Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics
(IMACM)

Preprint BUW-IMACM 20/29

Bálint Farkas and Martin Friesen and Barbara Rüdiger and Dennis Schroers

On a class of stochastic partial differential equations with multiple invariant measures

August 17, 2020

ON A CLASS OF STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS WITH MULTIPLE INVARIANT MEASURES

BÁLINT FARKAS, MARTIN FRIESEN, BARBARA RÜDIGER*, AND DENNIS SCHROERS

Abstract. In this work we investigate the long-time behavior for Markov processes obtained as the unique mild solution to stochastic partial differential equations in a Hilbert space. We analyze the existence and characterization of invariant measures as well as convergence of transition probabilities. While in the existing literature typically uniqueness of invariant measures is studied, we focus on the case where the uniqueness of invariant measures fails to hold. Namely, introducing a *generalized dissipativity condition* combined with a decomposition of the Hilbert space, we prove the existence of multiple limiting distributions in dependence of the initial state of the process and study the convergence of transition probabilities in the Wasserstein 2-distance. Finally, we apply our results to Lévy driven Ornstein-Uhlenbeck processes, the Heath-Jarrow-Morton-Musiela equation as well as to stochastic partial differential equations with delay.

1. Introduction

Stochastic partial differential equations arise in the modelling of applications in mathematical physics (e.g. Navier-Stokes equations [22, 18, 9, 37] or stochastic non-linear Schrödinger equations [4, 13]), biology (e.g. catalytic branching processes [12, 30]), and finance (e.g. forward prices [24, 38, 16]). While the construction of solutions to the underlying stochastic equations is an important mathematical issue, having applications in mind it is indispensable to also study their specific properties. Among them, an investigation of the long-time behavior of solutions, that is existence and uniqueness of invariant measures and convergence of transition probabilities, are often important and at the same time also challenging mathematical topics. In this work we investigate the long-time behavior of mild solutions to the stochastic partial differential equation of the form

$$dX_t = (AX_t + F(X_t))dt + \sigma(X_t)dW_t + \int_E \gamma(X_t, \nu)\widetilde{N}(dt, d\nu), \qquad t \geq 0 \tag{1.1}$$

on a separable Hilbert space H, where $(A, D(A))$ is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H, $(W_t)_{t\geq 0}$ is a Q-Wiener process and $\widetilde{N}(dt, d\nu)$ denotes a compensated Poisson random measure. The precise conditions need to be imposed on these objects will be formulated in the subsequent sections.

In the literature the study on the existence and uniqueness of invariant measures often relies on different variants of a *dissipativity condition*. The simplest form of such

Date: August 17, 2020.

2010 *Mathematics Subject Classification.* Primary 37L40, 60J25; Secondary 37A30, 37L55.

Key words and phrases. Stochastic Partial Differential Equations, multiple invariant measures, long time behaviour, dissipativity condition, semigroups, projection operators, Wasserstein metrics, Yosida approximation, Heath-Jarrow-Morton-Musiela equation, delay equations.

*Corresponding Author: Barbara Rüdiger, School of Mathematics and Natural Sciences, Bergische Universität Wuppertal, Gaußstraße 20, 42119 Wuppertal.

a dissipativity condition is: There exists $\alpha > 0$ such that

$$\langle Ax - Ay, x - y\rangle_H + \langle F(x) - F(y), x - y\rangle_H \leq -\alpha\|x - y\|_H^2, \qquad x, y \in D(A). \quad (1.2)$$

Indeed, if (1.2) is satisfied, σ and γ are globally Lipschitz-continuous, and α is large enough, then there exists a unique invariant measure for the Markov process obtained from (1.1), see, e.g., [32, Section 16], [10, Chapter 11, Section 6], and [36] where such a condition was formulated for the Yosida approximations of the operator $(A, D(A))$. Note that (1.2) is satisfied, if F is globally Lipschitz continuous and $(A, D(A))$ satisfies for some $\beta > 0$ large enough the inequality $\langle Ax, x\rangle_H \leq -\beta\|x\|_H^2$, $x \in D(A)$, i.e. $(A, D(A))$ is the generator of a strongly continuous semigroup satisfying $\|S(t)\|_{L(H)} \leq \mathrm{e}^{-\beta t}$. Here and below we denote by $L(H)$ the space of bounded linear operators from H to H and by $\|\cdot\|_{L(H)}$ its operator norm. For weaker variants of the dissipativity condition (e.g. cases where (1.2) only holds for $\|x\|_H, \|y\|_H \geq R$ for some $R > 0$), in general one can neither guarantee the existence nor uniqueness of an invariant measure. Hence, to treat such cases, additional arguments, e.g. coupling methods, are required. Such arguments have been applied to different stochastic partial differential equations on Hilbert spaces in [33, 34, 35] where existence and, in particular, uniqueness of invariant measures was studied. We also mention [7, 23] for an extension of Harris-type theorems for Wasserstein distances, and [25, 21] for extensions of coupling methods.

In contrast to the aforementioned methods and applications, several stochastic models exhibit phase transition phenomena where uniqueness of invariant measures fails to hold. For instance, the generator $(A, D(A))$ and drift F appearing in the Heath-Jarrow-Morton-Musiela equation do not satisfy (1.2), but instead F is globally Lipschitz continuous and the semigroup generated by $(A, D(A))$ satisfies

$$\|S(t)x - Px\|_H \leq \mathrm{e}^{-\alpha t}\|x - Px\|_H$$

for some projection operator P. Based on this property it was shown in [38, 36] that the Heath-Jarrow-Morton-Musiela equation has infinitely many invariant measures parametrized by the initial state of the process, see also Section 6. Another example is related to stochastic Volterra equations as studied, e.g., in [6]. There, using a representation of stochastic Volterra equations via SPDEs and combined with some arguments originated from the study of the Heath-Jarrow-Morton-Musiela equation, the authors studied existence of limiting distributions allowing, in particular, that these distributions depend on the initial state of the process.

In this work we provide a general and unified approach for the study of multiple invariant measures and, moreover, we show that with dependence on the initial distribution the law of the mild solution of (1.1) is governed in the limit $t \to \infty$ by one of the invariant measures. In particular, we show that the methods developed in [38, 36, 6] can be embedded as a special case of a general framework where one replaces (1.2) by a weaker dissipativity condition, which we call hereinafter *generalized dissipativity condition*:

(GDC) There exists a projection operator P_1 on the Hilbert space H and there exist constants $\alpha > 0, \beta \geq 0$ such that, for $x, y \in D(A)$, one has:

$$\begin{aligned}\langle Ax - Ay, x - y\rangle_H &+ \langle F(x) - F(y), x - y\rangle_H \\ &\leq -\alpha\|x - y\|_H^2 + (\alpha + \beta)\,\|P_1x - P_1y\|_H^2.\end{aligned}$$

Note that for the special case $P_1 = 0$ condition (GDC) contains the classical dissipativity condition. However, when $P_1 \neq 0$, the additional term $\|P_1x - P_1y\|_H^2$ describes the influence of the non-dissipative part of the drift. Sufficient conditions and additional remarks on this condition are collected in the end of Section 2 while particular examples are discussed in Sections 5 – 7.

We will show that under condition (GDC) and additional restrictions on the projected coefficients P_1F, $P_1\sigma$, and $P_1\gamma$, the Markov process obtained from (1.1) has for each initial data $X_0 = x$ a limiting distribution π_x depending only on P_1x. Moreover, the transition probabilities converge exponentially fast in the Wasserstein 2-distance to this limiting distribution. In order to prove this result, we first decompose the Hilbert space H according to

$$H = H_0 \oplus H_1, \qquad x = P_0x + P_1x, \quad P_0 := I - P_1,$$

where I denotes the identity operator on H, and then investigate the components P_0X_t and P_1X_t separately. Based on an technique from [39], we construct, for each $\tau \geq 0$, a coupling of X_t and $X_{t+\tau}$. This coupling will be then used to efficiently estimate the Wasserstein 2-distance for the solution started at two different points.

This work is organized as follows. In Section 2 we first discuss the special case where $F \equiv 0$ and σ, γ are independent of X. In such a case X is an Ornstein-Uhlenbeck type process and the collection of invariant measures can be easily characterized by its characteristic function. This section can be seen as a motivation for our more general results discussed in the subsequent sections. Afterward, we investigate in Sections 3 – 5 the general case for which the methods from Section 2 can not be applied. More precisely, after having introduced and discussed in Section 3 the *generalized dissipativity condition* (GDC), we state in Section 4 the precise conditions imposed on the coefficients of the SPDE (1.1), discuss some properties of the solution and then provide sufficient conditions for the generalized dissipativity condition (GDC). Based on condition (GDC) we derive in Section 4 an estimate on the trajectories of the process when started at two different initial points, i.e. we estimate the L^2-norm of $X_t^x - X_t^y$ when $x \neq y$. Based on this estimate, we then state and prove our main results in Section 5. Examples are then discussed in the subsequent Sections 6 and 7. Namely, the Heath-Jarrow-Morton-Musiela equation is considered in Section 6 for which we first show that the main results of Section 5 contain [38, 36], and then extend these results by characterizing its limiting distributions more explicitly. Finally, we apply our results in Section 7 to an SPDE with delay.

2. Ornstein-Uhlenbeck process in a Hilbert space

Let H be a separable Hilbert space and let $(Z_t)_{t\geq 0}$ be a H-valued Lévy process with Lévy triplet (b, Q, μ) defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with the usual conditions. This has characteristic exponent Ψ of Lévy-Khinchine form, i.e.

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u, Z_t\rangle_H}\right] = \mathrm{e}^{t\Psi(u)}, \qquad u \in H, \quad t > 0,$$

with Ψ given by

$$\Psi(u) = \mathrm{i}\langle b, u\rangle_H - \frac{1}{2}\langle Qu, u\rangle_H + \int_H \left(\mathrm{e}^{\mathrm{i}\langle u,z\rangle_H} - 1 - \mathrm{i}\langle u, z\rangle_H \mathbb{1}_{\{\|z\|_H \leq 1\}}\right) \mu(dz),$$

where $b \in H$ denotes the drift, Q denotes the covariance operator being a positive, symmetric, trace-class operator on H, and μ is a Lévy measure on H (see e.g. [27], [3], [32], [28]). Let $(S(t))_{t\geq 0}$ be a strongly continuous semigroup on H. The Ornstein-Uhlenbeck process driven by $(Z_t)_{t\geq 0}$ is the unique mild solution to

$$dX_t^x = AX_t^x dt + dZ_t, \qquad X_0^x = x \in H, \quad t \geq 0,$$

where $(A, D(A))$ denotes the generator of $(S(t))_{t\geq 0}$, i.e. $(X_t^x)_{t\geq 0}$ satisfies

$$X_t^x = S(t)x + \int_0^t S(t-s)dZ_s, \qquad t \geq 0.$$

The characteristic function of $(X_t^x)_{t\geq 0}$ is given by

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u, X_t^x\rangle_H}\right] = \exp\left(\mathrm{i}\langle S(t)x, u\rangle_H + \int_0^t \Psi(S(r)^*u)dr\right), \qquad u \in H, \quad t \geq 0.$$

See e.g. the review article [3] where also sufficient conditions for the existence and for the uniqueness as well as properties of invariant measures are discussed. It is well-known that the Ornstein-Uhlenbeck process has a unique invariant measure provided that $(S(t))_{t\geq 0}$ is uniformly exponentially stable, that is

$$\exists \alpha > 0, \ M \geq 1: \qquad \|S(t)\|_{L(H)} \leq M\mathrm{e}^{-\alpha t}, \qquad t \geq 0,$$

and the Lévy measure μ satisfies a log-integrability condition for its big jumps

$$\int_{\{\|z\|_H > 1\}} \log(1 + \|z\|_H)\mu(dz) < \infty. \tag{2.1}$$

Below we show that for a uniformly convergent semigroup $(S(t))_{t\geq 0}$ the corresponding Ornstein-Uhlenbeck process may admit multiple invariant measures parameterized by the range of the limiting projection operator of the semigroup.

Theorem 2.1. *Suppose that $(S(t))_{t\geq 0}$ is uniformly exponentially convergent, i.e. there exists a projection operator P on H and constants $M \geq 1$, $\alpha > 0$ such that*

$$\|S(t)x - Px\|_H \leq M\|x\|_H \mathrm{e}^{-\alpha t}, \qquad t \geq 0, x \in H. \tag{2.2}$$

Suppose that the Lévy process satisfies the following conditions:

(i) The drift b satisfies $Pb = 0$.
(ii) The covariance operator Q satisfies $PQu = 0$ for all $u \in H$.
(iii) The Lévy measure μ is supported on $\ker(P)$ and satisfies (2.1).

Then for each $x \in H$ it holds

$$X_t^x \longrightarrow Px + X_\infty^0, \qquad t \to \infty$$

in law, where X_∞^0 is an H-valued random variable determined by

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u, X_\infty^0\rangle_H}\right] = \exp\left(\int_0^\infty \Psi(S(r)^*u)dr\right).$$

In particular, the set of all limiting distributions for the Ornstein-Uhlenbeck process $(X_t^x)_{t\geq 0}$ is given by $\{\delta_x * \mu_\infty \mid x \in \mathrm{ran}(P)\}$, where μ_∞ denotes the law of X_∞^0.

Proof. We first prove the existence of a constant $C > 0$ such that

$$\int_0^\infty |\Psi(S(r)^*u)|dr \leq C(\|u\|_H + \|u\|_H^2), \qquad u \in H, \tag{2.3}$$

where $S(r)^*$ denotes the adjoint operator to $S(r)$ on $L(H)$. To do so we estimate

$$\begin{aligned}
|\Psi(S(r)^*u)| &\leq |\langle b, S(r)^*u\rangle| + |\langle QS(r)^*u, S(r)^*u\rangle| \\
&\quad + \int_{\{\|z\|_H \leq 1\}} \left|\mathrm{e}^{\mathrm{i}\langle S(r)^*u, z\rangle} - 1 - \mathrm{i}\langle S(r)^*u, z\rangle\right| \mu(dz) \\
&\quad + \int_{\{\|z\|_H > 1\}} \left|\mathrm{e}^{\mathrm{i}\langle S(r)^*u, z\rangle} - 1\right| \mu(dz) \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

We find by (2.2) that $\|S(r)x\|_H \le M\mathrm{e}^{-\alpha r}\|x\|_H$ for all $x \in \ker(P)$ and hence

$$I_1 = |\langle S(r)b, u\rangle| \le \|u\|_H \|S(r)b\|_H \le \|u\|_H M \mathrm{e}^{-\alpha r}\|b\|_H.$$

For the second term I_2 we use $\mathrm{ran}(Q) \subset \ker(P)$ so that

$$\|S(r)Qu\|_H \le M\mathrm{e}^{-\alpha r}\|Qu\|_H \le \mathrm{e}^{-\alpha r}\|Q\|_{L(H)}\|u\|_H.$$

This yields $\|QS(r)^*\|_{L(H)} = \|S(r)Q\|_{L(H)} \le M\mathrm{e}^{-\alpha r}\|Q\|_{L(H)}$ and hence

$$\begin{aligned} I_2 &= |\langle QS(r)^*u, S(r)^*u\rangle| \\ &\le \|QS(r)^*u\|_H \|S(r)^*u\|_H \\ &\le M\|u\|_H \|QS(r)^*u\|_H \\ &\le M\|u\|_H^2 \|Q\|_{L(H)} M\mathrm{e}^{-\alpha r}. \end{aligned}$$

For the third term I_3 we obtain

$$\begin{aligned} I_3 &\le C \int_{\{\|z\|_H \le 1\}} |\langle S(r)^*u, z\rangle|^2 \mu(dz) \\ &= C \int_{\{\|z\|_H \le 1\} \cap \ker(P)} |\langle u, S(r)z\rangle|^2 \mu(dz) \\ &\le C\|u\|_H^2 \mathrm{e}^{-\alpha r} \int_{\{\|z\|_H \le 1\}} \|z\|_H^2 \mu(dz), \end{aligned}$$

where $C > 0$ is a generic constant. Proceeding similarly for the last term, we obtain

$$\begin{aligned} I_3 &\le C \int_{\{\|z\|_H > 1\}} \min\{1, |\langle S(r)^*u, z\rangle|\} \, \mu(dz) \\ &\le C \int_{\{\|z\|_H > 1\} \cap \ker(P)} \min\left\{1, \|u\|_H \mathrm{e}^{-\alpha r}\|z\|_H\right\} \mu(dz) \\ &\le C\|u\|_H \mathrm{e}^{-\alpha r} \left(\mu(\{\|z\|_H > 1\}) + \int_{\{\|z\|_H > 1\}} \log(1 + \|z\|_H)\mu(dz) \right), \end{aligned}$$

where we have used, for $a = \|u\|_H \mathrm{e}^{-\alpha r}$, $b = \|z\|_H$, the elementary inequalities

$$\begin{aligned} \min\{1, ab\} &\le C\log(1 + ab) \\ &\le C\min\{\log(1+a), \log(1+b)\} + C\log(1+a)\log(1+b) \\ &\le Ca\left(1 + \log(1+b)\right), \end{aligned}$$

see [19, appendix]. Combining the estimates for I_1, I_2, I_3, I_4 we conclude that (2.3) is satisfied. Hence, using

$$\lim_{t\to\infty} \langle S(t)x, u\rangle = \langle Px, u\rangle$$

we find that

$$\lim_{t\to\infty} \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u, X_t^x\rangle}\right] = \exp\left(\mathrm{i}\langle Px, u\rangle + \int_0^\infty \Psi(S(r)^*u)dr\right). \tag{2.4}$$

Since, in view of (2.3), $u \longmapsto \int_0^\infty \Psi(S(r)^*u)dr$ is continuous at $u = 0$, the assertion follows from Lévy's continuity theorem combined with the particular form of (2.4). $\square$

Below we briefly discuss an application of this result to a stochastic perturbation of the Kolmogorov equation associated with a symmetric Markov semigroup. Let E be a Polish space and η a Borel probability measure on E. Let $(A, D(A))$ be the generator of a symmetric Markov semigroup $(S(t))_{t\geq 0}$ on $H := L^2(E,\eta)$. Then there exists, for each $f \in D(A)$, a unique solution to the Kolmogorov equation (see, e.g., [31])

$$\frac{dv(t)}{dt} = Av(t), \qquad v(0) = f.$$

Below we consider an additive stochastic perturbation of this equation in the sense of Itô, i.e. the stochastic partial differential equation

$$dv(t) = Av(t)dt + dZ_t, \qquad v(0) = f, \tag{2.5}$$

where $(Z_t)_{t\geq 0}$ is an $L^2(E,\eta)$-valued Lévy process with characteristic function Ψ. Let $(v(t);f))_{t\geq 0}$ be the unique mild solution to this equation.

