Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 123 (2013) 1563–1587

stochastic processes and their applications

www.elsevier.com/locate/spa

L^p and almost sure convergence of a Milstein scheme for stochastic partial differential equations[☆]

Andrea Barth[a,b], Annika Lang[a,c,*]

[a] ETH Zürich, Seminar für Angewandte Mathematik, Rämistrasse 101, CH–8092 Zürich, Switzerland
[b] Centre of Mathematics for Applications (CMA), University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway
[c] Fakultät für Mathematik und Informatik, Universität Mannheim, D–68131 Mannheim, Germany

Received 12 July 2011; received in revised form 2 July 2012; accepted 2 January 2013
Available online 9 January 2013

Abstract

In this paper, L^p convergence and almost sure convergence of the Milstein approximation of a partial differential equation of advection–diffusion type driven by a multiplicative continuous martingale is proven. The (semidiscrete) approximation in space is a projection onto a finite dimensional function space. The considered space approximation has to have an order of convergence fitting to the order of convergence of the Milstein approximation and the regularity of the solution. The approximation of the driving noise process is realized by the truncation of the Karhunen–Loève expansion of the driving noise according to the overall order of convergence. Convergence results in L^p and almost sure convergence bounds for the semidiscrete approximation as well as for the fully discrete approximation are provided.

Keywords: Stochastic partial differential equation; L^p convergence; Almost sure convergence; Milstein scheme; Galerkin method; Finite Element method; Backward Euler scheme; Advection–diffusion equation

1. Introduction

The numerical study of stochastic partial differential equations is a relatively recent topic. This is in contrast with the abundance of research (see e.g. [24]) that has been conducted for

☆ This research was supported in part by the European Research Council under grant ERC AdG 247277. The authors wish to express many thanks to Pao-Liu Chow, Arnulf Jentzen, Jürgen Potthoff, Markus Riedle, and Florian Werner for fruitful discussions and helpful comments.

* Corresponding author at: ETH Zürich, Seminar für Angewandte Mathematik, Rämistrasse 101, CH–8092 Zürich, Switzerland.

E-mail addresses: andrea.barth@sam.math.ethz.ch (A. Barth), annika.lang@sam.math.ethz.ch (A. Lang).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2013.01.003

real-valued stochastic differential equations or partial differential equations (e.g. [5,13,36]). In contrast to partial differential equations, for stochastic partial differential equations we have different notions of convergence. Pathwise convergence plays a central role in filtering theory and other phenomena in physics. For instance, the strong convergence of the second moment gives a bound for the expected error. The strong convergence of higher moments, and not only of the variance, is for pathwise approximations essential.

For a numerical treatment of stochastic partial differential equations, approximation has to be done in space, in time, and possibly of the driving noise process. In this paper, we study a Milstein scheme for the time approximation of the solution of a stochastic partial differential equation of the form

$$dX(t) = (A + B)X(t)\,dt + G(X(t))\,dM(t), \qquad X(0) = X_0. \tag{1.1}$$

Here, M is a continuous, square integrable martingale with values in a separable Hilbert space U. Probably the most popular example of such stochastic processes are Wiener processes. The linear operators A and B act on a dense subset of a separable Hilbert space H and the linear operator G is a mapping from H into the linear operators from U to H (detailed definitions and properties of A, B, and G are given in the next section).

Approximation schemes, like the Euler–Maruyama or Milstein scheme, are approximations of the stochastic integral of a stochastic differential equation which are derived from the Itô–Taylor expansion (see [24]). The Euler–Maruyama scheme has strong convergence of order $\mathrm{O}(\sqrt{k_n})$, where k_n denotes the time discretization step size, while the corresponding Milstein scheme converges with order $\mathrm{O}(k_n)$. In [28,29], a Milstein scheme was derived for a stochastic partial differential equation as introduced here, but driven by a Q-Wiener process W. The authors showed that the approximation, which was obtained by recursive insertion of the mild form of the stochastic partial differential equation, converges in L^2 and almost surely of order $\mathrm{O}(k_n)$. They derive a Milstein scheme which has two more terms than the Euler–Maruyama scheme. Here, we show that only one additional term is needed to derive the Milstein scheme from the Euler–Maruyama scheme. The same estimates apply to the calculations in [28,29] as remarked in [20]. Like in the case of a real-valued stochastic differential equation, a term treating the iterated stochastic integrals has to be added. In this paper, we use a similar scheme for the time discretization, where the driving noise is a continuous, square integrable martingale.

For the approximation in space, we project the solution on a finite dimensional subspace of the infinite dimensional solution space H. This approach can be numerically realized by a Galerkin projection. These approximations are typically implemented as Finite Element methods. So far Galerkin methods are mainly used for partial differential equations (cf. [36,13,35]) but first applications to stochastic partial differential equations have been done e.g. in [2,6,8,9,26] and references therein. The approximation of mild solutions with colored noise has been treated e.g. in [2,14,17,18,25,27,28,30,38] and references therein. First approaches to higher order approximation schemes using Taylor expansions were done e.g. in [19] with additive space–time white noise. Galerkin methods lead to pathwise approximations, also called strong approximations. Here, we combine this type of discretization with a higher order time discretization. Those approximations exhibit high order of convergence of the fully discrete Milstein approximation, while the regularity assumptions are minimal.

In most of these references, parabolic equations with (possibly) nonlinear terms are studied. Here we study an advection–diffusion type equation. An additive nonlinearity would not give rise to any additional difficulty in the approximation, as long as certain linear growth and Lipschitz conditions are fulfilled and the driving noise process is a continuous martingale.

The main result in this paper is: assume that Eq. (1.1) is approximated by a scheme which converges for the corresponding homogeneous, parabolic, deterministic problem with accuracy $\mathrm{O}((h+k_n^{1/2})^\alpha)$, for $\alpha \in \mathbb{N}$, to the solution of the homogeneous problem. Here, h denotes the space discretization step size and α is a regularity parameter. Then, the approximated stochastic partial differential equation converges with order $\mathrm{O}(h^\alpha + k_n^{\min(\alpha/2,1)})$ in L^p. Further, it converges almost surely to the mild solution of Eq. (1.1) with order $\mathrm{O}((h^2+k_n)^{1-\epsilon})$ for any $\epsilon > 0$ and the optimal choice $\alpha = 2$. Namely, we prove convergence results with minimal regularity assumptions on the initial condition. Higher regularity leads to higher order of convergence up to a convergence of order $\mathrm{O}(k_n)$ in time, which is the maximal convergence of a Milstein approximation.

The advection–diffusion type of the equation studied in this paper appears, among various phenomena in physics, in the study of Zakai's equation (cf. [39]). The stochastic partial differential equation of Zakai type, which was introduced by Zakai for a nonlinear filtering problem, reads, extended to continuous square integrable martingales,

$$du_t(x) = L^* u_t(x)\, dt + G(u_t(x))\, dM_t(x). \tag{1.2}$$

In the framework of this paper, the equation is considered on a bounded domain $D \subset \mathbb{R}^d$, with zero Dirichlet boundary conditions on the Lipschitz boundary ∂D and initial condition $u_0(x) = v(x)$. L^* is a second order elliptic differential operator of the form

$$L^* u = \frac{1}{2} \sum_{i,j=1}^{d} \partial_i \partial_j a_{ij} u - \sum_{i=1}^{d} \partial_i f_i u,$$

for $u \in C_c^2(D)$ and it can be split into the operators A and B in Eq. (1.1) for convenience of possible simulations. Originally, the operator G in Eq. (1.2) denotes a pointwise multiplication with a suitable function $g \in H$. This setting is included in the more general assumptions on G in Eq. (1.1), which we introduce in detail in the next section.

This work is organized as follows: Section 2 sets up the framework of this paper and contains a detailed analysis of Eq. (1.1). The introduction of the used discretization schemes for the space, time, and noise approximation is given in Section 3. In Section 4, a proof that the semidiscrete approximation (discretized in space) converges in L^p of order $\mathrm{O}(h^\alpha)$ and almost surely of order $\mathrm{O}(h^{\alpha-\epsilon})$ is provided. To have a more general result, we derive convergence rates in dependence of a regularity parameter α. Finally, L^p convergence of order $\mathrm{O}(h^\alpha + k_n^{\min(\alpha/2,1)})$ of the fully discrete Milstein type scheme including the noise approximation is proven in Section 5, as well as almost sure convergence of order $\mathrm{O}((h^2 + k_n)^{1-\epsilon})$.

2. Framework

Let H denote the Hilbert space $L^2(D)$, where $D \subset \mathbb{R}^d, d \in \mathbb{N}$, is a bounded domain with piecewise smooth boundary ∂D and let the subspaces H^α be the corresponding Sobolev spaces for $\alpha \in \mathbb{N}$ and H_0^α those with elements that satisfy zero Dirichlet boundary conditions, respectively. To simplify the notation we set for $\alpha = 0$, $H^0 = H$. We are interested in developing a numerical algorithm to estimate the solution of the stochastic partial differential equation

$$dX(t) = (A + B)X(t)\, dt + G(X(t))\, dM(t) \tag{2.1}$$

on the time interval $\tau := [0, T]$, where $T < +\infty$, with initial condition $X(0) = X_0$ and zero Dirichlet boundary conditions on ∂D. M is a continuous, square integrable martingale on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which satisfies the "usual conditions", with values in a separable Hilbert space $(U, (\cdot,\cdot)_U)$. The space of all continuous, square integrable martingales on U with respect to $(\mathcal{F}_t)_{t\geq 0}$ is denoted by $\mathcal{M}_c^2(U)$. We restrict ourselves to the following class of square integrable martingales

$$\begin{aligned}\mathcal{M}_{b,c}^2 &:= \{M \in \mathcal{M}_c^2(U) : \exists\, Q \in L_1^+(U) \text{ s.t. } \forall t \geq s \geq 0,\\ &\qquad \langle\!\langle M, M\rangle\!\rangle_t - \langle\!\langle M, M\rangle\!\rangle_s \leq (t-s)Q\},\end{aligned}$$

where $L_1^+(U)$ denotes the space of all linear, nuclear, symmetric, nonnegative-definite operators acting on U. The operator angle bracket process $\langle\!\langle M, M\rangle\!\rangle_t$ is defined as

$$\langle\!\langle M, M\rangle\!\rangle_t := \int_0^t Q_s \, d\langle M, M\rangle_s,$$

where $\langle M, M\rangle_t$ denotes the unique angle bracket process from the Doob–Meyer decomposition. The process $(Q_s,\ s \geq 0)$ is called the martingale covariance.

