WEAK ERROR ANALYSIS FOR SEMILINEAR STOCHASTIC VOLTERRA EQUATIONS WITH ADDITIVE NOISE

ADAM ANDERSSON, MIHÁLY KOVÁCS, AND STIG LARSSON

Abstract. We prove a weak error estimate for the approximation in space and time of a semilinear stochastic Volterra integro-differential equation driven by additive space-time Gaussian noise. We treat this equation in an abstract framework, in which parabolic stochastic partial differential equations are also included as a special case. The approximation in space is performed by a standard finite element method and in time by an implicit Euler method combined with a convolution quadrature. The weak rate of convergence is proved to be twice the strong rate, as expected. Our weak convergence result concerns not only the solution at a fixed time but also integrals of the entire path with respect to any finite Borel measure. The proof does not rely on a Kolmogorov equation. Instead it is based on a duality argument from Malliavin calculus.

1. Introduction

Let \((S_t)_{t \in [0,T]}\) be an evolution family of bounded, self-adjoint, linear operators on a separable Hilbert space \((H, \| \cdot \|, \langle \cdot, \cdot \rangle)\), not necessarily enjoying the semigroup property. Related to \((S_t)_{t \in [0,T]}\) is a densely defined, linear, self-adjoint, positive definite operator \(A: D(A) \subset H \to H\) with compact inverse. Let \((A^{\alpha})_{\alpha \in \mathbb{R}}\) denote the fractional powers of \(A\), which are well defined, let \((\dot{H}^\alpha)_{\alpha \in \mathbb{R}}\) denote the spaces \(\dot{H}^\alpha = D(A^\alpha)\) for \(\alpha \geq 0\) with dual spaces \(\dot{H}^{-\alpha} = (\dot{H}^\alpha)^*\). We assume that \((S_t)_{t \in [0,T]}\) is strongly differentiable with derivative \((\dot{S}_t)_{t \in [0,T]}\) and that there exist \(\rho \in [1,2)\) and constants \((L_s)_{s \in [0,2]}\) so that

\[
(1.1) \quad \| A^{\min(1,\alpha)} S_t x \| + \| A^{-\alpha \over 2} \dot{S}_t x \| \leq L_s t^{-s} \| x \|, \quad t \in (0,T], \ x \in H, \ s \in [0,2].
\]

If \((S_t)_{t \in [0,T]}\) is the analytic semigroup generated by \(-A\), then (1.1) holds with \(\rho = 1\). If \((S_t)_{t \in [0,T]}\) is the solution operator \(S_t x = Y^x_t\) of the Volterra equation

\[
\dot{Y}^x_t + \int_0^t b_{t-s} A Y^x_s \, ds = 0, \quad t \in (0,T]; \ Y^x_0 = x,
\]

where \(b: (0, \infty) \to \mathbb{R}\) is the Riesz kernel \(b_t = t^{\rho - 2} / \Gamma(\rho - 1)\) for some \(\rho \in (1,2)\), then \((S_t)_{t \in [0,T]}\) satisfies (1.1). The latter example is the main motivation of the present paper. In Subsection 5.2 we verify (1.1) for slightly more general kernels \(b\).

The main object of study in this paper is the stochastic evolution equation

\[
(1.2) \quad X_t = S_t x_0 + \int_0^t S_{t-s} F(X_s) \, ds + \int_0^t S_{t-s} \, dW_s, \quad t \in [0,T].
\]

2010 Mathematics Subject Classification. 60H15, 60H07, 65C30, 65M60.
Key words and phrases. Stochastic Volterra equations, finite element method, backward Euler, convolution quadrature, strong and weak convergence, Malliavin calculus, regularity, duality.
Under this assumption $X_t \in \dot{H}^3$, $P$-almost surely. The smoothest case $\beta = 1/\rho$ corresponds to trace class noise as (1.3) reduces to $\|Q_\beta\|_{L_2} = \sqrt{\text{Tr}(Q)} < \infty$.

For Hilbert spaces $U$, $V$ the space $C_b^k(U; V)$ consists of all, not necessarily bounded, functions $\phi: U \to V$, whose Fréchet derivatives of orders $1, \ldots, k$ are bounded. The non-linear drift $F: H \to H$ is assumed to satisfy, for some $\delta \in [0, 2/\rho)$,

$$(1.4)\quad F \in C^1_b(H; H) \cap C^2_b(H; \dot{H}^{-\delta}).$$

This assumption includes interesting cases where $F \not\in C^2_b(H; H)$, e.g., Nemytskii operators on $H = L^2(D)$ for a spatial domain $D \subset \mathbb{R}^d$, with $\delta > d/2$. The initial value $x_0$ is deterministic and, for simplicity, quite smooth:

$$(1.5)\quad x_0 \in \dot{H}^3 := \mathcal{D}(A^{\frac{1}{2}}).$$

In the present paper we study weak convergence of approximations of the solution of (1.2). Our main example is the mild solution of the stochastic Volterra integro-differential equation

$$(1.6)\quad dX_t + \left(\int_0^t b_{t-s}AX_s \, ds\right) \, dt = F(X_t) \, dt + dW_t, \quad t \in [0, T]; \quad X_0 = x_0,$$

where $b_t = t^{\rho-2}/\Gamma(\rho - 1)$ as above or slightly more general. Discretization in time is performed by the backward Euler method and the convolution integral is approximated by a convolution quadrature. For spatial approximation either spectral or finite element approximation is considered. In the papers [14], [15], strong, respectively weak, convergence of numerical approximations were proven, for linear stochastic Volterra equations ($F = 0$). The deterministic error analysis needed for the present paper will be cited from these papers.

Another example to which our results apply is the mild solution of the parabolic stochastic evolution equation

$$(1.7)\quad dX_t + AX_t \, dt = F(X_t) \, dt + dW_t, \quad t \in [0, T]; \quad X_0 = x_0.$$

Approximation in time is performed by the backward Euler method and the same spatial approximation is considered as for (1.6). Weak convergence analysis for (1.7) is well studied [1], [2], [4], [5], [6], [8], [11], [12], [23], [24], [25]. In contrast to [1], by means of the weaker assumption $F \in C^1_b(H; \dot{H}^{-\delta})$ in (1.4) and the new Lemma 4.5, we allow the nonlinear drift $F$ to be a Nemytskii operator not only in one space dimension but also in two and three dimensions and we no longer assume that the finite element mesh family is quasi-uniform. We also consider a more general form of the weak error, see (1.8) below. We thus present some new results also for (1.7).
Let \( \varphi : H \to \mathbf{R} \) be a twice Fréchet differentiable mapping of polynomial growth and \( \nu \) a finite Borel measure on \( [0, T] \). We consider the error
\[
e_{\varphi, \nu}(X, Y) = \left| \mathbf{E} \left[ \varphi \left( \int_0^T X_t \, d\nu_t \right) - \varphi \left( \int_0^T Y_t \, d\nu_t \right) \right] \right|,
\]
where \( X, Y \in L^2_{\nu}(0, T; L^2(\Omega; H)) \). In all the works we are aware of, \( (1.8) \) is considered with \( \nu = \delta_\tau \), where \( \delta_\tau \) is the Dirac measure concentrated \( \tau \), for fixed \( \tau \in (0, T] \). In that case \( \mathbf{E}[\varphi(X_\tau)] \) is the solution to a Kolmogorov PDE, which is used in the analysis. Unfortunately, this is not true for \( \mathbf{E}[\varphi(\int_0^T X_t \, d\nu_t)] \). Moreover, Volterra equations are non-Markovian, so there is no Kolmogorov equation available for the analysis. Instead, we use another approach to analyze \( (1.8) \) that was recently introduced in \([1]\). The approach relies on a duality argument with a Gelfand triple of refined Sobolev-Malliavin spaces. In \([1]\) the technique was demonstrated in the Markovian setting of \( (1.7) \) and \( \nu = \delta_\tau \). In the present paper we apply it in a setting where no other known approach applies.

The paper is organized as follows: In Subsection 2.1 we fix the basic notation and in Subsection 2.2 we recall the theory of refined Sobolev-Malliavin spaces from \([1]\). In Section 3 we discuss existence and uniqueness of solutions of \( (1.2) \) and prove temporal Hölder regularity in the classical \( L^p(\Omega; H) \)-sense and in the weaker sense of a dual Sobolev-Malliavin norm. In Section 4 we present an abstract approximation scheme for \( (1.2) \) and prove our main result on weak convergence, Theorem 4.7. In addition, we prove strong convergence, which is used to establish Malliavin regularity for the solution to \( (1.2) \) by a limiting procedure. In Section 5 we verify our abstract assumptions for semilinear parabolic stochastic partial differential equations and stochastic Volterra integro-differential equations.

2. Preliminaries

2.1. Spaces of functions and operators. Let \((U, \| \cdot \|_U, \langle \cdot, \cdot \rangle_U), (V, \| \cdot \|_V, \langle \cdot, \cdot \rangle_V)\) be separable Hilbert spaces. Let \( \mathcal{L}(U; V) \) be the Banach space of all bounded linear operators \( U \to V \). We use the abbreviations \( \mathcal{L}(U) = \mathcal{L}(U; U) \) and \( \mathcal{L} = \mathcal{L}(H) \), where \( H \) is the Hilbert space introduced in Section 1. By \( \mathcal{L}_2(U; V) \subset \mathcal{L}(U; V) \) we denote the subspace of all Hilbert-Schmidt operators. It is a Hilbert space endowed with the norm and inner product
\[
\| T \|_{\mathcal{L}_2(U; V)} = \left( \sum_{j \in \mathbf{N}} \| Tu_j \|_V^2 \right)^{\frac{1}{2}}, \quad \langle S, T \rangle_{\mathcal{L}_2(U; V)} = \sum_{j \in \mathbf{N}} \langle Su_j, Tu_j \rangle_V.
\]
Both are independent of the specific choice of ON-basis \((u_j)_{j \in \mathbf{N}} \subset U\).

