Appl Math Optim (2012) 66:387–413
DOI 10.1007/s00245-012-9176-y

Milstein Approximation for Advection-Diffusion Equations Driven by Multiplicative Noncontinuous Martingale Noises

Andrea Barth · Annika Lang

Published online: 10 August 2012

Abstract In this paper, the strong approximation of a stochastic partial differential equation, whose differential operator is of advection-diffusion type and which is driven by a multiplicative, infinite dimensional, càdlàg, square integrable martingale, is presented. A finite dimensional projection of the infinite dimensional equation, for example a Galerkin projection, with nonequidistant time stepping is used. Error estimates for the discretized equation are derived in L^2 and almost sure senses. Besides space and time discretizations, noise approximations are also provided, where the Milstein double stochastic integral is approximated in such a way that the overall complexity is not increased compared to an Euler–Maruyama approximation. Finally, simulations complete the paper.

Keywords Finite element method · Stochastic partial differential equation · Martingale · Galerkin method · Zakai equation · Advection-diffusion PDE · Milstein scheme · Karhunen–Loève expansion · Nonequidistant time stepping

1 Introduction

The numerical study and simulation of stochastic partial differential equations has been an active field of research for the last fifteen years. Within the last decades, the

The authors wish to express many thanks to Fred Espen Benth, Claude J. Gittelson, Arnulf Jentzen, Siddhartha Mishra, Santiago Moreno Bromberg, Jürgen Potthoff for fruitful discussions and comments as well as two anonymous referees for very helpful comments. This work was supported in part by ERC AdG no. 247277.

A. Barth (✉) · A. Lang
Seminar für Angewandte Mathematik, ETH, Rämistrasse 101, 8092 Zürich, Switzerland
e-mail: andrea.barth@sam.math.ethz.ch

A. Lang
e-mail: annika.lang@sam.math.ethz.ch

extension of partial differential equations to stochastic partial differential equations has become increasingly more important in applications, especially in engineering (image analysis, surface analysis, and filtering, etc. [26, 32, 38, 40, 45]). On the other hand, in finance, finite dimensional systems of stochastic differential equations have been extended to infinite dimensional ones, i.e., to stochastic partial differential equations (see e.g., [1, 17]). In most cases, it is not possible to obtain explicit solutions to these problems. It is therefore natural to study numerical solutions of these stochastic partial differential equations.

We aim at approximating (mild) solutions of the stochastic partial differential equation given by

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

Here M is a càdlàg, square integrable martingale taking values in a separable Hilbert space U. Probably the most popular examples of such stochastic processes are Wiener processes and jump processes which are square integrable martingales. The operators A and B are second and first order differential operators on the Hilbert space $H = L^2(D)$ for a bounded domain $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$. The operator G is a mapping from H into the linear operators from U to H. The initial condition X_0 is an H-valued random variable that is independent of the driving noise process and we consider the solution on a finite time interval $\tau = [0, T]$, $T < +\infty$.

The type of equation studied in this paper appears, besides geophysical models, in the study of Zakai's equation (cf. [47]). The stochastic partial differential equation of Zakai type, which was introduced by Zakai for a nonlinear filtering problem, reads, extended to square integrable martingales,

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

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 $u_0(x) = v(x)$, for $x \in D$. In the original filtering problem, 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 is explicitly split into the operators A and B in Eq. (1.1) in Sect. 2. In the original filtering problem, 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 are discussed in detail in Sect. 2.

When it comes to strong approximations of Hilbert-space-valued stochastic differential equations, approximation has to be performed in space and time; moreover, it is likely that the noise must also be approximated. In this paper, we study for the space approximation a projection of the original problem onto a finite dimensional subspace of H, which could be done, for example, by use of a Galerkin method. Further, we employ a Milstein approach with nonequidistant time stepping for the time approximation of the solution of the stochastic partial differential equation (1.1). We derive mean square and almost sure convergence results for our approximation scheme which are of special interest for all path dependent problems, e.g., in our case (among others) in filtering problems.

So far, Galerkin methods have been mainly used for partial differential equations (cf. [19, 20, 44, 46]), but some recent applications to stochastic partial differential equations have been performed e.g., in [5, 8, 10, 12, 13, 29–31]. One can find the approximation of mild solutions with colored noise e.g., in [4, 22, 31, 32] and references therein. First approaches to higher order approximation schemes using Taylor expansions were treated e.g., in [23, 39], and [24] with additive, space-time white noise and with multiplicative, colored Wiener noise in [25] and colored, continuous martingale noise in [6]. In [25], a Milstein scheme for stochastic partial differential equations driven by Wiener noise is derived and L^2 convergence of order $1-\epsilon$, for $\epsilon > 0$, in the time discretization is shown. In most of these references, parabolic equations with (possibly) nonlinear terms are studied. Here, we treat a larger class of (possibly) noncontinuous noises and study an advection-diffusion type equation. For Wiener noise, fully discrete approximations of the solution of Eq. (1.1) were already studied in [13], while higher order schemes were presented in [35, 36] for a time approximation. Furthermore, in [6] a (semidiscrete) space approximation and a fully discrete approximation using a Galerkin method in space and a backward Euler approach in time were introduced. A space approximation for an equation driven by a—not necessarily continuous—square integrable martingale was done in [34].

Here, we combine and extend results from [6, 33, 35, 36], and [34] and derive L^2 and almost sure convergence for an approximation scheme with a not necessarily equidistant time discretization. The increased convergence of order one in the time discretization is derived by adding an extra term to the well-known Euler–Maruyama scheme, which itself just leads to convergence of order $\mathrm{O}(k^{1/2})$. In [6], L^p convergence of order $\mathrm{O}(h^2+k)$ for a space discretization of width h and a time step of size k was shown and used to prove almost sure convergence of order $\mathrm{O}((h^2+k)^{1-\epsilon})$. For a noncontinuous, square integrable martingale, problems arise in the proof of almost sure convergence, as presented in [6] due to the missing time regularity of the solution of the stochastic partial differential equation. Namely, $X(t)-X(s)$ converges with order $(t-s)^{1/p}$ in L^p for $t \to s$ and this cannot be improved (see [34]). Therefore, the optimal order of almost sure convergence cannot be achieved with an argumentation based on the Borel–Cantelli lemma, which was used in the proof of almost sure convergence in [6]. Since the proof in [6] was done for $p > 2$, and then transferred to L^2 by Hölder's inequality, even the order of convergence in L^2 would not be optimal with this strategy. Here, we combine the proof given in [6] with the arguments in [33] in order to preserve L^2 convergence of order $\mathrm{O}(h^2+k)$, even for noncontinuous martingale noises, given that the equation fulfills sufficient smoothness assumptions. Further, we derive almost sure convergence of the approximation scheme.

Besides space and time discretization, we also address the problem of the approximation of the (noncontinuous) noise. As shown in [7], the appropriate truncation of the Karhunen–Loève expansion of the noise preserves the overall order of convergence for the Euler–Maruyama term. The corresponding result for the Milstein term (with its iterated stochastic integrals) exhibits further difficulties. Fast simulation of Wiener noise can be done, for example, with Fourier techniques in $O(N \log N)$, where N is the number of space discretization points, as shown in [37]. This is essentially the same computational cost as an optimal Finite Element solver needs for a discretized homogeneous (elliptic) problem. In this paper, we prove that for a given complexity of the Euler–Maruyama term, the overall order of computational work does not increase when a Milstein term is added, although multi dimensional iterated stochastic integrals have to be calculated. The order of computational work remains the same, due to the fact that the number of terms of the Karhunen–Loève expansion, which is needed to keep the overall order of convergence, is the square root of terms needed for the Euler–Maruyama term.

More precisely, the main result of this paper is the following: Assume that Eq. (1.1) is approximated by the projected stochastic partial differential equation onto a finite dimensional subspace of H, and discretized in time with an adapted Milstein scheme. Furthermore, suppose that the approximation of the corresponding homogeneous, parabolic, deterministic problem

$$\frac{\partial}{\partial t} u = Au$$

converges with order $O(h^\alpha + k^{\alpha/2})$, for $\alpha \in \mathbb{N}$, to the solution of the homogeneous problem. Then, the approximated stochastic partial differential equation converges with order $O(h^\alpha + k^{\min(\alpha/2,1)})$ in L^2. It also converges almost surely to the mild solution of Eq. (1.1) with at least order $O(k^{(1-\epsilon)/2})$ for any $\epsilon > 0$ and for $h^2 = O(k)$. Furthermore, conditions are given such that the approximation of the noise preserves this order of convergence.

This work is organized as follows: In Sect. 2, the framework and the properties of the stochastic partial differential equation and its solution are given. Section 3 introduces the space and time approximation and its L^2 and almost sure convergences. The noise is approximated in Sect. 4, and conditions are given so that the overall order of convergence that was proven in Sect. 3 is preserved. Finally, Sect. 5 provides an example of a compensated Poisson process, which illustrates the limit of the order of convergence in the time domain. Further, this section contains simulations of the mean square convergence of the Euler–Maruyama and Milstein term.

