

## STOCHASTIC EVOLUTION EQUATIONS ON HILBERT SPACES WITH PARTIALLY OBSERVED RELAXED CONTROLS AND THEIR NECESSARY CONDITIONS OF OPTIMALITY

N.U. AHMED

*University of Ottawa*

### Abstract

In this paper we consider the question of optimal control for a class of stochastic evolution equations on infinite dimensional Hilbert spaces with controls appearing in both the drift and the diffusion operators. We consider relaxed controls (measure valued random processes) and briefly present some results on the question of existence of mild solutions including their regularity followed by a result on existence of partially observed optimal relaxed controls. Then we develop the necessary conditions of optimality for partially observed relaxed controls. This is the main topic of this paper. Further we present an algorithm for computation of optimal policies followed by a brief discussion on regular versus relaxed controls. The paper is concluded by an example of a non-convex problem which is readily solvable by our approach.

**Keywords:** differential equations, Hilbert spaces, relaxed controls, optimal control, necessary conditions of optimality.

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

### 1. INTRODUCTION

Let  $E$  be a separable Hilbert space and consider the stochastic evolution equation on  $E$  of the form:

$$(1) \quad \begin{aligned} dx &= Axdt + b(t, x, u_t)dt + \sigma(t, x, u_t)dW, x(0) = x_0 \in E, \\ t &\in I \equiv [0, T], T < \infty, \end{aligned}$$

where  $A$  is the infinitesimal generator of a  $C_0$ -semigroup on  $E$ , and  $b : I \times E \times U \rightarrow E$  and  $\sigma : I \times E \times U \rightarrow \mathcal{L}(H, E)$  are Borel measurable maps. Let

$(\Omega, \mathcal{F}, \mathcal{F}_{t \geq 0}, P)$  be a complete probability space with the filtration  $\mathcal{F}_t \subset \mathcal{F}$ . The process  $W = \{W(t), t \geq 0\}$  is an  $\mathcal{F}_t$ -adapted  $H$  valued Brownian motion with incremental covariance operator  $Q \in \mathcal{L}_1(H)$ , the space of nuclear operators on  $H$ . In case of cylindrical Wiener process  $Q = I$  (identity operator) and clearly it is not nuclear. The set  $U$  is a compact Polish space. Let  $M_0(U)$  denote the class of regular probability measures on  $\mathcal{B}(U)$ , the sigma algebra of Borel subsets of the set  $U$ . Let  $L_\infty^\alpha(I, M_0(U))$  denote the class of weak star measurable  $\mathcal{G}_t \subset (\mathcal{F}_t)$ -adapted random processes with values in the space of probability measures  $M_0(U)$ . We consider  $\mathcal{U}_{ad} \equiv L_\infty^\alpha(I, M_0(U))$  as the set of admissible controls. The problem is to find a control that minimizes the cost functional

$$(2) \quad J(u) \equiv \mathbf{E} \left\{ \int_I \ell(t, x(t), u_t) dt + \Phi(x(T)) \right\},$$

where  $\ell : I \times E \times U \longrightarrow \bar{R}$  is a Borel measurable (extended) real valued function and  $\Phi : E \longrightarrow \bar{R}$  is a Borel measurable map. Precise assumptions will be given later.

For recent advances in optimal control theory for infinite dimensional systems the reader is referred to [5] and [6, Chapter 3] and the references therein. There is an abundance of literature on stochastic minimum principle for finite dimensional problems [9, 10, 15], see also the extensive references therein. In [9] we considered the question of existence of optimal relaxed controls for finite dimensional partially observed stochastic systems driven by continuous as well as jump diffusion processes. Further, we presented also the necessary conditions of optimality. In [10] the author proves maximum principle of Pontryagin type for finite dimensional neutral stochastic systems. But to the best of knowledge of the author, there is hardly any literature on relaxed controls for partially observed stochastic systems on infinite dimensional Hilbert spaces. In a recent paper [2] we proved the existence of optimal relaxed controls of many standard and non-standard control problems for a class of neutral systems. But there are not much activities on relaxed control for partially observed infinite dimensional systems, in particular their necessary conditions of optimality. In this paper we wish to present such necessary conditions of optimality.

The rest of the paper is organized as follows. In Section 2, we introduce some basic background materials and notations. In Sections 3–5 we present briefly some important results on the questions of existence and regularity of mild solutions, continuous dependence of solutions on relaxed controls endowed with the weak star topology, and existence of optimal relaxed controls. The major concern of this paper is to develop the necessary conditions of optimality using relaxed controls. This is done in Section 6. In Subsection 6.1, we present a conceptual computational algorithm based on these necessary conditions. We have also given the proof of convergence of the algorithm. In Section 7, we compare

regular controls with relaxed ones and discuss the impact of these controls on the questions of existence of optimal controls and necessary conditions of optimality. We conclude the paper with an example where relaxed control is the only choice.

## 2. BASIC MATERIALS AND NOTATIONS

All random processes mentioned below are based on the complete filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_{t \geq 0}, P)$  without further notice. Let  $B_\infty^a(I, E) \subset L_{\infty,2}^a(I \times \Omega, E) = L_\infty^a(I, L_2(\Omega, E))$  denote the space of  $\mathcal{F}_t$ -adapted random processes with values in the Hilbert space  $E$  satisfying

$$\sup \{ \mathbf{E} |x(t)|_E^2, t \in I \} < \infty \}.$$

The space  $B_\infty^a(I, E)$ , equipped with the norm topology  $\sup \{ \sqrt{\mathbf{E} |x(t)|_E^2}, t \in I \}$ , is a closed subspace of the Banach space  $L_{\infty,2}^a(I \times \Omega, E)$  and hence itself a Banach space.

As regards control, first let us consider the space  $M(U)$  of signed Borel measures defined on the Borel field of subsets of the set  $U$ , where  $U$  is a compact Polish space. Let  $L_\infty^\alpha(I, M(U))$  denote the space of  $\mathcal{G}_t$  adapted random processes with values in the space of signed measures  $M(U)$ . With respect to the norm topology, given by  $ess - \sup \{ |u_t|_{\tau_v}, (t, \omega) \in I \times \Omega \}$ , this is a Banach space. We are not interested in the norm topology. For controls, we consider weak star topology on  $L_\infty^\alpha(I, M(U))$  and for admissible controls we choose an appropriate subset  $\mathcal{U}_{ad}$  (described later) of the set  $L_\infty^\alpha(I, M_0(U))$  where  $M_0(U) \subset M(U)$  denotes the space of regular probability measures defined on the Borel subsets of  $U$ . This space is endowed with the weak star topology denoted by  $\tau_w$ . It follows from Alaoglu's theorem that it is a weak star compact subset of  $L_\infty^\alpha(I, M(U))$ . Let  $(D, \geq)$  be a directed index set. With respect to the weak star topology, a net (or a generalized sequence)  $u^\gamma \xrightarrow{\tau_w} u^o$  ( $\gamma \in D$ ) if and only if for every  $\varphi \in L_1^\alpha(I, C(U))$

$$\mathbf{E} \int_I \int_U \varphi(t, \xi) u_t^\gamma(d\xi) dt \longrightarrow \mathbf{E} \int_I \int_U \varphi(t, \xi) u_t^o(d\xi) dt.$$

Let  $E_1$  and  $E_2$  be a pair of Banach spaces. The space of linear operators from  $E_1$  to  $E_2$  is denoted by  $L(E_1, E_2)$  and the space of bounded linear operators by  $\mathcal{L}(E_1, E_2)$ . In case  $E_1 = E_2 = E$ ,  $\mathcal{L}(E)$  denotes the space of bounded linear operators in  $E$ . The space of nuclear operators is denoted by  $\mathcal{L}_1(E)$ . In case  $E$  is a Hilbert space, the class of positive self adjoint operators in  $E$  is denoted by  $\mathcal{L}_s^+(E)$ . For  $Q \in \mathcal{L}_1^+(H)$  let us denote by  $L_{2,Q}^a(I, L(H, E))$  the Hilbert space of  $\mathcal{F}_t$  adapted  $L(H, E)$ -valued random processes with the scalar product given by

$\mathbf{E} \int_I \text{Tr}(\sigma_1(t) Q \sigma_2^*(t)) dt$ . The associated norm is given by the square root of

$$\mathbf{E} \int_I \text{Tr}(\sigma(t) Q \sigma^*(t)) dt \equiv \mathbf{E} \int_I |\sigma(t)|_Q^2 dt.$$

Completion of  $L_{2,Q}^a(I, L(H, E))$  with respect to the scalar product (and the corresponding norm) is a Hilbert space which is again denoted by  $L_{2,Q}^a(I, L(H, E))$  for convenience. Note that this is a larger space than  $\mathcal{L}_{2,Q}^a(I, L(H, E))$ . Further notations will be introduced when required.

### 3. EXISTENCE OF MILD SOLUTIONS AND THEIR REGULARITY

Before we can consider the control problem, we must study the question of existence, uniqueness and regularity properties of solutions. Let us consider the evolution equation (1). Here we have used the notation

$$\begin{aligned} b(t, x, u_t) &\equiv \int_U b(t, x, \xi) u_t(d\xi) \\ \sigma(t, x, u_t) &\equiv \int_U \sigma(t, x, \xi) u_t(d\xi). \end{aligned}$$

Letting  $S(t), t \geq 0$ , denote the semigroup generated by  $A$ , we can formally use Dhumels formula and integration by parts to convert the evolution equation (1) into the following integral equation

$$\begin{aligned} (3) \quad x(t) &= S(t)x_0 + \int_0^t S(t-r)b(r, x(r), u_r)dr \\ &\quad + \int_0^t S(t-r)\sigma(r, x(r), u_r) dW(r). \end{aligned}$$

Throughout the rest of the paper, we let  $L_2^0(\Omega, E)$  denote the class of  $\mathcal{F}_0$  measurable  $E$ -valued random variables having finite second moment. That is, for  $x_0 \in L_2^0(\Omega, E)$ , we have  $\mathbf{E}|x_0|_E^2 \equiv \int_\Omega |x_0|_E^2 P(d\omega) < \infty$ .