Denote by \( \mathcal{C}(U; V) \) the space of all continuous mappings. Let \( \mathcal{C}^k(U; V) \subset \mathcal{C}(U; V) \) be the subspace of all \( k \)-times continuously Fréchet differentiable mappings \( U \to V \). When \( V = \mathbf{R} \) we can identify the first derivative of \( \phi \in \mathcal{C}^1(U; \mathbf{R}) \) with its gradient \( \phi'(u) \in U^* = U \), by the Riesz Representation Theorem. For integers \( 0 \leq k \leq m \) and \( \phi \in \mathcal{C}^k(U; V) \), let
\[
|\phi|_{C^k_p(U; V)} = \sup_{u_1, \ldots, u_k \in U} \| \phi^{(k)}(u) \cdot (u_1, \ldots, u_k) \|_V \left( 1 + \| u \|^{m-k} \| u_1 \| \cdots \| u_k \| \right),
\]
and let \( \mathcal{C}^k_p(U; V) \) be the space of \( \phi \in \mathcal{C}^k(U; V) \) such that \( |\phi|_{C^k_p(U; V)} < \infty \) for \( 0 \leq l \leq k \), i.e., the space of functions with polynomially bounded derivatives. Let \( \mathcal{C}^\infty_p(U; V) \) be the space of all infinitely many times differentiable mappings.
\( \phi : U \to V \) such that \( \phi \) and all its derivatives satisfy a polynomial bound. Let \( C^k_b(U; V) \) denote the space of \( \phi \in C^k(U; V) \) such that
\[
|\phi|_{C^k_b(U; V)} = \sup_{u, u_1, \ldots, u_l \in U} \frac{||\phi^{(l)}(u) \cdot (u_1, \ldots, u_l)||_V}{||u_1||_V \cdots ||u_l||_V} < \infty, \quad 1 \leq l \leq k.
\]

Recall that the Mean Value Theorem for \( \phi \in C^1(U; V) \) reads as
\[ (2.3) \quad \phi(x) = \phi(y) + \int_0^1 \phi'(y + \lambda(x - y)) \cdot (x - y) \, d\lambda, \quad x, y \in U. \]

By \( \mathcal{M}_T \) we denote the space of all finite Borel measures on the interval \([0, T]\). For \( \nu \in \mathcal{M}_T \) we write \( |\nu| = \nu([0, T]) \) and for a Banach space \( V \) we let \( L^p(\nu) \) be the Bochner space of \( \nu \)-measurable mappings \( f : [0, T] \to V \) such that
\[
\|Z\|_{L^p(\nu)} = \left( \int_0^T \|Z_t\|_V^p \, d\nu_t \right)^{\frac{1}{p}} < \infty,
\]
with the usual modification for \( p = \infty \). When \( \nu \) is Lebesgue measure we write \( L^p(0, T; V) \).

The next lemma is used in the proof of Malliavin regularity in Proposition 4.4 by a limiting procedure.

**Lemma 2.1.** Let \( \mathcal{X}, \mathcal{Y} \) be separable Hilbert spaces such that the embedding \( \mathcal{X} \subset \mathcal{Y} \) is continuous. If \( x \in \mathcal{Y} \) and \((x_n)_{n \in \mathbb{N}} \subset \mathcal{X} \) are such that \( x_n \to x \) weakly in \( \mathcal{Y} \) as \( n \to \infty \) and \( \sup_{n \in \mathbb{N}} \|x_n\|_{\mathcal{X}} < \infty \), then \( x \in \mathcal{X} \).

**Proof.** Any closed ball in \( \mathcal{X} \) is weakly compact and since \((x_n)_{n \in \mathbb{N}} \) is a bounded sequence in \( \mathcal{X} \), there exists a subsequence \((x_{n_k})_{k \in \mathbb{N}} \) and \( \bar{x} \in \mathcal{X} \) such that \( x_{n_k} \to \bar{x} \) weakly in \( \mathcal{X} \). Therefore \( x_{n_k} \to \bar{x} \) also in the weak topology of \( \mathcal{Y} \) because \( \mathcal{Y}^* \subset \mathcal{X}^* \) is continuous. By assumption \( x_n \to x \) weakly in \( \mathcal{Y} \), so \( x = \bar{x} \in \mathcal{X} \). \( \square \)

We cite the following version of Gronwall’s lemma [9, Lemma 7.1].

**Lemma 2.2.** Let \( T > 0, N \in \mathbb{N}, k = T/N, \) and \( t_n = nk \) for \( 0 \leq n \leq N \). If \( \varphi_1, \ldots, \varphi_N \geq 0 \) satisfy for some \( M_0, M_1 \geq 0 \) and \( \mu, \nu > 0 \) the inequality
\[
\varphi_n \leq M_0 (1 + t_n^{1+\mu}) + M_1 \sum_{j=1}^{n-1} t_{n-j}^{1+\nu} \varphi_j, \quad 1 \leq n \leq N,
\]
then there exists a constant \( M_2 = M_2(\mu, \nu, M_1, T) \) such that
\[
\varphi_n \leq M_0 M_2 (1 + t_n^{1+\mu}), \quad 1 \leq n \leq N.
\]

### 2.2. The Wiener integral and Malliavin calculus

Let \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})\) be a filtered probability space, with Bochner spaces \( L^p(\Omega; V) = L^p(\Omega, \mathcal{F}, \mathbb{P}; V) \), \( p \in [1, \infty], V \) being a Banach space. In the case \( V = \mathbb{R} \) we write \( L^p(\Omega) = L^p(\Omega, \mathbb{R}) \). Recall that \( Q \in \mathcal{L}(H) \) is a linear, self-adjoint and positive semidefinite operator. Let \( H_0 = Q^{\frac{1}{2}}(H) \) be the Hilbert space endowed with inner product \( \langle u, v \rangle_{H_0} = \langle Q^{-\frac{1}{2}} u, Q^{-\frac{1}{2}} v \rangle \), where \( Q^{-\frac{1}{2}} \) denotes the pseudoinverse of \( Q^{\frac{1}{2}} \) if it is not injective. By \( \mathcal{L}_2(\mathbb{H}_0; H) \) we denote the space of Hilbert-Schmidt operators \( H_0 \to H \). Let \( W \) be a cylindrical Q-Wiener process on \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})\), i.e., \( W \in \mathcal{L}(H_0; C(0, T; L^2(\Omega))) \) and \((W_t u)_{t \in [0, T]} \) is an \((\mathcal{F}_t)_{t \in [0, T]}\)-adapted real-valued Brownian motion for every \( u \in H_0 \) with
\[
\mathbb{E}[W_s u W_t v] = \min(s, t) \langle u, v \rangle_{H_0}, \quad u, v \in H_0, \ s, t \in [0, T].
\]
The stochastic Wiener integral
\[ \int_0^T \Phi_t \, dW_t, \quad \Phi \in L^2(0, T; L^0_2), \]
is a random variable in \( L^p(\Omega; H), \ p \in [2, \infty) \). It can be defined in various ways and its basic properties are not hard to derive, we refer to [7, 19, 22]. We cite the following consequence of the Burkholder inequality [7, Lemma 7.2], for deterministic integrands and \( p \geq 2 \),
\[ (2.4) \quad \left\| \int_0^T \Phi_t \, dW_t \right\|_{L^p(\Omega; H)} \leq \frac{p(p-1)}{2} \left\| \Phi \right\|_{L^2(0, T; L^0_2)}^p, \quad \Phi \in L^2(0, T; L^0_2). \]
By taking \( H = \mathbb{R} \) and noting the isomorphisms \( H_0 \cong H_0^0 \cong L_2(H_0; \mathbb{R}) \) we see that a function \( \phi \in L^2(0, T; H_0) \) defines an integrand in \( L^2(0, T; L_2(H_0; \mathbb{R})) \) for the stochastic integral and the integral \( \int_0^T \phi_t \, dW_t \in L^2(\Omega) \) is real-valued. As \( L^p(0, T; H_0) \subset L^2(0, T; H_0) \) for \( p \geq 2 \) the stochastic integral is well defined for \( \phi \in L^p(0, T; H_0) \).

We now recall some concepts from Malliavin calculus introduced in [1]. For \( q \in [2, \infty] \) let \( S^q(\mathbb{R}) \) be the class of smooth cylindrical random variables of the form
\[ F = f \left( \int_0^T \phi_{1,s} \, dW_s, \ldots, \int_0^T \phi_{n,s} \, dW_s \right), \]
\[ f \in C_0^\infty(\mathbb{R}^n; \mathbb{R}), \ (\phi_k)_{k=1}^n \subset L^q(0, T; H_0), \ n \in \mathbb{N}. \]
For \( F \in S^q(\mathbb{R}) \) with the above representation, we define the Malliavin derivative
\[ (D_t F)_{t \in [0, T]} = \left( \sum_{j=1}^n \partial_j f \left( \int_0^T \phi_{1,s} \, dW_s, \ldots, \int_0^T \phi_{n,s} \, dW_s \right) \otimes \phi_{j,t} \right)_{t \in [0, T]}. \]
Let \( V \) be a separable Hilbert space. We define \( S^q(V) \) to be the space of all \( V \)-valued random variables of the form \( Y = \sum_{i=1}^m v_i \otimes F_i \) with \( (v_i)_{i=1}^m \subset V, \ (F_i)_{i=1}^m \subset S^q(\mathbb{R}), \ m \in \mathbb{N} \). The Malliavin derivative of \( Y \in S^q(V) \) of the above form is given by \( D_t Y = \sum_{i=1}^m v_i \otimes (D_t F_i) \). As \( (D_t F_i)_{t \in [0, T]} \) is an \( H_0 \)-valued process, \( (D_t Y)_{t \in [0, T]} \) is a \( V \otimes H_0 = L_2(H_0; V) \)-valued process.

