# DUALITY IN REFINED SOBOLEV-MALLIAVIN SPACES AND WEAK APPROXIMATION OF SPDE

ADAM ANDERSSON, RAPHAEL KRUSE, AND STIG LARSSON

Abstract. We introduce a new family of refined Sobolev-Malliavin spaces that capture the integrability in time of the Malliavin derivative. We consider duality in these spaces and derive a Burkholder type inequality in a dual norm.

The theory we develop allows us to prove weak convergence with essentially optimal rate for numerical approximations in space and time of semilinear parabolic stochastic evolution equations driven by Gaussian additive noise. In particular, we combine a standard Galerkin finite element method with backward Euler timestepping. The method of proof does not rely on the use of the Kolmogorov equation or the Itō formula and is therefore non-Markovian in nature. Test functions satisfying polynomial growth and mild smoothness assumptions are allowed, meaning in particular that we prove convergence of arbitrary moments with essentially optimal rate.

## 1. Introduction

The classical Sobolev-Malliavin spaces capture the integrability in the chance parameter of a random variable and its Malliavin derivatives. In many situations, where Malliavin calculus is used, in particular, for stochastic evolution equations, the Malliavin derivative is a stochastic process. One purpose of this paper is to introduce a refined family of Sobolev-Malliavin spaces that capture the integrability properties of the Malliavin derivative with respect to its time parameter. It turns out that the Malliavin derivative of the solution to a parabolic stochastic evolution equation has, depending on the regularity of the noise, good integrability properties in time and, in the case of trace class noise, it is even bounded. However, the main purpose of the new feature is not to measure regularity in a refined way, but to exploit that the corresponding dual norms are weaker with respect to integrability in time.

Let $(H, \|\cdot\|, \langle\cdot,\cdot\rangle)$ be a separable Hilbert space and $Q \in \mathcal{L}(H)$ be a selfadjoint positive semidefinite linear operator on $H$. We define the space $H_0 = Q^{\frac{1}{2}}(H)$ and let $\mathcal{L}_2^0 = \mathcal{L}_2(H_0, H)$ be the space of Hilbert-Schmidt operators from $H_0$ to $H$. We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbf{P})$ on which an $L^2([0,T], H_0)$-isonormal process is defined. For a differentiable random variable $X$ the Malliavin derivative $DX = (D_t X)_{t\in[0,T]}$ with respect to the isonormal process, is an $\mathcal{L}_2^0$-valued stochastic process. We introduce, for $p, q \geq 2$, the refined Sobolev-Malliavin

2010 *Mathematics Subject Classification.* 60H15, 60H07, 65C30, 65M60.

*Key words and phrases.* SPDE, finite element method, backward Euler, weak convergence, convergence of moments, Malliavin calculus, duality, spatio-temporal discretization.

spaces $\mathbf{M}^{1,p,q}(H)$ of random variables $X \in L^2(\Omega, H)$ such that

$$\|X\|_{\mathbf{M}^{1,p,q}(H)} = \Big(\|X\|^p_{L^p(\Omega,H)} + \|DX\|^p_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))}\Big)^{\frac{1}{p}} < \infty.$$

The classical Sobolev-Malliavin spaces are obtained for $q = 2$. We use the refined spaces in a duality argument based on the Gelfand triple

$$\mathbf{M}^{1,p,q}(H) \subset L^2(\Omega, H) \subset \mathbf{M}^{1,p,q}(H)^*.$$

A key ingredient is the following inequality for the $H$-valued stochastic Itō-integral $\int_0^T \Phi \, \mathrm{d}W$ in the dual norm of $\mathbf{M}^{1,p,q}(H)$, where $W$ is a cylindrical $Q$-Wiener process and $\Phi \in L^p(\Omega, L^2([0,T], \mathcal{L}_2^0))$ is a predictable stochastic process. In Theorem 3.5 we show

$$\Big\| \int_0^T \Phi(t) \, \mathrm{d}W(t) \Big\|_{\mathbf{M}^{1,p,q}(H)^*} \le \big\|\Phi\big\|_{L^{p'}(\Omega, L^{q'}([0,T],\mathcal{L}_2^0))}, \tag{1.1}$$

where $p', q'$ are the conjugate exponents to $p, q \ge 2$. We apply this inequality in situations, where one usually relies on the Burkholder-Davis-Gundy inequality, see Lemma 2.2. There the $L^2(\Omega, H)$-norm of the stochastic integral is bounded in terms of the $L^p(\Omega, L^2([0,T], \mathcal{L}_2^0))$-norm of $\Phi$, whereas here the dual norm of the integral is bounded by the $L^{p'}(\Omega, L^{q'}([0,T], \mathcal{L}_2^0))$-norm of $\Phi$. Since $q' \le 2$, this allows stronger singularities with respect to $t$.

In defining the spaces $\mathbf{M}^{1,p,q}(H)$ some care needs to be taken. For $q \ge 2$ we define the Malliavin derivative on a non-standard core $\mathcal{S}^q(H)$, see (3.2), (3.3), of smooth and cylindrical random variables, more regular than in the classical theory in which $q = 2$. By proving that the operator $D \colon \mathcal{S}^q(H) \to L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))$ is well defined and closable, we show that $\mathbf{M}^{1,p,q}(H)$ are Banach spaces. The proofs are rather elementary and rely to a large extent on existing results for the case $q = 2$. The spaces are new to the best of our knowledge, although there are similarities with the Hida and Kondratiev spaces, see, e.g., [4] or [23].

The motivation for introducing the spaces described above is found in our aim to develop new methods for the analysis of the weak error of numerical approximations of semilinear parabolic stochastic partial differential equations of the form

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = F(X(t))\,\mathrm{d}t + \mathrm{d}W(t),\ t \in (0,T]; \quad X(0) = X_0. \tag{1.2}$$

Both space-time white noise and trace class noise are considered and the nonlinearity $F$ is allowed to be a Nemytskii operator. See Assumption 2.3 below for precise conditions on $A$, $F$, $W$, $X_0$. We treat discretizations in space and time, allowing for any spatial discretization scheme that satisfies the abstract Assumption 2.4 below. We verify this assumption for piecewise linear finite element approximations of the heat equation. Discretization in time is performed by the semi-implicit backward Euler method. Our main result, weak convergence of essentially optimal rate, is stated in Theorem 4.4.

Weak convergence for linear stochastic evolution equations was studied in [14], [16], [19], [28], [29], [31], [35] and the works [5], [6], [7], [21], [22], [27, Chapt. 5], [44], [45], [46] treat semilinear equations with additive noise. Of these [27, Chapt. 5] is unique in that it treats a nonglobal Lipschitz drift term. In [8], [9] the authors study weak convergence for stochastic ordinary delay differential equations. Most of these works are based on Itō's formula and Kolmogorov's equation. It becomes apparent while reading the literature that proving weak convergence of optimal order is a challenging task. Semilinear equations with multiplicative noise was treated in [2], [12], [15] but only [12] covers noise more general than linear. No results are known for multiplicative noise in the form of a nonlinear Nemytskii operator. As in [5], [7], [22], [27, Chapt. 5], [44], [45] we allow $F$ to be a nonlinear Nemytskii operator.

Let $X, Y \in L^2(\Omega, H)$ and $\varphi\colon H \to \mathbf{R}$ have two continuous Fréchet derivatives of polynomial growth. Our technique relies on the following linearization of the weak error

$$\mathbf{E}\big[\varphi(X) - \varphi(Y)\big] = \mathbf{E}\big[\big\langle \tilde{\varphi}, X - Y \big\rangle\big], \quad \text{where} \quad \tilde{\varphi} = \int_0^1 \varphi'(\varrho X + (1-\varrho)Y)\,\mathrm{d}\varrho,$$

introduced in [10] and [31] independently. The paper [10] then proceeds by using an adjoint problem. Our method is the following: If $V \subset L^2(\Omega, H) \subset V^*$ is a Gelfand triple such that $\tilde{\varphi} \in V$, then we obtain by duality

$$\big|\mathbf{E}\big[\varphi(X) - \varphi(Y)\big]\big| \leq \big\|\tilde{\varphi}\big\|_V \big\|X - Y\big\|_{V^*}.$$

With a good choice of $V$, the error converges in the $V^*$-norm with twice the rate of convergence in the $L^2(\Omega, H)$-norm, which is the expected rate of weak convergence. For linear equations we prove that $V = \mathbf{M}^{1,p,p}(H)$ is a good choice for some $p > 2$. The main part of the error $X - Y$ is then a stochastic convolution $\int_0^T E(T-t)\,\mathrm{d}W(t)$. Bounding the error operator $E(T-t)$ in the apropriate norm yields convergence at the price of a singularity at $t = T$. By using the inequality (1.1) on this integral with sufficiently large $p = q > 2$, we may integrate a stronger singularity and obtain a higher rate of convergence. For semilinear equations the main difference is that a term involving $F(X) - F(Y)$ appears. We then use $V = \mathbf{G}^{1,p}(H) = \mathbf{M}^{1,p,p}(H) \cap L^{2p}(\Omega, H)$. In Lemma 3.9 we show that $F\colon V^* \to V^*$ is locally Lipschitz with a constant depending on $\|X\|_{\mathbf{M}^{1,2p,p}(H)}$, $\|Y\|_{\mathbf{M}^{1,2p,p}(H)}$. The choice of a stronger $V$-norm is necessary in order to control the nonlinearity in this way. After bounding these norms, we may use a standard Gronwall argument to bound $\|X - Y\|_{V^*}$.

As our method does not rely on the use of Kolmogorov's equation or Itō's formula, it extends to non-Markovian equations. In the work [1] our method is used to prove weak convergence for semilinear stochastic Volterra equations driven by additive noise. Such equations suffer from the lack of a Kolmogorov equation and therefore the classical proof is not feasible. We hope that our method will enable weak error analysis for other non-Markovian equations such as for instance random evolution PDEs. In this context we mention the work [8] in which non-Markovian stochastic ordinary delay equations with delay in the diffusion is treated with a completely different method, relying on the tame Itō formula from the anticipating stochastic calculus. For a discussion of the difficulties that arise in connection with a possible extension to multiplicative noise, see Subsection 4.3 below.

An additional advantage of the present work is that we only require the test function $\varphi$ to have two continuous Fréchet derivatives of polynomial growth. This means, in particular, that we prove convergence of arbitrary moments with the higher rate. Except in [31] for the case of linear equations, the test function in the previous weak error analysis is assumed to have bounded derivatives and convergence of moments is treated separately, for example, in [11]. In addition, our weak error estimate in Theorem 4.4 is uniform with respect to the time partition unlike earlier results in the literature.

The paper is organized as follows. In Section 2 we present preliminary material and our basic assumptions on the stochastic partial differential equation and the numerical scheme. The core of the paper is Section 3, which contains our extensions of the Malliavin calculus. In 3.1 we introduce the refined Sobolev-Malliavin spaces and prove that they are well defined. Duality of our new spaces is treated in 3.2, with the inequality (1.1) and a local Lipschitz bound as the main results. In 3.3 and 3.4 regularity in terms of the new spaces is proved for the solution to the stochastic evolution equation and its approximation, respectively. Section 4 contains the weak convergence analysis. In 4.1 we restrict the discussion to approximations of the stochastic convolution and in 4.2 we treat semilinear equations. Finally, in Section 5 we verify our assumption on the numerical method for a standard finite element approximation of the heat equation.

## 2. Setting and preliminaries

2.1. **Analytic preliminaries.** Let $(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U)$ and $(V, \|\cdot\|_V, \langle\cdot,\cdot\rangle_V)$ be separable Hilbert spaces and let $\mathcal{L}(U,V)$ be the Banach space of all bounded linear operators $U \to V$. If $U = V$, then we write $\mathcal{L}(U) = \mathcal{L}(U,U)$ and if $U = H$, we abbreviate $\mathcal{L} = \mathcal{L}(H)$. We denote by $\mathcal{L}_2(U,V) \subset \mathcal{L}(U,V)$ the subspace of all Hilbert-Schmidt operators endowed with the standard norm and inner product

$$\|T\|_{\mathcal{L}_2(U,V)} = \Big(\sum_{j\in\mathbf{N}} \|Tu_j\|_V^2\Big)^{\frac{1}{2}}, \quad \langle S,T\rangle_{\mathcal{L}_2(U,V)} = \sum_{j\in\mathbf{N}} \langle Su_j, Tu_j\rangle_V,$$

where both are independent of the particular choice of ON-basis $(u_j)_{j\in\mathbf{N}} \subset U$.

For separable Hilbert spaces $U_1,\dots,U_m$, $m \in \mathbf{N}$, we denote by $\mathcal{L}^{[m]}(U_1 \times \cdots \times U_m, V)$ the space of multi-linear operators $b\colon U_1 \times \cdots \times U_m \to V$. We use the notation $b\cdot(u_1,\dots,u_m) = b(u_1,\dots,u_m)$ for $u_i \in U_i$, $i = 1,\dots,m$, to emphasize that $b$ is multi-linear. If $U = U_1 = \ldots = U_m$ we abbreviate $\mathcal{L}^{[m]}(U\times\cdots\times U, V) = \mathcal{L}^{[m]}(U,V)$. The norm $\|b\|_{\mathcal{L}^{[m]}(U_1\times\cdots\times U_m,V)}$ is the smallest constant $C$ such that

$$\|b\cdot(u_1,\dots,u_m)\|_V \le C\|u_1\|_{U_1}\cdots\|u_m\|_{U_m}, \quad \forall u_i \in U_i,\ i = 1,\dots,m. \tag{2.1}$$

By $\mathcal{C}^m(U,V)$ we denote the space of all mappings $\phi\colon U \to V$ with continuous Fréchet derivatives of order $m$, $\mathcal{C}^m_{\mathrm{b}}(U,V)$ is the subspace with $m \ge 1$ bounded derivatives $\phi',\dots,\phi^{(m)}$ (note that $\phi$ needs not be bounded), and $\mathcal{C}^m_{\mathrm{p}}(U,V)$ denotes the analogous space with derivatives of polynomial growth. On $\mathcal{C}^m_{\mathrm{b}}(U,V)$ we use the natural seminorm $|\phi|_{\mathcal{C}^m_{\mathrm{b}}} = \sup_{x\in U}\|\phi^{(m)}(x)\|_{\mathcal{L}^{[m]}(U,V)}$. We define $\mathcal{C}^0_{\mathrm{b}}(U,V)$ to be all bounded continuous mappings $U \to V$, endowed with the uniform norm. The first derivative of $\phi \in \mathcal{C}^1(U,V)$ is an operator $\phi'(x) \in \mathcal{L}(U,V) = \mathcal{L}^{[1]}(U,V)$ for every $x \in U$. When $V = \mathbf{R}$ we may identify $\phi'(x) \in \mathcal{L}(U,\mathbf{R}) = U^*$ with its gradient $\phi'(x) \in U$ via $\phi'(x)\cdot u = \langle\phi'(x),u\rangle_U$ by the Riesz representation theorem.

Similarly, for $\phi \in \mathcal{C}^2(U, \mathbf{R})$ we will sometimes identify $\phi''(x) \in \mathcal{L}^{[2]}(U, \mathbf{R})$ with an operator $\phi''(x) \in \mathcal{L}(U)$ via $\phi''(x) \cdot (u_1, u_2) = \langle \phi''(x)u_1, u_2 \rangle_U$. By the mean value theorem we have, for $\phi \in \mathcal{C}^1(U, V)$,

$$\phi(x) = \phi(y) + \int_0^1 \phi'(y + \rho(x-y)) \cdot (x-y) \, \mathrm{d}\rho, \quad x, y \in U. \tag{2.2}$$

We will use the following version of Gronwall's Lemma, for a proof see [17, Lemma 7.1].

**Lemma 2.1.** *Let $T > 0$, $N \in \mathbf{N}$, $k = \frac{T}{N}$, and $t_n = nk$ for $0 \le n \le N$. If $(\varphi_j)_{j=1}^N$ are nonnegative real numbers with*

$$\varphi_n \le C_1 \, (1 + t_n^{-1+\mu}) + C_2 \, k \sum_{j=0}^{n-1} t_{n-j}^{-1+\nu} \varphi_j, \quad 1 \le n \le N,$$

*for some constants $C_1, C_2 \ge 0$ and $\mu, \nu > 0$, then there exists a constant $C = C(\mu, \nu, C_2, T)$ such that*

$$\varphi_n \le C \, C_1 \, (1 + t_n^{-1+\mu}), \quad 1 \le n \le N.$$

We sometimes write $a \lesssim b$ to denote $a \le Cb$ for some constant $C > 0$. Constants arising from the estimates (2.3), (2.4), (2.10) and (2.12), as well as trivial numerical constants, will be suppressed with this symbol.