2 Framework

Let H denote the Hilbert space $L^2(D)$ with Lebesgue measure, where $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is a bounded domain with Lipschitz boundary ∂D, and let the subspaces H^α be the Sobolev spaces for a smoothness parameter $\alpha \in \mathbb{N}$, and H_0^α the closure of $C_c^\infty(D)$ in H^α. Here, $C_c^\infty(D)$ is the space of all infinitely often differentiable functions with compact support. Consequently, H_0^1 denotes the space of all weakly differentiable functions that vanish at the boundary. We are interested in the development of a numerical approximation scheme to generate paths of the solution of the stochastic partial differential equation

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

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

$$\mathcal{M}_b^2(U) = \big\{M \in \mathcal{M}^2(U), \exists Q \in L_1^+(U) \text{ s.t. } \forall t \geq s \geq 0,\\ \langle\langle M, M\rangle\rangle_t - \langle\langle M, M\rangle\rangle_s \leq (t-s)Q\big\},$$

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 $\langle M, M\rangle_t$ is the unique angle bracket process from the Doob–Meyer decomposition. The process $(Q_s, s \geq 0)$ is called the *martingale covariance*. Examples of such processes are square integrable Lévy martingales, i.e., those Lévy martingales with Lévy measure ν that satisfies, for $\varphi \in U$,

$$\int_U \|\varphi\|_U^2\, \nu(d\varphi) < +\infty.$$

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 spectral 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 $Q^{-1/2}$ denotes the pseudo inverse of $Q^{1/2}$. Let $(\mathcal{H}, (\cdot,\cdot)_{\mathcal{H}})$ be the Hilbert space defined by $\mathcal{H} = Q^{1/2}(U)$ and endowed with the inner product $(\psi, \phi)_{\mathcal{H}} = (Q^{-1/2}\psi, Q^{-1/2}\phi)_U$ 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.

By Proposition 8.16 in [42] we have

$$\mathbb{E}\left(\left\|\int_0^t \Psi(s)\, dM(s)\right\|_H^2\right) \leq \mathbb{E}\left(\int_0^t \|\Psi(s)\|^2_{L_{HS}(\mathcal{H},H)}\, ds\right), \tag{2.2}$$

for $t \in \tau = [0, T]$ with $T < +\infty$, $M \in \mathcal{M}_b^2(U)$, and a locally bounded, predictable process $\Psi : \tau \to L_{HS}(\mathcal{H}, H)$ with

$$\mathbb{E}\left(\int_0^T \|\Psi(s)\|^2_{L_{HS}(\mathcal{H},H)}\, ds\right) < +\infty.$$

For an introduction to Hilbert-space-valued stochastic differential equations we refer the reader to [11, 16, 42, 43].

The operators A and B in Eq. (2.1) are defined as follows: We assume that the functions a_{ij}, for $i, j = 1, \dots, d$, 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, $i = 1, \dots, d$, 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 to be 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 $i, j = 1, \dots, d$, 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 $\mathbb{E}(\|X_0\|_{H^\alpha}^2) < +\infty$,

(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 $\phi \in H^\beta$

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

Assumption 2.1(b) implies that the operator A is dissipative, see e.g., [28]. Then, by the Lumer–Phillips theorem, e.g., [18], A generates a strongly continuous contraction semigroup on H which we denote by $S=(S(t),t\geq 0)$. Furthermore, by Corollary 2 in [27], S is analytic in the right half-plane. Therefore, fractional powers of $-A$ are well-defined, cf. [18], 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 [3].

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*

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

for all $\phi\in H_0^1$.

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

$$\|\Phi\|_{H,L^2}=\big(\mathbb{E}\big(\|\Phi\|_H^2\big)\big)^{1/2}.$$

Furthermore, we abbreviate the norm in $C(\tau;L^2(\Omega;H))$ with

$$\|\Psi\|_{H,L^2,\infty_\tau}=\sup_{t\in\tau}\big\|\Psi(t)\big\|_{H,L^2},$$

for a stochastic process $\Psi=(\Psi(t),t\in\tau)$ with finite second moment for all $t\in\tau$.

With these notations, Assumption 2.1 also implies, by results in Chap. 9 in [42], that Eq. (2.1) has a unique mild solution X in H^α, i.e.,

$$\|X\|_{H^\alpha,L^2,\infty_\tau}<+\infty,$$

and X admits for $t\in(0,T]$ the (mild) form

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

Furthermore, it is clear that the solution is in H^β, for all $0\leq\beta\leq\alpha$.

We remark that we could have added a nonlinearity $F(X(t))\,dt$ as done in [25]. With some smoothness assumptions, the results in the subsequent sections would still hold for continuous martingales, where F has to be treated as in [25]. If the additional smoothness assumptions are fulfilled, the additive nonlinearity does not affect the choice of approximation. For example, Galerkin methods could be used for nonlinear, parabolic equations.

3 Approximation Scheme and Order of Convergence

In this section we derive a fully discrete approximation scheme for Eq. (2.1) and prove the convergence properties of this scheme.

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 $(\mathcal{S}_h, h > 0)$ 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 such that $\mathcal{S}_h \to H$ for $h \to 0$ and furthermore $\mathcal{S}_h \subset H_0^1(D)$ 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 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

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

Furthermore, we assume that the speed of convergence is specified by

$$\big\|(P_h - \mathbb{1})\phi\big\|_H \le C h^\alpha \|\phi\|_{H^\alpha}, \tag{3.1}$$

for $\phi \in H^\alpha$. The Finite Element spaces $(\mathcal{S}_h, h > 0)$ satisfy this inequality for $\alpha \le 2$. Moreover, Eq. (3.1) is satisfied for the space of piecewise polynomials of degree at most $\alpha - 1$ on a quasi-uniform triangulation (c.f. Theorem 4.28 in [20] and Satz 6.4 in [9]).

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

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

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 $S_h = (S_h(t), t \ge 0)$ defined formally by $S_h(t) = \exp(tA_h)$, for $t \ge 0$. The càdlàg, semidiscrete mild solution is then given by

$$X_h(t) = S_h(t)P_h X_0 + \int_0^t S_h(t-s)P_h B X_h(s)\,ds + \int_0^t S_h(t-s)P_h G\big(X_h(s)\big)\,dM(s). \tag{3.2}$$

By Assumption 2.1, S_h is self-adjoint, positive-semidefinite on H and positive-definite on V_h. We assume that for $\alpha \ge \beta \ge 0$ with $\phi \in H^\beta$ and $t \in \tau$, we have that

$$\big\|\big(S(t) - S_h(t)P_h\big)\phi\big\|_H \le C h^\alpha t^{-(\alpha-\beta)/2} \|\phi\|_{H^\beta}. \tag{3.3}$$

This is for example satisfied by the Finite Element spaces $(\mathcal{S}_h, h > 0)$ as introduced before for $\alpha = 2$ (see Theorem 3.5 in [46]). In the more general setting of piecewise polynomials of degree at most $\alpha - 1$, Theorem 5.7 in [20] as well as Proposition 11.2.2 in [44] imply Eq. (3.3).

The proposed space-discretized equation converges uniformly, almost surely with order $O(h^{\alpha-\epsilon})$ and with order $O(h^{\alpha})$ in L^p for $p>0$ to the mild solution of Eq. (2.1), which was shown in [34].

For the time discretization, we propose a similar scheme to [35, 36] and a simplified version of [7], which is a combination of a linearized nonequidistant time discretization, i.e., a linear-implicit backward Euler approach, and a Milstein scheme but with one approximation term less than in [35, 36]. We introduce the following framework:

We shall always consider a finite time interval $\tau=[0,T]$ with $T<+\infty$. Let $\mathbb{T}=(\mathbb{T}_n, n\in\mathbb{N})$ be a sequence of partitions $\mathbb{T}_n$, $n\in\mathbb{N}$, of the interval τ whose mesh width tends to zero as n tends to $+\infty$. We set $\mathbb{T}_n=\{t_0^n,t_1^n,\dots,t_{l_n}^n\}$ with $l_n\in\mathbb{N}$, $0=t_0^n<t_1^n<\cdots<t_{l_n}^n=T$, step size $k_n(i)=t_{i+1}^n-t_i^n$, and maximal step size k_n in $\mathbb{T}_n$ given by

$$k_n=\max\bigl\{k_n(i), i=0,\dots,l_n-1\bigr\}.$$

For $n\in\mathbb{N}$, we define the map $\pi_n:\tau\to\{t_i^n, i=0,\dots,l_n\}$ by $\pi_n(s)=t_i^n$, if $t_i^n\le s<t_{i+1}^n$. Furthermore, we set $\iota_n(j)=t_j^n$ for $j=0,\dots,n$. 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 $\mathbb{T}_n$.

In [7], the rational approximation of the semigroup was done by $r(k_n(j)A_h)$ with $r(\lambda)=(1+\lambda/2)/(1-\lambda/2)$, $\lambda\in\mathbb{R}\backslash\{2\}$, which resembles a Crank–Nicolson method. Here, we simplify this scheme to a backward Euler scheme, i.e., $r(\lambda)=(1-\lambda)^{-1}$ for $\lambda\in\mathbb{R}\backslash\{1\}$ and generalized in the standard way to compact operators. For an equidistant time discretization, this approximation is the same for all $j=0,\dots,l_n-1$. In our approach with a variable step size, the approximations are not necessarily the same at all discretization times. To simplify the notation, we set

$$R_j^{n,h}=r\bigl(k_n(j)A_h\bigr).$$

Furthermore, for $0\le i<j\le l_n$ we denote

$$R_{(j-1:i)}^{n,h}=R_{j-1}^{n,h}R_{j-2}^{n,h}\cdots R_{i+1}^{n,h}R_i^{n,h}.$$

In case of an equidistant partition we have $R_{(j-1:i)}^{n,h}=r(k_nA_h)^{j-i+1}$. The recursive approximation scheme, which was derived in [6], reads then

$$\begin{aligned}X_j^n&=R_{j-1}^{n,h}X_{j-1}^n+\int_{t_{j-1}^n}^{t_j^n}R_{j-1}^{n,h}P_hBX_{j-1}^n\,ds\\&\quad+\int_{t_{j-1}^n}^{t_j^n}R_{j-1}^{n,h}P_hG\bigl(X_{j-1}^n\bigr)\,dM(s)\\&\quad+\int_{t_{j-1}^n}^{t_j^n}\biggl(R_{j-1}^{n,h}P_hG\biggl(\int_{t_{j-1}^n}^{s}G\bigl(X_{j-1}^n\bigr)\,dM(r)\biggr)\biggr)\,dM(s)\end{aligned}\qquad(3.4)$$

and can be rewritten as

$$X_j^n = R^{n,h}_{(j-1:0)} P_h X_0 + \int_0^{t_j^n} R^{n,h}_{(j-1:\kappa_n(s))} P_h B X^n_{\kappa_n(s)}\, ds$$
$$+ \int_0^{t_j^n} R^{n,h}_{(j-1:\kappa_n(s))} P_h G\big(X^n_{\kappa_n(s)}\big)\, dM(s)$$
$$+ \int_0^{t_j^n} \bigg(R^{n,h}_{(j-1:\kappa_n(s))} P_h G\bigg(\int_{\pi_n(s)}^s G\big(X^n_{\kappa_n(s)}\big)\, dM(r)\bigg)\bigg)\, dM(s). \quad (3.5)$$

The approximation scheme in Eq. (3.5) is not limited to a backward Euler approach, but any other time stepping scheme which fulfills the following assumption can be used.