For \( p \in [2, \infty), \ q \in [2, \infty], \ S^q(V) \subset L^p(\Omega; V) \) is dense by [1, Lemma 3.1] and the operator \( D: S^q(V) \to L^p(\Omega; L^q(0, T; L_2(H_0; V))) \) is closable by [1, Lemma 3.2]. Let \( M_1^{1-p,q}(V) \) denote the closure of \( S^q(V) \) with respect to the norm
\[ \|Y\|_{M_1^{1-p,q}(V)} = \left( \|Y\|_{L^p(\Omega; V)}^p + \|DY\|_{L^p(\Omega; L^q(0, T; L_2(H_0; V)))}^p \right)^{\frac{1}{p}}. \]
The spaces \( M_1^{1-p,q}(V) \) are Banach spaces, densely embedded into \( L^2(\Omega; V) \). Thus, \( M_1^{1-p,q}(V) \subset L^2(\Omega; V) \subset M_1^{1-p,q}(V)^* \) is a Gelfand triple. By [1, Theorem 3.5] the following inequality holds for \( p \in [2, \infty), \ q \in [2, \infty] \) with \( \frac{1}{q} + \frac{1}{q'} = 1 \):
\[ (2.5) \quad \left\| \int_0^T \Phi_t \, dW_t \right\|_{M_1^{1-p,q}(V)} \leq \left\| \Phi \right\|_{L^{q'}(0, T; L_2(H_0; V))}^{q'}, \quad \Phi \in L^2(0, T; L_2(H_0; V)). \]
What makes this duality theory useful is the possibility of taking \( q' \) close to 1, c.f., (2.4) where the exponent is 2. We only need (2.4) and (2.5) for deterministic integrands but remark that [1, Theorem 3.5] allows \( \Phi \) to be random and only Skorohod integrability is required. Following [1] we refer to \( M_1^{1-p,q}(H) \) for \( q > 2 \) as refined Sobolev-Malliavin spaces. The spaces \( M_1^{1-p,q}(V) \) are classical.
Sobolev-Malliavin spaces, often denoted $D^{1,p}(V)$. As in \cite{1} we also define the spaces $G^{1,p}(V) = M^{1,p}(V) \cap L^{2p}(\Omega; V)$, $p \geq 2$, equipped with the norm
\[
\| Y \|_{G^{1,p}(V)} = \max (\| Y \|_{L^{2p}(\Omega; V)}, \| Y \|_{M^{1,p}(V)}),
\]
and the corresponding Gelfand triple $G^{1,p}(V) \subset L^2(\Omega; V) \subset G^{1,p}(V)^\ast$. We next cite \cite[Lemma 3.9]{1}. It provides a local Lipschitz bound that enables us prove an error estimate in the $G^{1,p}(H)^\ast$-norm by a Gronwall argument in Lemma 4.6 below.

**Lemma 2.3.** Let $U, V$ be separable Hilbert spaces, $\sigma \in C^2_b(U; V)$, and $p \in [2, \infty)$. For $Y^1, Y^2 \in M^{1,2p}(U; V)$ it holds
\[
\| \sigma(Y^1) - \sigma(Y^2) \|_{G^{1,p}(V)} \leq \max (\| \sigma \|_{C^1_b(U; V)} \| \sigma \|_{C^2_b(U; V)}) \times \left( \sum_{i=1}^2 \| Y^i \|_{M^{1,2p}(U; V)} \right) \| Y^1 - Y^2 \|_{G^{1,p}(V)}.
\]

The next lemma is useful in the linearization step of our weak convergence proof. It can be extracted from \cite[Lemma 3.3]{1}, but we present a proof for the convenience of the reader.

**Lemma 2.4.** Let $U$ be a separable Hilbert space, let $p \in [2, \infty)$, $m \geq 2$, $\varphi \in C^2(U; R)$, and $Y \in M^{1,2(m-1)p}(U)$. Then $\varphi'(Y) \in G^{1,p}(U)$ and
\[
\| \varphi'(Y) \|_{G^{1,p}(U)} \leq 3 \max \left( \| \varphi \|_{C_b^1(U; R)}, \| \varphi \|_{C_b^2(U; R)} \right) \left( 1 + \| Y \|_{M^{1,2(m-1)p}(U)}^{m-1} \right).
\]

**Proof.** We must bound the norm of $\varphi'(Y)$ in $L^p(\Omega; H)$, $L^{2p}(\Omega; H)$, and the norm of $D\varphi'(Y)$ in $L^p(\Omega; L^p(0,T; L^2(H_0; U)))$. First, by (2.2), we have
\[
\| \varphi'(Y) \|_{L^p(\Omega; H)} \leq \| \varphi'(Y) \|_{L^{2p}(\Omega; H)} \leq \| \varphi \|_{C_b^1(U; R)} \left( 1 + \| Y \|_{L^{2p(m-1)}(\Omega; U)}^{m-1} \right)
\]
\[
\leq \| \varphi \|_{C_b^1(U; R)} \left( 1 + \| Y \|_{M^{1,2(m-1)p}(U; H)}^{m-1} \right).
\]

With $V = U$, $\sigma = \varphi$, $r = m - 1$ in \cite[Lemma 3.3]{1} it follows that $D\varphi'(Y) = \varphi''(Y)DY$. To bound $D\varphi'(Y)$ we use Hölder’s inequality with exponents $(m - 1)/(m - 2)$ and $m - 1$, and use $(1 + a^{m-2}a) \leq 1 + 2a^{m-1}$ for $a \geq 0$, to get
\[
\| D\varphi'(Y) \|_{L^p(\Omega; L^p(0,T; L^2(H_0; U)))}
\leq \left( E \left[ \| \varphi''(Y)DY \|_{L^p(0,T; L^2(H_0; U))}^p \right] \right)^{\frac{1}{p}}
\leq \| \varphi \|_{C_b^2(U; R)} \left( 1 + \| Y \|_{M^{1,2(m-1)p}(U; H)}^{m-1} \right) \| Y \|_{L^{2p(m-1)}(\Omega; L^p(0,T; L^2(H_0; U)))}
\leq 2 \| \varphi \|_{C_b^2(U; R)} \left( 1 + \| Y \|_{M^{1,2(m-1)p}(U; H)}^{m-1} \right) \| Y \|_{L^{2p(m-1)}(\Omega; L^p(0,T; L^2(H_0; U)))}
\leq 2 \| \varphi \|_{C_b^2(U; R)} \left( 1 + \| Y \|_{M^{1,2(m-1)p}(U; H)}^{m-1} \right) \| Y \|_{L^{2p(m-1)}(\Omega; L^p(0,T; L^2(H_0; U)))}.
\]

Combining the two bounds yields
\[
\| \varphi'(Y) \|_{M^{1,2p}(U; V)} \leq \max (\| \varphi \|_{C_b^1(U; R)}, \| \varphi \|_{C_b^2(U; R)}) \left( 2 + 3 \| Y \|_{M^{1,2(m-1)p}(U; V)}^{m-1} \right).
\]
Lemma 2.5. Let $p \in [2, \infty)$, $q \in [2, \infty]$. Then for all $S \in \mathcal{L}(H)$, $Y \in L^2(\Omega; H)$ it holds that
\[
\|SY\|_{M^{1,p,q}(H)^*} \leq \|S\|_{\mathcal{L}(H)}\|Y\|_{M^{1,p,q}(H)^*}.
\]

Proof. We compute by duality
\[
\|SY\|_{M^{1,p,q}(H)^*} = \sup_{\|Z\|_{M^{1,p,q}(H)} \leq 1} \langle SY, Z \rangle_{L^2(\Omega; H)}
\]
\[
= \|S^*\|_{\mathcal{L}(M^{1,p,q}(H)^*)} \sup_{\|Z\|_{M^{1,p,q}(H)} \leq 1} \left\langle Y, \frac{S^* Z}{\|S^*\|_{\mathcal{L}(M^{1,p,q}(H)^*)}} \right\rangle_{L^2(\Omega; H)}
\]
\[
\leq \|S^*\|_{\mathcal{L}(M^{1,p,q}(H)^*)} \sup_{\|Z\|_{M^{1,p,q}(H)} \leq 1} \langle Y, Z \rangle_{L^2(\Omega; H)}.
\]

Finally, we note that $\|S^*\|_{\mathcal{L}(M^{1,p,q}(H)^*)} \leq \|S^*\|_{\mathcal{L}(H)} = \|S\|_{\mathcal{L}(H)}$, because $DSY = SDY$ for $Y \in M^{1,p,q}(H)$. We omit the details. \qed

3. Existence, uniqueness and regularity

Throughout this section we assume that (1.1), (1.3)–(1.5) hold with $\rho \in [1,2)$, $\beta \in (0, 1/\rho]$. We begin by proving existence, uniqueness, and Malliavin regularity of the solution of (1.2). Recall that two stochastic processes $X^1, X^2$ are modifications of each other if for all $t \in [0,T]$ it holds that $\mathbb{P}(X^1_t \neq X^2_t) = 0$.

Proposition 3.1. There exists an, up to modification, unique stochastic process $X: [0, T] \times \Omega \to H$ such that $X \in \mathcal{C}(0,T; L^p(\Omega, H))$ for $p \in [2, \infty)$ and such that $X \in \mathcal{C}(0,T; M^{1,p,q}(H))$ for $p \in [2, \infty)$, $q \in [2, (2-\frac{2}{1-\rho})]$, and which satisfies equation (1.2) $\mathbb{P}$-a.s.

Proof. Existence is proved by a standard application of Banach’s Fixed Point Theorem, see, e.g., [13, Theorem 1] or [3, Theorem 3.3]. We note that for proving existence and uniqueness in $\mathcal{C}(0,T; L^p(\Omega, H))$ it is not crucial whether $(S_t)_{t \in [0,T]}$ is a semigroup or not. For the $\mathcal{C}(0,T; M^{1,p,q}(H))$ regularity, see Proposition 4.4 below. \qed

The next proposition states the temporal Hölder regularity of $X$ in the $L^p(\Omega, H)$- and $M^{1,p,q}(H)^*$-norms. Note that the Hölder exponent in the (weaker) $M^{1,p,q}(H)^*$-norm is twice that in the (stronger) $L^p(\Omega, H)$-norm.

