WONG-ZAKAI APPROXIMATIONS OF STOCHASTIC EVOLUTION EQUATIONS

GIANMARIO TESSITORE AND JERZY ZABCZYK

Abstract. Theorems on weak convergence of the laws of the Wong-Zakai approximations for evolution equation

$$dX(t) = (AX(t) + F(X(t)))dt + G(X(t))dW(t)$$
$$X(0) = x \in H$$

are proved. The operator A in the equation generates an analytic semigroup of linear operators on a Hilbert space H. The tightness of the approximating sequence is established using the stochastic factorization formula. Applications to strongly damped wave and plate equations as well as to stochastic invariance are discussed.

1. Introduction

Let $W(t) = (\beta^1(t), \ldots, \beta^d(t))$, be a d-dimensional standard Wiener process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with filtration $(\mathcal{F}_t)$ satisfying the usual conditions. A natural way to solve numerically the stochastic Ito equations in R^m:

$$dX(t) = F(X(t))dt + \sum_{j=1}^{d} G^j(X(t))d\beta^j(t), \quad X(0) = x, \tag{1.1}$$

is to replace in (1.1)the Wiener process W by its polygonal approximation

$$W_n(t) = W(t_n) + 2^n(t - t_n)(W(t_n^+) - W(t_n)), \tag{1.2}$$

where

$$t_n = [2^n t]/2^n \quad \text{and} \quad t_n^+ = ([2^n t] + 1)/2^n. \tag{1.3}$$

The equation (1.1) becomes an ordinary differential equation with random coefficients:

$$\frac{dy_n(t)}{dt} = F(y_n(t)) + \sum_{j=1}^{d} G^j(y_n(t))\frac{d\beta_n^j(t)}{dt}, \quad y_n(0) = x. \tag{1.4}$$

Approximations y_n of the described form have been first considered by E. Wong and M. Zakai in [39] and [40], in the one dimensional case, and by D. W. Stroock and S. R. S. Varadhan in [38], in the general $m-$ dimensional case. It turned out that the processes y_n converge, in a proper sense, to a solution of the modified equation:

$$dy(t) = \left(F(y(t)) + \frac{1}{2}\sum_{j=1}^{d} \nabla G^j(y(t))G^j(y(t)) \right) dt + \sum_{j=1}^{d} G^j(y(t))d\beta^j(t), \tag{1.5}$$
$$y(t) = x,$$

where ∇G stands for the gradient of the mapping G. The expression $\frac{1}{2}\sum_{j=1}^{d} \nabla G^j(y(t))G^j(y(t))$, is called the $Wong - Zakai$ correction term.

The results by E. Wong and M. Zakai and D. W. Stroock and S. R. S. Varadhan were extended and generalized in various directions in particular to stochastic parabolic equations which are the main subject of the present paper.

Typical examples to which the theory developed in the paper is applicable are *non-linear, stochastic heat equations*:

$$\frac{\partial u}{\partial t}(t,\xi) = \Delta_\xi u(t,\xi) + f(u(t,\xi)) + g(u(t,\xi))\frac{\partial \beta}{\partial t}, \tag{1.6}$$

$$u(0,\xi) = x(\xi),\ \xi \in \mathcal{O}, \quad u(t,\xi) = 0,\ t > 0,\ \xi \in \partial\mathcal{O}, \tag{1.7}$$

and *strongly damped, non-linear, stochastic wave equations*:

$$\frac{\partial^2 u}{\partial t^2}(t,\xi) = \Delta_\xi u(t,\xi) + \rho\Delta_\xi \frac{\partial u}{\partial t}(t,\xi) + f(u(t,\xi)) + g(u(t,\xi))\frac{\partial \beta}{\partial t} \tag{1.8}$$

$$u(t,\xi) = 0, \quad t > 0,\ \xi \in \partial\mathcal{O}, \tag{1.9}$$

$$u(0,\xi) = x_0(\xi), \quad \frac{\partial u}{\partial t}(0,\xi) = x_1(\xi), \quad \xi \in \mathcal{O}. \tag{1.10}$$

In the above equations $\mathcal{O}$ stands for a domain in R^m, Δ_ξ is the Laplace operator and ρ a positive constant. The modified limiting equations are of the form:

$$\frac{\partial u}{\partial t}(t,\xi) = \Delta_\xi u(t,\xi) + (f + \frac{1}{2}g'g)(u(t,\xi)) + g(u(t,\xi))\frac{\partial \beta}{\partial t}, \tag{1.11}$$

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(t,\xi) &= \Delta_\xi u(t,\xi) + \rho\Delta_\xi \frac{\partial u}{\partial t}(t,\xi) + (f + \frac{1}{2}g'g)(u(t,\xi)) \\ &+ g(u(t,\xi))\frac{\partial \beta}{\partial t}. \end{aligned} \tag{1.12}$$

We are concerned with Wong-Zakai approximations of weak solutions of general stochastic evolution equation, in a separable Hilbert space H, of the form

$$\begin{aligned} &dX(t) = (AX(t) + F(X(t)))dt + G(X(t))dW(t), \\ &X(0) = x \in H, \end{aligned} \tag{1.13}$$

where the operator A generates an analytic semi-group of linear operators $S(t)$ on H. Let $y_n(t)$, $t \geq 0$ be the weak solution of the equation:

$$\frac{dy_n(t)}{dt} = Ay_n(t) + F(y_n(t)) + G(y_n(t))\dot{W}_n(t), y_n(0) = x \tag{1.14}$$

called the Wong-Zakai approximation of the solution X of (1.13). To emphasize the dependence of the solution on the initial data we will also write $y_n(t,x)$ instead of $y_n(t)$ only.

The content of the paper can be described as follows.

In Preliminaries we gather basic notations and results related to analytic semi-groups and evolution equations needed in the sequel. Section 3 is devoted to the proof of Theorem 3.1 stating conditions under which the Wong- Zakai approximations are bounded in $p-$ moments in some Sobolev type norms. Tightness of the approximations is studied in Section 4. It turns out, see Theorem 4.1, that for tightness an additional condition of compactness of the operators $S(t), t > 0$, is needed. The convergence of the approximations, see Theorem 5.1, is established in the second half of Chapter 5. In the first part we prove an auxiliary result of independent interest, see Theorem 5.2, that martingale problem associated with the equation has a unique solution. Applications of the results to stochastic invariance are sketched in Section 6.

The main difficulty of the present generalization is due to the unboundedness of the operator A and the lack of local compactness of the Hilbert space H. Moreover in order to allow applications to the case in which F and G are Nemytskii (evaluation) operators over spaces of summable functions, see §2.1 and in particular Remark 2.3, we just assume that F and G are once Gâteaux differentiable (and never require any Fréchet differentiability). For simplicity of the presentation we cover only finite dimensional noise but the techniques used in the paper allow some generalizations to infinite dimensional noise as well. In the proofs we are using basic properties of analytic semi-groups and interpolation spaces and the factorization formula for stochastic convolutions.To some extent we follow the scheme developed for finite dimensional equations in the paper [38].

There exists a substantial number of publications devoted to the Wong- Zakai approximations of stochastic evolution equations. Among the earliest one should mention papers by P. Acquistapace and B. Terreni, [1], Z. Brzezniak, M. Capinski, and F. Flandoli [6], I. Gyöngy [9], and Gyöngy and T. Pröhle [13]. Important recent contributions are due to V. Bally, A. Millet and M. Sanz-Solé [5], I. Gyöngy, D. Nualart and M. Sanz- Solé [12], A. Millet and M. Sanz-Solé [22], [23] and [24]. Those papers investigate either linear stochastic evolution equations or stochastic partial differential equations in one space variable. They often prove path convergence of the approximations and in the present paper we are concerned with the boundedness and the weak convergence of the laws only.

The method we choose here is based on the reformulation of the stochastic partial differential equations as a Hilbert-valued stochastic evolution equation. It, provides abstract results that, due to the generality of our assumptions, can be applied to several concrete cases. Some of them are already treated in literature others, such as stochastic partial differential equations in many space variables, systems of reaction-diffusion equations and equations with elliptic operators A of order higher than 2 and systems, are apparently new. Approximations for some stochastic evolution equations in Hilbert spaces with the operator A generating a strongly continuous semi-group were studied by K. Twardowska, [37], [36], under strong assumptions on A and G. Recently, in T. Nakayama [27], the L^p convergence of the Wong-Zakai approximations of a stochastic evolutions equation similar to (1.13) was obtained in the case in which A is the infinitesimal generator of a C_0 semigroup (not necessarily analytic) but coefficients F and G are twice Fréchet differentiable. On one side first order (and a fortriori second order) Fréchet differentiability of coefficients is, in general, a restrictive assumption that, for instance, is not verified by any of the examples we present here, see Remark 2.3 and Examples 6.1, 6.2 and 6.3. On the other side the results in [27] have applications to important financial models such as Heath-Jarrow-Morton equation for the evolution of forward rate curves, in which the coefficients F and G are constant and the differential operator A is of the first order, that do not fit the present framework, see [28].

Preliminary version of the paper appeared as Warwick Preprint:9/2001, May 2001 under the title *Wong- Zakai approximations of stochastic evolution equations.*

2. Preliminaries

2.1. **Analytic semigroups.** To state our results we need to fix some notations. The norm and the scalar products on H are denoted by $|\cdot|$ and $\langle\cdot,\cdot\rangle$. We fix constants M and a such that

$$|S(t)| \leq Me^{at}, \quad t \geq 0.$$

By V_α, $\alpha \in (0,1)$, we denote the domains $D(\lambda I - A)^\alpha$ of the fractional powers $(\lambda I - A)^\alpha$ where λ is any fixed number greater than ω. The set V_α is a separable Hilbert space with the norm

$$|x|_\alpha = (|x|^2 + |(\lambda I - A)^\alpha x|^2)^{1/2}, \quad x \in V_\alpha.$$

We will frequently use the following estimates for analytic semigroups valid for suitable constants C_α, $C_{\alpha,\beta}$ and $t \in]0,T]$:

$$\begin{aligned} &|S(t)x|_\alpha \leq C_{\alpha,\beta} t^{\beta-\alpha}|x|_\beta \quad \text{for all } 0 \leq \beta \leq \alpha \leq 1,\ x \in V_\beta \\ &|S(t)x - x| \leq C_\alpha t^\alpha |x|_\alpha \quad \text{for all } \alpha \in [0,1],\ x \in V_\alpha \end{aligned} \tag{2.1}$$

2.2. **A class of Gateâux differentiable functions.** Stochastic evolution equations which one often meets in applications have coefficients which are not Fréchet differentiable. Following [8] and [42] we introduce a class of maps acting among Banach spaces, possessing regularity properties weaker than Fréchet differentiability. This class is sufficiently large and includes operators commonly used as Nemytskii (evaluation) operators. It is well known that the Nemytskii operators are Fréchet differentiable only in trivial cases.

Let U, V, Z denote Banach spaces. We recall that for a mapping $\Phi : U \to V$ the directional derivative at point $x \in U$ in the direction $h \in U$ is defined as

$$\nabla\Phi(x;h) = \lim_{s\to 0} \frac{\Phi(x+sh)-\Phi(x)}{s},$$

whenever the limit exists in the topology of V. The mapping Φ is called Gâteaux differentiable at point x if it has directional derivative in every direction at point x and there exists an element of $L(U,V)$, denoted $\nabla\Phi(x)$ and called Gâteaux derivative, such that $\nabla\Phi(x;h) = \nabla\Phi(x)h$ for every $h \in U$. We say that a mapping $\Phi : U \to V$ *belongs to the class* $\mathcal{G}^1(U;V)$ if it is continuous, Gâteaux differentiable on U, and $\nabla\Phi : U \to L(U,V)$ is strongly continuous.
The last requirement of the definition means that for every $h \in U$ the map $\nabla\Phi(\cdot)h : U \to V$ is continuous. Note that $\nabla\Phi : U \to L(U,V)$ is, in general, not continuous if $L(U,V)$ is endowed with the norm operator topology. If this happens then Φ is Fréchet differentiable on U. Some features of the class $\mathcal{G}^1(U,V)$ are collected below.

Lemma 2.1. *Suppose* $\Phi \in \mathcal{G}^1(U,V)$. *Then*

- **(i):** $(x,h) \mapsto \nabla\Phi(x)h$ *is continuous from* $U \times U$ *to* V*;*
- **(ii):** *If* $\Psi \in \mathcal{G}^1(V,Z)$ *then* $\Psi(\Phi) \in \mathcal{G}^1(U,Z)$ *and*

$$\nabla(\Psi(\Phi))(x) = \nabla\Psi(\Phi(x))\nabla\Phi(x).$$

- **(iii):** *For all* $x, h \in U$ *it holds* $\Phi(x+h) = \Phi(x) + \int_0^1 \nabla\Phi(x+\theta h)h d\theta.$

That a map belongs to $\mathcal{G}^1(U,V)$ may be often checked by an application of the following lemma.

Lemma 2.2. *A map* $\Phi : U \to V$ *belongs to* $\mathcal{G}^1(U,V)$ *provided the following conditions hold:*

- **(i):** *the directional derivatives* $\nabla\Phi(x;h)$ *exist at every point* $x \in U$ *and in every direction* $h \in U$*;*
- **(ii):** *the mapping* $\nabla\Phi(\cdot;\cdot) : U \times U \to V$ *is continuous.*

The proofs of the above lemmas are left to the reader

Remark 2.3. *Let* $\mathcal{O}$ *be a bounded open subset of* $\mathbb{R}^m$ *and let* $H = L^2(\mathcal{O})$. *Moreover let* $\psi \in C^1(\mathbb{R})$ *bounded and Lipschitz and define* $\Psi : H \to H$ *by* $\Psi(x)(\xi) = \psi(x(\xi))$, $x \in H$, $\xi \in \mathcal{O}$. *It is immediate to check that* Ψ *is of class* $\mathcal{G}^1(H)$ *with* $[\nabla\Psi(x)h](\xi) = \psi'(x(\xi))h(\xi)$ *but it is never Fréchet differentiable (unless* ψ *is affine). Moreover* Ψ *is never twice Gâteaux differentiable (again unless* ψ *is affine).*

2.3. **Probabilistic estimate.** We will need also the following lemma.

