\{T\beta(T)\} < \infty.$$

In view of the above analysis, it is clear that the following limits are well defined

$$\begin{aligned}
&z \equiv s - \lim_{\varepsilon\downarrow 0} z^\varepsilon = s - \lim_{\varepsilon\downarrow 0}(1/\varepsilon)(x^\varepsilon - x^o)\ \text{ in } \Lambda_2\\
&\nu \equiv s - \lim_{\varepsilon\downarrow 0} \nu^\varepsilon = s - \lim_{\varepsilon\downarrow 0}(1/\varepsilon)(\mu^\varepsilon - \mu^o)\ \text{ in } C(I, \mathcal{M}^s_{\gamma^2}(X))\\
&\Lambda^{u-u^o} \equiv s - \lim_{\varepsilon\downarrow 0} \Lambda_\varepsilon^{u-u^o}\ \text{ in } \mathcal{SM}_2^c(I, X).
\end{aligned}$$

Thus, letting $\varepsilon \downarrow 0$ in the identities (44) and (46), it follows from assumption (A4) including ((A2), (A3)) and dominated convergence theorem and strong convergence of semimartingales with the strong convergence of their intensities, that

$$
\begin{aligned}
z(t) = & \int_0^t S(t-s) f_x(s, x^o(s), \mu^o(s), u_s^o) z(s) ds \\
& + \int_0^t S(t-s) f_\mu(s, x^o(s), \mu^o(s), u_s^o) \nu(s) ds \\
& + \int_0^t S(t-s) \sigma_x(s, x^o(s), \mu^o(s), u_s^o; z(s)) dW(s) \\
& + \int_0^t S(t-s) \sigma_\mu(s, x^o(s), \mu^o(s), u_s^o; \nu(s)) dW(s) + V_t^{u-u^o}, \ t \in I,
\end{aligned} \tag{51}
$$

where

$$
\begin{aligned}
V_t^{u-u^o} & = \int_0^t S(t-s) f(s, x^o(s), \mu^o(s), u_s - u_s^o) ds \\
& \quad + \int_0^t S(t-s) \sigma(s, x^o(s), \mu^o(s), u_s - u_s^o) dW(s) \\
& = \int_0^t S(t-s) d\Lambda^{u-u^o}(s), \ t \in I.
\end{aligned} \tag{52}
$$

It follows from our assumptions (A2) and (A3) that $\{\mathbf{\Lambda}^{u-u^o}(t), t \in I\}$ is an X-valued (norm) square integrable continuous semi-martingale. For simplicity of notation, we introduce the following abbreviations:

$$
\begin{aligned}
& F_1(t) \equiv f_x(t, x^o(t), \mu^o(t), u_t^o), F_2(t) \equiv f_\mu(t, x^o(t), \mu^o(t), u_t^o), \\
& \Sigma_1(t; z(t)) \equiv \sigma_x(t, x^o(t), \mu^o(t), u_t^o; z(t)), \Sigma_2(t; \nu(t)) \equiv \sigma_\mu(t, x^o(t), \mu^o(t), u_t^o; \nu(t)).
\end{aligned}
$$

These operators are all evaluated along the optimal path $\{x^o, \mu^o, u^o\}$. It follows from assumption (A4) that they are all bounded in operator norm uniformly on I and therefore, for all $t \in I$, $F_1(t) \in \mathcal{L}(X)$, $F_2(t) \in \mathcal{L}(M_{\gamma^2}^s(X), X)$, $\Sigma_1(t; \cdot) \in \mathcal{L}(X, \mathcal{L}_{\mathcal{R}}(H, X))$ and $\Sigma_2(t; \cdot) \in \mathcal{L}(M_{\gamma^2}^s(X), \mathcal{L}_{\mathcal{R}}(H, X))$ respectively. Using these notations in the integral equation (51) we obtain

$$
\begin{aligned}
z(t) = & \int_0^t S(t-s) F_1(s) z(s) ds + \int_0^t S(t-s) F_2(s) \nu(s) ds \\
& + \int_0^t S(t-s) \Sigma_1(s; z(s)) dW(s) + \int_0^t S(t-s) \Sigma_2(s; \nu(s)) dW(s) \\
& + \int_0^t S(t-s) d\mathbf{\Lambda}^{u-u^o}(s), \ t \in I.
\end{aligned} \tag{53}
$$

