Weak error analysis for semilinear stochastic Volterra equations with additive noise

Citation for the original published paper (version of record):
Andersson, A., Kovacs, M., Larsson, S. (2016). Weak error analysis for semilinear stochastic Volterra equations with additive noise. Journal of Mathematical Analysis and Applications, 437(2): 1283-1304. http://dx.doi.org/10.1016/j.jmaa.2015.09.016

J. Math. Anal. Appl. 437 (2016) 1283–1304

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Weak error analysis for semilinear stochastic Volterra equations with additive noise

Adam Andersson [a,*], Mihály Kovács [b], Stig Larsson [c]

[a] Technische Universität Berlin, Institut für Mathematik, Secr. MA 5-3, Strasse des 17. Juni 136, DE-10623 Berlin, Germany
[b] Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin, 9054, New Zealand
[c] Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 2 June 2015
Available online 29 January 2016
Submitted by H. Frankowska

Keywords:
Stochastic Volterra equations
Finite element method
Backward Euler
Convolution quadrature
Strong and weak convergence
Malliavin regularity

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 convergence result concerns not only functionals of the solution at a fixed time but also more complicated functionals of the entire path and includes convergence of covariances and higher order statistics. 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 $A\colon \mathcal{D}(A)\subset H\to H$ with compact inverse. Let $(A^\alpha)_{\alpha\in\mathbf{R}}$ denote the fractional powers of A, which are well defined, let $(\dot H^\alpha)_{\alpha\in R}$ denote the spaces $\dot H^\alpha=\mathcal{D}(A^\alpha)$ for $\alpha\ge 0$ 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 $(L_s)_{s\in[0,2]}$ such that

$$\big\|A^{\frac{\min(1,s)}{\rho}}S_tx\big\|+\big\|A^{\frac{s-1}{\rho}}\dot S_tx\big\|\le L_st^{-s}\|x\|,\quad t\in(0,T],\ x\in H,\ s\in[0,2]. \tag{1.1}$$

* Corresponding author.
E-mail addresses: andersson@math.tu-berlin.de (A. Andersson), mkovacs@maths.otago.ac.nz (M. Kovács), stig@chalmers.se (S. Larsson).

http://dx.doi.org/10.1016/j.jmaa.2015.09.016
0022-247X/

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 $S_t x = Y_t^x$ of the Volterra equation

$$\dot{Y}_t^x + \int_0^t b_{t-s} A Y_s^x \, \mathrm{d}s = 0, \quad t \in (0,T]; \quad Y_t^x = x,$$

where $b\colon (0,\infty) \to \mathbf{R}$ 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

$$X_t = S_t x_0 + \int_0^t S_{t-s} F(X_s) \, \mathrm{d}s + \int_0^t S_{t-s} \, \mathrm{d}W_s, \quad t \in [0,T]. \tag{1.2}$$

The noise is generated by a cylindrical Q-Wiener process W on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbf{P})$ with positive semidefinite self-adjoint covariance operator $Q \in \mathcal{L}(H)$, where the latter is the space of bounded linear operators on H. Let $H_0 = Q^{\frac{1}{2}}(H)$, and let $\mathcal{L}_2$ and $\mathcal{L}_2^0$ denote the spaces of Hilbert–Schmidt operators $H \to H$ and $H_0 \to H$, respectively. The regularity of the noise is measured by a parameter $\beta \in (0, 1/\rho]$, by assuming

$$\big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0} = \big\|A^{\frac{\beta\rho-1}{2\rho}} Q^{\frac{1}{2}}\big\|_{\mathcal{L}_2} < \infty. \tag{1.3}$$

Under this assumption $X_t \in \dot{H}^\beta$, $\mathbf{P}$-almost surely. The smoothest case $\beta = 1/\rho$ corresponds to trace class noise as (1.3) reduces to $\|Q^{\frac{1}{2}}\|_{\mathcal{L}_2} = \sqrt{\mathrm{Tr}(Q)} < \infty$.

For Hilbert spaces U, V the space $\mathcal{G}_{\mathrm{b}}^k(U;V)$ consists of all, not necessarily bounded, functions $\phi\colon U \to V$, whose Gâteaux derivatives of orders $1,\ldots,k$ are bounded, symmetric and strongly continuous. The nonlinear drift $F\colon H \to H$ is assumed to satisfy, for some $\delta \in [0, 2/\rho)$,

$$F \in \mathcal{G}_{\mathrm{b}}^1(H;H) \cap \mathcal{G}_{\mathrm{b}}^2(H;\dot{H}^{-\delta}). \tag{1.4}$$

This assumption includes interesting cases where $F \notin \mathcal{G}_{\mathrm{b}}^2(H;H)$, e.g., Nemytskii operators on $H = L^2(\mathrm{D})$ for a spatial domain $\mathrm{D} \subset \mathbf{R}^d$, with $\delta > d/2$. The initial value x_0 is deterministic and satisfies

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

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

$$\mathrm{d}X_t + \Big(\int_0^t b_{t-s} A X_s \, \mathrm{d}s\Big) \, \mathrm{d}t = F(X_t) \, \mathrm{d}t + \mathrm{d}W_t, \ t \in [0,T]; \quad X_0 = x_0, \tag{1.6}$$

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 [15, 16], 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

$$\mathrm{d}X_t + AX_t\,\mathrm{d}t = F(X_t)\,\mathrm{d}t + \mathrm{d}W_t,\ t \in [0,T]; \quad X_0 = x_0. \tag{1.7}$$

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 [1,2,4–6,8,11–13,24–26]. In contrast to [1] we allow the nonlinear drift F to be a Nemytskii operator not only in one space dimension but also in two and three space dimensions, without imposing restrictions on the choice of the spatial approximation. 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 $K \in \mathbf{N}$ and $\varphi_i\colon H \to \mathbf{R}$, $i = 1,\dots,K$, be twice Gâteaux differentiable mappings of polynomial growth and $\nu_1,\dots,\nu_K$ finite Borel measures on $[0,T]$. We consider the generalized weak error

$$\Big|\mathbf{E}\Big[\Phi(X) - \Phi(Y)\Big]\Big|, \quad \text{with} \quad \Phi(Z) = \prod_{i=1}^{K} \varphi_i\Big(\int_0^T Z_t\,\mathrm{d}\nu_{i,t}\Big), \tag{1.8}$$

for $X, Y, Z \in \cap_{i=1}^K L^1_{\nu_i}(0,T; L^p(\Omega;H))$ with a suitable exponent $p \geq 2$. In all the works we are aware of, (1.8) is considered with $K = 1$, $\nu_1 = \nu = \delta_\tau$, where δ_τ is the Dirac measure concentrated at τ, 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 $\mathbf{E}[\varphi(\int_0^T X_t\,\mathrm{d}\nu_t)]$. 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.

Our main result, Theorem 4.7, shows convergence of the weak error of the form (1.8) for abstractly defined approximations of the solution X to (1.2). The general form of the functional Φ allows us to prove convergence of approximations of covariances

$$\mathrm{Cov}\big(\langle X_{t_1}, \phi_1\rangle, \langle X_{t_2}, \phi_2\rangle\big), \quad \phi_1, \phi_2 \in H,\ t_1, t_2 \in (0,T],$$

in Corollary 4.8. The generalization to higher order statistics is straightforward and omitted.

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 abstractly defined approximation scheme for (1.2) and prove our main result on weak convergence, Theorem 4.7. In addition, we prove strong convergence in Theorem 4.2, which is then 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

$$\|T\|_{\mathcal{L}_2(U;V)} = \Big(\sum_{j\in\mathbf{N}} \|Tu_j\|_V^2 \Big)^{\frac{1}{2}}, \quad \langle S,T\rangle_{\mathcal{L}_2(U;V)} = \sum_{j\in\mathbf{N}} \langle Su_j, Tu_j\rangle_V. \tag{2.1}$$

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

For $k \geq 1$, let $\mathcal{L}^{[k]}(U;V)$ be the Banach space of all multilinear mappings $b\colon U^k \to V$, equipped with the norm

$$\|b\|_{\mathcal{L}^{[k]}(U;V)} = \sup_{u_1,\dots,u_k\in U} \frac{\|b\cdot(u_1,\dots,u_k)\|_V}{\|u_1\|_U \cdots \|u_k\|_U}.$$

It is clear that $\mathcal{L}^{[1]}(U;V) = \mathcal{L}(U;V)$.