Lemma 2.4. *Let* K *be a Hilbert space and fix* $r \geq 1$. *There exists* c_r *such that for all* $\phi \in L^r(\Omega, L^2([a,b],K))$ *and* $0 \leq a < b$ *with* $\phi(s)$, $\mathcal{F}_{s_n}$*-measurable for* $s \in [s_n, s_n^+[$ *it holds*

$$\mathbb{E}\left|\int_a^b \phi(s)\dot\beta_n(s)ds\right|_K^{2r} \leq c_r \mathbb{E}\left(\int_a^b |\phi(s)|_K^2 ds\right)^r$$

Proof. We have:

$$\mathbb{E}\left|\int_a^b \phi(s)\dot\beta_n(s)ds\right|_K^{2r} = \mathbb{E}\left|\sum_{k=[2^na]}^{[2^nb]}\left(\int_{2^{-n}k\vee a}^{2^{-n}(k+1)\wedge b}\phi(s)ds\right)\dot\beta_n(2^{-n}k\vee a)ds\right|_K^{2r} \le$$
$$\le c_p\mathbb{E}\left(2^n\sum_{k=[2^na]}^{[2^nb]}\left|\int_{2^{-n}k\vee a}^{2^{-n}(k+1)\wedge b}\phi(s)ds\right|_K^2\right)^r \le$$
$$\le c_r\mathbb{E}\left(\sum_{k=[2^na]}^{[2^nb]}\int_{2^{-n}k\vee a}^{2^{-n}(k+1)\wedge b}|\phi(s)|_K^2ds\right)^r \le c_r\mathbb{E}\left(\int_a^b|\phi(s)|_K^2ds\right)^r$$

□

3. Boundedness of the approximations

The approximating sequence y_n can be regarded as a sequence of random variables taking values in some function spaces. In this section we show that the sequence is bounded in $p-$ moments in appropriate function spaces of regular trajectories. To state our first theorem we formulate first the required assumptions.

Let us recall that if F is a mapping from H into H then its Gateâux derivative at point $x \in H$ will be denoted by $\nabla F(x)$. If for each $h \in H$, the mapping $\nabla F(x)h$, $x \in H$ is continuous then, by definition, *F belongs to the class $\mathcal{G}^1(H)$.*

We will need the following conditions.

(A.1) F and $G^1, \dots, G^d$ belong to $\mathcal{G}^1(H)$ and

$$\sup_{x\in H}\left(|F(x)| + |\nabla F(x)| + \sum_{j=1}^d\left(|G^j(x)| + |\nabla G^j(x)|\right)\right) < +\infty\,.$$

(A.2) There exist $\alpha \in (\frac{1}{2}, 1)$, $\epsilon \in (\frac{\alpha}{2}, \frac{1}{2}]$ and $c > 0$ such that

$$|G^j(x)|_\epsilon \le c(1 + |x|_\alpha), \quad x \in V_\alpha, \quad j = 1, \dots, d.$$

(A.3) There exists a set $\Gamma \subset H$ and a constant $c > 0$ such that

$$G^j(x) \in \Gamma,\ \nabla G^j(x)G^j(x) \in \Gamma,\ \text{for } x \in H,\ j = 1, \dots, d$$

and

$$\sum_{j=1}^d |\nabla G^j(x)u - \nabla G^j(y)u| \le c|x - y|$$

for all $x, y \in H$ and $u \in \Gamma$.

Here is the main result of the present section.

Theorem 3.1. *Assume that the operator A generates an analytic C_0-semi-group and that conditions (A.1), (A.2) and (A.3) are satisfied. Then for all $p > 0$ and all $R > 0$*

$$\sup\left\{\mathbb{E}\left(t^{\alpha p}|y_n(t,x)|_\alpha^p\right);\ |x| \le R,\ n \in \mathbb{N},\ t \in [0,T]\right\} < +\infty.$$

Let us remark that the boundedness is formulated in stronger norms then the basic norm in H. To treat all initial conditions in H, also those which are not in V_α we had to introduce the mollifier $t^{\alpha p}$.

To simplify presentation we assume that $d = 1$ and identify W with β.

It is well known that under the assumptions of the theorem for all initial datum $x \in H$ there exists a unique adapted process $y \in L^2(\Omega; \mathcal{C}([0,T],H))$ solving equation (1.13) in the usual weak, or equivalently, mild sense that is

$$y(t) = S(t)x + \int_0^t S(t-s)F(y(s))ds + \int_0^t S(t-s)G(y(s))d\beta(s)$$

for all $t \in [0,T]$, $\mathbb{P}$-a.s.

Lemma 3.2. *Fix $n \in \mathbb{N}$. For all $x \in H$ there exists a unique measurable process $y_n : \Omega \to C([0,T],H)$, such that for all $t \in [0,T]$, $\mathbb{P}$-a.s.:*

$$y_n(t,\omega) = S(t)x + \int_0^t S(t-s)\left[F(y_n(s,\omega)) + G(y_n(s,\omega))\dot{\beta}_n(s,\omega)\right] ds. \tag{3.1}$$

Moreover $\sup_{t\in(0,T]} t^\alpha |y_n(t,\omega)|_\alpha < +\infty$, $\mathbb{P}$*-a.s. for all $\alpha \in (0,1)$. Finally for all $p > 0$*

$$\mathbb{E}\big(\sup_{t\in(0,T]} t^{\alpha p}|y_n(t,\omega)|_\alpha^p\big) < +\infty. \tag{3.2}$$

Proof. The existence of y_n verifying (3.1) and its measurability is a straight-forward consequence of standard fixed point arguments in $C([0,T],H)$, because F and G are uniformly Lipschitz. Moreover by (2.1) we get for a suitable constant C:

$$|y_n(t,\omega)|_\alpha \le t^{-\alpha}|x| + C\int_0^t (t-s)^{-\alpha}(1 + |\dot{\beta}_n(s,\omega)|)ds.$$

If we choose $p > (1-\alpha)^{-1}$ and $1/p + 1/q = 1$ then by Holder's inequality:

$$|y_n(t,\omega)|_\alpha^p \le t^{-\alpha p}|x| + Ct^{(1/q-\alpha)p}\int_0^T (1 + |\dot{\beta}_n(s,\omega)|)^p ds.$$

Since $\mathbb{E}\left(\int_0^T (1+|\dot{\beta}_n(s,\omega)|)^p ds\right) < +\infty$, the required inequality holds. It is also clear that if the claim holds for some $p > 0$ then it holds for all $0 < p' \le p$. □

Proof of Theorem 3.1. We start by noticing again that if the claim holds for some $p > 0$ then it holds for all $0 < p' \le p$. We introduce the following notation: for all $\gamma, p > 0$, and all $\phi \in L^p(\Omega, C(]0,T], V_\alpha))$ let

$$\|\phi\|_{\gamma,\alpha,p} = \sup\left\{t^\alpha e^{-\gamma t}\left(\mathbb{E}(|\phi(t)|_\alpha^p)\right)^{1/p} : t \in]0,T]\right\}, \tag{3.3}$$

and

$$\|\phi\|_{\alpha,p} = \sup\left\{t^\alpha \left(\mathbb{E}(|\phi(t)|_\alpha^p)\right)^{1/p} : t \in]0,T]\right\}. \tag{3.4}$$

By Lemma 3.2, $\|y_n\|_{\gamma,\alpha,p} < +\infty$. We claim that for all p large enough there exist γ large enough and a constant $\ell_{\gamma,\alpha,p,R}$ such that $\|y_n\|_{\gamma,\alpha,p} \le \ell_{\gamma,\alpha,p,R}$ for all $n \in \mathbb{N}$ and all $x \in H$ with $|x| \le R$. Relation (3.1) can be rewritten as:

$$\begin{aligned} y_n(t) &= S(t)x + \int_0^t S(t-s)F(y_n(s))ds + \int_0^{2^{-n}\wedge t} S(t-s)G(y_n(s))\dot{\beta}_n(0)ds+ \\ &\quad + \int_{2^{-n}\wedge t}^t S(t-s)G(y_n(s_n))\dot{\beta}_n(s)ds \\ &\quad + \int_{2^{-n}\wedge t}^t S(t-s)\left[G(y_n(s)) - G(y_n(s_n))\right]\dot{\beta}_n(s)ds \\ &:= S(t)x + I_n^1(t) + I_n^2(t) + I_n^3(t) + I_n^4(t) \end{aligned}$$

It is evident that $sup_{|x|\le R}\|S(\cdot)x\|_{\alpha,\gamma,p} < \infty$.

We start estimating the terms I_n^i, $i = 1, 2, 3, 4$.

In the following by c we denote a generic constant which value depends only on F, G, x, α, T and p but not on γ, t and n. Its value can change from line to line.

To start with by (2.1) we immediately get:

$$|I_n^1|_\alpha \leq c \int_0^t (t-s)^{-\alpha} ds \tag{3.5}$$

Coming now to I_2 we get:

$$\begin{aligned} \mathbb{E}|I_n^2(t)|_\alpha^p &\leq ct^{\alpha p}\mathbb{E}(|\dot{\beta}_n(0)|^p)\left(\int_0^{2^{-n}\wedge t}(t-s)^{-\alpha}ds\right)^p \\ &\leq ct^{\alpha p}2^{np/2}\left(\int_0^{2^{-n}\wedge t}(t-s)^{-\alpha}ds\right)^p \end{aligned}$$

If $t \leq 2^{-n+1}$ then:

$$t^{\alpha p}2^{np/2}\left(\int_0^{2^{-n}\wedge t}(t-s)^{-\alpha}ds\right)^p \leq t^{\alpha p}2^{np/2}\left(\int_0^{t}s^{-\alpha}ds\right)^p \leq c2^{-np/2}.$$

Moreover if $t > 2^{-n+1}$ then $t/(t-s) \leq 2$ for $s \leq 2^{-n}$ and therefore:

$$\begin{aligned} t^{\alpha p}2^{np/2}\left(\int_0^{2^{-n}\wedge t}(t-s)^{-\alpha}ds\right)^p &\leq c2^{np/2}\left(\int_0^{2^{-n}}\left(\frac{t}{t-s}\right)^\alpha ds\right)^p \\ &\leq c2^{np/2}\left(2^{-n}2^\alpha\right)^p \leq c2^{-np/2}2^{\alpha p}. \end{aligned}$$

We can therefore conclude that:

$$\|I_n^2\|_{\alpha,\gamma,p}^p \leq c \tag{3.6}$$

We examine now I_n^3. Clearly $I_n^3(t) = 0$ if $t \leq 2^{-n}$ We therefore assume that $t > 2^{-n}$. By Lemma 2.4 we get:

$$\begin{aligned} t^{\alpha p}e^{-\gamma pt}\mathbb{E}|I_n^3(t)|_\alpha^p &\leq t^{\alpha p}e^{-\gamma pt}\mathbb{E}\left(\int_{2^{-n}}^t |S(t-s)G(y_n(s_n)|_\alpha^2 ds\right)^{p/2} \leq \\ &\leq ct^{\alpha p}e^{-\gamma pt}\left(\int_0^{t/2}(t-s)^{-2\alpha}ds\right)^{p/2} \\ &\quad + ct^{\alpha p}e^{-\gamma pt}\mathbb{E}\left(\int_{t/2\vee 2^{-n}}^t (t-s)^{-2(\alpha-\epsilon)}(1+|y_n(s_n)|_\alpha)^2 ds\right)^{p/2} \leq \\ &\leq ct^{\alpha p}e^{-\gamma pt}t^{(1-2\alpha)p/2} \\ &\quad + ct^{\alpha p}e^{-\gamma pt}\mathbb{E}\left(\int_{t/2\vee 2^{-n}}^t (t-s)^{-2(\alpha-\epsilon)}(1+|y_n(s_n)|_\alpha^2)ds\right)^{p/2} \leq \\ &\leq ct^{p/2}e^{-\gamma pt} + ct^{\alpha p}e^{-\gamma pt}\left(\int_{t/2\vee 2^{-n}}^t (t-s)^{-2(\alpha-\epsilon)})ds\right)^{p/2} + \\ &\quad + ct^{\alpha p}e^{-\gamma pt}\mathbb{E}\left(\int_{t/2\vee 2^{-n}}^t (t-s)^{-2(\alpha-\epsilon)}|y_n(s_n)|_\alpha^2)ds\right)^{p/2} \end{aligned}$$

Clearly $t^{\alpha p}(t/2)^{-\alpha p}t^{p/2}$ is uniformly bounded when $t \in [0,T]$. Moreover since $\alpha - \epsilon < 1/2$ we have: $e^{-\gamma pt}t^{\alpha p}\left(\int_{t/2}^t (t-s)^{-2(\alpha-\epsilon)}ds\right)^{p/2} \leq c$.

