WEAK SOLUTIONS OF STOCHASTIC REACTION DIFFUSION EQUATIONS AND THEIR OPTIMAL CONTROL

N. U. Ahmed *
EECS, University of Ottawa
Ottawa, K1N6N5, Canada

Abstract. In this paper we consider a class of stochastic reaction diffusion equations with polynomial nonlinearities. We prove existence and uniqueness of weak solutions and their regularity properties. We introduce a suitable topology on the space of stochastic relaxed controls and prove continuous dependence of solutions on controls with respect to this topology and the norm topology on the natural space of solutions. Also we prove that the attainable set of measures induced by the weak solutions is weakly compact. Then we consider some optimal control problems, including the Bolza problem, and some target seeking problems in terms of the attainable sets in the space of measures and prove existence of optimal controls. In the concluding section we present briefly some extensions of the results presented here.

1. Introduction. Reaction diffusion equation arises in the study of many physical problems such as chemical kinetics, population biology, demography, Ecology, Micro-biology, Medicine (immunology). In chemical kinetics many different chemicals react to produce new chemicals or desired materials. In population biology, ecology, microbiology many different species interact leading to dynamic changes in the demography. Deterministic mathematical models for such systems are extensively studied in the literature [9], [10], [12], [13]. The authors of [10], [13] considered spatial version of the well known Lotka-Volterra model and studied many interesting properties of solutions. The lumped parameter model has been used in the study of optimal pest control in the field of agriculture [4]. Stochastic version of reaction diffusion equation has been studied in [6] where the author considers the question of existence of explosive solutions in the mean of order \( p > 1 \). In a recent paper [2], we studied the question of existence of mild solutions for a system of reaction diffusion equations with polynomial nonlinearities and considered optimal control problems proving existence of optimal controls. Also we developed the necessary conditions of optimality for the system with possible application to immunotherapy in cancer treatment. Here in this paper we study optimal control problems for a system of stochastic reaction-diffusion equations having weak solutions and prove existence of optimal controls with potential application to biology and medicine.

First we introduce the mathematical model for stochastic reaction-diffusion equation. In case there are \( n \) different species of population or reactants, the dynamics is
governed by a system of $n$ coupled partial differential equations describing diffusion and interaction of population within a given bounded open domain $\Sigma \subset \mathbb{R}^d$ (called the habitat) with sufficiently smooth boundary as follows:

$$\begin{align*}
\partial_t \varphi(t, x) &= D \Delta \varphi(t, x) + f(x, \varphi) + \sigma(t, x) \partial_t W(t, x), (t, x) \in (0, T) \times \Sigma, \\
\varphi(0, x) &= y_0(x), t = 0, x \in \Sigma \\
\partial \varphi/\partial \nu &= 0, (t, x) \in (0, T) \times \partial \Sigma
\end{align*}$$

(1)

where $D$ is a diagonal matrix with positive entries, $f$ is the interaction term, and $W$ is the space time Gaussian random field perturbing the system. We reformulate the interaction of population within a given bounded open domain $\Sigma \subset \mathbb{R}^n$ governed by a system of coupled partial differential equations describing diffusion and interaction of population within a given bounded open domain $\Sigma \subset \mathbb{R}^n$.

\[\begin{align*}
\partial_t \varphi(t, x) &= D \Delta \varphi(t, x) + f(x, \varphi) + \sigma(t, x) \partial_t W(t, x), (t, x) \in (0, T) \times \Sigma, \\
\varphi(0, x) &= y_0(x), t = 0, x \in \Sigma \\
\partial \varphi/\partial \nu &= 0, (t, x) \in (0, T) \times \partial \Sigma
\end{align*}\]

We are interested in the control of this system and we assume that control appears in the drift term $f = f(y, u)$ giving the controlled system

\[\begin{align*}
\dot{y} &= Ay + f(y) + \sigma(t) \dot{W}, \\
y(0) &= y_0.
\end{align*}\]

(2)

where $\sigma$ is an operator valued function with values in $\mathcal{L}(E, H)$ with $E$ being another separable Hilbert space and $W$ is an $E$-valued $Q$-Brownian motion supported on a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. For any $e \in E$, $(W(t), e)_E$ is a scalar Brownian motion with mean $\mathbb{E}((W(t), e)_E = 0$ and variance $\mathbb{E}(W(t), e)^2_E = t(Qe, e)$. The correct interpretation of equation (2) is given in terms of Itô stochastic differential equation written in the form

\[\begin{align*}
dy &= Aydt + f(y)dt + \sigma(t)dW(t), y(0) = y_0, t \in I.
\end{align*}\]

(3)

We are interested in the control of this system and we assume that control appears in the drift term $f = f(y, u)$ giving the controlled system

\[\begin{align*}
dy &= Aydt + f(y, u)dt + \sigma(t)dW(t), y(0) = y_0, t \in I,
\end{align*}\]

(4)

where $u$ is an $\mathcal{F}_t$-adapted random process with values in a compact Polish space $U$. For Convenience of presentation, we denote the class of admissible controls by $U_{ad}$ to be fully characterized later in the sequel. Our main objective is to consider control problem of the system (4) with the cost functional given by
2. Some notations. For any Banach space $X$ and any $r \in [1, \infty)$ we use the notation $L^r(I, X)$ to denote the class of all $X$-valued $\mathcal{F}_t$-adapted random processes whose $X$-norms are $r$-th power integrable, that is, $\xi \in L^r(I, X)$ if

$$
\| \xi \|_{L^r(I, X)} \equiv \left( \mathbb{E} \int_I |\xi(t)|_X^r dt \right)^{1/r} < \infty.
$$

For $r = \infty$, $L^\infty(I, X) = L^\infty(I, L^2(\Omega, X))$ denotes the class of $X$ valued $\mathcal{F}_t$ adapted random processes with essentially bounded second moments. That is,

$$
\text{ess.sup}\{ (\mathbb{E} |\xi(t)|_{L^2(X)}^2)^{1/2}, t \in I \} < \infty.
$$

With respect to the norm topology these are Banach spaces. For convenience of notation we use $L^2(\Omega, H)$, instead of $L^2(\Omega, H; X)$, to denote the class of $\mathcal{F}_0$ measurable $H$ valued random variables having finite second moments.

For any $1 \leq p < \infty$, let $X_p = L_{2p}(\Sigma, R^n)$ denote the standard Banach space with the dual $X_p^* = L_{2p/(2p-1)}(\Sigma, R^n)$. Let $C^1(I)$ denote the class of $C^1$ functions defined on $I$ and vanishing at $T$.

**Definition 2.1.** A process $y \in L^\infty(I, H) \cap L^2(I, V) \cap L^2_p(I, X_p)$ is said to be a weak solution of equation (4) if for every $\nu \in V$ and $v \in C^1_T$, $y$ satisfies the following identity $P$-a.s.

$$
(y_0, \varphi(0)v) + \int_I (y(t), \varphi_v)_{H} dt + \int_I < y(t), A^* v >_{V^*, V} \varphi(t) dt
$$

$$
+ \int_I < f(y(t), u_t), \varphi(t)v >_{X_p^*, X_p} dt + \int_I (\varphi(t)\sigma^*(t)v, dW(t)) = 0, \quad (6)
$$

3. Existence and uniqueness of solutions. In this section we consider the question existence and uniqueness of weak solutions (in the PDE sense not in the martingale sense). We shall use the classical approach [3] due to J.L. Lions based on the Gelfand triple, $\{V, H, V^*\}$. Let $H \equiv L_2(\Sigma, R^n)$ be the Hilbert space as introduced above and, for the homogeneous Neumann boundary condition, we take $V \equiv H^1(\Sigma, R^n) = W^{1,2}(\Sigma, R^n)$ and denote its topological dual by $V^*$. Then it is well known that the embeddings $V \hookrightarrow H \hookrightarrow V^*$ are continuous and dense. Let $\{v_i\}$ denote a basis of the triple $\{V, H, V^*\}$ orthogonal in $V$ and $V^*$ and orthonormal in $H$.

For any $1 \leq p < \infty$, let $X_p = L_{2p}(\Sigma, R^n)$ denote the standard Banach space with the dual $X_p^* = L_{2p/(2p-1)}(\Sigma, R^n)$. We introduce the following assumptions on the nonlinear operator $f$. The operator $f : X_p \times U \rightarrow X_p^*$ is continuous in all its arguments.