Denote by $\mathcal{C}(U;V)$ the space of all continuous mappings $U \to V$ and further by $\mathcal{C}_{\mathrm{str}}(U;\mathcal{L}^{[k]}(U;V))$ the space of strongly continuous mappings $U \to \mathcal{L}^{[k]}(U;V)$, i.e., mappings $B\colon U \to \mathcal{L}^{[k]}(U;V)$ such that for $u_1,\dots,u_k \in U$, the mapping

$$U \ni x \mapsto B(x)\cdot(u_1,\dots,u_k) \in V,$$

is continuous. A function $\phi\colon U \to V$ is said to be k times Gâteaux differentiable if the recursively defined derivatives, $\phi^{(l)}\colon U^{l+1} \to V$, $l \in \{1,\dots,k\}$,

$$\phi^{(l)}(x)\cdot(u_1,\dots,u_l) = \lim_{\epsilon\to 0} \frac{\phi^{(l-1)}(x+\epsilon u_l)\cdot(u_1,\dots,u_{l-1}) - \phi^{(l-1)}(x)\cdot(u_1,\dots,u_{l-1})}{\epsilon},$$

exist for $u_1,\dots,u_l,x \in U$, $l \in \{1,\dots,k\}$, as limits in V, where $\phi^{(0)} = \phi$. This class of functions is large and fails to have natural properties, e.g., Gâteaux differentiability does not imply continuity and the multilinear mapping $\phi^{(l)}(x)$ may not be symmetric. We therefore introduce a smaller class, with useful properties. For $k \geq 1$, let $\mathcal{G}^k(U;V) \subset \mathcal{C}(U;V)$ be the subset of all k-times Gâteaux differentiable mappings $\phi \in \mathcal{C}(U;V)$, whose derivatives $\phi^{(l)} \in \mathcal{C}_{\mathrm{str}}(U;\mathcal{L}^{[l]}(U;V))$, $l \in \{1,\dots,k\}$, are symmetric. This is a weaker assumption than requiring $\phi^{(l)} \in \mathcal{C}(U;\mathcal{L}^{[l]}(U;V))$, $l \in \{1,\dots,k\}$, which would be the same as assuming Fréchet differentiability. For integers $k \in \{0,\dots,m\}$ and $\phi \in \mathcal{G}^k(U;V)$, let

$$|\phi|_{\mathcal{G}_{\mathrm{p}}^{k,m}(U;V)} = \sup_{u\in U} \frac{\|\phi^{(k)}(u)\|_{\mathcal{L}^{[k]}(U;V)}}{(1+\|u\|_U^{m-k})}, \tag{2.2}$$

and let $\mathcal{G}_{\mathrm{p}}^{k,m}(U;V)$ be the space of $\phi \in \mathcal{G}^k(U;V)$ such that $|\phi|_{\mathcal{G}_{\mathrm{p}}^{l,m}(U;V)} < \infty$ for $l \in \{1,\dots,k\}$. Let $\mathcal{G}_{\mathrm{p}}^{\infty}(U;V)$ be the space of all infinitely many times differentiable mappings $\phi\colon U \to V$ such that ϕ and all its derivatives satisfy a polynomial bound. Let $\mathcal{G}_{\mathrm{b}}^k(U;V)$ denote the space of $\phi \in \mathcal{G}^k(U;V)$ such that

$$|\phi|_{\mathcal{G}_{\mathrm{b}}^l(U;V)} = \sup_{u\in U} \|\phi^{(l)}(u)\|_{\mathcal{L}^{[l]}(U;V)} < \infty, \quad l \in \{1,\dots,k\}.$$

For $\phi \in \mathcal{G}^1(U;\mathbf{R})$ we can identify the derivative with the gradient $\phi'(u) \in U^* = U$, by the Riesz Representation Theorem. For $m \geq 1$, $\phi \in \mathcal{G}_{\mathrm{p}}^{1,m}(U;V)$, the map $[0,1] \ni \lambda \mapsto \phi'(y+\lambda(x-y))\cdot(x-y) \in V$ is continuous and Bochner integrable and therefore

$$\phi(x) = \phi(y) + \int_0^1 \phi'(y+\lambda(x-y))\cdot(x-y)\,\mathrm{d}\lambda, \quad x,y \in U. \tag{2.3}$$

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 $L^p_\nu(0,T;V)$ be the Bochner space of ν-measurable mappings $Z\colon [0,T] \to V$ such that

$$\|Z\|_{L^p_\nu(0,T;V)} = \Big(\int_0^T \|Z_t\|_V^p \, \mathrm{d}\nu_t \Big)^{\frac{1}{p}} < \infty,$$

with the usual modification for $p = \infty$. When ν is Lebesgue measure we write $L^p(0,T;V)$.

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

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 $(x_n)_{n\in\mathbf{N}} \subset \mathcal{X}$ satisfies $x_n \to x$ weakly in $\mathcal{Y}$ as $n \to \infty$ and $\sup_{n\in\mathbf{N}} \|x_n\|_{\mathcal{X}} < \infty$, then $x \in \mathcal{X}$.*

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

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

Lemma 2.2. *Let $T > 0$, $N \in \mathbf{N}$, $k = T/N$, and $t_n = nk$ for $0 \le n \le N$. If $\varphi_1, \dots, \varphi_N \ge 0$ satisfy for some $M_0, M_1 \ge 0$ and $\mu, \nu > 0$ the inequality*

$$\varphi_n \le M_0 \,(1 + t_n^{-1+\mu}) + M_1\, k \sum_{j=1}^{n-1} t_{n-j}^{-1+\nu} \varphi_j, \quad 1 \le n \le N,$$

then there exists a constant $M_2 = M_2(\mu, \nu, M_1, T)$ such that

$$\varphi_n \le M_0 M_2 \,(1 + t_n^{-1+\mu}), \quad 1 \le n \le N.$$

2.2. The Wiener integral and Malliavin calculus

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbf{P})$, be a filtered probability space, with Bochner spaces $L^p(\Omega;V) = L^p((\Omega,\mathcal{F},\mathbf{P});V)$, $p \in [1,\infty]$, V being a Banach space. In the case $V = \mathbf{R}$ we write $L^p(\Omega) = L^p(\Omega;\mathbf{R})$. 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 $\mathcal{L}_2^0 = \mathcal{L}_2(H_0;H)$ we denote the space of Hilbert–Schmidt operators $H_0 \to H$. Let W be a cylindrical Q-Wiener process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbf{P})$, i.e., $W \in \mathcal{L}(H_0; \mathcal{C}(0,T;L^2(\Omega)))$ and $((Wu)_t)_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-adapted real-valued Brownian motion for every $u \in H_0$ with

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

The stochastic Wiener integral

$$\int_0^T \Phi_t \, \mathrm{d}W_t, \quad \Phi \in L^2(0,T;\mathcal{L}_2^0),$$

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 [7,20,23]. We cite the following consequence of the Burkholder inequality [7, Lemma 7.2], for deterministic integrands and $p \geq 2$,

$$\Big\| \int_0^T \Phi_t \, \mathrm{d}W_t \Big\|_{L^p(\Omega;H)} \leq \frac{p(p-1)}{2} \|\Phi\|_{L^2(0,T;\mathcal{L}_2^0)}, \quad \Phi \in L^2(0,T;\mathcal{L}_2^0). \tag{2.4}$$

By taking $H = \mathbf{R}$ and noting the isomorphisms $H_0 \cong H_0^* \cong \mathcal{L}_2(H_0; \mathbf{R})$ we see that a function $\phi \in L^2(0,T;H_0)$ defines an integrand in $L^2(0,T;\mathcal{L}_2(H_0;\mathbf{R}))$ for the stochastic integral and the integral $\int_0^T \phi_t \, \mathrm{d}W_t \in L^2(\Omega)$ 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 $\mathcal{S}^q(\mathbf{R})$ be the class of smooth cylindrical random variables of the form

$$F = f\Big(\int_0^T \phi_{1,s} \, \mathrm{d}W_s, \ldots, \int_0^T \phi_{n,s} \, \mathrm{d}W_s \Big),$$
$$f \in \mathcal{G}_\mathrm{p}^\infty(\mathbf{R}^n;\mathbf{R}),\ (\phi_k)_{k=1}^n \subset L^q(0,T;H_0),\ n \in \mathbf{N}.$$

For $F \in \mathcal{S}^q(\mathbf{R})$ with the above representation, we define the Malliavin derivative

$$(D_t F)_{t\in[0,T]} = \Bigg(\sum_{j=1}^n \partial_j f\Big(\int_0^T \phi_{1,s} \, \mathrm{d}W_s, \ldots, \int_0^T \phi_{n,s} \, \mathrm{d}W_s \Big) \otimes \phi_{j,t}, \Bigg)_{t\in[0,T]}.$$

Let V be a separable Hilbert space. We define $\mathcal{S}^q(V)$ to be the space of all V-valued random variables of the form $Y = \sum_{i=1}^m v_i \otimes F_i$ with $(v_i)_{i=1}^m \subset V$, $(F_i)_{i=1}^m \subset \mathcal{S}^q(\mathbf{R})$, $m \in \mathbf{N}$. The Malliavin derivative of $Y \in \mathcal{S}^q(V)$ 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 $V \otimes H_0 = \mathcal{L}_2(H_0;V)$-valued process.

For $p \in [2,\infty)$, $q \in [2,\infty]$, $\mathcal{S}^q(V) \subset L^p(\Omega;V)$ is dense by [1, Lemma 3.1] and the operator $D\colon \mathcal{S}^q(V) \to L^p(\Omega; L^q(0,T;\mathcal{L}_2(H_0;V)))$ is closable by [1, Lemma 3.2]. Let $\mathbf{M}^{1,p,q}(V)$ denote the closure of $\mathcal{S}^q(V)$ with respect to the norm

$$\|Y\|_{\mathbf{M}^{1,p,q}(V)} = \Big(\|Y\|_{L^p(\Omega;V)}^p + \|DY\|_{L^p(\Omega;L^q(0,T;\mathcal{L}_2(H_0;V)))}^p \Big)^{\frac{1}{p}}.$$

We also use the corresponding seminorm $|Y|_{\mathbf{M}^{1,p,q}(V)} = \|DY\|_{L^p(\Omega;L^q(0,T;\mathcal{L}_2(H_0;V)))}$. The spaces $\mathbf{M}^{1,p,q}(V)$ are Banach spaces, densely embedded into $L^2(\Omega;V)$. Thus, $\mathbf{M}^{1,p,q}(V) \subset L^2(\Omega;V) \subset \mathbf{M}^{1,p,q}(V)^*$ is a Gelfand triple. By [1, Theorem 3.5] the following inequality holds for $p \in [2,\infty)$, $q \in [2,\infty]$ with $\frac{1}{q} + \frac{1}{q'} = 1$:

$$\Big\| \int_0^T \Phi_t \, \mathrm{d}W_t \Big\|_{\mathbf{M}^{1,p,q}(V)^*} \leq \|\Phi\|_{L^{q'}(0,T;\mathcal{L}_2(H_0;V))}, \quad \Phi \in L^2(0,T;\mathcal{L}_2(H_0;V)). \tag{2.5}$$

What makes this duality theory useful is the possibility of taking q' close to 1, cf., (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 Φ to be random and only Skorohod integrability is required. Following [1] we refer to $\mathbf{M}^{1,p,q}(H)$ for $q > 2$ as refined Sobolev–Malliavin spaces. The spaces $\mathbf{M}^{1,p,2}(V)$ are classical Sobolev–Malliavin spaces, often denoted $\mathbf{D}^{1,p}(V)$. For $p = q$ we write $\mathbf{M}^{1,p}(V) := \mathbf{M}^{1,p,p}(V)$.