Assumption 3.1 Assume that the approximation of the semigroup $(R^{n,h}_{(j-1:i)}, 0 \le i < j \le l_n)$ is stable, i.e., there exists a constant C such that for all $n \in \mathbb{N}$, $h > 0$, and $0 \le i < j \le l_n$

$$\big\| R^{n,h}_{(j-1:i)} P_h \big\|_{L(H)} \le C,$$

and that there exists a constant C such that for all $n \in \mathbb{N}$, $h > 0$, $0 \le i < j \le l_n$, fixed $\alpha \in \mathbb{N}$, $\beta \in \{0, 1\}$, and $\phi \in H^{\alpha-\beta}$

$$\big\|\big(S\big(t_j^n - t_i^n\big) - R^{n,h}_{(j-1:i)} P_h\big)\phi\big\|_H \le C\big(h + k_n^{1/2}\big)^\alpha \big(t_j^n - t_i^n\big)^{-\beta/2} \|\phi\|_{H^{\alpha-\beta}}. \quad (3.6)$$

This is especially met by a backward Euler scheme, which is shown similarly to Theorem 7.7 in [46] with Theorems 7.3 and 3.5 in the same book.

The order of convergence is proven in the following theorem.

Theorem 3.2 *Let Assumption 3.1 be satisfied. Then, the approximation $X^n = (X_j^n, j = 0, \dots, l_n)$ defined by Eq. (3.5) converges in mean square to the mild solution X of the stochastic partial differential equation (2.1) and satisfies for constants C_1 and C_2 that depend on T*

$$\sup_{0 \le j \le l_n} \big\|X\big(t_j^n\big) - X_j^n\big\|_{H,L^2} \le C_1\big(h^\alpha + k_n^{\alpha/2}\big)\|X\|_{H^\alpha,L^2,\infty_\tau} + C_2 k_n \|X\|_{H^1,L^2,\infty_\tau}.$$

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

$$\sup_{0 \le j \le l_n} \big\|X\big(t_j^n\big) - X_j^n\big\|_{H,L^2} = \mathrm{O}\big(h^2 + k_n\big).$$

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

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

$$X(t) = S(t)X_0 + \int_0^t S(t-s) B S\big(s - \pi_n(s)\big) X\big(\pi_n(s)\big)\, ds$$

$$+\int_0^t \left(S(t-s)B\int_{\pi_n(s)}^s S(s-r)BX(r)\,dr\right)ds$$
$$+\int_0^t \left(S(t-s)B\int_{\pi_n(s)}^s S(s-r)G\big(X(r)\big)\,dM(r)\right)ds$$
$$+\int_0^t S(t-s)G\big(S\big(s-\pi_n(s)\big)X\big(\pi_n(s)\big)\big)\,dM(s)$$
$$+\int_0^t \left(S(t-s)G\left(\int_{\pi_n(s)}^s S(s-r)BX(r)\,dr\right)\right)dM(s)$$
$$+\int_0^t \left(S(t-s)G\left(\int_{\pi_n(s)}^s S(s-r)G\big(X(r)\big)\,dM(r)\right)\right)dM(s)$$

similarly to Eq. (3.5) as done in [6, 35]. We remark that the third, the fourth, and the sixth term on the right hand side are not approximated in scheme (3.5) because they (for themselves) converge as fast as the overall approximation scheme.

For fixed $n \in \mathbb{N}$, the difference of the mild solution and the fully discrete approximation (3.5) is split into the initial condition, the Bochner integral and the Itô integral terms

$$X\big(t_j^n\big) - X_j^n = \big(S\big(t_j^n\big) - R_{(j-1:0)}^{n,h} P_h\big)X_0 + \xi^n(j) + \eta^n(j).$$

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

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

with

$$\xi_1^n(j) = \int_0^{t_j^n} \big(S\big(t_j^n - s\big)BS\big(s-\pi_n(s)\big)X\big(\pi_n(s)\big) - R_{(j-1:\kappa_n(s))}^{n,h} P_h B X_{\kappa_n(s)}^n\big)\,ds,$$
$$\xi_2^n(j) = \int_0^{t_j^n} \left(S\big(t_j^n - s\big)B\int_{\pi_n(s)}^s S(s-r)BX(r)\,dr\right)ds,$$
$$\xi_3^n(j) = \int_0^{t_j^n} \left(S\big(t_j^n - s\big)B\int_{\pi_n(s)}^s S(s-r)\,G\big(X(r)\big)\,dM(r)\right)ds.$$

Similarly, the stochastic integral is decomposed into

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

with

$$\eta_1^n(j) = \int_0^{t_j^n} \big(S\big(t_j^n - s\big)G\big(S\big(s-\pi_n(s)\big)X\big(\pi_n(s)\big)\big) - R_{(j-1:\kappa_n(s))}^{n,h} P_h G\big(X_{\kappa_n(s)}^n\big)\big)\,dM(s),$$
$$\eta_2^n(j) = \int_0^{t_j^n} \left(S\big(t_j^n - s\big)G\left(\int_{\pi_n(s)}^s S(s-r)BX(r)\,dr\right)\right)dM(s),$$

$$\eta_3^n(j) = \int_0^{t_j^n} \biggl(S\bigl(t_j^n - s\bigr) G\biggl(\int_{\pi_n(s)}^s S(s-r) G\bigl(X(r)\bigr)\, dM(r) \biggr) \\ - R^{n,h}_{(j-1:\kappa_n(s))} P_h G\biggl(\int_{\pi_n(s)}^s G\bigl(X^n_{\kappa_n(s)}\bigr)\, dM(r) \biggr) \biggr)\, dM(s).$$

The estimates for the seven terms are a combination of those presented in [6] and in [34]. Therefore, we further split three of the terms. We may write

$$\xi_1^n(j) = \int_0^{t_j^n} S\bigl(t_j^n - s\bigr) B\bigl(S\bigl(s - \pi_n(s)\bigr) - \mathbb{1}\bigr) X\bigl(\pi_n(s)\bigr)\, ds \\ + \int_0^{t_j^n} \bigl(S\bigl(t_j^n - s\bigr) - S\bigl(t_j^n - \pi_n(s)\bigr)\bigr) B X\bigl(\pi_n(s)\bigr)\, ds \\ + \int_0^{t_j^n} \bigl(S\bigl(t_j^n - \pi_n(s)\bigr) - R^{n,h}_{(j-1:\kappa_n(s))} P_h\bigr) B X\bigl(\pi_n(s)\bigr)\, ds \\ + \int_0^{t_j^n} R^{n,h}_{(j-1:\kappa_n(s))} P_h B\bigl(X\bigl(\pi_n(s)\bigr) - X^n_{\kappa_n(s)}\bigr)\, ds,$$

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

$$\eta_1^n(j) = \int_0^{t_j^n} S\bigl(t_j^n - s\bigr) G\bigl(\bigl(S\bigl(s - \pi_n(s)\bigr) - \mathbb{1}\bigr) X\bigl(\pi_n(s)\bigr)\bigr)\, dM(s) \\ + \int_0^{t_j^n} \bigl(S\bigl(t_j^n - s\bigr) - S\bigl(t_j^n - \pi_n(s)\bigr)\bigr) G\bigl(X\bigl(\pi_n(s)\bigr)\bigr)\, dM(s) \\ + \int_0^{t_j^n} \bigl(S\bigl(t_j^n - \pi_n(s)\bigr) - R^{n,h}_{(j-1:\kappa_n(s))} P_h\bigr) G\bigl(X\bigl(\pi_n(s)\bigr)\bigr)\, dM(s) \\ + \int_0^{t_j^n} R^{n,h}_{(j-1:\kappa_n(s))} P_h G\bigl(X\bigl(\pi_n(s)\bigr) - X^n_{\kappa_n(s)}\bigr)\, dM(s)$$

