A Milstein scheme for SPDEs

Arnulf Jentzen[1] and Michael Röckner[2]

[1]Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544-1000, USA, e-mail: ajentzen@math.princeton.edu

[2]Faculty of Mathematics, Bielefeld University, 33501 Bielefeld, Germany, e-mail: roeckner@math.uni-bielefeld.de

January 27, 2012

Abstract

This article studies an infinite dimensional analog of Milstein's scheme for finite dimensional stochastic ordinary differential equations (SODEs). The Milstein scheme is known to be impressively efficient for SODEs which fulfill a certain commutativity type condition. This article introduces the infinite dimensional analog of this commutativity type condition and observes that a certain class of semilinear stochastic partial differential equation (SPDEs) with multiplicative trace class noise naturally fulfills the resulting infinite dimensional commutativity condition. In particular, a suitable infinite dimensional analog of Milstein's algorithm can be simulated efficiently for such SPDEs and requires less computational operations and random variables than previously considered algorithms for simulating such SPDEs. The analysis is supported by numerical results for a stochastic heat equation and stochastic reaction diffusion equations showing significant computational savings.

1 Introduction

In this article an infinite dimensional analog of Milstein's scheme for finite dimensional stochastic ordinary differential equations (SODEs) is studied. In order to get a better understanding of this Milstein type scheme in infinite dimensions, we first briefly review Milstein's method for finite dimensional SODEs and then concentrate on the case of infinite dimensional stochastic partial differential equations (SPDEs) in the rest of this introductory section.

Let $T \in (0, \infty)$ be a real number, let $d, m \in \mathbb{N} := \{1, 2, \ldots\}$ be natural numbers, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a normal filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and let $w = (w^1, \ldots, w^m)\colon [0, T] \times \Omega \to \mathbb{R}^m$ be an m-dimensional standard $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion. Moreover, let $x_0 \in \mathbb{R}^d$ and let $\mu = (\mu_1, \ldots, \mu_d)\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma = (\sigma_{i,j})_{i\in\{1,\ldots,d\},j\in\{1,\ldots,m\}}\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be two smooth functions satisfying suitable Lipschitz assumptions (see condition (3.21) in Theorem 10.3.5 in P. E. Kloeden and E. Platen [34] for details). The SODE

$$dX_t = \mu(X_t)\, dt + \sigma(X_t)\, dw_t, \qquad X_0 = x_0 \tag{1}$$

for all $t \in [0, T]$ then admits a unique solution. More precisely, there exists an up to indistinguishability unique adapted stochastic process $X\colon [0, T] \times \Omega \to \mathbb{R}^d$ with continuous sample paths which satisfies

$$\begin{aligned} X_t &= x_0 + \int_0^t \mu(X_s)\, ds + \int_0^t \sigma(X_s)\, dw_s \\ &= x_0 + \int_0^t \mu(X_s)\, ds + \sum_{i=1}^m \int_0^t \sigma_i(X_s)\, dw_s^i \end{aligned} \tag{2}$$

$\mathbb{P}$-a.s. for all $t \in [0, T]$. Here $\sigma_i\colon \mathbb{R}^d \to \mathbb{R}^d$ is given by $\sigma_i(x) = (\sigma_{1,i}(x), \ldots, \sigma_{d,i}(x))$ for all $x \in \mathbb{R}^d$ and all $i \in \{1, \ldots, m\}$. Milstein's method (see, e.g., (3.3) in Section 10.3 in P. E. Kloeden and E. Platen [34] and also G. N. Milstein's original article [45]) applied to the SODE (1) is then given by $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable mappings $y_n^N\colon \Omega \to \mathbb{R}^d$, $n \in \{0, 1, \ldots, N\}$, $N \in \mathbb{N}$, with $y_0^N = x_0$ and

$$\begin{aligned} y_{n+1}^N = y_n^N &+ \frac{T}{N} \cdot \mu(y_n^N) + \sum_{i=1}^m \sigma_i(y_n^N) \cdot \left(w^i_{\frac{(n+1)T}{N}} - w^i_{\frac{nT}{N}} \right) \\ &+ \sum_{i,j=1}^m \sum_{k=1}^d \left(\frac{\partial}{\partial x_k} \sigma_i \right)(y_n^N) \cdot \sigma_{k,j}(y_n^N) \cdot \int_{\frac{nT}{N}}^{\frac{(n+1)T}{N}} \int_{\frac{nT}{N}}^{s} dw_u^j\, dw_s^i \end{aligned} \tag{3}$$

$\mathbb{P}$-a.s. for all $n \in \{0, 1, \ldots, N-1\}$ and all $N \in \mathbb{N}$. Although Milstein's scheme is known to converge significantly faster than many other methods such as the Euler-Maruyama scheme, it is only of limited use due to difficult simulations of the iterated stochastic integrals $\int_{\frac{nT}{N}}^{\frac{(n+1)T}{N}} \int_{\frac{nT}{N}}^{s} dw_u^j \, dw_s^i$ for $i, j \in \{1, \ldots, m\}$ with $i \neq j$, $n \in \{0, 1, \ldots N-1\}$ and $N \in \mathbb{N}$ in (3). In the special situation of so called *commutative noise* (see (3.13) in Section 10.3 in [34]), i.e.,

$$\sum_{k=1}^{d} \left(\frac{\partial}{\partial x_k} \sigma_i \right)(x) \cdot \sigma_{k,j}(x) = \sum_{k=1}^{d} \left(\frac{\partial}{\partial x_k} \sigma_j \right)(x) \cdot \sigma_{k,i}(x) \tag{4}$$

for all $x \in \mathbb{R}^d$ and all $i, j \in \{1, \ldots, m\}$, the Milstein scheme can be simplified and complicated iterated stochastic integrals in (3) can be avoided. More precisely, in case (4), Milstein's scheme (3) reduces to

$$\begin{aligned} y_{n+1}^N = y_n^N &+ \frac{T}{N} \cdot \mu(y_n^N) + \sum_{i=1}^{m} \sigma_i(y_n^N) \cdot \left(w^i_{\frac{(n+1)T}{N}} - w^i_{\frac{nT}{N}} \right) \\ &+ \frac{1}{2} \sum_{i,j=1}^{m} \sum_{k=1}^{d} \left(\frac{\partial}{\partial x_k} \sigma_i \right)(y_n^N) \cdot \sigma_{k,j}(y_n^N) \cdot \left(w^i_{\frac{(n+1)T}{N}} - w^i_{\frac{nT}{N}} \right) \cdot \left(w^j_{\frac{(n+1)T}{N}} - w^j_{\frac{nT}{N}} \right) \\ &- \frac{T}{2N} \sum_{i=1}^{m} \sum_{k=1}^{d} \left(\frac{\partial}{\partial x_k} \sigma_i \right)(y_n^N) \cdot \sigma_{k,i}(y_n^N) \end{aligned} \tag{5}$$

$\mathbb{P}$-a.s. for all $n \in \{0, 1, \ldots, N-1\}$ and all $N \in \mathbb{N}$ (see (3.16) in Section 10.3 in [34]). For instance, in the case $d = m = 1$, condition (4) is obviously fulfilled and the Milstein scheme (5) can then be written as

$$\begin{aligned} y_{n+1}^N = y_n^N &+ \frac{T}{N} \cdot \mu(y_n^N) + \sigma(y_n^N) \cdot \left(w_{\frac{(n+1)T}{N}} - w_{\frac{nT}{N}} \right) \\ &+ \frac{1}{2} \cdot \sigma'(y_n^N) \cdot \sigma(y_n^N) \cdot \left(\left(w_{\frac{(n+1)T}{N}} - w_{\frac{nT}{N}} \right)^2 - \frac{T}{N} \right) \end{aligned} \tag{6}$$

$\mathbb{P}$-a.s. for all $n \in \{0, 1, \ldots, N-1\}$ and all $N \in \mathbb{N}$ (see (3.1) in Section 10.3 in [34]). Of course, (6) can be simulated very efficiently. However, (4) is in the case of a multidimensional SODE seldom fulfilled and even if it is fulfilled, Milstein's method (5) becomes less useful if $d, m \in \mathbb{N}$ are large. For example, if $d = m = 20$ holds, then the middle term in (5) contains $20^3 = 8000$ summands. So, more than 8000 additional arithmetic operations are needed to compute y_{n+1}^N from y_n^N for $n \in \{0, 1, \ldots, N-1\}$, $N \in \mathbb{N}$ via (5) in the case $d = m = 20$ in general, which makes Milstein's scheme less efficient. This suggests that there is no hope to expect that an infinite dimensional analog of Milstein's method can be simulated efficiently in the case of infinite dimensional state spaces such as $L^2((0,1), \mathbb{R})$ instead of $\mathbb{R}^d$ and $\mathbb{R}^m$ respectively. One purpose of this article is to demonstrate that this is not true in the case of a suitable class of semilinear SPDEs with multiplicative trace class noise. We now illustrate this in more detail.

Let $H = L^2((0,1), \mathbb{R})$ be the $\mathbb{R}$-Hilbert space of equivalence classes of Lebesgue square integrable functions from $(0,1)$ to $\mathbb{R}$ and let $f, b \colon (0,1) \times \mathbb{R} \to \mathbb{R}$ be two smooth functions satisfying appropriate Lipschitz assumptions (see (45) and (46) for details). As usual we do not distinguish between a Lebesgue square integrable function from $(0,1)$ to $\mathbb{R}$ and its equivalence class in H. Moreover, let $\kappa \in (0, \infty)$ be a real number, let $\xi \colon [0,1] \to \mathbb{R}$ with $\xi(0) = \xi(1) = 0$ be a smooth function and let $W \colon [0,T] \times \Omega \to H$ be a standard Q-Wiener process with respect to $\mathcal{F}_t$, $t \in [0,T]$, where $Q \colon H \to H$ is a trace class operator (see, for instance, Definition 2.1.12 in [52]). It is a classical result (see, e.g., Proposition 2.1.5 in [52]) that the covariance operator $Q \colon H \to H$ of the Wiener process $W \colon [0,T] \times \Omega \to H$ has an orthonormal basis $g_j \in H$, $j \in \mathbb{N}$, of eigenfunctions with summable eigenvalues $\eta_j \in [0, \infty)$, $j \in \mathbb{N}$. In order to have a more concrete example, we consider the choice $g_j(x) = \sqrt{2} \sin(j \pi x)$ and $\eta_j = \frac{1}{j^2}$ for all $x \in (0,1)$ and all $j \in \mathbb{N}$ in the following and refer to Section 2 for our general setting and to Section 4 for further possible examples. Then we consider the SPDE

$$dX_t(x) = \left[\kappa \frac{\partial^2}{\partial x^2} X_t(x) + f(x, X_t(x)) \right] dt + b(x, X_t(x)) \, dW_t(x) \tag{7}$$

with $X_t(0) = X_t(1) = 0$ and $X_0(x) = \xi(x)$ for $x \in (0,1)$ and $t \in [0,T]$ on H. Under the assumptions above the SPDE (7) has a unique mild solution. Specifically, there exists an up to indistinguishability unique adapted stochastic process $X \colon [0,T] \times \Omega \to H$ with continuous sample paths which satisfies

$$X_t = e^{At} \xi + \int_0^t e^{A(t-s)} F(X_s) \, ds + \int_0^t e^{A(t-s)} B(X_s) \, dW_s \tag{8}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$, where $A \colon D(A) \subset H \to H$ is the Laplacian with Dirichlet boundary conditions times the constant $\kappa \in (0, \infty)$ and where $F \colon H \to H$ and $B \colon H \to HS(U_0, H)$ are given by $(F(v))(x) = f(x, v(x))$

and $(B(v)u)(x) = b(x, v(x)) \cdot u(x)$ for all $x \in (0,1)$, $v \in H$ and all $u \in U_0$. Here $U_0 = Q^{1/2}(H)$ with $\langle v, w \rangle_{U_0} = \langle Q^{-1/2}v, Q^{-1/2}w \rangle_H$ for all $v, w \in U_0$ is the image $\mathbb{R}$-Hilbert space of $Q^{1/2}$ (see Appendix C in [52]). (Note that A und Q commutate and even more satisfy $-\kappa\pi^2 A^{-1} = Q$ in our example SPDE (8) in this introductory section but our general setting below does not require that these conditions are fulfilled; see Section 2 and Subsection 4.2.)

Then our goal is to solve the strong approximation problem (see Section 9.3 in [34]) of the SPDE (7). More precisely, we want to compute an $\mathcal{F}/\mathcal{B}(H)$-measurable numerical approximation $Y\colon \Omega \to H$ such that

$$\left(\mathbb{E}\left[\int_0^1 |X_T(x) - Y(x)|^2 \, dx\right]\right)^{1/2} < \varepsilon \tag{9}$$

for a given precision $\varepsilon > 0$ with the least possible computational effort (number of computational operations and independent standard normal random variables needed to compute $Y\colon \Omega \to H$). A computational operation is here an arithmetic operation (addition, subtraction, multiplication, division), a trigonometric operation (sine, cosine) or an evaluation of $f\colon (0,1) \times \mathbb{R} \to \mathbb{R}$, $b\colon (0,1) \times \mathbb{R} \to \mathbb{R}$ or the exponential function.

In order to be able to simulate such a numerical approximation on a computer both the time interval $[0,T]$ and the infinite dimensional space $H = L^2((0,1),\mathbb{R})$ have to be discretized. While for temporal discretizations the linear implicit Euler scheme (see [9, 10, 11, 18, 23, 24, 25, 58, 59, 62]) and the linear implicit Crank-Nicolson scheme (see [23, 24, 58, 59]) are often used, spatial discretizations are usually achieved with finite elements (see [1, 2, 10, 11, 14, 25, 33, 36, 37, 38, 43, 59, 62]), finite differences (see [17, 22, 44, 51, 54, 55, 56, 58, 60]) and spectral Galerkin methods (see [16, 24, 28, 30, 35, 41, 42, 47, 49, 50]). For instance, the linear implicit Euler scheme combined with spectral Galerkin methods which we denote by $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $Z_n^N\colon \Omega \to H$, $n \in \{0,1,\ldots,N^3\}$, $N \in \mathbb{N}$, is given by $Z_0^N = P_N(\xi)$ and

$$Z_{n+1}^N = P_N\left(I - \frac{T}{N^3}A\right)^{-1}\left(Z_n^N + \frac{T}{N^3} \cdot f(\cdot, Z_n^N) + b(\cdot, Z_n^N) \cdot \left(W^N_{\frac{(n+1)T}{N^3}} - W^N_{\frac{nT}{N^3}}\right)\right) \tag{10}$$

$\mathbb{P}$-a.s. for all $n \in \{0,1,\ldots,N^3-1\}$ and all $N \in \mathbb{N}$. Here the bounded linear operators $P_N\colon H \to H$, $N \in \mathbb{N}$, and the Wiener processes $W^N\colon [0,T] \times \Omega \to H$, $N \in \mathbb{N}$, are given by

$$(P_N(v))(x) = \sum_{n=1}^N 2\sin(n\pi x) \int_0^1 \sin(n\pi y)\, v(y)\, dy \tag{11}$$

for all $x \in (0,1)$, $v \in H$, $N \in \mathbb{N}$ and by $W_t^N(\omega) = P_N(W_t(\omega))$ for all $t \in [0,T]$, $\omega \in \Omega$, $N \in \mathbb{N}$. Moreover, the notations $v \cdot w\colon (0,1) \to \mathbb{R}$, $v^2\colon (0,1) \to \mathbb{R}$ and $\varphi(\cdot, v)\colon (0,1) \to \mathbb{R}$ given by

$$\big(v \cdot w\big)(x) = v(x) \cdot w(x), \qquad \big(v^2\big)(x) = \big(v(x)\big)^2, \qquad \big(\varphi(\cdot, v)\big)(x) = \varphi(x, v(x)) \tag{12}$$

for all $x \in (0,1)$ and all functions $v, w\colon (0,1) \to \mathbb{R}$, $\varphi\colon (0,1) \times \mathbb{R} \to \mathbb{R}$ are used here and below. In (10) the infinite dimensional $\mathbb{R}$-Hilbert space H is projected down to the N-dimensional $\mathbb{R}$-Hilbert space $P_N(H)$ with $N \in \mathbb{N}$ and the infinite dimensional Wiener process $W\colon [0,T] \times \Omega \to H$ is approximated by the finite dimensional processes $W^N\colon [0,T] \times \Omega \to H$, $N \in \mathbb{N}$, for the spatial discretization. For the temporal discretization in the scheme Z_n^N, $n \in \{0,1,\ldots,N^3\}$, above the time interval $[0,T]$ is divided into N^3 subintervals, i.e., N^3 time steps are used, for $N \in \mathbb{N}$. The exact solution $X\colon [0,T] \times \Omega \to H$ of the SPDE (7) has values in $D((-A)^\gamma)$ and satisfies $\mathbb{E}\big[\|(-A)^\gamma X_T\|_H^2\big] < \infty$ for all $\gamma \in (0, \frac{3}{4})$ (see Section 4.3 in [32]). This shows

$$\begin{aligned}\left(\mathbb{E}\Big[\|X_T - P_N(X_T)\|_H^2\Big]\right)^{1/2} &\leq \left(\mathbb{E}\Big[\|(-A)^\gamma X_T\|_H^2\Big]\right)^{1/2} \big\|(-A)^{-\gamma}(I - P_N)\big\|_{L(H)} \\ &\leq \left(\mathbb{E}\Big[\|(-A)^\gamma X_T\|_H^2\Big]\right)^{1/2} \left(1 + \kappa^{-1}\right) N^{-2\gamma} < \infty\end{aligned}$$

for all $N \in \mathbb{N}$ and all $\gamma \in (0, \frac{3}{4})$. So, $P_N(X_T)$ converges in the root mean square sense to X_T with order $\frac{3}{2}-$ as N goes to infinity. (For a real number $\delta \in (0,\infty)$, we write $\delta-$ for the convergence order if the convergence order is higher than $\delta - r$ for all arbitrarily small $r \in (0,\delta)$.) Additionally, the solution process of the SPDE (7) is known to be $\frac{1}{2}$-Hölder continuous in the root mean square sense (see, e.g., Theorem 1 in [32]) and therefore, the linear implicit Euler scheme converges temporally in the root mean square to the exact solution of the SPDE (7) with order $\frac{1}{2}$ (see, e.g., Theorem 1.1 in [62]). Combining the convergence rate $\frac{3}{2}-$ for the spatial discretization and the convergence rate $\frac{1}{2}$ for the temporal discretization indicates that it is asymptotically optimal to use the cubic number N^3 of time steps in the linear implicit Euler scheme Z_n^N, $n \in \{0,1,\ldots,N^3\}$, above.

We now review how efficiently the numerical method (10) solves the strong approximation problem (9) of the SPDE (7). Standard results in the literature (see, for instance, Theorem 2.1 in [24]) yield the existence of real numbers $C_r > 0$, $r \in (0, \frac{3}{2})$, such that

$$\left(\mathbb{E}\left[\int_0^1 \left|X_T(x) - Z_{N^3}^N(x)\right|^2 dx\right]\right)^{1/2} \leq C_r \cdot N^{(r - \frac{3}{2})} \tag{13}$$

for all $N \in \mathbb{N}$ and all arbitrarily small $r \in (0, \frac{3}{2})$. The linear implicit Euler approximation $Z^N_{N^3}$ thus converges in the root mean square sense to X_T with order $\frac{3}{2}-$ as N goes to infinity. Moreover, since $P_N(H)$ is N-dimensional and since N^3 time steps are used in (10), $O(N^4 \log(N))$ computational operations and random variables are needed to compute $Z^N_{N^3}$. The logarithmic term in $O(N^4 \log(N))$ arises due to computing the nonlinearities f and b with fast Fourier transform where aliasing errors are neglected here and below. Combining the computational effort $O(N^4 \log(N))$ and the convergence order $\frac{3}{2}-$ in (13) shows that the linear implicit Euler scheme needs about $O(\varepsilon^{-\frac{8}{3}})$ computational operations and independent standard normal random variables to achieve the desired precision $\varepsilon > 0$ in (9). In fact, we have demonstrated that Euler's method (10) needs $O(\varepsilon^{-(\frac{8}{3}+r)})$ computational operations and independent standard normal random variables to solve (9) for all arbitrarily small $r \in (0, \infty)$. However, we write about $O(\varepsilon^{-\frac{8}{3}})$ computational operations and independent standard normal random variables for simplicity here and below.

