Dynamic Systems and Applications 25 (2016) 61-88

SYSTEMS GOVERNED BY MEAN-FIELD STOCHASTIC EVOLUTION EQUATIONS ON HILBERT SPACES AND THEIR OPTIMAL CONTROL

N. U. AHMED

Department of EECS, University of Ottawa,
Ottawa, Ontario, CANADA

ABSTRACT. In this paper we consider a general class of controlled McKean-Vlasov mean-field stochastic evolution equations on Hilbert spaces. We prove existence, uniqueness and regularity properties of mild solutions of these equations. Relaxed controls, covering regular controls, adapted to a current of sub-sigma algebras generated by observable processes and taking values from a Polish space, are used. An appropriate metric topology, based on weak star convergence, is introduced. We prove continuous dependence of solutions on controls with respect to this topology. These results are then used to prove existence of optimal controls for Bolza problem. Then we develop the necessary conditions of optimality using semi-martingale representation theory and show that the adjoint processes arising from the necessary conditions can be constructed from the mild solution of certain backward stochastic mean field evolution equation (BSMEE). The paper is concluded with some applications to mean-field linear quadratic regulator problems.

AMS (MOS) Subject Classification: 49J27, 60H15,93E20

Key Words: McKean-Vlasov mean-field Stochastic Differential Equation, Hilbert Spaces, Relaxed Controls, Existence of Optimal Controls, Necessary conditions of optimality.

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 its mean. This class of systems without control have been studied extensively in the literature [12, 16, 17, 22, 23, 26] after McKean introduced this model in [23]. Since then control problems involving this general model have been studied in [8, 9, 10] with control appearing only in the drift. Here in this paper we consider more general system model that admits mean-field and control in both the drift and diffusion operators.

Received January 28, 2016 1056-2176 $15.00 ©Dynamic Publishers, Inc.

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, 6, 7, 11, 13, 18, 20, 21, 27]. See also the extensive references given therein. Control of McKean-Vlasov type stochastic differential equations were studied in [8, 9, 10, 22, 28]. However, necessary conditions of optimality were not considered in these papers. In [20] 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 class of mean-field evolution equations on infinite dimensional Hilbert spaces.

In [2, 3, 7, 11] we considered semi-linear stochastic evolution equations (including equations of neutral type) with controls in the drift and the diffusion operators and presented necessary conditions of optimality. In [18] Duncan and Pasic-Duncan considered linear stochastic differential equations on Hilbert spaces with exponential-quadratic cost functionals giving differential operator Ricatti equations. They presented also several interesting examples from initial boundary value problems. Shen and Siu [20] presented maximum principle for a class of finite dimensional jump-diffusion stochastic differential equations. The cost functional is of Bolza type. In [21] Hu and Peng developed some fundamental results on the question of existence and uniqueness of solutions for a class of backward stochastic evolution equations (BSDE) on Hilbert spaces. In [27] Zhou developed necessary conditions of optimality (maximum principle) for a class of linear non-degenerate (strictly elliptic) second order partial differential equations on a d-dimensional space with all the coefficients containing control. In [8] we considered control of McKean-Vlasov equations and presented existence of optimal controls where control appears only in the drift. In [10] we considered McKean-Vlasov equations on finite dimensional spaces and developed HJB equations. The author is not aware of any literature where stochastic necessary conditions of optimality for McKean-Vlasov type or mean-field stochastic evolution equations on infinite dimensional Hilbert spaces with control in the drift as well as diffusion have been considered. This is what motivates us to consider optimal control of mean-field stochastic evolution equations on infinite dimensional spaces and develop necessary conditions of optimality thereof. We present existence of optimal controls and also necessary conditions of optimality. These are new results (including the existence theorem 3.1) considered as major contributions of this paper. For recent development on this topic involving more general McKean-Vlasov equation see [28].

The paper is organized as follows. In section 2, we present the mathematical model of the system and state the optimal control problems. In section 3, after basic assumptions are introduced, we prove the existence and regularity properties of mild solutions. Existence of optimal control is proved in section 4. In section 5, we present the necessary conditions of optimality. For illustration of the abstract results, section 6 is devoted to some examples of linear quadratic regulator problems of meanfield type.

2. **SYSTEM MODEL**

Let E 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 $\mathcal{R}$ being nuclear.

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 finite 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, and $\mathcal{G}_t \subset \mathcal{F}_t$, $t \geq 0$, denote a current of nondecreasing family of sub-sigma algebras of the sigma-algebra $\mathcal{F}_t$. We use $L^{\alpha}_{\infty}(I, \mathcal{M}_1(U))$ to denote the class of weak star measurable $\mathcal{G}_t$-adapted $\mathcal{M}_1(U)$ valued random processes. That is, for each $u \in L^{\alpha}_{\infty}(I, \mathcal{M}_1(U))$, and for every $\varphi \in C(U)$, the process $t \longrightarrow u_t(\varphi)$ is a measurable random process adapted to the current of sigma algebras $\mathcal{G}_t$.

Throughout the rest of the paper, the space of bounded linear operators $\mathcal{L}(E_1, E_2)$ from Banach space E_1 to a Banach space E_2 will be assumed to be equipped with the uniform operator topology τ_{uo}, unless otherwise stated. And functions defined on I, with values in $\mathcal{L}(E_1, E_2)$, are assumed to be measurable in the uniform operator topology (equivalently Bochner measurable) and the norm is Lebesgue integrable. By Bochner measurability of $T : I \longrightarrow \mathcal{L}(E_1, E_2)$ one means that there exists a sequence of simple functions $\{T_n\}$ with values in $\mathcal{L}(E_1, E_2)$ such that $T_n(t) \xrightarrow{\tau_{uo}} T(t)$ a.e on I.

Now we are prepared to introduce the system considered in this paper. It is governed by the following mean-field controlled evolution equation on the Hilbert space E driven by the H-Brownian motion W:

$$dx = Axdt + \overline{f}(t, x, \overline{x}, u)dt + \overline{\sigma}(t, x, \overline{x}, u)dW, \quad t \in I \equiv [0, T],\ x(0) = x_0, \tag{2.1}$$

where A is the infinitesimal generator of a C_0-semigroup $S(t)$, $t \in I$, on E and $\overline{f}$ is a Borel measurable map from $I \times E \times E \times \mathcal{M}_1(U)$ to E and $\overline{\sigma}$ is also a Borel measurable map from $I \times E \times E \times \mathcal{M}_1(U)$ to $\mathcal{L}(H, E)$, the space of bounded linear operators from H to E equipped with the uniform operator topology, and x_0 is the initial state. The drift $\overline{f}$ and the diffusion $\overline{\sigma}$ are not only dependent on the state x but also its mean $\overline{x} \equiv \int_E z\mu(dz)$ where μ is the measure induced by the E-valued random variable x.

Here, by the notation $\overline{f}(t,x,\bar{x},u)$ and $\overline{\sigma}(t,x,\bar{x},u)$, we mean

$$\overline{f}(t,x,\overline{x},u) \equiv \int_U f(t,x,\overline{x},\xi)u(d\xi), \overline{\sigma}(t,x,\overline{x},u) \equiv \int_U \sigma(t,x,\overline{x},\xi)u(d\xi)$$

for any $u \in \mathcal{M}_1(U)$ where f and σ are Borel measurable maps from $I \times E \times E \times U$ to E and $\mathcal{L}(H,E)$ respectively. In case both E 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. BASIC ASSUMPTIONS AND EXISTENCE OF SOLUTIONS

Now we are prepared to introduce the basic assumptions. In order to study control problems involving the system (2.1) we must now define 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 E satisfying

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

(A2): The function $f : I \times E \times E \times U \longrightarrow E$ 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

$$|f(t,x,y,\xi)|_E^2 \leq K^2\{1 + |x|_E^2 + |y|_E^2\}, \quad \forall\, x,y \in E$$

$$|f(t,x_1,y_1,\xi) - f(t,x_2,y_2,\xi)|_E^2 \leq K^2\{|x_1 - x_2|_E^2 + |y_1 - y_2|_E^2\}, \quad \forall\, x_1,x_2,y_1,y_2 \in E,$$

uniformly with respect to $t \in I$, $\xi \in U$.

(A3): The incremental covariance of the Brownian motion W denoted by $\mathcal{R} \in \mathcal{L}(H)$ (is positive nuclear). The diffusion $\sigma : I \times E \times E \times U \longrightarrow \mathcal{L}(H,E)$ is measurable (in the uniform operator topology) 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

$$|\sigma(t,x,y,\xi)|_{\mathcal{R}}^2 \leq K_{\mathcal{R}}^2\{1 + |x|_E^2 + |y|_E^2\}, \quad \forall\, x,y \in E$$

$$|\sigma(t,x_1,y_1,\xi) - \sigma(t,x_2,y_2,\xi)|_{\mathcal{R}}^2 \leq K_{\mathcal{R}}^2\{|x_1 - x_2|_E^2 + |y_1 - y_2|_E^2\} \quad \forall\, x_1,x_2,y_1,y_2 \in E$$

uniformly with respect to $(t,\xi) \in I \times U$, where $|\sigma|_{\mathcal{R}}^2 = tr(\sigma\mathcal{R}\sigma^*)$.

Let $\mathcal{L}_{\mathcal{R}}(H,E)$ denote the completion of the space of linear operators from H to E with respect to the inner product $\langle K,L\rangle \equiv Tr(K\mathcal{R}L^*)$ and norm $|K|_{\mathcal{R}} \equiv \sqrt{Tr(K\mathcal{R}K^*)}$. Clearly this is a Hilbert space.

For proof of existence, uniqueness and regularity properties of solutions of equation (2.1), we must introduce the appropriate function spaces where they may reside.

Let $L_2(\Omega, E)$ denote the Hilbert space of second order random variables with values in E equipped with the standard norm topology. For $1 \leq p \leq \infty$, let $L_p^{\mathcal{F}}(I, L_2(\Omega, E))$ denote the Banach space of $\mathcal{F}_t$-adapted second order random processes with values in E furnished with the norm topology

$$\| y \|_{L_p^{\mathcal{F}}(I,L_2(\Omega,E))} \equiv \left(\int_I \left(\int_\Omega |y(t,\omega)|_E^2 P(d\omega) \right)^{p/2} dt \right)^{1/p}$$

for $y \in L_p^{\mathcal{F}}(I, L_2(\Omega, E))$. For $p = \infty$, as usual, the norm is given by

$$\| y \|_{L_\infty^{\mathcal{F}}(I,L_2(\Omega,E))} \equiv ess - \sup\{(E|y(t)|_E^2)^{1/2}, t \in I\}.$$

For convenience of notation we shall denote these spaces by $L_p^a(I, L_2(\Omega, E))$, $1 \leq p \leq \infty$.

Let $B_\infty^a(I, E)$ denote the vector space of E valued $\mathcal{F}_t$-adapted random processes having square integrable norms (with respect to the measure P) which are bounded on I. Furnished with the norm topology,

$$\| x \|_{B_\infty^a(I,E)} \equiv (\sup\{\mathbf{E}|x(t)|_E^2, t \in I\})^{1/2},$$

$B_\infty^a(I, E)$ is a closed subspace of the Banach space $L_\infty^a(I, L_2(\Omega, E))$ and hence it is a Banach space. For admissible controls, let $\mathcal{G}_t, t \geq 0$, denote a nondecreasing family of sub-sigma 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. In general, for admissible controls we may choose any closed subset $\mathcal{U}_{ad} \subseteq L_\infty^\alpha(I, \mathcal{M}_1(U)) \subset L_\infty^\alpha(I, \mathcal{M}(U))$ which consist of $\mathcal{G}_t$-adapted $\mathcal{M}_1(U)$-valued random processes, endowed with the weak star topology. More precise characterization of admissible controls is given in section 4. With this preparation we prove the following existence result.