We next state a modified version of [1, Lemma 3.10]. It provides a local Lipschitz bound that enables us to prove an error estimate in the $\mathbf{M}^{1,p}(H)^*$-norm by a Gronwall argument in Lemma 4.6 below. More precisely, [1, Lemma 3.10] is a local Lipschitz bound from $\mathbf{G}^{1,p}(U)^*$ to $\mathbf{G}^{1,p}(V)^*$ for mappings $\sigma \in \mathcal{G}_{\mathrm{b}}^2(U;V)$, where $\mathbf{G}^{1,p}(U) = \mathbf{M}^{1,p}(U) \cap L^{2p}(\Omega;U)$. The Lipschitz constant depends on the $\mathbf{M}^{1,2p,p}(U)$-norms of the random variables. By restriction to random variables in $\mathbf{M}^{1,p}(U)$ with Malliavin derivative bounded over Ω, Lemma 2.3 provides a more natural bound, obviating the need for the spaces $\mathbf{G}^{1,p}(V)$. The Lipschitz constant now depends on the $\mathbf{M}^{1,\infty,p}(U)$-seminorm. It is proved in the same way as [1, Lemma 3.10], by application of a modified version of [1, Lemma 3.8], based on Hölder's inequality with exponents 1, ∞ instead of 2, 2. We omit the details. In the following Lemma 2.4 we cite parts of [1, Lemma 3.9].

Lemma 2.3. *Let U, V be separable Hilbert spaces, $\sigma \in \mathcal{G}_{\mathrm{b}}^2(U;V)$, and $p \in [2,\infty)$. For $Y^1, Y^2 \in \mathbf{M}^{1,p}(U)$ with $DY^1, DY^2 \in L^\infty(\Omega; L^p(0,T;\mathcal{L}(H_0;U)))$, it holds that*

$$\big\|\sigma(Y^1) - \sigma(Y^2)\big\|_{\mathbf{M}^{1,p}(V)^*} \le \max\big(|\sigma|_{\mathcal{G}_{\mathrm{b}}^1(U;V)}, |\sigma|_{\mathcal{G}_{\mathrm{b}}^2(U;V)}\big)\Big(1 + \sum_{i=1}^{2} \big|Y^i\big|_{\mathbf{M}^{1,\infty,p}(U)}\Big)\big\|Y^1 - Y^2\big\|_{\mathbf{M}^{1,p}(U)^*}.$$

Lemma 2.4. *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 $\|SY\|_{\mathbf{M}^{1,p,q}(H)^*} \le \|S\|_{\mathcal{L}(H)}\|Y\|_{\mathbf{M}^{1,p,q}(H)^*}$.*

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 $\mathbf{P}(X_t^1 \ne X_t^2) = 0$.

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

Proof. Existence is proved by a standard application of Banach's Fixed Point Theorem, see, e.g., [14, 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 $\mathcal{C}(0,T;\mathbf{M}^{1,p,q}(H))$ regularity, see Proposition 4.4 below. □

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

Proposition 3.2. *Let X be the solution to (1.2). For $\gamma \in (0,\beta)$, $p \ge 2$, $q = \frac{2}{1-\rho\gamma}$, there exists $C > 0$ such that*

$$\big\|X_{t_2} - X_{t_1}\big\|_{L^p(\Omega;H)} \le C\big|t_2 - t_1\big|^{\frac{\rho\gamma}{2}}, \quad t_1, t_2 \in [0,T],$$
$$\big\|X_{t_2} - X_{t_1}\big\|_{\mathbf{M}^{1,p,q}(H)^*} \le C\big|t_2 - t_1\big|^{\rho\gamma}, \quad t_1, t_2 \in [0,T].$$

Proof. Fix $\gamma \in (0,\beta)$, $p \ge 2$. In order to treat both cases simultaneously we define $V_2 = L^p(\Omega;H)$, $c_{p,2} = p(p-1)/2$, and $V_r = \mathbf{M}^{1,p,r}(H)^*$, $c_{p,r} = 1$ for $r \in (2,\infty]$. In view of (2.4) and (2.5) it holds that

$$\Big\|\int_0^T \Phi_t \,\mathrm{d}W_t\Big\|_{V_r} \le c_{p,r}\big\|\Phi\big\|_{L^{r'}(0,T;\mathcal{L}_2^0)}, \quad \Phi \in L^2(0,T;\mathcal{L}_2^0),\ r \in [2,\infty], \tag{3.1}$$

where $\frac{1}{r} + \frac{1}{r'} = 1$. Let $t_2 > t_1$. The difference $X_{t_2} - X_{t_1}$ can be written in the form

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

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

$$\begin{aligned}\|X_{t_2} - X_{t_1}\|_{V_r} \le{}& \|(S_{t_2} - S_{t_1})x_0\| \\ &+ \Big\|\int_0^{t_1} (S_{t_2-s} - S_{t_1-s})F(X_s)\,\mathrm{d}s\Big\|_{L^p(\Omega;H)} + \Big\|\int_{t_1}^{t_2} S_{t_2-s}F(X_s)\,\mathrm{d}s\Big\|_{L^p(\Omega;H)} \\ &+ \Big\|\int_0^{t_1} (S_{t_2-s} - S_{t_1-s})\,\mathrm{d}W_s\Big\|_{V_r} + \Big\|\int_{t_1}^{t_2} S_{t_2-s}\,\mathrm{d}W_s\Big\|_{V_r}.\end{aligned}$$

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

$$\|(S_{t_2} - S_{t_1})x_0\| = \Big\|\int_{t_1}^{t_2} \dot{S}_t A^{-\frac{1}{\rho}} A^{\frac{1}{\rho}} x_0\,\mathrm{d}t\Big\| \le L_0 \|A^{\frac{1}{\rho}} x_0\|(t_2 - t_1).$$

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

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

$$\begin{aligned}\Big\|\int_0^{t_1} (S_{t_2-s} - S_{t_1-s})\,\mathrm{d}W_s\Big\|_{V_r} &\le c_{p,r}\Big(\int_0^{t_1} \|(S_{t_2-s} - S_{t_1-s})A^{\frac{1-\beta\rho}{2\rho}}\|_{\mathcal{L}}^{r'} \|A^{\frac{\beta\rho-1}{2\rho}}\|_{\mathcal{L}_2^0}^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}} \\ &\le c_{p,r}\|A^{\frac{\beta\rho-1}{2\rho}}\|_{\mathcal{L}_2^0}\Big(\int_0^{t_1}\Big(\int_{t_1}^{t_2} \|\dot{S}_{t-s}A^{\frac{(3-\beta\rho)/2-1}{\rho}}\|_{\mathcal{L}}\,\mathrm{d}t\Big)^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}} \\ &\le c_{p,r}\|A^{\frac{\beta\rho-1}{2\rho}}\|_{\mathcal{L}_2^0} L_{\frac{3-\beta\rho}{2}}\Big(\int_0^{t_1}\Big(\int_{t_1}^{t_2} (t-s)^{-\frac{3-\beta\rho}{2}}\,\mathrm{d}t\Big)^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}}.\end{aligned}$$

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

$$\begin{aligned}\Big(\int_0^{t_1}\Big(\int_{t_1}^{t_2} (t-s)^{-\frac{3-\beta\rho}{2}}\,\mathrm{d}t\Big)^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}} &\le \Big(\int_0^{t_1}\Big((t_1-s)^{-\frac{1-(\beta-2\eta)\rho}{2}}\int_{t_1}^{t_2} (t-t_1)^{-1+\eta\rho}\,\mathrm{d}t\Big)^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}} \\ &= \frac{(t_2-t_1)^{\eta\rho}}{\eta\rho}\Big(\int_0^{t_1} (t_1-s)^{-\frac{r}{r-1}\frac{1-(\beta-2\eta)\rho}{2}}\,\mathrm{d}s\Big)^{\frac{r-1}{r}}.\end{aligned}$$

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

$$\frac{r}{r-1}\frac{1-(\beta-2\eta)\rho}{2} = \frac{1-\beta\rho+2\eta\rho}{1+\rho\gamma} < 1.$$