**Theorem 3.1.** *Suppose  $A$  is the infinitesimal generator of a  $C_0$ -semigroup  $S(t), t \geq 0$ , on  $E$  satisfying  $\sup\{\|S(t)\|_{\mathcal{L}(E)}, t \in I\} \leq M$ . The drift and the diffusion operators satisfy the following assumptions:  $b : I \times E \times U \rightarrow E$  and  $\sigma : I \times E \times U \rightarrow \mathcal{L}(H, E)$  are Borel measurable maps satisfying the following properties:*

(A1): *there exists a constant  $K > 0$  such that for all  $(t, x, \xi) \in I \times E \times U$*

$$|b(t, x, \xi) - b(t, z, \xi)|_E^2 \leq K^2|x - z|_E^2, \quad |b(t, x, \xi)|^2 \leq K^2(1 + |x|_E^2).$$

(A2): there exists a constant  $K_Q > 0$  such that for all  $(t, x, \xi) \in I \times E \times U$

$$\begin{aligned} |\sigma(t, x, \xi) - \sigma(t, z, \xi)|_Q^2 &\equiv \text{Tr}([\sigma(t, x, \xi) - \sigma(t, z, \xi)]Q[\sigma(t, x, \xi) - \sigma(t, z, \xi)]^*) \\ &\leq K_Q^2 |x - z|_E^2, \\ |\sigma(t, x, \xi)|_Q^2 &\leq K_Q^2 (1 + |x|_E^2). \end{aligned}$$

The incremental covariance operator  $Q$  of the Wiener process  $W$  is either an identity operator in  $H$  (in which case  $W$  is a cylindrical Brownian motion) or nuclear in which case  $Q \in \mathcal{L}_1(H)$ . Then, for each  $x_0 \in L_2^0(\Omega, E)$  and control  $u \in \mathcal{U}_{ad} \equiv L_\infty^\alpha(I, M_0(U))$ , the integral equation (3) has a unique solution  $x \in B_\infty^a(I, E)$  and further it has a continuous modification.

**Proof.** For any fixed control  $u \in \mathcal{U}_{ad}$ , define the operator  $G$  by

$$(Gx)(t) \equiv S(t)x_0 + \int_0^t S(t-r)b(r, x(r), u_r)dr + \int_0^t S(t-r)\sigma(r, x(r), u_r)dW(r),$$

for  $t \in I$ . Under the given assumptions it is easy to verify that

$$G : B_\infty^a(I, E) \longrightarrow B_\infty^a(I, E).$$

Again under the given assumptions one can easily verify that

$$\begin{aligned} &\mathbf{E}|(Gx)(t) - (Gy)(t)|_E^2 \\ &\leq (KM)^2 t \int_0^t \mathbf{E}|x(r) - y(r)|_E^2 dr + (K_Q M)^2 \int_0^t \mathbf{E}|x(r) - y(r)|_E^2 dr. \end{aligned}$$

For any  $t \in I$  define

$$\|x - y\|_t \equiv \sqrt{\sup_{0 \leq s \leq t} |x(s) - y(s)|_E^2}.$$

Then it follows from the previous inequality that

$$\|Gx - Gy\|_t \leq \left( \sqrt{(KM)^2 t^2 + (K_Q M)^2 t} \right) \|x - y\|_t.$$

Thus for  $T_1 \in [0, T]$  sufficiently small, the operator  $G$  is a contraction on  $B_\infty^a([0, T_1], E)$ , and thus by Banach fixed point theorem, it has a unique fixed point. Since  $I$  is a compact interval, it can be covered by a finite number of closed intervals such as  $\{[0, T_1], [T_1, 2T_1], \dots\}$ . Hence the solution can be continued to cover the whole interval and hence for each control  $u \in \mathcal{U}_{ad}$  the integral equation (3) has a unique solution  $x \in B_\infty^a(I, E)$  which is the mild solution of the evolution equation (1). This completes the proof. ■

In view of the above result we have the following corollary.

**Corollary 3.2.** *Under the assumptions of Theorem 3.1, the solution set  $\{x^u, u \in \mathcal{U}_{ad}\}$  is a bounded subset of the Banach space  $B_\infty^a(I, E)$ .*

**Proof.** Let  $x^u \in B_\infty^a(I, E)$  denote the solution of the integral equation (3). Then taking the expected value of the norm square on either side and using the assumptions of Theorem 3.1, it is easy to verify that

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

where

$$C_1 = 2^2 \{M^2 \mathbf{E}|x_0|^2 + (MKT)^2 + (MK_Q)^2 T\}$$

and  $C_2 = (MK)^2 T + (MK_Q)^2$ . Since the constants  $C_1$  and  $C_2$  hold independently of the controls, the proof follows from Gronwall inequality.  $\blacksquare$

#### 4. CONTINUOUS DEPENDENCE OF SOLUTIONS ON CONTROLS

In the study of optimal control we need the continuity of the map  $u \rightarrow x$ , that is, control to solution map. This is crucial in 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 already the norm (or metric) topology on  $B_\infty^a(I, E)$  as seen in Section 3. So we must consider the control space. Let  $\mathcal{G}_t, t \geq 0$ , denote a current of complete sub-sigma algebras of the current of sigma algebras  $\mathcal{F}_t, t \geq 0$ . Let  $U$  be a compact (not necessarily convex) subset of a Polish space and  $C(U)$  the space of real valued continuous functions defined on  $U$ . Recall that for admissible controls we have chosen the so called relaxed controls (probability measure valued processes) which are  $\mathcal{G}_t$ -adapted in the sense that for any  $\varphi \in C(U)$

$$t \rightarrow u_t(\varphi) \equiv \int_U \varphi(\xi) u_t(d\xi)$$

is a real valued  $\mathcal{G}_t$ -adapted random process  $P$ -a.s essentially bounded on  $I$ . We have denoted this space by  $L_\infty^\alpha(I, M_0(U))$  and endowed this with the weak star topology. Let  $(D, \geq)$  denote any directed set. Then a net (a generalized sequence)  $\{u^\gamma, \gamma \in D\}$  is said to converge to  $u^o$  in this topology if, for each  $\psi \in L_1^\alpha(I, C(U))$ ,

$$\mathbf{E} \int_{I \times U} \psi(t, \xi) u_t^\gamma(d\xi) dt \rightarrow \mathbf{E} \int_{I \times U} \psi(t, \xi) u_t^o(d\xi) dt.$$

This topology is called the vague or weak star topology and we denote this topology by  $\tau_w$  and the above convergence by simply writing  $u^\gamma \xrightarrow{\tau_w} u^o$ . Since  $U$  is a compact polish space (a complete separable metric space), it is well known that  $M_0(U)$  is compact [13]. Thus it follows from the well known Alaoglu's theorem that  $L_\infty^\alpha(I, M_0(U))$  is a weak star compact subset of  $L_\infty^\alpha(I, M(U))$ , the dual of the Banach space  $L_1^\alpha(I, C(U))$ . For admissible controls this topology is rather too weak; we consider a slightly stronger topology.

Consider the measure space  $(I \times \Omega, \mathcal{B}(I) \times \mathcal{F}, \lambda(dt) \times P(d\omega))$  where  $\lambda$  denotes the Lebesgue measure. Let  $\mathcal{P}$  denote the sigma algebra of predictable subsets of the set  $I \times \Omega$  with respect to the filtration  $\mathcal{G}_{t \geq 0} \subseteq \mathcal{F}_{t \geq 0}$ . Let  $\mu = \mu(dt \times d\omega)$  denote the restriction of the product measure  $\lambda(dt) \times P(d\omega)$  to the predictable sigma field  $\mathcal{P}$ . Recall that  $L_\infty^\alpha(I, M_0(U))$  denotes the probability measure valued random processes adapted to the filtration  $\mathcal{G}_{t \geq 0}$ . Consider the measure space  $(I \times \Omega, \mathcal{P}, \mu)$  and introduce the following topological space  $\Xi((I \times \Omega, \mathcal{P}, \mu), M_0(U))$  of weak star  $\mathcal{P}$  measurable  $M_0(U)$  valued random processes. We can introduce a suitable metric topology on this space and turn this into a complete metric space as follows.

**Lemma 4.1.** *The space  $\Xi((I \times \Omega, \mathcal{P}, \mu), M_0(U))$  is metrizable with a metric  $d$  given by*

$$d(u, v) \equiv \sum_{n=1}^{\infty} (1/2^n) \frac{\mu\{(t, \omega) \in I \times \Omega : u_{t,\omega}(\varphi_n) \neq v_{t,\omega}(\varphi_n)\}}{1 + \mu\{(t, \omega) \in I \times \Omega : u_{t,\omega}(\varphi_n) \neq v_{t,\omega}(\varphi_n)\}},$$

where  $\{\varphi_n, n \in N\} \subset C(U)$  is a dense set. With respect to this metric topology it is a complete metric space. We denote this metric space by  $(\mathcal{M}, d)$ .