2.2. **Stochastic preliminaries.** Let $(H, \|\cdot\|, \langle \cdot, \cdot \rangle)$ be a separable Hilbert space and let $Q \in \mathcal{L} = \mathcal{L}(H)$ be a selfadjoint, positive semidefinite operator on $H$ and $Q^{\frac{1}{2}}$ its unique positive square root. The space $H_0 = Q^{\frac{1}{2}}(H)$ is a Hilbert space with scalar product $\langle u, v \rangle_{H_0} = \langle Q^{-\frac{1}{2}} u, Q^{-\frac{1}{2}} v \rangle$. We denote by $\mathcal{L}_2^0 = \mathcal{L}_2(H_0, H)$ the space of Hilbert-Schmidt operators $H_0 \to H$. We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbf{P})$ and the corresponding Bochner spaces $L^p(\Omega, V) = L^p((\Omega, \mathcal{F}, \mathbf{P}), V)$, $p \in [1, \infty]$, $V$ a Banach space. We abbreviate $L^2(\Omega) = L^2(\Omega, \mathbf{R})$. We assume that $(W(t))_{t \in [0,T]}$ is a cylindrical $Q$-Wiener process, meaning that $W \in \mathcal{C}([0,T], \mathcal{L}(H_0, L^2(\Omega)))$ is such that $t \mapsto W(t)u$ is an $\mathcal{F}_t$-predictable real-valued Brownian motion for every $u \in H_0$ and

$$\mathbf{E}\big[W(s)u \, W(t)v\big] = \min(s,t) \langle u, v \rangle_{H_0}, \quad u, v \in H_0, \; s, t \in [0,T].$$

For predictable $\Phi \in L^2([0,T] \times \Omega, \mathcal{L}_2^0)$ the $H$-valued stochastic Itō-integral

$$\int_0^T \Phi(t) \, \mathrm{d}W(t) \in L^2(\Omega, H),$$

is a well defined random variable. For details on the construction of cylindrical Wiener processes and the corresponding stochastic integral we refer to [13, 39, 42]. For technical reasons we assume that the $\sigma$-field $\mathcal{F}$ is generated by $(W(t))_{t \in [0,T]}$ and the filtration $(\mathcal{F}_t)_{t \in [0,T]}$ is the natural filtration associated with $(W(t))_{t \in [0,T]}$.

We cite the following special case of Burkholder's inequality [13, Lemma 7.2].

**Lemma 2.2.** *Let $(\Phi(t))_{t\in[0,T]}$ be a predictable and $\mathcal{L}_2^0$-valued process such that $\|\Phi\|_{L^p(\Omega,L^2([0,T],\mathcal{L}_2^0))} < \infty$ for some $p \geq 2$. Then there exists a constant $C_p$, such that*

$$\Big\| \int_0^T \Phi(s)\,\mathrm{d}W(s) \Big\|_{L^p(\Omega,H)} \leq C_p \|\Phi\|_{L^p(\Omega,L^2([0,T],\mathcal{L}_2^0))}.$$

2.3. **The stochastic equation.** We study equation (1.2) under the following assumption and recall that the solution $X$ takes values in $H$.

**Assumption 2.3.** (i) *Let $(A,\mathcal{D}(A))$ be a linear operator on $H$ such that $A^{-1} \in \mathcal{L}(H)$ exists and $-A$ is the generator of an analytic semigroup $(S(t))_{t\geq 0}$ of bounded linear operators $S(t) = \mathrm{e}^{-tA}$ on $H$.*

(ii) *The initial value $X_0$ is deterministic and satisfies $X_0 \in \dot{H}^{2\beta}$, for some $\beta \in (0,1]$, where $\dot{H}^\alpha \subset H$ denotes the domain of $A^{\frac{\alpha}{2}}$.*

(iii) *The covariance operator $Q$ satisfies $\|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0} = \|A^{\frac{\beta-1}{2}} Q^{\frac12}\|_{\mathcal{L}_2} < \infty$, for the same $\beta$ as in* (ii).

(iv) *The drift $F\colon H \to H$ is assumed to be twice differentiable in the sense $F \in \mathcal{C}_{\mathrm{b}}^1(H,H) \cap \mathcal{C}_{\mathrm{b}}^2(H,\dot{H}^{-1})$, where $\dot{H}^{-1}$ is defined below.*

Under Assumption 2.3 (i) the fractional powers $A^{\frac{r}{2}}$ for $r \in \mathbf{R}$ are well defined, see [38, Section 2.6]. We define the norms $\|v\|_r = \|A^{\frac{r}{2}} v\|$ and let $\dot{H}^r = \mathcal{D}(A^{\frac{r}{2}})$ for $r \geq 0$. For $r < 0$ we define $\dot{H}^r$ as the closure of $H$ under the norm $\|v\|_r$. The spaces $\dot{H}^r \subset H \subset \dot{H}^{-r}$ form a Gelfand triple for $r > 0$.

The analytic semigroup $(S(t))_{t\geq0}$ generated by $-A$ satisfies, see [38, Section 2.6],

$$\|A^\varrho S(t)\|_{\mathcal{L}} \leq C_\varrho t^{-\varrho}, \quad t > 0,\ \varrho \geq 0, \tag{2.3}$$

$$\|(S(t) - I)A^{-\varrho}\|_{\mathcal{L}} \leq C_\varrho t^{\varrho}, \qquad t \geq 0,\ 0 < \varrho \leq 1. \tag{2.4}$$

Under Assumption 2.3, the stochastic equation (1.2) has a mild solution $X \in \mathcal{C}([0,T], L^p(\Omega,H))$, for every $p \geq 2$, in the sense that it satisfies the integral equation

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))\,\mathrm{d}s + \int_0^t S(t-s)\,\mathrm{d}W(s), \quad t \in [0,T], \tag{2.5}$$

and

$$\sup_{t\in[0,T]} \|X(t)\|_{L^p(\Omega,H)} \leq C(1 + \|X_0\|). \tag{2.6}$$

For every $\gamma \in [0,\beta)$ the solution satisfies $X(t) \in \dot{H}^\gamma$, $\mathbf{P}$-a.s., for all $t \in [0,T]$. For more details we refer to [13], [25], [32], and the references therein.

In [2] and [15] the authors assume $F \in \mathcal{C}_{\mathrm{b}}^2(H,H)$, which works well for the analysis but has the following disadvantage: If $D \subset \mathbf{R}^d$, $d = 1,2,3$, $H = L^2(D)$ and $F\colon H \to H$ is a Nemytskii operator, i.e., a mapping in the form $g \mapsto F(g) = f(g(\cdot))$, where $f \in \mathcal{C}_{\mathrm{b}}^2(\mathbf{R},\mathbf{R})$, then $F \in \mathcal{C}_{\mathrm{b}}^1(H,H)$ but in general $F \notin \mathcal{C}_{\mathrm{b}}^2(H,H)$. This disqualifies the most interesting examples of nonlinearities $F$. On the other hand by the Sobolev embedding theorem $F \in \mathcal{C}_{\mathrm{b}}^2(H,\dot{H}^{-\frac{d}{2}+\epsilon})$ for $\epsilon > 0$ and hence Assumption 2.3 admits Nemytskii operators for $d = 1$. See [44, Example 5.1] for a verification. For $d = 2,3$ one needs to assume $F \in \mathcal{C}_{\mathrm{b}}^2(H,\dot{H}^{-s})$ with $s > 1$, which works for spectral Galerkin approximations but not for the finite element method due to the restriction on $\varrho$ in (2.11) below. In [1] this restriction is removed, allowing for finite element discretization also for $d = 2,3$. Papers that include Nemytskii operators are [5], [7], [22], [44], [45] and our Assumption 2.3 (iv) is a reformulation of [44, Assumption 5.1].

2.4. **Approximation of the solution.** We approximate equation (1.2) in finite-dimensional approximation spaces $V_h \subseteq H$, $h \in (0,1]$. The parameter $h \in (0,1]$ is a refinement parameter. We denote by $P_h \colon H \to V_h$ the orthogonal projector onto $V_h$ and by $(A_h)_{h\in(0,1]}$ a family of operators $A_h \colon V_h \to V_h$ approximating $A$. The assumptions on $(V_h)_{h\in(0,1]}$, and $(A_h)_{h\in(0,1]}$ are given in Assumption 2.4 below.

For the time discretization let $k \in (0,1)$ be the constant step size. We define the discrete time points by $t_n = nk$, $n = 0,\dots,N$, where $N = N(k) \in \mathbf{N}$ is determined by $t_N \le T < t_N + k$. We define the operator $S_{h,k} = (I + kA_h)^{-1}P_h$ and notice that $S_{h,k}Q^{\frac12} \in \mathcal{L}_2(H)$, since $S_{h,k}$ is a finite rank operator. Hence, it is a valid integrand for the stochastic integral. Our completely discrete scheme is to find the recursive sequence $(X^n_{h,k})_{n=0}^N \subset V_h$ given by the semi-implicit Euler-Maruyama method:

$$
\begin{aligned}
&X^{n+1}_{h,k} = S_{h,k}X^n_{h,k} + kS_{h,k}F(X^n_{h,k}) + \int_{t_n}^{t_{n+1}} S_{h,k}\,\mathrm{d}W(s), \quad n = 0,\dots,N-1;\\
&X^0_{h,k} = P_hX_0.
\end{aligned} \tag{2.7}
$$

By iterating (2.7) we obtain the discrete analog of (2.5)

$$
\begin{aligned}
X^n_{h,k} = S^n_{h,k}P_hX_0 &+ k\sum_{j=0}^{n-1} S^{n-j}_{h,k}F(X^j_{h,k})\\
&+ \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} S^{n-j}_{h,k}\,\mathrm{d}W(t), \quad n = 0,\dots,N.
\end{aligned} \tag{2.8}
$$

Further, we define the error operators $E^n_{h,k}$, $h,k \in (0,1]$, by

$$
E^n_{h,k} := S(nk) - S^n_{h,k}. \tag{2.9}
$$

We now state our assumption on the numerical discretization.

**Assumption 2.4.** *The linear operators $A_h \colon V_h \to V_h$ and the orthogonal projectors $P_h \colon H \to V_h$, $h \in (0,1]$, satisfy*

$$
\|A_h^{\varrho}S^n_{h,k}\|_{\mathcal{L}} \le Ct_n^{-\varrho}, \quad n = 1,\dots,N, \quad \varrho \ge 0, \tag{2.10}
$$

$$
\|A_h^{-\varrho}P_hA^{\varrho}\|_{\mathcal{L}} \le C, \quad 0 \le \varrho \le \tfrac12, \tag{2.11}
$$

*uniformly in $h,k \in (0,1]$, and, for $0 \le \theta \le 2$, $-\theta \le \varrho \le \min(1, 2-\theta)$,*

$$
\|E^n_{h,k}A^{\frac{\varrho}{2}}\|_{\mathcal{L}} \le C\big(h^{\theta} + k^{\frac{\theta}{2}}\big)t_n^{-\frac{\theta+\varrho}{2}}, \quad n = 1,\dots,N. \tag{2.12}
$$

We emphasize that the restriction $\varrho \le \frac12$ in (2.11) is dictated by our desire to include standard finite element spaces, for which $V_h \subset \dot{H}^1$, and no better. We remark that the error estimate (2.12) is non-standard, due to the low regularity regime we consider. In fact, when $\varrho \ge 0$, it corresponds to an error estimate for the deterministic linear equation with rough initial data, i.e., $S(t)X_0 = S(t)A^{\frac{\varrho}{2}}x$

with $x \in H$, so that $X_0 = A^{\frac{\varrho}{2}} x \in \dot{H}^{-\varrho}$. We verify (2.12) in Section 5 for the finite element method and the heat equation by means of interpolation techniques, using already established results from [30, 31]. By [30, Example 3.4], spectral Galerkin approximations also fit under our Assumption 2.4.

Finally, for future reference, we formulate an important consequence of the smoothing properties (2.3) and (2.10), (2.11), respectively, in conjunction with the assumption on the covariance operator in Assumption 2.3 (iii).

**Lemma 2.5.** *Let Assumptions 2.3 and 2.4 hold with $\beta \in [0,1]$. Let $q \in [2, \frac{2}{1-\beta})$ with $q = \infty$ allowed if $\beta = 1$. Then*

$$\|S\|_{L^q([0,T],\mathcal{L}_2^0)} \le C \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0}$$

*and*

$$\Big( k \sum_{j=1}^{N} \|S_{h,k}^j\|_{\mathcal{L}_2^0}^q \Big)^{1/q} \le C \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0}.$$

*Proof.* Let first $q < \infty$. By (2.3) with $\varrho = \frac{1-\beta}{2}$ we get

$$\begin{aligned}
\|S\|_{L^q([0,T],\mathcal{L}_2^0)}^q &= \int_0^T \|S(t)\|_{\mathcal{L}_2^0}^q \,\mathrm{d}t \le \int_0^T \|A^{\frac{1-\beta}{2}} S(t)\|_{\mathcal{L}}^q \,\mathrm{d}t \, \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0}^q \\
&\le C \int_0^T t^{-q\frac{1-\beta}{2}} \,\mathrm{d}t \, \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0}^q \le C \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0}^q.
\end{aligned}$$

For the second inequality we use instead (2.10), (2.11) with $\varrho = \frac{1-\beta}{2}$ to get

$$\begin{aligned}
\|S_{h,k}^j\|_{\mathcal{L}_2^0} &\le \|A_h^{\frac{1-\beta}{2}} S_{h,k}^j\|_{\mathcal{L}} \|A_h^{\frac{\beta-1}{2}} P_h\|_{\mathcal{L}_2^0} \\
&\le \|A_h^{\frac{1-\beta}{2}} S_{h,k}^j\|_{\mathcal{L}} \|A_h^{\frac{\beta-1}{2}} P_h A^{\frac{1-\beta}{2}}\|_{\mathcal{L}} \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0} \\
&\le C t_j^{-\frac{1-\beta}{2}} \|A^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0},
\end{aligned}$$

which can be summed as desired. The case when $q = \infty$, $\beta = 1$ is now obvious. □

## 3. Malliavin calculus

The papers [20] and [34] are the earliest works to treat Malliavin calculus for stochastic evolution equations in the Hilbert space framework. Later it was used in several papers related to optimal control of stochastic partial differential equations, in particular, in connection with backward stochastic differential equations [18] and backward stochastic Volterra integral equations in Hilbert spaces [3]. Malliavin differentiability of solutions to stochastic evolution equations is proved in [18]. There are also works using the Malliavin calculus for specific equations outside the setting of the present paper and it is more extensively developed for equations studied in the framework of [43], see the book [40]. We mention also the papers [2], [5], [6], [8], [10], [15], [21], [22], [26], [27, Chapt. 5], [46], where the Malliavin calculus is applied to the problem of proving weak convergence. Below we take a new direction and introduce in Subsection 4.1 a family of refined Sobolev-Malliavin spaces. We show in Subsection 4.2 that these spaces are particularly useful in connection with duality.

3.1. **Refined Sobolev-Malliavin spaces.** Let $I\colon L^2([0,T],H_0) \to L^2(\Omega)$ be the mapping given by

$$I(\phi) = \int_0^T \phi(t)\,\mathrm{d}W(t), \quad \phi \in L^2([0,T],H_0),$$

where we identify $L^2([0,T],H_0) \cong L^2([0,T],\mathcal{L}_2(H_0,\mathbf{R}))$. This identification is important since an $\mathbf{R}$-valued stochastic integral has an $L^2([0,T],\mathcal{L}_2(H_0,\mathbf{R}))$-valued integrand. Fix an ON-basis $(\phi_j)_{j\in\mathbf{N}} \subset L^2([0,T],H_0)$, let $\mathcal{P}_n$ be the set of random variables given by $n$:th order polynomials of the random variables $(I(\phi_j))_{j\in\mathbf{N}}$. The set $\mathcal{P} = \cup_{n\in\mathbf{N}}\mathcal{P}_n$ is independent of the choice of basis, see [24], and

$$\mathcal{P} \subset L^p(\Omega) \text{ is dense for } 1 \leq p < \infty. \tag{3.1}$$

Let $2 \leq q \leq \infty$ and let the mapping $i\colon L^q([0,T],H_0) \to L^2([0,T],H_0)$ denote the canonical embedding. Let $\mathcal{S}^q$ be the set of random variables $F$ of the form

$$\begin{aligned} &F = f(I(i(\phi_1)),\dots,I(i(\phi_n))),\\ &f \in \mathcal{C}^1_{\mathrm{p}}(\mathbf{R}^n,\mathbf{R}),\ (\phi_j)_{j=1}^n \subset L^q([0,T],H_0),\ n \in \mathbf{N}. \end{aligned} \tag{3.2}$$

The class $\mathcal{S}^2$ is standard in Malliavin calculus and is usually denoted by $\mathcal{S}$. Our definition coincides with that in [31] but in the standard work [36] and many other works $\mathcal{C}^\infty_{\mathrm{p}}(\mathbf{R}^n,\mathbf{R})$ is used instead of $\mathcal{C}^1_{\mathrm{p}}(\mathbf{R}^n,\mathbf{R})$. The classes $\mathcal{S}^q$ for $q > 2$ are new to our knowledge.