Having reviewed Euler's method (10), we now derive and study an infinite dimensional analog of Milstein's scheme. In the finite dimensional SODE case, the Milstein scheme (3) is derived by applying Itô's formula to the integrand process $\sigma(X_t)$, $t \in [0, T]$, in (2). This approach is based on the fact that the diffusion coefficient σ is a smooth test function and that the solution process of (1) is an Itô process. This strategy is not directly available in infinite dimensions since (7) does in general not admit a strong solution to which the standard Itô formula could be applied. Recently, in [31] in the case of additive noise and in [29] in the general case, this problem has been overcome by first applying Taylor's formula in Banach spaces to the diffusion coefficient B in the mild integral equation (8) and by then inserting a lower order approximation recursively (see Section 4.3 in [29]). More formally, using $F(X_s) \approx F(X_0)$ and $B(X_s) \approx B(X_0) + B'(X_0)(X_s - X_0)$ for small $s \in [0, T]$ in (8) shows

$$\begin{aligned} X_t &\approx e^{At}\xi + \int_0^t e^{A(t-s)} F(X_0)\, ds + \int_0^t e^{A(t-s)} B(X_0)\, dW_s + \int_0^t e^{A(t-s)} B'(X_0)(X_s - X_0)\, dW_s \\ &\approx e^{At} \left(X_0 + t \cdot F(X_0) + \int_0^t B(X_0)\, dW_s + \int_0^t B'(X_0)(X_s - X_0)\, dW_s \right) \end{aligned}$$

for small $t \in [0, T]$. (We would like to remark that B is, in general, not Fréchet differentiable on H but on a suitable dense subspace of H; see Assumption 4 below for details. In this introductory section we simply write B' for this Fréchet derivative on a suitable dense subspace of H and refer to Section 2 below for the precise handling of this issue.) The estimate $X_s \approx X_0 + \int_0^s B(X_0)\, dW_u$ for small $s \in [0, T]$ then gives

$$X_t \approx e^{At}\left(X_0 + t \cdot F(X_0) + \int_0^t B(X_0)\, dW_s + \int_0^t B'(X_0) \left(\int_0^s B(X_0)\, dW_u \right) dW_s \right) \tag{14}$$

for small $t \in [0, T]$. Using Itô's formula this temporal approximation has already been obtained in (1.12) in [46] under additional smoothness assumptions of the driving noise process of the SPDE (8) (see Assumption C in [46]) which guarantee the existence of a strong solution and thus allow the application of the standard Itô formula. Combining the temporal approximation (14) and the spatial discretization in (10) suggests the numerical scheme with $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $Y^N_n \colon \Omega \to H$, $n \in \{0, 1, \ldots, N^2\}$, $N \in \mathbb{N}$, given by $Y^N_0 = P_N(\xi)$ and

$$\begin{aligned} Y^N_{n+1} = P_N\, e^{A\frac{T}{N^2}} \Bigg(& Y^N_n + \frac{T}{N^2} \cdot F(Y^N_n) + B(Y^N_n) \left(W^N_{\frac{(n+1)T}{N^2}} - W^N_{\frac{nT}{N^2}} \right) \\ & + \int_{\frac{nT}{N^2}}^{\frac{(n+1)T}{N^2}} B'(Y^N_n) \left(\int_{\frac{nT}{N^2}}^{s} B(Y^N_n)\, dW^N_u \right) dW^N_s \Bigg) \end{aligned} \tag{15}$$

$\mathbb{P}$-a.s. for all $n \in \{0, 1, \ldots, N^2 - 1\}$ and all $N \in \mathbb{N}$. Now, we are at a stage similar to the finite dimensional case (3): a higher order Milstein type method seems to be derived which nevertheless seems to be of limited use due to the iterated high dimensional stochastic integral in (15). However, a key observation here is the formula

$$\begin{aligned} &\int_{\frac{nT}{N^2}}^{\frac{(n+1)T}{N^2}} B'(Y^N_n) \left(\int_{\frac{nT}{N^2}}^{s} B(Y^N_n)\, dW^N_u \right) dW^N_s \\ &\qquad = \frac{1}{2} \left(\frac{\partial}{\partial y} b \right)(\cdot, Y^N_n) \cdot b(\cdot, Y^N_n) \cdot \left(\left(W^N_{\frac{(n+1)T}{N^2}} - W^N_{\frac{nT}{N^2}} \right)^2 - \frac{T}{N^2} \sum_{i=1}^{N} \eta_i (g_i)^2 \right) \end{aligned} \tag{16}$$

$\mathbb{P}$-a.s. for all $n \in \{0, 1, \ldots, N^2 - 1\}$ and all $N \in \mathbb{N}$ (see Subsection 5.7 for the proof of the iterated integral identity (16) and see below for a heuristic explanation of this fact). So, the iterated high dimensional stochastic integral in (15) reduces to a simple product of functions. The function $\frac{\partial}{\partial y} b \colon (0, 1) \times \mathbb{R} \to \mathbb{R}$ is here the partial derivative $(\frac{\partial}{\partial y}b)(x,y)$ for $x \in (0,1)$ and $y \in \mathbb{R}$. Using (16) the numerical scheme (15) thus reduces to

$$\begin{aligned} Y_{n+1}^N = P_N\, e^{A\frac{T}{N^2}} \Bigg(& Y_n^N + \frac{T}{N^2} \cdot f(\cdot, Y_n^N) + b(\cdot, Y_n^N) \cdot \left(W^N_{\frac{(n+1)T}{N^2}} - W^N_{\frac{nT}{N^2}} \right) \\ & + \frac{1}{2} \left(\frac{\partial}{\partial y} b \right)(\cdot, Y_n^N) \cdot b(\cdot, Y_n^N) \cdot \left(\left(W^N_{\frac{(n+1)T}{N^2}} - W^N_{\frac{nT}{N^2}} \right)^2 - \frac{T}{N^2} \sum_{i=1}^N \eta_i (g_i)^2 \right) \Bigg) \end{aligned} \tag{17}$$

$\mathbb{P}$-a.s. for all $n \in \{0, 1, \ldots, N^2 - 1\}$ and all $N \in \mathbb{N}$. Note that only increments of the finite dimensional Wiener processes $W^N \colon [0,T] \times \Omega \to H$, $N \in \mathbb{N}$, are used in (17). Moreover, observe that, as in the case of (10), the infinite dimensional $\mathbb{R}$-Hilbert space H is projected down to the N-dimensional $\mathbb{R}$-Hilbert space $P_N(H)$ with $N \in \mathbb{N}$ and the infinite dimensional Wiener process $W \colon [0,T] \times \Omega \to H$ is approximated by the finite dimensional Wiener processes $W^N \colon [0,T] \times \Omega \to H$, $N \in \mathbb{N}$, for the spatial discretization in (17). For the temporal discretization in the scheme Y_n^N, $n \in \{0, 1, \ldots, N^2\}$, above the time interval $[0,T]$ is divided into N^2 subintervals, i.e., N^2 instead of N^3 time steps are used in (17), for $N \in \mathbb{N}$. In the following we explain why it is crucial to use N^2 time steps in (17) instead of N^3 time steps in the case of the linear implicit Euler scheme (10).

More formally, we now illustrate how efficiently the Milstein type algorithm (17) solves the strong approximation problem (9) of the SPDE (7). Theorem 1 (see Section 3 below) gives the existence of real numbers $C_r > 0$, $r \in (0, \frac{3}{2})$, such that

$$\left(\mathbb{E}\left[\int_0^1 \left| X_T(x) - Y_{N^2}^N(x) \right|^2 dx \right] \right)^{1/2} \leq C_r \cdot N^{(r - \frac{3}{2})} \tag{18}$$

for all $N \in \mathbb{N}$ and all arbitrarily small $r \in (0, \frac{3}{2})$. The approximation $Y_{N^2}^N$ thus converges in the root mean square sense to X_T with order $\frac{3}{2}-$ as N goes to infinity. The expression

$$\frac{1}{2} \left(\frac{\partial}{\partial y} b \right)(\cdot, Y_n^N) \cdot b(\cdot, Y_n^N) \cdot \left(\left(W^N_{\frac{(n+1)T}{N^2}} - W^N_{\frac{nT}{N^2}} \right)^2 - \frac{T}{N^2} \sum_{i=1}^N \eta_i (g_i)^2 \right) \tag{19}$$

for $n \in \{0, 1, \ldots, N^2 - 1\}$ and $N \in \mathbb{N}$ in the Milstein type approximation (17) contains additional information of the solution process of (7) and this allows us to use less time steps, N^2 in (17) instead of N^3 in (10), to achieve the same convergence rate as the linear implicit Euler scheme (10) (compare (13) and (18)). Nonetheless, (19) and hence the numerical method (17) can be simulated easily. The function $\frac{T}{N^2} \sum_{i=1}^N \eta_i \, (g_i)^2$ in (19) can be computed once in advance for which $O(N^2)$ computational operations are needed. Having computed $\frac{T}{N^2} \sum_{i=1}^N \eta_i \, (g_i)^2$, $O(N \log(N))$ further computational operations and random variables are needed to compute (19) from Y_n^N for one fixed $n \in \left\{0, 1, \ldots, N^2 - 1\right\}$ by using fast Fourier transform. Since $O(N \log(N))$ computational operations and independent standard normal random variables are needed for one time step and since N^2 time steps are used in (17), $O(N^3 \log(N))$ computational operations and random variables are needed to compute $Y_{N^2}^N$. Combining the computational effort $O(N^3 \log(N))$ and the convergence order $\frac{3}{2}-$ in (18) shows that the Milstein type method (17) needs about $O(\varepsilon^{-2})$ computational operations and independent standard normal random variables to achieve the desired precision $\varepsilon > 0$ in (9). To sum up, the Milstein type algorithm (17) requires about $O(\varepsilon^{-2})$ and the linear-impicit Euler scheme (10) requires about $O(\varepsilon^{-\frac{8}{3}})$ computational operations and independent standard normal random variables for solving the strong approximation problem (9) for the SPDE (7).

The convergence rates $O(\varepsilon^{-\frac{8}{3}})$ and $O(\varepsilon^{-2})$ are both asymptotic results as $\varepsilon > 0$ tends to zero. Therefore, from a more practical point of view, one may ask whether the Milstein type algorithm (17) solves the strong approximation problem (9) more efficiently than the linear implicit Euler scheme (10) for a given concrete $\varepsilon > 0$ and a given example of the form (7). In order to study this question we compare both methods in the case of a simple stochastic reaction diffusion equation. More formally, let $\kappa = \frac{1}{100}$, let $\xi \colon [0,1] \to \mathbb{R}$ be given by $\xi(x) = 0$ for all $x \in [0,1]$ and suppose that $f, b \colon (0,1) \times \mathbb{R} \to \mathbb{R}$ are given by $f(x,y) = 1 - y$ and $b(x,y) = \frac{1-y}{1+y^2}$ for all $x \in (0,1)$, $y \in \mathbb{R}$. The SPDE (7) thus reduces to

$$dX_t(x) = \left[\frac{1}{100} \frac{\partial^2}{\partial x^2} X_t(x) + 1 - X_t(x) \right] dt + \frac{1 - X_t(x)}{1 + X_t(x)^2} \, dW_t(x) \tag{20}$$

with $X_t(0) = X_t(1) = 0$ and $X_0 = 0$ for $x \in (0,1)$ and $t \in [0,1]$ (see also Section 4.1 for more details concerning this example). Additionally, assume that (9) for the SPDE (20) should be solved with the precision of say three decimals, i.e., with the precision $\varepsilon = \frac{1}{1000}$ in (9). In Figure 1 the approximation error in the sense of (9) of the linear implicit Euler approximation $Z_{N^3}^N$ (see (10)) and of the approximation $Y_{N^2}^N$ (see (17)) is plotted against the precise number of independent standard normal random variables needed to compute the corresponding approximation for $N \in \{2, 4, 8, 16, 32, 64, 128\}$: It turns out that $Z^{128}_{128^3}$ in the case of the linear implicit Euler scheme (10) and that $Y^{128}_{128^2}$ in the case of the Milstein type algorithm (17) achieve the desired precision $\varepsilon = \frac{1}{1000}$ in (9) for the SPDE (20). The MATLAB codes for simulating $Z^{128}_{128^3}$ via (10) and $Y^{128}_{128^2}$ via (17) for the SPDE (20) are presented below in Figure 2 and Figure 3 respectively. The differences of the codes and the additional code needed for the Milstein type algorithm (17) are printed bold in Figure 3. The MATLAB code in Figure 2 requires on our INTEL PENTIUM D running at 3.0 GHz a CPU time of about 15 minutes and 25.03 seconds (925.03 seconds) while the code in Figure 3 requires a CPU time of about 8.93 seconds to be evaluated on the same computer. So, on the above computer the Milstein type algorithm (17) is for the SPDE (20) about hundred times faster than the linear implicit Euler scheme (10) in order to achieve a precision of three decimals in (9). Further numerical examples for the Milstein type algorithm (17) can be found in Section 4 below.

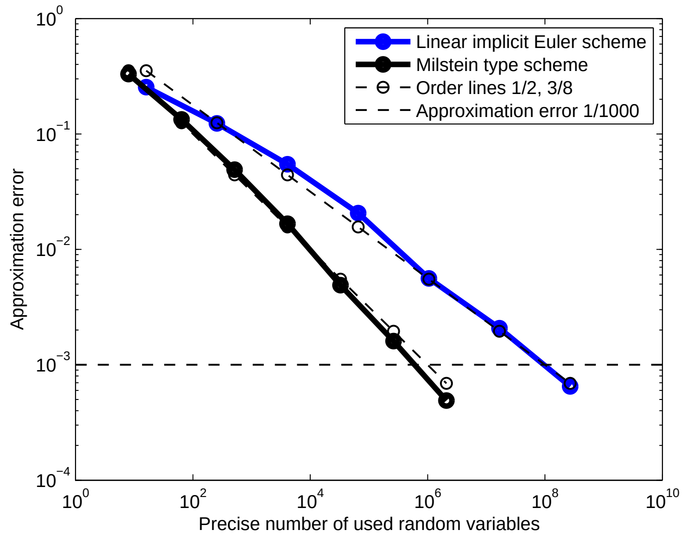

Figure 1: SPDE (20): Approximation error in the sense of (9) of the linear implicit Euler approximation $Z^N_{N^3}$ (see (10)) and of the Milstein type approximation $Y^N_{N^2}$ (see (17)) against the precise number of independent standard normal random variables needed to compute the corresponding approximation for $N \in \{2, 4, 8, 16, 32, 64, 128\}$.

```
N = 128; M = N^3; A = -pi^2*(1:N).^2/100; Y = zeros(1,N);
mu = (1:N).^-2; f = @(x) 1-x; b = @(x) (1-x)./(1+x.^2);
for m=1:M
  y = dst(Y) * sqrt(2);
  dW = dst( randn(1,N) .* sqrt(mu*2/M) );
  y = y + f(y)/M + b(y).*dW;
  Y = idst( y ) / sqrt(2) ./ ( 1 - A/M );
end
plot( (0:N+1)/(N+1), [0,dst(Y)*sqrt(2),0] );
```

Figure 2: MATLAB code for simulating the linear implicit Euler approximation $Z^N_{N^3}$ with $N = 128$ (see (10)) for the SPDE (20).

Having illustrated the efficiency of the method (17), we now take a short look at the literature of numerical analysis for SPDEs. First, it should be mentioned that any combination of finite elements, finite differences or spectral Galerkin methods for the spatial discretization and the linear implicit Euler scheme or also the linear implicit Crank-Nicolson scheme for the temporal discretization do not reduce the computational effort $O(\varepsilon^{-\frac{8}{3}})$ for the problem (9) in case of the SPDE (7). However, the splitting-up method (see [6, 7, 13, 19, 20, 21, 27] and the references therein) converges with a higher temporal order than the linear implicit Euler scheme. The key idea of the splitting-up method is to split the considered SPDE into appropriate subequations that are easier to solve than the original SPDE, e.g., that can be solved explicitly. The applicability of the splitting-up method thus essentially depends on the simplicity of the resulting subequations and can therefore in general not be used efficiently for nonlinear SPDEs. Nonetheless, in the case of an appropriate class of linear SPDEs, the splitting-up method and the Milstein type scheme here require nearly the same computational effort for solving the strong approximation problem (see Section 4.3 for a more detailed comparison of the splitting-up

```
N = 128; M = N^2; A = -pi^2*(1:N).^2/100; Y = zeros(1,N);
mu = (1:N).^-2; f = @(x) 1-x; b = @(x) (1-x)./(1+x.^2);
bb = @(x) (1-x).*(x.^2-2*x-1)/2./(1+x.^2).^3; g = zeros(1,N);
for n=1:N
  g = g+2*sin(n*(1:N)/(N+1)*pi).^2*mu(n)/M;
end
for m=1:M
  y = dst(Y) * sqrt(2);
  dW = dst( randn(1,N) .* sqrt(mu*2/M) );
  y = y + f(y)/M + b(y).*dW + bb(y).*(dW.^2 - g);
  Y = exp( A/M ) .* idst( y ) / sqrt(2);
end
plot( (0:N+1)/(N+1), [0,dst(Y)*sqrt(2),0] );
```

Figure 3: MATLAB code for simulating the Milstein type approximation $Y^N_{N^2}$ with $N = 128$ (see (17)) for the SPDE (20).

method and the Milstein type algorithm here). Additionally, in the case $f = 0$ in (7), T. Müller-Gronbach and K. Ritter proposed a new scheme which reduces the number of independent standard normal random variables needed for solving a similar problem as (9) from about $O(\varepsilon^{-\frac{8}{3}})$ to about $O(\varepsilon^{-2})$ (see [47] and also [48]). Nonetheless, the number of computational operations needed and thus the overall computational effort could not be reduced by the algorithm in [47]. Moreover, Milstein type schemes for SPDEs have been considered in [16, 35, 39, 46]. In [16], W. Grecksch and P. E. Kloeden proposed a Milstein like scheme for an SPDE driven by a scalar one-dimensional Brownian motion (see also [35]). In view of (6), the Milstein type scheme in [16, 35] can be simulated efficiently since the driving noise process is one-dimensional. Furthermore, in the case of a suitable linear SPDE, A. Lang, P.-L. Chow and J. Potthoff constructed in the interesting article [39] a scheme similar to (15) but with an additional term. (The additional term may be useful for decreasing the error constant but turns out not to be needed in order to achieve the higher approximation order due to Theorem 1 here.) In order to simulate the iterated stochastic integral in their scheme, they then suggest to omit the summands in the double sum in (5) for which $i \neq j$ holds (see (10) in [39] and also [40]). Their idea thus yields a scheme that can be simulated very efficiently, but does in general not converge with a higher order anymore, except for a linear SPDE driven by a scalar one-dimensional Brownian motion. Finally, based on Itô's formula, Y. S. Mishura and G. M. Shevchenko proposed in [46] the temporal approximation (14) under additonal smoothness assumptions of the driving noise process of the SPDE (8) which garantuee the existence of a strong solution and thus allow the application of Itô's formula (see Assumption C in [46]). The simulation of the iterated stochastic integrals in the Milstein type approximation in [46] remained an open question (see Remark 1.1 in [46]). To sum up, in the general setting of the possibly nonlinear SPDE (7), the Milstein type algorithm (17) is – to the best of our knowledge – the first numerical approximation method which has been shown to require asymptotically less computational operations and independent standard normal random variables than the required $O(\varepsilon^{-\frac{8}{3}})$ of the linear-implicit Euler scheme in order to solve the strong approximation problem (9) of the SPDE (7).

The rest of this article is organized as follows. In Section 2 the setting and the assumptions used are formulated. The numerical method and its convergence result (Theorem 1) are presented in Section 3. In Section 4 several examples of Theorem 1 including a stochastic heat equation and stochastic reaction diffusion equations in one and two dimensions are considered. The proof of Theorem 1 is postponed to Section 5.

Next let us add some concluding remarks. There are a number of directions for further research arising from this work. One is to analyze whether the exponential term in (17) can be replaced by a simpler mollifier such as $(I - \frac{T}{N^2}A)^{-1}$ for $N \in \mathbb{N}$. This would make the scheme even simpler to simulate. A second direction is to combine the temporal approximation in (17) with other spatial discretizations such as finite elements. This makes it possible to handle more complicated multidimensional domains on which the eigenfunctions of the Laplacian are not known explicitly. A third direction is to reduce the Lipschitz assumptions in Section 2 in order to handle further classes of SPDEs with non-globally Lipschitz nonlinearities such as stochastic Burgers equations, stochastic porous medium equations and hyperbolic SPDEs. Next a combination of the multilevel Monte Carlo approach (see [26, 15]) with the Milstein type algorithm here should result in a faster approximation of statistical quantities of the solution process of the SPDE (7). After a first preprint of this work has appeared, a few research articles related to this work have appeared; see [4, 3, 61, 5, 8]. Some of the above outlined future research directions and other issues such as Runge-Kutta type schemes for SPDEs based on the Milstein type scheme here and Milstein type schemes for SPDEs driven by non-Gaussian noise have been investigated in these articles.

Finally, let us point out limitations of the Milstein type algorithm presented here. There are essentially two assumptions which need to be fulfilled so that the above outlined Milstein type algorithm can be applied.

First, the noise must be of *trace class* type. Indeed, the covariance operator $Q\colon H \to H$ of the Wiener process is assumed to be a trace class operator here and we do not know how to treat the case where $Q\colon H \to H$ is bounded but without a finite trace with the above outlined Milstein type algorithm. Second, the diffusion coefficient B must satisfy a certain *commutativity type condition*; see Remark 1 below for details. In the setting of the SPDE (7) this condition is always fulfilled and no restriction is required even when the domain of the SPDE becomes multi-dimensional dimensional (see, e.g., Section 4.3) but in the more general setting of systems of SPDEs this results in a serious restriction on the diffusion coefficient of the considered system of SPDEs. In addition to these two essential restrictions, several further restrictive assumptions such as the semilinear structure of the SPDE, the explicit knowledge of the eigenfunctions of A and Q as well as Lipschitz assumptions on the nonlinearities F and B are used in this article. These further assumptions are, however, no serious restrictions since they can be relaxed or have already been relaxed (see [61, 4, 3, 8]).

2 Setting and assumptions

Throughout this article suppose that the following setting and the following assumptions are fulfilled. Let $T \in (0, \infty)$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a normal filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and let $(H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ and $(U, \langle\cdot,\cdot\rangle_U, \|\cdot\|_U)$ be two separable $\mathbb{R}$-Hilbert spaces. Moreover, let $Q\colon U \to U$ be a trace class operator and let $W\colon [0,T] \times \Omega \to U$ be a standard Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

Assumption 1 (Linear operator A)**.** *Let $\mathcal{I}$ be a finite or countable set and let $(\lambda_i)_{i\in\mathcal{I}} \subset (0,\infty)$ be a family of real numbers with $\inf_{i\in\mathcal{I}} \lambda_i \in (0,\infty)$. Moreover, let $(e_i)_{i\in\mathcal{I}}$ be an orthonormal basis of H and let $A\colon D(A) \subset H \to H$ be a linear operator with*

$$Av = \sum_{i\in\mathcal{I}} -\lambda_i \langle e_i, v\rangle_H \, e_i \tag{21}$$

for all $v \in D(A)$ and with $D(A) = \big\{w \in H\colon \sum_{i\in\mathcal{I}} |\lambda_i|^2 \, |\langle e_i, w\rangle_H|^2 < \infty\big\}$.

By $H_r := D((-A)^r)$ equipped with the norm $\|v\|_{H_r} := \|(-A)^r v\|_H$ for all $v \in H_r$ and all $r \in [0,\infty)$ we denote the $\mathbb{R}$-Hilbert spaces of domains of fractional powers of the linear operator $-A : D(A) \subset H \to H$.

Assumption 2 (Drift coefficient F)**.** *Let $\beta \in [0,1)$ be a real number and let $F\colon H_\beta \to H$ be a twice continuously Fréchet differentiable mapping with $\sup_{v\in H_\beta} \|F'(v)\|_{L(H)} < \infty$ and $\sup_{v\in H_\beta} \|F''(v)\|_{L^{(2)}(H_\beta,H)} < \infty$.*

For formulating the assumption on the diffusion coefficient of our SPDE, denote by $\big(U_0, \langle\cdot,\cdot\rangle_{U_0}, \|\cdot\|_{U_0}\big)$ the separable $\mathbb{R}$-Hilbert space $U_0 := Q^{1/2}(U)$ with $\langle v, w\rangle_{U_0} = \big\langle Q^{-1/2}v, Q^{-1/2}w\big\rangle_U$ for all $v, w \in U_0$ (see, for example, Section 2.3.2 in [52]). For an arbitrary bounded linear operator $S \in L(U)$, we denote by $S^{-1}\colon \mathrm{im}(S) \subset U \to U$ the pseudo inverse of S (see, for instance, Appendix C in [52]).

Assumption 3 (Diffusion coefficient B)**.** *Let $B\colon H_\beta \to HS(U_0, H)$ be a twice continuously Fréchet differentiable mapping with $\sup_{v\in H_\beta}\|B'(v)\|_{L(H,HS(U_0,H))} < \infty$ and $\sup_{v\in H_\beta}\|B''(v)\|_{L^{(2)}(H_\beta,HS(U_0,H))} < \infty$. Moreover, let $\alpha, c \in (0,\infty)$, $\delta, \vartheta \in (0, \frac{1}{2})$, $\gamma \in [\max(\delta,\beta), \delta + \frac{1}{2})$ be real numbers, let $B(H_\delta) \subset HS(U_0, H_\delta)$ and suppose that*

$$\|B(u)\|_{HS(U_0,H_\delta)} \leq c\left(1 + \|u\|_{H_\delta}\right), \tag{22}$$

$$\|B'(v)\,B(v) - B'(w)\,B(w)\|_{HS^{(2)}(U_0,H)} \leq c\,\|v - w\|_H\,, \tag{23}$$

$$\big\|(-A)^{-\vartheta}B(v)Q^{-\alpha}\big\|_{HS(U_0,H)} \leq c\left(1 + \|v\|_{H_\gamma}\right) \tag{24}$$

for all $u \in H_\delta$ and all $v, w \in H_\gamma$. Finally, let the bilinear Hilbert-Schmidt operator $B'(v)B(v) \in HS^{(2)}(U_0, H)$ be symmetric for all $v \in H_\beta$.

We now add some comments concerning Assumption 3. First, we note that Assumption 3 implies $\beta \leq \delta + \frac{1}{2}$. Indeed, $\beta > \delta + \frac{1}{2}$ implies $[\max(\delta,\beta), \delta + \frac{1}{2}) = \emptyset$, which contradicts to $\gamma \in [\max(\delta,\beta), \delta + \frac{1}{2})$ in Assumption 3. Furthemore, we observe that the above assumption $\sup_{v\in H_\beta} \|B'(v)\|_{L(H,HS(U_0,H))} < \infty$ and the fact that H_β is dense in H imply that $B\colon H_\beta \to HS(U_0,H)$ can be continuously extended to a globally Lipschitz continuous mapping $\tilde{B}\colon H \to HS(U_0,H)$ from H to $HS(U_0,H)$. Here and below we do not distinguish between $B\colon H_\beta \to HS(U_0,H)$ and its extension $\tilde{B}\colon H \to HS(U_0,H)$ for simplicity of presentation. Additionally, we note that the operator $B'(v)B(v)\colon U_0 \times U_0 \to H$ given by

$$\Big(B'(v)B(v)\Big)(u,\tilde{u}) = \Big(B'(v)\big(B(v)u\big)\Big)(\tilde{u}) \tag{25}$$

for all $u, \tilde{u} \in U_0$ is a bilinear Hilbert-Schmidt operator in

$$HS^{(2)}(U_0,H) \cong HS(\overline{U_0 \otimes U_0}, H) \tag{26}$$

for all $v \in H_\beta$. Next we add a remark on the symmetry assumption on these bilinear Hilbert-Schmidt operators in Assumption 3.

