UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR MATHEMATIK

Derivative-Free Numerical Schemes for Stochastic Partial Differential Equations

Dissertation for Fulfillment of
Requirements for the Doctoral Degree
of the University of Lübeck

from the Institute of Mathematics

Submitted by

Claudine Leonhard

from Radolfzell am Bodensee

Lübeck, May 16, 2016

First referee:	Prof. Dr. Andreas Rößler
Second referee:	Prof. Dr. Peter E. Kloeden
Date of oral examination:	August 3, 2016
Approved for printing.	Lübeck, March 7, 2017

Zusammenfassung

In vielen Fällen sind analytische Lösungen von stochastischen partiellen Differentialgleichungen nicht explizit berechenbar, weshalb das Ziel dieser Arbeit die Entwicklung numerischer Verfahren zur Lösung dieser Gleichungen ist. Insbesondere liegt der Fokus auf ableitungsfreien Verfahren höherer Ordnung. Diese Methoden haben im Allgemeinen, verglichen mit Approximationsverfahren welche mit der Ableitung des Diffusionsoperators arbeiten, einen geringeren Rechenaufwand bei gleicher Konvergenzordnung.
In dieser Arbeit wird die Konvergenz im quadratischen Mittel der vorgestellten numerischen Verfahren bewiesen und die Ergebnisse werden durch Simulationen bestätigt. Außerdem wird der Rechenaufwand der einzelnen Verfahren analysiert. Die Minimierung des Fehlers für gegebenen Aufwand liefert die effektive Konvergenzordnung. Dieser Wert wird für die verschiedenen Verfahren bestimmt und verglichen. Es wird gezeigt, dass die ableitungsfreien Verfahren im Allgemeinen eine höhere effektive Konvergenzordnung erreichen.
Bei der Approximation von nicht-kommutativen Gleichungen müssen zusätzlich iterierte stochastische Integrale simuliert werden. Es werden Verfahren vorgestellt um diese Integrale zu approximieren und die Gleichungen zu lösen. Die theoretischen Ergebnisse werden mit numerischen Simulationen veranschaulicht. Auch für diese Klasse von stochastischen partiellen Differentialgleichungen erreichen die ableitungsfreien Verfahren in vielen Fällen eine höhere effektive Konvergenzordnung.

Summary

Analytical solutions to stochastic partial differential equations are in most cases not explicitly computable. Therefore, the goal of this work is to derive numerical schemes to solve these equations. Particularly, we focus on schemes with higher orders of convergence that are free of derivatives. These numerical methods involve, in general, less computational effort with the same high order of convergence compared to schemes that include the derivative of the diffusion operator.

The convergence of the numerical schemes is proved analytically and the computational costs of these schemes are analyzed. We derive the effective order of convergence for various schemes by minimizing the mean-square error given some fixed computational cost. In general, this number is higher for the derivative-free methods.

In the approximation of stochastic partial differential equations that are not commutative, iterated stochastic integrals have to be simulated. We present and analyze numerical schemes to complete this task. These schemes are incorporated into a derivative-free numerical method to approximate the mild solution of such equations. Moreover, the effective order of convergence of these algorithms is derived and compared to the order of established approximation methods.

The theoretical results are illustrated and confirmed with numerical simulations for both types of equation.

Acknowledgment

I want to express my honest gratitude towards my supervisor Prof. Dr. Andreas Rößler for giving me all the freedom to work in my own fashion as well as for taking the time to discuss my results anytime I needed a second opinion. The discussions with him helped a lot to form new ideas and he created a great working environment for me. Andreas Rößler owns a good share at making the last years such an exciting time.

It goes without saying that there are a few more people who I want to thank for their support. They know about their importance to me.

Moreover, I am grateful to the Graduate School for Computing in Medicine and Life Sciences, funded by Germany's Excellence Initiative, for their financial support.

Contents

Motivation

From medicine over biology and geology to finance - stochastic partial differential equations are a powerful modeling tool in numerous applications. As an example, we present the transfer of information in the human brain, that is, the modeling of the evolution of action potentials between neurons.
Let us focus on one neuron with resting membrane potential V_0 and denote by V the membrane potential of this cell, that is, the difference between the voltage inside and outside of the neuron. One model to describe changes in the membrane potential is the model by FitzHugh and Nagumo [18, 19, 52], which is a simplification of the seminal model developed by Hodgkin and Huxley [28]. We assume that the neuron can be represented by a cable of length l and radius r and restrict this representation to the one-dimensional case for simplicity. The model comprises a system of two (stochastic) partial differential equations

$$\begin{aligned}
\frac{\partial V}{\partial t} &= D_1 \frac{\partial^2 V}{\partial x^2} + \kappa V(V-a)(1-V) - \lambda Z + I(x,t,V), \\
\frac{\partial Z}{\partial t} &= D_2 \frac{\partial^2 Z}{\partial x^2} + \varepsilon(V - pZ + b), \quad t > 0,\ x \in (0,l),
\end{aligned}$$

where Z is the recovery variable responsible for obtaining the equilibrium potential after an action potential occurred. Moreover, we need to specify boundary conditions and initial values for V and Z. For the parameters, it holds $D_1, D_2, \kappa, a, \lambda, \varepsilon, p > 0$, $b \in \mathbb{R}$. Here, $I(x,t,V)$, $t > 0$, $x \in [0,l]$, is the input current which results from extern or cell intern sources. This term can be random and accounts for input from other neurons, the variability in the interspike intervals, the opening and closing of ion channels or random postsynaptic potentials, see [68]. For more details on this or similar neuronal models, we refer the reader to [11, 64, 66, 67, 68, 69].

1
Introduction

Let $T \in (0, \infty)$ and let $(\Omega, \mathcal{F}, P)$ denote a probability space endowed with some filtration $(\mathcal{F}_t)_{t\in[0,T]}$ fulfilling the usual conditions. In this work, we investigate semilinear, parabolic stochastic partial differential equations (SPDEs) on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$. These equations are of the following general form

$$\mathrm{d}X_t = \big(AX_t + F(X_t)\big)\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t, \quad t \in (0, T], \qquad X_0 = \xi. \tag{1.1}$$

Here, the solution process $(X_t)_{t\in[0,T]}$ is an element of some separable Hilbert space H_γ for some suitable $\gamma \in [0, 1)$ and $(W_t)_{t\in[0,T]}$ is a Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. More details on the operators, spaces, and processes in this equation are given in Chapter 2.