**Lemma 3.1.** *For $1 \leq p < \infty$ and $2 \leq q \leq \infty$, $\mathcal{S}^q \subset L^p(\Omega)$ is dense.*

*Proof.* Without causing confusion we also let $i$ denote the canonical embedding from $L^q([0,T],\mathbf{R})$ to $L^2([0,T],\mathbf{R})$. We notice the isomorphism $L^2([0,T],H_0) \cong L^2([0,T],\mathbf{R}) \otimes H_0$.

Since there even exists a bounded ON-basis of the space $L^2([0,T],\mathbf{R})$ we clearly find a sequence $(f_n)_{n\in\mathbf{N}} \subset L^q([0,T],\mathbf{R})$ such that $(i(f_n))_{n\in\mathbf{N}}$ is an ON-basis for $L^2([0,T],\mathbf{R})$. If $(h_n)_{n\in\mathbf{N}}$ is an ON-basis for $H_0$, then $(i(f_m)\otimes h_n)_{m,n\in\mathbf{N}}$ is an ON-basis for $L^2([0,T],\mathbf{R}) \otimes H_0$. In particular, we have that $i(f_m \otimes h_n) = i(f_m) \otimes h_n$.

Since the result (3.1) is independent of the choice of the basis, we conclude our assertion by using the sequence $(I(i(f_m \otimes h_n)))_{m,n\in\mathbf{N}}$. □

For $1 \leq p < \infty$ and $2 \leq q \leq \infty$ we define the action of the Malliavin derivative $D\colon \mathcal{S}^q \to L^p(\Omega, L^q([0,T],H_0))$ on a random variable $F$ of the form (3.2) by

$$D_t F = \sum_{j=1}^n \partial_j f(I(i(\phi_1)),\dots,I(i(\phi_n))) \otimes \phi_j(t), \quad t \in [0,T].$$

This is well defined because $\phi_1,\dots,\phi_n \in L^q([0,T],H_0)$, the random variables $I(\phi_1),\dots,I(\phi_n)$ are Gaussian with all existing moments and since $f$ has polynomial growth. By a direct modification of [31, Proposition 4.2] it does not depend on the specific representation of $F$.

We remark that for $q = 2$ the linear operator $D\colon \mathcal{S}^2 \to L^p(\Omega, L^2([0,T],H_0))$ is the standard Malliavin derivative. Technically speaking, we have restricted the domain of the Malliavin derivative to $\mathcal{S}^q \subset \mathcal{S}^2$ for $2 < q \leq \infty$. By this we have ensured that $D|_{\mathcal{S}^q}$ maps into the smaller space $L^p(\Omega, L^q([0,T],H_0)) \subset L^p(\Omega, L^2([0,T],H_0))$.

We define the Malliavin derivative for $H$-valued random variables as in [31, Chapt. 4], [36, Chapt. 1]. For this we denote by $\mathcal{S}^q(H)$ the collection of all $H$-valued smooth random variables of the form

$$X = \sum_{j=1}^{n} h_j \otimes F_j, \quad h_1, \dots, h_n \in H, \quad F_1, \dots, F_n \in \mathcal{S}^q, \; n \in \mathbf{N}. \tag{3.3}$$

Since $H$ is separable and by Lemma 3.1 it follows that $\mathcal{S}^q(H)$ is dense in $L^p(\Omega, H)$ for all $1 \leq p < \infty$. The Malliavin derivative $D \colon \mathcal{S}^q(H) \to L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))$ acts in the following way:

$$D_t X = D_t \sum_{j=1}^{n} h_j \otimes F_j = \sum_{j=1}^{n} h_j \otimes D_t F_j, \quad t \in [0,T].$$

Here we did the identifications

$$H \otimes L^p(\Omega, L^q([0,T], H_0)) \cong L^p(\Omega, H \otimes L^q([0,T], H_0)) \cong L^p(\Omega, L^q([0,T], \mathcal{L}_2^0)).$$

We write $D_t^u X = D_t X u \in L^2(\Omega, H)$ for the derivative in the direction $u \in H_0$.

In the final step of its construction we extend the domain of the Malliavin derivative to its closure with respect to the graph norm. For this we recall that an unbounded operator $A \colon U \to V$ is *closable* if and only if for every $(u_n)_{n \in \mathbf{N}} \subset U$ such that $\lim_{n\to\infty} u_n = 0$ and $\lim_{n\to\infty} A u_n = v$, we have $v = 0$.

**Lemma 3.2.** *The Malliavin derivative $D \colon \mathcal{S}^q(H) \to L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))$ is closable for $1 < p < \infty$ and $2 \leq q \leq \infty$.*

*Proof.* We will use the fact that $D \colon \mathcal{S}^2(H) \to L^p(\Omega, L^2([0,T], \mathcal{L}_2^0))$ is closable for $p > 1$, [31, Proposition 4.4]. Let $(X_n)_{n \in \mathbf{N}} \subset \mathcal{S}^q(H) \subset \mathcal{S}^2(H)$ be a sequence satisfying $\lim_{n\to\infty} X_n = 0$ in $L^p(\Omega, H)$ such that $\lim_{n\to\infty} D X_n = Z$ in $L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))$ and hence also in $L^p(\Omega, L^2([0,T], \mathcal{L}_2^0))$. By the closability we have $Z = 0$ in $L^p(\Omega, L^2([0,T], \mathcal{L}_2^0))$ and hence also in $L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))$. $\square$

For $1 < p < \infty$ and $2 \leq q \leq \infty$ we can therefore consider the closure $\mathbf{M}^{1,p,q}(H)$ of $\mathcal{S}^q(H)$ with respect to the norm

$$\|X\|_{\mathbf{M}^{1,p,q}(H)} = \Big( \|X\|_{L^p(\Omega,H)}^p + \|DX\|_{L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))}^p \Big)^{\frac{1}{p}}.$$

Clearly, the spaces $\mathbf{M}^{1,p,2}(H)$, $p > 1$, coincide with the classical Sobolev-Malliavin spaces of the Malliavin calculus, which are usually denoted by $\mathbf{D}^{1,p}(H)$. The standard Malliavin derivative is uniquely extended to an operator from $\mathbf{M}^{1,p,2}(H)$ to $L^p(\Omega, L^2([0,T], \mathcal{L}_2^0))$. In addition it holds $\mathbf{M}^{1,p,q_1}(H) \subset \mathbf{M}^{1,p,q_2}(H)$ for all $\infty \geq q_1 \geq q_2 \geq 2$ and from Lemma 3.2 it follows that the restriction of the standard Malliavin derivative $D|_{\mathbf{M}^{1,p,q}(H)}$ is a well-defined operator from $\mathbf{M}^{1,p,q}(H)$ to $L^p(\Omega, L^q([0,T], \mathcal{L}_2^0))$. If $p = q$, we abbreviate $\mathbf{M}^{1,p}(H) = \mathbf{M}^{1,p,p}(H)$.

The space $\mathbf{M}^{1,2}(H)$ is a Hilbert space and it has a well developed theory of Malliavin calculus. The adjoint of the Malliavin derivative $D \colon \mathbf{M}^{1,2}(H) \subset L^2(\Omega, H) \to L^2([0,T] \times \Omega, \mathcal{L}_2^0)$ is called the *divergence* operator or the *Skorohod* integral and denoted by $\delta\colon L^2([0,T]\times\Omega,\mathcal{L}_2^0)\to L^2(\Omega,H)$ with domain $\mathcal{D}(\delta)$. The duality reads

$$\big\langle X,\delta\Phi\big\rangle_{L^2(\Omega,H)}=\big\langle DX,\Phi\big\rangle_{L^2([0,T]\times\Omega,\mathcal{L}_2^0)},\quad X\in\mathbf{M}^{1,2}(H),\ \Phi\in\mathcal{D}(\delta). \tag{3.4}$$

We refer to this as the Malliavin integration by parts formula. It is well known that for predictable $\Phi\in\mathcal{D}(\delta)$ the action of $\delta$ coincides with that of the $H$-valued Itō integral, i.e., $\delta\Phi=\int_0^T\Phi(t)\,\mathrm{d}W(t)$, [31, Proposition 4.12].

In the remainder of this subsection we state a modification of the chain rule from [31, Lemma 4.7] and a product rule for the Malliavin derivative.

**Lemma 3.3.** *Let $U,V$ be two separable Hilbert spaces and let $\gamma\in\mathcal{C}^1(U,V)$, be such that there exist constants $C$ and $r\geq 0$ with*

$$\|\gamma(u)\|_V\leq C\big(1+\|u\|_U^{1+r}\big),\quad \|\gamma'(u)\|_{\mathcal{L}(U,V)}\leq C\big(1+\|u\|_U^r\big),$$

*for all $u\in U$. Then, for $1<p<\infty$, $2\leq q\leq\infty$ and $X\in\mathbf{M}^{1,(1+r)p,q}(U)$, it follows that $\gamma(X)\in\mathbf{M}^{1,p,q}(V)$ with $\|\gamma(X)\|_{\mathbf{M}^{1,p,q}(V)}\lesssim(1+\|X\|^{1+r}_{\mathbf{M}^{1,(1+r)p,q}(U)})$ and*

$$D_t(\gamma(X))=\gamma'(X)\cdot D_tX,\quad t\in[0,T]. \tag{3.5}$$

*Proof.* Let $p>1$ be arbitrary. For $q=2$ the result follows directly from [31, Lemma 4.7]. Therefore, it suffices to show that $\|\gamma(X)\|_{\mathbf{M}^{1,p,q}(V)}<\infty$ if $X\in\mathbf{M}^{1,(1+r)p,q}(U)$ for $q>2$. Indeed, from the polynomial growth condition it follows that

$$\big\|\gamma(X)\big\|_{L^p(\Omega,V)}\leq C\big(1+\|X\|^{1+r}_{L^{(1+r)p}(\Omega,U)}\big)\leq C\big(1+\|X\|^{1+r}_{\mathbf{M}^{1,(1+r)p,q}(U)}\big).$$

Moreover, it holds

$$\begin{aligned}
&\big\|D\gamma(X)\big\|_{L^p(\Omega,L^q([0,T],\mathcal{L}_2(H_0,V)))}\\
&\quad=\big(\mathbf{E}\big[\big\|\gamma'(X)\cdot DX\big\|^p_{L^q([0,T],\mathcal{L}_2(H_0,V))}\big]\big)^{\frac1p}\\
&\quad\lesssim\big(\mathbf{E}\big[\big(1+\big\|X\big\|_U^r\big)^p\|DX\|^p_{L^q([0,T],\mathcal{L}_2(H_0,U))}\big]\big)^{\frac1p}\\
&\quad\leq\big(1+\|X\|^r_{L^{(1+r)p}(\Omega,U)}\big)\big\|DX\big\|_{L^{(1+r)p}(\Omega,L^q([0,T],\mathcal{L}_2(H_0,U)))}\\
&\quad\lesssim\big(1+\|X\|^{1+r}_{\mathbf{M}^{1,(1+r)p,q}(U)}\big),
\end{aligned}$$

where we applied the polynomial growth condition on $\gamma'$ and Hölder's inequality with exponents $(r+1)/r$ and $r+1$. This completes the proof. □

**Lemma 3.4.** *Let $U_1,U_2,V$ be separable Hilbert spaces and $1<p<\infty$, $2\leq q\leq\infty$. For $\sigma\in\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))\cap\mathcal{C}^1_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))$ it holds $\sigma(X)\cdot Y\in\mathbf{M}^{1,p,q}(V)$ for all $X\in\mathbf{M}^{1,2p,q}(U_1)$ and $Y\in\mathbf{M}^{1,2p,q}(U_2)$. In addition, we have*

$$D_t(\sigma(X)\cdot Y)=\sigma'(X)\cdot(D_tX,Y)+\sigma(X)\cdot D_tY,\quad t\in[0,T]. \tag{3.6}$$

*Proof.* The proof is done by an application of the chain rule. For this define the mapping $\gamma\colon U_1\times U_2\to V$ given by $\gamma(x,y)=\sigma(x)\cdot y$. Certainly, it holds $\gamma\in\mathcal{C}^1(U_1\times U_2,V)$ and we have $\|\gamma(x,y)\|_V=\|\sigma(x)\cdot y\|_V\leq|\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\|y\|_{U_2}$ for all $(x,y)\in U_1\times U_2$. Further, it holds

$$\gamma'(x,y)\cdot(z_1,z_2)=\sigma'(x)\cdot(z_1,y)+\sigma(x)\cdot z_2,$$

for all $(x, y), (z_1, z_2) \in U_1 \times U_2$. Therefore,

$$\begin{aligned}\big\|\gamma'(x,y)\cdot(z_1,z_2)\big\|_V &\le |\sigma|_{\mathcal{C}_\mathrm{b}^1(U_1,\mathcal{L}(U_2,V))}\|z_1\|_{U_1}\|y\|_{U_2} + |\sigma|_{\mathcal{C}_\mathrm{b}^0(U_1,\mathcal{L}(U_2,V))}\|z_2\|_{U_2}\\ &\le \max\{|\sigma|_{\mathcal{C}_\mathrm{b}^0(U_1,\mathcal{L}(U_2,V))}, |\sigma|_{\mathcal{C}_\mathrm{b}^1(U_1,\mathcal{L}(U_2,V))}\}\\ &\quad\times\big(1+\|y\|_{U_2}\big)\big(\|z_1\|_{U_1}+\|z_2\|_{U_2}\big).\end{aligned}$$

Hence, $\gamma$ satisfies the assumption of Lemma 3.3 with $r = 1$. Thus, the result follows from an application of Lemma 3.3. □

3.2. **Duality.** For any $2 \le p < \infty$, $2 \le q \le \infty$ the inclusion $\mathbf{M}^{1,p,q}(H) \subset L^2(\Omega, H)$ is dense and continuous and hence the spaces

$$\mathbf{M}^{1,p,q}(H) \subset L^2(\Omega, H) \subset \mathbf{M}^{1,p,q}(H)^*,$$

define a Gelfand triple, where we identify $L^2(\Omega, H) \cong L^2(\Omega, H)^*$ by the Riesz Representation Theorem. We denote the dual pairing of $\mathbf{M}^{1,p,q}(H)^*$ and $\mathbf{M}^{1,p,q}(H)$ by $[Z, Y]$ for $Z \in \mathbf{M}^{1,p,q}(H)^*$, $Y \in \mathbf{M}^{1,p,q}(H)$. The inclusion $L^2(\Omega, H) \subset \mathbf{M}^{1,p,q}(H)^*$ is realized through the definition $[Z, Y] = \langle Z, Y\rangle_{L^2(\Omega,H)}$ for all $Z \in L^2(\Omega, H)$, $Y \in \mathbf{M}^{1,p,q}(H)$, with the norm

$$\|Z\|_{\mathbf{M}^{1,p,q}(H)^*} = \sup_{Y\in\mathbf{M}^{1,p,q}(H)} \frac{\langle Y, Z\rangle_{L^2(\Omega,H)}}{\|Y\|_{\mathbf{M}^{1,p,q}(H)}}, \quad Z \in L^2(\Omega, H). \tag{3.7}$$

The Burkholder type inequality in Lemma 2.2 gives an estimate of the norm of a stochastic integral that is $L^2$ in time. We will now prove a similar inequality with respect to the $\mathbf{M}^{1,p,q}(H)^*$-norm, which is $L^{q'}$ in time, where $q'$ is the conjugate exponent to $q$ given by $\frac{1}{q} + \frac{1}{q'} = 1$ if $q < \infty$ and $q' = 1$ otherwise. Since $q \in [2, \infty]$, and hence $q' \in [1, 2]$, this admits worse singularities than in Lemma 2.2.