We are left with:

$$t^{\alpha p}e^{-\gamma pt}\mathbb{E}\left(\int_{t/2\vee 2^{-n}}^{t}(t-s)^{-2(\alpha-\epsilon)}|y_n(s_n)|_\alpha^2)ds\right)^{p/2} =$$
$$t^{\alpha p}\mathbb{E}\left(\int_{t/2\vee 2^{-n}}^{t}(t-s)^{-2(\alpha-\epsilon)}s^{-2\alpha}e^{-2\gamma(t-s)}(s_n^{2\alpha}e^{-2\gamma s_n}|y_n(s_n)|_\alpha^2)ds\right)^{p/2}$$

where we have used the fact that, for all $s \geq 2^{-n}$, $s(s_n)^{-1} \leq 2$. Applying Holder inequality with exponents $p/2$ and q^* with $2/p + 1/q^* = 1$ we get:

$$\begin{aligned}
&t^{\alpha p}\mathbb{E}\left(\int_{t/2\vee 2^{-n}}^{t}(t-s)^{-2(\alpha-\epsilon)}s^{-2\alpha}e^{-2\gamma(t-s)}(s_n^{2\alpha}e^{-2\gamma s_n}|y_n(s_n)|_\alpha^2)ds\right)^{p/2} \leq \\
&\leq t^{\alpha p}\left(\int_{t/2}^{t}(t-s)^{-2q^*(\alpha-\epsilon)}s^{-2q^*\alpha}ds\right)^{p/2q^*}\|y_n\|_{\alpha,\gamma,p}\int_{t/2}^{t}e^{-\gamma p(t-s)}ds \\
&\leq c(\gamma p)^{-1}t^{p/2q^*(1-2q^*(-\alpha+\epsilon))}\|y_n\|_{\alpha,\gamma,p} \leq \frac{c}{\gamma}\|y_n\|_{\alpha,\gamma,p}
\end{aligned}$$

where we have chosen p large enough so that $2q^*(\alpha-\epsilon) < 1$ (notice that by our assumptions $\alpha - \epsilon \leq \alpha/2 \leq 1/2$) Summarizing we again get, for p large enough:

$$\|I_n^3\|_{\alpha,\gamma,p}^p \leq c\left[1 + \gamma^{-1}\|y_n\|_{\alpha,\gamma,p}^p\right] \tag{3.7}$$

Finally we have to deal with I_n^4. Since $G \in \mathcal{G}^1$ we can write, for $t > 2^{-n}$

$$I_n^4(t) = \int_{2^{-n}}^{t}S(t-s)\Gamma_n(s)\left(y_n(s)-y_n(s_n)\right)\dot{\beta}_n(s_n)ds$$

where

$$\Gamma_n(s) = \int_0^1 \nabla G(y_n(s_n) + \theta(y_n(s)-y_n(s_n))d\theta. \tag{3.8}$$

Notice that for all $\omega \in \Omega$, $\Gamma_n(\cdot,\omega)$ is a strongly continuous map with values in $\mathcal{L}(H)$ such that $\sup\left\{|\Gamma_n(t,\omega)|_{\mathcal{L}(H)} : t \in [0,T], n \in \mathbb{N}, \omega \in \Omega\right\} < \infty$. Using the definition of mild solution of equation (1.14) with initial datum $y_n(s_n)$ at initial time s_n we get:

$$\begin{aligned}
I_n^4(t) &= \int_{2^{-n}}^{t}S(t-s)\Gamma_n(s)\left(S(s-s_n)-I\right)y_n(s_n)\dot{\beta}_n(s_n)ds+ \\
&\quad + \int_{2^{-n}}^{t}S(t-s)\Gamma_n(s)\left(\int_{s_n}^{s}S(s-\sigma)F(y_n(\sigma)d\sigma\right)\dot{\beta}_n(s_n)ds+ \\
&\quad + \int_{2^{-n}}^{t}S(t-s)\Gamma_n(s)\left(\int_{s_n}^{s}S(s-\sigma)G(y_n(\sigma)d\sigma\right)\left(\dot{\beta}_n(s_n)\right)^2 ds \\
&:= I_n^{4.1}(t) + I_n^{4.2}(t) + I_n^{4.3}(t)
\end{aligned}$$

We start again estimating $I^{4.1}$:

$$\begin{aligned}
&e^{-\gamma pt}t^{\alpha p}\mathbb{E}\left|I^{4.1}(t)\right|_\alpha^p \leq \\
&\leq ct^{\alpha p}\mathbb{E}\left(\int_{2^{-n}}^{t}((t-s)s_n)^{-\alpha}(s-s_n)^{\alpha}e^{\gamma(s_n-t)}\left[s_n^{\alpha}e^{-\gamma s_n}|y_n(s_n)|_\alpha\right]|\dot{\beta}_n(s_n)|ds\right)^p
\end{aligned}$$

Applying again Holder inequality with $1/p + 1/q + 1/r = 1$ and choosing q such that $\alpha q < 1$ and recalling again that $s/s_n \le 2$ we get:

$$\begin{aligned} & e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left| I^{4.1}(t) \right|_\alpha^p \le \\ & \le c t^{\alpha p} \left(\int_0^t ((t-s)s)^{-\alpha q} ds \right)^{p/q} \gamma^{-p/r} \times \\ & \qquad \times \mathbb{E} \int_{2^{-n}}^t (s-s_n)^{\alpha p} \left[s_n^{\alpha p} e^{-\gamma s_n p} |y_n(s_n)|_\alpha^p \right] \left(\dot\beta_n(s_n) \right)^p ds. \end{aligned}$$

Recalling that random variables $\dot\beta_n(s_n)$ and $y_n(s_n)$ are independent:

$$\begin{aligned} e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left| I^{4.1}(t) \right|_\alpha^p & \le c t^{(1/q-\alpha)p} \gamma^{-p/r} 2^{np/2} \|y_n\|_{\alpha,\gamma,p}^p \int_0^t (s-s_n)^{\alpha p} ds \\ & \le c \gamma^{-\frac{p}{r}} t^{1+(1/q-\alpha)p} 2^{-np(\alpha-1/2)} \|y_n\|_{\alpha,\gamma,p}^p. \end{aligned}$$

Choosing again p, q and r so that $1/q \ge \alpha$ and $r \le p$ we can conclude:

$$e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left| I^{4.1}(t) \right|_\alpha^p \le c \gamma^{-1} \|y_n\|_{\alpha,\gamma,p}^p \tag{3.9}$$

Moreover:

$$\begin{aligned} e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left| I^{4.2}(t) \right|_\alpha^p & \le e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left(\int_{2^{-n}}^t (t-s)^{-\alpha} (s-s_n) |\dot\beta_n(s_n)| ds \right)^p \le \\ & \le c t^{\alpha p} 2^{np/2} \left(\int_0^t (t-s)^{-\alpha q} \right)^{p/q} \int_0^t (s-s_n)^p ds \le c t^{p/q} 2^{-np/2} \end{aligned} \tag{3.10}$$

Finally:

$$\begin{aligned} e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left| I^{4.3}(t) \right|_\alpha^p & \le e^{-\gamma p t} t^{\alpha p} \mathbb{E} \left(\int_{2^{-n}}^t (t-s)^{-\alpha} (s-s_n) \left(\dot\beta_n(s_n) \right)^2 ds \right)^p \le \\ & \le c t^{\alpha p} 2^{np} \left(\int_0^t (t-s)^{-\alpha q} \right)^{p/q} \int_0^t (s-s_n)^p ds \le c t^{p/q} \end{aligned} \tag{3.11}$$

By (3.5), (3.6), (3.7), (3.9), (3.10) and (3.11) we get:

$$\|y_n\|_{\alpha,\gamma,p}^p \le c \left(1 + \gamma^{-1} \|y_n\|_{\alpha,\gamma,p}^p \right)$$

where c does not depend on γ. The claim follows choosing γ large enough. □

4. Tightness of the approximations

We pass now to the tightness of the laws of the approximations and show the following theorem.

Theorem 4.1. *Assume in addition to the conditions of Theorem 3.1 that for each $t > 0$, $S(t)$ is a compact operator from H into H. Then for each x, the laws $\mathcal{L}(y_n(\cdot, x))$, $n = 1, 2, \dots$ are tight on $C(0, T; H)$. Moreover*

$$\sup_n \mathbb{E} \left(\sup_{t \in [0,T]} |y_n(t,x)| \right) < +\infty$$

Proof of Theorem 4.1. Let $C_x([0,T], H)$ be the space of continuous functions $\phi : [0,T] \to H$ such that $\phi(0) = x$.

We write y_n as in (3.1). We only show that the sequence of laws of the processes:

$$\tilde y_n(t) = \int_0^t S(t-s) G(y_n(s)) \dot\beta_n(s) ds$$

is tight in $C_x([0,T], H)$ and

$$\sup_n \mathbb{E} \left(\sup_{t \in [0,T]} \left| \int_0^t S(t-s) G(y_n(s)) \dot\beta_n(s) ds \right|^p \right) < \infty.$$

The proof of the same property for the deterministic convolution

$$\int_0^t S(t-s)F(y_n(s))ds$$

is similar and easier.

We use the factorization formula, see [7], and write $\tilde{y}_n(t)$ as:

$$\tilde{y}_n(t) = \Lambda_{\tilde{\alpha}} Y_{n,\tilde{\alpha}}(t), \tag{4.1}$$

where for all $\tilde{\alpha} \in (0,1)$ and all $f \in L^p([0,T],H)$ with $1/p < \tilde{\alpha}$:

$$\Lambda_{\tilde{\alpha}} f(t) = \int_0^t (t-s)^{\tilde{\alpha}-1} S(t-s) f(s) ds$$

and

$$Y_{n,\tilde{\alpha}}(t) = \int_0^t (t-s)^{-\tilde{\alpha}} S(t-s) G(y_n(s)) \dot{\beta}_n(s) ds$$

(the proof of (4.1) follows by Fubini-Tonelli theorem, see [7]). Since, see Proposition 8.4 in [7], operators $\Lambda_{\tilde{\alpha}}$ are compact from $L^p([0,T],H)$ to $C([0,T],H)$ it is enough to show that there exist p, $\tilde{\alpha}$ and ℓ with $1/p < \tilde{\alpha} < 1/2$ and:

$$\mathbb{E} \int_0^T |Y_{n,\tilde{\alpha}}(t)|^p dt \le \ell \qquad \forall n \in \mathbb{N}$$

Again we divide

$$\begin{aligned} Y_{n,\tilde{\alpha}}(t) &= \int_0^{2^{-n}\wedge t} (t-s)^{-\tilde{\alpha}} S(t-s) G(y_n(s)) \dot{\beta}_n(0) ds \\ &\quad + \int_{2^{-n}\wedge t}^t (t-s)^{-\tilde{\alpha}} S(t-s) G(y_n(s_n)) \dot{\beta}_n(s) ds \\ &\quad + \int_{2^{-n}\wedge t}^t (t-s)^{-\tilde{\alpha}} S(t-s) \left[G(y_n(s_n)) - G(y_n(s)) \right] \dot{\beta}_n(s) ds \\ &= I_n^1(t) + I_n^2(t) + I_n^3(t) \end{aligned}$$

In the following by ℓ we denote a generic constant and its value can change from line to line. We have:

$$\begin{aligned} \mathbb{E} \int_0^T |I_n^1(t)|^p dt &\le \ell \int_0^T \left(2^{n/2} \int_0^{2^{-n}\wedge t} (t-s)^{-\tilde{\alpha}} ds \right)^p dt \\ &\le \ell \int_0^T \left(2^{n/2} \int_{(t-2^{-n})\vee 0}^t \sigma^{-\tilde{\alpha}} d\sigma \right)^p dt \\ &\le \ell T \left(2^{n/2} \int_0^{2^{-n}} \sigma^{-\tilde{\alpha}} d\sigma \right)^p \le \ell T 2^{np/2} 2^{-n(1-\tilde{\alpha})p} \le \ell \end{aligned}$$

since $1 - \tilde{\alpha} > 1/2$. Moreover by Lemma 2.4

$$\mathbb{E} \int_{2^{-n}}^T |I_n^2(t)|^p dt \le \ell \mathbb{E} \int_0^T \left(\int_{2^{-n}}^t (t-s)^{-2\tilde{\alpha}} ds \right)^{p/2} \le \ell \int_0^T t^{(1-2\tilde{\alpha})p/2} dt \le \ell$$

We come to the estimate of I_n^3. We choose $\tilde\alpha$ such that $\tilde\alpha < 1-\alpha$ and $p > \tilde\alpha^{-1} \vee \overline{p}$ with $\overline{p}$ selected similarly as in the proof of Theorem 3.1. We again split the expression:

$$I_n^3(t) = \int_{2^{-n}}^t (t-s)^{-\tilde\alpha} S(t-s)\Gamma_n(s)\left(S(s-s_n)-I\right) y_n(s_n)\dot\beta_n(s_n)ds+$$
$$+\int_{2^{-n}}^t (t-s)^{-\tilde\alpha} S(t-s)\Gamma_n(s)\left(\int_{s_n}^s S(s-\sigma)F(y_n(\sigma)d\sigma\right)\dot\beta_n(s_n)ds+$$
$$+\int_{2^{-n}}^t (t-s)^{-\tilde\alpha} S(t-s)\Gamma_n(s)\left(\int_{s_n}^s S(s-\sigma)G(y_n(\sigma)d\sigma\right)\left(\dot\beta_n(s_n)\right)^2 ds$$
$$:= I_n^{3.1}(t) + I_n^{3.2}(t) + I_n^{3.3}(t)$$

We proceed as in the proof of Theorem 3.1.

$$\mathbb{E}|I_n^{3.1}(t)|^p \le \ell\mathbb{E}\left(\int_{2^{-n}}^t (t-s)^{-\tilde\alpha}(s-s_n)^\alpha s_n^{-\alpha}\left[s_n^\alpha|y_n(s_n)|_\alpha\right]|\dot\beta_n(s_n)|ds\right)^p$$
$$\le \ell\left(\int_{2^{-n}}^t (t-s)^{-\tilde\alpha q}s_n^{-\alpha q}ds\right)^{p/q}\int_{2^{-n}}^t (s-s_n)^{\alpha p}\mathbb{E}\left[s_n^{\alpha p}|y_n(s_n)|_\alpha^p|\dot\beta_n(s_n)|^p\right]ds$$

Since for $s \ge 2^{-n}$, $s/s_n \le 2$ and $y_n(s_n)$ is independent from $\dot\beta_n(s_n)$ we get by Theorem 3.1 and definition (3.4)

$$\begin{aligned}\mathbb{E}|I_n^{3.1}(t)|^p &\le \ell\left(\int_{2^{-n}}^t (t-s)^{-\tilde\alpha q}s^{-\alpha q}\right)^{p/q}\left(\int_{2^{-n}}^t (s-s_n)^{\alpha p}ds\right)2^{np/2}\|y_n\|_{\alpha,p}^p\\ &\le c\ell t^{p/q}t^{-(\alpha+\tilde\alpha)p}t2^{-n\alpha p}2^{np/2} \le c\ell t^{p(1-\alpha-\tilde\alpha)}.\end{aligned}$$

Since $\alpha+\tilde\alpha \le 1$ this implies $\mathbb{E}|I_n^{3.1}(t)|^p \le \ell$ for all $t \in [0,T]$. To complete the proof we estimate $\int_0^T \mathbb{E}|I_n^{3.3}(t)|^p dt$. The estimate of the term $\int_0^T \mathbb{E}|I_n^{3.2}(t)|^p dt$ is similar and easier.
It holds

$$\begin{aligned}\int_0^T \mathbb{E}|I_n^{3.3}(t)|^p dt &\le \ell\mathbb{E}\left(\int_{2^{-n}}^t (t-s)^{-\tilde\alpha}(s-s_n)|\dot\beta_n(s_n)|^2ds\right)^p\\ &\le \left(\int_{2^{-n}}^t (t-s)^{-\tilde\alpha q}ds\right)^{p/q}\left(\int_{2^{-n}}^t (s-s_n)^p ds\right)2^{np}\\ &\le \ell t^{1+(1-\tilde\alpha q)p/q} \le \ell\end{aligned}$$

since $\tilde\alpha \le 1/2 \le 1/q = 1-1/p$ and this completes the proof. □

5. Convergence to the solution of the modified equation

Here is the main result of the paper

Theorem 5.1. *Under the conditions of Theorem 4.1, the laws $\mathcal{L}(y_n(\cdot,x))$ converge weakly on $C(0,T;H)$ to the law of the solution of the following evolution equation*

$$dy(t) = \left(Ay(t) + F(y(t)) + \frac{1}{2}\sum_{j=1}^d \nabla G^j(y(t))G^j(y(t))\right)dt + \sum_{j=1}^d G^j(y(t))d\beta^j(t),\quad y(t) = x. \tag{5.1}$$

To prove the theorem we need a result on the solution of the martingale problem in infinite dimensions of independent interest.