Remark 1 (Commutative noise in infinite dimensions)**.** *The assumed symmetry of the bilinear Hilbert-Schmidt operator $B'(v)B(v) \in HS^{(2)}(U_0, H)$ in Assumption 3 reads as*

$$\Big(B'(v)\big(B(v)u\big)\Big)(\tilde{u}) = \Big(B'(v)\big(B(v)\tilde{u}\big)\Big)(u) \tag{27}$$

for all $u, \tilde{u} \in U_0$ and all $v \in H_\beta$. Note that (27) *is the abstract possibly infinite dimensional coordinate free analog of* (4)*. More formally, if $H = \mathbb{R}^d$, $U = \mathbb{R}^m$ and $Q = I$ with $d, m \in \mathbb{N}$ holds, then* (27) *reduces to* (4) *(with σ replaced by B).*

In Section 4 below we describe a natural class of examples of SPDEs which satisfy the commutativity condition (27) (see (47) and (51) for details). Finally, we emphasize that we do not assume that the linear operators $A\colon D(A) \subset H \to H$ and $Q\colon U \to U$ are simultaneously diagonalizable.

Assumption 4 (Initial value ξ)**.** *Let $\xi\colon \Omega \to H_\gamma$ be an $\mathcal{F}_0/\mathcal{B}\left(H_\gamma\right)$-measurable mapping with $\mathbb{E}\big[\|\xi\|^4_{H_\gamma}\big] < \infty$.*

These assumptions suffice to ensure the existence of an up to modifications unique solution of the SPDE (28).

Proposition 1 (Existence, uniqueness and regularity of solutions)**.** *Let Assumptions 1-4 in Section 2 be fulfilled. Then there exists an up to modifications unique predictable stochastic process $X\colon [0,T] \times \Omega \to H_\gamma$ which fulfills $\sup_{t\in[0,T]} \mathbb{E}\big[\|X_t\|^4_{H_\gamma}\big] < \infty$, $\sup_{t\in[0,T]} \mathbb{E}\big[\|B(X_t)\|^4_{HS(U_0,H_\delta)}\big] < \infty$ and*

$$X_t = e^{At}\xi + \int_0^t e^{A(t-s)} F(X_s)\, ds + \int_0^t e^{A(t-s)} B(X_s)\, dW_s \tag{28}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$. Moreover, we have

$$\sup_{\substack{t_1,t_2\in[0,T]\\ t_1\neq t_2}} \frac{\big(\mathbb{E}\big[\|X_{t_2} - X_{t_1}\|^4_{H_r}\big]\big)^{\frac{1}{4}}}{|t_2 - t_1|^{\min(\gamma - r, \frac{1}{2})}} < \infty \tag{29}$$

for all $r \in [0,\gamma]$. Additionally, the solution process X_t, $t \in [0,T]$, is continuous with respect to $\big(\mathbb{E}\big[\|\cdot\|^p_{H_\gamma}\big]\big)^{1/p}$.

Proposition 1 immediately follows from Theorem 1 in [32].

3 Numerical scheme and main result

In this section the numerical method is introduced and its convergence result is stated. To this end let $\mathcal{J}$ be a set, let $(g_j)_{j\in\mathcal{J}} \subset U$ be an orthonormal basis of eigenfunctions of $Q\colon U \to U$ and let $(\eta_j)_{j\in\mathcal{J}} \subset [0,\infty)$ be the corresponding family of eigenvalues (such an orthonormal basis of eigenfunctions exists since $Q\colon U \to U$ is a trace class operator, see, e.g., Proposition 2.1.5 in [52]). In particular, we have

$$Qu = \sum_{j\in\mathcal{J}} \eta_j \left\langle g_j, u\right\rangle_U g_j \tag{30}$$

for all $u \in U$. Additionally, let $(\mathcal{I}_N)_{N\in\mathbb{N}}$ and $(\mathcal{J}_K)_{K\in\mathbb{N}}$ be sequences of finite subsets of $\mathcal{I}$ and $\mathcal{J}$ respectively. Then we define the linear projection operators $P_N\colon H \to H$, $N \in \mathbb{N}$, by $P_N(v) := \sum_{i\in\mathcal{I}_N} \left\langle e_i, v\right\rangle_H e_i$ for all $v \in H$ and all $N \in \mathbb{N}$. Furthermore, we define Wiener processes $W^K\colon [0,T]\times\Omega \to U_0$, $K \in \mathbb{N}$, by

$$W^K_t(\omega) := \sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} \left\langle g_j, W_t(\omega)\right\rangle_U g_j \tag{31}$$

for all $t \in [0,T]$, $\omega \in \Omega$ and all $K \in \mathbb{N}$. We also use the $\mathcal{F}/\mathcal{B}(U_0)$-measurable mappings $\Delta W^{M,K}_m\colon \Omega \to U_0$, $m \in \{0,1,\dots,M-1\}$, $M,K \in \mathbb{N}$, given by $\Delta W^{M,K}_m(\omega) := W^K_{\frac{(m+1)T}{M}}(\omega) - W^K_{\frac{mT}{M}}(\omega)$ for all $\omega \in \Omega$, $m \in \{0,1,\dots,M-1\}$ and all $M,K \in \mathbb{N}$. The numerical scheme which we denote by $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $Y^{N,M,K}_m\colon \Omega \to H_N$, $m \in \{0,1,\dots,M\}$, $N,M,K \in \mathbb{N}$, is then given by $Y^{N,M,K}_0 := P_N(\xi)$ and

$$\begin{aligned} Y^{N,M,K}_{m+1} := P_N\, e^{A\frac{T}{M}} \Bigg(& Y^{N,M,K}_m + \frac{T}{M}\cdot F\big(Y^{N,M,K}_m\big) + B\big(Y^{N,M,K}_m\big)\Delta W^{M,K}_m \\ & + \frac{1}{2} B'\big(Y^{N,M,K}_m\big)\Big(B\big(Y^{N,M,K}_m\big)\Delta W^{M,K}_m\Big)\Delta W^{M,K}_m \\ & - \frac{T}{2M} \sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} \eta_j B'\big(Y^{N,M,K}_m\big)\Big(B\big(Y^{N,M,K}_m\big)\, g_j\Big) g_j \Bigg) \end{aligned} \tag{32}$$

for all $m \in \{0, 1, \ldots, M-1\}$ and all $N, M, K \in \mathbb{N}$. Note that only increments of the Wiener processes $W^K \colon [0,T] \times \Omega \to U_0$, $K \in \mathbb{N}$, are used in the scheme above and we emphasize that for many SPDEs the method (33) is easy to simulate and to implement (see Sections 1 and 4 for a few examples). Moreover, observe that the scheme (33) can also be written as

$$\begin{aligned} Y_{m+1}^{N,M,K} = P_N\, e^{A\frac{T}{M}} \Bigg(& Y_m^{N,M,K} + \frac{T}{M} \cdot F\big(Y_m^{N,M,K}\big) + B\big(Y_m^{N,M,K}\big)\Delta W_m^{M,K} \\ & + \frac{1}{2} B'\big(Y_m^{N,M,K}\big)\Big(B\big(Y_m^{N,M,K}\big)\Delta W_m^{M,K}\Big)\Delta W_m^{M,K} \\ & - \frac{1}{2}\mathbb{E}\Big[B'\big(Y_m^{N,M,K}\big)\Big(B\big(Y_m^{N,M,K}\big)\Delta W_m^{M,K}\Big)\Delta W_m^{M,K} \,\big|\, \mathcal{F}_{\frac{mT}{M}}\Big]\Bigg) \end{aligned} \tag{33}$$

for all $m \in \{0, 1, \ldots, M-1\}$ and all $N, M, K \in \mathbb{N}$. In the next step the convergence result for the scheme (33) is presented.

Theorem 1 (Main result)**.** *Let Assumptions 1-4 in Section 2 be fulfilled. Then there is a real number $C \in (0, \infty)$ such that*

$$\left(\mathbb{E}\Big[\big\|X_{\frac{mT}{M}} - Y_m^{N,M,K}\big\|_H^2\Big]\right)^{\frac{1}{2}} \leq C\left(\left(\inf_{i \in \mathcal{I}\setminus\mathcal{I}_N} \lambda_i\right)^{-\gamma} + \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)}\right) \tag{34}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$.

We now explain the result of Theorem 1 in more detail. The strong root mean square difference $\big(\mathbb{E}\big[\|X_{\frac{mT}{M}} - Y_m^{N,M,K}\|_H^2\big]\big)^{1/2}$ for $m \in \{0, 1, \ldots, M\}$ and for $N, M, K \in \mathbb{N}$ of the exact solution of the SPDE (28) and of the numerical solution (33) is estimated in (34) by a constant times the sum of three terms. The first term, i.e., $(\inf_{i \in \mathcal{I}\setminus\mathcal{I}_N} \lambda_i)^{-\gamma}$ for $N \in \mathbb{N}$, arises due to discretizing the exact solution spatially, i.e., due to $\mathbb{E}\big[\|X_t - P_N(X_t)\|_H^2\big]$ for $N \in \mathbb{N}$ and $t \in [0,T]$. The second expression, i.e., $(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j)^{\alpha}$ for $K \in \mathbb{N}$, occurs due to discretizing the noise spatially, i.e., due to $\mathbb{E}\big[\|W_t - W_t^K\|_H^2\big]$ for $K \in \mathbb{N}$ and $t \in [0,T]$. If $U_0 \subset U$ is finite dimensional we choose $\mathcal{J}_K := \big\{j \in \mathcal{J} \big| \eta_j \neq 0\big\}$ for all $K \in \mathbb{N}$ and obtain $(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j)^{\alpha} = 0$ for all $K \in \mathbb{N}$ in that case. The third term, i.e., $M^{-\min(2(\gamma-\beta),\gamma)}$ for $M \in \mathbb{N}$, corresponds to the temporal discretization error and converges to zero as the number of time steps $M \in \mathbb{N}$ goes to infinity.

4 Examples

In this section Theorem 1 is illustrated with various examples. To this end let $d \in \{1, 2, 3\}$ and let $H = U = L^2((0,1)^d, \mathbb{R})$ be the $\mathbb{R}$-Hilbert space of equivalence classes of Lebesgue square integrable functions from $(0,1)^d$ to $\mathbb{R}$. As usual we do not distinguish between a Lebesgue square integrable function from $(0,1)^d$ to $\mathbb{R}$ and its equivalence class in H. The scalar product and the norm in H and U are given by

$$\langle v, w\rangle_H = \langle v, w\rangle_U = \int_{(0,1)^d} v(x) \cdot w(x)\, dx \tag{35}$$

and

$$\|v\|_H = \|v\|_U = \left(\int_{(0,1)^d} |v(x)|^2\, dx\right)^{\frac{1}{2}} \tag{36}$$

for all $v, w \in H = U$. Additionally, the notations

$$\|v\|_{C((0,1)^d,\mathbb{R})} := \sup_{x \in (0,1)^d} |v(x)| \in [0, \infty] \tag{37}$$

and

$$\|v\|_{C^r((0,1)^d,\mathbb{R})} := \sup_{x \in (0,1)^d} |v(x)| + \sup_{\substack{x,y \in (0,1)^d \\ x \neq y}} \frac{|v(x) - v(y)|}{\|x - y\|_{\mathbb{R}^d}^r} \in [0, \infty] \tag{38}$$

are used troughout this section for all functions $v \colon (0,1)^d \to \mathbb{R}$ and all $r \in (0,1]$. Here and below we use the Euclidean norms $\|x\|_{\mathbb{R}^n} := (|x_1|^2 + \ldots + |x_n|^2)^{1/2}$ for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and all $n \in \mathbb{N}$. Concerning the Wiener process $W \colon [0,T] \times \Omega \to U$ we assume that the eigenfunctions $g_j \in U$, $j \in \mathcal{J}$, of the covariance operator $Q \colon U \to U$ are continuous and satisfy

$$\sup_{j \in \mathcal{J}} \|g_j\|_{C((0,1)^d,\mathbb{R})} < \infty \qquad \text{and} \qquad \sum_{j \in \mathcal{J}} \Big(\eta_j\, \|g_j\|_{C^\rho((0,1)^d,\mathbb{R})}^2\Big) < \infty \tag{39}$$

for some $\rho \in (0,1)$ in this section. We will give some concrete examples for the $(g_j)_{j\in\mathcal{J}}$ that fulfill (39) later. Additionally, we denote by $s_n(D) \in [0,\infty)$, $n \in \mathbb{N}$, the sequence of characteristic numbers of a compact operator $D\colon H \to H$ (see, e.g., Section 9 in Chapter XI in [12]). Finally, we define the Schatten norms by

$$\|D\|_{S_p(H)} := \left(\sum_{n=1}^{\infty} |s_n(D)|^p \right)^{\frac{1}{p}} \in [0,\infty] \tag{40}$$

for all compact operators $D\colon H \to H$ and all $p \in [1,\infty)$ (see also the above named reference).

We now present a prominent example of **the linear operator A in Assumption 1**. Let $\mathcal{I} = \mathbb{N}^d$ and let $e_i \in H$ for $i \in \mathcal{I}$ be given by

$$e_i(x) = 2^{\frac{d}{2}} \sin(i_1 \pi x_1) \cdot \ldots \cdot \sin(i_d \pi x_d) \tag{41}$$

for all $x = (x_1, \ldots, x_d) \in (0,1)^d$ and all $i = (i_1, \ldots, i_d) \in \mathbb{N}^d$. Additionally, let $\kappa \in (0,\infty)$ be a fixed real number and let $(\lambda_i)_{i\in\mathcal{I}}$ be given by

$$\lambda_i = \kappa\, \pi^2 \left((i_1)^2 + \ldots + (i_d)^2 \right) \tag{42}$$

for all $i = (i_1, \ldots, i_d) \in \mathbb{N}^d$. Hence, the linear operator $A\colon D(A) \subset H \to H$ in Assumption 1 reduces to the Laplacian with Dirichlet boundary conditions times the constant $\kappa \in (0,\infty)$, i.e.,

$$Av = \kappa \cdot \Delta v = \kappa \left\{ \left(\frac{\partial^2}{\partial x_1^2} \right) v + \ldots + \left(\frac{\partial^2}{\partial x_d^2} \right) v \right\} \tag{43}$$

for all $v \in D(A)$ in this section. Furthermore, let $(\mathcal{I}_N)_{N\in\mathbb{N}}$ be given by $\mathcal{I}_N = \{1, \ldots, N\}^d$ for all $N \in \mathbb{N}$.

In order to describe a natural candidate for **the drift coefficient in Assumption 2**, let $\beta = \frac{d}{5}$ and let $f\colon (0,1)^d \times \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function with $\int_{(0,1)^d} |f(x,0)|^2 \, dx < \infty$ and $\sup_{x\in(0,1)^d} \sup_{y\in\mathbb{R}} \left| \left(\frac{\partial^n}{\partial y^n} f \right)(x,y) \right| < \infty$ for all $n \in \{1,2\}$. Then the (in general nonlinear) operator $F\colon H_\beta \to H$ given by

$$\big(F(v)\big)(x) = f(x, v(x)) \tag{44}$$

for all $x \in (0,1)^d$ and all $v \in H_\beta$ satisfies Assumption 2 since $H_\beta = H_{\frac{d}{5}} \subset L^5((0,1)^d, \mathbb{R})$ continuously (see Remark 6.94 in [53]).

In the next step a natural example for **the diffusion coefficient in Assumption 3** is given. Let $b\colon (0,1)^d \times \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function with

$$|b(x,0)| \leq q, \qquad \left| \left(\frac{\partial^n}{\partial y^n} b \right)(x,y) \right| \leq q, \qquad \left\| \left(\frac{\partial}{\partial x} b \right)(x,y) \right\|_{L(\mathbb{R}^d,\mathbb{R})} \leq q \tag{45}$$

and

$$\left| \left(\frac{\partial}{\partial y} b \right)(x,y) \cdot b(x,y) - \left(\frac{\partial}{\partial y} b \right)(x,z) \cdot b(x,z) \right| \leq q\, |y - z| \tag{46}$$

for all $x \in (0,1)^d$, $y, z \in \mathbb{R}$, $n \in \{1,2\}$ and some given $q \in (0,\infty)$. We refer to Subsections 4.1-4.3 below for concrete functions $b\colon (0,1)^d \times \mathbb{R} \to \mathbb{R}$ satisfying (45) and (46). Now let $B\colon H_\beta \to HS(U_0, H)$ be the (in general nonlinear) operator

$$\Big(B(v)u\Big)(x) = b(x, v(x)) \cdot u(x) \tag{47}$$

for all $x \in (0,1)^d$, $v \in H_\beta$ and all $u \in U_0 \subset U = H$. We now check step by step that $B\colon H_\beta \to HS(U_0, H)$ given by (47) satisfies Assumption 3. First of all, B is well defined. More precisely, we have

$$\begin{aligned} \|B(v)\|_{HS(U_0,H)}^2 &= \sum_{j\in\mathcal{J}} \left\| B(v) \sqrt{\eta_j} g_j \right\|_H^2 = \sum_{j\in\mathcal{J}} \eta_j \, \|B(v) g_j\|_H^2 \\ &= \sum_{j\in\mathcal{J}} \eta_j \left(\int_{(0,1)^d} |b(x, v(x)) \cdot g_j(x)|^2 \, dx \right) \leq \sum_{j\in\mathcal{J}} \eta_j \left(\int_{(0,1)^d} |b(x, v(x))|^2 \, dx \right) \left(\sup_{x\in(0,1)^d} |g_j(x)|^2 \right) \end{aligned}$$

and hence

$$\begin{aligned} \|B(v)\|_{HS(U_0,H)}^2 &\leq \sum_{j\in\mathcal{J}} \eta_j \left(\int_{(0,1)^d} \left(|b(x, v(x)) - b(x,0)| + |b(x,0)| \right)^2 dx \right) \|g_j\|_{C((0,1)^d,\mathbb{R})}^2 \\ &\leq \sum_{j\in\mathcal{J}} q^2 \eta_j \left(\int_{(0,1)^d} \left(|v(x)| + 1 \right)^2 dx \right) \|g_j\|_{C((0,1)^d,\mathbb{R})}^2 \end{aligned}$$

and finally

$$\begin{aligned}
\|B(v)\|_{HS(U_0,H)} &\leq q\left(\|v\|_H+1\right)\left(\sum_{j\in\mathcal{J}}\eta_j\|g_j\|^2_{C((0,1)^d,\mathbb{R})}\right)^{\frac{1}{2}}\\
&\leq q\left(\|v\|_H+1\right)\left(\sum_{j\in\mathcal{J}}\eta_j\right)^{\frac{1}{2}}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)\\
&= q\sqrt{\mathrm{Tr}(Q)}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)\left(\|v\|_H+1\right)<\infty
\end{aligned} \tag{48}$$

for all $v\in H_\beta$ which indeed shows that B is well defined. Moreover, B is twice continuously Fréchet differentiable and we have

$$\begin{aligned}
\|B'(v)u\|^2_{HS(U_0,H)} &= \sum_{j\in\mathcal{J}}\left\|\left(B'(v)u\right)\sqrt{\eta_j}g_j\right\|^2_H = \sum_{j\in\mathcal{J}}\eta_j\left\|\left(B'(v)u\right)g_j\right\|^2_H\\
&= \sum_{j\in\mathcal{J}}\eta_j\left(\int_{(0,1)^d}\left|\left(\frac{\partial}{\partial y}b\right)(x,v(x))\cdot u(x)\cdot g_j(x)\right|^2dx\right)\leq\sum_{j\in\mathcal{J}}q^2\eta_j\left(\int_{(0,1)^d}|u(x)\cdot g_j(x)|^2\,dx\right)
\end{aligned}$$

and hence

$$\begin{aligned}
\|B'(v)u\|_{HS(U_0,H)} &\leq\left(\sum_{j\in\mathcal{J}}q^2\eta_j\|u\|^2_H\|g_j\|^2_{C((0,1)^d,\mathbb{R})}\right)^{1/2}\\
&\leq q\|u\|_H\left(\sum_{j\in\mathcal{J}}\eta_j\right)^{1/2}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)=q\sqrt{\mathrm{Tr}(Q)}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)\|u\|_H
\end{aligned}$$

for all $u,v\in H_\beta$ which shows

$$\sup_{v\in H_\beta}\|B'(v)\|_{L(H,HS(U_0,H))}\leq q\sqrt{\mathrm{Tr}(Q)}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)<\infty. \tag{49}$$

Additionally, we have

$$\begin{aligned}
\|B''(v)(u,w)\|^2_{HS(U_0,H)} &= \sum_{j\in\mathcal{J}}\left\|B''(v)(u,w)\sqrt{\eta_j}g_j\right\|^2_H\\
&= \sum_{j\in\mathcal{J}}\eta_j\left(\int_{(0,1)^d}\left|\left(\frac{\partial^2}{\partial y^2}b\right)(x,v(x))\cdot u(x)\cdot w(x)\cdot g_j(x)\right|^2dx\right)\\
&\leq\sum_{j\in\mathcal{J}}q^2\eta_j\left(\int_{(0,1)^d}|u(x)\cdot w(x)|^2\,dx\right)\|g_j\|^2_{C((0,1)^d,\mathbb{R})}
\end{aligned}$$

and using $L^5\big((0,1)^d,\mathbb{R}\big)\subset L^4\big((0,1)^d,\mathbb{R}\big)$ continuously shows

$$\begin{aligned}
&\|B''(v)(u,w)\|_{HS(U_0,H)}\\
&\leq q\sqrt{\mathrm{Tr}(Q)}\left(\int_{(0,1)^d}|u(x)|^4\,dx\right)^{\frac{1}{4}}\left(\int_{(0,1)^d}|w(x)|^4\,dx\right)^{\frac{1}{4}}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)\\
&\leq q\sqrt{\mathrm{Tr}(Q)}\left(\sup_{j\in\mathcal{J}}\|g_j\|_{C((0,1)^d,\mathbb{R})}\right)\left(\int_{(0,1)^d}|u(x)|^5\,dx\right)^{\frac{1}{5}}\left(\int_{(0,1)^d}|w(x)|^5\,dx\right)^{\frac{1}{5}}
\end{aligned}$$

for all $u,v,w\in H_\beta$. Therefore, $H_\beta\subset L^5\left((0,1)^d,\mathbb{R}\right)$ continuously shows

$$\sup_{v\in H_\beta}\|B''(v)\|_{L^{(2)}(H_\beta,HS(U_0,H))}<\infty \tag{50}$$

due to (39) and hence, it remains to establish (22)-(24) and the symmetry of $B'(v)B(v)\in HS^{(2)}(U_0.H)$ for all $v\in H_\beta$. For the latter one, note that

$$\Big(\big(B'(v)B(v)\big)(u,\tilde{u})\Big)(x)=\Big(B'(v)\big(B(v)u\big)\tilde{u}\Big)(x)=\left(\frac{\partial}{\partial y}b\right)(x,v(x))\cdot b(x,v(x))\cdot u(x)\cdot\tilde{u}(x) \tag{51}$$

for all $x \in (0,1)^d$, $u, \tilde{u} \in U_0$ and all $v \in H_\beta$ which immediately shows that $B'(v)B(v) \in HS^{(2)}(U_0, H)$ is symmetric for all $v \in H_\beta$ (see (27)). Moreover, we have

$$\|B'(v)B(v) - B'(w)B(w)\|^2_{HS^{(2)}(U_0,H)} = \sum_{j,k \in \mathcal{J}} \eta_j \eta_k \, \|B'(v)\,(B(v)g_j)\,g_k - B'(w)\,(B(w)g_j)\,g_k\|^2_H$$

$$\leq \sum_{j,k\in\mathcal{J}} \eta_j\eta_k \left(\int_{(0,1)^d} \left| \left(\frac{\partial}{\partial y} b\right)(x, v(x)) \cdot b(x, v(x)) - \left(\frac{\partial}{\partial y} b\right)(x, w(x)) \cdot b(x, w(x)) \right|^2 dx \right) \left[\sup_{l \in \mathcal{J}} \|g_l\|^4_{C((0,1)^d,\mathbb{R})} \right]$$

and using (46) yields

$$\begin{aligned} \|B'(v)B(v) - B'(w)B(w)\|_{HS^{(2)}(U_0,H)} &\leq q\,\|v-w\|_H \left(\sum_{j,k\in\mathcal{J}} \eta_j\eta_k \right)^{\frac{1}{2}} \left(\sup_{i\in\mathcal{J}} \|g_j\|^2_{C((0,1)^d,\mathbb{R})} \right) \\ &= q\mathrm{Tr}(Q) \left(\sup_{i\in\mathcal{J}} \|g_j\|^2_{C((0,1)^d,\mathbb{R})} \right) \|v-w\|_H \end{aligned}$$

for all $v, w \in H_\beta$ which shows that (23) indeed holds. Estimates (22) and (24) will be verified in the more concrete examples in Subsections 4.1-4.3 below.

