Nonlinear Functional Analysis and Applications
Vol. 18, No. 1 (2013), pp. 103-126

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm

NECESSARY CONDITIONS OF OPTIMALITY FOR OUTPUT FEEDBACK CONTROL LAW FOR A CLASS OF UNCERTAIN INFINITE DIMENSIONAL STOCHASTIC SYSTEMS

N. U. Ahmed

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

Abstract. In this paper we consider a class of partially observed semilinear stochastic evolution equations on Hilbert space subject to measurement uncertainty. The control is based on output feedback with noisy measurement. Using the space of bounded linear operator valued functions, furnished with the Tychonoff product topology, as feedback control laws we present the necessary conditions of optimality.

1. Introduction

Optimal control theory for finite and infinite dimensional systems with open loop control is extensively studied in the literature by many authors like Cesari [9], Fattorini [11] and Ahmed, Teo, Xiang [1]-[8], (see also the references therein). In applications of control theory, there are many problems in physical sciences and engineering where open loop control is not feasible. One must use feedback control based on imperfect measurement data. In the case of stochastic systems, the classical approach is to use Bellman's optimality principle to construct the HJB (Hamilton-Jacob-Bellman) equation which is a nonlinear PDE on a Hilbert space with solution, if one exists, giving the value function. The feedback control law is then constructed from the value function provided it is sufficiently smooth. This is valid only for fully observed problem.

[0]Received October 13, 2012. Revised January 25, 2013.

[0]2000 Mathematics Subject Classification: 49J24, 49J27, 49K24, 49K35, 93E20, 93B52, 35R60, 35R70, 34G25, 34H05, 47A62.

[0]Keywords: Partially observed uncertain systems, operator valued functions, feedback operators, necessary conditions of optimality in the presence uncertainty, saddle points.

For details the reader is referred to [16]-[19] and the references therein. For partially observed problems this is not possible. In this paper we follow a different path, we present the necessary conditions of optimality on the space of bounded linear operator valued functions in the presence of measurement uncertainty thereby directly providing optimal linear feedback operator (control law).

The rest of the paper is organized as follows. In section 2 we present some typical notations. In section 3, we present the mathematical model describing the system and formulate the control problem considered in the paper. The basic assumptions used are given in section 4 followed by a brief review of the question of existence of solutions and their continuous dependence on feedback operator and disturbance [1]. In section 5, necessary conditions of optimality of feedback operators in the presence of uncertainty are presented. The paper is concluded with some comments on point wise necessary conditions of optimality.

2. Some Notations

Let $\{X, Y, U\}$ denote a triple of real separable Hilbert spaces representing the state space, the output (measurement) space and the control space respectively. Let $I = [0, T]$ denote any closed bounded interval. For any separable reflexive Banach space Z, we let $L_1(I, Z)$ denote the space of Bochner integrable functions with values in Z, and its dual by $L_\infty(I, Z^*)$. Let Z_1, Z_2 be any pair real separable reflexive Banach spaces and $\mathcal{L}(Z_1, Z_2)$ the Banach space of bounded linear operators from Z_1 to Z_2. The topological dual of $\mathcal{L}(Z_1, Z_2)$ is given by the class of nuclear operators $\mathcal{L}_1(Z_1^*, Z_2^*)$. A continuous linear functional ℓ on $\mathcal{L}(Z_1, Z_2)$, denoted by $\ell(L)$, has the representation $\ell(L) = Tr(L^* S)$ for some $S \in \mathcal{L}_1(Z_1^*, Z_2^*)$. It is easy to verify that there exists a constant $c > 0$ such that

$$|\ell(L)| \leq c \parallel L \parallel_{\mathcal{L}(Z_1, Z_2)}$$

where $c \leq |Tr(S)|$. We use $B_1(Z)$ to denote the closed unit ball in any Banach space Z. An operator $C \in \mathcal{L}(Z_1, Z_2)$ is said to be compact if $C(B_1(Z_1))$ is a relatively compact subset of Z_2.

Let $B_\infty(I, \mathcal{L}(Z_1, Z_2))$ denote the space of operator valued functions $\{T\}$ which are measurable in the strong operator topology and uniformly bounded on the interval I in the sense that

$$\sup\{\parallel T(t) \parallel_{\mathcal{L}(Z_1, Z_2)}, t \in I\} < \infty.$$

Suppose this is furnished with the topology of strong convergence (convergence in the strong operator topology) uniformly on I in the sense that, given $T_n, T \in$

$B_\infty(I, \mathcal{L}(Z_1, Z_2))$, $T_n \xrightarrow{\tau_{so}} T$ in this topology iff for every $z \in Z_1$,

$$\sup\{|T_n(t)z - T(t)z|_{Z_2}, t \in I\} \to 0$$

as $n \to \infty$. Let $\mathcal{K}(Z_1, Z_2))$ denote the class of compact operators from Z_1 to Z_2. It is well known that this is a closed linear subspace of $\mathcal{L}(Z_1, Z_2)$ in the uniform operator topology and hence a Banach space. Let Γ be a closed convex subset of $\mathcal{K}(Z_1, Z_2)$. We are interested in the set $B_\infty(I, \Gamma) \subset B_\infty(I, \mathcal{K}(Z_1, Z_2))$ endowed with the relative topology of convergence in the strong operator topology of the space $\mathcal{L}(Z_1, Z_2)$ point wise in $t \in I$. Later in the sequel, this is used as the set of admissible feedback operator valued functions. Throughout the rest of the paper we consider only Hilbert spaces.

3. Uncertain system and Problem Formulation

Let X, Y, U, E be real separable Hilbert spaces, with X denoting the state space, Y denoting the output space, and U the space where controls take their values from. Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ denote a complete filtered probability space where $\{\mathcal{F}_t, t \geq 0\}$ is an increasing family of subsigma algebras of the σ-algebra $\mathcal{F}$. For any random variable z, $\mathcal{E}(z) \equiv \int_\Omega z(\omega)P(d\omega)$ denotes the expected value (average) of the random variable z. Let $\{W(t), t \geq 0\}$ denote the E-valued Brownian motion with $P\{W(0) = 0\} = 1$, and covariance operator $Q \in \mathcal{L}_1^+(E)$, the space of positive nuclear operators on E. The complete system is governed by the following system of equations:

$$dx = Axdt + F(x)dt + B(t)udt + \sigma(x)dW, \; x(0) = x_0 \text{ in } X, \quad (3.1)$$

$$y = L(t)x + \xi \text{ in } Y, \quad (3.2)$$

$$u = K(t)y \text{ in } U, \quad (3.3)$$

where the first equation describes the dynamics of the system in the state space X giving the state $x(t)$ at any time $t \geq 0$, the second equation describes the measurement process (sensor) that observes the status of the system in a noisy environment characterized by the uncertain process ξ and delivers the output $y(t), t \geq 0$, with values from the Hilbert space Y. In order to regulate the system (3.1), the third equation provides the control $u(t)$ with values in the Hilbert space U. It is the output of a linear operator valued function K (to be chosen) with the measurement process y being its input.

In general the operator A is an unbounded linear operator with domain and range in X. The operator F is a nonlinear map in X, the operator valued function B takes values from $\mathcal{L}(U, X)$ and $\sigma : X \longrightarrow \mathcal{L}(E, X)$ is a nonlinear operator. The operator L, representing the sensor (or measurement system), takes values from $\mathcal{L}(X, Y)$ and the output feedback control operator K is an operator valued function taking values from the space $\mathcal{L}(Y, U)$. The process

$\xi(t), t \geq 0$, represents the uncertainty in the measurement data and takes values from the Hilbert space Y. For most practical situations, it is reasonable to assume that the measurement uncertainty is bounded. And so without any loss of generality, we may assume that the process ξ is any strongly measurable function taking values from the closed unit ball $B_1(Y)$ centered at the origin. We denote this class of disturbance processes by $\mathcal{D}$. Let $\mathcal{F}_{ad}$, whose precise characterization is given later, denote the class of admissible feedback operator valued functions $\{K(t), t \geq 0\}$ with values in $\mathcal{L}(Y, U)$.

The performance of the system over the time horizon $I \equiv [0, T]$ is measured by the following functional (called cost functional)

$$J(K, \xi) \equiv \mathcal{E}\left\{\int_I \ell(t, x(t))dt + \Phi(x(T))\right\} \tag{3.4}$$

where $\ell : I \times X \longrightarrow [0, \infty]$ and $\Phi : X \longrightarrow [0, \infty]$. Clearly the cost functional depends on the choice of the control law K in the presence of disturbance ξ. Our objective is to find a bounded strongly measurable operator valued function K that minimizes the cost functional taking into account the worst case measurement uncertainty. In other words, we want a feedback law that minimizes the maximum risk. This problem can be formulated as min-max problem as stated below

$$\min_{K \in \mathcal{F}_{ad}} \max_{\xi \in \mathcal{D}} J(K, \xi). \tag{3.5}$$

4. Regularity of solutions

In order to solve the problem as stated above, we introduce the following basic assumptions:

(A1) The operator $A : D(A) \subset X \longrightarrow X$ is the infinitesimal generator of a C_0-semigroup of operators $S(t), t \geq 0$, on X.