**Proof.** By assumption,  $U$  is a compact polish space and therefore the Banach space  $C(U)$ , furnished with usual norm topology, is separable. Let  $\{\varphi_n\}$  be any countable set dense in the closed unit ball of the  $B$ -space  $C(U)$  and  $\mu$  the measure defined on the predictable sigma algebra  $\mathcal{P}$  as introduced above. The reader can easily verify that  $d$  as given above defines a metric on the topological space  $\Xi((I \times \Omega, \mathcal{P}, \mu), M_0(U))$ . Completeness of this metric topology follows from that of the topology of almost everywhere convergence from classical measure theory. This completes the proof. ■

Now we are prepared to consider the question of continuity of the control to solution map. Let  $\mathcal{U}_{ad}$  be a compact subset of the metric space  $(\mathcal{M}, d)$  and consider this as the class of admissible controls for the rest of the paper. We shall denote this metric topology by  $\tau_{w\mu}$  in order to emphasize that this metric is equivalent to the weak star convergence in  $M_0(U)$  for  $\mu - a.e (t, \omega) \in I \times \Omega$ .

**Theorem 4.3.** *Consider the system (1) (or equivalently (3)) driven by the control  $u \in \mathcal{U}_{ad}$ , where  $\mathcal{U}_{ad}$  is a  $\tau_{w\mu}$  compact subset of  $(\mathcal{M}, d)$  and suppose the assumptions of Theorem 3.1 hold. Then the control to solution map  $u \rightarrow x$  is continuous with respect to the relative  $\tau_{w\mu}$  topology on  $\mathcal{U}_{ad}$  and the strong (norm) topology on  $B_\infty^a(I, E)$ .*

**Proof.** Let  $u^\alpha, u^o \in \mathcal{U}_{ad}$  and suppose  $u^\alpha \xrightarrow{\tau_{w\mu}} u^o$ . Let  $\{x^\alpha, x^o\} \in B_\infty^a(I, E)$  denote the corresponding solutions of the integral equation (3). Then it is easy to verify that

$$(4) \quad \begin{aligned} x^o(t) - x^\alpha(t) &= e_1^\alpha(t) + e_2^\alpha(t) + \int_0^t S(t-r)(b(r, x^o(r), u_r^\alpha) - b(r, x^\alpha(r), u_r^\alpha))dr \\ &\quad + \int_0^t S(t-r)(\sigma(r, x^o(r), u_r^\alpha) - \sigma(r, x^\alpha(r), u_r^\alpha))dW(r) \end{aligned}$$

where

$$(5) \quad e_1^\alpha(t) \equiv \int_0^t S(t-r)(b(r, x^o(r), u_r^o) - b(r, x^o(r), u_r^\alpha))dr$$

$$(6) \quad e_2^\alpha(t) \equiv \int_0^t S(t-r)(\sigma(r, x^o(r), u_r^o) - \sigma(r, x^o(r), u_r^\alpha))dW(r).$$

Using the Lipschitz properties (A1) and (A2) and taking the expected value of the norm square, it follows from (4) that

$$(7) \quad \mathbf{E}|x^o(t) - x^\alpha(t)|_E^2 \leq 2^3 \left\{ \mathbf{E}|e_1^\alpha(t)|_E^2 + \mathbf{E}|e_2^\alpha(t)|_E^2 + C \int_0^t \mathbf{E}|x^o(r) - x^\alpha(r)|_E^2 dr \right\}$$

where  $C = (KM)^2T + (K_QM)^2$ . Computing  $\mathbf{E}|e_1^\alpha(t)|_E^2$  and  $\mathbf{E}|e_2^\alpha(t)|_E^2$  we obtain

$$(8) \quad \mathbf{E}|e_1^\alpha(t)|_E^2 \leq TM^2 \mathbf{E} \int_0^T |b(r, x^o(r), u_r^o) - b(r, x^o(r), u_r^\alpha)|_E^2 dr, t \in I.$$

$$(9) \quad \mathbf{E}|e_2^\alpha(t)|_E^2 \leq (M)^2 \mathbf{E} \int_0^T |\sigma(r, x^o(r), u_r^o) - \sigma(r, x^o(r), u_r^\alpha)|_Q^2 dr, t \in I.$$

It follows from the assumptions (A1) and (A2), in particular the growth conditions, that the integrands in (8) and (9) are dominated by integrable processes (functions). Since  $u^\alpha \xrightarrow{\tau_{w\mu}} u^o$ , the integrands also converge to zero  $\mu$ -a.e. Thus it follows from Lebesgue dominated convergence theorem that

$$h_\alpha(t) \equiv \mathbf{E}|e_1^\alpha(t)|_E^2 + \mathbf{E}|e_2^\alpha(t)|_E^2 \rightarrow 0$$

uniformly on  $I$ . Now it follows from Gronwall inequality applied to (7) that

$$(10) \quad \mathbf{E}|x^o(t) - x^\alpha(t)|_E^2 \leq 2^3 h_\alpha(t) + 2^3 C \int_0^t \exp\{2^3 C(t-s)\} h_\alpha(s) ds, t \in I.$$

From this inequality it readily follows that  $\sup\{\mathbf{E}|x^o(t) - x^\alpha(t)|_E, t \in I\} \rightarrow 0$ . Thus we have proved that as  $u^\alpha \xrightarrow{\tau_w \mu} u^o$  in  $\mathcal{U}_{ad}$ ,  $x^\alpha \xrightarrow{s} x^o$  in the Banach space  $B_\infty^a(I, E)$ . This proves the continuity as stated. ■

## 5. EXISTENCE OF OPTIMAL CONTROLS

In this section we consider a typical control problem and present a result on existence of optimal relaxed controls. We consider the Bolza problem,

$$(11) \quad J(u) \equiv \mathbf{E} \left\{ \int_0^T \ell(t, x(t), u_t) dt + \Phi(x(T)) \right\},$$

where  $u \in \mathcal{U}_{ad} \subset (\mathcal{M}, d)$  and  $x = x^u \in B_\infty^a(I, E)$  is the mild solution of the evolution equation (1) or equivalently the solution of the corresponding functional equation (3). As usual in this paper we define  $\ell(t, e, u) \equiv \int_U \ell(t, e, \xi) u(d\xi)$  for  $u \in M_0(U)$ . Our objective is to find a control  $u \in \mathcal{U}_{ad}$  that minimizes the functional (11) subject to the dynamic constraint (3).

**Theorem 5.1.** *Consider the system (1) driven by relaxed controls  $\mathcal{U}_{ad}$  with the cost functional (11) and suppose the assumptions of Theorem 4.3 hold. Further, suppose that*

(a1):  $\ell(t, e, \xi)$  is measurable in  $t \in I$ , continuous in  $e \in E$  uniformly with respect to  $\xi \in U$ , and continuous in  $\xi \in U$  for each  $(t, e) \in I \times E$  satisfying

$$|\ell(t, e, \xi)| \leq \ell_0(t) + \alpha_1 |e|_E^2, \text{ for some } \ell_0 \in L_1^+(I), \text{ and } \alpha_1 \geq 0$$

for all  $(t, e, \xi) \in I \times E \times U$ ; and

(a2):  $\Phi$  is lower semicontinuous on  $E$  satisfying for all  $e \in E$

$$|\Phi(e)| \leq \alpha_2 + \alpha_3 |e|_E^2 \text{ for some } \alpha_2, \alpha_3 \geq 0.$$

Then there exists an optimal control.

**Proof.** Using the continuity result of Theorem 4.3 and the assumptions on  $\ell$  and Fatou's Lemma one can verify that the functional

$$u \rightarrow J_1(u) \equiv \mathbf{E} \int_I \ell(t, x^u(t), u_t) dt$$

is lower semi-continuous on  $\mathcal{U}_{ad}$  with respect to the  $\tau_{w\mu}$  topology. By Theorem 3.1, the solutions of the integral equation (3) have continuous modifications. Thus the functional  $J_2(u) \equiv \mathbf{E}\Phi(x^u(T))$  is well defined. Then using the assumption (a2) and Fatou's Lemma once again one can verify that the functional  $u \rightarrow J_2(u)$  is also lower semi-continuous with respect to the  $\tau_{w\mu}$  topology. Hence the functional given by the sum,  $u \rightarrow J(u)$ , is lower semicontinuous on  $\mathcal{U}_{ad}$ . Since  $\mathcal{U}_{ad}$  is  $\tau_{w\mu}$  compact and  $J$  is lower semicontinuous on it with respect to this topology,  $J$  attains a minimum on  $\mathcal{U}_{ad}$ . Hence an optimal control exists. ■

In the next section we are mainly interested in developing necessary conditions of optimality for the Bolza problem (11) as stated above.

## 6. NECESSARY CONDITIONS OF OPTIMALITY

One of the main difficulties in developing necessary conditions of optimality for infinite dimensional systems arises from the presence of the unbounded operator  $A$  in the evolution equation (1). However, by use of the Yosida approximation of  $A$  giving a sequence  $\{A_n\}$  of bounded operators converging to  $A$  on  $D(A) \subset E$  in the strong operator topology, we can prove necessary conditions of optimality. Let  $\rho(A)$  denote the resolvent set of the operator  $A$  and  $R(\lambda, A)$  denote the resolvent of  $A$  corresponding to  $\lambda \in \rho(A)$ . Then define

$$A_n \equiv nAR(n, A), \text{ for } n \in \rho(A) \cap \mathbb{N}.$$

This is the Yosida approximation of  $A$  and it is well known from semigroup theory, see Ahmed [7], that  $\{A_n\}$  is a sequence of bounded operators in  $E$  and that it converges to  $A$  on  $D(A)$  in the strong operator topology, that is, for each  $x \in D(A)$ ,  $A_n x \rightarrow Ax$  strongly in  $E$ . For proof see [7, Lemma 2.2.5, p. 225]. Further, it is also well known from semigroup theory [7, Remark 2.2.8 and Corollary 2.2.9, p. 30] that the semigroups  $\{S_n(t), t \geq 0\}$ , generated by the sequence of operators  $\{A_n\}$ , converge in the strong operator topology in  $\mathcal{L}(E)$  to the semigroup  $S(t), t \geq 0$ , generated by the operator  $A$ . This convergence is uniform in  $t$  on any closed bounded interval  $I \subset \mathbb{R}$ . These results are valid not only for contraction semigroups but also for general  $C_0$ -semigroups on Banach spaces.