Corollary 2.2. *Suppose that the semigroup generated by $(A, D(A))$ on $L^2(E,\eta)$ satisfies* (2.2) *with the projection operator*

$$Pv = \int_E v(x)\eta(dx),$$

and $H = L^2(E,\eta)$. Assume that the Lévy process $(Z_t)_{t\geq 0}$ satisfies the conditions (i) – (iii) of Theorem 2.1. Then

$$v(t;f) \longrightarrow \int_E f(x)\eta(dx) + v(\infty), \qquad t \to \infty$$

in law, where $v(\infty)$ is a random variable whose characteristic function is given by

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u, v(\infty)\rangle_{L^2}}\right] = \exp\left(\int_0^\infty \Psi(S(r)^* u)dr\right).$$

We close this section with an example of a semigroup $(S(t))_{t\geq 0}$ for which this corollary can be applied.

Example 2.3. *Let $(X_t)_{t\geq 0}$ be a Feller process on a separable Hilbert space E and let $(p_t)_{t\geq 0}$ be its transition semigroup acting on $C_b(E)$. Suppose that $(X_t)_{t\geq 0}$ has a unique invariant measure η. Then, by Yensen inequality, $(p_t)_{t\geq 0}$ can be uniquely extended to a strongly continuous semigroup on $L^2(E,\eta)$ which is for simplicity again denoted by $(p_t)_{t\geq 0}$. Suppose that this semigroup is L^2-exponentially convergent in the sense that*

$$\lim_{t\to\infty} \int_E \left(p_t f - \int_E f(x)\eta(dx)\right)^2 d\eta = 0, \qquad \forall f \in L^2(E,\eta).$$

Then $(p_t)_{t\geq 0}$ satisfies (2.2) *with projection operator $Pv = \int_E v(x)\eta(dx)$.*

3. Preliminaries

3.1. **Framework and notation.** Here and throughout this work, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ is a filtered probability space satisfying the usual conditions. Let U be a separable Hilbert space and $W = (W_t)_{t\geq 0}$ be a Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$, where $Q : U \to U$ is a non-negative, symmetric, trace class operator. Let E be a Polish space, $\mathcal{E}$ the Borel-σ-field on E, and μ a σ-finite measure on $(E,\mathcal{E})$. Let $N(dt,d\nu)$ be a $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure with compensator $dt\mu(d\nu)$ and denote by $\widetilde{N}(dt,d\nu) = N(dt,d\nu) - dt\mu(d\nu)$ the corresponding compensated Poisson random measure. Suppose that the random objects $(W_t)_{t\geq 0}$ and $N(dt,d\nu)$ are mutually independent.

In this work we investigate the long-time behavior of mild solutions to the stochastic partial differential equation (1.1) with initial condition $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, that is

$$dX_t = (AX_t + F(X_t))dt + \sigma(X_t)dW_t + \int_E \gamma(X_t^x, \nu)\widetilde{N}(dt, d\nu), \qquad t \geq 0, \tag{3.1}$$

where $(A, D(A))$ is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H, $H \ni x \mapsto F(x) \in H$ and $H \ni x \mapsto \sigma(x) \in L_2^0$ are Borel measurable mappings, and $(x, \nu) \mapsto \gamma(x, \nu)$ is measurable from $(H \times E, \mathcal{B}(H) \otimes \mathcal{E})$ to $(H, \mathcal{B}(H))$. Here $\mathcal{B}(H)$ denotes the Borel-σ-algebra on H, and $L_2^0 := L_2^0(H)$ is the Hilbert space of all Hilbert-Schmidt operators from U_0 to H, where $U_0 := Q^{1/2}U$ is a separable Hilbert space endowed with the scalar product

$$\langle x, y\rangle_0 := \langle Q^{-1/2}x, Q^{-1/2}y\rangle_U = \sum_{k\in\mathbb{N}} \tfrac{1}{\lambda_k}\langle x, e_k\rangle_U \langle e_k, y\rangle_U, \qquad \forall x, y \in U_0,$$

and $Q^{-1/2}$ denotes the pseudoinverse of $Q^{1/2}$. Here $(e_j)_{j\in\mathbb{N}}$ denotes an orthogonal basis of eigenvectors of Q in U with corresponding eigenvalues $(\lambda_j)_{j\in\mathbb{N}}$. For comprehensive introductions to integration concepts in infinite dimensional settings we refer e.g. to [10] for the case of Q-Wiener processes and e.g. to [3], [32], [28] for compensated Poisson random measures as integrators. Throughout this work we suppose that the coefficients F, σ, γ are Lipschitz continuous. More precisely:

(A1) There exist constants $L_F, L_\sigma, L_\gamma \geq 0$ such that for all $x, y \in H$

$$\begin{aligned} \|F(x) - F(y)\|_H^2 &\leq L_F\|x - y\|_H^2, \\ \|\sigma(x) - \sigma(y)\|_{L_2^0(H)}^2 &\leq L_\sigma\|x - y\|_H^2, \\ \int_E \|\gamma(x, \nu) - \gamma(y, \nu)\|_H^2 \mu(d\nu) &\leq L_\gamma\|x - y\|_H^2. \end{aligned} \tag{3.2}$$

Moreover we suppose that

$$\int_E \|\gamma(0, \nu)\|_H^2 \mu(d\nu) < \infty. \tag{3.3}$$

Note that condition (3.3) implies that the jumps satisfy the usual growth conditions, i.e.

$$\begin{aligned} \int_E \|\gamma(x, \nu)\|_H^2 \mu(d\nu) &\leq 2\int_E \|\gamma(x, \nu) - \gamma(0, \nu)\|_H^2 \mu(d\nu) + 2\int_E \|\gamma(0, \nu)\|_H^2 \mu(d\nu) \\ &\leq 2\max\left\{L_\gamma, \int_E \|\gamma(0, \nu)\|_H^2 \mu(d\nu)\right\}(1 + \|x\|_H^2). \end{aligned}$$

Moreover, it follows from (GDC) and (3.2) that

$$\langle Ax, x\rangle_H \leq \left(\beta + \sqrt{L_F}\right)\|x\|_H^2, \qquad x \in D(A).$$

Hence $A - (\beta + \sqrt{L_F})$ is dissipative and thus by the Lumer-Phillips theorem the semigroup $(S(t))_{t\geq 0}$ generated by $(A, D(A))$ is quasi-contractive, i.e.

$$\|S(t)x\|_H \leq \mathrm{e}^{\left(\beta + \sqrt{L_F}\right)t}\|x\|_H, \quad x \in H. \tag{3.4}$$

Then, under conditions (GDC) and (A1), for each initial condition $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ there exists a unique càdlàg, $(\mathcal{F}_t)_{t\geq 0}$-adapted, mean square continuous, mild solution

$(X_t)_{t\geq 0}$ to (3.1) such that, for each $T>0$, there exists a constant $C(T)>0$ satisfying

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|X_t\|_H^2\right] \leq C(T)\left(1+\mathbb{E}\left[\|X_0\|_H^2\right]\right) \tag{3.5}$$

This means that $(X_t)_{t\geq 0}$ satisfies $\mathbb{P}$-a.s.

$$\begin{aligned} X_t &= S(t)X_0 + \int_0^t S(t-s)F(X_s)ds + \int_0^t S(t-s)\sigma(X_s)dW_s \qquad (3.6)\\ &\quad + \int_0^t\int_E S(t-s)\gamma(X_s,\nu)\widetilde{N}(ds,d\nu), \qquad t\geq 0, \end{aligned}$$

where all (stochastic) integrals are well-defined, see, e.g., [1], [28], and [17]. Moreover, for each $X_0, Y_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$, the corresponding unique solutions $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ satisfy

$$\mathbb{E}\left[\|X_t - Y_t\|_H^2\right] \leq C(T)\mathbb{E}\left[\|X_0-Y_0\|_H^2\right], \qquad t\in[0,T]. \tag{3.7}$$

If $X_0 \equiv x \in H$, then we denote by $(X_t^x)_{t\geq 0}$ the corresponding solution to (3.1). Such solution constitutes a Markov process whose transition probabilities $p_t(x,dy) = \mathbb{P}[X_t^x \in dy]$ are measurable with respect to x. By slight abuse of notation we denote by $(p_t)_{t\geq 0}$ its transition semigroup, i.e., for each bounded measurable function $f: H \longrightarrow \mathbb{R}$, $p_t f$ is given by

$$p_t f(x) = \mathbb{E}\left[f(X_t^x)\right] = \int_H f(y)p_t(x,dy), \qquad t\geq 0, \quad x\in H.$$

Using the continuous dependence on the initial condition, see (3.7), it can be shown that $p_t f \in C_b(H)$ for each $f\in C_b(H)$, i.e. the transition semigroup is C_b-Feller.

In this work we investigate the the existence of invariant measures and convergence of the transition probabilities towards these measures for the Markov process $(X_t^x)_{t\geq 0}$ with particular focus on the cases where uniqueness of invariant measures fails to hold. We denote by p_t^* the adjoint operator to p_t defined by

$$p_t^*\rho(dx) = \int_H p_t(y,dx)\rho(dy), \qquad t\geq 0.$$

Recall that a probability measure π on $(H,\mathcal{B}(H))$ is called *invariant measure* for the semigroup $(p_t)_{t\geq 0}$ if and only if $p_t^*\pi = \pi$ holds for each $t\geq 0$. Let $\mathcal{P}_2(H)$ be the space of Borel probability measures ρ on $(H,\mathcal{B}(H))$ with finite second moments. Recall that $\mathcal{P}_2(H)$ is separable and complete when equipped with the *Wasserstein-2-distance*

$$\mathrm{W}_2(\rho,\widetilde{\rho}) = \inf_{G\in\mathcal{H}(\rho,\widetilde{\rho})}\left(\int_{H\times H}\|x-y\|_H^2 G(dx,dy)\right)^{\frac{1}{2}}, \qquad \rho,\widetilde{\rho}\in\mathcal{P}_2(H). \tag{3.8}$$

Here $\mathcal{H}(\rho,\widetilde{\rho})$ denotes the set of all couplings of $(\rho,\widetilde{\rho})$, i.e. Borel probability measures on $H\times H$ whose marginals are given by ρ and $\widetilde{\rho}$, respectively, see [40, Section 6] for a general introduction to couplings and Wasserstein distances.

3.2. **Discussion of generalized dissipativity condition.** In this section we briefly discuss the condition

$$\langle Ax,x\rangle_H \leq -\lambda_0\|x\|_H^2 + (\lambda_0+\lambda_1)\|P_1x\|_H^2, \qquad x\in D(A), \tag{3.9}$$

where $\lambda_0>0$ and $\lambda_1\geq 0$. Note that, if (3.9) and condition (3.1) are satisfied, then

$$\begin{aligned} &\langle Ax-Ay,x-y\rangle_H + \langle F(x)-F(y),x-y\rangle_H \qquad (3.10)\\ &\leq \langle Ax-Ay,x-y\rangle_H + \sqrt{L_F}\|x-y\|_H^2 \end{aligned}$$

$$\leq -\left(\lambda_0 - \sqrt{L_F}\right)\|x-y\|_H^2 + (\lambda_0+\lambda_1)\|P_1x - P_1y\|_H^2,$$

i.e. the generalized dissipativity condition (GDC) is satisfied for $\alpha = \lambda_0 - \sqrt{L_F}$ and $\beta = \lambda_1 + \sqrt{L_F}$, provided that $\lambda_0 > \sqrt{L_F}$.

Proposition 3.1. *Suppose that there exists an orthogonal decomposition $H = H_0 \oplus H_1$ of H into closed linear subspaces $H_0, H_1 \subset H$ such that $(S(t))_{t\geq 0}$ leaves H_0 and H_1 invariant and there exist constants $\lambda_0 > 0$ and $\lambda_1 \geq 0$ satisfying*

$$\|S(t)x_0\|_H \leq \mathrm{e}^{-\lambda_0 t}\|x_0\|_H, \qquad \|S(t)x_1\|_H \leq \mathrm{e}^{\lambda_1 t}\|x_1\|_H, \qquad \forall t \geq 0.$$

for all $x_0 \in H_0$ and $x_1 \in H_1$. Then (3.9) *holds for P_1 being the orthogonal projection operator onto H_1.*

Proof. Let P_0 be the orthogonal projection operator onto H_0. Since $(S(t))_{t\geq 0}$ leaves the closed subspace H_0 invariant, its restriction $(S(t)|_{H_0})_{t\geq 0}$ onto H_0 is a strongly continuous semigroup of contractions on H_0 with generator $(A_0, D(A_0))$ being the H_0 part of A, that is

$$A_0x = Ax, \qquad x \in D(A_0) = \{y \in D(A) \cap H_0 \mid Ay \in H_0\}.$$

Since H_0 is closed and $S(t)$ leaves H_0 invariant, it follows that $Ay = \lim_{t\to 0}\frac{S(t)y-y}{t} \in H_0$ for $y \in D(A) \cap H_0$, i.e. $D(A_0) = D(A) \cap H_0$ and $P_0 : D(A) \to D(A_0)$. Arguing exactly in the same way shows that the restriction $(S(t)|_{H_1})_{t\geq 0}$ is a strongly continuous semigroup of contractions on H_1 with generator $(A_1, D(A_1))$ given by $A_1x = Ax$ and $x \in D(A_1) = D(A) \cap H_1$ so that $P_1 : D(A) \to D(A_1)$. Since $S(t)$ leaves H_0 and H_1 invariant, we obtain $P_0S(t) = S(t)P_0$, $P_1S(t) = S(t)P_1$ from which we conclude that $AP_1x = P_1Ax$ and $AP_0x = P_0Ax$ for $x \in D(A)$.

Since $(\mathrm{e}^{\lambda_0 t}S(t)|_{H_0})_{t\geq 0}$ is a strongly continuous semigroup of contractions on H_0 with generator $A_0 + \lambda_0 I$, and $(\mathrm{e}^{-\lambda_1 t}S(t)|_{H_1})_{t\geq 0}$ is a strongly continuous semigroup of contractions on H_1 with generator $A_1 - \lambda_1 I$, we have by the Lumer-Phillips theorem (see [31, Theorem 4.3])

$$\langle A_0x_0, x_0\rangle_H \leq -\lambda_0\|x_0\|_H^2 \ \text{ and } \ \langle A_1x_1, x_1\rangle_H \leq \lambda_1\|x_1\|_H^2, \qquad x_0 \in H_0, \quad x_1 \in H_1.$$

Hence we find that

$$\begin{aligned}
\langle Ax, x\rangle_H &= \langle Ax, P_0x\rangle_H + \langle Ax, P_1x\rangle_H \\
&= \langle P_0Ax, P_0x\rangle_H + \langle P_1Ax, P_1x\rangle_H \\
&= \langle A_0P_0x, P_0x\rangle_H + \langle A_1P_1x, P_1x\rangle_H \\
&\leq -\lambda_0\|P_0x\|_H^2 + \lambda_1\|P_1x\|_H^2 \\
&= -\lambda_0\|x\|_H^2 + (\lambda_0+\lambda_1)\|P_1x\|_H^2,
\end{aligned}$$

where the last equality follows from $H_0 \perp H_1$. This proves the assertion. □

At this point it is worthwhile to mention that Onno van Gaans has investigated in [39] ergodicity for a class of Lévy driven stochastic partial differential equations where the semigroup $(S(t))_{t\geq 0}$ was supposed to be hyperbolic. Proposition 3.1 can be also applied for hyperbolic semigroups provided that the hyperbolic decomposition is orthogonal. The conditions of previous proposition are satisfied whenever $(S(t))_{t\geq 0}$ is a symmetric, uniformly convergent semigroup.

Remark 3.2. *Suppose that $(S(t))_{t\geq 0}$ is a strongly continuous semigroup on H and there exists an orthogonal projection operator P on H and $\lambda_0 > 0$ such that*

$$\|S(t)x - Px\|_H \leq \mathrm{e}^{-\lambda_0 t}\|x - Px\|_H, \qquad t \geq 0, \quad x \in H. \tag{3.11}$$

Then the conditions of Proposition 3.1 are satisfied for $H_0 = \ker(P)$ *and* $H_1 = \operatorname{ran}(P)$ *with* $\lambda_0 > 0$ *and* $\lambda_1 = 0$*. In particular,* $(S(t))_{t\geq 0}$ *is a semigroup of contractions.*

The following example shows that (3.9) can also be satisfied for non-symmetric and non-convergent semigroups.

Example 3.3. *Let* $H = \mathbb{R}^2$*,* $H_0 = \mathbb{R}\times\{0\}$*,* $H_1 = \{0\}\times\mathbb{R}$*, and denote by* P_0, P_1 *the projection operators onto* H_0 *and* H_1*, respectively. Let* A *be given by* $A = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$*. Then*

$$\begin{aligned}\left\langle \begin{pmatrix} x \\ y \end{pmatrix}, A\begin{pmatrix} x \\ y \end{pmatrix} \right\rangle &= -x^2 + xy + y^2 \\ &\leq -\frac{1}{2}(x^2+y^2) + 2y^2 \\ &= -\frac{1}{2}\|(x,y)\|_H^2 + 2\|P_1(x,y)\|_H^2,\end{aligned}$$

i.e. (3.9) *holds for* $\lambda_0 = \frac{1}{2}$ *and* $\lambda_1 = \frac{3}{2}$*. Since* $\mathrm{e}^{tA} = \begin{pmatrix} \mathrm{e}^{-t} & \frac{\mathrm{e}^{t}-\mathrm{e}^{-t}}{2} \\ 0 & \mathrm{e}^{t} \end{pmatrix}$*, it is clear that neither the conditions of Proposition 3.1 nor of Remark 3.2 are satisfied.*

3.3. **Key stability estimate.** Define, for $x, y \in D(A)$, the function

$$\begin{aligned}\mathcal{L}(\|\cdot\|_H^2)(x,y) &:= 2\langle A(x-y) + F(x) - F(y), x-y\rangle_H + \|\sigma(x)-\sigma(y)\|_{L_2^0(H)}^2 \\ &\quad + \int_E \|\gamma(x,\nu)-\gamma(y,\nu)\|_H^2 \mu(d\nu).\end{aligned}$$

Remark that if (1.1) has a strong solution, then the function

$$\mathcal{L}(\|\cdot\|_H^2)(z) := 2\langle A(z) + F(z), z\rangle_H + \|\sigma(z)\|_{L_2^0(H)}^2 + \int_E \|\gamma(z,\nu)\|_H^2 \mu(d\nu).$$

is simply the generator $\mathcal{L}$ applied to the unbounded function $\|z\|_H^2$, see, e.g,. [2, equation (3.4)]). Since we work with mild solutions instead, all computations given below require to use additionally Yosida approximations for the mild solution of (1.1).