Theorem 3.1. Consider the system (2.1) and suppose the assumptions **(A1)–(A3)** hold. Further, suppose that W is the H Brownian motion with incremental covariance operator $\mathcal{R} \in \mathcal{L}_1^+(H)$. Then, for every $\mathcal{F}_0$ measurable E valued random variable $x_0 \in L_2(\Omega, E)$, and control $u \in \mathcal{U}_{ad}$, the stochastic evolution equation has a unique mild solution $x \in B_\infty^a(I, E)$ in the sense that it satisfies the following stochastic integral equation:

$$x(t) \equiv S(t)x_0 + \int_0^t S(t-\tau)\overline{f}(\tau, x(\tau), \bar{x}(\tau), u_\tau)d\tau + \int_0^t S(t-\tau)\overline{\sigma}(\tau, x(\tau), \bar{x}(\tau), u_\tau)dW(\tau) \quad t \in I. \tag{3.1}$$

Further the solution has a continuous modification.

Proof. Take any $z \in B_\infty^a(I, E)$ and consider the evolution equation

$$dx = Axdt + \overline{f}(t, x, \bar{z}(t), u_t)dt + \overline{\sigma}(t, x, \bar{z}(t), u_t)dW(t), \quad x(0) = x_0, \quad t \in I, \tag{3.2}$$

where $\bar{z}(t) \equiv \mathbf{E}z(t)$. First we show that this equation has a unique mild solution $x \in B^a_\infty(I,E)$ corresponding to the given $z \in B^a_\infty(I,E)$. Let Ψ denote the map $z \longrightarrow x \equiv \Psi(z)$. It suffices to prove that Ψ has a unique fixed point in the Banach space $B^a_\infty(I,E)$. First we must show that equation (3.2) has a unique mild solution for the given z. Clearly this solution is given by the solution of the integral equation

$$x(t) = S(t)x_0 + \int_0^t S(t-r)\overline{f}(r,x(r),\bar{z}(r),u_r)dr + \int_0^t S(t-r)\overline{\sigma}(r,x(r),\bar{z}(r),u_r)dW(r), \quad t \in I. \tag{3.3}$$

For the fixed but arbitrary initial state x_0 and control u, define the integral operator F_z by

$$(F_z x)(t) \equiv S(t)x_0 + \int_0^t S(t-r)\overline{f}(r,x(r),\bar{z}(r),u_r)dr + \int_0^t S(t-r)\overline{\sigma}(r,x(r),\bar{z}(r),u_r)dW(r), \quad t \in I. \tag{3.4}$$

Thus the question of existence of a solution of the integral equation (3.3) is equivalent to the question of existence of a fixed point of the operator F_z, that is an $x \in B^a_\infty(I,E)$ so that $x = F_z(x)$. Since $W(t)$ is $\mathcal{F}_t$-adapted and $x(t), z(t)$ are $\mathcal{F}_t$ -adapted and u_t is $\mathcal{G}_t(\subset \mathcal{F}_t)$-adapted in the weak star sense, we conclude that $(F_z x)(t)$ is $\mathcal{F}_t$-adapted. Then, by use of the growth assumptions in (A1) and (A2), it is easy to verify that $F_z : B^a_\infty(I,E) \longrightarrow B^a_\infty(I,E)$. Now we verify that F_z has a unique fixed point in $B^a_\infty(I,E)$. Let $x_1, x_2 \in B^a_\infty(I,E)$ and compute

$$\begin{aligned} &F_z(x_1)(t) - F_z(x_2)(t) \\ &= \int_0^t S(t-r)\big(\overline{f}(r,x_1(r),\bar{z}(r),u_r) - \overline{f}(r,x_2(r),\bar{z}(r),u_r)\big)dr \\ &\quad + \int_0^t S(t-r)\big(\overline{\sigma}(r,x_1(r),\bar{z}(r),u_r) - \overline{\sigma}(r,x_2(r),\bar{z}(r),u_r)\big)dW(r). \end{aligned} \tag{3.5}$$

By straightforward computation using the Lipschitz assumptions (A2) and (A3), we arrive at the following inequality,

$$\mathbf{E}|F_z(x_1)(t) - F_z(x_2)(t)|^2_E \leq 2tM^2K^2 \int_0^t \mathbf{E}|x_1(r) - x_2(r)|^2_E dr + 2M^2K^2_{\mathcal{R}} \int_0^t \mathbf{E}|x_1(r) - x_2(r)|^2_E dr \quad t \in I. \tag{3.6}$$

For any $\tau \in I$, let $I_\tau \equiv [0,\tau]$. Clearly, it follows from the above inequality that

$$\sup\{\mathbf{E}|F_z(x_1)(t) - F_z(x_2)(t)|^2_E, t \in I_\tau\} \leq \alpha(\tau)\sup\{\mathbf{E}|x_1(s) - x_2(s)|^2_E, s \in I_\tau\}, \tag{3.7}$$

where $\alpha(\tau) \equiv 2M^2\tau(K^2\tau + K_{\mathcal{R}}^2)$. Hence

$$\| F_z(x_1) - F_z(x_2) \|_{B_\infty^a(I_\tau,E)} \leq \sqrt{\alpha(\tau)} \| x_1 - x_2 \|_{B_\infty^a(I_\tau,E)} . \tag{3.8}$$

Thus for $\tau = \tau_1 \in I$, sufficiently small, $\alpha(\tau_1) < 1$ and therefore the operator F_z is a contraction on the Banach space $B_\infty^a(I_{\tau_1}, E)$ and hence by Banach fixed point theorem it has a unique fixed point, say $x^1 \in B_\infty^a(I_{\tau_1}, E)$. Then it follows from the well known factorization technique due to Da Prato and Zabczyk [15] that it has a continuous modification which is also denoted by x^1. Clearly $x^1(\tau_1)$ is $\mathcal{F}_{\tau_1}$ measurable and belongs to $L_2(\Omega, E)$. Taking this as the initial state for the evolution equation

$$\begin{aligned} dx &= Axdt + \overline{f}(t, x, \bar{z}(t), u_t)dt + \overline{\sigma}(t, x, \bar{z}(t), u_t)dW(t), \\ x(\tau_1) &= x^1(\tau_1), \quad t \in [\tau_1, T], \end{aligned} \tag{3.9}$$

the associated integral equation is given by

$$\begin{aligned} x(t) = S(t-\tau_1)x^1(\tau_1) &+ \int_{\tau_1}^t S(t-r)\overline{f}(r, x(r), \bar{z}(r), u_r)dr \\ &+ \int_{\tau_1}^t S(t-r)\overline{\sigma}(r, x(r), \bar{z}(r), u_r)dW(r), \quad t \in [\tau_1, T]. \end{aligned} \tag{3.10}$$

Thus, following the same procedure as given above one can verify that there exists a $\tau_2 \in (\tau_1, T]$ such that $\alpha(\tau_2 - \tau_1) < 1$ and the operator F_z (restricted to the interval $[\tau_1, \tau_2]$) has a unique fixed point $x^2 \in B_\infty^a([\tau_1, \tau_2], E)$ satisfying $x^2(\tau_1) = x^1(\tau_1)$. Since I is a compact interval it can be covered by a finite number of such intervals. Hence the mild solution of the evolution equation (3.2) can be constructed by concatenation of this finite sequence of solutions $\{x^1, x^2, \dots\}$ as shown above. Thus we conclude that for every given $z \in B_\infty^a(I, E)$ the evolution equation (3.2) has a unique mild solution $x \in B_\infty^a(I, E)$. Let Ψ denote the map $z \longrightarrow x$ from $B_\infty^a(I, E)$ to $B_\infty^a(I, E)$ written as $x = \Psi(z)$. It is clear that if Ψ has a fixed point $x^* \in B_\infty^a(I, E)$ then x^* itself is the mild solution of the evolution equation (2.1). Thus it suffices to prove that Ψ has a unique fixed point in the Banach space $B_\infty^a(I, E)$. For the fixed initial state x_0 and control u, let $x_1 \equiv \Psi(z_1) \in B_\infty^a(I, E)$ be the mild solution of the evolution equation (3.2) corresponding to $z_1 \in B_\infty^a(I, E)$ and $x_2 \equiv \Psi(z_2) \in B_\infty^a(I, E)$ the mild solution corresponding to $z_2 \in B_\infty^a(I, E)$. Then

(3.11)

$$\begin{aligned} x_1(t) - x_2(t) = &\int_0^t S(t-r)[\overline{f}(r, x_1(r), \bar{z}_1(r), u_r) - \overline{f}(r, x_2(r), \bar{z}_2(r), u_r)]dr \\ &+ \int_0^t S(t-r)[\overline{\sigma}(r, x_1(r), \bar{z}_1(r), u_r) - \overline{\sigma}(r, x_2(r), \bar{z}_2(r), u_r)]dW(r), \quad t \in I. \end{aligned}$$

Recall that $|\bar{z}_1(t) - \bar{z}_2(t)|_E^2 \leq \mathbf{E}|z_1(t) - z_2(t)|_E^2$ for $t \in I$. Hence, using the equation (3.11) and the assumptions (A1)–(A3) and repeating the same procedure, it is easy to verify that

$$\text{(3.12)}\quad \begin{aligned} &\mathbf{E}|x_1(t)-x_2(t)|_E^2 \\ &\leq 2M^2(tK^2+K_{\mathcal{R}}^2)\left\{\int_0^t \mathbf{E}|x_1(r)-x_2(r)|_E^2 dr + \int_0^t \mathbf{E}|z_1(r)-z_2(r)|_E^2 dr\right\}. \end{aligned}$$

Thus for any $\tau \in I$, we have

$$\text{(3.13)}\quad \sup_{t\in I_\tau}\mathbf{E}|x_1(t)-x_2(t)|_E^2 \leq \alpha(\tau)\left\{\sup_{t\in I_\tau}\mathbf{E}|x_1(t)-x_2(t)|_E^2 + \sup_{t\in I_\tau}\mathbf{E}|z_1(t)-z_2(t)|_E^2\right\}.$$

In other words,

$$\text{(3.14)}\qquad \parallel x_1 - x_2 \parallel_{B_\infty^a(I_\tau,E)}^2 \leq \alpha(\tau)\left\{\parallel x_1 - x_2 \parallel_{B_\infty^a(I_\tau,E)}^2 + \parallel z_1 - z_2 \parallel_{B_\infty^a(I_\tau,E)}^2\right\}.$$

Since $\alpha(\tau)$ is a continuous and increasing function of its argument starting from $\alpha(0) = 0$, we can choose a $\tau_1 \in I$ such that $\alpha(\tau_1) \leq (1/3)$. For this choice, it is evident that

$$\text{(3.15)}\qquad \parallel x_1 - x_2 \parallel_{B_\infty^a(I_{\tau_1},E)} \leq \sqrt{(1/2)} \parallel z_1 - z_2 \parallel_{B_\infty^a(I_{\tau_1},E)}$$

and hence we have proved that

$$\text{(3.16)}\qquad \parallel \Psi(z_1) - \Psi(z_2) \parallel_{B_\infty^a(I_{\tau_1},E)} \leq \sqrt{(1/2)} \parallel z_1 - z_2 \parallel_{B_\infty^a(I_{\tau_1},E)} .$$

Thus the map Ψ is a contraction on the Banach space $B_\infty^a(I_{\tau_1}, E)$ and hence, again by Banach fixed point theorem, it has a unique fixed point say $x^* \in B_\infty^a(I_{\tau_1}, E)$. Since I is a compact interval, it can be covered by a finite number of such intervals and therefore by concatenation we can construct the solution for the entire interval I. Thus we conclude that Ψ has a unique fixed point $x \in B_\infty^a(I, E)$. This proves that for any given $\mathcal{F}_0$ measurable initial state $x_0 \in L_2(\Omega, E)$ and any $u \in L_\infty^\alpha(I, \mathcal{M}_1(U))$ the evolution equation (2.1) has a unique mild solution having continuous modification. This completes the proof. □

For any fixed $\mathcal{F}_0$ measurable random variable $x_0 \in L_2(\Omega, E)$, let $x(u) \in B_\infty^a(I, E)$ denote the solution of the integral equation (3.1) corresponding to the control $u \in \mathcal{U}_{ad}$. Then we have the following result as a corollary of Theorem 3.1.

