Pure and Applied Functional Analysis
Volume 8, Number 1, 2023, 1–26

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

OPTIMAL CONTROL OF MEASURES INDUCED BY STOCHASTIC DIFFERENTIAL EQUATIONS ON UMD BANACH SPACES

N. U. AHMED

Abstract. In this paper we consider stochastic evolution equations on Banach spaces and their optimal control. We present briefly some recent results on existence and regularity properties of solutions of stochastic differential equations on UMD Banach spaces. We use these results in control theory and prove existence of optimal relaxed controls for partially observed control problems. Also we study the properties of measure valued functions induced by the solution trajectories and prove weak compactness of reachable set of measures. Then we prove existence of optimal controls for several control problems involving functionals of induced measure valued functions and problems related to first escape (blow up) time.

1. Introduction

In this paper we consider control problems of stochastic evolution equations on UMD Banach spaces. Our work is based on some fundamental results on existence, uniqueness and regularity properties of solutions of stochastic differential equations on UMD Banach spaces due to Neerven, Verrar and Weis [17, 19] and Brzeźniak [7, 11]. It is clear from their work that, it is the UMD Banach spaces introduced by Burkholder [10, 13] which play the crucial role in the development of stochastic integration in Banach spaces like the Iˆto integration in Hilbert spaces as seen in Da Prato and Zabczyk [12]. Further, in the Banach space setting, γ-Radonifying operators play the same role as the Hilbert-Schmidt operators do in the Hilbert space setting [12]. We use these results as our starting point for study of control problems on UMD Banach spaces. It is interesting to note that most of the work on optimal control of stochastic systems are based on Hilbert space setting as seen in [3, 6] and the references therein. Recently Brzeźniak and Serrano [8] have considered control problems for a class of stochastic systems on UMD (type-2) Banach space proving existence of optimal controls for Bolza problem. They prove the existence of optimal relaxed controls under the assumptions that the sum of the principal operator A (more precisely, its restriction to a Banach space B continuously embedded in the UMD-type-2 Banach space E) and the nonlinear drift F, giving A_B+F, is dissipative while F is continuous in all its arguments. The (diffusion) operator G is assumed to

2020 *Mathematics Subject Classification.* 35K10, 35L10, 93C20.
Key words and phrases. Stochastic evolution equations on Banach spaces, existence of solutions, existence of optimal controls, reachable set of measures, control of measures, first escape time.
This work was partially supported by the National Science and Engineering Research Council of Canada.

be given by $(-A)^{\sigma}G_0$, $\sigma \in (0,1)$, where the map $G_0 : I \times B \longrightarrow \gamma(H,E)$ is bounded and continuous with respect the state variable. Further, it is also assumed in [8] that the operator G is independent of the control variable. The technique used by Brzeźniak and Serrano [8] to prove existence of optimal relaxed controls is based on an extension of Skorohod representation theorem where the probability space is not fixed a priori. In [4,5] we considered relaxed control problems in Hilbert space setting under somewhat similar assumptions requiring the nonlinear drift F to be dissipative admitting polynomial growth and A to be the generator of an analytic semigroup. But only additive noise is admitted.

Here in this paper we assume that A is the infinitesimal generator of an analytic semigroup and that both the drift and the diffusion operators F and G are Lipschitz in the state variable from E to $E_{-\theta_1}$ and $E_{-\theta_2}$, $\theta_1, \theta_2 \in (0,1)$, respectively. This implies that both F and G are unbounded nonlinear operators on the state space E. But both contain control variables unlike in [9]. Our approach is different. We introduce a metric topology on the space of relaxed controls adapted to a current of subsigma algebras of the parent sigma algebra $\mathcal{F}$, and then prove continuity of the control to solution map. Using this result we prove the existence of optimal relaxed controls for Bolza problem. Further on, we prove weak compactness of the reachable set of measures induced by the mild solutions. Using the reachable set we formulate several interesting control problems involving the induced measures, and present several results on existence of optimal relaxed controls for such problems. Such problems are not considered in [9]. Thus there are two fundamental differences between our paper and the paper of Brzeźniak-Serrano [9]. We do not impose dissipativity condition on the drift but we demand Lipschitz property for both drift and diffusion. However, we believe the Lipschitz assumption can be relaxed to local Lipschitz property using stopping time arguments. In [9], the authors assume the diffusion to be free of control while we assume that both the drift and diffusion operators are explicitly control dependent. Another interesting difference is that Brzeźniak-Serrano paper uses weak formulation based on generalized Skorohod representation, and in contrast, we use strong formulation. Thus the results of [9] and this paper are complementary. To admit non-Lipschitz, and non-dissipative vector fields one is required to extend the notion of mild solutions to relaxed solutions or measure valued solutions as seen in [3, 5].

The rest of the paper is organized as follows. In section 2, we state some well-known facts on γ-Radonifying operators, the class of UMD Banach spaces, and the notions of type and cotype of Banach spaces. Further, we discuss the question of integrability of Banach space valued (operator valued) stochastic processes with respect to cylindrical Brownian motion. In section 3, we introduce stochastic differential equations on Banach spaces and present some special spaces in which the solutions are expected to reside. In section 4, we present the basic assumptions on the operators describing the stochastic system and consider the question of existence and uniqueness of mild solutions after presenting some basic properties of γ-Radonifying operators and stochastic convolutions. In section 5, we consider control problems and introduce a metric topology on the space of relaxed controls. We prove continuity of the control to solution map with respect to the metric topology on the set of admissible controls and the norm topology on the path space of solutions. Then we prove the existence of optimal relaxed controls for Meyer's problem. Further, in section 6, we present some compactness properties of induced measures. In particular, we prove weak compactness of the reachable set of measures and then consider several interesting control problems involving objective functionals which are suitable functions of induced measures and measure valued functions. The author is not aware of any paper dealing with such problems in the context of SDE on Banach spaces.

2. UMD spaces and γ-radonifying operators

For study of stochastic differential equations on Banach spaces we need some basic concepts and definitions not encountered in the study of stochastic differential equations on Hilbert spaces. These are UMD-Banach spaces, γ-Radonifying operators and type and co-type of Banach spaces. We present these materials here briefly. For details the reader is referred to the literature [7, 10, 11, 13, 16, 17].

Let H be a real separable Hilbert space and E a real Banach space and $\{\gamma_n\}$ a sequence of mutually independent, centered, zero mean, standard Gaussian random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. Let $\{h_n\}$ be a complete orthonormal basis of H and let $\mathcal{F}(H, E)$ denote the class of (linear) finite rank operators from H to E. An element $L \in \mathcal{F}(H, E)$ has the representation $L \equiv \sum_i^k h_i \otimes e_i$ with $Lh \equiv \sum_{i=1}^k (h, h_i)e_i$ where $\{e_i\} \in E$ is an arbitrary sequence from the Banach space E and k is any finite positive integer. Let $\mathcal{F}(H, E)$ be given the norm topology

$$\| L \|_{\gamma(H,E)} \equiv \left(\mathbf{E} \| \sum_{i=1}^k \gamma_i L h_i \|_E^2 \right)^{1/2} \equiv \left(\mathbf{E} \| \sum_{i=1}^k \gamma_i \ e_i \|_E^2 \right)^{1/2}.$$

Completion of $\mathcal{F}(H, E)$ with respect to this norm topology is a Banach space which is denoted by $\gamma(H, E)$ and it is called the space of γ-Radonifying operators. Since finite rank operators $\mathcal{F}(H, E)$ are compact, and the γ-Radonifying operators are given by the limits of finite rank operators with respect to the above norm topology, γ-Radonifying operators are also compact. If E is also a Hilbert space, the space $\gamma(H, E)$ coincides with the space of Hilbert-Schmidt operators denoted by $\mathcal{L}_2(H, E)$ and the Hilbert-Schmidt norm equals the γ-Radonifying norm. In other words, when both H and E are Hilbert spaces, the space $\gamma(H, E)$ is isometrically isomorphic to $\mathcal{L}_2(H, E)$ and this property is symbolized by the expression $\gamma(H, E) \cong \mathcal{L}_2(H, E)$.

Let E be any real Banach space and let $FON(H)$ denote the class of all finite subsets of any system of complete orthonormal basis of the Hilbert space H. It is known, Neerven [17] and Brzeźniak [7, 11], that for $L \in \gamma(H, E)$

$$\| L \|_{\gamma(H,E)}^2 = \sup_{h \in FON(H)} \mathbf{E} \| \sum \gamma_i L h_i \|_E^2,$$

where the supremum is taken over all finite orthonormal systems in the Hilbert space H. In view of this, the class of γ-Radonifying operators is formally defined as follows.

Definition 2.1 (γ-Radonifying Operators)**.** Let E be a Banach space and H a separable Hilbert space. A bounded linear operator $L \in \mathcal{L}(H, E)$ is said to be a γ-Radonifying operator if there exists a constant $C_L > 0$ such that

$$\mathbf{E} \parallel \sum \gamma_n L h_n \parallel_E^2 \leq \ C_L \tag{2.1}$$

for every complete orthonormal system $\{h_n\}$ of H. In other words, the sum converges in $L_2(\Omega, E)$ independently of the choice of the orthonormal system.

Another definition of γ-Radonifying operators is more intuitive. Let γ denote the cylindrical Wiener measure on the Hilbert space H, and E a separable Banach space. An operator $L \in \mathcal{L}(H, E)$ is said to be γ-Radonifying if

$$\parallel L \parallel_{\gamma(H,E)} \equiv \left(\int_H \parallel Lh \parallel_E^2 d\gamma(h) \right)^{1/2} = \left(\int_E \parallel z \parallel_E^2 d\mu(z) \right)^{1/2} < \infty,$$

where $\mu \equiv \gamma o L^{-1}$ is a Borel measure defined on the Borel algebra $\mathcal{B}(E)$ of subsets of the Banach space E. It is interesting to mention that any $L \in \gamma(H, E)$ maps the cylindrical Wiener measure γ on the Hilbert space H to a Radon measure μ defined on $\mathcal{B}(E)$. This means that the γ-Radonifying operators are sufficiently regular or smoothing. An excellent illustration of γ-Radonifying operators, based on special pairs of Hilbert and Banach spaces (H, E), in terms of integral operators in classical Banach spaces $L_p, 1 < p < \infty$, can be found in Brzeźniak in [11].

Definition 2.2 (UMD Space)**.** A Banach space E is called an UMD-space (Unconditional Martingale differences) if for each $1 < p < \infty$ and every E valued L_p martingale difference sequence $\{d_i\}$, there exists a constant $\alpha > 0$ such that for any $\varepsilon \in \{-1, 1\}^n$ the following inequality holds

$$\mathbf{E} \parallel \sum_{i=1}^{n} \varepsilon_i d_i \parallel_E^p \leq \alpha^p \ \mathbf{E} \parallel \sum_{i=1}^{n} d_i \parallel_E^p \tag{2.2}$$

for every $n \in N$.

It is known that the UMD property is independent of $p \in (1, \infty)$. If E is a Hilbert space, it is easy to verify that $\alpha = 1$. It is also known that, for all $p \in (1, \infty)$, all L_p spaces (over sigma finite measure spaces) are UMD spaces and clearly they are also reflexive Banach spaces. In general UMD spaces are reflexive Banach spaces but the converse is false. For details on UMD spaces see [7, 9–11, 13, 16, 17].

Definition 2.3 (Type and Co-type [Tzafriri, 18])**.** A Banach space E is said to be of type $p \in [1, 2]$ if, and only if, there exists a constant $C > 0$ such that for any $n \in N$ and for any sequence $\{e_i\}_{i=1}^n \in E$ and any symmetric i.i.d random variables $\{\zeta_i\}$ with values $\{-1, 1\}$ the following inequality holds

$$\mathbf{(T1)}: \left(\mathbf{E} \parallel \sum_{\mathbf{i=1}}^{\mathbf{n}} \zeta_{\mathbf{i}} \mathbf{e_i} \parallel_{\mathbf{E}}^{\mathbf{2}} \right)^{\mathbf{1/2}} \leq \mathbf{C} \left(\sum_{\mathbf{i=1}}^{\mathbf{n}} \parallel \mathbf{e_i} \parallel_{\mathbf{E}}^{\mathbf{p}} \right)^{\mathbf{1/p}}.$$

The smallest constant C for which the above inequality holds is called the type p-constant of E and denoted by $C_p(E)$.