Analytical properties of stochastic differential equations in infinite dimensions have been analyzed extensively, consult, for example, [12, 13, 34, 36, 42, 47, 55, 69]. Different notions of solutions for SPDEs have been introduced and properties of the solution process, such as its regularity, have been studied, [7, 13, 34, 42, 79]. Existence and uniqueness of solutions have been proved by different means. Walsh introduced the martingale approach, see [69]; another idea is the variational approach, which is employed in [53] or [63], for example. In this work, we chose the semigroup approach to prove the existence and uniqueness of a mild solution, as described in [13] or [34].
SPDEs are, however, a complex class of differential equations barely allowing for an analytical solution. As for stochastic ordinary differential equations (SODEs) and partial differential equations (PDEs), there is thus a need for numerical schemes to approximate the solution process. Particularly, we need tools from both the fields of SODEs and PDEs in the approximation of SPDEs. Research in this direction includes [1, 2, 6, 21, 22, 25, 26, 30, 31, 44, 48, 49, 51, 70, 78],

to name only a few works.
In order to approximate the solution process of a SPDE numerically, one has to discretize the infinite dimensional stochastic process next to the time and space domain. Concerning the space domain most methods work with a spectral Galerkin method or a finite element discretization to obtain a finite dimensional system of stochastic differential equations in the projection space, see [1, 35, 41, 70, 78], for example.
So far, schemes with higher order of convergence in the temporal direction remain rare and efficient algorithms are restricted to equations under very specific assumptions, see [3, 4, 5, 20, 35, 44, 45], or [71].
In [32], Jentzen and Kloeden derived Taylor approximations for the mild solution of SPDE (1.1), which is given by

$$X_t = e^{At}\,\xi + \int_0^t e^{A(t-s)}\,F(X_s)\,\mathrm{d}s + \int_0^t e^{A(t-s)}\,B(X_s)\,\mathrm{d}W_s \qquad P\text{-a.s.}$$

for $t \in [0,T]$. These approximations allow for the schematic derivation of numerical schemes. Based upon this work, the authors proposed the exponential Euler scheme, see [31]. Furthermore, higher order schemes such as the Milstein scheme in [35] and a Wagner-Platen type scheme in [5] were developed for SPDEs driven by a Q-Wiener process with trace class that fulfill certain commutativity conditions. In [71], the authors derived a derivative-free version of the Milstein scheme based on similar assumptions. Moreover, further Milstein-type schemes were developed in [3, 4, 44]. All these higher order schemes are, however, only efficiently applicable to a specific class of SPDEs with an operator B that is pointwise multiplicative in the Q-Wiener process.

The main novelty in this work is the development of efficient derivative-free approximation schemes with computational cost of optimal order for SPDEs of type (1.1).
We start from the Milstein scheme for n-dimensional stochastic differential equations for some fixed $n \in \mathbb{N}$ to emphasize the crux. Let $(W_t)_{t\in[0,T]}$ be a k-dimensional Brownian motion with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ for some fixed $k \in \mathbb{N}$ and assume $a\colon \mathbb{R}^n \to \mathbb{R}^n$ and $b = (b_1,\ldots,b_k)\colon \mathbb{R}^n \to \mathbb{R}^{n\times k}$ with $b_j(x) = (b_{1,j}(x),\ldots,b_{n,j}(x))^T$, $j \in \{1,\ldots,k\}$, $x \in \mathbb{R}^n$, to be Lipschitz continuous functions. In this setting, the system of SODEs

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \sum_{j=1}^{k} b_j(X_t)\,\mathrm{d}W_t^j$$

for $t \in (0,T]$ with initial value $X_0 = \xi \in \mathbb{R}^n$ allows for a unique solution, [38]. For simplicity, let the step size $h = \frac{T}{M}$ for some $M \in \mathbb{N}$ be constant such that we obtain an equidistant discretization $t_m = m \cdot h$, $m \in \{0,\ldots,M\}$, of the time interval $[0,T]$.
Now, define $\Delta W_m^j = W_{t_{m+1}}^j - W_{t_m}^j$ for all $j \in \{1,\ldots,k\}$, $m \in \{0,\ldots,M-1\}$, $M \in \mathbb{N}$. It is well known that the following expression holds for the stochastic double integrals

$$\int_{t_m}^{t_{m+1}} \int_{t_m}^{s} \mathrm{d}W_u^j\,\mathrm{d}W_s^i + \int_{t_m}^{t_{m+1}} \int_{t_m}^{s} \mathrm{d}W_u^i\,\mathrm{d}W_s^j = \Delta W_m^i\,\Delta W_m^j \tag{1.2}$$

for $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $m \in \{0, \ldots, M-1\}$, $M \in \mathbb{N}$, see [38]. If we assume the SODE to be commutative, that is,

$$\sum_{r=1}^{n} b_{r,j} \frac{\partial b_{l,i}}{\partial x_r} = \sum_{r=1}^{n} b_{r,i} \frac{\partial b_{l,j}}{\partial x_r} \tag{1.3}$$

holds for all $l \in \{1, \ldots, n\}$ and $i, j \in \{1, \ldots, k\}$, the Milstein scheme can be easily simulated due to expression (1.2).