**Lemma 6.1.** *Consider the control system (1) with the initial state  $x_0 \in L_2^0(\Omega, E)$  and any control  $u \in \mathcal{U}_{ad}$  fixed, and suppose the basic assumptions of Theorem 3.1 hold. Let  $S(t), t \geq 0$ , denote the semigroup corresponding to the generator  $A$  and  $x^o \in B_\infty^a(I, E)$  denote the corresponding mild solution of (1). Let  $x^n$  denote the mild solution of equation (1) with the unbounded operator  $A$  replaced by its*

*Yosida approximation*  $A_n$  generating the sequence of semigroups  $\{S_n(t), t \geq 0\}$ . Then  $x^n \xrightarrow{s} x^o$  in  $B_\infty^a(I, E)$ .

**Proof.** Using the expression (3), it follows from Theorem 3.1 that  $\{x^o, x^n\}$  satisfy the following integral equations

$$(12) \quad \begin{aligned} x^o(t) = & S(t)x_0 + \int_0^t S(t-r)b(r, x^o(r), u_r)dr \\ & + \int_0^t S(t-r)\sigma(r, x^o(r), u_r) dW(r). \end{aligned}$$

$$(13) \quad \begin{aligned} x^n(t) = & S_n(t)x_0 + \int_0^t S_n(t-r)b(r, x^n(r), u_r)dr \\ & + \int_0^t S_n(t-r)\sigma(r, x^n(r), u_r) dW(r). \end{aligned}$$

Subtracting (13) from (12) term by term we have on the left-hand side

$$(14) \quad L \equiv (x^o(t) - x^n(t))$$

and on the right-hand side we have several terms grouped as shown below

$$(15) \quad R1 \equiv (S(t) - S_n(t))x_0$$

$$(16) \quad \begin{aligned} R2 \equiv & \int_0^t S(t-s)b(s, x^o(s), u_s) - \int_0^t S_n(t-s)b(s, x^n(s), u_s)ds \\ & = \int_0^t (S(t-s) - S_n(t-s))b(s, x^o(s), u_s)ds \\ & + \int_0^t S_n(t-s)(b(s, x^o(s), u_s) - b(s, x^n(s), u_s))ds. \end{aligned}$$

$$(17) \quad \begin{aligned} R3 \equiv & \int_0^t \{S(t-s)\sigma(s, x^o(s), u_s) - S_n(t-s)\sigma(s, x^n(s), u_s)\}dW(s) \\ & = \int_0^t (S(t-s) - S_n(t-s))\sigma(s, x^o(s), u_s)dW(s) \end{aligned}$$

$$(18) \quad + \int_0^t S_n(t-s)(\sigma(s, x^o(s), u_s) - \sigma(s, x^n(s), u_s))dW(s).$$

Using the expressions (14)–(18), we derive the following estimate

$$(19) \quad \mathbf{E}|x^o(t) - x^n(t)|_E^2 \leq h_n(t) + C \int_0^t \mathbf{E}|x^o(s) - x^n(s)|_E^2 ds$$

where the function  $h_n$  is given by

$$(20) \quad \begin{aligned} h_n(t) \equiv & \mathbf{E}|(S(t) - S_n(t))x_0|_E^2 + \mathbf{E} \int_0^t |(S(t-s) - S_n(t-s))b(s, x^o(s), u_s)|_E^2 ds \\ & + \mathbf{E} \int_0^t |(S(t-s) - S_n(t-s))\sigma(s, x^o(s), u_s)|_Q^2 ds. \end{aligned}$$

The constant  $C$  in the expression (19) is dependent on the basic parameters of Theorem 3.1 such as  $\{M, T, K, K_Q\}$ . Recall that  $S_n(t) \xrightarrow{\tau_{so}} S(t)$  uniformly on  $I$ , and  $x_0 \in L_2^0(\Omega, E)$ , and further, it follows from the basic properties of  $\{b, \sigma\}$  as stated in Theorem 3.1 that  $b(\cdot, x^o(\cdot), u_\cdot) \in L_2^a(I, E)$  and  $\sigma(\cdot, x^o(\cdot), u_\cdot) \in L_{2,Q}^a(I, L(H, E))$ . Hence the reader can easily verify that all the integrands in the expression (20) converge to zero  $dt \times dP$  a.e. and further they are all dominated by integrable functions. Thus it follows from Lebesgue dominated convergence theorem that  $h_n(t) \rightarrow 0$  uniformly on  $I$  and hence it follows from generalized Gronwall lemma applied to the inequality (19) that

$$(21) \quad \lim_{n \rightarrow \infty} \sup \{ \mathbf{E}|x^o(t) - x^n(t)|_E^2, t \in I \} = 0.$$

In other words we have shown that  $x^n \xrightarrow{s} x^o$  in the Banach space  $B_\infty^a(I, E)$  whenever  $A_n \xrightarrow{\tau_{so}} A$  on  $D(A)$ . This completes the proof. ■

We make crucial use of the Lemma 6.1 for developing the necessary conditions of optimality. Even though we may write as differential equations, we always mean the mild solutions thereof, that is the solutions of the corresponding integral equations. Another way is to replace  $A$  by its Yosida approximation  $A_n$  and carry out all the derivations of the necessary conditions of optimality and then take the limit. This gives rise to cumbersome notations with suffix  $n$  for every symbol. However this works exactly because of Lemma 6.1. We can avoid this by simply using the Lemma 6.1 and recalling that we always mean mild solutions.

**Theorem 6.2.** *Consider the system (1) and suppose the basic assumptions of Theorem 3.1, Theorem 4.3 and Lemma 6.1 hold. Further, suppose that the first Gateaux derivatives of  $b$  and  $\sigma$  with respect to  $x \in E$  are uniformly bounded on  $I \times E \times U$  and linear in any direction  $\eta \in E$ . The cost integrands  $\ell$  and  $\Phi$  satisfy the following properties*

$$|\ell_x(t, e, u)|_E \leq \beta_1(t) + \beta_2|e| \quad \forall e \in E,$$

$$|\Phi_x(e)|_E \leq \beta_3 + \beta_4|e|_E, \quad \forall e \in E$$

where  $\beta_1(\geq 0) \in L_1^+(I)$  and  $\{\beta_2, \beta_3, \beta_4 \geq 0\}$ . Then, the necessary conditions for a control  $u^o \in \mathcal{U}_{ad}$  and the corresponding solution  $x^o \in B_\infty^a(I, E)$  of (1) to be optimal, it is necessary that there exists a pair  $\{\psi, \mathcal{R}\} \in B_\infty^a(I, E) \times L_{2,Q}^a(I, \mathcal{L}(H, E))$  such that the following equation

$$(22) \quad dx^o(t) = Ax^o(t)dt + b(t, x^o(t), u_t^o)dt + \sigma(t, x^o(t), u_t^o)dW(t), x^o(0) = x_0, t \in I;$$

and the inequality

$$(23) \quad \begin{aligned} dJ(u^o; u - u^o) = \mathbf{E} \left\{ \int_0^T \left( \langle \psi(t), b(t, x^o(t), u_t - u_t^o) \rangle_E \right. \right. \\ \left. \left. + \text{Tr}(\mathcal{R}(t)Q\sigma^*(t, x^o(t), u_t - u_t^o)) \right. \right. \\ \left. \left. + \ell(t, x^o(t), u_t - u_t^o) \right) dt \right\} \geq 0, \quad \forall u \in \mathcal{U}_{ad} \end{aligned}$$

hold.

**Proof.** Let  $u^o \in \mathcal{U}_{ad}$  denote the optimal control and  $u \in \mathcal{U}_{ad}$  any other control. Since  $M_0(U)$  is the space of probability measures, the set  $\mathcal{U}_{ad}$  is convex. Thus for any  $\varepsilon \in [0, 1]$ , it follows from convexity of the set  $\mathcal{U}_{ad}$  that  $u^\varepsilon \equiv u^o + \varepsilon(u - u^o) \in \mathcal{U}_{ad}$ . Hence  $J(u^\varepsilon) \geq J(u^o)$  for all  $\varepsilon \in [0, 1]$  and any  $u \in \mathcal{U}_{ad}$ . By definition, the Gateaux differential of  $J$  at  $u^o$  in the direction  $u - u^o$  is given by

$$dJ(u^o; u - u^o) = \lim_{\varepsilon \downarrow 0} (1/\varepsilon)(J(u^\varepsilon) - J(u^o)).$$

So by optimality of  $u^o$  we have

$$(24) \quad dJ(u^o; u - u^o) \geq 0 \quad \forall u \in \mathcal{U}_{ad}.$$

Let  $x^\varepsilon \in B_\infty^a(I, E)$  denote the solution of equation (3) corresponding to the control  $u^\varepsilon$  while  $x^o \in B_\infty^a(I, E)$  is the solution of (3) corresponding to the control  $u^o$ . Using the expression (2) for  $J$  and the expression (24), for its Gateaux derivative, and the differentiability assumptions on  $\ell$  and  $\Phi$ , it is easy to verify that

$$(25) \quad \begin{aligned} dJ(u^o; u - u^o) = \mathbf{E} \left\{ \int_0^T \langle \ell_x(t, x^o(t), u_t^o), z(t) \rangle_E dt + \langle \Phi_x(x^o(T)), z(T) \rangle_E \right. \\ \left. + \int_0^T \ell(t, x^o(t), u_t - u_t^o) dt \right\} \geq 0 \quad \forall u \in \mathcal{U}_{ad} \end{aligned}$$

where the process  $z$  is given by the limit

$$(26) \quad z(t) = \lim_{\varepsilon \downarrow 0} (1/\varepsilon)(x^\varepsilon(t) - x^o(t)), \quad t \in I.$$