**Theorem 3.5.** *Let $p \in [2, \infty)$, $q \in [2, \infty]$ and $p'$, $q'$ denote the conjugate exponents. If $\Phi \in L^2([0, T] \times \Omega, \mathcal{L}_2^0)$ is predictable, then*

$$\Big\|\int_0^T \Phi(t)\,\mathrm{d}W(t)\Big\|_{\mathbf{M}^{1,p,q}(H)^*} \le \|\Phi\|_{L^{p'}(\Omega,L^{q'}([0,T],\mathcal{L}_2^0))}.$$

*Proof.* We use the fact that the stochastic integral of $\Phi$ equals $\delta\Phi$. By (3.7), (3.4), and Hölder's inequality, we get

$$\begin{aligned}\big\|\delta\Phi\big\|_{\mathbf{M}^{1,p,q}(H)^*} &= \sup_{Y\in\mathbf{M}^{1,p,q}(H)} \frac{\big\langle Y, \delta\Phi\big\rangle_{L^2(\Omega,H)}}{\|Y\|_{\mathbf{M}^{1,p,q}(H)}} = \sup_{Y\in\mathbf{M}^{1,p,q}(H)} \frac{\big\langle DY, \Phi\big\rangle_{L^2([0,T]\times\Omega,\mathcal{L}_2^0)}}{\|Y\|_{\mathbf{M}^{1,p,q}(H)}}\\ &\le \sup_{Y\in\mathbf{M}^{1,p,q}(H)} \frac{\|DY\|_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))}\big\|\Phi\big\|_{L^{p'}(\Omega,L^{q'}([0,T],\mathcal{L}_2^0))}}{\|Y\|_{\mathbf{M}^{1,p,q}(H)}}\\ &\le \|\Phi\|_{L^{p'}(\Omega,L^{q'}([0,T],\mathcal{L}_2^0))},\end{aligned}$$

which finishes the proof. □

**Remark 3.6.** Since the inequality in Lemma 2.2 is actually double-sided, one may ask whether this is true also for Theorem 3.5. In fact we can prove the reverse inequality for deterministic $\Phi \in L^2([0, T], \mathcal{L}_2^0)$. Since $\mathcal{H}_1^q(H) := \{\delta\Psi : \Psi \in$

$L^q([0,T],\mathcal{L}_2^0)\} \subset \mathbf{M}^{1,p,q}(H)$ we get an inequality in (3.7) by taking the supremum over $\mathcal{H}_1^q(H)$ instead of $\mathbf{M}^{1,p,q}(H)$:

$$\begin{aligned}\big\|\delta\Phi\big\|_{\mathbf{M}^{1,p,q}(H)^*} &= \sup_{Y\in\mathbf{M}^{1,p,q}(H)} \frac{\big\langle Y,\delta\Phi\big\rangle_{L^2([0,T]\times\Omega,\mathcal{L}_2^0)}}{\|Y\|_{\mathbf{M}^{1,p,q}(H)}} \\ &\geq \sup_{Y\in\mathcal{H}_1^q(H)} \frac{\big\langle DY,\Phi\big\rangle_{L^2([0,T]\times\Omega,\mathcal{L}_2^0)}}{\|Y\|_{\mathbf{M}^{1,p,q}(H)}} \\ &= \sup_{\Psi\in L^q([0,T],\mathcal{L}_2^0)} \frac{\big\langle D\delta\Psi,\Phi\big\rangle_{L^2([0,T]\times\Omega,\mathcal{L}_2^0)}}{\Big(\|\delta\Psi\|^p_{L^p(\Omega,H)} + \|D\delta\Psi\|^p_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))}\Big)^{\frac{1}{p}}}.\end{aligned}$$

We next use the fact that $D\delta\Psi = \Psi + \delta D\Psi = \Psi$ for deterministic $\Psi \in L^q([0,T],\mathcal{L}_2^0)$. By Burkholder's inequality Lemma 2.2 and Hölder's inequality we get

$$\begin{aligned}\big\|\delta\Phi\big\|_{\mathbf{M}^{1,p,q}(H)^*} &\geq \sup_{\Psi\in L^q([0,T],\mathcal{L}_2^0)} \frac{\big\langle \Psi,\Phi\big\rangle_{L^2([0,T],\mathcal{L}_2^0)}}{\Big(C_p^p\|\Psi\|^p_{L^2([0,T],\mathcal{L}_2^0)} + \|\Psi\|^p_{L^q([0,T],\mathcal{L}_2^0)}\Big)^{\frac{1}{p}}} \\ &\geq \frac{1}{\big(C_p^pT^{\frac{q}{q-2}}+1\big)^{\frac{1}{p}}} \sup_{\Psi\in L^q([0,T],\mathcal{L}_2^0)} \frac{\big\langle \Psi,\Phi\big\rangle_{L^2([0,T],\mathcal{L}_2^0)}}{\|\Psi\|_{L^q([0,T],\mathcal{L}_2^0)}} \\ &= \frac{1}{\big(C_p^pT^{\frac{q}{q-2}}+1\big)^{\frac{1}{p}}}\|\Phi\|_{L^{q'}([0,T],\mathcal{L}_2^0)}.\end{aligned}$$

The proof relies on the fact that $D\Psi = 0$. For random $\Phi$ one needs random $\Psi \in L^p(\Omega, L^q([0,T],\mathcal{L}_2^0))$ and, since $\delta D\Psi \neq 0$ in this case, this proof does not work.

**Remark 3.7.** One consequence of Theorem 3.5 is that the stochastic integral can be extended in $\mathbf{M}^{1,p,q}(H)^*$ to integrands in $L^{p'}(\Omega, L^{q'}([0,T],\mathcal{L}_2^0))$. The elements of $\mathbf{M}^{1,p,q}(H)^*$ are distributions defined by their action on random variables in $\mathbf{M}^{1,p,q}(H)$. One can show that the solution of the linear stochastic heat equation driven by space-time white noise in two space dimensions is a stochastic process $X \in \mathcal{C}([0,T],\mathbf{M}^{1,p,q}(H)^*)$ for every $p \geq 2$ and $q > 2$. In three space dimensions the same is valid for every $p \geq 2$ and $q > 4$. In higher space dimensions than three the solution is not $\mathbf{M}^{1,p,q}(H)^*$-valued since this would force $q' < 1$. Hölder continuity in time in the $\mathbf{M}^{1,p,q}(H)^*$-norms can be shown for the solution in two and three space dimensions for the $p,q$ for which the solution is defined. See Lemma 3.9 below for the regular case. Solutions defined in a distributional sense with respect to $\Omega$ is not a new concept. This is the heart of the white noise approach to SPDE, see, e.g., [4], [23].

Theorem 3.5 is a key result in the present work. But to be able to perform error estimates for semilinear equations we also need an intermediate space between $\mathbf{M}^{1,p,p}(H)$ and $\mathbf{M}^{1,2p,p}(H)$. For $2 \leq p < \infty$ we define

$$\mathbf{G}^{1,p}(H) = \mathbf{M}^{1,p,p}(H) \cap L^{2p}(\Omega, H).$$

It is a Banach space equipped with the norm

$$\|Y\|_{\mathbf{G}^{1,p}(H)} = \max\Big(\|Y\|_{\mathbf{M}^{1,p,p}(H)}, \|Y\|_{L^{2p}(\Omega,H)}\Big).$$

We have $\mathbf{M}^{1,2p,p}(H) \subset \mathbf{G}^{1,p}(H) \subset \mathbf{M}^{1,p,p}(H)$ and we obtain a new Gelfand triple

$$\mathbf{G}^{1,p}(H) \subset L^2(\Omega,H) \subset \mathbf{G}^{1,p}(H)^*.$$

The next lemma is a slightly modified version of Lemma 3.4, which is necessary to prove the local Lipschitz bound in Lemma 3.9.

**Lemma 3.8.** *Let $U_1, U_2, V$ be separable Hilbert spaces. For $\sigma \in \mathcal{C}^0_{\mathrm{b}}(U_1, \mathcal{L}(U_2,V)) \cap \mathcal{C}^1_{\mathrm{b}}(U_1, \mathcal{L}(U_2,V))$, $X \in \mathbf{M}^{1,2p,p}(U_1)$, $Y \in \mathbf{G}^{1,p}(U_2)$, $2 < p < \infty$, it holds $\sigma(X)\cdot Y \in \mathbf{G}^{1,p}(V)$. In addition, we have*

$$\begin{aligned}&\|\sigma(X)\cdot Y\|_{\mathbf{G}^{1,p}(V)}\\ &\quad\le \max\big(|\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}, |\sigma|_{\mathcal{C}^1_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\big)\Big(1+\|X\|_{\mathbf{M}^{1,2p,p}(U_1)}\Big)\|Y\|_{\mathbf{G}^{1,p}(U_2)}.\end{aligned}$$

*Proof.* It particularly holds $X \in \mathbf{M}^{1,p,p}(U_1)$, $Y \in \mathbf{M}^{1,p,p}(U_2)$, $p > 2$, and, hence, we directly obtain from Lemma 3.4 that $\sigma(X)\cdot Y \in \mathbf{M}^{1,\frac{p}{2},p}(V)$. In addition, we get

$$\|\sigma(X)\cdot Y\|_{L^{2p}(\Omega,V)} \le |\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\|Y\|_{L^{2p}(\Omega,U)} \le |\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\|Y\|_{\mathbf{G}^{1,p}(U)}.$$

Further, by (3.6) we have

$$\begin{aligned}&\big\|D(\sigma(X)\cdot Y)\big\|_{L^p(\Omega,L^p([0,T],\mathcal{L}_2(H_0,V)))}\\ &\quad= \big\|\sigma'(X)\cdot\big(DX,Y\big) + \sigma(X)\cdot DY\big\|_{L^p(\Omega,L^p([0,T],\mathcal{L}_2(H_0,V)))}\\ &\quad\le |\sigma|_{\mathcal{C}^1_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\Big(\mathbf{E}\Big[\|DX\|^p_{L^p([0,T],\mathcal{L}_2(H_0,U_1))}\,\|Y\|^p_{U_2}\Big]\Big)^{\frac{1}{p}}\\ &\qquad+|\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\|DY\|_{L^p(\Omega,L^p([0,T],\mathcal{L}_2(H_0,U_2)))}\\ &\quad\le |\sigma|_{\mathcal{C}^1_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\|DX\|_{L^{2p}(\Omega,L^p([0,T],\mathcal{L}_2(H_0,U_1)))}\|Y\|_{L^{2p}(\Omega,U_2)}\\ &\qquad+|\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\|DY\|_{L^p(\Omega,L^p([0,T],\mathcal{L}_2(H_0,U_2)))}\\ &\quad\le \max\big(|\sigma|_{\mathcal{C}^1_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}, \|\sigma\|_{\mathcal{C}^0_{\mathrm{b}}(U_1,\mathcal{L}(U_2,V))}\big)\Big(1+\|X\|_{\mathbf{M}^{1,2p,p}(U_1)}\Big)\|Y\|_{\mathbf{G}^{1,p}(U_2)}.\end{aligned}$$

These bounds show that $\sigma'(X)\cdot Y \in \mathbf{G}^{1,p}(V)$ as well as the desired bound. $\square$

Our next key result is stated in Lemma 3.9 below. It establishes a local Lipschitz bound in the $\mathbf{G}^{1,p}(H)^*$-norm. This allows us to perform a Gronwall argument in this norm in Section 4.2.

**Lemma 3.9.** *Let $U, V$ be separable Hilbert spaces, $\eta \in \mathcal{C}^2_{\mathrm{b}}(U,V)$, and $2 < p < \infty$. Then, for all $X_1, X_2 \in \mathbf{M}^{1,2p,p}(U)$,*

$$\begin{aligned}&\big\|\eta(X_1)-\eta(X_2)\big\|_{\mathbf{G}^{1,p}(V)^*}\\ &\quad\le \max\big(|\eta|_{\mathcal{C}^1_{\mathrm{b}}(U,V)}, |\eta|_{\mathcal{C}^2_{\mathrm{b}}(U,V)}\big)\Big(1+\sum_{i=1}^{2}\|X_i\|_{\mathbf{M}^{1,2p,p}(U)}\Big)\big\|X_1-X_2\big\|_{\mathbf{G}^{1,p}(U)^*}.\end{aligned}$$

*Proof.* In view of (2.2) it suffices to show

$$\begin{aligned}&\|\eta'(X)\cdot Y\|_{\mathbf{G}^{1,p}(V)^*}\\&\quad\le \max\left(|\eta|_{\mathcal{C}^1_{\mathrm{b}}(U,V)},|\eta|_{\mathcal{C}^2_{\mathrm{b}}(U,V)}\right)\Big(1+\|X\|_{\mathbf{M}^{1,2p,p}(U)}\Big)\|Y\|_{\mathbf{G}^{1,p}(U)^*},\end{aligned}$$

for all $X,Y\in\mathbf{M}^{1,2p,p}(U)$. We have

$$\big\|\eta'(X)\cdot Y\big\|_{\mathbf{G}^{1,p}(V)^*}\le\big\|\eta'(X)\big\|_{\mathcal{L}(\mathbf{G}^{1,p}(U)^*,\mathbf{G}^{1,p}(V)^*)}\big\|Y\big\|_{\mathbf{G}^{1,p}(U)^*}.$$

Since $\|\eta'(X)\|_{\mathcal{L}(\mathbf{G}^{1,p}(U)^*,\mathbf{G}^{1,p}(V)^*)}=\|\eta'(X)^*\|_{\mathcal{L}(\mathbf{G}^{1,p}(V),\mathbf{G}^{1,p}(U))}$, it suffices to give a bound of the latter term. For this we define $\sigma\colon U\to\mathcal{L}(V,U)$ by

$$\sigma(x):=\eta'(x)^*.$$

Then $\sigma\in\mathcal{C}^0_{\mathrm{b}}(U,\mathcal{L}(V,U))\cap\mathcal{C}^1_{\mathrm{b}}(U,\mathcal{L}(V,U))$ with $|\sigma|_{\mathcal{C}^0_{\mathrm{b}}(U,\mathcal{L}(V,U))}=|\eta|_{\mathcal{C}^1_{\mathrm{b}}(U,V)}$ and $|\sigma|_{\mathcal{C}^1_{\mathrm{b}}(U,\mathcal{L}(V,U))}=|\eta|_{\mathcal{C}^2_{\mathrm{b}}(U,V)}$. Hence, the assertion follows directly from an application of Lemma 3.8. □

3.3. **Regularity of the solution.** Here we prove regularity in terms of the Malliavin derivative, as well as Hölder continuity in the $\mathbf{M}^{1,p,q}(H)^*$-norm, of the solution $X$ to (2.5) under Assumption 2.3. For suitably chosen $p$ and $q$ the Hölder exponent turns out to be twice as high as in the $L^2(\Omega,H)$-norm. By combining these results with a duality argument we show Hölder continuity of the Markov semigroup. The Hölder exponent is later, in Theorem 4.4, shown to coincide with the rate of weak convergence, which is natural.

The Malliavin derivative $D_rX(t)$ of $X(t)$ at time $r\in[0,T]$ satisfies the equation, see [18, Proposition 3.5 (ii)],

$$D_rX(t)=\begin{cases}S(t-r)+\displaystyle\int_r^t S(t-s)F'(X(s))D_rX(s)\,\mathrm{d}s, & t\in(r,T],\\ 0, & t\in[0,r].\end{cases}\tag{3.8}$$

The next result can be verified by using (3.11) of [18, Proposition 3.5 (ii)] and holds for multiplicative noise, as well. For completeness we present a proof in the simpler case of additive noise that we consider here.

**Proposition 3.10.** *Let Assumption 2.3 hold and let $X$ be the solution of (2.5). If $\beta\in(0,1)$, then*

$$\sup_{t\in[0,T]}\big\|X(t)\big\|_{\mathbf{M}^{1,p,q}(H)}<\infty,$$

*for $2\le p<\infty$ and $2\le q<\frac{2}{1-\beta}$. If $\beta=1$, then the same holds for $2\le p<\infty$ and $2\le q\le\infty$.*

*Proof.* We remark that the case $p=q=2$ was already proved in [18]. The moment estimate (2.6) implies that $\sup_{t\in[0,T]}\|X(t)\|_{L^p(\Omega,H)}<\infty$ for $2\le p<\infty$. Next we take norms in (3.8) and use Minkowski's inequality on the convolution term.