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,

$$\begin{aligned}\Big\| \int_{t_1}^{t_2} S_{t_2-s}\,\mathrm{d}W_s \Big\|_{V_r} &\le c_{p,r}\Big(\int_{t_1}^{t_2} \|S_{t_2-s}A^{\frac{1-\beta\rho}{2\rho}}\|_{\mathcal{L}}^{r'} \|A^{\frac{\beta\rho-1}{2\rho}}\|_{\mathcal{L}_2^0}^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}} \\ &\le c_{p,r}L_{\frac{1-\beta\rho}{2}}\|A^{\frac{\beta\rho-1}{2\rho}}\|_{\mathcal{L}_2^0}^{r'}\Big(\int_{t_1}^{t_2}(t_2-s)^{-\frac{r}{r-1}\frac{1-\beta\rho}{2}}\,\mathrm{d}s\Big)^{\frac{r-1}{r}} \\ &\le (t_2-t_1)^{\frac{r-1}{r}-\frac{1-\beta\rho}{2}}.\end{aligned}$$

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

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

and for $r = 2$ the Hölder exponents equals $\beta\rho/2 > \gamma\rho/2$. □

4. Weak and strong convergence

This section contains our main result and its proof. Theorem 4.7 states a weak error estimate for abstractly defined approximations of quantities of the form $\mathbf{E}[\Phi(X)] = \mathbf{E}[\prod_{i=1}^{K}\varphi_i(\int_0^T X_t\,\mathrm{d}\nu_t^i)]$ for $(\nu^i)_{i=1}^K \subset \mathcal{M}_T$, $(\varphi_i)_{i=1}^K \subset \mathcal{G}_{\mathrm{p}}^{2,m}(H;\mathbf{R})$, $m \ge 2$, and X being the solution to (1.2). Theorem 4.2 provides a strong error estimate for approximations of X. For parabolic problems, weak convergence, more precisely, convergence of approximations of $\mathbf{E}[\varphi(X_t)]$ for fixed $t \in [0,T]$ has been considered [1], and for Volterra equations in [16] but only in the linear case $F = 0$. To the best of our knowledge the more general convergence in Theorem 4.7 is new in both cases. The rate of convergence for $\mathbf{E}[\Phi(X)]$ is twice the strong rate as expected. We begin by presenting a family of abstractly defined approximations.

4.1. Approximation

Assume that (1.1), (1.3)–(1.5) hold. Let $(V_h)_{h\in(0,1)}$ be a family of finite-dimensional subspaces of H and let $P_h\colon H \to V_h$ be the orthogonal projector. Let $k \in (0,1)$ and $t_n = nk$, $n = 0,\dots,N$, where $t_N < T \le t_N + k$. Let $(B^{h,k})_{h,k\in(0,1)}$ be a family of operator-valued functions $B^{h,k}\colon \{0,\dots,N\} \to \mathcal{L}(H;V_h)$ such that $B_n^{h,k} = B_n^{h,k}P_h$, and let $(A_h)_{h\in(0,1)}$ be a collection of linear operators $A_h\colon V_h \to V_h$ such that for $n = 1,\dots,N$ it holds that

$$\|A_h^{\frac{s}{\rho}}B_n^{h,k}x\| \le L_s t_n^{-s}\|x\|, \quad x \in H,\ 0 \le s \le 1, \tag{4.1}$$

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

$$\|E_n^{h,k}x\| \le K_\epsilon\big(h^\sigma + k^{\frac{\sigma}{2}}\big)\|x\|_{\dot H^{\sigma(1+\epsilon)}}, \quad 0 \le \sigma \le 2,\ \epsilon > 0, \tag{4.2}$$

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

$$\big\|A^{\frac{s}{2\rho}} E_n^{h,k} x\big\| \le R_s\big(h^{\frac{\sigma}{\rho}} + k^{\frac{\sigma}{2}}\big) t_n^{-\frac{\sigma+s}{2}} \|x\|, \quad 0 \le \sigma \le 2,\ 0 \le s \le 1 - \sigma/2, \tag{4.3}$$

$$\big\|\big(e^{-tA} - e^{-tA_h} P_h\big)x\big\| \le R_0 h^\sigma t^{-\frac{\sigma}{2}} \|x\|, \quad 0 \le \sigma \le 2, \tag{4.4}$$

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 $\tilde{E}^{h,k}\colon [0,T] \to \mathcal{L}$ given by $\tilde{E}_t^{h,k} = S_t - B_n^{h,k}$ for $t \in [t_n, t_{n+1})$ and $n = 0, \dots N-1$. By (1.1) and (4.2) the family $(\tilde{E}_t^{h,k})_{t\in[0,T]}$ satisfies for $t \in (0,T]$ the bound

$$\big\|A^{\frac{s}{2\rho}} \tilde{E}_t^{h,k}\big\|_{\mathcal{L}} \le R_s\big(h^{\frac{\sigma}{\rho}} + k^{\frac{\sigma}{2}}\big) t^{-\frac{\sigma+s}{2}}, \quad 0 \le \sigma \le 2,\ 0 \le s \le 1 - \sigma/2. \tag{4.5}$$

The discrete and continuous stochastic convolutions are defined by

$$W_t^S = \int_0^t S_{t-s}\,\mathrm{d}W_s, \quad t \in [0,T]; \quad W_n^{B^{h,k}} = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k}\,\mathrm{d}W_t, \quad n = 1, \dots, N.$$

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

$$X_n^{h,k} = B_n^{h,k} x_0 + k \sum_{j=1}^{n-1} B_{n-j}^{h,k} F(X_j^{h,k}) + W_n^{B^{h,k}}, \quad n = 1, \dots, N. \tag{4.6}$$

4.2. Strong convergence

Boundedness in the $L^p(\Omega; H)$-sense of the approximate family $(X_n^{h,k})_{n=0}^N$ 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 Section 4.1 hold. For $p \ge 2$ it holds that*

$$\sup_{h,k\in(0,1)} \max_{n\in\{0,\dots,N\}} \big\|X_n^{h,k}\big\|_{L^p(\Omega;H)} < \infty.$$

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 Section 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 $C > 0$ such that*

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

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

$$X_{t_n} - X_n^{h,k} = \big(S_{t_n} - B_n^{h,k}\big)x_0 + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \big(S_{t_n - t} - B_{n-j}^{h,k}\big) F(X_t)\,\mathrm{d}t$$
$$+ \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B_{n-j}^{h,k} \big(F(X_t) - F(X_j^{h,k})\big)\,\mathrm{d}t + W_{t_n}^S - W_n^{B^{h,k}}. \tag{4.7}$$

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 $\tilde{E}^{h,k}_t = S_t - B^{h,k}_n$ for $t \in [t_n, t_{n+1})$ and $n = 0, \dots, N-1$. We get

$$\begin{aligned}\|X_{t_n} - X^{h,k}_n\|_{L^p(\Omega;H)} \le \|E^{h,k}_n x_0\|_H &+ \Big\| \int_0^{t_n} \tilde{E}^{h,k}_{t_n-t} F(X_t)\,\mathrm{d}t \Big\|_{L^p(\Omega;H)} \\ &+ \Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B^{h,k}_{n-j}\big(F(X_t) - F(X^{h,k}_j)\big)\,\mathrm{d}t \Big\|_{L^p(\Omega;H)} \\ &+ \|W^S_{t_n} - W^{B^{h,k}}_n\|_{L^p(\Omega;H)}.\end{aligned}$$

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

$$\max_{n\in\{0,\dots,N\}} \|E^{h,k}_n x_0\| \le K_{\frac{3-2\gamma\rho}{2\gamma\rho}} \big(h^{2\rho\gamma} + k^{\rho\gamma}\big) \|x_0\|_{\dot{H}^3}. \tag{4.8}$$

By Proposition 3.1, (1.4), (4.5) it holds that

$$\begin{aligned}\Big\| \int_0^{t_n} \tilde{E}^{h,k}_{t_n-t} F(X_t)\,\mathrm{d}t \Big\|_{L^p(\Omega;H)} &\le \int_0^{t_n} \|\tilde{E}^{h,k}_{t_n-t}\|_{\mathcal{L}} \|F(X_t)\|_{L^p(\Omega;H)}\,\mathrm{d}t \\ &\le R_0\big(h^{2\gamma} + k^{\rho\gamma}\big)|F|_{\mathcal{G}^1_{\mathrm{b}}(H;H)}\Big(1 + \sup_{t\in[0,T]} \|X_t\|_{L^p(\Omega;H)}\Big) \int_0^{t_n} (t_n - t)^{-\rho\gamma}\,\mathrm{d}t \\ &\lesssim h^{2\gamma} + k^{\rho\gamma}.\end{aligned} \tag{4.9}$$

Using (1.4), (2.3), (4.1), and Proposition 3.2 yields

$$\begin{aligned}&\Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B^{h,k}_{n-j}\big(F(X_t) - F(X^{h,k}_j)\big)\,\mathrm{d}t \Big\|_{L^p(\Omega;H)} \\ &\quad\le |F|_{\mathcal{G}^1_{\mathrm{b}}(H;H)} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|B^{h,k}_{n-j}\|_{\mathcal{L}} \|X_t - X^{h,k}_j\|_{L^p(\Omega;H)}\,\mathrm{d}t \\ &\quad\le L_0|F|_{\mathcal{G}^1_{\mathrm{b}}(H;H)} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \Big(\|X_t - X_{t_j}\|_{L^p(\Omega;H)} + \|X_{t_j} - X^{h,k}_j\|_{L^p(\Omega;H)}\Big)\,\mathrm{d}t \\ &\quad\le L_0|F|_{\mathcal{G}^1_{\mathrm{b}}(H;H)}\Big(CTk^{\frac{\rho\gamma}{2}} + k\sum_{j=0}^{n-1} \|X_{t_j} - X^{h,k}_j\|_{L^p(\Omega;H)}\Big).\end{aligned}$$