The process  $z$  is the unique mild solution of the stochastic variational equation

$$(27) \quad \begin{aligned} dz &= Azdt + b_x(t, x^o(t), u_t^o)z(t)dt + \sigma_x(t, x^o(t), u_t^o; z(t))dW + dM_{u-u^o}^o(t) \\ z(0) &= z_0 = 0 \end{aligned}$$

driven by the control dependent semi martingale  $\{M_{u-u^o}^o(t), t \in I\}$ . This semi martingale is governed by the following stochastic differential expression,

$$(28) \quad dM_v^o(t) = b(t, x^o(t), v_t)dt + \sigma(t, x^o(t), v_t)dW,$$

for  $v \in L_\infty^\alpha(I, M(U))$ . These are easily proved by following the definition (26) using the integral equation (3) corresponding to the controls  $u^\varepsilon$  and  $u^o$  respectively and Lebesgue dominated convergence theorem. The reader may like to carry out the details. Note that by our assumption the Gateaux derivative of  $\sigma$  in  $x \in E$  is uniformly bounded on  $I \times E \times M_0(U)$  and directionally linear in the sense that  $\eta \rightarrow \sigma_x(t, x^o(t), u_t^o; \eta)$  is a linear map from  $E$  to  $\mathcal{L}(H, E)$  and that there exists a constant  $c > 0$  such that P-a.s

$$\|\sigma_x(t, x^o(t), u_t^o; \eta)\|_{\mathcal{L}(H, E)} \leq c|\eta|_E, \forall t \in I.$$

Consider the functional

$$(29) \quad L(z) \equiv \mathbf{E} \left\{ \int_0^T \langle \ell_x(t, x^o(t), u_t^o), z(t) \rangle_E dt + \langle \Phi_x(x^o(T)), z(T) \rangle_E \right\}.$$

It follows from our assumption on  $\ell$  that  $\ell_x(\cdot, x^o(\cdot), u^o) \in L_1^a(I, L_2(\Omega, E))$  and on the other hand we have  $z \in B_\infty^a(I, E) \subset L_\infty^a(I, L_2(\Omega, E))$ . Thus the first integral in (29) is well defined as a duality product in the spaces indicated. Since  $z \in B_\infty^a(I, E)$  and it has continuous modification,  $z(T) \in L_2(\Omega, E)$  and it is  $\mathcal{F}_T$ -adapted and by our assumption  $\Phi_x(x^o(T)) \in L_2(\Omega, E)$ . Hence the last scalar product is also well defined. Thus the functional  $z \rightarrow L(z)$ , given by (29), is a well defined continuous linear functional in  $z \in B_\infty^a(I, E)$ . Now note that the variational equation (27) is a linear nonhomogeneous stochastic differential equation on the Hilbert space  $E$  driven by the semi martingale  $M_{u-u^o}^o$ . Let  $\mathcal{SM}_T^2(E)$  denote the space of  $E$ -valued square integrable semi martingales defined on  $I \equiv [0, T]$  with initial value zero. It follows from our basic assumptions on  $b$  and  $\sigma$  that for every  $u \in \mathcal{U}_{ad}$  the process  $M_{u-u^o}^o$ , given by (28) for  $v = u - u^o$ , is a continuous  $E$ -valued square integrable semi-martingale satisfying  $\sup\{\mathbf{E}|M_{u-u^o}^o(t)|_E^2, t \in I\} < \infty$ . Considering the homogeneous part of the variational equation (27) and recalling that by our assumption  $b_x(t, x^o(t), u_t^o) \in \mathcal{L}(E)$ ,  $\sigma_x(t, x^o(t), u_t^o) \in \mathcal{L}(E, \mathcal{L}(H, E))$  and that they are uniformly bounded P-a.s, it follows from perturbation theory of semigroups that there exists an  $\mathcal{F}_t$ -adapted P-a.s essentially bounded evolution operator  $\Psi(t, s), 0 \leq s < t \leq T$ ,

taking values from  $\mathcal{L}(E)$  such that

$$(30) \quad z(t) = \Psi(t, 0)z_0 + \int_0^t \Psi(t, s) dM_{u-u^o}(s).$$

It follows from the above observations that the composition map  $M \rightarrow z \rightarrow L(z)$  is a continuous linear functional on the space  $\mathcal{SM}_T^2(E)$  of square integrable and continuous  $E$ -valued semi martingales. Thus it follows from classical semi-martingale representation theorem that there exists a pair  $(\psi, \mathcal{R}) \in L_2^a(I, E) \times L_{2,Q}^a(I, L(H, E))$  possibly dependent on the optimal pair  $\{u^o, x^o\}$  such that

$$(31) \quad L(z) = \mathbf{E} \left\{ \int_0^T \langle \psi(t), b(t, x^o(t), u_t - u_t^o) \rangle_E dt + \int_0^T \text{Tr}(\mathcal{R}(t) Q \sigma^*(t, x^o(t), u_t - u_t^o)) dt \right\}.$$

Hence it follows from the expression (25) that

$$(32) \quad \begin{aligned} dJ(u^o; u - u^o) = \mathbf{E} \left\{ \int_0^T \left( \langle \psi(t), b(t, x^o(t), u_t - u_t^o) \rangle_E \right. \right. \\ \left. \left. + \text{Tr}(\mathcal{R}(t) Q \sigma^*(t, x^o(t), u_t - u_t^o)) \right. \right. \\ \left. \left. + \ell(t, x^o(t), u_t - u_t^o) \right) dt \right\} \geq 0 \quad \forall \quad u \in \mathcal{U}_{ad}. \end{aligned}$$

This proves the necessary condition (23). Equation (22) is the state equation corresponding to the optimal control so nothing to prove. This completes the proof of the necessary conditions as stated.  $\blacksquare$

Note that the previous theorem asserts that for a control state pair  $\{u^o, x^o\} \in \mathcal{U}_{ad} \times B_\infty^a(I, E)$  to be optimal it is necessary that there exists a pair  $(\psi, \mathcal{R}) \in B_\infty^a(I, E) \times L_{2,Q}^a(I, L(H, E))$  satisfying (23). But it does not say how to construct it. In the following theorem we do just that. For convenience of notations we set

$$b_x^o(t) \equiv b_x(t, x^o(t), u_t^o), \sigma_x^o(z) \equiv \sigma_x(t, x^o(t), u_t^o; z), \ell_x^o(t) \equiv \ell_x(t, x^o(t), u_t^o)$$

and note that  $\sigma_x^o \in \mathcal{L}(E, \mathcal{L}(H, E))$  and it is P-a.s uniformly bounded on  $I$ .

**Theorem 6.3.** *Suppose the assumptions of Theorem 6.2 hold. Then a pair  $(\varphi, \Gamma) \in B_\infty^a(I, E) \times L_{2,Q}^a(I, L(H, E))$ , equivalent to the semi-martingale intensity pair  $(\psi, \mathcal{R})$  whose existence was proved in Theorem 6.2, can be constructed from the solution of the backward stochastic differential equation*

$$\begin{aligned}
-d\varphi(t) &= A^* \varphi(t) dt + (b_x^o)^* \varphi dt - \Sigma^o(t) \varphi(t) dt + \ell_x(t, x^o(t), u_t^o) dt \\
(33) \quad &+ \sigma_x^o(\varphi(t)) dW(t)
\end{aligned}$$

$$\varphi(T) = \Phi_x(x^o(T)),$$

where the operator valued process  $\Gamma(t) = -\sigma_x^o(\varphi(t))$ ,  $t \in I$ , and  $\Sigma^o(t)$ ,  $t \in I$ , is given by the bilinear form

$$(\Sigma^o(t)\xi, \zeta)_E \equiv (\Sigma(t, x^o(t), u_t^o)\xi, \zeta) = Tr(\sigma_x^o(\xi)Q(\sigma_x^o(\zeta))^*), \xi, \zeta \in E.$$

**Proof.** In view of the variational equation (27), we may take without any loss of generality the form of the adjoint equation as

$$(34) \quad d\varphi = -A^* \varphi dt + \dots \text{ (bounded variation terms) } dt \dots + \Gamma(t) dW,$$

with  $\Gamma \in L_{2,Q}^a(I, L(H, E))$ . Let  $z$  denote the mild solution of the variational equation (27) corresponding to the optimal control state pair as indicated and  $u \in \mathcal{U}_{ad}$ . Now using the well known Itô differential rule we compute the Itô differential of the scalar product  $\langle \varphi(t), z(t) \rangle_E$ . It is well known that

$$(35) \quad d \langle \varphi, z \rangle_E = \langle d\varphi, z \rangle_E + \langle \varphi, dz \rangle_E + \langle \langle d\varphi, dz \rangle \rangle$$

where the last term represents the quadratic variation. Using the notations introduced above, the variational equation (27) can be compactly written as

$$\begin{aligned}
(36) \quad dz &= Azdt + b_x^o(t)z(t)dt + \sigma_x^o(z(t))dW + dM_{u-u^o}^o(t) \\
z(0) &= z_0 = 0.
\end{aligned}$$