(A1): There exist $\delta \geq 0, \beta > 0$ such that

$$
< -f(y,u), y >_{X_p^*, X_p} + \delta |y|_{H}^2 \geq \beta \| y \|_{X_p}^2, \quad \forall u \in U \quad \text{and} \quad \forall y \in X_p.
$$

(A2): There exist constants $C_1, C_2 > 0$, possibly dependent on the set $U$ such that

$$
\| f(y,u) \|_{X_p} \leq C_1 + C_2 \| y \|_{X_p}^{2p-1}, \quad \forall y \in X_p \quad \text{and} \quad \forall u \in U,
$$

(A3): For all $u \in U$, $y \rightarrow f(y,u)$ is hemicontinuous and dissipative satisfying

$$
< f(y,u) - f(z,u), y - z >_{X_p^*, X_p} \leq 0, \quad \forall u \in U.
$$
Theorem 3.1. Consider the evolution equation (4) (representing the initial boundary value problem). Suppose $D$ appearing in the definition of the operator $A$ is a strictly positive diagonal matrix with the smallest element (smallest eigen value) $\alpha > 0$, the nonlinear operator $f$ is $m$-dissipative satisfying the assumptions $(A1)$-$(A3)$, and $\sigma$ is a uniformly bounded (in $t \in I$) operator valued function with values in $\mathcal{L}(E, H)$. And suppose the embeddings

$$V \hookrightarrow X_p \hookrightarrow H \hookrightarrow X_p^* \hookrightarrow V^*$$

are continuous. Then for any $\mathcal{F}_0$ measurable initial state $y_0 \in L^2_0(\Omega, H)$ and control $u \in \mathcal{U}_{ad}$, the evolution equation (4) has a unique weak solution $y \in L^2_0(I, H) \cap L^2_2(I, V) \cap L^2_p(I, X_p)$ and it is continuously dependent on the initial data.

**Proof.** The proof is an extension of the classical approach based on Galerkin approximation as in the deterministic systems [Ahmed and Teo, [3]]. By scalar multiplying the equation

$$dy = Ay dt + f(y, u) dt + \sigma(t) dW$$

with $y$ and integrating by parts and using the assumption $(A1)$, one can deduce that

$$E|y(t)|^2_H + 2\alpha \int_0^t E \| y(s) \|_V^2 ds + 2\beta \int_0^t E \| y(s) \|_{X_p^*}^{2p} ds$$

$$\leq E|y_0|^2_H + 2\delta \int_0^t E|y(s)|^2_H ds, \forall t \in I,$$

where the expected value of the stochastic term $\int_0^t (\sigma^*(t)y, dW)$ equals zero. This is justified by the fact that

$$E(\int_0^t (\sigma^*(t)y, dW))^2 = E \int_0^t (Q\sigma^*(t)y(t), \sigma^*(t)y(t))_E dt$$

$$= \int_0^t E|\sigma^*(t)y(t)|^2_E dt < \infty$$

By use of Gronwall Lemma, we conclude from the above inequality that $y \in L^2_0(I, H)$. Based on this fact, again it follows from the above inequality that $y \in L^2_0(I, V) \cap L^2_p(I, X_p)$. Now letting $\{v_i\}$ denote a basis for the Gelfand triple $(V, H, V^*)$, and using the Galerkin approach we project the infinite dimensional system to a sequence of Itô differential equations of finite dimension $n \in N$. This is given by the following system

$$(dy^n, v_i) =< Ay^n, v_i >_V dt + < f(y^n, u_i), v_i >_{X_p^*} dt$$

$$+ (\sigma(t)^* v_i, dW^n)_E, 1 \leq i \leq n.$$  \hfill (9)

where $y^n(t) \equiv \sum_{j=1}^n z^n_j(t)v_j$, $t \in I$ and $y^n(0) = \sum_{j=1}^n (y_0, v_j)v_j = \sum_{j=1}^n z^n_0(v_j)$, and

$$W^n(t) \equiv \sum_{i=1}^n (W(t), e_i)_E e_i, t \in I,$$

with $\{e_i\}$ being an orthonormal basis of the Hilbert space $E$. The process

$$z^n \equiv (z^n_1, z^n_2, \cdots, z^n_n)'$$

is given by the solution of the following system of finite $(n)$ dimensional stochastic differential equations,

$$dz^n = Az^n dt + F(z^n, u_i) dt + \Gamma(t) dB^n, t \in I,$$  \hfill (10)
where $\mathcal{A}$ is a $n \times n$ matrix, $F$ is a $n$-vector and $\Gamma$ is a $n \times n$ matrix with elements as shown below,

$$\begin{cases}
    a_{ij} \equiv <Av_i,v_j>_{V^*,V},
    F_i(z^n,u) \equiv <f(\sum_j z^n_j v_j,u),v_i >_{X^*_p,X_p},
    
    \gamma_{i,j}(t) \equiv \sqrt{\lambda_j} (v_i,\sigma(t)e_j)_H, i,j=1,2,\cdots ,n
\end{cases}$$

(11)

with $B^n$ being the $n$-dimensional standard Brownian motion with stochastically independent components and $\{\lambda_i\}$ are the (nonnegative) eigen values of the covariance operator $Q \in L^+_s(E)$. Since by assumption $f$ is $m$-dissipative from $X_p$ to its dual $X^*_p$, the operator $F$ is $m$-dissipative on $R^n$ and therefore for any $\delta > 0,$ $(I - \delta F)^{-1}$ is a continuous bounded (nonlinear) operator in $R^n$ and it follows from Crandal-Liggett generation theorem [5], Theorem 4.7, p121 that it generates a nonlinear semigroup. We use this result to justify that the sequence of approximate solutions of equation (10) constructed by using implicit difference scheme converges to the solution of equation (10). For convenience of notation we omit the superscript $n$ on $z$. To start we consider any finite (uniform) partition of the interval $I$, $\pi_m \equiv \{0 = t_0^n < t_1^n < t_2^n, \cdots t_m^n = T\}$, with $(t_i^n - t_{i+1}^n) = \delta_m > 0$ for all $0 \leq i \leq m - 1$ giving $m\delta_m = T$ for all $m \in N$. Then we use the implicit difference scheme, as indicated above, to construct a sequence of approximate solutions of equation (10) denoted by $\{z_m(t), t \in I\}$ which is given by linear interpolation of the nodes as defined below,

$$z_m(t_i^{n+1}) \equiv (I - \delta_m F)^{-1} [z_m(t_i^n) + \delta_m A z_m(t_i^n) + \Gamma(t_i^n)(B^n(t_i^n) - B^n(t_i^{n+1})) \right]$$

$$i = 0,1,\cdots ,m-1.$$  

(12)

Now it follows from Crandal-Liggett generation theorem that as $m \to \infty$, $z_m$ converges almost surely uniformly on $I$ to the solution $z = z^n$ of equation (10). Using this we construct $y^n = \sum_i z_i^n v_i$ giving the solution of equation (9). Recall that $C^1_V(I)$ denotes the class of $C^1$ functions defined on $I$ and vanishing at $T$. Multiplying equation (9) by any $\varphi \in C^1_V(I)$, and integrating by parts, one obtains the following identity

$$\begin{align}
(y^n_0,\varphi(0)v_i) &+ \int_I (y^n(t),\phi v_i)_H d(t) + \int_I <y^n(t),A^*v_i >_{V^*,V} \varphi(t) dt \\
&+ \int_I f(y^n(t),u_i),\varphi(t)v_i >_{X^*_p,X_p} dt \\
&+ \int_I (\varphi(t)\sigma^*(t)v_i, dW^n(t))_E = 0,
\end{align}$$

(13)

for $1 \leq i \leq n$. This sequence $\{y^n\}$ also satisfies the a-priori estimate (7) and therefore it is contained in a bounded subset of $L^\infty(I, H) \cap L^2(I, V) \cap L^{2p}_2(I, X_p)$. By Alaoglu’s theorem, a bounded subset of $L^\infty(I, H)$ is relatively $w$-compact (weak star compact). Further the spaces $L^2(I, V)$ and $L^{2p}_2(I, X_p)$ are reflexive Banach spaces and so a bounded subset of such spaces is relatively weakly compact. In view of these facts we conclude that there exists a subsequence of the sequence $\{y^n\}$, relabeled as $\{y^n\}$, and $y^* \in L^\infty(I, H) \cap L^2(I, V) \cap L^{2p}_2(I, X_p)$ such that

$$y^n \overset{w^*}{\longrightarrow} y^* \text{ in } L^\infty(I, H)$$

(14)

$$y^n \overset{w}{\longrightarrow} y^* \text{ in } L^2(I, V)$$

(15)
We want to prove that $y^*$ satisfies the identity (6) P-a.s. In other words $y^*$ is a weak solution of the evolution equation (4). Let $\zeta \in L_\infty(\Omega)$ be any arbitrary $\mathcal{F}$ measurable random variable. Multiplying equation (13) by $\zeta$ and taking the expectation we obtain

$$
\mathbb{E}\{\zeta(y^n_0, \varphi(0)v_i)\} + \mathbb{E}\{\zeta \int_I (y^n(t), \dot{\varphi}(t)v_i)_H dt \}
$$

$$
+ \mathbb{E}\{\zeta \int_I < y^n(t), A^*v_i >_{V,V^*} \varphi(t) dt \}
$$

$$
+ \mathbb{E}\{\zeta \int_I < f(y^n(t), u_t), \varphi(t)v_i >_{X^*_p,X_p} dt \}
$$

$$
+ \mathbb{E}\{\zeta \int_I (\varphi(t)\sigma^i(t)v_i, dW^n(t))_E \} = 0, \ i = 1, 2, \cdots n. \quad (17)
$$