Since $Q \in L_1^+(U)$, there exists an orthonormal basis $(e_n,\ n \in \mathbb{N})$ of U consisting of eigenvectors of Q. Therefore we have the representation $Qe_n = \gamma_n e_n$, where $\gamma_n \geq 0$ is the eigenvalue corresponding to e_n. The square root of Q is defined as

$$Q^{1/2}\psi := \sum_n (\psi, e_n)_U \, \gamma_n^{1/2} \, e_n,$$

for $\psi \in U$, and $Q^{-1/2}$ is the pseudo inverse of $Q^{1/2}$. Let us denote by $(\mathcal{H}, (\cdot,\cdot)_{\mathcal{H}})$ the Hilbert space defined by $\mathcal{H} := Q^{1/2}(U)$ endowed with the inner product given by $(\psi, \phi)_{\mathcal{H}} = (Q^{-1/2}\psi, Q^{-1/2}\phi)_U$, for $\psi, \phi \in \mathcal{H}$. Let $L_{HS}(\mathcal{H}, H)$ refer to the space of all Hilbert–Schmidt operators from $\mathcal{H}$ to H and $\|\cdot\|_{L_{HS}(\mathcal{H},H)}$ denote the corresponding norm. The canonical example of a process that belongs to $\mathcal{M}_{b,c}^2$ is a Q-Wiener process, but, in general, a stochastic covariance process would be possible.

In what follows, we introduce a Burkholder–Davis–Gundy type inequality as a generalization of the Itô isometry for square integrable martingales of class $\mathcal{M}_{b,c}^2$. Let $\mathbb{L}_{\mathcal{H},\tau}^2(H) := L^2(\Omega \times \tau;\ L_{HS}(\mathcal{H}, H))$ be the space of integrands, defined over the measure space $(\Omega \times \tau,\ \mathcal{P}_\tau,\ \mathbb{P} \otimes d\lambda)$, where $\mathcal{P}_\tau$ denotes the σ-field of predictable sets in $\Omega \times \tau$ and $d\lambda$ is the Lebesgue measure. Then, by Eq. (1.6) in [16], we have as a generalization of Proposition 8.16 in [33], for $p > 0$ and for every $\Psi \in \mathbb{L}_{\mathcal{H},\tau}^2(H)$, a Burkholder–Davis–Gundy type inequality

$$\mathbb{E}\left(\sup_{t\in\tau}\left\|\int_0^t \Psi(s)\, dM(s)\right\|_H^p\right) \leq C_p\, \mathbb{E}\left(\left(\int_0^T \|\Psi(s)\|_{L_{HS}(\mathcal{H},H)}^2\, ds\right)^{p/2}\right). \tag{2.2}$$

For a full introduction to Hilbert-space-valued stochastic differential equations, we refer the reader to [33,11,7,34].

The operators A and B in Eq. (2.1) are defined as follows. We assume that the functions a_{ij}, for $i, j = 1, \ldots, d$, are twice continuously differentiable on D with continuous extension to the closure $\bar{D}$. The operator A is the unique self-adjoint extension of the differential operator

$$\frac{1}{2}\sum_{i,j=1}^d \partial_i(a_{ij}\, \partial_j u), \quad u \in C_c^2(D).$$

B is a first order differential operator given by

$$Bu := \sum_{i=1}^{d} \partial_i (b_i u), \quad u \in C_c^1(D),$$

with elements b_i that are defined as

$$b_i := \frac{1}{2} \sum_{j=1}^{d} \partial_j a_{ij} - f_i,$$

where the functions f_i, $i = 1, \ldots, d$, are continuously differentiable on D with continuous extension to $\bar{D}$. Defined this way, we also include the differential operator L^* in Eq. (1.2).

With the following assumptions, the right hand side of Eq. (2.1) is well defined and its solution has certain regularity properties which are shown later. From here on, let the smoothness parameter $\alpha \in \mathbb{N}$ be fixed.

Assumption 2.1. The coefficients of A and B, the operator G, and the initial condition X_0 satisfy the following conditions:

(a) for $i, j = 1, \ldots, d$, the elements a_{ij} belong to $C_b^{\alpha+1}(D)$ and f_i to $C_b^{\alpha}(D)$ with continuous extensions to $\bar{D}$,
(b) there exists $\delta > 0$ such that for all $x \in D$ and $\xi \in \mathbb{R}^d$

$$\sum_{i,j=1}^{d} a_{ij}(x) \xi_i \xi_j \geq \delta \|\xi\|_{\mathbb{R}^d}^2,$$

(c) X_0 is $\mathcal{F}_0$-measurable and $\mathbb{E}(\|X_0\|_{H^\alpha}^p) < +\infty$ for chosen $p > 0$,
(d) G is a linear mapping from H into $L(U, H)$ that satisfies for $C > 0$ that for $0 \leq \beta \leq \alpha$ and $\phi \in H^\beta$

$$\|G(\phi)\|_{L_{HS}(\mathcal{H}, H^\beta)} \leq C \|\phi\|_{H^\beta}.$$

Assumption 2.1(b) implies that the operator A is dissipative; see e.g. [23]. Then, by the Lumer–Phillips theorem, e.g. [12], A generates a strongly continuous contraction semigroup on H which we denote by $S = (S(t),\ t \geq 0)$. Furthermore, by Corollary 2 in [22], S is analytic in the right half-plane. Therefore, fractional powers of $-A$ are well defined, cf. [12], and we denote for simplicity $A_{-\beta} = (-A)^{-\beta}$ and $A_\beta = A_{-\beta}^{-1}$ for $\beta > 0$.

In this context we shall make use of the following lemma – whose statement is also known as *Kato's conjecture* – which was proven in [1].

Lemma 2.2. *The domain of $A_{1/2}$ is $\mathcal{D}(A_{1/2}) = H_0^1$ and the norm $\|A_{1/2} \cdot\|_H$ is equivalent to $\|\cdot\|_{H^1}$, i.e., there exists $C > 0$ such that*

$$\|A_{1/2}\phi\|_H \leq C \|\phi\|_{H^1} \quad \text{and} \quad \|\phi\|_{H^1} \leq C \|A_{1/2}\phi\|_H,$$

for all $\phi \in H^1$.

To simplify the notation in the preceding, we introduce the following norm for an H-valued random variable Φ with finite p-th moment

$$\|\Phi\|_{H,L^p} := \Big(\mathbb{E}\big(\|\Phi\|_H^p\big)\Big)^{1/p}.$$

Furthermore, we abbreviate for $p > 0$ the norm in $C(\tau; L^p(\Omega; H))$ with

$$\|\Psi\|_{H,L^p,\infty_\tau} := \sup_{t\in\tau} \|\Psi(t)\|_{H,L^p}$$

and the one in $L^p(\Omega; C(\tau; H))$ with

$$\|\Psi\|_{H,\infty_\tau,L^p} := \mathbb{E}\big(\sup_{t\in\tau} \|\Psi(t)\|_H^p\big)^{1/p},$$

for a stochastic process $\Psi = (\Psi(t),\ t \in \tau)$ with finite p-th moment for all $t \in \tau$. For $\phi : \tau \to H$, we set

$$\|\phi\|_{H,\infty_\tau} := \sup_{t\in\tau} \|\phi(t)\|_H,$$

accordingly. We refer to subintervals of τ by $\tau_s := [0, s]$ for $s \leq T$.

Assumption 2.1 also implies by results in Chapter 9 in [33], that Eq. (2.1) has a unique mild solution in H^α, i.e.,

$$\|X(t)\|_{H^\alpha,L^2,\infty_{[0,T]}} < +\infty,$$

for all $T \in (0, +\infty)$, and the stochastic partial differential equation can be rewritten for all $t > 0$ in mild form

$$X(t) = S(t)X_0 + \int_0^t S(t-s)BX(s)\,ds + \int_0^t S(t-s)G(X(s))\,dM(s). \tag{2.3}$$

Those assumptions even ensure that Eq. (2.1) has a unique strong solution in H^2 by Theorem 2.3 in [31]. Furthermore, we have similarly to [15,37,16] that Eq. (2.2) implies for all $\Psi \in \mathbb{L}^2_{\mathcal{H},\tau}(H)$

$$\left\|\int_0^t S(t-s)\Psi(s)\,dM(s)\right\|^p_{H,\infty_\tau,L^p} \leq C_p\,\mathbb{E}\left(\left(\int_0^T \|\Psi(s)\|_H^2\,ds\right)^{p/2}\right). \tag{2.4}$$

Before we introduce the approximation schemes, we provide two lemmas on the properties of the solution of Eq. (2.1) that are needed in later proofs. The first gives some insight on the space regularity of the mild solution. Under certain regularity assumptions on the initial condition of the stochastic partial differential equation, we have regularity of the mild solution. The second lemma gives a regularity result for the mild solution in time, i.e., Hölder continuity of order $1/2$ is shown. This result is necessary for the convergence proof of the approximation schemes.

Lemma 2.3. *Under Assumption 2.1, the mild solution satisfies for $p > 0$ and $\beta \leq \alpha$*

$$\|X\|_{H^\beta,L^p,\infty_\tau} \leq \|X\|_{H^\beta,\infty_\tau,L^p} < +\infty.$$

Proof. From here on, C denotes varying constants depending on p and T. We consider $p > 2$, for $p \leq 2$ the result follows by Hölder's inequality. We estimate

$$\begin{aligned}\|X\|^p_{H^\beta,\infty_\tau,L^p} = \Big\| S(t)X_0 &+ \int_0^t S(t-s)BX(s)\,ds\\ &+ \int_0^t S(t-s)G(X(s))\,dM(s)\Big\|^p_{H^\beta,\infty_\tau,L^p}\end{aligned}$$

$$
\begin{aligned}
&\le C\Bigg(\|X_0\|^p_{H^\beta,L^p} + \bigg\|\int_0^t S(t-s)BX(s)\,ds\bigg\|^p_{H^\beta,\infty_\tau,L^p} \\
&\quad + \bigg\|A_{\beta/2}\int_0^t S(t-s)G(X(s))\,dM(s)\bigg\|^p_{H,\infty_\tau,L^p}\Bigg) \\
&\le C\left(\|X_0\|^p_{H^\beta,L^p} + \mathbb{E}\left(\sup_{t\in\tau}\left(\int_0^t \|S(t-s)BX(s)\|_{H^\beta}\,ds\right)^p\right)\right. \\
&\quad \left. + \mathbb{E}\left(\left(\int_0^T \|A_{\beta/2}G(X(s))\|^2_{L_{HS}(\mathcal{H},H)}\,ds\right)^{p/2}\right)\right) \\
&\le C\left(\|X_0\|^p_{H^\beta,L^p} + \mathbb{E}\left(\sup_{t\in\tau}\left(\int_0^t (t-s)^{-1/2}\|X(s)\|_{H^\beta}\,ds\right)^p\right)\right. \\
&\quad \left. + \mathbb{E}\left(\left(\int_0^T \|X(s)\|^2_{H^\beta}\,ds\right)^{p/2}\right)\right) \\
&\le C\left(\|X_0\|^p_{H^\beta,L^p} + 2\int_0^T \|X\|^p_{H^\beta,\infty_{\tau_s},L^p}\,ds\right),
\end{aligned}
$$

where we used the boundedness of the contraction semigroup in the first and Eq. (2.4) in the second step, Lemma 2.2, Theorem 6.13 in [32], and the definition of the Bochner integral in the third one, and Hölder's inequality in the fourth. Finally, we apply Gronwall's inequality which yields

$$\|X\|^p_{H^\beta,\infty_\tau,L^p} \le C\,\|X_0\|^p_{H^\beta,L^p}\exp(2CT) < +\infty,$$

since $\|X_0\|_{H^\beta,L^p}$ is finite by Assumption 2.1(c). □

Lemma 2.4. *If X is the mild solution of Eq. (2.1), then for $p > 2$ and $0 \le r \le R \le T$*

$$\|X(R) - X(r)\|^p_{H,L^p} \le C\,\|X\|^p_{H^1,L^p,\infty_\tau}\,(R-r)^{p/2}.$$

Proof. We employ Assumption 2.1, Theorem 6.13 in [32], Eq. (2.4), and Lemma 2.2 to estimate

$$
\begin{aligned}
\|X(R) - X(r)\|^p_{H,L^p} &\le 3^{p-1}\Bigg(\|(S(R-r)-\mathbb{1})X(r)\|^p_{H,L^p} \\
&\quad + \bigg\|\int_r^R S(R-s)BX(s)\,ds\bigg\|^p_{H,L^p} \\
&\quad + \bigg\|\int_r^R S(R-s)G(X(s))\,dM(s)\bigg\|^p_{H,L^p}\Bigg) \\
&\le C\left(\|A_{1/2}X(r)\|^p_{H,L^p} + 2\cdot\|X\|^p_{H,L^p,\infty_\tau}\right)(R-r)^{p/2} \\
&\le C\,\|X\|^p_{H^1,L^p,\infty_\tau}\,(R-r)^{p/2}. \quad \square
\end{aligned}
$$