Considering the right hand side of the expression (35) and using the variational equation (36) and the form of the adjoint equation (34), we obtain

$$\begin{aligned}
(37) \quad RHS(35) &= \langle d\varphi, z \rangle + \langle \varphi, Azdt + b_x^o(t)zdt + \sigma_x^o(z)dW + dM_{u-u^o}^o \rangle \\
&+ \langle \langle \sigma_x^o(z)dW + dM_{u-u^o}^o, \Gamma(t)dW \rangle \rangle.
\end{aligned}$$

Using (28) for  $M^o$  at  $u - u^o$  and computing the quadratic variation term this takes the following form

$$\begin{aligned}
(38) \quad RHS(35) &= \langle d\varphi, z \rangle + \langle \varphi, Azdt + b_x^o(t)zdt + \sigma_x^o(z)dW + dM_{u-u^o}^o \rangle \\
&+ Tr(\Gamma(t)Q(\sigma_x^o)^*(z))dt + Tr(\Gamma(t)Q\sigma^*(t, x^o(t), u_t - u_t^o))dt.
\end{aligned}$$

Since the process  $\{Tr(\Gamma(t)Q(\sigma_x^o)^*(z(t))), t \in I\}$  is linear in  $z$  and is an integrable  $\mathcal{F}_t$ -adapted process, there exists a process  $\varrho_\Gamma \in L_2^a(I, L_2(\Omega, E))$  such that

$$(39) \quad Tr(\Gamma(t)Q(\sigma_x^o)^*(z(t))) \equiv \langle \varrho_\Gamma(t), z(t) \rangle_E.$$

Using this expression in (38) we arrive at the following expression

$$(40) \quad \begin{aligned} RHS(35) = & \langle d\varphi, z \rangle + \langle \varphi, Azdt + b_x^o(t)zdt + \sigma_x^o(z)dW + dM_{u-u^o}^o \rangle \\ & + \langle \varrho_\Gamma(t), z(t) \rangle dt + Tr(\Gamma(t)Q\sigma^*(t, x^o(t), u_t - u_t^o))dt. \end{aligned}$$

Since we consider solutions of all our evolution equations only in the mild sense, by virtue of Lemma 6.1 it is justified to rewrite (40) using the adjoint operations as follows

$$(41) \quad \begin{aligned} RHS(35) = & \langle d\varphi, z \rangle + \langle A^*\varphi dt + (b_x^o)^*\varphi dt + \varrho_\Gamma(t)dt + \sigma_x^o(\varphi)dW, z \rangle \\ & + \langle \varphi, dM_{u-u^o}^o \rangle + Tr(\Gamma(t)Q\sigma^*(t, x^o(t), u_t - u_t^o))dt. \end{aligned}$$

Now integrating either side of the expression (35) we have

$$(42) \quad \mathbf{E} \int_0^T d\langle \varphi, z \rangle = \mathbf{E} \int_0^T \{ \langle d\varphi, z \rangle_E + \langle \varphi, dz \rangle_E + \langle \langle d\varphi, dz \rangle \rangle \}.$$

Computing the expression on the left hand side and recalling that  $z(0) = 0$ , we have

$$(43) \quad LHS(42) = \mathbf{E} \int_0^T d\langle \varphi, z \rangle = \mathbf{E} \langle \varphi(T), z(T) \rangle_E.$$

Now setting

$$(44) \quad \begin{aligned} d\varphi + A^*\varphi dt + (b_x^o)^*\varphi dt + \varrho_\Gamma(t)dt + \sigma_x^o(\varphi)dW &= -\ell_x(t, x^o(t), u_t^o)dt \\ \varphi(T) &= \Phi_x(x^o(T)), \end{aligned}$$

in the expression (41) and (43) and integrating the former it follows from the expression (42) that

$$(45) \quad \begin{aligned} & \mathbf{E} \left\{ \langle \Phi_x(x^o(T)), z(T) \rangle_E + \int_0^T \langle \ell_x(t, x^o(t), u_t^o), z(t) \rangle_E dt \right\} \\ &= \mathbf{E} \left\{ \int_0^T \langle \varphi(t), dM_{u-u^o}^o(t) \rangle + \int_0^T Tr(\Gamma(t)Q\sigma^*(t, x^o(t), u_t - u_t^o))dt \right\}. \end{aligned}$$

Using the representation (28) for the semi-martingale  $M_{u-u^o}$  in the second term on the righthand side of the above expression, we obtain

$$\begin{aligned}
 & \mathbf{E} \left\{ \langle \Phi_x(x^o(T)), z(T) \rangle_E + \int_0^T \langle \ell_x(t, x^o(t), u_t^o), z(t) \rangle_E dt \right\} \\
 (46) \quad &= \mathbf{E} \left\{ \int_0^T \langle \varphi(t), b(t, x^o(t), u_t - u_t^o) \rangle dt \right. \\
 & \quad \left. + \int_0^T \text{Tr}(\Gamma(t) Q \sigma^*(t, x^o(t), u_t - u_t^o)) dt \right\},
 \end{aligned}$$

where we have used stopping time argument to verify that

$$\mathbf{E} \int_0^T (\sigma^*(t, x^o(t), u_t - u_t^o) \varphi(t), dW(t)) = 0.$$

Now comparing the form of the adjoint equation (34) with equation (44) we observe that

$$\Gamma(t) = -\sigma_x^o(\varphi(t)) \equiv -\sigma_x(t, x^o(t), u_t^o; \varphi(t)), t \in I.$$

Then using the definition of the process  $\varrho_\Gamma$  given by the expression (39) we have

$$\begin{aligned}
 \langle \varrho_\Gamma(t), z(t) \rangle &\equiv \text{Tr}(\Gamma(t) Q (\sigma_x^o)^*(z(t))) \\
 (47) \quad &= -\text{Tr}(\sigma_x^o(\varphi(t)) Q (\sigma_x^o)^*(z(t))) \\
 &\equiv -(\Sigma^o(t, x^o(t), u_t^o) \varphi(t), z(t))_E \equiv -(\Sigma^o(t) \varphi(t), z(t))_E.
 \end{aligned}$$

Since the Gateaux derivatives  $\sigma_x^o(\cdot)$  are uniformly bounded it follows from the above expression that the operator valued process  $\Sigma^o$  is  $P$ -a.s bounded on  $I$  and that it is a positive self adjoint operator in  $E$ . Thus finally the adjoint system (44) takes the following form

$$\begin{aligned}
 (48) \quad & d\varphi + A^* \varphi dt + ((b_x^o)^* - \Sigma^o) \varphi dt + \sigma_x^o(\varphi) dW + \ell_x^o dt = 0 \\
 & \varphi(T) = \Phi_x(x^o(T)),
 \end{aligned}$$

and this is the system (33) as stated in the theorem. It follows from a general result on the backward stochastic evolution equation due to Hu and Peng [14, Theorem 3.1, p. 455] that the evolution equation (48) has a unique mild solution giving  $(\varphi, \Gamma) \in B_\infty^a(I, E) \times L_{2,Q}^a(I, L(H, E))$  with  $\Gamma$  given by

$$\Gamma(t) = -\sigma_x^o(\varphi(t)) \equiv -\sigma_x(t, x^o(t), u_t^o; \varphi(t)), t \in I.$$

Now it is very important to note that the left hand side of the identity (46) is precisely the linear functional  $L(z)$  given by the expression (29) and having the representation (31) in terms of the semi-martingale intensity pair  $(\psi, \mathcal{R})$  as seen in Theorem 6.2. This shows that the same linear functional  $L(z)$  is also given in terms of the solution pair  $(\varphi, \Gamma) = (\varphi, -\sigma_x^o(\varphi))$  of the adjoint evolution equation (48). In this sense the solution pair  $(\varphi, \Gamma) = (\varphi, -\sigma_x^o(\varphi))$  of the backward adjoint evolution equation (48) is perfectly equivalent to the pair  $(\psi, \mathcal{R})$ , as given by the semi-martingale representation theory. This completes the proof of all the statements of the Theorem. ■

**Remark 6.4.** In view of the above theorem the inequality (23) is equivalent to

$$(49) \quad dJ(u^o, u - u^o) = \mathbf{E} \left\{ \int_0^T \left( \langle \varphi(t), b(t, x^o(t), u_t - u_t^o) \rangle_E \right. \right. \\ \left. \left. - \text{Tr}(\sigma_x^o(\varphi(t))Q\sigma^*(t, x^o(t), u_t - u_t^o)) \right. \right. \\ \left. \left. + \ell(t, x^o(t), u_t - u_t^o) \right) dt \right\} \geq 0, \quad \forall u \in \mathcal{U}_{ad},$$

written in terms of the  $\mathcal{F}_t$ -adapted mild solution of the backward adjoint evolution equation (33). The inequality (49) can be written compactly as follows. We see later that there are some conceptual and computational advantages with this compactification. Define the (Hamiltonian like) functional  $H : I \times E \times E \times M_0(U) \longrightarrow \overline{R}$  as follows:

$$H(t, x, \varphi, u) \equiv \langle \varphi, b(t, x, u) \rangle_E - \text{Tr}(\sigma_x(\varphi)Q\sigma^*(t, x, u)) + \ell(t, x, u)$$

where  $\sigma_x(\varphi) \equiv \sigma_x(t, x, u; \varphi)$  is the Gateaux derivative of  $\sigma$  with respect to the state in the direction  $\varphi$  and note that

$$H(t, x, \varphi, u) \equiv \int_U H(t, x, \varphi, \xi) u(d\xi)$$

for  $u \in M_0(U)$ . Using this notation we can rewrite the necessary condition (23), equivalent to (49), in the form of Pontryagin minimum principle,

$$(50) \quad \mathbf{E} \int_0^T H(t, x^o(t), \varphi(t), u_t) dt \geq \mathbf{E} \int_0^T H(t, x^o(t), \varphi(t), u_t^o) dt \quad \forall u \in \mathcal{U}_{ad},$$

where  $x^o \in B_\infty^a(I, E)$  is the mild solution of the evolution equation (1) corresponding to the control  $u^o$  and  $\varphi$  is the mild solution of the adjoint evolution equation (48) determined by the pair  $(u^o, x^o) \in \mathcal{U}_{ad} \times B_\infty^a(I, E)$ .