Now we want to let $n \to \infty$ in the expression (17) and determine the limits. Starting with the first term we note that since $V$ is dense in $H$ it is clear that $L^2_0(\Omega, V)$ is also dense in $L^2_0(\Omega, H)$ and hence $y^n_0 \to y_0$ in $L^2_0(\Omega, H)$. Clearly the random variable $\{\zeta(y^n_0, \varphi(0)v_i)\}$ is $\mathcal{F}$ measurable and integrable with respect $P$ measure and $\varphi(0)v_i \in L_2(\Omega, V) \subset L_2(\Omega, H)$. Thus letting $n \to \infty$ we have

$$
\mathbb{E}\{\zeta(y^n_0, \varphi(0)v_i)\} \to \mathbb{E}\{\zeta(y_0, \varphi(0)v_i)\}. \quad (18)
$$

Considering the second term, it is easy to verify that there exists a constant

$$
a_1 = a_1(\| \zeta \|_\infty, \| \varphi \|_{L_2(I), \lambda(I)}) > 0,
$$

dependent on the parameters shown such that

$$
|\mathbb{E}\{\zeta \int_I (y^n(t), \dot{\varphi}(t)v_i)_H dt \}| \leq a_1 \sup\{\mathbb{E}|y^n(t)|_{\mathcal{H}}^2, t \in I\}^{1/2} = a_1 \| y^n \|_{L^2_\infty(I,H)}.
$$

Since \( \dot{\varphi}v_i \) is deterministic, it is clear $\dot{\varphi}v_i \in L_1^2(I,V) \subset L_2^2(I,H)$. Thus letting $n \to \infty$, it follows from (14) that

$$
\mathbb{E}\{\zeta \int_I (y^n(t), \dot{\varphi}(t)v_i)_H dt \} \to \{\zeta \int_I (y^*(t), \dot{\varphi}(t)v_i)_H dt \}. \quad (19)
$$

Considering the third term, it follows from Hölder inequality that

$$
|\mathbb{E}\{\zeta \int_I (y^n(t), A^*v_i \dot{\varphi}(t))_H dt \}| \leq a_2 \| y^n \|_{L^2_2(I,V)} \quad (20)
$$

where the constant $a_2 = a_2(\| \zeta \|_\infty, \| A \|_{L_2(V,V^*), \lambda(I)}) > 0$, is dependent on the parameters displayed. Since $A \in \mathcal{L}(V,V^*)$, $\varphi \in C^1$ and $v_i \in V$ are all deterministic it is clear that $\varphi A^*v_i \in L^2_2(I,V^*)$. Hence it follows from (15) that

$$
\mathbb{E}\{\zeta \int_I < y^n(t), \varphi(t)A^*v_i >_{V,V^*} dt \} \to \mathbb{E}\{\zeta \int_I < y^*(t), \varphi(t)A^*v_i >_{V,V^*} dt \}. \quad (21)
$$

Considering the fourth term, it follows from the assumption (A2) that for all $u \in U$, $f$ maps $X_p$ to $X^*_p$. Thus, for $y^n \in L_{2p}^2(I,X_p)$, it follows from Hölder and Cauchy inequalities that there exist constants $a_3, a_4 > 0$ such that

$$
\| f(y^n, u) \|_{L_{2p/(2p-1)}^2(I,X_p^*)} \leq a_3 + a_4 \| y^n \|_{L_{2p}^2(I,X_p)}^{2p-1} \quad (22)
$$

where the constants $\{a_3, a_4\}$ are dependent on the growth parameters $\{C_1, C_2, p \geq 1\}$ of the nonlinear operator $f$ (see assumption (A2)). By our assumption the embedding $V \hookrightarrow X_p$ is continuous. Thus there exists a constant $c > 0$ such that
\[ \|v_i\|_{X_p} \leq c \|v_i\|_{V}. \]

Using this fact and the above inequality we obtain the following upper bound for the fourth term of the expression (17),

\[ |E\{\zeta \int_I <f(y^n, u), \varphi v_i >_{X^*_p, X_p} dt\}| \leq (a_5 + a_6 \|y^n\|_{L_{2p}^p(I, X_p)}^{2p-1}) \quad (23) \]

where the constants \(a_5, a_6\) depend on the parameters \(a_1, a_2, c\) and the norms of \(\varphi, \zeta\). Since the sequence \(\{y^n\}\) also satisfies the apriori bound given by the inequality (7), it follows from Gronwall inequality that the sequence \(\{y^n\}\) is contained in a bounded subset of \(L_{2p}^p(I, X_p)\). Hence it follows from inequality (22) that the sequence \(\{f^n(\cdot)\} \equiv \{f(y^n(\cdot), u(\cdot))\}\) is contained in a bounded subset of the dual \((L_{2p}(I, X_p))^* = L_{2p/2p-1}^a(I, X_p^*)\) uniformly with respect to \(u \in U_{ad}\).

Since, for \(\infty > p \geq 1\), these are reflexive Banach spaces, there exists a subsequence of the sequence \(\{f^n\}\), relabeled as the original sequence, and an element \(g \in L_{2p/(2p-1)}^a(I, X_p^*)\) such that

\[ f^n \xrightarrow{w*} g \text{ in } L_{2p/(2p-1)}^a(I, X_p^*). \quad (24) \]

Hence, by similar duality argument we have the following weak star \((w*)\) convergence

\[ E\{\zeta \int_I <f(y^n(t), u_t), \varphi(t) v_i >_{X^*_p, X_p} dt\} \rightarrow E\{\zeta \int_I <g(t), \varphi(t) v_i >_{X^*_p, X_p} dt\}. \quad (25) \]

Recall that, for reflexive Banach spaces, weak and weak star topologies are equivalent. So no distinction is necessary between the weak and weak star convergence. Now we can use the assumption (A3) (hemicontinuity and dissipativity) to verify that

\[ g(t) = f(y^*(t), u_t), t \in I. \]

For any \(\xi \in L_{2p}^a(I, X_p)\), it follows from the dissipativity of \(f\) (see (A3)) that for any \(n \in N\),

\[ \int_I <f(\xi(t), u_t) - f(y^n(t), u_t), \xi(t) - y^n(t) >_{X^*_p, X_p} dt \leq 0, \quad \forall \xi \in L_{2p}^a(I, X_p) \quad (26) \]

with probability one. It follows from Mazur’s theorem that there exists a sequence \(\{\eta^n\} \subset L_{2p}^a(I, X_p)\), given by a proper convex combination of the sequence \(y^n\), such that \(\eta^n \xrightarrow{a.s.} y^*\) in \(L_{2p}^a(I, X_p)\). Thus with probability one,

\[ \int_I <f(\xi(t), u_t) - f(\eta^n(t), u_t), \xi(t) - \eta^n(t) >_{X^*_p, X_p} dt \leq 0, \quad \forall \xi \in L_{2p}^a(I, X_p), \quad (27) \]

and, upon letting \(n \to \infty\) in the above inequality, it follows from (24) that

\[ \int_I <f(\xi(t), u_t) - g(t), \xi(t) - y^*(t) >_{X^*_p, X_p} dt \leq 0, \quad \forall \xi \in L_{2p}^a(I, X_p) \quad P - a.s. \quad (28) \]

Since this holds for all \(\xi \in L_{2p}^a(I, X_p)\), we can choose \(\xi = y^* + \varepsilon w\) for any \(w \in L_{2p}^a(I, X_p)\) and any \(\varepsilon > 0\) and obtain the following inequality

\[ \int_I <f(y^*(t) + \varepsilon w(t), u_t) - g(t), \varepsilon w(t) >_{X^*_p, X_p} dt \leq 0, \]