5.1. **Martingale problem and weak solutions of evolution equations.** We recall first basic definitions related to the martingale problem and establish a uniqueness result needed to prove the convergence Theorem 5.1.

Denote by $\nu(t)$, $t \geq 0$, the canonical process on $C_x([0,T],H)$:

$$\nu(t,\omega) = \omega(t), \quad t \geq 0, \quad \omega \in C_x([0,T],H),$$

and let $\mathcal{F}_t = \sigma\{\nu(\tau) : \tau \leq t\}$, $t \geq 0$, $\mathcal{F} = \sigma\{\nu(\tau) : \tau \geq 0\}$ be the canonical $\sigma-$ field. Let $F : H \to H$ and $G : H \to L(\mathbb{R}^d, H)$ be given measurable functions and A the infinitesimal generator of a C_0-semi-group $S(t)$, $t \geq 0$, on H. Coordinate functions of G are denoted respectively as $G^1, \ldots, G^d$.

We say that $\varphi : C_x([0,T],H) \to \mathbb{R}$ is a *regular, cylindrical function* if there exist a natural number m, a C^∞-function ψ on $\mathbb{R}^m$ with compact support, and elements $a_1, a_2, \ldots, a_m \in D(A^*)$ such that

$$\varphi(x) = \psi\big(\langle x, a_1\rangle, \ldots, \langle x, a_m\rangle\big). \tag{5.2}$$

If φ is a regular, cylindrical function, then we set

$$\begin{aligned} &\mathcal{L}\varphi(x) \\ &= \langle x, A^* D_x\varphi(x)\rangle + \langle F(x), D_x\varphi(x)\rangle + \frac{1}{2}\operatorname{Trace} G^*(u) D^2_{xx}\varphi(u) G(x) \\ &= \sum_{k=1}^m \frac{\partial\psi}{\partial\xi_k}\big(\langle x, a_1\rangle, \ldots, \langle x, a_d\rangle\big)\big[\langle x, A^* a_k\rangle + \langle F(x), a_k\rangle\big] \\ &\quad + \frac{1}{2}\sum_{n=1}^d \sum_{k,l=1}^m \frac{\partial^2\psi}{\partial\xi_k\,\partial\xi_l}\big(\langle x, a_1\rangle, \ldots, \langle x, a_d\rangle\big)\langle a_k, G^n(x)\rangle\langle a_l, G^n(x)\rangle \end{aligned} \tag{5.3}$$

A probability measure $\mathbb{Q}$ on $\big(C_x([0,T],H), \mathcal{F}\big)$ is a solution to martingale problem with parameters (x, A, F, GG^*) if the process

$$\varphi(\nu(t)) - \int_0^t \mathcal{L}\varphi(\nu(s))\,ds, \quad t \geq 0, \tag{5.4}$$

is an $\mathcal{F}_t$-martingale on $\big(C_x([0,T],H), \mathcal{F}, \mathbb{Q}\big)$ for an arbitrary regular cylindrical function.

In the formulation of the theorem below, $\nu^d(t)$, $t \geq 0$, is the canonical process on $C_0([0,T],\mathbb{R}^d)$:

$$\nu^d(t,\omega^d) = \omega^d(t), t \geq 0, \omega^d \in C_0([0,T],\mathbb{R}^d),$$

and $\mathcal{G}_t = \sigma\{\nu^d(\tau) : \tau \leq t\}$, $\mathcal{G} = \sigma\{\nu^d(\tau) : \tau \geq 0\}$.

Theorem 5.2. *Assume that $F, G^1, \ldots, G^d$ are Lipschitz functions and a probability measure $\mathbb{Q}$ on $\big(C_x([0,T],H), \mathcal{F}\big)$ is a solution to the martingale problem (x, A, F, GG^*). Then there exists a d-dimensional Wiener process $\hat{W}(t)$, $t \geq 0$, on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{Q}})$, $\hat{\Omega} = C_x([0,T],H) \times C_0([0,T],\mathbb{R}^d)$, $\hat{\mathcal{F}} = \mathcal{F} \otimes \mathcal{G}$, $\hat{\mathbb{Q}} = \mathbb{Q} \times W$, such that the process $\nu(t)$, $t \geq 0$, extended from Ω to $\hat{\Omega}$ by the formula*

$$\nu(t, (\omega, \omega^d)) = \nu(t, \omega), \quad t \geq 0, \ (\omega, \omega^d) \in \hat{\Omega},$$

is a weak solution of the stochastic Itô equation

$$d\nu(t) = \big(A\nu(t) + F(\nu(t))\big)\,dt + G(\nu(t))\,d\hat{W}(t), \quad \nu(0) = x. \tag{5.5}$$

It is well known that under the conditions of Theorem 5.2 the equation (5.5) has a unique weak solution, see e.g. [7], therefore the measure $\mathbb{Q}$ is identical with the law of the solution and we have the following corollary.

Theorem 5.3. *Under the conditions of Theorem 5.2 the solution to the martingale problem (x, A, F, GG^*) is unique.*

Proof of Theorem 5.2. Let us choose an orthonormal and complete basis (e_m) in H composed of vectors from $D(A^*)$ and such that the set $\{e_m : m \in \mathbb{N}\}$ is linearly dense in $D(A^*)$ equipped with the graph norm. It is enough to construct a Wiener process $\hat{\omega}(t)$, $t \geq 0$, such that for each m

$$d\langle e_m, \nu(t)\rangle = \big[\langle A^* e_m, \nu(t)\rangle + \langle e_m, F(\nu(t)\rangle\big]\, dt + \langle e_m, G(\nu(t))\, d\hat{\omega}(t)\rangle.$$

Define $x_m = \langle e_m, x\rangle$, $\nu_m(t) = \langle e_m, \nu(t)\rangle$,

$$M_m(t) = \nu_m(t) - x_m - \int_0^t \big[\langle A^* e_m, \nu(s)\rangle + \langle e_m, F(\nu(s))\rangle\big]\, ds,$$
$$n = 1, 2, \ldots, \quad t \geq 0,$$

and let $\psi_R \in C_0^\infty(\mathbb{R})$ be such that

$$\psi_R(\sigma) = \sigma, \quad \text{if } |\sigma| \leq R.$$

Considering martingales (5.4) for $\varphi_1(x) = \psi_R\big(\langle e_m, x\rangle\big)$, $\varphi_2(x) = \psi_R\big(\langle e_m, x\rangle\big)\psi_R\big(\langle e_l, x\rangle\big)$, $x \in H$, and taking into account that ,

$$\varphi_1(\nu(t)) = \nu_m(t), \quad \varphi_2(\nu(t)) = \nu_m(t)\nu_l(t),$$
$$\text{for } t \leq \tau_R = \inf\{t \geq 0 : |\nu(t)| \geq R\},$$

one can easily show that processes M_m, $m = 1, 2, \ldots$, are local martingales, with the quadratic covariation given by the formulae

$$\langle\langle M_m, M_l\rangle\rangle_t = \int_0^l \langle G^*(\nu(s))e_m, G^*(\nu(s))e_l\rangle\, ds, \quad t \geq 0.$$

For arbitrary natural $k \leq l$ define

$$M_{k,l}(t) = \sum_{j=k}^{m} M_j(t) e_j, \quad t \geq 0.$$

Then $M_{k,l}$ is an H-valued local martingale with continuous paths. By Doob's maximal inequality

$$\mathbb{E} \sup_{0\leq t\leq T} \big|M_{k,l}(t \wedge \tau_R)\big|^2 \leq 4 \sup_{0\leq t\leq T} \mathbb{E}\big|M_{k,l}(t \wedge \tau_R)\big|^2$$
$$\leq 4\mathbb{E} \int_0^T \left(\sum_{j=k}^{l} \big|G^*(\nu(s \wedge \tau_R))e_j\big|^2\right) ds, \quad k \leq l.$$

However,

$$\mathbb{E} \int_0^T \left(\sum_{j=k}^{\infty} \big|G^*(\nu(s \wedge \tau_R)e_j)\big|^2\right) ds \leq T \sup_{y\in H} \|G(y)\|^2_{L(H,H)} < +\infty.$$

Consequently,

$$\lim_{k,l\to\infty} \mathbb{E} \sup_{0\leq t\leq T} \big|M_{k,l}(t \wedge \tau_R)\big|^2 \to 0$$

and therefore the series

$$M(t) = \sum_{j=1}^{+\infty} M_j(t) e_j, \quad t \geq 0,$$

defines an H-valued, continuous local martingale. Since the processes $M_{k,l}(t \wedge \tau_R)$, $t \geq 0$, are square integrable martingales, the process $M(t)$, $t \geq 0$, is a continuous, local square integrable martingale with the square bracket

$$\langle\langle M, M\rangle\rangle_t = \int_0^t \left(\sum_{k=1}^{d} G^k(\nu(s)) \otimes G^k(\nu(s))\right) ds, \quad t \geq 0,$$

where $a \otimes b$ is the one-dimensional operator $a \otimes b(y) = a\langle b, y\rangle$, $y \in H$.

The Wiener process $\hat{W}$ from the formulation of the theorem is constructed similarly to the classical finite-dimensional case, see vol. 2, V.20 of [32]. Let, for each $y \in H$, $G^{-1}(y)$ be the pseudoinverse of $G(y)$ acting from $\operatorname{Range} G(y)$ into $\mathbb{R}^d$ and $\pi(y)$ the orthogonal projection of H onto $\operatorname{Range} G(y)$. Since the range of $G(y)$ is a finite-dimensional space, the operator $\Theta(y) = G^{-1}(y)\pi(y)$ is well defined, linear bounded operator for each $y \in H$. Moreover, the operators

$$G(y)\Theta(y) = G(y)G^{-1}(y)\pi(y) = \pi(y) \tag{5.6}$$

$$\Theta(y)G(y) = G^{-1}(y)\pi(y)G(y) = \pi_1(y) \tag{5.7}$$

are orthogonal projectors and

$$G(y)\Theta(y)G(y) = \pi(y)G(y) = G(y) \tag{5.8}$$

$$\Theta(y)G(y)\Theta(y) = \Theta(y)\pi(y) = \Theta(y), \quad y \in H. \tag{5.9}$$

Define

$$\rho(y) = I - \pi_1(y), \quad y \in H. \tag{5.10}$$

Then $\rho(y)$ is an orthogonal projector as well.
We finally define

$$\hat{W}(t) = \int_0^t \Theta(\nu(s))\, dM(s) + \int_0^t \rho(\nu(s))\, dW(s)$$

where W is a d-dimensional, canonical Wiener process on $\big(C_0([0,T],\mathbb{R}^d), \mathcal{G}\big)$ trivially extended to $\hat{\Omega}$. The same calculations as in finite dimensions based on (5.6)–(5.10) show that

$$\langle\langle \hat{W}, \hat{W}\rangle\rangle_t = I,$$

$$\Big\langle\Big\langle M(\cdot) - \int_0^\cdot G(\nu(s))\, d\hat{W}(s), M(\cdot) - \int_0^\cdot G(\nu(s))\, d\hat{W}(s)\Big\rangle\Big\rangle_t = 0, \quad t \geq 0,$$

and therefore $\hat{W}$ is a d-dimensional Wiener process. Moreover,

$$\int_0^t G(\nu(s))\, d\hat{W}(s) = \int_0^t G(\nu(s))\Theta(\nu(s))\, dM(s) + \int_0^t G(\nu(s))\rho(\nu(s))\, dW(s) = M(t).$$

This completes the proof of the theorem. □

5.2. **Convergence of the approximations.** We go back to the proof of Theorem 5.1.
Let $\mathbb{Q}_n$ be the measure induced on $C_x([0,T],H)$ by $y_n(\cdot)$ and assume that $\mathbb{Q}_n \rightharpoonup \mathbb{Q}$ weakly on $C_x([0,T],H)$ as $n \to \infty$. To identify the limiting measure $\mathbb{Q}$ we compute

$$\mathbb{E}^{\mathbb{Q}}\left(h \cdot (f(\nu(t)) - f(\nu(s)))\right) \quad \text{with } t > s > 0$$

where $\nu(t)$, $t \geq 0$ is the canonical process in $C_x([0,T],H)$, h is a generic bounded, continuous, $\mathcal{F}_s = \sigma\{\nu(\tau) : \tau \leq s\}$ measurable function on $C_x([0,T],H)$ and we assume that f is of the form

$$f(y) = \varphi(\langle v_1, y\rangle, \ldots, \langle v_r, y\rangle); \quad v_1, \ldots, v_r \in D(A^*),\ \varphi \in C_c^\infty(\mathbb{R}^r). \tag{5.11}$$

In the following proofs we denote by $\mathbb{E}^{\mathbb{Q}}$, respectively $\mathbb{E}^{\mathbb{Q}_n}$, the expectation with respect to measure $\mathbb{Q}$, respectively $\mathbb{Q}_n$, on $C_x([0,T],H)$ and by $\mathbb{E}$ the expectation with respect to the original probability measure$\mathbb{P}$.