For the error of the stochastic convolution we write the difference in the form

$$W^S_{t_n} - W^{B^{h,k}}_n = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \big(S_{t_n-t} - B^{h,k}_{n-j}\big)\,\mathrm{d}W_t \tag{4.10}$$

$$= \int_0^{t_n} \tilde{E}^{h,k}_{t_n-t}\,\mathrm{d}W_t = \int_0^{t_n} \tilde{E}^{h,k}_t\,\mathrm{d}W_t. \tag{4.11}$$

By (2.4) and (4.5) with $\sigma = \gamma\rho$, and $s = 1 - \beta\rho$, we obtain the estimate

$$\|W^S_{t_n} - W^{B^{h,k}}_n\|_{L^p(\Omega;H)} \le \Big(\frac{p(p-1)}{2}\int_0^{t_n} \|A^{\frac{\beta\rho-1}{2\rho}}\|^2_{\mathcal{L}^0_2}\|A^{\frac{1-\beta\rho}{2\rho}}\tilde{E}^{h,k}_t\|^2_{\mathcal{L}}\,\mathrm{d}t\Big)^{\frac{1}{2}}$$

$$\lesssim R_{1-\beta\rho}\Big(\int_0^{t_n} t^{\rho(\beta-\gamma)-1}\,\mathrm{d}t\Big)^{\frac{1}{2}}\big(h^\gamma + k^{\frac{\rho\gamma}{2}}\big) \lesssim h^\gamma + k^{\frac{\rho\gamma}{2}}.$$

Collecting the estimates yields that, for all $n = 0, \dots, N$, it holds

$$\|X_{t_n} - X^{h,k}_n\|_{L^p(\Omega;H)} \lesssim h^\gamma + k^{\frac{\rho\gamma}{2}} + k\sum_{j=0}^{n-1}\|X_{t_j} - X^{h,k}_j\|_{L^p(\Omega;H)}.$$

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 $\mathbf{M}^{1,p}(H)^*$ 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 in a more general setting. We begin by proving the Malliavin differentiability of $X^{h,k}$.

Proposition 4.3. *Let the setting of Section 4.1 hold, and let $X^{h,k}$ be the solution to (4.6). For $p \in [2, \infty)$, $q \in [2, \frac{2}{1-\rho\beta})$, it holds that*

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

Sketch of proof. Note first that $DX^{h,k}_0 = 0$ as $X^{h,k}_0$ is deterministic. Therefore it follows inductively that $X^{h,k}_j$, $j = 0, \dots, N$, are differentiable and the derivative satisfies the equation

$$D_r X^{h,k}_n = k\sum_{j=0}^{n-1} B^{h,k}_{n-j}F'(X^{h,k}_j)D_r X^{h,k}_j + \sum_{j=0}^{n-1}\chi_{[t_j,t_{j+1})}(r)B^{h,k}_{n-j}. \tag{4.12}$$

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 Section 4.1 hold and let X be the solution to (1.2). For $p \in [2, \infty)$, $q \in [2, \frac{2}{1-\rho\beta})$, it holds that $X \in \mathcal{C}(0, T; \mathbf{M}^{1,p,q}(H))$, and moreover it holds that*

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

Proof. Let $\tilde{X}^{h,k}_t = X^{h,k}_n$ for $t \in [t_n, t_{n+1})$, $n = 0, \dots, N-1$, $h, k \in (0, 1)$. 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 $\mathcal{X} = L^2(0, T; \mathbf{M}^{1,2,2}(H))$, and by Theorem 4.2 it holds that $\tilde{X}^{h,k} \to X$ as $h, k \to 0$ in the Hilbert space $\mathcal{Y} = L^2(0, T; L^2(\Omega; H))$. Lemma 2.1 applies and ensures that $X \in \mathcal{X} = L^2(0, T; \mathbf{M}^{1,2,2}(H))$.

By [10, Lemma 3.6] it holds that also $\int_0^\cdot S_{\cdot-s}F(X_s)\,\mathrm{d}s \in L^2(0,T;\mathbf{M}^{1,2,2}(H))$ with $D_r\int_0^t S_{t-s}F(X_s)\,\mathrm{d}s = \int_r^t S_{t-s}F'(X_s)D_rX_s\,\mathrm{d}s$, for $0\le r\le t\le T$, and $\int_0^\cdot S_{\cdot-s}\,\mathrm{d}W_s \in L^2(0,T;\mathbf{M}^{1,2,2}(H))$ with $D_r\int_0^t S_{t-s}\,\mathrm{d}W_s = S_{t-r}$, for $0\le r\le t\le T$. 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

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

A straightforward analysis of this equation, by a Gronwall argument, remove as in the proof of [1, Proposition 3.10] completes the proof. □

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

$$\|A_h^{-\frac{\delta}{2}}P_hx\| \le \|A_h^{\frac{\delta}{2}}P_hA^{-\frac{\delta}{2}}\|_{\mathcal{L}}\|A^{-\frac{\delta}{2}}x\| \le C\|A^{-\frac{\delta}{2}}x\|, \tag{4.13}$$

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 (4.13). In [22, Lemma 5.3] it is shown that for finite element discretization and for $\delta=0,1,2$ it holds

$$\|A_h^{-\frac{\delta}{2}}P_hx\| \le C\big(\|A^{-\frac{\delta}{2}}x\| + h^\delta\|x\|\big), \quad x\in H.$$

The next lemma is a generalization of this result, assuming the availability of a non-smooth data estimate of the form (4.4). It will be used in the proof of Lemma 4.6 below with $\mathcal{X}=\mathbf{M}^{1,p}(H)^*$ for a certain p. By using it we need not rely on (4.13) and in this way we include finite element discretization under the same generality as spectral approximations.

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

$$\big\|A_h^{-\frac{\kappa}{2}}P_hY\big\|_{\mathcal{X}} \le \big\|A^{-\frac{\kappa}{2}}Y\big\|_{\mathcal{X}} + Ch^\sigma\big\|Y\big\|_{L^2(\Omega;H)}, \quad h\in(0,1).$$

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

$$\begin{aligned}\big\|A_h^{-\frac{\kappa}{2}}P_hY\big\|_{\mathcal{X}} &\le \big\|A^{-\frac{\kappa}{2}}Y\big\|_{\mathcal{X}} + \big\|\big(A_h^{-\frac{\kappa}{2}}P_h - A^{-\frac{\kappa}{2}}\big)Y\big\|_{\mathcal{X}}\\ &\lesssim \big\|A^{-\frac{\kappa}{2}}Y\big\|_{\mathcal{X}} + \big\|A_h^{-\frac{\kappa}{2}}P_h - A^{-\frac{\kappa}{2}}\big\|_{\mathcal{L}}\big\|Y\big\|_{L^2(\Omega;H)}.\end{aligned}$$

By [19, Chapter 2, (6.9)] we have

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

Therefore, by (4.4),

$$\big\|A_h^{-\frac{\kappa}{2}}P_h - A^{-\frac{\kappa}{2}}\big\|_{\mathcal{L}} \le \frac{1}{\Gamma(\kappa/2)}\int_0^\infty t^{\frac{\kappa}{2}-1}\big\|e^{-tA_h}P_h - e^{-tA}\big\|_{\mathcal{L}}\,\mathrm{d}t$$

$$\lesssim \int_0^{h^{-2}} t^{\frac{\kappa}{2}-1} \|e^{-tA_h}P_h - e^{-tA}\|_{\mathcal{L}} \,\mathrm{d}t + \int_{h^{-2}}^{\infty} t^{\frac{\kappa}{2}-1} \|e^{-tA_h}P_h - e^{-tA}\|_{\mathcal{L}} \,\mathrm{d}t$$

$$\lesssim h^{\frac{\kappa+\sigma}{2}} \int_0^{h^{-2}} t^{\frac{\kappa-\sigma}{4}-1} \,\mathrm{d}t + h^2 \int_{h^{-2}}^{\infty} t^{\frac{\kappa}{2}-2} \,\mathrm{d}t = \frac{4h^\sigma}{\kappa-\sigma} + \frac{2}{2-\kappa} h^2 h^{2-\frac{\kappa}{2}} \lesssim h^\sigma. \qquad \square$$

The next result is a strong error estimate in the $\mathbf{M}^{1,p}(H)^*$ 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 Section 4.1 hold, and let X and $X^{h,k}$ be the solutions to (1.2) and (4.6), respectively. For $\gamma \in [0,\beta)$, $p = \frac{2}{1-\rho\gamma}$, there exists $C > 0$ such that*

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

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

$$\begin{aligned}
\|X_{t_n} - X_n^{h,k}\|_{\mathbf{M}^{1,p}(H)^*} &\le \|E_n^{h,k} x_0\|_H + \Big\| \int_0^{t_n} \tilde{E}^{h,k}_{t_n-t} F(X_t) \,\mathrm{d}t \Big\|_{L^p(\Omega;H)} \\
&\quad + \Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B^{h,k}_{n-j} \big(F(X_t) - F(X_j^{h,k})\big) \,\mathrm{d}t \Big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\quad + \|W^S_{t_n} - W^{B^{h,k}}_n\|_{\mathbf{M}^{1,p}(H)^*}.
\end{aligned}$$