\[ \forall \varepsilon > 0, \quad \forall \ w \in L_{2p}^a(I, X_p). \quad (29) \]
Dividing the above expression by \( \varepsilon > 0 \) and letting \( \varepsilon \downarrow 0 \) it follows from the hemicontinuity (A3) of \( f \) (in its first argument) that
\[
\int_{I} < f(y^*(t), u_t) - g(t), w(t) >_{X_1^*, X_p} dt \leq 0, \quad \forall \ w \in L^2_{2p}(I, X_p).
\] (30)
This is possible if and only if \( g(\cdot) = f(y^*(\cdot), u(\cdot)) \) almost everywhere \( P \)-a.s. Hence, it follows from (25) that
\[
\mathbb{E}\{\zeta \int_{I} < f(y^n(t), u_t), \varphi(t) v_i >_{X_1^*, X_p} dt \}
\rightarrow \mathbb{E}\{\zeta \int_{I} < f(y^*(t), u_t), \varphi(t) v_i >_{X_1^*, X_p} dt \}.
\] (31)
Since \( W^n \) is simply the \( n \)-dimensional projection of the infinite dimensional Brownian motion \( W \) with values in \( E \), it is straightforward to verify that the last term of the expression (17) converges to the limit as shown below
\[
\mathbb{E}\left\{\zeta \int_{I} (\varphi(t) \sigma^*(t) v_i, dW^n(t))_E \right\} \rightarrow \left\{\zeta \int_{I} (\varphi(t) \sigma^*(t) v_i, dW(t))_E \right\}
\] (32)
as \( n \to \infty \). Therefore, on the basis of the above analysis, letting \( n \to \infty \) in the expression (17), it follows from (18), (19), (21), (31) and (32) that
\[
\mathbb{E}\{\zeta(y_0, \varphi(0)v_i)\} + \mathbb{E}\{\zeta \int_{I} (y^*(t), \varphi(t) v_i)_H \}
\]
\[+ \mathbb{E}\{\zeta \int_{I} < y^*(t), A^* v_i >_{V_1^*, \varphi(t)} dt \}
\]
\[+ \mathbb{E}\{\zeta \int_{I} < f(y^*(t), u_t), \varphi(t) v_i >_{X_1^*, X_p} dt \}
\]
\[+ \mathbb{E}\{\zeta \int_{I} (\varphi(t) \sigma^*(t) v_i, dW(t))_E \} = 0, \quad \forall \ i \in N = \{1, 2, 3, \cdots \}.
\] (33)
This holds for every \( \zeta \in L_{\infty}(\Omega) \) and hence the following identity holds with probability one
\[
(y_0, \varphi(0)v_i) + \int_{I} (y^*(t), \varphi(t) v_i)_H + \int_{I} < y^*(t), A^* v_i >_{V_1^*, \varphi(t)} dt
\]
\[+ \int_{I} < f(y^*(t), u_t), \varphi(t) v_i >_{X_1^*, X_p} dt
\]
\[+ \int_{I} (\varphi(t) \sigma^*(t) v_i, dW(t))_E = 0, \quad \forall \ i \in N.
\] (34)
Since \( \{v_i\} \) is a basis for \( V \), it is clear that the above identity holds for all \( v \in V \) leading to the following identity
\[
(y_0, \varphi(0)v) + \int_{I} (y^*(t), \varphi(t)v)_H + \int_{I} < y^*(t), A^* v >_{V_1^*, \varphi(t)} dt
\]
\[+ \int_{I} < f(y^*(t), u_t), \varphi(t) v >_{X_1^*, X_p} dt
\]
\[+ \int_{I} (\varphi(t) \sigma^*(t)v, dW(t))_E = 0.
\] (35)
This proves that \( y^* \in L_{\infty}(I, H) \cap L^2(I, V) \cap L^2_p(I, X_p) \) is a weak solution of the system (4). To verify uniqueness of solution, we use the facts that the operator \( A \) is negative self-adjoint and \( f \) is dissipative. Let \( x, y \in L_{\infty}(I, H) \cap L^2(I, V) \cap L^2_p(I, X_p) \)
denote the weak solutions corresponding to the initial conditions \(x_0, y_0 \in L^0_2(\Omega, H)\) and any fixed control \(u \in \mathcal{U}_{ad}\). Then it follows from integration by parts and the property of \(A\) that

\[
E|x(t) - y(t)|_H^2 + 2\alpha E \int_0^t \|x(s) - y(s)\|_{V'}^2 \, ds \leq E|x_0 - y_0|_H
\]

\[
+2E \int_0^t <f(x(s), u_s) - f(y(s), u_s), x(s) - y(s)>_{X'_p, X_p} \, ds. \quad (36)
\]

Since \(f\) is dissipative, it follows from (A3) and the above inequality that

\[
E|x(t) - y(t)|_H^2 + 2\alpha E \int_0^t \|x(s) - y(s)\|_{V'}^2 \, ds \leq E|x_0 - y_0|_H \quad (37)
\]

Thus the solution is continuously dependent on the initial state. Setting \(x_0 = y_0\), we find that \(x = y\). This proves that for any given initial state \(y_0 \in L^0_2(\Omega, H)\) and control \(u \in \mathcal{U}_{ad}\), the evolution equation (4) has unique weak solution. This completes the proof. 

3.1. Admissibility of polynomial nonlinearities. (E1): The degree of polynomial nonlinearity admissible is determined by the Sobolev embedding theorem, in particular, for \(\Sigma \subset \mathbb{R}^d\) with smooth boundary, the embeddings

\[
W^{m,q}(\Sigma) \hookrightarrow L_p(\Sigma), \quad \text{for } mq = d, q \leq p < \infty
\]

are continuous. If \(\mathbb{R}^d = \mathbb{R}^2\) and \(W^{m,q} = W^{1,2} = V\), then \(2 \leq p < \infty.\) Thus for \(\Sigma \subset \mathbb{R}^2\), polynomial nonlinearity of \(f\) of any finite degree \((2p - 1)\) for integer \(p \in [1, \infty)\) is admissible. This implies that the embedding \(V \hookrightarrow X_p\) is continuous for all \(p \geq 1\).

(E2): Similarly, for \(mq < d\), we have \(W^{m,q} \hookrightarrow L_p\), for all \(q \leq p \leq (dq/d - mq)\). In our problem \(m = 1, q = 2\) and, for \(\Sigma \subset \mathbb{R}^d, d = 3\); and hence \(W^{1,2} = V \hookrightarrow L_p\) for \(2 \leq p \leq 6\). Hence the embedding \(V \hookrightarrow X_p\) is continuous for \(1 \leq p \leq 3\).

Under the given conditions (E1) and (E2), the embeddings \(V \hookrightarrow X_p \hookrightarrow H\) are also compact. For more on Sobolev embedding theorems see Adams [1], Theorem 5.4, p97.

In view of (E1) and (E2) we see that Theorem 3.1 is applicable to reaction diffusion equations defined on two dimensional spatial domain with polynomial nonlinearity of any order \((2p - 1)\) for \(1 \leq p < \infty\) while, for three dimensional spatial domain, the admissible order is \((2p - 1)\) only for \(1 \leq p \leq 3\).

4. Continuity of solutions with respect to controls. Let \(U\) be a compact metric space and \(C(U)\) denote the space of continuous functions on \(U\) with the usual sup-norm topology. The topological dual of \(C(U)\) is given by the space of regular Borel measures on \(U\) denoted by \(\mathcal{M}(U)\). Let \(\mathcal{M}_1(U) \subset \mathcal{M}(U)\) denote the space of probability measures on \(U\). Consider the complete separable filtered probability space \((\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{P})\) with \(\mathcal{F}_t\) being right continuous having left limits and let \(\mathcal{G}_t\) be a nondecreasing family of subsigma algebras of the sigma algebra \(\mathcal{F}_t\), also right continuous having left limits. Let \(\mathcal{P}\) denote the \(\mathcal{G}_t\) predictable subsigma field of the product sigma field \(\mathcal{B}(I) \times \mathcal{F}\) and \(\mu\) the restriction of the product measure \(\text{Leb.} \times \mathcal{P}\) on it. Note that the measure space \((I \times \Omega, \mathcal{P}, \mu)\) is separable and, since \(U\) is a compact metric space, the Banach space \(C(U)\) is also separable. Thus the Lebesgue-Bochner space \(L_1(\mu, C(U))\) is a separable Banach space and hence it has
a countable dense set \{\varphi_n\}. Without loss of generality, we may assume that they are normalized so that \{\varphi_n\} ∈ \mathcal{B}_1(L_1(\mu, C(U))) the closed unit ball. Since \mathcal{M}(U) does not satisfy RNP [7], Diezel and Uhl.Jr., the dual of \(L_1(\mu, C(U))\) is not given by \(L_\infty(\mu, \mathcal{M}(U))\). It is given by the weak star measurable functions with values in \(\mathcal{M}(U)\) which we denote by \(L_\infty^*(\mu, \mathcal{M}(U))\). Consider the subset \(L_\infty^*(\mu, \mathcal{M}_1(U)) \subset L_\infty^*(\mu, \mathcal{M}(U))\). An element \(u \in L_\infty^*(\mu, \mathcal{M}_1(U))\) is a \(\mathcal{G}_t\) adapted random process with values in the space of probability measures on \(U\). For convenience of notation, an element of this set is denoted by \(u_t\) or \(u_{t,\omega}\). The first natural topology is the weak star topology. By Alaoglu’s theorem [8], the set \(L_\infty^*(\mu, \mathcal{M}_1(U))\) is weak star compact and since \(L_1(\mu, C(U))\) is separable this topology is metrizable. The metric is given by

\[
\rho(u, v) \equiv \sum_{n=1}^{\infty} 1/2^n \min \left\{ 1, \left| \int_{I \times \Omega} [u(\varphi_n) - v(\varphi_n)] \, d\mu \right| \right\}.
\]

Unfortunately, this metric topology is too weak for nonlinear stochastic problems considered here. So we consider a stronger metric topology given by

\[
d(u, v) \equiv \sum_{n=1}^{\infty} 1/2^n \int_{I \times \Omega} \min \left\{ 1, |u(\varphi_n) - v(\varphi_n)| \right\} \, d\mu,
\]

where \(u(\varphi)(t, \omega) \equiv \int_U \varphi(t, \omega, \xi) u_{t,\omega}(d\xi)\). Consider the set \((L_\infty^*(\mu, \mathcal{M}_1(U)), d)\) endowed with the metric topology \(d\). Let \(\{u^n\}\) be a Cauchy sequence in this metric topology. It follows from the definition of the metric \(d\) that, for every \(\varphi \in L_1(\mu, C(U))\), \(u^n(\varphi)\) is a Cauchy sequence in \(\mu\) measure. Since \((I \times \Omega, \mathcal{P}, \mu)\) is a finite measure space and the space of measurable functions is complete with respect to the topology of convergence in measure, there exists an element \(u^*(\varphi)\) such that \(u^n(\varphi) \to u^*(\varphi)\) in \(\mu\) measure. This holds for every \(\varphi \in L_1(\mu, C(U))\). Thus the metric space \((L_\infty^*(\mu, \mathcal{M}_1(U)), d)\) is complete. So a set in this metric space is compact if it is totally bounded. Thus we assume that \(U_{ad}\) is a totally bounded subset of the complete metric space \((L_\infty^*(\mu, \mathcal{M}_1(U)), d)\) and hence compact. We choose this set as the set of admissible controls.