Corollary 3.2. Suppose the assumptions of Theorem 3.1 hold with the admissible controls $\mathcal{U}_{ad}$. Then the solution set $\Xi \equiv \{x(u), u \in \mathcal{U}_{ad}\}$ is a bounded subset of $B_\infty^a(I, E)$.

Proof. We present a brief outline. Let $x(u) \in B_\infty^a(I, E)$ denote the solution of the integral equation (3.1) corresponding to the control $u \in \mathcal{U}_{ad}$. For any fixed $x, y \in E$, it follows from the first part of the assumptions (A2)–(A3) that both f and σ are uniformly bounded with respect to controls. Hence, using the integral equation (3.1), it is easy to verify that for every $u \in \mathcal{U}_{ad}$ we have

$$\mathbf{E}|x(u)(t)|_E^2 \leq b^2 + c^2 \int_0^t \mathbf{E}|x(u)(s)|_E^2 ds, t \in I, \tag{3.17}$$

where

$$b^2 \equiv 3M^2 \mathbf{E}|x_0|_E^2 + 3M^2 T(TK^2 + K_{\mathcal{R}}^2) \text{ and } c^2 \equiv 6M^2(TK^2 + K_{\mathcal{R}}^2).$$

Since the constants b, c are independent of control, the conclusion follows from Gronwall inequality applied to the expression (3.17). This completes the proof. □

Remark 3.3. In the proof of existence we have used the quadratic function $\alpha(\Delta\tau) \equiv a(\Delta\tau)^2 + b\Delta\tau$, with $a = 2M^2K^2$ and $b = 2M^2K_{\mathcal{R}}^2$. For the first part of the proof, based on contraction principle, we required that $0 < \alpha(\Delta\tau) < 1$. This puts an upper bound on the length of the subintervals used in the partition. Since I is a compact interval, it can be covered by a finite number of such intervals of positive length. For the second part of the proof, we used $0 < \alpha(\Delta\tau) \leq 1/3$ and this demands a smaller upper limit on the length of the partition interval. In any case one can choose a partition of uniform size for all the subintervals satisfying the second inequality.

Remark 3.4. In Theorem 3.1, 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.

4. EXISTENCE OF OPTIMAL CONTROLS

For study of optimal controls we use the continuity of the map $u \longrightarrow x$, that is, the control to solution map. This is crucial for the proof of existence of optimal controls. 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 the norm topology on $B_\infty^a(I, E)$ as seen in section 3. So we must consider an admissible topology for the control space. In sections 2 and 3, we introduced formally the set $L_\infty^\alpha(I, \mathcal{M}_1(U))$, or any closed subset thereof, as the candidate for the space of admissible controls. This is the class of weak star measurable $\mathcal{G}_t$-adapted random processes with values in $\mathcal{M}_1(U)$. To be more precise, let $\lambda \times P$ denote the product of Lebesgue measure and the probability measure on the Cartesian product $I \times \Omega$, and let $\mathcal{P}$ denote the sigma algebra of $\mathcal{G}_t$-predictable subsets of the set $I \times \Omega$ and μ denote the restriction of the measure $\lambda \times P$ on to $\mathcal{P}$. We consider the finite measure space $(I \times \Omega, \mathcal{P}, \mu)$ to be complete and separable. Let $L_1(\mu, C(U)) \equiv L_1((I \times \Omega, \mathcal{P}, \mu), C(U))$ denote the space of μ-measurable Bochner integrable functions on $I \times \Omega$ with values in the Banach space $C(U)$. Since the dual $\mathcal{M}(U)$ of the space $C(U)$ does not satisfy the Radon-Nikodym property the topological dual of the space $L_1(\mu, C(U))$ is not given by $L_\infty(\mu, \mathcal{M}(U))$. However, by virtue of the theory of "lifting" [24, Theorems 7, 9, pp. 94–97], its topological dual is given by $L_\infty^{w^*}(\mu, \mathcal{M}(U))$ which consists of w^*-μ-measurable $\mathcal{M}(U)$ valued functions defined on $I \times \Omega$. We denote this space by $L_\infty^\alpha(\mu, \mathcal{M}(U))$. By measurability here we mean the function

$$(t,\omega) \longrightarrow u_{t,\omega}(\varphi) \equiv \int_U \varphi(\xi) u_{t,\omega}(d\xi)$$

is μ-measurable for each $\varphi \in C(U)$. Equivalently, one may consider the measurability with respect to the Borel sigma algebra generated by the weak star open (or closed) subsets of $\mathcal{M}(U)$. Thus, for any continuous linear functional ℓ on the space $L_1(\mu, C(U))$, there exists a unique $\nu \in L_\infty^\alpha(\mu, \mathcal{M}(U))$ (determined solely by ℓ) such that

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

This is the natural duality pairing. Occasionally, for convenience of notation we may also write this as

$$\ell(\varphi) = \mathbf{E} \int_{I\times U} \varphi(t,\xi)\nu_t(d\xi)dt.$$

For admissible controls, it is possible to choose any closed bounded convex subset of $L_\infty^\alpha(\mu, \mathcal{M}(U))$ since, by Alaoglu's theorem, such a set is weak star compact. However, for reasons of compatibility with the popular view of controls as measurable functions with values in U, it is more appropriate to choose $L_\infty^\alpha(\mu, \mathcal{M}_1(U))$ as the set of admissible controls. By virtue of Alaoglu's theorem the set $L_\infty^\alpha(\mu, \mathcal{M}_1(U))$ is weak star compact. Since the measure space $(I \times \Omega, \mathcal{P}, \mu)$ is separable and, by compactness of the Polish space U, $C(U)$ is also separable, the Lebesgue-Bochner space $L_1(\mu, C(U))$ is separable. Thus it follows from a well known theorem [see Dunford-Schwartz, 19, Theorem V.5.1, p. 426] that $L_\infty^\alpha(\mu, \mathcal{M}_1(U))$ is metrizable. Let $\{g_n\}$ be a dense subset of $L_1(\mu, C(U))$ and $u, v \in L_\infty^\alpha(\mu, \mathcal{M}_1(U))$. Define

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

where $u_{t,\omega}(g_n) \equiv \int_U g_n(t,\omega,\xi) u_{t,\omega}(d\xi) \equiv \overline{g}_n(t,\omega,u_{t,\omega})$. The reader can easily verify that d defines a metric on $L_\infty^\alpha(\mu, \mathcal{M}_1(U))$ and that, with respect to this metric topology, it is a compact metric space and this metric topology is equivalent to its original weak star topology. We denote this metric space by $\mathcal{U}_d$ and for admissible controls we choose $\mathcal{U}_{ad} = \mathcal{U}_d$.

Now we present a result on continuity of the control to solution map.

Theorem 4.1. Consider the control system (2.1) (or equivalently (3.1)) with admissible controls $\mathcal{U}_{ad} \equiv \mathcal{U}_d$ equipped with the metric topology d. Suppose the assumptions of Theorem 3.1 hold and that the semigroup $S(t)$, $t > 0$, is compact. Then, the control to solution map $u \longrightarrow x$ is continuous with respect to the metric topology d on $\mathcal{U}_{ad}$ and the strong (norm) topology on $B_\infty^a(I, E)$.

Proof. Let $\{u^n, u^o\} \in \mathcal{U}_{ad}$ be any sequence and suppose $u^n \xrightarrow{d} u^o$. Let $\{x^n, x^o\} \in B^a_\infty(I, E)$, with $x^n(0) = x^o(0) = x_0$, denote the solutions of the integral equation (3.1) corresponding to the controls $\{u^n, u^o\}$ respectively. We show that $x^n \xrightarrow{s} x^o$ in $B^a_\infty(I, E)$. Clearly, it follows from equation (3.1) corresponding to the controls $\{u^n, u^o\}$ that

$$(4.1)\; x^n(t) - x^o(t) = \int_0^t S(t-s)\big(\overline{f}(s, x^n(s), \overline{x^n(s)}, u^n_s) - \overline{f}(s, x^o(s), \overline{x^o(s)}, u^o_s)\big)\, ds$$
$$+ \int_0^t S(t-s)\big(\overline{\sigma}(s, x^n(s), \overline{x^n(s)}, u^n_s) - \overline{\sigma}(s, x^o, \overline{x^o(s)}, u^o_s)\big)\, dW(s), \quad t \in I.$$

Following similar computations as in the proof of Theorem 3.1, it follows from (4.1) that

$$(4.2) \quad \mathbf{E}|x^o(t) - x^n(t)|^2_E \le 4M^2K^2 t \int_0^t \{\mathbf{E}|x^o(s) - x^n(s)|^2_E + |\overline{x^o(s)} - \overline{x^n(s)}|^2_E\} ds$$
$$+ 4M^2K^2_{\mathcal{R}} \int_0^t \{\mathbf{E}|x^o(s) - x^n(s)|^2_E + |\overline{x^o(s)} - \overline{x^n(s)}|^2_E\} ds$$
$$+ 4\mathbf{E}|e^n_1(t)|^2_E + 4\mathbf{E}|e^n_2(t)|^2_E, \quad t \in I,$$

where

$$(4.3) \qquad e^n_1(t) = \int_0^t S(t-s)\overline{f}(s, x^o(s), \overline{x^o(s)}, u^o_s - u^n_s) ds, \quad t \in I,$$

$$(4.4) \qquad e^n_2(t) = \int_0^t S(t-s)\overline{\sigma}(s, x^o(s), \overline{x^o(s)}, u^o_s - u^n_s) dW(s), \quad t \in I.$$

Recalling that

$$|\overline{x^o(s)} - \overline{x^n(s)}|^2_E \le \mathbf{E}|x^o(s) - x^n(s)|^2, \quad s \in I,$$

it follows from the expression (4.2) that

$$(4.5) \qquad \mathbf{E}|x^o(t) - x^n(t)|^2_E \le C \int_0^t \mathbf{E}|x^o(s) - x^n(s)|^2_E ds + h_n(t), \quad t \in I,$$

where $C \equiv C(T) = 8M^2K^2T + 8M^2K^2_{\mathcal{R}}$, and the function h_n is given by

$$h_n(t) \equiv 4\mathbf{E}|e^n_1(t)|^2_E + 4\mathbf{E}|e^n_2(t)|^2_E, \quad t \in I.$$