Clearly, this is the integral equation corresponding to the following stochastic evolution equation,

$$
\begin{aligned}
dz &= Azdt + F_1(t)zdt + F_2(t)\nu(t)dt + \Sigma_1(t; z(t))dW(t) \\
&\quad + \Sigma_2(t; \nu(t))dW(t) + d\mathbf{\Lambda}_t^{u-u^o}, \quad z(0) = 0, \quad t \in I,
\end{aligned} \tag{54}
$$

which is the compact form of equation (42). We note that $\nu(t)$ is not the probability law of $z(t)$ and $\nu(t)(X) = 0$ for all $t \in I$. However, for each $t \in I$, the signed measure $\nu(t)$ is linearly related to $z(t)$. Indeed, for any $\Psi \in C_\rho(X)$,

$$< \Psi, \nu(t) >_{C_\rho(X), M^s_{\gamma^2}(X)} = \mathbf{E} < D_x\Psi(x^o(t)), z(t) >_X .$$

Thus equation (54) is in fact a linear integral equation in z driven by the semimartingale $\Lambda^{u-u^o} \in \mathcal{SM}_2^c(I, X)$. It follows from assumption (A4) that, along the process $\{x^o, \mu^o, u^o\}$, the operator valued functions $\{F_1, F_2\}$ and $\{\Sigma_1, \Sigma_2\}$ are all uniformly bounded. Thus it follows, as a special case of Theorem 4.2, that the integral equation (54) has a unique solution. In other words, the variational equation (54), equivalently (42), driven by the semimartingale $\mathbf{\Lambda}^{u-u^o}(t)$, $t \in I$, has a unique mild solution $z \in \Lambda_2$ with $\nu \in C(I, \mathcal{M}^s_{\gamma^2}(X))$. For the second part of the Lemma, first note that $|\nu(t)|^2_{\mathcal{M}^s_{2,\rho}(X)} \leq \mathbf{E}|z(t)|^2_X$ for all $t \in I$. For convenience of notation, define

$$
\begin{aligned}
f_o(t) &\equiv f_o^{u-u^o}(t) \equiv f(t, x^o(t), \mu^o(t), u_t - u_t^o), \text{ and} \\
\sigma_o(t) &\equiv \sigma_o^{u-u^o}(t) \equiv \sigma(t, x^o(t), \mu^o(t), u_t - u_t^o), \; t \in I.
\end{aligned}
$$

Then it follows from the expression (54) and the above inequality and the uniform boundedness of the operators $\{F_1, F_2, \Sigma_1, \Sigma_2\}$ that there exist constants $a_1, a_2 \geq 0$, dependent on the bounds of the operators $\{F_1, F_2, \Sigma_1, \Sigma_2\}$ and $\{M, T\}$, such that

$$\mathbf{E}|z(t)|^2_X \leq a_1 \int_0^t \mathbf{E}|z(s)|^2_X ds + a_2 \parallel \Lambda^{u-u^o} \parallel^2_{\mathcal{SM}_2^c(I,X))}, \; t \in I, \tag{55}$$

where

$$\parallel \Lambda^{u-u^o} \parallel^2_{\mathcal{SM}_2^c(I,X)} \equiv \mathbf{E} \int_0^T |f_o^{u-u^o}|^2_X ds + \mathbf{E} \int_0^T Tr\big(\sigma_o^{u-u^o} \mathcal{R} (\sigma_o^{u-u^o})^*\big) ds.$$

By virtue of Gronwall inequality, it follows from (55) that

$$\parallel z \parallel^2_{\Lambda_2} \equiv \sup\{\mathbf{E}|z(t)|^2_X, t \in I\} \leq (a_2 \exp a_1 T) \parallel \Lambda^{u-u^o} \parallel^2_{\mathcal{SM}_2^c(I,X)} . \tag{56}$$

Clearly, it follows from this inequality and the linearity of equation (54) that, the map $\Lambda^{u-u^o} \longrightarrow z$ is a continuous linear operator from $\mathcal{SM}_2^c(I, X)$ to Λ_2. This completes the proof. ■

Now we are prepared to prove the necessary conditions of optimality.

Theorem 6.2. *Consider the system* (1) *with the admissible controls* $\mathcal{U}_{ad}$ *and the cost functional* (36), *and suppose the assumption* (A5) *and those of Lemma* 6.1 *hold. Then, in order that* $u^o \in \mathcal{U}_{ad}$ *be optimal, with* $\{x^o, \mu^o\}$ *being the corresponding mild solutions of the evolution equation* (1), *it is necessary that there exists a pair* $(\psi, Q) \in \Lambda_2 \times L_2^{\mathcal{F}}(I, \mathcal{L}_{\mathcal{R}}(H, X))$ *such that the following inequality holds,*

$$\begin{aligned} &\mathbf{E}\left\{\int_I \ell(t, x^o(t), \mu^o(t), u_t - u_t^o)dt + (\psi(t), f(t, x^o(t), \mu^o(t), u_t - u_t^o))_X dt \right. \\ &\left. + \int_I Tr[Q(t)\mathcal{R}\sigma^*(t, x^o(t), \mu^o(t), u_t - u_t^o)]dt\right\} \geq 0, \quad \forall\ u \in \mathcal{U}_{ad}. \end{aligned} \tag{57}$$

