Sufficient conditions for the eventual strong Feller property for degenerate stochastic evolutions

Joris Bierkens *

November 23, 2010

Abstract

The strong Feller property is an important quality of Markov semigroups which helps for example in establishing uniqueness of invariant measure. Unfortunately degenerate stochastic evolutions, such as stochastic delay equations, do not possess this property. However the *eventual* strong Feller property is sufficient in establishing uniqueness of invariant probability measure.

In this paper we provide operator theoretic conditions under which a stochastic evolution equation with additive noise posseses the eventual strong Feller property. The results are used to establish uniqueness of invariant probability measure for stochastic delay equations and stochastic partial differential equations with delay, with an application in neural networks.

Keywords: Stochastic evolution equations, Malliavin calculus, stochastic delay equations

AMS Subject Classification (2010): 34K50, 60H07, 47D07

A stochastic delay differential equation with additive noise can be modeled (see [7], [2]) as a stochastic Cauchy problem in some Hilbert space H of the form

$$\begin{cases} dX(t) = [AX + F(X)]\ dt + G\ dW(t) & t \geq 0, \text{a.s.} \\ X(0) = x \quad \text{a.s.} \end{cases} \tag{1}$$

where A is the generator of the delay semigroup, F a sufficiently smooth function (e.g. Lipschitz), and G a linear operator mapping the Wiener process W into H. It is well known that under the mentioned assumptions, existence and uniqueness of solutions is guaranteed.

So far however, the ergodic behaviour of these systems was less well understood. An important notion in this respect is that of *invariant probability measure*, i.e. a positive finite Borel measure μ on H with $\mu(H) = 1$ such that if the initial condition x has law μ, then the solution $X(t; x)$ has law μ for all $t \geq 0$. Recently the existence of an invariant (probability) measure was established for a sufficiently broad class of stochastic Cauchy problems to include the case of finite dimensional stochastic delay differential equations [2].

Apart from the existence of an invariant probability measure, its uniqueness is an important issue. When an invariant probability measure is unique, the *ergodic property* 'time average equals spatial average' holds:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi(X(t; x))\ dt = \int_H \varphi\ d\mu, \quad \varphi \in B_b(H),$$

where $B_b(H)$ are the bounded Borel measurable functions on H.

Just as the problem of existence of invariant measures, also the problem of uniqueness of the invariant probability measure of stochastic delay differential equations were open for some time.

*Joris Bierkens, Centrum Wiskunde & Informatica Amsterdam, j.bierkens@cwi.nl

A partial solution to this problem was proposed by using the dissipativity properties of the delay semigroup (see [8] and [1]).

In [19] general conditions for the uniqueness of the invariant probability measure are established for the nondegenerate noise case. However, the noise that perturbs delay equations can influence only the present of the process and not the past and is therefore essentially degenerate, so these results do not apply here. In [4] results are obtained for degenerate noise, but these do not include the case of delay equations.

Often uniqueness of invariant probability measure is proved using Doob's theorem (see e.g. [8], Theorem 5.2.1). This requires irreducibility and the strong Feller property of solutions. In [8] the eventual strong Feller property for systems of the form (1) was conjectured. This property states that $P(t)\varphi$ is continuous and bounded for any $\varphi \in B_b$ and is important in establishing the uniqueness of the invariant probability measure. It is not immediate that the strong Feller property holds, because usually some kind of non-degeneracy assumption on the noise is required. However, in the case of stochastic delay differential equations, the noise is intrinsically degenerate because it can only work on the 'present' of the process, while the state space also contains the 'past' of the stochastic evolution.

In [19] uniqueness of an invariant probability measure was established for nondegenerate diffusions in Hilbert spaces, and in [4] for degenerate diffusions. However, in the latter, only the immediate strong Feller property was established which is too strong for our purposes: the delay semigroup can never be immediately strong Feller. In [14] and [16] an overview is given of results on uniqueness of invariant probability measures and on strong Feller diffusions, respectively. In [9] and [10] the immediate strong Feller property and irreducibility are proven for (possibly degenerate) diffusions, by applying Malliavin calculus. Their result does not apply to stochastic delay equations since these can only be eventually strong Feller. Uniqueness of invariant probability measure in Banach spaces is discussed in [20].

In this paper we establish conditions that are sufficient to establish uniqueness of the invariant probability measure for degenerate stochastic Cauchy problems of the form (1). We combine methods from the now classical semigroup approach initiated by Da Prato and Zabczyk [7], and from Malliavin calculus, inspired by succesful applications in e.g. [12], to obtain the eventualy strong Feller property, uniqueness of invariant probability measure and eventual irreducibility. In [15] the eventual strong Feller property for delay equations with additive noise is established by a probabilistic method. However, we think the operator theoretic conditions established by our method are easier to verify in practice. Very recently the uniqueness of invariant probability measure for general stochastic delay equations with multiplicative noise was established in [11].

Our main result is stated in Section 1. The proof is split into two parts, discussed in Sections 2 (strong Feller property) and 3 (irreducibility). The result is applied to stochastic delay differential equations and a stochastic partial differential equation with delay (with application to the field of neural networks) in Section 4.

1 Main result

We will study differential equation (1) under the following assumptions. See [7] for necessary definitions.

Hypothesis 1.1. *(i) H is a Hilbert space;*

(ii) $(S(t))_{t\geq 0}$ is a strongly continuous semigroup acting on H with generator A;

(iii) W is a cylindrical Wiener process with RKHS $\mathcal{H}$;

(iv) $F : H \to V \subset \mathrm{im}\ (G)$, with V a closed subspace of H, is globally Lipschitz;

(v) $G \in L(\mathcal{H}; H)$ *(the linear space of bounded linear operators from* $\mathcal{H}$ *into* H*) and a mapping* $G^{-1} \in L(V, \mathcal{H})$ *exists such that* $GG^{-1} = I$ *on* V*;*

In many cases it is convenient to take $V = \mathrm{im}\ (G)$. However if F maps into a strict subspace of $\mathrm{im}\ (G)$ the condition of pseudoinvertibility of G can be relaxed by letting $V \subsetneq \mathrm{im}\ (G)$.

We will assume throughout this section that for any $x \in H$, there exists a unique mild solution $(X(t;x))_{t\geq 0}$ of (1). Sufficient conditions for this to hold are that $G \in L_{\mathrm{HS}}(\mathcal{H}; H)$ (see [7]).

We need the notions of null controllability and approximate controllability, which we will define now.