Below we first prove a Lyapunov-type estimate for $\mathcal{L}(\|\cdot\|_H^2)$ and then deduce from that by an application of the generalized Itô-formula A.2 to (3.1) an estimate for the L^2-norm of $X_t^x - X_t^y$.

Lemma 3.4. *Assume that condition (GDC) and (A1) are satisfied. Then*

$$\mathcal{L}(\|\cdot\|_H^2)(x,y) \leq -\left(2\alpha - L_\sigma - L_\gamma\right)\|x-y\|_H^2 + 2(\alpha+\beta)\|P_1x - P_1y\|_H^2 \tag{3.12}$$

holds for $x, y \in D(A)$*.*

Proof. Using first (A1) and then (GDC) we find that

$$\begin{aligned}\mathcal{L}(\|\cdot\|_H^2)(x,y) &\leq (L_\sigma + L_\gamma)\|x-y\|_H^2 \\ &\qquad + 2\langle Ax - Ay, x-y\rangle_H + 2\langle F(x)-F(y), x-y\rangle_H \\ &\leq -\left(2\alpha - L_\sigma - L_\gamma\right)\|x-y\|_H^2 + 2\left(\alpha+\beta\right)\|P_1x - P_1y\|_H^2.\end{aligned}$$

This proves the asserted inequality. □

The following is our key stability estimate.

Proposition 3.5. *Suppose that (GDC) and (A1) are satisfied, that*

$$\varepsilon := 2\alpha - L_\sigma - L_\gamma > 0, \tag{3.13}$$

and suppose that

$$\sup_{x\in H}\int_E \|\gamma(x,\nu)\|^4\mu(d\nu) < \infty. \tag{3.14}$$

Then, for each $X_0, Y_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ *and all* $t \geq 0$,

$$\begin{aligned}&\mathbb{E}\left[\|X_t - Y_t\|_H^2\right]\\ \leq &\mathrm{e}^{-\varepsilon t}\mathbb{E}\left[\|X_0 - Y_0\|_H^2\right] + 2(\alpha+\beta)\int_0^t \mathrm{e}^{-\varepsilon(t-s)}\mathbb{E}\left[\|P_1X_s - P_1Y_s\|_H^2\right]ds,\end{aligned} \tag{3.15}$$

where $(X_t)_{t\geq 0}$ *and* $(Y_t)_{t\geq 0}$ *denote the unique solutions to* (3.1)*, respectively.*

Proof. Let $(X_t^n)_{t\geq 0}$ and $(Y_t^n)_{t\geq 0}$ be the strong solutions to the corresponding Yosida-approximation systems

$$\begin{cases} dX_t^n = AX_t^n + R_nF(X_t^n)dt + R_n\sigma(X_t^n)dW_t + \int_E R_n\gamma(X_t^n,\nu)\widetilde{N}(dt,d\nu), \\ X_0^n = R_nX_0, \quad t \geq 0 \end{cases}$$

and

$$\begin{cases} dY_t^n = AY_t^n + R_nF(Y_t^n)dt + R_n\sigma(Y_t^n)dW_t + \int_E R_n\gamma(Y_t^n,\nu)\widetilde{N}(dt,d\nu), \\ Y_0^n = R_nY_0, \quad t \geq 0 \end{cases}$$

where $R_n = n(n-A)^{-1}$ for $n \in \mathbb{N}$ with $n > \alpha + \beta + \sqrt{L_F} =: \lambda$. By (3.4) we find for each $n \geq 1 + \lambda$ the inequality

$$\|R_nz\|_H \leq \frac{n}{n-\lambda}\|z\|_H \leq (1+\lambda)\|z\|_H.$$

By classical properties of the resolvent (see [31, Lemma 3.2]), one clearly has $R_nz \to z$ as $n \to \infty$ in H . Moreover, by properties of the Yosida approximation of mild solutions of SPDEs (compare e.g. with Appendix A2 in [28] or Section 2 in [2]) we have

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}\|X_t^n - X_t\|_H^2 + \sup_{t\in[0,T]}\|Y_t^n - Y_t\|_H^2\right] = 0, \qquad \forall T > 0$$

and hence there exists a subsequence (which is again denoted by n) such that $X_t^n \longrightarrow X_t$ and $Y_t^n \longrightarrow Y_t$ hold a.s. for each $t \geq 0$. Following a method proposed in [2] we verify that sufficient conditions are satisfied to apply the generalized Itô-formula from Theorem A.2 to the function $F(t,z) := \mathrm{e}^{\varepsilon t}\|z\|_H^2$, where $\varepsilon = 2\alpha - L_\sigma - L_\gamma$ is given by (3.13):

$$\begin{aligned}X_t^n - Y_t^n = &R_n(X_0 - Y_0) + \int_0^t \{A(X_s^n - Y_s^n) + R_n(F(X_s^n) - F(Y_s^n))\}\,ds \\ &+ \int_0^t R_n(\sigma(X_s^n) - \sigma(Y_s^n))dW_s + \int_0^t\int_E R_n(\gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu))\widetilde{N}(ds,d\nu).\end{aligned}$$

Observe that, by condition (A1) and (3.14), one has

$$\begin{aligned}&\int_0^t\int_E \|R_n(\gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu))\|_H^2\mu(d\nu)ds \\ &\qquad + \int_0^t\int_E \|R_n(\gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu))\|_H^4\mu(d\nu)ds \\ &\leq (1+\lambda)^2\int_0^t\int_E \|\gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu)\|_H^2\mu(d\nu)ds\end{aligned}$$

$$
\begin{aligned}
&+ 8(1+\lambda)^4 \int_0^t \int_E \left(\|\gamma(X_s^n,\nu)\|_H^4 + \|\gamma(Y_s^n,\nu)\|_H^4\right)\mu(d\nu)ds \\
&\leq L_\gamma(1+\lambda)^2 \int_0^t \|X_s^n - Y_s^n\|_H^2 ds \\
&\qquad + 16(1+\lambda)^4 t \sup_{z\in H}\int_E \|\gamma(z,\nu)\|_H^4 \mu(d\nu) < \infty.
\end{aligned}
$$

Thus we can apply the generalized Itô-formula from Theorem A.2 and obtain (similar to (3.5) in [2])

$$
\begin{aligned}
&\mathrm{e}^{\varepsilon t}\|X_t^n - Y_t^n\|_H^2 - \|R_n(X_0 - Y_0)\|_H^2 \\
&= \int_0^t \langle 2\mathrm{e}^{\varepsilon s}(X_s^n - Y_t^n), R_n(\sigma(X_s^n) - \sigma(Y_s^n))dW_s\rangle_H \\
&+ \int_0^t \mathrm{e}^{\varepsilon s}\left[\varepsilon\|X_s^n - Y_s^n\|_H^2 + \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n)\right] ds \\
&+ \int_0^t \int_E \mathrm{e}^{\varepsilon s}\left[\|X_s^n - Y_s^n + R_n(\gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu))\|_H^2 - \|X_s^n - Y_s^n\|_H^2\right]\widetilde{N}(ds,d\nu),
\end{aligned}
\tag{3.16}
$$

where we used, for $z, w \in D(A)$, the notation

$$
\begin{aligned}
\mathcal{L}_n(\|\cdot\|_H^2)(z,w) &:= 2\langle z - w, A(z-w) + R_n(F(z) - F(w))\rangle_H + \|R_n(\sigma(z) - \sigma(w))\|_{L_2^0(H)}^2 \\
&\quad + \int_E \|R_n(\gamma(z,\nu) - \gamma(w,\nu))\|_H^2 \mu(d\nu).
\end{aligned}
$$

Taking expectations in (3.16) yields

$$
\begin{aligned}
&\mathrm{e}^{\varepsilon t}\mathbb{E}\left[\|X_t^n - Y_t^n\|_H^2\right] - \mathbb{E}\left[\|R_n(X_0 - Y_0)\|_H^2\right] \\
&= \mathbb{E}\left[\int_0^t \mathrm{e}^{\varepsilon s}\left(\varepsilon\|X_s^n - Y_s^n\|_H^2 + \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n)\right) ds\right].
\end{aligned}
\tag{3.17}
$$

Lemma 3.4 yields

$$
\begin{aligned}
&\mathrm{e}^{\varepsilon t}\mathbb{E}\left[\|X_t^n - Y_t^n\|_H^2\right] - \mathbb{E}\left[\|R_n(x-y)\|_H^2\right] - 2(\alpha+\beta)\int_0^t \mathrm{e}^{\varepsilon s}\mathbb{E}\left[\|P_1 X_s^n - P_1 Y_s^n\|_H^2\right] ds \\
&\leq \mathbb{E}\left[\int_0^t \mathrm{e}^{\varepsilon s}(-\mathcal{L}(\|\cdot\|_H^2)(X_s^n, Y_s^n) + \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n))ds\right].
\end{aligned}
$$

Below we prove that the right-hand-side tends to zero as $n \to \infty$, which would imply the assertion of this theorem. To prove the desired convergence to zero we apply the generalized Lebesgue Theorem (see [28, Theorem 7.1.8]). For this reason we have to prove that

$$
\mathcal{L}(\|\cdot\|_H^2)(X_s^n, Y_s^n) - \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n) \to 0 \tag{3.18}
$$

holds a.s. for each $s > 0$ as $n \to \infty$ and, moreover, there exists a constant $C > 0$ such that

$$
|\mathcal{L}(\|\cdot\|_H^2)(X_s^n, Y_s^n) - \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n)| \leq C\|X_s^n - Y_s^n\|_H^2. \tag{3.19}
$$

We start with the proof of (3.18). Denote $F_s^n := F(X_s^n) - F(Y_s^n)$, $\sigma_s^n := \sigma(X_s^n) - \sigma(Y_s^n)$ and $\gamma_s^n(\nu) := \gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu)$ and analogously $F_s := F(X_s) - F(Y_s)$, $\sigma_s := \sigma(X_s) - \sigma(Y_s)$ and $\gamma_s(\nu) := \gamma(X_s,\nu) - \gamma(Y_s,\nu)$ for each $n \in \mathbb{N}$, $s \geq 0$ and $\nu \in E$. Then

$$
|(\mathcal{L}(\|\cdot\|_H^2)(X_s^n, Y_s^n) - \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n))|
$$

$$\begin{aligned}&\leq 2|\langle X_s^n - Y_s^n, F_s^n - R_nF_s^n\rangle_H| + |\|\sigma_s^n\|_{L_2^0}^2 - \|R_n\sigma_s^n\|_{L_2^0}^2| \\ &+ \left|\int_E \|\gamma_s^n(\nu)\|_H^2 - \|R_n\gamma_s^n(\nu)\|_H^2\mu(d\nu)\right| \\ &=: I_1 + I_2 + I_3.\end{aligned}$$

For the first term I_1 we estimate

$$\begin{aligned}I_1 &\leq 2\|X_s^n - Y_s^n\|_H\|F_s^n - R_nF_s^n\|_H \\ &\leq 2\|X_s^n - Y_s^n\|_H\left(\|F_s^n - F_s\|_H + \|F_s - R_nF_s\|_H + \|R_nF_s - R_nF_s^n\|_H\right) \\ &\leq 2\|X_s^n - Y_s^n\|_H\left(\|F_s^n - F_s\|_H + \|F_s - R_nF_s\|_H + (1+\lambda)\|F_s - F_s^n\|_H\right).\end{aligned}$$

Using that $X_s^n \to X_s$ and $Y_s^n \to Y_s$ as a.s. for some subsequence (also denoted by n), we easily find that the right-hand side tends to zero. The convergence of the second term follows from

$$\begin{aligned}I_2 &= \left|\|\sigma_s^n\|_{L_2^0} - \|R_n\sigma_s^n\|_{L_2^0}\right|\left(\|\sigma_s^n\|_{L_2^0} + \|R_n\sigma_s^n\|_{L_2^0}\right) \\ &\leq (2+\lambda)\sqrt{L_\sigma}\|\sigma_s^n - R_n\sigma_s^n\|_{L_2^0}\|X_s^n - Y_s^n\|_H \\ &\leq (2+\lambda)^2\sqrt{L_\sigma}\|X_s^n - Y_s^n\|_H\left(\|\sigma_s^n - \sigma_s\|_{L_2^0} + \|\sigma_s - R_n\sigma_s\|_{L_2^0} + \|\sigma_s - \sigma_s^n\|_{L_2^0}\right).\end{aligned}$$

It remains to show the convergence of the third term. First, observe

$$\begin{aligned}I_3 &\leq (2+\lambda)\int_E \|\gamma_s^n(\nu) - R_n\gamma_s^n(\nu)\|_H\|\gamma_s^n(\nu)\|_H\mu(d\nu) \\ &\leq (2+\lambda)\int_E \Big(\|\gamma_s^n(\nu) - \gamma_s(\nu)\|_H + \|\gamma_s(\nu) - R_n\gamma_s(\nu)\|_H \\ &\qquad\qquad + \|R_n\gamma_s(\nu) - R_n\gamma_s^n(\nu)\|_H\Big)\|\gamma_s^n(\nu)\|_H\mu(d\nu) \\ &\leq (2+\lambda)\left(\int_E \|\gamma_s^n(\nu)\|_H^2\mu(d\nu)\right)^{\frac{1}{2}}\left[\left(\int_E \|\gamma_s^n(\nu) - \gamma_s(\nu)\|_H^2\mu(d\nu)\right)^{\frac{1}{2}}\right. \\ &\qquad + \left.\left(\int_E \|\gamma_s(\nu) - R_n\gamma_s(\nu)\|_H^2\mu(d\nu)\right)^{\frac{1}{2}} + \left(\int_E \|R_n\gamma_s(\nu) - R_n\gamma_s^n(\nu)\|_H^2\mu(d\nu)\right)^{\frac{1}{2}}\right] \\ &\leq \sqrt{2}(2+\lambda)^2L_\gamma\|X_s^n - Y_s^n\|_H\left(\|X_s^n - X_s\|_H + \|Y_s^n - Y_s\|_H\right) \\ &\quad + (2+\lambda)\sqrt{L_\gamma}\|X_s^n - Y_s^n\|_H\left(\int_E \|\gamma_s(\nu) - R_n\gamma_s(\nu)\|_H^2\mu(d\nu)\right)^{\frac{1}{2}} \\ &= I_3^1 + I_3^2\end{aligned}$$

where the last inequality follows from condition (A1) combined with the inequality

$$\begin{aligned}&\|R_n\gamma_s(\nu) - R_n\gamma_s^n(\nu)\|_H^2 \\ &\leq (1+\lambda)^2\|\gamma_s(\nu) - \gamma_s^n(\nu)\|_H^2 \\ &\leq 2(1+\lambda)^2\left(\|\gamma(X_s,\nu) - \gamma(Y_s,\nu)\|_H^2 + \|\gamma(X_s^n,\nu) - \gamma(Y_s^n,\nu)\|_H^2\right).\end{aligned}$$

The first expression I_1^1 clearly tends to zero as $n \to \infty$. For the second expression I_3^2 we use the inequality $\|\gamma_s(\nu) - R_n\gamma_s(\nu)\|_H^2 \leq 2(2+\lambda)^2\|\gamma_s(\nu)\|_H^2$ so that dominated convergence theorem is applicable, which shows that $I_3^2 \to 0$ as $n \to \infty$ a.s.. This proves (3.18). Concerning (3.19), we find that

$$|(\mathcal{L}(\|\cdot\|_H^2)(X_s^n, Y_s^n) - \mathcal{L}_n(\|\cdot\|_H^2)(X_s^n, Y_s^n))|$$

$$\begin{aligned}
&\leq 2|\langle X_s^n - Y_s^n, F_s^n - R_n F_s^n\rangle_H| + |\|\sigma_s^n\|^2_{L_2^0(H)} - \|R_n\sigma_s^n\|^2_{L_2^0(H)}| \\
&\quad + \left|\int_E \|\gamma_s^n(\nu)\|_H^2 - \|R_n\gamma_s^n(\nu)\|_H^2\mu(d\nu)\right| \\
&\leq 2(2+\lambda)\|X_s^n - Y_s^n\|_H\|F_s^n\|_H + \left(1+(1+\lambda)^2\right)\left[\|\sigma_s^n\|^2_{L_2^0(H)} + \int_E \|\gamma_s^n(\nu)\|_H^2\mu(d\nu)\right] \\
&\leq 2(2+\lambda)L_F\|X_s^n - Y_s^n\|_H^2 + \left(1+(1+\lambda)^2\right)(L_\sigma + L_\gamma)\|X_s^n - Y_s^n\|_H^2.
\end{aligned}$$

Hence the generalized Lebesgue Theorem is applicable, and thus the assertion of this theorem is proved. □

Note that condition (3.14) is used to guarantee that the Itô-formula A.2 for Hilbert space valued jump diffusions can be applied for $(x,t) \to \mathrm{e}^{t\varepsilon}\|x\|_H^2$. The assertion of Proposition 3.5 is also true when $\varepsilon \leq 0$, but will be only applied for the case when $\varepsilon > 0$.

4. Convergence to limiting distribution

4.1. **The strongly dissipative case.** As a consequence of our key stability estimate we can provide a simple proof for the existence and uniqueneness of a unique limiting distribution.

Theorem 4.1. *Assume that conditions (GDC), (A1), and* (3.14) *are satisfied. Suppose that*

$$\delta := \varepsilon + 2(\alpha+\beta) = 2\beta - L_\sigma - L_\gamma > 0 \tag{4.1}$$

Then there exists a constant $C > 0$ such that

$$\mathrm{W}_2(p_t^*\rho, p_t^*\widetilde{\rho}) \leq C\mathrm{W}_2(\rho,\widetilde{\rho})e^{-\delta/2t}, \qquad t \geq 0,$$

for any $\rho, \widetilde{\rho} \in \mathcal{P}_2(H)$. In particular, the Markov process determined by (3.1) *has a unique invariant measure π. This measure has finite second moments and it holds that*

$$\mathrm{W}_2(p_t^*\rho, \pi) \leq C\mathrm{W}_2(\rho,\pi)e^{-\delta/2t}, \qquad t \geq 0,$$

for each $\rho \in \mathcal{P}_2(H)$.