Then, by virtue of Gronwall inequality, it follows from (4.5) that

$$(4.6) \qquad \mathbf{E}|x^o(t) - x^n(t)|^2_E \le h_n(t) + Ce^{CT} \int_0^T h_n(s) ds.$$

We prove that the expression on the righthand side of the above inequality converges to zero uniformly on I. Consider the expressions given by (4.3) and (4.4). Since the semigroup is compact, and the integrands are dominated by an integrable process (due to Corollary 3.2 and growth properties (A2)-(A3)), and $u^n \xrightarrow{d} u^o$ it is clear that for each fixed $t \in I$, both $e^n_1(t)$ and $e^n_2(t)$ converge strongly in E to zero P-as. Moreover, computing the expected value of the square of the norm and using the conditional expectation with respect to the current of sigma algebras $\mathcal{G}_t, t \geq 0$, we obtain the following inequalities:

$$\text{(4.7)} \qquad \mathbf{E}|e_1^n(t)|_E^2 \leq t\mathbf{E}\int_0^t \mathbf{E}\big\{|S(t-s)\overline{f}(s, x^o(s), \overline{x}^o(s), u_s^o - u_s^n)|_E^2 |\mathcal{G}_s\big\} ds$$
$$\leq T\int_{I\times\Omega} \chi_t(s)\mathbf{E}\big\{|S(t-s)\overline{f}(s, x^o(s), \overline{x}^o(s), u_s^o - u_s^n)|_E^2 |\mathcal{G}_s\big\}\ d\mu$$

$$\text{(4.8)} \qquad \mathbf{E}|e_2^n(t)|_E^2 = \mathbf{E}\int_0^t \mathbf{E}\big\{|S(t-s)\overline{\sigma}(s, x^o(s), \overline{x}^o(s), u_s^o - u_s^n)|_{\mathcal{R}}^2 |\mathcal{G}_s\big\} ds$$
$$\leq \int_{I\times\Omega} \chi_t(s)\mathbf{E}\big\{|S(t-s)\overline{\sigma}(s, x^o(s), \overline{x}^o(s), u_s^o - u_s^n)|_{\mathcal{R}}^2 |\mathcal{G}_s\big\}\ d\mu,$$

where $\chi_t(\cdot)$ denotes the characteristic function of the interval $[0,t]$. It follows from the assumption (A1) and the growth properties in assumptions **(A2)** and **(A3)** that the integrands in (4.7) and (4.8), are dominated by the following μ-integrable processes $\{\zeta_1, \zeta_2\}$ given by the conditional expectations (with respect to the current of sigma algebras $\mathcal{G}_s \subset \mathcal{F}_s$, $s \in I$) as shown below,

$$\text{(4.9)} \qquad \zeta_1(s) \equiv M^2K^2 s\mathbf{E}\big\{\big(1 + |x^o(s)|_E^2 + |\overline{x}^o(s)|_E^2\big)|_{\mathcal{G}_s}\big\}, \quad s \in I,$$

$$\text{(4.10)} \qquad \zeta_2(s) \equiv M^2K_{\mathcal{R}}^2\mathbf{E}\big\{\big(1 + |x^o(s)|_E^2 + |\overline{x}^o(s)|_E^2\big)|_{\mathcal{G}_s}\big\}, \quad s \in I.$$

Since the semigroup $S(t)$, $t > 0$, is compact and the integrands in (4.7) and (4.8) are dominated by the μ-integrable processes given by (4.9)–(4.10), and $u^n \xrightarrow{d} u^o$, it follows from the inequalities (4.7)–(4.8) that for each $t \in I$, both $\mathbf{E}|e_1^n(t)|_E^2 \longrightarrow 0$ and $\mathbf{E}|e_2^n(t)|_E^2 \longrightarrow 0$ as $n \to \infty$. Since by Corollary 3.2 these functions are also bounded above, it follows from Lebesgue bounded (dominated) convergence theorem that, as $n \to \infty$, both the integrals $\int_I \mathbf{E}|e_n^1(t)|_E^2 dt \to 0, \int_I \mathbf{E}|e_n^2(t)|_E^2 dt \to 0$. This shows that the integral $\int_I h_n(t)dt$, as defined above, converges to zero. Further, by virtue of continuity of the semigroup (in the strong operator topology) and the fact that the integrand is dominated by an integrable function, we conclude that $t \longrightarrow e_1^n(t)$ is continuous P-a.s. Again, using Da Prato-Zabczyk factorization technique [18], one can verify that $t \longrightarrow e_2^n(t)$ is also continuous (more precisely has continuous modification). Thus both $\mathbf{E}|e_1^n(t)|_E^2$ and $\mathbf{E}|e_2^n(t)|_E^2$ are continuous and bounded and converge to zero for all $t \in I$. Hence $h_n(t) \longrightarrow 0$ uniformly on I and therefore, the expression on the righthand side of the inequality (4.6) converges to zero uniformly on I. Clearly, it follows from these facts that $x^n \longrightarrow x^o$ in the norm topology of $B_\infty^a(I, E)$ proving continuity as stated. □

Now we are prepared to consider control problems. Consider the system (2.1) with the cost functional given by

$$\text{(4.11)} \qquad J(u) \equiv \mathbf{E}\bigg\{\int_0^T \ell(t, x(t), \overline{x(t)}, u_t)dt + \Phi(x(T), \overline{x(T)})\bigg\}$$

and admissible controls $\mathcal{U}_{ad}$ as described at the beginning of this section. Our objective is to find a control from the admissible class that minimizes the cost functional (4.11). In the following theorem we prove the existence of optimal control.

Theorem 4.2. Consider the system (2.1) with admissible controls $\mathcal{U}_{ad}$ and the cost functional (4.11). Suppose the assumptions of Theorem 4.1 hold and that ℓ and Φ are Borel measurable real valued functions on $I \times E \times E \times \mathcal{M}_1(U)$ and $E \times E$ respectively and satisfy the following growth conditions:

$$(H1): \ |\ell(t,x,y,u)| \leq C_1(1+|x|_E^2+|y|_E^2) \ \forall \ u \in \mathcal{M}_1(U)$$
$$(H2): \ |\Phi(x,y)| \leq C_2(1+|x|_E^2+|y|_E^2))$$

for some constants $C_1, C_2 > 0$. Further, suppose Φ is lower semicontinuous on $E \times E$ and ℓ is lower semicontinuous in the second and third argument in the norm topology of $E \times E$; and in the fourth argument with respect to the w^* (weak star) topology on $\mathcal{M}_1(U)$. Then there exists an optimal control $u^o \in \mathcal{U}_{ad}$.

Proof. Since $\mathcal{U}_{ad}$ is compact with respect to the metric topology d it suffices to prove that J is lower semicontinuous on $\mathcal{U}_{ad}$ with respect to this (metric) topology. Let $\{u^n\} \subset \mathcal{U}_{ad}$ be a sequence such that $u^n \xrightarrow{d} u^o \in \mathcal{U}_{ad}$. Let $\{x^n\}$ and x^o denote the mild solutions of the evolution equation (2.1) corresponding to the controls $\{u^n\}$ and u^o respectively. Then it follows from theorem 4.1 that $x^n \xrightarrow{s} x^o$ in $B_\infty^a(I,E)$. This implies that, for each $t \in I$, there exists a subsequence of $\{x^n(t)\}$ that converges to $x^o(t)$ P-a.s strongly in E. However, since the Lebesgue measure $\lambda(I) < \infty$, it is clear that the embedding $B_\infty^a(I,E) \hookrightarrow L_2^a(I,E) \equiv L_2^{\mathcal{F}}(I \times \Omega, E)$ is continuous. Thus whenever $x^n \xrightarrow{s} x^o$ in $B_\infty^a(I,E)$, it is clear that $x^n \xrightarrow{s} x^o$ in $L_2^a(I,E)$ also and hence x^n converges to x^o in measure and therefore there exists a subsequence $\{x^{n_k}\} \subset \{x^n\}$, independent of $(t,\omega) \in I \times \Omega$, such that $x^{n_k}(t,\omega) \xrightarrow{s} x^o(t,\omega)$ in E, $\lambda \times P - a.e.$ It is evident that $\| x^{n_k} - x^o \|_{B_\infty^a(I,E)} \longrightarrow 0$ as $k \to \infty$. For convenience of notation we continue to relabel the subsequence as the original sequence. it is easy to verify that there exists a subsequence of the sequence $\{x^n\}$ (independent of (t,ω)), relabeled as $\{x^n\}$, that converges strongly in $L_2^{\mathcal{F}}(I \times \Omega, E)$ to the same limit x^o. This follows from the fact that the norm topology of $B_\infty^a(I,E)$ is stronger than that of the Hilbert space $L_2^{\mathcal{F}}(I \times \Omega, E)$. Clearly,

$$|\overline{x^n(t)} - \overline{x^o(t)}|_E = |\mathbf{E}x^n(t) - \mathbf{E}x^o(t)|_E \leq \sqrt{\mathbf{E}|x^n(t) - x^o(t)|_E^2}.$$

Hence $\overline{x^n(t)} \xrightarrow{s} \overline{x^o(t)}$ in E for each $t \in I$. Recalling that $\{x^n, x^o\}$ have continuous modifications, it follows from lower semicontinuity assumption of ℓ and Φ that

$$\ell(t, x^o(t), \overline{x^o(t)}, u_t^o) \leq \liminf \ell(t, x^n(t), \overline{x^n(t)}, u_t^n), \quad t \in I,$$
$$\Phi(x^o(T), \overline{x^o(T)}) \leq \liminf \Phi(x^n(T), \overline{x^n(T)})$$

P-a.s. By virtue of the hypothesis (H1) and (H2) and Corollary 3.2, it is easy to see that the processes $\{\ell(t, x^n(t), \overline{x^n(t)}, u_t^n)\}$ and $\{\Phi(x^n(T), \overline{x^n(T)})\}$ are dominated by integrable processes. Hence it follows from generalized Fatou's Lemma that

$$\mathbf{E}\int_0^T \ell(t, x^o(t), \overline{x^o(t)}, u_t^o)\, dt \le \liminf \mathbf{E}\int_0^T \ell(t, x^n(t), \overline{x^n(t)}, u_t^n)\, dt \tag{4.12}$$

$$\mathbf{E}\{\Phi(x^o(T), \overline{x^o(T)})\} \le \liminf \mathbf{E}\{\Phi(x^n(T), \overline{x^n(T)})\}. \tag{4.13}$$

Since the sum of lower semicontinuous functionals is lower semicontinuous we conclude that J is lower semicontinuous on $\mathcal{U}_{ad}$ in the metric topology d. That is,

$$J(u^o) \le \liminf J(u^n).$$

Thus there exists a control $u^* \in \mathcal{U}_{ad}$ at which J attains its minimum. This proves the existence of an optimal control. □

Remark 4.3. It is well known that the class of regular controls (denoted by) $\mathcal{U}^r$, consisting of bounded measurable $\mathcal{G}_t$ adapted U-valued random processes (furnished with the topology of convergence in μ measure), is a subclass of relaxed controls $\mathcal{U}_{ad}$. This follows from the fact that, for every $u \in \mathcal{U}^r$, the Dirac measure $\delta_{u(t)}, u \in \mathcal{U}^r$, concentrated along the path process u, is a relaxed control. In other words, for any test function $\varphi \in L_1(\mu, C(U))$ the duality product is given by

$$\mathbf{E}\int_{I\times U} \varphi(t,\xi) u_t(d\xi)dt = \mathbf{E}\int_{I\times U} \varphi(t,\xi)\delta_{u(t)}(d\xi)dt = \mathbf{E}\int_I \varphi(t, u(t))dt.$$

Clearly, the embedding $\mathcal{U}^r \hookrightarrow \mathcal{U}_{ad}$ is continuous. The advantage of using relaxed controls is that the set U need not be convex, can be even a discrete set of points, where as, if one uses regular controls, one can show that in the absence of convexity no optimal control exists. This is a well known fact even for deterministic systems with measurable controls [Ahmed, [29], example 6.2.22, p. 276].