3. Approximation schemes

In this section, we derive a semidiscrete and a fully discrete approximation scheme for Eq. (2.1). The convergence properties of these schemes are proven in Sections 4 and 5.

To derive a semidiscrete form of Eq. (2.1) first, we project H onto a finite dimensional subspace V_h of H, for instance a Finite Element space. This can for example be done by first discretizing D by a triangulation defined over a finite number of points. Then, let $(\mathbb{S}_h,\ h > 0)$ denote a family of Finite Element spaces, consisting of piecewise linear, continuous polynomials with respect to the family of triangulations $(\mathcal{T}_h,\ h > 0)$ of D, with mesh width h, such that $\mathbb{S}_h \to H$ for $h \to 0$ and furthermore $\mathbb{S}_h \subset H_0^1(D)$ for $h > 0$. In the general framework, let $\mathcal{V} := (V_h,\ h > 0)$ be a family of finite dimensional subspaces of H_0^1 with refinement sizes h, H-orthogonal projection P_h and norm derived from H. For $h \to 0$ the sequence $\mathcal{V}$ is supposed to be dense in H in the following sense: for all $\phi \in H$ it holds that

$$\lim_{h\to 0} \|P_h\phi - \phi\|_H = 0.$$

The semidiscrete problem is to find $X_h(t) \in V_h$ such that for $t \in \tau$

$$dX_h(t) = (A_h + P_h B)X_h(t)\,dt + P_h G(X_h(t))\,dM(t), \qquad X_h(0) = P_h X_0.$$

Here, we define the approximate operator $A_h : V_h \to V_h$ through the bilinear form

$$(-A_h\varphi_h, \psi_h)_H = B_A(\varphi_h, \psi_h) := \sum_{i,j=1}^{d} (a_{ij}\partial_j\varphi_h, \partial_i\psi_h)_H,$$

for all $\varphi_h, \psi_h \in V_h$. The operator A_h is the generator of an analytic semigroup $S_h = (S_h(t),\ t \geq 0)$ defined formally by $S_h(t) = \exp(tA_h)$, for $t \geq 0$. The semidiscrete mild solution is then given by

$$X_h(t) = S_h(t)P_h X_0 + \int_0^t S_h(t-s)P_h B X_h(s)\,ds + \int_0^t S_h(t-s)P_h G(X_h(s))\,dM(s). \tag{3.1}$$

By Assumption 2.1, S_h is self-adjoint, positive-semidefinite on H and positive-definite on V_h. We prove in Section 4 that Eq. (3.1) converges in L^p and almost surely to the mild solution of Eq. (2.3) with order $\mathrm{O}(h^\alpha)$ resp. $\mathrm{O}(h^{\alpha-\epsilon})$, for any $\epsilon > 0$.

For the time approximation, we introduce a combination of a first order time stepping method, e.g., a backward Euler approximation, and a Milstein scheme. To this end, we consider, for $n \in \mathbb{N}$, equidistant partitions $0 = t_0^n < \cdots < t_n^n = T$ of the interval τ with step size $k_n := T/n$. We set $T^n = \{t_j^n,\ j = 0, \ldots, n\}$ and refer to the norm $C(T^n; H)$ with the subscript ∞_{T^n}. For $i < n$, the subset $\{t_j^n,\ j = 0, \ldots, i\}$ of T^n is denoted by T_i^n. For $n \in \mathbb{N}$, we define the map $\pi_n : \tau \to \{t_j^n,\ j = 0, \ldots, n\}$ by $\pi_n(s) = t_j^n$, if $t_j^n \leq s < t_{j+1}^n$. Furthermore, we set $\iota_n(j) = t_j^n$ for $j = 0, \ldots, n$. Then, ι_n is a bijective map and $\kappa_n = \iota_n^{-1} \circ \pi_n$ is well defined and gives for $t \in \tau$ the index of the next smaller grid point in T^n. The approximations introduced in the following and its convergence results also apply to nonequidistant partitions as used in [3], but for the sake of simplicity, we present here an equidistant time stepping. Inserting Eq. (2.3) recursively into itself yields

$$X(t_{j+1}^n) = S(t_{j+1}^n - t_j^n)X(t_j^n) + \int_{t_j^n}^{t_{j+1}^n} S(t_{j+1}^n - s)B\Bigg(S(s - t_j^n)X(t_j^n)$$
$$+ \int_{t_j^n}^{s} S(s-r)BX(r)\,dr + \int_{t_j^n}^{s} S(s-r)G(X(r))\,dM(r)\Bigg)\,ds$$
$$+ \int_{t_j^n}^{t_{j+1}^n} S(t_{j+1}^n - s)G\Bigg(S(s - t_j^n)X(t_j^n) + \int_{t_j^n}^{s} S(s-r)BX(r)\,dr$$
$$+ \int_{t_j^n}^{s} S(s-r)G(X(r))\,dM(r)\Bigg)\,dM(s). \tag{3.2}$$

To provide some intuition regarding the structure of the approximation, we analyze the following deterministic partial differential equation with source term following [36]. We demonstrate the method for a backward Euler time stepping scheme. We remark that we are not restricted to this time stepping scheme; any scheme fulfilling certain approximation properties, as specified in Eqs. (4.1), resp. (5.1) could be used. For simplicity, we omit details on the boundary or initial conditions, since the following are just heuristics. Consider

$$\frac{dX(t)}{dt} = AX(t) + f(X(t)).$$

The time derivative is approximated by

$$\frac{dX(t)}{dt} \approx \frac{X_{j+1}^n - X_j^n}{k_n}$$

and $AX(t)$ on the right hand side by

$$AX(t) \approx AX_{j+1}^n,$$

where $X_j^n := X(t_j^n)$, for $j = 0, \ldots, n$. The source term is approximated by

$$f(X(t)) \approx f(X_j^n),$$

which is called linearization. Overall the scheme takes the following form:

$$\frac{X_{j+1}^n - X_j^n}{k_n} = AX_{j+1}^n + f(X_j^n),$$

which can be transformed into

$$X_{j+1}^n = r(k_n A)X_j^n + k_n\, r(k_n A) f(X_j^n).$$

Here, r denotes the rational approximation of the semigroup which is given by $r(\lambda) := (1-\lambda)^{-1}$, for $\lambda \neq 1$. If we apply this approximation scheme in projected form to Eq. (3.2), we may write

$$X_{j+1}^n = r(k_n A_h)X_j^n + \int_{t_j^n}^{t_{j+1}^n} r(k_n A_h)P_h B X_j^n\,ds + \int_{t_j^n}^{t_{j+1}^n} r(k_n A_h)P_h G(X_j^n)\,dM(s)$$
$$+ \int_{t_j^n}^{t_{j+1}^n} \left(r(k_n A_h)P_h G\left(\int_{t_j^n}^{s} G(X_j^n)\,dM(r)\right)\right) dM(s). \tag{3.3}$$

The three terms from Eq. (3.2)

$$\int_{t_j^n}^{t_{j+1}^n} S(t_{j+1}^n - s) B \int_{t_j^n}^{s} S(s-r) G(X(r))\, dM(r)\, ds,$$
$$\int_{t_j^n}^{t_{j+1}^n} S(t_{j+1}^n - s) B \int_{t_j^n}^{s} S(s-r) B X(r)\, dr\, ds,$$
$$\int_{t_j^n}^{t_{j+1}^n} S(t_{j+1}^n - s) G\left(\int_{t_j^n}^{s} S(s-r) B X(r)\, dr\right) dM(s)$$

have been omitted since they for themselves converge as fast as the overall achieved convergence rate of the approximation scheme, which is shown in Section 5. There, we prove that this approximation converges in L^p and almost surely to the mild solution of Eq. (2.1) with order $\mathrm{O}(h^\alpha + k_n^{\min\{\alpha/2,1\}})$ resp. $\mathrm{O}((h^2+k_n)^{1-\epsilon})$, for all $\epsilon > 0$ and the optimal choice $\alpha = 2$. Eq. (3.3) can be rewritten with respect to the functions π_n and κ_n, which were introduced with the time discretization, by

$$\begin{aligned}
X^n_{\kappa_n(t)} &= r(k_n A_h) X^n_{\kappa_n(t)-1} + \int_{\pi_n(t)-k_n}^{\pi_n(t)} r(k_n A_h) P_h B X^n_{\kappa_n(t)-1}\, ds \\
&\quad + \int_{\pi_n(t)-k_n}^{\pi_n(t)} r(k_n A_h) P_h G(X^n_{\kappa_n(t)-1})\, dM(s) \\
&\quad + \int_{\pi_n(t)-k_n}^{\pi_n(t)} \left(r(k_n A_h) P_h G\left(\int_{\pi_n(t)-k_n}^{s} G(X^n_{\kappa_n(t)-1})\, dM(r)\right)\right) dM(s) \\
&= r(k_n A_h)^{\kappa_n(t)} P_h X_0 + \int_0^{\pi_n(t)} r(k_n A_h)^{\kappa_n(t)-\kappa_n(s)} P_h B X^n_{\kappa_n(s)}\, ds \\
&\quad + \int_0^{\pi_n(t)} r(k_n A_h)^{\kappa_n(t)-\kappa_n(s)} P_h G(X^n_{\kappa_n(s)})\, dM(s) \\
&\quad + \int_0^{\pi_n(t)} \left(r(k_n A_h)^{\kappa_n(t)-\kappa_n(s)} P_h G\left(\int_{\pi_n(s)}^{s} G(X^n_{\kappa_n(s)})\, dM(r)\right)\right) dM(s),
\end{aligned}$$

for all $t \in [k_n, T]$.