Proposition 3.2. Let $X$ be the solution to (1.2). For $\gamma \in (0, \beta)$, $p \geq 2$, $q = \frac{2}{1-\rho\gamma}$, there exists $C$ such that
\[
\|X_{t_2} - X_{t_1}\|_{L^p(\Omega, H)} \leq C|t_2 - t_1|^\frac{\rho\gamma}{p}, \quad t_1, t_2 \in [0, T],
\]
\[
\|X_{t_2} - X_{t_1}\|_{M^{1,p,q}(H)^*} \leq C|t_2 - t_1|^{\rho\gamma}, \quad t_1, t_2 \in [0, T].
\]

Proof. Fix $\gamma \in (0, \beta)$, $p \geq 2$. In order to treat both norms simultaneously, we define $V_2 = L^p(\Omega, H)$, $c_{p,2} = p(p - 1)/2$, and $V_r = M^{1,p,r}(H)^*$, $c_{p,r} = 1$ for $r \in (2, \infty)$. In view of (2.4) and (2.5) it holds that
\[
\int_0^T \Phi_t dW_t \leq c_{p,r} \|\Phi\|_{L^p(0,T; L^2_r)}, \quad \Phi \in L^2(0,T; L^2_r), \quad r \in [2, \infty],
\]
and
where $\frac{1}{r} + \frac{1}{q} = 1$. Let $t_2 > t_1$. The difference $X_{t_2} - X_{t_1}$ can be written in the form

$$X_{t_2} - X_{t_1} = (S_{t_2} - S_{t_1})x_0 + \int_0^{t_1} (S_{t_2-s} - S_{t_1-s})F(X_s)\,ds + \int_{t_1}^{t_2} S_{t_2-s}F(X_s)\,ds + \int_0^{t_1} (S_{t_2-s} - S_{t_1-s})\,dW_s + \int_{t_1}^{t_2} S_{t_2-s}\,dW_s.$$

Taking $V_r$-norms, using the continuous embeddings $H \subset L^p(\Omega; H) \subset L^2(\Omega; H) \subset M^{1,p,r}(H)^*$, yields

$$\|X_{t_2} - X_{t_1}\|_{V_r} \leq \|(S_{t_2} - S_{t_1})x_0\| + \left\| \int_0^{t_1} (S_{t_2-s} - S_{t_1-s})F(X_s)\,ds \right\|_{L^p(\Omega; H)} + \left\| \int_{t_1}^{t_2} S_{t_2-s}F(X_s)\,ds \right\|_{L^p(\Omega; H)} + \left\| \int_0^{t_1} (S_{t_2-s} - S_{t_1-s})\,dW_s \right\|_{V_r} + \left\| \int_{t_1}^{t_2} S_{t_2-s}\,dW_s \right\|_{V_r}.$$

First, by (1.1) and (1.5), we obtain

$$\|(S_{t_2} - S_{t_1})x_0\| = \left\| \int_{t_1}^{t_2} \dot{S}_t A^{-\frac{\beta}{2}} A^{\frac{\beta}{2}} x_0 \,dt \right\| \leq L_0 \left\| A^{\frac{\beta}{2}} x_0 \right\| (t_2 - t_1).$$

It is straightforward to show that the terms containing $F$ are bounded up to a constant by $|t_2 - t_1|^{1-r}$, and $|t_2 - t_1|$ respectively, for every $\epsilon \in (0, 1)$. For the case $p = 1$, see the proof of [1, Proposition 3.11]; the remaining case is treated similarly.

By (3.1), (1.3), and (1.1) we get

$$\left\| \int_0^{t_1} (S_{t_2-s} - S_{t_1-s})\,dW_s \right\|_{V_r} \leq c_{p,r} \left( \int_0^{t_1} \left\| (S_{t_2-s} - S_{t_1-s}) A^{\frac{\beta}{2p}} \right\|^r \left\| A^{\frac{\beta}{2p}} \right\|^{r'} \,ds \right)^{\frac{1}{r'}} \leq c_{p,r} \left\| A^{\frac{\beta-1}{2p}} \right\|_{L^p} \left( \int_0^{t_1} \left( \int_{t_1}^{t_2} \dot{S}_s A^{(3-\beta\rho)/2-1} \,ds \right)^{r'} \,ds \right)^{\frac{1}{r'}} \leq c_{p,r} \left\| A^{\frac{\beta-1}{2p}} \right\|_{L^p} L^{2-\beta\rho} \left( \int_0^{t_1} \left( \int_{t_1}^{t_2} (t - s)^{\frac{3-\beta\rho}{2}} \,ds \right)^{r'} \,ds \right)^{\frac{1}{r'}}.$$

Bounding the integrals yields, for $\eta \in (0, 1/p)$ to be chosen,

$$\left( \int_0^{t_1} \left( \int_{t_1}^{t_2} (t - s)^{-\frac{3-\beta\rho}{2}} \,ds \right)^{r'} \,ds \right)^{\frac{1}{r'}} \leq \left( \int_0^{t_1} (t_1 - s)^{-\frac{1-(\beta-2\eta)\rho}{2}} \int_{t_1}^{t_2} (t - t_1)^{-1+\eta\rho} \,ds \right)^{\frac{1}{r'}} = (t_2 - t_1)^{\eta\rho} \left( \int_0^{t_1} (t_1 - s)^{-\frac{1-(\beta-2\eta)\rho}{2}} \,ds \right)^{\frac{1}{r'}}.$$

For $r = q = 2/(1 - \gamma \rho)$ and $\eta < (\beta + \gamma)/2$, the exponent is

$$\frac{r}{r - 1} \frac{1 - (\beta - 2\eta)\rho}{2} = 1 - \frac{\beta \rho + 2\eta \rho}{1 + \rho \gamma} < 1.$$
In particular, we can take $\eta = \gamma$ as required since $\gamma < \beta$. For $r = 2$, the analogous condition is $\eta < \beta/2$ and we can take $\eta = \gamma/2$. Next, similarly,

$$\left\| \int_{t_1}^{t_2} S_{t_2-s} \, dW_s \right\|_{V_n} \leq c_{p,r} \left( \int_{t_1}^{t_2} \left\| S_{t_2-s} A^{\frac{\beta_2}{\alpha_2}} \right\|_2 \left\| A^{\frac{\beta_1-1}{\alpha_1}} r' \right\|_{L^2_0} \, ds \right)^{\frac{2}{r}}$$

$$\leq c_{p,r} L \left\| A^{\frac{\beta_2}{\alpha_2}} \right\|_2 \left( \int_{t_1}^{t_2} (t_2 - s)^{-\frac{1 - \beta_2}{\alpha_1}} \, ds \right)^{\frac{r-1}{r}} \leq (t_2 - t_1)^{\frac{2}{r} - \frac{1 - \beta_2}{\alpha_1}}.$$

For $r = q = 2/(1 - \gamma \rho)$ we have the Hölder exponent

$$\frac{r-1}{r} - \frac{1 - \beta_2}{2} = \frac{\rho(\beta + \gamma)}{2} > \gamma \rho,$$

and for $r = 2$ the Hölder exponent equals $\beta \rho/2 > \gamma \rho/2$. \hfill \Box

4. WEAK AND STRONG CONVERGENCE

This section contains our main result and its proof. Theorem 4.7 states a weak error estimate for approximations of $\int_0^T X_t \, d\nu_t$ for $\nu \in \mathcal{M}_T$, where $X$ is the solution to (1.2), and Theorem 4.2 provides a strong error estimate for approximations of $X$. For parabolic problems weak convergence of approximations of $X$ for $t \in [0, T]$ has been considered [1], and for Volterra equations in [15] but only in the linear case $F = 0$. To the best of our knowledge, weak convergence of $\int_0^T X_t \, d\nu_t$ is new in both cases. The weak rate is twice the strong rate as expected.

4.1. Approximation. The following assumptions are justified in Section 5. Assume that (1.1), (1.3)–(1.5) hold. Let $(V_n)_{n \in (0,1)}$ be a family of finite-dimensional subspaces of $H$ and let $P_h : H \rightarrow V_h$ be the orthogonal projector. Let $k \in (0, 1)$ and $t_n = nk, n = 0, \ldots, N$, where $t_N < T \leq t_N + k$. Let $(B_{h,k})_{h,k \in (0,1)}$ be a family of operator-valued functions $B_{h,k} : \{0, \ldots, N\} \rightarrow \mathcal{L}(H; V_h)$ such that $B_{h,k}^n = B_{h,k}^0 P_h$, and let $(A_{h,k})_{h,k \in (0,1)}$ be a collection of linear operators $A_h : V_h \rightarrow V_h$ such that for $n = 1, \ldots, N$ it holds

$$\left\| A_{h,k}^n B_{h,k}^n x \right\| \leq L_s t_n^{-s} \| x \|, \quad x \in H, \quad 0 \leq s \leq 1,$$

with the same constants $(L_s)_{s \in [0,1]}$ as in (1.1). For other constants $(K_\epsilon)_{\epsilon \in (0,\infty)}$ and $(R_\epsilon)_{\epsilon \in [0,1]}$, let the corresponding error operator $(E_{h,k})_{h,k \in (0,1)}$, given by $E_{h,k} = S_{t_n} - B_{h,k}^n$ for $n = 0, \ldots, N$, satisfy the smooth data error estimate

$$\left\| E_{h,k} x \right\| \leq K_\epsilon (h^s + k^{2s}) \| x \|_{H^{s(1+\epsilon)}}, \quad 0 \leq s \leq 2, \quad \epsilon > 0,$$

and the non-smooth data error estimates, for $n = 1, \ldots, N, \tau > 0$,

$$\left\| A_{h,k}^\tau E_{h,k}^n x \right\| \leq R_s (h_\tau + k_\tau) t_n^{-\frac{\alpha_\tau}{\alpha_1}}, \quad 0 \leq s \leq 2, \quad 0 \leq \tau \leq 1 - s/2,$$

$$\left\| (e^{-tA} - e^{-tA_h} P_h) x \right\| \leq R_0 h^s t^{-\frac{\alpha_\tau}{\alpha_1}} \| x \|, \quad 0 \leq s \leq 2,$$

where $(e^{-tA})_{t \in [0,\infty)}$ and $(e^{-tA_h})_{t \in [0,\infty)}$ are the analytic semigroups generated by $-A$ and $-A_h$, respectively. We introduce the piecewise continuous operator function $E_{h,k}^n : [0, T] \rightarrow \mathcal{L}$ given by $E_{h,k}^n = S_t - B_{h,k}^n$ for $t \in [t_n, t_{n+1})$ and $n = 0, \ldots, N - 1$. By (1.1) and (4.3), the family $(E_{h,k}^n)_{t \in [0,T]}$ satisfies for $t \in (0, T]$ the bound

$$\left\| A_{h,k}^\tau E_{h,k}^n \right\| \leq R_s (h_\tau + k_\tau) t_n^{-\frac{\alpha_\tau}{\alpha_1}}, \quad 0 \leq s \leq 2, \quad 0 \leq \tau \leq 1 - s/2.$$

The discrete and continuous stochastic convolutions are defined by

\[ W^S_t = \int_0^t S_{t-s} \, dW_s, \quad t \in [0,T]; \quad W^B_{n,h,k} = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} \, dW_t, \quad n = 1, \ldots, N. \]

We now define approximations of equation (1.2). For \( h, k \in (0,1) \), let \( (X^{h,k}_n)^N_{n=0} \) be the solution to the equation

\[ X^{h,k}_n = B_{n,h,k} x_0 + k \sum_{j=1}^{n-1} B_{n-j}^{h,k} F(X^{h,k}_j) + W^B_{n,h,k}, \quad n = 1, \ldots, N. \]

### 4.2. Strong convergence

Boundedness in the \( L^p(\Omega; H) \)-sense of the approximate family \( (X^{h,k}_n)^N_{n=0} \) is stated in the next proposition. For a proof in the parabolic case, i.e., for \( \rho = 1 \), see [1, Proposition 3.15]. The general case is proved in the same way but using the different smoothing property in (4.1).