A Banach E is said to have co-type $p \in [2,\infty)$ if the reverse inequality holds, that is, there exists a constant $\tilde{C}$ such that

$$\mathbf{(T2)}: \ \Big(\sum_{\mathbf{i=1}}^{\mathbf{n}} \parallel \mathbf{e_i} \parallel_{\mathbf{E}}^{\mathbf{p}}\Big)^{\mathbf{1/p}} \leq \mathbf{\tilde{C}}\Big(\mathbf{E} \parallel \sum_{\mathbf{i=1}}^{\mathbf{n}} \zeta_{\mathbf{i}}\mathbf{e_i} \parallel_{\mathbf{E}}^{\mathbf{2}}\Big)^{\mathbf{1/2}}.$$

It is known that all normed spaces have type 1 and co-type ∞, and that the type for Banach spaces always lies in the interval $1 \leq p \leq 2$ and co-type lies in the interval $2 \leq p < \infty$. All Hilbert spaces have type 2. The type and also co-type is a kind of measure of disparity or distance of a Banach space with respect to Hilbert space. For Hilbert spaces the parallelogram identity,

$$\parallel x+y \parallel^2 + \parallel x-y \parallel^2 = 2\big(\parallel x \parallel^2 + \parallel y \parallel^2\big),$$

holds, and in fact this distinguishes Hilbert spaces from general Banach spaces. The inequality $\mathbf{(T1)}$ plays somewhat similar role for Banach spaces as the parallelogram identity does for Hilbert spaces. For type 2 Banach space E, compare $\mathbf{(T1)}$ with the above parallelogram identity. If a Banach space has type 2 and co-type 2 then the inequality turns into an equality with $C = 2$ and in this case E must be isomorphic to a Hilbert space. Thus Hilbert space is at the center with type p Banach spaces sitting on the left and co-type p Banach spaces sitting on the right. It is known that the L_p spaces, for $1 \leq p \leq 2$, have type p and co-type 2; while the spaces $L_p, 2 \leq p < \infty$, have co-type p and type 2. It is known that if a Banach space X has type p then its dual X^* has co-type q for $1/p + 1/q = 1$. For more on type and co-type of Banach spaces see Tzafriri [19].

Stochastic Integrability: One crucial question that arises in the study of stochastic differential equations on Banach spaces is the question of integrability of Banach space valued random processes with respect to Wiener process. Here we present a brief outline of some crucial points. Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ denote a complete filtered probability space equipped with the filtration $\mathcal{F}_t, t \geq 0$, which is right continuous having left limits. Consider the stochastic integral,

$$z \equiv \int_0^T \Phi(t) dW_H, \tag{2.3}$$

where Φ is an $\mathcal{F}_t$-adapted operator valued process with values in $\mathcal{L}(H,E)$, and W_H is the H-cylindrical Brownian motion defined on the above probability space. It is well known that if both H and E are Hilbert spaces, the integral (2.3) is well defined as an Itô integral provided Φ is an $\mathcal{F}_t$ adapted random process taking values in the space of Hilbert-Schmidt operators, that is, $\Phi \in L_2^a(I \times \Omega, \mathcal{L}_2(H,E))$ where $\mathcal{L}_2(H,E)$ denotes the space of Hilbert-Schmidt operators from H to E. The superscript "a" indicates that the elements of this class are $\mathcal{F}_t$-adapted. Under these conditions $z \in L_2(\Omega, E)$ and

$$\mathbf{E} \parallel z \parallel_E^2 = \mathbf{E} \int_0^T Tr(\Phi^*(t)\Phi(t)) dt = \mathbf{E} \int_0^T \parallel \Phi(t) \parallel_{\mathcal{L}_2(H,E)}^2 dt.$$

Under the same conditions, the process $\{\xi(t), t \in I\}$, given by

$$\xi(t) \equiv \int_0^t \Phi(s) dW_H(s), t \in I,$$

is a well defined E valued stochastic process having continuous modification yielding $\xi \in L_2(\Omega, C(I,E))$ satisfying the well known Doob's martingale inequality,

$$\mathbf{E}\{\sup_{t\in I} \parallel \xi(t) \parallel_E^p\} \leq (p/p-1)^p \ \mathbf{E} \parallel \xi(T) \parallel_E^p,$$

for all $p \in (1,\infty)$.

In contrast, in the case of Banach spaces, the stochastic integral is well defined if E is a UMD space and the operator valued $\mathcal{F}_t$-adapted process Φ is γ-Radonifying, that is, $\Phi \in \gamma(L_2(I,H),E)$ P-a.s. The question, why we need UMD spaces and γ-Radonifying operators for stochastic integration in Banach spaces, can be clearly understood from the pioneering work of Neerven, Veraar and Weis [16, 17]. Here we present only an intuitive understanding of the subject by writing the integral (2.3) for step functions Φ. Consider the interval $I \equiv [0,T]$ and, for each $n \in N$, the partition

$$\Pi_n \equiv \{\Delta_j^n \equiv (t_{j-1}, t_j] = ((j-1)T/n, jT/n], j = 1,2,\cdots,n\}$$

of I into disjoint intervals $\{\Delta_j^n, j = 1,2,\cdots,n\}$ giving $I = \bigcup_{j=1}^n \Delta_j^n$ with the understanding that $t_0 = 0$. Letting d denote the diameter of the partition Π_n, it is clear that $d(\Pi_n) \to 0$ as $n \to \infty$. Let $\Phi(\cdot)$ be a step function or simple process with values in $\mathcal{L}(H,E)$ P-a.s and let $\mathcal{F}_{j-1} \equiv \mathcal{F}_{t_{j-1}} \subset \mathcal{F}$ denote the subsigma algebra corresponding to the time index $t_{j-1} \equiv (j-1)T/n, j = 1,2,\cdots n$, with $\Phi(t) = \Phi_{j-1}, t \in \Delta_j^n$, with $\Phi_{j-1} \in \mathcal{L}(H,E)$ being $\mathcal{F}_{j-1}$ measurable. Let us define the random element (variable) z_n as given by

$$z_n \equiv \sum_{j=1}^n \Phi_{j-1}[W_H(jT/n) - W_H((j-1)T/n))]. \tag{2.4}$$

Since H is separable, it has a complete orthonormal basis $\{h_k\}$ and we can define a sequence of standard and mutually independent Brownian motions as $\{\beta_k(t) \equiv (W_H(t), h_k), t \geq 0\}$. For each $k \in N$, define the sequence

$$\gamma_k(j-1) \equiv (\sqrt{(}n/T))[\beta_k(jT/n) - \beta_k((j-1)T/n)], j = 1,2,\cdots,n; k \geq 1.$$

Clearly, for each fixed $j \in \{1,2,\cdots,n\}$, this is a sequence of independent Gaussian random variables with mean zero and variance one. In fact, for each $j \in \{1,2,\cdots n\}$, the sequence $\{\gamma_k(j-1), k \in N\}$ is a martingale difference sequence. Thus the expression (2.4) can be rewritten as

$$z_n \equiv \sum_{j=1}^n \sum_{k=1}^\infty \gamma_k(j-1)\Phi_{j-1}h_k \equiv \sum_{j=1}^n d_j, \tag{2.5}$$

where for each $j \in \{1,2,\cdots,n\}$,

$$d_j \equiv \sum_{k=1}^\infty \gamma_k(j-1)\Phi_{j-1}h_k. \tag{2.6}$$

For convergence of the series (2.5) in the Banach space E, first and foremost, it is clearly necessary that, for each $j \in [1,2,\cdots,n]$, the series (2.6) converges in probability unconditionally in the Banach space E. Note that $\{d_j\}$ in the summation, $z_n \equiv \sum_{j=1}^n d_j$, is a martingale difference sequence. We want this sequence to converge strongly in $L_p(\Omega, E)$ for some $p \in [1, \infty)$. This requires that the Banach space E be a UMD space and that, for each j, $\{\Phi_j, j = 0, 1, \cdots, n-1\}$ be a γ-Radonifying operator belonging to $L_p(\Omega, \gamma(H, E))$ for some $1 \leq p < \infty$. This ensures that the sum on the right-hand side of equation (2.6) converges in $L_p(\Omega, E)$ and hence $d_j \in L_p(\Omega, E)$. Thus for each $n \in N$, z_n is a well defined random element belonging to $L_p(\Omega, E)$. Now assuming that Φ, as a random process, belongs to $L_p(\Omega, \gamma(L_2(I, H), E))$, and that the class of simple processes is dense in $L_p(\Omega, \gamma(L_2(I, H), E))$, one proves that the sequence z_n converges strongly to an element z in the Banach space $L_p(\Omega, E)$. This is the intuitive explanation of the stochastic integral (2.3) with values in E.

It is easy to verify that the covariance operator of the random element z is given by

$$< Qe^*, e^* >_{E^{**}, E^*} = \mathbf{E} \int_I < \Phi(t)\Phi^*(t)e^*, e^* >_{E, E^*} dt,$$

for $e^* \in E^*$. The random element z determines a regular probability measure on $\mathcal{B}(E)$ if, and only if, Q is nuclear belonging to the space of nuclear operators $\mathcal{L}_1(E^*, E) \subset \mathcal{L}_1(E^*, E^{**})$.

Remark 2.4. Using the expression for the norm topology of the Banach space $\gamma(H, E)$, it is easy to verify that it is both a left and right ideal in the Banach space of bounded linear operators endowed with the uniform operator topology. In particular, for any pair of UMD spaces $\{E, F\}$ and separable Hilbert space H, and for any $B \in \mathcal{L}(E, F)$, and $L \in \gamma(H, E)$, we have $BL \in \gamma(H, F)$ and $\| BL \|_{\gamma(H,F)} \leq \| B \|_{\mathcal{L}(E,F)} \| L \|_{\gamma(H,E)}$. Similarly, for any separable Hilbert space $\mathcal{H}$, and $C \in \mathcal{L}(\mathcal{H}, H)$, we have $LC \in \gamma(\mathcal{H}, E)$ and

$$\| LC \|_{\gamma(\mathcal{H},E)} \leq \| C \|_{\mathcal{L}(\mathcal{H},H)} \| L \|_{\gamma(H,E)} .$$

As an application of the ideal property, let us consider the stochastic convolution,

$$y(t) \equiv \int_0^t S(t-s)\Phi(s)dW_H(s), t \in I,$$

where $S(t), t \geq 0$, is a general C_0-semigroup Ahmed [1] on the Banach space E, and Φ is an $\mathcal{F}_t$-adapted stochastic process satisfying $\Phi \in L_2^a(\Omega, L_2(I, \gamma(H, E)))$ and E is a UMD Type-2 Banach space. Then, it follows from the preceding discussion on stochastic integration and ideal property that

$$\mathbf{E} \| y(t) \|_E^2 \leq c(M)\mathbf{E} \int_0^t \| \Phi(s) \|_{\gamma(H,E)}^2 ds < \infty, t \in I,$$

where $c(M)$ is a positive constant dependent on $M = \sup\{\| S(t) \|_{\mathcal{L}(E)}, t \in I\}$. Thus $y \in B_0^a(I, L_2(\Omega, E))$, the space of $\mathcal{F}_t$ adapted E valued strongly measurable random processes having uniformly bounded second moments. In contrast, if $S(t), t \geq 0$, is an analytic semigroup, the process y has better spatial regularity. As discussed in the next section, its infinitesimal generator A has the property that $-A$ has fractional powers, and using these fractional powers one can construct scales of Banach spaces $E_\eta \hookrightarrow E \hookrightarrow E_{-\eta}$ for any $\eta \in [0, 1)$ with E as the pivot space. Let

$\eta, \theta \in [0,1)$ satisfying $0 \leq \eta + \theta < 1/2$, and suppose $\Phi \in B_0^a(I, L_2(\Omega, \gamma(H, E_{-\theta})))$. Recall that, for analytic semigroups, there exists a constant $M_{\eta,\theta}$ such that

$$\| (-A)^{\eta+\theta} S(t) \|_{\mathcal{L}(E)} \leq M_{\eta,\theta}/t^{\eta+\theta}, t > 0.$$

For proof see Ahmed [1]. Again, let E be a UMD-type-2 Banach space. Then using the above properties and Fubini's theorem one can easily verify that

$$\begin{aligned}\mathbf{E} \| y(t) \|_{E_\eta}^2 &\equiv \mathbf{E} \| (-A)^\eta y(t) \|_E^2 \\ &\leq C_{\eta,\theta} \int_0^t (t-s)^{-2(\eta+\theta)} \mathbf{E} \| \Phi(s) \|_{\gamma(H, E_{-\theta})}^2 \, ds, \\ &\leq C_{\eta,\theta} T^{1-2(\eta+\theta)} \| \Phi \|_{B_0^a(I, L_2(\Omega, \gamma(H, E_{-\theta})))}^2,\end{aligned}$$

for all $t \in [0,T]$ where $C_{\eta,\theta}$ is a positive constant dependent on $M_{\eta,\theta}$. This shows that, if the semigroup is analytic we have $y \in B_0^a(I, L_2(\Omega, E_\eta))$ possessing better spatial regularity. Temporal regularity has been studied by Brzeźniak using the well known DaPrato-Kwapien-Zabczyk factorization technique asserting that the process $y \in C(I, E_\eta)$ P-a.s. For details see Brzeźniak [7, Theorem 2.4].