Proof. Let $u^o \in \mathcal{U}_{ad}$ be the optimal control with the pair $\{x^o, \mu^o\} \in \Lambda_2 \times C(I, \mathcal{M}_{2,\rho}(X))$ being the corresponding unique mild solution of the evolution equation (1). Let $u \in \mathcal{U}_{ad}$ and define $u^\varepsilon \equiv u^o + \varepsilon(u - u^o)$ for $\varepsilon \in (0,1)$. Clearly, by convexity of $\mathcal{U}_{ad}$, $u^\varepsilon \in \mathcal{U}_{ad}$ and by optimality of u^o, $J(u^\varepsilon) \geq J(u^o)$ for all $\varepsilon \in (0,1)$ and $u \in \mathcal{U}_{ad}$. Hence the Gâteaux differential of J at u^o, in the direction $(u - u^o)$, denoted by $dJ(u^o; u - u^o)$ satisfies

$$dJ(u^o, u - u^o) \geq 0 \ \ \forall \ \ u \in \mathcal{U}_{ad}. \tag{58}$$

Again, for simplicity of notations, we denote

$$\begin{aligned} &\ell_1^o(t) \equiv \ell_x(t, x^o(t), \mu^o(t), u_t^o), \quad \ell_2^o(t) \equiv \ell_\mu(t, x^o(t), \mu^o(t), u_t^o) \\ &\Phi_1^o(T) \equiv \Phi_x(x^o(t), \mu^o(T)), \quad \Phi_2^o(T) \equiv \Phi_\mu(x^o(t), \mu^o(T)). \end{aligned}$$

Under the assumption (A5), it follows from Corollary 4.3 that $\ell_2^o(t) \in C_\rho(X) \subset C_{\gamma^2}(X)$ P-a.s for all $t \in I$, and $\Phi_2^o(T) \in C_\rho(X)$ P-a.s. Hence the Gâteaux differential of J, denoted by dJ, is well defined and it is given by

$$dJ(u^o, u - u^o) = L(z) + \mathbf{E}\int_0^T \ell(t, x^o(t), \mu^o(t), u_t - u_t^o)dt. \tag{59}$$

where

$$\begin{aligned} L(z) \equiv \mathbf{E}\bigg\{\int_I \{&< \ell_1^o(t), z(t) >_X + < \ell_2^o(t), \nu(t) >_{C_{\gamma^2}(X), M^s_{\gamma^2}(X)}\}dt \\ &+ < \Phi_1^o(T), z(T) >_X + < \Phi_2^o(T), \nu(T) >_{C_{\gamma^2}(X), M^s_{\gamma^2}(X)}\bigg\}. \end{aligned} \tag{60}$$

Since, for each $t \in I$, $\ell_2^o(t) \in C_\rho(X)$ P-a.s, and $\Phi_2^o(T) \in C_\rho(X)$ P-a.s, using Lagrange formula and Fubini's theorem it is easy to verify that the above expression is equivalent to the following one

$$L(z) = \mathbf{E}\left\{ \int_I \{< \ell_1^o(t) + D_x\ell_2^o(t), z(t) >_X\}dt + < \Phi_1^o(T) + D_x\Phi_2^o(T), z(T) >_X \right\}. \tag{61}$$

By virtue of the growth properties **(C1)**–**(C2)** of ℓ and Φ, and the growth property **(A5)** of their Gâteaux differentials, we have $\ell_1^o + D_x\ell_2^o \in L_2^{\mathcal{F}}(I, X)$ and $\Phi_1^o(T) + D_x\Phi_2^o(T) \in L_2^{\mathcal{F}_T}(\Omega, X)$. Thus it follows from (61) that $z \longrightarrow L(z)$ is a continuous linear functional on Λ_2. On the other hand, it follows from Lemma 6.1 that $\Lambda^{u-u^o} \longrightarrow z$ is a continuous linear operator from the Hilbert space $\mathcal{SM}_2^c(I, X)$ to the Banach space Λ_2. Hence the composition map

$$\Lambda^{u-u^o} \longrightarrow z \longrightarrow L(z) \equiv \tilde{L}(\Lambda^{u-u^o})$$

is a continuous linear functional on $\mathcal{SM}_2^c(I, X)$. Thus, it follows from semimartingale representation theorem that there exists a unique pair $(\psi, Q) \in \Lambda_2 \times L_2^{\mathcal{F}}(I, \mathcal{L}_{\mathcal{R}}(H, X))$ such that

$$L(z) = \tilde{L}(\Lambda^{u-u^o}) = \mathbf{E}\left\{ \int_0^T (\psi(t), f(t, x^o(t), \mu^o(t), u_t - u_t^o))_X + \int_0^T Tr[Q(t)\mathcal{R}\sigma^*(t, x^o(t), \mu^o(t), u_t - u_t^o)]dt \right\}. \tag{62}$$