**Proposition 4.1.** Let the setting of Subsection 4.1 hold. For \( p \geq 2 \) it holds

\[ \sup_{h,k \in (0,1)} \max_{n \in \{0, \ldots, N\}} \|X^{h,k}_n\|_{L^p(\Omega; H)} < \infty. \]

We next prove strong convergence. This is interesting in itself, but it is also used in our proof of the Malliavin regularity of \( X \) in Proposition 4.4.

**Theorem 4.2.** Let the setting of Subsection 4.1 hold, let \( X \) be the solution to (1.2) and let \( (X^{h,k})_{h,k \in [0,1]} \) be the solutions to (4.6). For \( \gamma \in [0, \beta] \), \( p \in [2, \infty) \), there exists \( C \) such that

\[ \max_{n \in \{0, \ldots, N\}} \|X_n - X^{h,k}_n\|_{L^p(\Omega; H)} \leq C(\|\gamma\| + k^2), \quad h, k \in (0,1). \]

**Proof.** We take the difference of (1.2) and (4.6) to obtain the equation for the error,

\[ X_{t_n} - X^{h,k}_n = (S_{t_n} - B_{n,h,k}) x_0 + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} (S_{t_n-t} - B_{n-j}^{h,k}) F(X_t) \, dt \]

\[ + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} (F(X_t) - F(X^{h,k}_j)) \, dt + W^S_{t_n} - W^B_{n,h,k}. \]

The deterministic nature of the first two terms allows us to obtain twice the rate of convergence compared to the other terms. This will be used later in the proof of Lemma 4.6. Recall that \( \tilde{E}^{h,k}_t = S_t - B_{n,h}^{h,k} \) for \( t \in [t_n, t_{n+1}] \) and \( n = 0, \ldots, N - 1 \). We get

\[ \|X_{t_n} - X^{h,k}_n\|_{L^p(\Omega; H)} \leq \|E_{t_n}^{h,k} x_0\|_H + \|\int_0^{t_n} \tilde{E}_{t_n-t}^{h,k} F(X_t) \, dt\|_{L^p(\Omega; H)} \]

\[ + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} (F(X_t) - F(X^{h,k}_j)) \, dt\|_{L^p(\Omega; H)} \]

\[ + \|W^S_{t_n} - W^B_{n,h,k}\|_{L^p(\Omega; H)}. \]

Using (1.5), (4.2) with \( \sigma = 2\rho_\gamma, \epsilon = (3 - 2\gamma\rho)/2\gamma\rho \) we obtain

\[ \max_{n \in \{0, \ldots, N\}} \|E_{t_n}^{h,k} x_0\| \leq K_2 \cdot 2^{2\rho_\gamma} (h^{2\rho_\gamma} + k^{\rho_\gamma}) \|x_0\|_H^3. \]
By (1.4), (4.5), and Proposition 3.1, it holds that
\[
\left\| \int_0^{t_n} \tilde{E}_{t_n-t}^{h,k} F(X_t) \, dt \right\|_{L^p(\Omega; H)} \\
\leq \int_0^{t_n} \left\| \tilde{E}_{t_n-t}^{h,k} \right\|_\mathcal{L} \left\| F(X_t) \right\|_{L^p(\Omega; H)} \, dt \\
\leq R_0 (h^{2\gamma} + k^{\rho \gamma}) |F|_{C^\gamma_0(H;H)} \left( 1 + \sup_{t \in [0,T]} \| X_t \|_{L^p(\Omega; H)} \right) \int_0^{t_n} (t_n - t)^{-\rho \gamma} \, dt \\
\lesssim h^{2\gamma} + k^{\rho \gamma}.
\]
Using (1.4), (2.3), (4.1), and Proposition 3.2 yields
\[
\left\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} (F(X_t) - F(X_j^{h,k})) \, dt \right\|_{L^p(\Omega; H)} \\
\leq \| F \|_{C^\gamma_0(H;H)} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \left\| B_{n-j}^{h,k} \right\|_\mathcal{L} \left\| X_t - X_j^{h,k} \right\|_{L^p(\Omega; H)} \, dt \\
\leq L_0 \| F \|_{C^\gamma_0(H;H)} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \left( \left\| X_t - X_j \right\|_{L^p(\Omega; H)} + \left\| X_t - X_j^{h,k} \right\|_{L^p(\Omega; H)} \right) \, dt \\
\leq L_0 \| F \|_{C^\gamma_0(H;H)} \left( C T k^{\frac{2\gamma}{p}} + k \sum_{j=0}^{n-1} \left\| X_t - X_j^{h,k} \right\|_{L^p(\Omega; H)} \right).
\]
For the error of the stochastic convolution, we write the difference in the form
\[
W^S_{t_n} - W^B_{t_n} = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} (S_{t_n-t} - B_{n-j}^{h,k}) \, dW_t \\
= \int_0^{t_n} \tilde{E}_{t_n-t}^{h,k} \, dW_t = \int_0^{t_n} \tilde{E}_{t}^{h,k} \, dW_t.
\]
By (2.4) and (4.5) with \( \sigma = \gamma \rho \), and \( s = 1 - \beta \rho \), we obtain the estimate
\[
\left\| W^S_{t_n} - W^B_{t_n} \right\|_{L^p(\Omega; H)} \leq \left( p(p-1) - \frac{1}{2} \int_0^{t_n} \left\| A^{\frac{\beta \rho - 1}{2}} \right\|^2_\mathcal{L} \left\| A^{\frac{1-\beta \rho}{2}} \tilde{E}_t^{h,k} \right\|^2_\mathcal{L} \, dt \right)^{\frac{1}{2}} \\
\lesssim R_{1-\beta \rho} \left( \int_0^{t_n} t^{p(\beta - \gamma) - 1} \, dt \right)^{\frac{1}{2}} (h^{\gamma} + k^{\frac{2\gamma}{p}}) \lesssim h^{\gamma} + k^{\frac{2\gamma}{p}}.
\]
Collecting the estimates yields that, for all \( n = 0, \ldots, N \), it holds
\[
\left\| X_{t_n} - X_n^{h,k} \right\|_{L^p(\Omega; H)} \lesssim h^{\gamma} + k^{\frac{2\gamma}{p}} + k \sum_{j=0}^{n-1} \left\| X_{t_j} - X_j^{h,k} \right\|_{L^p(\Omega; H)}.
\]
The proof is completed by Gronwall’s lemma.

4.3. Regularity and weak convergence. Here we state and prove our main result on weak convergence. It is based on a strong error estimate in the \( \mathbf{G}^{1,p}(H)^* \)-norm combined with boundedness of \( X \) and \( X^{h,k} \) in \( \mathbf{M}^{1,p,q}(H) \) for suitable \( p, q \). The methodology was introduced in [1], but here we exploit it further to obtain more general type of convergence, namely in the \( e_{\varphi,\psi} \)-distances defined in (1.8). We begin by proving the Malliavin differentiability of \( X^{h,k} \).
Proposition 4.3. Let the setting of Subsection 4.1 hold, and let $X^{h,k}$ be the solution to (4.6). For $p \in [2, \infty)$, $q \in [2, \frac{2}{1+p/2q})$, it holds

$$\sup_{h,k \in (0,1)} \max_{n \in \{0, \ldots, N\}} \|X^{h,k}_n\|_{M^{1,p,q}(H)} < \infty.$$ 

Sketch of proof. Note first that $DX^0_{0} = 0$ as $X^0_{0}$ is deterministic. Therefore it follows inductively that $X^j_{0}$, $j = 0, \ldots, N$, are differentiable and the derivative satisfies the equation

$$D_rX^{h,k}_{n} = k \sum_{j=0}^{n-1} B_{n-j}^{h,k}(X^{h,k}_{j})D_rX_{j}^{h,k} + \sum_{j=0}^{n-1} \chi_{(t_j, t_{j+1})}(r)B_{n-j}^{h,k}.$$ 

The proof is performed by straightforward analysis of this equation using the discrete Gronwall’s lemma, see [1, Proposition 3.16] for details in the parabolic case $\rho = 1$. The general case is treated analogously. □

The Malliavin regularity of $X$ is next obtained by a limiting procedure.

Proposition 4.4. Let the setting of Subsection 4.1 hold and let $X$ be the solution to (1.2). For $p \in [2, \infty)$, $q \in [2, \frac{2}{1+p/2q})$, it holds that $X \in \mathcal{C}(0, T; M^{1,p,q}(H))$.

Proof. Let $\tilde{X}^{h,k}_n = X^{h,k}_n$ for $t \in [t_n, t_{n+1})$, $n = 0, \ldots, N-1$, $h,k \in (0,1)$. By Proposition 4.3 it holds in particular, that the family $(\tilde{X}^{h,k}_{h,k \in (0,1)})$ is bounded in the Hilbert space $\mathcal{X} = L^2(0, T; M^{1,2,2}(H))$, and by Theorem 4.2 it holds that $\tilde{X}^{h,k} \rightarrow X$ as $h,k \rightarrow 0$ in the Hilbert space $\mathcal{Y} = L^2(0, T; L^2(\Omega; H))$. Lemma 2.1 applies and ensures that $X \in \mathcal{X} = L^2(0, T; M^{1,2,2}(H))$.