3. The system model and special Banach spaces

The system without control is governed by a class of nonlinear stochastic differential equations on a suitable Banach space. It is given by

$$dx = Axdt + F(t,x)dt + G(t,x)dW_H, t \in I \equiv (0,T], x(0) = x_0, \tag{3.1}$$

where A is the infinitesimal generator of a C_0-semigroup [1] in a Banach space E, and F and G are suitable Borel measurable maps to be clarified later and $\{W_H(t), t \geq 0\}$ is H-cylindrical Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$.

Throughout the rest of the paper we assume that the operator A is the infinitesimal generator of an analytic semigroup $S(t), t \geq 0$, and without loss of generality we may also assume that $0 \in \rho(A)$, the resolvent set of A, so that $-A$ has fractional powers [1]. Using the fractional powers one can introduce the scale of Banach spaces

$$E_\eta \hookrightarrow E \hookrightarrow E_{-\eta},$$

for $0 \leq \eta < 1$, where $E_\eta \equiv \{x \in E : (-A)^\eta x \in E\}$. The space E_η endowed with the norm topology, $\| x \|_{E_\eta} \equiv \| (-A)^\eta x \|_E$, is a Banach space. The space $E_{-\eta}$ is the completion of E with respect to the norm topology $\| x \|_{E_{-\eta}} \equiv \| (-A)^{-\eta} x \|$.

For study of SDE on UMD spaces, Neerven et all [17, 19] introduced several special Banach spaces. Let $(S, \mathcal{B}_S, \mu) \equiv S_\mu$ be a finite measure space. The space $L_2^\gamma(S_\mu, E_\eta)$ is defined as

$$L_2^\gamma(S_\mu, E_\eta) \equiv L_2(S_\mu, E_\eta) \cap \gamma(L_2(S_\mu), E_\eta).$$

Endowed with the norm topology,

$$\| \Psi \|_{L_2^\gamma} \equiv \| \Psi \|_{L_2(S_\mu, E_\eta)} + \| \Psi \|_{\gamma(L_2(S_\mu), E_\eta)},$$

the space $L_2^\gamma(S_\mu, E_\eta)$ is a Banach space. Note that the first one is the L_2-norm in the sense of Bochner and the second one is the γ-radonifying norm. Potentially there are several Banach spaces on which one can consider the question of existence of solutions of the evolution equation (3.1). We present a brief description of these spaces.

For $\alpha \in (0, 1/2)$, $\eta \in [0, 1)$, and $1 < p < \infty$, let us denote the space of all $\mathcal{F}_{t\geq 0}$-adapted E_η valued random processes defined on the interval $I \equiv [0, T]$ by $V^p_{\alpha,\infty}(I \times \Omega; E_\eta)$ which is endowed with the following norm topology,

$$\parallel \varphi \parallel_{V^p_{\alpha,\infty}} \equiv \left(\mathbf{E} \parallel \varphi \parallel^p_{C(I,E_\eta)} \right)^{1/p} + \sup_{t\in I} \left(\mathbf{E} \parallel (\mathbf{t} - \cdot)^{-\alpha} \chi_{[\mathbf{0},\mathbf{t}]}(\cdot) \varphi(\cdot) \parallel^{\mathbf{p}}_{\gamma(\mathbf{L_2}[\mathbf{0},\mathbf{t}],\mathbf{E}_\eta)} \right)^{\mathbf{1/p}}, \tag{3.2}$$

where $\chi_{[0,t]}(\cdot)$ denotes the indicator function of the set $[0, t]$. Closely related to this space, there are two other spaces $V^p_{\alpha,p}(I \times \Omega, E_\eta)$ and $\tilde{V}^p_{\alpha,\infty}(I \times \Omega; E_\eta)$ where the first one is given the norm topology

$$\parallel \varphi \parallel_{V^p_{\alpha,p}} \equiv \left(\mathbf{E} \parallel \varphi \parallel^p_{C(I,E_\eta)} \right)^{1/p} + \left(\int_I \mathbf{E} \parallel (\mathbf{t} - \cdot)^{-\alpha} \chi_{[\mathbf{0},\mathbf{t}]}(\cdot) \varphi(\cdot) \parallel^{\mathbf{p}}_{\gamma(\mathbf{L_2}[\mathbf{0},\mathbf{t}],\mathbf{E}_\eta)} \, \mathbf{dt} \right)^{\mathbf{1/p}}. \tag{3.3}$$

These are the two classes of path-wise continuous $\mathcal{F}_t$-adapted processes. The last one is given the following norm topology

$$\parallel \varphi \parallel_{\tilde{V}^p_{\alpha,\infty}} \equiv \left(\mathbf{E} \parallel \varphi \parallel^p_{B_0(I,E_\eta)} \right)^{1/p} + \sup_{t\in I} \left(\mathbf{E} \parallel (\mathbf{t} - \cdot)^{-\alpha} \chi_{[\mathbf{0},\mathbf{t}]}(\cdot) \varphi(\cdot) \parallel^{\mathbf{p}}_{\gamma(\mathbf{L_2}[\mathbf{0},\mathbf{t}],\mathbf{E}_\eta)} \right)^{\mathbf{1/p}} \tag{3.4}$$

where $B_0(I, E_\eta)$ denotes the Banach space of E_η-valued path-wise bounded measurable functions furnished with the standard sup norm topology. Clearly, $V^p_{\alpha,\infty}$ is a closed subspace of $\tilde{V}^p_{\alpha,\infty}$. It is known that, with respect to the above norm topologies, these are Banach spaces. It was shown in Neerven, Veraar and Weis [17] that under certain assumptions one can prove the existence and uniqueness of mild solutions of the evolution equation (3.1) in all of the three spaces introduced above. The proof is largely similar.

Throughout the rest of the paper we use the Banach space $V^p_{\alpha,\infty}$ endowed with the norm given by the expression (3.2) and, for convenience of notation, we denote it by V. For convenience of presentation, we shall also use the notation $V([s, t] \times \Omega, E_\eta)$, and when E_η is understood simply $V_{[s,t]}$, for the restriction of the elements of the Banach space $V^p_{\alpha,\infty} \equiv V$ to the interval $[s, t], 0 \leq s < t \leq T$.

4. Existence and uniqueness of solutions

Now we introduce the following basic assumptions used for study of existence of solutions of the stochastic differential equation (3.1) and their regularity properties.

Basic Assumptions:

(A1) The Banach space E is a UMD space of type $\tau \in [1,2)$ and A is the infinitesimal generator of an analytic semigroup [1] $S(t), t \geq 0$, on E. For details on semigroup theory see Ahmed [1].

(A2) There exist $\eta \in [0,1), \theta_1 \in [0,1)$ such that $F : I \times \Omega \times E_\eta \longrightarrow E_{-\theta_1}$ is Borel measurable and there exist constants $C_1 \geq 0, L_1 > 0$ such that
(1) $\| F(t,\omega,x) \|_{E_{-\theta_1}} \leq C_1(1+ \| x \|_{E_\eta}),\ \forall\, \omega \in \Omega$
(2) $\| F(t,\omega,x) - F(t,\omega,y) \|_{E_{-\theta_1}} \leq L_1 \| x - y \|_{E_\eta},\ \forall\, \omega \in \Omega.$
Further, for each $x \in E_\eta$, $(t,\omega) \longrightarrow F(t,\omega,x)$ is an $\mathcal{F}_t$-adapted $E_{-\theta_1}$ valued strongly measurable function.

(A3) There exists $\theta_2 \in [0,1)$ such that $G : I \times \Omega \times E_\eta \longrightarrow \gamma(H, E_{-\theta_2}) \subset \mathcal{L}(H, E_{-\theta_2})$ is H-strongly measurable and there exist constants $C_2 \geq 0, L_2 > 0$ such that for every $x, y \in L_2^\gamma(I_\mu, E_\eta)$ and for all $\omega \in \Omega$
(1) $\| G(\cdot,\omega,x) \|_{\gamma(L_2(I_\mu,H),E_{-\theta_2})} \leq C_2\big(1+ \| x \|_{L_2^\gamma(I_\mu,E_\eta)}\big)$
(2) $\| G(\cdot,\omega,x) - G(\cdot,\omega,y) \|_{\gamma(L_2(I_\mu,H),E_{-\theta_2})} \leq L_2\ (\| x - y \|_{L_2^\gamma(I_\mu,E_\eta)}).$
Further, for each $x \in E_\eta$, $(t,\omega) \longrightarrow G(t,\omega,x)$ is an $\mathcal{F}_t$-adapted $\gamma(H, E_{-\theta_2})$ valued strongly measurable function defined on $I \times \Omega$.

Before we consider the question of existence, uniqueness and other regularity properties of mild solutions of SDE (3.1), we need some preparatory materials as presented below. As stated earlier, we use the Banach space $V \equiv V^p_{\alpha,\infty}(I \times \Omega, E_\eta)$.

Definition 4.1. An $\mathcal{F}_t$-adapted process x, defined on the interval I and taking values in the Banach space E_η, is said to be a mild solution of the evolution equation (3.1) if it satisfies the following conditions:
(i) $x \in V \equiv V^p_{\alpha,\infty}(I \times \Omega, E_\eta)$ and
(ii) x satisfies the following integral equation,

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \int_0^t S(t-s)G(s,x(s))dW_H(s), t \in I. \tag{4.1}$$

To consider the question of existence of solutions of the above integral equation we introduce the family of operators, $\{\Upsilon_r, r \in I\}$ on V, where Υ_r is given by the following integral operator,

$$(\Upsilon_r x)(t) \equiv S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \int_0^t S(t-s)G(s,x(s))dW_H(s), t \in [0,r], \tag{4.2}$$

for $x \in V$. It is clear that the question of existence of a solution of the integral equation (4.1) is equivalent to the question of existence of a fixed point of the operator Υ_T. We prove that Υ_T has a unique fixed point in V.

We quote two fundamental estimates for deterministic and stochastic convolutions due to Neerven, Veraar and Weis [17, 19]. These are used in the study of existence and regularity properties of solutions of the integral equation (4.1). For proof see [17, Lemma 4.2, p17; Proposition 6.1, p26]

Deterministic Convolution:

Lemma 4.2. *Let* $\alpha \in (0, 1/2), \eta \geq 0, \theta_1 \geq 0, p > 2$ *satisfy* $\eta + \theta_1 < 1$. *Let* A *be the infinitesimal generator of an analytic semigroup* $\{S(t), t \geq 0\}$ *on a UMD Banach space* E *of Type* $\tau \in [1, 2]$, *and let* $\Psi : I \times \Omega \longrightarrow E_{-\theta_1}$ *be measurable and* $\mathcal{F}_t$*-adapted and belong to* $L_p^a(\Omega, B_0(I, E_{-\theta_1}))$. *Then there exists a constant* $c_1 > 0$ *such that the process* $z_1 \equiv \{z_1(t), t \in I\}$, *given by*

$$z_1(t) \equiv \int_0^t S(t-s)\Psi(s)ds, t \in I,$$

satisfies the estimate $\| z_1 \|_V \leq c_1 T^{(1/2-\alpha)\wedge(1-\eta-\theta_1)} \| \Psi \|_{L_p(\Omega, B_0(I, E_{-\theta_1}))}$.