The first two terms was already estimated as desired in (4.8) and (4.9). Choose κ so that $\max(\delta, 2\gamma) < \kappa < 2/\rho$, where δ is the parameter in (1.4). Since $\rho\kappa < 2$, we have, by Lemma 2.4 and (4.1) with $s = \rho\kappa/2$, that

$$\begin{aligned}
&\Big\| \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} B^{h,k}_{n-j} \big(F(X_t) - F(X_j^{h,k})\big) \,\mathrm{d}t \Big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\quad \le \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|B^{h,k}_{n-j} A_h^{\frac{\kappa}{2}} P_h\|_{\mathcal{L}} \|A_h^{-\frac{\kappa}{2}} P_h \big(F(X_t) - F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*} \,\mathrm{d}t \\
&\quad \le L_{\frac{\kappa\rho}{2}} \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} t_{n-j}^{-\frac{\kappa\rho}{2}} \|A_h^{-\frac{\kappa}{2}} P_h \big(F(X_t) - F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*} \,\mathrm{d}t.
\end{aligned}$$

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

$$\begin{aligned}
&\|A_h^{-\frac{\kappa}{2}} P_h \big(F(X_t) - F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*} \\
&\quad \le C h^{2\gamma} \|F(X_t) - F(X_j^{h,k})\|_{L^2(\Omega;H)} + \|A^{-\frac{\kappa}{2}} \big(F(X_t) - F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*}.
\end{aligned}$$

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

$$\sup_{t\in[0,T]}\max_{j\in\{0,\dots,N\}}\|F(X_t)-F(X_j^{h,k})\|_{L^2(\Omega;H)}$$
$$\leq |F|_{\mathcal{G}_{\mathrm{b}}^1(H;H)}\Big(\sup_{t\in[0,T]}\|X_t\|_{L^2(\Omega;H)}+\max_{j\in\{0,\dots,N\}}\|X_j^{h,k}\|_{L^2(\Omega;H)}\Big)<\infty.$$

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

$$\begin{aligned}\|Y\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})^*}&=\sup_{Z\in\mathbf{M}^{1,p}(\dot H^{-\delta})}\frac{\langle Z,Y\rangle_{L^2(\Omega;\dot H^{-\delta})}}{\|Z\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})}}\\&=\sup_{Z\in\mathbf{M}^{1,p}(\dot H^{-\delta})}\frac{\langle A^{-\frac{\delta}{2}}Z,A^{-\frac{\delta}{2}}Y\rangle_{L^2(\Omega;H)}}{\|Z\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})}}\\&=\sup_{Z\in\mathbf{M}^{1,p}(\dot H^{-\delta})}\frac{\langle Z,A^{-\frac{\delta}{2}}Y\rangle_{L^2(\Omega;H)}}{\|A^{\frac{\delta}{2}}Z\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})}}\\&=\sup_{Z\in\mathbf{M}^{1,p}(H)}\frac{\langle Z,A^{-\frac{\delta}{2}}Y\rangle}{\|Z\|_{\mathbf{M}^{1,p}(H)}}=\|A^{-\frac{\delta}{2}}Y\|_{\mathbf{M}^{1,p}(H)^*}.\end{aligned}$$

Therefore, by Lemma 2.4 and Lemma 2.3 applied with $U=H$, $V=\dot H^{-\delta}$, $\sigma=F$ we get

$$\begin{aligned}&\|A^{-\frac{\kappa}{2}}\big(F(X_t)-F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*}\\&\leq\|A^{-\frac{\kappa-\delta}{2}}\|_{\mathcal{L}}\|A^{-\frac{\delta}{2}}\big(F(X_t)-F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*}\\&=\|A^{-\frac{\kappa-\delta}{2}}\|_{\mathcal{L}}\|F(X_t)-F(X_j^{h,k})\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})^*}\\&\leq\|A^{-\frac{\kappa-\delta}{2}}\|_{\mathcal{L}}\max\Big(|F|_{\mathcal{G}_{\mathrm{b}}^1(H;\dot H^{-\delta})},|F|_{\mathcal{G}_{\mathrm{b}}^2(H;\dot H^{-\delta})}\Big)\\&\quad\times\Big(\sup_{j\in\{0,\dots,N\}}|X_j^{h,k}|_{\mathbf{M}^{1,\infty,p}(H)}+\sup_{t\in[0,T]}|X_t|_{\mathbf{M}^{1,\infty,p}(H)}\Big)\\&\quad\times\Big(\|X_t-X_{t_j}\|_{\mathbf{M}^{1,p}(H)^*}+\|X_{t_j}-X_j^{h,k}\|_{\mathbf{M}^{1,p}(H)^*}\Big).\end{aligned}$$

By Propositions 3.2 and 4.3 and Proposition 4.4, we conclude

$$\|A^{-\frac{\kappa}{2}}\big(F(X_t)-F(X_j^{h,k})\big)\|_{\mathbf{M}^{1,p}(H)^*}\lesssim k^{\rho\gamma}+\|X_{t_j}-X_j^{h,k}\|_{\mathbf{M}^{1,p}(H)^*}.$$

Thus,

$$\Big\|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}B_{n-j}^{h,k}\big(F(X_t)-F(X_j^{h,k})\big)\,\mathrm{d}t\Big\|_{\mathbf{M}^{1,p}(H)^*}$$
$$\lesssim h^{2\gamma}+k^{\rho\gamma}+k\sum_{j=0}^{n-1}t_{n-j}^{-\frac{\kappa\rho}{2}}\|X_{t_j}-X_j^{h,k}\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})^*}.$$

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

$$\begin{aligned}
&\big\|W^S_{t_n} - W^{B^{h,k}}_n\big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\le \Big(\int_0^{t_n} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|^{\frac{2}{1+\rho\gamma}}_{\mathcal{L}_2^0} \big\|A^{\frac{1-\beta\rho}{2\rho}} \tilde{E}^{h,k}_t\big\|^{\frac{2}{1+\rho\gamma}}_{\mathcal{L}} \,\mathrm{d}t\Big)^{\frac{1+\rho\gamma}{2}} \\
&\le R_{1-\beta\rho} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0} \Big(\int_0^{t_n} t^{\frac{\rho(\beta-\gamma)}{1+\rho\gamma}-1} \,\mathrm{d}t\Big)^{\frac{1+\rho\gamma}{2}} \big(h^{2\gamma} + k^{\rho\gamma}\big).
\end{aligned}$$

Altogether we have that for every $n = 1, \ldots, N$ it holds that

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

Lemma 2.2 finishes the proof. □

We next state our main result on weak convergence. We remark that to the best of our knowledge all previous weak convergence results concern convergence of $|\mathbf{E}[\varphi(X^{h,k}_\tau) - \varphi(X_\tau)]|$ for fixed $\tau \in [0, T]$, which is a special case of the following theorem.

Theorem 4.7. *Let X and $X^{h,k}$ be the solutions to (1.2) and (4.6), respectively. Let $\tilde{X}^{h,k}_t = X^{h,k}_n$, for $t \in [t_n, t_{n+1})$, $n \in \{0, \ldots, N-1\}$ and $\tilde{X}^{h,k}_t = X^{h,k}_N$, for $t \in [t_N, T]$. For $K \ge 1$, $m_1, \ldots, m_K \ge 2$, $\varphi_i \in \mathcal{G}^{2,m_i}_{\mathrm{p}}(H; \mathbf{R})$, $\nu_i \in \mathcal{M}_T$, $i = 1, \ldots, K$, $\Phi(Z) = \prod_{i=1}^K \varphi_i(\int_0^T Z_t \,\mathrm{d}\nu_{i,t})$, $\gamma \in [0, \beta)$, there exists $C > 0$ such that*

$$\big|\mathbf{E}\big[\Phi(X) - \Phi(\tilde{X}^{h,k})\big]\big| \le C\big(h^{2\gamma} + k^{\rho\gamma}\big), \quad h, k \in (0, 1).$$

Proof. We start by observing that by (2.3) we have

$$\begin{aligned}
&\prod_{i=1}^K \varphi_i(x_i) - \prod_{i=1}^K \varphi_i(y_i) \\
&= \sum_{l=1}^K \prod_{i=1}^{l-1} \varphi_i(x_i) \prod_{j=l+1}^K \varphi_j(y_j)\big(\varphi_l(x_l) - \varphi_l(y_l)\big) \\
&= \sum_{l=1}^K \Big\langle \prod_{i=1}^{l-1} \varphi_i(x_i) \prod_{j=l+1}^K \varphi_j(y_j) \int_0^1 \varphi_l'(y_l + \lambda(x_l - y_l)) \,\mathrm{d}\lambda, x_l - y_l \Big\rangle \\
&=: \sum_{l=1}^K \langle \gamma_l(x_1, \ldots, x_l, y_l, \ldots, y_K), x_l - y_l\rangle.
\end{aligned}$$

Here we use the convention that an empty product equals 1. We get

$$\big|\mathbf{E}\big[\Phi(X) - \Phi(\tilde{X}^{h,k})\big]\big| = \Big|\sum_{l=1}^K \Big\langle \gamma_l(Y^{h,k}_l), \int_0^T \big(X_t - \tilde{X}^{h,k}_t\big) \,\mathrm{d}\nu_{l,t}\Big\rangle_{L^2(\Omega;H)}\Big|,$$

where

$$Y_l^{h,k} = \Big(\int_0^T X_t \, \mathrm{d}\nu_{1,t}, \ldots, \int_0^T X_t \, \mathrm{d}\nu_{l,t}, \int_0^T \tilde{X}_t^{h,k} \, \mathrm{d}\nu_{l,t}, \ldots, \int_0^T \tilde{X}_t^{h,k} \, \mathrm{d}\nu_{K,t} \Big).$$