Remark 4.4. Practical realization of relaxed controls is difficult. However, it follows from Krien-Millman theorem that $clco\{ext(\mathcal{M}_1(U))\} = \mathcal{M}_1(U)$ where $ext(\mathcal{M}_1(U))$ denotes the set of extreme points of the set $\mathcal{M}_1(U)$. The set of extreme points, $ext(\mathcal{M}_1(U))$, is just the set of Dirac measures $\{\delta_u, u \in U\}$. And hence regular controls under the embedding is dense in the relaxed controls. Thus relaxed controls can be approximated by regular controls as closely as necessary.

5. NECESSARY CONDITIONS OF OPTIMALITY

In this section we develop the necessary (and possibly sufficient) conditions of optimality whereby one can develop computational algorithm to determine the optimal control. To construct necessary conditions of optimality one requires more regularity properties for the drift and the diffusion operators. For this reason we introduce the following additional assumptions:

(A4): The drift $\overline{f} = \overline{f}(t, x, y, u)$ and the diffusion operator $\overline{\sigma} = \overline{\sigma}(t, x, y, u)$ are continuously once Fréchet differentiable in their second and third argument and the Fréchet derivatives are uniformly bounded on $I \times E \times E \times \mathcal{M}_1(U)$.

(A5): The cost integrand $\ell = \ell(t, x, y, u)$ and $\Phi = \Phi(x, y)$ are continuously Gâteaux differentiable with respect to the arguments $x, y \in E$, and there exist constants $C_1, C_2 > 0$ so that their Gâteaux derivatives satisfy the following growth conditions:

$$|\ell_x(t, x, y, u)|_E \leq C_1(1 + |x|_E + |y|_E) \ \forall \ (t, x, y, u) \in I \times E \times E \times \mathcal{M}_1(U);$$
$$|\ell_y(t, x, y, u)|_E \leq C_1(1 + |x|_E + |y|_E) \ \forall \ (t, x, y, u) \in I \times E \times E \times \mathcal{M}_1(U)$$
$$|\Phi_x(x, y)|_E \leq C_2(1 + |x|_E + |y|_E) \ \forall \ (x, y) \in E \times E$$
$$|\Phi_y(x, y)|_E \leq C_2(1 + |x|_E + |y|_E) \ \forall \ (x, y) \in E \times E.$$

In order 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 (2.1) with respect to controls $u \in \mathcal{U}_{ad}$. We present this in the following lemma.

Lemma 5.1. Suppose the assumptions **(A1)–(A4)** and those of Theorem 4.1 hold. Then for any pair $\{u^o, u\} \in \mathcal{U}_{ad}$, there exists a $z \in B^a_\infty(I, E) \subset L^a_2(I, E)$ which is the unique mild solution of the following variational equation

$$(5.1)\quad \begin{aligned} dz = {}& Azdt + \overline{f}_x(t, x^o(t), \overline{x^o(t)}, u^o_t)zdt + \overline{f}_y(t, x^o(t), \overline{x^o(t)}, u^o_t)\overline{z}dt \\ & + \overline{\sigma}_x(t, x^o(t), \overline{x^o(t)}, u^o_t; z)dW(t) + \overline{\sigma}_y(t, x^o(t), \overline{x^o(t)}, u^o_t; \overline{z})dW(t) + d\Lambda^{u-u^o}_t, \\ & z(0) = 0, \quad t \in I, \end{aligned}$$

where Λ is the semi-martingale given by

$$d\Lambda^{u-u^o}_t = \overline{f}(t, x^o(t), \overline{x^o(t)}, u_t - u^o_t)dt + \overline{\sigma}(t, x^o(t), \overline{x^o(t)}, u_t - u^o_t)dW(t),$$

starting from $\Lambda^{u-u^o}_0 = 0$. The solution $z = s - \lim_{\varepsilon \downarrow 0} \{z^\varepsilon \equiv (1/\varepsilon)(x^\varepsilon - x^o)\}$ where $\{x^\varepsilon, x^o\} \subset B^a_\infty(I, E)$ are the solutions of the integral equation (3.1) corresponding to the controls $\{u^\varepsilon, u^o\} \in \mathcal{U}_{ad}$ respectively for $u^\varepsilon \equiv u^o + \varepsilon(u - u^o), \varepsilon \in (0, 1)$.

Proof. For simplicity of notation let us denote by $F_1(t), F_2(t), \sigma_1(t; \cdot), \sigma_2(t; \cdot)$ the operators $\{\overline{f}_x, \overline{f}_y, \overline{\sigma}_x, \overline{\sigma}_y\}$ respectively all evaluated at $(t, x^o(t), \overline{x^o(t)}, u^o_t)$. Using this notation, equation (5.1) can be compactly written as

$$(5.2)\quad \begin{aligned} dz = {}& Azdt + F_1(t)zdt + F_2(t)\overline{z}dt + \sigma_1(t; z)dW(t) + \sigma_2(t; \overline{z})dW(t) + d\Lambda^{u-u^o}_t, \\ & z(0) = 0, \quad t \in I. \end{aligned}$$

It follows from the assumption **(A4)** that $\{F_1, F_2\}$ and $\{\sigma_1, \sigma_2\}$ are uniformly bounded $\mathcal{F}_t$-adapted operator valued functions with values in $\mathcal{L}(E)$ and $\mathcal{L}(E, \mathcal{L}(H, E))$ respectively. This is a linear stochastic evolution equation on E and it is a special case of the nonlinear equation (2.1). Thus it follows from Theorem 3.1 that this equation has a unique mild solution $z \in B^a_\infty(I, E)$ having continuous modification. We must verify that this is given by the limit

$$\lim_{\varepsilon \to 0} z^\varepsilon(t) \equiv \lim_{\varepsilon \to 0}(1/\varepsilon)\{x^\varepsilon(t) - x^o(t)\} = z(t) \tag{5.3}$$

in the norm topology of $B^a_\infty(I, E)$, where x^ε and x^o are the unique mild solutions of the SDE (2.1) corresponding to controls $u^\varepsilon \equiv u^o + \varepsilon(u - u^o)$ and u^o respectively. Recall that the set of relaxed controls is naturally convex and hence $u^\varepsilon \in \mathcal{U}_{ad}$. It is tedious but straightforward to verify (5.3). For the sake of the reader we present a broad outline. We use Lagrange formula. Let X, Y be any pair of real Banach spaces and $F : X \longrightarrow Y$ once continuously Gâteaux differentiable. Then for any pair $\{x^o, x\} \in X$ one has

$$F(x) = F(x^o) + \int_0^1 d\theta \ F_x(x^o + \theta(x - x^o)) \cdot [x - x^o]$$

where F_x denotes the Gâteaux differential of F evaluated along the line segment $x^o + \theta(x - x^o), \theta \in (0, 1)$, as shown and it is an operator valued function with values in $\mathcal{L}(X, Y)$. In case $Y = R$, this reduces to

$$F(x) = F(x^o) + \int_0^1 d\theta \langle F_x(x^o + \theta(x - x^o)), [x - x^o]\rangle_{X^*,X}$$

where now $F_x \in X^*$, the dual of X. Define the following operator valued functions:

$$\begin{aligned} G^\varepsilon_1(s) &\equiv \int_0^1 d\theta \overline{f}_x(s, x^o(s) + \theta(x^\varepsilon(s) - x^o(s)), \overline{x^o(s)} + \theta(\overline{x^\varepsilon(s)} - \overline{x^o(s)}), u^o_s), \quad s \in I; \end{aligned} \tag{5.4}$$

$$\begin{aligned} G^\varepsilon_2(s) &\equiv \int_0^1 d\theta \Big\{ \overline{f}_x(s, x^o(s) + \theta(x^\varepsilon(s) - x^o(s)), \overline{x^o(s)} + \theta(\overline{x^\varepsilon(s)} - \overline{x^o(s)}), u^o_s) \\ &\quad - \overline{f}_x(s, x^o(s), \overline{x^o(s)}, u^o_s) \Big\}, \quad s \in I; \end{aligned} \tag{5.5}$$

$$\begin{aligned} G^\varepsilon_3(s; e) &\equiv \int_0^1 d\theta \ \overline{\sigma}_x(s, x^o(s) + \theta(x^\varepsilon(s) - x^o(s)), \overline{x^o(s)} + \theta(\overline{x^\varepsilon(s)} - \overline{x^o(s)}), u^o_s; e), \\ &\quad e \in E, \quad s \in I; \end{aligned} \tag{5.6}$$

$$(5.7)\quad G_4^\varepsilon(s;e) \equiv \int_0^1 d\theta\Big\{\overline{\sigma}_x(s, x^o(s)+\theta(x^\varepsilon(s)-x^o(s)), \overline{x^o(s)}+\theta(\overline{x^\varepsilon(s)}-\overline{x^o(s)}), u_s^o; e) \quad \overline{\sigma}_x(s, x^o(s), \overline{x^o(s)}, u_s^o; e)\Big\}, \quad e\in E, \quad s\in I.$$

It is clear that $\{G_1^\varepsilon(t), G_2^\varepsilon(t)\}$ and $\{G_3^\varepsilon(t;\cdot), G_4^\varepsilon(t;\cdot)\}$ are $\mathcal{F}_t$-adapted random processes taking values in $\mathcal{L}(E)$ and $\mathcal{L}(E,\mathcal{L}(H,E))$ respectively and by assumption **(A4)** are uniformly norm bounded on I. Let these bounds be denoted by b for $\{G_1^\varepsilon, G_2^\varepsilon\}$ and $b_{\mathcal{R}}$ for $\{G_3^\varepsilon, G_4^\varepsilon\}$. Further, since by Theorem 4.1, $x^\varepsilon \longrightarrow x^o$ in $B_\infty^a(I,E)$, it follows from continuity of the Frèchet derivatives $\{\overline{f}_x, \overline{\sigma}_x\}$ that

$$(5.8)\qquad \lim_{\varepsilon\to 0} G_2^\varepsilon(t) \longrightarrow 0 \text{ and } \lim_{\varepsilon\to 0} G_4^\varepsilon(t;e) \longrightarrow 0 \text{ for } a.e\ t\in I, \quad P.a.s\ \forall e\in E.$$

Defining $\eta^\varepsilon(t)\equiv z^\varepsilon(t)-z(t)$ and using the operators introduced above one obtains the following integral equation for η^ε