Let $\mathcal{H}, H$ be Hilbert spaces. Consider the controlled Cauchy problem

$$\begin{cases} \dot{x}(t) = Ax(t) + f(x(t)) + Gu(t), & t \geq 0 \\ x(0) = x_0 \end{cases} \tag{2}$$

with A the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H, $f : H \to H$ globally Lipschitz, $G \in L(\mathcal{H}; H)$ and where $u \in L^2([0,T]; \mathcal{H})$ is called the *control*.

Definition 1.2 (null controllability)**.** *The system (2) is* null controllable *in time* $t > 0$ *if for any* $x_0 \in H$ *there exists a control* $u \in L^2([0,t]; \mathcal{H})$ *such that* $x(t; x_0) = 0$*.*

The pair (A, G) *is called* null controllable *in time* $t > 0$ *if (2) with* $f \equiv 0$ *is null controllable in time* $t > 0$*.*

It is well known (see [7], Section B.3) that null controllability of (A, G) in time $t > 0$ is equivalent to

$$\mathrm{im}\ \ S(t) \subset \mathrm{im}\ \ Q_t^{1/2}, \tag{3}$$

where the *controllability Gramian* $Q_t \in L(H)$ is defined by

$$Q_t x := \int_0^t S(s)GG^* S^*(s)x \, ds, \quad x \in H. \tag{4}$$

Furthermore, since the linear operator $Q_t^{-1/2} S(t) : H \to H$ is closed and defined everywhere on H, by the closed graph theorem it is bounded.

Definition 1.3 (approximate controllability)**.** *The system (2) is said to be* approximately controllable *in time* $t > 0$ *if, for arbitrary* $x_0, z \in H$ *and* $\varepsilon > 0$*, there exists a control* $u \in L^2([0,t]; \mathcal{H})$ *such that* $|x(t; x_0, u) - z| < \varepsilon$*.*

The pair (A, G) *is said to be* approximately controllable *in time* $t > 0$ *if (2) with* $f \equiv 0$ *is approximately controllable in time* $t > 0$*.*

Any property of an evolution that holds for some fixed time $t > 0$, but not at time $t = 0$, is said to hold *eventually*. Note that eventual approximate controllability is not implied by eventual null controllability, as illustrated by the following example.

Example 1.4. Consider the nilpotent shift semigroup $(S(t))_{t\geq 0}$ on $H = L^2([0,1])$, given by

$$S(t)f(\sigma) = \begin{cases} f(t+\sigma), & 0 \leq \sigma \leq 1-t, \\ 0 & 1-t < \sigma \leq 1. \end{cases}$$

Let A denote the infinitesemal generator of $(S(t))_{t\geq 0}$. Then the (deterministic) evolution, given by

$$\dot{x}(t) = Ax(t),$$

considered as a control system with $G = 0$, is null controllable in time 1. Indeed, $S(1)f = 0$ for all $f \in H$. However, it is clearly not approximately controllable: 0 is the only reachable state.

Note that in this case, there is a unique and strongly mixing invariant probability measure, namely the Dirac measure on 0. In general, for linear equations (i.e. of the form (1) with $F = 0$), null controllability is sufficient to ensure regularity and hence uniqueness of invariant probability measure, which then is strongly mixing. See [8], Theorem 7.2.1. $\diamond$

Before we can state the main result of this paper, we recall some more notions. The *(Markov) transition semigroup* associated to a Markov process $(X(t;x))$ is defined as the family of operators $(P(t))_{t\geq 0}$ acting on $B_b(H)$, defined by

$$P(t)\varphi(x) = \mathbb{E}[\varphi(X(t;x))], \quad \varphi \in B_b(H), x \in H, t \geq 0.$$

The transition semigroup $(P(t))_{t\geq 0}$ is called *strong Feller* at $t > 0$ if $P(t)\varphi \in C_b(H)$ for all $\varphi \in B_b(H)$, and *irreducible* at $t > 0$ if $P(t)\mathbb{1}_\Gamma(x) > 0$ for any open, non-empty $\Gamma \subset H$, $x \in H$. A positive Borel measure μ on H is said to be *invariant* for $(P(t))_{t\geq 0}$ if

$$\int_H P(t)\varphi \, d\mu = \int_H \varphi \, d\mu \quad \text{for all } \varphi \in B_b(H), t \geq 0.$$

If furthermore $\mu(H) = 1$ then μ is called *invariant probability measure.*

An invariant measure μ is called *strongly mixing* if

$$\lim_{t\to\infty} P(t)\varphi(x) = \int_H \varphi \, d\mu, \quad \text{for all } \varphi \in B_b(H), x \in H.$$

The following theorem is the main result of this paper.

Theorem 1.5. *Suppose the assumptions of Hypothesis 1.1 hold and the pair (A, G) is eventually null controllable. Then the transition semigroup corresponding to (1) is eventually strong Feller and there exists at most one invariant probability measure for (1). Furthermore, if the pair (A, G) is eventually approximately controllable, then the transition semigroup is eventually irreducible. In case (A, G) is both null controllable and approximately controllable then the unique invariant probability measure is strongly mixing, in case it exists.*

PROOF: In case of eventual null controllability, by Theorem 2.5 the transition semigroup associated to (1) is eventually strong Feller. Furthermore, by Theorem 2.6, there exists at most one invariant probability measure for (1). By Corollary 3.2 it is eventually irreducible in case of approximate controllability of (A, G). Hence by Khas'minskii's theorem ([8], Theorem 4.1.1), in the combined case, the transition semigroup of (1) is regular at time $2T$. Then the strongly mixing property follows from Doob's theorem ([8], Theorem 4.2.1). $\square$

2 Null controllability and the strong Feller property

In this section, we will see that Hypothesis 1.1 and the null controllability of (A, G) are together sufficient to prove the strong Feller property and uniqueness of invariant probability measure of (1).

2.1 Linearized flow

We will make use of the notioin of the Fréchet differential. Suppose H, K are Hilbert spaces and $F : H \to K$ is Fréchet differentiable. We then denote the Fréchet differential of F by $dF : H \to L(H; K)$. Let V denote a closed subspace of K containing im (F) and note that $dF : H \to L(H; V)$.

We are interested in dependence of the solution $(X(t;x))_{t\geq 0}$ of (1) on the initial condition x. Therefore we define, for arbitrary directions $\xi \in H$, the derivative processes $J_{0,t}\xi := d_x X(t;x)\xi$, where $d_x X(t;x)$ is the Fréchet differential of $X(t;x)$ with respect to x. Assume for now that $F : H \to \text{im }(G)$ is continuously Fréchet differentiable, with $||dF||_\infty < \infty$.