By duality in the Gelfand triple $\mathbf{M}^{1,p}(H) \subset L^2(\Omega; H) \subset \mathbf{M}^{1,p}(H)^*$ we obtain

$$\begin{aligned} &\big| \mathbf{E}\big[\Phi(X) - \Phi(\tilde{X}^{h,k})\big] \big| \\ &\quad \le \sum_{l=1}^K \big\|\gamma_l(Y_l^{h,k})\big\|_{\mathbf{M}^{1,p}(H)} \Big\| \int_0^T \big(X_t - \tilde{X}_t^{h,k}\big) \, \mathrm{d}\nu_{l,t} \Big\|_{\mathbf{M}^{1,p}(H)^*} \\ &\quad \le \sum_{l=1}^K \Big(\sup_{h,k \in (0,1)} \big\|\gamma_l(Y_l^{h,k})\big\|_{\mathbf{M}^{1,p}(H)} \Big) \big\|X - \tilde{X}^{h,k}\big\|_{L^1_{\nu_l}(0,T;\mathbf{M}^{1,p}(H)^*)}. \end{aligned}$$

Here $\gamma_l \in \mathcal{G}_{\mathrm{p}}^{1,r}(H^{K+1}; H)$ and $Y_l^{h,k} \in \mathbf{M}^{1,rp}(H^{K+1})$ with $r = \sum_{i=1}^K m_i - 1$. Therefore [1, Lemma 3.3] applied with $U = H^{K+1}$ and $V = H$ gives for $l \in \{1, \ldots, K\}$ the bound

$$\sup_{h,k \in (0,1)} \big\|\gamma_l(Y_l^{h,k})\big\|_{\mathbf{M}^{1,p}(H)} \le C_l \Big(1 + \sup_{h,k \in (0,1)} \big\|Y_l^{h,k}\big\|^r_{\mathbf{M}^{1,rp}(H^{K+1})} \Big).$$

Propositions 4.3 and 4.4 ensure that

$$\begin{aligned} &\sup_{h,k \in (0,1)} \big\|\gamma_l(Y_l^{h,k})\big\|_{\mathbf{M}^{1,p}(H)} \\ &\quad \le \tilde{C}_l \Big(1 + \sum_{i=1}^K \Big(\|X\|^r_{L^1_{\nu_i}(0,T;\mathbf{M}^{1,rp,p}(H))} + \sup_{h,k \in (0,1)} \big\|\tilde{X}^{h,k}\big\|^r_{L^1_{\nu_i}(0,T;\mathbf{M}^{1,rp,p}(H))} \Big) \Big) < \infty. \end{aligned}$$

Let $\tilde{X}$ be the process $\tilde{X}_t = X_{t_n}$ for $t \in [t_n, t_{n+1})$, $n \in \{0, \ldots, N-1\}$. Proposition 3.2 and Lemma 4.6 give, for $l \in \{1, \ldots, K\}$,

$$\begin{aligned} &\big\|X - \tilde{X}^{h,k}\big\|_{L^1_{\nu_l}(0,T;\mathbf{M}^{1,p}(H)^*)} \\ &\quad \le \big\|X - \tilde{X}\big\|_{L^1_{\nu_l}(0,T;\mathbf{M}^{1,p}(H)^*)} + \big\|\tilde{X} - \tilde{X}^{h,k}\big\|_{L^1_{\nu_l}(0,T;\mathbf{M}^{1,p}(H)^*)} \lesssim h^{2\gamma} + k^{\rho\gamma}. \end{aligned}$$

This completes the proof. □

Finally, we formulate a corollary of Theorem 4.7 that can be used to prove convergence of covariances and higher order statistics of approximate solutions. We demonstrate this for covariances; higher order statistics can be treated in a similar way.

Corollary 4.8. *Let X and $X^{h,k}$ be the solutions to (1.2) and (4.6), respectively. Let $\tilde{X}_t^{h,k} = X_n^{h,k}$, 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 $K \ge 1$, $\phi_1, \ldots, \phi_K \in H$, $t_1, \ldots, t_K \in (0,T]$, $\gamma \in [0, \beta)$, there exists $C > 0$ such that*

$$\Big| \mathbf{E}\Big[\prod_{i=1}^K \langle X_{t_i}, \phi_i \rangle - \prod_{i=1}^K \langle \tilde{X}_{t_i}^{h,k}, \phi_i \rangle \Big] \Big| \le C\big(h^{2\gamma} + k^{\rho\gamma}\big), \quad h,k \in (0,1).$$

In particular, for $\phi_1, \phi_2 \in H$, $t_1, t_2 \in (0,T]$, it holds that

$$\begin{aligned} &\big| \mathrm{Cov}\big(\langle X_{t_1}, \phi_1 \rangle, \langle X_{t_2}, \phi_2 \rangle\big) - \mathrm{Cov}\big(\langle \tilde{X}_{t_1}^{h,k}, \phi_1 \rangle, \langle \tilde{X}_{t_2}^{h,k}, \phi_2 \rangle\big) \big| \\ &\quad \le C\big(h^{2\gamma} + k^{\rho\gamma}\big), \quad h,k \in (0,1). \end{aligned}$$

Proof. The first statement follows from Theorem 4.7 by setting $\varphi_i = \langle \phi_i, \cdot \rangle$, $\nu_i = \delta_{t_i}$, $i \in \{1, \dots, K\}$, where δ_{t_i} is the Dirac measure concentrated at t_i. The second is a consequence of the first and the fact that

$$\begin{aligned}
&\operatorname{Cov}\big(\langle X_{t_1}, \phi_1\rangle, \langle X_{t_2}, \phi_2\rangle\big) - \operatorname{Cov}\big(\langle \tilde{X}^{h,k}_{t_1}, \phi_1\rangle, \langle \tilde{X}^{h,k}_{t_2}, \phi_2\rangle\big) \\
&= \mathbf{E}\big[\langle X_{t_1}, \phi_1\rangle\langle X_{t_2}, \phi_2\rangle\big] - \mathbf{E}\big[\langle \tilde{X}^{h,k}_{t_1}, \phi_1\rangle\langle \tilde{X}^{h,k}_{t_2}, \phi_2\rangle\big] \\
&\quad - \mathbf{E}\big[\langle X_{t_1}, \phi_1\rangle - \langle \tilde{X}^{h,k}_{t_1}, \phi_1\rangle\big]\mathbf{E}\big[\langle X_{t_2}, \phi_2\rangle\big] \\
&\quad - \mathbf{E}\big[\langle \tilde{X}^{h,k}_{t_1}, \phi_1\rangle\big]\mathbf{E}\big[\langle X_{t_2}, \phi_2\rangle - \langle \tilde{X}^{h,k}_{t_2}, \phi_2\rangle\big]. \qquad \square
\end{aligned}$$

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 $\mathcal{D} \subset \mathbf{R}^d$ 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 $f \in \mathcal{G}^2_{\mathrm{b}}(\mathbf{R}; \mathbf{R})$. We consider the stochastic partial differential equation:

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

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 $H = L^2(\mathcal{D})$, $A \colon \mathcal{D}(A) \subset H \to H$ be given by $A = -\Delta$ with $\mathcal{D}(A) = H^1_0(\mathcal{D}) \cap H^2(\mathcal{D})$. 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 $F \colon H \to H$ is the Nemytskii operator determined by the action $(F(g))(x) = f(g(x))$, $x \in \mathcal{D}$, $g \in H$. Assumption (1.4) for F is verified in [25] for $\delta = \frac{d}{2} + \epsilon$.

Let $(\mathcal{T}_h)_{h \in (0,1)}$ denote a family of regular triangulations of $\mathcal{D}$ 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 $(\mathcal{T}_h)_{h \in (0,1)}$ and $P_h \colon H \to V_h$ be the orthogonal projector. The operators $A_h \colon V_h \to V_h$ 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 $\mathcal{D}$ is such that the pairs of eigenvalues and eigenfunctions $(\lambda_n, e_n)_{n \in \mathbf{N}}$ of A are known, e.g., $\mathcal{D} = [0,1]^d$, then instead of finite element discretization one can consider a spectral Galerkin approximation. Let the eigenvalues be ordered in increasing order so that $\lambda_n \le \lambda_{n+1}$ for every $n \in \mathbf{N}$. Further, let $h = \lambda_{N+1}^{-\frac{1}{2}}$ and $V_h = \operatorname{span}\{\phi_n : n \le N\}$. By $P_h \colon H \to V_h$ we denote the orthogonal projector and we define $A_h = AP_h = P_h A = P_h A P_h$.

We discretize in time by a semi-implicit Euler–Maruyama method. By defining $B^{h,k}_1 = (I + kA_h)^{-1} P_h$ and $B^{h,k}_n = (B^{h,k}_1)^n$ for $n \ge 1$, the discrete solutions $(X^{h,k}_n)_{n=0}^N$ are recursively given by

$$X_{n+1}^{h,k} = B_1^{h,k} X_n^{h,k} + k B_1^{h,k} F(X_n^{h,k}) + \int_{t_n}^{t_{n+1}} B_1^{h,k} \, \mathrm{d}W_s, \quad n = 0, \ldots, N-1,$$

$$X_0^{h,k} = P_h x_0.$$

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., [22]. For a proof of (4.5), see [1, Lemma 5.1].