Proof. See expression 6.5 [17, Lemma 6.1, p27]. □

Stochastic Convolution:

Lemma 4.3. *Let* $\alpha \in (0, 1/2), \lambda \geq 0, \eta \geq 0, \theta_2 \geq 0, p > 2$ *satisfy* $\lambda+\eta+\theta_2 < \alpha-1/p$. *Let* A *be the infinitesimal generator of an analytic semigroup* $\{S(t), t \geq 0\}$ *on a UMD Banach space* E *of Type* $\tau \in [1, 2]$ *and let* $\Phi : I \times \Omega \longrightarrow \mathcal{L}(H, E_{-\theta_2})$ *be* $H-$*strongly measurable and* $\mathcal{F}_t$*-adapted satisfying*

$$\sup_{t \in I} \Big\{ \mathbf{E} \| \chi_{[0,t]}(\cdot)(t - \cdot)^{-\alpha}\Phi(\cdot) \|^p_{\gamma(L_2(0,t;H), E_{-\theta_2})} \Big\} < \infty.$$

Then there exist constants $\varepsilon > 0$ *(dependent on* $\{\alpha, \eta, \theta_2\}$*), and* $c_2 > 0$ *such that the process* $z_2 \equiv \{z_2(t), t \in I\}$, *given by the following stochastic convolution*

$$z_2(t) \equiv \int_0^t S(t-s)\Phi(s)dW(s), t \in I,$$

satisfies the following estimate

$$\| z_2 \|_V \leq c_2 T^{\varepsilon\wedge(1/2-\eta-\theta_2)} \left(\sup_{t \in I} \mathbf{E} \| \chi_{[0,t]}(\cdot)(t - \cdot)^{-\alpha}\Phi(\cdot) \|^p_{\gamma(L_2(0,t;H), E_{-\theta_2})} \right)^{1/p}.$$

Proof. See expression 6.8 [17, Lemma 6.1, p28]. □

Next we present some important properties of the family of integral operators $\{\Upsilon_r, r \in I\}$ defined by the expression (4.2). In particular, we present some important estimates which are used to prove existence of solutions of the integral equation (4.1). These estimates are derived from Lemmas 4.2 and 4.3.

Theorem 4.4. *Let* E *be an UMD space with type* $\tau \in [1, 2)$ *and suppose the assumptions* (**A1**)-(**A3**) *hold and further the parameters* $\{\tau, p, \alpha, \eta, \theta_1, \theta_2\}$ *satisfy*

(i) $0 \leq \eta + \theta_1 < 3/2 - 1/\tau$
(ii) $0 \leq \eta + \theta_2 < 1/2$
(iii) $p > 2, \alpha \in (0, 1/2)$ *such that* $\eta + \theta_2 < \alpha - 1/p$.

Then, for every $x_0 \in L_p(\Omega, \mathcal{F}_0, E_\eta)$, *the operator* Υ_T *is well defined on the space* V. *Further, there exist constants* $C_3 > 0$ *and* $\Theta_r > 0$ *for* $r \in (0, T]$, *with the property:* $\lim_{r\downarrow 0} \Theta_r = 0$, *such that for all* $\varphi, \varphi_1, \varphi_2 \in V_{[0,r]}$, *restriction of* V *to the interval* $[0, r]$, *we have*

(S1) $\| \Upsilon_r(\varphi) \|_{V_{[0,r]}} \leq C_3(1+ \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}) + \Theta_r \| \varphi \|_{V_{[0,r]}}$

(S2) $\| \Upsilon_r(\varphi_1) - \Upsilon_r(\varphi_2) \|_{V_{[0,r]}} \leq \Theta_r \| \varphi_1 - \varphi_2 \|_{V_{[0,r]}}$.

Proof. See Neerven-Veraar-Weis [17, Proposition 6.1, p26]. □

We note that the map $r \longrightarrow \Theta_r$ used above is strictly a function of the measure of length of any interval $[a, a + r]$ for all $a, r \geq 0$ such that $a, a + r \in I$. In other words, Θ is actually a nonnegative measure absolutely continuous with respect to the Lebesgue measure.

Using the above estimates one can prove the following result on existence and uniqueness of solution of the integral equation (4.1). The existence result given below is originally due to Neerven-Veraar-Weis [17]. For some degree of completeness we present a brief outline of its proof.

Theorem 4.5. *Consider the integral equation (4.1) and suppose the assumptions of Lemma 4.2, Lemma 4.3 and Lemma 4.4 hold. Then, for every* $x_0 \in L_p(\Omega, \mathcal{F}_0, E_\eta), p > 2$, *the integral equation has a unique solution* $x \in V \equiv V^p_{\alpha,\infty}(I \times \Omega, E_\eta)$ *and there exists a constant* $\hat{C}$ *such that*

$$\| x \|_V \leq \hat{C}\big(1+ \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}\big). \tag{4.3}$$

Proof. We present a brief outline of the proof. It is based on classical Banach fixed point theorem. For any $r \in I$, let $V_{[0,r]} \equiv V^p_{\alpha,\infty}([0, r] \times \Omega, E_\eta)$ denote the restriction of V to the interval $[0, r]$ as stated above. It follows from Lemma 4.4, in particular (**S1**) and (**S2**), that for any $r \in I$, Υ_r (the restriction of the operator Υ_T to $V_{[0,r]}$) maps $V_{[0,r]}$ into itself and it follows from the property of Θ_r that, for r sufficiently small, $\Theta_r \leq 1/2$. Thus for such a choice of r, the operator Υ_r is a contraction on the Banach space $V_{[0,r]}$ and therefore it has a unique fixed point in it. Hence the integral equation (4.1) has a unique solution over the interval $[0, r]$. Since $x \in V_{[0,r]}$, we have $x \in C([0, r], E_\eta)$ $P - a.s$, and hence $x(r)$ is well defined. Thus, starting with $x(r)$ as the initial state, and repeating the above procedure, one can verify that the operator Υ_T restricted to $V_{[r,2r]}$ has a unique fixed point. Since $r > 0$, and $T(> 0)$ is finite, continuing this process for a finite number of times, one can cover the entire interval $I \equiv [0, T]$ and conclude that Υ_T has a unique fixed point in V and hence the integral equation has a unique solution in V. This completes the outline of our proof. □

5. Optimal control

In this section we consider control of the following stochastic system

$$dx = Axdt + \hat{F}(t, x, u_t)dt + \hat{G}(t, x, u_t)dW_H(t), t \in I, \tag{5.1}$$

where u denotes the control and $\hat{F}$ and $\hat{G}$ are suitable functions to be defined shortly. The cost functional is given by

$$J(u) \equiv \mathbf{E}\left\{\int_0^T \hat{\ell}(t, x(t), u_t)dt + \Phi(x(T))\right\}. \tag{5.2}$$

The objective is to find a control policy $u \in \mathcal{U}_{ad}$ that minimizes the functional (5.2). Before we can do so we must introduce the set of admissible controls $\mathcal{U}_{ad}$.

We consider relaxed controls. Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ be a complete separable filtered probability space. Let U be a compact Polish space and $\mathcal{M}_1(U)$ the space of regular probability measures on $\mathcal{B}(U)$, the class of Borel subsets of the set U. Let $\mathcal{G}_{t\geq 0}$ be a nondecreasing family of subsigma algebras of the sigma algebra $\mathcal{F}_{t\geq 0}$ and let $\mathcal{P}$ denote the $\mathcal{G}_t$-predictable subsigma algebra of the product sigma algebra $\mathcal{B}(I) \times \mathcal{F}$ and let ν denote the restriction of the product measure $dt \times dP$ on $\mathcal{P}$ and introduce the Lebesgue-Bochner space $L_1(\nu, C(U))$ as ν measurable Bochner integrable processes with values in $C(U)$. Since the dual of $C(U)$ is given by the space of regular Borel measures $\mathcal{M}_B(U)$ and the later space does not satisfy Radon-Nikodym property (RNP), the dual of $L_1(\nu, C(U))$ is not given by $L_\infty(\nu, \mathcal{M}_B(U))$. However, it follows from the "theory of Lifting" [16, Theorem 7, p94] that the dual is given by $L_\infty^w(\nu, \mathcal{M}_B(U))$ which consists of weak star $\mathcal{P}$-measurable essentially bounded random processes with values in $\mathcal{M}_B(U)$. Recall that $\mathcal{M}_1(U) \subset \mathcal{M}_B(U)$ denotes the space of regular probability measures on $\mathcal{B}(U)$. Let $M_\infty^w(\nu, \mathcal{M}_1(U))$ denote the class of weak-star ν-measurable processes with values in the space of probability measures $\mathcal{M}_1(U)$. For admissible controls one may like to choose the set $M_\infty^w(\nu, \mathcal{M}_1(U))$. Clearly this is a closed bounded convex subset of $L_\infty^w(\nu, \mathcal{M}_B(U))$. Thus by Alaoglu's theorem this set is weak star compact. But this vague topology is rather too weak. We introduce a slightly stronger topology. Since U is a compact Polish space, $C(U)$ is a separable Banach space, and since the probability space is assumed to be separable, the Banach space $L_1(\nu, C(U))$ is separable. Therefore, it follows from Dunford-Schwartz [14, Theorem V.5.1, p426] that the set $M_\infty^w(\nu, \mathcal{M}_1(U))$ is metrizable. Let $\{\varphi_i\}$ be a set dense in $L_1(\nu, C(U))$. For $u, v \in M_\infty^w(\nu, \mathcal{M}_1(U))$ define the metric

$$d(u, v) \equiv \sum_{i=1}^{\infty}(1/2^i) \int_{I\times\Omega} \min\{1, |\varphi_i(u) - \varphi_i(v)|\} \, d\nu, \tag{5.3}$$

where $\varphi(u) \equiv \varphi(u)(t, \omega) \equiv \int_U \varphi(t, \omega, \xi)u_{t,\omega}(d\xi), (t, \omega) \in I \times \Omega$. This is a complete metric space. A sequence $\{u^n\} \in \mathcal{U}_{ad}$, converging in the metric topology d to u^o, is equivalent to convergence of $\varphi(u^n)$ to $\varphi(u^o)$ in ν-measure on $I \times \Omega$ for any $\varphi \in L_1(\nu, C(U))$. That is, for each $\varphi \in L_1(\nu, C(U))$,

$$\begin{aligned} I \times \Omega \ni (t, \omega) \longrightarrow \int_U &\varphi(t, \omega, \xi)[u_{t,\omega}^n(d\xi) - u_{t,\omega}^o(d\xi)] \\ &\equiv \varphi(u^n) - \varphi(u^o) = \varphi(u^n - u^o) \longrightarrow 0 \end{aligned}$$

in ν measure on $I \times \Omega$ as $n \to \infty$. We denote this metric space by (M, d) and note that it is a complete metric space. For the set of admissible controls we take any compact (closed and totally bounded) subset of the metric space (M,d), and denote it by $\mathcal{U}_{ad}$.

As stated earlier, for control problems (5.1) and (5.2) we have chosen the Banach space $V \equiv V^p_{\alpha,\infty}(I\times\Omega, E_\eta)$. In this case we need slight modification of the basic assumptions (**A1**), (**A2**), (**A3**). They are replaced by (**B1**), (**B2**), (**B3**) as follows:

(B1) = **(A1)**.