By [8], Theorem 5.4.1, $J_{0,t}\xi$ is a mild solution to

$$\begin{cases} \frac{d}{dt} J_{0,t}\xi = AJ_{0,t}\xi + dF(X(t;x))J_{0,t}\xi & \text{a.s., } t \geq 0 \\ J_{0,s} = \xi \quad \text{a.s.,} \end{cases}$$

and there exists a constant $C > 0$ independent of ξ such that

$$\sup_{t\in[0,T]} \mathbb{E}|J_{0,t}\xi|^2 \leq C|\xi|^2. \tag{5}$$

More generally define for $s \geq 0$ and $t \geq s$ the linear, stochastic operators $J_{s,t}$ as the pathwise solutions of

$$J_{s,t}\xi = S(t-s)\xi + \int_s^t S(t-r)dF(X(r;x))J_{s,r}\xi \, dr \tag{6}$$

for $\xi \in H$.

We set out to express the dependence of $X(T;x)$ on the initial condition x in terms the dependence of $X(T;x)$ on the noise process W. For this we need the notion of Malliavin derivative.

2.2 Malliavin calculus

Our exposition of the Malliavin calculus is based on [3], Chapter 5.

Let W be a cylindrical Brownian motion with reproducing kernel Hilbert space $\mathcal{H}$ and let K be a separable Hilbert space.

We first define the Malliavian derivative of smooth variables. A random variable $X \in L^2(\Omega; K)$ is called *smooth* if X has the form

$$X = \psi(W(\Phi_1), \ldots, W(\Phi_n)),$$

with $\psi : \mathbb{R}^n \to K$ infinitely often differentiable, $\Phi_1, \ldots, \Phi_n \in L^2([0,T]; \mathcal{H})$ and

$$W(\Phi) := \int_0^T \langle \Phi(t), \, dW(t) \rangle, \quad \Phi \in L^2([0,T]; \mathcal{H}).$$

We denote all smooth K-valued random variables by $\mathcal{S}(K)$.

For $X \in \mathcal{S}(K)$ we define the *Malliavin derivative* DX of X as the $K \otimes L^2([0,T]; \mathcal{H})$-valued random variable

$$DX = \sum_{i=1}^n \frac{\partial}{\partial x_i} \psi(W(\Phi_1), \ldots, W(\Phi_n)) \otimes \Phi_i.$$

Note that we may identify the range of D with $L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$, so we can (and will) interpret DX as a (possibly non-adapted) stochastic process $(D_t X)_{t\in[0,T]}$ with values in $L_{\mathrm{HS}}(\mathcal{H}; K)$.

The mapping $X \mapsto DX : \mathcal{S}(K) \to L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$ is closable ([3], Proposition 5.1), and we call its closure $D : \mathbb{H}(K) \to L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$ the *Malliavin derivative*, where the domain $\mathbb{H}(K)$ of D is a linear subspace of $L^2(\Omega; K)$.

For $v \in L^2(\Omega \times [0,T]; \mathcal{H})$ we define the *Malliavin derivative in the direction* v pointwise almost everywhere on Ω as the K-valued square integrable random variable

$$D^v X := \int_0^T D_t X \circ v(t) \, dt.$$

Remark 2.1. Intuitively, DX is the stochastic process which, when integrated with respect to W over $[0,T]$, results in the random variable X. As such, DX represents the dependence of X on the noise process W, and $D^v X$ indicates the infinitesemal change in X if we perturb W infinitesemally in the direction of v. Note that this interpretation makes sense only if $(D_t X)_{t\in[0,T]}$ is adapted; see however the Skorohod integral below. $\diamond$

We will use the following version of the *chain rule for the Malliavin derivative* (which holds more generally, see [3], Proposition 5.2): Suppose K_1 and K_2 are separable Hilbert spaces and assume $\varphi : K_1 \to K_2$ is Fréchet differentiable with uniformly bounded Fréchet derivative $d\varphi$. Then for $X \in \mathbb{H}(K_1)$, we have $\varphi(X) \in \mathbb{H}(K_2)$ and

$$D\varphi(X) = (d\varphi(X))(DX). \tag{7}$$

The adjoint operator $\delta : \mathfrak{D}(\delta, K) \to L^2(\Omega; K)$ of D is defined by the duality

$$\mathbb{E}\langle DX, \Phi\rangle_{L^2([0,T];L_{\mathrm{HS}}(\mathcal{H};K))} = \mathbb{E}\langle X, \delta\Phi\rangle_K,$$

for $X \in \mathbb{H}(K)$ and $\Phi \in \mathfrak{D}(\delta, K) \subset L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H};K))$ and is called the *Skorohod integral*, also denoted by

$$\delta\Phi = \int_0^T \Phi(t)\ \delta W(t).$$

If Φ is a predictable process in $L_{\mathrm{HS}}(\mathcal{H};K)$ such that

$$\mathbb{E}\int_0^T ||\Phi(t)||^2_{L_{\mathrm{HS}}(\mathcal{H};K)}\ dt < \infty,$$

then $\Phi \in \mathfrak{D}(\delta, K)$ and the Skorohod integral and the Itô integral coincide ([3], Theorem 5.1):

$$\int_0^T \Phi(t)\ \delta W(t) = \int_0^T \Phi(t)\ dW(t).$$

We therefore have, for predictable Φ, the integration by parts formula

$$\mathbb{E}\langle DX, \Phi\rangle_{K\otimes L^2([0,T];\mathcal{H})} = \mathbb{E}\left\langle X, \int_0^T \Phi(t)\ dW(t)\right\rangle_K$$

and in particular

$$\mathbb{E}\left[D^v X\right] = \mathbb{E}\left[X \int_0^T v(s)\ dW(s)\right]$$

where $X \in \mathbb{H}(K)$ and $v \in L^2(\Omega \times [0,T]; \mathcal{H})$ predictable.

We conclude our summary of Malliavin calculus with a commutation rule for the Malliavin derivative and the Skorohod integral (a straightforward extension to the infinite-dimensional case of [17], Proposition 1.3.2):

$$D^v \delta\Phi = \int_0^T \langle v(t), \Phi(t)\rangle_K\ dt + \delta(D^v\Phi), \tag{8}$$

which holds for (deterministic) $v \in L^2([0,T];\mathcal{H})$ and $\Phi \in \mathfrak{D}(\delta, K)$ such that $D^v\Phi \in \mathfrak{D}(\delta)$.