Concerning **the initial value in Assumption 4**, let $x_0\colon [0,1]^d \to \mathbb{R}$ be a twice continuously differentiable function with $x_0|_{\partial(0,1)^d} \equiv 0$. Then the $\mathcal{F}_0/\mathcal{B}(H_\gamma)$-measurable mapping $\xi\colon \Omega \to H_\gamma$ given by $\xi(\omega) = x_0$ for all $\omega \in \Omega$ fulfills Assumption 4 for all $\gamma \in (0,1)$.

Having constructed examples for Assumptions 1-4, we now formulate the SPDE (28) in the setting of this section. More precisely, in the setting above the SPDE (28) reduces to

$$dX_t(x) = \Big[\kappa \Delta X_t(x) + f(x, X_t(x))\Big]\,dt + b(x, X_t(x))\,dW_t(x) \tag{52}$$

with $X_{t\,|\,\partial(0,1)^d} \equiv 0$ and $X_0(x) = x_0(x)$ for $t \in [0,T]$ and $x \in (0,1)^d$. Moreover, we define a family $\beta^j\colon [0,T]\times\Omega \to \mathbb{R}$, $j \in \{k \in \mathcal{J}\colon \eta_k \neq 0\}$, of independent standard Brownian motions by

$$\beta^j_t(\omega) := \frac{1}{\sqrt{\eta_j}}\,\langle g_j, W_t(\omega)\rangle_U$$

for all $\omega \in \Omega$, $t \in [0,T]$ and all $j \in \mathcal{J}$ with $\eta_j \neq 0$. Using this notation, the SPDE (52) can be written as

$$dX_t(x) = \Big[\kappa\Delta X_t(x) + f(x, X_t(x))\Big]\,dt + \sum_{\substack{j\in\mathcal{J} \\ \eta_j\neq 0}} \Big[b(x, X_t(x))\sqrt{\eta_j}\,g_j(x)\Big]\,d\beta^j_t \tag{53}$$

with $X_{t\,|\,\partial(0,1)^d} \equiv 0$ and $X_0(x) = x_0(x)$ for $t \in [0,T]$ and $x \in (0,1)^d$. The Milstein type algorithm (33) applied to the SPDE (52) then reduces to $Y_0^{N,M,K} = P_N(x_0)$ and

$$\begin{aligned} Y^{N,M,K}_{m+1} = P_N\, e^{A\frac{T}{M}} \Bigg(& Y^{N,M,K}_m + \frac{T}{M}\cdot f(\cdot, Y^{N,M,K}_m) + b(\cdot, Y^{N,M,K}_m)\cdot \Delta W^{M,K}_m \\ & + \frac{1}{2}\left(\frac{\partial}{\partial y} b\right)(\cdot, Y^{N,M,K}_m)\cdot b(\cdot, Y^{N,M,K}_m)\cdot \left(\left(\Delta W^{M,K}_m\right)^2 - \frac{T}{M}\sum_{j\in\mathcal{J}_K} \eta_j (g_j)^2 \right) \Bigg) \end{aligned} \tag{54}$$

for all $m \in \{0, 1, \ldots, M-1\}$ and all $N, M, K \in \mathbb{N}$. Finally, Theorem 1 shows the existence of a real number $C \in (0,\infty)$ such that

$$\left(\mathbb{E}\left[\int_{(0,1)^d} \left| X_T(x) - Y^{N,M,K}_M(x) \right|^2 dx \right] \right)^{1/2} \leq C \left(N^{-2\gamma} + \left(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j \right)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)} \right) \tag{55}$$

for all $N, M, K \in \mathbb{N}$. We now illustrate estimate (55) in the following three more concrete examples. We begin with the introductory example from Section 1 (see (7) and (20)).

4.1 A one-dimensional stochastic reaction diffusion equation

In this subsection let $d = 1$, $T = 1$, $\kappa = \frac{1}{100}$, let $x_0\colon [0,1] \to \mathbb{R}$ be given by $x_0(x) = 0$ for all $x \in [0,1]$, let $f, b\colon (0,1)\times\mathbb{R} \to \mathbb{R}$ be given by $f(x,y) = 1 - y$ and $b(x,y) = \frac{1-y}{1+y^2}$ for all $x \in (0,1)$, $y \in \mathbb{R}$, let $\mathcal{J} = \mathbb{N}$, let $\mathcal{J}_K = \{1, 2, \ldots, K\}$ for all $K \in \mathbb{N}$, let $\eta_j = \frac{1}{j^2}$ and let $g_j = e_j$ for all $j \in \mathbb{N}$. The SPDE (53) thus reduces to

$$dX_t(x) = \left[\frac{1}{100}\frac{\partial^2}{\partial x^2} X_t(x) + 1 - X_t(x)\right] dt + \sum_{j=1}^{\infty} \frac{1 - X_t(x)}{1 + X_t(x)^2}\,\frac{\sqrt{2}}{j}\,\sin(j\pi x)\,d\beta^j_t \tag{56}$$

with $X_t(0) = X_t(1) = 0$ and $X_0(x) = 0$ for $x \in (0,1)$ and $t \in [0,1]$. The SPDE (56) is nothing else than equation (20) in the introduction. In order to apply Theorem 1 it remains to verify (22) and (24). Estimate (22) is fulfilled for all $\delta \in (0, \frac{1}{4})$ here due to Subsection 4.3 in [32]. In order to establish (24) several preparations are needed. More formally, let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ be a further probability space on which a sequence $\chi_i \colon \tilde{\Omega} \to \mathbb{R}$, $i \in \mathbb{N}$, of $\tilde{\mathcal{F}}/\mathcal{B}(\mathbb{R})$-measurable independent standard normal random variables is defined. Then we define the $\tilde{\mathcal{F}}/\mathcal{B}(U_0)$-measurable mappings $\chi^{K,\vartheta} \colon \tilde{\Omega} \to U_0$ by

$$\chi^{K,\vartheta}(\omega, x) := \sum_{i=1}^{K} \chi_i(\omega)\, (\lambda_i)^{-\vartheta}\, e_i(x) \tag{57}$$

for all $\omega \in \tilde{\Omega}$, $x \in (0,1)$, $K \in \mathbb{N}$ and all $\vartheta \in (0, \frac{1}{2})$. It will be essential to estimate $\tilde{\mathbb{E}}\big[|\chi^{K,\vartheta}(x)|^2\big]$ and $\tilde{\mathbb{E}}\big[|\chi^{K,\vartheta}(x) - \chi^{K,\vartheta}(y)|^2\big]$ for $x, y \in (0,1)$, $K \in \mathbb{N}$ and $\vartheta \in (0, \frac{1}{2})$ in order to check (24). (Here $\tilde{\mathbb{E}}[Z] := \int_{\tilde{\Omega}} Z(\tilde{\omega})\, \tilde{\mathbb{P}}(\tilde{\omega}) \in [0, \infty]$ for every $\tilde{\mathcal{F}}/\mathcal{B}([0,\infty))$-measurable mapping $Z \colon \tilde{\Omega} \to [0, \infty)$.) To this end note that

$$\tilde{\mathbb{E}}\Big[\big|\chi^{K,\vartheta}(x)\big|^2\Big] = \sum_{i=1}^{K} (\lambda_i)^{-2\vartheta}\, |e_i(x)|^2 \le 2 \sum_{i=1}^{K} \big(\kappa \pi^2 i^2\big)^{-2\vartheta} \le 2\left(1 + \kappa^{-1}\right) \left(\sum_{i=1}^{\infty} i^{-4\vartheta}\right) \tag{58}$$

for all $x \in (0,1)$, $K \in \mathbb{N}$ and all $\vartheta \in (0, \frac{1}{2})$. Moreover, we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\Big[\big|\chi^{K,\vartheta}(x) - \chi^{K,\vartheta}(y)\big|^2\Big] = \sum_{i=1}^{K} (\lambda_i)^{-2\vartheta}\, |e_i(x) - e_i(y)|^2 \\
&\le \sum_{i=1}^{K} (\lambda_i)^{-2\vartheta}\, |e_i(x) - e_i(y)|^{2s} \left(|e_i(x)| + |e_i(y)|\right)^{2(1-s)} \le \sum_{i=1}^{K} (\lambda_i)^{-2\vartheta} \left(2\pi^2 i^2\right)^s 8^{(1-s)}\, |x-y|^{2s} \\
&\le 8 \left(\sum_{i=1}^{K} \left(\kappa \pi^2 i^2\right)^{-2\vartheta} \left(\pi^2 i^2\right)^s\right) |x-y|^{2s} \le \frac{8}{\kappa^{2\vartheta}} \left(\sum_{i=1}^{K} (\pi i)^{(2s - 4\vartheta)}\right) |x-y|^{2s} \\
&\le 3\left(1 + \kappa^{-1}\right) \left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right) |x-y|^{2s}
\end{aligned} \tag{59}$$

for all $x, y \in (0,1)$, $K \in \mathbb{N}$, $\vartheta \in (\frac{s}{2} + \frac{1}{4}, \frac{1}{2})$ and all $s \in (0, \frac{1}{2})$. We also use the notation

$$\|v\|_{W^{r,2}} := \left(\int_0^1 |v(x)|^2\, dx + \int_0^1 \int_0^1 \frac{|v(x) - v(y)|^2}{|x-y|^{(1+2r)}}\, dx\, dy\right)^{\frac{1}{2}} \in [0, \infty]$$

for all $\mathcal{B}((0,1))/\mathcal{B}(\mathbb{R})$-measurable mapping $v : (0,1) \to \mathbb{R}$ and all $r \in (0, \infty)$. Then we obtain

$$\begin{aligned}
\mathbb{E}\Big[\big\|B(v)\chi^{K,\vartheta}\big\|^2_{W^{r,2}((0,1),\mathbb{R})}\Big] &= \int_0^1 \mathbb{E}\Big[\big|b(x, v(x)) \cdot \chi^{K,\vartheta}(x)\big|^2\Big]\, dx \\
&\quad + \int_0^1 \int_0^1 \frac{\mathbb{E}\Big[\big|b(x, v(x)) \cdot \chi^{K,\vartheta}(x) - b(y, v(y)) \cdot \chi^{K,\vartheta}(y)\big|^2\Big]}{|x-y|^{(1+2r)}}\, dx\, dy \\
&\le 2 \int_0^1 |b(x, v(x))|^2\, \mathbb{E}\Big[\big|\chi^{K,\vartheta}(x)\big|^2\Big]\, dx \\
&\quad + 2 \int_0^1 \int_0^1 \frac{|b(x, v(x))|^2\, \mathbb{E}\Big[\big|\chi^{K,\vartheta}(x) - \chi^{K,\vartheta}(y)\big|^2\Big]}{|x-y|^{(1+2r)}}\, dx\, dy \\
&\quad + 2 \int_0^1 \int_0^1 \frac{|b(x, v(x)) - b(y, v(y))|^2\, \mathbb{E}\Big[\big|\chi^{K,\vartheta}(y)\big|^2\Big]}{|x-y|^{(1+2r)}}\, dx\, dy
\end{aligned}$$

and using (58) and (59) shows

$$\begin{aligned}\mathbb{E}\Big[\big\|B(v)\chi^{K,\vartheta}\big\|^2_{W^{r,2}((0,1),\mathbb{R})}\Big] &\leq 4\left(1+\kappa^{-1}\right)\left(\sum_{i=1}^{\infty} i^{-4\vartheta}\right)\|b(\cdot,v)\|^2_{W^{r,2}((0,1),\mathbb{R})}\\
&\quad + 6\left(1+\kappa^{-1}\right)\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\int_0^1\int_0^1 \frac{|b(x,v(x))|^2}{|x-y|^{(1+2r-2s)}}\,dx\,dy\\
&\leq 4\left(1+\kappa^{-1}\right)\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\|b(\cdot,v)\|^2_{W^{r,2}((0,1),\mathbb{R})}\\
&\quad + 12\left(1+\kappa^{-1}\right)\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\|b(\cdot,v)\|^2_{H}\int_0^1 y^{(2s-2r-1)}\,dy\\
&\leq \frac{10\left(1+\kappa^{-1}\right)}{(s-r)}\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\|b(\cdot,v)\|^2_{W^{r,2}((0,1),\mathbb{R})}\end{aligned}$$

for all $v \in H$, $\vartheta \in (\frac{s}{2}+\frac{1}{4},\frac{1}{2})$, $s \in (r,\frac{1}{2})$, $K \in \mathbb{N}$ and all $r \in (0,\frac{1}{2})$. Therefore, inequality (25) in Section 4 in [32] gives

$$\begin{aligned}\left(\sup_{K\in\mathbb{N}}\mathbb{E}\Big[\big\|B(v)\chi^{K,\vartheta}\big\|^2_{W^{r,2}((0,1),\mathbb{R})}\Big]\right)^{\frac{1}{2}} &\leq \frac{4\left(1+\kappa^{-1}\right)}{(s-r)}\sqrt{\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}}\,\|b(\cdot,v)\|_{W^{r,2}((0,1),\mathbb{R})}\\
&\leq \frac{4\left(1+\kappa^{-1}\right)}{(s-r)}\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\frac{3qC_{\frac{r}{2}}}{(1-r)}\left(1+\|v\|_{H_{\frac{r}{2}}}\right)\\
&\leq \frac{12C_{\frac{r}{2}}q\left(1+\kappa^{-1}\right)}{(s-r)^2}\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\left(1+\|v\|_{H_{\frac{r}{2}}}\right)<\infty\end{aligned}\tag{60}$$

for all $\vartheta \in (\frac{s}{2}+\frac{1}{4},\frac{1}{2})$, $s \in (r,\frac{1}{2})$, $v \in H_{\frac{r}{2}}$ and all $r \in (0,\frac{1}{2})$. Moreover, we have

$$\begin{aligned}&\left\|(-A)^{-\vartheta}B(v)Q^{-\alpha}\right\|_{HS(U_0,H)} = \left\|(-A)^{-\vartheta}B(v)Q^{(\frac{1}{2}-\alpha)}\right\|_{HS(H)} = \left\|Q^{(\frac{1}{2}-\alpha)}B(v)(-A)^{-\vartheta}\right\|_{HS(H)}\\
&= \left(\kappa\pi^2\right)^{(\frac{1}{2}-\alpha)}\left\|\left(\frac{Q}{\kappa\pi^2}\right)^{(\frac{1}{2}-\alpha)}B(v)(-A)^{-\vartheta}\right\|_{HS(H)} = \left(\kappa\pi^2\right)^{(\frac{1}{2}-\alpha)}\left\|B(v)(-A)^{-\vartheta}\right\|_{HS(H,H_{(\alpha-1/2)})}\end{aligned}$$

and using inequality (20) in Section 4 in [32] and estimate (60) in this article then yields

$$\begin{aligned}\left\|(-A)^{-\vartheta}B(v)Q^{-\alpha}\right\|_{HS(U_0,H)} &= \left(\kappa\pi^2\right)^{(\frac{1}{2}-\alpha)}\left(\sup_{K\in\mathbb{N}}\mathbb{E}\Big[\big\|B(v)\chi^{K,\vartheta}\big\|^2_{H_{(\alpha-1/2)}}\Big]\right)^{\frac{1}{2}}\\
&\leq C_{(\alpha-\frac{1}{2})}\left(1+\kappa^{-1}\right)\left(\sup_{K\in\mathbb{N}}\mathbb{E}\,\big\|B(v)\chi^{K,\vartheta}\big\|^2_{W^{2\alpha-1,2}((0,1),\mathbb{R})}\right)^{\frac{1}{2}}\\
&\leq \frac{12C^2_{(\alpha-\frac{1}{2})}q\left(1+\kappa^{-1}\right)^2}{(s+1-2\alpha)^2}\left(\sum_{i=1}^{\infty} i^{(2s-4\vartheta)}\right)\left(1+\|v\|_{H_{(\alpha-\frac{1}{2})}}\right)<\infty\end{aligned}$$

for all $\vartheta \in (\frac{s}{2}+\frac{1}{4},\frac{1}{2})$, $s \in (2\alpha-1,\frac{1}{2})$, $v \in H_{(\alpha-\frac{1}{2})}$ and all $\alpha \in (\frac{1}{2},\frac{3}{4})$. Therefore, estimate (24) is satisfied for all $\alpha \in (0,\frac{3}{4})$ and all $\gamma \in (\frac{1}{2},\frac{3}{4})$. This finally shows that Assumptions 1-4 are fullfilled for the SPDE (56) for all $\alpha \in (0,\frac{3}{4})$, $\beta = \frac{1}{5}$ and all $\gamma \in (\frac{1}{2},\frac{3}{4})$.

Theorem 1 therefore yields the existence of real numbers $C_r \in (0,\infty)$, $r \in (0,\frac{3}{4})$, such that

$$\left(\mathbb{E}\left[\int_0^1\left|X_T(x)-Y_M^{N,M,K}(x)\right|^2dx\right]\right)^{\frac{1}{2}} \leq C_r\left(N^{(r-\frac{3}{2})}+K^{(r-\frac{3}{2})}+M^{(r-\frac{3}{4})}\right)\tag{61}$$

for all $N,M,K \in \mathbb{N}$ and all arbitrarily small $r \in (0,\frac{3}{4})$. In order to balance the error terms on the right hand side of (61) we choose $N^2 = K^2 = M$ in (61) and obtain the existence of real numbers $C_r \in (0,\infty)$, $r \in (0,\frac{3}{2})$, such that

$$\left(\mathbb{E}\left[\int_0^1\left|X_T(x)-Y_{N^2}^{N,N^2,N}(x)\right|^2dx\right]\right)^{\frac{1}{2}} \leq C_r\cdot N^{(r-\frac{3}{2})}\tag{62}$$

for all $N \in \mathbb{N}$ and all arbitrarily small $r \in (0,\frac{3}{2})$. Estimate (62) is nothing else than inequality (18) in the introduction. We also refer to Figure 1 in the introduction for a numerical result illustrating (62).

4.2 A one-dimensional stochastic reaction diffusion equation with $AQ \neq QA$

In Subsection 4.1 we assumed that the eigenfunctions of the dominating linear operator A and of the covariance operator Q of the driving Wiener process $W\colon [0,T] \times \Omega \to H$ of the SPDE (52) coincide and in particular, we assumed in Subsection 4.1 that

$$AQv = QAv \tag{63}$$

holds for all $v \in D(A)$. However, our general setting in Section 2 does not need condition (63) to be fulfilled. To illustrate this fact we consider in this subsection an example in which (63) fails to hold. More precisely, in this subsection let $d = 1$, $T = 1$, $\kappa = \frac{1}{20}$, let $x_0\colon [0,1] \to \mathbb{R}$ be given by $x_0(x) = 0$ for all $x \in [0,1]$, let $f, b\colon (0,1) \times \mathbb{R} \to \mathbb{R}$ be given by $f(x,y) = 1 - y$ and $b(x,y) = \frac{y}{1+y^2}$ for all $x \in (0,1)$, $y \in \mathbb{R}$, let $\mathcal{J} = \{0,1,2,\ldots\}$, let $\mathcal{J}_K = \{0,1,\ldots,K\}$ for all $K \in \mathbb{N}$, let $\eta_0 = 0$, $\eta_j = \frac{1}{j^3}$ and let $g_j\colon (0,1) \to \mathbb{R}$ be given by $g_0(x) = 1$, $g_j(x) = \sqrt{2}\cos(j\pi x)$ for all $x \in (0,1)$ and all $j \in \mathbb{N}$. The SPDE (52) thus reduces to

$$dX_t(x) = \left[\frac{1}{20}\frac{\partial^2}{\partial x^2} X_t(x) + 1 - X_t(x)\right] dt + \frac{X_t(x)}{1 + X_t(x)^2}\, dW_t(x) \tag{64}$$

with $X_t(0) = X_t(1) = 0$ and $X_0(x) = 0$ for $x \in (0,1)$ and $t \in [0,1]$. Of course, the SPDE (64) can also be written as

$$dX_t(x) = \left[\frac{1}{20}\frac{\partial^2}{\partial x^2} X_t(x) + 1 - X_t(x)\right] dt + \sum_{j=1}^{\infty} \frac{X_t(x)}{1 + X_t(x)^2} \frac{\sqrt{2}}{j^{1.5}} \cos(j\pi x)\, d\beta_t^j$$

with $X_t(0) = X_t(1) = 0$ and $X_0(x) = 0$ for $x \in (0,1)$ and $t \in [0,1]$. Estimate (22) is here fulfilled for all $\delta \in (0, \frac{1}{2})$ due to Subsection 4.2 in [32]. Moreover, as in Subsection 4.1 it can be shown that inequality (24) holds for all $\alpha \in (0, \frac{2}{3})$ and all $\gamma \in (\frac{1}{2}, 1)$. This finally shows that Assumptions 1-4 are fullfilled for the SPDE (64) for all $\alpha \in (0, \frac{2}{3})$, $\beta = \frac{1}{5}$ and all $\gamma \in (\frac{1}{2}, 1)$.

Theorem 1 therefore yields the existence of real numbers $C_r \in (0,\infty)$, $r \in (0,1)$, such that

$$\left(\mathbb{E}\left[\int_0^1 \left|X_T(x) - Y_M^{N,M,K}(x)\right|^2 dx\right]\right)^{\frac{1}{2}} \leq C_r \left(N^{(r-2)} + K^{(r-2)} + M^{(r-1)}\right) \tag{65}$$

for all $N, M, K \in \mathbb{N}$ and all arbitrarily small $r \in (0,1)$. Choosing $N^2 = K^2 = M$ in (65) hence gives the existence of real numbers $C_r \in (0,\infty)$, $r \in (0,2)$, such that

$$\left(\mathbb{E}\left[\int_0^1 \left|X_T(x) - Y_{N^2}^{N,N^2,N}(x)\right|^2 dx\right]\right)^{\frac{1}{2}} \leq C_r \cdot N^{(r-2)} \tag{66}$$

for all $N \in \mathbb{N}$ and all arbitrarily small $r \in (0,2)$. The Milstein type approximation $Y_{N^2}^{N,N^2,N}$ thus converges in the root mean square sense to X_T with order $2-$ as N goes to infinity. Since $P_N(H) \subset H$ is N-dimensional and since N^2 time steps are used to simulate $Y_{N^2}^{N,N^2,N}$, $O(N^3 \log(N))$ computational operations and random variables are needed to simulate $Y_{N^2}^{N,N^2,N}$ here. Combining the computational effort $O(N^3 \log(N))$ and the convergence order $2-$ in (66) shows that the Milstein type algorithm (54) with $N^2 = K^2 = M$ needs about $O(\varepsilon^{-\frac{3}{2}})$ computational operations and random variables to achieve a root mean square precision $\varepsilon > 0$.