Proof. Since the proof is rather standard we give only a sketch of proof. Namely, for given $x, y \in H$ we find by Proposition 3.5

$$\mathbb{E}[\|X_t^x - X_t^y\|_H^2] \leq e^{-\varepsilon t}\|x-y\|_H^2 + 2(\alpha+\beta)\int_0^t e^{-\varepsilon(t-s)}\mathbb{E}[\|X_s^x - X_s^y\|_H^2]ds.$$

from which we readily deduce that

$$\mathbb{E}[\|X_t^x - X_t^y\|_H^2] \leq \|x-y\|^2 - \delta\int_0^t \mathbb{E}[\|X_s^x - X_s^y\|_H^2]ds.$$

This implies that

$$\mathbb{E}[\|X_t^x - X_t^y\|_H^2] \leq \|x-y\|^2 e^{-\delta t}, \qquad t \geq 0.$$

The assertion can be now deduced by standard arguments. □

The condition $\delta > 0$ requires that the drift is strong enough. It can be seen as an analogue of the conditions introduced in [32, Section 16], [10, Chapter 11, Section 6], and [36], where a similar statement was derived.

Opposite to this case, in this work we focus on the study of multiple invariant measures. For this purpose we will assume that $\varepsilon > 0$ which is weaker than condition (4.1).

4.2. **The Case of Vanishing Coefficients.** While Proposition 3.5 provides an estimate on the L^2-norm of the difference $X_t^x - X_t^y$, such an estimate alone does neither imply the existence nor uniqueness of an invariant distribution. However, if the coefficients F, σ, γ vanish at H_1, then we may characterize the limiting distributions in L^2.

Theorem 4.2. *Suppose that (GDC) holds with a projection operator P_1, (A1),* (3.14), (3.13) *are satisfied, that $(S(t))_{t\geq 0}$ leaves $H_0 := \operatorname{ran}(I-P_1)$ invariant, and that* $\operatorname{ran}(P_1) \subset \ker(A)$. *Moreover, assume that*

$$P_1 F \equiv 0, \ \ P_1\sigma \equiv 0, \ \ P_1\gamma \equiv 0. \tag{4.2}$$

Given any $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ which satisfies

$$F(P_1 x) = 0, \ \ \sigma(P_1 x) = 0, \ \ \gamma(P_1 x, \cdot) = 0, \qquad a.s., \tag{4.3}$$

then the inequality

$$\mathbb{E}\left[\|X_t - P_1 X_0\|_H^2\right] \leq \mathrm{e}^{-\varepsilon t}\mathbb{E}\left[\|(I-P_1)X_0\|_H^2\right]$$

holds. In particular, let ρ be the law of $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ and ρ_1 be the law of $P_1 X_0$, respectively. Then ρ_1 is an invariant measure.

Proof. Fix $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ with property (4.3) and set $P_0 = I - P_1$. Since $\operatorname{ran}(P_1) \subset \ker(A)$ we find that $S(t)P_1 = P_1$ for $t \geq 0$ and hence $P_0 S(t) P_1 = 0$. Moreover, since $(S(t))_{t\geq 0}$ leaves H_0 invariant, we obtain $P_0 S(t) = P_0 S(t) P_0 + P_0 S(t) P_1 = P_0 S(t) P_0 = S(t) P_0$. Hence, using (4.2) we find that

$$P_1 X_t = P_1 S(t) X_0 = P_1 S(t) P_0 X_0 + P_1 S(t) P_1 X_0 = P_1 X_0.$$

From this we conclude that $(P_0 X_t)_{t\geq 0}$ satisfies

$$\begin{aligned}
P_0 X_t = P_0 S(t) X_0 &+ \int_0^t P_0 S(t-s) F(X_s) ds + \int_0^t P_0 S(t-s)\sigma(X_s) dW_s \\
&+ \int_0^t \int_E P_0 S(t-s)\gamma(X_s)\widetilde{N}(ds, d\nu) \\
= S(t) P_0 X_0 &+ \int_0^t S(t-s) P_0 F(P_1 X_0 + P_0 X_t) ds + \int_0^t S(t-s) P_0 \sigma(P_1 X_0 + P_0 X_s) dW_s \\
&+ \int_0^t \int_E S(t-s) P_0 \gamma(P_1 X_0 + P_0 X_s)\widetilde{N}(ds, d\nu) \\
= S(t) P_0 X_0 &+ \int_0^t S(t-s)\widetilde{F}(P_0 X_t) ds + \int_0^t S(t-s)\widetilde{\sigma}(P_0 X_s) dW_s \\
&\int_0^t \int_E S(t-s)\widetilde{\gamma}(P_0 X_s)\widetilde{N}(ds, d\nu),
\end{aligned}$$

where we have set $\widetilde{F}(y) := P_0 F(P_1 X_0 + y)$, $\widetilde{\sigma}(y) := P_0\sigma(P_1 X_0 + y)$ and $\widetilde{\gamma}(y,\nu) := P_0\gamma(P_1 X_0 + y, \nu)$ for all $y \in H_0$ and $\nu \in E$. Since these coefficients share the same Lipschitz estimates as F, σ and γ, are $\mathcal{F}_0$-measurable and the noise terms are independent of $\mathcal{F}_0$, we can apply Proposition 3.5 (conditionally on $\mathcal{F}_0$) to the process $(P_0 X_t)_{t\geq 0}$ obtained from the above auxiliary SPDE and obtain

$$\mathbb{E}[\|X_t - P_1 X_0\|_H^2] = \mathbb{E}[\|P_0 X_t\|_H^2] = \mathbb{E}[\|P_0 X_t - P_0 Y_t^0\|_H^2] \leq \mathrm{e}^{-\varepsilon t}\mathbb{E}[\|P_0 X_0\|_H^2],$$

where we have used that $P_0 Y_t = 0$ for the unique solution with $Y_0 = 0$ due to (4.3). $\square$

This theorem can be applied, for instance, to the Heath-Jarrow-Morton-Musiela equation, see Section 5.

4.3. **Main result: The General Case.** In Theorem 4.2 we have assumed (4.2), (4.3), and that $(S(t))_{t\geq 0}$ leaves H_0 invariant. Below we continue with the more general case. Namely, for the projection operator P_1 given by condition (GDC) we set $P_0 = I - P_1$ and suppose that:

(A2) The semigroup $(S(t))_{t\geq 0}$ leaves $H_1 := \mathrm{ran}(I - P_0)$ invariant, one has

$$P_1\sigma = P_1\gamma = 0 \quad \text{and} \quad P_1F(x) = P_1F(P_1x), \qquad x \in H.$$

Let us briefly comment on this condition. Let $(X_t)_{t\geq 0}$ be the unique solution to (3.6) and decompose the process X_t according to $X_t = P_0X_t + P_1X_t$. Then condition (A2) simply implies that P_1X_t is $\mathcal{F}_0$-measurable and satisfies ω-wisely the deterministic equation

$$f(t;x) = P_1S(t)x + \int_0^t P_1S(t-s)P_1F(P_1f(s;x))ds, \quad f(0,x) = x \in H, \tag{4.4}$$

i.e. $P_1X_t = f(t;x)$ with $f(0,x) = x = X_0$ holds a.s. Our next condition imposes a control on this component:

(A3) For each $x \in H_1 = \mathrm{ran}(P_1)$ there exists $\widetilde{f}(x) \in H_1$ and constants $C(x) > 0$, $\delta(x) > 0$ such that

$$\|f(t;x) - \widetilde{f}(x)\|_H^2 \leq C(x)\mathrm{e}^{-\delta(x)t}, \qquad t \geq 0.$$

Without loss of generality we will always suppose that $\delta(x) \in (0, |\varepsilon|)$. Such assumption will simplify our arguments later on. Note that, if $P_1F(P_1\cdot) = 0$ then condition (A3) reduces to a condition on the limiting behavior of the semigroup $(S(t))_{t\geq 0}$ when restricted to $H_1 = \mathrm{ran}(P_1)$. In such a case condition (A3) is, for instance, satisfied if $\mathrm{ran}(P_1) \subset \ker(A)$. Recall that condition (GDC) was formulated in the introduction and that (A1), (3.14) and (3.13) were formulated in Section 3. The following is our main result in this Section.

Theorem 4.3. *Suppose that condition (GDC) holds for some projection operator P_1, that conditions (A1) – (A3),* (3.14) *and* (3.13) *are satisfied. Then the following assertions hold:*

(a) For each $x \in H$ there exists an invariant measure $\pi_{\delta_x} \in \mathcal{P}_2(H)$ for the Markov semigroup $(p_t)_{t\geq 0}$ and a constant $K(\alpha, \beta, \varepsilon, h) > 0$ such that

$$\mathrm{W}_2(p_t(x,\cdot), \pi_{\delta_x}) \leq K(\alpha, \beta, \varepsilon, x)\mathrm{e}^{-\frac{\delta(x)}{2}t}, \qquad t \geq 0.$$

(b) Suppose, in addition to the conditions of (A3), that there are constants δ and C, such that

$$\delta(x) \geq \delta > 0 \quad \text{and} \quad C(x) \leq C(1 + \|x\|_H)^4, \qquad x \in H. \tag{4.5}$$

Then, for each $\rho \in \mathcal{P}_2(H)$, there exists an invariant measure $\pi_\rho \in \mathcal{P}_2(H)$ for the Markov semigroup $(p_t)_{t\geq 0}$ and a constant $K(\alpha, \beta, \varepsilon) > 0$ such that

$$\mathrm{W}_2(p_t^*\rho, \pi_\rho) \leq K(\alpha, \beta, \varepsilon) \int_H (1 + \|x\|_H)^2 \rho(dx)\mathrm{e}^{-\frac{\delta}{2}t}, \qquad t \geq 0.$$

The proof of this theorem relies on the key stability estimate formulated in Proposition 3.5 and is given at the end of this section. So far we have stated the existence of invariant measures parametrized by the initial state of the process. However, under the given conditions it can also be shown that π_{δ_x} as well as π_ρ depend only on the H_1 part of x or ρ, respectively.

Corollary 4.4. *Suppose that condition (GDC) holds for some projection operator P_1, that conditions (A1) – (A3),* (3.14) *and* (3.13) *are satisfied. Then the following assertions hold:*

(a) Let $x, y \in H$ be such that $P_1 x = P_1 y$. Then $\pi_{\delta_x} = \pi_{\delta_y}$.

(b) Suppose, in addition, that (4.5) *holds. Let $\rho, \widetilde{\rho} \in \mathcal{P}_2(H)$ be such that $\rho \circ P_1^{-1} = \widetilde{\rho} \circ P_1^{-1}$. Then $\pi_\rho = \pi_{\widetilde{\rho}}$.*

Let us briefly compare the conditions imposed in Theorem 4.2 with those imposed in Theorem 4.3. In Theorem 4.3 we have weakened (4.2) with respect to F by replacing $P_1 F = 0$ by $P_1 F(x) = P_1 F(P_1 x)$. Moreover, we have replaced $\mathrm{ran}(P_1) \subset \ker(A)$ by condition (A3). Finally note that condition (4.3) is not assumed in Theorem 4.3. Below we provide a counter example showing that, in general, condition (A3) cannot be omitted.

Example 4.5. *Let $H = \mathbb{R}^2$ and $(W_t)_{t\geq 0}$ be a 2-dimensional standard Brownian motion. Let $Y_t = (Y_t^1, Y_t^2) \in H = \mathbb{R}^2$ be the solution of*

$$dY_t = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} Y_t dt + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} dW_t.$$

Then condition (A1) holds for $F = 0$, $\gamma = 0$ and clearly $\sigma(x) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Example 3.3 shows that (GDC) holds with P_1 being the projection onto the second coordinate. Moreover, (4.2) *and hence (A2) holds. However, since*

$$Y_t^2 = \mathrm{e}^t Y_0^2 + \int_0^t \mathrm{e}^{t-s} dW_s^2$$

it is clear that condition (A3) is not satisfied. Moreover, Y_t^2 does not have a limiting distribution and hence also Y_t cannot have a limiting distribution.

Next we turn to a proof of Theorem 4.3 and Corollary 4.4.

4.4. **Construction of a coupling.** Let $x \in H$ and let $(X_t^x)_{t\geq 0}$ be the unique mild solution to (3.6). Below we construct for given $\tau \geq 0$ a coupling for the law of $(X_t^x, X_{t+\tau}^x)$. Let $(Y_t^{x,\tau})_{t\geq 0}$ be the unique mild solution to the SPDE

$$\begin{aligned} Y_t^{x,\tau} &= S(t)x + \int_0^t S(t-s)F(Y_s^{x,\tau})ds + \int_0^t S(t-s)\sigma(Y_s^{x,\tau})dW_s^\tau \qquad (4.6) \\ &+ \int_0^t \int_E S(t-s)\gamma(Y_s^{x,\tau}, \nu)\widetilde{N}^\tau(ds, d\nu), \qquad t \geq 0, \end{aligned}$$

where $W_s^\tau = W_{\tau+s} - W_\tau$ is a Q-Wiener process, and $\widetilde{N}^\tau(ds, d\nu)$ defined by

$$\widetilde{N}^\tau((0,t] \times A) := \widetilde{N}((\tau, \tau + t] \times A)$$

for $t > 0$ and $A \in \mathcal{E}$ is a Poisson random measure with respect to the filtration $(\mathcal{F}_s^\tau)_{s\geq 0}$ defined by $\mathcal{F}_s^\tau = \mathcal{F}_{s+\tau}$.

Lemma 4.6. *Suppose that (GDC), (A1),* (3.14) *and* (3.13) *are satisfied. Then for each $x \in H$ and $t, \tau \geq 0$ the following assertions hold:*

(a) $Y_t^{x,\tau}$ has the same law as X_t^x.

(b) It holds that

$$\mathbb{E}\left[\|Y_t^{x,\tau} - X_{t+\tau}^x\|_H^2\right] \leq \mathrm{e}^{-\varepsilon t}\mathbb{E}\left[\|x - X_\tau^x\|_H^2\right]$$

$$+ 2(\alpha+\beta)\int_0^t \mathrm{e}^{-\varepsilon(t-s)}\mathbb{E}\left[\|P_1Y_s^{x,\tau} - P_1X_{s+\tau}^x\|_H^2\right] ds.$$

Proof. (a) Since (3.6) has a unique solution it follows from the Yamada-Watanabe Theorem (see [26]) that also uniqueness in law holds for this equation. Since the driving noises N^τ and W^τ in (4.6) have the same law as N and W from (3.6), it follows that the unique solution to (4.6) has the same law as the solution to (3.6). This proves the assertion.

(b) Set $X_t^{x,\tau} := X_{t+\tau}^x$, then by direct computation we find that

$$\begin{aligned}
X_t^{x,\tau} &= S(t)S(\tau)x + \int_0^{t+\tau} S(t+\tau-s)F(X_s^x)ds + \int_0^{t+\tau} S(t+\tau-s)\sigma(X_s^x)dW_s \\
&\quad + \int_0^{t+\tau}\int_E S(t+\tau-s)\gamma(X_s^x,\nu)\widetilde{N}(ds,d\nu) \\
&= S(t)S(\tau)x + S(t)\int_0^{\tau} S(\tau-s)F(X_s^x)ds + S(t)\int_0^{\tau} S(\tau-s)\sigma(X_s^x)dW_s \\
&\quad + S(t)\int_0^{\tau}\int_E S(\tau-s)\gamma(X_s^x,\nu)\widetilde{N}(ds,d\nu) \\
&\quad + \int_\tau^{t+\tau} S(t+\tau-s)F(X_s^x)ds + \int_\tau^{t+\tau} S(t+\tau-s)\sigma(X_s^x)dW_s \\
&\quad + \int_\tau^{t+\tau}\int_E S(t+\tau-s)\gamma(X_s^x,\nu)\widetilde{N}(ds,d\nu) \\
&= S(t)X_0^{x,\tau} + \int_0^t S(t-s)F(X_s^{x,\tau})ds + \int_0^t S(t-s)\sigma(X_s^{x,\tau})dW_s^\tau \\
&\quad + \int_0^t\int_E S(t-s)\gamma(X_s^{x,\tau},\nu)\widetilde{N}^\tau(ds,d\nu),
\end{aligned}$$

where in the last equality we have used, for appropriate integrands $\Phi(s,\nu)$ and $\Psi(s)$, that

$$\int_\tau^{\tau+t}\Psi(s)dW_s = \int_0^t \Psi(s+\tau)dW_s^\tau,$$
$$\int_\tau^{\tau+t}\int_E \Phi(s,\nu)\widetilde{N}(ds,d\nu) = \int_0^t\int_E \Phi(s+\tau,\nu)\widetilde{N}^\tau(ds,d\nu).$$

Hence $(X_t^{x,\tau})_{t\geq 0}$ also solves (4.6) with $\mathcal{F}_0^\tau = \mathcal{F}_\tau$ and initial condition $X_0^{x,\tau} = X_\tau^x$. Consequently, the assertion follows from Proposition 3.5 applied to $X_t^{x,\tau}$ and $Y_t^{x,\tau}$. □

4.5. **Proof of Theorem 4.3.**

Proof of Theorem 4.3. Fix $x \in H$ and recall that $p_t(x,\cdot)$ denotes the transition probabilities of the Markov process obtained from (3.6). Below we prove that $(p_t(x,\cdot))_{t\geq 0} \subset \mathcal{P}_2(H)$ is a Cauchy sequence with respect to the Wasserstein distance W_2. Fix $t,\tau \geq 0$. We treat the cases $\tau \in (0,1]$ and $\tau > 1$ separately.