Now we are prepared to prove the continuity of the control to solution map. For continuity, however, we do not need compactness. We use this property later for existence of optimal controls.

**Theorem 4.1.** Under the assumptions of Theorem 3.1, the map \(u \to y\) is continuous from \(U_{ad}\) to \(L_\infty^*(I, H) \cap L_2^p(I, V)\) with respect to the metric topology \(d\) on \(U_{ad}\) and the norm topologies on \(L_\infty^*(I, H) \cap L_2^p(I, V)\).

**Proof.** Let \(\{u^n\}\) be any sequence from \(U_{ad}\) and suppose \(u^n \to u^o\) and let \(\{y^n, y^o\}\) denote the corresponding weak solutions of the evolution equation (4) for the same initial state \(y^n(0) = y^o(0) = y_0\). We show that \(y^n\) converges to \(y^o\) in the norm topologies of \(L_\infty^*(I, H) \cap L_2^p(I, V)\). Since by our assumption, the operator \(f\) satisfies the properties (A1)-(A3) uniformly with respect to the set \(U\), the sequence \(\{y^n, y^o\}\) is contained in a bounded subset of \(L_\infty^*(I, H) \cap L_2^p(I, V) \cap L_2^p(I, X_p)\). Hence there exists a finite positive number \(b\) so that

\[
\max\{\|y^n\|_{L_2^p(I, X_p)}, \|y^o\|_{L_2^p(I, X_p)}\}, n \in N \leq b.
\]  

Clearly, it follows from equation (4) that the difference \(y^n - y^o\) satisfies the following equation

\[
d(y^n - y^o) = A(y^n - y^o)dt + (f(y^n, u^n) - f(y^o, u^o))dt, y^n(0) - y^o(0) = 0
\]  

\(\text{(39)}\)
in the weak sense. Scalar multiplying on both sides of the above equation by \( y^n - y^o \) and integrating by parts and following similar steps as in the derivation of the apriori bound we find that

\[
E|y^n(t) - y^o(t)|_H^2 + 2\alpha E \int_0^t \| y^n(s) - y^o(s) \|_V^2 \, ds \\
\leq 2E \int_0^t < f(y^n, u^n) - f(y^o, u^o), y^n(s) - y^o(s) >_{X^*_p \times X_p} \, ds \\
\leq 2E \int_0^t < f(y^n, u^n) - f(y^o, u^o), y^n - y^o >_{X^*_p \times X_p} \, ds \\
+ 2E \int_0^t < f(y^o, u^n) - f(y^o, u^o), y^n - y^o >_{X^*_p \times X_p} \, ds. \tag{40}
\]

It follows from dissipativity property of \( f \) (see (A3)) (which holds uniformly with respect to \( U \)) that the third line in the above expression is equal to or less than zero. Thus

\[
E|y^n(t) - y^o(t)|_H^2 + 2\alpha E \int_0^t \| y^n(s) - y^o(s) \|_V^2 \, ds \\
\leq 2E \int_0^t < f(y^o, u^n) - f(y^o, u^o), y^n - y^o >_{X^*_p \times X_p} \, ds. \tag{41}
\]

We show that the expression on the right hand side of the inequality converges to zero as \( n \to \infty \). Let us denote this expression by \( Z_n(t) \),

\[
Z_n(t) \equiv 2E \int_0^t < f(y^o, u^n) - f(y^o, u^o), y^n - y^o >_{X^*_p \times X_p} \, ds.
\]

Since \( f \) is linear with respect to the relaxed controls, it is clear that

\[
Z_n(t) = 2E \int_0^t < f(y^o, u^n - u^o), y^n - y^o >_{X^*_p \times X_p} \, ds.
\]

For convenience of notation let us set \( f(y^o(s), u^n_s - u^o_s) \equiv f^o(s, u^n_s - u^o_s) \). Using Hölder inequality applied to the last scalar product and using the upper bound given by (38) we obtain

\[
Z_n(t) = 2E \int_0^t f^o(s, u^n_s - u^o_s) \| y^n(s) - y^o(s) \|_{X^*_p \times X_p} \, ds \\
\leq 2 \left( E \int_0^t \| f^o(s, u^n_s - u^o_s) \|_{X^*_p}^{2p/2p-1} \, ds \right)^{(2p-1)/2p} \left( E \int_0^t \| y^n - y^o \|_{X_p}^{2p} \, ds \right)^{1/2p} \\
\leq 4b \left( E \int_0^t \| f^o(s, u^n_s - u^o_s) \|_{X^*_p}^{2p/2p-1} \, ds \right)^{(2p-1)/2p}, \ \forall \ t \in I.
\]

Hence

\[
\sup \{|Z_n(t)|, \ t \in I\} \leq 4b \left( E \int_0^T \| f^o(s, u^n_s - u^o_s) \|_{X^*_p}^{2p/2p-1} \, ds \right)^{(2p-1)/2p}. \tag{42}
\]

Now it follows from (41) and (42) that, for all \( t \in I \), we have

\[
E|y^n(t) - y^o(t)|_H^2 + 2\alpha E \int_0^t \| y^n(s) - y^o(s) \|_V^2 \, ds
\]
\[
\leq 4b \left( E \int_0^t \| f^o(s, u^n_s - u^o_s) \|_{X^*_p}^{2p/2p-1} \, ds \right)^{(2p-1)/2p}
\]
\[
\leq 4b \left( \int_{I \times \Omega} \| f^o(s, u^n_s - u^o_s) \|_{X^*_p}^{2p/2p-1} \, d\mu \right)^{(2p-1)/2p}.
\]

By virtue of assumption (A2), \( f(y^o(\cdot), u_\cdot) \in L^p_{2p/2p-1}(I, X^*_p) \) uniformly with respect to \( u \in U_{ad} \). Thus the above integral is well defined. Note that convergence of \( u^n \) to \( u^o \) in the metric topology \( d \) is equivalent to convergence of \( u^n(\varphi) \) to \( u^o(\varphi) \) in measure for each \( \varphi \in L_1(\mu, C(U)) \). In fact this holds also for every \( \varphi \in L_1(\mu, C(U, X^*_p)) \) in the sense that
\[
\lim_{n \to \infty} \mu \{(t, \omega) \in I \times \Omega : \| u^n(\varphi) - u^o(\varphi) \|_{X^*_p} > \varepsilon \} = 0.
\]

It is well known that a sequence that converges in measure has a subsequence that converges a.e. Thus there exists a subsequence \( \{u^{n_k}(\varphi)\} \) of the sequence \( \{u^n(\varphi)\} \) such that, along the subsequence, the integrand on the righthand side of the expression (43) converges to zero \( \mu \) a.e. Further, as \( y^o \in L^2_{2p}(I, X_p) \), it follows from the assumption (A2) that the integrand is dominated by an integrable random process uniformly with respect to the set of admissible controls \( U_{ad} \). Thus, it follows from Lebesgue dominated convergence theorem that the integral on the righthand side of the inequality (43) converges to zero as \( n \to \infty \). Hence, letting \( n \to \infty \), it follows from the above inequality that
\[
\lim_{n \to \infty} \sup_{t \in I} \left\{ E|y^n(t) - y^o(t)|_{H}^2 + 2\alpha E \int_0^t \| y^n(s) - y^o(s) \|_{V}^2 \, ds \right\} = 0 \quad (44)
\]

This proves the continuity of the map \( u \to y \) (the control to solution map) as stated in the theorem. \( \square \)

**Remark 4.2.** It is interesting to note that the integral defining the function \( Z_n \) can be expressed equivalently as follows:
\[
Z_n(t) = 2E \int_0^t \Phi^o(s, u^n_s - u^o_s, y^n(s) - y^o(s) > X^*_p, X_p) \, ds
\]
\[
= \int_{I \times \Omega} \chi(s) < \varphi_n(s), u^n_s - u^o_s >_{C(U), M_1(U)} \, d\mu.
\]

where the process \( \varphi_n \) is given by the following conditional expectation
\[
\varphi_n(t) \equiv 2E\{ < f^o(t, \cdot), y^n(t) - y^o(t) >_{X^*_p, X_p} | G_t \}
\]

for \( t \in I \), and it belongs to \( L_1(\mu, C(U)) \). Since \( u^n \) converges weakly to \( u^o \), and \( Z_n(t) \) converges to zero uniformly on \( I \), it is necessary that \( \varphi_n \) converges strongly to zero in \( L_1(\mu, C(U)) \). This implies that the set \( \{ \varphi_n \} \) is a relatively compact subset of \( L_1(\mu, C(U)) \) which is compatible.