The necessary condition (57) then follows from (58), (59) and (62). This completes the proof. ■

We note that Theorem 6.2 asserts the existence of the pair

$$(\psi, Q) \in \Lambda_2 \times L_2^{\mathcal{F}}(I, \mathcal{L}_{\mathcal{R}}(H, X))$$

as the necessary condition for optimality. However, it does not say how one can construct such a pair. Here we present a constructive procedure. The operators appearing in the following theorem are defined in the body of its proof.

Theorem 6.3. *Suppose the assumptions of Theorem* 6.2 *hold and further* σ *is uniformly bounded on* $I \times X \times M_1(X) \times U$. *Then the pair* (ψ, Q) *is given by the* $\mathcal{F}_t$*-adapted mild solution of the following adjoint evolution equation* (*backward stochastic evolution equation*),

$$-d\varphi = A^*\varphi dt + F_1^*(t)\varphi dt + B_1^*(t)\varphi dt + \Upsilon(t)\varphi dt + [\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)]dW + (\ell_1^o(t) + D_x\ell_2^o(t))dt \tag{63}$$

satisfying the terminal condition,

$$\varphi(T) = \varphi^o(T) \equiv \Phi_1^o(T) + D_x\Phi_2^o(T) \tag{64}$$

giving $\psi(t) = \varphi(t)$ *and* $Q(t) = \Sigma_o(t) \equiv \tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)$ *for* $t \in I$; *where the operators are identified in the body of the proof. Further, the backward stochastic evolution equation* (63) *with the terminal condition* (64) *has a unique mild solution* $\varphi \in \Lambda_2 \subset L_2^{\mathcal{F}}(I, X)$.

Proof. We prove that the necessary condition given by Theorem 6.2 leads to a backward stochastic evolution equation of the form

$$\begin{aligned} &d\varphi = -A^*\varphi dt + (BV \text{ terms})dt + \Sigma_o(t)dW \\ &\varphi(T) = \varphi^o(T), \ t \in I. \end{aligned} \tag{65}$$

Using the notations presented above the inequality (55), we rewrite the variational equation (54) as follows:

$$\begin{aligned} dz &= Azdt + F_1(t)zdt + F_2(t)\nu(t)dt + f_o(t)dt \\ &\quad + \Sigma_1(t;z(t))dW(t) + \Sigma_2(t;\nu(t))dW(t) + \sigma_o(t)dW, \\ z(0) &= 0, \ t \in I. \end{aligned} \tag{66}$$

Keeping in mind that we are only interested in mild solutions, we can formally compute the Itô differential of the scalar product $(\varphi(t), z(t))_X$. This can be justified rigorously using Yosida approximation A_n of A and taking limits using the fact that the corresponding semigroups $S_n(t), t \geq 0$, converge in the strong operator topology to $S(t), t \geq 0$, uniformly on compact intervals [8, Theorem 4.5.4, p. 133]. Thus we have

$$d(\varphi, z) = (d\varphi, z) + (\varphi, dz) + << d\varphi, dz >> \tag{67}$$

where $<< \cdot, \cdot >>$ denotes the quadratic variation. Integrating this over the interval $[0, T]$ we have

$$\mathbf{E}(\varphi(T), z(T)) = \mathbf{E}\int_0^T (d\varphi, z) + \mathbf{E}\int_0^T (\varphi, dz) + \mathbf{E}\int_0^T << d\varphi, dz >> . \tag{68}$$

Considering the second term on the righthand side and using equation (66) and necessary adjoint operations, we have

$$\begin{aligned} &\mathbf{E}\int_0^T (\varphi, dz) = \mathbf{E}\int_0^T (A^*\varphi dt + F_1^*(t)\varphi dt, z) \\ &+ \mathbf{E}\int_0^T (\varphi, \Sigma_1(t;z)dW) + \mathbf{E}\int_0^T (\varphi, F_2(t)\nu)dt + \mathbf{E}\int_0^T (\varphi, \Sigma_2(t;\nu)dW) \\ &+ \mathbf{E}\int_0^T (\varphi, f_0)dt + \mathbf{E}\int_0^T (\varphi, \sigma_0(t)dW). \end{aligned} \tag{69}$$