To show that $X \in \mathcal{C}(0, T; M^{1,p,q}(H))$, we argue as follows. By [10, Lemma 3.6] it holds that also $\int_0^t S_{t-s} F(X_s) \,ds \in L^2(0, T; M^{1,2,2}(H))$ with $D_r \int_0^t S_{t-s} F(X_s) \,ds = \int_0^t S_{t-s} F'(X_s)D_rX_s \,ds$, for $0 \leq r \leq t \leq T$, and $\int_0^t S_{t-s} dW_s \in L^2(0, T; M^{1,2,2}(H))$ with $D_r \int_0^t S_{t-s} dW_s = S_{t-r}$, for $0 \leq r \leq t \leq T$. We remark that [10, Lemma 3.6] is formulated for semigroups, but the semigroup property is not used in the proof. We have thus proved that we can differentiate the equation for $X$ term by term, and obtain the equation

$$D_rX_t = \begin{cases} S_{t-r} + \int_0^t S_{t-s} F'(X_s)D_rX_s \,ds, & t \in (r, T], \\ 0, & t \in [0, r]. \end{cases}$$

A straightforward analysis of this equation, by a Gronwall argument, as in the proof of [1, Proposition 3.10] completes the proof that $X \in \mathcal{C}(0, T; M^{1,p,q}(H))$. □

In the proof of [1, Lemma 4.6], which is the analogue of Lemma 4.6 below, a bound

$$\|A^{1/2}_h P_h x\| \leq \|A^{1/2}_h P_h A^{-1/2} \|_E \|A^{-1/2} x\| \leq C \|A^{-1/2} x\|,$$

was used in the special case $\delta = 1$. This estimate is true for all $\delta \in [0, 1]$ for both the finite element method and for spectral approximation. For $\delta > 1$ it holds only for spectral approximation. In this paper we need $\delta \in [0, 2/\rho)$ and therefore we cannot rely on (4.12). In [21, Lemma 5.3] it is shown that for finite element discretization and for $\delta = 0, 1, 2$ it holds

$$\|A^{1/2}_h P_h x\| \leq C(\|A^{-1/2} x\| + h^\delta \|x\|), \quad x \in H.$$
The next lemma is a generalization of this result, assuming the availability of a non-smooth data error estimate for spatial approximation of the semigroup generated by $-A$, see (4.4). It will be used in the proof of Lemma 4.6 below with $X = G^{1,p}(H)^*$. In this way we need not rely on (4.12) and we include finite element discretization under the same generality as spectral approximations.

**Lemma 4.5.** Let the setting of Subsection 4.1 hold and let $X$ be a Banach space such that the embedding $L^2(\Omega; H) \subset X$ is continuous. For $\kappa \in [0, 2]$, $\sigma \in [0, \kappa)$, there exists $C$ such that for $Y \in L^2(\Omega; H)$ it holds

$$\|A_h^{-\frac{\kappa}{2}} P_h Y\|_X \leq \|A^{-\frac{\kappa}{2}} Y\|_X + C h^\sigma \|Y\|_{L^2(\Omega; H)}, \quad h \in (0, 1).$$

**Proof.** By the continuous embedding $L^2(\Omega; H) \subset X$ we get that

$$\|A_h^{-\frac{\kappa}{2}} P_h Y\|_X \leq \|A^{-\frac{\kappa}{2}} Y\|_X + \|(A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}}) Y\|_X \lesssim \|A^{-\frac{\kappa}{2}} Y\|_X + \|A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}}\|_{L^\infty} \|Y\|_{L^2(\Omega; H)}.$$  

By [18, Chapter 2, (6.9)] we have

$$A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}} = \frac{1}{\Gamma(\kappa/2)} \int_0^\infty t^{\frac{\kappa}{2}-1} (e^{-tA_h} P_h - e^{-tA}) dt.$$  

Therefore, by (4.4),

$$\|A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}}\|_{L^\infty} \leq \frac{1}{\Gamma(\kappa/2)} \int_0^\infty t^{\frac{\kappa}{2}-1} \|e^{-tA_h} P_h - e^{-tA}\|_{L^\infty} dt \lesssim \int_0^{h^{-2}} t^{\frac{\kappa}{2}-1} \|e^{-tA_h} P_h - e^{-tA}\|_{L^\infty} dt + \int_h^{\infty} t^{\frac{\kappa}{2}-1} \|e^{-tA_h} P_h - e^{-tA}\|_{L^\infty} dt \lesssim h^{\frac{\kappa}{2}+\kappa-1} dt + h^2 \int_h^{\infty} t^{\frac{\kappa}{2}-2} dt = \frac{Ah^\sigma}{\kappa - \sigma} + \frac{2}{2 - \kappa} h^2 h^{\frac{\kappa}{2}-\frac{\kappa}{2}} \lesssim h^\sigma.$$  

\[\square\]

The next result is a strong error estimate in the (weak) $G^{1,p}(H)^*$-norm. Together with the regularity stated in Propositions 4.3 and 4.4 it is the key to the proof of Theorem 4.7 below on weak convergence.

**Lemma 4.6.** Let the setting of Subsection 4.1 hold, and let $X$ and $X^{h,k}$ be the solutions to (1.2) and (4.6), respectively. For $\gamma \in [0, \beta]$, $p = \frac{2}{1 - \gamma}$, there exists $C$ such that

$$\max_{n \in \{0, \ldots, N\}} \|X_{t_n} - X_{t_n}^{h,k}\|_{G^{1,p}(H)^*} \leq C (h^{2\gamma} + k^{\rho\gamma}), \quad h, k \in (0, 1).$$

**Proof.** The proof is similar to that of Theorem 4.2. By (4.7) and the continuous embeddings $H \subset L^p(\Omega; H) \subset L^2(\Omega; H) \subset M^{1,p}(H)^* \subset G^{1,p}(H)^*$, it follows that

$$\|X_{t_n} - X_{t_n}^{h,k}\|_{G^{1,p}(H)^*} \leq \|\tilde{E}_{t_n}^{h,k} X_0\|_{H} + \left\| \int_0^{t_n} \tilde{E}_{t_n-t}^{h,k} F(X_t) dt \right\|_{L^p(\Omega; H)} + \left\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} (F(X_t) - F(X_{t_j}^{h,k})) dt \right\|_{G^{1,p}(H)^*} + \left\| W_{t_n}^{h} - W_{t_n}^{h,k} \right\|_{M^{1,p}(H)^*}.$$
The first two terms was already estimated as desired in (4.8) and (4.9). Choose \( \kappa \) so that \( \max(\delta, 2\gamma) < \kappa < \frac{2}{\rho} \), where \( \delta \) is the parameter in (1.4). Since \( \rho \kappa < 2 \), we have, by Lemma 2.5 and (4.1) with \( s = \rho \kappa/2 \), that

\[
\left\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{-h,k} (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H^*)} \\
\leq \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \left\| B_{n-j}^{-h,k} \tilde{A} \delta P_h \right\|_L \left\| A_h^{-\frac{\sigma}{2}} P_h (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H^*)} dt \\
\leq L_{n-h} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \left\| A_h^{-\frac{\sigma}{2}} P_h (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H^*)} dt.
\]

Applying Lemma 4.5 with \( X = G^{1,p}(H^*) \) and \( \sigma = 2\gamma < \kappa \) yields

\[
\left\| A_h^{-\frac{\sigma}{2}} P_h (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H^*)} \leq C h^{2\gamma} \left\| F(X_t) - F(X_j^{h,k}) \right\|_{L^2(\Omega; H)} \\
+ \left\| A^{-\frac{\sigma}{2}} (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H^*)}.
\]

For the first term we get by (1.4), Propositions 3.1, and 4.1 that

\[
\sup_{t \in [0,T]} \max_{j \in \{0, \ldots, N\}} \left\| F(X_t) - F(X_j^{h,k}) \right\|_{L^2(\Omega; H)} \\
\leq \left\| F \right\|_{c_{1}(\Omega; H^*)} \left( \sup_{t \in [0,T]} \left| X_t \right|_{L^2(\Omega; H)} + \max_{j \in \{0, \ldots, N\}} \left| X_j^{h,k} \right|_{L^2(\Omega; H)} \right) < \infty.
\]

By duality in the Gelfand triple \( G^{1,p}(H^{-\delta}) \subset L^2(\Omega; H^{-\delta}) \subset G^{1,p}(H^{-\delta})^{*} \) we compute that for \( Y \in L^2(\Omega; H^{-\delta}) \),

\[
\left\| Y \right\|_{G^{1,p}(H^{-\delta})^*} = \sup_{Z \in G^{1,p}(H^{-\delta})} \left\langle Z, Y \right\rangle_{L^2(\Omega; H^{-\delta})} = \sup_{Z \in G^{1,p}(H^{-\delta})} \left\langle Z, A^{-\frac{\sigma}{2}} Y \right\rangle_{L^2(\Omega; H)} \\
= \sup_{Z \in G^{1,p}(H)} \left\| A^{-\frac{\sigma}{2}} Z \right\|_{G^{1,p}(H^{-\delta})^*} = \sup_{Z \in G^{1,p}(H)} \left\| Z \right\|_{G^{1,p}(H)} = \left\| A^{-\frac{\sigma}{2}} Y \right\|_{G^{1,p}(H)^*}.
\]

Therefore, by Lemma 2.5 and Lemma 2.3 applied with \( U = H, V = H^{-\delta}, \sigma = F \), and by the continuous embedding \( M^{1,p}(H)^* \subset G^{1,p}(H)^* \) we get

\[
\left\| A^{-\frac{\sigma}{2}} (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H)^*} \\
\leq \left\| A^{-\frac{\sigma}{2}} \right\|_L \left\| A^{-\frac{\sigma}{2}} (F(X_t) - F(X_j^{h,k})) \right\|_{G^{1,p}(H)^*} \\
= \left\| A^{-\frac{\sigma}{2}} \right\|_L \left\| F(X_t) - F(X_j^{h,k}) \right\|_{G^{1,p}(H^{-\delta})^*} \\
\leq \left\| A^{-\frac{\sigma}{2}} \right\|_L \max \left( \left\| F \right\|_{c_1(\Omega; H^{-\delta})}, \left\| F \right\|_{c_2(\Omega; H^{-\delta})} \right) \\
\times \left( \sup_{j \in \{0, \ldots, N\}} \left\| X_j^{h,k} \right\|_{M^{1,2,p}(H)} + \sup_{t \in [0,T]} \left\| X_t \right\|_{M^{1,2,p}(H)} \right) \\
\times \left( \left\| X_t - X_{t_j} \right\|_{M^{1,1,p}(H)} + \left\| X_{t_j} - X_j^{h,k} \right\|_{G^{1,p}(H)^*} \right).
\]
By Propositions 3.2 and 4.3 and Proposition 4.4, we conclude
\[ \| A^{-\frac{2}{3}} (F(X_t) - F(X_{j,h,k})) \|_{\mathbf{G}^{1, p(H^*)}} \lesssim h^{2\gamma} + \| X_{t_j} - X_{j,h,k} \|_{\mathbf{G}^{1, p(H^*)}}. \]