Theorem 5.1 is an immediate corollary of the following proposition.

Proposition 5.4. *For all f of the form (5.11), all $0 \le s \le t$ and all h bounded and $\mathcal{F}_s$ measurable it holds:*

$$\mathbb{E}^{\mathbb{Q}}\left[h\cdot(f(\nu(t))-f(\nu(s)))\right]=\frac{1}{2}\mathbb{E}^{\mathbb{Q}}\left[h\cdot\int_s^t\langle D^2 f(\nu(\tau))G(\nu(\tau)),G(\nu(\tau))\rangle d\tau\right]$$
$$+\mathbb{E}^{\mathbb{Q}}\Big[h\cdot\int_s^t\sum_{i=1}^r\frac{\partial\varphi}{\partial v_i}(\langle v_1,\nu(\tau)\rangle,\ldots,\langle v_r,\nu(\tau)\rangle)\times$$
$$\times\left(\langle A^* v_i,\nu(\tau)\rangle+\langle v_i,F(\nu(\tau))\rangle\right)d\tau\Big]$$
$$+\quad\frac{1}{2}\mathbb{E}^{\mathbb{Q}}\left[h\cdot\int_s^t Df(\nu(\tau))\nabla G(\nu(\tau))G(\nu(\tau))d\tau\right].$$

Thus $\mathbb{Q}$ is the law of the unique solution of equation (5.1)

The following lemma will be frequently used in the proof of the proposition.

Lemma 5.5. *If $\mu_n \rightharpoonup \mu$ weakly and $\exists c>0$ such that $\int_H |x|^{1+\epsilon}\mu_n(dx)\le c$, $n=1,2,\ldots$ then $\forall\psi\in C(H)$ with $|\psi(x)|\le C(1+|x|)$, $x\in H$, it holds:*

$$\int_H\psi(x)\mu_n(dx)\to\int_H\psi(x)\mu(dx).$$

Proof. Note first that $\int_H |x|^{1+\epsilon}\mu(dx)\le c$. Define $\psi_N=\psi I_{\{|x|\le N\}}$. Then

$$\int_H\psi(x)\mu_m(dx)-\int_H\psi(x)\mu(dx)=$$
$$=\int_H\psi_N\left(\mu_m-\mu\right)(dx)+\int_H(\psi-\psi_N)\mu_m(dx)+\int_H(\psi_N-\psi)\mu(dx)$$

The first and the third term converge to 0. The first, for all N fixed, as $m\to\infty$ and the third as $N\to\infty$. We have to prove that the second term goes to 0 as $N\to\infty$ uniformly in m.

$$\left|\int_H(\psi_N-\psi)\mu_m(dx)\right|\le C\int_H|x|I_{\{|x|\ge N\}}\mu_m(dx)\le C\left[\mu_m(\{|x|>N\})\right]^{1/q}$$

with $\frac{1}{1+\epsilon}+\frac{1}{q}=1$. However $\mu_m\{|x|>N\}\le\frac{1}{N}\int|x|\mu_m(dx)\to 0$, as $N\to\infty$ uniformly in m, and the result follows. □

Proof. We pass to the proof of Proposition 5.4.

Let $0\le s\le t$,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\quad&\left[h\cdot(f(\nu(t))-f(\nu(s)))\right]\\
&=\lim_{n\to\infty}\mathbb{E}^{\mathbb{Q}_n}\left[h\cdot(f(\nu(t))-f(\nu(s)))\right]\\
&=\lim_{n\to\infty}\mathbb{E}\left[h\cdot\int_s^t Df(y_n(\tau))\dot{y}^{(n)}(\tau)d\tau\right]\\
&=\lim_{n\to\infty}\mathbb{E}\Big[h\cdot\Big[\int_s^t\sum_{i=1}^r\frac{\partial\varphi}{\partial v_i}(\langle v_1,y_n(\tau)\rangle,\ldots,\langle v_r,y_n(\tau)\rangle)\langle A^*v_i,y_n(\tau)\rangle d\tau\\
&\qquad+\int_s^t\sum_{i=1}^r\frac{\partial\varphi}{\partial v_i}(\langle v_1,y_n(\tau)\rangle,\ldots,\langle v_r,y_n(\tau)\rangle)\langle v_i,F(y_n(\tau))\rangle d\tau\\
&\qquad+\int_s^t Df(y_n(\tau))G(y_n(\tau))\dot{\beta}_n(\tau)d\tau\Big]\Big]\end{aligned}$$

By Lemma 5.5 the first two terms converge to

$$
\begin{aligned}
= \mathbb{E}^{\mathbb{Q}}\Big[h \cdot \Big[\int_s^t \sum_{i=1}^r \frac{\partial \varphi}{\partial v_i}(\langle v_1, \nu(\tau)\rangle, \ldots, \langle v_r, \nu(\tau)\rangle)\langle A^* v_i, \nu(\tau)\rangle d\tau \\
+ \int_s^t \sum_{i=1}^r \frac{\partial \varphi}{\partial v_i}(\langle v_1, \nu(\tau)\rangle, \ldots, \langle v_r, \nu(\tau)\rangle)\langle v_i, F(\nu(\tau))\rangle d\tau \Big]\Big].
\end{aligned}
$$

We are therefore left with the limit of

$$
\begin{aligned}
\mathbb{E}\Big[h \cdot & \int_s^t Df(y_n(\tau))G(y_n(\tau))\dot{\beta}_n(\tau)d\tau\Big] \\
&= \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))G(y_n(\tau_n))\dot{\beta}_n(\tau)d\tau\right] \\
&\quad + \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))\left(G(y_n(\tau)) - G(y_n(\tau_n))\right)\dot{\beta}_n(\tau)d\tau\right] \\
&=: J_n^1 + J_n^2.
\end{aligned}
$$

But

$$
J_n^1 = \mathbb{E}\left[h \cdot \int_s^t \left[Df(y_n(\tau)) - Df(y_n(\tau_n))\right] G(y_n(\tau_n))\dot{\beta}_n(\tau)d\tau\right],
$$

because: $\mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau_n))G(y_n(\tau_n))\dot{\beta}_n(\tau)d\tau\right] = 0$. Therefore

$$
\begin{aligned}
J_n^1 \quad &= \mathbb{E}\Big[h \cdot \int_s^t \int_{\tau_n}^{\tau} \sum_{ij=1}^r \frac{\partial^2 \varphi}{\partial v_i \partial v_j}(\langle v_1, y_n(\zeta)\rangle, \ldots, \langle v_r, y_n(\zeta)\rangle)\langle v_i, G(y_n(\tau_n))\rangle \\
&\qquad \Big[\langle A^* v_j, y_n(\zeta)\rangle + \langle v_j, F(y_n(\zeta))\rangle \dot{\beta}_n(\tau_n)\Big] d\zeta d\tau\Big] \\
&+ \mathbb{E}\Big[h \cdot \int_s^t \int_{\tau_n}^{\tau} \sum_{ij=1}^r \frac{\partial^2 \varphi}{\partial v_i \partial v_j}(\langle v_1, y_n(\zeta)\rangle, \ldots, \langle v_r, y_n(\zeta)\rangle) \\
&\qquad \langle v_i, G(y_n(\tau_n))\rangle \langle v_j, G(y_n(\zeta))(\dot{\beta}_n(\sigma_n))^2 d\zeta d\tau\Big] \\
&=: J_n^{1.1} + J_n^{1.2}.
\end{aligned}
$$

We notice that

$$
J_n^{1.1} = \mathbb{E}h \cdot \int_s^t \left(\int_{\tau_n}^{\tau} \ell_n(\zeta, \tau_n) d\zeta\right) \dot{\beta}_n(\tau_n) d\tau
$$

and

$$
\begin{aligned}
|\ell_n(\zeta, \tau_n)| \quad &= \Big| \sum_{ij=1}^r \frac{\partial^2 \varphi}{\partial v_i \partial v_j}(\langle v_1, y_n(\zeta)\rangle, \ldots, \langle v_r, y_n(\zeta)\rangle) \\
&\qquad \langle v_i, G(y_n(\tau_n))\rangle \left[\langle A^* v_j, y_n(\zeta)\rangle + \langle v_j, F(y_n(\zeta))\rangle\right] \Big| \\
&\le c(1 + |y_n(\zeta)|).
\end{aligned}
$$

Consequently

$$\begin{aligned}
|J_n^{1.1}| &\le c\mathbb{E}\left(\int_s^t \left(\int_{\tau_n}^\tau (1+|y_n(\zeta)|)d\zeta\right)|\dot\beta_n(\tau_n)|d\tau\right)\\
&\le c\int_s^t \left(\mathbb{E}\left(\int_{\tau_n}^\tau (1+|y_n(\zeta)|)d\zeta\right)^p\right)^{1/p}\left(\mathbb{E}|\dot\beta_n(\tau_n)|^q\right)^{1/q}d\tau\\
&\le c2^{n/2}\int_s^t \left(\mathbb{E}\left(\int_{\tau_n}^\tau (1+|y_n(\zeta)|)d\zeta\right)^p\right)^{1/p}d\tau\\
&\le c2^{n/2}\int_s^t \left(\mathbb{E}\left((\tau-\tau_n)^{1/q}\left(\int_{\tau_n}^\tau (1+|y_n(\zeta)|)^p d\zeta\right)^{1/p}\right)^p\right)^{1/p}d\tau\\
&\le c2^{n/2}\int_s^t (\tau-\tau_n)^{1/q}\left(\int_{\tau_n}^\tau \mathbb{E}(1+|y_n(\zeta)|)^p d\zeta\right)^{1/p}d\tau\\
&\le c2^{n/2}\int_s^t (\tau-\tau_n)^{1+1/q}d\tau\\
&\le c2^{n/2}2^n(t-s)\left(\frac{1}{2^n}\right)^{2+\frac1q}=o(1)
\end{aligned}$$

where by $o(1)$ we denote a generic function of n, infinitesimal as $n\to\infty$.

$$\begin{aligned}
J_n^{1.2} &= \mathbb{E}\Big[h\cdot\int_s^t\int_{\tau_n}^\tau \sum_{i,j=1}^r \frac{\partial^2\varphi}{\partial v_i\partial v_j}(\langle v_1,y_n(\zeta)\rangle,\dots,\langle v_r,y_n(\zeta)\rangle)\\
&\qquad\qquad \langle v_i, G(y_n(\zeta))\rangle\langle v_j,G(y_n(\tau_n))\rangle(\dot\beta_n(\tau_n))^2 d\zeta d\tau\Big]\\
&= \mathbb{E}\left[h\cdot\int_s^t\int_{\tau_n}^\tau \sum_{i,j=1}^r \mu_{i,j}(y_n(\zeta))\nu_j(y_n(\tau_n))(\dot\beta_n(\tau_n))^2 d\zeta d\tau\right]
\end{aligned}$$

were $\mu_{i,j}(y)=\frac{\partial^2\varphi}{\partial v_i\partial v_j}(\langle v_1,y\rangle,\dots,\langle v_r,y\rangle)\langle v_i,G(y)\rangle$ and $\nu_j(y)=\langle v_j,G(y)\rangle$ are bounded and Lipschitz. But

$$\begin{aligned}
J_n^{1.2} &= \sum_{i,j=1}^r \mathbb{E}\left[h\cdot\int_s^t\int_{\tau_n}^\tau \sum_{i,j=1}^r \mu_{i,j}(y_n(\tau_n))\nu_j(y_n(\tau_n))(\dot\beta_n(\tau_n))^2 d\zeta d\tau\right]\\
&\quad +\mathbb{E}\left[h\cdot\int_s^t\int_{\tau_n}^\tau \left(\mu(y_n(\zeta))-\mu(y_n(\tau_n))\right)\nu(y_n(\tau_n))(\dot\beta_n(\tau_n))^2 d\zeta d\tau\right]\\
&=: J_n^{1.2.1}+J_n^{1.2.2}
\end{aligned}$$

We start from the second term:

$$\begin{aligned}
|J_n^{1.2.2}| &\le \int_s^t\int_{\tau_n}^\tau \mathbb{E}\left(|y_n(\zeta)-y_n(\zeta_n)|(\dot\beta_n(\tau_n))^2\right)d\zeta d\tau\\
&\le \int_s^t\int_{\tau_n}^\tau \mathbb{E}\Big(\left|\left(e^{(\zeta-\zeta_n)A}-I\right)y_n(\zeta_n)\right|(\dot\beta_n(\tau_n))^2\Big)d\zeta d\tau\\
&\quad+\int_s^t\int_{\tau_n}^\tau \mathbb{E}\left(\left|\int_{\zeta_n}^\zeta e^{(\zeta-\sigma)A}\Big(F(y_n(\sigma))+G(y_n(\sigma))\dot\beta_n(\sigma)\Big)d\sigma\right|\times\right.\\
&\qquad\qquad\left.\times(\dot\beta_n(\tau_n))^2\right)d\zeta d\tau\\
&=: J_n^{1.2.2.1}+J_n^{1.2.2.2}
\end{aligned}$$

$$
\begin{aligned}
J_n^{1.2.2.1} &\le c\int_{s\wedge 2^{-n}}^{2^{-n}}\left(\int_0^\tau 1d\zeta\right)2^n d\tau \\
&\quad + c\mathbb{E}\left[\int_{s\vee 2^{-n}}^t\int_{\tau_n}^\tau (\zeta-\tau_n)^\alpha \tau_n^{-\alpha}(\tau_n^\alpha|y_n(\tau_n)|_\alpha)(\dot\beta_n(\tau_n))^2 d\zeta d\tau\right] \\
&\le c2^n 2^{-2n} + c\mathbb{E}\left[\int_{2^{-n}}^t (\tau-\tau_n)^{1+\alpha}\tau_n^{-\alpha}(\tau_n^\alpha|y_n(\tau_n)|_\alpha)(\dot\beta_n(\tau_n))^2 d\tau\right] \\
&\le c2^{-n} + c\left(\int_{2^{-n}}^t \tau_n^{-\alpha q}d\tau\right)^{1/q} \\
&\quad \times\left(\mathbb{E}\left[\int_{2^{-n}}^t(\tau-\tau_n)^{p(1+\alpha)}(\tau_n^\alpha|y_n(\tau_n)|_\alpha)^p(\dot\beta_n(\tau_n))^{2p}d\tau\right]\right)^{1/p} \\
&\le c2^{-n} + c\left(t^{1-\alpha q}\right)^{1/q}\left(2^{np}\int_{2^{-n}}^t(\tau-\tau_n)^{p(1+\alpha)}d\tau\right)^{1/p} \\
&\le c2^{-n} + c2^{np+1}2^{-n[(1+\alpha)p+1]} \le c2^{-n}+c2^{-np\alpha} = o(1)
\end{aligned}
$$