### 6.1. A conceptual algorithm

We use the compact form (50) of the necessary conditions of optimality to develop a computational algorithm. Towards this goal, we note that since the controls are adapted to the family of sigma algebras  $\mathcal{G}_t, t \geq 0$ , which are sub-sigma algebras of the family of sigma algebras  $\mathcal{F}_t, t \geq 0$ , we can use the properties of iterated conditional expectations to rewrite the inequality in a more convenient form. First note that

$$\mathbf{E} \int_0^T H(t, x^o(t), \varphi(t), u_t) dt = \mathbf{E} \int_0^T \langle H(t, x^o(t), \varphi(t), \cdot), u_t \rangle_{C(U), M(U)} dt$$

where we have used the duality pairing between the Banach space  $C(U)$  and its dual  $M(U)$ . Since  $u_t$  is  $\mathcal{G}_t$  adapted (in the weak star sense), using the conditional expectation we have

$$\begin{aligned} & \mathbf{E} \int_0^T \langle H(t, x^o(t), \varphi(t), \cdot), u_t \rangle_{C(U), M(U)} dt \\ &= \mathbf{E} \int_0^T \langle \mathbf{E}\{H(t, x^o(t), \varphi(t), \cdot) | \mathcal{G}_t\}, u_t \rangle dt \\ &= \mathbf{E} \int_0^T \langle H_o(t), u_t \rangle_{C(U), M(U)} dt, \end{aligned}$$

where  $\{H_o(t), t \geq 0\}$  is a  $C(U)$ -valued  $\mathcal{G}_t$ -adapted random process given by

$$H_o(t, \xi) \equiv \mathbf{E}\{H(t, x^o(t), \varphi(t), \xi) | \mathcal{G}_t\}.$$

Clearly,  $H_o$  is dependent on the optimal state and the associated adjoint process  $\{x^o, \varphi\}$ . Thus the necessary inequality (optimality condition) (50) can be rewritten as

$$(51) \quad \mathbf{E} \int_0^T \langle H_o(t), u_t \rangle_{C(U), M(U)} dt \geq \mathbf{E} \int_0^T \langle H_o(t), u_t^o \rangle_{C(U), M(U)} dt$$

for all  $u \in \mathcal{U}_{ad}$ . This is the inequality we are going to use to develop a conceptual algorithm.

**Proposition 6.5.** *Under the assumptions of Theorem 6.2 and 6.3, for every initial choice of control  $u_0 \in \mathcal{U}_{ad}$ , there exists a weak star convergent sequence of controls  $\{u^n\} \subset \mathcal{U}_{ad}$  so that  $\{J(u^n)\}$  is monotonically decreasing to a local minimum.*

**Proof.** Suppose at the  $n$ -th stage of iteration we have reached  $u^n \in \mathcal{U}_{ad}$  starting from  $u_0$ . We construct at the next iteration a control  $u^{n+1}$  satisfying  $J(u^{n+1}) \leq$

$J(u^n)$  as shown below. Let  $x^n \in B_\infty^a(I, E)$  denote the mild solution of the evolution equation (22) corresponding to the control  $u^n$  in place of  $u^o$ . Using the pair  $\{u^n, x^n\}$  in the adjoint equation (33) (in place of  $\{u^o, x^o\}$ ) we obtain the adjoint process  $\varphi^n \in B_\infty^a(I, E)$ . Note that the superscript  $o$  in the equations (22) is replaced by  $n$ . Define the  $C(U)$ -valued process  $H_n$  with values  $H_n(t), t \in I$ , given by

$$H_n(t)(\xi) = H_n(t, \xi) \equiv H(t, x^n(t), \varphi^n(t), \xi), \quad t \in I, \xi \in U.$$

Using this compact form, we deduce from the expression (49) that the Gateaux differential of  $J$  at  $u^n$  in the direction  $u^{n+1} - u^n$  is given by

$$(52) \quad dJ(u^n; u^{n+1} - u^n) = \mathbf{E} \int_0^T \langle H_n(t), u_t^{n+1} - u_t^n \rangle_{CU, M(U)} dt.$$

Using this expression, we can write

$$(53) \quad \begin{aligned} J(u^{n+1}) &= J(u^n) + dJ(u^n; u^{n+1} - u^n) + o(e_n) \\ &= J(u^n) + E \int_0^T \langle H_n(t), u_t^{n+1} - u_t^n \rangle dt + o(e_n) \end{aligned}$$

where  $o(e_n)$  denotes the small order of  $\|e_n\| \equiv \|u^{n+1} - u^n\|$ . At this stage we need the duality map. Let  $X$  be any Banach space with (topological) dual  $X^*$ . Define the duality map  $\Lambda : X \rightarrow 2^{X^*}$  by

$$\Lambda(x) \equiv \{x^* \in X^* : (x^*, x) = |x|_X^2 = |x^*|_{X^*}^2\}.$$

In general this is a multi-valued map. If  $X^*$  is strictly convex then  $\Lambda$  is single valued demicontinuous. In our case  $X = C(U)$  and  $X^* = M(U) \supset M_0(U)$ . The dual  $M(U)$  is not strictly convex. Therefore the duality map is multivalued. It is not difficult to show that for any  $x \in X$ ,  $\Lambda(x)$  is a convex subset of  $X^*$ . By use of Hahn-Banach theorem, one can show that  $\Lambda(x)$  is a weak star closed subset of  $X^*$ . Further, it is scalarly weak star continuous in the sense that whenever  $x_n \xrightarrow{s} x_0$  in  $X$ ,  $\Lambda(x_n)(\xi) \rightarrow \Lambda(x_0)(\xi)$  in  $2^{X^*}$  for every  $\xi \in X$ . More precisely, if  $\xi_n \in \Lambda(x_n)$  and  $x_n \xrightarrow{s} x_0$  in  $X$  and  $\xi_n \xrightarrow{w^*} \xi_0$  in  $X^*$  then  $\xi_0 \in \Lambda(x_0)$ . Considering now the multivalued map  $t \rightarrow \Lambda(H_n(t))$  on  $I$ , it follows from the scalar continuity of the duality map  $x \rightarrow \Lambda(x)$  and Borel measurability of  $t \rightarrow H_n(t)$  that the composition map  $t \rightarrow \Lambda(H_n(t)) \in 2^{M(U)} \setminus \emptyset$  is measurable in the sense of multifunctions. Since, by our assumption,  $U$  is a compact Polish space,  $M(U)$  is a compact Polish space. Hence it follows from a well known measurable selection theorem [13, Theorem 2.1, p. 154], or the Yankov-Von

Neumann-Aumann selection theorem [13, Theorem 2.14, p. 158] that there exists a measurable selection  $v^n$  so that  $v_t^n \in \Lambda(H_n(t))$  for almost all  $t \in I$  and that

$$\mathbf{E} \int_0^T \langle H_n(t), v_t^n \rangle dt = \mathbf{E} \int_0^T \|H_n(t)\|_{C(U)}^2 dt = \mathbf{E} \int_0^T \|v_t^n\|_{M(U)}^2 dt.$$

Thus for a choice of  $\varepsilon > 0$  sufficiently small, so that the control  $u^{n+1} \equiv u^n - \varepsilon v^n \in \mathcal{U}_{ad}$ , it follows from (53) that

$$(54) \quad J(u^{n+1}) = J(u^n) - \varepsilon E \int_0^T \|H_n(t)\|_{C(U)}^2 dt + o(\varepsilon).$$

Hence, for such an  $\varepsilon$ ,  $J(u^{n+1}) < J(u^n)$ . This proves that we can construct a sequence of controls  $\{u^n\} \in \mathcal{U}_{ad}$  such that the corresponding sequence of cost functionals  $\{J(u^n)\}$  is monotone decreasing to a possibly local minimum. This completes the proof.  $\blacksquare$

## 7. REGULAR VS RELAXED CONTROLS

Recall that the relaxed controls  $\mathcal{U}_{ad}$  are probability measure valued  $\mathcal{G}_t$ -adapted random processes defined on  $I \equiv [0, T]$  while the regular controls  $\mathcal{U}_r$  are Borel measurable functions defined on  $I$  and adapted to the filtration  $\mathcal{G}_t$  and taking values in  $U$ . It is interesting to mention that the necessary conditions of optimality based on relaxed controls are very general. It was shown in [8, Corollary 8.3.7, p. 278] that the Pontryagin minimum principle follows immediately from the necessary conditions based on relaxed controls and its proof is very simple. In view of this fact, if indeed an optimal control exists in the class  $\mathcal{U}_r$ , we can replace the measure valued processes  $\{\{u_t^o, u_t\}, t \in I\}$  by regular controls  $\{u^o(t), u(t)\}, t \in I\}$ . Then by use of Lebesgue density argument, one can easily derive from (51) the pointwise necessary conditions of optimality as presented below:

$$H(t, x^o(t), \varphi(t), u) \geq H(t, x^o(t), \varphi(t), u^o(t)) \quad \mu - a.e \quad (t, \omega) \in I \times \Omega \quad \& \quad \forall u \in U$$

where  $\{x^o, \varphi\}$  are the mild solutions of equations (22)–(33) corresponding to the regular control  $\{u^o(t), t \in I\}$ .