(A2) The vector field $F : X \longrightarrow X$ is once continuously Fréchet differentiable with the Frechèt derivative uniformly bounded on X and measurable in the uniform operator topology along any bounded trajectory.

(A3) Both B and L are measurable in the uniform operator topology, and $B \in L_2^{\ell oc}([0, \infty), \mathcal{L}(U, X)), L \in L_\infty^{\ell oc}([0, \infty), \mathcal{L}(X, Y)$.

For the admissible feedback control laws represented by the operator valued function K, we introduce the following assumption.

(A4) Let $\Gamma \subset \mathcal{L}(Y, U)$ be a nonempty closed bounded convex set and

$$\mathcal{F}_{ad} \equiv \{K \in L_\infty^{\ell oc}([0, \infty), \mathcal{L}(Y, U)) : K(t) \in \Gamma \ a.e\}.$$

(A5) The disturbance (measurement uncertainty) process $\xi : [0, \infty) \longrightarrow Y$, is any measurable function taking values from the closed ball $B_1(Y)$

of the B-space Y. We denote this family by $\mathcal{D}$. This represents the uncertainty without any probabilistic structure.

(A6) The nonlinear diffusion operator $\sigma : X \longrightarrow \mathcal{L}(E, X)$ is once continuously Fréchet differentiable with the first Fréchet derivative being uniformly bounded on X.

(A7) The integrand $\ell : [0, \infty) \times X \longrightarrow (-\infty, \infty]$ is measurable in the first argument and once continuously Fréchet differentiable in the second satisfying

$$|\ell(t, x)| \leq g(t) + c_1 |x|_X^2, x \in X, t \geq 0$$

with $0 \leq g \in L_1^{\ell oc}([0, \infty))$ and $c_1 \geq 0$. The function Φ is once continuously Fréchet differentiable and there exist constants $c_2, c_3 \geq 0$ such that

$$|\Phi(x)| \leq c_2 + c_3 |x|_X^2.$$

By substituting the equations (3.2) and (3.3) into equation (3.1) we obtain the following stochastic feedback system

$$dx = Axdt + F(x)dt + BKLxdt + BK\xi dt + \sigma(x)dW, x_0 \in X \text{ (fixed)} , \quad (4.1)$$

subject to the (unstructured) disturbance $\xi \in \mathcal{D}$ and $K \in \mathcal{F}_{ad}$. Note that, because of uncertainty, the stochastic differential equation (4.1) is completely equivalent to the following differential inclusion

$$dx(t) \in (Ax + F(x) + BKLx + BKD(t))dt + \sigma(x)dW, x_0 \in X$$

for $K \in \mathcal{F}_{ad}$ where $D : I \longrightarrow 2^{B_1(Y)} \setminus \emptyset$ is a measurable multifunction, for example the constant multi $D(t) = B_1(Y)$ for all $t \geq 0$.

We introduce the vector space $B_\infty^a(I, L_2(\Omega, X))$ consisting of X valued second order progressively measurable random process adapted to the current of sigma algebras $\{\mathcal{F}_t, t \geq 0\}$. This is endowed with the norm topology $\parallel z \parallel$ derived from the following expression

$$\parallel z \parallel^2 \equiv \sup\{\mathcal{E}|z(t)|_X^2, t \in I\}.$$

It is easy to verify that, with respect to this topology, the normed space $B_\infty^a(I, L_2(\Omega, X))$ is a Banach space.

Before we conclude this section we present the following fundamental result on the existence and regularity of solutions of the feedback system.

Theorem 4.1. *Consider the stochastic uncertain feedback system given by* (4.1) *over any finite time horizon* $I \equiv [0, T]$, *and suppose the assumptions* (A1)-(A6) *hold. Then for every* $\mathcal{F}_0$*-measurable initial state* $x(0) = x_0 \in L_2(\Omega, X)$, *and any feedback law* $K \in \mathcal{F}_{ad}$ *and disturbance* $\xi \in \mathcal{D}$, *the system*

(4.1) *has a unique mild solution* $x \in B^a_\infty(I, L_2(\Omega, X))$. *Further, the solution set*

$$\mathcal{X} \equiv \left\{ x(\cdot, K, \xi) \in B^a_\infty(I, L_2(\Omega, X)) : K \in \mathcal{F}_{ad}, \xi \in \mathcal{D} \right\}$$

is a bounded subset of $B^a_\infty(I, L_2(\Omega, X))$.

Proof. See [1]. □

Using the factorization technique due to Da Prato and Zabczyk [10] we can show that each member of $\mathcal{X}$ has a continuous modification.

For characterization of feedback control operators, we use the following result due to Mayoral [12]. This result characterizes relatively compact subsets of $\mathcal{K}(Y, U)$. See also [Serrano, Pineiro, Delgado 13] for similar results.

Lemma 4.2. [Mayoral [12], Theorem 1, p.79] *If the B-space Y does not contain a copy of ℓ_1, a set $\Gamma \subset \mathcal{K}(Y, U)$ is relatively compact iff* (i): *Γ is uniformly completely continuous (ucc) and* (ii): *for every $y \in Y$, the y section of Γ denoted by $\Gamma(y) \equiv \{L(y), L \in \Gamma\}$ is relatively compact in U.*

Since reflexive Banach spaces do not contain a copy of ℓ_1, the above proposition holds for such spaces. We need the following assumptions.

(H1): Let $\{Y, U\}$ be any pair of real separable Hilbert spaces and $\Gamma \subset \mathcal{K}(Y, U)$ compact (Note that since Y is Hilbert, it does not contain a copy of ℓ_1). The set of admissible feedback control laws, denoted by $\mathcal{F}_{ad} \equiv B_\infty(I, \Gamma) \subset B_\infty(I, \mathcal{K}(Y, U))$, is furnished with the Tychonoff product topology τ_T so that it is a compact topological space.

(H2): For the class of disturbance $\mathcal{D}$, we choose the set $L_\infty(I, B_1(Y))$ of (essentially norm bounded) strongly measurable functions defined on I and taking values from the closed unit ball $B_1(Y)$ of the Hilbert space Y. By Alaoglu's theorem, this set is a weak star compact subset of $L_\infty(I, Y)$ which is the dual of $L_1(I, Y^*)$. This topology will be denoted by τ_w.

Theorem 4.3. *Consider the feedback system* (4.1) *and suppose the assumptions* (A1)-(A6) *and* (H1)-(H2) *hold. Then the solution map*

$$(K, \xi) \longrightarrow x(K, \xi) = \{x(t, K, \xi), t \in I\}$$

is jointly continuous on $\mathcal{F}_{ad} \times \mathcal{D}$ to $B^a_\infty(I, L_2(\Omega, X)$ with respect to the product topology $\tau_T \times \tau_w$ on $\mathcal{F}_{ad} \times \mathcal{D}$ and the norm topology on $B^a_\infty(I, L_2(\Omega, X))$. Further, if (A7) *holds then the functional J is jointly continuous on $\mathcal{F}_{ad} \times \mathcal{D}$ with respect to the product topology $\tau_T \times \tau_w$.*

Proof. See [1, Theorem 5.2]. □

5. Necessary Conditions of Optimality

Here we are mainly interested in developing a set of necessary conditions of optimality. In order for $K^o \in \mathcal{F}_{ad}$ to be optimal in the face of adversity it is necessary that there exists a $\xi^o \in \mathcal{D}$ so that the pair $\{K^o, \xi^o\} \in \mathcal{F}_{ad} \times \mathcal{D}$ satisfies the following inequalities,

$$J(K^o, \xi) \leq J(K^o, \xi^o) \leq J(K, \xi^o) \; \forall \; (K, \xi) \in \mathcal{F}_{ad} \times \mathcal{D}, \tag{5.1}$$

called the saddle point inequality. In other words $\{K^o, \xi^o\}$ is a saddle point for J. Throughout the rest of this paper we assume that a saddle point exists. For existence, interested readers may see [1,20]. The inequalities (5.1) signify that the system (4.1) with feedback control law K^o will operate well in the most adverse situation and even better in any other situation. Using this we can construct necessary conditions of optimality. We use the right hand inequality to characterize the optimal $K^o \in \mathcal{F}_{ad}$ and the left hand inequality to characterize the worst disturbance $\xi^o \in \mathcal{D}$. Combining the two results we obtain the optimal policy. For the characterization of the optimal feedback operator K^o we consider the system (4.1) subject to the worst case disturbance $\xi^o \in \mathcal{D}$ giving

$$dx = Axdt + F(x)dt + BKLxdt + BK\xi^o dt + \sigma(x)dW, x_0 \in X, \tag{5.2}$$

for $K \in \mathcal{F}_{ad}$. For the necessary conditions, we assume that the set Γ, defining the admissible feedback operators $\mathcal{F}_{ad}$, is a closed convex subset of $\mathcal{K}(Y, U)$. For the proof of necessary conditions we shall use Yosida approximation of the unbounded operator A. Let $\rho(A)$ denote the resolvent set of A and $R(\lambda, A) \equiv (\lambda I - A)^{-1}$ denote the resolvent of A corresponding to $\lambda \in \rho(A)$. It is well known that

$$\{A_\lambda \equiv \lambda A R(\lambda, A), \lambda \in \rho(A)\}$$

is a family of bounded linear operators in X and that A_λ converges strongly to A on $D(A)$. Let $\{S_\lambda(t), S(t), t \geq 0\}$, denote the semigroups of operators taking values in $\mathcal{L}(X)$ generated by the pair $\{A_\lambda, A\}$ respectively. It is well known [2] that $S_\lambda(t) \xrightarrow{\tau_{so}} S(t)$ uniformly on any bounded set $I \equiv [0, T]$. With this preparation, now we can present the following result.