We have used that $\sup_{n,\tau}\mathbb{E}\left(\tau^{\alpha p}|y_n(\tau)|_\alpha^p\right) \le c$. Moreover

$$
\begin{aligned}
J_n^{1.2.2.2} &\le c\mathbb{E}\int_s^t\int_{\tau_n}^\tau\left[\left(\int_{\zeta_n}^\zeta(1+|\dot\beta_n(\sigma)|)d\sigma\right)(\dot\beta_n(\tau_n))^2\right]d\zeta d\tau \\
&\le c\mathbb{E}\int_s^t\int_{\tau_n}^\tau(\zeta-\zeta_n)\left(\dot\beta_n(\tau_n)+|\dot\beta_n(\tau_n)|^3\right)d\zeta d\tau \\
&\le c2^{\frac32 n}\int_s^t(\tau-\tau_n)^2 d\tau \le c2^{\frac32 n+n}2^{-3n} = o(1).
\end{aligned}
$$

Therefore $J_n^{1.2.2} = o(1)$. Coming back to $J_n^{1.2.1}$ we have:

$$
\begin{aligned}
J_n^{1.2.1} = \mathbb{E}\Big[h\cdot\int_s^t\int_{\tau_n}^\tau\sum_{ij=1}^r\frac{\partial^2\varphi}{\partial v_i\partial v_j}(\langle v_1,y_n(\zeta_n)\rangle,\dots,\langle v_r,y_n(\zeta_n)\rangle)\times \\
\times\langle v_i,G(y_n(\tau_n))\rangle\langle v_j,G(y_n(\tau_n))\rangle 2^{2n}(\beta(\tau_n^+)-\beta(\tau_n))^2 d\zeta d\tau\Big].
\end{aligned}
$$

Define

$$\ell(z) = \langle D^2 f(z)G(z), G(z)\rangle,\ \ z\in H.$$

Then

$$
\begin{aligned}
J_n^{1.2.1} &= \mathbb{E}\left[h\cdot 2^n\int_s^t(\tau-\tau_n)\ell(y_n(\sigma_n))d\sigma\right] = \frac12\mathbb{E}\left[h2^{-n}(\sum_{k=[2^n s]+1}^{[2^n t]-1}\ell(y_n(\frac{k}{2^n}))\right] \\
&\quad +\frac12\mathbb{E}\left[h2^n(s-s_n)^2\ell(y_n(s_n))\right]+\frac12\mathbb{E}\left[h2^n(t-t_n)^2\ell(y_n(t_n))\right] \\
&= \frac12\mathbb{E}\left[h\int_{s_n^+}^{t_n}\ell(y_n(\sigma_n))d\sigma + h2^n[(s-s_n)^2\ell(y_n(s_n))+(t-t_n)^2\ell(y_n(t_n))]\right] \\
&= \frac12\mathbb{E}h\int_s^t\ell(y_n(\sigma_n))d\sigma \\
&\quad +\frac12\mathbb{E}\left[h[(2^n(t-t_n)^2-(t-t_n))\ell(y_n(t_n))+(2^n(s-s_n)^2-(s-s_n))\ell(y_n(s_n))]\right] \\
&= \frac12\mathbb{E}h\int_s^t\ell(y_n(\sigma_n))d\sigma+\epsilon_n,
\end{aligned}
$$

where

$$\epsilon_n = \frac12\mathbb{E}\left[h[(2^n(t-t_n)^2-(t-t_n))\ell(y_n(t_n))+(2^n(s-s_n)^2-(s-s_n))\ell(y_n(s_n))]\right].$$

It is clear that $\epsilon_n \to 0$ as $n \to \infty$. Consequently

$$\begin{aligned}\mathbb{E}\left[h \cdot 2^n \int_s^t (\tau - \tau_n)\ell(y_n(\sigma_n))d\sigma\right] &= \frac{1}{2}\mathbb{E}\left[h \cdot \int_s^t \ell(y_n(\tau_n))d\tau\right] + \epsilon_n \\ = \frac{1}{2}\mathbb{E}^{\mathbb{Q}_n}\left[h \cdot \int_s^t \ell(\nu(\tau_n))d\tau\right] &+ \epsilon_n \to \frac{1}{2}\mathbb{E}^{\mathbb{Q}}\left[h \cdot \int_s^t \ell(\nu(\tau))d\tau\right] \\ &= \frac{1}{2}\mathbb{E}^{\mathbb{Q}}\left[h \cdot \int_s^t \langle D^2 f(\nu(\tau))G(\nu(\tau)), G(\nu(\tau))\rangle d\tau\right],\end{aligned}$$

by the lemma below.

Lemma 5.6. *Assume that a sequence (μ_n) of probability measures on a Polish space E converges weakly to a probability measure μ. If a uniformly bounded sequence (φ_n) of continuous functions on E converges to a function φ uniformly on any compact set then*

$$\lim_n \int_E \varphi_n(x)\mu_n(dx) = \int_E \varphi(x)\mu(dx).$$

We pass now to the term J_n^2

$$\begin{aligned}J_n^2 &= \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))\left[G(y_n(\tau)) - G(y_n(\tau_n))\right]\dot{\beta}_n(\tau_n)d\tau\right] \\ &= \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))\Gamma_n(\tau)\left(y_n(\tau) - y_n(\tau_n)\right)\dot{\beta}_n(\tau_n)d\tau\right] \\ &= \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))\Gamma_n(\tau)\left(e^{(\tau-\tau_n)A}y_n(\tau_n) - y_n(\tau_n)\right)\dot{\beta}_n(\tau_n)d\tau\right] \\ &\quad + \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))\Gamma_n(\tau)\left(\int_{\tau_n}^{\tau} e^{(\tau-\sigma)A}F(y_n(\sigma))d\sigma\right)\dot{\beta}_n(\tau_n)d\tau\right] \\ &\quad + \mathbb{E}\left[h \cdot \int_s^t Df(y_n(\tau))\Gamma_n(\tau)\left(\int_{\tau_n}^{\tau} e^{(\tau-\sigma)A}G(y_n(\sigma))d\sigma\right)(\dot{\beta}_n(\tau_n))^2 d\tau\right] \\ &=: J_n^{2.0} + J_n^{2.1} + J_n^{2.2},\end{aligned}$$

where

$$\Gamma_n(\tau) = \int_0^1 \nabla G(y_n(\tau_n) + \theta(y_n(\tau) - y_n(\tau_n)))\nu d\theta.$$

But

$$\begin{aligned}|J_n^{2.0}| &\le c2^{-n}\mathbb{E}\left(|\dot{\beta}_n(0)|\right) + \int_{2^{-n}}^t \mathbb{E}(\tau - \tau_n)^{\alpha}\tau_n^{-\alpha}|\tau_n^{\alpha}y_n(\tau_n)|_{\alpha}|\dot{\beta}_n(\tau_n)|d\tau \\ &\le c2^{-n/2} + \left(\int_{2^{-n}}^t (\tau - \tau_n)^{\alpha p}\mathbb{E}\left(|\dot{\beta}_n(\tau_n)|^p\right)\mathbb{E}\left(|\tau_n^{\alpha p}y_n(\tau_n)|_{\alpha}^p\right)d\tau\right)^{1/p} \times \\ &\qquad \times \left(\int_{2^{-n}}^t \tau_n^{-\alpha q}d\tau\right)^{1/q} \\ &\le c2^{-n/2} + c\left(t^{1-\alpha q}\right)^{1/q}\left[2^n t 2^{np/2}2^{-n(1+\alpha p)}\right]^{1/p} \\ &\le c2^{-n/2} + ct^{\frac{1}{q}-\alpha}t^{\frac{1}{p}}2^{n(\frac{1}{2}-\alpha)} = o(1) \quad \text{if } \alpha > \frac{1}{2}.\end{aligned}$$

Moreover

$$|J_n^{2.1}| \le \int_s^t c(\tau - \tau_n)\mathbb{E}|\dot{\beta}_n(\tau_n)| \le c2^{n/2}2^n 2^{-2n} = o(1)$$

Finally

$$
\begin{aligned}
J_n^{2.2} &= \mathbb{E}\left[h \cdot \int_s^t (\tau-\tau_n) Df(y_n(\tau_n))\nabla G(y_n(\tau_n))G(y_n(\tau_n))\left(\dot{\beta}_n(\tau_n)\right)^2 d\tau\right] \\
&\quad + \mathbb{E}\Bigg[h \cdot \int_s^t Df(y_n(\tau))\Gamma_n(\tau)\times \\
&\qquad \times \left[\int_{\tau_n}^{\tau}\left\{e^{(\tau-\sigma)A}G(y_n(\sigma)) - G(y_n(\sigma))\right\} d\sigma\right]\left(\dot{\beta}_n(\tau_n)\right)^2 d\tau\Bigg] \\
&\quad + \mathbb{E}\Bigg[h \cdot \int_s^t Df(y_n(\tau))\Gamma_n(\tau)\times \\
&\qquad \times \left[\int_{\tau_n}^{\tau}\left\{G(y_n(\sigma)) - G(y_n(\sigma_n))\right\} d\sigma\right]\left(\dot{\beta}_n(\tau_n)\right)^2 d\tau\Bigg] \\
&\quad + \mathbb{E}\Bigg[h \cdot \int_s^t [Df(y_n(\tau))\Gamma_n(\tau) - Df(y_n(\tau_n))\nabla G(\tau_n)] \\
&\qquad \times(\tau-\tau_n)G(y_n(\tau_n))d\sigma\left(\dot{\beta}_n(\tau_n)\right)^2 d\tau\Bigg] \\
&=: J_n^{2.2.1} + J_n^{2.2.2} + J_n^{2.2.3} + J_n^{2.2.4}
\end{aligned}
$$

However, by Lemma 5.6

$$
\begin{aligned}
J_n^{2.2.1} &= \frac{1}{2}\frac{1}{2^{2n}}2^n \mathbb{E}\left[h \cdot \sum_{k=[2^n s]}^{[2^n t]} Df(y_n(k))\nabla G(y_n(k))G(y_n(k))\right] \\
&= \frac{1}{2}\mathbb{E}^{\mathbb{Q}}\left[h \cdot \int_s^t Df(\nu(\tau_n))\nabla G(\nu(\tau_n))G(\nu(\tau_n))d\tau\right] \\
&\to \frac{1}{2}\mathbb{E}^{\mathbb{Q}}\left[h \cdot \int_s^t Df(\nu(\tau))\nabla G(\nu(\tau))G(\nu(\tau))d\tau\right] \qquad \text{as } n \to \infty.
\end{aligned}
$$

To conclude the proof we have to show that $|J_n^{2.2.2}| + |J_n^{2.2.3}| + |J_n^{2.2.4}| = o(1)$. We start from

$$
|J_n^{2.2.2}| \le c\int_s^t \mathbb{E}\left(\left|\int_{\tau_n}^{\tau}\left(e^{(\tau-\sigma)A}G(y_n(\sigma)) - G(y_n(\sigma))\right)d\sigma\right|(\dot{\beta}_n(\tau_n))^2\right)d\tau.
$$

Since

$$
\begin{aligned}
&\mathbb{E}\left(\left|\int_{\tau_n}^{\tau}\left(e^{(\tau-\sigma)A}G(y_n(\sigma)) - G(y_n(\sigma))\right)d\sigma\right|(\dot{\beta}_n(\tau_n))^2\right) \\
&\le c(\tau-\tau_n)\mathbb{E}\left((\dot{\beta}_n(\tau_n))^2\right) \le c2^{-n}2^n = c
\end{aligned}
$$

it is enough to prove that $\forall \tau > 0$ fixed

$$
\mathbb{E}\left[\,\left|\int_{\tau_n}^{\tau}\left(e^{(\tau-\sigma)A} - I\right)G(y_n(\sigma))d\sigma\right|(\dot{\beta}_n(\tau_n))^2\right] \to 0.
$$