(B2) There exist $\eta\in[0,1), \theta_1\in[0,1)$ such that $F: I\times\Omega\times E_\eta\times U \longrightarrow E_{-\theta_1}$ is a Borel measurable $\mathcal{F}_{t\geq 0}$ adapted map, continuous in the last two arguments, and there exist constants $C_1\geq 0, L_1>0$ such that

(1) $\parallel F(t,\omega,x,\xi)\parallel_{E_{-\theta_1}} \leq C_1(1+\parallel x\parallel_{E_\eta}), \forall\,\omega\in\Omega$, uniformly with respect to $\xi\in U$,

(2) $\parallel F(t,\omega,x,\xi)-F(t,\omega,y,\xi)\parallel_{E_{-\theta_1}} \leq L_1\parallel x-y\parallel_{E_\eta}, \forall\,\omega\in\Omega$, uniformly with respect to $\xi\in U$.

Further, for each $x\in E_\eta$, and $\xi\in U$, $(t,\omega)\longrightarrow F(t,\omega,x,\xi)$ is an $\mathcal{F}_t$-adapted $E_{-\theta_1}$ valued strongly measurable function.

(B3) There exists $\theta_2\in[0,1)$ such that $G: I\times\Omega\times E_\eta\times U\longrightarrow \gamma(H,E_{-\theta_2})\subset \mathcal{L}(H,E_{-\theta_2})$ is H-strongly Borel measurable $\mathcal{F}_{t\geq 0}$ adapted map, continuous in the last two arguments, and there exist constants $C_2\geq 0, L_2>0$ such that, for every $x,y\in L^\gamma_2(I_\mu,E_\eta)$,

(1) $\parallel G(\cdot,\omega,x,\xi)\parallel_{\gamma(L_2(I_\mu,H),E_{-\theta_2})} \leq C_2\big(1+\parallel x\parallel_{L^\gamma_2(I_\mu,E_\eta)}\big)$, $\forall\,\omega\in\Omega$, uniformly in $\xi\in U$,

(2) $\parallel G(\cdot,\omega,x,\xi)-G(\cdot,\omega,y,\xi)\parallel_{\gamma(L_2(I_\mu,H),E_{-\theta_2})} \leq L_2\big(\parallel x-y\parallel_{L^\gamma_2(I_\mu,E_\eta)}\big)$, $\forall\,\omega\in\Omega$, uniformly in $\xi\in U$.

Further, for each $x\in E_\eta, \xi\in U$, $(t,\omega)\longrightarrow G(t,\omega,x,\xi)$ is an $\mathcal{F}_t$-adapted $\gamma(H,E_{-\theta_2})$ valued H-strongly measurable function.

The system is now given by the following controlled stochastic differential equation,

$$dx = Axdt + \hat{F}(t,x,u_t)dt + \hat{G}(t,x,u_t)dW_H,\ t\in I\equiv[0,T], \quad x(0)=x_0, \tag{5.4}$$

where for each $u\in\mathcal{U}_{ad}$, $\hat{\ell}(t,x,u_t)\equiv\int_U \ell(t,x,\xi)u_t(d\xi)$, and

$$\hat{F}(t,x,u_t)\equiv\int_U F(t,x,\xi)u_t(d\xi),\ \hat{G}(t,x,u_t)\equiv\int_U G(t,x,\xi)u_t(d\xi). \tag{5.5}$$

Recall that $\mathcal{U}_{ad}$ consists of relaxed controls convexifying nonconvex control problems. For example, U may be nonconvex; it may consist of a finite or a countable set of discrete points in a Polish space etc. In the absence of convexity optimal control may not exist in the class of regular controls ($\mathcal{P}$-measurable random processes with values in U). However relaxed controls may exist. We prove in this section that optimal controls exist in the class of relaxed controls.

Corollary 5.1. *Suppose the assumptions* (**B1**)-(**B3**) *hold. Then, for each* $x_0\in L_p(\Omega,\mathcal{F}_0,P)$, $\infty>p>2$, *and* $u\in\mathcal{U}_{ad}$, *the control system (5.4) has a unique mild solution* $x\equiv x(u)(\cdot)\in V$. *Further, the solution set* $\mathcal{S}\equiv\{x(u),u\in\mathcal{U}_{ad}\}$ *is a bounded subset of* V.

Proof. Under the assumptions **(B1)-(B3)**, for each given control $u \in \mathcal{U}_{ad}$ the assumptions **(A1)-(A3)** hold. Thus the statement on existence of solution follows as a corollary of Theorem 4.5. It remains to prove that the solution set $\mathcal{S}$ is bounded. Note that U is a compact Polish space and both F and G are Lipschitz with respect to the state variable $x \in E_\eta$ uniformly with respect to $\xi \in U$, and posses linear growth property in the state variable. Thus both $\hat{F}$ and $\hat{G}$ are also Lipschitz in the state variable uniformly with respect to the controls $u \in \mathcal{U}_{ad}$. This is what is used to prove the bound. Let $u \in \mathcal{U}_{ad}$ and let $x(u) \in V$ denote the corresponding solution. Then considering the restriction of V to $V[0,t]$ for $t \in I \equiv [0,T]$, it follows from Lemma 4.4, in particular the inequality **(S1)**, that for all $u \in \mathcal{U}_{ad}$, we have

$$\| x(u) \|_{V[0,t]} \leq C_3 \big[1 + \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}\big] + \Theta_t \| x(u) \|_{V[0,t]}, t \in I,$$

where the constant C_3 and the function Θ are independent of $u \in \mathcal{U}_{ad}$. It follows from the property of the function $t \longrightarrow \Theta_t$, as stated in Lemma 4.4, that there exists a $t_1 \in (0,T]$ such that $\Theta_{t_1} \leq (1/2)$. Thus, for $t = t_1$, it follows from the above inequality that

$$\| x(u) \|_{V[0,t_1]} \leq 2C_3 \big[1 + \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}\big], \forall u \in \mathcal{U}_{ad}.$$

Since $x \in V$ implies $x \in C(I, E_\eta)$ $P - a.s$, we conclude that $x(u)(t_1)$ is well defined and belongs to $L_p(\Omega, \mathcal{F}_{t_1}, E_\eta)$. Thus following similar steps we obtain the following inequality

$$\begin{aligned} x(u) \|_{V[t_1,t]} \leq C_3 \big[1 + \| x(u)(t_1) \|_{L_p(\Omega,\mathcal{F}_{t_1},E_\eta)}\big] \\ + \Theta_{t-t_1} \| x(u) \|_{V[t_1,t]}, t \in [t_1, T]. \end{aligned}$$

Again it follows from the property of the function Θ, that there exists $t_2 \in (t_1, T]$ such that $\Theta_{t_2-t_1} \leq 1/2$. Thus following similar steps we arrive at the following inequality

$$\| x(u) \|_{V[t_1,t_2]} \leq 2C_3 \big[1 + \| x(u)(t_1) \|_{L_p(\Omega,\mathcal{F}_{t_1},E_\eta)}\big].$$

Clearly, $\| x(u)(t_1) \|_{L_p(\Omega,\mathcal{F}_{t_1},E_\eta)} \leq \| x(u) \|_{V_{[0,t_1]}}$. Using this in the above inequality we arrive at the following bound

$$\| x(u) \|_{V_{[t_1,t_2]}} \leq 2C_3 + (2C_3)^2 \big[1 + \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}\big], \ \forall u \in \mathcal{U}_{ad}.$$

Continuing this process for a finite number of times, say m, so that $t_m = T$, we find that

$$\begin{aligned} \| x(u) \|_{V_{[t_{m-1},t_m]}} \leq \ & \big\{2C_3 + (2C_3)^2 + \cdots + (2C_3)^{m-1}\big\} \\ & + (2C_3)^m \big[1 + \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}\big], \ \forall u \in \mathcal{U}_{ad}. \end{aligned}$$

From the above estimates it is clear that there exists a finite positive number b dependent on the parameters $\{C_3, T, \| x_0 \|_{L_p(\Omega,\mathcal{F}_0,E_\eta)}\}$ such that

$$\sup\{\| x(u) \|_{V[0,T]}, u \in \mathcal{U}_{ad}\} \leq b.$$

This completes the proof. □

For the proof of existence of optimal controls we need the continuity of the control to solution map $u \to x(u)$. We prove this in the following theorem.

Theorem 5.2. *Consider the control system (5.4) with the admissible controls $\mathcal{U}_{ad}$ and suppose the assumptions of Corollary 5.1 hold. Then the control to solution map $u \longrightarrow x(u)$ is continuous with respect to the metric topology on $\mathcal{U}_{ad}$ and the norm topology on V.*

Proof. Recall the Banach space $V^p_{\alpha,\infty}$ with the norm given by the expression (3.2). For convenience of presentation, as stated above, we denote this by simply V. Let $\{x^n, x^o\} \in V$ denote the solutions of the following integral equation,

$$x(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s, x(s), u_s)ds + \int_0^t S(t-s)\hat{G}(s, x(s), u_s)dW_H(s), t \in I,$$

corresponding to the controls $\{u^n, u^o\} \in \mathcal{U}_{ad}$ respectively. That is, $\{x^n, x^o\}$ satisfy, respectively, the following integral equations,

$$x^n(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s, x^n(s), u^n_s)ds + \int_0^t S(t-s)\hat{G}(s, x^n(s), u^n_s)dW_H(s), t \in I, \tag{5.6}$$

and

$$x^o(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s, x^o(s), u^o_s)ds + \int_0^t S(t-s)\hat{G}(s, x^o(s), u^o_s)dW_H(s), t \in I. \tag{5.7}$$

Subtracting the expression (5.7) from the expression (5.6) we obtain

$$x^n(t) - x^o(t) = \int_0^t S(t-s)[\hat{F}(s, x^n(s), u^n_s) - \hat{F}(s, x^o(s), u^o_s)]ds + \int_0^t S(t-s)[\hat{G}(s, x^n(s), u^n_s) - \hat{G}(s, x^o(s), u^o_s)]dW_H(s), t \in I. \tag{5.8}$$

By suitably rearranging the terms appearing on the righthand side of the above identity one can easily verify that

$$x^n(t) - x^o(t) = Z_1(t) + Z_2(t) + e_{1,n}(t) + e_{2,n}(t) \ \forall\ t \in I, \tag{5.9}$$

where, for all $t \in I$, the processes $\{Z_1(t), Z_2(t), e_{1,n}(t), e_{2,n}(t)\}$ are given by

$$Z_1(t) = \int_0^t S(t-s)\big(\hat{F}(s, x^n(s), u^n_s) - \hat{F}(s, x^o(s), u^n_s)\big)ds, \tag{5.10}$$

$$Z_2(t) \equiv \int_0^t S(t-s)\big(\hat{G}(s, x^n(s), u^n_s) - \hat{G}(s, x^o(s), u^n_s)\big)dW_H(s), \tag{5.11}$$

and

$$e_{1,n}(t) \equiv \int_0^t S(t-s)\big(\hat{F}(s,x^o(s),u_s^n) - \hat{F}(s,x^o(s),u_s^o)\big)ds, \tag{5.12}$$

$$e_{2,n}(t) \equiv \int_0^t S(t-s)\big(\hat{G}(s,x^o(s),u_s^n) - \hat{G}(s,x^o(s),u_s^o)\big)dW_H(s). \tag{5.13}$$