Thus,
\[
\left\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} (F(X_t) - F(X_{j,h,k})) \, dt \right\|_{\mathbf{G}^{1, p(H^*)}} \\
\lesssim h^{2\gamma} + k^{\rho \gamma} + k \sum_{j=0}^{n-1} t_{n-j}^{-\frac{a p}{2}} \| X_{t_j} - X_{j,h,k} \|_{\mathbf{G}^{1, p(H^*)}}.
\]

By (4.10), (2.5), and (4.5) with \( s = 1 - \beta \rho, \sigma = 2 \gamma \rho \), and since \( p = \frac{2}{1+\rho \gamma} \) and \( p' = \frac{2}{1+\rho \gamma} \), we get
\[
\left\| W_{t_n}^S - W_n^{h,k} \right\|_{\mathbf{M}^{1, p, p'(H)}} \leq \left( \int_0^{t_n} \left\| A^{\frac{3p-1}{2p}} \| A^{\frac{1-\rho \gamma}{2p}} E_t^{h,k} \|_H^{\frac{1+\rho \gamma}{2p}} \, dt \right\|_{L_2^p} \right)^{\frac{1}{2}} \\
\leq R_{1-\beta \rho} \left\| A^{\frac{3p-1}{2p}} \| A^{\frac{1-\rho \gamma}{2p}} \|_H^{\frac{1+\rho \gamma}{2p}} \right\|_{L_2^p} \left( \int_0^{t_n} t^{\frac{a p}{2} - 1} \, dt \right)^{\frac{1}{2}} (h^{2\gamma} + k^{\rho \gamma}).
\]

Altogether we have that for every \( n = 1, \ldots, N \) it holds that
\[
\left\| X_{t_n} - X_n^{h,k} \right\|_{\mathbf{G}^{1, p(H^*)}} \lesssim h^{2\gamma} + k^{\rho \gamma} + k \sum_{j=0}^{n-1} t_{n-j}^{-\frac{a p}{2}} \| X_{t_j} - X_{j,h,k} \|_{\mathbf{G}^{1, p(H^*)}}.
\]

The proof is finished by an application of Lemma 2.2.

We next state our main result on weak convergence. Recall the definition of the distance \( e_{\varphi, \nu} \) in (1.8). We remark that to our knowledge, all previous weak convergence results concern the case \( \nu = \delta_s \) for fixed \( \tau \in (0, T) \), corresponding to convergence of \( \| E[\varphi(X_{t}^{h,k}) - \varphi(X_t)] \| \).

**Theorem 4.7.** Let \( X \) be the solution to (1.2) and \( (X_{h,k})_{h,k \in (0,1)} \) be the family of the solutions to (4.6). Define \( \tilde{X}_{t}^{h,k} = X_n^{h,k} \) for \( t \in [t_n, t_{n+1}) \), \( n \in \{0, \ldots, N-1\} \) and \( \tilde{X}_{t}^{h,k} = X_n^{h,k} \) for \( t \in [t_N, T] \). For \( m \geq 2 \), \( \varphi \in C^{2,m}_p(H; \mathbb{R}), \nu \in \mathcal{M}_T, \gamma \in [0, \beta) \), there exists \( C \) such that
\[
e_{\varphi, \nu}(X, \tilde{X}_{t}^{h,k}) \leq C (h^{2\gamma} + k^{\rho \gamma}), \quad h, k \in (0, 1).
\]

**Proof.** By the Mean Value Theorem (2.3) we get
\[
e_{\varphi, \nu}(X, \tilde{X}_{t}^{h,k}) \\
= \left( \int_0^T \varphi' \left( \int_0^T [\tilde{X}_{t}^{h,k} + \lambda (X_t - \tilde{X}_{t}^{h,k})] \, d\nu_t \right) \, d\lambda_t \int_0^T (X_t - \tilde{X}_{t}^{h,k}) \, d\nu_t \right)_{L_2^p(\Omega; H^*)}.
\]
By duality in the Gelfand triple $G^{1,p}(H) \subset L^2(\Omega ; H) \subset G^{1,p}(H)^*$ we obtain

$$|e_{\varphi,\nu}(X, \tilde{X}^{h,k})| \leq \left\| \int_0^1 \varphi' \left( \int_0^T [\tilde{X}_t^{h,k} + \lambda (X_t - \tilde{X}_t^{h,k})] \, dt \right) \, d\lambda \right\|_{G^{1,p}(H)}$$

$$\times \left\| \int_0^T (X_t - \tilde{X}_t^{h,k}) \, dt \right\|_{G^{1,p}(H)^*} \leq \sup_{\lambda \in [0,1]} \left\| \varphi' \left( \int_0^T [\tilde{X}_t^{h,k} + \lambda (X_t - \tilde{X}_t^{h,k})] \, dt \right) \right\|_{G^{1,p}(H)}$$

$$\times \left\| X - \tilde{X}^{h,k} \right\|_{L^1(0,T;G^{1,p}(H)^*)}.$$

By Lemma 2.4, Proposition 4.3, and Proposition 4.4 it holds

$$\sup_{\lambda \in [0,1]} \left\| \varphi' \left( \int_0^T [\tilde{X}_t^{h,k} + \lambda (X_t - \tilde{X}_t^{h,k})] \, dt \right) \right\|_{G^{1,p}(H)}$$

$$\lesssim 1 + \left\| X \right\|_{L^{1,m-1}(0,T;M^{1,2(m-1)p,p}(H))}^{m-1} + \left\| \tilde{X}^{h,k} \right\|_{L^{1,m-1}(0,T;M^{1,2(m-1)p,p}(H))} < \infty.$$

Finally, Proposition 3.2 and Lemma 4.6 give

$$\left\| X - \tilde{X}^{h,k} \right\|_{L^1(0,T;G^{1,p}(H)^*)} \leq \left\| X - \tilde{X} \right\|_{L^1(0,T;G^{1,p}(H)^*)} + \left\| \tilde{X} - \tilde{X}^{h,k} \right\|_{L^1(0,T;G^{1,p}(H)^*)} \lesssim h^{2\gamma} + k^{\rho\gamma}.$$

This completes the proof.

5. Examples

In this section we consider two different types of equations and write them in the abstract form of Section 1. We verify the abstract assumptions in both cases. Numerical approximation by the finite element method and suitable time discretization schemes are proved to satisfy the assumptions of Section 4. We start with parabolic stochastic partial differential equations and continue with Volterra equations in a separate subsection.

5.1. Stochastic parabolic partial differential equations. Let $D \subset \mathbb{R}^d$ for $d = 1, 2, 3$ be a convex polygonal domain. Let $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ be the Laplace operator and $f \in C_0^1(\mathbb{R}; \mathbb{R})$. We consider the stochastic partial differential equation:

$$\begin{align*}
\dot{u}(t, x) &= \Delta u(t, x) + f(u(t, x)) + \dot{\eta}(t, x), \quad (t, x) \in (0, T) \times D, \\
u(t, x) &= 0, \quad (t, x) \in (0, T) \times \partial D, \\
u(0, x) &= u_0(x), \quad x \in D.
\end{align*}$$

The noise $\dot{\eta}$ is not well defined as a function, as it is written, but makes sense as a random measure. We will study this equation in the abstract framework of Section 1. Let $H = L^2(D)$, $A : D(A) \subset H \rightarrow H$ be given by $A = -\Delta$ with $D(A) = H^2_0(D) \cap H^1(D)$. Let $(S_t)_{t \in [0,T]}$ denote the analytic semigroup $S_t = e^{-tA}$ of bounded linear operators generated by $-A$. Assumption 1.1 is satisfied with $\rho = 1$, as is easily seen by a spectral argument. The drift $F : H \rightarrow H$ is the Nemytskii operator determined by the action $(F(g))(x) = f(g(x))$, $x \in D$, $g \in H$. Assumption 1.4 for $F$ is verified in [24] for $\delta = \frac{\gamma}{2} + \epsilon$.

Let $(T_h)_{h \in (0,1)}$ denote a family of regular triangulations of $D$ where $h$ denotes the maximal mesh size. Let $(V_h)_{h \in (0,1)}$ be the finite element spaces of continuous
piecewise linear functions with respect to \((T_h)_{h \in (0,1)}\) and \(P_h : H \to V_h\) be the orthogonal projector. The operators \(A_h : V_h \to V_h\) are uniquely determined by
\[
\langle A_h \phi_h, \psi_h \rangle = \langle \nabla \phi_h, \nabla \psi_h \rangle, \quad \forall \phi_h, \psi_h \in V_h \subset \dot{H}^1.
\]

**Remark 5.1.** If the domain \(D\) is such that the pairs of eigenvalues and eigenfunctions \((\lambda_n, \varphi_n)_{n \in \mathbb{N}}\) of \(A\) are known, e.g., \(D = [0,1]^d\), then instead of finite element discretization one can consider a spectral Galerkin approximation. Let the eigenvalues be ordered in increasing order so that \(\lambda_n \leq \lambda_{n+1}\) for every \(n \in \mathbb{N}\). Further, let \(h = \lambda_{N+1}^{-1}\) and \(V_h = \text{span}\{\varphi_n : n \leq N\}\). By \(P_h : H \to V_h\) we denote the orthogonal projector and we define \(A_h = AP_h = P_hA = P_hAP_h\).

We discretize in time by a semi-implicit Euler-Maruyama method. By defining \(B_h^{1,k} = (I + kA_h)^{-1}P_h\) and \(B_n^{h,k} = (B_1^{h,k})^n\) for \(n \geq 1\), the discrete solutions \((X_n^{h,k})_{n=0}^N\) are recursively given by
\[
X_n^{h,k} = B_1^{h,k}X_0^{h,k} + kB_1^{h,k}F(X_n^{h,k}) + \int_{t_n}^{t_{n+1}} B_1^{h,k} dW_s, \quad n = 0, \ldots, N-1,
\]
\[
X_0^{h,k} = P_hx_0.
\]

Iterating the scheme gives the discrete variation of constants formula (4.6). For both finite element and spectral approximation the assumptions (4.1), (4.2), (4.3), (4.4), are valid, see, e.g., [21]. For a proof of (4.5), see [1, Lemma 5.1]. We remark that we need not assume that the mesh family is quasi-uniform due to the use of Lemma 4.5.