Note that all random variables involved in Eq. (3.3) can be simulated in the following way. If $U = H$ and V_h contains a finite subset of the eigenbasis of M, the noise is automatically finite dimensional. Otherwise this approximation might not be suitable for simulations. In [4], it is shown for a class of Lévy processes which choices of noise approximations imply that the overall order of convergence is preserved. We follow this approach here. Therefore, let

$$\langle\langle M, M\rangle\rangle_t - \langle\langle M, M\rangle\rangle_s = (t-s)Q,$$

i.e., M is a Q-Wiener process (see [33,21]). Let us denote the Itô integral to be simulated by

$$\int_a^b P_h \Psi(s)\, dM(s)$$

with $a < b, a, b \in \tau$ and $\Psi \in \mathbb{L}^2_{\mathcal{H},\tau}(H)$. This expression can be rewritten using the Karhunen–Loève representation of M, to

$$\sum_{i=1}^{\infty} \sqrt{\gamma_i} \int_a^b P_h \Psi(s) e_i \, dM_i(s).$$

Here, the elements γ_i denote the eigenvalues of the covariance operator Q and e_i the corresponding eigenfunctions. To evaluate this expression, we might have to simulate an infinite number of continuous one-dimensional martingales M_i. One possibility to overcome this problem is to approximate the infinite dimensional process by a truncation of the series expansion, i.e., set

$$M^\kappa(t) := \sum_{i=1}^{\kappa} \sqrt{\gamma_i} \, M_i(t) \, e_i.$$

Let $(M^\kappa, \ \kappa \in \mathbb{N})$ be the sequence of truncated series expansions with covariance Q^κ that converges almost surely to the martingale M with covariance Q. We set

$$M^{c\kappa}(t) := M(t) - M^\kappa(t) = \sum_{i=\kappa+1}^{\infty} \sqrt{\gamma_i} \, M_i(t) \, e_i$$

with covariance $Q^{c\kappa} := Q - Q^\kappa$, which converges almost surely to zero. This implies for the Itô integral of $\Psi \in \mathbb{L}^2_{\mathcal{H},\tau}(H)$ that

$$\int_a^b \Psi(s) \, dM(s) - \int_a^b \Psi(s) \, dM^\kappa(s) = \int_a^b \Psi(s) \, dM^{c\kappa}(s). \tag{3.4}$$

This difference converges to zero depending on the decay of the eigenvalues $(\gamma_i, \ i \in \mathbb{N})$, which is shown in the following lemma. We omit the proof, since it is equivalent to Lemma 3.1 in [4].

Lemma 3.1. *If $\|\Psi\|_{L(U,H),\infty_{[a,b]},L^p} < +\infty$ and there exist constants $C_\nu, C, \mu > 0$ and $\nu > 1$ such that the eigenvalues satisfy $\gamma_i \le C_\nu \, i^{-\nu}$ and $\kappa \ge C \, h^{-\mu}$, then*

$$\mathbb{E}\left(\sup_{t \in [a,b]} \left\| \int_a^t \Psi(s) \, dM(s) - \int_a^t \Psi(s) \, dM^\kappa(s) \right\|_H^p \right) \le C_p \, h^{\frac{\mu(\nu-1)p}{2}},$$

for a constant C_p.

We use Lemma 3.1 to derive an error bound for the approximation of the Milstein term in Eq. (3.3)

$$\int_{t_j^n}^{t_{j+1}^n} \left(r(k_n A_h) P_h G \left(\int_{t_j^n}^{s} P_h G(X_j^n) \, dM(r) \right) \right) dM(s). \tag{3.5}$$

To simplify the notation we introduce the separable Hilbert spaces H and U. The Hilbert space H is for example $L^2(D)$ or some approximation space V_h. Further, we consider a linear map $\Gamma : H \to L(U, H)$ satisfying Assumption 2.1(d) for $\beta = 0$ and the norm in $L(U, H)$. In addition, we have a bounded map $\sigma : \tau \to L(H, H)$. For $0 \le a < b \le T$ and an H-valued adapted stochastic process $\psi = (\psi(t), \ t \in \tau)$, we rewrite Eq. (3.5) more generally as

$$\int_a^b \sigma(a) \Gamma \left(\int_a^s \Gamma(\psi(a)) \, dM(r) \right) dM(s).$$

The following error bound is proven similarly to Lemma 4.2 in [3].

Lemma 3.2. *For $n \in \mathbb{N}$, let $\sigma : T^n \to L(H,H)$, $\Gamma : H \to L_{HS}(\mathcal{H}, H)$ be linear and satisfy Assumption 2.1(d) for $\beta = 0$ and the norm in $L(U,H)$ as well as*

$$\Gamma(\Gamma(\psi)e_i)e_j = \Gamma(\Gamma(\psi)e_j)e_i,$$

for $i, j \in \mathbb{N}$. Further, let $\psi = (\psi(t),\ t \in T^n)$ be an adapted H-valued stochastic process. If

$$\mathbb{E}\left(\int_0^T \|\psi(\pi_n(s))\|_H^p \, ds\right) < +\infty$$

and there exist constants $C_\nu, C > 0$ such that the eigenvalues of the covariance operator Q of M satisfy $\gamma_i \le C_\nu i^{-\nu}$, for some $\nu > 1$ and all $i \in \mathbb{N}$, and $\kappa \ge C h^{-\mu}$, for some $\mu > 0$, then there exists a constant C_p such that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in\tau}\left\| \int_0^t \sigma(\pi_n(s))\, \Gamma\left(\int_{\pi_n(s)}^s \Gamma(\psi(\pi_n(s)))\, dM(r)\right) dM(s)\right.\right.\\
&\qquad \left.\left. - \int_0^t \sigma(\pi_n(s))\, \Gamma\left(\int_{\pi_n(s)}^s \Gamma(\psi(\pi_n(s)))\, dM^\kappa(r)\right) dM^\kappa(s)\right\|_H^p\right)\\
&\quad \le C_p\, (k_n\, h^{\mu(\nu-1)})^{p/2}.
\end{aligned}$$

To get optimal convergence rates, the noise approximation should have the same order of convergence as the spacial and temporal approximations. We couple all error contributions in Section 5. In the next section, we derive error bounds for the semidiscrete approximation.

4. Convergence of the semidiscrete approximation

In this section, we present an L^p and an almost sure convergence result for the semidiscrete approximation. We assume that for $\alpha \ge \beta \ge 0$ with $\phi \in H^\alpha$ and $t \in \tau$, we have that

$$\|(S(t) - S_h(t)P_h)\phi\|_H \le C\, h^\alpha t^{-\beta/2} \|\phi\|_{H^{\alpha-\beta}}. \tag{4.1}$$

This is for example satisfied by the Finite Element spaces $(\mathbb{S}_h, h > 0)$ as introduced before for $\alpha = 2$ (see Theorem 3.5 in [36]). In the more general setting of piecewise polynomials of degree at most $\alpha - 1$, Theorem 5.7 in [13] as well as Proposition 11.2.2 in [35] imply Eq. (4.1). We note that in the proofs of Theorems 4.1 and 4.2, Eq. (4.1) just has to be satisfied for $\beta = 0$ and $\beta = 1$. If it holds only for $\beta = 0$, the theorems stay true when the mild solution satisfies $\|X\|_{H^{\alpha+1}, L^p, \infty_\tau} < +\infty$.

The proposed space discretized equation converges uniformly, almost surely with order $\mathrm{O}(h^{\alpha-\epsilon})$ and with order $\mathrm{O}(h^\alpha)$ in L^p to the mild solution of Eq. (2.1), which is stated in the following two theorems.

Theorem 4.1. *The sequence of semidiscrete mild solutions $(X_h, h > 0)$ defined in Eq. (3.1) converges in L^p to the mild solution X of Eq. (2.1) of order $\mathrm{O}(h^\alpha)$, i.e., for all $p > 0$*

$$\|X - X_h\|_{H,\infty_\tau,L^p} \le C_p\, h^\alpha\, \|X\|_{H^\alpha,\infty_\tau,L^p}.$$

Proof. We first assume that $p > 2$. It holds that

$$\|X - X_h\|^p_{H,\infty_\tau,L^p} \leq 3^{p-1}\Bigg(\|(S - S_h P_h)X_0\|^p_{H,\infty_\tau,L^p} + \mathbb{E}\Bigg(\sup_{t\in\tau}\Bigg\|\int_0^t S(t-s)BX(s)\,ds - \int_0^t S_h(t-s)P_h BX_h(s)\,ds\Bigg\|^p_H\Bigg) + \mathbb{E}\Bigg(\sup_{t\in\tau}\Bigg\|\int_0^t S(t-s)G(X(s))\,dM(s) - \int_0^t S_h(t-s)P_h G(X_h(s))\,dM(s)\Bigg\|^p_H\Bigg)\Bigg), \tag{4.2}$$

where we applied Hölder's inequality. The first term satisfies for $\beta = 0$ by Eq. (4.1)

$$\|(S - S_h P_h)X_0\|^p_{H,\infty_\tau,L^p} \leq C\, h^{p\alpha}\, \|X_0\|^p_{H^\alpha,L^p}.$$

The second one is split into

$$\begin{aligned}&\Bigg\|\int_0^t S(t-s)BX(s)\,ds - \int_0^t S_h(t-s)P_h BX_h(s)\,ds\Bigg\|^p_{H,\infty_\tau,L^p}\\ &\leq 2^{p-1}\Bigg(\Bigg\|\int_0^t (S(t-s) - S_h(t-s)P_h)BX(s)\,ds\Bigg\|^p_{H,\infty_\tau,L^p} + \Bigg\|\int_0^t S_h(t-s)P_h B(X(s) - X_h(s))\,ds\Bigg\|^p_{H,\infty_\tau,L^p}\Bigg).\end{aligned}$$

The first of these expressions is bounded by the properties of the Bochner integral, Eq. (4.1) for $\beta = 1$, Hölder's inequality, Fubini's theorem, and Lemma 2.2 by

$$\begin{aligned}&\Bigg\|\int_0^t (S(t-s) - S_h(t-s)P_h)BX(s)\,ds\Bigg\|^p_{H,\infty_\tau,L^p}\\ &\leq Ch^{p\alpha}\sup_{t\in\tau}\left(\int_0^t (t-s)^{-p/2(p-1)}\,ds\right)^{p-1}\|BX\|^p_{H^{\alpha-1},L^p,\infty_\tau} \leq Ch^{p\alpha}\,\|X\|^p_{H^\alpha,L^p,\infty_\tau}.\end{aligned}$$

Furthermore, the second term satisfies

$$\begin{aligned}&\Bigg\|\int_0^t S_h(t-s)P_h B(X(s) - X_h(s))\,ds\Bigg\|^p_{H,\infty_\tau,L^p}\\ &\leq C\,\mathbb{E}\left(\sup_{t\in\tau}\left(\int_0^t (t-s)^{-1/2}\|X(s) - X_h(s)\|_H\,ds\right)^p\right)\end{aligned}$$

by the properties of the Bochner integral and Theorem 6.13 in [32]. Hölder's inequality for $p > 2$ leads to

$$\Bigg\|\int_0^t S_h(t-s)P_h B(X(s) - X_h(s))\,ds\Bigg\|^p_{H,\infty_\tau,L^p} \leq C\int_0^T \|X - X_h\|^p_{H,\infty_{\tau_s},L^p}\,ds.$$