By virtue of assumption (A4), F_2 and Σ_2 are uniformly bounded (and uniformly measurable) linear operator valued functions with values $F_2(t) \in \mathcal{L}(M^s_{\gamma^2}(X), X)$ and $\Sigma_2(t;\cdot) \in \mathcal{L}(M^s_{\gamma^2}(X), \mathcal{L}_{\mathcal{R}}(H,X))$ respectively. Recall that the signed measure ν is linearly related to the process z. In particular, we have (with a slight abuse of notation)

$$
\begin{aligned}
F_2(t)\nu(t) &= \int_X F_2(t)(\xi)\nu(t)(d\xi) \\
&= \lim_{\varepsilon\downarrow 0}(1/\varepsilon)\left\{\int_X F_2(t)(\xi)(\mu^\varepsilon(t)(d\xi) - \mu^o(t)(d\xi))\right\} \\
&= \mathbf{E}\lim_{\varepsilon\downarrow 0}(1/\varepsilon)[F_2(t)(x^\varepsilon(t)) - F_2(t)(x^o(t))] \\
&= \mathbf{E}D_xF_2(t)(x^o(t);z(t)) \equiv \mathbf{E}(D_xF_2(t;z(t))).
\end{aligned}
\tag{70}
$$

The reader can easily verify this using Lagrange formula, Fubini's theorem and dominated convergence theorem. Similarly, we have $\Sigma_2(t;\nu(t)) = \mathbf{E}D_x\Sigma_2(t, x^o(t); z(t)) \equiv \mathbf{E}D_x\Sigma_2(t;z(t))$; and both are linear in z. Thus the middle three terms of equation (69), denoted by $M_3(69)$, can be written as

$$
\begin{aligned}
M_3(69) = \mathbf{E}\int_0^T(\varphi, \Sigma_1(t;z)dW) + \mathbf{E}\int_0^T(\varphi, \mathbf{E}(D_xF_2(t;z(t)))dt \\
+ \mathbf{E}\int_0^T(\varphi, \mathbf{E}(D_x\Sigma_2(t;z(t)))dW).
\end{aligned}
\tag{71}
$$

Define the following multi linear forms:

$$
\begin{aligned}
&b_0(t,z,\varphi,h) \equiv (\varphi, \Sigma_1(t;z)h) \\
&b_1(t,\varphi,z) \equiv (\varphi, \mathbf{E}(D_xF_2(t;z(t))) \\
&b_2(t,\varphi,z,h) \equiv (\varphi, \mathbf{E}(D_x\Sigma_2(t;z(t)))h), h \in \mathcal{H}_t.
\end{aligned}
\tag{72}
$$

Since the operators $\{F_1, F_2, \Sigma_1, \Sigma_2\}$ are bounded and linear, these forms are also bounded. For each $t \in I$, let $\mathcal{X}_t$ ($\mathcal{H}_t$) denote the space of all $\mathcal{F}_t$-measurable X valued (H-valued) norm-square integrable random variables. Clearly, for each $t \in I$, b_0 is a trilinear form on $\mathcal{X}_t \times \mathcal{X}_t \times \mathcal{H}_t$, and therefore by Riesz representation theorem on Hilbert spaces there exists a unique operator valued function $\tilde{\Sigma}_1$ such that b_0 has an equivalent representation given by

$$
b_0(t,z,\varphi,h) \equiv (\varphi, \Sigma_1(t;z)h)_{\mathcal{X}_t} = (z, \tilde{\Sigma}_1(t;\varphi)h)_{\mathcal{X}_t}.
$$

Similarly, b_1 is a bilinear form on $\mathcal{X}_t \times \mathcal{X}_t$ and therefore, again by Riesz representation theorem, there exists a unique $\mathcal{F}_t$-adapted bounded linear operator valued function $B_1(t) \in \mathcal{L}(\mathcal{X}_t)$ such that $b_1(t,\varphi,z) = (\varphi, B_1(t)z)_{\mathcal{X}_t}$. Again by the same representation theorem, the trilinear form b_2 admits two equivalent representations

$$b_2(t,\varphi,z,h) = (\varphi, B_2(t;z)h)_{\mathcal{X}_t} = (z, \tilde{B}_2(t;\varphi)h)_{\mathcal{X}_t}$$

where $B_2(t;\cdot) \in \mathcal{L}(\mathcal{X}_t, \mathcal{L}(\mathcal{H}_t, \mathcal{X}_t))$ and $\tilde{B}_2(t;\cdot) \in \mathcal{L}(\mathcal{X}_t, \mathcal{L}(\mathcal{H}_t, \mathcal{X}_t))$. Based on these representations, we conclude that the expression (71) is equivalent to the following one,

$$\begin{aligned} M_3(69) = \mathbf{E}\int_0^T (z(t), \tilde{\Sigma}_1(t;\varphi)dW) + \mathbf{E}\int_0^T (z(t), B_1^*(t)\varphi(t))dt \\ + \mathbf{E}\int_0^T (z(t), \tilde{B}_2(t;\varphi)dW). \end{aligned} \tag{73}$$