The linear implicit Euler scheme combined with spectral Galerkin methods which we denote by $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $Z_n^N\colon \Omega \to H$, $n \in \{0,1,\ldots,N^4\}$, $N \in \mathbb{N}$, is given by $Z_0^N = 0$ and

$$Z_{n+1}^N = P_N\left(I - \frac{T}{N^4}A\right)^{-1}\left(Z_n^N + \frac{T}{N^4} \cdot f(\cdot, Z_n^N) + b(\cdot, Z_n^N) \cdot \Delta W_n^{N^4,N}\right) \tag{67}$$

for all $n \in \{0,1,\ldots,N^4 - 1\}$ and all $N \in \mathbb{N}$ here.

In Figure 4 the root mean square approximation error $\left(\mathbb{E}\left[\|X_T - Z_{N^4}^N\|_H^2\right]\right)^{1/2}$ of the linear implicit Euler approximation $Z_{N^4}^N$ (see (67)) and the root mean square approximation error $\left(\mathbb{E}\left[\|X_T - Y_{N^2}^{N,N^2,N}\|_H^2\right]\right)^{1/2}$ of the Milstein type approximation $Y_{N^2}^{N,N^2,N}$ (see (33) and (54)) is plotted against the precise number of independent standard normal random variables needed to compute the corresponding approximation for $N \in \{4,8,16,32\}$.

4.3 A two-dimensional stochastic heat equation and splitting-up approximations

In this subsection the Milstein type algorithm (33) is compared and related to certain splitting-up type approximations in the case of a two-dimensional linear stochastic heat equation with multiplicative noise. More formally, in this subsection let $d = 2$, $T = 1$, $\kappa = \frac{1}{50}$, let $x_0\colon [0,1]^2 \to \mathbb{R}$ be given by $x_0(x_1,x_2) = 2\sin(\pi x_1)\sin(\pi x_2)$ for all $x_1, x_2 \in [0,1]$, let $f, b\colon (0,1)^2 \times \mathbb{R} \to \mathbb{R}$ be given by $f(x_1,x_2,y) = 0$ and $b(x_1,x_2,y) =$

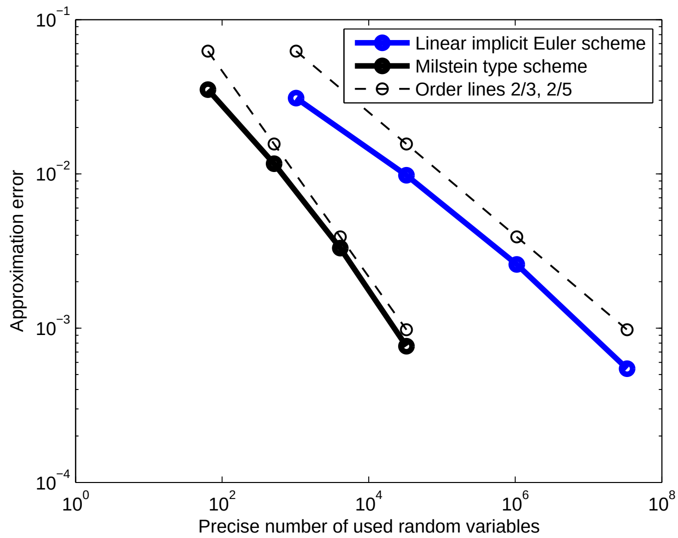

Figure 4: SPDE (64): Root mean square approximation error $\big(\mathbb{E}\big[\|X_T - Z^N_{N^4}\|^2_H\big]\big)^{1/2}$ of the linear implicit Euler approximation $Z^N_{N^4}$ (see (67)) and root mean square approximation error $\big(\mathbb{E}\big[\|X_T - Y^{N,N^2,N}_{N^2}\|^2_H\big]\big)^{1/2}$ of the Milstein type approximation $Y^{N,N^2,N}_{N^2}$ (see (33) and (54)) against the precise number of independent standard normal random variables needed to compute the corresponding approximation for $N \in \{4, 8, 16, 32\}$.

y for all $x_1, x_2 \in (0,1)$, $y \in \mathbb{R}$, let $\mathcal{J} = \mathbb{N}^2$, let $\mathcal{J}_K = \{1, 2, \dots, K\}^2$ for all $K \in \mathbb{N}$, let $\eta_{(j_1,j_2)} = (j_1 + j_2)^{-4}$ and let $g_{(j_1,j_2)} = e_{(j_1,j_2)}$ for all $j_1, j_2 \in \mathbb{N}$. The SPDE (52) thus reduces to

$$dX_t(x_1,x_2) = \left[\frac{1}{50}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) X_t(x_1,x_2)\right] dt + X_t(x_1,x_2)\, dW_t(x_1,x_2) \tag{68}$$

with $X_t|_{\partial(0,1)^2} \equiv 0$ and $X_0(x_1,x_2) = 2\sin(\pi x_1)\sin(\pi x_2)$ for $x_1, x_2 \in (0,1)$ and $t \in [0,1]$. In view of (53) the SPDE (68) can also be written as

$$dX_t(x_1,x_2) = \left[\frac{1}{50}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) X_t(x_1,x_2)\right] dt + \sum_{j_1,j_2=1}^{\infty} \frac{X_t(x_1,x_2)}{(j_1+j_2)^2} 2\sin(j_1\pi x_1)\sin(j_2\pi x_2)\, d\beta_t^{(j_1,j_2)} \tag{69}$$

with $X_t|_{\partial(0,1)^2} \equiv 0$ and $X_0(x_1,x_2) = 2\sin(\pi x_1)\sin(\pi x_2)$ for $x_1, x_2 \in (0,1)$ and $t \in [0,1]$.

Due to Subsection 4.3 in [32] inequality (22) holds for all $\delta \in (0, \frac{1}{2})$ here. In order to verify (24) the notation

$$\|v\|_{L^\infty((0,1)^2,\mathbb{R})} := \inf\big\{R \in [0,\infty] \colon \lambda\big(\big\{x \in (0,1)^2 \colon v(x) > R\big\}\big) = 0\big\} \in [0,\infty]$$

is used for all $\mathcal{B}((0,1)^2)/\mathcal{B}(\mathbb{R})$-measurable mappings $v\colon (0,1)^2 \to \mathbb{R}$ in this subsection. Then

$$\begin{aligned} \|v\|_{L^\infty((0,1)^2,\mathbb{R})} &\leq \sum_{i\in\mathbb{N}^2} |\langle e_i, v\rangle_H|\, \|e_i\|_{C((0,1)^2,\mathbb{R})} \\ &\leq 2\left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{-2r}\right)^{\frac{1}{2}} \left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{2r} |\langle e_i, v\rangle_H|^2\right)^{\frac{1}{2}} = 2\left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{-2r}\right)^{\frac{1}{2}} \|v\|_{H_r} \end{aligned} \tag{70}$$

for all $v \in H_r$ and all $r \in (\frac{1}{2}, \infty)$. Moreover, we have

$$\begin{aligned} \left\|(-A)^{-\vartheta} B(v) Q^{-\alpha}\right\|_{HS(U_0,H)} &= \left\|(-A)^{-\vartheta} B(v) Q^{(\frac{1}{2}-\alpha)}\right\|_{HS(H)} = \left\|(-A)^{-\vartheta} B(v) Q^{(\frac{1}{2}-\alpha)}\right\|_{S_2(H)} \\ &\leq \left\|(-A)^{-\vartheta}\right\|_{S_{\frac{1}{\alpha}}(H)} \|B(v)\|_{L(H)} \left\|Q^{(\frac{1}{2}-\alpha)}\right\|_{S_{\frac{2}{(1-2\alpha)}}(H)} \\ &\leq \left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{-\frac{\vartheta}{\alpha}}\right)^{\alpha} \|b(\cdot,v)\|_{L^\infty((0,1)^2,\mathbb{R})} \left(\sum_{j\in\mathcal{J}} \eta_j^{(\frac{1}{2}-\alpha)\frac{2}{(1-2\alpha)}}\right)^{\frac{(1-2\alpha)}{2}} \end{aligned}$$

and using (70) shows

$$\begin{aligned}
\left\|(-A)^{-\vartheta}\, B(v) Q^{-\alpha}\right\|_{HS(U_0,H)} &\leq \left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{-\frac{\vartheta}{\alpha}}\right)^{\alpha} \left(q\,\|v\|_{L^\infty((0,1)^2,\mathbb{R})} + q\right) (\operatorname{Tr}(Q))^{\left(\frac{1}{2}-\alpha\right)} \\
&\leq q\,(1+\operatorname{Tr}(Q)) \left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{-\frac{\vartheta}{\alpha}}\right)^{\alpha} \left(1 + \|v\|_{L^\infty((0,1)^2,\mathbb{R})}\right) \\
&\leq 2q\,(1+\operatorname{Tr}(Q)) \left(\sum_{i\in\mathbb{N}^2} (\lambda_i)^{-\frac{\vartheta}{\alpha}}\right)^{\alpha} \left(1 + \sum_{i\in\mathbb{N}^2} (\lambda_i)^{-2\gamma}\right) \left(1 + \|v\|_{H_\gamma}\right) < \infty
\end{aligned}$$

for all $\vartheta \in (\alpha, \frac{1}{2})$, $\alpha \in (0, \frac{1}{2})$, $v \in H_\gamma$ and all $\gamma \in (\frac{1}{2}, 1)$. Inequality (24) thus holds for all $\alpha \in (0, \frac{1}{2})$ and all $\gamma \in (\frac{1}{2}, 1)$ here. This finally shows that Assumptions 1-4 are fulfilled for the SPDE (68) for all $\alpha \in (0, \frac{1}{2})$, $\beta = \frac{2}{5}$ and all $\gamma \in (\frac{1}{2}, 1)$.

Theorem 1 therefore yields the existence of real numbers $C_r \in (0,\infty)$, $r \in (0,1)$, such that

$$\left(\mathbb{E}\left[\int_0^1 \int_0^1 \left|X_T(x_1,x_2) - Y_M^{N,M,K}(x_1,x_2)\right|^2 dx_1\, dx_2\right]\right)^{1/2} \leq C_r \left(N^{(r-2)} + K^{(r-2)} + M^{(r-1)}\right) \tag{71}$$

holds for all $N, M, K \in \mathbb{N}$ and all arbitrarily small $r \in (0,1)$. In order to balance the error terms on the right hand side of (71) we choose $M = N^2 = K^2$ in (71) and obtain the existence of real numbers $C_r \in (0,\infty)$, $r \in (0,2)$, such that

$$\left(\mathbb{E}\left[\int_0^1 \int_0^1 \left|X_T(x_1,x_2) - Y_{N^2}^{N,N^2,N}(x_1,x_2)\right|^2 dx_1\, dx_2\right]\right)^{1/2} \leq C_r \cdot N^{(r-2)} \tag{72}$$

holds for all $N \in \mathbb{N}$ and all arbitrarily small $r \in (0,2)$. The approximation $Y_{N^2}^{N,N^2,N}$ thus converges in the root mean square sense to X_T with order $2-$ as N goes to infinity. The numerical approximations $Y_n^{N,N^2,N}\colon \Omega \to H$, $n \in \{0,1,\dots,N^2\}$, $N \in \mathbb{N}$, (see (33) and (54)) are here given by $Y_0^{N,N^2,N} = x_0$ and

$$Y_{n+1}^{N,N^2,N} = P_N\, e^{A\frac{T}{N^2}} \left(\left[1 + \Delta W_n^{N^2,N} + \frac{1}{2}\left(\Delta W_n^{N^2,N}\right)^2 - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K} \eta_j (g_j)^2\right] \cdot Y_n^{N,N^2,N}\right) \tag{73}$$

for all $n \in \{0,1,\dots,N^2-1\}$ and all $N \in \mathbb{N}$. Since $P_N(H) \subset H$ is N^2-dimensional here and since N^2 time steps are used to simulate $Y_{N^2}^{N,N^2,N}$, $O(N^4\log(N))$ computational operations and random variables are needed to simulate $Y_{N^2}^{N,N^2,N}$. Combining the computational effort $O(N^4\log(N))$ and the convergence order $2-$ in (72) shows that the algorithm (73) in this article needs about $O(\varepsilon^{-2})$ computational operations and random variables to achieve a root mean square precision $\varepsilon > 0$.

The linear implicit Euler scheme combined with spectral Galerkin methods which we denote by $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $Z_n^N\colon \Omega \to H$, $n \in \{0,1,\dots,N^4\}$, $N \in \mathbb{N}$, is given by $Z_0^N = x_0$ and

$$Z_{n+1}^N = P_N\left(I - \frac{T}{N^4}A\right)^{-1}\left(\left[1 + \Delta W_n^{N^4,N}\right] \cdot Z_n^N\right) \tag{74}$$

for all $n \in \{0,1,\dots,N^4-1\}$ and all $N \in \mathbb{N}$ here.

Moreover, since the SPDE (68) is linear here, the splitting-up method can be used in order to solve (68) approximatively. The idea of the splitting-up approach is to split the SPDE (68) into the explicit solvable subequations

$$d\tilde{X}_t(x_1,x_2) = \left[\frac{1}{50}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right)\tilde{X}_t(x_1,x_2)\right] dt, \qquad \tilde{X}_t|_{\partial(0,1)^2} \equiv 0 \tag{75}$$

and

$$d\tilde{\tilde{X}}_t(x_1,x_2) = \tilde{\tilde{X}}_t(x_1,x_2)\, dW_t(x_1,x_2) \tag{76}$$

for $t \in [0,1]$ and $x_1, x_2 \in (0,1)$. For the solution processes $\tilde{X}, \tilde{\tilde{X}}\colon [0,T]\times\Omega \to H$ of (75) and (76) we obtain $\tilde{X}_t = e^{At}\tilde{X}_0$ and $\tilde{\tilde{X}}_t = e^{\left(W_t - \frac{t}{2}\sum_{j\in\mathcal{J}}\eta_j(g_j)^2\right)} \cdot \tilde{\tilde{X}}_0$ $\mathbb{P}$-a.s. for all $t \in [0,1]$. This suggests the splitting-up approximation

$$X_t \approx e^{At}\left(e^{\left(W_t - \frac{t}{2}\sum_{j\in\mathcal{J}}\eta_j(g_j)^2\right)} \cdot X_0\right)$$

for $t \in [0,1]$ where $X\colon [0,T] \times \Omega \to H$ is the solution process of the SPDE (68). The resulting splitting-up method which we denote by $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $\tilde{Z}_n^N\colon \Omega \to H$, $n \in \{0,1,\dots,N^2\}$, $N \in \mathbb{N}$, is then given by $\tilde{Z}_0^N = x_0$ and

$$\tilde{Z}_{n+1}^N = P_N\, e^{A\frac{T}{N^2}} \left(e^{\left(\Delta W_n^{N^2,N} - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K}\eta_j(g_j)^2\right)} \cdot \tilde{Z}_n^N \right) \tag{77}$$

for all $n \in \{0,1,\dots,N^2-1\}$ and all $N \in \mathbb{N}$. Using the Talyor approximation $e^x \approx 1 + x + \frac{x^2}{2}$ for all $x \in \mathbb{R}$ then yields

$$\begin{aligned}
& e^{\left(\Delta W_n^{N^2,N} - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K}\eta_j(g_j)^2\right)} \\
& \approx 1 + \Delta W_n^{N^2,N} - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K}\eta_j(g_j)^2 + \frac{1}{2}\left(\Delta W_n^{N^2,N} - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K}\eta_j(g_j)^2\right)^2 \\
& \approx 1 + \Delta W_n^{N^2,N} + \frac{1}{2}\left(\Delta W_n^{N^2,N}\right)^2 - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K}\eta_j(g_j)^2
\end{aligned} \tag{78}$$

for all $n \in \{0,1,\dots,N^2-1\}$ and all $N \in \mathbb{N}$. Using approximation (78) in (77) finally shows

$$\tilde{Z}_{n+1}^N \approx P_N\, e^{A\frac{T}{N^2}} \left(\left[1 + \Delta W_n^{N^2,N} + \frac{1}{2}\left(\Delta W_n^{N^2,N}\right)^2 - \frac{T}{2N^2}\sum_{j\in\mathcal{J}_K}\eta_j(g_j)^2\right] \cdot \tilde{Z}_n^N \right)$$

for all $n \in \{0,1,\dots,N^2-1\}$ and all $N \in \mathbb{N}$ which is nothing else than the recursion for $Y_n^{N,N^2,N}$, $n \in \{0,1,\dots,N^2\}$, $N \in \mathbb{N}$, in (73). So, in the case of the linear SPDE (68) an alternative way for deriving the Milstein type algorithm (73) is to apply an appropriate Taylor approximation for the exponential function (see (78) for details) to the splitting-up approximation (77). More results on splitting-up methods can be found in A. Bensoussan, R. Glowinski and A. Rascanu [6, 7], P. Florchinger and F. Le Gland [13], I. Gyöngy and N. Krylov [19, 20, 21] and K. Ito and B. L. Rozovskii [27] and the references therein.

In Figure 5 the root mean square approximation error $\left(\mathbb{E}\left[\|X_T - Z_{N^4}^N\|_H^2\right]\right)^{1/2}$ of the linear implicit Euler approximation $Z_{N^4}^N$ (see (74)), the root mean square approximation error $\left(\mathbb{E}\left[\|X_T - Y_{N^2}^{N,N^2,N}\|_H^2\right]\right)^{1/2}$ of the Milstein type approximation $Y_{N^2}^{N,N^2,N}$ (see (73)) and the root mean square approximation error $\left(\mathbb{E}\left[\|X_T - \tilde{Z}_{N^2}^N\|_H^2\right]\right)^{1/2}$ of the splitting-up approximation $\tilde{Z}_{N^2}^N$ (see (77)) is plotted against the precise number of independent standard normal random variables needed to compute the corresponding approximation for $N \in \{2,4,8,16,32\}$: It turns out that $Z_{32^4}^{32}$ ($32^6 = 1\,073\,741\,824$ random variables) in the case of the linear implicit Euler scheme (74), that $Y_{32}^{32,32^2,32}$ ($32^4 = 1\,048\,576$ random variables) in the case of the algorithm (73) and that $\tilde{Z}_{32^2}^{32}$ ($32^4 = 1\,048\,576$ random variables) in the case of the splitting-up method (77) achieve a root mean square precision $\varepsilon = \frac{1}{1000}$ for the SPDE (68).

5 Proof of Theorem 1

Throughout this section the notation

$$\|Z\|_{L^p(\Omega;E)} := \left(\mathbb{E}\left[\|Z\|_E^p\right]\right)^{1/p} \in [0,\infty] \tag{79}$$

is used for an $\mathbb{R}$-Banach space $(E, \|\cdot\|_E)$, an $\mathcal{F}/\mathcal{B}(E)$-measurable mapping $Z\colon \Omega \to E$ and a real number $p \in [1,\infty)$. We also use the following simple lemma (see, e.g., Theorem 37.5 in [57]).

Lemma 1. *Let Assumptions 1-4 in Section 2 be fulfilled. Then*

$$\left\|(-tA)^r e^{At}\right\|_{L(H)} \leq 1 \qquad \text{and} \qquad \left\|(-tA)^{-r}(e^{At} - I)\right\|_{L(H)} \leq 1 \tag{80}$$

for all $t \in (0,\infty)$ and all $r \in [0,1]$.

We now prove Theorem 1. First of all, note that the exact solution of the SPDE (28) satisfies

$$\begin{aligned}
X_{mh} &= e^{Amh}\xi + \int_0^{mh} e^{A(mh-s)} F(X_s)\, ds + \int_0^{mh} e^{A(mh-s)} B(X_s)\, dW_s \\
&= e^{Amh}\xi + \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(mh-s)} F(X_s)\, ds + \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(mh-s)} B(X_s)\, dW_s
\end{aligned} \tag{81}$$

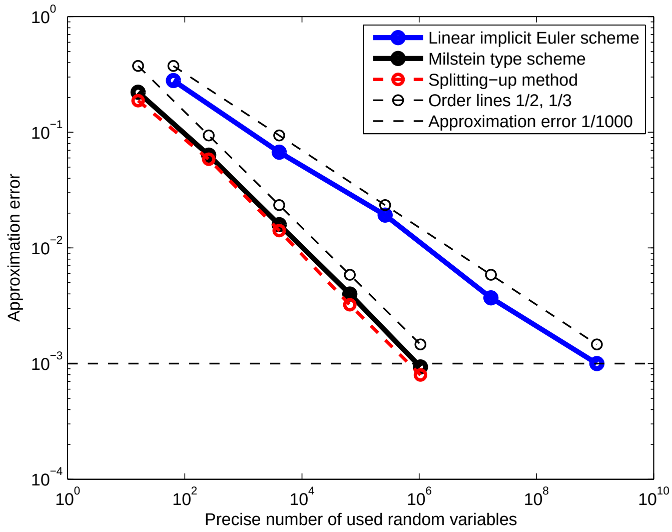

Figure 5: SPDE (68): Root mean square approximation error $\big(\mathbb{E}\big[\|X_T - Z^N_{N^4}\|_H^2\big]\big)^{1/2}$ of the linear implicit Euler approximation $Z^N_{N^4}$ (see (74)), root mean square approximation error of the linear implicit Euler approximation $Z^N_{N^4}$ (see (74)), the root mean square approximation error $\big(\mathbb{E}\big[\|X_T - Y^{N,N^2,N}_{N^2}\|_H^2\big]\big)^{1/2}$ of the Milstein type approximation $Y^{N,N^2,N}_{N^2}$ (see (73)) and root mean square approximation error $\big(\mathbb{E}\big[\|X_T - \tilde{Z}^N_{N^2}\|_H^2\big]\big)^{1/2}$ of the splitting-up approximation $\tilde{Z}^N_{N^2}$ (see (77)) against the precise number of independent standard normal random variables needed to compute the corresponding approximation for $N \in \{2, 4, 8, 16, 32\}$.