5. **Optimal control.** In this section we consider the question of existence of optimal control.

**Theorem 5.1.** Consider the system (4) with the cost functional (5) and admissible controls \( U_{ad} \), a compact subset of \( L^\infty_{2p}(I, M_1(U)) \) endowed with the metric topology \( d \). Suppose the assumptions of Theorem 4.1 hold, and further the functions \( \{ \ell, \Phi \} \) satisfy the following conditions:
\[ \ell = \ell(t,x,\xi) : I \times V \times U \rightarrow R^+ = [0,\infty] \text{ is measurable in the first argument, continuous in the second uniformly with respect to } \xi \in U, \text{ and continuous and bounded with respect to the third argument satisfying the following inequality} \]
\[ \ell(t,x,\xi) \leq \alpha(t) + \beta \| x \|^2, \forall \xi \in U \]
for some \( \alpha \in L^1_t(I) \) and \( \beta > 0; \)
\( \Phi : H \rightarrow R^+ \) is lower semicontinuous on \( H \) and there exist \( \gamma \geq 0 \) and \( \delta > 0 \) such that \( \Phi(x) \leq \gamma + \delta|x|^2_H. \)

Then there exists an optimal control minimizing the functional \( J(u). \)

**Proof.** Since \( \mathcal{U}_{ad} \) is compact in the metric topology \( d \) it suffices to prove that \( J \)

is lower semi-continuous with respect to this topology. Let \( u^n \xrightarrow{d} u^o \) in \( \mathcal{U}_{ad}, \) and let \( \{y^n, y^o\} \in L^2(I,H) \cap L^2(I,V) \cap L^2(I,X_p) \) denote the corresponding weak solutions of the stochastic evolution equation (4). Consider the first term of the cost functional,

\[ J_1(u) \equiv E \int_I \ell(t,y,u)dt. \]

We prove that it is lower semi-continuous. Corresponding to the pair \( \{u^o, u^n\}, \) we write this as a sum of three terms.

\[ J_1(u^o) = E \int_I [\ell(t,y^o(t),u^o) - \ell(t,y^n(t),u^n)]dt \]
\[ + E \int_I [\ell(t,y^o(t),u^n) - \ell(t,y^n(t),u^n)]dt + J_1(u^n). \]

Since \( u^n \xrightarrow{d} u^o, \) for every \( \varepsilon > 0, \) there exists an \( n_{1,\varepsilon} \in N \) such that the absolute value of the first term on the right of the above expression is equal to or less than \( \varepsilon/2 \) for all \( n \geq n_{1,\varepsilon}. \) Now we consider the second component on the right hand side of the expression (45). Using the basic property of conditional expectation, we can rewrite this term in the following form,

\[ E \int_I [\ell(t,y^o(t),u^n) - \ell(t,y^n(t),u^n)]dt \]
\[ = E \int_I E \{\ell(t,y^o(t),\xi) - \ell(t,y^n(t),\xi)\}G_t u^n d\xi dt. \]

It follows from Theorem 4.1 that \( y^n \xrightarrow{a.s} y^o \) in \( L^2(I,V). \) Hence there exists a subsequence of the sequence \( \{y^n\}, \) relabeled as \( \{y^n\}, \) such that \( y^n(t) \xrightarrow{a.s} y^o(t) \) in \( V \) for almost all \( t \in I \) and \( P-a.s. \) By virtue of the growth property of \( \ell \) (See (a1)), and the apriori bound (7), it is clear that the integrands are dominated by \( \mu \)-integrable functions. Further, by assumption, \( \ell = \ell(t,x,\xi) \) is continuous on \( V \) in its second argument uniformly with respect \( \xi \in U \) for almost all \( t \in I. \) Since the controls are probability measure valued \( G_t \)-adapted random processes, it is clear from the above facts that, for any \( \varepsilon > 0, \) there exists \( n_{2,\varepsilon} \in N \) such that for all \( n \geq n_{2,\varepsilon} \)

\[ \left| E \int_I [\ell(t,y^o(t),u^n) - \ell(t,y^n(t),u^n)]dt \right| \]
\[ = \left| E \int_I E \{\ell(t,y^o(t),\xi) - \ell(t,y^n(t),\xi)\}G_t u^n d\xi dt \right| \leq \varepsilon/2 \]

uniformly with respect to controls \( u \in \mathcal{U}_{ad}. \) Thus it follows from the above results that, for any \( \varepsilon > 0, \) we have

\[ J_1(u^o) \leq \varepsilon + J_1(u^n), \forall n > (n_{1,\varepsilon} \lor n_{2,\varepsilon}). \]
Since, $\varepsilon > 0$, is arbitrary and the left hand side of the above inequality is independent of $n \in N$, we arrive at the following inequality

$$J_1(u^n) \leq \lim_{n \to \infty} J_1(u^n),$$

proving the lower semi-continuity of $J_1$. Next, we prove that the second term $J_2(u) = E\Phi(y(T))$ is also lower semi-continuous. Let $L^2_{\mathcal{F}_t}(\Omega, H)$ denote the $L_2$ space of $\mathcal{F}_t$ measurable random variables taking values in the Hilbert space $H$ and having square integrable norms. In view of Theorem 4.1 (see expression (44)), it is clear that $y^n(t)$ converges to $y^\alpha(t)$ in $L^2_{\mathcal{F}_t}(\Omega, H)$ for each $t \in I$, whenever $u^n \to u^\alpha$. Thus it follows from lower semi-continuity of $\Phi$ that

$$\Phi(y^n(T)) \leq \lim_{n \to \infty} \Phi(y^n(T)) \quad P - a.s.$$

By assumption (a2) we have

$$\Phi(y^n(T)) \leq \gamma + \delta |y^n(T)|^2_{H} \quad P - a.s.$$

Since, by virtue of the apriori bound (see (7)), $\sup\{E|y^n(T)|^2_{H}, n \in N\}$ is finite, it is clear from the above inequality that

$$\sup\{E\Phi(y^n(T)), n \in N\} < \infty.$$

Thus it follows from Fatou’s lemma and the definition of $J_2$ that

$$J_2(u^n) = E\Phi(y^n(T)) \leq \lim_{n \to \infty} E\Phi(y^n(T)) = \lim_{n \to \infty} J_2(u^n)$$

proving lower semi-continuity of $J_2$. It is well known that the sum of a finite number of lower semi-continuous functions is lower semi-continuous. So we conclude that $J$ is lower semi-continuous in the sense stated in the Theorem. Since $U_{ad}$ is compact in the metric topology $d$ and $J$ is lower semi-continuous in this topology, it attains its minimum on it. This proves the existence of an optimal control.

In the study of mass transport and reachability problems, among others, it is important to consider the evolution of probability laws directly. Let $\mathcal{B}(H)$ denote the sigma algebra of Borel sets in the Hilbert space $H$ and let $\mathcal{M}_1(H)$ denote the space of regular Borel probability measures on $\mathcal{B}(H)$. We have seen in Theorem 3.1 that for each $u \in U_{ad}$ the system (4) has a unique weak solution which we denote by $y^u \in L^2_{\infty}(I, H) \cap L^2_p(I, V) \cap L^2_{2p}(I, X_p)$. In fact, it is clear from the apriori bound (7) that $y^u \in \mathcal{B}^2_{\infty}(I, H)$, which consists of $\mathcal{F}_t$-adapted $H$ valued random processes having bounded second moments. In other words, $y^u(t)$ has bounded second moment for each $t \in I$. Thus the corresponding probability measures are well defined as functions of $t \in I$ with values in $\mathcal{M}_1(H)$. For each $u \in U_{ad}$ and $t \in I$, we can define the measure valued function $\mu^u_t$ as follows,

$$\mu^u_t(\Gamma) = \text{Prob.}\{y^u(t) \in \Gamma\}, \quad \text{for any } \Gamma \in \mathcal{B}(H).$$

Using this we can introduce the attainable (or reachable) sets corresponding to the evolution equation (4) and the admissible controls $U_{ad}$ as follows

$$\mathcal{A}(t) \equiv \left\{ \mu^u_t, u \in U_{ad} \right\}, \quad t > 0.$$

**Theorem 5.2.** Consider the system (4) with the admissible controls $U_{ad}$ equipped with metric topology $d$. Suppose the assumptions of Theorem 4.1 hold. Then, for each $t \in I$, the attainable set $\mathcal{A}(t)$ is a weakly compact subset of the space of probability measures $\mathcal{M}_1(H)$. 