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

$$\eta_3^n(j) \\ = \int_0^{t_j^n} S\bigl(t_j^n - s\bigr) G\biggl(\int_{\pi_n(s)}^s \bigl(S(s-r) - \mathbb{1}\bigr) G\bigl(X(r)\bigr)\, dM(r) \biggr)\, dM(s) \\ + \int_0^{t_j^n} S\bigl(t_j^n - s\bigr) G\biggl(\int_{\pi_n(s)}^s G\bigl(X(r) - X\bigl(\pi_n(s)\bigr)\bigr)\, dM(r) \biggr)\, dM(s) \\ + \int_0^{t_j^n} \bigl(S\bigl(t_j^n - s\bigr) - S\bigl(t_j^n - \pi_n(s)\bigr)\bigr) G\biggl(\int_{\pi_n(s)}^s G\bigl(X\bigl(\pi_n(s)\bigr)\bigr)\, dM(r) \biggr)\, dM(s) \\ + \int_0^{t_j^n} \bigl(S\bigl(t_j^n - \pi_n(s)\bigr) - R^{n,h}_{(j-1:\kappa_n(s))} P_h\bigr) G\biggl(\int_{\pi_n(s)}^s G\bigl(X\bigl(\pi_n(s)\bigr)\bigr)\, dM(r) \biggr)\, dM(s)$$

$$+\int_0^{t_j^n} R^{n,h}_{(j-1:\kappa_n(s))} P_h G\left(\int_{\pi_n(s)}^s G\big(X\big(\pi_n(s)\big) - X^n_{\kappa_n(s)}\big)\, dM(r)\right) dM(s).$$

The initial condition is bounded by Assumption 3.1 for $\beta = 0$ by

$$\big\|\big(S\big(t_j^n\big) - R^{n,h}_{(j-1:0)} P_h\big) X_0\big\|^2_{H,L^2} \le C\big(h + k_n^{1/2}\big)^{2\alpha} \|X_0\|^2_{H^\alpha,L^2}.$$

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 [41] that

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

Hölder's inequality and Fubini's theorem imply that

$$\begin{aligned}
\big\|\xi^n_{1,1}(j)\big\|^2_{H,L^2} &\le Ck_n^\alpha \int_0^{t_j^n} \big(t_j^n - s\big)^{-1/2} ds \int_0^{t_j^n} \big(t_j^n - s\big)^{-1/2} \big\|X\big(\pi_n(s)\big)\big\|^2_{H^\alpha,L^2}\, ds \\
&\le Ck_n^\alpha\, (2\sqrt{T})^2 \|X\|^2_{H^\alpha,L^2,\infty_\tau}.
\end{aligned}$$

The property of the semigroup with similar estimates leads to

$$\begin{aligned}
&\big\|\xi^n_{1,2}(j)\big\|^2_{H,L^2} + \big\|\eta^n_{1,1}(j)\big\|^2_{H,L^2} + \big\|\eta^n_{1,2}(j)\big\|^2_{H,L^2} + \big\|\eta^n_{3,1}(j)\big\|^2_{H,L^2} + \big\|\eta^n_{3,3}(j)\big\|^2_{H,L^2} \\
&\quad \le Ck_n^\alpha \|X\|^2_{H^\alpha,L^2,\infty_\tau}.
\end{aligned}$$

The convergence properties of the approximate semigroup in Assumption 3.1 imply for $\xi^n_{1,3}(j)$ for $\beta = 1$ with similar estimates as before concerning B

$$\begin{aligned}
\big\|\xi^n_{1,3}(j)\big\|^2_{H,L^2} &\le C\frac{1}{4} 2\sqrt{T}\big(h^2 + k_n\big)^\alpha \int_0^{t_j^n} \big(t_j^n - s\big)^{-1/2} \big\|BX\big(\pi_n(s)\big)\big\|^2_{H^{\alpha-1},L^2}\, ds \\
&\le CT\big(h^2 + k_n\big)^\alpha \|X\|^2_{H^\alpha,L^2,\infty_\tau}.
\end{aligned}$$

These estimates are also applied to the following two terms and give

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

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

$$\big\|\xi^n_{1,4}(j)\big\|^2_{H,L^2} \le C\mathbb{E}\left(\left(\int_0^{t_j^n} \big(t_j^n - \pi_n(s)\big)^{-1/2} \big\|X\big(\pi_n(s)\big) - X^n_{\kappa_n(s)}\big\|_H\, ds\right)^2\right)$$

$$\leq C2\sqrt{T}\sum_{i=0}^{j-1} k_n(i)\big(t_j^n - t_i^n\big)^{-1/2}\big\|X\big(t_i^n\big) - X_i^n\big\|_{H,L^2}^2,$$

where we used Eq. (4.2) in [31]. The stability of the semigroup approximation for $\eta_{1,4}^n(j)$ and $\eta_{3,5}^n(j)$ leads to

$$\begin{aligned}\big\|\eta_{1,4}^n(j)\big\|_{H,L^2}^2 + \big\|\eta_{3,5}^n(j)\big\|_{H,L^2}^2 &\leq C\sum_{i=0}^{j-1} k_n(i)\big(1+k_n(i)\big)\big\|X\big(t_i^n\big) - X_i^n\big\|_{H,L^2}^2 \\ &\leq C(1+T)\sum_{i=0}^{j-1} k_n(i)\big\|X\big(t_i^n\big) - X_i^n\big\|_{H,L^2}^2.\end{aligned}$$

The remaining 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 2.5 in [33] and Eq. (2.2), combined with previous estimates

$$\begin{aligned}\big\|\eta_{3,2}^n(j)\big\|_{H,L^2}^2 &\leq C\int_0^{t_j^n}\int_{\pi_n(s)}^s \big\|X(r) - X\big(\pi_n(s)\big)\big\|_{H,L^2}^2\,dr\,ds \\ &\leq Ck_n^2\|X\|_{H^1,L^2,\infty_\tau}^2.\end{aligned}$$

The convergence for two of the remaining terms that were not approximated in Eq. (3.5) results from the upper and lower limit of the inner integral, i.e., we have

$$\big\|\xi_2^n(j)\big\|_{H,L^2}^2 + \big\|\eta_2^n(j)\big\|_{H,L^2}^2 \leq Ck_n^2\|X\|_{H^1,L^2,\infty_\tau}^2.$$

Finally, to give estimates on $\xi_3^n(j)$, we may write

$$\begin{aligned}\big\|\xi_3^n(j)\big\|_{H,L^2}^2 &= \left\|\sum_{i=1}^{j}\int_{t_{i-1}^n}^{t_i^n} S\big(t_j^n - s\big)B\int_{t_{i-1}^n}^s S(s-r)G\big(X(r)\big)\,dM(r)\,ds\right\|_{H,L^2}^2 \\ &= \sum_{i,k=1}^{j}\mathbb{E}\bigg(\bigg(\int_{t_{i-1}^n}^{t_i^n} S\big(t_j^n - s\big)B\int_{t_{i-1}^n}^s S(s-r)G\big(X(r)\big)\,dM(r)\,ds, \\ &\qquad \int_{t_{k-1}^n}^{t_k^n} S\big(t_j^n - s\big)B\int_{t_{k-1}^n}^s S(s-r)G\big(X(r)\big)\,dM(r)\,ds\bigg)_H\bigg).\end{aligned}$$

For $i \neq k$ the inner product is zero, since one term is independent and the other measurable with respect to the filtration at the smaller of the two time points t_i^n, t_k^n. This implies

$$\big\|\xi_3^n(j)\big\|_{H,L^2}^2 = \sum_{i=1}^{j}\left\|\int_{t_{i-1}^n}^{t_i^n} S\big(t_j^n - s\big)B\int_{t_{i-1}^n}^s S(s-r)G\big(X(r)\big)\,dM(r)\,ds\right\|_{H,L^2}^2.$$

Hölder's inequality and similar estimates as before lead to

$$\big\|\xi_3^n(j)\big\|_{H,L^2}^2 \le Ck_n \int_0^{t_j^n} \int_{\pi_n(s)}^s \big\|X(r)\big\|_{H^1,L^2}^2 \,dr\,ds \le Ck_n^2 \|X\|_{H^1,L^2,\infty_\tau}^2.$$

This concludes the estimates of the terms, and overall we have for $0 < j \le l_n$

$$\begin{aligned}\big\|X\big(t_j^n\big) - X_j^n\big\|_{H,L^2}^2 &\le C_1\big(\big(h^{2\alpha} + k_n^\alpha\big)\|X\|_{H^\alpha,L^2,\infty_\tau}^2 + k_n^2\|X\|_{H^1,L^2,\infty_\tau}^2\big)\\ &\quad + C_2\sum_{i=0}^{j-1} k_n(i)\big(1 + \big(t_j^n - t_i^n\big)^{-1/2}\big)\big\|X\big(t_i^n\big) - X_i^n\big\|_{H,L^2}^2.\end{aligned}$$

A discrete version of Gronwall's inequality (cf. [14]) implies

$$\begin{aligned}\big\|X\big(t_j^n\big) - X_j^n\big\|_{H,L^2}^2 &\le C_1\big(\big(h^{2\alpha} + k_n^\alpha\big)\|X\|_{H^\alpha,L^2,\infty_\tau}^2 + k_n^2\|X\|_{H^1,L^2,\infty_\tau}^2\big)\\ &\quad \times \prod_{i=0}^{j-1}\big(1 + C_2\,k_n(i)\,\big(1 + \big(t_j^n - t_i^n\big)^{-1/2}\big)\big)\\ &\le C_1\big(\big(h^{2\alpha} + k_n^\alpha\big)\|X\|_{H^\alpha,L^2,\infty_\tau}^2 + k_n^2\|X\|_{H^1,L^2,\infty_\tau}^2\big)\\ &\quad \times \exp\big(C_2(T + 2\sqrt{T})\big),\end{aligned}$$

which concludes the proof. □

This theorem entails two remarks. The first comments on the choice of α.

Remark 3.3 The result of Theorem 3.2 implies that k_n is a sharp bound for the convergence rate in time due to the properties of the mild solution of Eq. (2.1) while the convergence in space depends on the regularity of the solution. If the mild solution is in H^α, the approximation converges with h^α. Especially, if the solution is in H^1, the approximation scheme still converges with order $h + k_n^{1/2}$, however in this case an Euler–Maruyama scheme, which leads to the same error bound, would be simpler and faster to simulate.