We note that $D_r X(s) = 0$ for $s \leq r$ because $X(s)$ is $\mathcal{F}_r$-measurable, so that the convolution term can be written $\int_0^t \dots \, \mathrm{d}s$. We get

$$\begin{aligned}
\big\|DX(t)\big\|_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))} &= \big\|DX(t)\big\|_{L^p(\Omega,L^q([0,t],\mathcal{L}_2^0))} \\
&\leq \big\|S(t-\cdot)\big\|_{L^q([0,t],\mathcal{L}_2^0)} + \Big\| \int_0^t S(t-s)F'(X(s))DX(s)\,\mathrm{d}s \Big\|_{L^p(\Omega,L^q([0,t],\mathcal{L}_2^0))} \\
&\leq \big\|S\big\|_{L^q([0,t],\mathcal{L}_2^0)} + \int_0^t \big\|S(t-s)F'(X(s))DX(s)\big\|_{L^p(\Omega,L^q([0,t],\mathcal{L}_2^0))} \,\mathrm{d}s \\
&\leq \big\|S\big\|_{L^q([0,T],\mathcal{L}_2^0)} + \|S\|_{L^\infty([0,T],\mathcal{L})} |F|_{\mathcal{C}_\mathrm{b}^1} \int_0^t \big\|DX(s)\big\|_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))} \,\mathrm{d}s.
\end{aligned}$$

We conclude by using Lemma 2.5 and the standard Gronwall lemma. $\square$

We next consider Hölder continutity in the $\mathbf{M}^{1,p,q}(H)^*$-norm. For comparison we recall that the Hölder exponent in the $L^2(\Omega, H)$-norm is $\gamma < \beta/2$ under Assumption 2.3. Here we have $\gamma < \beta$, if $q$ is sufficiently large.

**Proposition 3.11.** *Let Assumption 2.3 hold with $\beta \in (0,1]$ and denote by $X$ the solution to (2.5). Let $2 \leq p < \infty$, $\gamma \in [0, \beta)$, and set $q = \frac{2}{1-\gamma}$. Then there exists a constant $C = C_\gamma$ such that*

$$\big\|X(t_2) - X(t_1)\big\|_{\mathbf{M}^{1,p,q}(H)^*} \leq C\Big(1 + \big\|X_0\big\|_{\dot{H}^{2\beta}}\Big)\big|t_2 - t_1\big|^\gamma, \quad t_1, t_2 \in [0,T].$$

*Proof.* Without loss of generality we assume $t_2 > t_1 > 0$. From (2.5) we then get

$$\begin{aligned}
X(t_2) - X(t_1) = {}& \big(S(t_2 - t_1) - I\big)S(t_1)X_0 \\
&+ \big(S(t_2 - t_1) - I\big)\int_0^{t_1} S(t_1 - s)F(X(s))\,\mathrm{d}s \\
&+ \big(S(t_2 - t_1) - I\big)\int_0^{t_1} S(t_1 - s)\,\mathrm{d}W(s) \\
&+ \int_{t_1}^{t_2} S(t_2 - s)F(X(s))\,\mathrm{d}s + \int_{t_1}^{t_2} S(t_2 - s)\,\mathrm{d}W(s).
\end{aligned}$$

In the following we study the $\mathbf{M}^{1,p,q}(H)^*$-norms of these five summands. For the first, second, and fourth terms we use the fact that $\|Z\|_{\mathbf{M}^{1,p,q}(H)^*} \leq \|Z\|_{L^2(\Omega,H)}$.

For the first summand, we use (2.4) with $\varrho = \gamma$ and (2.3) with $\varrho = 0$ as well as Assumption 2.3 (ii). This yields

$$\begin{aligned}
\big\|\big(S(t_2 - t_1) - I\big)S(t_1)X_0\big\|_{\mathbf{M}^{1,p,q}(H)^*} &\leq \big\|\big(S(t_2 - t_1) - I\big)A^{-\gamma}S(t_1)A^\gamma X_0\big\|_{L^2(\Omega,H)} \\
&\lesssim \big|t_2 - t_1\big|^\gamma \|A^\gamma X_0\| \lesssim \big|t_2 - t_1\big|^\gamma \|X_0\|_{\dot{H}^{2\beta}}.
\end{aligned}$$

The estimate of the second summand is done by applying Assumption 2.3 (iv) and the same arguments as for the first term. More precisely, we use that $F \in$

$\mathcal{C}^1_{\mathrm{b}}(H,H)$ implies linear growth, to get

$$\begin{aligned}
&\Big\|\big(S(t_2-t_1)-I\big)\int_0^{t_1} S(t_1-s)F(X(s))\,\mathrm{d}s\Big\|_{\mathbf{M}^{1,p,q}(H)^*}\\
&\quad\le \big\|\big(S(t_2-t_1)-I\big)A^{-\gamma}\big\|_{\mathcal{L}}\int_0^{t_1}\big\|A^{\gamma}S(t_1-s)\big\|_{\mathcal{L}}\big\|F(X(s))\big\|_{L^2(\Omega,H)}\,\mathrm{d}s\\
&\quad\lesssim \big|t_2-t_1\big|^{\gamma}\int_0^{t_1}(t_1-s)^{-\gamma}\,\mathrm{d}s\Big(1+\sup_{s\in[0,T]}\|X(s)\|_{L^2(\Omega,H)}\Big)\lesssim \big|t_2-t_1\big|^{\gamma},
\end{aligned}$$

where we also used (2.6) and that $\gamma<\beta\le 1$.

We now turn to the third term. We recall that $q=2/(1-\gamma)$ and $q'=2/(1+\gamma)$. Since $\gamma<\beta$, we have

$$q'\frac{2\gamma+1-\beta}{2}=\frac{2\gamma+1-\beta}{1+\gamma}=1-\frac{\beta-\gamma}{1+\gamma}<1. \tag{3.9}$$

We apply Theorem 3.5 to the third summand. Then by (2.3), (2.4), Assumption 2.3 (iii), and (3.9), we obtain

$$\begin{aligned}
&\Big\|\big(S(t_2-t_1)-I\big)\int_0^{t_1} S(t_1-s)\,\mathrm{d}W(s)\Big\|_{\mathbf{M}^{1,p,q}(H)^*}\\
&\quad\le \big\|\big(S(t_2-t_1)-I\big)S(t_1-\cdot)\big\|_{L^{p'}(\Omega,L^{q'}([0,t_1],\mathcal{L}_2^0))}\\
&\quad\le \big\|\big(S(t_2-t_1)-I\big)A^{-\gamma}\big\|_{\mathcal{L}}\Big(\int_0^{t_1}\big\|A^{\gamma}A^{\frac{1-\beta}{2}}S(t_1-s)A^{\frac{\beta-1}{2}}\big\|^{q'}_{\mathcal{L}_2^0}\,\mathrm{d}s\Big)^{\frac{1}{q'}}\\
&\quad\lesssim \big|t_2-t_1\big|^{\gamma}\Big(\int_0^{t_1}(t_1-s)^{-q'\frac{2\gamma+1-\beta}{2}}\,\mathrm{d}s\,\big\|A^{\frac{\beta-1}{2}}\big\|^{q'}_{\mathcal{L}_2^0}\Big)^{\frac{1}{q'}}\lesssim \big|t_2-t_1\big|^{\gamma}.
\end{aligned}$$

Next we turn to the fourth term. By applying the same arguments as for the second summand, we derive the bound

$$\begin{aligned}
\Big\|\int_{t_1}^{t_2} S(t_2-s)F(X(s))\,\mathrm{d}s\Big\|_{\mathbf{M}^{1,p,q}(H)^*}&\le\int_{t_1}^{t_2}\big\|S(t_2-s)F(X(s))\big\|_{L^2(\Omega,H)}\,\mathrm{d}s\\
&\lesssim |t_2-t_1|\Big(1+\sup_{s\in[0,T]}\big\|X(s)\big\|_{L^2(\Omega,H)}\Big).
\end{aligned}$$

Finally, a further application of Theorem 3.5 and (2.3) with $\varrho=\frac{1-\beta}{2}$ yields for the fifth summand

$$\begin{aligned}
\Big\|\int_{t_1}^{t_2} S(t_2-s)\,\mathrm{d}W(s)\Big\|_{\mathbf{M}^{1,p,q}(H)^*}&\le\Big(\int_{t_1}^{t_2}\big\|S(t_2-s)A^{\frac{1-\beta}{2}}\big\|^{q'}_{\mathcal{L}}\big\|A^{\frac{\beta-1}{2}}\big\|^{q'}_{\mathcal{L}_2^0}\,\mathrm{d}s\Big)^{\frac{1}{q'}}\\
&\lesssim\Big(\int_{t_1}^{t_2}(t_2-s)^{-\frac{1-\beta}{2}}\,\mathrm{d}s\Big)^{\frac{1}{q'}}\lesssim|t_2-t_1|^{\frac{1}{q'}-\frac{1-\beta}{2}}.
\end{aligned}$$

By inserting $q'=2/(1+\gamma)$ and $\beta>\gamma$, we see that the exponent is

$$\frac{1}{q'}-\frac{1-\beta}{2}=\frac{1+\gamma}{2}-\frac{1-\beta}{2}=\frac{\gamma+\beta}{2}>\gamma.$$

This completes the proof. □

As a consequence of Propositions 3.10 and 3.11 we now show Hölder continuity of the Markov semigroup $(P(t))_{t\in[0,T]}$ related to $X$. This will not be used in the sequel but it is a neat application of the duality argument. Define for $(t,x)\in[0,T]\times H$, $(P(t)\varphi)(x)=\mathbf{E}[\varphi(X(t,x))]$, where $X(t,x)$ denotes the solution to equation (2.5) with initial value $X_0=x\in\dot H^{2\beta}$.

**Corollary 3.12.** *Let Assumption 2.3 hold with $\beta\in(0,1]$ and let $\varphi\in\mathcal{C}^2_{\mathrm{p}}(H,\mathbf{R})$. For every $\gamma\in[0,\beta)$ there is a constant $C$ such that*

$$\big|(P(t_2)\varphi)(x)-(P(t_1)\varphi)(x)\big|\le C\big(1+\|x\|_{\dot H^{2\beta}}\big)\big|t_2-t_1\big|^{\gamma},\quad t_1,t_2\in[0,T],\ x\in\dot H^{2\beta}.$$

*Proof.* We fix $x$ and suppress it from the notation. Applying (2.2) yields

$$\begin{aligned}\big|(P(t_2)\varphi)(x)-(P(t_1)\varphi)(x)\big|&=\big|\mathbf{E}\big[\varphi(X(t_2))-\varphi(X(t_1))\big]\big|\\&=\Big|\Big\langle\int_0^1\varphi'\big(\varrho X(t_2)+(1-\varrho)X(t_1)\big)\,\mathrm{d}\varrho,X(t_2)-X(t_1)\Big\rangle_{L^2(\Omega,H)}\Big|.\end{aligned}$$

For arbitrary $p\in[2,\infty)$ we obtain by duality

$$\begin{aligned}&\big|(P(t_2)\varphi)(x)-(P(t_1)\varphi)(x)\big|\\&\quad\le\Big\|\int_0^1\varphi'\big(\varrho X(t_2)+(1-\varrho)X(t_1)\big)\,\mathrm{d}\varrho\Big\|_{\mathbf{M}^{1,p,p}(H)}\big\|X(t_2)-X(t_1)\big\|_{\mathbf{M}^{1,p,p}(H)^*}.\end{aligned}$$

Now take $p=\frac{2}{1-\gamma}$. The first factor is finite by Proposition 3.10 and the chain rule; for details see the proof of Lemma 4.2 below. Proposition 3.11 applies to the second factor and this completes the proof. □

**Remark 3.13.** Proposition 3.11 can be proved without additional difficulties in the case of multiplicative noise and so can Proposition 3.10, due to the comment right before its statement. Therefore, Corollary 3.12 holds for multiplicative noise.

**Remark 3.14.** We end this section with a comment on implications to stochastic ordinary differential equations. This corresponds to the case $A=0$, $\beta=1$, and multiplicative noise with diffusion coefficient $G\in\mathcal{C}^2_{\mathrm{b}}(H,\mathcal{L}^0_2)$, i.e., we consider the equation

$$\mathrm{d}X(t)=F(X(t))\,\mathrm{d}t+G(X(t))\,\mathrm{d}W(t),\ t\in(0,T];\quad X(0)=X_0. \tag{3.10}$$

In this case one can prove Proposition 3.11 with $p\ge2$, $q=\infty$, and $\gamma=1$, meaning that the solution is Lipschitz continuous in time in the $\mathbf{M}^{1,p,\infty}(H)^*$-norm for every $p\ge2$. For $\beta=1$ the covariance operator $Q$ is of trace class and the cylindrical Wiener process $W$ is well defined as an $H$-valued Brownian motion. We see that also $W$ is Lipschitz continuous in $\mathbf{M}^{1,p,\infty}(H)^*$ by Proposition 3.5. Indeed,

$$\begin{aligned}\big\|W(t_2)-W(t_1)\big\|_{\mathbf{M}^{1,p,\infty}(H)^*}&=\Big\|\int_{t_1}^{t_2}\mathrm{d}W(t)\Big\|_{\mathbf{M}^{1,p,\infty}(H)^*}\\&\le\big\|\chi_{[t_1,t_2]}\big\|_{L^{p'}(\Omega,L^1([0,T],\mathcal{L}^0_2))}=\mathrm{Tr}(Q)\big|t_2-t_1\big|,\quad t_1,t_2\in[0,T].\end{aligned}$$

This suggests that $\mathrm{d}X(t)=\dot X(t)\,\mathrm{d}t$ and $\mathrm{d}W(t)=\dot W(t)\,\mathrm{d}t$, where $\dot X$ and $\dot W$ are $\mathbf{M}^{1,p,\infty}(H)^*$-valued functions on $[0,T]$. This further suggests that (3.10) might be written in the form

$$\dot{X}(t) = F(X(t)) + G(X(t))\dot{W}(t).$$

If this formulation is useful or fully makes sense is an open question. There seems to be a connection to the functional white noise approach of stochastic differential equations, see [37], that remains to be understood. In this approach the time derivative of Brownian motion is well defined in the space of Hida distributions and the corresponding product of $G$ and $\dot{W}$ is the Wick product.

3.4. **Regularity of the numerical solution.** Here we first show a bound on the $p$:th-moment of the discrete solutions $X_{h,k}$ to (2.7), uniformly in $h,k \in (0,1]$, and then we prove a discrete analog of Proposition 3.10.