Let us focus on the commutative SODE for now. In this setting, the Milstein scheme can be written as $Y_0^M = \xi$ and

$$\begin{aligned} Y_{m+1}^M = Y_m^M + h\, a(Y_m^M) + \sum_{j=1}^{k} b_j(Y_m^M)\, \Delta W_m^j \\ + \frac{1}{2} \sum_{i,j=1}^{k} \left(\frac{\partial b_{l,i}}{\partial x_r}(Y_m^M) \right)_{1 \leq l,r \leq n} b_j(Y_m^M) \left(\Delta W_m^i \cdot \Delta W_m^j \right) \\ - \frac{h}{2} \sum_{j=1}^{k} \left(\frac{\partial b_{l,j}}{\partial x_r}(Y_m^M) \right)_{1 \leq l,r \leq n} b_j(Y_m^M) \end{aligned}$$

for $m \in \{0, \ldots, M-1\}$, $M \in \mathbb{N}$. We refer to [38] for more details. The implementation of this scheme is straightforward as no stochastic double integrals have to be simulated. The Milstein scheme achieves a strong order of convergence of 1.0, which is an improvement compared to the Euler-Maruyama scheme for SODEs with rate 0.5, [38]. The computation of the Milstein scheme is costly, however. In each time step one has to evaluate the drift and diffusion functions a, b_j, $j \in \{1, \ldots, k\}$, which are $(k+1) \cdot n$ evaluations of scalar (nonlinear) functions in each step. The same terms have to be computed in the Euler-Maruyama scheme as well. Moreover, for fixed $m \in \{0, \ldots M\}$, $M \in \mathbb{N}$, one has to calculate the Jacobian $\left(\frac{\partial b_{l,i}}{\partial x_r}(Y_m^M) \right)_{1 \leq l,r \leq n}$ for all $i \in \{1, \ldots, k\}$ in the Milstein scheme, which results in $n^2 k$ function evaluations. In total, we obtain a computational effort of $\mathcal{O}(nkM)$ for the Euler-Maruyama scheme and $\mathcal{O}(n^2 kM)$ for the Milstein scheme. The number of time steps employed to simulate one path with a prescribed accuracy and to obtain X_T with the Milstein scheme is, however, smaller than for the Euler-Maruyama scheme as it attains a higher order of convergence, in general.

In order to reduce the computational cost while keeping the strong order of convergence at the same level, Rößler introduced approximation schemes for SODEs that are free of derivatives in [60, 61, 62].

For some $N, K \in \mathbb{N}$, let $P_N : H \to H_N$ and $P_K : V \to V_K$ denote some projection operators for the spatial discretization and the approximation of the Q-Wiener process, respectively - more details are given in Section 3.3. In the setting of infinite dimensional stochastic differential equations, the computational cost of the Milstein scheme is cubic in the dimensions N and K of the projection spaces H_N and V_K; precisely, the computational effort of the scheme is of order $\mathcal{O}(N^2 K)$ in each time step. Since these dimensions need to increase to obtain a higher accuracy of the approximation, the reduction of the computational cost becomes even more important for SPDEs.

Jentzen and Röckner developed a Milstein scheme for SPDEs that are commutative in [35]; the commutativity condition for the setting of SPDEs is specified in Section 3.3. They solved the problem of high computational effort by restricting their examples to the case of an operator B that is pointwise multiplicative in the Q-Wiener process, that is, $(B(y)v)(x) = b(x, y(x)) \cdot v(x)$ for all $x \in (0,1)^d$, $y \in H = V = L^2((0,1)^d, \mathbb{R})$, $v \in V_0 \subset V$, $b\colon (0,1)^d \times \mathbb{R} \to \mathbb{R}$ and $d = 1, 2, 3$. Therewith, they avoid the evaluation of terms that result in cubic computational costs. In addition, their scheme is applicable to more general equations as well; in this case, the computational effort is of order $\mathcal{O}(N^2KM)$, however. For general commutative equations of type (1.1), the scheme reads as $Y_0^{N,K,M} = P_N\xi$ and

$$\begin{aligned} Y_{m+1}^{N,K,M} = P_N\bigg(e^{Ah}\Big(&Y_m^{N,K,M} + hF(Y_m^{N,K,M}) + B(Y_m^{N,K,M})\Delta W_m^{K,M} \\ &+ \frac{1}{2}B'(Y_m^{N,K,M})\big(B(Y_m^{N,K,M})\Delta W_m^{K,M}, \Delta W_m^{K,M}\big) \\ &- \frac{h}{2}\sum_{j\in\mathcal{J}_K}\eta_j\, B'(Y_m^{N,K,M})\big(B(Y_m^{N,K,M})\tilde{e}_j, \tilde{e}_j\big)\Big)\bigg) \end{aligned}$$

for all $m \in \{0, 1, \ldots, M-1\}$, $M, N, K \in \mathbb{N}$. Details on the notation can be found in Section 3.3.

Another approach to reduce the computational cost is the derivative-free version of the Milstein scheme for SPDEs derived by Wang and Gan in [71]. This scheme is only applicable if the operator B is pointwise multiplicative in the Q-Wiener process, the setting considered in [35] as well, and cannot be employed to solve equation (1.1) in general. The Runge-Kutta type scheme has the following general form $Y_0^{N,K,M} = P_N\xi$ and

$$\begin{aligned} Y_{m+1}^{N,K,M} = P_N\bigg(e^{Ah}\Big(&Y_m^{N,K,M} + hF(Y_m^{N,K,M}) + B(Y_m^{N,K,M})\Delta W_m^{K,M} \\ &+ \frac{1}{2}BB(Y_m^{N,K,M}, h)(\Delta W_m^{K,M}, \Delta W_m^{K,M}) - \frac{h}{2}\sum_{j\in\mathcal{J}_K}\eta_j BB(Y_m^{N,K,M}, h)(\tilde{e}_j, \tilde{e}_j)\Big)\bigg) \end{aligned}$$

for all $m \in \{0, 1, \ldots, M-1\}$, $M, N, K \in \mathbb{N}$. The bilinear operator BB has to fulfill the following assumptions. There exists a constant C, independent of $h > 0$, such that

$$\|BB(v,h) - BB(w,h)\|^2_{L_{HS}^{(2)}(V_0,H)} \leq \frac{C}{h}\|v - w\|_H^2 \tag{1.4}$$

$$\|BB(v,h) - B'(v)B(v)\|^2_{L_{HS}^{(2)}(V_0,H)} \leq Ch\left(1 + \|v\|^4_{H_\beta}\right) \tag{1.5}$$

for all $v, w \in H_\beta$ and some $\beta \in [0,1)$. Under these conditions, the strong order of the Milstein scheme is preserved.