Lemma 5.1. *Consider the system* (5.2) *and suppose the assumptions* (A1)-(A6) *and* (H1) *hold. Let x^o denote the (mild) solution of* (5.2) *corresponding to the pair $\{A, K\}$ and $\{x^n\}$ the solution of* (5.2) *corresponding to the pair $\{A_n, K_n\}$ for $n \in \rho(A) \cap N$ and $K_n \in \mathcal{F}_{ad}$. Then, as $A_n \to A$ strongly on $D(A)$ and $K_n \xrightarrow{\tau_T} K$, we have $x^n \xrightarrow{s} x^o$ in $B^a_\infty(I, L_2(\Omega, X))$.*

Proof. Note that x^o is the solution of the integral equation

$$x^o(t) = S(t)x_0 + \int_0^t S(t-r)F(x^o(r))dr + \int_0^t S(t-r)B(KLx^o + \xi^o)dr + \int_0^t S(t-r)\sigma(x^o(r))dW(r) \tag{5.3}$$

and x^n is the solution of the integral equation

$$x^n(t) = S_n(t)x_0 + \int_0^t S_n(t-r)F(x^n(r))dr + \int_0^t S_n(t-r)BK_n(Lx^n + \xi^o)dr + \int_0^t S_n(t-r)\sigma(x^n(r))dW(r). \tag{5.4}$$

Subtracting equation (5.4) from (5.3) and rearranging terms we have

$$x^o(t) - x^n(t) = e_n^1(t) + e_n^2(t) + \int_0^t S_n(t-r)[F(x^o(r)) - F(x^n(r))]dr + \int_0^t S_n(t-r)BK_nL(x^o - x^n)dr + \int_0^t S_n(t-r)(\sigma(x^o) - \sigma(x^n))dW(r) \tag{5.5}$$

where e_n^1, e_n^2 are given by

$$e_n^1(t) = (S(t) - S_n(t))x_0 + \int_0^t (S(t-r) - S_n(t-r))F(x^o(r))dr + \int_0^t (S(t-r) - S_n(t-r))BK(Lx^o + \xi^o)dr + \int_0^t S_n(t-r)B(K - K_n)[Lx^o + \xi^o)dr \tag{5.6}$$

and

$$e_n^2(t) = \int_0^t (S(t-r) - S_n(t-r))\sigma(x^o(r))dW(r) \tag{5.7}$$

respectively. Since the sequence of semigroups S_n converges to S in the strong operator topology uniformly on I, there exists a finite positive number M such that

$$\sup\{\| S_n(t) \|_{\mathcal{L}(X)}, \| S(t) \|_{\mathcal{L}(X)}, t \in I\} \le M.$$

It follows from assumptions (A2) and (A6) that there exists a positive constant β such that the Fréshet derivatives of F and σ are bounded above by β. By assumptions (A3)-(A4) there exist positive numbers $\{L_0, \gamma\}$ representing the essential sup of the operator norm of L and the bound of the set Γ respectively. Thus taking the expected value of the square of the X norm of $(x^o(t) - x^n(t))$, it follows from equation (5.5) that

$$\mathcal{E}|x^o(t) - x^n(t)|_X^2 \leq \eta_n(t) + \int_0^t h(r)\mathcal{E}|x^o(r) - x^n(r)|_X^2 dr \tag{5.8}$$

where the function h is given by

$$h(t) = 25\{(M\beta)^2 + (M\gamma \parallel B(t) \parallel)^2 + (M\beta)^2 trQ\}$$

and the sequence η_n is given by

$$\eta_n(t) \equiv 25\mathcal{E}\{|e_n^1(t)|_X^2 + |e_n^2(t)|_X^2\}, t \in I.$$

Consider the sequence $\{e_n^1\}$. Since $S_n(t) \stackrel{\tau_{so}}{\longrightarrow} S(t)$ in $\mathcal{L}(X)$ uniformly on I, the first three terms converge to zero P-a.s uniformly in t on I and they are also dominated by square integrable random processes. So it follows from dominated convergence theorem that the square of their X norms converges to zero uniformly in $t \in I$. Note that $y^o \equiv Lx^o + \xi^o \in B_\infty^a(I, L_2(\Omega, Y))$ and therefore, considering the fourth term of e_n^1, we have

$$\begin{aligned}
&\mathcal{E}|\int_0^t S_n(t-r)B(r)(K(r) - K_n(r))y^o(r)dr|_X^2 \\
&\leq TM^2 \int_0^T \parallel B(r) \parallel_{\mathcal{L}(U,X)}^2 \mathcal{E}\{\parallel (K(r) - K_n(r))y^o(r)|_U^2\}dr. \quad (5.9)
\end{aligned}$$

Since K_n converges to K in the Tychonoff product topology on $\mathcal{F}_{ad} \equiv B_\infty(I, \Gamma)$, the integrand converges to zero for almost all $t \in I$ P-a.s. It is also dominated by an integrable function. Hence by the dominated convergence theorem, the fourth term of e_n^1 also converges to zero and hence $\mathcal{E}|e_n^1(t)|_X^2 \to 0$ uniformly on I. Considering the process $\{e_n^2\}$ it is easy to verify that

$$\begin{aligned}
&\mathcal{E}\{|e_n^2(t)|_X^2\} \\
&= \mathcal{E}\int_0^t tr\{(S(t-r) - S_n(t-r))\sigma(x^o(r))Q\sigma^*(x^o(r)) \\
&\quad \times (S^*(t-r) - S_n^*(t-r))\}dr. \\
&= \mathcal{E}\int_0^t \parallel (S(t-r) - S_n(t-r))\sigma(x^o(r))\sqrt{Q} \parallel_{\mathcal{L}(E,X)}^2 dr. \quad (5.10)
\end{aligned}$$

By assumption Q is nuclear so compact and therefore the composition operator $\sigma(x^o(r))\sqrt{Q}$ is compact P almost surely. Thus it follows from the convergence of $S_n(t)$ to $S(t)$ in the strong operator topology uniformly on I that the integrand converges to zero for almost all $r \in [0,t]$ with probability one. Since $x^o \in B^a_\infty(I, L_2(\Omega, X))$, it follows from the assumption (A6) that the integrand in (5.10) is dominated by an integrable random process. Thus it follows from Lebesgue dominated convergence theorem that the integral converges to zero uniformly on I. Hence we conclude that the sequence $\{\eta_n\}$ is bounded on I and that it converges to zero uniformly on I. Since Q is nuclear, it follows from the expression for h given above and the assumption (A3) that $h \in L_1^+(I)$. Hence by virtue of Gronwall inequality it follows from (5.8) that $x_n \xrightarrow{s} x^o$ in the Banach space $B^a_\infty(I, L_2(\Omega, R^n))$. This completes the proof. □

Now we are prepared to develop the necessary conditions of optimality. First, we consider the Yosida regularized system and present the necessary conditions of optimality. Later, through limiting process, we derive from this result the necessary conditions for the original system. For this purpose we need the Fréchet derivatives of F and σ denoted by F_x and σ_x. These are given by

$$\begin{aligned}&(F(\zeta+h) - F(\zeta) - F_x(\zeta)h) = o(h), \text{ for } \zeta, h \in X\\&(\sigma(\zeta+h)e - \sigma(\zeta)e - \sigma_x(\zeta,h)e) = o(h) \text{ for } \zeta, h \in X \text{and } e \in E,\end{aligned}$$

where $\lim_{|h|_X \to 0}(|o(h)|_X/|h|_X) = 0$. Note that $h \longrightarrow \sigma_x(\zeta, h)$ is linear. By our assumptions (A2) and (A6), it is clear from the above expressions that there exists a finite number $\beta > 0$ such that for all $\zeta \in X$

$$\begin{aligned}&|F_x(\zeta)h|_X \le \beta|h|_X\\&|\sigma_x(\zeta,h)e|_X \le \beta|h|_X|e|_E.\end{aligned}$$

Thus $F_x(\zeta) \in \mathcal{L}(X)$ and $\sigma_x(\zeta, h) \in \mathcal{L}(E, X)$ for all $\zeta, h \in X$. Now returning to our problem, recall that for any $n \in \rho(A) \cap N$, the Yosida regularized version of system (5.2) is given by

$$dx = A_n x dt + F(x)dt + BKLxdt + BK\xi^o dt + \sigma(x)dW, x_0 \in X. \quad (5.11)$$

The objective functional is the same as (3.4) repeated here for convenience,

$$J_0(K) \equiv J(K, \xi^o) \equiv \mathcal{E}\left\{\int_I \ell(t, x(t))dt + \Phi(x(T))\right\}. \quad (5.12)$$

The problem is to find a $K^o \in \mathcal{F}_{ad}$ that minimizes the functional (5.12). The following result characterizes the optimal K^o.