Lemma 2.2. *Suppose* $(X(t;x))_{t\geq 0}$ *is the solution of*

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]\ dt + G\ dW(t) & t \in [0,T] \\ X(0) = x, \end{cases}$$

with A *the generator of a strongly continuous semigroup* $(S(t))_{t\geq 0}$*,* $F : H \to H$ *Fréchet differentiable, and* $G \in L(\mathcal{H}; H)$*.*

Then, for $v \in L^2(\Omega \times [0,T]; \mathcal{H})$,

$$D^v X(t;x) = \int_0^t J_{s,t} G v(s)\ ds, \quad \textit{almost surely}, t \in [0,T]$$

where $J_{s,t}$ *is defined by (6).*

PROOF: For $t > 0$ we have that $X(t;x) \in \mathbb{H}(H)$ by [3], Lemma 5.3. We have

$$D^v X(t;x) = D^v S(t)x + D^v \int_0^t S(t-s)F(X(s;x))\ ds + D^v \int_0^t S(t-s)G\ dW(s).$$

The first term disappears since $S(t)x$ is deterministic. By the chain rule of Malliavin calculus,

$$D \int_0^t S(t-s)F(X(s;x))\ ds = \int_0^t S(t-s)dF(X(s;x))DX(s;x)\ ds,$$

and hence

$$\begin{aligned}
&D^v \int_0^t S(t-s)F(X(s;x))\ ds \\
&= \int_0^T \int_0^t S(t-s)dF(X(s;x))D_r X(s;x)\ ds \circ v(r)\ dr \\
&= \int_0^T \int_0^t S(t-s)dF(X(s;x))D_r X(s;x) \circ v(r)\ dsdr \\
&= \int_0^t S(t-s)dF(X(s;x))D^v X(s;x)\ ds.
\end{aligned}$$

Finally by (8) for v deterministic

$$\begin{aligned}
D^v \int_0^t S(t-s)G\ dW(s) &= D^v \delta(S(t-s)G\mathbb{1}_{s\le t})_{s\in[0,T]} \\
&= \int_0^T S(t-s)G\mathbb{1}_{s\le t} v(s)\ ds = \int_0^t S(t-s)Gv(s)\ ds.
\end{aligned}$$

Hence for simple functions $v = \sum_{i=1}^n v_i \mathbb{1}_{E_i}$, with $E_i \in \mathcal{F}$ and $v_i \in L^2([0,T]; \mathcal{H})$,

$$D^v \int_0^t S(t-s)G\ dW(s) = \int_0^t S(t-s)Gv(s)\ ds, \quad \text{almost surely.} \tag{9}$$

We obtain (9) for general $v \in L^2(\Omega \times [0,T]; \mathcal{H})$ by approximating v by simple functions.

Hence

$$\begin{aligned}
&D^v X(t;x) \\
&= \int_0^t S(t-s)dF(X(s;x))D^v X(s;x)\ ds + \int_0^t S(t-s)Gv(s)\ ds \quad \text{a.s.},
\end{aligned}$$

or equivalently, using the definition of $(J_{s,t})_{t\ge s}$ in (6),

$$D^v X(t;x) = \int_0^t J_{s,t} G v(s)\ ds, \quad t \in [0,T].$$

□

Lemma 2.3. *Assume Hypothesis 1.1, (A, G) is null controllable in time $T > 0$ and that $F : H \to$ im (G) is Fréchet differentiable, with $||dF||_\infty < \infty$.*

Then for all $\xi \in H$ there exists a stochastic process $v = v_\xi \in L^2(\Omega \times [0, T]; \mathcal{H})$, adapted to $(\mathcal{F}_t)_{t\in[0,T]}$, such that

$$J_{0,T}\xi = \int_0^T J_{s,T} G v(s)\ ds,$$

and there exists a constant M, independent of ξ and the initial value x of $X(t; x)$, such that

$$\mathbb{E} \int_0^T |v(s)|^2\ ds \leq M|\xi|^2.$$

Proof: By [7], (B.26), there exists $u_1 \in L^2([0, T]; \mathcal{H})$ such that

$$S(T)(-\xi) + \int_0^T S(T - s)Gu_1(s)\ ds = 0,$$

and

$$\int_0^T |u_1(s)|^2\ ds = |Q_T^{-1/2} S(T)\xi|^2 \leq ||Q_T^{-1/2} S(T)||^2|\xi|^2, \tag{10}$$

where Q_T is the controllability Gramian defined by (4).

Let $(\zeta(t))_{t\in[0,T]}$ be the solution of the pathwise inhomogeneous Cauchy problem

$$\begin{cases} \dot\zeta(t) = A\zeta(t) + dF(X(t; x))J_{0,t}\xi + Gu_1(t), & t \geq 0, \\ \zeta(0) = 0. \end{cases} \tag{11}$$

Then

$$\begin{aligned} \zeta(T) &= \int_0^T S(T - s)dF(X(s, x))J_{0,s}\xi\ ds + \int_0^T S(T - s)Gu_1(s)\ ds \\ &= \int_0^T S(T - s)dF(X(s, x))J_{0,s}\xi\ ds + S(T)\xi = J_{0,T}\xi. \end{aligned}$$

Define

$$u_2(t) := G^{-1}dF(X(t; x))[J_{0,t}\xi - \zeta(t)], \quad t \in [0, T], \text{ a.s.} \tag{12}$$

We see that $\zeta(t)$ also satisfies almost surely the inhomogeneous Cauchy problem

$$\begin{cases} \dot\zeta(t) = A\zeta(t) + dF(X(t; x))\zeta(t) + Gu_1(t) + Gu_2(t), & t \geq 0, \\ \zeta(0) = 0, \end{cases}$$

or, using variation of constants,

$$\zeta(t) = \int_0^t J_{s,t} G v(s)\ ds,$$

where $v : \Omega \times [0, T] \to \mathcal{H}$ is defined by $v(t) := u_1(t) + u_2(t)$, $t \in [0, T]$. From (5), (10), (11) and (12) we see, using the Gronwall inequality, that $\mathbb{E} \int_0^t |v(s)|^2\ ds \leq M|\xi|^2$ for some $M > 0$ independent of ξ and x. □

The following corollary is a direct consequence of Lemma 2.2 and Lemma 2.3.