Here, we choose an alternative way to handle the issue of high dimensionality. We devise numerical schemes which are applicable to a general class of equations of type (1.1) and free of derivatives.

Particularly, we specify an approximation operator of the term

$$\begin{aligned}&\frac{1}{2}B'(Y_m^{N,K,M})\big(B(Y_m^{N,K,M})\Delta W_m^{K,M},\Delta W_m^{K,M}\big)\\&=\frac{1}{2}\sum_{i,j\in\mathcal{J}_K}\sqrt{\eta_i}\sqrt{\eta_j}B'(Y_m^{N,K,M})\big(B(Y_m^{N,K,M})\tilde{e}_i,\tilde{e}_j\big)\Delta\beta_m^i\Delta\beta_m^j,\end{aligned}$$

which has to be computed in the Milstein scheme for some $N, M, K \in \mathbb{N}$ and all $m \in \{0,\dots,M\}$; this approximation can be obtain with reduced computational cost. The idea is based on the work by Rößler for finite dimensional SODEs, see [60, 61, 62], for example, and allows to lower the computational cost to the effort involved in the Euler-Maruyama scheme, that is, $\mathcal{O}(nkM)$. In the setting of SPDEs, this reduction in the computational cost (CC) is even more powerful as the effective order of convergence can be improved.
The effective order of convergence is defined as the rate that we obtain by solving the optimization problem

$$\min_{N,M,K}\Big(\sup_{m\in\{0,\dots,M\}}\mathrm{E}\Big[\big\|X_{t_m}-Y_m^{M,N,K}\big\|_H^2\Big]\Big)^{\frac{1}{2}}\quad\text{such that}\quad \mathrm{CC}=\bar{c}$$

for some $\bar{c} > 0$. This concept is introduced in Section 3.2.
The formulation of the approximation method as well as its proof can, however, not be transferred from the finite dimensional setting directly. The number of independent Brownian motions $K \in \mathbb{N}$ has to increase to obtain a higher accuracy in the approximation of SPDEs, but the constant in the error estimate for SODEs is not independent of K. Therefore, a different approach is needed in order to prove convergence. We elaborate on this issue in Chapter 3.
Commutative SPDEs allow for an expression of stochastic double integrals of the form

$$\int_s^t B'(X_s)\Big(\int_s^r B(X_s)\,\mathrm{d}W_u^K\Big)\,\mathrm{d}W_r^K,\quad s,t\in[0,T],\ s\le t,\ K\in\mathbb{N},$$

in terms of increments of the Q-Wiener process; thus their simulation is straightforward. This simplification is not possible if the commutativity does not hold. In Section 4.1, we propose two algorithms to approximate the stochastic double integrals in this case. These approximation methods are based on the schemes by Wiktorsson, [39], and Kloeden, Platen, and Wright, [75], developed for finite dimensional SODEs. We transfer the methods to the setting of SPDEs and obtain error estimates which differ from the expressions derived for SODEs.

This work is composed as follows. In Chapter 2, we lay the foundation to work with stochastic partial differential equations. We explain and illustrate the equation of interest and present stochastic calculus in Hilbert spaces. Furthermore, we elaborate on the theory of existence and uniqueness of a mild solution and some important properties of this process to some extent. The focus of this chapter is on stating results that are crucial for the numerical analysis later on. In the main part of this work, Chapter 3 and Chapter 4, we develop derivative-free schemes to solve equation (1.1) efficiently. We devise different schemes depending on the assumption of commutativity. These schemes solve the issue of high computational costs related to a higher order of convergence, as described above. We introduce the concept of the effective order of convergence for comparability of the schemes. Moreover, we prove the convergence of these methods theoretically and analyze their computational effort.

Chapter 3 is concerned about commutative equations. We design a derivative-free Milstein scheme and prove its strong convergence; this result is stated in Theorem 3.1. Besides, we analyze the effective order of convergence for various schemes and compare these values for different types of equations. A summary can be found in Table 3.2. Finally, we illustrate our findings with numerical simulations.

In Chapter 4, we turn our attention to SPDEs of type (1.1) that are not commutative. We introduce two algorithms to approximate the iterated stochastic integrals and prove convergence results as specified in Theorem 4.1 and Theorem 4.2, respectively. These schemes are incorporated in a numerical method which is free of derivatives and attains a higher effective order of convergence than the exponential Euler scheme for a large number of equations. The results on the convergence are stated in Theorem 4.3 and a comparison of the schemes can be found in Table 4.1.

Eventually, we state some ideas on the connection of local and global errors and close with a discussion of our results.

2

Stochastic Differential Equations in Infinite Dimensions

There are different ways to look at a stochastic partial differential equation - on the one hand, we can describe it as a stochastic differential equation in infinite dimensions, on the other, it defines an evolution equation where a stochastic term is added. It is therefore obvious, that the study of stochastic differential equations in infinite dimensions employs tools from various disciplines. We need to introduce some definitions and theorems from the theory of semigroups; this is necessary due to the notion of a solution, the mild solution, that we focus on. We prove the existence and uniqueness of the solution, similar as for partial differential equations (PDEs), by means of semigroups. Next to ideas from the analysis of PDEs and some definitions from functional analysis, we introduce stochastic analysis in Hilbert spaces, as the solution to SPDE (1.1) is a Hilbert-space valued stochastic process.

First, we present the general setting in which the stochastic evolution equations are analyzed in the following. Let $(H, \langle\cdot,\cdot\rangle_H)$ and $(V, \langle\cdot,\cdot\rangle_V)$ be separable real-valued Hilbert spaces. In particular, these spaces have countable orthonormal bases; $\{e_i,\ i \in \mathcal{I}\}$ denotes the basis of H and $\{\tilde{e}_j,\ j \in \mathcal{J}\}$ the basis of V, where $\mathcal{I}, \mathcal{J}$ are countable index sets. More details on these bases are given in Section 2.1 and Section 2.2.
Let $T \in (0,\infty)$ and define a probability space $(\Omega, \mathcal{F}, P)$ with filtration $(\mathcal{F}_t)_{t\in[0,T]}$. We assume $(\mathcal{F}_t)_{t\in[0,T]}$ to be right-continuous and that $\mathcal{F}_0$ contains all sets $A \in \mathcal{F}$ with measure zero.
We are interested in solving semilinear, parabolic equations on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ of the general form

$$\mathrm{d}X_t = (AX_t + F(X_t))\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t, \quad t \in (0,T], \qquad X_0 = \xi, \tag{2.1}$$

where the solution process belongs to a Hilbert space H_γ, $\gamma \in [0,1)$, and $(W_t)_{t\in[0,T]} \in V$ is a Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. Details and properties of the operators, processes, and spaces will be given in Section 2.1 to Section 2.3.