Theorem 5.2. *Suppose the assumptions* (A1)-(A7) *hold and that* Γ *is convex. Then, in order that the operator* $K^o \in \mathcal{F}_{ad}$*, with the associated solution* $x^o \in B^a_\infty(I, L_2(\Omega, X))$*, be optimal it is necessary that there exists a (uniformly bounded) linear operator* $\Upsilon(x^o) \in \mathcal{L}(X)$*, dependent on* x^o*, and an associated*

$\psi \in B^a_\infty(I, L_2(\Omega, X))$ satisfying the following inequality and the pair of evolution equations:

$$\mathcal{E} \int_I < B(K - K^o)(Lx^o + \xi^o), \psi >_{X,X^*} dt \geq 0 \; \forall \; K \in \mathcal{F}_{ad}, \tag{5.13}$$

$$dx^o = [A_n x^o + F(x^o) + BK^o(Lx^o + \xi^o)]dt + \sigma(x^o)dW, x(0) = x_0, \tag{5.14}$$

$$\begin{aligned} -d\psi &= [A_n^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi + \Upsilon(x^o)\psi + \ell_x(t, x^o)]dt \\ &\quad + \sigma_x^*(x^o, \psi)dW, \\ \psi(T) &= \Phi_x(x^o(T)), \end{aligned} \tag{5.15}$$

where $F_x(x^o) \in \mathcal{L}(X)$ denotes the Fréchet derivative of F evaluated along the path $x^o = x^o(t)$ at time t and $F_x^*(x^o)$ its adjoint; $\sigma_x^*(x^o, \psi)$ is the adjoint of the Fréchet derivative of σ evaluated at x^o in the direction ψ.

Proof. Since, for each $n \in \rho(A) \cap N$, $A_n \in \mathcal{L}(X)$ and F, σ are Lipschitz, the problem (5.11) has a unique strong solution $x \in B^a_\infty(I, L_2(\Omega, X))$. For simplicity of notation, here we temporarily suppress the display of dependence on n. Let $K^o \in \mathcal{F}_{ad}$ denote the optimal feedback operator for the Problem (5.11) and (5.12) and let x^o denote the corresponding solution. Let $K \in \mathcal{F}_{ad}$, $\varepsilon \in [0, 1]$ and consider the control (operator valued function) $K^\varepsilon = K^o + \varepsilon(K - K^o)$. By convexity of Γ, it is clear that $K^\varepsilon \in \mathcal{F}_{ad}$. Let $x^\varepsilon \in B^a_\infty(I, L_2(\Omega, X))$ denote the (strong = mild) solution of the system (5.11) corresponding to $K = K^\varepsilon$. Clearly by optimality of K^o,

$$J_0(K^\varepsilon) - J_0(K^o) \geq 0 \tag{5.16}$$

for all $\varepsilon \in [0, 1]$ and all $K \in \mathcal{F}_{ad}$. Define

$$z(t) \equiv \lim_{\varepsilon \downarrow 0} (1/\varepsilon)\big(x^\varepsilon(t) - x^o(t)\big).$$

By straightforward algebra, it is easy to verify that z exists and satisfies the following evolution equation

$$\begin{aligned} dz &= A_n z dt + F_x(x^o) z dt + BK^oLzdt \\ &\quad + B(K - K^o)(Lx^o + \xi^o)dt + \sigma_x(x^o, z))dW, \\ z(0) &= 0, \end{aligned} \tag{5.17}$$

where we have suppressed the t variable. This is a linear SDE with a non homogeneous term given by $B(K - K^o)(Lx^o + \xi^o)$. It is easy to verify that as $K \to K^o$ in the τ_T (Tychonoff) topology, $z \xrightarrow{s} 0$ in $B^a_\infty(I, L_2(\Omega, X))$. Also note that under the assumptions (A3)-(A5), $B(K - K^o)(Lx^o + \xi^o) \in L^a_2(I, L_2(\Omega, X))$. By Theorem 4.1, the equation (5.17) has a unique strong (= mild) solution $z \in B^a_\infty(I, L_2(\Omega, X))$. Clearly

$$f \equiv B(K - K^o)(Lx^o + \xi^o) \longrightarrow z \tag{5.18}$$

is a continuous linear map from $L_2^a(I, L_2(\Omega, X))$ to $B_\infty^a(I, L_2(\Omega, X))$. Thus there exists a bounded strongly measurable stochastic evolution operator $\Psi(t, s), 0 \leq s \leq t \leq T$ such that the linear integral operator,

$$z(t) = \int_0^t \Psi(t, s) f(s) ds, t \in I,$$

maps $L_2^a(I, L_2(\Omega, X))$ into $B_\infty^a(I, L_2(\Omega, X))$. By assumption (A7) both ℓ and Φ are continuously Fŕchet differentiable. Hence along the path x^o, $\ell_x(\cdot, x^o(\cdot)) \in L_1^a(I, L_2(\Omega, X))$ and $\Phi_x(x^o(T)) \in L_2(\Omega, X)$. Dividing the inequality (5.16) by $\varepsilon \in (0, 1]$ and letting $\varepsilon \to 0$, it follows from the properties of $\{\ell, \Phi\}$ mentioned above that the Gateaux derivative of J_0 at K^o in the direction $K - K^o$ is well defined and it satisfies the following inequality

$$\begin{aligned} & dJ_0(K^o, K - K^o) \\ & = \mathcal{E} \int_I (\ell_x(t, x^o(t)), z(t))_X dt + (\Phi_x(x^o(T)), z(T))_X \geq 0 \end{aligned} \tag{5.19}$$

for all $K \in \mathcal{F}_{ad}$. In view of the properties of ℓ_x and Φ_x, it is clear that

$$z \longrightarrow L(z) \equiv \mathcal{E} \int_I (\ell_x(t, x^o(t)), z(t))_X dt + (\Phi_x(x^o(T)), z(T))_X \tag{5.20}$$

is a continuous linear functional. Thus the composition map

$$f \equiv B(K - K^o)(Lx^o + \xi^o) \longrightarrow z \longrightarrow L(z)$$

is a continuous linear functional on $L_2^a(I, L_2(\Omega, X))$. Hence by Riesz representation theorem, there exists a $\psi \in L_2^a(I, L_2(\Omega, X))$ such that

$$L(z) = \mathcal{E} \int_I (B(K - K^o)(Lx^o + \xi^o), \psi)_X dt. \tag{5.21}$$

Hence it follows from (5.19)-(5.21) that

$$\begin{aligned} & dJ_o(K^o, K - K^o) \\ & = \mathcal{E} \int_I (B(K - K^o)(Lx^o + \xi^o), \psi)_X dt \geq 0 \ \forall \ K \in \mathcal{F}_{ad}. \end{aligned} \tag{5.22}$$

This proves the necessary condition (5.13). The necessary condition (5.14) giving x^o is just the solution of the state equation corresponding to A_n and K^o, so nothing to prove. It remains to verify (5.15). Using the variational equation (5.17) into the expression (5.21), we obtain

$$\begin{aligned} & L(z) \\ & = \mathcal{E}\Big\{ \int_I (dz, \psi) - \big(z, [A_n^* \psi + F_x^*(x^o)\psi + (BK^o L)^* \psi] dt + \sigma_x^*(x^o, \psi) dW\big)\Big\}. \end{aligned} \tag{5.23}$$

Since all the operators are bounded, Itô differential rule applies. Using the Itô differential rule given by

$$d(z,\psi) = (dz,\psi) + (z,d\psi) + < dz, d\psi >,$$

where the last term denotes the quadratic variation term, it follows from (5.23) that

$$L(z) = \mathcal{E}\Big\{ \int_I d(z,\psi) - \big(z, d\psi + [A_n^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi]dt + \sigma_x^*(x^o,\psi)dW\big) - \int_I < dz, d\psi > \Big\}. \tag{5.24}$$