$\mathbb{P}$-a.s. for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. Here and below h is the time stepsize $h = h_M = \frac{T}{M}$ with $M \in \mathbb{N}$. In particular, (81) shows

$$P_N(X_{mh}) = e^{Amh} P_N(\xi) + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(mh-s)} F(X_s)\, ds\right) + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(mh-s)} B(X_s)\, dW_s\right) \tag{82}$$

$\mathbb{P}$-a.s. for all $m \in \{0, 1, \ldots, M\}$ and all $N, M \in \mathbb{N}$. In order to estimate the difference of the exact solution (81) and the numerical solution (33) we rewrite the numerical method (33) in some sense. More precisely, the identity

$$\begin{aligned} &\frac{1}{2} B'\big(Y^{N,M,K}_m\big)\Big(B\big(Y^{N,M,K}_m\big)\Delta W^{M,K}_m\Big)\Delta W^{M,K}_m - \frac{T}{2M} \sum_{\substack{j \in \mathcal{J}_K \\ \eta_j \neq 0}} \eta_j B'\big(Y^{N,M,K}_m\big)\Big(B\big(Y^{N,M,K}_m\big)\, g_j\Big) g_j \\ &= \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\big(Y^{N,M,K}_m\big)\left(\int_{\frac{mT}{M}}^{s} B\big(Y^{N,M,K}_m\big)\, dW^K_u\right) dW^K_s \end{aligned} \tag{83}$$

$\mathbb{P}$-a.s. holds for all $m \in \{0, 1, \ldots, M-1\}$ and all $N, M, K \in \mathbb{N}$. The proof of (83) can be found in Subsection 5.7. Using (83) shows that the numerical solution (33) fulfills

$$\begin{aligned} Y^{N,M,K}_{m+1} = P_N\, e^{A\frac{T}{M}} \Bigg(& Y^{N,M,K}_m + \frac{T}{M} \cdot F\big(Y^{N,M,K}_m\big) + \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B\big(Y^{N,M,K}_m\big)\, dW^K_s \\ &+ \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\big(Y^{N,M,K}_m\big)\left(\int_{\frac{mT}{M}}^{s} B\big(Y^{N,M,K}_m\big)\, dW^K_u\right) dW^K_s \Bigg) \end{aligned} \tag{84}$$

$\mathbb{P}$-a.s. for all $m \in \{0,1,\dots,M-1\}$ and all $N,M,K \in \mathbb{N}$. Therefore, the numerical solution (33) satisfies

$$
\begin{aligned}
Y_m^{N,M,K} &= e^{Amh} P_N(\xi) + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} F\Big(Y_l^{N,M,K}\Big)\, ds\right) \\
&\quad + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} B\Big(Y_l^{N,M,K}\Big)\, dW_s^K\right) \\
&\quad + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} B'\Big(Y_l^{N,M,K}\Big)\left(\int_{lh}^{s} B\Big(Y_l^{N,M,K}\Big)\, dW_u^K\right) dW_s^K\right)
\end{aligned} \tag{85}
$$

$\mathbb{P}$-a.s. for all $m \in \{0,1,\dots,M\}$ and all $N,M,K \in \mathbb{N}$. In order to estimate $\mathbb{E}\big[\|X_{mh} - Y_m^{N,M,K}\|_H^2\big]$ for $m \in \{0,1,\dots,M\}$ and $N,M,K \in \mathbb{N}$, we define the $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $Z_m^{N,M,K} \colon \Omega \to H$, $m \in \{0,1,\dots,M\}$, $N,M,K \in \mathbb{N}$, by

$$
\begin{aligned}
Z_m^{N,M,K} &:= e^{Amh} P_N(\xi) + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} F(X_{lh})\, ds\right) \\
&\quad + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} B(X_{lh})\, dW_s^K\right) \\
&\quad + P_N\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} B'(X_{lh})\left(\int_{lh}^{s} B(X_{lh})\, dW_u^K\right) dW_s^K\right)
\end{aligned} \tag{86}
$$

$\mathbb{P}$-a.s. for all $m \in \{0,1,\dots,M\}$ and all $N,M,K \in \mathbb{N}$. The inequality

$$
(a_1 + \ldots + a_n)^2 \leq n\left((a_1)^2 + \ldots + (a_n)^2\right) \tag{87}
$$

for all $a_1,\dots,a_n \in \mathbb{R}$ and all $n \in \mathbb{N}$ then shows

$$
\begin{aligned}
&\mathbb{E}\big[\|X_{mh} - Y_m^{N,M,K}\|_H^2\big] \\
&\leq 3 \cdot \mathbb{E}\big[\|X_{mh} - P_N(X_{mh})\|_H^2\big] + 3 \cdot \mathbb{E}\big[\|P_N(X_{mh}) - Z_m^{N,M,K}\|_H^2\big] + 3 \cdot \mathbb{E}\big[\|Z_m^{N,M,K} - Y_m^{N,M,K}\|_H^2\big]
\end{aligned} \tag{88}
$$

for all $m \in \{0,1,\dots,M\}$ and all $N,M,K \in \mathbb{N}$. In order to estimate the expressions $\mathbb{E}\big[\|X_{mh} - P_N(X_{mh})\|_H^2\big]$, $\mathbb{E}\big[\|P_N(X_{mh}) - Z_m^{N,M,K}\|_H^2\big]$ and $\mathbb{E}\big[\|Z_m^{N,M,K} - Y_m^{N,M,K}\|_H^2\big]$ for $m \in \{0,1,\dots,M\}$ and $N,M,K \in \mathbb{N}$, a real number $R \in (0,\infty)$ satisfying

$$
\begin{gathered}
\mathbb{E}\big[\|B(X_t)\|_{HS(U_0,H_\delta)}^2\big] \leq R, \qquad \|F'(v)\|_{L(H)} \leq R, \qquad \|F''(v)\|_{L^{(2)}(H_\beta,H)} \leq R, \\
\mathbb{E}\big[\|F(X_t)\|_H^2\big] \leq R, \;\; \|B'(v)\|_{L(H,HS(U_0,H))} \leq R, \;\; \|B''(v)\|_{L^{(2)}(H_\beta,HS(U_0,H))} \leq R, \\
\mathbb{E}\big[\|(-A)^\gamma X_t\|_H^2\big] = \mathbb{E}\big[\|X_t\|_{H_\gamma}^2\big] \leq R, \;\; \mathbb{E}\big[\|X_{t_2} - X_{t_1}\|_{H_\beta}^4\big] \leq R\,|t_2 - t_1|^{\min(4(\gamma-\beta),2)}, \\
c + \frac{1}{(1-\gamma)} + \frac{1}{(1-2\vartheta)} + \frac{1}{(1-2\delta)} + T + \left\|A^{-1}\right\|_{L(H)} \leq R
\end{gathered}
$$

for all $v \in H_\beta$ and all $t, t_1, t_2 \in [0,T]$ is used throughout this proof. Due to Assumptions 1-4 in Section 2 and Proposition 1 such a real number indeed exists. For the spatial discretization error $\mathbb{E}\big[\|X_{mh} - P_N(X_{mh})\|_H^2\big]$ we then obtain

$$
\begin{aligned}
\mathbb{E}\Big[\|X_{mh} - P_N(X_{mh})\|_H^2\Big] &= \mathbb{E}\Big[\|(I - P_N)\, X_{mh}\|_H^2\Big] = \mathbb{E}\Big[\Big\|(-A)^{-\gamma}\,(I - P_N)\,(-A)^{\gamma}\, X_{mh}\Big\|_H^2\Big] \\
&\leq \Big\|(-A)^{-\gamma}\,(I - P_N)\Big\|_{L(H)}^2\, \mathbb{E}\Big[\|X_{mh}\|_{H_\gamma}^2\Big] \leq R\,(r_N)^2
\end{aligned} \tag{89}
$$

for all $m \in \{0,1,\dots,M\}$ and all $N,M \in \mathbb{N}$ where here and below the real numbers $(r_N)_{N\in\mathbb{N}} \subset \mathbb{R}$ are given by

$$
r_N := \Big\|(-A)^{-\gamma}\,(I - P_N)\Big\|_{L(H)} = \left(\inf_{i \in \mathcal{I}\setminus\mathcal{I}_N} \lambda_i\right)^{-\gamma} \tag{90}
$$

for all $N \in \mathbb{N}$. The rest of this proof is then divided into six parts. In the first part (see Subsection 5.1) we establish

$$
\mathbb{E}\left[\left\|\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \Big(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh})\Big)\, ds\right\|_H^2\right] \leq \frac{36R^8}{M^{\min(4(\gamma-\beta),2\gamma)}} \tag{91}
$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. We show

$$\mathbb{E}\left[\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(mh-s)} B(X_s)\, d\big(W_s - W_s^K\big) \right\|_H^2 \right] \leq 4R^5 \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} \tag{92}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M, K \in \mathbb{N}$ in the second part (see Subsection 5.2) and

$$\mathbb{E}\left[\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \Big(e^{A(mh-s)} - e^{A(m-l)h}\Big) B(X_s)\, dW_s^K \right\|_H^2 \right] \leq \frac{3R^4}{M^{(1+2\delta)}} \tag{93}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M, K \in \mathbb{N}$ in the third part (see Subsection 5.3). The fourth part (see Subsection 5.4) gives

$$\mathbb{E}\left[\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} \int_0^1 B''(X_{lh} + r\,(X_s - X_{lh}))\,(X_s - X_{lh}, X_s - X_{lh})\,(1-r)\, dr\, dW_s^K \right\|_H^2 \right] \leq \frac{R^6}{M^{\min(4(\gamma-\beta),2)}} \tag{94}$$

and in the fifth part (see Subsection 5.5) we obtain

$$\mathbb{E}\left[\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} B'(X_{lh}) \left(X_s - X_{lh} - \int_{lh}^{s} B(X_{lh})\, dW_u^K \right) dW_s^K \right\|_H^2 \right] \leq \frac{20R^{13}}{M^{\min(4(\gamma-\beta),2\gamma)}} + 20R^{11} \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} \tag{95}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M, K \in \mathbb{N}$. The inequalities (91)-(95) are used below to estimate $\mathbb{E}\big[\|P_N(X_{mh}) - Z_m^{N,M,K}\|_H^2\big]$ for $m \in \{0, 1, \ldots, M\}$ and $N, M, K \in \mathbb{N}$ in (88). In the sixth part (see Subsection 5.6) we estimate

$$\mathbb{E}\Big[\big\|Z_m^{N,M,K} - Y_m^{N,M,K}\big\|_H^2\Big] \leq \frac{9R^4}{M} \left(\sum_{l=0}^{m-1} \mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big] \right) \tag{96}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$ by using the global Lipschitz continuity of the coefficients $F \colon H_\beta \to H$ (see Assumption 2) and $B \colon H_\beta \to HS(U_0, H)$ (see Assumption 3). Combining (88), (89) and (96) then yields

$$\begin{aligned} &\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] \\ &\leq 3R\,(r_N)^2 + 3 \cdot \mathbb{E}\Big[\big\|P_N(X_{mh}) - Z_m^{N,M,K}\big\|_H^2\Big] + \frac{27R^4}{M} \left(\sum_{l=0}^{m-1} \mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big] \right) \end{aligned} \tag{97}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$. Hence, (82), (86), (87) and the fact $\|P_N(v)\|_H \leq \|v\|_H$ for all $v \in H$ show

$$\begin{aligned} \mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] &\leq 9 \cdot \mathbb{E}\left[\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \Big(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh})\Big)\, ds \right\|_H^2 \right] \\ &\quad + 9 \cdot \mathbb{E}\left[\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(mh-s)} B(X_s)\, d\big(W_s - W_s^K\big) \right\|_H^2 \right] \\ &\quad + 9 \cdot \mathbb{E}\Bigg[\Bigg\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \Big(e^{A(mh-s)} B(X_s) - e^{A(m-l)h} B(X_{lh})\Big)\, dW_s^K \\ &\qquad - \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} B'(X_{lh}) \left(\int_{lh}^{s} B(X_{lh})\, dW_u^K \right) dW_s^K \Bigg\|_H^2 \Bigg] \\ &\quad + 3R\,(r_N)^2 + \frac{27R^4}{M} \left(\sum_{l=0}^{m-1} \mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big] \right) \end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$. Therefore, (91) and (92) yield

$$\begin{aligned}\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] &\leq \frac{324R^8}{M^{\min(4(\gamma-\beta),2\gamma)}} + 36R^5\bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\bigg)^{2\alpha} \\ &+ 18\cdot\mathbb{E}\Bigg[\bigg\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\Big(e^{A(mh-s)} - e^{A(m-l)h}\Big)\,B(X_s)\,dW_s^K\bigg\|_H^2\Bigg] \\ &+ 18\cdot\mathbb{E}\Bigg[\bigg\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}\,(B(X_s) - B(X_{lh}))\,dW_s^K \\ &\quad - \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}B'(X_{lh})\left(\int_{lh}^{s}B(X_{lh})\,dW_u^K\right)dW_s^K\bigg\|_H^2\Bigg] \\ &+ 3R\,(r_N)^2 + \frac{27R^4}{M}\left(\sum_{l=0}^{m-1}\mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big]\right)\end{aligned}$$

and (93) shows

$$\begin{aligned}\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] &\leq \frac{324R^8}{M^{\min(4(\gamma-\beta),2\gamma)}} + 36R^5\bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\bigg)^{2\alpha} + \frac{54R^4}{M^{(1+2\delta)}} \\ &+ 18\cdot\mathbb{E}\Bigg[\bigg\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}\,(B(X_s) - B(X_{lh}))\,dW_s^K \\ &\quad - \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}B'(X_{lh})\left(\int_{lh}^{s}B(X_{lh})\,dW_u^K\right)dW_s^K\bigg\|_H^2\Bigg] \\ &+ 3R\,(r_N)^2 + \frac{27R^4}{M}\left(\sum_{l=0}^{m-1}\mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big]\right)\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$. The fact

$$B(X_s) - B(X_{lh}) = B'(X_{lh})(X_s - X_{lh}) + \int_0^1 B''(X_{lh} + r(X_s - X_{lh}))(X_s - X_{lh}, X_s - X_{lh})(1-r)\,dr \quad (98)$$

for all $s \in [lh, (l+1)h]$, $l \in \{0, 1, \ldots, M-1\}$ and all $M \in \mathbb{N}$ then yields

$$\begin{aligned}\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] &\leq \frac{324R^8}{M^{\min(4(\gamma-\beta),2\gamma)}} + 36R^5\bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\bigg)^{2\alpha} + \frac{54R^4}{M^{(1+2\delta)}} \\ &+ 36\cdot\mathbb{E}\Bigg[\bigg\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}B'(X_{lh})\Big(X_s - X_{lh} - \int_{lh}^{s}B(X_{lh})\,dW_u^K\Big)\,dW_s^K\bigg\|_H^2\Bigg] \\ &+ 36\cdot\mathbb{E}\Bigg[\bigg\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}\int_0^1 B''(X_{lh} + r\,(X_s - X_{lh}))\big(X_s - X_{lh}, X_s - X_{lh}\big)\,(1-r)\,dr\,dW_s^K\bigg\|_H^2\Bigg] \\ &+ 3R\,(r_N)^2 + \frac{27R^4}{M}\left(\sum_{l=0}^{m-1}\mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big]\right)\end{aligned}$$

for all $m \in \{0, 1, \ldots, M-1\}$ and all $N, M, K \in \mathbb{N}$. Therefore, (94) and (95) give

$$\begin{aligned}\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] &\leq \frac{324R^8}{M^{\min(4(\gamma-\beta),2\gamma)}} + 756R^{11}\bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\bigg)^{2\alpha} + \frac{54R^4}{M^{(1+2\delta)}} \\ &+ \frac{720R^{13}}{M^{\min(4(\gamma-\beta),2\gamma)}} + \frac{36R^6}{M^{\min(4(\gamma-\beta),2)}} + 3R\,(r_N)^2 + \frac{27R^4}{M}\left(\sum_{l=0}^{m-1}\mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big]\right)\end{aligned}$$

and hence

$$\begin{aligned}\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] &\leq \big(324R^8 + 54R^4 + 720R^{13} + 36R^6\big)\,\frac{1}{M^{\min(4(\gamma-\beta),2\gamma)}} \\ &+ 756R^{11}\bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\bigg)^{2\alpha} + 3R\,(r_N)^2 + \frac{27R^4}{M}\left(\sum_{l=0}^{m-1}\mathbb{E}\Big[\big\|X_{lh} - Y_l^{N,M,K}\big\|_H^2\Big]\right)\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$. Gronwall's lemma thus shows

$$\begin{aligned}
&\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big] \\
&\le e^{27R^4}\left(\frac{\left(324R^8 + 54R^4 + 720R^{13} + 36R^6\right)}{M^{\min(4(\gamma-\beta),2\gamma)}} + 756R^{11}\left(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} + 3R\,(r_N)^2\right) \\
&\le 1134R^{13}e^{27R^4}\left((r_N)^2 + \left(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} + M^{-\min(4(\gamma-\beta),2\gamma)}\right)
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$. Finally, we obtain

$$\begin{aligned}
\left(\mathbb{E}\Big[\big\|X_{mh} - Y_m^{N,M,K}\big\|_H^2\Big]\right)^{\frac{1}{2}} &\le 34R^7e^{14R^4}\left(\left(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\right)^{-\gamma} + \left(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)}\right) \\
&\le e^{20R^4}\left(\left(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\right)^{-\gamma} + \left(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)}\right)
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$.

5.1 Temporal discretization error: Proof of (91)

First of all, we have

$$\begin{aligned}
&\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left(e^{A(mh-s)}F(X_s) - e^{A(m-l)h}F(X_{lh})\right)ds\right\|_{L^2(\Omega;H)} \\
&\le \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left\|\left(e^{A(mh-s)} - e^{A(m-l)h}\right)F(X_s)\right\|_{L^2(\Omega;H)}ds \\
&\quad + \left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}e^{A(m-l)h}\left(F(X_s) - F(X_{lh})\right)ds\right\|_{L^2(\Omega;H)}
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. Using

$$F(X_s) - F(X_{lh}) = F'(X_{lh})\,(X_s - X_{lh}) + \int_0^1 F''(X_{lh} + r\,(X_s - X_{lh}))\,(X_s - X_{lh}, X_s - X_{lh})\,(1-r)\,dr \quad (99)$$

for all $s \in [lh, (l+1)h]$, $l \in \{0, 1, \ldots, M-1\}$ and all $M \in \mathbb{N}$ then shows

$$\begin{aligned}
&\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left(e^{A(mh-s)}F(X_s) - e^{A(m-l)h}F(X_{lh})\right)ds\right\|_{L^2(\Omega;H)} \\
&\le \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left\|e^{A(mh-s)} - e^{A(m-l)h}\right\|_{L(H)}\|F(X_s)\|_{L^2(\Omega;H)}\,ds \\
&+ \left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}e^{A(m-l)h}F'(X_{lh})\,(X_s - X_{lh})\,ds\right\|_{L^2(\Omega;H)} \\
&+ \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\int_0^1\|F''(X_{lh} + r\,(X_s - X_{lh}))\,(X_s - X_{lh}, X_s - X_{lh})\|_{L^2(\Omega;H)}\,(1-r)\,dr\,ds
\end{aligned}$$

and hence

$$\begin{aligned}
&\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left(e^{A(mh-s)}F(X_s) - e^{A(m-l)h}F(X_{lh})\right)ds\right\|_{L^2(\Omega;H)} \\
&\le R\left(2h + \sum_{l=0}^{m-2}\int_{lh}^{(l+1)h}\left\|e^{A(mh-s)} - e^{A(m-l)h}\right\|_{L(H)}ds\right) \\
&\quad + \left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}e^{A(m-l)h}F'(X_{lh})\,(X_s - X_{lh})\,ds\right\|_{L^2(\Omega;H)} \\
&\quad + \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\int_0^1\left\|R\,\|X_s - X_{lh}\|_{H_\beta}^2\right\|_{L^2(\Omega;\mathbb{R})}(1-r)\,dr\,ds
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. Therefore, we obtain

$$\begin{aligned}
&\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)} \\
&\leq R \left(2h + \sum_{l=0}^{m-2} \int_{lh}^{(l+1)h} \left\| A e^{A(mh-s)} \right\|_{L(H)} \left\| A^{-1} \left(e^{A(s-lh)} - I \right) \right\|_{L(H)} ds \right) \\
&\quad + \left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} F'(X_{lh}) \left(X_s - X_{lh} \right) ds \right\|_{L^2(\Omega;H)} \\
&\quad + \frac{R}{2} \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(\mathbb{E} \left\| X_s - X_{lh} \right\|_{H_\beta}^4 \right)^{\frac{1}{2}} ds \right)
\end{aligned}$$

and Lemma 1 gives

$$\begin{aligned}
&\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)} \\
&\leq R \left(2h + \sum_{l=0}^{m-2} \int_{lh}^{(l+1)h} \frac{(s-lh)}{(mh-s)} \, ds \right) \\
&\quad + \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| e^{A(m-l)h} F'(X_{lh}) \left(\left(e^{A(s-lh)} - I \right) X_{lh} \right) \right\|_{L^2(\Omega;H)} ds \\
&\quad + \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| e^{A(m-l)h} F'(X_{lh}) \left(\int_{lh}^{s} e^{A(s-u)} F(X_u) \, du \right) \right\|_{L^2(\Omega;H)} ds \\
&\quad + \left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h} F'(X_{lh}) \left(\int_{lh}^{s} e^{A(s-u)} B(X_u) \, dW_u \right) ds \right\|_{L^2(\Omega;H)} \\
&\quad + \frac{R}{2} \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(R \, (s-lh)^{\min(4(\gamma-\beta),2)} \right)^{\frac{1}{2}} ds \right)
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. This shows

$$\begin{aligned}
&\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)} \\
&\leq R \left(2h + \sum_{l=0}^{m-2} \int_{lh}^{(l+1)h} \frac{(s-lh)}{(m-l-1)h} \, ds \right) \\
&\quad + \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| F'(X_{lh}) \left(\left(e^{A(s-lh)} - I \right) X_{lh} \right) \right\|_{L^2(\Omega;H)} ds \\
&\quad + \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| F'(X_{lh}) \left(\int_{lh}^{s} e^{A(s-u)} F(X_u) \, du \right) \right\|_{L^2(\Omega;H)} ds \\
&\quad + \left\{ \sum_{l=0}^{m-1} \mathbb{E} \left\| \int_{lh}^{(l+1)h} e^{A(m-l)h} F'(X_{lh}) \left(\int_{lh}^{s} e^{A(s-u)} B(X_u) \, dW_u \right) ds \right\|_H^2 \right\}^{\frac{1}{2}} \\
&\quad + \frac{R^2}{2} \left(\sum_{l=0}^{m-1} h^{(1+\min(2(\gamma-\beta),1))} \right)
\end{aligned}$$

and

$$\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)}$$

$$\leq R \left(2h + \sum_{l=0}^{m-2} \frac{h}{2(m-l-1)} \right) + \frac{1}{2} R^2 T h^{\min(2(\gamma-\beta),1)}$$

$$+ R \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| \left(e^{A(s-lh)} - I \right) X_{lh} \right\|_{L^2(\Omega;H)} ds \right)$$

$$+ R \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| \int_{lh}^{s} e^{A(s-u)} F(X_u) \, du \right\|_{L^2(\Omega;H)} ds \right)$$

$$+ \sqrt{h} \left\{ \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \mathbb{E} \left\| F'(X_{lh}) \left(\int_{lh}^{s} e^{A(s-u)} B(X_u) \, dW_u \right) \right\|_H^2 ds \right\}^{\frac{1}{2}}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. Hence, we have

$$\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)}$$

$$\leq R \left(2h + \frac{h}{2} \left(\sum_{l=1}^{m-1} \frac{1}{l} \right) \right) + \frac{1}{2} R^3 h^{\min(2(\gamma-\beta),1)}$$

$$+ R \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\| (-A)^{-\gamma} \left(e^{A(s-lh)} - I \right) \right\|_{L(H)} \left\| (-A)^{\gamma} X_{lh} \right\|_{L^2(\Omega;H)} ds \right)$$

$$+ R \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \int_{lh}^{s} \| F(X_u) \|_{L^2(\Omega;H)} \, du \, ds \right)$$

$$+ R\sqrt{h} \left\{ \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \mathbb{E} \left\| \int_{lh}^{s} e^{A(s-u)} B(X_u) \, dW_u \right\|_H^2 ds \right\}^{\frac{1}{2}}$$

and Lemma 1 shows

$$\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)}$$

$$\leq R \left(2h + \frac{h}{2} \left(1 + \ln(M) \right) \right) + \frac{1}{2} R^4 M^{-\min(2(\gamma-\beta),1)}$$

$$+ R^2 \left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} (s - lh)^{\gamma} \, ds \right) + \frac{1}{2} R^2 M h^2$$

$$+ R\sqrt{h} \left\{ \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \int_{lh}^{s} \mathbb{E} \left\| e^{A(s-u)} B(X_u) \right\|_{HS(U_0,H)}^2 du \, ds \right\}^{\frac{1}{2}}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. Therefore, we obtain

$$\begin{aligned}
&\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)} \\
&\leq Rh \left(\frac{5}{2} + \frac{1}{2} \ln(M) \right) + \frac{1}{2} R^4 M^{-\min(2(\gamma-\beta),1)} + R^2 M h^{(1+\gamma)} + \frac{1}{2} R^3 h \\
&\quad + R\sqrt{h} \left\{ \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \int_{lh}^{s} \mathbb{E} \left\| B(X_u) \right\|^2_{HS(U_0,H)} du\, ds \right\}^{\frac{1}{2}} \\
&\leq R^2 M^{-1} \left(\frac{5}{2} + \frac{1}{2} \ln(M) \right) + \frac{1}{2} R^4 M^{-\min(2(\gamma-\beta),1)} + R^4 M^{-\gamma} + \frac{1}{2} R^4 M^{-1} \\
&\quad + R\sqrt{h} \left\{ \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \int_{lh}^{s} \left\| (-A)^{-\delta} \right\|^2_{L(H)} \mathbb{E} \left\| (-A)^{\delta} B(X_u) \right\|^2_{HS(U_0,H)} du\, ds \right\}^{\frac{1}{2}} \\
&\leq \frac{5}{2} R^2 M^{-1} \left(1 + \ln(M) \right) + \frac{1}{2} R^4 M^{-\min(2(\gamma-\beta),1)} + R^4 M^{-\gamma} + \frac{1}{2} R^4 M^{-1} + R^2 \sqrt{h} \left(\frac{1}{2} M h^2 \right)^{\frac{1}{2}}
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$. The estimate

$$1 + \ln(x) = 1 + \int_1^x \frac{1}{s}\, ds \leq 1 + \int_1^x \frac{1}{s^{(1-r)}}\, ds = 1 + \frac{(x^r - 1)}{r} = \frac{x^r}{r} - \frac{(1-r)}{r} \leq \frac{x^r}{r}$$

for all $r \in (0, 1]$ and all $x \in [1, \infty)$ then shows

$$\begin{aligned}
&\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)} \\
&\leq \frac{5}{2} R^2 \frac{M^{(1-\gamma)}}{M(1-\gamma)} + \frac{R^4}{2M^{\min(2(\gamma-\beta),1)}} + \frac{R^4}{M^\gamma} + \frac{R^4}{2M} + R^2 \sqrt{h} \left(Th \right)^{\frac{1}{2}} \\
&\leq \frac{5R^4}{2M^\gamma} + \frac{R^4}{2M^{\min(2(\gamma-\beta),1)}} + \frac{R^4}{M^\gamma} + \frac{R^4}{2M} + \frac{R^4}{M}
\end{aligned}$$

and finally

$$\begin{aligned}
&\left\| \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} F(X_s) - e^{A(m-l)h} F(X_{lh}) \right) ds \right\|_{L^2(\Omega;H)} \\
&\leq \left(\frac{5}{2} + \frac{1}{2} + 1 + \frac{1}{2} + 1 \right) \frac{R^4}{M^{\min(2(\gamma-\beta),\gamma)}} \leq \frac{6R^4}{M^{\min(2(\gamma-\beta),\gamma)}}
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M \in \mathbb{N}$.