Case $0 < \tau \leq 1$: Then using the coupling lemma 4.6.(b) yields

$$\begin{aligned}
\mathrm{W}_2(p_{t+\tau}(x,\cdot),p_t(x,\cdot)) &\leq \left(\mathbb{E}\left[\|Y_t^{x,\tau} - X_{t+\tau}^x\|_H^2\right]\right)^{1/2} \\
&\leq \mathrm{e}^{-\frac{\varepsilon}{2}t}\left(\mathbb{E}\left[\|X_\tau^x - x\|_H^2\right]\right)^{1/2} \\
&\quad + \sqrt{2(\alpha+\beta)}\left(\int_0^t \mathrm{e}^{-\varepsilon(t-s)}\mathbb{E}\left[\|P_1Y_s^{x,\tau} - P_1X_{s+\tau}^x\|_H^2\right] ds\right)^{1/2}
\end{aligned}$$

$$=: I_1 + I_2.$$

The first term I_1 can be estimated by

$$I_1 \le \mathrm{e}^{-\frac{\varepsilon}{2}t} \sup_{s\in[0,1]} \left(\mathbb{E}\left[\|X_s^x - x\|_H^2\right]\right)^{1/2}.$$

To estimate the second term I_2 we first observe that by condition (A2) we have $P_1 Y_s^{x,\tau} = P_1 X_s^x = f(s;x)$ being deterministic and hence by condition (A3) one has for each $s \ge 0$ that

$$\begin{aligned}\mathbb{E}\left[\|P_1 Y_s^{x,\tau} - P_1 X_{s+\tau}^x\|_H^2\right] &\le 2\|P_1 Y_s^{x,\tau} - \widetilde{f}(x)\|_H^2 + 2\|P_1 X_{s+\tau}^x - \widetilde{X}_\infty^x\|_H^2 \\ &\le 4C(x)\mathrm{e}^{-\delta(x)s}. \end{aligned} \tag{4.7}$$

This readily yields

$$\begin{aligned}&\int_0^t \mathrm{e}^{-\varepsilon(t-s)}\mathbb{E}\left[\|P_1 Y_s^{x,\tau} - P_1 X_{s+\tau}^x\|_H^2\right] ds \\ &\le 4C(x)\int_0^t \mathrm{e}^{-\varepsilon(t-s)}\mathrm{e}^{-\delta(x)s} ds \\ &= 4C(x)\mathrm{e}^{-\varepsilon t}\frac{\mathrm{e}^{(\varepsilon-\delta(x))t} - 1}{\varepsilon - \delta(x)} \\ &\le 4C(x)\frac{\mathrm{e}^{-\delta(x)t}}{\varepsilon - \delta(x)}. \end{aligned} \tag{4.8}$$

Inserting this into the definition of I_2 gives

$$I_2 \le 2\sqrt{\frac{(\alpha+\beta)C(x)}{\varepsilon - \delta(x)}}\mathrm{e}^{-\frac{\delta(x)}{2}t}.$$

Case $\tau > 1$: Fix some $N \in \mathbb{N}$ with $\tau < N < 2\tau$ and define a sequence of numbers $(a_n)_{n=0,\dots,N}$ by

$$a_n := \frac{\tau}{N}n, \qquad n = 0,\dots,N.$$

Then $a_0 = 0$, $a_N = \tau$ and $a_n - a_{n-1} = \frac{\tau}{N} =: \varkappa \in (\frac{1}{2}, 1)$ for $n = 1,\dots,N$. Hence we obtain from the coupling Lemma 4.6.(b)

$$\begin{aligned}&\mathrm{W}_2(p_{t+\tau}(x,\cdot), p_t(x,\cdot)) \\ &\le \sum_{n=1}^N \mathrm{W}_2(p_{t+a_n}(x,\cdot), p_{t+a_{n-1}}(x,\cdot)) \\ &\le \sum_{n=1}^N \left(\mathbb{E}\left[\|Y_{t+a_{n-1}}^{x,\varkappa} - X_{t+a_{n-1}+\varkappa}^x\|_H^2\right]\right)^{1/2} \\ &\le \sum_{n=1}^N \mathrm{e}^{-\frac{\varepsilon}{2}(t+a_{n-1})}\left(\mathbb{E}\left[\|X_\varkappa^x - x\|_H^2\right]\right)^{1/2} \\ &\quad + \sqrt{2(\alpha+\beta)}\sum_{n=1}^N \left(\int_0^{t+a_{n-1}} \mathrm{e}^{-\varepsilon(t+a_{n-1}-s)}\mathbb{E}\left[\|P_1 Y_s^{x,\varkappa} - P_1 X_{s+\varkappa}^x\|_H^2\right] ds\right)^{1/2} \\ &=: J_1 + J_2.\end{aligned}$$

For the first term J_1 we use $\varkappa > \frac{1}{2}$ so that

$$\sum_{n=1}^{N} \mathrm{e}^{-\frac{\varepsilon}{2}\varkappa(n-1)} \leq \sum_{n=0}^{\infty} \mathrm{e}^{-\frac{\varepsilon}{4}n} = \left(1 - \mathrm{e}^{-\frac{\varepsilon}{4}}\right)^{-1},$$

from which we obtain

$$\begin{aligned} J_1 &= \mathrm{e}^{-\frac{\varepsilon}{2}t} \sup_{s\in[0,1]} \left(\mathbb{E}[\|X_s^x - x\|_H^2]\right)^{\frac{1}{2}} \sum_{n=1}^{N} \mathrm{e}^{-\frac{\varepsilon}{2}\varkappa(n-1)} \\ &\leq \sup_{s\in[0,1]} \left(\mathbb{E}[\|X_s^x - x\|_H^2]\right)^{\frac{1}{2}} \left(1 - \mathrm{e}^{-\frac{\varepsilon}{4}}\right)^{-1} \mathrm{e}^{-\frac{\varepsilon}{2}t}. \end{aligned}$$

To estimate the second term J_2 we first observe that by condition (A2) we have $P_1 Y_s^{x,\tau} = P_1 X_s^x = f(s;x)$ being deterministic and hence by condition (A3), one has for $s \geq 0$

$$\begin{aligned} \mathbb{E}\left[\|P_1 Y_s^{x,\varkappa} - P_1 X_{s+\varkappa}^x\|_H^2\right] &\leq 2\|P_1 Y_s^{x,\varkappa} - \widetilde{f}(x)\|_H^2 + 2\|P_1 X_{s+\varkappa}^x - \widetilde{f}(x)\|_H^2 \\ &\leq 4C(x)\mathrm{e}^{-\delta(x)s}. \end{aligned}$$

Hence we find that

$$\begin{aligned} &\int_0^{t+a_{n-1}} \mathrm{e}^{-\varepsilon(t+a_{n-1}-s)} \mathbb{E}\left[\|P_1 Y_s^{x,\varkappa} - P_1 X_{s+\varkappa}^x\|_H^2\right] ds \\ &\leq 4C(x) \int_0^{t+a_{n-1}} \mathrm{e}^{-\varepsilon(t+a_{n-1}-s)} \mathrm{e}^{-\delta(x)s} ds \\ &= 4C(x)\mathrm{e}^{-\varepsilon(t+a_{n-1})} \frac{\mathrm{e}^{(\varepsilon-\delta(x))(t+a_{n-1})} - 1}{\varepsilon - \delta(x)} \\ &\leq 4C(x)\frac{\mathrm{e}^{-\delta(x)t}}{\varepsilon - \delta(x)} \mathrm{e}^{-\delta(x)a_{n-1}} \\ &\leq 4C(x)\frac{\mathrm{e}^{-\delta(x)t}}{\varepsilon - \delta(x)} \mathrm{e}^{-\frac{\delta(x)}{2}(n-1)} \end{aligned}$$

where the last inequality follows from $a_{n-1} = \varkappa(n-1) \geq \frac{1}{2}(n-1)$. From this we readily derive the estimate

$$J_2 \leq 2\sqrt{\frac{(\alpha+\beta)C(x)}{\varepsilon - \delta(x)}} \left(1 - \mathrm{e}^{-\frac{\delta(x)}{4}}\right)^{-1} \mathrm{e}^{-\frac{\delta(x)}{2}t}.$$

Hence, using also (3.5) we obtain

$$\mathrm{W}_2(p_{t+\tau}(x,\cdot), p_t(x,\cdot)) \leq K(\alpha,\beta,\varepsilon,x)\mathrm{e}^{-\frac{\delta(x)}{2}t}, \qquad t,\tau \geq 0, \tag{4.9}$$

where the constant $K(\alpha,\beta,\varepsilon,x) > 0$ is given by

$$K(\alpha,\beta,\varepsilon,x) = K(\varepsilon)(1 + \|x\|_H) + 2\sqrt{\frac{(\alpha+\beta)C(x)}{\varepsilon - \delta(x)}} \left(1 - \mathrm{e}^{-\frac{\delta(x)}{4}}\right)^{-1}$$

with another constant $K(\varepsilon) > 0$. This implies that, for each $x \in H$, $(p_t(x,\cdot))_{t\geq 0}$ has a limit in $\mathcal{P}_2(H)$. Denote this limit by π_{δ_x}. Assertion (a) now follows by taking the limit $\tau \to \infty$ in (4.9) and using the fact that $K(\alpha,\beta,\varepsilon,x)$ is independent of τ.

It remains to prove assertion (b). First observe that, using $\delta(x) \geq \delta > 0$ and $C(x) \leq C(1 + \|x\|_H)^4$, we have

$$K(\alpha,\beta,\varepsilon,x) \leq (1 + \|x\|_H)^2 \widetilde{K}(\alpha,\beta,\varepsilon)$$

for some constant $\widetilde{K}(\alpha,\beta,\varepsilon)$. Note that

$$p_t^*\rho(dy) = \int_H p_t(z,dy)\rho(dz) \quad \text{and} \quad p_{t+\tau}^*\rho(dy) = \int_H p_{t+\tau}(z,dy)\rho(dz).$$

Hence using first the convexity of the Wasserstein distance and then (4.9) we find that

$$\begin{aligned} \mathrm{W}_2(p_{t+\tau}^*\rho, p_t^*\rho) &\le \int_H \mathrm{W}_2(p_{t+\tau}(x,\cdot), p_t(x,\cdot))\rho(dx) \\ &\le \widetilde{K}(\alpha,\beta,\varepsilon)\int_H (1+\|x\|_H)^2\rho(dx)\cdot \mathrm{e}^{-\frac{\delta}{2}t}. \end{aligned}$$

Since $\rho \in \mathcal{P}_2(H)$, the assertion is proved. □

4.6. **Proof of Corollary 4.4.**

Proof of Corollary 4.4. Recall that, by condition (A2) the process $P_1X_t^x$ solves

$$P_1X_t^x = P_1S(t)P_1x + \int_0^t P_1S(t-s)F(P_1X_s^x)ds.$$

Since F is globally Lipschitz continuous by condition (A1), it follows that this equation has for each $x \in H$ a unique solution and is deterministic. From this we readily conclude that $P_1X_t^x = P_1X_t^y$ holds for all $t \ge 0$, provided that $P_1x = P_1y$. Hence Proposition 3.5 yields for such x, y

$$\mathbb{E}\left[\|X_t^x - X_t^y\|_H^2\right] \le \mathrm{e}^{-\varepsilon t}\|x-y\|_H^2, \qquad \forall t \ge 0. \tag{4.10}$$

Then for each $x, y \in H$ with $P_1x = P_1y$ and each $t \ge 0$ we obtain

$$\begin{aligned} \mathrm{W}_2(\pi_{\delta_x}, \pi_{\delta_y}) &\le \mathrm{W}_2(\pi_{\delta_x}, p_t(x,\cdot)) + \mathrm{W}_2(p_t(x,\cdot), p_t(y,\cdot)) + \mathrm{W}_2(p_t(y,\cdot), \pi_{\delta_y}) \\ &\le \mathrm{W}_2(\pi_{\delta_x}, p_t(x,\cdot)) + \mathrm{e}^{-\frac{\varepsilon}{2}t}\|x-y\|_H + \mathrm{W}_2(p_t(y,\cdot), \pi_{\delta_y}). \end{aligned}$$

Letting $t \to \infty$ yields $\pi_{\delta_x} = \pi_{\delta_y}$ and hence assertion (a) is proved.

To prove assertion (b), let $\rho, \widetilde{\rho} \in \mathcal{P}_2(H)$ be such that $\rho \circ P_1^{-1} = \widetilde{\rho} \circ P_1^{-1}$. Then

$$\mathrm{W}_2(\pi_\rho, \pi_{\widetilde{\rho}}) \le \mathrm{W}_2(\pi_\rho, p_t^*\rho) + \mathrm{W}_2(p_t^*\rho, p_t^*\widetilde{\rho}) + \mathrm{W}_2(p_t^*\widetilde{\rho}, \pi_{\widetilde{\rho}})$$

Again, by letting $t \to \infty$, it suffices to prove that

$$\limsup_{t\to\infty} \mathrm{W}_2(p_t^*\rho, p_t^*\widetilde{\rho}) = 0. \tag{4.11}$$

Let G be a coupling of $(\rho, \widetilde{\rho})$. Using the convexity of the Wasserstein distance and Proposition 3.5 gives

$$\begin{aligned} &\mathrm{W}_2(p_t^*\rho, p_t^*\widetilde{\rho}) \\ &\le \int_{H\times H} \mathrm{W}_2(p_t(x,\cdot), p_t(y,\cdot))G(dx,dy) \\ &\le \int_{H\times H} \left(\mathbb{E}\left[\|X_t^x - X_t^y\|_H^2\right]\right)^{1/2} G(dx,dy) \\ &\le \int_{H\times H} \mathrm{e}^{-\frac{\varepsilon}{2}t}\|x-y\|_H G(dx,dy) \\ &\quad + \sqrt{2(\alpha+\beta)}\int_{H\times H}\left(\int_0^t \mathrm{e}^{-\varepsilon(t-s)}\mathbb{E}\left[\|P_1X_s^x - P_1X_s^y\|_H^2\right]ds\right)^{1/2} G(dx,dy) \\ &=: I_1 + I_2. \end{aligned}$$

The first term I_1 satisfies

$$I_1 \leq \left(2 + \int_H \|x\|_H^2 \rho(dx) + \int_H \|y\|_H^2 \widetilde{\rho}(dy)\right) \mathrm{e}^{-\frac{\varepsilon}{2}t}.$$

For the second term we first use (A2) so that $P_1 X_s^x = P_1 X_s^{P_1 x}$, $P_1 X_s^y = P_1 X_s^{P_1 y}$ and hence we find for each $T > 0$ a constant $C(T) > 0$ such that for $t \in [0, T]$

$$\begin{aligned} I_2 &= \sqrt{2(\alpha+\beta)} \int_{H_1 \times H_1} \left(\int_0^t \mathrm{e}^{-\varepsilon(t-s)} \|P_1 X_s^x - P_1 X_s^y\|_H^2 ds \right)^{1/2} G(dx, dy) \\ &\leq C(T) \left(\int_{H \times H} \|P_1 x - P_1 y\|_H^2 G(dx, dy) \right)^{1/2}. \end{aligned}$$

Let us choose a particular coupling G as follows: By disintegration we write $\rho(dx) = \rho(x_1, dx_0)(\rho \circ P_1^{-1})(dx_1)$, $\widetilde{\rho}(dx) = \widetilde{\rho}(x_1, dx_0)(\widetilde{\rho} \circ P_1^{-1})(dx_1) = \widetilde{\rho}(x_1, dx_0)(\rho \circ P_1^{-1})(dx_1)$ where $\rho(x_1, dx_0)$, $\widetilde{\rho}(x_1, dx_0)$ are conditional probabilities defined on $\mathcal{B}(H_0)$ and we have used that $(\rho \circ P_1^{-1})(dx_1) = (\widetilde{\rho} \circ P_1^{-1})(dx_1)$. Then G is, for $A, B \in \mathcal{B}(H)$, given by

$$G(A \times B) := \int_{H \times H} \mathbb{1}_A(x_0, x_1) \mathbb{1}_B(y_0, y_1) \rho(x_1, dx_0) \widetilde{\rho}(y_1, dy_0) \widetilde{G}(dx_1, dy_1),$$

where $\widetilde{G}$ is a probability measure on H_1^2 given, for $A_1, B_1 \in \mathcal{B}(H_1)$, by

$$\widetilde{G}(A_1 \times B_1) = (\rho \circ P_1^{-1})(A_1 \cap B_1) = \rho\left(\{x \in H \mid P_1 x \in A_1 \cap B_1\}\right).$$

For this particular choice of G we find that

$$\begin{aligned} \int_{H \times} \|P_1 x - P_1 y\|_H^2 G(dx, dy) &= \int_{H_1 \times H_1} \int_{H_0^2} \|x_1 - y_1\|_H^2 \rho(x_1, dx_0) \widetilde{\rho}(y_1, dy_0) \widetilde{G}(dx_1, dy_1) \\ &= \int_{H_1 \times H_1} \|x_1 - y_1\|_H^2 \widetilde{G}(dx_1, dy_1) = 0 \end{aligned}$$

and hence $I_2 = 0$, since $\widetilde{G}$ is supported on the diagonal of $H_1 \times H_1$. This proves (4.11) and completes the proof. □

5. The Heath-Jarrow-Mortion-Musiela equation

The Heath-Jarrow-Morton-Musiela equation (HJMM-equation) describes the term structure of interest rates in terms of its forward rate dynamics modelled, for $\beta > 0$ fixed, on the separable Hilbert space of forward curves

$$H_\beta = \{h : \mathbb{R}_+ \to \mathbb{R} : h \text{ is absolutely continuous and } \|h\|_\beta < \infty\}, \tag{5.1}$$
$$\langle h, g \rangle_\beta = h(\infty)g(\infty) + \int_0^\infty h'(x) g'(x) \mathrm{e}^{\beta x} dx$$

with norm $\|h\|_\beta^2 = \langle h, h \rangle_\beta$. Such space was first motivated and introduced by Filipovic [15]. Note that $h(\infty) := \lim_{x \to \infty} h(x)$ exists, whenever $\int_0^\infty (h'(x))^2 \mathrm{e}^{\beta x} dx < \infty$. It is called the *long rate* of the forward curve h. The HJMM-equation on H_β is given by

$$\begin{cases} dX_t = (AX_t + F_{HJMM}(\sigma, \gamma)(X_t))\, dt + \sigma(X_t) dW_t + \int_E \gamma(X_t, \nu) \widetilde{N}(dt, d\nu), \\ X_0 = h_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H_\beta) \end{cases} \tag{5.2}$$

where $(W_t)_{t \geq 0}$ is a Q-Wiener process, $\widetilde{N}(dt, d\nu)$ is a compensated Poisson random measure on E with compensator $dt\mu(d\nu)$ as defined in Section 2 for $H := H_\beta$, and

(i) A is the infinitesimal generator of the shift semigroup $(S(t))_{t \in \mathbb{R}_+}$ on H_β, that is $S(t)h(x) := h(x+t)$ for all $t, x \geq 0$.

(ii) $h \mapsto \sigma(h)$ is a $\mathcal{B}(H_\beta)/\mathcal{B}(L_2^0)$-measurable mapping from H_β into $L_2^0(H_\beta)$ and $(h,\nu) \mapsto \gamma(h,\nu)$ is $\mathcal{B}(H_\beta) \otimes \mathcal{E}/\mathcal{B}(H_\beta)$-measurable mapping from $H_\beta \times E$ into H_β.