We want to emphasize that the following sections are not supposed to be complete but simply aim to provide the theory and tools employed in the main part of this work.

2.1 Semigroups

As we make use of the theory of semigroups in the proof of existence and uniqueness of a mild solution to SPDE (2.1), we need to collect some facts on semigroups on Hilbert spaces. We prove estimates that are essential in the analyses that we conduct later on. The definitions, theorems, and proofs in this part are taken from the fundamental works [54], [57], and [65].

First, we look at a deterministic Cauchy problem, [57].

Example 2.1
Consider

$$\begin{aligned} \mathrm{d}X(t) &= AX(t)\,\mathrm{d}t, \quad t \in (0,T], \\ X(0) &= x_0, \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$, $x_0 \in \mathbb{R}^n$, *and* $n \in \mathbb{N}$.
The solution to this equation is known to equal

$$X(t) = x_0 e^{At}.$$

In this example, the operator A is bounded and the definition of e^{At}, $t \in [0,T]$, *is straightforward.*

In the case of evolution equations in infinite dimensions, the linear operator A does, however, not have to be bounded; consider the Laplace operator $A = \Delta$, for example. Thus, the questions we need to answer are under which conditions on A the operator e^{At}, $t \in [0,T]$, is defined and how e^{At}, $t \in [0,T]$, can be represented in this case. It will turn out that we need A to generate some type of semigroup - an analytic semigroup in our setting - for a reasonable definition of this expression. We start with semigroups of bounded linear operators and move step by step to the definition of analytic semigroups and its properties.

Definition 2.1 (Semigroup of Bounded Linear Operators and C_0-Semigroup)
$S(t)$ is defined to be a semigroup of bounded linear operators on H if $S(t) \in L(H)$, for all $t > 0$, and

- $S(0) = I$
- $S(t)S(s) = S(t+s), \quad s,t \in [0,\infty).$

We call $S(t)$, $t \geq 0$, a C_0-semigroup if $S(t)$, $t \geq 0$, is a semigroup of bounded linear operators on H and

$$\lim_{t \to 0+} S(t)w = w \qquad \textit{for every } w \in H.$$

[54, Chapter 1, Definitions 1.1,2.1]

In the study of evolution equations we are, however, given the linear operator A instead of the semigroup $S(t)$, $t \geq 0$. Therefore, we are interested in the relation of the operator A and the corresponding semigroup as well as the properties that this operator has to fulfill such that $S(t)$, $t \geq 0$, is a C_0- or analytic semigroup, respectively.

Definition 2.2 (Infinitesimal Generator)
Let $S(t)$, $t \geq 0$, be a C_0-semigroup on H. The infinitesimal generator A of $S(t)$, $t \geq 0$, is defined by

$$Aw := \lim_{h \to 0^+} \frac{S(h) - I}{h} w = \left. \frac{\mathrm{d}^+ (S(t))w}{\mathrm{d}t} \right|_{t=0}$$

for $w \in D(A)$. The domain $D(A)$, in turn, is defined as

$$D(A) := \left\{ w \in H : \lim_{h \to 0^+} \frac{S(h) - I}{h} w \textit{ exists} \right\}.$$

[54, p.1]

The following assumptions on the operator A, stated in the Theorem by Hille-Yosida, guarantee that A generates a C_0-semigroup.

Theorem 2.1 (Hille-Yosida)
A linear operator A is the infinitesimal generator of a C_0-semigroup with $\|S(t)\|_{L(H)} \leq 1$ for all $t \geq 0$ if and only if

i) A is a closed operator and $\overline{D(A)} = H$

ii) $\rho(A) := \{\lambda \in \mathbb{C} : (\lambda I - A) \textit{ is invertible}\}$, called the resolvent set of A, contains $\mathbb{R}^+$ and

$$\|R_\lambda(A)\|_{L(H)} = \|(\lambda I - A)^{-1}\|_{L(H)} \leq \frac{1}{\lambda}$$

for all $\lambda > 0$.

[54, Chapter 1, Theorem 3.1]

Proof. The proof of this important theorem can be found in [54, p.8,9]. ☐

In the analysis of the existence and uniqueness of a solution to SPDE (2.1) and the investigation of the numerical schemes in Chapter 3 and Chapter 4, we need the operator A not only to generate a C_0-semigroup but an analytic semigroup instead.

Definition 2.3 (Analytic Semigroup)
An analytic semigroup $S(t)$, $t \geq 0$, is a C_0-semigroup which fulfills the following additional requirements

- $S(t) \in L(H)$ *can be extended to* $t \in \Delta_\phi = \{0\} \cup \{t \in \mathbb{C} |\, |\arg t| < \phi\}$ *for some* $\phi \in (0, \frac{\pi}{2})$ *and Definition 2.1 holds for all* $t \in \Delta_\phi$,
- $S(t)$ *is analytic in* t *for* $t \in \Delta_\phi \setminus \{0\}$ *(in the uniform operator topology).*

[57, Definition 11.30]

As above, we are interested in the properties of the operator A that guarantee that the corresponding semigroup is analytic.