So overall, we have for the second term on the right hand side in Eq. (4.2)

$$\left\| \int_0^t S(t-s) B X(s)\, ds - \int_0^t S_h(t-s) P_h B X_h(s)\, ds \right\|^p_{H,\infty_\tau,L^p}$$
$$\leq C \left(h^{p\alpha} \|X\|^p_{H^\alpha,L^p,\infty_\tau} + \int_0^t \|X - X_h\|^p_{H,\infty_{\tau_s},L^p}\, ds \right).$$

The third expression on the right hand side of Eq. (4.2) is split into the two following terms

$$\left\| \int_0^t S(t-s) G(X(s))\, dM(s) - \int_0^t S_h(t-s) P_h G(X_h(s))\, dM(s) \right\|^p_{H,\infty_\tau,L^p}$$
$$\leq 2^{p-1} \left(\left\| \int_0^t (S(t-s) - S_h(t-s) P_h) G(X(s))\, dM(s) \right\|^p_{H,\infty_\tau,L^p} \right.$$
$$\left. + \left\| \int_0^t S_h(t-s) P_h (G(X(s)) - G(X_h(s)))\, dM(s) \right\|^p_{H,\infty_\tau,L^p} \right).$$

The first of these expressions satisfies by Lemma 4.3, which is proven afterwards, and the properties of G

$$\left\| \int_0^t (S(t-s) - S_h(t-s) P_h) G(X(s))\, dM(s) \right\|^p_{H,\infty_\tau,L^p} \leq C h^{p\alpha} \|X\|^p_{H^\alpha,\infty_\tau,L^p}.$$

Eq. (2.4) yields with Hölder's inequality and Fubini's theorem for the other term

$$\left\| \int_0^t S_h(t-s) P_h (G(X(s)) - G(X_h(s)))\, dM(s) \right\|^p_{H,\infty_\tau,L^p}$$
$$\leq C \int_0^T \|G(X) - G(X_h)\|^p_{L_{HS}(\mathcal{H},H),\infty_{\tau_s},L^p}\, ds,$$

and the properties of G imply that

$$\|G(X) - G(X_h)\|^p_{L_{HS}(\mathcal{H},H),\infty_{\tau_s},L^p} \leq C \|X - X_h\|^p_{H,\infty_{\tau_s},L^p}.$$

So overall, we have due to the finiteness of $\|X\|_{H^\alpha,\infty_\tau,L^p}$ with Gronwall's inequality

$$\|X - X_h\|^p_{H,\infty_\tau,L^p} \leq C_1 h^{p\alpha} + C_2 \int_0^T \|X - X_h\|^p_{H,\infty_{\tau_s},L^p}\, ds \leq C h^{p\alpha},$$

for constants C_1, C_2, and C depending on the regularity of the mild solution, T, and p.

Finally, for $p \leq 2$, we have for any $p' > 2$ by Hölder's inequality

$$\|X - X_h\|_{H,\infty_\tau,L^p} \leq \|X - X_h\|_{H,\infty_\tau,L^{p'}} = \mathrm{O}(h^\alpha). \quad \square$$

This theorem implies almost sure convergence as stated in the next theorem.

Theorem 4.2. *For every $\epsilon > 0$ and for $h > 0$ small enough such that h decays to zero with order $\mathrm{O}(n^{-\lambda})$, for $n \in \mathbb{N}$ and fixed $\lambda > 0$,*

$$\|X - X_h\|_{H,\infty_\tau} \leq h^{\alpha-\epsilon}, \quad \mathbb{P}\text{-a.s.},$$

i.e., the family of approximations $(X_h, h > 0)$ introduced in Eq. (3.1) converges uniformly, almost surely to X of order $\mathrm{O}(h^{\alpha-\epsilon})$, for $h \to 0$.

Proof. To show almost sure convergence, we use Theorem 4.2 and Chebyshev's inequality in the following way:

$$\mathbb{P}\big(\|X_h - X\|_{H,\infty_\tau} \geq h^{\alpha-\epsilon}\big) \leq h^{-(\alpha-\epsilon)p} \, \|X_h - X\|^p_{H,\infty_\tau,L^p} \leq C_p \, h^{p\epsilon}.$$

Since $h = \mathrm{O}(n^{-\lambda})$, the corresponding series is convergent for any $p > (\epsilon\lambda)^{-1}$ and therefore, by the Borel–Cantelli lemma, we have asymptotically

$$\|X - X_h\|_{H,\infty_\tau} \leq h^{\alpha-\epsilon}, \quad \mathbb{P}\text{-a.s.},$$

which proves the theorem. □

The proof of Theorem 4.1 required a Burkholder–Davis–Gundy type result on the convergence of the approximated semigroup in combination with a stochastic integral. In this case, Eq. (2.4) cannot be applied, since this leads to a lower order of convergence.

Lemma 4.3. *For $p > 2$ and $\Psi \in \mathbb{L}^2_{\mathcal{H},\tau}(H)$*

$$\left\| \int_0^t (S_h(t-s)P_h - S(t-s))\,\Psi(s)\,dM(s) \right\|_{H,\infty_\tau,L^p} \leq C\,h^\alpha \,\|\Psi\|_{L_{HS}(\mathcal{H},H^\alpha),\infty_\tau,L^p}.$$

Proof. We closely follow the proof of Theorem 5.12 in [11]. For $0 < \nu < 1$, the following identity holds:

$$\int_s^t (t-r)^{\nu-1}(r-s)^{-\nu}dr = \frac{\pi}{\sin \nu\pi}.$$

It follows from Fubini's theorem and the semigroup property that

$$\begin{aligned}\int_0^t S(t-s)\,\Psi(s)\,dM(s) &= \frac{\sin\nu\pi}{\pi} \int_0^t \left(\int_s^t (t-r)^{\nu-1}(r-s)^{-\nu}\,dr \right) \\ &\quad \times S(t-s)\,\Psi(s)\,dM(s) \\ &= \frac{\sin\nu\pi}{\pi} \int_0^t (t-r)^{\nu-1} S(t-r) \\ &\quad \times \int_0^r (r-s)^{-\nu} S(r-s)\,\Psi(s)\,dM(s)\,dr \\ &= \frac{\sin\nu\pi}{\pi} \int_0^t (t-r)^{\nu-1} S(t-r) Y(r)\,dr,\end{aligned}$$

where $Y(r) = \int_0^r (r-s)^{-\nu} S(r-s)\,\Psi(s)\,dM(s)$. Similar calculations for the semidiscrete version lead to

$$\int_0^t S_h(t-s)P_h\,\Psi(s)dM(s) = \frac{\sin\nu\pi}{\pi} \int_0^t (t-r)^{\nu-1} S_h(t-r)P_h Y_h(r)\,dr.$$

Note that since P_h is a projection $P_h = P_h^2$. We decompose the equation to be verified in the following way

$$\begin{aligned}&\left\| \int_0^t (S_h(t-s)P_h - S(t-s))\,\Psi(s)\,dM(s) \right\|^p_{H,\infty_\tau,L^p} \\ &\quad \leq C \left| \frac{\sin\nu\pi}{\pi} \right|^p \left(\left\| \int_0^t (t-r)^{\nu-1}(S_h(t-r)P_h - S(t-r))Y(r)\,dr \right\|^p_{H,\infty_\tau,L^p} \right.\end{aligned}$$

$$
\begin{aligned}
&+ \left\| \int_0^t (t-r)^{\nu-1} S_h(t-r) P_h (Y_h(r) - Y(r))\, dr \right\|_{H,\infty_\tau,L^p}^p \Bigg) \\
&=: C \left| \frac{\sin \nu\pi}{\pi} \right|^p (\mathrm{I} + \mathrm{II}).
\end{aligned}
$$

For ease of readability C and $C_{\nu,p}$ denote varying constants that are independent of h. We approximate the two expressions separately. By the definition of the Bochner integral and Hölder's inequality, we obtain for term I and $\nu > 1/p$

$$
\begin{aligned}
\mathrm{I} &\le \mathbb{E}\left(\sup_{t\in\tau} \left(\int_0^t \|(t-r)^{\nu-1}(S_h(t-r)P_h - S(t-r))Y(r)\|_H \, dr \right)^p \right) \\
&\le C_{\nu,p}\, \mathbb{E}\left(\sup_{t\in\tau} \int_0^t \|(S_h(t-r)P_h - S(t-r))Y(r)\|_H^p \, dr \right) \\
&\le C_{\nu,p}\, h^{p\alpha} \int_0^T \|Y(t)\|_{H^\alpha,L^p}^p \, dt,
\end{aligned}
$$

where we used Eq. (4.1) in the third step. Moreover, considering $\nu < 1/2$, Eq. (2.4), Assumption 2.1, the closed graph theorem, and the commutativity of the operator and the semigroup yield

$$
\begin{aligned}
\|Y(t)\|_{H^\alpha,L^p}^p &\le C\, \mathbb{E}\left(\left\| A_{\alpha/2} \int_0^t (t-s)^{-\nu} S(t-s)\, \Psi(s)\, dM(s) \right\|_H^p \right) \\
&\le C\, \mathbb{E}\left(\left\| \int_0^t (t-s)^{-\nu} S(t-s) A_{\alpha/2}\, \Psi(s)\, dM(s) \right\|_H^p \right) \\
&\le C\, \mathbb{E}\left(\int_0^t (t-s)^{-2\nu} \|\Psi(s)\|_{L_{HS}(\mathcal{H},H^\alpha)}^2 \, ds \right)^{p/2} \\
&\le C_{\nu,p}\, \|\Psi\|_{L_{HS}(\mathcal{H},H^\alpha),\infty_\tau,L^p}^p .
\end{aligned}
$$

Altogether, we obtain

$$
\mathrm{I} \le C_{\nu,p}\, h^{p\alpha}\, \|\Psi\|_{L_{HS}(\mathcal{H},H^\alpha),\infty_\tau,L^p}^p .
$$

For term II, Hölder's inequality for $\nu > 1/p$ and the fact that $S_h(t)P_h$ is bounded imply

$$
\begin{aligned}
\mathrm{II} &\le \mathbb{E}\left(\sup_{0\le t\le T} \left(\int_0^t (t-r)^{\nu-1} \|S_h(t-r)P_h(Y_h(r) - Y(r))\|_H \, dr \right)^p \right) \\
&\le C_{\nu,p} \int_0^T \|Y_h(r) - Y(r)\|_{H,L^p}^p \, dr.
\end{aligned}
$$

We further approximate

$$
\begin{aligned}
&\|Y_h(r) - Y(r)\|_{H,L^p}^p = \mathbb{E}\left(\left\| \int_0^r (r-s)^{-\nu} (S_h(r-s)P_h - S(r-s))\, \Psi(s)\, dM(s) \right\|_H^p \right) \\
&\le \mathbb{E}\left(\int_0^r (r-s)^{-2\nu} \|(S_h(r-s)P_h - S(r-s))\, \Psi(s)\|_{L_{HS}(\mathcal{H},H)}^2 ds \right)^{p/2} \\
&\le C_{\nu,p}\, h^{p\alpha}\, \|\Psi\|_{L_{HS}(\mathcal{H},H^\alpha),\infty_\tau,L^p}^p ,
\end{aligned}
$$

where Eqs. (2.2) and (4.1) are used. Altogether this gives for term II

$$\mathrm{II} \le C_{\nu,p}\, h^{p\alpha}\, \|\Psi\|^p_{L_{HS}(\mathcal{H},H^\alpha),\infty_\tau,L^p}.$$

Choosing any $1/p < \nu < 1/2$, we finally get

$$\left\| \int_0^t (S_h(t-s)P_h - S(t-s))\,\Psi(s)\,dM(s) \right\|^p_{H,\infty_\tau,L^p} \le C\, h^{p\alpha}\, \|\Psi\|^p_{L_{HS}(\mathcal{H},H^\alpha),\infty_\tau,L^p},$$

which concludes the proof. □

5. Convergence of the fully discrete approximation

In this section, we prove L^p and almost sure convergence of the Milstein scheme introduced in Section 3. With an Euler–Maruyama scheme, in general, only convergence of rate $\mathrm{O}(\sqrt{k_n})$ in time can be achieved, whereas a Milstein scheme converges at a rate of order $\mathrm{O}(k_n)$.