(iii) The drift is of the form

$$F_{HJMM}(\sigma,\gamma)(h) = \sum_{j\in\mathbb{N}} \sigma^j(h)\Sigma^j(h) - \int_E \gamma(h,\nu)\left(\mathrm{e}^{\Gamma(h,\nu)} - 1\right)\mu(d\nu),$$

with $\sigma^j(h) = \sqrt{\lambda_j}\sigma(h)e_j$,

$$\Sigma^j(h)(t) = \int_0^t \sigma^j(h)(s)ds \quad \text{and} \quad \Gamma(h,\nu)(t) = -\int_0^t \gamma(h,\nu)(s)ds.$$

The special form of the drift stems from mathematical finance and is sufficient for the absence of arbitrage opportunities. We denote the space of all forward rates with long rate equal to zero by

$$H_\beta^0 = \{h \in H_\beta : h(\infty) = 0\}.$$

For the construction of a unique mild solution to (5.2) the following conditions have been introduced in [11]:

(B1) $\sigma : H_\beta \to L_2^0(H_\beta^0)$, $\gamma : H_\beta \times E \to H_{\beta'}^0$ are Borel measurable for some $\beta' > \beta$.

(B2) There exists a function $\Phi : E \to \mathbb{R}_+$ such that $\Phi(\nu) \geq |\Gamma(h,\nu)(t)|$ for all $h \in H_\beta$, $\nu \in E$ and $t \geq 0$.

(B3) There is an $M \geq 0$ such that, for all $h \in H_\beta$, and some $\beta' > \beta$

$$\|\sigma(h)\|_{L_2^0(H_\beta)} \leq M, \qquad \int_E \mathrm{e}^{\Phi(\nu)} \max\{\|\gamma(h,\nu)\|_{\beta'}^2, \|\gamma(h,\nu)\|_{\beta'}^4\}\mu(d\nu) \leq M.$$

(B4) The function $F_2 : H_\beta \to H_\beta^0$ defined by

$$F_2(h) = -\int_E \gamma(h,\nu)\left(\mathrm{e}^{\Gamma(h,\nu)} - 1\right)\mu(d\nu)$$

has the weak derivative given by

$$\frac{d}{dx}F_2(h) = \int_E \gamma(h,\nu)^2 \mathrm{e}^{\Gamma(h,\nu)}\mu(d\nu) - \int_E \left(\frac{d}{dx}\gamma(h,\nu)\right)\left(\mathrm{e}^{\Gamma(h,\nu)} - 1\right)\mu(d\nu).$$

(B5) There are constants $L_\sigma, L_\gamma > 0$ such that, for all $h_1, h_2 \in H_\beta$, we have

$$\|\sigma(h_1) - \sigma(h_2)\|_{L_2^0(H_\beta)}^2 \leq L_\sigma \|h_1 - h_2\|_\beta^2,$$

$$\int_E \mathrm{e}^{\Phi(\nu)} \|\gamma(h_1,\nu) - \gamma(h_2,\nu)\|_{\beta'}^2 \mu(d\nu) \leq L_\gamma \|h_1 - h_2\|_\beta^2.$$

The following is the basic existence and uniqueness result for the Heath-Jarrow-Morton-Musiela equation (5.2).

Theorem 5.1. [11] *Suppose that conditions (B1) – (B5) are satisfied. Then* $F_{HJMM} : H_\beta \longrightarrow H_\beta^0$ *and there exists a constant* $L_F > 0$ *such that, for each* $h_1, h_2 \in H_\beta$,

$$\|F_{HJMM}(h_1) - F_{HJMM}(h_2)\|_\beta^2 \leq L_F \|h_1 - h_2\|_\beta^2. \tag{5.3}$$

This constant can be choosen as

$$L_F = \frac{\max(L_\sigma, L_\gamma)\sqrt{M}}{\beta}\left(\sqrt{6M\sqrt{2}} + \sqrt{\frac{8}{\beta^3} + \frac{16}{\beta}} + \sqrt{\frac{16(1+\frac{1}{\sqrt{\beta}})^2 + 48}{(\beta' - \beta)}}\right). \tag{5.4}$$

Moreover, for each initial condition $h \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H_\beta)$ *there is a unique adapted, cádlág mild solution* $(X_t)_{t\geq 0}$ *to (5.2).*

Proof. This result can be found essentially in [11], where the bound on L_F is an immediate result from its derivation. $\square$

Using the space of all functions with zero long rate we obtain the decomposition

$$H_\beta = H_\beta^0 \oplus \mathbb{R}, \qquad h = (h - h(\infty)) + h(\infty),$$

where $h(\infty) \in \mathbb{R}$ is identified with a constant function. Denote by

$$P_0 h = h - h(\infty) \ \text{ and } \ P_1 h = h(\infty)$$

the corresponding projections onto H_β^0 and $\mathbb{R}$, respectively. Such a decomposition of H_β was first used in [38] to study invariant measures for the HJMM-equation driven by a Q-Wiener process. An extension to the Lévy driven HJMM-equation was then obtained in [36]. The proof of the next theorem shows that the results of Section 4 imply the stability properties of the HJMM-equation as a particular case.

Theorem 5.2. *Suppose that conditions (B1) – (B5) are satisfied. If*

$$\beta > 2\sqrt{L_F} + L_\sigma + L_\gamma, \tag{5.5}$$

then for each initial distribution ρ on H_β with finite second moments there exists an invariant measure π_ρ and it holds that

$$\mathrm{W}_2(p_t^*\rho, \pi_\rho) \leq K \left(1 + \int_{H_\beta} \|h\|_{H_\beta}^2 \rho(dh) + \int_{H_\beta} \|h\|_{H_\beta}^2 \pi_\rho(dh) \right) \mathrm{e}^{-\frac{\beta - 2\sqrt{L_F} - L_\sigma - L_\gamma}{2} t} \tag{5.6}$$

for some constant $K = K(\beta, \sigma, \gamma) > 0$. Moreover, given $\rho, \widetilde{\rho}$ such that $\rho \circ P_1^{-1} = \widetilde{\rho} \circ P_1^{-1}$, then $\pi_\rho = \pi_{\widetilde{\rho}}$.

Proof. Observe that the assertion is an immediate consequence of Theorem 4.3 and Corollary 4.4. Below we briefly verify the assumptions given in these statements. Condition (A1) follows from (B1), (B5), and (5.3). The growth condition (3.14) is satisfied by (B3) and the fact that $\|\cdot\|_\beta \leq \|\cdot\|_{\beta'}$ for $\beta < \beta'$. It is not difficult to see that

$$\|S(t)h - P_1 h\|_\beta \leq \mathrm{e}^{-\frac{\beta}{2}t} \|h - P_1 h\|_\beta, \qquad t \geq 0$$

and that $(S(t))_{t\geq 0}$ leaves H_β^0 as well as $\mathbb{R} \subset H_\beta$ invariant. Hence Remark 3.2 yields that

$$\langle Ah, h \rangle \leq -\frac{\beta}{2}\|h\|_\beta^2 + \frac{\beta}{2}\|P_1 h\|_\beta^2, \qquad h \in D(A).$$

It follows from the considerations in Section 2 (see (3.10)) that (GDC) is satisfied for $\alpha = \frac{\beta}{2} - \sqrt{L_F}$. Consequently, $\varepsilon = \beta - 2\sqrt{L_F} - L_\sigma - L_\gamma$ and (3.13) holds due to (5.5). Since the coefficients map into H_β^0 and $S(t)P_1 h = h(\infty) = P_1 h$, conditions (A2), (A3) and (4.5) are trivially satisfied. The particular form of the estimate (5.6) follows from the proof of Theorem 4.3. $\square$

Comparing our result with [38, 36], we allow for a more general jump noise and prove convergence in the stronger Wasserstein distance with an exponential rate. Moreover, assuming that the volatilities map constant functions onto zero, i.e.

$$\sigma(c) \equiv 0, \quad \gamma(c, \nu) \equiv 0, \qquad \forall c \in \mathbb{R} \subset H_\beta, \ \nu \in E \tag{5.7}$$

shows that $F(c) \equiv 0$ and hence also (4.3) is satisfied. Hence we may apply Theorem 4.2 to characterize these invariant measures more explicitly.

Corollary 5.3. *Suppose that conditions (B1) – (B5) are satisfied, that* (5.5) *and* (5.7) *hold. Then*

$$\mathbb{E}\left[\|X_t - X_0(\infty)\|_\beta^2\right] \le \mathbb{E}\left[\|X_0 - X_0(\infty)\|_\beta^2\right] \mathrm{e}^{-\left(\beta - 2\sqrt{L_F} - L_\sigma - L_\gamma\right)t}$$

for each $X_0 \in L^2(\Omega, \mathbb{F}_0, \mathbb{P}, H)$.

We close this section by applying our results for the particular example discussed before also in [36].

Example 5.4. *Take*

$$\sigma^1(h)(x) := \int_x^\infty \min\left(\mathrm{e}^{-\beta y},\ |h'(y)|\right) dy$$

and $\sigma^j \equiv 0$ *for* $j \ge 2$. *Then*

$$\|\sigma(h)\|_{L_2^0}^2 = \|\sigma^1(h)\|_\beta^2 \le \int_0^\infty (\mathrm{e}^{-2\beta x})\mathrm{e}^{\beta x} dx = \frac{1}{\beta} =: M$$

and since $\min(a, b_1) - \min(a, b_2) \le |b_1 - b_2|$ *for* $a, b_1, b_2 \in \mathbb{R}_+$, *we also have*

$$\begin{aligned}\|\sigma(h_1) - \sigma(h_2)\|_{L_2^0}^2 &= \|\sigma^1(h) - \sigma^1(h_2)\|_\beta^2 \\ &= \int_0^\infty (\min(\mathrm{e}^{-\beta x}, |h_1'(x)|) - \min(\mathrm{e}^{-\beta x}, |h_2'(x)|))^2 \mathrm{e}^{\beta x} dx \\ &\le \int_0^\infty (h_1'(x) - h_2'(x))^2 \mathrm{e}^{\beta x} dx \\ &\le \|h_1 - h_2\|_\beta^2.\end{aligned}$$

Consequently, by taking $\gamma \equiv 0$, *the conditions (B1) – (B5) are satisfied with* $L_\sigma = 1$ *and* $L_\gamma = 0$ *and* $M = \frac{1}{\beta}$ *for the Lipschitz and growth constants. By* (5.4) *we get*

$$L_F = \frac{1}{\sqrt{\beta^3}}\left(\sqrt{\frac{6\sqrt{2}}{\beta}} + \sqrt{\frac{8}{\beta^3} + \frac{16}{\beta}} + \sqrt{\frac{16(1 + \frac{1}{\sqrt{\beta}})^2 + 48}{(\beta' - \beta)}}\right),$$

for all $\beta' > \beta$. *Choosing* $\beta \ge 3$ *and* $\beta' > \beta$ *large enough such that* $L_F < 1$, *we find that*

$$2\sqrt{L_F} + L_\sigma + L_\gamma < 3 = \beta,$$

i.e. (5.5) *is satisfied. It is clear that* $\sigma(c) \equiv 0$ *for each constant function* c. *Hence Corollary 5.3 is applicable.*

6. Stochastic partial differential equations with delay

6.1. **Description of the model.** Let H be a separable Hilbert space and $(W_t)_{t\ge0}$ a Q-Wiener process on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ with the usual conditions. Below we investigate invariant measures for the stochastic delay equation

$$\begin{cases} dX_t &= (AX_t + G(X_{t+\cdot}))\, dt + \sigma(X_t, X_{t+\cdot}) dW_t, \qquad t > 0 \\ X_0 &= \phi_0, X_{0+\cdot} = \phi, \end{cases} \tag{6.1}$$

where $\phi_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, $\phi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; L^2([-1,0]; H))$ and for $t \ge 1$ $X_{t+\cdot}$ denotes the past segment of the trajectory, i.e.

$$\begin{aligned} X_{t+\cdot} : \quad [-1, 0] &\longrightarrow \quad H \\ s &\longmapsto \quad X_{t+s}, \end{aligned}$$

and for $t \in [0,1)$

$$\begin{aligned} X_{t+\cdot}: \quad [-1,0] &\longrightarrow \quad H \\ s &\longmapsto \quad \phi(t+s)\mathbf{1}_{[-1,-t)}(s) + X_{t+s}\mathbf{1}_{[-t,0]}(s), \end{aligned}$$

and

(i) $(A, D(A))$ is the infinitesimal generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H.
(ii) $(\psi_0, \psi) \mapsto \sigma(\psi_0, \psi)$ is measurable from $H \times L^2([-1,0];H)$ to $L_2^0(H)$.
(iii) $G : W^{1,2}([-1,0];H) \to H$ is a continuous linear operator given by the Riemann-Stieltjes integral

$$G\phi := \int_{-1}^{0} \eta(ds)\phi(s)$$

where $\eta : [-1,0] \to L(H)$ is of bounded variation.

Such an equation is usually studied in an extended Hilbert space which also takes the evolution of the past segment $(X_{t+\cdot})_{t\geq 0}$ into account, see [8]. Below we follow this approach. Namely, introduce the new Hilbert space

$$\mathcal{H} = H \times L^2([-1,0];H), \qquad \|(\phi_0, \phi)\|_{\mathcal{H}} = \left(\|\phi_0\|_H^2 + \|\phi\|^2_{L^2([-1,0];H)}\right)^{1/2}. \tag{6.2}$$

Define the operator

$$\mathcal{A}_0 := \begin{pmatrix} A & 0 \\ 0 & \frac{d}{ds} \end{pmatrix} \quad D(\mathcal{A}) = \{(\phi_0, \phi)^T \in D(A) \times W^{1,2}([0,1];H) : \phi(0) = \phi_0\},$$

which generates a strongly continuous semigroup $(\mathcal{S}_0(t))_{t\geq 0}$ on $\mathcal{H}$, given by

$$\mathcal{S}_0(t) := \begin{pmatrix} S(t) & 0 \\ S_t & T_0(t) \end{pmatrix} \tag{6.3}$$

due to [5, Theorem 3.25]. Here $(T_0(t))_{t\geq 0}$ is the nilpotent left shift semigroup on $L^2([-1,0];H)$ and

$$S_t\phi_0(\tau) := \begin{cases} S(t+\tau)\phi_0, & -t < \tau \leq 0 \\ 0, & -1 \leq \tau \leq -t, \end{cases}$$

It then follows from [5, Theorem 3.29] that the operator $\mathcal{A}$ with domain $D(\mathcal{A}) = D(\mathcal{A}_0)$ given by

$$\mathcal{A} := \begin{pmatrix} A & G \\ 0 & \frac{d}{ds} \end{pmatrix} = \mathcal{A}_0 + \begin{pmatrix} 0 & G \\ 0 & 0 \end{pmatrix} \tag{6.4}$$

is the generator of a strongly continuous semigroup $(\mathcal{S}(t))_{t\geq 0}$ on $\mathcal{H}$. Thus, we can formally identify (6.1) with the $\mathcal{H}$-valued SPDE

$$\begin{cases} d\mathcal{X}_t = \mathcal{A}\mathcal{X}_t dt + \Sigma(\mathcal{X}_t)dW_t \\ \mathcal{X}_0 = (\phi_0, \phi)^T \quad t \geq 0, \end{cases} \qquad \Sigma(\phi_0, \phi) := \begin{pmatrix} \sigma(\phi_0, \phi) & 0 \\ 0 & 0 \end{pmatrix}. \tag{6.5}$$

6.2. **Main results for** (6.5)**.** Next we proceed to apply the results of this work to the SPDE (6.5). For this purpose we we make the following assumption:

(C1) There exists an $L_\sigma > 0$ such that

$$\|\sigma(\phi_0, \phi) - \sigma(\psi_0, \psi)\|^2_{L_0^2(H)} \leq L_\sigma \left(\|\phi_0 - \psi_0\|_H^2 + \|\phi - \psi\|^2_{L^2([-1,0];H)}\right)$$

holds for all $(\phi, \phi_0), (\psi_0, \psi) \in \mathcal{H}$.
(C2) The operator $(A, D(A))$ satisfies (GDC) with projection operators P_0, P_1 and constants $\alpha > 0$, $\beta \geq 0$.

We will see that condition (C1) implies (A1), condition (C2) will be used to prove that $\mathcal{A}$ also satisfies (GDC) with respect to a (possibly equivalent) scalar product on $\mathcal{H}$.

Proposition 6.1. *Suppose that conditions (C1), (C2) are satisfied, that η has a jump at -1 and that one of the following conditions hold:*

(i) G is bounded on $L^2([-1,0];H)$ or

(ii) $(S(t))_{t\geq 0}$ leaves $H_0 = \operatorname{ran}(P_0)$ and $H_1 = \operatorname{ran}(P_1)$ invariant, H_0, H_1 are orthogonal, $\operatorname{ran}(G) \subset H_1$ and GP_0 extends to a bounded linear operator on $L^2([-1,0];H)$.

Then for each initial condition $(\phi_0, \phi) \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathcal{H})$ there exists a unique mild solution $(\mathcal{X}_t)_{t\geq 0} \subset L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{H})$ to (6.5).

Proof. Under condition (i) we work on the Hilbert space $\mathcal{H}$ while under condition (ii) we work on the Hilbert space $\mathcal{H}^\tau$ given by $\mathcal{H}$ equipped with the equivalent norm given by

$$\|(\phi_0, \phi)\|_{\mathcal{H}^\tau}^2 := \|\phi_0\|_H^2 + \int_{-1}^0 \|P_0\phi(s)\|_H^2 ds + \int_{-1}^0 \|P_1\phi(s)\|_H^2 \tau(s) ds. \tag{6.6}$$

where

$$\tau(r) = \int_{-1}^r \|\eta(dr)\|_{L(H)}, \qquad r \in [-1, 0] \tag{6.7}$$

denotes the variation of η. Note that due to a result of Webb (see [41] and Remark 6.4 below) this norm is, indeed, equivalent to the original norm on $\mathcal{H}$. For condition (A1) we first observe that $L_F = L_\gamma = 0$ and if assumption (i) holds, then

$$\begin{aligned}\|\Sigma(\phi_0, \phi) - \Sigma(\psi_0, \psi)\|_{L_0^2(\mathcal{H})}^2 &\leq \|\sigma(\phi_0, \phi) - \sigma(\psi_0, \psi)\|_{L_0^2(H)}^2 \\ &\leq L_\sigma \left(\|\phi_0 - \psi_0\|_H^2 + \|\phi - \psi\|_{L^2([-1,0];H)}^2 \right) \\ &= L_\sigma \|(\phi_0, \phi)^T - (\psi_0, \psi)^T\|_{\mathcal{H}}^2.\end{aligned}$$

If condition (ii) holds, then analogously we obtain

$$\begin{aligned}\|\Sigma(\phi_0, \phi) - \Sigma(\psi_0, \psi)\|_{L_0^2(\mathcal{H}^\tau)}^2 &\leq L_\sigma \|(\phi_0, \phi)^T - (\psi_0, \psi)^T\|_{\mathcal{H}}^2 \\ &\leq \max\{1, \tau(0)\} L_\sigma \|(\phi_0, \phi)^T - (\psi_0, \psi)^T\|_{\mathcal{H}^\tau}^2.\end{aligned}$$

This shows that condition (A1) is satisfied. Finally, it follows from Proposition 6.5 below that the operator $(\mathcal{A}, D(\mathcal{A}))$ satisfies condition (GDC). ☐

We proceed to formulate our main results on invariant measures for (6.5). For this purpose we introduce the following additional condition:

(C3) For each $(\phi_0, \phi) \in \mathcal{H}$ there exist $M(\phi_0, \phi) \geq 1$, $\delta(\phi_0, \phi) > 0$ and an element $\widetilde{f}(\phi_0, \phi) \in \mathcal{H}$ such that

$$\|\mathcal{S}(t)(P_1\phi_0, \phi) - \widetilde{f}(\phi_0, \phi)\|_{\mathcal{H}} \leq M(\phi_0, \phi) \mathrm{e}^{-t\delta(\phi_0, \phi)}, \qquad t \geq 0.$$

Observe that (C3) is precisely condition (A3). This is trivially satisfied, if $(\mathcal{S}(t))_{t\geq 0}$ is exponentially stable which is for example the case in the setting of [5, Corollary 5.9].