We show that $\| x^n - x^o \|_V \longrightarrow 0$ as $n \to \infty$. Consider Z_1, and let $\delta \equiv \min\{1/2 - \alpha, 1-\eta-\theta_1\}$ where $1-(\eta+\theta_1) > (1/\tau)-(1/2)$ and type $\tau \in [1,2)$. Then using the assumption **(B2)** and the estimate for the deterministic convolution (See Lemma 4.2) we obtain,

$$\| Z_1 \|_{V([0,t]\times\Omega,E_\eta)} \leq (CL_1 t^\delta) \| x^n - x^o \|_{V([0,t]\times\Omega,E_\eta)}, t \in I, \tag{5.14}$$

for a constant C which may depend on $M \equiv \sup\{\| S(t) \|_{\mathcal{L}(E)}, t \in I\}$. Under the assumption **(B2)**, the estimate (5.14) holds uniformly with respect to the set $\mathcal{U}_{ad}$. Next, consider Z_2 and define $\beta \equiv \min\{\varepsilon, 1/2 - \eta - \theta_2\}, \varepsilon \in (0,1)$. Using the assumption **(B3)** and the estimate for the stochastic convolution (see Lemma 4.3), we obtain the following estimate for all $t \in I$,

$$\| Z_2(\cdot) \|_{V([0,t]\times\Omega,E_\eta)} \leq (CL_2 t^\beta) \| x^n(\cdot) - x^o(\cdot) \|_{V([0,t]\times\Omega,E_\eta)}, \tag{5.15}$$

for the same constant C as above. We carry out similar computations and estimates for the processes $\{e_{1,n}, e_{2,n}\}$. Considering $e_{1,n}$ and using the assumption **(B2)** and following similar steps one can verify that

$$\begin{aligned} \| e_{1,n}(\cdot) \|_{V([0,t]\times\Omega,E_\eta)} &\leq C_1 t^\delta \| \int_U F(\cdot,x^o(\cdot),\xi)[u_\cdot^n d(\xi) - u_\cdot^o(d\xi)] \|_{L_p(\Omega,B_0([0,t],E_{-\theta_1}))} \\ &= C_1 t^\delta \| \hat{F}(\cdot,x^o(\cdot),u^n) - \hat{F}(\cdot,x^o(\cdot),u_\cdot^o \|_{L_p(\Omega,B_0([0,t],E_{-\theta_1}))} \\ &= C_1 t^\delta \| \hat{F}(\cdot,x^o(\cdot),u_\cdot^n - u_\cdot^o \|_{L_p(\Omega,B_0([0,t],E_{-\theta_1}))}, \end{aligned} \tag{5.16}$$

where, for simplicity of notation, we have again used the abbreviation $\hat{\Psi}(\nu) \equiv \int_U \Psi(\xi)\nu(d\xi)$. We use this abbreviation throughout the rest of the paper without further notice. Define the measure $\mu_{t,\alpha}$ on $\mathcal{B}(0,t)$ as follows:

$$\mu_{t,\alpha}(J) \equiv \int_0^t (t-s)^{-2\alpha}\chi_J(s)ds \text{ for any} J \in \mathcal{B}(0,t),$$

where χ_J denotes the indicator function of the set $J \in \mathcal{B}(0,t)$. Considering $e_{2,n}$ and using the assumption **(B3)** and following similar steps one can verify that

$$\begin{aligned} &\| e_{2,n}(\cdot) \|_{V([0,t]\times\Omega,E_\eta)} \\ &\leq C_2 t^\beta \| \hat{G}(\cdot,x^o(\cdot),u_\cdot^n - u_\cdot^o) \|_{L_p(\Omega,\gamma(L_2([0,t],\mu_{t,\alpha},H),E_{-\theta_2}))} . \end{aligned} \tag{5.17}$$

Defining

$$C_5(t) \equiv C\big(L_1 t^\delta + L_2 t^\beta\big) \text{ and } C_6(t) \equiv (C_1 t^\delta + C_2 t^\beta), t \in I,$$

and combining the above estimates, and using triangle inequality applied to the expression (5.9), one can verify that

$$\| x^n(\cdot) - x^o(\cdot) \|_{V([0,t]\times\Omega,E_\eta)} \leq C_5(t) \| x^n(\cdot) - x^o(\cdot) \|_{V([0,t]\times\Omega,E_\eta)} + +C_6(t)\Big\{\| \hat{F}(\cdot, x^o(\cdot), u^n - u^o) \|_{L_p(\Omega,B_0([0,t],E_{-\theta_1}))} + \| \hat{G}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \|_{L_p(\Omega,\gamma(L_2([0,t],\mu_{t,\alpha},H),E_{-\theta_2}))}\Big\}, \ \forall\, t \in I. \tag{5.18}$$

Thus it follows from the definition of $C_5 \equiv C_5(t)$ that we can choose a $t_1 \in I$, sufficiently small, so that $C_5(t_1) \leq 1/2$. So for this choice, it follows from the inequality (5.18) that

$$\| x^n(\cdot) - x^o(\cdot) \|_{V([0,t_1]\times\Omega,E_\eta)} \leq 2C_6(t_1)\Big\{\| \hat{F}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \|_{L_p(\Omega,B_0([0,t_1],E_{-\theta_1}))} + \| \hat{G}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \|_{L_p(\Omega,\gamma(L_2([0,t_1],\mu_{t_1,\alpha},H),E_{-\theta_2}))}\Big\}. \tag{5.19}$$

Recall the expressions (5.5) and note that the controls are measure valued processes and hence $\hat{F}$ and $\hat{G}$, which are integrals with respect to these measure valued processes, are linear in these variables. Since $u^n \xrightarrow{d} u^o$ (in the metric topology d), it follows from the property of the metric topology and the continuity of the map F in its third argument, that

$$\hat{F}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \longrightarrow 0 \text{ in } \nu \text{ measure}$$

and therefore there exists a subsequence, relabeled as the original sequence, such that $\hat{F}(s, x^o(s), u^n_s - u^o_s) \longrightarrow 0$ ν-a.e on $I \times \Omega$. Further, it follows from **(B2)** that $\hat{F}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \in L^a_p(\Omega, B_0(I, E_{-\theta_1}))$ independently of $n \in N$, and hence it follows from Lebesgue dominated convergence theorem that for every $t_1 \in I$,

$$\lim_{n\to\infty} \| \hat{F}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \|_{L_p(\Omega,B_0([0,t_1],E_{-\theta_1}))} = 0. \tag{5.20}$$

Using the assumption **(B3)** and following similar argument one can verify that

$$\lim_{n\to\infty} \| \hat{G}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \|_{L_p(\Omega,\gamma(L_2([0,t_1],\mu_{t_1,\alpha},H),E_{-\theta_2}))} = 0 \tag{5.21}$$

for every $t_1 \in I$. By virtue of (5.20) and (5.21) it follows from the expression (5.19) that

$$\lim_{n\to\infty} \| x^n(\cdot) - x^o(\cdot) \|_{V([0,t_1]\times\Omega,E_\eta)} = 0. \tag{5.22}$$

To continue this process beyond time t_1, let us note that for any $t \in [t_1, T]$, it follows from the semigroup property that

$$x^n(t) - x^o(t) = S(t-t_1)[x^n(t_1) - x^o(t_1)] + +\Big\{\int_{t_1}^t S(t-s)[\hat{F}(s, x^n(s), u^n_s) - \hat{F}(s, x^o(s), u^o_s)]ds + \int_{t_1}^t S(t-s)[\hat{G}(s, x^n(s), u^n_s) - \hat{G}(s, x^o(s), u^o_s)]dW(s)\Big\}, t \geq t_1.$$

Note that the second term on the righthand side of the above expression is similar to the one given by the expression (5.8). Now taking $t_2 \in (t_1, T]$, so that $C_5(t_2 - t_1) \leq (1/2)$, and considering t_1 as the starting time, and using the uniform bound of the semigroup on the interval I, we obtain an expression very similar to the expression (5.19) as presented below

$$\parallel x^n(\cdot) - x^o(\cdot) \parallel_{V([t_1,t_2]\times\Omega,E_\eta)} \leq 2M \parallel x^n(t_1) - x^o(t_1) \parallel_{E_\eta} + +2C_6(t_2 - t_1)\Big\{ \parallel \hat{F}(\cdot, x^o(\cdot), u^n - u^o) \parallel_{L_p(\Omega,B_0([t_1,t_2],E_{-\theta_1}))} + \parallel \hat{G}(\cdot, x^o(\cdot), u^n_\cdot - u^o_\cdot) \parallel_{L_p(\Omega,\gamma(L_2([t_1,t_2],\mu_{t_2,\alpha},H),E_{-\theta_2}))} \Big\}. \tag{5.23}$$

By virtue of the norm topology of V given by the expression (3.2), the elements of V are continuous almost surely. Hence it follows from (5.22) that

$$x^n(t_1) \xrightarrow{s} x^o(t_1) \text{ in } E_\eta,\ P - a.s.$$

Thus, again using similar argument as seen above, we obtain

$$\lim_{n\to\infty} \parallel x^n(\cdot) - x^o(\cdot) \parallel_{V([t_1,t_2]\times\Omega,E_\eta)} = 0.$$

Since T is finite, continuing this process step by step for a finite number of times, we conclude that

$$\lim_{n\to\infty} \parallel x^n(\cdot) - x^o(\cdot) \parallel_{V((0,T]\times\Omega,E_\eta)} = 0.$$

Thus we have proved that the control to solution map, $u \longrightarrow x(u)$, from $\mathcal{U}_{ad}$ to $V \equiv V([0,T] \times \Omega, E_\eta)$ is continuous with respect to the metric topology on $\mathcal{U}_{ad}$ and the norm topology on V. This completes the proof. □

Now we are prepared to consider the question of existence of optimal controls.

Theorem 5.3. *Consider the control system (5.4) with the cost functional J given by (5.2). Suppose $\Phi : E_\eta \longrightarrow R$ is lower semi-continuous, and $\ell : I \times E_\eta \times U \longrightarrow R \equiv [0, \infty]$ is Borel measurable in all the variables, and lower semi-continuous in the second and continuous in the third argument, and there exist $g \in L_1^+(I)$ and constants $a, b, c \geq 0$ such that*

$$|\ell(t, x, \xi)| \leq g(t) + a \parallel x \parallel^p_{E_\eta}, \infty > p > 2,\ \forall\ x \in E_\eta, \tag{5.24}$$

$$|\Phi(x)| \leq b + c \parallel x \parallel^p_\eta,\ \forall\ x \in E_\eta. \tag{5.25}$$

Then there exists a control $u^o \in \mathcal{U}_{ad}$ minimizing the cost functional J.

Proof. Let $u^n \in \mathcal{U}_{ad}$ be any sequence. Since the set $\mathcal{U}_{ad}$ is compact in the metric topology d, there exists a subsequence, relabeled as the original sequence, that converges in the metric topology to some $u^o \in \mathcal{U}_{ad}$. It follows from Theorem 5.2 that the sequence of corresponding mild solutions $\{x^n\}$ of the system (5.4) converges in the norm topology of V to an element $x^o \in V$ which is the mild solution of equation (5.4) corresponding to the control u^o. Also, it follows from the definition of the norm topology of V (see equation (3.2)) that $x^n(\cdot) \longrightarrow x^o(\cdot)$ in the sup norm topology of

$C(I, E_\eta)$ P-a.s as $u^n \xrightarrow{d} u^o$ in $\mathcal{U}_{ad}$. Thus it follows from lower semi-continuity of $\hat{\ell}$ in the second argument and continuity in its third argument that

$$\hat{\ell}(t, x^o(t), u^o_t) \leq \underline{\lim}\, \hat{\ell}(t, x^n(t), u^n_t) \text{ for } \nu \ a.e \text{ on } I \times \Omega. \tag{5.26}$$

It follows from Corollary 5.1 that the solution set $\mathcal{S}$ is a bounded subset of $L^a_p(I \times \Omega, E_\eta)$ and thus by virtue of the inequality (5.24) we conclude that $\hat{\ell}$ is bounded from bellow (and above) by an integrable process. Hence it follows from generalized Fatou's Lemma that

$$\mathbf{E} \int_I \hat{\ell}(t, x^o(t), u^o_t)\ dt \leq \underline{\lim}\, \mathbf{E} \int_I \hat{\ell}(t, x^n(t), u^n_t)\ dt. \tag{5.27}$$

Considering the terminal cost, since Φ is also lower semicontinuous on E_η and bounded from below by an integrable random variable, again it follows from generalized Fatou's Lemma that

$$\mathbf{E}\Phi(x^o(T)) \leq \underline{\lim}\, \mathbf{E}\Phi(x^n(T)). \tag{5.28}$$

It is well known that the sum of a finite number of lower semi-continuous functionals is lower semi-continuous. Combing the above results we conclude that the functional, $u \longrightarrow J(u)$, given by (5.2), is lower semi-continuous on $\mathcal{U}_{ad}$ in the metric topology d. The set $\mathcal{U}_{ad}$ is compact (in the metric topology d) and thus J attains its minimum on it proving the existence of an optimal control. This completes the proof. □

6. Induced measures and their optimal control

In this section we consider some control problems related to the measure valued functions induced by the solution set $\mathcal{S} \subset V$ as introduced in Corollary 5.1. Let $\mathcal{M}_0(E_\eta)$ denote the space of regular Borel probability measures defined on the Borel algebra $\mathcal{B}$ of subsets of the space E_η. Let $x \in \mathcal{S}$ and define, for each $t \in I$, the measure $\mu^x_t(D) \equiv P\{x(t) \in D\}$ for any set $D \subset E_\eta$, $D \in \mathcal{B}$. Then, with slight abuse of notation, we can introduce the reachable set of measures at time $t \in I$ by

$$\mathcal{R}(t) \equiv \{\mu^x_t(\cdot), x \in \mathcal{S}\} = \{\mu^{x(u)}_t(\cdot) \equiv \mu^u_t(\cdot), u \in \mathcal{U}_{ad}\} \subset \mathcal{M}_0(E_\eta).$$

For convenience of presentation, we use the notation $\mathcal{L}(z)$ to denote the (probability) law of the random element z. In the following theorem we prove an importan property of the reachable set.