5.2. Stochastic Volterra integro-differential equations

Consider the semi-linear stochastic Volterra type equation

$$\begin{aligned} \dot{u}(t,x) &= \int_0^t b(t-s)\Delta u(t,x)\,\mathrm{d}s + f(u(t,x)) + \dot{\eta}(t,x), && (t,x) \in (0,T] \times \mathcal{D}, \\ u(t,x) &= 0, && (t,x) \in (0,T] \times \partial\mathcal{D}, \\ u(0,x) &= u_0, && x \in \mathcal{D}. \end{aligned} \tag{5.1}$$

We assume that the kernel $b \in L^1_{\mathrm{loc}}(\mathbf{R}_+)$ 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

$$\limsup_{t\to 0,\infty} \Big(\frac{1}{t}\int_0^t s b(s)\,\mathrm{d}s\Big) \Big/ \Big(\int_0^t -s\dot{b}(s)\,\mathrm{d}s\Big) < +\infty. \tag{5.2}$$

In this case it follows from [21, Proposition 3.10] that the parameter ρ in Assumption 4.1 is given by

$$\rho = 1 + \frac{2}{\pi} \sup\{|\arg \hat{b}(\lambda)| : \operatorname{Re}\lambda > 0\} \in (1,2), \tag{5.3}$$

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 $b(t) = \frac{1}{\Gamma(\rho-1)} t^{\rho-2} e^{-\eta t}$, for some $\rho \in (1,2)$ and $\eta \geq 0$. When $\eta = 0$, then the corresponding equation can be viewed as a fractional-in-time stochastic equation.

We write the equation in the abstract Itō form (1.6) with A, F, W, x_0 as in Subsection 5.1. Here one needs $\delta = \frac{d}{2} + \epsilon < \frac{2}{\rho}$ and this requires $\rho < \frac{4}{3}$ and ϵ small in the case $d = 3$ but causes no restrictions in the case $d = 1, 2$. Under the above assumptions there exists a resolvent family of operators $(S_t)_{t\in[0,T]}$ defined by the strong operator limit

$$S_t = \sum_{j=1}^{\infty} s_{j,t}\,(e_j \otimes e_j); \quad \dot{s}_{j,t} + \lambda_j \int_0^t b(t-r) s_{j,r}\,\mathrm{d}r = 0,\ t > 0; \quad s_{j,0} = 1. \tag{5.4}$$

Here $(\lambda_j, e_j)_{j\in\mathbf{N}}$ 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

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

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

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

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 [17,18]. Let $(\omega_j^k)_{j\in\mathbf{N}}$ be weights determined by

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

Then we use the approximation

$$\sum_{j=1}^{n} \omega_{n-j}^k f(t_j) \sim \int_0^{t_n} b(t_n - s) f(s) \, \mathrm{d}s, \quad f \in \mathcal{C}(0,T;\mathbf{R}).$$

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:

$$X_{n+1}^{h,k} - X_n^{h,k} + k \sum_{j=1}^{n+1} \omega_{n+1-j}^k A_h X_j^{h,k} = k P_h F(X_n^{h,k}) + \int_{t_n}^{t_{n+1}} P_h \, \mathrm{d}W_t, \quad n = 0, \dots, N-1,$$

$$X_0^{h,k} = P_h x_0.$$

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 [15] that one has the explicit representation

$$B_n^{h,k} = \int_0^{\infty} S_{ks}^h P_h \frac{e^{-s} s^{n-1}}{(n-1)!} \, \mathrm{d}s, \quad n \geq 1,$$

where

$$S_t^h = \sum_{j=1}^{N_h} s_{j,t}^h \, (e_j^h \otimes e_j^h) P_h; \quad \dot{s}_{j,t}^h + \lambda_j^h \int_0^t b(t-r) s_{j,r}^h \, \mathrm{d}r = 0, \; t > 0; \quad s_{j,0}^h = 1,$$

and $(\lambda_j^h, e_j^h)_{j=1}^{N_h}$ are the eigenpairs corresponding to A_h. The stability (4.1) holds by [16, Theorem 3.1] and the smooth data error estimate (4.2) was proved in [15, Remark 5.3]. It remains to verify (4.3). By [16, Theorem 3.1] there exists $\tilde{C}$ such that

$$\big\|E_n^{h,k}\big\|_{\mathcal{L}} \leq \tilde{C} t_n^{-\frac{\delta}{2}} \big(h^{\frac{\delta}{\rho}} + k^{\frac{\delta}{2}}\big), \quad 0 \leq \delta \leq 2, \; n = 1, \dots, N.$$

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

$$\begin{aligned}\left\|A^{\frac{s}{2\rho}} E_{n,\theta}\right\|_{\mathcal{L}} &\leq \left\|E_n^{h,k}\right\|_{\mathcal{L}}^{1-s} \left\|A^{\frac{1}{2\rho}} E_n^{h,k}\right\|_{\mathcal{L}}^{s} \\ &\leq \left\|E_n^{h,k}\right\|_{\mathcal{L}}^{1-s} \Big(\left\|A^{\frac{1}{2\rho}} S_{t_n}\right\|_{\mathcal{L}} + \left\|A^{\frac{1}{2\rho}} B_n^{h,k}\right\|_{\mathcal{L}} \Big)^s \\ &\leq \Big(\tilde{C} t_n^{-\frac{\delta}{2}} \big(h^{\frac{\delta}{\rho}} + k^{\frac{\delta}{2}} \big) \Big)^{1-s} \big(2L_{\frac{1}{2}} t_n^{-\frac{1}{2}} \big)^s \\ &\leq \tilde{C}^{1-s} (2L_{\frac{1}{2}})^s t_n^{-\frac{\delta(1-s)+s}{2}} \big(h^{\frac{\delta(1-s)}{\rho}} + k^{\frac{\delta(1-s)}{2}} \big).\end{aligned}$$

Setting $\sigma = \delta(1-s)$ and $R_s = \tilde{C}^{1-s}(2L_{\frac{1}{2}})^s$ yields the estimate

$$\left\|A^{\frac{s}{2\rho}} E_n^{h,k}\right\|_{\mathcal{L}} \leq R_s t_n^{-\frac{\sigma+s}{2}} \big(h^{\frac{\sigma}{\rho}} + k^{\frac{\sigma}{2}} \big), \quad 0 \leq \sigma \leq 2,\ 0 \leq s \leq 1 - \frac{\sigma}{2},$$

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

Acknowledgments

We thank Arnaud Debussche for suggesting the consideration of covariances and higher order statistics. This motivated us to extend Theorem 4.7 from its original statement with $K = 1$ to $K \geq 1$ and also to include Corollary 4.8. The second author was partially supported by the Marsden Fund project number UOO1418.

References

[1] A. Andersson, R. Kruse, S. Larsson, Duality in refined Sobolev–Malliavin spaces and weak approximation of SPDE, Stoch. Partial Differential Equations: Anal. Comput. (2015), http://dx.doi.org/10.1007/s40072-015-0065-7, online first.
[2] A. Andersson, S. Larsson, Weak convergence for a spatial approximation of the nonlinear stochastic heat equation, Math. Comp. (2015), http://dx.doi.org/10.1090/mcom/3016, online first.
[3] B. Baeumer, M. Geissert, M. Kovács, Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise, J. Differential Equations (258) (2014) 535–554, 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 Anal. 40 (2014) 1–40 (in English).
[5] C.-E. Bréhier, M. Kopec, Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme, arXiv:1311.7030, 2013.
[6] D. Conus, A. Jentzen, R. Kurniawan, Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients, arXiv:1408.1108, 2014.
[7] G. Da Prato, 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, 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, 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, in: Stochastic Analysis and Related Topics VIII, in: Progress in Probability, vol. 53, Birkhäuser, Basel, 2003, pp. 111–128.
[12] E. Hausenblas, Weak approximation of the stochastic wave equation, J. Comput. Appl. Math. 235 (2010) 33–58.
[13] A. Jentzen, R. Kurniawan, Weak convergence rates for Euler-type approximations of semilinear stochastic evolution equations with nonlinear diffusion coefficients, arXiv:1501.03539, 2015.
[14] A. Jentzen, M. Röckner, Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise, J. Differential Equations 252 (2012) 114–136.
[15] M. Kovács, J. Printems, Strong order of convergence of a fully discrete approximation of a linear stochastic Volterra type evolution equation, Math. Comp. 83 (2014) 2325–2346.
[16] M. Kovács, J. Printems, 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.
[17] C. Lubich, Convolution quadrature and discretized operational calculus. I, Numer. Math. 52 (1988) 129–145.
[18] C. Lubich, Convolution quadrature and discretized operational calculus. II, Numer. Math. 52 (1988) 413–425.

[19] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Applied Mathematical Sciences, vol. 44, Springer, New York, 1983.
[20] C. Prévôt, M. Röckner, A Concise Course on Stochastic Partial Differential Equations, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
[21] J. Prüss, Evolutionary Integral Equations and Applications, Modern Birkhäuser Classics, Birkhäuser/Springer Basel AG, Basel, 1993, 2012 reprint of the 1993 edition.
[22] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, second ed., Springer Series in Computational Mathematics, vol. 25, Springer, Berlin, 2006.
[23] J.M.A.M. van Neerven, Stochastic Evolution Equations, ISEM Lecture Notes, 2008.
[24] X. Wang, An exponential integrator scheme for time discretization of nonlinear stochastic wave equation, arXiv:1312.5185, 2013.
[25] X. Wang, Weak error estimates of the exponential Euler scheme for semi-linear SPDEs without Malliavin calculus, arXiv:1408.0713, 2014.
[26] X. Wang, 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.