Since it is only the martingale terms that contribute to the quadratic variation, we have

$$\mathcal{E}\int_I < dz, d\psi > = \mathcal{E}\int_I tr(\sigma_x(x^o,z)Q\sigma_x^*(x^o,\psi))dt \tag{5.25}$$

where $Q \in \mathcal{L}_1^+(E)$ (space of positive nuclear operators on the Hilbert space E) is the incremental covariance of the Wiener process W. Note that, given $x^o(t)$,

$$(\zeta,\eta) \longrightarrow tr(\sigma_x(x^o(t),\zeta)Q\sigma_x^*(x^o(t),\eta))$$

is a symmetric bilinear map from $X \times X$ to R. Thus there exists a bounded linear operator $\Upsilon_Q(x^o(t)) \in \mathcal{L}(X)$ (dependent on x^o) such that

$$(\zeta,\eta) \longrightarrow tr(\sigma_x(x^o(t),\zeta)Q\sigma_x^*(x^o(t),\eta)) = (\Upsilon_Q(x^o(t))\zeta,\eta)_X.$$

It follows from assumption (A6) that the map $x \longrightarrow \Upsilon(x)$ is continuous and uniformly bounded on X. Using this operator, the expression (5.24) can be rewritten as

$$L(z) = \mathcal{E}\Big\{ \int_I d(z,\psi) - \big(z, d\psi + [A_n^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi]dt + \sigma_x^*(x^o,\psi)dW\big) - (z, \Upsilon_Q(x^o)\psi)dt \Big\}. \tag{5.26}$$

Now integrating by parts the first term on the right hand side of equation (5.26) and setting

$$\begin{aligned} -d\psi &= [A_n^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi + \Upsilon_Q(x^o(t))\psi + \ell_x(t,x^o)]dt \\ &\quad + \sigma_x^*(x^o,\psi)dW \\ \psi(T) &= \Phi_x(x^o(T)), \end{aligned} \tag{5.27}$$

we obtain

$$L(z) = \mathcal{E}\Big\{ \int_I (z(t), \ell_x(t,x^o(t))dt + (z(T), \Phi_x(x^o(T)) \Big\}. \tag{5.28}$$

This is precisely the functional (5.20) which was originally obtained through the variation of the cost functional. This verifies consistency. Thus we have obtained the stochastic differential equation (5.27) which coincides with the necessary condition given by the adjoint equation (5.15). This completes the proof. □

To proceed further we must now indicate the dependence of the optimal variables $\{K^o, x^o, \psi\}$ on $n \in N$ associated with the regularized problem (5.11) and (5.12). From hereon these will be denoted by $\{K_n^o, x_n^o, \psi_n\}$. In other words, the logical notation for the complete optimality system given by (5.13)-(5.15) is as follows:

$$\mathcal{E}\int_I < B(K - K_n^o)(Lx_n^o + \xi^o), \psi_n >_X dt \geq 0 \;\forall\; K \in \mathcal{F}_{ad}, \tag{5.29}$$

$$dx_n^o = [A_n x_n^o + F(x_n^o) + BK_n^o(Lx_n^o + \xi^o)]dt + \sigma(x_n^o)dW, x(0) = x_0, \tag{5.30}$$

$$\begin{aligned} -d\psi_n &= [A_n^*\psi + F_x^*(x_n^o)\psi_n + (BK_n^o L)^*\psi_n + \Upsilon_Q(x_n^o)\psi_n \\ &\quad + \ell_x(t, x_n^o)]dt + \sigma_x^*(x_n^o, \psi_n)dW, \\ \psi_n(T) &= \Phi_x(x_n^o(T)). \end{aligned} \tag{5.31}$$

We must now prove that the optimality system for the original problem (5.2) with the cost functional (5.12) is given by the following theorem.

Theorem 5.3. *Suppose the assumptions* (A1)-(A7) *hold and that* Γ *is a convex subset of* $\mathcal{L}(Y, U)$ *and* $\mathcal{F}_{ad}$ *the corresponding class of admissible feedback operators. Then, in order that the operator* $K^o \in \mathcal{F}_{ad}$, *with the associated solution* $x^o \in B_\infty^a(I, L_2(\Omega, X))$, *be optimal it is necessary that there exists a (uniformly bounded) linear operator* $\Upsilon_Q(x^o) \in \mathcal{L}(X)$, *dependent on* x^o, *and an associated* $\psi \in B_\infty^a(I, L_2(\Omega, X))$ *satisfying the following inequality and the pair of evolution equations:*

$$\mathcal{E}\int_I < B(K - K^o)(Lx^o + \xi^o), \psi >_X dt \geq 0 \;\forall\; K \in \mathcal{F}_{ad}, \tag{5.32}$$

$$dx^o = [Ax^o + F(x^o) + BK^o(Lx^o + \xi^o)]dt + \sigma(x^o)dW, x(0) = x_0, \tag{5.33}$$

$$\begin{aligned} -d\psi &= [A^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi + \Upsilon_Q(x^o)\psi + \ell_x(t, x^o)]dt \\ &\quad + \sigma_x^*(x^o, \psi)dW, \\ \psi(T) &= \Phi_x(x^o(T)), \end{aligned} \tag{5.34}$$

where $F_x(x^o) \in \mathcal{L}(X)$ *denotes the Fréchet derivative of* F *evaluated along the path* $x^o = x^o(t)$ *and* $F_x^*(x^o)$ *its adjoint;* $\sigma_x^*(x^o, \psi)$ *is the Fréchet derivative of* $\sigma^*(x)$ *evaluated at* x^o *in the direction* ψ.

Proof. We use the sequence of optimality systems (5.29)-(5.31) corresponding to the Yosida regularized problem and prove that in the limit they satisfy the inequality (5.32) and the evolution equations (5.33)-(5.34). Recall that by Theorem 4.1, the evolution equation (5.33) has a unique mild solution $x^o \in B^a_\infty(I, L_2(\Omega, X))$ and further the solution set corresponding to the admissible set $\mathcal{F}_{ad}$ is also a bounded subset of $B^a_\infty(I, L_2(\Omega, X))$. We start with the sequence $\{K^o_n, x^o_n, \psi_n\} \in \mathcal{F}_{ad} \times B^a_\infty(I, L_2(\Omega, X)) \times B^a_\infty(I, L_2(\Omega, X))$ which is optimal for the sequence of Yosida regularized system (5.29)-(5.31). Since $\mathcal{F}_{ad}$ is compact in the τ_T topology there exists a generalized subsequence (subnet), relabeled as the original sequence, and an element $K^o \in \mathcal{F}_{ad}$ such that

$$K^o_n \xrightarrow{\tau_T} K^o \text{ in } \mathcal{F}_{ad} \tag{5.35}$$

as $n \to \infty$. Let $x^o \in B^a_\infty(I, L_2(\Omega, X))$ denote the mild solution of equation (5.2) (equivalently (5.33)) corresponding to the pair $\{A, K^o\}$. Then it follows from Lemma 5.1 that

$$x^o_n \xrightarrow{s} x^o \text{ in } B^a_\infty(I, L_2(\Omega, X)). \tag{5.36}$$

Similarly, considering the sequence of adjoint systems (5.31), we can prove that, along a subsequence if necessary,

$$\psi_n \xrightarrow{s} \psi \text{ in } B^a_\infty(I, L_2(\Omega, X)) \tag{5.37}$$

where ψ is the mild solution of the adjoint evolution equation (5.34). The fact that it is $\mathcal{F}_t$ progressively measurable follows from the Riesz representation theorem leading to the expression (5.21). The proof is very similar to that of Lemma 5.1. We present only a brief outline. The mild solution of equation (5.34) is given by the solution of the backward stochastic integral equation

$$\begin{aligned}
\psi(t) = S^*(T-t)\Phi_x(x^o(T)) &+ \int_t^T S^*(T-s)F_x(x^o(s))\psi(s)ds \\
&+ \int_t^T S^*(T-s)(BK^oL)(s)^*\psi(s)ds \\
&+ \int_t^T S^*(T-s)\Upsilon(x^o(s))\psi(s)ds \\
&+ \int_t^T S^*(T-s)\ell_x(s, x^o(s))ds \\
&+ \int_t^T S^*(T-s)\sigma^*_x(x^o(s), \psi(s))dW(s).
\end{aligned} \tag{5.38}$$

Similarly the mild solution of equation (5.31) is given by the solution of the integral equation

$$\psi_n(t) = S_n^*(T-t)\Phi_x(x_n^o(T)) + \int_t^T S_n^*(T-s)F_x(x_n^o(s))\psi_n(s)ds$$
$$+ \int_t^T S_n^*(T-s)(BK_n^o L)(s)^*\psi_n(s)ds$$
$$+ \int_t^T S_n^*(T-s)\Upsilon(x_n^o(s))\psi_n(s)ds$$
$$+ \int_t^T S^*(T-s)\ell_x(s, x_n^o(s))ds$$
$$+ \int_t^T S_n^*(T-s)\sigma_x^*(x_n^o(s), \psi_n(s))dW(s). \tag{5.39}$$

We have seen in the course of the proof of Theorem 5.2 that by virtue of Riesz representation theorem the adjoint process ψ belongs to the class $L_2^a(I, L_2(\Omega, X))$ and further it follows from (5.37) that $\psi \in B_\infty^a(I, L_2(\Omega, X))$. Thus the identities in (5.38) and (5.39) are to be understood in the sense that all the terms on the righthand side of these identities are the projections on to the sigma algebra $\mathcal{F}_t$ for each $t \in I$. For notational convenience we can avoid this since the proof is independent of the additional notations. Now subtracting (5.39) from (5.38) and carrying out long but straight forward algebra and following, by and large, similar steps as in Lemma 5.1, we arrive at the conclusion

$$\psi_n \xrightarrow{s} \psi \text{ in } B_\infty^a(I, L_2(\Omega, X)). \tag{5.40}$$