Theorem 6.1. *Suppose the assumptions of Corollary 5.1 and Theorem 5.2 hold. Then for each $t \in I$, the reachable set $\mathcal{R}(t)$ is a weakly compact subset of the space of regular Borel probability measures $\mathcal{M}_0(E_\eta)$.*

Proof. Let $\mu^n \in \mathcal{R}(t)$. Then there exists a sequence $x^n \in \mathcal{S}$ such that for $t \in I$, $\mu^n = \mathcal{L}(x^n(t))$, the (probability) law of $x^n(t)$. Corresponding to the sequence $\{x^n\}$ there exists a sequence of controls $\{u^n\}$ such that $x^n = x(u^n)$. Since the set $\mathcal{U}_{ad}$ is compact in the metric topology d, there exists a subsequence of the sequence $\{u^n\}$, relabeled as $\{u^n\}$, and a $u^o \in \mathcal{U}_{ad}$, such that $u^n \xrightarrow{d} u^o$. Let $\{x^n, x^o\}$ denote the mild solutions of the evolution equation (5.4) corresponding to the controls $\{u^n, u^o\} \in \mathcal{U}_{ad}$ respectively. It follows from Theorem 5.2 that along a subsequence, if necessary, $x^n \xrightarrow{s} x^o$ in V, that is, x^n converges to x^o in the norm topology of the Banach space V. Thus $x^n \xrightarrow{s} x^o$ in the space $L_p^a(\Omega, C(I, E_\eta))$ also. Hence for each $t \in I$, $x^n(t) \xrightarrow{s} x^o(t)$ in $L_p(\Omega, E_\eta)$. Let $BC(E_\eta)$ denote the space of bounded continuous real valued functions defined on the Banach space E_η endowed with the sup-norm topology. Then, for any $\varphi \in BC(E_\eta)$, $\varphi(x^n(t)) \longrightarrow \varphi(x^o(t))$ in measure and this is equivalent to convergence of μ_t^n to μ_t^o in distribution. Since convergence in distribution is equivalent to weak convergence, we conclude that μ_t^n converges weakly to μ_t^o where $\mu_t^o = \mathcal{L}(x^o(t))$ and hence $\mu_t^o \in \mathcal{R}(t)$. This proves the weak compactness of the set $\mathcal{R}(t)$ as a subset of $\mathcal{M}_0(E_\eta)$ □

Remark 6.2. It follows from Corollary 5.1 that whenever the initial state $x_0 \in L_p(\Omega, \mathcal{F}_0, P)$ for any $p \in (2, \infty)$, the (mild) solutions of equation (5.4) have p-th moments for the same p, and therefore the reachable set $\mathcal{R}(t) \subset \mathcal{M}_p(E_\eta)$ where $\mathcal{M}_p(E_\eta)$ denotes the space of regular Borel probability measures on E_η having p-th moment.

Remark 6.3. Let $B_0(E_\eta)(\supset BC(E_\eta))$ denote the Banach space of bounded Borel measurable real valued functions defined on E_η equipped with the standard supnorm topology. Suppressing the time variable let $\mathcal{A}_u$, for each $u \in \mathcal{U}_{ad}$, denote the forward Kolmogorov operator corresponding to the controlled stochastic differential equation 5.4, and let $\mathcal{L}(x_0) = \vartheta \in \mathcal{M}_0(E_\eta)$ denote the probability law of the initial state. Then, under the given assumptions, one can verify that the measure valued function, $t \longrightarrow \mu_t^u$, $t \in I$, is weakly differentiable and satisfies in the weak sense the following forward Kolmogorov equation,

$$\frac{d}{dt}\mu_t(\varphi) = \mu_t(\mathcal{A}_u\varphi), \mu_0(\varphi) = \vartheta(\varphi), t \in I,$$

for every $\varphi \in BC(E_\eta)$ for which $\mathcal{A}_u(\varphi) \in B_0(E_\eta)$, where

$$\mu_t(\varphi) \equiv \int_{E_\eta} \varphi(x)\mu_t(dx).$$

For some control problems, it is required to find a control policy $u \in \mathcal{U}_{ad}$ so that the corresponding measure valued function $\mu_t^u, t \in I$, induced by the solution process $x(u) \in V$, is close to a desired measure valued function $m \in M_\infty^w(I, \mathcal{M}_0(E_\eta))$. Since the topology of weak convergence of measures on separable metric spaces is equivalent to convergence in the Prokhorov metric, this problem can be formulated as follows. Let ρ denote the Prokhorov metric on $\mathcal{M}_0(E_\eta)$ and define the cost functional as

$$J_1(u) \equiv \int_I \rho(\mu_t^u, m_t)\lambda(dt) \tag{6.1}$$

where λ is a positive finite Borel measure on I. The problem is to find a control policy $u^o \in \mathcal{U}_{ad}$ that minimizes the functional $J_1(u)$.

Corollary 6.4. *Consider the control system (5.4) with the admissible controls $\mathcal{U}_{ad}$ and cost functional J_1 given by the expression (6.1). Suppose E is separable and the assumptions of Theorem 6.1 hold. Then there exists a control $u^o \in \mathcal{U}_{ad}$ that minimizes the functional $J_1(u)$ on $\mathcal{U}_{ad}$.*

Proof. Let $u^n \in \mathcal{U}_{ad}$ be a minimizing sequence, that is,

$$\lim_{n\to\infty} J_1(u^n) = \inf\{J_1(u).,u \in \mathcal{U}_{ad}\}.$$

By compactness of $\mathcal{U}_{ad}$ in its metric topology d, there exists a control $u^o \in \mathcal{U}_{ad}$ such that, along a subsequence if necessary, $u^n \xrightarrow{d} u^o \in \mathcal{U}_{ad}$. By Theorem 5.2, $x^n \equiv x(u^n) \xrightarrow{s} x(u^o) \equiv x^o$ in V and hence also in $C(I, E_\eta)$ P-a.s. Then it follows from Theorem 6.1 that, along a subsequence if necessary, for each $t \in I$,

$$\mu_t^{u^n} \equiv \mu_t^n \xrightarrow{w} \mu_t^o \equiv \mu_t^{u^o}(\text{weakly}).$$

Recall that, for separable metric spaces, weak convergence (of measures) is equivalent to convergence in the Lévy-Prokhorov metric ρ. Hence, for almost all $t \in I$, $\rho(\mu_t^n, \mu_t^o) \to 0$. Thus considering the integrand of the functional (6.1), we conclude that $\rho(\mu_t^n, m_t) \longrightarrow \rho(\mu_t^o, m_t)$ for almost all $t \in I$. Since $\rho(\mu_t^n, m_t) \leq 2$, for all $t \in I$ and $n \in N$, it follows from Lebesgue bounded convergence theorem that

$$\lim_{n\to\infty} J_1(u^n) = J_1(u^o).$$

Thus J_1 is continuous on $\mathcal{U}_{ad}$ in its metric topology d. Since $\mathcal{U}_{ad}$ is compact in this topology J_1 attains its infimum on $\mathcal{U}_{ad}$ at the point u^o proving the existence of an optimal control. □

Let $C(t), t \in I$, be a nonempty closed bounded set valued function with values $C(t) \subset E_\eta$ for all $t \in I$, and possibly continuous in the Hausdorff metric. The problem is to find a control that maximizes the functional

$$J_2(u) \equiv \int_0^T \mu_t^u(C(t))dt. \tag{6.2}$$

The physical significance of this problem is tracking or following a moving set valued target, $C(t), t \in I$, in the state space E_η as closely as possible under the given control constraints.

Corollary 6.5. *Consider the control system (5.4) with the admissible controls $\mathcal{U}_{ad}$ and the objective functional J_2 given by (6.2). Suppose the assumptions of Theorem 6.1 hold and E is separable. Then there exists a control $u^o \in \mathcal{U}_{ad}$ at which J_2 attains its maximum.*

Proof. Using the Prokhorov metric ρ on $\mathcal{M}_0(E_\eta)$, we introduce the metric $D(\mu, \nu)$ on the space of measure valued functions $B_0(I, \mathcal{M}_0(E_\eta))$ as follows:

$$D(\mu, \nu) \equiv \sup\{\rho(\mu_t, \nu_t), t \in I\}.$$

Let $\{u^n\} \subset \mathcal{U}_{ad}$ be a maximizing sequence for the functional J_2 given by (6.2) in the sense that

$$\lim_{n\to\infty} J_2(u^n) = \sup\{J_2(u), u \in \mathcal{U}_{ad}\} \equiv M_o. \tag{6.3}$$

Since $\mathcal{U}_{ad}$ is compact in the metric topology d, there exists a control $u^o \in \mathcal{U}_{ad}$ such that, along a subsequence if necessary, relabeled as the original sequence, $u^n \xrightarrow{d} u^o$. Let x^n and x^o denote the (mild) solutions of equation (5.4) corresponding to the controls $\{u^n, u^o\} \in \mathcal{U}_{ad}$ respectively. Then it follows from Theorem 5.2 that $x^n \xrightarrow{s} x^o$ in V and hence in $L_p^a(\Omega, C(I, E_\eta))$. Let $\{\mu^n, \mu^o\}$ denote the corresponding probability measure valued functions. Then it follows from Theorem 6.1 that for each $t \in I$, $\mu_t^n \xrightarrow{w} \mu_t^o$ and hence (by the equivalence of Prokhorov metric topology and the topology of weak convergence) $D(\mu^n, \mu^o) \to 0$ as $n \to \infty$. In other words the weak convergence is uniform in $t \in I$. Since, for each $t \in I$, $C(t)$ is a closed subset of E_η, it follows from Parthasarathy [15, Theorem 6.1, p.40] that $\overline{\lim}\, \mu_t^n(C(t)) \leq \mu_t^o(C(t)), t \in I$. Hence we conclude that

$$\overline{\lim} \int_0^T \mu_t^n(C(t))\ dt \leq \int_0^T \overline{\lim}\, \mu_t^n(C(t))\ dt \leq \int_0^T \mu_t^o(C(t))\ dt.$$

Thus $\overline{\lim}\, J_2(u^n) \leq J_2(u^o)$ and it follows from (6.3) that

$$M_o = \lim_{n\to\infty} J_2(u^n) \leq \overline{\lim}\, J_2(u^n) \leq J_2(u^o). \tag{6.4}$$

Since $u^o \in \mathcal{U}_{ad}$, again it follows from (6.3) that $J_2(u^o) \leq M_o$. Combining the above inequalities we conclude that $J_2(u^o) = M_o$ and hence u^o is the optimal control. □

Another interesting problem is time optimal control related to stability and residence time. Let $\mu_0 = \nu \in \mathcal{M}_0(E_\eta)$ denote the probability measure corresponding to the initial state x_0. Suppose it is supported on a closed bounded subset $C_0 \subset E_\eta$. Let

$$d_\eta(x,y) \equiv \parallel x - y \parallel_{E_\eta}, x, y \in E_\eta,$$

denote the metric induced by the norm topology of E_η. For $r > 0$, let $C_r \equiv \{x \in E_\eta : d_\eta(x, C_0) \leq r\}$ denote the closed r-neighbourhood of C_0. The objective is to find a control that maximizes the residence time of the state in the set C_r containing the set C_0. This can be formulated as follows: for sufficiently small $\delta \in (0,1)$ (as small as desired), find a control that maximizes the escape time $\tau_\delta(u) \equiv \inf\{t \geq 0 : \mu_t^u(C_r) < 1 - \delta\}$. So we define the pay-off functional as

$$J_3(u) \equiv \tau_\delta(u) \equiv \inf\{t \geq 0 : \mu_t^u(C_r) < 1 - \delta\}. \tag{6.5}$$

If the underlying set is empty, we set $\tau_\delta(u) = T$. Clearly, such a control has a stabilizing effect on the system. We prove the existence of such a control.