5.2. **Stochastic Volterra integro-differential equations.** Consider the semilinear stochastic Volterra type equation
\[
\begin{align*}
\dot{u}(t,x) &= \int_0^t b(t-s) \Delta u(t,x) \, ds + f(u(t,x)) + \dot{\eta}(t,x), \quad (t,x) \in (0,T] \times D, \\
u(t,x) &= 0, \quad (t,x) \in (0,T] \times \partial D, \\
u(0,x) &= u_0, \quad x \in D.
\end{align*}
\]

We assume that the kernel \(b \in L^1_{\text{loc}}(\mathbb{R}_+)\) is 4-monotone; that is, \(b\) is twice continuously differentiable on \((0,\infty)\), \((-1)^n b^{(n)}(t) \geq 0\) for \(t > 0\), \(0 \leq n \leq 2\), and \(b^{(2)}\) is non-increasing and convex. We suppose further that \(\lim_{t \to \infty} b(t) = 0\) and
\[
\limsup_{t \to 0, \infty} \left( \frac{1}{T} \int_0^t sb(s) \, ds \right) / \left( \int_0^t -\dot{b}(s) \, ds \right) < +\infty.
\]

In this case it follows from [20, Proposition 3.10] that the parameter \(\rho\) in Assumption 4.1 is given by
\[
\rho = 1 + \frac{2}{\pi} \sup\{|\arg \hat{b}(\lambda)| : \text{Re} \lambda > 0\} \in (1,2),
\]
where \(\hat{b}\) denotes the Laplace transform of \(b\). Finally, in order to be able to use non-smooth data estimates for the deterministic problem we suppose that \(\hat{b}\) can be extended to an analytic function in a sector \(\Sigma_\theta = \{z \in \mathbb{C} : |\arg z| < \theta\}\) with \(\theta > \frac{\pi}{2}\) and \(|\hat{b}^{(k)}(z)| \leq C|z|^{1-\rho-k}\), \(k = 0,1\), \(z \in \Sigma_\theta\). An important example is the kernel \(b(t) = \frac{1}{(t^{\rho-1})} t^{\rho-2} e^{-\eta t}\), for some \(\rho \in (1,2)\) and \(\eta \geq 0\). When \(\eta = 0\), the corresponding equation can be viewed as a fractional-in-time stochastic equation.
We write the equation in the abstract Itô form (1.6) with \( A, F, W, x_0 \) as in Subsection 5.1. Here one needs \( \delta = \frac{d}{2} + \epsilon < \frac{3}{2} \) and this requires \( \rho < \frac{4}{3} \) and \( \epsilon \) small in the case \( d = 3 \) but causes no restrictions in the case \( d = 1, 2 \). Under the above assumptions there exist a resolvent family of operators \( (S_t)_{t \in [0, T]} \) defined by the strong operator limit

\[
S_t = \sum_{j=1}^{\infty} s_{j,t} (e_j \otimes e_j); \quad \hat{s}_{j,t} + \lambda_j \int_0^t b(t-r)s_{j,r} \, dr = 0, \quad t > 0; \quad s_{j,0} = 1.
\]

Here \( (\lambda_j, e_j)_{j \in \mathbb{N}} \) are the eigenpairs of \( A \). The operator family \( (S_t)_{t \in [0, T]} \) does not possess the semigroup property because of the presence of the memory term. It is the solution operator to the abstract linear homogeneous problem

\[
\dot{Y}_t + \int_0^t b(t-s)AY_s \, ds = 0, \quad t \in [0, T]; \quad Y_0 = y_0,
\]

i.e., \( Y_t = S_t y_0 \). The inhomogeneous problem with right hand side \( g(t) \) for Bochner integrable \( g: [0, T] \to H \) is solved by the variation of constants formula

\[
Y_t = S_t y_0 + \int_0^t S_{t-s} g(s) \, ds, \quad t \in [0, T].
\]

By [3, Lemma 4.4] condition (1.1) holds for \( S \). Thus the setting of Section 1 is applicable.

We now turn our attention to the numerical approximation by presenting the convolution quadrature that we use, which was introduced by Lubich [16, 17]. Let \( (\omega_j^k)_{j \in \mathbb{N}} \) be weights determined by

\[
\hat{b}\left(\frac{1-z}{k}\right) = \sum_{j=0}^{\infty} \omega_j^k z^j, \quad |z| < 1.
\]

Then we use the approximation

\[
\sum_{j=1}^{n} \omega_{n-j}^k f(t_j) \sim \int_0^{t_n} b(t_n-s)f(s) \, ds, \quad f \in \mathcal{C}(0, T; \mathbb{R}).
\]

To discretize the time derivative we use a backward Euler method, which is explicit in the semilinear term \( F \). Our fully discrete scheme then reads:

\[
X_{n+1}^{h,k} - X_n^{h,k} + k \sum_{j=1}^{n+1} \omega_{n+1-j}^k A_h X_j^{h,k} = kP_h F(X_n^{h,k}) + \int_{t_n}^{t_{n+1}} P_h \, dW_t, \quad n \geq 0,
\]

\[
X_0^{h,k} = P_h x_0.
\]

It is possible to write \( (X_n^{h,k})_{n=0}^N \) as a variation of constants formula (4.6). Indeed, it is shown in [14] that one has the explicit representation

\[
B_n^{h,k} = \int_0^\infty S_{k,s}^h P_h e^{-s \ominus_n-1} \, ds, \quad n \geq 1,
\]

where

\[
S_t^h = \sum_{j=1}^{N_h} s_{j,t}^h (e_j^h \otimes e_j^h)P_h; \quad \hat{s}_{j,t}^h + \lambda_j^h \int_0^t b(t-r)\hat{s}_{j,r}^h \, dr = 0, \quad t > 0; \quad \hat{s}_{j,0}^h = 1.
\]
and \((A_j^{h}, e_j^{h})^N_{j=1}\) are the eigenpairs corresponding to \(A_h\). The stability (4.1) holds by [15, Theorem 3.1] and the smooth data error estimate (4.2) was proved in [14, Remark 5.3]. It remains to verify (4.3). By [15, Theorem 3.1] there exists \(\tilde{C}\) so that

\[
\left\| \omega_{n}^{h,k} \right\|_{\mathcal{L}} \leq \tilde{C} \left( \delta + k \right), \quad 0 \leq \delta \leq 2, \ n = 1, \ldots, N.
\]

Let \(0 \leq \delta \leq 2\). Interpolation with \(0 \leq s \leq 1\) yields

\[
\left\| A^h \omega_{n}^{h,k} \right\|_{\mathcal{L}} \leq \left\| \omega_{n}^{h,k} \right\|_{\mathcal{L}}^{1-s} \left\| A^h B^{h,k} \right\|_{\mathcal{L}}^s
\]

\[
\leq \left( \tilde{C} \left( \delta + k \right) \right)^{1-s} \left( 2L t^{-k/2} \right)^s
\]

\[
\leq \tilde{C}^{1-s} \left( 2L \right)^s t^{-k/2} \left( \delta + k \right).
\]

Setting \(\sigma = \delta(1-s)\) and \(R_s = \tilde{C}^{1-s} \left( 2L \right)^s\) yields the estimate

\[
\left\| A^h \omega_{n}^{h,k} \right\|_{\mathcal{L}} \leq R_s t^{-k/2} \left( \delta + k \right), \quad 0 \leq \sigma \leq 2, \ 0 \leq s \leq 1 - \frac{\sigma}{2},
\]

for \(n = 1, \ldots, N\). Therefore (4.3) holds.

References

[1] A. Andersson, R. Kruse, and S. Larsson, Duality in refined Sobolev-Malliavin spaces and weak approximation of SPDE, ArXiv Preprint, arXiv:1312.5893 (2013).
[2] A. Andersson and S. Larsson, Weak convergence for a spatial approximation of the nonlinear stochastic heat equation, ArXiv Preprint, arXiv:1212.5893 (2012). To appear in Math. Comp.
[3] B. Baeumer, M. Geissert, and M. Kovács, Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise, J. Differential Equations (2014). http://dx.doi.org/10.1016/j.jde.2014.09.020.
[4] C.-E. Bréhier, Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise, Potential Analysis 40 (2014), 1–40 (English).
[5] C.-E. 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).
[6] 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).
[7] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.
[8] A. Debussche, Weak approximation of stochastic partial differential equations: the nonlinear case, Math. Comp. 80 (2011), 89–117.
[9] 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), 603–630, S33–S36.
[10] 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.
[11] E. Hausenblas, Weak approximation for semilinear stochastic evolution equations, Stochastic analysis and related topics VIII, 2003, pp. 111–128.
[12] , Weak approximation of the stochastic wave equation, J. Comput. Appl. Math. 235 (2010), 33–58.
[13] A. Jentzen and M. Röckner, Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise, J. Differential Equations 252 (2012), 114–136.
[14] M. Kovács and J. Printems, Strong order of convergence of a fully discrete approximation of a linear stochastic Volterra type evolution equation, Math. Comp. 83 (2014), 2925–2946.
Weak convergence of a fully discrete approximation of a linear stochastic evolution equation with a positive-type memory term, J. Math. Anal. Appl. 413 (2014), 939–952.

C. Lubich, Convolution quadrature and discretized operational calculus. I, Numer. Math. 52 (1988), 129–145.

C. Lubich, Convolution quadrature and discretized operational calculus. II, Numer. Math. 52 (1988), 413–425.

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

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

J. Prüss, Evolutionary Integral Equations and Applications, Modern Birkhäuser Classics, Birkhäuser/Springer Basel AG, Basel, 1993. [2012] reprint of the 1993 edition.

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

J. M. A. M. van Neerven, Stochastic Evolution Equations, 2008. ISEM lecture notes.

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

X. Wang, Weak error estimates of the exponential Euler scheme for semi-linear SPDEs without Malliavin calculus, ArXiv Preprint, arXiv:1408.0713 (2014).

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), 151–169.

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

Mihaly Kovács, Department of mathematics and statistics, University of Otago, PO Box 56, Dunedin, 9054, New Zealand
E-mail address: mkovacs@maths.otago.ac.nz

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