5.2 Noise discretization error: Proof of (92)

We have

$$\begin{aligned}
&\mathbb{E} \left\| \int_s^t e^{A(t-u)} B(X_u)\, d\big(W_u - W_u^K\big) \right\|_H^2 = \mathbb{E} \left\| \sum_{\substack{j \in \mathcal{J} \setminus \mathcal{J}_K \\ \eta_j \neq 0}} \int_s^t e^{A(t-u)} B(X_u) g_j\, d\langle g_j, W_u \rangle_U \right\|_H^2 \\
&= \sum_{\substack{j \in \mathcal{J} \setminus \mathcal{J}_K \\ \eta_j \neq 0}} \eta_j \int_s^t \mathbb{E} \left\| e^{A(t-u)} B(X_u) g_j \right\|_H^2 du = \sum_{\substack{j \in \mathcal{J} \setminus \mathcal{J}_K \\ \eta_j \neq 0}} \eta_j \int_s^t \mathbb{E} \left\| e^{A(t-u)} B(X_u) Q^{-\alpha} \left(Q^\alpha g_j \right) \right\|_H^2 du \\
&= \sum_{\substack{j \in \mathcal{J} \setminus \mathcal{J}_K \\ \eta_j \neq 0}} (\eta_j)^{(1+2\alpha)} \left(\int_s^t \mathbb{E} \left\| e^{A(t-u)} B(X_u) Q^{-\alpha} g_j \right\|_H^2 du \right)
\end{aligned}$$

for all $s, t \in [0, T]$ with $s \leq t$ and all $K \in \mathbb{N}$. This shows

$$\begin{aligned}
&\mathbb{E}\left\|\int_s^t e^{A(t-u)} B(X_u)\, d\big(W_u - W_u^K\big)\right\|_H^2 \\
&\leq \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\sum_{\substack{j \in \mathcal{J}\setminus\mathcal{J}_K \\ \eta_j \neq 0}} \eta_j \left(\int_s^t \mathbb{E}\left\|e^{A(t-u)} B(X_u) Q^{-\alpha} g_j\right\|_H^2 du\right)\right) \\
&\leq \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\sum_{j \in \mathcal{J}} \eta_j \int_s^t \mathbb{E}\left\|e^{A(t-u)} B(X_u) Q^{-\alpha} g_j\right\|_H^2 du\right) \\
&\leq \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\int_s^t \mathbb{E}\left\|e^{A(t-u)} B(X_u) Q^{-\alpha}\right\|_{HS(U_0,H)}^2 du\right) \\
&\leq \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\int_s^t (t-u)^{-2\vartheta}\, \mathbb{E}\left\|(-A)^{-\vartheta} B(X_u) Q^{-\alpha}\right\|_{HS(U_0,H)}^2 du\right)
\end{aligned}$$

for all $s, t \in [0, T]$ with $s \leq t$ and all $K \in \mathbb{N}$. Therefore, we obtain

$$\begin{aligned}
&\mathbb{E}\left\|\int_s^t e^{A(t-u)} B(X_u)\, d\big(W_u - W_u^K\big)\right\|_H^2 \\
&\leq c^2 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\int_s^t (t-u)^{-2\vartheta}\, \mathbb{E}\left[\Big(1 + \|X_u\|_{H_\gamma}\Big)^2\right] du\right) \\
&\leq 2c^2 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\int_s^t (t-u)^{-2\vartheta} \left(1 + \mathbb{E}\|X_u\|_{H_\gamma}^2\right) du\right) \\
&\leq 4R^3 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\int_s^t (t-u)^{-2\vartheta}\, du\right) = 4R^3 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left(\int_0^{(t-s)} u^{-2\vartheta}\, du\right)
\end{aligned}$$

for all $s, t \in [0, T]$ with $s \leq t$ and all $K \in \mathbb{N}$. Hence, we have

$$\begin{aligned}
\mathbb{E}\left\|\int_s^t e^{A(t-u)} B(X_u)\, d\big(W_u - W_u^K\big)\right\|_H^2 &\leq 4R^3 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \left[\frac{u^{(1-2\vartheta)}}{(1-2\vartheta)}\right]_{u=0}^{u=(t-s)} \\
&\leq 4R^4 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} (t-s)^{(1-2\vartheta)} \leq 4R^5 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha} \qquad (100)
\end{aligned}$$

for all $s, t \in [0, T]$ with $s \leq t$ and all $K \in \mathbb{N}$. In particular, we obtain

$$\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(mh-s)} B(X_s)\, d\big(W_s - W_s^K\big)\right\|_H^2 \\
&\qquad = \mathbb{E}\left\|\int_0^{mh} e^{A(mh-s)} B(X_s)\, d\big(W_s - W_s^K\big)\right\|_H^2 \leq 4R^5 \left(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{2\alpha}
\end{aligned}$$

for all $m \in \{0, 1, \ldots, M\}$ and all $M, K \in \mathbb{N}$ which shows (92).

5.3 Temporal discretization error: Proof of (93)

We have

$$\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left(e^{A(mh-s)} - e^{A(m-l)h}\right) B(X_s)\, dW_s^K\right\|_H^2 \\
&\leq \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \mathbb{E}\left\|\left(e^{A(mh-s)} - e^{A(m-l)h}\right) B(X_s)\right\|_{HS(U_0,H)}^2 ds \\
&\leq \sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \left\|(-A)^{-\delta}\left(e^{A(mh-s)} - e^{A(m-l)h}\right)\right\|_{L(H)}^2 \mathbb{E}\left\|(-A)^{\delta} B(X_s)\right\|_{HS(U_0,H)}^2 ds
\end{aligned}$$

and hence

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left(e^{A(mh-s)}-e^{A(m-l)h}\right)B(X_s)\,dW_s^K\right\|_H^2\\
&\leq R\left(\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left\|(-A)^{-\delta}\left(e^{A(mh-s)}-e^{A(m-l)h}\right)\right\|_{L(H)}^2 ds\right)\\
&\leq R\int_{(m-1)h}^{mh}\left\|(-A)^{-\delta}\left(e^{A(mh-s)}-e^{Ah}\right)\right\|_{L(H)}^2 ds\\
&+R\left(\sum_{l=0}^{m-2}\int_{lh}^{(l+1)h}\left\|(-A)^{-1}\left(e^{A(s-lh)}-I\right)\right\|_{L(H)}^2\left\|(-A)^{(1-\delta)}e^{A(mh-s)}\right\|_{L(H)}^2 ds\right)
\end{aligned}
$$

for all $m\in\{0,1,\ldots,M\}$ and all $M,K\in\mathbb{N}$. Therefore, we obtain

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left(e^{A(mh-s)}-e^{A(m-l)h}\right)B(X_s)\,dW_s^K\right\|_H^2\\
&\leq R\int_{(m-1)h}^{mh}\left\|(-A)^{-\delta}\left(e^{A(s-(m-1)h)}-I\right)\right\|_{L(H)}^2 ds\\
&\quad+Rh^2\left(\sum_{l=0}^{m-2}\int_{lh}^{(l+1)h}\left\|(-A)^{(1-\delta)}e^{A(mh-s)}\right\|_{L(H)}^2 ds\right)\\
&\leq R\int_{(m-1)h}^{mh}(s-(m-1)h)^{2\delta}\,ds+Rh^2\left(\sum_{l=0}^{m-2}\int_{lh}^{(l+1)h}(mh-s)^{2(\delta-1)}\,ds\right)\\
&\leq Rh^{(1+2\delta)}+Rh^3\left(\sum_{l=0}^{m-2}(m-l-1)^{2(\delta-1)}\,h^{2(\delta-1)}\right)
\end{aligned}
$$

for all $m\in\{0,1,\ldots,M\}$ and all $M,K\in\mathbb{N}$. This implies

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left(e^{A(mh-s)}-e^{A(m-l)h}\right)B(X_s)\,dW_s^K\right\|_H^2\leq Rh^{(1+2\delta)}+Rh^{(1+2\delta)}\left(\sum_{l=1}^{m-1}l^{2(\delta-1)}\right)\\
&\leq Rh^{(1+2\delta)}\left(2+\sum_{l=2}^{\infty}l^{2(\delta-1)}\right)\leq Rh^{(1+2\delta)}\left(2+\int_1^\infty s^{2(\delta-1)}ds\right)\\
&\leq Rh^{(1+2\delta)}\left(2+\left[\frac{s^{(2\delta-1)}}{(2\delta-1)}\right]_{s=1}^{s=\infty}\right)=Rh^{(1+2\delta)}\left(2+\frac{1}{(1-2\delta)}\right)\leq 3R^2h^{(1+2\delta)}\leq\frac{3R^4}{M^{(1+2\delta)}}
\end{aligned}
$$

for all $m\in\{0,1,\ldots,M\}$ and all $M,K\in\mathbb{N}$.

5.4 Temporal discretization error: Proof of (94)

We have

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}e^{A(m-l)h}\int_0^1 B''(X_{lh}+r\,(X_s-X_{lh}))\,(X_s-X_{lh},X_s-X_{lh})\,(1-r)\,dr\,dW_s^K\right\|_H^2\\
&\leq\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\int_0^1\mathbb{E}\left\|B''(X_{lh}+r\,(X_s-X_{lh}))(X_s-X_{lh},X_s-X_{lh})\right\|_{HS(U_0,H)}^2 dr\,ds\\
&\leq\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\mathbb{E}\left[\left(R\|X_s-X_{lh}\|_{H_\beta}^2\right)^2\right]ds=R^2\left(\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\mathbb{E}\|X_s-X_{lh}\|_{H_\beta}^4\,ds\right)\\
&\leq R^2\left(\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}R\,(s-lh)^{\min(4(\gamma-\beta),2)}\,ds\right)\leq R^3\left(\sum_{l=0}^{m-1}h^{(1+\min(4(\gamma-\beta),2))}\right)\\
&\leq R^3Mh^{(1+\min(4(\gamma-\beta),2))}=R^3Th^{\min(4(\gamma-\beta),2)}\leq\frac{R^6}{M^{\min(4(\gamma-\beta),2)}}
\end{aligned}
$$

for all $m\in\{0,1,\ldots,M\}$ and all $M,K\in\mathbb{N}$.

5.5 Temporal discretization error: Proof of (95)

In order to show (95), we first estimate

$$\mathbb{E}\left\|X_t - X_s - \int_s^t B(X_s)\,dW_u^K\right\|_H^2$$

for all $s,t \in [0,T]$ with $s \le t$ and all $K \in \mathbb{N}$. More precisely, we have

$$\begin{aligned}
&\mathbb{E}\left\|X_t - X_s - \int_s^t B(X_s)\,dW_u^K\right\|_H^2 \le 5 \cdot \mathbb{E}\left\|\left(e^{A(t-s)} - I\right) X_s\right\|_H^2 \\
&\quad + 5 \cdot \mathbb{E}\left\|\int_s^t e^{A(t-u)} F(X_u)\,du\right\|_H^2 + 5 \cdot \mathbb{E}\left\|\int_s^t e^{A(t-u)} B(X_u)\,d\big(W_u - W_u^K\big)\right\|_H^2 \\
&\quad + 5 \cdot \mathbb{E}\left\|\int_s^t \left(e^{A(t-u)} - I\right) B(X_u)\,dW_u^K\right\|_H^2 + 5 \cdot \mathbb{E}\left\|\int_s^t \left(B(X_u) - B(X_s)\right) dW_u^K\right\|_H^2
\end{aligned}$$

and using (100) shows

$$\begin{aligned}
&\mathbb{E}\left\|X_t - X_s - \int_s^t B(X_s)\,dW_u^K\right\|_H^2 \le 5\left\|(-A)^{-\gamma}\left(e^{A(t-s)} - I\right)\right\|_{L(H)}^2 \mathbb{E}\|(-A)^{\gamma} X_s\|_H^2 \\
&\quad + 5\,(t-s)\left(\int_s^t \mathbb{E}\left\|e^{A(t-u)} F(X_u)\right\|_H^2 du\right) + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} \\
&\quad + 5\left(\int_s^t \mathbb{E}\left\|\left(e^{A(t-u)} - I\right) B(X_u)\right\|_{HS(U_0,H)}^2 du\right) \\
&\quad + 5\left(\int_s^t \mathbb{E}\,\|B(X_u) - B(X_s)\|_{HS(U_0,H)}^2\, du\right)
\end{aligned}$$

for all $s,t \in [0,T]$ with $s \le t$ and all $K \in \mathbb{N}$. This implies

$$\begin{aligned}
&\mathbb{E}\left\|X_t - X_s - \int_s^t B(X_s)\,dW_u^K\right\|_H^2 \\
&\le 5R\,(t-s)^{2\gamma} + 5\,(t-s)\left(\int_s^t \mathbb{E}\,\|F(X_u)\|_H^2\, du\right) + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} \\
&\quad + 5\left(\int_s^t \left\|(-A)^{-\delta}\left(e^{A(t-u)} - I\right)\right\|_{L(H)}^2 \mathbb{E}\left\|(-A)^{\delta} B(X_u)\right\|_{HS(U_0,H)}^2 du\right) \\
&\quad + 5R^2\left(\int_s^t \mathbb{E}\,\|X_u - X_s\|_H^2\, du\right) \\
&\le 5R\,(t-s)^{2\gamma} + 5R\,(t-s)^2 + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} \\
&\quad + 5\left(\int_s^t (t-u)^{2\delta}\,\mathbb{E}\left\|(-A)^{\delta} B(X_u)\right\|_{HS(U_0,H)}^2 du\right) \\
&\quad + 5R^2\left(\int_s^t \left\|(-A)^{-\beta}\right\|_{L(H)}^2 \mathbb{E}\,\|X_u - X_s\|_{H_\beta}^2\, du\right)
\end{aligned}$$

for all $s,t \in [0,T]$ with $s \le t$ and all $K \in \mathbb{N}$. Therefore, we obtain

$$\begin{aligned}
&\mathbb{E}\left\|X_t - X_s - \int_s^t B(X_s)\,dW_u^K\right\|_H^2 \le 5R\,(t-s)^{2\gamma} + 5R\,(t-s)^2 \\
&\quad + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} + 5R\left(\int_s^t (t-u)^{2\delta}\, du\right) + 5R^4\left(\int_s^t \mathbb{E}\,\|X_u - X_s\|_{H_\beta}^2\, du\right) \\
&\le 10R^3\,(t-s)^{2\gamma} + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} + 5R\,(t-s)^{(1+2\delta)} + 5R^4\left(\int_s^t \mathbb{E}\,\|X_u - X_s\|_{H_\beta}^2\, du\right) \\
&\le 15R^3\,(t-s)^{2\gamma} + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} + 5R^4\left(\int_s^t \left(\mathbb{E}\,\|X_u - X_s\|_{H_\beta}^4\right)^{\frac{1}{2}} du\right) \\
&\le 15R^3\,(t-s)^{2\gamma} + 20R^5\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha} + 5R^4\left(\int_s^t \left(R\,(u-s)^{\min(4(\gamma-\beta),2)}\right)^{\frac{1}{2}} du\right)
\end{aligned}$$

and hence

$$
\begin{aligned}
&\mathbb{E}\left\|X_t - X_s - \int_s^t B(X_s)\, dW_u^K\right\|_H^2 \\
&\leq 15R^3\,(t-s)^{2\gamma} + 20R^5\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha} + 5R^5\left(\int_s^t (u-s)^{\min(2(\gamma-\beta),1)}\,du\right) \\
&\leq 15R^3\,(t-s)^{2\gamma} + 20R^5\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha} + 5R^5\,(t-s)^{(1+\min(2(\gamma-\beta),1))} \\
&\leq 15R^3\,(t-s)^{2\gamma} + 20R^5\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha} + 5R^6\,(t-s)^{\min(4(\gamma-\beta),2)} \\
&\leq 20R^8\,(t-s)^{\min(4(\gamma-\beta),2\gamma)} + 20R^5\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha}
\end{aligned}
\tag{101}
$$

for all $s,t\in[0,T]$ with $s\leq t$ and all $K\in\mathbb{N}$. Now we prove (95). To this end we note that

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}B'(X_{lh})\left(X_s - X_{lh} - \int_{lh}^s B(X_{lh})\,dW_u^K\right)dW_s^K\right\|_H^2 \\
&\leq \sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\mathbb{E}\left\|B'(X_{lh})\left(X_s - X_{lh} - \int_{lh}^s B(X_{lh})\,dW_u^K\right)\right\|_{HS(U_0,H)}^2 ds \\
&\leq R^2\left(\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\mathbb{E}\left\|X_s - X_{lh} - \int_{lh}^s B(X_{lh})\,dW_u^K\right\|_H^2 ds\right)
\end{aligned}
$$

holds for all $m\in\{0,1,\ldots,M\}$ and all $M,K\in\mathbb{N}$. Hence, (101) yields

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h} e^{A(m-l)h}B'(X_{lh})\left(X_s - X_{lh} - \int_{lh}^s B(X_{lh})\,dW_u^K\right)dW_s^K\right\|_H^2 \\
&\leq 20R^{10}\left(\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}\left((s-lh)^{\min(4(\gamma-\beta),2\gamma)} + \Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha}\right)ds\right) \\
&\leq 20R^{10}\left(\sum_{l=0}^{m-1}\int_{lh}^{(l+1)h}(s-lh)^{\min(4(\gamma-\beta),2\gamma)}\,ds\right) + 20R^{10}T\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha} \\
&\leq 20R^{10}Mh^{(1+\min(4(\gamma-\beta),2\gamma))} + 20R^{11}\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha} \\
&\leq 20R^{11}h^{\min(4(\gamma-\beta),2\gamma)} + 20R^{11}\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha} \leq \frac{20R^{13}}{M^{\min(4(\gamma-\beta),2\gamma)}} + 20R^{11}\Bigg(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Bigg)^{2\alpha}
\end{aligned}
$$

for all $m\in\{0,1,\ldots,M\}$ and all $M,K\in\mathbb{N}$.