We define the approximation $X^n = (X^n_j,\ j = 0,\dots,n)$ of Eq. (2.3) by the Milstein scheme introduced in Eq. (3.3). For the convergence of the approximated semigroup $r(k_n A_h)$ we assume that it is stable and there exists a constant C such that for all $0 < j \le n$ and fixed $\alpha \in \mathbb{N}$ and $\beta \in \{0, 1\}$

$$\|(S(t^n_j) - r(k_n A_h)^j P_h)\phi\|_H \le C\,(h + k_n^{1/2})^\alpha (t^n_j)^{-\beta/2} \|\phi\|_{H^{\alpha-\beta}}. \tag{5.1}$$

This is especially met by a backward Euler scheme, which is shown similarly to Theorem 7.7 in [36] with Theorems 7.3 and 3.5 in the same book. Assumption (5.1) implies similarly to Lemma 4.3 the convergence of the rational approximation of the semigroup in combination with a stochastic integral.

Lemma 5.1. *For $p > 2$, it holds that*

$$\left\| \int_0^t \big(S(t-\pi_n(s)) - r(k_n A_h)^{\kappa_n(t)-\kappa_n(s)} P_h\big) G(X(s))\,dM(s) \right\|^p_{H,\infty_{T^n},L^p} \le C\,(h + k_n^{1/2})^{p\alpha}\, \|X\|^p_{H^\alpha,\infty_\tau,L^p}.$$

Proof. Except for the fact that one applies Eq. (5.1) instead of Eq. (4.1), this proof is identical to that of Lemma 4.3 and therefore we omit it. □

The order of convergence of the fully discrete approximation to the mild solution is stated in the following theorem.

Theorem 5.2. *For $p > 0$, the sequence of approximations $(X^n,\ n \in \mathbb{N})$ defined by Eq. (3.3) converges in p-th moment to the mild solution X of Eq. (2.1) and satisfies for constants C_1 and C_2 that depend on T*

$$\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n},L^p} \le C_1 (h + k_n^{1/2})^\alpha \|X\|_{H^\alpha,\infty_\tau,L^p} + C_2\, k_n \|X\|_{H^1,L^p,\infty_\tau}.$$

Especially, for $\alpha = 2$ and $X \in H^2$, it holds that

$$\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n},L^p} = \mathrm{O}(h^2 + k_n).$$

Proof. The proof of the theorem involves numerous estimates, where the same techniques are used many times. Therefore, we derive the terms to be estimated and choose one of each type to show the techniques that are employed.

Eq. (2.3) can be rewritten for $t \in \tau$ as

$$\begin{aligned}X(t) &= S(t)X_0 + \int_0^t S(t-s)BS(s-\pi_n(s))X(\pi_n(s))\,ds\\&\quad+\int_0^t \left(S(t-s)B\int_{\pi_n(s)}^s S(s-r)BX(r)\,dr\right)ds\\&\quad+\int_0^t \left(S(t-s)B\int_{\pi_n(s)}^s S(s-r)G(X(r))\,dM(r)\right)ds\\&\quad+\int_0^t S(t-s)G\big(S(s-\pi_n(s))X(\pi_n(s))\big)\,dM(s)\\&\quad+\int_0^t \left(S(t-s)G\left(\int_{\pi_n(s)}^s S(s-r)BX(r)\,dr\right)\right)dM(s)\\&\quad+\int_0^t \left(S(t-s)G\left(\int_{\pi_n(s)}^s S(s-r)G(X(r))\,dM(r)\right)\right)dM(s)\end{aligned}$$

similarly to Eq. (3.2) as done in [28,29]. We remark that the third, the fourth, and the sixth term on the right hand side are not approximated in scheme (3.3) because they (for themselves) converge as fast as the overall approximation scheme.

For fixed $n \in \mathbb{N}$, the difference of the mild solution of Eq. (2.1) and the fully discrete approximation (3.3) is split into the initial condition, the Bochner integral, and the Itô integral terms

$$X(t_j^n) - X_j^n = (S(t_j^n) - r(k_nA_h)^jP_h)X_0 + \xi^n(j) + \eta^n(j).$$

The Bochner integral part ξ^n is split again into three parts

$$\xi^n := \xi_1^n + \xi_2^n + \xi_3^n$$

with

$$\begin{aligned}\xi_1^n(j) &:= \int_0^{t_j^n} \left(S(t_j^n-s)\,B\,S(s-\pi_n(s))X(\pi_n(s)) - r(k_nA_h)^{j-\kappa_n(s)}P_hBX^n_{\kappa_n(s)}\right)ds,\\ \xi_2^n(j) &:= \int_0^{t_j^n} \left(S(t_j^n-s)\,B\int_{\pi_n(s)}^s S(s-r)\,BX(r)\,dr\right)ds,\\ \xi_3^n(j) &:= \int_0^{t_j^n} \left(S(t_j^n-s)\,B\int_{\pi_n(s)}^s S(s-r)\,G(X(r))\,dM(r)\right)ds.\end{aligned}$$

Similarly, the stochastic integral is decomposed into

$$\eta^n := \eta_1^n + \eta_2^n + \eta_3^n$$

with

$$\eta_1^n(j) := \int_0^{t_j^n} \Big(S(t_j^n - s)\, G\big(S(s-\pi_n(s))X(\pi_n(s))\big) \\ - r(k_n A_h)^{j-\kappa_n(s)} P_h G(X^n_{\kappa_n(s)})\Big)\, dM(s),$$

$$\eta_2^n(j) := \int_0^{t_j^n} \left(S(t_j^n - s) G\left(\int_{\pi_n(s)}^s S(s-r) B X(r)\, dr \right)\right) dM(s),$$

$$\eta_3^n(j) := \int_0^{t_j^n} \left(S(t_j^n - s) G\left(\int_{\pi_n(s)}^s S(s-r) G(X(r))\, dM(r) \right) \right. \\ \left. - r(k_n A_h)^{j-\kappa_n(s)} P_h G\left(\int_{\pi_n(s)}^s G(X^n_{\kappa_n(s)})\, dM(r)\right)\right) dM(s).$$

We further split three of the terms. We may write

$$\begin{aligned} \xi_1^n(j) &= \int_0^{t_j^n} S(t_j^n - s) B(S(s-\pi_n(s)) - \mathbb{1}) X(\pi_n(s))\, ds \\ &\quad + \int_0^{t_j^n} (S(t_j^n - s) - S(t_j^n - \pi_n(s))) B X(\pi_n(s))\, ds \\ &\quad + \int_0^{t_j^n} (S(t_j^n - \pi_n(s)) - r(k_n A_h)^{j-\kappa_n(s)} P_h) B X(\pi_n(s))\, ds \\ &\quad + \int_0^{t_j^n} r(k_n A_h)^{j-\kappa_n(s)} P_h B(X(\pi_n(s)) - X^n_{\kappa_n(s)})\, ds, \end{aligned}$$

and we refer to the terms on the right hand side by $\xi_{1,i}^n(j)$ for $i = 1, \ldots, 4$. Similarly, $\eta_1^n(j)$ is split into the following four terms

$$\begin{aligned} \eta_1^n(j) &= \int_0^{t_j^n} S(t_j^n - s)\, G\big((S(s-\pi_n(s)) - \mathbb{1}) X(\pi_n(s))\big)\, dM(s) \\ &\quad + \int_0^{t_j^n} \big(S(t_j^n - s) - S(t_j^n - \pi_n(s))\big)\, G(X(\pi_n(s)))\, dM(s) \\ &\quad + \int_0^{t_j^n} \big(S(t_j^n - \pi_n(s)) - r(k_n A_h)^{j-\kappa_n(s)} P_h\big)\, G(X(\pi_n(s)))\, dM(s) \\ &\quad + \int_0^{t_j^n} r(k_n A_h)^{j-\kappa_n(s)} P_h G(X(\pi_n(s)) - X^n_{\kappa_n(s)})\, dM(s), \end{aligned}$$

and $\eta_3^n(j)$ into five terms

$$\begin{aligned} \eta_3^n(j) &= \int_0^{t_j^n} S(t_j^n - s) G\left(\int_{\pi_n(s)}^s (S(s-r) - \mathbb{1}) G(X(r))\, dM(r)\right) dM(s) \\ &+ \int_0^{t_j^n} S(t_j^n - s) G\left(\int_{\pi_n(s)}^s G(X(r) - X(\pi_n(s)))\, dM(r)\right) dM(s) \\ &+ \int_0^{t_j^n} \big(S(t_j^n - s) - S(t_j^n - \pi_n(s))\big) G\left(\int_{\pi_n(s)}^s G(X(\pi_n(s)))\, dM(r)\right) dM(s) \end{aligned}$$

$$+\int_0^{t_j^n} (S(t_j^n - \pi_n(s)) - r(k_n A_h)^{j-\kappa_n(s)} P_h) G\left(\int_{\pi_n(s)}^s G(X(\pi_n(s)))\, dM(r)\right) dM(s)$$

$$+\int_0^{t_j^n} r(k_n A_h)^{j-\kappa_n(s)} P_h G\left(\int_{\pi_n(s)}^s G(X(\pi_n(s)) - X^n_{\kappa_n(s)})\, dM(r)\right) dM(s).$$

For now, let $p > 2$. For better readability, we add t_j^n resp. j in the terms to be estimated, although it is not necessary in the norm. The initial condition is bounded by Eq. (5.1) for $\beta = 0$

$$\|(S(t_j^n) - r(k_n A_h)^j P_h) X_0\|^p_{H,\infty_{T^n},L^p} \le C(h + k_n^{1/2})^{p\alpha} \|X_0\|^p_{H^\alpha,L^p}.$$