Corollary 2.4. *Under Hypothesis 1.1, if (A, G) is null controllable in time $T > 0$ and if $F : H \to H$ is Fréchet differentiable with uniformly bounded Fréchet derivative, then, for $\xi \in H$ and $v = v_\xi$ associated to ξ by Lemma 2.3, we have*

$$d_x X(T; x)\xi = J_{0,T}\xi = \int_0^T J_{s,T} G v(s)\ ds = D^v X(T; x). \tag{13}$$

In (other) words: we have expressed the dependence of $X(T;x)$ on its initial condition x in terms of the dependence of $X(T;x)$ on the noise process W.

We can now give a short proof, as in [12], of the following theorem.

Theorem 2.5. *Under the conditions of Hypothesis 1.1 and if (A, G) is null controllable at time $T > 0$, then the transition semigroup associated to* (1) *is strong Feller at time T.*

PROOF: Suppose for the moment $\varphi \in C_b^1(H)$ and $F \in C_b^1(H;H)$. Let $(P(t))_{t\geq 0}$ denote the transition semigroup associated to (1). We have, using (13), the chain rule and integration by parts for the Malliavin derivative, that

$$\begin{aligned} dP(T)\varphi(x)\xi &= \mathbb{E}\left[d\varphi(X(T;x))J_{0,T}\xi\right] = \mathbb{E}\left[d\varphi(X(T;x)D^v X(T;x)\right] \\ &= \mathbb{E}\left[D^v\varphi(X(T;x))\right] = \mathbb{E}\left[\varphi(X(T;x))\int_0^T \langle v(s),\ dW(s)\rangle\right] \\ &\leq ||\varphi||_\infty \left(\mathbb{E}\int_0^T |v(s)|^2\ ds\right)^{1/2}, \end{aligned}$$

where v is as described in Lemma 2.3, so that $\mathbb{E}\int_0^T |v(s)|^2\ ds \leq M^2|\xi|^2$ for some $M > 0$, independent of x and ξ. Hence

$$||dP(T)\varphi(x)||_{H^*} \leq M||\varphi||_\infty, \quad \text{for all } \varphi \in C_b^1(H), x \in H.$$

It follows that

$$|P(T)\varphi(x) - P(T)\varphi(y)| \leq M||\varphi||_\infty|x-y|_H, \quad \varphi \in C_b^1(H), x, y \in H.$$

We can extend this estimate to $\varphi \in B_b(H)$ and Lipschitz F by approximating φ by a sequence $(\varphi_n) \subset C_b^1(H)$, and F by a sequence $(F_n) \subset C_b^1(H;H)$ with $||dF_n||_\infty < [F]_{\mathrm{Lip}}$ (see the proof of [8], Theorem 7.1.1). ☐

We can now establish uniqueness of invariant probability measure under these conditions (we thank the reviewer for pointing this out).

Theorem 2.6. *Under the conditions of Hypothesis 1.1 and if (A, G) is eventually null controllable, then there exists at most one invariant probability measure for equation* (1).

PROOF: It is sufficient to prove that $0 \in \mathrm{supp}(\mu)$, i.e. $\mu(U) > 0$ for all open environments U of 0 in H, for all invariant measures μ (see [6], Proposition 7.8, or [12], Corollary 3.17).

Let $(X(t;x))_{t\geq 0}$ denote the solution of (1) with initial condition $x \in H$. By Girsanov's theorem (see e.g. [7], Theorem 10.18) which may be applied because of Hypothesis 1.1 (v), the law of $(X(t;x))_{t\geq 0}$ is equivalent to the law of the solution to

$$d\widetilde{X}(t) = A\widetilde{X}(t)\ dt + G\ dW(t), \quad \widetilde{X}(0) = x. \tag{14}$$

Let B_ε denote the sphere of radius $\varepsilon > 0$ in H. Let us fix the time of null controllability of (A, G) at $T > 0$, as usual. Now

$$\mathbb{P}(\widetilde{X}(T;x) \in B_\varepsilon) = \mathbb{P}\left(S(T)x + \int_0^T S(T-s)GdW(s) \in B_\varepsilon\right).$$

Let $Z = \int_0^T S(T-s)GdW(s)$ and note that $Z \sim N(0, Q_T)$, with Q_T defined by (4). Let $M = \overline{\mathrm{im}(Q_T^{1/2})}$, and let P denote the orthogonal projection on M. Then $PZ = Z$, and by (3), $PS(T)x = S(T)x$. Hence

$$\mathbb{P}(\widetilde{X}(T) \in B_\varepsilon) = \mathbb{P}\left(P(S(T)x + Z) \in B_\varepsilon\right) = \mathbb{P}\left(P(S(T)x + Z) \in B_\varepsilon \cap M\right) > 0,$$

where the last inequality follows since the probability measure of Z is full on M ([6], Proposition 1.25). Hence

$$\mathbb{P}(X(T;x) \in B_\varepsilon) > 0.$$

In particular, if we now consider the evolution of $(X(t))_{t\geq 0}$ governed by

$$dX(t) = [AX(t) + F(X(t))]\ dt + G\ dW(t), \quad X(0) \sim \mu,$$

where μ is an invariant measure for (1), then note that $X(T) \sim \mu$ by the definition of invariant measure, and therefore, for any $\varepsilon > 0$,

$$\mu(B_\varepsilon) = \mathbb{P}(X(T) \in B_\varepsilon) = \int_H P(X(T;x) \in B_\varepsilon)\ \mu(dx) > 0.$$

□

3 Approximate controllability and irreducibility

Hypothesis 1.1 and the approximate controllability of (A, G) will be seen to be sufficient to prove the irreducibility of (1).

Proposition 3.1. *Suppose the assumptions of Hypothesis 1.1 hold and (A, G) is approximately controllable in time $T > 0$. Then the system (2), with $f = F$, is approximately controllable in time T.*

Proof: Let $x, z \in H$ and $\varepsilon > 0$. Since (A, G) is approximately controllable, there exists a control $u_1 \in L^2([0,T]; \mathcal{H})$ such that $\eta \in L^2([0,T]; H)$ defined by

$$\eta(t) := S(t)x + \int_0^t S(t-s)Gu_1(s)\ ds \tag{15}$$

satisfies $|\eta(T) - z| < \varepsilon$.