Proof. Let \( u^n \in A(t) \) be a generalized sequence. Then, by definition, there exists a sequence \( \{u^n\} \subset U_{ad} \) such that
\[
\nu^n(\cdot) = \mu^n_t(\cdot) = \text{Prob}\{y^n(t) \in (\cdot)\}.
\]
Since \( U_{ad} \) is compact in the metric topology \( d \), there exists a subsequence \( \{u^{n_k}\} \) of the sequence \( \{u^n\} \), and an \( u^o \in U_{ad} \) such that \( u^{n_k} \stackrel{d}{\longrightarrow} u^o \). Let \( \{y^{n_k}\} \) denote the weak solutions of equation (4) corresponding to the controls \( \{u^{n_k}\} \), and \( y^o \) the weak solution corresponding to the control \( u^o \). Then it follows from Theorem 4.1 that \( y^{n_k}(t) \stackrel{a.s.}{\longrightarrow} y^o(t) \) in \( L^2(\Omega, H) \). Hence there exists a subsequence of the sequence \( \{y^{n_k}\} \), relabeled as is, such that \( y^{n_k}(t) \stackrel{a.s.}{\longrightarrow} y^o(t) \) in \( H \)-P-a.s. Let \( BC(H) \) denote the space of continuous and bounded real valued functions on \( H \) equipped with the standard sup-norm topology. Clearly, this is a Banach space. Then, for any \( \varphi \in BC(H) \), it is clear that
\[
\varphi(y^{n_k}(t)) \longrightarrow \varphi(y^o(t)), P - a.s.
\]
Since \( \varphi \) is uniformly bounded on \( H \), it follows from dominated convergence theorem that
\[
E\varphi(y^{n_k}(t)) \longrightarrow E\varphi(y^o(t)).
\]
But this is precisely equivalent to
\[
\int_H \varphi(\xi)\mu^{n_k}_t(\xi)\,d\xi \longrightarrow \int_H \varphi(\xi)\mu^o_t(\xi)\,d\xi.
\]
Thus we have shown that any generalized sequence from the attainable set \( A(t) \), for any \( t \in I \), has a generalized subsequence that converges weakly to an element in \( A(t) \). This proves that the attainable \( A(t) \) is weakly compact. \( \square \)

Here, we consider an application of the above theorem. Consider the attainable set at time \( T \), \( A(T) \). Let \( \mu_0 \) denote the probability law of the initial state of the system (4) with support \( C_0 \), a closed subset of \( H \). Let \( C_1 \) be another closed subset of \( H \), possibly with \( C_1 \cap C_0 = \emptyset \), which we call the target set. The problem is to find a control that maximizes the content \( \mu^o_T(C_1) \). So for this problem we take the objective functional as
\[
J(u) \equiv \mu^o_T(C_1) \longrightarrow \sup.
\]

Corollary 5.3. Consider the system (4) with the objective functional (49) and suppose the assumptions of Theorem 5.2 hold. Then, for any closed set \( C_1 \subset H \), there exists a control \( u^o \in U_{ad} \) at which \( J \) attains its maximum.

Proof. Since by Theorem 5.2, the attainable set \( A(T) \) is a weakly compact subset of \( M_1(H) \), it suffices to prove that the functional
\[
\bar{J}(\mu) \equiv \mu(C_1)
\]
is weakly upper semi-continuous on the attainable set \( A(T) \). Let \( D \) be any directed set and let \( \{\mu_\alpha, \alpha \in D\} \) be a net or a generalized sequence. Since \( A(T) \) is weakly compact there exists a subnet (generalized subsequence), relabeled as the original net, such that \( \mu_\alpha \stackrel{w}{\longrightarrow} \mu_o \in A(T) \). Since \( C_1 \) is a closed set and \( A(T) \) is weakly compact, it follows from \([11\text{, Theorem 6.1}]\) that \( \liminf \mu_\alpha(C_1) \leq \mu_o(C_1) \) which is equivalent to
\[
\liminf \bar{J}(\mu_\alpha) \leq \bar{J}(\mu_o)
\]
This shows that the functional \( \tilde{J} \) is weakly upper semi-continuous on \( \mathcal{A}(T) \) and therefore it attains its maximum on it. Hence we conclude that there exists an optimal control \( u^* \) such that \( J(u^*) \geq J(u) \) \( \forall u \in \mathcal{U}_{ad} \).

Considering the system (4) with the set of admissible controls \( \mathcal{U}_{ad} \), another interesting problem is to find a control that produces a measure close to a given measure \( \nu \in \mathcal{M}_1(H) \). Let \( \rho \) denote the Prohorov metric on \( \mathcal{M}_1(H) \) and let \( \nu \in \mathcal{M}_1(H) \).

Using the metric \( \rho \) this problem can be formulated as follows: find a control that minimizes the following functional

\[
J_2(u) \equiv \rho(\mu^u_\tau, \nu) \rightarrow \inf.
\]  

(51)

**Corollary 5.4.** Consider the system (4) with the cost functional (51) and suppose the assumptions of Theorem 5.2 hold. Then there exists an optimal control \( u^0 \in \mathcal{U}_{ad} \) at which \( J_2 \) attains its minimum.

**Proof.** First we observe that the problem (51) is equivalent to the problem of seeking the minimum of the functional \( \Psi(\mu) \equiv \rho(\mu, \nu) \) on the attainable set \( \mathcal{A}(T) \). Thus it suffices to find a \( \mu^0 \in \mathcal{A}(T) \) at which the functional \( \Psi \) attains its minimum.

It is well known that the topology of weak convergence of probability measures is equivalent to the (metric) topology induced by the Prohorov metric \( \rho \). Thus if any generalized sequence \( \mu^n \in \mathcal{A}(T) \) converges weakly to \( \mu^* \), \( \rho(\mu^n, \nu) \rightarrow \rho(\mu^*, \nu) \). Hence the functional \( \Psi \) is weakly continuous and, since the attainable set \( \mathcal{A}(T) \) is weakly compact, \( \Psi \) attains its minimum on \( \mathcal{A}(T) \) at some point \( \mu^0 \in \mathcal{A}(T) \). Thus there exists a control \( u^0 \in \mathcal{U}_{ad} \) such that \( \mu^0 = \mu^u_\tau \) and \( J_2(u^0) = \rho(\mu^0, \nu) \leq J_2(u) \) for all \( u \in \mathcal{U}_{ad} \). This completes the proof. \( \square \)

Another interesting problem is the time optimal control problem. For example, let \( C \) be a closed set in \( H \). The problem is to find a control that minimizes the first hitting time of the set \( C \) with probability greater or equal to \( \alpha \) for a given \( \alpha \in (0,1) \). Define

\[
\tau(u) \equiv \inf\{t \geq 0 : \mu^u_t(C) \geq \alpha\}.
\]

If the above set is empty we set \( \tau(u) = \infty \).

**Corollary 5.5.** Consider the system (4) and suppose the assumptions of Theorem 5.2 hold. Then there exists a time optimal control.

**Proof.** We show that \( u \rightarrow \tau(u) \) is a lower semicontinuous functional on \( \mathcal{U}_{ad} \). Let \( \{u^n\} \in \mathcal{U}_{ad} \) be a sequence that converges in the metric topology \( d \) to \( u^0 \). Then it follows from Theorem 5.2 that, along a (generalized) subsequence, if necessary, \( \mu_t^{u^n} \xrightarrow{w} \mu_t^{u^0} \) for each \( t \geq 0 \). Since \( C \) is a closed set in \( H \), and, for each \( t \geq 0 \), \( \mathcal{A}(t) \) is weakly compact we have

\[
\liminf \mu_t^{u^n}(C) \leq \mu_t^{u^0}(C).
\]

Thus it follows from the above inequality that

\[
\{t \geq 0 : \mu_t^{u^0}(C) \geq \alpha\} \supset \{t \geq 0 : \liminf \mu_t^{u^n}(C) \geq \alpha\}.
\]

From the above inclusion and the definition of hitting time we can readily conclude that \( \tau(u^0) \leq \lim \tau(u^n) \). Thus the functional \( \tau \) is lower semi-continuous with respect to the metric topology \( d \) on \( \mathcal{U}_{ad} \) which is compact. Hence there exists a control \( u^0 \) at which \( \tau \) attains its minimum. \( \square \)
6. **An extension of Theorem 3.1.** In section 3, we considered the system

\[ dy = Aydt + f(y, u)dt + \sigma(t)dW \]

with \( \sigma \) independent of the state \( y \). Here we let \( \sigma \) depend on \( y \) giving

\[ dy = Aydt + f(y, u)dt + \sigma(y)dW \]

(52)

with the same initial state \( y(0) = y_0 \). We introduce the following assumption. The operator \( \sigma : H \to \mathcal{L}(E, H) \) is continuous and uniformly bounded, that is, there exists a positive number \( r \) such that

\[ \sup\{ \| \sigma(h) \|_{\mathcal{L}(E, H)}, h \in H \} \leq r. \]

Let \( \mathcal{L}(E, H) \) be endowed with the weak operator topology giving the locally convex topological vector space \( (\mathcal{L}(E, H), \tau_{wo}) = \mathcal{L}_{wo}(E, H) \). Consider the closed ball \( B_r(\mathcal{L}_{wo}(E, H)) \) of radius \( r \) around the origin. It is well known [8, Dunford and Schwartz] that the ball \( B_r(\mathcal{L}(E, H)) \) is compact in the weak operator topology. Thus \( B_r(\mathcal{L}_{wo}(E, H)) \) is a compact topological space. Let \( H_w \) denote the Hilbert space equipped with the weak topology. Consider the product space \( H_w \times \mathcal{L}_{wo}(E, H) \). The graph of the operator \( \sigma \) is given by

\[ \mathcal{G}(\sigma) = \{(h, L) \in H_w \times \mathcal{L}_{wo}(E, H) : L = \sigma(h)\}. \]

**Theorem 6.1.** Consider the control system (52). Suppose the operators \( A \) and \( f \) satisfy the assumptions of Theorem 3.1. The operator \( \sigma : H_w \to \mathcal{L}_{wo}(E, H) \) is continuous having closed graph \( \mathcal{G}(\sigma) \) and uniformly bounded satisfying

\[ \sup\{ \| \sigma(h) \|_{\mathcal{L}(E, H)}, h \in H \} \leq r < \infty. \]

(53)

Then, for any initial state \( y_0 \in L^p_2(\Omega, H) \equiv L^p_2(\Omega, H) \), equation (52) has at least one weak solution \( y \in L^p_\infty(I, H) \cap L^2(I, V) \cap L^2_2(I, X_p) \).