Theorem 2.2 (Generator of Analytic Semigroup)
Let A be a linear, closed, and densely defined operator in H. A is the generator of an analytic semigroup if and only if there exists some $w \in \mathbb{R}$ such that $\{\lambda : Re\, \lambda > w\} \subset \rho(A)$ and it exists a constant C such that

$$\|R_\lambda(A)\|_{L(H)} \leq \frac{C}{\lambda - w}$$

for $Re\, \lambda > w$.
Under these assumptions, it holds $\{\lambda : |\arg(\lambda - w)| < \frac{\pi}{2} + \delta\} \subset \rho(A)$ for some $\delta > 0$ and the semigroup can be represented as

$$e^{At} = \frac{1}{2\pi i} \int_\Gamma e^{\lambda t} (\lambda I - A)^{-1} \,\mathrm{d}\lambda,$$

where Γ is a curve from $e^{-i\varphi}\infty$ to $e^{i\varphi}\infty$ such that $\Gamma \subset \{|\arg(\lambda - w)| \leq \varphi\}$ for $\frac{\pi}{2} < \varphi < \frac{\pi}{2} + \delta$.

[57, Theorem 11.31]

Proof. For a proof, we refer to [57, p. 412-414]. □

Moreover, A generates an analytic semigroup if we assume $(-A)$ to be self adjoint and bounded below with compact resolvent, [65, Theorem 32.1].

Next, we want to list and illustrate some basic features of analytic semigroups. These are important in the analysis of the solvability of equation (2.1) and later on to prove the convergence of the numerical schemes. First, we define the fractional power of the generator A of an analytic semigroup. Therefore, we need to define the spectrum of the operator A

$$\sigma(A) := \mathbb{C} \setminus \rho(A),$$

see [57, Definition 7.39] for details.

Definition 2.4 (Fractional Power)
Assume A to be the generator of an analytic semigroup and that the spectrum of A lies in the open left half-plane. We define

$$(-A)^{-\alpha} = -\frac{1}{2\pi i}\int_\Gamma \lambda^{-\alpha}(\lambda I + A)^{-1}\,\mathrm{d}\lambda, \tag{2.2}$$

where Γ is again a curve from $e^{-i\varphi}\infty$ to $e^{i\varphi}\infty$, with $\frac{\pi}{2}-\delta<\varphi<\pi$ and δ as specified in Theorem 2.2, such that the origin lies to the left of Γ and the spectrum of $-A$ to its right. $\lambda^{-\alpha}$ is taken to be positive on the positive real axis. The integral converges in the uniform operator topology for every $\alpha>0$.

[57, p. 415]

For $\alpha>0$ the integral in (2.2) can be expressed as

$$(-A)^{-\alpha} = \frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}e^{At}\,\mathrm{d}t.$$

The simple reformulations can be found in [54] or [57].
Now, we state some estimates which are of importance in our analysis in various steps. We give a proof of these estimates as they are essential in this work.

Theorem 2.3 (Estimates of Analytic Semigroups)
Let A be the generator of an analytic semigroup $S(t)$, $t\geq 0$, with spectrum that lies entirely in the open left half-plane. Then, it holds for some $C>0$, $\delta>0$, and all $t\geq 0$

a) $\|S(t)\|_{L(H)} \leq Ce^{-\delta t}$,

b) $\|(-A)^{-\alpha}\|_{L(H)} \leq C$, $0\leq\alpha\leq 1$,

c) $\|(-A)^{\alpha}e^{At}\|_{L(H)} \leq C_\alpha t^{-\alpha}$, $\alpha\geq 0$, $t>0$,

d) $\|(-A)^{-\alpha}(e^{At}-I)\|_{L(H)} \leq C_\alpha t^{\alpha}$, $0<\alpha\leq 1$.

[54, Chapter 2, Lemma 6.3,6.13]

Proof. The proofs are mainly taken from [54] and [65].

a) For C_0-semigroups, it holds $\|S(t)\|_{L(H)} \leq Ce^{wt}$ with constants $C\geq 1$, $w\geq 0$, [54, Chapter 1, Theorem 2.2]. As the spectrum of A lies in the open left half-plane, this estimate is true for $w=0$. Moreover, we can choose some $\delta>0$ such that $A+\delta$ is again the generator of an analytic semigroup with

$$\|e^{(A+\delta)t}\|_{L(H)} \leq C \Rightarrow \|e^{At}\|_{L(H)} \leq Ce^{-\delta t}$$

for all $t\geq 0$.

b) Let $\alpha \in (0,1)$, with part a) it holds

$$\begin{aligned}\|(-A)^{-\alpha}\|_{L(H)} &= \Big\|\frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}e^{At}\,\mathrm{d}t\Big\|_{L(H)}\\ &\le \frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}\|e^{At}\|_{L(H)}\,\mathrm{d}t \le \frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}Ce^{-\delta t}\,\mathrm{d}t\\ &= \frac{C}{\Gamma(\alpha)}\int_0^\infty \Big(\frac{u}{\delta}\Big)^{\alpha-1}e^{-u}\frac{1}{\delta}\,\mathrm{d}u = \frac{C\Gamma(\alpha)}{\Gamma(\alpha)\delta^{\alpha-2}} = C.\end{aligned}$$

We get $\|(-A)^{-1}\|_{L(H)} \le C$ by similar computations.

c) First, note that $(-A)^\alpha$ is a closed and densely defined operator for $\alpha \ge 0$, [54, Chapter 2, Theorem 6.8]. So, $(-A)^\alpha S(t)$, $t>0$, is closed and everywhere defined and the Closed Graph Theorem, [73, Theorem IV.4.5], implies $(-A)^\alpha S(t) \in L(H)$ for all $t>0$, $\alpha \ge 0$.

Let $k > \alpha > k-1$, $k \in \mathbb{N}$, $t > 0$, we obtain

$$\begin{aligned}\|(-A)^\alpha e^{At}\|_{L(H)} &= \|(-A)^{\alpha-k}A^k e^{At}\|_{L(H)}\\ &= \Big\|\frac{1}{\Gamma(k-\alpha)}\int_0^\infty s^{k-\alpha-1}A^k e^{A(t+s)}\,\mathrm{d}s\Big\|_{L(H)}\\ &\le \frac{1}{\Gamma(k-\alpha)}\int_0^\infty s^{k-\alpha-1}\|A^k e^{A(t+s)}\|_{L(H)}\,\mathrm{d}s.\end{aligned}$$