Introduce the subspaces

$$\mathcal{H}_0 := H_0 \times \{0\} \quad \text{and} \quad \mathcal{H}_1 := H_1 \times L^2([-1,0];H),$$

which yield an orthogonal decomposition of $\mathcal{H}$ with projection operators

$$\begin{aligned}\mathcal{P}_0 &: \mathcal{H} \longrightarrow \mathcal{H}_0, \quad (\phi_0, \phi) \longmapsto (P_0\phi_0, 0), \\ \mathcal{P}_1 &: \mathcal{H} \longrightarrow \mathcal{H}_1, \quad (\phi_0, \phi) \longmapsto (P_1\phi_0, \phi).\end{aligned}$$

The following is our main result for this section.

Theorem 6.2. *Suppose that conditions (C1) – (C3) hold, that $P_1\sigma(\phi_0,\phi) = 0$ for all $(\phi_0,\phi) \in \mathcal{H}$, and that one of the following conditions are satisfied:*

(i) G is bounded on $L^2([-1,0];H)$, $GP_1 = P_1G$, $(S(t))_{t\geq 0}$ commutes with P_1, and
$$\alpha > 1/2 + L_\sigma/2;$$

(ii) $(S(t))_{t\geq 0}$ leaves $H_0 = \mathrm{ran}(P_0)$ and $H_1 = \mathrm{ran}(P_1)$ invariant, H_0, H_1 are orthogonal, $\mathrm{ran}(G) \subset H_1$, GP_0 extends to a bounded linear operator on $L^2([-1,0];H)$, and
$$\alpha > 1/2 + \max\{1,\tau(0)\}L_\sigma/2.$$

Then the assertions of Theorem 4.3 and Corollary 4.4 are applicable. In particular, for each $(\phi_0,\phi) \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathcal{H})$ there exists an invariant measure $\pi_{Law(h)}$ for the Markov process $(\mathcal{X}_t)_{t\geq 0}$, and this measure satisfies $\pi_{Law(h)} = \pi_{Law(\mathcal{P}_1 h)}$.

Proof. Let us first show condition (A2) is satisfied, i.e. that $\mathcal{S}(t)$ leaves $\mathcal{H}_1$ invariant and $\mathcal{P}_1\Sigma = 0$. It follows from Lemma 6.6 below that $\mathcal{P}_1$ commutes with the semigroup $(\mathcal{S}(t))_{t\geq 0}$. Moreover, one has
$$\mathcal{P}_1\Sigma(\phi_0,\phi) = \begin{pmatrix} P_1\sigma(\phi_0,\phi) & 0 \\ 0 & 0 \end{pmatrix} = 0$$
due to $P_1\sigma = 0$. This shows that condition (A2) is satisfied. Condition (A3) is immediate by assumption (C3) while (3.13) reduces under condition (i) to
$$\varepsilon = 2\left(\alpha - \frac{1}{2}\right) - L_\sigma > 0,$$
and under condition (ii) to
$$\varepsilon = 2\left(\alpha - \frac{1}{2}\right) - \max\{1,\tau(0)\}L_\sigma > 0.$$
Altogether we conclude that Theorems 4.3 and 4.4 apply, which proves the assertion. ☐

Remark 6.3. *Condition (ii) is slightly more restrictive on the semigroup and the projection operators than condition (i). In contrast to the latter, condition (ii) contains delay operators like point evaluations in H_1 , that is $G = \delta_{-1}P_1$ for $\delta_{-1}\phi = \phi(-1)$ for $\phi \in W^{1,2}([-1,0];H_1)$.*

6.3. **Some technical results.** Let us first provide a sufficient and easy to check condition for the operator $\mathcal{A}$ to satisfy the generalized dissipativity condition (GDC), and afterward we state our main result on invariant measures for this stochastic delay equation. As a first step we recall a result from [41].

Remark 6.4 (An equivalent scalar product)**.** *Let τ be defined as in* (6.7) *and suppose that η has a jump at -1. Suppose that there exists $c \in \mathbb{R}$ such that $A - c$ is dissipative. Then the Hilbert space norm defined by*
$$\|(\phi_0,\phi)\|^2_{\mathcal{H}^\tau} := \|\phi_0\|_H^2 + \int_{-1}^0 \|\phi(s)\|_H^2 \tau(s)ds$$
is equivalent to the original one on $\mathcal{H}$. Moreover, $\mathcal{A} - \gamma I$ is dissipative for every $\gamma \geq \max\{0, c+\tau(0)\}$ with respect to this norm, i.e.
$$\langle \mathcal{A}(\phi_0,\phi)^T, (\phi_0,\phi)\rangle_{\mathcal{H}^\tau} \leq \gamma \|(\phi_0,\phi)^T\|_{\mathcal{H}^\tau}, \qquad \forall (\phi_0,\phi) \in D(\mathcal{A}).$$

Based on this observation we can now provide sufficient conditions for $(\mathcal{A}, D(\mathcal{A}))$ to satisfy (GDC).

Proposition 6.5. *Suppose that A satisfies (GDC) with constants $\alpha, \beta \geq 0$.*

(i) *If G extends to a continuous linear operator on $L^2([-1,0];H)$ and $\alpha > 1/2$, then $\mathcal{A}$ satisfies (GDC), i.e.,*

$$\langle \mathcal{A}(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}} \leq -\widetilde{\alpha}\|(\phi_0,\phi)\|_{\mathcal{H}}^2 + \left(\widetilde{\alpha} + \widetilde{\beta}\right)\|\mathcal{P}_1(\phi_0,\phi)^T\|_{\mathcal{H}}^2$$

where $\varepsilon > 0$ is such that $\varepsilon < \sqrt{2\alpha - 1}$, and

$$\widetilde{\alpha} := \alpha - \frac{1+\varepsilon^2}{2} \quad \text{and} \quad \widetilde{\beta} := \beta + \alpha + \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2 + \frac{\varepsilon^2}{2}.$$

(ii) *Assume that H_0, H_1 provide an orthogonal decomposition of H such the semigroup generated $(S(t))_{t\geq 0}$ generated by $(A, D(A))$ leaves H_0 and H_1 invariant. Moreover, suppose that $\operatorname{ran}(G) \subseteq H_1$, and that $W^{1,2}([-1,0];H) \ni \phi \mapsto GP_0\phi \in H_1$ extends to a continuous linear operator $GP_0 : L^2([-1,0];H) \to H_1$ with operator norm denoted by $\|GP_0\|$. Define an equivalent Hilbert space norm by (6.6). If $\alpha > 1/2$, then $\mathcal{A}$ satisfies (GDC) with respect to this norm, i.e., it holds that*

$$\begin{aligned}\langle \mathcal{A}(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}^\tau} \leq& -\left(\alpha - \frac{1}{2}\right)\|(\phi_0,\phi)\|_{\mathcal{H}^\tau} \\ &+ \left(\left(\alpha - \frac{1}{2}\right) + \beta + \tau(0) + \frac{\|GP_0\|}{2}\right)\|\mathcal{P}_1(\phi_0,\phi)\|_{\mathcal{H}^\tau}^2\end{aligned}$$

Proof. (i) For $(\phi_0,\phi)^T \in D(\mathcal{A}_0)$ we have

$$\begin{aligned}\langle \mathcal{A}_0(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}} &= \langle A\phi_0, \phi_0\rangle_H + \int_{-1}^0 \left\langle \frac{d}{ds}\phi(s), \phi(s)\right\rangle_H ds \\ &= \langle A\phi_0, \phi_0\rangle_H + \int_{-1}^0 \frac{1}{2}\frac{d}{ds}\|\phi(s)\|_H^2 ds \\ &= \langle A\phi_0, \phi_0\rangle_H + \frac{1}{2}(\|\phi(0)\|_H^2 - \|\phi(-1)\|_H^2) \\ &\leq \langle A\phi_0, \phi_0\rangle_H + \frac{1}{2}\|\phi_0\|_H^2,\end{aligned}$$

where we used the fact that $\phi_0 = \phi(0)$. Making further use of the fact that A satisfies (GDC) we find

$$\begin{aligned}\langle \mathcal{A}_0(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}} &\leq -\left(\alpha - \frac{1}{2}\right)\|\phi_0\|_H^2 + (\beta + \alpha)\|P_1\phi_0\|_H^2 \\ &\leq -\left(\alpha - \frac{1}{2}\right)\|(\phi_0,\phi)\|_{\mathcal{H}}^2 + \left(\beta + 2\alpha - \frac{1}{2}\right)\|(P_1\phi_0,\phi)^T\|_{\mathcal{H}}^2.\end{aligned}$$

To estimate the operator $\mathcal{A}$ we will use that

$$\begin{aligned}\langle G\phi, \phi_0\rangle_H &\leq \|G\phi\|_H\|\phi_0\|_H \\ &\leq \frac{1}{2\varepsilon^2}\|G\phi\|_H^2 + \frac{\varepsilon^2}{2}\|\phi_0\|_H^2 \\ &= \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2\|\phi\|_{L^2([-1,0];H)}^2 + \frac{\varepsilon^2}{2}\|\phi_0\|_H^2 \\ &\leq \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2\|\mathcal{P}_1(\phi_0,\phi)^T\|_{\mathcal{H}}^2 + \frac{\varepsilon^2}{2}\|(\phi_0,\phi)^T\|_{\mathcal{H}}^2\end{aligned}$$

where $\varepsilon > 0$. Thus we obtain

$$\begin{aligned}
&\langle \mathcal{A}(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}} \\
&= \langle \mathcal{A}_0(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}} + \langle G\phi, \phi_0\rangle_H \\
&\leq -\left(\alpha - \frac{1}{2}\right)\|(\phi_0,\phi)\|_{\mathcal{H}}^2 + \left(\beta + 2\alpha - \frac{1}{2}\right)\|\mathcal{P}_1(\phi_0,\phi)^T\|_{\mathcal{H}}^2 \\
&\quad + \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2\|\mathcal{P}_1(\phi_0,\phi)^T\|_{\mathcal{H}}^2 + \frac{\varepsilon^2}{2}\|(\phi_0,\phi)^T\|_{\mathcal{H}}^2 \\
&= -\left(\alpha - \frac{1+\varepsilon^2}{2}\right)\|(\phi_0,\phi)\|_{\mathcal{H}}^2 + \left(\beta + 2\alpha - \frac{1}{2} + \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2\right)\|\mathcal{P}_1(\phi_0,\phi)^T\|_{\mathcal{H}}^2.
\end{aligned}$$

Assuming ε is so small that $\varepsilon < \sqrt{2\alpha - 1}$, we obtain $\alpha - \frac{1+\varepsilon^2}{2} > 0$ and

$$\begin{aligned}
&\beta + 2\alpha - \frac{1}{2} + \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2 \\
&= \left(\alpha - \frac{1+\varepsilon^2}{2}\right) + \beta + \alpha + \frac{1}{2\varepsilon^2}\|G\|_{L(L^2([-1,0];H))}^2 + \frac{\varepsilon^2}{2} > 0
\end{aligned}$$

which proves the assertion.

(ii) As P_0, P_1 are complementary self-adjoint projections, they induce an orthogonal decomposition $H = H_0 \oplus H_1$. Thus, for $(\phi_0,\phi) \in \mathcal{H}$ we have $(\phi_0,\phi) = (P_0\phi_0, P_0\phi) + (P_1\phi_0, P_1\phi)$ which gives also an orthogonal decomposition

$$\mathcal{H} = \left(H_0 \times L^2([-1,0];H_0)\right) \oplus \left(H_1 \times L^2([-1,0];H_1)\right).$$

Applying Remark 6.4 to the Hilbert space $H_1 \times L^2([-1,0];H_1)$ we find that

$$\|P_1\phi_0\|_{H^2} + \int_{-1}^0 \|P_1\phi(s)\|_H^2 \tau(s)ds$$

gives rise to a norm on $H_1 \times L^2([-1,0];H_1)$ which is equivalent to the one given by (6.2) when applied to $(P_1\phi_0, P_1\phi)$. Thus, the norm defined in (6.6) is, indeed, equivalent to the original norm on $\mathcal{H}$. Let $(\phi_0,\phi) \in D(\mathcal{A})$, so that $\phi(0) = \phi_0$ and we can write:

$$\begin{aligned}
&\langle \mathcal{A}(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}^\tau} \\
&= \left\langle \left(A\phi_0 + G\phi, \frac{d}{ds}\phi\right)^T, (\phi_0,\phi)^T \right\rangle_{\mathcal{H}^\tau} \\
&= \langle A\phi_0, \phi_0\rangle_H + \langle G\phi, \phi_0\rangle_H \\
&\quad + \int_{-1}^0 \left\langle \frac{d}{ds}P_0\phi(s), P_0\phi(s)\right\rangle_H ds + \int_{-1}^0 \left\langle \frac{d}{ds}P_1\phi(s), P_1\phi(s)\right\rangle_H \tau(s)ds \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

For the first term I_1 we use $\phi_0 = P_0\phi_0 + P_1\phi_0$, then the fact that P_0, P_1 are self-adjoint projection operators and finally $P_0A = AP_0$, $P_1A = AP_1$ (similarly to the proof of Proposition 3.1) to find that

$$\begin{aligned}
I_1 &= \langle P_0A\phi_0, P_0\phi_0\rangle_H + \langle P_1A\phi_0, P_1\phi_0\rangle_H \\
&= \langle AP_0\phi_0, P_0\phi_0\rangle_H + \langle AP_1\phi_1, P_1\phi_0\rangle_H \leq -\alpha\|P_0\phi_0\|_H^2 + \langle AP_1\phi_1, P_1\phi_0\rangle_H,
\end{aligned}$$

where the last inequality follows from (GDC) combined with $P_1P_0\phi_0 = 0$. Likewise, for the second term we use that $\operatorname{ran}(G) \subset H_1$ so that $P_0G = 0$ to obtain

$$I_2 = \langle G\phi, P_1\phi_0\rangle_H$$

$$
\begin{aligned}
&= \langle GP_0\phi, P_1\phi_0\rangle_H + \langle GP_1\phi, P_1\phi_0\rangle_H \\
&\le \|GP_0\|\|P_0\phi\|_{L^2([-1,0];H)}\|P_1\phi_0\|_H + \langle GP_1\phi, P_1\phi_0\rangle_H \\
&\le \frac{\|GP_0\|}{2}\|P_0\phi\|^2_{L^2([-1,0];H)} + \frac{\|GP_0\|}{2}\|P_1\phi_0\|^2_H + \langle GP_1\phi, P_1\phi_0\rangle_H.
\end{aligned}
$$

For the third term we obtain

$$
I_3 = \frac{1}{2}\int_{-1}^0 \frac{d}{ds}\|P_0\phi(s)\|^2_H ds \le \frac{1}{2}\|P_0\phi_0\|^2_H.
$$

To summarize, we obtain

$$
\begin{aligned}
&\langle \mathcal{A}(\phi_0,\phi)^T, (\phi_0,\phi)^T\rangle_{\mathcal{H}^\tau} \\
&\le -\left(\alpha - \frac{1}{2}\right)\|P_0\phi_0\|^2_H + \frac{\|GP_0\|}{2}\left(\|P_1\phi_0\|^2_H + \|P_0\phi\|^2_{L^2([-1,0];H)}\right) \\
&\quad + \langle AP_1\phi_1, P_1\phi_0\rangle_H + \langle GP_1\phi, P_1\phi_0\rangle_H + \int_{-1}^0 \left\langle \frac{d}{ds}P_1\phi(s), P_1\phi(s)\right\rangle_H \tau(s)ds \\
&= -\left(\alpha - \frac{1}{2}\right)\|(\phi_0,\phi)^T\|^2_{\mathcal{H}^\tau} + \left(\alpha - \frac{1}{2}\right)\|P_1\phi_0\|^2_H + \left(\alpha - \frac{1}{2}\right)\int_{-1}^0 \|P_0\phi(s)\|^2_H ds \\
&\quad + \left(\alpha - \frac{1}{2}\right)\int_{-1}^0 \|P_1\phi(s)\|^2_H \tau(s)ds + \frac{\|GP_0\|}{2}\|P_1\phi_0\|^2_H + \frac{\|GP_0\|}{2}\int_{-1}^0 \|P_0\phi(s)\|^2_H ds \\
&\quad + \left\langle \mathcal{A}(P_1\phi_0, P_1\phi)^T, (P_1\phi_0, P_1\phi)^T\right\rangle_{\mathcal{H}^\tau} \\
&\le -\left(\alpha - \frac{1}{2}\right)\|(\phi_0,\phi)^T\|^2_{\mathcal{H}^\tau} + \left(\alpha - \frac{1}{2} + \gamma + \frac{\|GP_0\|}{2}\right)\|P_1\phi_0\|^2_H \\
&\quad + \left(\alpha - \frac{1}{2} + \frac{\|GP_0\|}{2}\right)\int_{-1}^0 \|P_0\phi(s)\|^2_H ds + \left(\alpha - \frac{1}{2} + \gamma\right)\int_{-1}^0 \|P_1\phi(s)\|^2_H \tau(s)ds \\
&\le -\left(\alpha - \frac{1}{2}\right)\|(\phi_0,\phi)^T\|^2_{\mathcal{H}^\tau} + \left(\alpha - \frac{1}{2} + \gamma + \frac{\|GP_0\|}{2}\right)\|\mathcal{P}_1(\phi_0,\phi)^T\|^2_{\mathcal{H}^\tau},
\end{aligned}
$$

where we have used the fact that $A - \beta I$ is dissipative so that by Remark 6.4 with $\gamma = \beta + \tau(0)$

$$
\begin{aligned}
\left\langle \mathcal{A}(P_1\phi_0, P_1\phi)^T, (P_1\phi_0, P_1\phi)^T\right\rangle_{\mathcal{H}^\tau} &\le \gamma\|(P_1\phi_0, P_1\phi)^T\|^2_{\mathcal{H}^\tau} \\
&= \gamma\|P_1\phi_0\|^2_H + \gamma\int_{-1}^0 \|P_1\phi(s)\|^2_H \tau(s)ds.
\end{aligned}
$$

This proves the assertion. □

The next result has been used in the previous proof.