For $0 \leq t \leq T$, choose

$$u_2(t) := -G^{-1}F(\eta(t)).$$

Then, for

$$u(t) := u_1(t) + u_2(t),$$

the solution of (2) is given by

$$\begin{aligned} y(t) &= S(t)x + \int_0^t S(t-s)F(y(s))\ ds + \int_0^t S(t-s)G(u_1(s) + u_2(s))\ ds \\ &= S(t)x + \int_0^t S(t-s)[F(y(s)) - F(\eta(s))]\ ds + \int_0^t S(t-s)Gu_1(s)\ ds. \end{aligned}$$

Suppose $(S(t))$ satisfies $||S(t)|| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$ and some $M, \omega \geq 0$. Let

$$\zeta(t) := y(t) - \eta(t).$$

Then

$$\begin{aligned} |\mathrm{e}^{-\omega t}\zeta(t)| &= \left|\mathrm{e}^{-\omega t} \int_0^t S(t-s)[F(y(s)) - F(\eta(s))]\ ds\right| \\ &\leq \int_0^t M[F]_{\mathrm{Lip}}|\mathrm{e}^{-\omega s}\zeta(s)|\ ds, \quad t \geq 0, \end{aligned}$$

so that by Gronwall $\zeta \equiv 0$.

Hence $|y(T) - z| < \varepsilon$. □

Corollary 3.2. *Suppose the assumptions of Hypothesis 1.1 hold and (A, G) is approximately controllable. Then the transition semigroup corresponding to the stochastic system (1) is irreducible in time T.*

PROOF: This follows immediately from the approximate controllability proven in Proposition 3.1 and Theorem 7.4.1 in [8], which states that approximate controllability implies irreducibility. $\square$

4 Examples

In this section we will give some illustrative examples to which the results of this paper apply.

4.1 Stochastic delay differential equations

Consider, similar to [8], Section 10.2, a stochastic delay equation in $\mathbb{R}^d$ of the form

$$\begin{cases} dY(t) = \left(BY(t) + \sum_{i=1}^N B_i Y(t+\theta_i) + \varphi(Y(t), Y_t)\right)\, dt + \psi\, dW(t), \qquad t \geq 0 \\ Y(0) = x, \\ Y(\theta) = f(\theta), \quad \theta \in [-r, 0], \end{cases} \tag{16}$$

where $N \in \mathbb{N}$, $B, B_1, \ldots, B_N \in L(\mathbb{R}^d)$, $-r = \theta_1 < \theta_2 < \ldots < \theta_N < 0$, $\psi \in L(\mathbb{R}^m; \mathbb{R}^d)$, $(W(t))_{t\geq 0}$ an m-dimensional standard Brownian motion and the initial condition $x \in \mathbb{R}^d$. The segment process $(Y_t)_{t\geq 0}$ is defined by $Y_t(\theta) := Y(t+\theta)$ for $t \geq 0$, $-r \leq \theta \leq 0$, and $f \in L^2([-r, 0]; \mathbb{R}^d)$ is the initial segment. The nonlinear perturbation $\varphi : \mathbb{R}^d \times L^2([-r, 0]; \mathbb{R}^d) \to \mathbb{R}^d$ is assumed to be Lipschitz.

As explained in [7], [2], we can cast this into the infinite dimensional framework (1) by choosing as Hilbert space $H := \mathbb{R}^d \times L^2([-r, 0]; \mathbb{R}^d)$, and letting the closed, densely defined operator A, described by

$$\mathfrak{D}(A) = \left\{ \begin{pmatrix} c \\ y \end{pmatrix} \in \mathbb{R}^d \times W^{1,2}([-r, 0]; \mathbb{R}^d) : y(0) = c \right\},$$
$$A \begin{pmatrix} c \\ y \end{pmatrix} := \begin{pmatrix} Bc + \sum_{i=1}^N B_i y(\theta_i) \\ \dot{y} \end{pmatrix}$$

denote the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$. (See e.g. [5], Section 2.4.)

As nonlinear perturbation $F : H \to H$ and noise factor $G \in L(\mathbb{R}^m; H)$ we take, respectively,

$$F \begin{pmatrix} c \\ y \end{pmatrix} := \begin{pmatrix} \varphi(c, y) \\ 0 \end{pmatrix}, \quad \text{and } G := \begin{pmatrix} \psi \\ 0 \end{pmatrix}.$$

For convenience, we recall the following result ([8], Theorem 10.2.3). See also [18] for the null controllability and [5], Theorem 4.2.10 for the approximate controllability.

Theorem 4.1. *The pair (A, G) is null controllable for all $t > r$ if and only if*

$$\operatorname{rank} \left[\lambda I - \sum_{i=1}^N e^{\lambda \theta_i} B_i, \psi \right] = d \tag{17}$$

for all $\lambda \in \mathbb{C}$.

The pair (A, G) is approximately controllable for all $t > r$ if and only if

$$\operatorname{rank} \left[\lambda I - B - \sum_{i=1}^N e^{\lambda \theta_i} B_i, \psi \right] = d, \quad \operatorname{rank}[B_1, \psi] = d \tag{18}$$

for all $\lambda \in \mathbb{C}$.

Remark 4.2. The above theorem is partly based on [18]. In this paper null controllability after some time $t > 0$ is established; however from the proof in this paper it is not clear whether null controllability holds for all $t > r$. This has no significant consequence since, without loss of generality, we may take $r > 0$ large enough so that we indeed have null controllability for all $t > r$.

We can now state the main result of this section.

Theorem 4.3. *Suppose conditions (17) and (18) are satisfied. Let* $\widetilde{V}$ *be a linear subspace of* $\mathbb{R}^d$ *such that* $\overline{\varphi(H)} \subset \widetilde{V}$. *Suppose that a mapping* $\psi^{-1} \in L(\widetilde{V}; \mathbb{R}^m)$ *of* ψ *exists, i.e.* $\psi\psi^{-1}v = v$ *for* $v \in \widetilde{V}$.

Then there exists at most one invariant probability measure for (16) on the state space H*, and if an invariant probability measure exists, it is strongly mixing.*

Proof: Define $G^{-1} \in L(\widetilde{V} \times \{0\}; \mathbb{R}^m)$ by $G^{-1}\begin{pmatrix} v \\ 0 \end{pmatrix} := \psi^{-1}v$. All the conditions of Hypotheses 1.1are satisfied, (A, G) is approximately controllable and null controllable in time $T > 0$, (with $T > r$ and $V = \widetilde{V} \times \{0\} \subset \mathbb{R}^d \times L^2([-r, 0]; \mathbb{R}^d)$), and by Theorem 2.6 we may deduce the uniqueness of an invariant probability measure and the strong mixing property of such a measure. □

Note that the conditions of Theorem 4.3 are not necessarily very restrictive:

Corollary 4.4. *Suppose that* $m \geq d$ *and* $\psi \in L(\mathbb{R}^m; \mathbb{R}^d)$ *is surjective.*

Then there exists at most one invariant probability measure for (16) on the state space H*, and if it exists, it is strongly mixing.*

Proof: Define the pseudoinverse ψ^{-1} by letting $\psi^{-1}v$ denote the element w of minimal norm in $\mathbb{R}^m$ such that $\psi w = v$. Then $\psi^{-1} \in L(\mathbb{R}^d; \mathbb{R}^m)$ is a linear operator. Since $m \geq d$ and ψ is surjective, we find that rank $\psi = d$ and hence (17) and (18) hold. The result follows now from Theorem 4.3, of which all conditions are satisfied (with $\widetilde{V} = \mathbb{R}^d$). □

For convenience we combine our result with a result of of [2] on the existence of invariant probability measures.