We compute an estimate of $\|A^k e^{At}\|_{L(H)}$, $t>0$, first. For now, let $k=1$. $S(t)$, $t>0$, is differentiable, see [54, Chapter 2, Theorem 5.2], and it holds

$$AS(t) = \frac{\mathrm{d}}{\mathrm{d}t}S(t) = \frac{1}{2\pi i}\int_\Gamma \lambda e^{\lambda t}(\lambda I - A)^{-1}\,\mathrm{d}\lambda.$$

Since $S(t)$, $t \ge 0$, is an analytic semigroup, we can shift Γ to the rays $\rho e^{i\varphi}$ and $\rho e^{-i\varphi}$ with $\rho \in (0,\infty)$ and φ as given in Theorem 2.2, [54, p.63]. Therewith, and as $\|(\lambda I - A)^{-1}\|_{L(H)} \le \frac{C}{\lambda}$ for $Re\,\lambda > 0$, we obtain

$$\|AS(t)\|_{L(H)} \le \frac{C}{\pi t|\cos(\varphi)|} = Ct^{-1}.$$

Now, we easily get for arbitrary $k \in \mathbb{N}$

$$\big\|A^k e^{At}\big\|_{L(H)} = \big\|\big(Ae^{A\frac{t}{k}}\big)^k\big\|_{L(H)} \le \Big(C\Big(\frac{t}{k}\Big)^{-1}\Big)^k = C_k t^{-k}.$$

With the substitution $r = \frac{s}{t}$, it follows for all $t>0$

$$\begin{aligned}\|(-A)^\alpha e^{At}\|_{L(H)} &\le \frac{1}{\Gamma(k-\alpha)}\int_0^\infty s^{k-\alpha-1}C_k(t+s)^{-k}\,\mathrm{d}s\\ &= \frac{C_k}{\Gamma(k-\alpha)}\int_0^\infty (rt)^{k-\alpha-1}t^{-k}(1+r)^{-k}t\,\mathrm{d}r\\ &= \frac{C_k}{\Gamma(k-\alpha)t^\alpha}\int_0^\infty r^{k-\alpha-1}(1+r)^{-k}\,\mathrm{d}r\end{aligned}$$

$$= \frac{C_k}{\Gamma(k-\alpha)t^\alpha}\Big(\int_0^1 \frac{r^{k-\alpha-1}}{(1+r)^k}\,\mathrm{d}r + \int_1^\infty \frac{r^{k-\alpha-1}}{(1+r)^k}\,\mathrm{d}r\Big).$$

As $k > \alpha > k-1$, we obtain for all $t > 0$

$$\begin{aligned}\|(-A)^\alpha e^{At}\|_{L(H)} &\le \frac{C_k}{\Gamma(k-\alpha)t^\alpha}\Big(\int_0^1 r^{k-\alpha-1}\,\mathrm{d}r + \int_1^\infty \frac{r^{k-\alpha-1}}{r^k}\,\mathrm{d}r\Big)\\ &\le \frac{C_k}{\Gamma(k-\alpha)t^\alpha}\Big(\frac{1}{k-\alpha}+\frac{1}{\alpha}\Big) = C_\alpha t^{-\alpha}.\end{aligned}$$

d) We estimate this term with the help of c), [54, Chapter 1, Theorem 2.4], and [54, Chapter 2, Theorem 6.13(b)]. For $w \in D\big((-A)^\alpha\big)$ and $0 < \alpha \le 1$, it holds

$$\begin{aligned}\|e^{At}w - w\|_H &= \Big\|\int_0^t Ae^{As}w\,\mathrm{d}s\Big\|_H = \Big\|\int_0^t (-A)^{1-\alpha}e^{As}(-A)^\alpha w\,\mathrm{d}s\Big\|_H\\ &\le \int_0^t \|(-A)^{1-\alpha}e^{As}\|_{L(H)}\|(-A)^\alpha w\|_H\,\mathrm{d}s \le C\int_0^t s^{\alpha-1}\|(-A)^\alpha w\|_H\,\mathrm{d}s\\ &= C_\alpha t^\alpha\|(-A)^\alpha w\|_H.\end{aligned}$$

□

By means of the fractional power of the operator $(-A)$, we define the solution spaces of SPDE (2.1). Let $(-A)$ be self-adjoint and positive with compact resolvent, then A generates an analytic semigroup and there exists an orthonormal basis of H consisting of eigenfunctions $\{e_i,\ i \in \mathbb{N}\}$ of $(-A)$ with eigenvalues λ_i, $i \in \mathbb{N}$, such that $\inf_{i\in\mathbb{N}} \lambda_i > 0$ and $\lim_{i\to\infty} \lambda_i = \infty$, [65, p.66]. Following [65], we define the interpolation spaces H_r, $r \in [0,\infty)$, as $H_r := D((-A)^r) \subset H$ with norm $\|x\|_{H_r} = \|(-A)^r x\|_H$ for all $x \in H_r$, $r \in [0,\infty)$. Let the inner product be defined as $\langle x, y\rangle_r := \sum_{i\in\mathbb{N}} \lambda_i^{2r} x_i y_i$ for $x, y \in H_r$; then, the spaces $(H_r, \langle\cdot,\cdot\rangle_r)$, $r \in [0,\infty)$, are separable Hilbert spaces as well. Moreover, the relation $H_s \subset H_p$ for $p < s$, $p, s \in [0,\infty)$ holds.

2.2 Stochastic Processes in Hilbert Spaces

We turn our attention to the analysis of stochastic processes in Hilbert spaces now. We clarify what we understand by $(W_t)_{t\in[0,T]}$ and explain in what sense the integral version of (2.1) is to be understood. The results and definitions in this section can be found in [13] and [55].

The first step towards understanding a Hilbert space-valued Brownian motion is to define a Gaussian measure on that Hilbert space.