**Proposition 3.15.** *Let Assumptions 2.3 and 2.4 hold with $\beta \in (0,1]$ and let $2 \leq p < \infty$. Then*

$$\max_{n\in\{0,\dots,N\}} \sup_{h,k\in(0,1]} \big\|X^n_{h,k}\big\|_{L^p(\Omega,H)} \leq C.$$

*Proof.* For $n \in \{1,\dots,N\}$ we recall the representation (2.8) of $X^n_{h,k}$. Hence, it follows that

$$\begin{aligned}\big\|X^n_{h,k}\big\|_{L^p(\Omega,H)} &\leq \big\|S^n_{h,k}P_hX_0\big\| + k\sum_{j=0}^{n-1}\big\|S^{n-j}_{h,k}F(X^j_{h,k})\big\|_{L^p(\Omega,H)}\\ &\quad + \Big\|\int_0^T \Big(\sum_{j=0}^{n-1}\chi_{[t_j,t_{j+1})}(t)S^{n-j}_{h,k}\Big)\,\mathrm{d}W(t)\Big\|_{L^p(\Omega,H)}.\end{aligned}$$

By (2.10) with $\varrho = 0$ we have

$$\sup_{n\in\{1,\dots,N\}} \big\|S^n_{h,k}\big\|_{\mathcal{L}} \lesssim 1, \tag{3.11}$$

so that $\|S^n_{h,k}P_hX_0\| \lesssim 1$. Therefore, by applying also Lemma 2.2,

$$\big\|X^n_{h,k}\big\|_{L^p(\Omega,H)} \lesssim 1 + k\sum_{j=0}^{n-1}\big\|F(X^j_{h,k})\big\|_{L^p(\Omega,H)} + \Big\|\sum_{j=0}^{n-1}\chi_{[t_j,t_{j+1})}S^{n-j}_{h,k}\Big\|_{L^2([0,T],\mathcal{L}_2^0)}.$$

By referring to Lemma 2.5 with $q=2$, we have

$$\Big\|\sum_{j=0}^{n-1}\chi_{[t_j,t_{j+1})}S^{n-j}_{h,k}\Big\|^2_{L^2([0,T],\mathcal{L}_2^0)} = k\sum_{j=0}^{n-1}\big\|S^{n-j}_{h,k}\big\|^2_{\mathcal{L}_2^0} \leq k\sum_{j=1}^{N}\big\|S^j_{h,k}\big\|^2_{\mathcal{L}_2^0} \lesssim 1.$$

Further, since the drift $F\colon H\to H$ satisfies a linear growth bound under Assumption 2.3 (iv), it follows that

$$\big\|X^n_{h,k}\big\|_{L^p(\Omega,H)} \lesssim 1 + k\sum_{j=0}^{n-1}\big\|X^j_{h,k}\big\|_{L^p(\Omega,H)}$$

and the proof is completed by an application of Gronwall's Lemma 2.1. $\square$

**Proposition 3.16.** *Let Assumptions 2.3 and 2.4 hold with $\beta \in (0,1]$. If $\beta \in (0,1)$, then*

$$\max_{n\in\{1,\dots,N\}} \sup_{h,k\in(0,1]} \big\|X_{h,k}^n\big\|_{\mathbf{M}^{1,p,q}(H)} < \infty,$$

*for $2 \le p < \infty$ and $2 \le q < \frac{2}{1-\beta}$. If $\beta = 1$, then the same holds for $2 \le p < \infty$ and $2 \le q \le \infty$.*

*Proof.* We mimic the proof of Proposition 3.10. The $L^p(\Omega,H)$-norm of $X_{h,k}$ is treated in Proposition 3.15 and it remains to bound $DX_{h,k}$.

By using the chain rule (3.5) and $D_r \int_{t_j}^{t_{j+1}} S_{h,k}^{n-j}\,\mathrm{d}W(s) = \chi_{[t_j,t_{j+1})}(r) S_{h,k}^{n-j}$, we apply the Malliavin derivative termwise to equation (2.8) and obtain

$$D_r X_{h,k}^n = k\sum_{j=0}^{n-1} S_{h,k}^{n-j} F'(X_{h,k}^j) D_r X_{h,k}^j + \sum_{j=0}^{n-1} \chi_{[t_j,t_{j+1})}(r) S_{h,k}^{n-j}. \tag{3.12}$$

Here we note that $D_r X_{h,k}^j = 0$ for $t_j \le r$, since $X_{h,k}^j$ is $\mathcal{F}_r$-measurable. Therefore,

$$D_r X_{h,k}^n = \sum_{i=0}^{n-1} \chi_{[t_i,t_{i+1})}(r)\Big(k\sum_{j=i+1}^{n-1} S_{h,k}^{n-j} F'(X_{h,k}^j) D_r X_{h,k}^j + S_{h,k}^{n-i}\Big)$$

in full analogy with (3.8). However, as in the proof of Proposition 3.10, it is more convenient to take norms in (3.12) and use Minkowski's inequality on the convolution term:

$$\begin{aligned}
&\big\|DX_{h,k}^n\big\|_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))} = \big\|DX_{h,k}^n\big\|_{L^p(\Omega,L^q([0,t_n],\mathcal{L}_2^0))} \\
&\quad\le \Big\|\sum_{j=0}^{n-1} \chi_{[t_j,t_{j+1})} S_{h,k}^{n-j}\Big\|_{L^q([0,t_n],\mathcal{L}_2^0)} \\
&\qquad + \Big\|k\sum_{j=0}^{n-1} S_{h,k}^{n-j} F'(X_{h,k}^j) D_r X_{h,k}^j\Big\|_{L^p(\Omega,L^q([0,t_n],\mathcal{L}_2^0))} \\
&\quad\le \Big(k\sum_{j=1}^{N} \|S_{h,k}^j\|_{\mathcal{L}_2^0}^q\Big)^{1/q} + \sup_{1\le j\le N} \big\|S_{h,k}^j\big\|_{\mathcal{L}} |F|_{\mathcal{C}_{\mathrm{b}}^1} k\sum_{j=0}^{n-1} \big\|DX_{h,k}^j\big\|_{L^p(\Omega,L^q([0,T],\mathcal{L}_2^0))}.
\end{aligned}$$

We conclude by using Lemma 2.5, (3.11), and the discrete Gronwall Lemma 2.1. □

## 4. Weak convergence by duality

Let $X$ be the solution to equation (2.5) and $X_{h,k}$ be the discretization given by the semi-implicit scheme (2.7) and take $\varphi \in \mathcal{C}^1(H,\mathbf{R})$. Our approach to weak convergence begins with an application of (2.2) to get

$$\mathbf{E}\big[\varphi(X(t_n)) - \varphi(X_{h,k}^n)\big] = \big\langle \Phi_{h,k}^n, X(t_n) - X_{h,k}^n\big\rangle_{L^2(\Omega,H)},$$

where

$$\Phi_{h,k}^n = \int_0^1 \varphi'(\Theta_{h,k}^n(\varrho))\,\mathrm{d}\varrho \quad\text{and}\quad \Theta_{h,k}^n(\varrho) = \varrho X(t_n) + (1-\varrho)X_{h,k}^n, \tag{4.1}$$

for $n \in \{1, \dots, N\}$. This linearization was first proposed in [10] for nonlinear stochastic ordinary differential equations. They proceed by a duality argument based on an adjoint equation.

This linearization was independently used in [31] for linear stochastic partial differential equations. Extending the idea of [31], we proceed as follows: choose a Gelfand triple $V \subset L^2(\Omega, H) \subset V^*$ such that $\Phi^n_{h,k} \in V$. By duality we have

$$\big|\mathbf{E}\big[\varphi(X(t_n)) - \varphi(X^n_{h,k})\big]\big| \le \Big(\sup_{h,k\in(0,1]} \big\|\Phi^n_{h,k}\big\|_V\Big)\big\|X(t_n) - X^n_{h,k}\big\|_{V^*}. \tag{4.2}$$

The proof of our weak convergence result in Theorem 4.4 then amounts to showing that we can find a suitable space $V$ such that, for $\gamma \in (0, \beta)$,

$$\begin{aligned} \max_{n\in\{1,\dots,N\}} \sup_{h,k\in(0,1]} \big\|\Phi^n_{h,k}\big\|_V &\le C, \\ \max_{n\in\{1,\dots,N\}} \big\|X(t_n) - X^n_{h,k}\big\|_{V^*} &\le C\big(h^{2\gamma} + k^{\gamma}\big), \quad h, k \in (0, 1]. \end{aligned} \tag{4.3}$$

In comparison, the strong error converges with half this rate, i.e., for $\gamma \in (0, \beta)$ there exists $C$ such that

$$\max_{n\in\{1,\dots,N\}} \|X(t_n) - X^n_{h,k}\|_{L^2(\Omega,H)} \le C(h^{\gamma} + k^{\frac{\gamma}{2}}), \quad h, k \in (0, 1].$$

In Corollary 4.7 we deduce this from (4.3) by an interpolation argument.

We explain our method by gradually choosing more sophisticated spaces $V$. We begin in the next subsection with the simpler problem of the weak approximation of the stochastic convolution. This problem is treated in [16], [19] [28], [29], [31], and to some extent in [48]. We show that in this case $V = L^2(\Omega, \dot H^{\gamma})$ and $V = \mathbf{M}^{1,p,p}(H)$ with $p = \frac{2}{1-\gamma}$ suffice with different degrees of success. The proofs are simpler than in the mentioned papers, except for [31] to which the present paper is an extension. We continue with a subsection containing our main result Theorem 4.4, which is concerned with semilinear equations with additive noise. Here we use the space $V = \mathbf{G}^{1,p}(H)$, whose dual norm allows for a Gronwall argument based on Lemma 3.9. Finally, we discuss multiplicative noise in Subsection 4.3 and illustrate why our approach is not yet sufficient for this generality.

We assume that test functions are taken from $\mathcal{C}^2_{\mathrm{p}}(H, \mathbf{R})$ with a precise definition in the following assumption. Recall the norm defined in (2.1).

**Assumption 4.1.** *The test function $\varphi \in \mathcal{C}^2_{\mathrm{p}}(H, \mathbf{R})$, i.e., $\varphi\colon H \to \mathbf{R}$ is twice continuously Fréchet differentiable and there exists an integer $m \ge 2$ and a constant $C$ such that*

$$\|\varphi^{(j)}(x)\|_{\mathcal{L}^{[j]}(H,\mathbf{R})} \le C\big(1 + \|x\|^{m-j}\big), \quad x \in H, \ j = 1, 2.$$

4.1. **The stochastic convolution.** We consider the stochastic convolution $W^A$ and its approximation $W^{A_h}_{h,k}$,

$$W^A(t_n) = \int_0^{t_n} S(t_n - s)\,\mathrm{d}W(s) \quad \text{and} \quad W^{A_h,n}_{h,k} = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S^{n-j}_{h,k}\,\mathrm{d}W(s)$$

for $n \in \{1,\dots,N\}$. For $\gamma \in (0,\beta)$, we consider first the Gelfand triple

$$L^2\big(\Omega, \dot H^{\gamma}\big) \subset L^2(\Omega, H) \subset L^2\big(\Omega, \dot H^{-\gamma}\big).$$

In order to have $\Phi^n_{h,k} \in L^2(\Omega, \dot H^{\gamma})$ we need to impose an extra assumption on $\varphi$, namely that, for some $m \geq 1$ and every $\gamma \in (0,\beta)$, it holds

$$\big\|\varphi'(x)\big\|_{\dot H^{\gamma}} \leq C\Big(1 + \|x\|^{m-1}_{\dot H^{\gamma}}\Big), \quad x \in \dot H^{\gamma}. \tag{4.4}$$

Then we first get by the Sobolev regularity of $W^A$ and $W^{A_h}_{h,k}$

$$\big\|\Phi^n_{h,k}\big\|_{L^2(\Omega,\dot H^{\gamma})} \lesssim \big\|W^A(t_n)\big\|^{m-1}_{L^{2(m-1)}(\Omega,\dot H^{\gamma})} + \big\|W^{A_h,n}_{h,k}\big\|^{m-1}_{L^{2(m-1)}(\Omega,\dot H^{\gamma})} \lesssim 1,$$

uniformly in $h,k \in (0,1]$. To prove convergence in $L^2(\Omega, \dot H^{-\gamma})$ we write the difference of the stochastic convolution and its numerical discretization in the form

$$W^A(t_n) - W^{A_h,n}_{h,k} = \int_0^{t_n} \tilde E_{h,k}(t_n - t)\,\mathrm{d}W(t), \tag{4.5}$$

where $\tilde E_{h,k} \colon (0,T) \to \mathcal{L}^0_2$ is given by

$$\tilde E_{h,k}(t) := S(t) - S^{j+1}_{h,k}, \qquad \text{for } t \in (t_j, t_{j+1}),\ j = 0,\dots,N-1. \tag{4.6}$$

Provided that this error operator satisfies

$$\big\|A^{-\frac{\gamma}{2}}\tilde E_{h,k}(t) A^{\frac{1-\beta}{2}}\big\|_{\mathcal{L}} \lesssim \big(h^{2\gamma} + k^{\gamma}\big)t^{\frac{-1+\beta-\gamma}{2}}, \quad t > 0, \tag{4.7}$$

then we obtain by the Itō isometry and Assumption 2.3 (iii)

$$\begin{aligned}
\big\|W^A(t_n) - W^{A_h,n}_{h,k}\big\|_{L^2(\Omega,\dot H^{-\gamma})} &= \Big(\int_0^{t_n} \big\|A^{-\frac{\gamma}{2}}\tilde E_{h,k}(t_n - t)\big\|^2_{\mathcal{L}^0_2}\,\mathrm{d}t\Big)^{\frac12}\\
&\leq \Big(\int_0^{t_n} \big\|A^{-\frac{\gamma}{2}}\tilde E_{h,k}(t_n - t)A^{\frac{1-\beta}{2}}\big\|^2_{\mathcal{L}}\,\big\|A^{\frac{\beta-1}{2}}\big\|^2_{\mathcal{L}^0_2}\,\mathrm{d}t\Big)^{\frac12}\\
&\lesssim \big(h^{2\gamma} + k^{\gamma}\big)\Big(\int_0^{t_n} (t_n - t)^{-1+\beta-\gamma}\,\mathrm{d}t\Big)^{\frac12} \lesssim h^{2\gamma} + k^{\gamma}.
\end{aligned}$$

The error estimate (4.7) is verified for Galerkin finite element approximations in Section 5 for $\gamma = 0$, see Lemma 5.1 with $\theta = \gamma$, $\varrho = 1-\beta$, but the case $\gamma > 0$ is not to be found in the literature, so for this particular choice of Gelfand triple we do not work in full rigor. An integrated version of (4.7) is found in [48], details in [47], and we find no reason to doubt the validity of (4.7). In view of (4.2), by assuming (4.4) and (4.7), we can prove weak convergence with the desired rate.

Actually, [48, Theorem 1.2] shows convergence of order $O(h^{2\beta}+k^{\beta})$ in $L^2(\Omega, \dot H^{-1})$ (except for a logarithmic factor). However, the fact that $L^2(\Omega, \dot H^{-1})$-convergence implies weak convergence for other than linear test functions was not realized in the early work [48]. Subsequent works except [31] rely on the use of Kolmogorov's equation. In the paper [19] this was done for test functions satisfying (4.4), while [16] only assumed $\varphi \in \mathcal{C}^2_{\mathrm{b}}(H,\mathbf{R})$. We also remark that the only technical ingredient used in the present proof is the Itō isometry. Therefore this proof carries over without additional difficulties to the case when the cylindrical $Q$-Wiener process $W$ is replaced by a square integrable martingale $M$, by just introducing the suitable notation. This gives a partial extension of the results in [35], in which impulsive noise was considered. In that paper the additional assumption (4.4) was not used but instead the test functions were assumed to be in $\mathcal{C}^2_{\mathrm{b}}(H,\mathbf{R})$.

Fix $\gamma\in(0,\beta)$ and let $p=\frac{2}{1-\gamma}$. We next consider the Gelfand triple

$$\mathbf{M}^{1,p,p}(H)\subset L^2(\Omega,H)\subset \mathbf{M}^{1,p,p}(H)^*.$$

With these spaces we need no assumption on the test function other than Assumption 4.1. We state the two parts of (4.3) as two separate lemmas. Notice that the first lemma is not restricted to the stochastic convolution.