For convenience of the reader we list only the ingredients used for the detailed proof. These are (p1): $A_n \xrightarrow{\tau_{so}} A$ on $D(A)$ implies $S_n(t) \xrightarrow{\tau_{so}} S(t)$ on X uniformly on I. Since X is a Hilbert space, the adjoint semigroups are also continuous in the strong operator topology so that $S_n^*(t) \xrightarrow{\tau_{so}} S^*(t)$ uniformly on I. (p2): By assumption (A7), $\zeta \longrightarrow \Phi_x(\zeta)$ is bounded and continuous on X. (p3): $\eta \longrightarrow F_x(\eta)$ is continuous on X and uniformly bounded as elements of $\mathcal{L}(X)$ and hence as $x_n^o(t) \xrightarrow{s} x^o(t)$ in X for each $t \in I$, P-a.s., we have $F_x(x_n^o(t)) \longrightarrow F_x(x^o(t))$ again P-a.s. (p4): $B_\infty(I, \Gamma) \subset B_\infty(I, \mathcal{K}(Y, U))$ is a bounded set and compact in the Tychonoff product topology and hence, along a generalized subsequence (subnet) if necessary, $K_n^o \xrightarrow{\tau_T} K^o$. Thus by the definition of Tychonoff product topology and Mayoral's compactness result (Lemma 4.2), we have $K_n^o(t) \xrightarrow{\tau_{so}} K^o(t)$ for each $t \in I$. (p5): Recalling the definition of the operator Υ_Q given by

$$(\zeta, \eta) \longrightarrow tr(\sigma_x(x^o(t), \zeta)Q\sigma_x^*(x^o(t), \eta)) = (\Upsilon_Q(x^o(t))\zeta, \eta)_X,$$

it follows from assumption (A6) that this (nonlinear) operator is uniformly bounded and $x \longrightarrow \Upsilon_Q(x)$ is continuous from X to $\mathcal{L}(X)$. (p6): $\zeta \longrightarrow \ell_x(t,\zeta)$ from X to X is continuous and it follows from our assumption (A7) that it is square integrable and dominated by such functions. (p7): By assumption (A6), $\parallel \sigma_x(z,h) \parallel_{\mathcal{L}(E,X)} \leq \beta |h|_X$ independently of $z \in X$ and that $z \longrightarrow \sigma_x(z,h)$ is continuous. (p8): Finally by assumption (A3), $\parallel L \parallel \equiv \{|L(t)|_{\mathcal{L}(X,Y)}, t \in I\} \equiv \hat{L} < \infty$. Using these facts in the analysis combined with Gronwall inequality we arrive at the conclusion stating (5.37). Thus $\{x^o, \psi\} \in B_\infty^a(I, L_2(\Omega, X))$ are the mild solutions of the evolution equations (5.33) and (5.34) respectively. Now it is easy to show that the inequality (5.32) follows from the inequality (5.29). Indeed define the function

$$H_o : I \times \mathcal{K}(Y,U) \times X \times X \longrightarrow R$$

by

$$H_o(t,K,\zeta,\eta) \equiv \big(B(t)K(L(t)\zeta + \xi^o(t)), \eta\big)_X.$$

Subtracting the expression on the left of (5.29) from the expression on the left of (5.32) we obtain

$$\begin{aligned} \mathcal{E}\Big\{ \int_I & H_o(t, K_n^o(t) - K^o(t), x^o(t), \psi(t))dt \\ & + H_o(t, K(t) - K_n^o(t), x^o(t) - x_n^o(t), \psi(t))dt \\ & + \int_I H_o(t, K(t) - K_n^o(t), x_n^o(t), \psi(t) - \psi_n(t))dt \Big\}, \end{aligned} \tag{5.41}$$

which is defined for all $K \in \mathcal{F}_{ad}$. In view of the boundedness of the solution set as seen in Lemma 4.1, which implies boundedness of the set $\{\psi_n\} \subset B_\infty^a(I, L_2(\Omega, X))$, and the assumption (A3) giving $B \in L_2(I, \mathcal{L}(U,X))$, the integrands are all dominated by integrable functions for almost all $t \in I$ and P-a.s. By virtue of (5.35) the first integrand converges to zero for almost all $t \in I$, P-a.s. Since Γ is a bounded set, it follows from (5.36) that the second integrand converges to zero for almost all $t \in I$, P-a.s. Similarly, by virtue of (5.37) and boundedness of the sequence $\{K_n^o, x_n^o\} \in \mathcal{F}_{ad} \times B_\infty^a(I, L_2(\Omega, X))$ the third integrand converges to zero for almost all $t \in I$, P-a.s. Thus by letting $n \to \infty$ it follows from the Lebesgue dominated convergence theorem that the expression (5.41) converges to zero. Since the inequality (5.29) holds for all $n \in N$ and $K \in \mathcal{F}_{ad}$, it follows from the above convergence results that the inequality (5.32) holds for all $K \in \mathcal{F}_{ad}$. This proves that the necessary conditions of optimality for the original problem are given by (5.32)-(5.35) . □

Remark 5.4. It is interesting to note that, for the necessary conditions of optimality, we can relax the Mayoral type [Lemma 4.2] compactness of the set $\Gamma \subset \mathcal{L}(Y,U)$ to mere compactness in the strong operator topology while retaining the Tychonoff product topology on $B_\infty(I,\Gamma)$.

We used the right hand inequality of (5.1) to characterize the optimality of the (feedback) operator valued function $K^o \in \mathcal{F}_{ad}$. Now we consider the problem of characterizing the worst case disturbance ξ^o and for this we use the left hand inequality of (5.1). Combining the two results later we will have the complete optimal policy. This is presented in the following theorem.

Theorem 5.5. *Consider the system* (4.1) *corresponding to a fixed* $K^o \in \mathcal{F}_{ad}$ *and* $\xi \in \mathcal{D}$ *arbitrary and suppose the assumptions* (A1)-(A7) *hold. Then, in order that* $\xi^o \in \mathcal{D}$, *with the corresponding solution* $\tilde{x}^o \in B_\infty^a(I, L_2(\Omega, X))$, *be the extremal (worst disturbance) it is necessary that there exists an adjoint process* $\varphi \in B_\infty^a(I, L_2(\Omega, X))$ *such that the triple* $\{\tilde{x}^o, \varphi, \xi^o\}$ *satisfy the following evolution equations,*

$$d\tilde{x}^o = A\tilde{x}^o dt + F(\tilde{x}^o)dt + BK^o(L\tilde{x}^o + \xi^o)dt + \sigma(\tilde{x}^o)dW, \tilde{x}^o(0) = x_0, \quad (5.42)$$

in X,

$$-d\varphi = \big(A^*\varphi + F_x^*(\tilde{x}^o)\varphi + (BK^oL)^*\varphi) + \Upsilon_Q(\tilde{x}^o)\varphi\big)dt + \ell_x(t,\tilde{x}^o)dt$$
$$+\sigma_x^*(\tilde{x}^o(t),\psi(t))dW, \ \varphi(T) = \Phi_x(\tilde{x}^o(T)) \text{ in } X, \quad (5.43)$$

and the inequality given by

$$\mathcal{E}\int_I < BK^o(\xi^o - \xi), \varphi >_X dt \geq 0 \ \forall \ \xi \in \mathcal{D}. \quad (5.44)$$

Proof. The proof of this result is entirely similar to that of Theorem 5.3 and again this is based on the Yosida regularization as seen in Theorem 5.2. To avoid the repetition we omit it. □

Now we are ready to present our final result. Combining the results of Theorem 5.3 and Theorem 5.5 we obtain the necessary conditions of optimality for our original problem, the min-max problem (3.5) subject to (3.1)-(3.4). This is stated in the following theorem.