Lemma 6.6. *Consider the setting of stochastic delay equation, i.e., let* $(\mathcal{S}_0(t))_{t\ge 0}$*,* $(\mathcal{S}(t))_{t\ge 0}$*,* $(\mathcal{A}_0, D(\mathcal{A}_0))$*,* $(\mathcal{A}, D(\mathcal{A}))$*,* G*,* P_1 *as in Sections 6.1, 6.2 and Theorem 6.2. In particular suppose that* $P_1S(t) = P_1S(t)$*, where* $S(t)$ *is given in* (6.3)*. Then*

$$
\mathcal{P}_1\mathcal{S}(t) = \mathcal{S}(t)\mathcal{P}_1, \qquad t \ge 0.
$$

Proof. We prove this statement in two different cases. Let us first consider the case where G satisfies assumption (i) from Proposition 6.5, i.e. G is bounded from $L^2([-1,0];H)$ to H. Since G is bounded we obtain from the bounded perturbation theorem (the

Dyson-Phillips series) the representation

$$\mathcal{S}(t) = \sum_{n=0}^{\infty} \mathcal{S}_0^{(n)}(t),$$

where the series converges in $L(\mathcal{H})$, and $\mathcal{S}_0^{(n)}(t)$ is inductively defined by

$$\mathcal{S}_0^{(0)}(t) = \mathcal{S}_0(t), \qquad \mathcal{S}_0^{(n+1)}(t) = \int_0^t \mathcal{S}_0^{(n)}(s) \begin{pmatrix} 0 & G \\ 0 & 0 \end{pmatrix} \mathcal{S}_0^{(}t-s)ds.$$

Thus it suffices to prove that

$$\mathcal{P}_1 \mathcal{S}_0^{(n)}(t) = \mathcal{S}_0^{(n)}(t)\mathcal{P}_1, \qquad n \geq 1, \quad t \geq 0. \tag{6.8}$$

For $n = 0$ we use the particular form of $\mathcal{P}_1$ and $\mathcal{S}_0(t)$ to find that

$$\begin{aligned} \mathcal{P}_1\mathcal{S}_0(t)(\phi_0,\phi)^T &= \mathcal{P}_1 \begin{pmatrix} S(t)\phi_0 \\ S_t\phi_0 + T_0(t)\phi \end{pmatrix} \\ &= \begin{pmatrix} P_1S(t)\phi_0 \\ S_t\phi_0 + T_0(t)\phi \end{pmatrix} = \begin{pmatrix} S(t)P_1\phi_0 \\ S_t\phi_0 + T_0(t)\phi \end{pmatrix} = \mathcal{S}_0(t)\mathcal{P}_1(\phi_0,\phi)^T, \end{aligned}$$

where we have used that $S(t)$ commutes with P_1. Now suppose that (6.8) holds for some $n \geq 0$. Then

$$\begin{aligned} \mathcal{P}_1\mathcal{S}_0^{(n+1)}(t) &= \mathcal{P}_1\mathcal{S}_0(t) + \int_0^t \mathcal{P}_1\mathcal{S}_0^{(n)}(s) \begin{pmatrix} 0 & G \\ 0 & 0 \end{pmatrix} \mathcal{S}_0^{(}t-s)ds \\ &= \mathcal{S}_0(t)\mathcal{P}_1 + \int_0^t \mathcal{S}_0^{(n)}(s) \begin{pmatrix} 0 & G \\ 0 & 0 \end{pmatrix} \mathcal{S}_0^{(n)}(t-s)\mathcal{P}_1 ds = \mathcal{S}_0^{(n+1)}(t)\mathcal{P}_1, \end{aligned}$$

where we have used that

$$\begin{aligned} \mathcal{P}_1 \begin{pmatrix} 0 & G \\ 0 & 0 \end{pmatrix} (\phi_0,\phi)^T &= \mathcal{P}_1(G\phi, 0)^T \\ &= (P_1G\phi, 0) = (GP_1\phi, 0) = \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix} \mathcal{P}_1(\phi_0,\phi)^T. \end{aligned}$$

This completes the proof for the case where $G : L^2([-1,0]; H) \longrightarrow H$ is bounded.

Let us now consider the case where condition (ii) from Proposition 6.5 holds. Following [5, Theorem 3.29] we know that the semigroup $(\mathcal{S}(t))_{t\geq 0}$ is constructed as a Miyadera-Voigt perturbation and hence has due to [14, Chapter III, Corollary 3.15] a series representation of the form

$$\mathcal{S}(t) = \sum_{n=0}^{\infty} \overline{V}^n \mathcal{S}_0(t),$$

where $\overline{V}$ denotes the closure of the operator

$$F \longmapsto VF(t) := \int_0^t F(s) \begin{pmatrix} 0 & G \\ 0 & 0 \end{pmatrix} \mathcal{S}_0(t-s)ds,$$

where $F \in C([0,t_0], L_s(\mathcal{H}))$ for some small $t_0 > 0$ and $L_s(\mathcal{H})$ denotes the space of bounded linear operators over $\mathcal{H}$ equipped with the strong operator topology. Following the same computations as in the first case, we can prove that $\mathcal{P}_1V^n\mathcal{S}_0(t) = V^n\mathcal{S}_0(t)\mathcal{P}_1$ and hence $\mathcal{P}_1\overline{V}^n\mathcal{S}_0(t) = \overline{V}^n\mathcal{S}_0(t)\mathcal{P}_1$. This proves the assertion also in this case. $\square$

Appendix A. Itô formula

Below we recall an Itô formula for Hilbert space valued semimartingales of the form

$$X(t) = X(0) + \int_0^t a(s)ds + \int_0^t \sigma(s)dW_s + \int_0^t \int_E \gamma(s,\nu)\widetilde{N}(ds,d\nu),$$

where a and σ are as before and $(\gamma(t,\nu))_{t\geq 0}$ is a predictable, H-valued stochastic process for each $\nu \in E$ such that

$$\mathbb{E}\left[\int_0^t \int_E \|\gamma(s,\nu)\|_H^2 \mu(d\nu)ds\right] < \infty$$

and

$$\mathbb{E}\left[\int_0^t \|\sigma(s)\|_{L_2^0}^2 ds\right] < \infty.$$

For this purpose we first introduce the class of quasi-sublinear functions.

Definition A.1 (Sublinear Functions)**.** *A continuous, non-decreasing function $h : \mathbb{R}_+ \to \mathbb{R}_+$ is called quasi-sublinear, if there exists a constant $C > 0$ such that*

$$h(x+y) \leq C(h(x)+h(y))$$
$$h(xy) \leq C(h(x)h(y))$$

for all $x, y \geq 0$.

The following Itô-Formula is a combination of [20] and [29].

Theorem A.2 (Generalized Itô-Formula)**.** *Let $F \in C^2(\mathbb{R}_+ \times H, \mathbb{R})$ and suppose there exist quasi-sublinear functions $h_1, h_2 : \mathbb{R}_+ \to \mathbb{R}_+$ such that for all $t \geq 0$ and $x \in H$*

$$\|F_x(t,x)\|_H \leq h_1(\|x\|_H), \qquad \|F_{xx}(t,x)\|_{L(H,L(H,\mathbb{R}))} \leq h_2(\|x\|_H)$$

and

$$\int_0^t \int_E \|\gamma(s,\nu)\|_H^2 \mu(d\nu)ds + \int_0^t \int_E h_1(\|\gamma(s,\nu)\|_H)^2 \|\gamma(s,\nu)\|_H^2 \mu(d\nu)ds$$
$$+ \int_0^t \int_E h_2(\|\gamma(s,\nu)\|_H)\|\gamma(s,\nu)\|_H^2 \mu(d\nu)ds < \infty$$

Then $\mathbb{P}$-almost surely for each $t \geq 0$:

$$\int_0^t \|F_t(s,X(s))\|_H ds + \int_0^t \int_E |F(s,X(s)+\gamma(s,\nu)) - F(s,X(s))|^2 \mu(d\nu)ds$$
$$+ \int_0^t \int_E |F(s,X(s)+\gamma(s,\nu)) - F(s,X(s)) - \langle F_x(s,X(s)),\gamma(s,\nu)\rangle_H| \mu(d\nu)ds < \infty.$$

Moreover, the generalized Itô-formula holds $\mathbb{P}$-almost surely for each $t \geq 0$ and

$$\begin{aligned} F(t,X(t)) = F(0,X(0)) &+ \int_0^t \mathcal{L}F(s,X(s))ds \\ &+ \int_0^t \langle F_x(s,X(s)), \sigma(s)dW_s\rangle_H \\ &+ \int_0^{t+} \int_E \{F(s,X(s-)+\gamma(s,\nu)) - F(s,X(s-))\}\, \widetilde{N}(ds,d\nu) \end{aligned}$$

where $\mathcal{L}F(x,X(s))$ is given by

$\mathcal{L}F(s,X(s))$

$$
\begin{aligned}
&= \int_0^t \{F_t(s, X(s)) + \langle F_x(s, X(s)), a(s)\rangle_H\} \, ds \\
&\quad + \frac{1}{2}\int_0^t \operatorname{tr}\left[F_{xx}(s, X(s))\sigma(s)Q\sigma(s)^*\right] ds \\
&\quad + \int_0^t \int_E \{F(s, X(s) + \gamma(s,\nu)) - F(s, X(s)) - \langle F_x(s, X(s)), \gamma(s,\nu)\rangle_H\} \, \mu(d\nu) ds
\end{aligned}
$$

Acknowledgements. Dennis Schroers was funded within the project STORM: Stochastics for Time-Space Risk Models, from the Research Council of Norway (RCN). Project number: 274410.

References

1. S. Albeverio, V Mandrekar, and B. Rüdiger, *Existence of mild solutions for stochastic differential equations and semilinear equations with non-gaussian lévy noise*, Stochastic Processes and their Applications **119** (2009), no. 3, 835 – 863.
2. Sergio Albeverio, Leszek Gawarecki, Vidyadhar Mandrekar, Barbara Rüdiger, and Barun Sarkar, *Itô Formula for Mild Solutions of SPDEs with Gaussian and non-Gaussian noise and Applications to Stability Properties*, Random Operators and Stochastic Equations **25** (2017), no. 2.
3. David Applebaum, *Infinite dimensional Ornstein-Uhlenbeck processes driven by Lévy processes*, Probab. Surv. **12** (2015), 33–54. MR 3385977
4. Viorel Barbu, Michael Röckner, and Deng Zhang, *Stochastic nonlinear Schrödinger equations*, Nonlinear Anal. **136** (2016), 168–194. MR 3474409
5. András Bátkai and Susanna Piazzera, *Semigroups for delay equations*, Research Notes in Mathematics, vol. 10, A K Peters, Ltd., Wellesley, MA, 2005. MR 2181405
6. Fred Espen Benth, Nils Detering, and Paul Kruehner, *Stochastic Volterra integral equations and a class of first order stochastic partial differential equations*, arXiv:1903.05045 (2019).
7. Oleg Butkovsky, *Subgeometric rates of convergence of Markov processes in the Wasserstein metric*, Ann. Appl. Probab. **24** (2014), no. 2, 526–552. MR 3178490
8. Sonja Cox and Mariusz Górajski, *Vector-valued stochastic delay equations - a semigroup approach*, Semigroup Forum **82** (2011), no. 3, 389–411.
9. Giuseppe Da Prato and Arnaud Debussche, *Ergodicity for the 3D stochastic Navier-Stokes equations*, J. Math. Pures Appl. (9) **82** (2003), no. 8, 877–947. MR 2005200
10. Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2014.
11. Stefan Tappe Damir Filipović and Josef Teichmann, *Term Structure Models Driven by Wiener Processes and Poisson Measures: Existence and Positivity*, SIAM Journal on Financial Mathematics **1** (2010), no. 1, 523–554.
12. Donald A. Dawson, Alison M. Etheridge, Klaus Fleischmann, Leonid Mytnik, Edwin A. Perkins, and Jie Xiong, *Mutually catalytic branching in the plane: infinite measure states*, Electron. J. Probab. **7** (2002), No. 15, 61. MR 1921744
13. A. de Bouard and A. Debussche, *The stochastic nonlinear Schrödinger equation in H^1*, Stochastic Anal. Appl. **21** (2003), no. 1, 97–126. MR 1954077
14. K.J. Engel, S. Brendle, R. Nagel, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, et al., *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Mathematics, Springer New York, 1999.
15. Damir Filipović, *Consistency Problems for HJM Interest Rate Models*, Springer, Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo, 2001.
16. Damir Filipović, *Term-Structure Models: A Graduate Course*, Springer Finance, Springer Berlin Heidelberg, 2009.
17. Damir Filipović, Stefan Tappe, and Josef Teichmann, *Jump-Diffusions in Hilbert Spaces: Existence, Stability and Numerics*, Stochastics **82** (2010), no. 5, 475–520.
18. Franco Flandoli, *Dissipativity and invariant measures for stochastic Navier-Stokes equations*, NoDEA Nonlinear Differential Equations Appl. **1** (1994), no. 4, 403–423. MR 1300150
19. Martin Friesen, Peng Jin, and Barbara Rüdiger, *Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes*, Ann. Appl. Prob. (2020), (to appear).

20. Leszek Gawarecki and Vidyadhar Mandrekar, *Stochastic Differential Equations in Infinite Dimensions: with Applications to Stochastic Partial Differential Equations*, Probability and Its Applications, Springer Berlin Heidelberg, 2010.
21. M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations*, Probab. Theory Related Fields **149** (2011), no. 1-2, 223–259. MR 2773030
22. Martin Hairer and Jonathan C. Mattingly, *Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) **164** (2006), no. 3, 993–1032. MR 2259251
23. ______, *Yet another look at Harris' ergodic theorem for Markov chains*, Seminar on Stochastic Analysis, Random Fields and Applications VI, Progr. Probab., vol. 63, Birkhäuser/Springer Basel AG, Basel, 2011, pp. 109–117. MR 2857021
24. David Heath, Robert Jarrow, and Andrew Morton, *Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation*, Econometrica **60** (1992), 77–105.
25. Alexei Kulik and Michael Scheutzow, *Generalized couplings and convergence of transition probabilities*, Probab. Theory Related Fields **171** (2018), no. 1-2, 333–376. MR 3800835
26. Thomas G. Kurtz, *The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities*, Electron. J. Probab. **12** (2007), 951–965. MR 2336594
27. Werner Linde, *Probability in Banach spaces—stable and infinitely divisible distributions*, second ed., A Wiley-Interscience Publication, John Wiley & Sons, Ltd., Chichester, 1986. MR 874529
28. Vidyadhar Mandrekar and Barbara Rüdiger, *Stochastic Integration in Banach Spaces: Theory and Applications (Probability Theory and Stochastic Modelling Book 73)*, Springer, 2014.
29. Vidyadhar Mandrekar, Barbara Rüdiger, and Stefan Tappe, *Itô's Formula for Banach-space-valued Jump Processes Driven by Poisson Random Measures*, Seminar on Stochastic Analysis, Random Fields and Applications VII, Springer Basel, 2013, pp. 171–186.
30. Leonid Mytnik, *Uniqueness for a mutually catalytic branching model*, Probab. Theory Related Fields **112** (1998), no. 2, 245–253. MR 1653845
31. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 710486
32. Szymon Peszat and Jerzy Zabczyk, *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2007.
33. Enrico Priola, Armen Shirikyan, Lihu Xu, and Jerzy Zabczyk, *Exponential ergodicity and regularity for equations with Lévy noise*, Stochastic Process. Appl. **122** (2012), no. 1, 106–133. MR 2860444
34. Enrico Priola, Lihu Xu, and Jerzy Zabczyk, *Exponential mixing for some SPDEs with Lévy noise*, Stoch. Dyn. **11** (2011), no. 2-3, 521–534. MR 2836539
35. Enrico Priola and Jerzy Zabczyk, *On linear evolution equations for a class of cylindrical Lévy noises*, Stochastic partial differential equations and applications, Quad. Mat., vol. 25, Dept. Math., Seconda Univ. Napoli, Caserta, 2010, pp. 223–242. MR 2985090
36. Anna Rusinek, *Mean Reversion for HJMM Forward Rate Models*, Advances in Applied Probability **42** (2010), no. 2, 371–391.
37. S. S. Sritharan and P. Sundar, *Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise*, Stochastic Process. Appl. **116** (2006), no. 11, 1636–1659. MR 2269220
38. Michael Tehranchi, *A Note on Invariant Measures for HJM Models*, Finance and Stochastics **9** (2005), no. 3, 389–398.
39. Onno van Gaans, *Invariant measures for stochastic evolution equations with Hilbert space valued Lévy noise*, 2005.
40. Cédric Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454
41. G. F. Webb, *Functional differential equations and nonlinear semigroups in L^p-spaces*, J. Differential Equations **20** (1976), no. 1, 71–89. MR 390422

(Bálint Farkas) School of Mathematics and Natural Sciences, University of Wuppertal, Germany
Email address, Bálint Farkas: farkas@math.uni-wuppertal.de

(Martin Friesen) School of Mathematics and Natural Sciences, University of Wuppertal, Germany
Email address, Martin Friesen: friesen@math.uni-wuppertal.de

(Barbara Rüdiger) School of Mathematics and Natural Sciences, University of Wuppertal, Germany
Email address, Barbara Rüdiger: ruediger@uni-wuppertal.de

(Dennis Schroers) Department of Mathematics, University of Oslo
Email address, Dennis Schroers: dennissc@math.uio.no