(5.9)

$$\eta^\varepsilon(t) = \int_0^t S(t-r)G_1^\varepsilon(r)\eta^\varepsilon(r)dr + \int_0^t S(t-r)G_1^\varepsilon(r)\ \overline{\eta^\varepsilon(r)}dr + \int_0^t S(t-r)G_3^\varepsilon(r;\eta^\varepsilon(r))dW(r) + \int_0^t S(t-r)G_3^\varepsilon(r;\overline{\eta^\varepsilon(r)})dW(r) + h^\varepsilon(t),$$

where $h^\varepsilon(t)\equiv\sum_{i=1}^6 e_i^\varepsilon(t)$ with $\{e_i^\varepsilon, i=1,\dots,6\}$ given by the following expressions:

$$(5.10)\quad e_1^\varepsilon(t) = \int_0^t S(t-r)G_2^\varepsilon(r)z(r)dr, \quad e_2^\varepsilon(t) = \int_0^t S(t-r)G_2^\varepsilon(r)\overline{z(r)}dr$$

$$(5.11)\quad e_3^\varepsilon(t) = \int_0^t S(t-r)G_4^\varepsilon(r;z(r))dW(r), \quad e_4^\varepsilon(t) = \int_0^t S(t-r)G_4^\varepsilon(r;\overline{z(r)})dW(r)$$

$$(5.12)\quad e_5^\varepsilon(t) = \int_0^t S(t-r)\big[\overline{f}(r,x^\varepsilon(r),\overline{x^\varepsilon(r)},u_r-u_r^o) - \overline{f}(r,x^o(r),\overline{x^o(r)},u_r-u_r^o)\big]dr$$

$$(5.13)\quad e_6^\varepsilon(t) = \int_0^t S(t-r)\big[\overline{\sigma}(r,x^\varepsilon,\overline{x^\varepsilon},u_r-u_r^o) - \overline{\sigma}(r,x^o,\overline{x^o},u_r-u_r^o)\big]dW(r).$$

Using the expression (5.9), computing the expected value of norm-square, and using Gronwall inequality and the fact that the operators G_1^ε and G_3^ε are uniformly bounded, one can easily verify that there exists a constant $C=C(M,b,b_{\mathcal{R}},T)$, dependent on the parameters displayed, such that

$$(5.14)\qquad \sup_{t\in I}\mathbf{E}|\eta^\varepsilon(t)|_E^2 \le C\sup_{t\in I}\mathbf{E}|h^\varepsilon(t)|_E^2 + Ce^{CT}\int_0^T \mathbf{E}|h^\varepsilon(t)|_E^2\ dt.$$

It remains to verify that the righthand side of (5.14) converges to zero as $\varepsilon\to 0$. Using the properties (5.8) and the fact that $z\in B_\infty^a(I,E)\subset L_2^a(I,E)$, and Lebesgue dominated convergence theorem, one can easily verify that

$$\lim_{\varepsilon\to 0}\Big\{\sup_{t\in I}\mathbf{E}|e_i^\varepsilon(t)|_E^2\Big\} = 0$$

for all $i = 1, 2, 3, 4$. Similarly, using the continuity and the growth assumptions in **(A2)** and **(A3)** and the fact that $x^\varepsilon(t) \xrightarrow{s} x^o(t)$ in E, P-a.s, once again it follows from Lebesgue dominated convergence theorem that

$$\lim_{\varepsilon\to 0}\left\{\sup_{t\in I}\mathbf{E}|e_i^\varepsilon(t)|_E^2\right\} = 0$$

for $i = 5, 6$. From these results it follows that the expression on the righthand side of the inequality (5.14) converges to zero as $\varepsilon \to 0$. Hence

$$\lim_{\varepsilon\to 0}\sup_{t\in I}\mathbf{E}|\eta^\varepsilon(t)|_E^2 = 0, \tag{5.15}$$

and we conclude that $\eta^\varepsilon \to 0$ in $B_\infty^a(I, E)$. Thus we have proved that $z^\varepsilon \to z$ in the norm topology of $B_\infty^a(I, E)$. This completes the proof. □

Now we are prepared to develop the necessary conditions of optimality. In order to do so we recall some basic properties of semimartingales. An E-valued norm-square integrable continuous $\mathcal{F}_t$-semimartingale with intensity parameters $\{m_0, \phi, \Sigma\}$ has the standard representation

$$m(t) \equiv m_0 + \int_0^t \phi(s)ds + \int_0^t \Sigma(s)dW(s), \quad t \in I,$$

where m_0 is $\mathcal{F}_0$ measurable belonging to $L_2(\Omega, E)$, $\phi \in L_2^a(I, E)$ and $\Sigma \in L_2^a(I, \mathcal{L}_\mathcal{R}(H, E))$. The scalar product of any two such semimartingales is given by

$$\mathbf{E}(m_1, m_2)_E \equiv \mathbf{E}(m_{1,0}, m_{2,0})_E + \mathbf{E}\int_0^T (\phi_1(s), \phi_2(s))_E ds + \mathbf{E}\int_0^T Tr(\Sigma_1(s)\mathcal{R}\Sigma_2^*(s))ds.$$

Completion of this space with respect to the above scalar product is a Hilbert space which we denote by $\mathcal{S}M^2(I, E)$. It is well known that any such semimartingale is uniquely determined by its intensity parameters and conversely. In the sequel, we need the space $\mathcal{S}M_0^2(I, E) \subset \mathcal{S}M^2(I, E)$ of semi-martingales starting from zero.

For convenience of notation we set $L_2^a(I, \mathcal{L}_\mathcal{R}(H, E)) \equiv L_{2,\mathcal{R}}^a(I, \mathcal{L}(H, E))$. From Lemma 5.1 we obtain the following corollary.

Corollary 5.2. Suppose the assumptions of Lemma 5.1 hold. Then the map $\Lambda \longrightarrow z$ is a continuous linear (hence bounded) operator from $\mathcal{S}M_0^2(I, E)$ to $B_\infty^a(I, E)$.

Proof. Using the semigroup and writing the integral equation corresponding to the variational equation (5.2) and using the fact that all the associated operators are uniformly bounded on I one can use Gronwall lemma to arrive at the inequality $\| z \|_{B_\infty^a(I,E)} \leq C \| \Lambda \|_{\mathcal{S}M_0^2(I,E)}$ for some constant $C > 0$. This concludes the proof. □

We present below the necessary conditions of optimality.

Theorem 5.3. Consider the control system (2.1) with the admissible controls $\mathcal{U}_{ad}$ and the cost functional (4.11). Suppose the assumptions of Theorem 4.2 and Lemma 5.1 hold and that ℓ and Φ satisfy the assumption (A5). Then, for a control $u^o \in \mathcal{U}_{ad}$ to be optimal, it is necessary that there exists a pair $(\psi, Q) \in L_2^a(I, E) \times L_{2,\mathcal{R}}^a(I, \mathcal{L}(H, E))$ such that the following inequality holds

$$\begin{aligned}(5.16) \qquad \mathbf{E}\int_I \langle \psi(t), \overline{f}(t, x^o(t), \overline{x^o(t)}, u_t - u_t^o\rangle_E dt \\ + \mathbf{E}\int_I Tr\big(Q(t)\mathcal{R}\overline{\sigma}^*(t, x^o(t), \overline{x^o(t)}, u_t - u_t^o)\big)dt \\ + \mathbf{E}\int_I \ell(t, x^o(t), \overline{x^o(t)}, u_t - u_t^o)\ dt \geq 0, \quad \forall\ u \in \mathcal{U}_{ad},\end{aligned}$$

where $x^o \in B_\infty^a(I, E)$ is the (mild) solution of the evolution equation (2.1) corresponding to the control $u^o \in \mathcal{U}_{ad}$.

Proof. Let $u^o \in \mathcal{U}_{ad}$ be an optimal control (minimizing the cost functional (4.11)) and $u \in \mathcal{U}_{ad}$ any other control. For any $\varepsilon \in (0, 1)$, it follows from the convexity of relaxed controls that $u^\varepsilon \equiv u^o + \varepsilon(u - u^o) \in \mathcal{U}_{ad}$. Thus $J(u^o) \leq J(u^\varepsilon)$ for all $\varepsilon \in (0, 1)$ and hence

$$(1/\varepsilon)(J(u^\varepsilon - J(u^o)) \geq 0\ \forall\ \varepsilon \in (0, 1).$$

Letting $\varepsilon \downarrow 0$ and using the assumption **(A5)**, it follows from the above expression that the Gâteaux differential of J at u^o in the direction $u - u^o$ must satisfy

$$(5.17) \quad dJ(u^o; u - u^o) = L(z) + \mathbf{E}\int_0^T \ell(t, x^o(t), \overline{x^o(t)}, u_t - u_t^o)dt \geq 0,\ \forall\ u \in \mathcal{U}_{ad},$$

where $L(z)$ is given by

$$\begin{aligned}L(z) \equiv \mathbf{E}\int_0^T \Big\{ \langle l_x(t, x^o(t), \overline{x^o(t)}, u_t^o), z(t)\rangle_E \\ + \langle l_y(t, x^o(t), \overline{x^o(t)}, u_t^o), \overline{z(t)}\rangle_E \Big\} dt \\ + \mathbf{E}\big\{ \langle \Phi_x(x^o(T), \overline{x^o(T)}), z(T)\rangle_E + \langle \Phi_y(x^o(T), \overline{x^o(T)}), \overline{z(T)}\rangle_E \big\}\end{aligned}$$

with z being the mild solution of the variational evolution equation (5.2). Clearly, for any two E-valued square integrable random variables $\{X, Y\}$ we have $\mathbf{E}(X, \overline{Y})_E = \mathbf{E}(\overline{X}, Y)_E$. Using this fact it follows from the above expression that

$$\begin{aligned}(5.18) \qquad L(z) \equiv \mathbf{E}\int_0^T \Big\{ \langle [l_x(t, x^o(t), \overline{x^o(t)}, u_t^o) + \mathbf{E} l_y(t, x^o(t), \overline{x^o(t)}, u_t^o)], z(t)\rangle_E \Big\} dt \\ + \mathbf{E}\big\{ \langle [\Phi_x(x^o(T), \overline{x^o(T)}) + \mathbf{E}\Phi_y(x^o(T), \overline{x^o(T)})], z(T)\rangle_E \big\}.\end{aligned}$$

Further, it follows from the assumption **(A5)** and Theorem 3.1 that the Gâteaux derivatives of ℓ with respect to x, y along the arguments indicated, are $\mathcal{F}_t$-adapted and that $\ell_x, \ell_y \in L_2^a(I, E)$. Due to the existence of continuous modification and assumption **(A5)**, it follows from identical arguments that $\Phi_x(x^o(T), \overline{x^o(T)}), \Phi_y(x^o(T), \overline{x^o(T)})$ are well defined and $\mathcal{F}_T$ measurable and belong to $L_2(\Omega, E)$. Thus it follows from the expression (5.18) that L is a continuous linear functional of $z \in B_\infty^a(I, E) \subset L_2^a(I, E)$. It follows from Corollary 5.2 that the mild solution z of the variational equation

(5.1/5.2) is continuously dependent on the semi-martingale $\Lambda \in \mathcal{SM}_0^2(I,E)$. In fact, $\Lambda \longrightarrow z$ is a continuous linear (hence bounded) operator from $\mathcal{S}M_0^2(I,E)$ to $B_\infty^a(I,E)$. Thus, we conclude that the composition map $\tilde{L}$, as defined below

$$\Lambda \longrightarrow z \longrightarrow L(z) \equiv \tilde{L}(\Lambda), \tag{5.19}$$

is a continuous linear functional on $\mathcal{SM}_0^2(I,E)$. Hence, by the semi-martingale representation theory on Hilbert spaces as discussed above, we conclude that there exists a pair $(\psi, Q) \in L_2^a(I,E) \times L_{2,\mathcal{R}}^a(I,\mathcal{L}(H,E))$ such that

$$\begin{aligned} L(z) = \tilde{L}(\Lambda) = \mathbf{E}\int_0^T \langle \psi(t), \overline{f}(t, x^o(t), \overline{x^o(t)}, u_t - u_t^o)\rangle_E dt \\ + \mathbf{E}\int_0^T Tr(Q\mathcal{R}\overline{\sigma}^*(t, x^o(t), \overline{x^o(t)}, u_t - u_t^o))dt. \end{aligned} \tag{5.20}$$