Corollary 4.5. *Suppose the solutions of (16) are bounded in probability on the state space* H*, and the conditions of Theorem 4.3 hold. Then there exists a unique, strongly mixing invariant probability measure for (16) on* H.

Proof: The existence of an invariant measure under these conditions is proven in [2]. The uniqueness follows from Theorem 4.3. □

4.2 Stochastic reaction-diffusion recurrent neural networks

In [13] the following stochastic partial differential equation in m dimensions with delay and noise is considered as an example of so-called *recurrent neural networks*.

$$
\begin{aligned}
dy_i(t) = \sum_{k=1}^{m} \frac{\partial}{\partial \xi_k}\left(D_i \frac{\partial y_i}{\partial \xi_k}\right)\, dt + \Bigg[& -c_i h_i(y_i(t,\xi)) + \sum_{j=1}^{n} a_{ij} f_j(y_j(t,\xi)) \\
& + \sum_{j=1}^{n} b_{ij} \int_{-\infty}^{t} \kappa_{ij}(t-s) g_j(y_j(s,\xi))\, ds + J_i \Bigg]\, dt \\
& + \sum_{l=1}^{\infty} \sigma_{il}(y_i(t,\xi))\, dw_{il}(t).
\end{aligned}
$$

We consider the following variant for n neurons in one dimension:

$$
\begin{aligned}
dy_i(t,\xi) = \Delta y_i(t,\xi)\, dt + \Bigg[& -c_i h_i(y_i(t,\xi)) + \sum_{j=1}^{n} a_{ij} f_j(y_j(t,\xi)) \\
& + \sum_{j=1}^{n} b_{ij} \int_{t-1}^{t} \kappa_{ij}(t-s) g_j(y_j(s,\xi))\, ds + J_i \Bigg]\, dt + (\Psi_i\, dW_i(t))(\xi),
\end{aligned} \tag{19}
$$

where

(i) Δ denotes the one-dimensional Laplacian $\frac{d^2}{d\xi^2}$ on $[0,\pi]$;

(ii) c_i, a_{ij}, b_{ij}, J_i are constants for $i = 1,\dots,n$, $j = 1,\dots,n$;

(iii) h_i, f_j and g_j are Lipschitz functions $\mathbb{R} \to \mathbb{R}$ for $i,j = 1,\dots,n$;

(iv) $\kappa_{ij} \in L^2([0,1])$ for $i,j = 1,\dots,n$;

(v) $\Psi_i \in L(\mathcal{H}_i, L^2([0,\pi]))$, where $\mathcal{H}_i$ is the RKHS of the cylindrical Wiener process W_i, $i = 1,\dots,n$.

Let $X = L^2([0,\pi]) \times \ldots \times L^2([0,\pi])$ and let the delay semigroup $(S(t))$ on $\mathcal{E}^2 = X \times L^2([-1,0];X)$ be generated by

$$
A = \begin{bmatrix} B & 0 \\ 0 & \frac{d}{d\sigma} \end{bmatrix} \tag{20}
$$

with

$$
B = \begin{bmatrix} \Delta & 0 & \dots & 0 \\ 0 & \Delta & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Delta \end{bmatrix}.
$$

Note that there is no dependence on the past in the generator of the delay semigroup. This will be in our advantage later on.

Denote typical elements of $\mathcal{E}^2$ by $z = \begin{pmatrix} u_1 & \dots & u_n \\ v_1 & \dots & v_n \end{pmatrix}$ with $u_i \in L^2([0,\pi])$ and $v_i \in L^2([-1,0]; L^2([0,\pi]))$, $i = 1,\dots,n$. Let $\varphi_i : X \times L^2([-1,0];X) \to L^2([0,\pi])$ be given by

$$
\begin{aligned}
&\varphi_i(u_1,\dots,u_n,v_1,\dots,v_n)[\xi] \\
&:= -c_i h_i(u_i(\xi)) + \sum_{j=1}^{n} a_{ij} f_j(u_j(\xi)) + \sum_{j=1}^{n} b_{ij} \int_{-1}^{0} \kappa_{ij}(-s) g_j(v_j(s,\xi))\, ds + J_i,
\end{aligned}
$$

and define $F : X\times L^2([-1,0];X) \to X\times L^2([-1,0];X)$ and $G \in L(\mathcal{H}_1\times\ldots\times\mathcal{H}_n; X\times L^2([-1,0];X))$ by

$$F(z) := \begin{pmatrix} \varphi_1(z) & \ldots & \varphi_n(z) \\ 0 & \ldots & 0 \end{pmatrix}, \quad z \in \mathcal{E}^2, \quad \text{and} \tag{21}$$

$$G\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} \Psi_1 w_1 & \ldots & \Psi_n w_n \\ 0 & \ldots & 0 \end{pmatrix}, \quad (w_1,\ldots,w_n) \in \mathcal{H}_1 \times \ldots \times \mathcal{H}_n. \tag{22}$$

Now (19) can be written in the form (1). By classical theory on stochastic evolutions (see [7]), there exists a unique solution to (1), with A, F and G as given in (20), (21) and (22).

We will now establish a sufficient condition for (A,G) to be eventually null controllable.

Theorem 4.6. *Suppose $\Psi_i \in L(\mathcal{H}_i;H)$ has a bounded inverse for all $i = 1,\ldots,n$. Then (A,G) is null controllable for $t > 1$.*

Proof: Since A and G are 'in diagonal form' it suffices to consider the case where $n = 1$, so let $\Psi \in L(\mathcal{H};X)$ be invertible with $\Psi^{-1} \in L(X;\mathcal{H})$. It is sufficient to establish the null controllability of (Δ,Ψ) on $X = L^2([0,\pi])$. Indeed, if (Δ,Ψ) is null controllable, then for any $t > 1$ and $x \in X \times L^2([-1,0];X)$ we may find a control which steers the first component of x to 0 in time $t-1$. By setting the control equal to zero after time $t-1$ the translation effect of the delay semigroup then ensures that x is steered to 0 in $X \times L^2([-1,0];X)$ in time t.