Furthermore, we remark on L^p convergence for $p > 2$ which was proven for continuous martingales in [6]. This type of convergence can also be proven for noncontinuous martingales but the order of convergence depends on p. Therefore, a proof of L^2 convergence by L^p convergence for $p > 2$ as done in [35, 36] and [6] would lead to a rate of $\mathrm{O}(k_n^{2/p})$.

Remark 3.4 The estimates in Theorem 3.2 can also be done in L^p for $p > 2$ with adjusted preliminaries, but the order of convergence is then dependent on p. This is due to the fact that for a stochastic integral of Itô type with respect to a noncontinuous, square integrable martingale a Burkholder–Davis–Gundy type inequality reads

$$\mathbb{E}\bigg(\bigg\|\int_r^t \Phi(s)\,dM(s)\bigg\|_H^p\bigg) \le C\mathbb{E}\bigg(\int_r^t \big\|\Phi(s)\big\|_{L(U,H)}^p\,ds\bigg),$$

while we have

$$\mathbb{E}\left(\left\|\int_r^t \Phi(s)\,dM(s)\right\|_H^p\right) \le C\mathbb{E}\left(\left(\int_r^t \|\Phi(s)\|_{L(U,H)}^2\,ds\right)^{p/2}\right)$$

for continuous martingales. This implies that, if the convergence results from the limits of the integral, we obtain $t-r$ for all $p \ge 2$ instead of $(t-r)^{p/2}$. This is illustrated in Sect. 5, where we derive error bounds for a compensated Poisson process.

Overall, Theorem 3.2 transforms for $p > 2$ to

$$\sup_{0\le j\le l_n} \left\|X\left(t_j^n\right) - X_j^n\right\|_{H,L^p} \le C_1\left(h^\alpha + k_n^{\alpha/2}\right)\|X\|_{H^\alpha,L^p,\infty_\tau} + C_2 k_n^{2/p}\|X\|_{H^1,L^p,\infty_\tau},$$

for some positive constants C_1 and C_2 that depend on α.

Theorem 3.2 implies almost sure convergence as stated in the next theorem. The rate of convergence in the time domain is at least $1/2-\epsilon$. To our knowledge, so far it is not known how to prove higher almost sure convergence rates for this type of Milstein scheme, which is e.g., of interest in filtering problems.

Theorem 3.5 *Let C_1 and C_2 be constants such that for all $n \in \mathbb{N}$, $k_n \le C_1 T/n$, and $h^2 = C_2 k_n$. Then, for $\alpha = 2$, $(X^n, n \in \mathbb{N})$ converges almost surely to X, i.e.,*

$$\lim_{n\to\infty} \sup_{0\le j\le l_n} \left\|X\left(t_j^n\right) - X_j^n\right\|_H = 0 \quad \mathbb{P}\text{-a.s.}$$

Proof Let $\epsilon > 0$, then Chebyshev's inequality implies with Theorem 3.2 for all $0 \le j \le l_n$ that

$$\mathbb{P}\left(\left\|X\left(t_j^n\right) - X_j^n\right\|_H \ge k_n^{(1-\epsilon)/2}\right) \le k_n^{-(1-\epsilon)}\left\|X\left(t_j^n\right) - X_j^n\right\|_{H,L^2}^2 \le C k_n^{1+\epsilon},$$

since $h^2 = C_2 k_n$. Furthermore, since $k_n \le C_1 T/n$, the corresponding series is convergent and therefore by the Borel–Cantelli lemma we get that for all $0 \le j \le l_n$ asymptotically

$$\left\|X\left(t_j^n\right) - X_j^n\right\|_H \le k_n^{(1-\epsilon)/2}, \quad \mathbb{P}\text{-a.s.},$$

i.e., there exists a $\mathbb{P}$-null set N_j such that for all ω in the complement N_j^c and $n \ge n_0(\omega)$ for some $n_0(\omega) \in \mathbb{N}$,

$$\left\|X\left(t_j^n, \omega\right) - X_j^n(\omega)\right\|_H \le k_n^{(1-\epsilon)/2}.$$

Since $l_n < +\infty$, $\bigcup_{0\le j\le l_n} N_j$ is a $\mathbb{P}$-null set and therefore, for all $\omega \in \bigcap_{0\le j\le l_n} N_j^c$ asymptotically

$$\sup_{0\le j\le l_n} \left\|X\left(t_j^n, \omega\right) - X_j^n(\omega)\right\|_H \le k_n^{(1-\epsilon)/2},$$

which proves the theorem. □

We continue with the approximation of M, since the presented scheme might still not be suitable for simulations although it is discretized in space and time.

4 Noise Approximation

In Sect. 3 we did not approximate the noise. Here, we derive an approximation of the driving noise term, in this case of a Lévy process, which preserves the overall order of convergence presented in Sect. 3. We assume that we are able to simulate the real-valued processes exactly, meaning that we do not take the error into account which stems from the approximation of small jumps by e.g., a Brownian motion. For the simulation and approximation of one dimensional Lévy processes we refer the reader to [2] and [15].

Let us assume in this section that $M \in \mathcal{M}_b^2(U)$ is a Lévy process L and therefore has a stationary covariance $Q \in L_1^+(U)$. Then, there exists an eigenbasis $(e_i \in U, i \in \mathbb{N})$ and a set of eigenvalues $(\gamma_i, i \in \mathbb{N})$ such that $Qe_i = \gamma_i e_i$ and $\gamma_i \geq 0$ for all $i \in \mathbb{N}$ as seen in Sect. 2. This implies that L admits the Karhunen–Loève expansion

$$L(t) = \sum_{i=1}^{\infty} \big(L(t), e_i\big)_U e_i = \sum_{i=1}^{\infty} \sqrt{\gamma_i} L_i(t) e_i, \tag{4.1}$$

where $(L_i, i \in \mathbb{N})$ is a family of real-valued, orthogonal Lévy processes (see Sect. 4.8 in [42]). Let us denote by L^κ the truncated process, i.e., for all $t \in \tau$ the stochastic process, which is given by

$$L^\kappa(t) = \sum_{i=1}^{\kappa} \sqrt{\gamma_i} L_i(t) e_i,$$

and its covariance is denoted by Q^κ. For $\kappa \to +\infty$, this process converges almost surely to L (see Sect. 4.8 in [42]). We set

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

with covariance $Q^{c\kappa} = Q - Q^\kappa$, which converges almost surely to zero. Let $\mathbb{L}^2_{\mathcal{H},\tau}(H) = L^2(\Omega \times \tau; L_{HS}(\mathcal{H}, H))$ be the space of integrands, defined over the probability 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, the Itô integral over $\Psi \in \mathbb{L}^2_{\mathcal{H},T}(H)$ satisfies that

$$\int_a^b \Psi(s)\, dL(s) - \int_a^b \Psi(s)\, dL^\kappa(s) = \int_a^b \Psi(s)\, dL^{c\kappa}(s). \tag{4.2}$$

In [7], it is shown, how the stochastic integral with respect to an L^κ consisting of independent Lévy processes $(L_i, i \in \mathbb{N})$ converges to a Hilbert-space-valued Lévy process L for $\kappa \to +\infty$. This result generalizes to the following lemma.

Lemma 4.1 *If for* $0 \leq a < b \leq T$ *and* $\Psi \in \mathbb{L}^2_{\mathcal{H},T}(H)$

$$\mathbb{E}\bigg(\int_a^b \|\Psi(s)\|^2_{L(U,H)}\, ds\bigg) < +\infty$$

and there exist constants $C_1, C_2 > 0$ *such that the eigenvalues of the covariance* Q *of* L *satisfy* $\gamma_i \le C_1 i^{-\delta}$ *for* $\delta > 1$, $i \in \mathbb{N}$ *and* $\kappa \ge C_2 h^{-\beta}$ *for some* $\beta > 0$, *then there exists a constant* $C(\delta)$ *such that*

$$\left\| \int_a^b \Psi(s)\, dL(s) - \int_a^b \Psi(s)\, dL^\kappa(s) \right\|_{H,L^2} \le C(\delta) \left(\mathbb{E}\left(\int_a^b \|\Psi(s)\|^2_{L(U,H)}\, ds \right) \right)^{1/2} h^{\frac{\beta(\delta-1)}{2}}.$$

Proof We first observe that

$$\left\| \int_a^b \Psi(s)\, dL(s) - \int_a^b \Psi(s)\, dL^\kappa(s) \right\|^2_{H,L^2} = \left\| \int_a^b \Psi(s)\, dL^{c\kappa}(s) \right\|^2_{H,L^2} = \mathbb{E}\left(\int_a^b \|\Psi(s)\|^2_{L_{HS}((Q^{c\kappa})^{1/2}U,H)}\, ds \right)$$

by Eqs. (4.2) and (2.2), where in this case equality holds (see Corollary 8.17 in [42]). Next, we calculate the Hilbert–Schmidt norm. We have that

$$\mathbb{E}\left(\int_a^b \|\Psi(s)\|^2_{L_{HS}((Q^{c\kappa})^{1/2}U,H)}\, ds \right) = \mathbb{E}\left(\int_a^b \sum_{i=\kappa+1}^\infty \gamma_i \|\Psi(s)e_i\|^2_H\, ds \right).$$

With the properties of Ψ it holds that

$$\mathbb{E}\left(\int_a^b \sum_{i=\kappa+1}^\infty \gamma_i \|\Psi(s)e_i\|^2_H\, ds \right) \le C\mathbb{E}\left(\int_a^b \|\Psi(s)\|^2_{L(U,H)}\, ds \right) \sum_{i=\kappa+1}^\infty \gamma_i,$$

and the decay of the eigenvalues and the assumptions on κ imply that

$$\sum_{i=\kappa+1}^\infty \gamma_i \le C_1 \sum_{i=\kappa+1}^\infty i^{-\delta} = C_1 \sum_{i=1}^\infty (i+\kappa)^{-\delta} \le C_1 \int_0^\infty (x+\kappa)^{-\delta} dx = C_1 C_2^{1-\delta} (\delta-1)^{-1} h^{\beta(\delta-1)}.$$

This proves the lemma. □

We remark that the estimates stay true, if Ψ also depends on the upper integration limit, i.e., the stochastic integral is a convolutional integral with respect to a semigroup, since Eq. (2.2) also holds for this type of integrands (see [21]).