In view of (5.19), using the expression (5.20) in (5.17) we obtain the necessary condition (5.16). This proves the necessary conditions of optimality. □

The question that arises now is how to find this pair

$$(\psi, Q) \in L_2^a(I,E) \times L_{2,\mathcal{R}}^a(I,\mathcal{L}(H,E)) = L_2^a(I,E) \times L_2^a(I,\mathcal{L}_\mathcal{R}(H,E)).$$

Under an additional assumption, we show in the following theorem that this is given by the mild solution of a backward stochastic differential equation. For economy of notations we write

$$\ell_x^o(t) \equiv \ell_x(t, x^o(t), \overline{x^o(t)}, u_t^o), \quad \ell_y^o(t) \equiv \ell_y(t, x^o(t), \overline{x^o(t)}, u_t^o),$$

$$F_1(t) \equiv \overline{f}_x(t, x^o(t), \overline{x^o(t)}, u_t^o), \quad F_2(t) \equiv \overline{f}_y(t, x^o(t), \overline{x^o(t)}, u_t^o),$$

$$\sigma_1(t;\cdot) \equiv \overline{\sigma}_x(t, x^o(t), \overline{x^o(t)}, u_t^o;\cdot), \quad \sigma_2(t;\cdot) \equiv \overline{\sigma}_y(t, x^o(t), \overline{x^o(t)}, u_t^o;\cdot).$$

Theorem 5.4. Suppose the assumptions of Theorem 5.3 hold and that the initial state x_0 of the system (2.1) is $\mathcal{F}_0$ measurable and belongs to $L_4(\Omega, E)$. Then the pair (ψ, Q) is given by the mild solution of the following BSDE:

$$\begin{aligned} -d\varphi =& A^*\varphi dt + F_1^*(t)\varphi dt + \Gamma_1(t)\varphi dt + \big(\mathbf{E}(F_2^*(t)\varphi) + \mathbf{E}(\Gamma_2(t)\varphi)\big)dt \\ &+ (\ell_x^o(t) + \mathbf{E}\ell_y^o(t))dt + \hat{\sigma}_1(t;\varphi)dW \end{aligned} \tag{5.21}$$

$$\varphi(T) = \Phi_x(x^o(T), \overline{x^o(T)}) + \mathbf{E}\Phi_y(x^o(T), \overline{x^o(T)}), \tag{5.22}$$

with $\psi(t) = \varphi(t)$, $t \in I$, and $Q(t) = \hat{\sigma}_1(t;\varphi(t))$, $t \in I$, where the operator valued processes $\{\hat{\sigma}_1(t;\cdot), \Gamma_1(t), \Gamma_2(t)\}$ are derived from the following multilinear forms:

$$(e_2, \sigma_1(t;e_1)h)_E \equiv (e_1, \hat{\sigma}_1(t;e_2)h)_E, \quad e_1, e_2 \in E \text{ and } h \in H,$$

$$Tr(-\hat{\sigma}_1(t;e_2)\mathcal{R}\sigma_1^*(t;e_1)) \equiv (\Gamma_1(t)e_2, e_1), \quad e_1, e_2 \in E,$$

$$Tr(-\hat{\sigma}_1(t;e_2)\mathcal{R}\sigma_2^*(t;e_1)) \equiv (\Gamma_2(t)e_2, e_1), \quad e_1, e_2 \in E.$$

Proof. The proof follows from similar arguments as in [2, Theorem 4.2, p. 72], [7, Theorem 6.3, p. 119]. We present a brief outline. Let $\Sigma_1 \in L_2^a(I, \mathcal{L}_{\mathcal{R}}(H, E))$ which is identified later as $Q = \hat{\sigma}_1(t;\varphi)$. Consider the stochastic system of the form

$$d\varphi = -A^*\varphi dt + (B.V)dt + \Sigma_1(t)dW \tag{5.23}$$

where "$(B.V)$" denotes all the bounded variation terms which are also identified in the body of the detailed proof. Recall the variational evolution equation given by (5.2) as reproduced below:

$$\begin{aligned} dz &= Azdt + F_1(t)zdt + F_2(t)\overline{z}dt + \sigma_1(t;z)dW(t) + \sigma_2(t;\overline{z})dW(t) + d\Lambda_t^{u-u^o}, \\ z(0) &= 0, \quad t \in I. \end{aligned} \tag{5.24}$$

Since it is the mild solution that matters, one can formally compute the Itô differential of the scalar product $(\varphi(t), z(t))_E$ giving

$$d(\varphi, z)_E = (d\varphi, z) + (\varphi, dz) + \ll d\varphi, dz \gg,$$

where the last term denotes the quadratic variation. Integrating this over the time interval I and computing the expected values and using the variational equation (5.24) one arrives at the same cost functional as given by (5.18) while identifying the adjoint evolution equation (5.21) along with the terminal condition (5.22). The formal computations are then justified by use of Yosida approximation of A and taking the limit. A direct proof of existence of an $\mathcal{F}_t$-adapted (mild) solution for equation (5.21)–(5.22) can be carried out using similar approach as given in Hu and Peng [21, Theorem 3.1, p. 455] slightly modified for the mean field components. This completes the brief outline of our proof. □

Remark 5.5. The preceding theorem is proved under the assumption that the initial state has fourth order moment. Clearly, if x_0 is deterministic, this assumption is automatically satisfied since in that case it has moments of all orders. This assumption can be replaced by an alternate assumption on the diffusion operator σ requiring that it is uniformly bounded.

Remark 5.6. Recently [28] we have also considered a more general class of controlled McKean-Vlasov equations where both the drift and the diffusion operators contain the state and the measure induced by it.

6. SOME EXAMPLES ON LQGR PROBLEMS

For illustration of the results presented above we consider several versions of linear quadratic control problems for a system governed by the following linear mean-field evolution equation,

$$dx = Axdt + F(t)\overline{x}(t)dt + B(t)u(t)dt + \sigma(t)dW \tag{6.1}$$
$$x(0) = x_0, \quad t \in I,$$

on the Hilbert space E. The operator A is the infinitesimal generator of a C_0-semigroup $S(t)$, $t \geq 0$, on E, F is a strongly measurable (in the strong operator topology) and uniformly bounded operator valued function with values in $\mathcal{L}(E)$ and B is also a strongly measurable uniformly bounded operator valued function with values in $\mathcal{L}(U, E)$ where U is any separable Hilbert space (so a Polish space) and σ is a strongly measurable operator valued function taking values in $\mathcal{L}_{\mathcal{R}}(H, E)$ with W being an H-Brownian motion with nuclear covariance $\mathcal{R}$. In view of Remark 5.5, if σ is a Bochner measurable uniformly bounded $\mathcal{L}(H, E)$ valued function, the restriction (see Theorem 5.4) requiring the initial state to have fourth order moment, can be removed. Since here the system is linear and the cost functionals are quadratic, the problem is convex, and hence optimal controls exist in the class of regular controls contained in relaxed controls. Further, the necessary conditions of optimality proved in the preceding section are also sufficient. For LQR and LQGR problems in finite dimensional spaces, the reader can refer to any graduate text on control theory, for example, [29, 30].

Example 1. Consider the cost functional of the form

$$J(u) \equiv (1/2)\mathbf{E}\bigg\{ \int_0^T \big[(Q_1(t)x, x)_E + (Q_2(t)\overline{x}, \overline{x})_E) + (R(t)u, u)_U\big] dt \tag{6.2}$$
$$+ (1/2)(M_1 x(T), x(T))_E + (M_2\overline{x}(T), \overline{x}(T))_E \bigg\}$$

where Q_1, Q_2 are strongly measurable operator valued functions taking values in the space of positive selfadjoint operators $\mathcal{L}_s^+(E)$, and R taking values from $\mathcal{L}_s^+(U)$ having continuous inverse. The operators $M_1, M_2 \in \mathcal{L}_s^+(E)$. Using Theorem 5.4 of the preceding section we observe that $\ell_x = Q_1 x$, $\mathbf{E}\ell_y = Q_2\overline{x}, \hat{\sigma}_1(t;\varphi) = 0$. Hence the adjoint evolution equation is given by

$$d\varphi = A^*\varphi dt + F^*(t)\overline{\varphi}dt + Q_1 x dt + Q_2\overline{x}dt \tag{6.3}$$
$$\varphi(T) = M_1 x(T) + M_2\overline{x}(T).$$

Note that it is an ordinary differential equation on the Hilbert space E with stochastic inputs and random boundary (terminal) condition. Clearly, φ is linear in x and $\overline{x}$ and hence it has the form

$$\varphi(t) = K_1(t)x(t) + K_2(t)\overline{x}(t) + r(t), \quad t \in I \equiv [0, T] \tag{6.4}$$

where K_1, K_2 are suitable operator valued functions taking values in $\mathcal{L}(E)$ satisfying the boundary conditions $K_1(T) = M_1$ and $K_2(T) = M_2$ and $r(T) = 0$ with r being an E-valued stochastic process identified below. Using the inequality (5.16) arising from the necessary condition, one can easily verify that the optimal control is given by

$$u^o(t) = -R^{-1}(t)B^*(t)\varphi(t) \tag{6.5}$$
$$= -R^{-1}(t)B^*(t)K_1(t)x(t) - R^{-1}(t)B^*(t)K_2(t)\overline{x(t)} - R^{-1}(t)B^*(t)r(t),$$

for $t \in I$. Following similar procedure as in the classical regulator problems, the reader can easily verify that the operators $\{K_1, K_2\}$ must satisfy the following system of coupled operator-Ricatti equations,

$$(d/dt)K_17 + (K_1A + A^*K_1) - K_1BR^{-1}B^*K_1 + Q_1 = 0, K_1(T) = M_1, \tag{6.6}$$

$$(d/dt)K_2 + K_2(A + F - BR^{-1}B^*K_1) + (A + F - BR^{-1}B^*K_1)^*K_2 \tag{6.7}$$
$$+ (K_1F + F^*K_1) - K_2BR^{-1}B^*K_2 + Q_2 = 0, K_2(T) = M_2,$$

on the Banach space $\mathcal{L}(E)$, and the process r satisfies the following BSDE on E,

$$dr + (A + F - BR^{-1}B^*K_1)^*rdt + K_1\sigma dW = 0, r(T) = 0. \tag{6.8}$$

These operator-Ricatti equations are solved in the weak sense. For example, considering equation (6.6), for any complete ortho-normal set $\{e_i\} \subset D(A) \subset E$, one solves the following system of equations

$$(\dot{K}_1e_i, e_j) + (Ae_i, K_1e_j) + (K_1e_i, Ae_j) - (BR^{-1}B^*K_1e_i, K_1e_j) + (Q_1e_i, e_j) = 0, \tag{6.9}$$
$$(K_1(T)e_i, e_j) = (M_1e_i, e_j), \quad i, j \in N.$$

Remark 6.1. If the mean field terms are omitted by setting $F(t) \equiv 0, Q_2(t) \equiv 0$, $M_2 = 0$, we find that the operator equation (6.7) reduces to

$$\dot{K}_2 + K_2(A - BR^{-1}B^*K_1) + (A - BR^{-1}B^*K_1)^*K_2 - K_2BR^{-1}B^*K_2 = 0, \tag{6.10}$$
$$K_2(T) = 0.$$

This is a homogeneous equation in K_2 with terminal condition zero. Hence it can have only the trivial solution $K_2(t) \equiv 0$ and we are left with the equations (6.6) and (6.8) (with $F \equiv 0$). Thus we have recovered the well known results for classical stochastic regulator problems. Further, in the absence of noise with $\sigma \equiv 0$, equation (6.8) turns into a homogeneous equation with $r(T) = 0$. In this case $r(t) \equiv 0$, $t \in I$, and we recover the classical results for deterministic linear quadratic regulator problems.