But

$$
\begin{aligned}
&\mathbb{E}\left[\,\left|\int_{\tau_n}^{\tau}\left(e^{(\tau-\sigma)A} - I\right)G(y_n(\sigma))d\sigma\right|(\dot{\beta}_n(\tau_n))^2\right] \\
&\le \mathbb{E}\left[\int_{\tau_n}^{\tau}\left|\left(e^{(\tau-\sigma)A} - I\right)G(y_n(\sigma))\right|(\dot{\beta}_n(\tau_n))^2 d\sigma\right] \\
&\le c\mathbb{E}\left[\int_{\tau_n}^{\tau}(\tau-\sigma)^{\epsilon}\left(|y_n(\sigma)|_\alpha + 1\right)(\dot{\beta}_n(\tau_n))^2 d\sigma\right] \\
&\le c2^n\int_{\tau_n}^{\tau}(\tau-\sigma)^{\epsilon}d\sigma + \\
&\quad c\left(\mathbb{E}\left(\dot{\beta}_n(\tau_n)\right)^{2q}\right)^{1/q}(\tau-\tau_n)^{1/q}\left(\int_{\tau_n}^{\tau}(\tau-\sigma)^{\epsilon p}\mathbb{E}|y_n(\sigma)|_\alpha^p\right)^{1/p} \\
&\le c2^n(2^{-n})^{1+\epsilon} + c2^n 2^{-n/q}2^{-n(1+\epsilon p)\frac{1}{p}} \le c2^{-n\epsilon} \to 0,
\end{aligned}
$$

because $\sup_{0\le\sigma\le T}\sigma^{\alpha p}\mathbb{E}\left(|y_n(\sigma)|_\alpha^p\right)<+\infty$.
Moreover G is Lipschitz and therefore

$$\begin{aligned}|J_n^{2.2.3}| &\le \mathbb{E}\Big[|h|\cdot\int_s^t |Df(y_n(\tau))||\Gamma_n(\tau)|\times \\ &\qquad\times\int_{\tau_n}^{\tau}|G(y_n(\sigma))-G(y_n(\tau_n))|\,d\sigma\left(\dot\beta_n(\tau_n)\right)^2 d\tau\Big] \\ &\le c\mathbb{E}\Big[|h|\cdot\int_s^t |Df(y_n(\tau))||\Gamma_n(\tau)|\times \\ &\qquad\times\int_{\tau_n}^{\tau}\Big|\left(e^{(\sigma-\tau_n)A}-I\right)y_n(\tau_n)\Big|\,d\sigma\left(\dot\beta_n(\tau_n)\right)^2 d\tau\Big] \\ &\quad +c\mathbb{E}\Big[|h|\cdot\int_s^t |Df(y_n(\tau))||\Gamma_n(\tau)|\left(\dot\beta_n(\tau_n)\right)^2\times \\ &\quad\times\int_{\tau_n}^{\tau}\Big|e^{(\zeta-\tau_n)A}\Big|\left(|F(y_n(\zeta))|+|G(y_n(\zeta))|\left|\dot\beta_n(\tau_n)\right|\right)d\zeta\,d\sigma\,d\tau\Big] \\ &=: J_n^{2.2.3.1}+J_n^{2.2.3.2}\end{aligned}$$

We notice that

$$|J_n^{2.2.3.1}|\le\int_s^t\int_{\tau_n}^{\tau}\mathbb{E}\Big|\left(e^{(\sigma-\tau_n)A}-I\right)y_n(\tau_n)\Big|\left(\dot\beta_n(\tau_n)\right)^2 d\sigma d\tau = o(1)$$

proceeding as in the estimate of the term $J_n^{1.2.2.1}$. Moreover

$$\begin{aligned}|J_n^{2.2.3.2}| &\le c\int_s^t\int_{\tau_n}^{\tau}\mathbb{E}\left(|\dot\beta_n(\tau_n)|^2+|\dot\beta_n(\tau_n)|^3\right)d\tau d\sigma \\ &\le c\int_s^t\int_{\tau_n}^{\tau}(\sigma-\tau_n)2^{\frac{3}{2}n}d\sigma \\ &\le c\int_s^t(\tau-\tau_n)^2 d\tau\cdot 2^{-\frac{3}{2}n} \\ &\le c2^n2^{-3n}2^{\frac{3}{2}n}=o(1)\end{aligned}$$

We are left with

$$\begin{aligned}J_n^{2.2.4} &= \mathbb{E}\Big[h\cdot\int_s^t\left(Df(y_n(\tau))\Gamma_n(\tau)-Df(y_n(\tau_n))\nabla G(y_n(\tau_n))\right) \\ &\qquad\qquad\times G(y_n(\tau_n))(\tau-\tau_n)\left(\dot\beta_n(\tau_n)\right)^2 d\tau\Big] \\ &= \mathbb{E}\Big[h\cdot\int_s^t Df(y_n(\tau))\left(\Gamma_n(\tau)-\nabla G(y_n(\tau_n))\right) \\ &\qquad\qquad\times G(y_n(\tau_n))(\tau-\tau_n)\left(\dot\beta_n(\tau_n)\right)^2 d\tau\Big] \\ &\quad +\mathbb{E}\Big[h\cdot\int_s^t\left(Df(y_n(\tau))-Df(y_n(\tau_n))\right)\nabla G(y_n(\tau_n)) \\ &\qquad\qquad\times G(y_n(\tau_n))(\tau-\tau_n)\left(\dot\beta_n(\tau_n)\right)^2 d\tau\Big] \\ &=: J_n^{2.2.4.1}+J_n^{2.2.4.2}\end{aligned}$$

Since

$$\begin{aligned}&\mathbb{E}\Big|Df(y_n(\tau))\left(\Gamma_n(\tau)-\nabla G(y_n(\tau_n))\right)G(y_n(\tau_n))(\tau-\tau_n)\left(\dot\beta_n(\tau_n)\right)^2\Big| \\ &\le c(\tau-\tau_n)\mathbb{E}\left(\left(\dot\beta_n(\tau_n)\right)^2\right)\le c.\end{aligned}$$

to prove that $|J_n^{2.2.4.1}| = o(1)$ it is enough, by dominated convergence theorem, to prove that for all $\tau > 0$

$$\begin{aligned}&\mathbb{E}\left|Df(y_n(\tau))\,(\Gamma_n(\tau) - \nabla G(y_n(\tau_n)))\,G(y_n(\tau_n))(\tau - \tau_n)\left(\dot{\beta}_n(\tau_n)\right)^2\right| \\ &\le \mathbb{E}\left|(\Gamma_n(\tau) - \nabla G(y_n(\tau_n)))\,G(y_n(\tau_n))(\tau-\tau_n)\right|^2 \\ &= \mathbb{E}\left|\int_0^1 (\nabla G(y_n(\tau_n) + \theta(y_n(\tau) - y_n(\tau_n))) - \nabla G(y_n(\tau)_n))\,G(y_n(\tau_n))d\theta\right|^2 \\ &\le c\int_0^1 \mathbb{E}\left|(\nabla G(y_n(\tau_n) + \theta(y_n(\tau) - y_n(\tau_n))) - \nabla G(y_n(\tau_n)))\,G(y_n(\tau_n))\right|^2 d\theta\end{aligned}$$

is infinitesimal as $n \to \infty$ so it is enough to prove that $\forall \theta \in [0,1]$

$$\mathbb{E}\Big[\,|(\nabla G(y_n(\tau_n) + \theta(y_n(\tau) - y_n(\tau_n))) - \nabla G(y_n(\tau_n)))\,G(y_n(\tau_n))|\,\Big]^2 \to 0 \tag{5.12}$$

We need the following

Lemma 5.7. $\sup_{\tau\in[t_0,T]} \mathbb{E}\,|y_n(\tau) - y_n(\tau_n)|^2 \to 0$, *for all* $t_0 \in]0,T]$.

Proof. For n sufficiently large, α as in Theorem 3.1

$$\sup_{\tau,n} \mathbb{E}\left(|y_n(\tau_n)|_\alpha^2\right) < +\infty$$

$$\begin{aligned}y_n(\tau) - y_n(\tau_n) &= \left(e^{(\tau-\tau_n)A} - I\right) y_n(\tau_n) + \int_{\tau_n}^{\tau} e^{(\sigma-\tau_n)A}F(y_n(\sigma))d\sigma \\ &\quad + \left(\int_{\tau_n}^{\tau} e^{(\sigma-\tau_n)A}G(y_n(\sigma))d\sigma\right)\dot{\beta}_n(\tau_n)\end{aligned}$$

Therefore

$$\begin{aligned}\mathbb{E}\left|y^{(n)}(\tau) - y^{(n)}(\tau_n)\right|^2 &\le (\tau-\tau_n)^{2\alpha}\mathbb{E}|y^{(n)}(\tau_n)|_\alpha^2 + (\tau-\tau_n)^2(1+2^n) \\ &\le c(2^{-2n\alpha} + 2^{-2n} + 2^{-2n}2^n) \to 0.\end{aligned}$$

□

Coming back to (5.12), by assumption (A.3),

$$\begin{aligned}&\mathbb{E}\left|[\nabla G(y_n(\tau_n) + \theta(y_n(\tau) - y_n(\tau_n))) - \nabla G(y_n(\tau_n))]\,G(y_n(\tau_n)\right|^2 \\ &\le c\theta^2\mathbb{E}|y_n(\tau) - y_n(\tau_n)|^2 \to 0.\end{aligned}$$

Exactly in the same way it can be proved that

$$\begin{aligned}J_n^{2.2.4.2} = \mathbb{E}\Big[h \cdot \int_s^t &(Df(y_n(\tau)) - Df(y_n(\tau_n)))\,\nabla G(y_n(\tau_n))\times \\ &\times G(y_n(\tau_n))(\tau-\tau_n)\left(\dot{\beta}_n(\tau_n)\right)^2 d\tau\Big] \to 0.\end{aligned}$$

and this completes the proof. □

6. Applications

6.1. **Specific examples.** We consider here specific examples to show how the general theory applies.

Example 6.1. Stochastic nonlinear heat equation

$$\frac{\partial u}{\partial t}(t,\xi) = \Delta_\xi u(t,\xi) + f(u(t,\xi)) + g(u(t,\xi))\frac{\partial \beta}{\partial t}, \tag{6.1}$$

$$u(0,\xi) = x(\xi),\ \xi \in \mathcal{O}, \quad u(t,\xi) = 0,\ t > 0,\ \xi \in \partial\mathcal{O},\ \mathcal{O} \subset \mathbb{R}^1. \tag{6.2}$$

Here we take $H = L^2(\mathcal{O})$, $A = \Delta$, $D(A) = H^2(\mathcal{O}) \cap H_0^1(\mathcal{O})$. It is known (see [33], [35]) that

$$\begin{aligned} V_\alpha &= H^{2\alpha}(\mathcal{O}), \text{ if } \alpha \in]0, \tfrac{1}{4}[\\ V_\alpha &= H_0^{2\alpha}(\mathcal{O}), \text{ if } \alpha \in]\tfrac{1}{4}, 1[. \end{aligned}$$

If $\beta \in (0, 1)$ then

$$\begin{aligned} \|x\|^2_{H^\beta(\mathcal{O})} &= \|x\|^2_{L^2(\mathcal{O})} \\ &+ \int_{\mathcal{O}} \int_{\mathcal{O}} \Big(\frac{|x(\xi) - x(\eta)|}{|\xi - \eta|^\beta} \Big)^2 \frac{1}{|\xi - \eta|^d} \, d\xi \, d\eta, \end{aligned}$$

and if $\beta \in (1, 2)$

$$\begin{aligned} \|x\|^2_{H^\beta(\mathcal{O})} &= \|x\|^2_{L^2(\mathcal{O})} + \|D_\xi x\|^2_{L^2(\mathcal{O}, \mathbb{R}^d)} \\ &+ \int_{\mathcal{O}} \int_{\mathcal{O}} \Big(\frac{|D_\xi x(\xi) - D_\xi x(\eta)|}{|\xi - \eta|^{\beta - 1}} \Big)^2 \frac{1}{|\xi - \eta|^d} \, d\xi \, d\eta. \end{aligned}$$

We set

$$F(x)(\xi) = f(x(\xi)), \quad G(x)(\xi) = g(x(\xi)), \quad \xi \in \mathcal{O}, \; x \in H.$$

Then

$$\nabla F(x) u(\xi) = f'(x(\xi)) u(\xi), \quad \nabla G(x) u(\xi) = g'(x(\xi)) u(\xi), \quad \xi \in \mathcal{O}, \; x, u \in H$$

and

$$|\nabla F(x)| = \sup_{z \in \mathbb{R}^1} |f'(z)|, \quad |\nabla G(x)| = \sup_{z \in \mathbb{R}^1} |g'(z)|.$$

Assumption (A.1) is therefore satisfied if f, f', g, g' are continuous and bounded functions. Assumption (A.2) is equivalent to existence of $c > 0$ such that

$$|G(x)|_\beta \le c(1 + |x|_\alpha), \quad x \in V_\alpha,$$

provided that $0 < \frac{\alpha}{2} < \beta < \frac{1}{2} < \alpha < 1$. For this it is enough that g' is continuous and bounded as then G transforms V_α into V_α.
Finally, the assumption (A.3) is satisfied if $|g|$ and $|g'|$ are functions bounded by a $\gamma > 0$, as then it is enough to define

$$\Gamma = \big\{ x : \operatorname{ess\,sup} |x(\xi)| \le (1 + \gamma)\gamma \big\}.$$

The heat semigroup $S(t)$, $t \ge 0$, generated by the operator A is analytic, and operators $S(t)$, $t > 0$, are compact and therefore all the assumptions of Theorems 3.1–5.1 hold so the Wong-Zakai approximations for the non-linear stochastic heat equation are bounded in the proper norms and converge weakly to the solution of the equation

$$\begin{aligned} \frac{\partial v}{\partial t}(t, \xi) &= \Delta_\xi (v(t, \xi) + \Big(f + \frac{1}{2} g g' \Big)(v(t, \xi)) + g(v(t, \xi)) \frac{\partial \beta}{\partial t}, \\ v(0, \xi) &= x(\xi), \qquad t > 0. \end{aligned}$$

Example 6.2. Strongly damped stochastic wave equation.
Let $\mathcal{O}$ be an open, bounded subset of $\mathbb{R}^m$ with a regular boundary. The equation we have in mind can be formally written as follows

$$\begin{aligned} &\frac{\partial^2 u}{\partial t^2}(t, \xi) = \Delta_\xi u(t, \xi) + \rho \Delta_\xi \frac{\partial u}{\partial t}(t, \xi) + f(u(t, \xi)) + g(u(t, \xi)) \frac{\partial \beta}{\partial t} \\ &u(t, \xi) = 0, \quad t > 0, \; \xi \in \partial \mathcal{O}, \\ &u(0, \xi) = x_0(\xi), \quad \frac{\partial u}{\partial t}(0, \xi) = x_1(\xi), \quad \xi \in \mathcal{O}, \end{aligned} \tag{6.3}$$

where ρ is a positive constant. Let $\Lambda = -\Delta$ with the domain $D(\Lambda) = H^2(\mathcal{O}) \cap H_0^1(\mathcal{O}) \subset L^2(\mathcal{O})$. Then Λ is a self-adjoint, positive definite operator in $L^2(\mathcal{O})$. The equation (6.3) can be written in an abstract form (1.13) in the Hilbert space $H = D(\Lambda^{1/2}) \oplus L^2(\mathcal{O})$ with the operator A:

$$D(A) = \left\{ \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \in H \ ; \ x_0 + \rho x_1 \in D(\Lambda) \right\}$$

$$A \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -\Lambda & -\rho\Lambda \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \quad \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \in D(A).$$

The operator A generates a strongly continuous analytic semigroup of contractions $S(t)$. Moreover

$$V_\alpha = D(\Lambda^\alpha) \oplus D(\Lambda^\alpha).$$

Since the embeddings $D(\Lambda^\alpha) \subset D(\Lambda^\beta) \subset L^2(\mathcal{O})$, $0 < \beta < \alpha < 1$, are compact and $S(t)H \subset V_\alpha$ for all $\alpha \in (0,1)$, the operators $S(t)$, $t > 0$ are compact as well.
Assume that functions f, g are continuously differentiable, bounded together with their first derivatives. We define

$$F \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} (\xi) = \begin{pmatrix} 0 \\ f(x_0(\xi)) \end{pmatrix}, \quad G \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} (\xi) = \begin{pmatrix} 0 \\ g(x_0(\xi)) \end{pmatrix}.$$

Then the transformations F and G satisfy all the assumptions of the theorems with $\epsilon = \frac{1}{2}$ and $\alpha > \frac{1}{2}$. As the set Γ one can take the product of $D(\Lambda^{1/2})$ with a ball in $L^\infty(\mathcal{O})$ of suitable radius.