For ξ^n and η^n we just give calculations for one term of each type of estimation to demonstrate the technique. The other terms are treated in a similar way. The first term of ξ_1^n satisfies by the properties of the Bochner integral, Lemma 2.2, and Theorem 6.13 in [32]

$$\begin{aligned}
&\|\xi^n_{1,1}(j)\|^p_{H,\infty_{T^n},L^p}\\
&\quad\le C\,\mathbb{E}\left(\sup_{0\le j\le n}\left(\int_0^{t_j^n} (t_j^n - s)^{-1/2} \|(S(s-\pi_n(s)) - \mathbb{1})X(\pi_n(s))\|_H\, ds\right)^p\right)\\
&\quad\le C\,\mathbb{E}\left(\sup_{0\le j\le n}\left(\int_0^{t_j^n} (t_j^n - s)^{-1/2} (s-\pi_n(s))^{\alpha/2} \|X(\pi_n(s))\|_{H^\alpha}\, ds\right)^p\right)\\
&\quad\le C\, k_n^{p\alpha/2}\,\mathbb{E}\left(\sup_{0\le j\le n}\left(\int_0^{t_j^n} (t_j^n - s)^{-1/2} \|X(\pi_n(s))\|_{H^\alpha}\, ds\right)^p\right).
\end{aligned}$$

Hölder's inequality and Fubini's theorem imply that

$$\begin{aligned}
\|\xi^n_{1,1}(j)\|^p_{H,\infty_{T^n},L^p} &\le C\, k_n^{p\alpha/2} \left(\int_0^T (T-s)^{-p/(p-1)2}\, ds\right)^{p-1} \int_0^T \|X(\pi_n(s))\|^p_{H^\alpha,L^p}\, ds\\
&\le C\, k_n^{p\alpha/2}\, T^{(p-2)/2} \|X\|^p_{H^\alpha,L^p,\infty_\tau}.
\end{aligned}$$

The property of the semigroup with similar estimates leads to

$$\begin{aligned}
&\|\xi^n_{1,2}(j)\|^p_{H,\infty_{T^n},L^p} + \|\eta^n_{1,1}(j)\|^p_{H,\infty_{T^n},L^p} + \|\eta^n_{1,2}(j)\|^p_{H,\infty_{T^n},L^p} + \|\eta^n_{3,1}(j)\|^p_{H,\infty_{T^n},L^p}\\
&\quad + \|\eta^n_{3,3}(j)\|^p_{H,\infty_{T^n},L^p} \le C_p\, k_n^{p\alpha/2}\, \|X\|^p_{H^\alpha,L^p,\infty_\tau},
\end{aligned}$$

where Eq. (2.4) is used for the terms labeled with η.

The convergence properties of the rational approximation of the semigroup in Eq. (5.1) imply for $\xi^n_{1,3}(j)$ for $\beta = 1$ with similar estimates as before concerning the operator B

$$\begin{aligned}
\|\xi^n_{1,3}(j)\|^p_{H,\infty_{T^n},L^p} &\le C\, T^{(p-2)/2}\, (h + k_n^{1/2})^{p\alpha} \int_0^T \|BX(\pi_n(s))\|^p_{H^{\alpha-1},L^p}\, ds\\
&\le C_p\, (h + k_n^{1/2})^{p\alpha}\, \|X\|^p_{H^\alpha,L^p,\infty_\tau}.
\end{aligned}$$

These estimates are also applied to the following terms and give with Lemma 5.1

$$\|\eta^n_{1,3}(j)\|^p_{H,\infty_{T^n},L^p} + \|\eta^n_{3,4}(j)\|^p_{H,\infty_{T^n},L^p} \le C(1 + k_n^{p/2})(h + k_n^{1/2})^{p\alpha}\, \|X\|^p_{H^\alpha,\infty_\tau,L^p}.$$

In the end, the difference of the mild solution and the approximation is estimated by their difference at previous time steps, which stems from the following calculation

$$\begin{aligned}\|\xi_{1,4}^n(j)\|_{H,\infty_{T^n},L^p}^p &\le C\,\mathbb{E}\left(\sup_{0\le j\le n}\left(\int_0^{t_j^n}(t_j^n-\pi_n(s))^{-1/2}\|X(\pi_n(s))-X_{\kappa_n(s)}^n\|_H\,ds\right)^p\right)\\ &\le C\,T^{(p-2)/2}\sum_{i=0}^{j-1}k_n\|X(t_i^n)-X_i^n\|_{H,\infty_{T_i^n},L^p}^p,\end{aligned}$$

where we used Eq. (4.2) in [26]. The stability of the semigroup approximation for $\eta_{1,4}^n(j)$ and $\eta_{3,5}^n(j)$ leads to

$$\begin{aligned}&\|\eta_{1,4}^n(j)\|_{H,\infty_{T^n},L^p}^p+\|\eta_{3,5}^n(j)\|_{H,\infty_{T^n},L^p}^p\\ &\quad\le C_p\sum_{i=0}^{j-1}k_n(1+k_n^{p/2})\,\|X(t_i^n)-X_i^n\|_{H,\infty_{T_i^n},L^p}^p\\ &\quad\le C_p\,(1+T^{p/2})\sum_{i=0}^{j-1}k_n\,\|X(t_i^n)-X_i^n\|_{H,\infty_{T_i^n},L^p}^p.\end{aligned}$$

The remaining of the approximated terms cannot be estimated with respect to α. For those, convergence is limited by the properties of the stochastic integral. We have with the regularity of the solution from Lemma 2.4, Eq. (2.4), Hölder's inequality, combined with previous estimates

$$\begin{aligned}\|\eta_{3,2}^n(j)\|_{H,\infty_{T^n},L^p}^p &\le C_p\int_0^T(s-\pi_n(s))^{(p-2)/2}\int_{\pi_n(s)}^s\|X(r)-X(\pi_n(s))\|_{H,L^p}^p\,dr\,ds\\ &\le C\,k_n^p\,\|X\|_{H^1,L^p,\infty_\tau}^p.\end{aligned}$$

The convergence for two of the remaining terms that were not approximated in Eq. (3.3) results from the upper and lower limits of the inner integral, i.e., we have

$$\|\xi_2^n(j)\|_{H,\infty_{T^n},L^p}^p+\|\eta_2^n(j)\|_{H,\infty_{T^n},L^p}^p\le C_p\,k_n^p\,\|X\|_{H^1,L^p,\infty_\tau}^p.$$

Finally, to give estimates on $\xi_3^n(j)$, we set $\Pi_n(r)=t_i^n$ for $t_{i-1}^n<r\le t_i^n$ and write

$$\begin{aligned}&\|\xi_3^n(j)\|_{H,\infty_{T^n},L^p}^p\\ &\quad=\left\|\int_0^{t_j^n}\int_0^{t_j^n}\mathbb{1}_{\{\pi_n(s)\le r\le s<\Pi_n(r)\}}S(t_j^n-s)BS(s-r)G(X(r))\,dM(r)\,ds\right\|_{H,\infty_{T^n},L^p}^p\\ &\quad=\left\|\int_0^{t_j^n}\int_0^{t_j^n}\mathbb{1}_{\{\pi_n(s)\le r\le s<\Pi_n(r)\}}S(t_j^n-s)BS(s-r)G(X(r))\,ds\,dM(r)\right\|_{H,\infty_{T^n},L^p}^p\end{aligned}$$

with a stochastic Fubini theorem (see Theorem 8.14 in [33]). Next, we apply Eq. (2.4), the properties of the Bochner integral, Hölder's inequality, and similar estimates as before to derive

$$\begin{aligned}\|\xi_3^n(j)\|_{H,\infty_{T^n},L^p}^p&=\left\|\int_0^{t_j^n}\int_r^{\Pi_n(r)}S(t_j^n-s)BS(s-r)G(X(r))\,ds\,dM(r)\right\|_{H,\infty_{T^n},L^p}^p\\ &=\left\|\int_0^{t_j^n}S(t_j^n-\Pi_n(r))\int_r^{\Pi_n(r)}S(\Pi_n(r)-s)BS(s-r)G(X(r))\,ds\,dM(r)\right\|_{H,\infty_{T^n},L^p}^p\end{aligned}$$

$$\leq C_p \int_0^T (\Pi_n(r) - r)^{p-1} \int_r^{\Pi_n(r)} \|X(r)\|^p_{H^1,L^p}\, ds\, dr$$
$$\leq C_p k_n^p \|X\|^p_{H^1,L^p,\infty_\tau}.$$

This concludes the estimates of the terms, and overall we have

$$\|X - X^n_{\kappa_n()}\|^p_{H,\infty_{T^n},L^p} \leq C_1\Big((h + k_n^{1/2})^{p\alpha}\|X\|^p_{H^\alpha,\infty_\tau,L^p} + k_n^p\|X\|^p_{H^1,L^p,\infty_\tau}\Big) + C_2\sum_{i=0}^{j-1} k_n \|X - X^n_{\kappa_n()}\|^p_{H,\infty_{T_i^n},L^p}.$$

A discrete version of Gronwall's inequality (cf. [10]) implies

$$\|X - X^n_{\kappa_n()}\|^p_{H,\infty_{T^n},L^p}$$
$$\leq C_1\Big((h + k_n^{1/2})^{p\alpha}\|X\|^p_{H^\alpha,\infty_\tau,L^p} + k_n^p\|X\|^p_{H^1,L^p,\infty_\tau}\Big)\cdot\prod_{i=0}^{j-1}\big(1 + C_2 k_n\big)$$
$$\leq C_1\Big((h + k_n^{1/2})^{p\alpha}\|X\|^p_{H^\alpha,\infty_\tau,L^p} + k_n^p\|X\|^p_{H^1,L^p,\infty_\tau}\Big)\cdot\exp(C_2 T),$$

which concludes the proof for $p > 2$. Finally, for $p \leq 2$, Hölder's inequality leads for $p' > 2$ to

$$\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n},L^p} \leq \|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n},L^{p'}} = \mathrm{O}((h + k_n^{1/2})^\alpha + k_n). \quad \square$$

Theorem 5.3. *For every $\epsilon > 0$*

$$\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n}} \leq (h^2 + k_n)^{1-\epsilon}, \quad \mathbb{P}\text{-a.s.},$$

asymptotically for h and k_n small enough such that there exists $\lambda > 0$ with $h^2 = \mathrm{O}(k_n^\lambda)$, i.e., the series of approximations $(X^n, n \in \mathbb{N})$ defined in Eq. (3.3) converges almost surely to X with order $\mathrm{O}((h^2 + k_n)^{1-\epsilon})$ for $h, k_n \to 0$.