5.6 Lipschitz estimates: Proof of (96)

Before we estimate $\mathbb{E}\left\|Z_m^{N,M,K} - Y_m^{N,M,K}\right\|_H^2$ for $m \in \{0,1,\ldots,M\}$ and for $N,M,K \in \mathbb{N}$, we need some preparations. More precisely, we have

$$\begin{aligned}
&\mathbb{E}\left\|B'(X_{lh})\int_{lh}^{s} B(X_{lh})\,dW_u^K - B'\Big(Y_l^{N,M,K}\Big)\int_{lh}^{s} B\Big(Y_l^{N,M,K}\Big)\,dW_u^K\right\|_{HS(U_0,H)}^2 \\
&= \mathbb{E}\Bigg\|B'(X_{lh})\Bigg(\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}}\int_{lh}^{s} B(X_{lh})\,g_j\,d\langle g_j, W_u\rangle_U\Bigg) \\
&\quad - B'\Big(Y_l^{N,M,K}\Big)\Bigg(\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}}\int_{lh}^{s} B\Big(Y_l^{N,M,K}\Big)\,g_j\,d\langle g_j, W_u\rangle_U\Bigg)\Bigg\|_{HS(U_0,H)}^2 \\
&= \mathbb{E}\Bigg\|B'(X_{lh})\Bigg(\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} B(X_{lh})\,g_j\,\langle g_j, W_s - W_{lh}\rangle_U\Bigg) \\
&\quad - B'\Big(Y_l^{N,M,K}\Big)\Bigg(\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} B\Big(Y_l^{N,M,K}\Big)\,g_j\,\langle g_j, W_s - W_{lh}\rangle_U\Bigg)\Bigg\|_{HS(U_0,H)}^2 \\
&= \mathbb{E}\Bigg\|\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}}\left\{B'(X_{lh})\,(B(X_{lh})\,g_j) - B'\Big(Y_l^{N,M,K}\Big)\Big(B\Big(Y_l^{N,M,K}\Big)\,g_j\Big)\right\}\langle g_j, W_s - W_{lh}\rangle_U\Bigg\|_{HS(U_0,H)}^2
\end{aligned}$$

for all $s \in [lh,(l+1)h]$, $l \in \{0,1,\ldots,M-1\}$ and all $M,K \in \mathbb{N}$. This implies

$$\begin{aligned}
&\mathbb{E}\left\|B'(X_{lh})\int_{lh}^{s} B(X_{lh})\,dW_u^K - B'\Big(Y_l^{N,M,K}\Big)\int_{lh}^{s} B\Big(Y_l^{N,M,K}\Big)\,dW_u^K\right\|_{HS(U_0,H)}^2 \\
&= \sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}}\mathbb{E}\left\|\left\{B'(X_{lh})\,(B(X_{lh})\,g_j) - B'\Big(Y_l^{N,M,K}\Big)\Big(B\Big(Y_l^{N,M,K}\Big)\,g_j\Big)\right\}\langle g_j, W_s - W_{lh}\rangle_U\right\|_{HS(U_0,H)}^2 \\
&= \sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}}\mathbb{E}\left\|B'(X_{lh})\,(B(X_{lh})\,g_j) - B'\Big(Y_l^{N,M,K}\Big)\Big(B\Big(Y_l^{N,M,K}\Big)\,g_j\Big)\right\|_{HS(U_0,H)}^2\cdot\mathbb{E}\left|\langle g_j, W_s - W_{lh}\rangle_U\right|^2
\end{aligned}$$

for all $s \in [lh,(l+1)h]$, $l \in \{0,1,\ldots,M-1\}$ and all $M,K \in \mathbb{N}$. Hence, we obtain

$$\begin{aligned}
&\mathbb{E}\left\|B'(X_{lh})\int_{lh}^{s} B(X_{lh})\,dW_u^K - B'\Big(Y_l^{N,M,K}\Big)\int_{lh}^{s} B\Big(Y_l^{N,M,K}\Big)\,dW_u^K\right\|_{HS(U_0,H)}^2 \\
&= \sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}}\eta_j\cdot\mathbb{E}\left\|B'(X_{lh})(B(X_{lh})\,g_j) - B'\Big(Y_l^{N,M,K}\Big)\Big(B\Big(Y_l^{N,M,K}\Big)\,g_j\Big)\right\|_{HS(U_0,H)}^2(s-lh) \\
&\leq \sum_{\substack{j,k\in\mathcal{J}\\ \eta_j,\eta_k\neq 0}}\eta_j\,\eta_k\,\mathbb{E}\left\|B'(X_{lh})(B(X_{lh})\,g_j)\,g_k - B'\Big(Y_l^{N,M,K}\Big)\Big(B\Big(Y_l^{N,M,K}\Big)\,g_j\Big)\,g_k\right\|_H^2(s-lh) \qquad (102)\\
&= (s-lh)\cdot\mathbb{E}\left\|B'(X_{lh})\,B(X_{lh}) - B'\Big(Y_l^{N,M,K}\Big)\,B\Big(Y_l^{N,M,K}\Big)\right\|_{HS^{(2)}(U_0,H)}^2 \\
&\leq c^2\cdot(s-lh)\cdot\mathbb{E}\left\|X_{lh} - Y_l^{N,M,K}\right\|_H^2 \leq R^3\cdot\mathbb{E}\left\|X_{lh} - Y_l^{N,M,K}\right\|_H^2
\end{aligned}$$

for all $s \in [lh, (l+1)h]$, $l \in \{0, 1, \ldots, M-1\}$ and all $M, K \in \mathbb{N}$. Additionally, (87) and the fact $\|P_N(v)\|_H \leq \|v\|_H$ for all $v \in H$ show

$$
\begin{aligned}
&\mathbb{E}\left\|Z_m^{N,M,K} - Y_m^{N,M,K}\right\|_H^2 \\
&\leq 3 \cdot \mathbb{E}\left\|\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h}\left(F(X_{lh}) - F\left(Y_l^{N,M,K}\right)\right) ds\right\|_H^2 \\
&\quad + 3 \cdot \mathbb{E}\left\|\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h}\left(B(X_{lh}) - B\left(Y_l^{N,M,K}\right)\right) dW_s^K\right\|_H^2 \\
&\quad + 3 \cdot \mathbb{E}\left\|\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} e^{A(m-l)h}\left(B'(X_{lh}) \int_{lh}^{s} B(X_{lh})\, dW_u^K - B'\left(Y_l^{N,M,K}\right) \int_{lh}^{s} B\left(Y_l^{N,M,K}\right) dW_u^K\right) dW_s^K\right\|_H^2 \\
&\leq 3Mh^2\left(\sum_{l=0}^{m-1} \mathbb{E}\left\|e^{A(m-l)h}\left(F(X_{lh}) - F\left(Y_l^{N,M,K}\right)\right)\right\|_H^2\right) \\
&\quad + 3\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \mathbb{E}\left\|e^{A(m-l)h}\left(B(X_{lh}) - B\left(Y_l^{N,M,K}\right)\right)\right\|_{HS(U_0,H)}^2 ds\right) \\
&\quad + 3\left(\sum_{l=0}^{m-1} \int_{lh}^{(l+1)h} \mathbb{E}\left\|B'(X_{lh}) \int_{lh}^{s} B(X_{lh})\, dW_u^K - B'\left(Y_l^{N,M,K}\right) \int_{lh}^{s} B\left(Y_l^{N,M,K}\right) dW_u^K\right\|_{HS(U_0,H)}^2 ds\right)
\end{aligned}
\tag{103}
$$

and due to (102) we finally obtain

$$
\begin{aligned}
\mathbb{E}\left\|Z_m^{N,M,K} - Y_m^{N,M,K}\right\|_H^2 &\leq 3Th\left(\sum_{l=0}^{m-1} \mathbb{E}\left\|F(X_{lh}) - F\left(Y_l^{N,M,K}\right)\right\|_H^2\right) \\
&\quad + 3h\left(\sum_{l=0}^{m-1} \mathbb{E}\left\|B(X_{lh}) - B\left(Y_l^{N,M,K}\right)\right\|_{HS(U_0,H)}^2\right) + 3R^3 h\left(\sum_{l=0}^{m-1} \mathbb{E}\left\|X_{lh} - Y_l^{N,M,K}\right\|_H^2\right) \\
&\leq 9R^3 h\left(\sum_{l=0}^{m-1} \mathbb{E}\left\|X_{lh} - Y_l^{N,M,K}\right\|_H^2\right) \leq \frac{9R^4}{M}\left(\sum_{l=0}^{m-1} \mathbb{E}\left\|X_{lh} - Y_l^{N,M,K}\right\|_H^2\right)
\end{aligned}
$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$.

5.7 Iterated integral identity: Proof of (83)

First of all, we have

$$
\begin{aligned}
&\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\left(Y_m^{N,M,K}\right)\left(\int_{\frac{mT}{M}}^{s} B\left(Y_m^{N,M,K}\right) dW_u^K\right) dW_s^K \\
&= \sum_{\substack{j,k \in \mathcal{J}_K \\ \eta_j, \eta_k \neq 0}} \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\left(Y_m^{N,M,K}\right)\left(\int_{\frac{mT}{M}}^{s} B\left(Y_m^{N,M,K}\right) g_k\, d\langle g_k, W_u\rangle_U\right) g_j\, d\langle g_j, W_s\rangle_U \\
&= \sum_{\substack{j,k \in \mathcal{J}_K \\ \eta_j, \eta_k \neq 0}} B'\left(Y_m^{N,M,K}\right)\left(B\left(Y_m^{N,M,K}\right) g_k\right) g_j \cdot \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} \int_{\frac{mT}{M}}^{s} d\langle g_k, W_u\rangle_U\, d\langle g_j, W_s\rangle_U
\end{aligned}
$$

$\mathbb{P}$-a.s. and hence

$$
\begin{aligned}
&\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\left(Y_m^{N,M,K}\right)\left(\int_{\frac{mT}{M}}^{s} B\left(Y_m^{N,M,K}\right) dW_u^K\right) dW_s^K \\
&= \sum_{\substack{j \in \mathcal{J}_K \\ \eta_j \neq 0}} B'\left(Y_m^{N,M,K}\right)\left(B\left(Y_m^{N,M,K}\right) g_j\right) g_j \cdot \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} \int_{\frac{mT}{M}}^{s} d\langle g_j, W_u\rangle_U\, d\langle g_j, W_s\rangle_U \\
&\quad + \sum_{\substack{j,k \in \mathcal{J}_K \\ \eta_j, \eta_k \neq 0 \\ j \neq k}} B'\left(Y_m^{N,M,K}\right)\left(B\left(Y_m^{N,M,K}\right) g_k\right) g_j \cdot \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} \int_{\frac{mT}{M}}^{s} d\langle g_k, W_u\rangle_U\, d\langle g_j, W_s\rangle_U
\end{aligned}
$$

$\mathbb{P}$-a.s. for all $m \in \{0,1,\dots,M-1\}$ and all $N,M,K \in \mathbb{N}$. Moreover, since the bilinear operator $B'\big(Y_m^{N,M,K}\big)\,B\big(Y_m^{N,M,K}\big) \in HS^{(2)}(U_0,H)$ is symmetric (see Assumption 3) and since

$$\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} \int_{\frac{mT}{M}}^{s} d\langle g_j, W_u\rangle_U \; d\langle g_j, W_s\rangle_U = \frac{1}{2}\left(\left(\langle g_j, \Delta W_m^{M,K}\rangle_U\right)^2 - \frac{T\eta_j}{M}\right)$$

$\mathbb{P}$-a.s. holds for all $j \in \mathcal{J}_K$, $m \in \{0,1,\dots,M-1\}$ and all $N,M,K \in \mathbb{N}$ (see (3.6) in Section 10.3 in [34]), we obtain

$$\begin{aligned}
&\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\big(Y_m^{N,M,K}\big)\left(\int_{\frac{mT}{M}}^{s} B\big(Y_m^{N,M,K}\big)\,dW_u^K\right)dW_s^K \\
&= \frac{1}{2}\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} B'\big(Y_m^{N,M,K}\big)\left(B\big(Y_m^{N,M,K}\big)\,g_j\right) g_j \left(\left(\langle g_j, \Delta W_m^{M,K}\rangle_U\right)^2 - \frac{T\eta_j}{M}\right) \\
&\quad + \frac{1}{2}\sum_{\substack{j,k\in\mathcal{J}_K\\ \eta_j,\eta_k\neq 0\\ j\neq k}} B'\big(Y_m^{N,M,K}\big)\left(B\big(Y_m^{N,M,K}\big)\,g_k\right) g_j \\
&\qquad \cdot \left(\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}}\int_{\frac{mT}{M}}^{s} d\langle g_k, W_u\rangle_U\, d\langle g_j, W_s\rangle_U + \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}}\int_{\frac{mT}{M}}^{s} d\langle g_j, W_u\rangle_U\, d\langle g_k, W_s\rangle_U\right)
\end{aligned}$$

$\mathbb{P}$-a.s. for all $m \in \{0,1,\dots,M-1\}$ and all $N,M,K \in \mathbb{N}$. The fact

$$\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}}\int_{\frac{mT}{M}}^{s} d\langle g_k, W_u\rangle_U\, d\langle g_j, W_s\rangle_U + \int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}}\int_{\frac{mT}{M}}^{s} d\langle g_j, W_u\rangle_U\, d\langle g_k, W_s\rangle_U = \langle g_k, \Delta W_m^{M,K}\rangle_U\,\langle g_j, \Delta W_m^{M,K}\rangle_U \tag{104}$$

$\mathbb{P}$-a.s. for all $j \in \mathcal{J}_K$, $m \in \{0,1,\dots,M-1\}$ and all $N,M,K \in \mathbb{N}$ (see (3.15) in Section 10.3 in [34]) then yields

$$\begin{aligned}
&\int_{\frac{mT}{M}}^{\frac{(m+1)T}{M}} B'\big(Y_m^{N,M,K}\big)\left(\int_{\frac{mT}{M}}^{s} B\big(Y_m^{N,M,K}\big)\,dW_u^K\right)dW_s^K \\
&= \frac{1}{2}\sum_{\substack{j,k\in\mathcal{J}_K\\ \eta_j,\eta_k\neq 0}} B'\big(Y_m^{N,M,K}\big)\left(B\big(Y_m^{N,M,K}\big)\,g_k\right) g_j\,\langle g_k, \Delta W_m^{M,K}\rangle_U\,\langle g_j, \Delta W_m^{M,K}\rangle_U \\
&\quad - \frac{T}{2M}\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} \eta_j B'\big(Y_m^{N,M,K}\big)\left(B\big(Y_m^{N,M,K}\big)\,g_j\right) g_j \\
&= \frac{1}{2}B'\big(Y_m^{N,M,K}\big)\left(B\big(Y_m^{N,M,K}\big)\,\Delta W_m^{M,K}\right)\Delta W_m^{M,K} - \frac{T}{2M}\sum_{\substack{j\in\mathcal{J}_K\\ \eta_j\neq 0}} \eta_j B'\big(Y_m^{N,M,K}\big)\left(B\big(Y_m^{N,M,K}\big)\,g_j\right) g_j
\end{aligned}$$

$\mathbb{P}$-a.s. for all $m \in \{0,1,\dots,M-1\}$ and all $N,M,K \in \mathbb{N}$ which shows (83).

Acknowledgment

This work has been supported by the BiBoS Research Center, by the research project "Numerical solutions of stochastic differential equations with non-globally Lipschitz continuous coefficients" funded by the German Research Foundation, by the Collaborative Research Centre 701 "Spectral Structures and Topological Methods in Mathematics" funded by the German Research Foundation and by the International Graduate School "Stochastics and Real World Models" funded by the German Research Foundation. The support of Issac Newton Institute for Mathematical Sciences in Cambridge is also gratefully acknowledged where part of this was done during the special semester on "Stochastic Partial Differential Equations". In addition, we are very grateful to Sebastian Becker for his help with the numerical simulations and to Carlo Marinelli for his helpful advice concerning Schatten norms.

References

[1] Allen, E. J., Novosel, S. J., and Zhang, Z. Finite element and difference approximation of some linear stochastic partial differential equations. *Stochastics Stochastics Rep. 64*, 1-2 (1998), 117–142.

[2] Babuška, I., Tempone, R., and Zouraris, G. E. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal. 42*, 2 (2004), 800–825 (electronic).

[3] Barth, A., and Lang, A. L^p and almost sure convergence of a Milstein scheme for stochastic partial differential equations. *SAM-Report 2011-15* (2011), 27 pages.

[4] Barth, A., and Lang, A. Milstein approximation for advection-diffusion equations driven by multiplicative noncontinuous martingale noises. *SAM-Report 2011-36* (2011), 25 pages.

[5] Barth, A., Lang, A., and Schwab, C. Multi-level Monte Carlo Finite Element method for parabolic stochastic partial differential equations. *Report 2011-30, SAM, ETH Zurich http://www.sam.math.ethz.ch/reports/2011/30* (2011), 24 pages.

[6] Bensoussan, A., Glowinski, R., and Răşcanu, A. Approximation of the Zakai equation by the splitting up method. *SIAM J. Control Optim. 28*, 6 (1990), 1420–1431.

[7] Bensoussan, A., Glowinski, R., and Răşcanu, A. Approximation of some stochastic differential equations by the splitting up method. *Appl. Math. Optim. 25*, 1 (1992), 81–106.

[8] Da Prato, G., Jentzen, A., and Röckner, M. A mild Itô formula for SPDEs. *arXiv:1009.3526v3* (2011), 30 pages.

[9] Debussche, A. Weak approximation of stochastic partial differential equations: the nonlinear case. *Math. Comp. 80*, 273 (2011), 89–117.

[10] Debussche, A., and Printems, J. Weak order for the discretization of the stochastic heat equation. *Math. Comp. 78*, 266 (2009), 845–863.

[11] Du, Q., and Zhang, T. Numerical approximation of some linear stochastic partial differential equations driven by special additive noises. *SIAM J. Numer. Anal. 40*, 4 (2002), 1421–1445 (electronic).

[12] Dunford, N., and Schwartz, J. T. *Linear operators. Part II.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. Spectral theory. Selfadjoint operators in Hilbert space, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1963 original, A Wiley-Interscience Publication.

[13] Florchinger, P., and Le Gland, F. Time-discretization of the Zakai equation for diffusion processes observed in correlated noise. *Stochastics Stochastics Rep. 35*, 4 (1991), 233–256.

[14] Geissert, M., Kovács, M., and Larsson, S. Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise. *BIT 49*, 2 (2009), 343–356.

[15] Giles, M. B. Multilevel Monte Carlo path simulation. *Oper. Res. 56*, 3 (2008), 607–617.

[16] Grecksch, W., and Kloeden, P. E. Time-discretised Galerkin approximations of parabolic stochastic PDEs. *Bull. Austral. Math. Soc. 54*, 1 (1996), 79–85.

[17] Gyöngy, I. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I. *Potential Anal. 9*, 1 (1998), 1–25.

[18] Gyöngy, I. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II. *Potential Anal. 11*, 1 (1999), 1–37.

[19] Gyöngy, I., and Krylov, N. On the rate of convergence of splitting-up approximations for SPDEs. In *Stochastic inequalities and applications*, vol. 56 of *Progr. Probab.* Birkhäuser, Basel, 2003, pp. 301–321.

[20] Gyöngy, I., and Krylov, N. On the splitting-up method and stochastic partial differential equations. *Ann. Probab. 31*, 2 (2003), 564–591.

[21] Gyöngy, I., and Krylov, N. An accelerated splitting-up method for parabolic equations. *SIAM J. Math. Anal. 37*, 4 (2005), 1070–1097 (electronic).

[22] Gyöngy, I., and Martínez, T. On numerical solution of stochastic partial differential equations of elliptic type. *Stochastics 78*, 4 (2006), 213–231.

[23] Hausenblas, E. Numerical analysis of semilinear stochastic evolution equations in Banach spaces. *J. Comput. Appl. Math. 147*, 2 (2002), 485–516.

[24] Hausenblas, E. Approximation for semilinear stochastic evolution equations. *Potential Anal. 18*, 2 (2003), 141–186.

[25] Hausenblas, E. Finite element approximation of stochastic partial differential equations driven by Poisson random measures of jump type. *SIAM J. Numer. Anal. 46*, 1 (2007/08), 437–471.

[26] Heinrich, S. Monte Carlo complexity of global solution of integral equations. *J. Complexity 14*, 2 (1998), 151–175.

[27] Ito, K., and Rozovskii, B. Approximation of the Kushner equation for nonlinear filtering. *SIAM J. Control Optim. 38*, 3 (2000), 893–915 (electronic).

[28] Jentzen, A. Pathwise numerical approximations of SPDEs with additive noise under non-global Lipschitz coefficients. *Potential Anal. 31*, 4 (2009), 375–404.

[29] Jentzen, A. Taylor expansions of solutions of stochastic partial differential equations. *Discrete Contin. Dyn. Syst. Ser. B 14*, 2 (2010), 515–557.

[30] Jentzen, A., and Kloeden, P. E. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 465*, 2102 (2009), 649–667.

[31] Jentzen, A., and Kloeden, P. E. Taylor expansions of solutions of stochastic partial differential equations with additive noise. *Ann. Probab. 38*, 2 (2010), 532–569.

[32] Jentzen, A., and Röckner, M. Regularity analysis of stochastic partial differential equations with nonlinear multiplicative trace class noise. *J. Differential Equations 252*, 1 (2012), 114–136.

[33] Katsoulakis, M. A., Kossioris, G. T., and Lakkis, O. Noise regularization and computations for the 1-dimensional stochastic Allen-Cahn problem. *Interfaces Free Bound. 9*, 1 (2007), 1–30.

[34] Kloeden, P. E., and Platen, E. *Numerical solution of stochastic differential equations*, vol. 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[35] Kloeden, P. E., and Shott, S. Linear-implicit strong schemes for Itô-Galerkin approximations of stochastic PDEs. *J. Appl. Math. Stochastic Anal. 14*, 1 (2001), 47–53. Special issue: Advances in applied stochastics.

[36] Kossioris, G. T., and Zouraris, G. E. Fully-discrete finite element approximations for a fourth-order linear stochastic parabolic equation with additive space-time white noise: Ii. 2d and 3d case. *arXiv:0906.1828v1* (2009), 23 pages.

[37] Kovács, M., Larsson, S., and Lindgren, F. Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise. *Numer. Algorithms 53*, 2-3 (2010), 309–320.

[38] Kovács, M., Larsson, S., and Saedpanah, F. Finite element approximation of the linear stochastic wave equation with additive noise. *SIAM J. Numer. Anal. 48*, 2 (2010), 408–427.

[39] Lang, A., Chow, P.-L., and Potthoff, J. Almost sure convergence for a semidiscrete Milstein scheme for SPDEs of Zakai type. *Stochastics 82*, 1-3 (2010), 315–326.

[40] Lang, A., Chow, P.-L., and Potthoff, J. Erratum: Almost sure convergence for a semidiscrete Milstein scheme for SPDEs of Zakai type. *Stochastics* (2011). http://www.tandfonline.com/doi/abs/10.1080/17442508.2011.618884.

[41] Lord, G. J., and Rougemont, J. A numerical scheme for stochastic PDEs with Gevrey regularity. *IMA J. Numer. Anal. 24*, 4 (2004), 587–604.

[42] Lord, G. J., and Shardlow, T. Postprocessing for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal. 45*, 2 (2007), 870–889 (electronic).

[43] Lord, G. J., and Tambue, A. A modified semi-implict Euler-Maruyama Scheme for finite element discretization of SPDEs. *arXiv:1004.1998v1* (2010), 23 pages.

[44] Millet, A., and Morien, P.-L. On implicit and explicit discretization schemes for parabolic SPDEs in any dimension. *Stochastic Process. Appl. 115*, 7 (2005), 1073–1106.

[45] Milstein, G. N. Approximate integration of stochastic differential equations. *Teor. Verojatnost. i Primenen. 19* (1974), 583–588.

[46] Mishura, Y. S., and Shevchenko, G. M. Approximation schemes for stochastic differential equations in a Hilbert space. *Teor. Veroyatn. Primen. 51*, 3 (2006), 476–495.

[47] Müller-Gronbach, T., and Ritter, K. An implicit Euler scheme with non-uniform time discretization for heat equations with multiplicative noise. *BIT 47*, 2 (2007), 393–418.

[48] Müller-Gronbach, T., and Ritter, K. Lower bounds and nonuniform time discretization for approximation of stochastic heat equations. *Found. Comput. Math. 7*, 2 (2007), 135–181.

[49] Müller-Gronbach, T., Ritter, K., and Wagner, T. Optimal pointwise approximation of a linear stochastic heat equation with additive space-time white noise. In *Monte Carlo and quasi-Monte Carlo methods 2006*. Springer, Berlin, 2007, pp. 577–589.

[50] Müller-Gronbach, T., Ritter, K., and Wagner, T. Optimal pointwise approximation of infinite-dimensional Ornstein-Uhlenbeck processes. *Stoch. Dyn. 8*, 3 (2008), 519–541.

[51] Pettersson, R., and Signahl, M. Numerical approximation for a white noise driven SPDE with locally bounded drift. *Potential Anal. 22*, 4 (2005), 375–393.

[52] Prévôt, C., and Röckner, M. *A concise course on stochastic partial differential equations*, vol. 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[53] Renardy, M., and Rogers, R. C. *An introduction to partial differential equations*, vol. 13 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1993.

[54] Roth, C. Difference methods for stochastic partial differential equations. *ZAMM Z. Angew. Math. Mech. 82*, 11-12 (2002), 821–830. 4th GAMM-Workshop "Stochastic Models and Control Theory" (Lutherstadt Wittenberg, 2001).

[55] Roth, C. A combination of finite difference and Wong-Zakai methods for hyperbolic stochastic partial differential equations. *Stoch. Anal. Appl. 24*, 1 (2006), 221–240.

[56] Roth, C. Weak approximations of solutions of a first order hyperbolic stochastic partial differential equation. *Monte Carlo Methods Appl. 13*, 2 (2007), 117–133.

[57] Sell, G. R., and You, Y. *Dynamics of evolutionary equations*, vol. 143 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002.

[58] Shardlow, T. Numerical methods for stochastic parabolic PDEs. *Numer. Funct. Anal. Optim. 20*, 1-2 (1999), 121–145.

[59] Walsh, J. B. Finite element methods for parabolic stochastic PDE's. *Potential Anal. 23*, 1 (2005), 1–43.

[60] Walsh, J. B. On numerical solutions of the stochastic wave equation. *Illinois J. Math. 50*, 1-4 (2006), 991–1018 (electronic).

[61] Wang, X. A Runge-Kutta type scheme for nonlinear stochastic partial differential equations with multiplicative trace class noise. *arXiv:1106.1889v1* (2011), 21 pages.

[62] Yan, Y. Galerkin finite element methods for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal. 43*, 4 (2005), 1363–1384 (electronic).