**Lemma 4.2.** *Let Assumptions 2.3, 2.4, and 4.1 hold with $\beta\in(0,1]$. For $\gamma\in(0,\beta)$, set $p=\frac{2}{1-\gamma}$. Then it holds*

$$\max_{n\in\{1,\dots,N\}}\sup_{h,k\in(0,1]}\big\|\Phi^n_{h,k}\big\|_{\mathbf{M}^{1,p,p}(H)}<\infty,$$

*where $\Phi^n_{h,k}$ is defined in (4.1).*

*Proof.* First note that $\varphi'$ satisfies the condition of the chain rule in Lemma 3.3 with $r=m-2$ by Assumption 4.1. Thus, it holds

$$\Phi^n_{h,k}=\int_0^1\varphi'(\Theta^n_{h,k}(\varrho))\,\mathrm{d}\varrho\in\mathbf{M}^{1,p,p}(H),$$

since $\Theta^n_{h,k}(\varrho)=\varrho X(t_n)+(1-\varrho)X^n_{h,k}\in\mathbf{M}^{1,(m-1)p,p}(H)$ by Propositions 3.10 and 3.16. Further, from Lemma 3.3 we also get

$$\begin{aligned}\big\|\Phi^n_{h,k}\big\|_{\mathbf{M}^{1,p,p}(H)}&\lesssim\big(1+\sup_{\varrho\in[0,1]}\big\|\Theta^n_{h,k}\big\|^{m-1}_{\mathbf{M}^{1,(m-1)p,p}(H)}\big)\\&\lesssim\big(1+\big\|X(t_n)\big\|^{m-1}_{\mathbf{M}^{1,(m-1)p,p}(H)}+\big\|X^n_{h,k}\big\|^{m-1}_{\mathbf{M}^{1,(m-1)p,p}(H)}\big).\end{aligned}$$

By Propositions 3.10 and 3.16, these are bounded independently of $h,k\in(0,1]$. □

**Lemma 4.3.** *Let Assumptions 2.3 and 2.4 hold with $\beta\in(0,1]$. For $\gamma\in(0,\beta)$, set $p=\frac{2}{1-\gamma}$. It holds*

$$\max_{n\in\{1,\dots,N\}}\big\|W^A(t_n)-W^{A_h,n}_{h,k}\big\|_{\mathbf{M}^{1,p,p}(H)^*}\le C\big(h^{2\gamma}+k^\gamma\big),\quad h,k\in(0,1].$$

*Proof.* By (4.5), Theorem 3.5, and Assumption 2.3 (iii), we get

$$\begin{aligned}\big\|W^A(t_n)-W^{A_h,n}_{h,k}\big\|_{\mathbf{M}^{1,p,p}(H)^*}&\le\Big(\int_0^{t_n}\big\|\tilde{E}_{h,k}(t_n-t)\big\|^{p'}_{\mathcal{L}^0_2}\,\mathrm{d}t\Big)^{\frac{1}{p'}}\\&\le\Big(\int_0^{t_n}\big\|\tilde{E}_{h,k}(t_n-t)A^{\frac{1-\beta}{2}}\big\|^{p'}_{\mathcal{L}}\big\|A^{\frac{\beta-1}{2}}\big\|^{p'}_{\mathcal{L}^0_2}\,\mathrm{d}t\Big)^{\frac{1}{p'}}.\end{aligned}$$

Recalling the error operator (2.9) we obtain for $t\in(t_j,t_{j+1})$, $j=0,\dots,n-1$,

$$\begin{aligned}&\big\|\tilde{E}_{h,k}(t_n-t)A^{\frac{1-\beta}{2}}\big\|_{\mathcal{L}}\\&\quad\le\big\|\big(S(t_n-t)-S(t_n-t_j)\big)A^{\frac{1-\beta}{2}}\big\|_{\mathcal{L}}+\big\|E^{n-j}_{h,k}A^{\frac{1-\beta}{2}}\big\|_{\mathcal{L}}\\&\quad\le\big\|\big(I-S(t-t_j)\big)A^{-\gamma}\big\|_{\mathcal{L}}\big\|S(t_n-t)A^{\frac{2\gamma+1-\beta}{2}}\big\|_{\mathcal{L}}+\big\|E^{n-j}_{h,k}A^{\frac{1-\beta}{2}}\big\|_{\mathcal{L}}\\&\quad\lesssim(t-t_j)^\gamma(t_n-t)^{-\frac{2\gamma+1-\beta}{2}}+\big(h^{2\gamma}+k^\gamma\big)(t_n-t_j)^{-\frac{2\gamma+1-\beta}{2}}\\&\quad\lesssim\big(h^{2\gamma}+k^\gamma\big)(t_n-t)^{-\frac{2\gamma+1-\beta}{2}},\end{aligned}\tag{4.8}$$

where we applied (2.3) with $\varrho = \gamma$ and (2.4), (2.12) with $\theta = 2\gamma$, $\varrho = 1-\beta$. By recalling (3.9), we conclude

$$\begin{aligned}\big\|W^A(t_n) - W_{h,k}^{A_h,n}\big\|_{\mathbf{M}^{1,p,p}(H)^*} &\lesssim \big(h^{2\gamma}+k^\gamma\big)\Big(\int_0^{t_n}(t_n-t)^{-p'\frac{2\gamma+1-\beta}{2}}\,\mathrm{d}t\Big)^{\frac{1}{p'}}\\ &\lesssim h^{2\gamma}+k^\gamma,\end{aligned}$$

which is the desired result. □

4.2. **Semilinear equation with additive noise.** Above we demonstrated that $V = \mathbf{M}^{1,p,p}(H)$ with $p$ large is suitable for the weak error analysis for the stochastic convolution. In order to treat semilinear equations we need a smaller space. Here we work with the Gelfand triple

$$\mathbf{G}^{1,p}(H) \subset L^2(\Omega,H) \subset \mathbf{G}^{1,p}(H)^*.$$

The line of proof is the same as above only that the convergence in the dual norm is more involved and relies on the local Lipschitz condition stated in Lemma 3.9, the Burkholder type inequality Lemma 3.5 and a classical Gronwall argument.

**Theorem 4.4.** *Let Assumptions 2.3 and 2.4 hold with $\beta \in (0,1]$. Let $X$ and $X_{h,k}$ be the solutions to equations (2.5) and (2.7), respectively. For every function $\varphi\colon H \to H$ that satisfies Assumption 4.1 and every $\gamma \in [0,\beta)$, we have for $h,k \in (0,1]$ the weak convergence*

$$\max_{n\in\{1,\dots,N\}}\big|\mathbf{E}\big[\varphi(X(t_n)) - \varphi(X_{h,k}^n)\big]\big| \le C\big(h^{2\gamma}+k^\gamma\big).$$

*Proof.* This is a direct consequence of (4.3) and Lemmas 4.5 and 4.6 below. □

**Lemma 4.5.** *Let the assumptions of Theorem 4.4 hold. For $\gamma \in (0,\beta)$, set $p = \frac{2}{1-\gamma}$. It holds*

$$\max_{n\in\{1,\dots,N\}}\sup_{h,k\in(0,1]}\big\|\Phi_{h,k}^n\big\|_{\mathbf{G}^{1,p}(H)} \le C.$$

*Proof.* By Lemma 4.2 we have $\|\Phi_{h,k}^n\|_{\mathbf{M}^{1,p,p}(H)} \le C$ uniformly in $n$ and $h,k$. In addition, by (2.6), Proposition 3.15, and Assumption 4.1, it holds $\|\Phi_{h,k}^n\|_{L^{2p}(\Omega,H)} \le C$ uniformly in $n$ and $h,k$. □

**Lemma 4.6.** *Let the assumptions of Theorem 4.4 hold. For $\gamma \in (0,\beta)$, set $p = \frac{2}{1-\gamma}$. Then there exists a constant $C$ independent of $h,k \in (0,1]$ such that*

$$\max_{n\in\{1,\dots,N\}}\big\|X(t_n) - X_{h,k}^n\big\|_{\mathbf{G}^{1,p}(H)^*} \le C\big(h^{2\gamma}+k^\gamma\big), \quad h,k\in(0,1].$$

*Proof.* Let $n \in \{1,\dots,N\}$ be arbitrary. By (2.5) and (2.8), we can write

$$\begin{aligned}X(t_n) - X_{h,k}^n &= \big(S(t_n) - S_{h,k}^n\big)X_0\\ &\quad+\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\big(S(t_n-t) - S_{h,k}^{n-j}\big)F(X(t))\,\mathrm{d}t\\ &\quad+\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}S_{h,k}^{n-j}\big(F(X(t)) - F(X_{h,k}^j)\big)\,\mathrm{d}t + W^A(t_n) - W_{h,k}^{A_h,n}.\end{aligned}$$

By recalling the error operators $E^n_{h,k}$ from (2.9) and $\tilde{E}_{h,k}(t)$ from (4.6), we obtain

$$
\begin{aligned}
\big\|X(t_n) - X^n_{h,k}\big\|_{\mathbf{G}^{1,p}(H)^*} &\leq \big\|E^n_{h,k}X_0\big\| \\
&\quad + \Big\| \int_0^{t_n} \tilde{E}_{h,k}(t_n - t)F(X(t))\,\mathrm{d}t\Big\|_{\mathbf{G}^{1,p}(H)^*} \\
&\quad + \Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S^{n-j}_{h,k}\big(F(X(t)) - F(X^j_{h,k})\big)\,\mathrm{d}t\Big\|_{\mathbf{G}^{1,p}(H)^*} \\
&\quad + \big\|W^A(t_n) - W^{A_h,n}_{h,k}\big\|_{\mathbf{G}^{1,p}(H)^*}.
\end{aligned}
\tag{4.9}
$$

By (2.12) with $\varrho = -\theta = -2\gamma$ and Assumption 2.3 (ii) we get

$$\big\|E^n_{h,k}X_0\big\| \leq \big\|E^n_{h,k}A^{-\gamma}\big\|_{\mathcal{L}}\big\|A^\gamma X_0\big\| \lesssim \big(h^{2\gamma} + k^\gamma\big)\big\|A^\gamma X_0\big\|.$$

For the second term in (4.9) we first use that $\|Z\|_{\mathbf{G}^{1,p}(H)^*} \leq \|Z\|_{L^2(\Omega,H)}$ for all $Z \in L^2(\Omega, H)$. Then by (4.8) with $\beta = 1$, the linear growth of $F$, and (2.6) we have

$$
\begin{aligned}
\Big\| \int_0^{t_n} \tilde{E}_{h,k}(t_n - t)F(X(t))\,\mathrm{d}t\Big\|_{\mathbf{G}^{1,p}(H)^*} &\leq \int_0^{t_n} \big\|\tilde{E}_{h,k}(t_n - t)\big\|_{\mathcal{L}}\,\big\|F(X(t))\big\|_{L^2(\Omega,H)}\,\mathrm{d}t \\
&\lesssim \big(h^{2\gamma} + k^\gamma\big)\int_0^{t_n} (t_n - t)^{-\gamma}\,\mathrm{d}t\,\Big(1 + \sup_{t\in[0,T]}\big\|X(t)\big\|_{L^2(\Omega,H)}\Big) \lesssim h^{2\gamma} + k^\gamma.
\end{aligned}
$$

For the third summand we first notice that Propositions 3.10 and 3.16 justify the use of Lemma 3.9 with $\eta = F$, $U = H$, $V = \dot{H}^{-1}$, $X_1 = X(t)$ and $X_2 = X^j_{h,k}$ with $t \in (t_j, t_{j+1}]$. We get

$$
\begin{aligned}
\big\|F(X(t)) - F(X^j_{h,k})\big\|_{\mathbf{G}^{1,p}(\dot{H}^{-1})^*} &\leq \max_{i\in\{1,2\}} |F|_{\mathcal{C}^i_{\mathrm{b}}(H,\dot{H}^{-1})} \\
&\quad\times \Big(1 + \|X(t)\|_{\mathbf{M}^{1,2p,p}(H)} + \|X^j_{h,k}\|_{\mathbf{M}^{1,2p,p}(H)}\Big)\|X(t) - X^j_{h,k}\|_{\mathbf{G}^{1,p}(H)^*} \\
&\lesssim \|X(t) - X^j_{h,k}\|_{\mathbf{G}^{1,p}(H)^*}.
\end{aligned}
$$

By (2.10), (2.11) with $\rho = \frac{1}{2}$, we get for the third term

$$
\begin{aligned}
&\Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S^{n-j}_{h,k}A_h^{\frac{1}{2}}A_h^{-\frac{1}{2}}P_hA^{\frac{1}{2}}A^{-\frac{1}{2}}\big(F(X(t)) - F(X^j_{h,k})\big)\,\mathrm{d}t\Big\|_{\mathbf{G}^{1,p}(H)^*} \\
&\quad\leq \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|S^{n-j}_{h,k}A_h^{\frac{1}{2}}\|_{\mathcal{L}}\|A_h^{-\frac{1}{2}}P_hA^{\frac{1}{2}}\|_{\mathcal{L}}\big\|F(X(t)) - F(X^j_{h,k})\big\|_{\mathbf{G}^{1,p}(\dot{H}^{-1})^*}\,\mathrm{d}t \\
&\quad\lesssim \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} t_{n-j}^{-\frac{1}{2}}\big(\big\|X(t) - X(t_j)\big\|_{\mathbf{G}^{1,p}(H)^*} + \big\|X(t_j) - X^j_{h,k}\big\|_{\mathbf{G}^{1,p}(H)^*}\big)\,\mathrm{d}t.
\end{aligned}
\tag{4.10}
$$

By Proposition 3.11, it holds $\|X(t) - X(t_j)\|_{\mathbf{G}^{1,p}(H)^*} \lesssim k^\gamma$ and therefore

$$\begin{aligned}&\Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S_{h,k}^{n-j} \big( F(X(t)) - F(X_{h,k}^j) \big) \,\mathrm{d}t \Big\|_{\mathbf{G}^{1,p}(H)^*} \\ &\qquad \lesssim k^{1+\gamma} \sum_{j=0}^{n-1} t_{n-j}^{-\frac{1}{2}} + k \sum_{j=0}^{n-1} t_{n-j}^{-\frac{1}{2}} \big\| X(t_j) - X_{h,k}^j \big\|_{\mathbf{G}^{1,p}(H)^*}.\end{aligned}$$

The fourth summand is estimated in Lemma 4.3. Altogether we conclude that

$$\big\| X(t_n) - X_{h,k}^n \big\|_{\mathbf{G}^{1,p}(H)^*} \lesssim \big( h^{2\gamma} + k^\gamma \big) + k \sum_{j=0}^{n-1} t_{n-j}^{-\frac{1}{2}} \big\| X(t_j) - X_{h,k}^j \big\|_{\mathbf{G}^{1,p}(H)^*}.$$

By the discrete Gronwall Lemma 2.1 the assertion follows. □

Weak approximation concerns the approximation of the Markov semigroup. In view of Theorem 4.4 and Corollary 3.12, we see that the rate of weak convergence in time coincides with the Hölder regularity in time for the Markov semigroup, which is intuitively to be expected for an Euler approximation. A similar connection to the discretization in space seems to be a more subtle issue.

The relationship between the strong and weak rate of convergence can also be seen in the view of duality. The following corollary deduces a strong convergence result from Lemma 4.6 and Propositions 3.10 and 3.16. It indicates why one often encounters the rule of thumb that the order of weak convergence is twice the order of strong convergence.

**Corollary 4.7.** *Let the assumptions of Theorem 4.4 hold. Let $X$ and $X_{h,k}$ denote the solutions to equations (2.5) and (2.7), respectively. Then for every $\gamma \in (0, \beta)$ there exists a constant $C$ such that*

$$\max_{n \in \{1,\dots,N\}} \|X(t_n) - X_{h,k}^n\|_{L^2(\Omega,H)} \le C(h^\gamma + k^{\frac{\gamma}{2}}), \quad h, k \in (0,1].$$

*Proof.* For arbitrary $n \in \{1, \dots, N\}$ we have by the duality argument with $p = \frac{2}{1-\gamma}$

$$\begin{aligned}\|X(t_n) - X_{h,k}^n\|_{L^2(\Omega,H)}^2 &= \big\langle X(t_n) - X_{h,k}^n, X(t_n) - X_{h,k}^n \big\rangle_{L^2(\Omega,H)} \\ &\le \big( \|X(t_n)\|_{\mathbf{G}^{1,p}(H)} + \|X_{h,k}^n\|_{\mathbf{G}^{1,p}(H)} \big) \|X(t_n) - X_{h,k}^n\|_{\mathbf{G}^{1,p}(H)^*}.\end{aligned}$$

The first factor is bounded independently of $n \in \{1, \dots, N\}$ by Propositions 3.10 and 3.16. For the second factor we apply Lemma 4.6 and since $(h^{2\gamma} + k^\gamma)^{\frac{1}{2}} \le (h^\gamma + k^{\frac{\gamma}{2}})$ for all $h, k \in (0,1]$ the result follows. □

4.3. **Multiplicative noise.** The choice $V = \mathbf{G}^{1,p}(H)$ of Subsection 4.2 works only for equations with additive noise. We demonstrate this here by considering the following equation with linear multiplicative noise

$$\mathrm{d}X(t) + AX(t) \,\mathrm{d}t = BX(t) \,\mathrm{d}W(t), \ t \in (0,T]; \quad X(0) = X_0.$$

Here $B \in \mathcal{L}(H, \mathcal{L}_2(H_0, \dot{H}^{\beta-1}))$. In order to perform the Gronwall argument in the $\mathbf{G}^{1,p}(H)^*$-norm for this equation, one would need a bound

$$\begin{aligned}\Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S_{h,k}^{n-j} B\big(X(t) - X_{h,k}^j\big) \, \mathrm{d}W(t) \Big\|_{\mathbf{G}^{1,p}(H)^*} \\ \lesssim \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \big\|X(t) - X_{h,k}^j\big\|_{\mathbf{G}^{1,p}(H)^*} \, \mathrm{d}t,\end{aligned} \tag{4.11}$$

cf. (4.10). Attempting to prove this, we integrate by parts and move the supremum inside the integral to get

$$\begin{aligned}&\Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S_{h,k}^{n-j} B\big(X(t) - X_{h,k}^j\big) \, \mathrm{d}W(t) \Big\|_{\mathbf{G}^{1,p}(H)^*} \\ &\quad = \sup_{Z \in \mathbf{G}^{1,p}(H)} \frac{1}{\|Z\|_{\mathbf{G}^{1,p}(H)}} \Big\langle Z, \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S_{h,k}^{n-j} B\big(X(t) - X_{h,k}^j\big) \, \mathrm{d}W(t) \Big\rangle_{L^2(\Omega,H)} \\ &\quad \leq \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \sup_{Z \in \mathbf{G}^{1,p}(H)} \frac{1}{\|Z\|_{\mathbf{G}^{1,p}(H)}} \big\langle B^* S_{h,k}^{n-j} D_t Z, X(t) - X_{h,k}^j \big\rangle_{L^2(\Omega,H)} \, \mathrm{d}t.\end{aligned}$$

If it would hold $B^* S_{h,k}^{n-j} D_t \in \mathcal{L}(\mathbf{G}^{1,p}(H))$, then the bound (4.11) would follow, but this is not the case as only $D_t \colon \mathbf{G}^{1,p}(H) \to L^p(\Omega, \mathcal{L}_2^0)$ for a.e. $t \in [0,T]$. We see no other natural choice of the space $V$ but it might be that the estimate (4.2) is too crude in order to treat multiplicative noise.