Example 6.3. Strongly damped plate equation.
In a similar way one can treat the equation

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(t,\xi) &= -\Delta_\xi^2 u(t,\xi) + \rho \Delta_\xi \frac{\partial u}{\partial t}(t,\xi) + f(u(t,\xi)) + g(u(t,\xi)) \frac{\partial \beta}{\partial t} \\ u(t,\xi) &= 0, \quad \Delta u(t,\xi) = 0, \quad t > 0, \ \xi \in \partial\mathcal{O}, \\ u(0,\xi) &= x_0(\xi), \quad \frac{\partial u}{\partial t}(t,\xi) = x_1(\xi), \quad \xi \in \mathcal{O}, \end{aligned}$$

where ρ is a positive constant. In this case $H = D(\Lambda) \oplus L^2(\mathcal{O})$ and $D(A) = D(\Lambda^2) \oplus D(\Lambda)$,

$$A \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -\Lambda^2 & -\rho\Lambda \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \quad \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \in D(A).$$

Again the operator A generates a strongly continuous analytic semigroup of compact contractions on H. Moreover

$$V_\alpha = D(\Lambda^{1+\alpha}) \oplus D(\Lambda^\alpha), \quad \alpha \in (0,1).$$

Moreover F and G are as in the previous example and still satisfy our assumptions.

6.2. **Stochastic invariance.** A set $K \subset H$ is said to be invariant for the stochastic equation (1.13) if for solutions $X(t,x)$, $t \geq 0$, $x \in H$ of (1.13) one has

$$\mathbb{P}(X(t,x) \in K) = 1, \quad \text{for all } t \geq 0, \, x \in K.$$

The literature on stochastic invariance is rather extensive, both for finite dimensional systems, see [14], [21], and [4] and references therein, and for infinite dimensional ones, see e.g. [34], [17] and [41]. for a discussion of recent results. In this section we deduce some sufficient conditions for stochastic invariance using results from deterministic theory. The following characterization was proved by Pavel [30, 31] in the case of compact operators $S(t)$, $t \geq 0$, and for general C_0-semigroups by Jachimiak [15].

Theorem 6.1. *Assume that the operator A generates a C_0-semigroup on a Banach space H and F is a Lipschitz transformation from H into H. A closed set $K \subset H$ is invariant for*

$$\frac{dy}{dt} = Ay(t) + F(y(t)), \quad y(0) = x \in H$$

if and only if for arbitrary $x \in K$

$$\liminf_{t \to 0} \frac{1}{t} \operatorname{dist}\left[S(t)x + tF(x), K\right] = 0. \tag{6.4}$$

If the set K is contained in the domain $D(A)$ of the generator A then the condition (6.4) can be replaced by the classical Nagumo's condition

$$\liminf_{t \to 0} \frac{1}{t} \operatorname{dist}\left[x + t(Ax + F(x)), K\right] = 0, \quad \text{for all } x \in K. \tag{6.5}$$

We have also the following invariance result for the stochastic equation (1.13).

Theorem 6.2. *Assume that (A.1), (A.2) and (A.3) hold for $G^1, \dots, G^d$ and for F replaced by $\tilde{F} = F - \frac{1}{2}\nabla G^j G^j$. Moreover assume that the conditions of Theorem 6.1 are satisfied and for arbitrary $x \in K$ and $u_1, \dots, u_d \in \mathbb{R}$*

$$\liminf_{t \to 0} \frac{1}{t} \operatorname{dist}\left[S(t)x + t\left[F(x) + \sum_{j=1}^{d}\left[u_j G^j(x) - \frac{1}{2}\nabla G^j(x) G^j(x)\right]\right], K\right] = 0. \tag{6.6}$$

Then the set K is invariant for (1.13).

Proof. If (6.6) holds then the set K is invariant for the deterministic equations

$$\begin{aligned}
&\frac{dz}{dt}(t) = Az(t) + F(z(t)) - \frac{1}{2}\sum_{j=1}^{d} \nabla G^j(z(t)) G^j(z(t)) + \sum_{j=1}^{d} G^j(z(t)) v_j(t), \\
&z(0) = x,
\end{aligned}$$

for any piecewise constant function $v_1, \dots, v_d$. Consequently the set K is also invariant for the solutions $\tilde{y}_n(t)$, $t \geq 0$ of the equations

$$\begin{aligned}
\frac{d\tilde{y}_n}{dt}(t) &= A\tilde{y}_n(t) + F(\tilde{y}_n(t)) - \frac{1}{2}\sum_{j=1}^{d} \nabla G^j(\tilde{y}_n(t)) G^j(\tilde{y}_n(t)) \\
&\quad + \sum_{j=1}^{d} G^j(\tilde{y}_n(t)) \dot{\beta}_n^j(t), \\
\tilde{y}_n(0) &= x.
\end{aligned}$$

But by Theorem 6.1 the laws of $\tilde{y}_n(\cdot)$ on $C(0,T;H)$ converge to the law of the solution X of (1.13). However the supports of the laws of $\tilde{y}_n$ are contained in $C(0,T;K)$ so the same is true for the law of X. This proves the result. □

Specific results can be obtained for sets

$$K = \left\{x \in L^2(\mathcal{O});\ x(\xi) \geq 0,\ \xi \in \mathcal{O}\right\}$$

or

$$K = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in L^2(\mathcal{O}) \times L^2(\mathcal{O});\ x(\xi) \geq y(\xi),\ \xi \in \mathcal{O}\right\}$$

leading to explicit condition for positivity of the solutions or to comparison like results. The analysis of the condition (6.6) is simplified if the set K is invariant for the semigroup $S(t)$, $t \geq 0$, (which is often the case).

References

[1] P. Acquistapace and B. Terreni, *An approach to linear equations in Hilbert spaces by approximation of white noise with colored noise.* Stoch. Anal. Appl., 2, 1984, 131–186.

[2] S. Aida, *Support theorem for diffusion processes on Hilbert space*, Publ. Res. Inst. Math. Sci., **26**, 1990, 947–965.

[3] J.P. Aubin, *Viability Theory*, Birkhäuser, Boston-Basel 1991.

[4] J.P. Aubin and G. Da Prato, *The viability theorem for stochastic differential inclusions.* Stochastic Analysis and Applications, **16**, 1998, 1–15.

[5] V. Bally, A. Millet and M. Sanz-Sole, *Approximation and support theorem in Holder norm for parabolic stochastic partial differential equations.* Annals of Probability, **23**, 1995, 178–222.

[6] Z. Brzezniak, M. Capinski and F. Flandoli, *A convergence result for stochastic partial differential equations.* Stochastics, **24**, 1988, 423–445.

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

[8] M. Fuhrman and G. Tessitore, *Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control*, Preprint, Politecnico di Milano, 2000.

[9] I. Gyöngy, *On the approximations of stochastic partial differential equations I*, Stochastics, 23(1988), n.3., 331- 352.

[10] I. Gyöngy, *The stability of stochastic partial differential equations and applications I*, Stochastics Stochastics Rep., **27**, 1989, 129–150.

[11] I. Gyöngy, *The stability of stochastic partial differential equations and applications II*, Stochastics Stochastics Rep., **27**, 1989, 189–233.

[12] I. Gyöngy, D. Nualart and M. Sanz- Sole
Approximations and support theorems in modulus spaces, PTRF, 101(1995), No.4, 495-509.

[13] I. Gyöngy and T. Pröhle, *On the approximation of stochastic differential equations and on Stroock-Varadhan support theorem*, Computers Math. Applic, **19**, 1990, 65–70.

[14] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, Amsterdam, Oxford, New York: North Holland; Tokyo: Kodansha 1981.

[15] W. Jachimiak, *A note on invariance for semilinear differential equations*, Bull. Pol. Sci., **44**, 1996, 179–183.

[16] W. Jachimiak, *Invariance problem for evolution equations*, PhD Thesis, Institute of Mathematics Polish Academy of Sciences, Warsaw 1998 (in Polish).

[17] W. Jachimiak, *Stochastic invariance in infinite dimensions*, Preprint 591, Institute of Mathematics Polish Academy of Sciences, Warsaw, October 1998.

[18] A. Lunardi, "Analytic semigroups and optimal regularity in parabolic problems", Birkhauser, 1995.

[19] V. Mackevicius, *On the support of the solution of stochastic differential equations*, Lietuvos Matematikow Rinkings, **36**(1), 1986, 91–98.

[20] A. Milian, *Nagumo's type theorems for stochastic equations*, PhD Thesis, Institute of Mathematics Polish Academy of Sciences, 1994.

[21] A. Milian, *Invariance for stochastic equations with regular coefficients*, Stochastic Analysis and Applications, **15**, 1997, 91–101.

[22] A. Millet and M. Sanz-Sole, *The support of the solution to a hyperbolic spde*, Probab. Th. Rel. Fields, **84**, 1994, 361–387.

[23] A. Millet and M. Sanz-Sole, *Approximation and support theorem for a two space-dimensional wave equation*, Mathematical Sciences Research Institute, Preprint No 1998-020, Berkeley, California.

[24] A. Millet and M. Sanz-Solé, *Approximation and support theorem for a wave equation in two space dimensions*, Bernoulli, **6**(5), 2000, 887–915.

[25] A. Millet and M. Sanz-Solé, *A simple proof of the support theorem for diffusion processes*, Séminaire de Probabilitiés, XXVIII, LNIM **1583**, Springer 1994, 36–48.

[26] A. Millet and M. Sanz-Solé, *The support of the solution to a hyperbolic SPDE*, Probab. Th. Rel. Fields, **98**(3), 1994, 361–387.

[27] T. Nakayama, *Support theorem for mild solutions of SDE's in Hilbert spaces*, J. Math. Sci. Univ. Tokyo, **11**(3), 2004, 245–311.

[28] T. Nakayama, *Viability theorem for SPDE's including HJM framework*, J. Math. Sci. Univ. Tokyo, **11**(3), 2004, 313–324.

[29] E. Pardoux, *Equations aux Dérivées Partielles Stochastiques Nonlinéaires Monotones, Étude de solutions fortes du type Itô*, Thèse, Paris-Sud, Orsay, 1975.

[30] N. Pavel, *Invariant sets for a class of semilinear equations of evolution*, Nonl. Anal. Theor., **1**, 1977, 187–196.

[31] N. Pavel, *Differential equations, flow invariance*, Pitman Lecture Notes, Boston 1984.

[32] L. C. G. Rogers and D. Williams, *Diffusion, Markov Processes and Martingales*, Cambridge University Press, 2002.

[33] T. Runst and W. Sickel, *Sobolev Spaces of Fractional Orders, Nemytskij Operators and Nonlinear Partial Differential Equations*, W. de Gruyter, 1996.

[34] G. Tessitore and J. Zabczyk, *Trotter's formula for evolution semigroups*, Semigroup Forum, **63**, vol 1, 2001

[35] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, North-Holland, 1978.

[36] K. Twardowska, *An approximation theorem of Wong-Zakai type for nonlinear stochastic partial differential equations*, Stoch. Anal. Appl., **13**, 1995, 601–626.

[37] K. Twardowska, *Approximation theorems of Wong-Zakai type for stochastic differential equations in infinite dimensions*, Disserationes Mathematicae, CCCXXV, 1993.

[38] D.W. Stroock and S.R.S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings 6th Berkeley Symposium Math. Statist. Probab., vol. 3, University of California Press, Berkeley 1972, 333–359.

[39] E. Wong and M. Zakai, *On the relationship between ordinary and stochastic differential equations*, Internat. J. Engin. Sci., **3**, 1965, 213- 229.

[40] E. Wong and M. Zakai, *Riemann- Stieltjes approximations of stochastic integrals*, Z. Wahrscheinlichkeitstheorie verw. Geb. **12**, 1969, 87- 97.

[41] J. Zabczyk, *Stochastic invariance and consistency of financial models*,Rendiconti Mat. Acc. Lincei, **9**, vol. 11, 2000, 67–80

[42] J. Zabczyk, Parabolic equations on Hilbert spaces, in *Stochastic PDE's and Kolmogorov Equations in Infinite Dimensions* G. Da Prato Ed., LNIM **1715**, Springer-Verlag 1999.

Gianmario Tessitore, Dipartimento di Matematica e Applicazioni, Università di Milano-Bicocca, via R. Cozzi 53 - Edificio U5, 20125 MILANO, Italy.

E-mail address: gianmario.tessitore@unimib.it

Jerzy Zabczyk: Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8, 00-950 Warszawa, Poland.

E-mail address: zabczyk@impan.gov.pl