Another interesting fact regarding relaxed controls is as follows. It is well known that if  $U$  is not convex, optimal controls may not exist in the class of regular controls  $\mathcal{U}_r$ . However, under fairly relaxed assumptions, optimal controls do exist in the class of relaxed controls  $\mathcal{U}_{ad}$  even if  $U$  is not convex as demonstrated in the following example. Then the important question one may ask: can the optimal relaxed control be approximated by a suitable regular control?. This

question can be answered by use of the well known Krein-Milman theorem [11, Theorem V.8.4, p. 440], [8, Theorem A.7.8] which states that any weakly compact convex set  $K$  in a Banach space has the property,  $clco(\text{ext}K) = K$  where  $\text{ext}K$  denotes the set of extreme points of the set  $K$ . Returning to our control problem, let  $\delta_v$  denote the Dirac measure concentrated at  $v \in U$ . Similarly, for any  $u \in \mathcal{U}_r$ , let  $\delta_{u(\cdot)}$  be the Dirac measure concentrated along the path  $\{u(t), t \in I\}$ . The reader can easily verify that the map  $u \xrightarrow{i} \delta_{u(\cdot)}$  is a continuous embedding of  $\mathcal{U}_r$  into  $\mathcal{U}_{ad}$  denoted by  $\mathcal{U}_r \hookrightarrow \mathcal{U}_{ad} \subset L_\infty^a(I, M_0(U))$ . Note that the set of extremals of the set  $\mathcal{U}_{ad}$  coincides with the set  $i(\mathcal{U}_r)$ . Hence it follows from the Krein-Milman theorem that  $cl^{w^*}co(i(\mathcal{U}_r)) = \mathcal{U}_{ad}$ . By virtue of this fact we conclude that the set of regular controls  $\mathcal{U}_r$  is  $w^*$ -dense in  $\mathcal{U}_{ad}$ . Thus it follows from continuity of the control to solution map (see Theorem 4.3) that optimal controls from the relaxed class can be approximated arbitrary closely by regular controls. Hence suboptimal controls can be found from the class of regular controls. For more on this topic see [8, 9].

### 7.1. An example of a non-convex problem

We consider the system (1) with an important class of admissible controls which is quite natural in physical applications. These controls can take only a finite number of values from a Polish space  $\mathcal{P}$ . Let this set be denoted by  $U \equiv \{u_i, 1 \leq i \leq n\}$  where  $\{u_i\} \in \mathcal{P}$ . Clearly, the set  $U$  is not convex but  $M_0(U)$  is. Thus the set  $\mathcal{U}_{ad}$ , the class of probability measure valued random processes adapted to the filtration  $\mathcal{G}_t, t \geq 0$ , is convex while the corresponding set of regular controls  $\mathcal{U}_r$  taking values in  $U$  is not. Thus optimal control may not exist in the class  $\mathcal{U}_r$  but does so in the class  $\mathcal{U}_{ad}$ . In this case the relaxed controls have the representation

$$(55) \quad u_t \equiv \sum_{i=1}^n \alpha_i(t) \delta_{u_i}$$

where  $\alpha_i(t) \geq 0$  and  $\sum_{i=1}^n \alpha_i(t) = 1$  for all  $t \in I$  and each process  $\alpha_i(t)$  is adapted to the filtration  $\mathcal{G}_t$ . By virtue of our existence theorem 5.1, an optimal control  $u^o \in \mathcal{U}_{ad}$  exists and it has the representation

$$(56) \quad u_t^o \equiv \sum_{i=1}^n \alpha_i^o(t) \delta_{u_i}$$

with  $\{\alpha_i^o(t), t \in I\}$  being a  $\mathcal{G}_t$ -adapted random process satisfying the constraints,  $\alpha_i^o(t) \geq 0$  and  $\sum_{i=1}^n \alpha_i^o(t) = 1$ . Substituting these controls in the necessary condition (49) we obtain

$$(57) \quad \mathbf{E} \int_I \sum_{i=1}^n H(t, x^o(t), \varphi(t), u_i) \alpha_i(t) dt \geq \mathbf{E} \int_I \sum_{i=1}^n H(t, x^o(t), \varphi(t), u_i) \alpha_i^o(t) dt$$

where  $\{x^o, \varphi\}$  are the mild solutions of (22) and (33) respectively corresponding to the control  $u^o$  given by (56). For each  $i \in \{1, 2, \dots, n\}$ , define

$$(58) \quad \beta_i^o(t) \equiv \mathbf{E}\{H(t, x^o(t), \psi(t), u_i) | \mathcal{G}_t\}.$$

Clearly,  $\beta_i^o$  is a  $\mathcal{G}_t$ -adapted random process. Recall that we want our controls to be adapted to the filtration  $\mathcal{G}_t, t \geq 0$ . Thus taking the conditional expectation with respect to the filtration  $\mathcal{G}_t, t \geq 0$ , of either side of (57) and using the expression (58) we arrive at the following necessary condition,

$$(59) \quad \mathbf{E} \int_I \langle \beta^o(t), \alpha(t) \rangle_{R^n} dt \geq \mathbf{E} \int_I \langle \beta^o(t), \alpha^o(t) \rangle_{R^n} dt,$$

where  $\alpha(t) \equiv \{\alpha_i(t), 1 \leq i \leq n\}$ . Note that for each  $t \in I$ ,  $\alpha(t), \alpha^o(t) \in \mathcal{S} \equiv \{\eta \in R^n : \eta_i \geq 0, \sum_{i=1}^n \eta_i = 1\}$ . Let  $L_1^{\mathcal{G}}(I \times \Omega, R^n) = L_1^{\alpha}(I, R^n)$  denote the space of  $R^n$  valued  $\mathcal{G}_t$ -adapted integrable random processes and its dual by  $L_{\infty}^{\mathcal{G}}(I \times \Omega, R^n) \equiv L_{\infty}^{\alpha}(I, R^n)$ . Under the basic assumptions on  $\{b, \sigma, Q, \ell, \Phi\}$ , the reader can easily verify that the Hamiltonian along any path  $\{x, \varphi\} \in B_{\infty}^a(I, E) \times B_{\infty}^a(I, E)$  is integrable. Thus  $\beta^o \in L_1^{\alpha}(I, R^n)$  and, by our choice of the admissible controls,  $\alpha \in L_{\infty}^{\mathcal{G}}(I \times \Omega, \mathcal{S}) \subset L_{\infty}^{\alpha}(I, R^n)$  and hence the functional

$$(60) \quad L_o(\alpha) \equiv \mathbf{E} \int_I \langle \beta^o(t), \alpha(t) \rangle_{R^n} dt$$

represents the natural duality product. Since  $\mathcal{S}$  is a compact simplex, by Alaoglu's theorem, the set  $L_{\infty}^{\mathcal{G}}(I \times \Omega, \mathcal{S}) \equiv L_{\infty}^{\alpha}(I, \mathcal{S})$  is  $w^*$ -compact. And the functional  $\alpha \rightarrow L_o(\alpha)$  is linear and  $w^*$ -continuous, and therefore it attains its minimum at some point in  $L_{\infty}^{\alpha}(I, \mathcal{S})$  as expected (Theorem 5.1).

### 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, *Deterministic and stochastic neutral systems on Banach spaces and their optimal feedback controls*, J. Nonlin. Syst. Appl. (2009) 151–160.
- [2] N.U. Ahmed, *Measure valued solutions for systems governed by neutral differential equations on Banach spaces and their optimal control*, Discuss. Math. DICO **33** (2013) 89–109. doi:10.7151/dmdico.1142.
- [3] N.U. Ahmed, *Relaxed solutions for stochastic evolution equations on Hilbert space with polynomial growth*, Publ. Math. Debrecen **54** (1–2) (1999) 75–101.

- [4] N.U. Ahmed, *Stochastic neutral evolution equations on Hilbert spaces with partially observed relaxed control and their necessary conditions of optimality*, Nonlin. Anal. TMA **101** (2014) 66–79.
- [5] N.U. Ahmed, Some Recent Developments in Systems and Control Theory on Infinite Dimensional Banach Spaces, Part 1 & 2, Proceedings of the 5th International Conference on Optimization and Control with Applications, (Edited by: K.L. Teo, H. Xu and Y. Zhang), Beijing, China, 2012; Publisher: Springer-Verlag (in Print).
- [6] N.U. Ahmed, 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-Wiley & Sons, Inc. New York, 1988.
- [7] 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.
- [8] N.U. Ahmed, Dynamic Systems and Control with Applications, World Scientific (New Jersey, London, Singapore, Beijing, Shanghai, Hong Kong, Taipei, Chennai, 2006). ISBN: 981-270-053-6
- [9] N.U. Ahmed and C.D. Charalambous, *Stochastic minimum principle for partially observed systems subject to continuous and jump diffusion processes and driven by relaxed controls*, SIAM J. Control and Optim. **51** (4) (2013) 3235–3257. doi:10.1137/120885656
- [10] S. Bahlali, *Necessary and sufficient optimality conditions for relaxed and strict control problems*, SIAM J. Control and Optim. **47** (2008) 2078–2095. doi:10.1137/070681053
- [11] N. Dunford and J.T. Schwartz, Linear Operators, Part 1 (Interscience Publishers, Inc., New York, 1958).
- [12] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions (Cambridge University Press, 1992). doi:10.1017/CBO9780511666223
- [13] S. Hu and N.S. Papageorgiou, Handbook of Multivalued Analysis, Vol. 1, Theory (Kluwer Academic publishers, Dordrecht/Boston/London, 1997).
- [14] 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
- [15] W. Wei, *Maximum principle for optimal control of neutral stochastic functional differential systems*, Science China Math. (to appear)

Received 3 February 2014