It remains to establish the null controllability of (Δ,Ψ) on $L^2([0,\pi])$. By [5], condition (4.12), this is equivalent to the the existence of a $\gamma(t) > 0$ for all $t > 0$ such that

$$\int_0^t ||\Psi^* T^*(t-s)z||_{\mathcal{H}}^2 \, ds \geq \gamma(t)||T^*(t)z||_X^2 \tag{23}$$

for all $z \in X$, where $(T(t))_{t\geq 0}$ is the semigroup generated by the Laplacian. Let $(e_n)_{n\in\mathbb{N}\cup\{0\}}$ be the orthonormal base of eigenvectors of the Laplacian with Neumann boundary conditions on $X = L^2([0,\pi])$, so

$$\begin{cases} e_0(\xi) = \frac{1}{\sqrt{\pi}}, \\ e_n(\xi) = \sqrt{\frac{2}{\pi}}\cos(n\xi), \qquad n \in \mathbb{N}, \end{cases}$$

and for $z \in X$ write $z_n := \langle z, e_n\rangle_X$ for $n \in \mathbb{N}\cup\{0\}$. We have, using selfadjointness of $(T(t))_{t\geq 0}$,

$$\begin{aligned} \int_0^t ||\Psi^* T^*(t-s)z||_{\mathcal{H}}^2 \, ds &\geq \frac{1}{||\Psi^{-1}||^2} \int_0^t ||T(t-s)z||_X^2 \, ds \\ &= \frac{1}{||\Psi^{-1}||^2} \int_0^t \left|\left| \sum_{n=0}^\infty T(t-s) z_n e_n \right|\right|_X^2 ds \\ &= \frac{1}{||\Psi^{-1}||^2} \int_0^t \left[z_0^2 + \sum_{n=1}^\infty \mathrm{e}^{-2n^2 s} z_n^2 \right] ds \\ &= \frac{1}{||\Psi^{-1}||^2} \left[t z_0^2 + \sum_{n=1}^\infty \frac{1}{2n^2}(1-\mathrm{e}^{-2n^2 t}) z_n^2 \right], \end{aligned}$$

which should be compared to

$$||T^*(t)z||_X^2 = z_0^2 + \sum_{n=1}^\infty \mathrm{e}^{-2n^2 t} z_n^2.$$

Using the basic inequality

$$\frac{1}{a}\left(1-\mathrm{e}^{-at}\right) \geq t\mathrm{e}^{-at}, \quad a, t > 0,$$

we find that (23) holds for

$$\gamma(t) = \frac{t}{||\Psi^{-1}||^2},$$

which establishes the null controllability of (Δ, Ψ). $\square$

Corollary 4.7. *The transition semigroup corresponding to* (1) *with* A*,* F *and* G *as given in* (20), (21) *and* (22)*, is eventually strong Feller, and there exists at most one invariant probability measure for* (1).

Proof: This is an immediate corollary of the previous theorem, Theorem 2.5 and Theorem 2.6, using the invertibility of Ψ_i, $i = 1, \ldots, n$. $\square$

Acknowledgements

The author would like to thank Onno van Gaans for his invaluable comments. Furthermore Anna Chojnowska-Michalik and the anonymous referee are kindly acknowledged for their suggestions for improvement.

References

[1] J. Bierkens. Dissipativity of the Delay Semigroup. *Submitted*, 2010.

[2] J. Bierkens, O. van Gaans, and S.M. Verduyn Lunel. Existence of an Invariant Measure for Stochastic Evolutions Driven by an Eventually Compact Semigroup. *Journal of Evolution Equations*, 9(4):771–786, 2009.

[3] R.A. Carmona and M.R. Tehranchi. *Interest rate models: an infinite dimensional stochastic analysis perspective*. Springer Finance. Springer, Berlin, 2006.

[4] A. Chojnowska-Michalik and B. Goldys. Existence, uniqueness and invariant measures for stochastic semilinear equations on Hilbert spaces. *Probability Theory and Related Fields*, 102(3):331–356, 1995.

[5] R.F. Curtain and H. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory*. Springer, 1995.

[6] G. Da Prato. *An introduction to infinite-dimensional analysis. Universitext*. Springer, 2006.

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

[8] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite Dimensional Systems*. Cambridge University Press, 1996.

[9] M. Fuhrman. Smoothing properties of nonlinear stochastic equations in Hilbert spaces. *NoDEA Nonlinear Differential Equations Appl.*, 3(4):445–464, 1996.

[10] M. Fuhrman. On a class of stochastic equations in Hilbert spaces: solvability and smoothing properties. *Stochastic Anal. Appl.*, 17(1):43–69, 1999.

[11] M. Hairer, J. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of harris' theorem with applications to stochastic delay equations. *Probability Theory and Related Fields*, pages 1–37, 2009.

[12] M. Hairer and J.C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Annals of Mathematics*, 164:993–1032, 2006.

[13] Y. Lv, W. Lv, and J. Sun. Convergence dynamics of stochastic reaction-diffusion recurrent neural networks in continuously distributed delays. *Nonlinear Anal. Real World Appl.*, 9(4):1590–1606, 2008.

[14] B. Maslowski and J. Seidler. Invariant measures for nonlinear SPDE's: uniqueness and stability. *Arch. Math. (Brno)*, 34(1):153–172, 1998. Equadiff 9 (Brno, 1997).

[15] B. Maslowski and J. Seidler. Probabilistic approach to the strong Feller property. *Probab. Theory Related Fields*, 118(2):187–210, 2000.

[16] B. Maslowski and J. Seidler. Strong Feller infinite-dimensional diffusions. In *Stochastic partial differential equations and applications (Trento, 2002)*, volume 227 of *Lecture Notes in Pure and Appl. Math.*, pages 373–387. Dekker, New York, 2002.

[17] D. Nualart. *The Malliavin Calculus And Related Topics.* Springer, second edition, 2006.

[18] A. W. Olbrot and L. Pandolfi. Null controllability of a class of functional-differential systems. *Internat. J. Control*, 47(1):193–208, 1988.

[19] S. Peszat and J. Zabczyk. Strong Feller property and irreducibility for diffusions on Hilbert spaces. *Ann. Probab.*, 23(1):157–172, 1995.

[20] J. van Neerven. Uniqueness of invariant measures for the stochastic Cauchy problem in Banach spaces. 127:491–517, 2001.