Theorem 5.6. *Consider the feedback system* (4.1) *corresponding to* $K \in \mathcal{F}_{ad}$ *and* $\xi \in \mathcal{D}$ *and suppose the assumptions* (A1)-(A7) *hold. Then, in order that the pair* $(K^o, \xi^o) \in \mathcal{F}_{ad} \times \mathcal{D}$, *with the corresponding solution* $x^o \in B_\infty^a(I, L_2(\Omega, X))$, *be optimal for the min-max problem* (3.5) *subject to* (3.1)-(3.4), *it is necessary that there exists an adjoint process* $\psi \in B_\infty^a(I, L_2(\Omega, X))$

such that the quadruple $\{x^o, \psi, K^o, \xi^o\}$ *satisfy the following evolution equations:*

$$dx^o = Ax^o dt + F(x^o)dt + BK^o(Lx^o + \xi^o)dt + \sigma(x^o)dW, x^o(0) = x_0, \quad (5.45)$$
$$\text{in } X,$$
$$-d\psi = \big(A^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi) + \Upsilon_Q(x^o)\psi\big)dt + \ell_x(t, x^o)dt$$
$$+\sigma_x^*(x^o(t), \psi(t))dW, \ \psi(T) = \Phi_x(x^o(T)) \text{ in } X, \quad (5.46)$$

and the following inequalities:

$$\mathcal{E}\int_I < B(K - K^o)(Lx^o + \xi^o), \psi >_X dt \geq 0 \ \forall \ K \in \mathcal{F}_{ad}, \quad (5.47)$$

$$\int_I < BK^o(\xi - \xi^o), \psi >_X dt \leq 0 \ \forall \ \xi \in \mathcal{D}. \quad (5.48)$$

Proof. Let $\{K^o, \xi^o\}$ be the optimal pair. Then comparing the state equations (5.33) and (5.42), it is clear that they are one and the same equation with the same initial condition. So by virtue of uniqueness, we have $x^o = \tilde{x}^o$ giving the state equation (5.45). Given this, comparing the adjoint equations (5.34) and (5.43) we observe that they are again one and the same equation with the same terminal condition. Thus $\psi = \varphi$ giving the adjoint equation (5.46). From this it is clear that the inequality (5.47) is the same as (5.32), and (5.48) coincides with (5.44). This proves all the necessary conditions of optimality as stated. □

Using the inequality (5.48), the necessary conditions of optimality given by Theorem 5.6 can be simplified as follows.

Corollary 5.7. *Consider the optimal feedback control problem* (3.4)-(3.5) *subject to the system dynamics* (4.1). *Then, for* $K^o \in \mathcal{F}_{ad}$ *to be optimal it is necessary that there exists a pair* $\{x^o, \psi\} \in B_\infty^a(I, L_2(\Omega, X))$ *satisfying the following evolution equations:*

$$dx^o = Ax^o dt + F(x^o)dt + BK^o(Lx^o + \eta((BK^o)^*\psi))dt + \sigma(x^o)dW, \quad (5.49)$$
$$x^o(0) = x_0, \ \text{in } X$$
$$-d\psi = \big(A^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi) + \Upsilon_Q(x^o)\psi\big)dt + \ell_x(t, x^o)dt$$
$$+\sigma_x^*(x^o(t), \psi(t))dW, \ \psi(T) = \Phi_x(x^o(T)) \text{ in } X, \quad (5.50)$$

and the following inequality:

$$\mathcal{E}\int_I < B(K - K^o)[Lx^o + \eta((BK^o)^*\psi)], \psi >_X dt \geq 0, \ \forall \ K \in \mathcal{F}_{ad}, \quad (5.51)$$

where

$$\eta(y) = \begin{cases} (y/|y|_Y) & \text{for } y(\neq 0) \in Y \\ 0 & \text{for } y = 0. \end{cases}$$

Proof. It follows from the inequality (5.48) that ξ^o must maximize the linear functional

$$\rho(\xi) \equiv \mathcal{E} \int_I < BK^o\xi, \psi >_X dt = \mathcal{E} \int_I < \xi, (BK^o)^*\psi >_Y dt \tag{5.52}$$

on $\mathcal{D} \equiv L_\infty(I, B_1(Y))$. Clearly, for almost all $t \in I$ and P-a.s, the functional

$$\xi \longrightarrow g_t(\xi) \equiv < \xi, (B(t)K^o(t))^*\psi(t) >_Y$$

is a continuous linear functional on Y and so weakly continuous. Since Y is a Hilbert space, the unit ball $B_1(Y)$ is weakly compact and hence g_t attains its maximum on the boundary $\partial B_1(Y)$ and the maximizer is given by $\zeta \equiv \eta((B(t)K^o(t))^*\psi(t))$. This is defined for almost all $t \in I$ and P-a.s. Thus we may define the function

$$\zeta(t) \equiv \eta((B(t)K^o(t))^*\psi(t)), \text{ for almost all } t \in I, P - a.s.$$

Clearly, it follows from the measurability of $\{B, K^o, \psi\}$ and continuity of the map η that the Y valued function $t \longrightarrow \zeta(t)$ is measurable. We take $\xi^o = \zeta$. Substituting this expression for ξ^o in equations (5.45)-(5.47) we obtain the necessary conditions (5.49)-(5.51). This completes the proof. □

Remark 5.8. Note that the state and the adjoint processes $\{x^o, \psi\}$ are given by the solutions of the two point stochastic boundary value problems (2PS-BVP) (5.49)-(5.50).

6. Pointwise Necessary Conditions & Computational Algorithm

From the necessary conditions of optimality given by Theorem 5.6, we can easily derive the following pointwise necessary conditions of optimality.

Corollary 6.1. *Suppose the assumptions of Theorem* 5.6 *hold. Then the necessary conditions of optimality given by* (5.47)-(5.48) *are equivalent to the necessary conditions given by the following inequalities:*

$$< B(t)(K - K^o(t))(L(t)x^o(t) + \xi^o(t)), \psi(t) >_X \geq 0, \ \forall \ K \in \Gamma, \tag{6.1}$$

$$< B(t)K^o(t)(\xi - \xi^o(t)), \psi(t) >_X \leq 0, \ \forall \ \xi \in B_1(Y) \tag{6.2}$$

which hold $dt \times dP$ a.e on predictable subsets of the set $I \times \Omega$ subject to the dynamic constraints (5.45)-(5.46). *Further, the inequalities* (6.1)-(6.2) *are equivalent to the single inequality*

$$< (K - K^o)[Lx^o + \eta((B(t)K^o(t))^*\psi(t))], B^*(t)\psi(t) >_U \geq 0 \ \forall \ K \in \Gamma. \tag{6.3}$$

Proof. Using Lebesgue density argument with respect to the $dt \times dP$ measure on the sigma algebra of $\mathcal{F}_t$ predictable subsets of the set $I \times \Omega$ one can easily derive the inequalities (6.1) and (6.2) from (5.47) and (5.48) respectively. In other words, the set of evolution equations (5.45)-(5.46) and the inequalities (6.1)-(6.2) constitute the point wise necessary conditions of optimality. Further, it follows from the inequality (6.2) that the worst case uncertainty ξ^o must be given by the expression $\xi^o(t) = \eta((B(t)K^o(t))^*\psi(t))$ where η is the retraction of the unit ball $B_1(Y)\setminus\{0\}$ (as given in Corollary 5.7). Substituting this in the inequality (6.1) we arrive at the expression (6.3). This completes the proof. □

A Conceptual algorithm:

Using the inequality (6.3) and the state and the adjoint equations (5.49)-(5.50) one can develop a computational algorithm and determine the extremal $K^o(t)$. But this will produce an operator valued random process K^o which is very difficult, if not impossible, to physically implement. Our objective here is to find an optimal deterministic operator valued function. To achieve this goal we use Fubinins theorem in (5.51) and Lebesgue density argument (now with respect to the dt measure only) to derive the following inequality

$$\mathcal{E}\big\{< (K - K^o(t))y(t, K^o(t), x^o(t), \psi^o(t)), B^*(t)\psi^o(t) >_U\big\} \geq 0, \quad (6.4)$$
$$\forall\ K \in \Gamma, a.e\ t \in I,$$

where

$$y(t, K^o(t), x^o(t), \psi^o(t)) \equiv [L(t)x^o(t) + \eta((B(t)K^o(t))^*\psi^o(t))]$$

and

$$\{x^o, \psi^o\}$$

are the mild solutions of the evolution equations (5.49)-(5.50). For any $y \in Y$ and $u \in U$ we define the linear operator $u\otimes y \in \mathcal{L}(Y, U)$ by setting $(u\otimes y)(z) = u\ (y, z)_Y$ for every $z \in Y$. Clearly it is a nuclear operator. Using the tensor product we define the operator valued random process

$$\Xi^o(t) \equiv \Xi(t, x^o, \psi^o, K^o) \equiv \big((B^*(t)\psi^o(t)) \otimes y(t, K^o(t), x^o(t), \psi^o(t))\big), t \in I.$$

It follows from the assumptions on the operator valued functions $\{B, L\}$ and the fact, $\{x^o, \psi^o\} \in B^a_\infty(I, L_2(\Omega, X))$, that $\Xi^o(t) \in \mathcal{L}_1(Y, U)$ (nuclear) for almost all $t \in I$ with probability one. In fact even more is true, $\Xi^o \in L^a_2(I, L_2(\Omega, \mathcal{L}_1(Y, U)))$. Using this operator we can rewrite the inequality (6.4) compactly as follows:

$$\mathcal{E}\big\{Tr\big((K - K^o(t))^*\ \Xi^o(t)\big)\big\} \geq 0\ \forall\ K \in \Gamma, t \in I. \quad (6.5)$$

Since our objective is to find an optimal deterministic feedback operator valued function, and the expectation and trace operations are both linear, we can rewrite the inequality (6.5) as

$$Tr\big((K - K^o(t))^*\mathcal{E}(\Xi^o(t))\big) \equiv Tr\big((K - K^o(t))^* R^o(t)\big) \geq 0, \ \forall\ K \in \Gamma, \quad (6.6)$$

where $R^o(t) \equiv \mathcal{E}(\Xi^o(t))$. Here the integration $\mathcal{E}(\cdot)$ is understood in the sense of Gelfand (weak star integral) where the Banach spaces in duality are $\mathcal{L}(Y, U)$ and $\mathcal{L}_1(Y, U)$, the later denoting the space of nuclear operators. Hence $R^o(t)$ is a well defined element of $\mathcal{L}_1(Y, U)$ for almost all $t \in I$.