## 5. Approximation by the finite element method

In this section we describe an explicit example for the linear operator $A$ and its corresponding numerical discretization by the finite element method.

For this we consider the Hilbert space $H = L^2(D)$, where $D \subset \mathbf{R}^d$, $d = 1, 2, 3$, is a bounded, convex, and polygonal domain. The linear operator $(A, \mathcal{D}(A))$ is defined to be $Au = -\nabla \cdot (a \nabla u) + cu$ with Dirichlet boundary conditions, where $a, c \colon D \to \mathbf{R}$ are sufficiently smooth with $c(\xi) \geq 0$ and $a(\xi) \geq a_0 > 0$ for $\xi \in D$. Then $A$ is an elliptic, selfadjoint, second order differential operator with compact inverse, see for instance [33, Chap. 6.1]. In particular, $A$ satisfies Assumption 2.3 (i).

We measure spatial regularity in terms of the abstract spaces $\dot{H}^\theta$, $\theta \in \mathbf{R}$, which now are related to the classical Sobolev spaces, for example $\dot{H}^1 = H_0^1(D)$ and $\dot{H}^2 = H_0^1(D) \cap H^2(D)$. For more details we refer to [31, App. B.2] and the references therein.

Let $(T_h)_{h \in (0,1]}$ be a regular family of triangulations of $D$ with maximal mesh size $h \in (0,1]$. We define a family of subspaces $(V_h)_{h \in (0,1]}$ of $\dot{H}^1$, consisting of continuous piecewise linear functions corresponding to $(T_h)_{h \in (0,1]}$. By equipping the space $\dot{H}^1$ with the inner product $\langle \cdot, \cdot \rangle_1 := \langle A^{\frac{1}{2}} \cdot, A^{\frac{1}{2}} \cdot \rangle$, we define $A_h \colon V_h \to V_h$, $h \in (0,1]$, to be the linear operators given by

$$\langle A_h v_h, u_h \rangle = \langle v_h, u_h \rangle_1, \quad \forall v_h, u_h \in V_h.$$

Now, from [31, (3.15)] we get $\|A_h^{-1}P_h x\| \le \|x\|_{-1}$ for all $x \in \dot{H}^{-1}$. Hence, it holds

$$\|A_h^{-\frac{1}{2}} P_h A^{\frac{1}{2}}\|_{\mathcal{L}} \le 1.$$

An interpolation between this and $\|P_h\|_{\mathcal{L}} \le 1$ yields (2.11) for $\varrho \in [0,1]$.

As in Subsection 2.3 we denote by $(S(t))_{t\ge 0}$ the semigroup generated by $-A$ and $S_{h,k} := (I + kA_h)^{-1}P_h$. The standard literature on finite element methods, for instance [41], provides error estimates for the approximation of the semigroup with smooth and nonsmooth initial data. More precisely, it holds for the error operator (4.6) that

$$\|\tilde{E}_{h,k}(t)x\| \le C\big(h^2 + k\big)t^{-\frac{2-q}{2}}\|x\|_{\dot{H}^q}, \quad x \in \dot{H}^q,\ q = 0, 2.$$

By interpolation this covers the smooth data case $-\theta \le \varrho \le 0$ of (2.12). For the purpose of the present work we need to extend this to less regular initial data. This is done by the next lemma, which is a consequence of [31, Lemma 3.12].

**Lemma 5.1.** *Under the above assumptions and for $0 \le \theta \le 2$ and $-\theta \le \varrho \le \min(1, 2-\theta)$, the following estimate holds true*

$$\|\tilde{E}_{h,k}(t)x\| \le C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-\frac{\theta+\varrho}{2}}\|x\|_{-\varrho}, \quad x \in \dot{H}^{-\varrho},\ t > 0,\ h, k \in (0, 1].$$

*Proof.* As noted above it remains to treat the case when $0 \le \varrho \le \min(1, 2-\theta)$. By [31, Lemma 3.12 (i)] the estimate

$$\|\tilde{E}_{h,k}(t)x\| \le C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-\frac{\theta}{2}}\|x\|, \quad t > 0,\ 0 \le \theta \le 2, \tag{5.1}$$

holds for all $h, k \in (0, 1]$. By [31, Lemma 3.12 (iii)] the error operator $\tilde{E}_{h,k}$ also satisfies, for $1 \le \theta \le 2$,

$$\|\tilde{E}_{h,k}(t)x\| \le C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-1}\|x\|_{-(2-\theta)}, \quad t > 0. \tag{5.2}$$

Interpolation of (5.1) and (5.2) with fixed $\theta \in [1, 2]$ gives that, for $\lambda \in [0, 1]$,

$$\begin{aligned}\|\tilde{E}_{h,k}(t)x\| &\le C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-(1-\lambda)\frac{\theta}{2}}t^{-\lambda}\|x\|_{-\lambda(2-\theta)}\\ &= C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-\frac{\theta}{2}-\frac{\lambda(2-\theta)}{2}}\|x\|_{-\lambda(2-\theta)}, \quad t > 0.\end{aligned}$$

If we let $\varrho = \lambda(2-\theta)$, then we get the following estimate: for $1 \le \theta \le 2$ and $0 \le \varrho \le 2 - \theta$,

$$\|\tilde{E}_{h,k}(t)x\| \le C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-\frac{\theta+\varrho}{2}}\|x\|_{-\varrho}, \quad t \ge 0. \tag{5.3}$$

By [31, Lemma 3.12 (ii)] it holds

$$\|\tilde{E}_{h,k}(t)x\| \le Ct^{-\frac{\varrho}{2}}\|x\|_{-\varrho}, \quad t > 0,\ 0 \le \varrho \le 1, \tag{5.4}$$

and using (5.3) with $\theta = 1$ and (5.4), both with the same $0 \le \varrho \le 1$, yields

$$\begin{aligned}\|\tilde{E}_{h,k}(t)x\| = \|\tilde{E}_{h,k}(t)x\|^\lambda\|\tilde{E}_{h,k}(t)x\|^{1-\lambda} &\le C\big(h + k^{\frac{1}{2}}\big)^\lambda t^{-\frac{\lambda+\varrho}{2}}\|x\|_{-\varrho}\\ &\le C\big(h^\lambda + k^{\frac{\lambda}{2}}\big)t^{-\frac{\lambda+\varrho}{2}}\|x\|_{-\varrho}, \quad t > 0,\ 0 \le \lambda \le 1.\end{aligned} \tag{5.5}$$

Combining (5.3) and (5.5) concludes the proof. □

Writing the statement of the lemma in operator form yields

$$\|\tilde{E}_{h,k}(t)A^{\frac{\varrho}{2}}\|_{\mathcal{L}} \le C\big(h^\theta + k^{\frac{\theta}{2}}\big)t^{-\frac{\theta+\varrho}{2}}, \quad t>0,\ 0\le\theta\le 2,\ -\theta\le\varrho\le\min(1,2-\theta).$$

This is (2.12) for the finite element method. To verify Assumption 2.4 it remains to show (2.10). By [31, (3.42)]

$$\|S_{h,k}^n x\| \le C t^{-\frac{1}{2}}\|x\|_{-1}.$$

Interpolating between this and $\|S_{h,k}^n x\| \le C\|x\|$ yields (2.10).

**Acknowledgement.** The authors wish to thank M. Kovács for fruitful discussions during the preparation of the work [1], which led to true improvements of the present paper. We also thank A. Lang and X. Wang for valuable comments on an earlier version of the manuscript.

The first two authors also acknowledge the kind support by W.-J. Beyn, B. Gentz, and the DFG-funded CRC 701 'Spectral Structures and Topological Methods in Mathematics' by making possible an inspiring research stay at Bielefeld University, where part of this work was written.

## References

[1] A. Andersson, M. Kovács, and S. Larsson, *Weak error analysis for semilinear stochastic Volterra equations with additive noise*, 2014. (in preparation).

[2] A. Andersson and S. Larsson, *Weak convergence for a spatial approximation of the nonlinear stochastic heat equation*, ArXiv Preprint, arXiv:1212.5564 (2012).

[3] V. V. Anh, W. Grecksch, and J. Yong, *Regularity of backward stochastic Volterra integral equations in Hilbert spaces*, Stoch. Anal. Appl. **29** (2011), no. 1, 146–168.

[4] F. E. Benth, T. Deck, and J. Potthoff, *A white noise approach to a class of non-linear stochastic heat equations*, J. Funct. Anal. **146** (1997), no. 2, 382–415.

[5] C.-É. Bréhier, *Approximation of the invariant measure with an Euler scheme for stochastic PDE's driven by space-time white noise*, ArXiv Preprint, arXiv:1202.2707 (2012).

[6] ———, *Strong and weak order in averaging for SPDEs*, Stoch. Proc. Appl. **122** (2012), no. 7, 2553–2593.

[7] C.-É. Bréhier and M. Kopec, *Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme*, ArXiv Preprint, arXiv:1311.7030 (2013).

[8] E. Buckwar, R. Kuske, S.-E. Mohammed, and T. Shardlow, *Weak convergence of the Euler scheme for stochastic differential delay equations*, LMS J. Comput. Math. **11** (2008), 60–99.

[9] E. Buckwar and T. Shardlow, *Weak approximation of stochastic differential delay equations*, IMA J. Numer. Anal. **25** (2005), no. 1, 57–86.

[10] E. Clément, A. Kohatsu-Higa, and D. Lamberton, *A duality approach for the weak approximation of stochastic differential equations*, Ann. Appl. Probab. **16** (2006), no. 3, 1124–1154.

[11] D. Cohen and M. Sigg, *Convergence analysis of trigonometric methods for stiff second-order stochastic differential equations*, Numer. Math. **121** (2012), no. 1, 1–29.

[12] D. Conus, A. Jentzen, and R. Kurniawan, *Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients*, ArXiv Preprint, arXiv:1408.1108 (2014).

[13] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.

[14] A. de Bouard and A. Debussche, *Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation*, Appl. Math. Optim. **54** (2006), 369–399.

[15] A. Debussche, *Weak approximation of stochastic partial differential equations: the nonlinear case*, Math. Comp. **80** (2011), no. 273, 89–117.

[16] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comp. **78** (2009), no. 266, 845–863.

[17] C. M. Elliott and S. Larsson, *Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation*, Math. Comp. **58** (1992), no. 198, 603–630, S33–S36.

[18] M. Fuhrman and G. Tessitore, *Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control*, Ann. Probab. **30** (2002), 1397–1465.

[19] M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT **49** (2009), 343–356.

[20] A. Grorud and É. Pardoux, *Intégrales Hilbertiennes anticipantes par rapport à un processus de Wiener cylindrique et calcul stochastique associé*, Appl. Math. Optim. **25** (1992), no. 1, 31–49.

[21] E. Hausenblas, *Weak approximation for semilinear stochastic evolution equations*, Stochastic analysis and related topics VIII, 2003, pp. 111–128. MR2189620 (2006k:60114)

[22] ———, *Weak approximation of the stochastic wave equation*, J. Comput. Appl. Math. **235** (2010), no. 1, 33–58.

[23] H. Holden, B. Øksendal, J. Ubøe, and T. Zhang, *Stochastic Partial Differential Equations*, Second, Universitext, Springer, New York, 2010. A modeling, white noise functional approach.

[24] S. Janson, *Gaussian Hilbert Spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997.

[25] A. Jentzen and M. Röckner, *Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise*, J. Differential Equations **252** (2012), no. 1, 114–136.

[26] A. Kohatsu-Higa, *Weak approximations. A Malliavin calculus approach*, Math. Comp. **70** (2001), no. 233, 135–172.

[27] M. Kopec, *Quelques contributions à l'analyse numérique d'équations stochastiques*, Ph.D. Thesis, 2014.

[28] M. Kovács, S. Larsson, and F. Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise*, BIT Numer. Math. **52** (2012), no. 1, 85–108.

[29] M. Kovács, S. Larsson, and F. Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes*, BIT Numer. Math. **53** (2013), no. 2, 497–525.

[30] R. Kruse, *Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise*, IMA J. Numer. Anal. **34** (2014), no. 1, 217–251.

[31] ———, *Strong and Weak Approximation of Stochastic Evolution Equations*, Lecture Notes in Mathematics, vol. 2093, Springer, 2014.

[32] R. Kruse and S. Larsson, *Optimal regularity for semilinear stochastic partial differential equations with multiplicative noise*, Electron. J. Probab. **17** (2012), no. 65, 1–19.

[33] S. Larsson and V. Thomée, *Partial Differential Equations with Numerical Methods*, Texts in Applied Mathematics, vol. 45, Springer-Verlag, Berlin, 2003.

[34] J. A. León and D. Nualart, *Stochastic evolution equations with random generators*, Ann. Probab. **26** (1998), no. 1, 149–186.

[35] F. Lindner and R. L. Schilling, *Weak order for the discretization of the stochastic heat equation driven by impulsive noise*, Potential Anal. **38** (2012), no. 2, 345–179.

[36] D. Nualart, *The Malliavin calculus and related topics*, Second, Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.

[37] B. Øksendal, *Stochastic Differential Equations*, Sixth, Universitext, Springer-Verlag, Berlin, 2003. An introduction with applications.

[38] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, vol. 44, Springer, New York, 1983.

[39] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.

[40] M. Sanz-Solé, *Malliavin Calculus: with Applications to Stochastic Partial Differential Equations*, 1. ed., Fundamental sciences : Mathematics, EPFL Press, Lausanne, Switzerland, 2005.

[41] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Second, Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 2006.

[42] J. M. A. M. van Neerven, *Stochastic Evolution Equations*, 2008. ISEM lecture notes.

[43] J. B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, XIV—1984, 1986, pp. 265–439.

[44] X. Wang, *An exponential integrator scheme for time discretization of nonlinear stochastic wave equation*, ArXiv Preprint, arXiv:1312.5185 (2013).

[45] ———, *Weak error estimates of the exponential Euler scheme for semi-linear SPDEs without Malliavin calculus*, ArXiv Preprint, arXiv:1408.0713 (2014).

[46] X. Wang and S. Gan, *Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise*, J. Math. Anal. Appl. **398** (2013), no. 1, 151–169.

[47] Y. Yan, *Error Analysis and Smoothing Properties of Discretized Deterministic and Stochastic Parabolic Problems*, Ph.D. Thesis, 2003. `http://www.math.chalmers.se/Math/Research/Preprints/Doctoral/2003/3.pdf`.

[48] ———, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), no. 4, 1363–1384.



Adam Andersson, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden

*E-mail address*: `adam.andersson@chalmers.se`

Raphael Kruse, Technische Universität Berlin, Institut für Mathematik, Sek. MA 5-3, Strasse des 17. Juni 136, DE-10623 Berlin, Germany

*E-mail address*: `kruse@math.tu-berlin.de`

Stig Larsson, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden

*E-mail address*: `stig@chalmers.se`