Thus the first two terms on the right hand side of equation (68) can be rewritten as follows:

$$\begin{aligned} &\mathbf{E}\int_0^T (d\varphi, z) + \mathbf{E}\int_0^T (\varphi, dz) \\ &= \mathbf{E}\int_0^T (d\varphi + A^*\varphi dt + F_1^*(t)\varphi dt + B_1^*(t)\varphi dt + [\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)]dW, z) \\ &+ \mathbf{E}\int_0^T (\varphi, f_o)dt + (\varphi, \sigma_o(t)dW). \end{aligned} \tag{74}$$

At this point we observe that the diffusion operator Σ_o of equation (65) can be identified as follows,

$$\Sigma_o(t) \equiv -(\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)),$$

which is linear in φ. Now considering the quadratic variation term $\mathbf{E}\int_0^T < d\varphi, dz >$ and recalling the representation of the trilinear form b_2 we have

$$\begin{aligned} &\mathbf{E}\int_0^T < d\varphi, dz >= \mathbf{E}\int_0^T Tr\big(\Sigma_o(t)\mathcal{R}(\Sigma_1(t;z) + \mathbf{E}D_x\Sigma_2(t;z) + \sigma_0(t))^*\big)dt \\ &= \mathbf{E}\int_0^T Tr\big(-[\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)]\mathcal{R}[\Sigma_1(t;z) + B_2(t;z)]^*\big)dt \\ &+ \mathbf{E}\int_0^T Tr(\Sigma_o(t)\mathcal{R}\sigma_o^*(t))dt. \end{aligned} \tag{75}$$

Considering the second line in the above expression, it follows from linearity and boundedness of the operators $\{\tilde{\Sigma}_1, \tilde{B}_2, \Sigma_1, B_2\}$ that the integrand is a (bounded) bilinear form on $\mathcal{X}_t$ uniformly in $t \in I$, and therefore there exists an uniformly bounded operator valued function Υ, with values $\Upsilon(t) \in \mathcal{L}(\mathcal{X}_t)$, such that

$$Tr\big(-[\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)]\mathcal{R}[\Sigma_1(t;z) + B_2(t;z)]^*\big) = (\Upsilon(t)\varphi(t), z(t)). \tag{76}$$

Using (74)–(75) and (76) in the expression (68) we obtain

$$
\begin{aligned}
\mathbf{E}(\varphi(T), z(T)) = \mathbf{E}\Big\{ \int_0^T (d\varphi + A^*\varphi dt + F_1^*(t)\varphi dt + B_1^*(t)\varphi dt + \Upsilon(t)\varphi dt \\
+ [\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t;\varphi)]dW, z)\Big\} \qquad (77) \\
+ \mathbf{E}\left\{ \int_0^T Tr(\Sigma_o(t)\mathcal{R}\sigma_o^*(t))dt + \int_0^T [(\varphi, f_o)dt + (\varphi, \sigma_o(t)dW]\right\}.
\end{aligned}
$$

Now requiring that φ is a mild solution of the following backward stochastic evolution equation,

$$
\begin{aligned}
-d\varphi = A^*\varphi dt + F_1^*(t)\varphi dt + B_1^*(t)\varphi dt + \Upsilon(t)\varphi dt + [\tilde{\Sigma}_1(t;\varphi) \\
+ \tilde{B}_2(t;\varphi)]dW + (\ell_1^o(t) + D_x\ell_2^o(t))dt,
\end{aligned} \tag{78}
$$

with the terminal condition, $\varphi(T) \equiv \Phi_1^o(T) + D_x\Phi_2^o(T)$, it follows from the identity (77) that

$$
\begin{aligned}
\mathbf{E}\int_0^T < \ell_1^o(t) + D_x\ell_2^o(t), z(t) > dt + \mathbf{E}\ \zeta(\Phi_1^o(T) + D_x\Phi_2^o(T), z(T)) \\
= \mathbf{E}\int_0^T Tr(\Sigma_o(t)\mathcal{R}\sigma_o^*(t))dt + \mathbf{E}\int_0^T (\varphi, f_o)dt + (\varphi, \sigma_o dW).
\end{aligned} \tag{79}
$$