Example 2. Consider the evolution equation (6.1) with the cost functional

$$J(u) \equiv (1/2)\mathbf{E}\bigg\{\int_0^T \{(Q(x - \overline{x}), x - \overline{x})_E + (Ru, u)_U\}dt \tag{6.11}$$
$$+ (M(x(T) - \overline{x}(T)), x(T) - \overline{x}(T))_E\bigg\},$$

where $Q(t) \in \mathcal{L}_s^+(E)$, $R(t) \in \mathcal{L}_s^+(U)$ for all $t \in I$, with $R^{-1}(t)$ continuous and bounded for all $t \in I$, and $M \in \mathcal{L}_s^+(E)$. Here the objective is to minimize the volatility or fluctuation around the mean. Following the necessary conditions of optimality, Theorem 5.3 and Theorem 5.4, one can verify that the adjoint evolution equation is given by

$$-d\varphi = A^*\varphi\, dt + F^*(t)\overline{\varphi}\, dt + Q(x - \overline{x})\, dt,\ \varphi(T) = M(x(T) - \overline{x}(T)). \tag{6.12}$$

Again, this is an ordinary (meanfield) linear differential equation on the Hilbert space E with stochastic input. Thus the solution is given by

$$\varphi(t) = K(t)(x(t) - \overline{x}(t)) + r(t)$$

for suitable operator valued function K and a process r. Following similar steps as in the first example, it is easy to verify that K satisfies the following operator Ricatti equation on $\mathcal{L}(E)$

$$\begin{aligned} &(d/dt)K + (KA + A^*K) - (KBR^{-1}B^*K) + Q = 0 \\ &K(T) = M \end{aligned} \tag{6.13}$$

in the weak sense and the process r is given by the (mild) solution of the following BSDE

$$dr + (A - BR^{-1}B^*K)^*rdt + (F + BR^{-1}B^*K)^*\overline{r}dt + K\sigma dW = 0, \quad r(T) = 0.$$

Using the above equation one can verify that $\overline{r}(t) \equiv 0$ and therefore it reduces to the following BSDE

$$dr + (A - BR^{-1}B^*K)^*rdt + K\sigma dW = 0, \quad r(T) = 0. \tag{6.14}$$

In this case the optimal feedback control is given by

$$u^o(t) = -R^{-1}(t)B^*(t)[K(t)(x(t) - \overline{x}(t)) + r(t)], \quad t \in I. \tag{6.15}$$

Example 3. Consider the evolution equation 6.1 with the cost functional given by

$$\begin{aligned} J(u) \equiv (1/2)\mathbf{E}\Big\{ &\int_0^T (Q(x - \overline{x}), x - \overline{x})_E - (C\overline{x}, \overline{x})_E + (Ru, u)_U dt \\ &+(M(x(T) - \overline{x}(T)), x(T) - \overline{x}(T))_E - (N\overline{x(T)}, \overline{x}(T))_E \Big\}, \end{aligned} \tag{6.16}$$

where $Q(t), C(t) \in \mathcal{L}_s^+(E)$, $R(t) \in \mathcal{L}_s^+(U)$ are bounded for all $t \in I$, with $R^{-1}(t)$ continuous and bounded for all $t \in I$, and $M, N \in \mathcal{L}_s^+(E)$. The objective here is to maximize the average yield as reflected in the quadratic forms $(C(t)\overline{x}(t), \overline{x}(t))_E$ and $(N\overline{x}(T), \overline{x}(T))_E$, and minimize fluctuation around the average yield. Following the necessary conditions of optimality (Theorem 5.3 and Theorem 5.4) one can verify that the adjoint evolution equation is given by

$$-d\varphi = A^*\varphi dt + F^*(t)\overline{\varphi}dt + Q(t)(x-\overline{x}) - C(t)\overline{x},$$
$$\varphi(T) = M(x(T)-\overline{x}(T)) - N\overline{x}(T). \tag{6.17}$$

Again following similar procedure as in examples 1 and 2, the optimal feedback control law is given by

$$u^o = -R^{-1}B^*[K_1(x-\overline{x}) + K_2\overline{x} + r] \tag{6.18}$$

where the operator valued functions $\{K_1, K_2\}$ and the process r are, respectively, the weak and mild solutions of the following system of evolution equations

$$(d/dt)K_1 + (K_1A + A^*K_1) - K_1BR^{-1}B^*K_1 + Q = 0, K_1(T) = M \tag{6.19}$$

$$(d/dt)K_2 + K_2(A+F) + (A+F)^*K_2 - K_2BR^{-1}B^*K_2 - C = 0,$$
$$K_2(T) = -N \tag{6.20}$$

$$dr + (A - BR^{-1}B^*K_1)^* r dt + K_1\sigma dW = 0, \; r(T) = 0, \tag{6.21}$$

on $\mathcal{L}(E)$ and E respectively. Note that if $C \equiv 0$ and $N = 0$, the operator equation (6.20) reduces to a homogeneous equation with only the trivial solution $K_2(t) \equiv 0$. Under these conditions equation (6.20) disappears and we obtain the results of Example 2.

7. ACKNOWLEDGEMENT

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

REFERENCES

[1] 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 Journal on Control and Optimization*, 51(4), (2013), 3235–3257.

[2] N. U. Ahmed, Stochastic Neutral Evolution Equations on Hilbert Spaces with Partially Observed Relaxed Controls and Their Necessary Conditions of Optimality, *Nonlinear Analysis Series A: Theory, Methods & Applications*, vol/issue. 101(2014), 66–79. DOI:10.1016/j.na.2014.01.019.

[3] N. U. Ahmed, Stochastic Neutral Evolution Equations on Hilbert Spaces and Their Partially Observed Optimal Relaxed Control, *Journal of Abstract Differential Equations and Applications* 5(1), (2014), 1–20.

[4] N. U. Ahmed, Some Recent Developments in Systems and Control Theory on Infinite Dimensional Banach Spaces, Part 1 & 2, *Springer Proceedings in Mathematics and Statistics, Optimization and Control Techniques and Applications* 978-3-662-43403-1, 319261-1-En (1). Vol. 86, (Eds. H. Xu, K. L. Teo and Y. Zhang)

[5] N. U. Ahmed,(1988), Optimization and Identification of Systems Governed by Evolution Equations on Banach Spaces, *Pitman research Notes in Mathematics Series*, Vol.184, Longman Scientific and Technical, U. K.; Co-published with John-Wiely & Sons, Inc. New York.

[6] N. U. Ahmed,(1991), Semigroup Theory with Applications to Systema 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.

[7] N. U. Ahmed, Stochastic Evolution Equations on Hilbert Spaces with Partially Observed Relaxed Controls and Their Necessary Conditions of Optimality, *Differential Inclusions, Control and Optimization* 34(2014), 105–129 doi:10.7151/dmdico.1153.

[8] N. U. Ahmed, Nonlinear Diffusion Governed by McKean-Vlasov Equation on Hilbert Space and Optimal Control, *SIAM J. Cotrol and Optim.* 46(1), (2007), 356–378.

[9] N. U. Ahmed and X. Ding, A semilinear McKean-Vlasov stochastic evolution equation in Hilbert space, *Stochastic Process. Appl.*, 60 (1995), 65–85.

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

[11] N. U. Ahmed, Stochastic initial boundary value problems subject to distributed and boundary noise and their optimal control, *Journal of Mathematical Analysis and Applications (JMAA)*, 421, (2015), 157-179, http://dx.doi.org/10.1016/j.jmaa.2014.06.078

[12] L. L. Bonilla, Stable nonequilibrium probability densities and phase transitions for mean-field models in the thermodynamic limit, *J. Stat. Phys.*, 46 (1987), 659–678.

[13] C. Clason, B. Kaltenbacher, and S. Veljovic. Boundary optimal control of the Westervelt and the Kuznetsov equation. *JMAA*, 356 (2009) 738–751, doi:10.1016/j.jmaa.2009.03.043.

[14] G. Da Prato and J. Zabczyk,(1992), Stochastic Equations in Infinite Dimensions, *Cambridge University Press.*

[15] G. Da Prato and J. Zabczyk, Ergodicity for Infinite Dimensional Systems, *London Math. Soc. Lecture Note Ser.* 229, Cambridge University Press, London, 1996.

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

[17] 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.

[18] T. E. Duncan and B. Pasik-Duncan, Linear-Exponential-Quadratic Gaussian Control for Stochastic Equations in a Hilbert Space, *Dynamic Systems and Applications*, Special Issue (Edited by N.U. Ahmed and S. Migorski), 21(2/3), 2012, 407–416.

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

[20] Yang Shen & Tak Kuen Siu, The Maximum Principle for a Jump-diffusion Mean-field Model and its Application to the Mean-variance Problem, *Nonlinear Analysis*, 86(2013), 58–73.

[21] Y. Hu and S. Peng, Adaptive Solution of a Backward Semilinear Stochastic Evolution Equation, *Stochastic Analysis and Applications*, 9(4), (1991), 445–459

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

[23] H. P. McKean, A class of Markov processes associated with nonlinear parabolic equations, *Proc. Natl. Acad. Sci.* USA, 56 (1966), 1907–1911.

[24] A. Ionescu Tulcea and C. Ionescu Tulcea, (1969), Topics in the Teory of Lifting, Springer-Verlag, Berlin, Heidelberg, New York.

[25] A. Pazy, (1983), *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, Berlin, Heidelberg Tokyo.

[26] M. Scheutzow, Periodic behavior of the stochastic Brusselator in the mean-field limit, *Probab. Theory Related Fields*, 72 (1986),425–462.

[27] X. Y. Zhou, On the Necessary Conditions of optimal controls for Stochastic partial Differential Equations . *SIAM J. Control Optim.* 31(6), 1462–1478, 1993.

[28] N. U. Ahmed, Optimal Control of General McKean-Vlasov Stochastic Evolution Equations on Hilbert Spaces and Necessary Conditions of Optimality, Discussiones Mathematicae, Differential Inclusions, Control and Optimization, 35 (2015), 165–195.

[29] N. U. Ahmed, (1988), Elements of Finite Dimensional Systems and Control Theory, Pitman Monographs and Surveys in Pure and Applied Mathematics, Vol. 37, Longman Scientic and Technical, U.K. and Co-published with John Wiely and Sons, New York.

[30] N. U. Ahmed, (2006), Dynamic Systems and Control with Applications, World Scientific Pub. Co. New Jersey, London, Singapore, Beijing, Shanghai, Hong Kong, Taipei, Chennai.