Proof. Let $\epsilon > 0$, then Chebyshev's inequality implies with Theorem 5.2 for all $0 \leq j \leq n$ that

$$\mathbb{P}\big(\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n}} \geq (h^2 + k_n)^{1-\epsilon}\big) \leq (h^2 + k_n)^{-(1-\epsilon)p}\,\|X - X^n_{\kappa_n()}\|^p_{H,\infty_{T^n},L^p}$$
$$\leq C_p\,(h^2 + k_n)^{p\epsilon}.$$

Since $k_n = \mathrm{O}(n^{-1})$ and $h^2 = \mathrm{O}(k_n^\lambda)$ for some $\lambda > 0$, the series

$$\sum_{n=1}^\infty \mathbb{P}\big(\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n}} \geq (h^2 + k_n)^{1-\epsilon}\big) \leq C\sum_{n=1}^\infty n^{-p\varepsilon(1\wedge\lambda)}$$

converges for any $p > \epsilon^{-1}(1 \wedge \lambda)^{-1}$ and therefore, by the Borel–Cantelli lemma

$$\|X - X^n_{\kappa_n()}\|_{H,\infty_{T^n}} \leq (h^2 + k_n)^{1-\epsilon}, \quad \mathbb{P}\text{-a.s.},$$

which concludes the proof. $\square$

As a final step we combine the approximation of the noise from Lemmas 3.1 and 3.2 with Theorem 5.2.

The fully approximated scheme reads then (see [3])

$$\tilde{X}^n_{j+1} = r(k_n A_h)\tilde{X}^n_j + \int_{t^n_j}^{t^n_{j+1}} r(k_n A_h) P_h B \tilde{X}^n_j \, ds + \int_{t^n_j}^{t^n_{j+1}} r(k_n A_h) P_h G(\tilde{X}^n_j) \, dM^{\kappa}(s)$$
$$+ \int_{t^n_j}^{t^n_{j+1}} \left(r(k_n A_h) P_h G\left(\int_{t^n_j}^{s} G(\tilde{X}^n_j) \, dM^{\sqrt{\kappa}}(r) \right) \right) dM^{\sqrt{\kappa}}(s).$$

To preserve the order of convergence for given $\nu > 1$, we require $\kappa_1 \geq C_2 h^{-4/(\nu-1)}$ for the Euler–Maruyama term and $\kappa_2 \geq C_2 h^{-2/(\nu-1)}$ for the Milstein term. For an equilibrated error we use the first κ terms of the Karhunen–Loève expansion for the Euler–Maruyama term and $\sqrt{\kappa}$ terms for the Milstein term. In this sense the simulation of the Milstein term is computationally not more expensive than the Euler term. For the Milstein term we have to sum over all mixed stochastic processes, i.e., κ_2^2 resp. $\kappa_2^2/2$ terms, if we use the symmetry of Γ. If the simulation of the Euler–Maruyama term needs computational effort $\mathrm{O}(\kappa_1)$ and $\kappa_1 = \kappa_2^2$, the overall work for the Milstein term is also $\mathrm{O}(\kappa_1)$. By adding the Milstein term, we increase the order of convergence, but with the correct truncation of the Karhunen–Loève expansion the overall work does not increase. Then, the next corollary is a consequence of Theorem 5.2 and Lemmas 3.1 and 3.2.

Corollary 5.4. *The sequence of fully discrete approximations $(\tilde{X}^n,\ n \in \mathbb{N})$ converges in L^p and almost surely to the mild solution X of Eq. (2.1) and satisfies for constants C_1 and C_2 that depend on T and for $\kappa \geq C\lceil h^{-2\max(\alpha,2)/(\nu-1)}\rceil$, for fixed $C > 0$, where $\nu > 1$ with $\gamma_i \leq C_\nu i^{-\nu}$, for $i \in \mathbb{N}$ and $C_\nu > 0$,*

$$\|X - \tilde{X}^n_{\kappa_n()}\|_{H,\infty_{T^n},L^p} \leq C_1 (h + k_n^{1/2})^{\alpha} \|X\|_{H^\alpha,\infty_\tau,L^p} + C_2 k_n \|X\|_{H^1,L^p,\infty_\tau}.$$

Especially for $\alpha = 2$ and $X \in H^2$, the error is bounded by

$$\|X - \tilde{X}^n_{\kappa_n()}\|_{H,\infty_{T^n},L^p} = \mathrm{O}(h^2 + k_n).$$

Furthermore, with the same prerequisites as in Theorem 5.3 it holds asymptotically for any $\epsilon > 0$ that

$$\|X - \tilde{X}^n_{\kappa_n()}\|_{H,\infty_{T^n}} \leq (h^2 + k_n)^{1-\epsilon}, \quad \mathbb{P}\text{-a.s.}$$

Similar results also hold in the semidiscrete case. In conclusion, we see that the approximation of the noise by an appropriate truncation of the Karhunen–Loève expansion does not affect the overall order of convergence of the approximation scheme. Otherwise the convergence of the noise approximation will dominate the errors (see [4]).

References

[1] P. Auscher, P. Tchamitchian, Square roots of elliptic second order divergence operators on strongly Lipschitz domains: L^2 theory, J. Anal. Math. 90 (2003) 1–12.

[2] A. Barth, A finite element method for martingale-driven stochastic partial differential equations, Commun. Stoch. Anal. 4 (2010) 355–375.

[3] A. Barth, A. Lang, Milstein approximation for advection–diffusion equations driven multiplicative noncontinuous martingale noises, Appl. Math. Opt. 66 (2012) 387–413.

[4] A. Barth, A. Lang, Simulation of stochastic partial differential equations using finite element methods, Stochastics 84 (2012) 217–231.

[5] D. Braess, Finite Elemente, third ed., Springer, Berlin, 2002.

[6] T. Caraballo, P.E. Kloeden, The pathwise numerical approximation of stationary solutions of semilinear stochastic evolution equations, Appl. Math. Optim. 54 (2006) 401–415.
[7] P.-L. Chow, Stochastic Partial Differential Equations, in: Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2007.
[8] P.-L. Chow, J.-L. Jiang, Almost sure convergence of some approximate solutions for random parabolic equations, in: Random Partial Differential Equations, Proc. Conf., Oberwolfach/Ger. 1989, in: ISNM, vol. 102, 1991, pp. 45–54.
[9] P.-L. Chow, J.-L. Jiang, J.-L. Menaldi, Pathwise convergence of approximate solutions to Zakai's equation in a bounded domain, in: G. Da Prato, et al. (Eds.), Stochastic Partial Differential Equations and Applications. Proceedings of the Third Meeting on Stochastic Partial Differential Equations and Applications Held at Villa Madruzzo, Trento, Italy, January 1990, in: Pitman Res. Notes Math. Ser., vol. 268, Longman Scientific & Technical, Harlow, 1992, pp. 111–123.
[10] D.S. Clark, Short proof of a discrete Gronwall inequality, Discrete Appl. Math. 16 (1987) 279–281.
[11] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, in: Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.
[12] K.-J. Engel, R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, in: Graduate Texts in Mathematics, vol. 194, Springer, Berlin, 2000.
[13] C. Grossmann, H.-G. Roos, M. Stynes, Numerical Treatment of Partial Differential Equations, in: Universitext, Springer, Berlin, 2007. Revised translation of the 3rd German edition of 'Numerische Behandlung partieller Differentialgleichungen' by Martin Stynes.
[14] E. Hausenblas, Approximation for semilinear stochastic evolution equations, Potential Anal. 18 (2003) 141–186.
[15] E. Hausenblas, J. Seidler, A note on maximal inequality for stochastic convolutions, Czechoslovak Math. J. 51 (2001) 785–790.
[16] E. Hausenblas, J. Seidler, Stochastic convolutions driven by martingales: maximal inequalities and exponential integrability, Stoch. Anal. Appl. 26 (2008) 98–119.
[17] A. Jentzen, Taylor expansions for stochastic partial differential equations, Ph.D. Thesis, Johann Wolfgang Goethe Universität Frankfurt am Main, 2009.
[18] A. Jentzen, P.E. Kloeden, The numerical approximation of stochastic partial differential equations, Milan J. Math. 77 (2009) 205–244.
[19] A. Jentzen, P.E. Kloeden, Taylor expansions of solutions of stochastic partial differential equations with additive noise, Ann. Probab. 38 (2010) 532–569.
[20] A. Jentzen, M. Röckner, A Milstein scheme for SPDEs. arXiv:math.NA.1001.2751, January 2010.
[21] O. Kallenberg, Foundations of Modern Probability, Springer, New York, 1997.
[22] S. Kantorovitz, On Liu's analyticity criterion for semigroups, Semigroup Forum 53 (1996) 262–265.
[23] T. Kato, Perturbation Theory for Linear Operators, Corr. printing of the second ed., in: Grundlehren der Mathematischen Wissenschaften, vol. 132, Springer, Berlin, Heidelberg, New York, 1980.
[24] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, in: Applications of Mathematics, vol. 23, Springer, Berlin, 1992.
[25] M. Kovács, S. Larsson, F. Lindgren, Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise, Numer. Algorithms 53 (2010) 309–320.
[26] R. Kruse, Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise, IMA J. Numer. Anal. (in press). arXiv:math.NA.1103.4504, March 2011.
[27] A. Lang, Simulation of stochastic partial differential equations and stochastic active contours, Ph.D. Thesis, Universität Mannheim, 2007.
[28] A. Lang, P.-L. Chow, J. Potthoff, Almost sure convergence of a semidiscrete Milstein scheme for SPDEs of Zakai type, Stochastics 82 (2010) 315–326.
[29] A. Lang, P.-L. Chow, J. Potthoff, Erratum: almost sure convergence of a semi-discrete Milstein scheme for SPDEs of Zakai type, Stochastics 84 (2012) 561.
[30] G.J. Lord, J. Rougemont, A numerical scheme for stochastic PDEs with Gevrey regularity, IMA J. Numer. Anal. 24 (2004) 587–604.
[31] E. Pardoux, Stochastic partial differential equation for the density of the conditional law of a diffusion process with boundary, in: Stochastic Analysis, Proc. int. Conf., Evanston/Ill. 1978, 1978, pp. 239–269.
[32] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, in: Applied Mathematical Sciences, vol. 44, Springer, New York, Berlin, Heidelberg, Tokyo, 1983.
[33] S. Peszat, J. Zabczyk, Stochastic Partial Differential Equations with Lévy Noise, in: Encyclopedia of Mathematics and its Applications, vol. 113, Cambridge University Press, Cambridge, 2007.

[34] C. Prévôt, M. Röckner, A Concise Course on Stochastic Partial Differential Equations, in: Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
[35] A. Quarteroni, A. Valli, Numerical Approximation of Partial Differential Equations, second corrected printing of the first ed., in: Springer Series in Computational Mathematics, vol. 23, Springer, Berlin, 1997.
[36] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, second ed., in: Springer Series in Computational Mathematics, vol. 25, Springer, Berlin, 2006.
[37] L. Tubaro, An estimate of Burkholder type for stochastic processes defined by the stochastic integral, Stoch. Anal. Appl. 2 (1984) 187–192.
[38] Y. Yan, Galerkin finite element methods for stochastic parabolic partial differential equations, SIAM J. Numer. Anal. 43 (2005) 1363–1384.
[39] M. Zakai, On the optimal filtering of diffusion processes, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 11 (1969) 230–243.