**Proof.** The proof is similar to that of Theorem 3.1 with some additional technical justifications required for the diffusion term only. The finite dimensional approximation (10) of the system (4) is replaced by the following system

\[ dz^n = A z^n dt + F(z^n, u_d)dt + \Gamma(z^n(t))dB^n, \quad z^n(0) = \{(y_{0j}, v_j) : j = 1, 2, \cdots, n\}, \]

(54)

where now the matrix \( \Gamma \) is state dependent. The proof of existence of solution of this equation is again based on implicit difference scheme and Crandall-Ligget generation theorem as before. Because \( \sigma \) is uniformly bounded on \( H \), the apriori estimate (7) remains unchanged. All the terms of equation (17), with the exception of the last term denoted by \( L_n \), remain the same. The last term is now given by

\[ L_n = \int_I (\varphi(t)\sigma^*(y^n(t))v_i) dW^n_E. \]

(55)

where \( \varphi \in C^1(I) \) and \( W^n = \sum_{i=1}^{n} \sqrt{\lambda_i}(W(t), e_i)_E e_i = \sum_{i=1}^{n} \sqrt{\lambda_i} e_i \beta_i(t) \) with \( \{\beta_i\} \) being a sequence of mutually independent standard scalar Brownian motions. Define \( \sigma_n(t) \equiv \sigma(y^n(t)), t \in I \). Since, by assumption (53), \( \sigma \) is uniformly bounded on \( H \), it is clear that \( \sigma_n(t) \in \mathcal{B}(\mathcal{L}_{wo}(E, H)) \) for all \( t \in I \) with probability one. Clearly, \( L_n \) can be written as

\[ L_n = \sum_{j=1}^{n} \int_I \sqrt{\lambda_j} \varphi(t)(v_i, \sigma_n(t)e_j)_H d\beta_j. \]

(56)
It is easy to verify that this random variable has finite second moment and this follows from the following estimate
\[ E|L_n|^2 \leq r^2 Tr(Q) \int I|\phi(t)|^2 dt < \infty \]
which holds uniformly with respect to \( n \in N \). Since the ball \( B_r(\mathcal{L}_{wo}(E, H)) \) is compact (in the weak operator topology), the set \( B^a_\infty(I, B_r(\mathcal{L}_{wo}(E, H))) \) is compact with respect to Tychonoff product topology \( T_{wo} \). Thus there exists a subsequence of the sequence \( \sigma_n \), relabeled as the original sequence, and a \( \sigma_o \in B^a_\infty(I, B_r(\mathcal{L}_{wo}(E, H))) \) such that \( \sigma_n \overset{wo}{\rightarrow} \sigma_o \in B^a_\infty(I, B_r(\mathcal{L}_{wo}(E, H))) \). Multiplying (56) by \( \zeta \in L_\infty(\Omega) \) and taking the expectation of the resulting product and letting \( n \to \infty \) we obtain
\[ \lim_{n \to \infty} E\{\zeta L_n\} = E\left\{ \zeta \sum_{j=1}^\infty \int_1^\infty \sqrt{\lambda_j} \phi(t)(v, \sigma_o(t)e_j)_{H\bar{H}} d\beta_j \right\} = E\left\{ \zeta \int_1^\infty \phi(t)(v, \sigma_o(t)dW(t))_H \right\} = E\{\zeta L\}. \quad (57) \]
Since by assumption \( \sigma \) has closed graph and the element \( \zeta \in L_\infty(\Omega) \) is arbitrary, it follows from (57) that
\[ L = \int_1^\infty \phi(t)(v, \sigma_o(t)dW(t))_H = \int_1^\infty \phi(t)(v, \sigma(y^o(t))dW(t))_H. \quad (58) \]
Thus the evolution equation (52) has at least one weak solution. This completes the outline of our proof. \( \square \)

**Remark 6.2.** In view of the above theorem the results of section 5 on existence of optimal controls remain valid.

**Remark 6.3.** Another useful extension is given by the following reaction-diffusion-transport equation
\[ dy = Aydt + Bydt + f(y, u)dt + \sigma(y)dW \]
where \( B \in \mathcal{L}(V, H) \) is the transport operator. In this case there is both diffusion and migration of population. Migration is a function of concentration gradient. Assuming that there exists a \( b > 0 \) such that \( |Bl|_H \leq b \| y \|_V \), the apriori bound given by the expression (7) remains valid with minor adjustment as follows:
\[ E|y(t)|^2_H + \alpha \int_0^t E \| y(s) \|^2_V ds + 2\beta \int_0^t E \| y(s) \|^2_{X_p} ds 
\leq E|y_0|^2_H + (2\delta + (b^2/\alpha)) \int_0^t E|y(s)|^2_H ds, \forall t \in I. \quad (59) \]
Thus all the results of sections 3-5 apply to this extended model.

**Remark 6.4.** Regular controls are strongly measurable functions with values in a weakly compact set \( U \subset Y \), where \( Y \) is any real Banach space, while relaxed controls are weak star measurable processes with values in \( M_1(U) \). In this paper we proved existence of optimal relaxed controls. It is well known that, in general (in the absence of convexity of \( U \) or convexity of the vector field), optimal controls may not exist in the class of regular controls \([3],[4]\). For relaxed controls, convexity of \( U \) is not required. However, if one sacrifices the generality of the reaction operator
f, one can prove existence of optimal control from the class of regular controls with $U$ convex. For example,

$$f(y, u) = F(y) + G(y)u,$$

where $F$ satisfies all the assumptions $(A1)$-$(A3)$ and $G : V \to \mathcal{L}(Y, H)$ satisfies the Lipschitz condition: $\|G(y_1) - G(y_2)\|_{\mathcal{L}(Y, H)} \leq \gamma \|y_1 - y_2\|_V$ for some $\gamma > 0$ and finite. In the proof we use the fact that the embedding $H \hookrightarrow V^*$ is compact.

**Open Problem.** With additional regularity assumptions for $f$ one can develop necessary conditions of optimality as in [2]. We leave it as an open problem.

**Acknowledgments.** This work was partially supported by the National Science and Engineering Research Council of Canada under the grant no. A7109.

**REFERENCES**

[1] R. A. Adams, *Sobolev Spaces*, Academic press, New York-London, 1975.

[2] N. U. Ahmed, Optimal Control of reaction diffusion equations with potential applications to biomedical systems, *Journal of Abstract Differential Equations and Applications*, 8 (2017), 48–70.

[3] N. U. Ahmed and K. L. Teo, *Optimal Control of Distributed Parameter Systems*, Elsevier, North Holland, New York, Oxford, 1981.

[4] N. U. Ahmed, *Elements of Finite Dimensional Systems and Control Theory*, Pitman Monographs and Surveys in Pure and Applied Mathematics, Vol. 37, Longman Scientific and Technical, U.K., 1988.

[5] N. U. Ahmed, *Optimization and Identification of Systems Governed by Evolution Equations on Banach Spaces*, Pitman Research Notes in Mathematics Series, 184. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1988.

[6] P. L. Chow, Explosive solutions of stochastic reaction diffusion equations in mean $L_p$-norm, *Journal of Differential equations*, 250 (2011), 2567–2580.

[7] J. Diestel and J. J. Uhl, Jr. *Vector Measures*, Mathematical Surveys and Monographs, AMS, Providence, R.I., 1977.

[8] N. Dunford and J. T. Schwartz, *Linear Operators, Part 1: General Theory*, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1988.

[9] X. He and W.-M. Ni, Global dynamics of the Lotka-Volterra competition-diffusion system: Diffusion and spatial heterogeneity 1, *Communication in Pure and Applied Mathematics*, 69 (2016), 981–1014.

[10] C. Neuhauser and S. W. Pacala, An explicitly spatial version of the Lotka-Volterra model with interspecific competition, *The Annals of Applied Probability*, 9 (1999), 1226–1259.

[11] K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York and London, 1967.

[12] J. Smoller, *Shock Waves and Reaction-Diffusion Equations,(2nd Ed.)*, Springer Science and Business Media, LLC, 1994.

[13] R. Temam, *Infinite Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo, 1988.

Received January 2017; revised March 2017.

E-mail address: nahmed@uottawa.ca