Next, we approximate the double stochastic integral, which is the additional term when extending an Euler–Maruyama scheme to a Milstein scheme. In Eq. (3.4) this term is

$$\int_{t_{j-1}^n}^{t_j^n} \left(R_{j-1}^{n,h} P_h G\left(\int_{t_{j-1}^n}^{s} G\big(X_{j-1}^n\big)\, dM(r) \right) \right) dM(s). \tag{4.3}$$

We discuss this term for Lévy processes but in the more general setting, where we introduce separable Hilbert spaces H and U. The Hilbert space H can then be set to $L^2(D)$ or some approximation space V_h to apply the theory to Theorem 3.2. 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. (4.3) more generally as

$$\int_a^b \sigma(a)\Gamma\left(\int_a^s \Gamma\big(\psi(a)\big)\,dL(r)\right)dL(s). \tag{4.4}$$

Using the Karhunen–Loève expansion of L given in Eq. (4.1), we have that

$$\begin{aligned}&\int_a^b \sigma(a)\Gamma\left(\int_a^s \Gamma\big(\psi(a)\big)\,dL(r)\right)dL(s)\\&\quad= \sum_{i,j=1}^{\infty} \sqrt{\gamma_i}\sqrt{\gamma_j}\sigma(a)\Gamma\big(\Gamma\big(\psi(a)\big)e_i\big)e_j \int_a^b\int_a^s dL_i(r)\,dL_j(s).\end{aligned}$$

With Itô's formula the iterated Itô integral is given, for $i = j$, by

$$\int_a^b\int_a^s dL_i(r)\,dL_i(s) = \frac{1}{2}\left(\big(L_i(b)-L_i(a)\big)^2-(b-a)-\sum_{a<s\le b}\big(L_i(s)-L_i(s-)\big)^2\right).$$

By the same argument, the mixed terms, i.e., for $i \neq j$, satisfy

$$\begin{aligned}&\int_a^b\int_a^s dL_i(r)\,dL_j(s)+\int_a^b\int_a^s dL_j(r)\,dL_i(s)\\&\quad= \big(L_i(b)-L_i(a)\big)\cdot\big(L_j(b)-L_j(a)\big)\\&\qquad-\sum_{a<s\le b}\big(L_i(s)-L_i(s-)\big)\cdot\big(L_j(s)-L_j(s-)\big)\\&\quad= \Delta L_{ij}.\end{aligned}$$

If

$$\Gamma\big(\Gamma(\psi)e_i\big)e_j = \Gamma\big(\Gamma(\psi)e_j\big)e_i, \tag{4.5}$$

for $i,j \in \mathbb{N}$—this is for example satisfied for multiplicative noise (see Eq. (27) in [25]) and especially in the case where Γ is a pointwise multiplication with a suitable function g, e.g. $g \in C_b^\alpha(D)$, $U = L^2(D)$, where the eigenbasis and eigenvalues of Q satisfy that $\sum_{i\in\mathbb{N}}\gamma_i\|e_i\|^2_{W^{2,\infty}(D)} < +\infty$—we rewrite Eq. (4.4) as

$$\begin{aligned}&\int_a^b \sigma(a)\Gamma\left(\int_a^s \Gamma(\psi)\,dL(r)\right)dL(s)\\&\quad= \frac{1}{2}\sum_{i,j=1}^{\infty}\sqrt{\gamma_i}\sqrt{\gamma_j}\sigma(a)\Gamma\big(\Gamma(\psi)e_i\big)e_j\big(\Delta L_{ij}-\delta_{ij}(b-a)\big).\end{aligned}$$

This implies that the stochastic integrals can be simulated, if we assume that we are able to simulate the real-valued processes exactly as discussed in the beginning of the section. Approximation of the small jumps would lead to an additional error contribution. Still, the number of processes to be simulated might be infinite. To approximate the series by a finite number of stochastic processes, we truncate the Karhunen–Loève expansion as in Lemma 4.1 and simulate

$$\int_a^b \sigma(a)\Gamma\left(\int_a^s \Gamma(\psi)\, dL^\kappa(r)\right) dL^\kappa(s) \\ = \frac{1}{2}\sum_{i,j=1}^{\kappa} \sqrt{\gamma_i}\sqrt{\gamma_j}\sigma(a)\Gamma\big(\Gamma(\psi)e_i\big)e_j\big(\Delta L_{ij} - \delta_{ij}(b-a)\big).$$

The resulting error is given in the following lemma.

Lemma 4.2 *For $n \in \mathbb{N}$, let $\sigma : \mathbb{T}_n \to L(H,H)$, $\Gamma : H \to L_{HS}(\mathcal{H}, H)$ be linear and satisfy Assumption 2.1(d) for $\beta = 0$ and the norm in $L(U,H)$ as well as Eq. (4.5). Further, let $\psi = (\psi(t), t \in \mathbb{T}_n)$ be an adapted, H-valued stochastic process. For $t \in \mathbb{T}_n$, if*

$$\mathbb{E}\left(\int_0^t \big\|\psi\big(\pi_n(s)\big)\big\|_H^2\, ds\right) < +\infty$$

and there exist constants $C_1, C_2 > 0$ such that the eigenvalues of the covariance Q of L satisfy $\gamma_i \le C_1 i^{-\delta}$ for $\delta > 1$, $i \in \mathbb{N}$ and $\kappa \ge C_2 h^{-\beta}$ for some $\beta > 0$, then there exists a constant C such that

$$\bigg\|\int_0^t \sigma\big(\pi_n(s)\big)\Gamma\left(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL(r)\right) dL(s) \\ - \int_0^t \sigma\big(\pi_n(s)\big)\Gamma\left(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL^\kappa(r)\right) dL^\kappa(s)\bigg\|_{H,L^2} \\ \le C \sup_{t\in\mathbb{T}_n}\big\|\sigma(t)\big\|_{L(H,H)}\left(\int_0^t \big\|\psi\big(\pi_n(s)\big)\big\|_{H,L^2}^2\, ds\right)^{1/2}\big(k_n h^{\beta(\delta-1)}\big)^{1/2}.$$

Proof We calculate the error using Eq. (2.2), the properties of σ and Γ, and in the last step the estimate in the proof of Lemma 4.1

$$\bigg\|\int_0^t \sigma\big(\pi_n(s)\big)\,\Gamma\left(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL(r)\right) dL(s) \\ - \int_0^t \sigma\big(\pi_n(s)\big)\Gamma\left(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL^\kappa(r)\right) dL^\kappa(s)\bigg\|_{H,L^2}^2 \\ \le 2\left(\bigg\|\int_0^t \sigma\big(\pi_n(s)\big)\Gamma\left(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL^{c\kappa}(r)\right) dL(s)\bigg\|_{H,L^2}^2\right.$$

$$
\begin{aligned}
&+ \left\| \int_0^t \sigma\big(\pi_n(s)\big)\Gamma\left(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\,dL^{\kappa}(r)\right) dL^{c\kappa}(s) \right\|_{H,L^2}^2 \Bigg) \\
&\le C \sup_{t\in\mathbb{T}_n} \left\|\sigma(t)\right\|_{L(H,H)}^2 \left(\int_0^t \left\| \int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\,dL^{c\kappa}(r) \right\|_{H,L^2}^2 ds \right. \\
&\quad \left. + \sum_{i=\kappa+1}^{\infty} \gamma_i \int_0^t \left\| \int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\,dL^{\kappa}(r) \right\|_{H,L^2}^2 ds \right) \\
&\le C \sup_{t\in\mathbb{T}_n} \left\|\sigma(t)\right\|_{L(H,H)}^2 \int_0^t \left\|\psi\big(\pi_n(s)\big)\right\|_{H,L^2}^2 ds k_n \sum_{i=\kappa+1}^{\infty} \gamma_i \\
&\le C \sup_{t\in\mathbb{T}_n} \left\|\sigma(t)\right\|_{L(H,H)}^2 \int_0^t \left\|\psi\big(\pi_n(s)\big)\right\|_{H,L^2}^2 ds k_n h^{\beta(\delta-1)}.
\end{aligned}
$$

□

Let $\delta > 1$ be given. For a convergence of h^2 of the Euler–Maruyama term in Lemma 4.1, we have to choose $\beta \geq 4/(\delta-1)$ and therefore $\kappa_1 = \kappa \geq C\,h^{-4/(\delta-1)}$ for some constant C. At the same time, a convergence of h^2 of the Milstein term is achieved if $\beta \geq 2/(\delta-1)$, which implies that $\kappa_2 = \kappa \geq C\,h^{-2/(\delta-1)}$ for some constant C. Therefore, the overall convergence of h^2 in Theorem 3.2 is preserved if the noise in the Euler–Maruyama and in the Milstein term are truncated according to κ_1 and κ_2. So, to balance the errors, we have to use the first κ_1 terms of the Karhunen–Loève expansion for the Euler–Maruyama term and $\sqrt{\kappa_1}$ terms for the Milstein term. With this observation we conclude that the simulation of the Milstein term is computationally not more expensive than the Euler–Maruyama 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 $\mathrm{O}(\kappa_1)$. Therefore, 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. We remark that the efficient simulation of the Milstein term in [25] is possible in $\mathrm{O}(N\log N)$, where N is the number of grid points of the underlying domain, since the special structure of the domain and the chosen discretization grid allow the use of fast Fourier methods. In the general case, when Finite Element methods on arbitrary bounded polyhedrons are used, this approach does not apply.