Since, by our assumption, σ is uniformly bounded, it is clear that $\mathbf{E}\int_0^T(\sigma_o^*\varphi, dW) = 0$. Thus it follows from the above expression that

$$
\begin{aligned}
\mathbf{E}\int_0^T < \ell_1^o(t) + D_x\ell_2^o(t), z(t) > dt + \mathbf{E}(\Phi_1^o(T) + D_x\Phi_2^o(T), z(T)) \\
= \mathbf{E}\int_0^T Tr(\Sigma_o(t)\mathcal{R}\sigma_o^*(t))dt + \mathbf{E}\int_0^T (\varphi, f_o)dt.
\end{aligned} \tag{80}
$$

The expression on the left hand side of the above identity coincides with the functional $L(z)$ given by the expression (60). Thus, if we identify $\Sigma_o(t)$ with $Q(t)$ and $\varphi(t)$ with $\psi(t)$ then the expression (80) coincides with the identity (62). Since, by assumption (A4), the Fréchet derivatives of f and σ are uniformly bounded, it is evident that the operator valued function Σ_o, given by $\Sigma_o(t) = -(\tilde{\Sigma}_1(t;\varphi) + \tilde{B}_2(t,\varphi))$, belongs to $L_2^{\mathcal{F}}(I, \mathcal{L}_{\mathcal{R}}(H,X))$ satisfying the necessary condition of Theorem 6.2. Thus the pair (ψ, Q), whose existence was guaranteed by the semimartingale representation theorem (see Theorem 6.2), can be actually computed by solving the backward stochastic evolution equation (63)–(64) provided this equation has an $\mathcal{F}_t$-adapted solution. Under the given assumptions, in particular (A4)–(A5), the operators $\{F_1^*(t), B_1^*(t), \Upsilon(t)\}$ and $[\tilde{\Sigma}_1(t;\cdot) + \tilde{B}_2(t;\cdot)]$

are all $\mathcal{F}_t$-adapted and uniformly bounded. Considering the system (78), the process η given by $\eta(t) \equiv (\ell_1^o(t) + D_x\ell_2^o(t))$ is an element of $L_2^{\mathcal{F}}(I, X)$. Thus, it follows from a result of Hu and Peng, [16, Theorem 3.1] on BSDE that the system (63)–(64) has a unique $\mathcal{F}_t$-adapted mild solution $\varphi \in L_2^{\mathcal{F}}(I, X)$. This completes the proof. ■

Under certain convexity assumptions, it is easy to verify that the necessary conditions are also sufficient. This is stated in the following corollary.

Corollary 6.4. *Suppose the assumptions of Theorem* 6.3 *hold and consider the real valued functional* $\mathcal{H}$ *given by*

$$\begin{aligned}\mathcal{H}(t, x, \mu, \psi, Q, u) &\equiv \ell(t, x, \mu, u) + < \psi, f(t, x, \mu, u) >_H \\ &+ Tr(Q\mathcal{R}\sigma^*(t, x, \mu, u))\end{aligned} \tag{81}$$

defined on $I \times X \times \mathcal{M}_{\gamma^2}(X) \times X \times \mathcal{L}_{\mathcal{R}}(H, X) \times \mathcal{M}_1(U)$. *Suppose* $\mathcal{H}$ *is Borel measurable in all the variables and, for each* $t \in I$, *it is continuous on* $X \times \mathcal{M}_{\gamma^2}(X) \times X \times \mathcal{L}_{\mathcal{R}}(H, X) \times U$ *and convex in the second and third argument. The terminal cost functional* Φ *is Borel measurable and convex in all its arguments. Then the necessary conditions of optimality given by Theorem 6.2 are also sufficient.*

Proof. Since the proof involves straightforward computation we present only a brief outline. Compute the Itô differential of the scalar product $(\psi(t), x(t) - x^o(t))_X$ where ψ is the mild solution of the BSDE (63)–(64) and $\{x^o, x\}$ are the mild solutions of the state equation (1) corresponding to controls $\{u^o, u\}$ respectively with $u \in \mathcal{U}_{ad}$ being arbitrary. Then taking the expected value of the integral over the interval $I \equiv [0, T]$ and using the necessary condition (inequality) (57) and the convexity assumptions one arrives at the inequality $J(u) - J(u^o) \geq 0$. This ends the outline. For rigorous justification of the steps one must use the Yosida approximation A_n of A, and use the Itô differential of the scalar product $(\psi_n, x_n - x_n^o)$ instead of $(\psi, x - x^o)$ where now ψ_n is the strong solutions of the BSDE (63)–(64) with A^* replaced by A_n^* and $\varphi^o(T)$ replaced by $\varphi_n^o(T) \equiv J_n(\varphi^o(T))$ with $J_n \equiv nR(n, A)$ where $R(n, A)$ denotes the resolvent of A corresponding to $n \in \rho(A)$. Similarly $\{x_n^o, x_n\}$ are the strong solutions of the state equation (1), corresponding to controls $\{u^o, u\}$ respectively, with A replaced by A_n and initial condition x_0 replaced by $J_n x_0$. By use of a well known result from semigroup theory [8, Theorem 4.5.4], one can easily verify that as $n \to \infty$, $(\psi_n, x_n - x_n^o) \longrightarrow (\psi, x - x^o)$. This completes the outline of our proof. ■

Remark 6.5. Theorems 6.2 and 6.3 provide the necessary conditions of optimality whereby one can compute the optimal controls.