Corollary 6.6. *Consider the control system (5.4) with the admissible controls $\mathcal{U}_{ad}$ and the objective functional $J_3(u) \equiv \tau_\delta(u)$ given by (6.5). Suppose the assumptions of Theorem 6.1 hold. Then there exists a control $u^o \in \mathcal{U}_{ad}$ at which J_3 attains its maximum.*

Proof. We prove that the functional J_3 is upper semicontinuous on $\mathcal{U}_{ad}$ with respect to the metric topology d. Let $\{u^n, u^o\} \in \mathcal{U}_{ad}$ and suppose $u^n \xrightarrow{d} u^o$. Let $\{\mu^n, \mu^o\}$ denote the measure valued functions induced by the (mild) solutions $\{x(u^n), x(u^o)\}$ of the system (5.4) corresponding to the controls $\{u^n, u^o\}$ respectively. It follows from Theorem 5.2 that, along a subsequence if necessary, $x(u^n) \xrightarrow{s} x(u^o)$ in V. By Theorem 6.1, for each $t \in I$, the reachable set $\mathcal{R}(t) \subset \mathcal{M}_0(E_\eta)$ is weakly compact. Thus for each $t \in I$, $\mu_t^n \xrightarrow{w} \mu_t^o$, and therefore, since C_r is a closed subset of E_η, again it follows from [15, Theorem 6.1, p.40] that

$$\overline{\lim}\, \mu_t^n(C_r) \leq \mu_t^o(C_r),\ \forall\ t \in I.$$

Clearly, given any $\varepsilon \in (0, \delta)$, we can find a $n_0 \in N$ such that

$$\mu_t^{n_0+k}(C_r) - \varepsilon \le \overline{\lim}\, \mu_t^n(C_r), \ \forall\ k \ge 1, t \in I.$$

In other words, the following inequality

$$\mu_t^{n_0+k}(C_r) \le \overline{\lim}\, \mu_t^n(C_r) + \varepsilon, \ \forall\ k \ge 1, t \in I$$

holds. From these inequalities one can readily deduce the following inclusions

$$\begin{aligned} &\big\{t \ge 0 : \mu_t^{n_0+k}(C_r) < 1 - \delta\big\} \\ &\supset \big\{t \ge 0 : \overline{\lim}\, \mu_t^n(C_r) + \varepsilon < 1 - \delta\big\} \\ &\supset \big\{t \ge 0 : \mu_t^o(C_r) + \varepsilon < 1 - \delta\big\}, \ \forall\ k \ge 1, \varepsilon \in (0, \delta). \end{aligned}$$

Clearly these inclusions imply the following inequalities,

$$\begin{aligned} &\inf\big\{t \ge 0 : \mu_t^{n_0+k}(C_r) < 1 - \delta\big\} \\ &\quad \le \inf\big\{t \ge 0 : \overline{\lim}\, \mu_t^n(C_r) + \varepsilon < 1 - \delta\big\} \\ &\quad \le \inf\big\{t \ge 0 : \mu_t^o(C_r) < 1 - (\delta - \varepsilon)\big\}, \ \forall\ k \ge 1, \varepsilon \in (0, \delta) \end{aligned}$$

and hence, by definition of $\tau_\delta(u)$ given by (6.5), we have

$$\tau_\delta(u^{n_0+k}) \le \tau_{\delta-\varepsilon}(u^o), \ \forall\ k \ge 1, \varepsilon \in (0, \delta).$$

Hence $\overline{\lim}\, \tau_\delta(u^n) = \overline{\lim}^k\, \tau_\delta(u^{n_o+k}) \le \tau_{\delta-\varepsilon}(u^o)$. Since this holds for any $\varepsilon \in (0, \delta)$, letting $\varepsilon \downarrow 0$ in this expression we conclude that $\overline{\lim}^n\, \tau_\delta(u^n) \le \tau_\delta(u^o)$. By definition of the objective functional (6.5) this is equivalent to $\overline{\lim}^n\, J_3(u^n) \le J_3(u^o)$ and hence J_3 is upper semicontinuos on $\mathcal{U}_{ad}$ in its metric topology d. Since $\mathcal{U}_{ad}$ is compact in this topology, J_3 attains its maximum on $\mathcal{U}_{ad}$. This proves the existence of a time optimal control. □

Another closely related problem is to find a control so as to avoid being in the neighbourhood of a forbidden zone for long period of time. This can be formulated as follows. Let D be a nonempty open subset of E_η, denoting the open neighbourhood of a forbidden zone $D_o \subsetneq D$. For a fixed $\delta \in (0, 1)$ define the set

$$\mathcal{S}_\delta(u) \equiv \{t \in I : \mu_t^u(D) > \delta\}$$

where $\mu_t^u \equiv \mathcal{L}(x(u)(t))$. Let λ denote the Lebesgue measure on the real line and define the objective functional as the Lebesgue measure of the set $\mathcal{S}_\delta(u)$:

$$J_4(u) \equiv \lambda(\mathcal{S}_\delta(u)). \tag{6.6}$$

The problem is to find a control policy that minimizes this functional.

Corollary 6.7. *Consider the control system (5.4) with the admissible controls $\mathcal{U}_{ad}$ and the objective functional J_4 given by (6.6). Suppose the assumptions of Theorem 6.1 hold. Then there exists a control $u^o \in \mathcal{U}_{ad}$ at which J_4 attains its minimum.*

Proof. The proof is largely similar to that of Corollary 6.5. If the underlying set is empty there is nothing to prove. So let us assume the contrary. Let D be an open set in E_η containing the forbidden zone D_o in its interior and suppose that for each $t \in I$, $\mu_t^n \xrightarrow{w} \mu_t^o$. Then it follows from Parthasarathy [15, Theorem 6.1, p.40] that $\mu_t^o(D) \le \underline{\lim}\, \mu_t^n(D)$. Using this result and following similar arguments as seen in the proof of Corollary 6.5, one can prove that $u \longrightarrow J_4(u)$ is lower semicontinuous on $\mathcal{U}_{ad}$. Since $\mathcal{U}_{ad}$ is compact, J_4 attains its minimum at some point $u^o \in \mathcal{U}_{ad}$ proving existence of an optimal control. □

Some Open Problems:

(1) It follows from the works of Van Neerven, Veraar and Weis [17, 19], and Brzeźniak [7, 11] that stochastic integration is well defined with respect to gamma radonifying operators from Hilbert spaces to UMD Banach spaces. Since UMD spaces are reflexive (but the converse is false), this is certainly a limitation. Thus, currently known theory does not cover all reflexive Banach spaces. Clearly stochastic integration theory extended to general Banach spaces will broaden the scope of applications. The author believes that this may be possible in some weak sense based on Pettis and Dunford integrals. Some results are known for linear stochastic differential equations on Banach spaces [5].

(2) The most interesting topic in the area of stochastic control theory is Optimal Feedback control [**?**, 8, 18]. This is a very challenging problem. Based on Bellman's principle of optimality one can formulate the optimal control problem leading to the HJB (Hamilton-Jacobi-Bellman) equation giving the value function from which one can determine the optimal feedback control law and the optimum cost. This is the method used in [8] for stochastic differential equations on Hilbert spaces. However, for practical application this requires solving nonlinear partial differential equations on infinite dimensional spaces, a formidable task indeed. It is theoretically interesting to investigate if the control problem considered here on UMD spaces can be formulated in the form of HJB equation under the same assumptions used here.

(3) We have not developed necessary conditions of optimality for the problem considered here. We believe that this is an interesting problem and can be treated following similar steps as seen in Ahmed [2].

References

[1] N. U. Ahmed, *Semigroup Theory With Applications to Systems and Control*, Pitman research Notes in Mathematics Series , vol. 246, Longman scientific and technical, U.K (copublished in U.S by John Wiley & Sons. Inc, New York), 1991.

[2] N. U. Ahmed, *Optimal relaxed controls for infinite dimensional stochastic systems of Zakai type*, SIAM J.Control. Optim. **34** (1996), 1592–1615.

[3] N. U. Ahmed, *Relaxed solutions for stochastic evolution equations on Hilbert spaces with polynomial nonlinearities*, Publicationes Mathematicae Debrecen (PMD) **54** (1999), 75–101.

[4] N. U. Ahmed, *Optimal control of infinite dimensional stochastic systems via generalized solutions of HJB equations*, Discussiones Mathematicae Differential Inclusions, Control and Optimization **21** (2001), 97–126.

[5] N. U. Ahmed, *Stochastic differential equations on Banach spaces and their optimal feedback control*, DICO **32** (2012), 87–109.

[6] N. U. Ahmed, *Stochastic evolution equations on Hilbert spaces with partially observed relaxed controls and their necessary conditions of optimality*, Discussiones Mathematicae, Differential Inclusions, Control and Optimization **34** (2014), 105–129.

[7] N. U. Ahmed, *Necessary conditions of optimality for a class of stochastic differential equations on UMD Banach spaces*, Discussiones Mathematicae Differential Inclusions, Control and Optimization **38** (2018),15–38.

[8] N. U. Ahmed, *Partially observed stochastic evolution equations on Banach spaces and their optimal Lipschitz feedback control law*, SIAM Journal on Control and Optimization **57** (2019), 3101–3117.

[9] Z. Brzeźniak, *On stochastic convolutions in Banach spaces and applications*, Stochastics and Stochastics Rep. **61** (1997), 245–295.

[10] Z. Brzeźniak and H. Long, *A Note on $\gamma-$Radonifying and Summing Operators, Stochastic Analysis, Banach Center Publications*, vol.105, Institute of Mathematics, Polish Academy of Sciences, Warsaw, 2015.

[11] Z. Brzeźniak and R. Serrano, *Optimal relaxed control of dissipative stochastic partial differential equations in Banach spaces*, SIAM J. Control Optim. **51** (2013), 2664–2703.

[12] D. L. Burkholder, *A geometrical characterization of Banach spaces in which Martingale differences are unconditional*, The Annals of Probability **9** (1981), 997–1011.

[13] D. L. Burkholder, *Martingales and Singular Integrals in Banach Spaces*, Handbook of the geometry of Banach Spaces, vol. 1, North Holland, Amsterdam, 2001, pp. 233–269.

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

[15] N. Dunford and J. T. Schwartz, *Linear Operators Part 1, General Theory*, Inter Science Publishers, Inc., New York, 1958.

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

[17] A. I. Tulcea and C. I. Tulcea, *Topics in The Theory of Lifting*, Springer-Verlag, Berli, Heidelberg, New York, 1969.

[18] L. Tzafriri, *On the type and Co-type of Banach spaces*, Israel Journal of Mathematics **32** (1979), 32–38.

[19] J. M. A. M. Van Neerven, M. C. Veraar and L. Weis, *Stochastic Evolution Equations in UMD Banach Spaces*, arXiv:0804.0932v1[math.FA] 6 April 2008.

[20] J. M. A. M. Van Neerven, M. C. Veraar and L. Weis, *Stochastic Integration in UMD Banach spaces*, The Annals of Probability **35** (2007), 1438–1478.

Manuscript received January 16 2021
revised March 20 2021

N. U. Ahmed
University of Ottawa, Canada
E-mail address: nahmed@uottawa.ca