Overall, the fully approximation scheme reads

$$
\begin{aligned}
\tilde{X}_j^n &= R_{(j-1:0)}^{n,h} P_h X_0 + \int_0^{t_j^n} R_{(j-1:\kappa_n(s))}^{n,h} P_h B \tilde{X}_{\kappa_n(s)}^n \,ds \\
&\quad + \int_0^{t_j^n} R_{(j-1:\kappa_n(s))}^{n,h} P_h G\big(\tilde{X}_{\kappa_n(s)}^n\big)\,dL^{\kappa}(s) \\
&\quad + \int_0^{t_j^n} \left(R_{(j-1:\kappa_n(s))}^{n,h} P_h G\left(\int_{\pi_n(s)}^s G\big(\tilde{X}_{\kappa_n(s)}^n\big)\,dL^{\sqrt{\kappa}}(r)\right)\right) dL^{\sqrt{\kappa}}(s),
\end{aligned}
$$

where we included beside a space and time discretization, the (according to Lemma 4.2) truncated driving noise process. Then, Lemma 4.1 and 4.2 together with Theorem 3.2 imply the following corollary, where we set $\sigma(t) = R^{n,h}_{(j-1:\kappa_n(t))} P_h$ and $\Gamma = G$.

Corollary 4.3 *Assume that $\kappa \geq C\lceil h^{-2\max(\alpha,2)/(\delta-1)}\rceil$ for some constant C, where $\delta > 1$ with $\gamma_i \leq \tilde{C}\, i^{-\delta}$ for $i \in \mathbb{N}$ and a fixed constant $\tilde{C}$. Then, the fully discrete approximation $\tilde{X}^n$ converges in mean square to the mild solution X of the stochastic partial differential equation (2.1) and satisfies for constants C_1 and C_2 that depend on T and δ that*

$$\sup_{0\leq j\leq l_n} \big\| X\big(t^n_j\big) - \tilde{X}^n_j \big\|_{H,L^2} \leq C_1\big(h^\alpha + k_n^{\alpha/2}\big)\|X\|_{H^\alpha,L^2,\infty_\tau} + C_2 k_n \|X\|_{H^1,L^2,\infty_\tau}.$$

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

$$\sup_{0\leq j\leq l_n} \big\| X\big(t^n_j\big) - \tilde{X}^n_j \big\|_{H,L^2} = \mathrm{O}\big(h^2 + k_n\big).$$

5 Examples and Simulation

A compensated Poisson process showcases that a noncontinuous, square integrable martingale exhibits the order of convergence in time described in Sect. 3. In general, for a noncontinuous martingale convergence of order k_n represents a sharp bound in L^2. The example is followed by simulations on the order of convergence for Euler–Maruyama and Milstein type terms.

Example 5.1 Let $L = (L(t), t \geq 0)$ be a Poisson process with intensity $\lambda > 0$, i.e., L has distribution

$$\mathbb{P}_{L(t)} = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \varepsilon_n.$$

Then, the corresponding compensated Poisson process

$$M(t) = L(t) - \lambda t$$

is a square integrable martingale. The quadratic variation $[M]$ of $M = (M(t), t \geq 0)$ is again the Poisson process L, i.e., we have $[M]_t = L(t)$ for all $t \geq 0$. The Burkholder–Davis–Gundy inequality for martingales implies

$$\begin{aligned}\mathbb{E}\bigg(\bigg\|\int_0^t dM(s)\bigg\|^p\bigg) &\leq C\mathbb{E}\bigg(\bigg(\int_0^t d[M]_s\bigg)^{p/2}\bigg) = C\mathbb{E}\bigg(\bigg(\int_0^t dL(s)\bigg)^{p/2}\bigg)\\ &= C\mathbb{E}\big(L(t)^{p/2}\big)\end{aligned}$$

and for $n \in \mathbb{N}$, the n-th moment of L is given by

$$\mathbb{E}\big(L(t)^n\big) = \sum_{k=0}^{n-1} \binom{n-1}{k} \lambda t \mathbb{E}\big(L(t)^k\big).$$

This implies for $t \downarrow 0$

$$\mathbb{E}\bigg(\bigg\| \int_0^t dM(s) \bigg\|^p\bigg) = \mathrm{O}(t)$$

and for $t \uparrow +\infty$

$$\mathbb{E}\bigg(\bigg\| \int_0^t dM(s) \bigg\|^p\bigg) = \mathrm{O}\big(t^{p/2}\big).$$

If we look at the (trivial) stochastic differential equation

$$dX(t) = dM(t)$$

with initial condition $X(0) = M(0)$, the regularity of the solution satisfies for $t > r$

$$\mathbb{E}\big(\|X(t) - X(r)\|^p\big) = \mathbb{E}\bigg(\bigg\| \int_r^t dM(s) \bigg\|^p\bigg) \le C \sum_{k=0}^{p/2-1} \binom{p/2-1}{k} \lambda(t-r) \mathbb{E}\big(L(t)^k\big).$$

So convergence in L^p for $r \to t$ is of order $(t-r)^{1/p}$ and this cannot be improved with the chosen methods. By Hölder's inequality, L^p convergence implies L^q convergence for $p \ge q$, since for $\phi \in L^p$

$$\|\phi\|_{H,L^q} \le \|\phi\|_{H,L^p}.$$

Therefore, using the estimate in L^p instead of a direct calculation in L^q for $p > q$, we obtain that

$$\big\|X(t) - X(r)\big\|_{H,L^q} \le (t-r)^{1/p},$$

i.e., the order of convergence reduces by a factor of q/p.

In the simulation we compare the convergence of the Euler–Maruyama and the Milstein term with domain $D = (0, 1)$ on the time interval $[0, 1]$, i.e., we approximate the integrals

$$\int_0^1 \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL(s) \quad \text{and} \quad \int_0^1 \Gamma\bigg(\int_{\pi_n(s)}^s \Gamma\big(\psi\big(\pi_n(s)\big)\big)\, dL(r)\bigg)\, dL(s)$$

and compare for κ chosen according to Lemma 4.1 and Lemma 4.2 the convergence rates with those proven in Sect. 4. Next, we introduce the parameters that were chosen for the simulation. The constructed Lévy process $L = (L(t), t \in \tau)$ is given by

$$L(t) = \sum_{i=1}^{\infty} \sqrt{\gamma_i} e_i L_i(t)$$

and

$$L_i(t) = W_i(t) + P_i(t),$$

where $(W_i(t), i \in \mathbb{N})$ is a family of independent, real-valued Wiener processes and $(P_i(t), i \in \mathbb{N})$ is a family of independent, real-valued compound Poisson processes, where the jump intensity is $\lambda = 1$ and the jump sizes are symmetric Gamma distributed with parameters 2 and 5, i.e., the jump size $J = Y \cdot Z$ is given by $Y \sim \Gamma(2, 5)$ and $Z \sim \mathcal{U}\{-1, 1\}$. We set the kernel of the covariance operator $\mathrm{Cov}(x, y) = \exp(-10|x - y|)$, for $x, y \in D$. This implies that the covariance operator Q has eigenvalues $(\gamma_k, k \in \mathbb{Z})$ given by

$$\gamma_k = \frac{20}{100 + 4\pi^2 k^2},$$

for $k \in \mathbb{Z}$, and corresponding eigenfunctions $(e_k, k \in \mathbb{Z})$ given by

$$e_k(x) = c_k^{-1}\left(\cos(2\pi kx) - \frac{10}{2\pi k}\sin(2\pi kx)\right),$$

for $k \in \mathbb{Z}$ and $x \in D$, where

$$c_k^2 = \frac{100}{8\pi^2 k^2} + \frac{1}{2}.$$

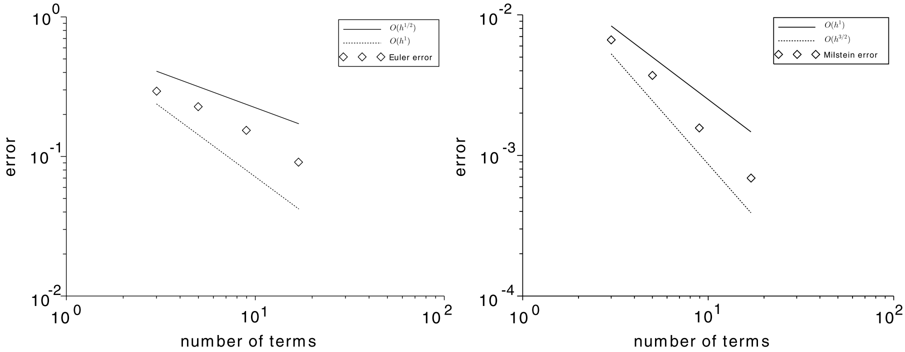
Therefore, the chosen covariance has the property that $\delta = 2$ in Lemma 4.1 and 4.2. Furthermore, we set

$$\big(\Gamma(\psi)\phi\big)(x) = g(x) \cdot \psi(x) \cdot \phi(x),$$

for $x \in D$, where $g(x) = x$ and $\psi \equiv 1$. Here, we can choose a constant ψ to simplify the simulation since otherwise, in both terms the same approximate solution of the stochastic partial differential equation would be plugged in. We choose $\beta = 1$ in Lemma 4.1 and in Lemma 4.2 and equal step sizes in space and time. The assumption that $h^2 = \mathrm{O}(k_n)$ is in this simulation superfluous and would just increase the computational costs in both simulations, since the convergence results in Lemma 4.1 and 4.2 do not depend on the relation of h and k_n. Then, Lemma 4.1 implies for the Euler–Maruyama term that it converges with order $\mathrm{O}(h^{1/2})$ and for the corresponding Milstein term, Lemma 4.2 leads to convergence of order $\mathrm{O}(h)$. The experimental results are shown in Fig. 1 and confirm the theory. For each plot, we simulated $N = 1000$ paths. As exact solution we chose the finest grid with 2^5 grid points in space and in time. We calculated the error

$$e_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \max_{j=0,\dots,2^n} \frac{1}{2^m}\sum_{k=1}^{2^m}\big(\hat{Y}_i(t_j, x_k) - Y_i(t_j, x_k)\big)^2},$$

for $n, m = 1, \dots, 4$. Here, $(Y_i, i = 1, \dots, N)$ is the set of simulated paths on a time grid $(t_j, j = 0, \dots, 2^n)$ and a space grid $(x_k, k = 0, \dots, 2^m)$, and $(\hat{Y}_i, i = 1, \dots, N)$ the family of simulated paths of the "exact" solution on the fine grid in time and space.