Acknowledgement

This work was partially supported by the National Science and Engineering Research Council of Canada under grant no. A7101.

References

[1] N.U. Ahmed, *Systems governed by mean-field stochastic evolution equations on Hilbert spaces and their optimal control*, Nonlinear Anal. (submitted).

[2] N.U. Ahmed, *Nonlinear diffusion governed by McKean-Vlasov equation on Hilbert space and optimal control*, SIAM J. Control and Optim. **46** (1) (2007), 356–378. doi:10.1137/050645944

[3] N.U. Ahmed and X. Ding, *A semilinear McKean-Vlasov stochastic evolution equation in Hilbert space*, Stoch. Proc. Appl. **60** (1995), 65–85. doi:10.1016/0304-4149(95)00050-X

[4] N.U. Ahmed and X. Ding, *Controlled McKean-Vlasov equations*, Commun. Appl. Anal. **5** (2001), 183–206.

[5] N.U. Ahmed, C.D. Charalambous, *Stochastic minimum principle for partially observed systems subject to continuous and jump diffusion processes and drviven by relaxed controls*, SIAM J. Control and Optim. **51** (4) (2013), 3235–3257. doi:10.1137/120885656

[6] N.U. Ahmed, *Stochastic neutral evolution equations on Hilbert spaces with partially observed relaxed controls and their necessary conditions of optimality*, Nonlinear Anal. (A) Theory, Methods & Applications **101** (2014), 66–79. doi:10.1016/j.na.2014.01.019

[7] N.U. Ahmed, *Stochastic neutral evolution equations on Hilbert spaces and their partially observed optimal relaxed control*, J. Abstract Diff. Equ. Appl. **5** (1) (2014), 1–20.

[8] N.U. Ahmed, Semigroup Theory with Applications to Systems and Control, Pitman Research Notes in Mathematics Series, Vol. 246, Longman Scientific and Technical, U.K. (Co-published with John-Wiley & Sons, Inc. New York, 1991).

[9] N.U. Ahmed, *Stochastic evolution equations on Hilbert spaces with partially observed relaxed controls and their necessary conditions of optimality*, Discuss. Math. Diff. Incl., Control and Optim. **34** (2014), 105–129. doi:10.7151/dmdico.1153

[10] N.U. Ahmed, *Stochastic initial boundary value problems subject to distributed and boundary noise and their optimal control*, JMAA **421** (2015), 157–179. doi:10.1016/j.jmaa.2014.06.078

[11] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions (Cambridge University Press, 1992). doi:10.1017/CBO9780511666223

[12] G. Da Prato and J. Zabczyk, Ergodicity for Infinite Dimensional Systems, London Math. Soc. Lecture Note Ser. 229 (Cambridge University Press, London, 1996). doi:10.1017/CBO9780511662829

[13] D.A. Dawson, *Critical dynamics and fluctuations for a mean-field model of cooperative behavior*, J. Stat. Phys. **31** (1983), 29–85. doi:10.1007/BF01010922

[14] D.A. Dawson and J. Gartner, Large Deviations, Free Energy Functional and Quasi-Potential for a Mean Field Model of Interacting Diffusions, Mem. Amer. Math. Soc. 398 (Providence, RI, 1989).

[15] N. Dunford and J.T. Schwartz, Linear Operators, Part 1 (Interscience Publishers, Inc., New York, 1958).

[16] Y. Hu and S. Peng, *Adaptive solution of a backward semilinear stochastic evolution equation*, Stoch. Anal. Appl. **9** (4) (1991), 445–459. doi:10.1080/07362999108809250

[17] N.I. Mahmudov and M.A. McKibben, *Abstract second order damped McKean-Vlasov stochastic evolution equations*, Stoch. Anal. Appl. **24** (2006), 303–328. doi:10.1080/07362990500522247

[18] H.P. McKean, *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Natl. Acad. Sci. USA **56** (1966), 1907–1911. doi:10.1073/pnas.56.6.1907

[19] Yang Shen and Tak Kuen Siu, *The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem*, Nonlin. Anal. **86** (2013), 58–73. doi:10.1016/j.na.2013.02.029

Received 3 November 2015