Based on the expression (6.6) one can construct a gradient type algorithm for approximately determining the optimal K. Given the $K_n \in \mathcal{F}_{ad}$ at the n-th stage, one obtains $\{x_n, \psi_n\}$ solving the coupled system of equations (5.49)-(5.50) corresponding to K_n. This gives the function $y_n(t) = y(t, K_n(t), x_n(t), \psi_n(t)), t \in I$, and hence the operator

$$R_n(t) \equiv \mathcal{E}(\Xi_n(t)) \equiv \mathcal{E}\{(B^*(t)\psi_n(t)) \otimes y_n(t)\}, t \in I.$$

The update for K at the $(n + 1)$-th stage is then given by

$$K_{n+1}(t) = K_n(t) - \varepsilon R_n(t), t \in I.$$

The choice of the step size $\varepsilon(\geq 0)$ is determined by the desired speed of convergence and the requirement that $K_{n+1}(t) \in \Gamma$ for $t \in I$. Thus at the n-th stage, ε may depend on n. Therefore, choosing $\varepsilon_n > 0$ sufficiently small, it follows from the definition of Gateaux derivative and (5.51) that

$$J_0(K_{n+1}) = J_0(K_n) - \varepsilon_n \int_I \parallel R_n(t) \parallel^2_{\mathcal{L}(U,Y)} dt + o(\varepsilon_n), \quad (6.7)$$

where J_0 is given by the functional (5.12) evaluated at

$$\xi^n(t) \equiv \eta((BK_n)^*(t)\psi_n(t))$$

replacing $\xi^o(t)$. From this it is evident that for a suitable choice of the step size $\varepsilon_n > 0$, the sequence $\{J_0(K_n)\}$ is monotone decreasing. Under the assumption (A7), $J_0(K) > -\infty$ for all $K \in \Gamma$. Thus $J_0(K_n)$ converges possibly to a local minimum. By our construction of the sequence $\{K_n(t), t \in I\} \subset \Gamma$ and by hypothesis (H1) of Theorem 4.3 and Mayoral's compactness (Lemma 4.2), there exists a subnet (generalized subsequence) of the sequence $\{K_n\}$ that converges to the optimal K^o in the Tychonoff's product topology on $B_\infty(I, \Gamma)$. Similar algorithm has been used successfully for finite dimensional problems illustrated by several numerical results presented in our paper [7].

Some Special Cases: (A) If the diffusion related operator valued function σ is independent of the state variable x, the necessary conditions given by the Corollary 5.7 simplify substantially. The adjoint equation reduces to the following proper differential equation on X

$$-(d\psi/dt) = \big(A^*\psi + F_x^*(x^o)\psi + (BK^oL)^*\psi\big) + \ell_x(t,x^o),$$
$$\psi(T) = \Phi_x(x^o(T)) \text{ in } X \tag{6.8}$$

with coefficients which are either operator or vector valued random processes such as $F_x^*(x^o(t))$ and $\ell_x(t,x^o(t))$ and the terminal condition $\Phi_x(x^o(T))$ etc. The necessary condition given by the inequality (5.51) however remains valid.

(B) Let us compare the above result with its deterministic counterpart where $\sigma \equiv 0$ and x_0 is deterministic. In this case the state equation (5.49) also reduces to a proper differential equation and the adjoint equation is again given by the equation (6.8) with deterministic coefficients. The necessary condition given by the inequality (5.51) remains unchanged. Therefore, the pointwise necessary condition given by the inequality (6.3) is valid for all these cases.

Some Open problems: In recent years fractional Brownian motion has been used in the study of filtering and control theory.

(A): Recently Ducan, Maslowski and Pasik-Duncan [15] have used fractional Brownian motion in their study of linear quadratic control problems on infinite dimensional Hilbert spaces. They assume the principal operator to be the infinitesimal generator of an analytic semigroup. This permits inclusion of unbounded operators for observation as well as control. In this paper we have considered general semilinear systems driven by Brownian motion (with the principal operator being the infinitesimal generator of a C_0-semigroup) on infinite dimensional Hilbert spaces. It would be interesting and useful to extend our results to systems driven by Fractional Brownian motion.

(B): In the study of linear filtering for finite dimensional systems, fractional Brownian motion, both for the dynamic system and the observation process, have been used [14]. To the best of knowledge of the author no extension of these results to infinite dimensional Hilbert spaces has been reported so far.

Acknowledgement. This research is partially supported by the National Science and Engineering Research Council of Canada under discovery grant no A7101.

References

[1] N.U. Ahmed, *Existence of Optimal Output Feedback Control Law for a Class of Uncertain Infinite Dimensional Stochastic Systems* : A Direct Approach, Evolution Equations and Control Theory (EECT), 1(2),(2012), 235-250.

[2] N.U. Ahmed, Semigroup Theory with Application to Systems and Control, Pitman research Notes in Mathematics series, 246, Longman Scientific and Technical, U.K; and Co-published with John-Wiely & Sons, Inc. New York (1991).

[3] N.U. Ahmed and X. Xiang, *Differential Inclusions on Banach Spaces and Their Optimal Control*, Nonlinear Funct. Anal.& Appl., **8(3)** (2003), 461–488.
[4] N.U. Ahmed, *Optimal Relaxed Controls for Systems Governed by Impulsive Differential Inclusions,* Nonlinear Funct. Anal.& Appl., **10(3)** (2005), 427–460.
[5] N.U. Ahmed and C.D. Charalambous, *Minimax Games for Stochastic Systems subject to Relative Entropy Uncertainty*: Applications to SDE's on Hilbert Spaces, Mathematics of Control, Signals and Systems, **19** (2007), 65–91.
[6] N.U. Ahmed, *Optimal Output Feedback Boundary Control for Systems Governed by Semilinear Parabolic Inclusions*: Uncertain Systems, Advances in Nonlinear Variational Inequalities, **11(1)** (2008), 61–79.
[7] N.U. Ahmed and Suruz Miah, *Optimal Feedback Control Law for a Class of Partially Observed Dynamic Systems: A Min-max Problem,* Dynamic Systems and Applications, **20** (2011), 149–168.
[8] N.U. Ahmed and K.L. Teo, Optimal Control of Distributed Parameter Systems, North Holland, New York, Oxford (1981).
[9] L. Cesari, Optimization Theory and Applications, Springer-Verlag (1983).
[10] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press (1992).
[11] H.O. Fattorini, Infinite Dimensional optimization and Control Theory, Encyclopedia of mathematics and its applications, 62, Cambridge University Press (1999).
[12] F. Mayoral, *Compact Sets of Compact Operators in Absence of* ℓ_1, Proc. Am. Math. Soc., **129(1)** (2001), 79–82.
[13] E. Serrano, C. Pineiro and J.M. Delgado, *Equicompact sets of Operators Defined on Banach Spaces,* Proc. of the American Mathematical Society, **134(3)** (2005), 689–695.
[14] N.U. Ahmed and C.D. Charalambos, *Filtering for Linear Systems Driven by Fractional Brownian Motion,* SIAM J. Control Optim., **41(1)** (2002), 313–330.
[15] T.E. Duncan, B. Maslowski and B. Pasik-Duncan, Linear-Quadratic Control for Stochastic Equations in a Hilbert Space with Fractional Brownain Motions, **50(1)** (2012), 507–531.
[16] F. Gozzi and E. Rouy, *Regular Solutions of Second-Order Stationary Hamilton-Jacobi Equations,* Journal of Differential Equations, **130** (1996), 210–234.
[17] N.U. Ahmed, *Generalized Solutions of HJB Equations Applied to Stochastic Control on Hilbert Space,* Nonlinear Analysis, **54** (2003), 495–523.
[18] N.U. Ahmed, *Optimal Relaxed Controls for Infinite Dimensional Stochastic Systems of Zakai Type*, SIAM J. Control and Optimization, **34(5)** (1996), 1592–1615.
[19] N.U. Ahmed, *Optimal Control of* ∞*-Dimensional Stochastic Systems Via Generalized Solutions of HJB Equations, Discussiones Mathematicae Differential Inclusions*, Control and Optimization, **21** (2001), 97–126.
[20] N.U. Ahmed, *Optimal Feedback Control for a Class of Partially Observed Uncertain Infinite Dimensional Dynamic Systems: A min-max Problem*, (submitted).