(a) Euler approximation (b) Milstein approximation

Fig. 1 Statistical error with 1000 sample paths

References

1. Andrea Barth, A., Benth, F.E.: The forward dynamics in energy markets—infinite dimensional modeling and simulation. In: Stochastic Partial Differential Equations: Approximations and Applications, vol. 916, pp. 77–96. University of Oslo, Oslo (2009). Chap. 5. Series of dissertations submitted to the Faculty of Mathematics and Natural Sciences. ISSN 1501-7710
2. Asmussen, S.R., Glynn, P.W.: Stochastic Simulation: Algorithms and Analysis. Stochastic Modelling and Applied Probability, vol. 57. Springer, New York (2007)
3. Auscher, P., Tchamitchian, P.: Square roots of elliptic second order divergence operators on strongly Lipschitz domains: L^2 theory. J. Anal. Math. **90**, 1–12 (2003)
4. Barth, A.: Stochastic Partial Differential Equations: Approximations and Applications. PhD thesis, University of Oslo, CMA (September 2009)
5. Barth, A.: A finite element method for martingale-driven stochastic partial differential equations. Commun. Stoch. Anal. **4**, 355–375 (2010)
6. Barth, A., Lang, A.: L^p and almost sure convergence of a Milstein scheme for stochastic partial differential equations. SAM report 2011-15, June (2009)
7. Barth, A., Lang, A.: Simulation of stochastic partial differential equations using finite element methods. Stochastics **84**, 217–231 (2012)
8. Bessaih, H., Schurz, H.: Upper bounds on the rate of convergence of truncated stochastic infinite-dimensional differential systems with h-regular noise. J. Comput. Appl. Math. **208**, 354–361 (2007)
9. Braess, D.: Finite Elemente, 3rd edn. Springer, Berlin (2002)
10. Caraballo, T., Kloeden, P.E.: The pathwise numerical approximation of stationary solutions of semilinear stochastic evolution equations. Appl. Math. Optim. **54**, 401–415 (2006)
11. Chow, P.-L.: Stochastic Partial Differential Equations. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton (2007)
12. Chow, P.-L., Jiang, J.-L.: Almost sure convergence of some approximate solutions for random parabolic equations. In: Hornung, U., Kotelenez, P., Papanicolaou, G. (eds.) Random Partial Differential Equations, Proc. Conf., Oberwolfach/Ger., 1989. ISNM, vol. 102, pp. 45–54 (1991)
13. Chow, P.-L., Jiang, J.-L., Menaldi, J.-L.: Pathwise convergence of approximate solutions to Zakai's equation in a bounded domain. In: Da Prato, G., Tubaro, L. (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. Pitman Res. Notes Math., vol. 268, pp. 111–123. Longman Scientific & Technical, Harlow (1992)
14. Clark, D.S.: Short proof of a discrete Gronwall inequality. Discrete Appl. Math. **16**, 279–281 (1987)
15. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall/CRC Financial Mathematics Series. Chapman and Hall/CRC, Boca Raton (2004)
16. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and Its Applications, vol. 44. Cambridge University Press, Cambridge (1992)

17. Ekeland, I., Taflin, E.: A theory of bond portfolios. Ann. Appl. Probab. **15**, 1260–1305 (2005)
18. Engel, K.-J., Nagel, R.: One-Parameter Semigroups for Linear Evolution Equations. Graduate Texts in Mathematics, vol. 194. Springer, Berlin (2000)
19. Fujita, H., Saito, N., Suzuki, T.: Operator Theory and Numerical Methods. Studies in Mathematics and Its Applications, vol. 30. North-Holland/Elsevier, Amsterdam (2001)
20. Grossmann, C., Roos, H.-G., Stynes, M.: Numerical Treatment of Partial Differential Equations. Revised translation of the 3rd German edition of 'Numerische Behandlung partieller Differentialgleichungen' by Martin Stynes. Universitext. Springer, Berlin (2007)
21. Hausenblas, E., Seidler, J.: Stochastic convolutions driven by martingales: Maximal inequalities and exponential integrability. Stoch. Anal. Appl. **26**, 98–119 (2008)
22. Jentzen, A.: Taylor Expansions for Stochastic Partial Differential Equations. PhD thesis, Johann Wolfgang Goethe Universität Frankfurt am Main (2009)
23. Jentzen, A., Kloeden, P.E.: Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space–time noise. Proc. R. Soc. A **465**, 649–667 (2009)
24. Jentzen, A., Kloeden, P.E.: Taylor expansions of solutions of stochastic partial differential equations with additive noise. Ann. Probab. **38**, 532–569 (2010)
25. Jentzen, A., Röckner, M.: A Milstein scheme for SPDEs. arXiv:1001.2751 [math.NA] (January 2012)
26. Juan, O., Keriven, R., Postelnicu, G.: Stochastic motion and the level set method in computer vision: Stochastic active contours. Int. J. Comput. Vis. **69**, 7–25 (2006)
27. Kantorovitz, S.: On Liu's analyticity criterion for semigroups. Semigroup Forum **53**, 262–265 (1996)
28. Kato, T.: Perturbation Theory for Linear Operators. Corr. printing of the 2nd edn., Grundlehren der mathematischen Wissenschaften, vol. 132. Springer, Berlin (1980)
29. Kloeden, P.E., Shott, S.: Linear-implicit strong schemes for Itô–Galerkin approximations of stochastic PDEs. J. Appl. Math. Stoch. Anal. **14**, 47–53 (2001)
30. Kovács, M., Larsson, S., Lindgren, F.: Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise. Numer. Algorithms **53**, 309–320 (2010)
31. Kruse, R.: Optimal error estimates of Galerkin Finite Element methods for stochastic partial differential equations with multiplicative noise. arXiv:1103.4504 [math.NA] (March 2011)
32. Lang, A.: Simulation of Stochastic Partial Differential Equations and Stochastic Active Contours. PhD thesis, Universität Mannheim (2007)
33. Lang, A.: Mean square convergence of a semidiscrete scheme for SPDEs of Zakai type driven by square integrable martingales. Proc. Comput. Sci. **1**, 1609–1617 (2010). ICCS 2010
34. Lang, A.: Almost sure convergence of a Galerkin approximation for SPDEs of Zakai type driven by square integrable martingales. J. Comput. Appl. Math. **236**, 1724–1732 (2012)
35. Lang, A., Chow, P.-L., Potthoff, J.: Almost sure convergence of a semidiscrete Milstein scheme for SPDEs of Zakai type. Stochastics **82**, 315–326 (2010)
36. Lang, A., Chow, P.-L., Potthoff, J.: Erratum: Almost sure convergence of a semi-discrete Milstein scheme for SPDEs of Zakai type. Stochastics (2011)
37. Lang, A., Potthoff, J.: Fast simulation of Gaussian random fields. Monte Carlo Methods Appl. **17**, 195–214 (2011)
38. Lou, Y., Christofides, P.: Nonlinear feedback control of surface roughness using a stochastic PDE. In: 45th IEEE Conference on Decision and Control, Dec. 2006, pp. 936–943 (2006)
39. Mishura, Y., Shevchenko, G.: Approximation schemes for stochastic differential equations in Hilbert space. Theory Probab. Appl. **51**, 442–458 (2007)
40. Ni, D., Christofide, P.D.: Construction of stochastic PDEs for feedback control of surface roughness in thin film deposition. In: Proceedings of the 2005 American Control Conference, vol. 4, June 2005, pp. 2540–2547 (2005)
41. Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences, vol. 44. Springer, New York (1983)
42. Peszat, S., Zabczyk, J.: Stochastic Partial Differential Equations with Lévy Noise. Encyclopedia of Mathematics and Its Applications, vol. 113. Cambridge University Press, Cambridge (2007)
43. Prévôt, C., Röckner, M.: A Concise Course on Stochastic Partial Differential Equations. Lecture Notes in Mathematics, vol. 1905. Springer, Berlin (2007)
44. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations, 2nd corrected printing. Springer Series in Computational Mathematics, vol. 23. Springer, Berlin (1997)

45. Shardlow, T.: Numerical simulation of stochastic PDEs for excitable media. J. Comput. Appl. Math. **175**, 429–446 (2005)
46. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, 2nd revised and expanded edn. Springer Series in Computational Mathematics, vol. 25. Springer, Berlin (2006)
47. Zakai, M.: On the optimal filtering of diffusion processes. Z. Wahrscheinlichkeitstheor. Verw. Geb. **11**, 230–243 (1969)