Definition 2.5 (Gaussian Measure on Hilbert Space)
A probability measure μ on $(V, \mathcal{B}(V))$ is called Gaussian if for all $v \in V$ there exist $a \in \mathbb{R}$ and $\sigma \ge 0$ such that

$$\mu(\{u \in V : \langle v, u\rangle_V \in B\}) = N_{a,\sigma}(B) \quad \textit{for all } B \in \mathcal{B}(\mathbb{R}).$$

If μ is Gaussian, there exist elements $m \in V$ and $Q \in L(V)$ such that

$$\langle m, u\rangle_V = \int_V \langle u, x\rangle_V \, \mu(\mathrm{d}x) \quad \forall u \in V,$$

and

$$\langle Qu, v\rangle_V = \int_V \langle u, x-m\rangle_V \langle v, x-m\rangle_V \, \mu(\mathrm{d}x) \quad \forall u, v \in V.$$

The characteristic function of μ is

$$\varphi(u) = \int_V e^{i\langle u,x\rangle_V} \, \mu(\mathrm{d}x) = e^{i\langle u,m\rangle_V - \frac{1}{2}\langle Qu,u\rangle_V}$$

for $u \in V$.
We call m the mean and Q the covariance operator of μ and denote the Gaussian measure μ by $N_{m,Q}$.

[13, p.46-48]

With this definition it is straightforward to define a Brownian motion which takes values in a Hilbert space. Let $Q \in L(V)$ be symmetric and nonnegative; this follows from Definition 2.5 if Q is the covariance operator of a Gaussian measure, see [13, p.47,48]. Moreover, we restrict our analysis to a class of operators Q with finite trace, that is,

$$\operatorname{tr} Q := \sum_{j\in\mathbb{N}} \langle Qg_j, g_j\rangle_V < \infty, \tag{2.3}$$

where $\{g_j,\ j \in \mathbb{N}\}$ is an arbitrary orthonormal basis of V.
Under these assumptions on Q, there exists an orthonormal basis $\{\tilde{e}_j,\ j \in \mathbb{N}\}$ of V, consisting of eigenvectors $\tilde{e}_j,\ j \in \mathbb{N}$, of Q such that

$$Q\tilde{e}_j = \eta_j \tilde{e}_j \tag{2.4}$$

for all $j \in \mathbb{N}$. Here, $\eta_j,\ j \in \mathbb{N}$, denote the corresponding eigenvalues and it holds $\eta_j \geq 0$ for all $j \in \mathbb{N}$ and $\eta_j \to 0$ for $j \to \infty$. This follows from the Hilbert-Schmidt Theorem, [57, Theorem 7.94], as the operator Q is compact, see [56, Thereom VI.21]. We fix this basis of V.
With this notation, we directly obtain

$$\operatorname{tr} Q = \sum_{j\in\mathbb{N}} \eta_j < \infty.$$

Next, we define the stochastic process $(W_t)_{t\in[0,T]}$ in SPDE (2.1).

Definition 2.6 (Q-Wiener Process)
A V-valued stochastic process $(W_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ is called a standard Q-Wiener process if

- *$W_0 = 0$ P-a.s.,*

- *the paths $t \mapsto W_t$ are P-a.s. continuous,*
- *the increments of $(W_t)_{t\in[0,T]}$ are independent, that is, the random variables*

$$W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}}$$

  are independent for all $0 \le t_1 < \dots < t_m \le T$, $m \in \mathbb{N}$,
- *the increments follow a Gaussian law*

$$W_t - W_s \sim N_{0,(t-s)Q}$$

  for all $0 \le s \le t \le T$.

[55, Definition 2.1.9]

We call $(W_t)_{t\in[0,T]}$ a Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ if W_t is $\mathcal{F}_t$-measurable and $W_t - W_s$ is independent of $\mathcal{F}_s$ for all $s, t \in [0,t]$, $s \le t$, [55, Definition 2.1.12].
An important connection between a Q-Wiener process and real-valued Brownian motions is illustrated in the following theorem; the Q-Wiener process can be represented by an infinite sum of real-valued Brownian motions. As this connection is crucial in the numerical simulations later on, we detail its proof here.

Theorem 2.4 (Representation by Real-valued Brownian Motions)
Let $\{\tilde{e}_j,\ j \in \mathbb{N}\}$ be an orthonormal basis of V consisting of eigenvectors of Q with corresponding eigenvalues η_j, $j \in \mathbb{N}$. Then a V-valued stochastic process $(W_t)_{t\in[0,T]}$ is a Q-Wiener process if and only if

$$W_t = \sum_{j\in\mathbb{N}} \sqrt{\eta_j}\beta_t^j \tilde{e}_j, \quad t \in [0,T], \tag{2.5}$$

where $(\beta_t^j)_{t\in[0,T]}$, $j \in \{n \in \mathbb{N} |\ \eta_n > 0\}$, are independent real-valued Brownian motions on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$. The series even converges in $L^2(\Omega, \mathcal{F}, P; C([0,T], V))$, and thus always has a P-a.s. continuous modification. In particular, for any Q as above there exists a Q-Wiener process on V.

[55, Proposition 2.1.10]

Proof. In the following we give, and detail, the proof from [55].
" $\Rightarrow$ "
Let $(W_t)_{t\in[0,T]}$ be a Q-Wiener process; this process can be expressed as $W_t = \sum_{j\in\mathbb{N}} \langle W_t, \tilde{e}_j\rangle_V \tilde{e}_j$ where $\langle W_t, \tilde{e}_j\rangle_V \sim N(0, \eta_j t)$ for all $t \in [0,T]$. Now, define for $j \in \mathbb{N}$, $t \in [0,T]$

$$\beta_t^j := \begin{cases} \frac{\langle W_t, \tilde{e}_j\rangle_V}{\sqrt{\eta_j}}, & \eta_j > 0 \\ 0, & else. \end{cases}$$

Then, we obtain $W_t = \sum_{j\in\mathbb{N}} \sqrt{\eta_j}\beta_t^j \tilde{e}_j$ and $\beta_t^j \sim N(0,t)$ for all $j \in \mathbb{N}$, $t \in [0,T]$.

First, we show that for fixed $t \in [0,T]$ the random variables $(\beta_t^j)_{j\in\mathbb{N}}$ are all independent. As $\beta_t^j \sim N(0,t)$ for all $j \in \mathbb{N}$, we only have to prove that β_t^j and β_t^k are uncorrelated for all $j,k \in \mathbb{N}$, $j \neq k$. This can easily be seen as

$$\begin{aligned} \mathrm{E}\big[\beta_t^j \beta_t^k\big] &= \mathrm{E}\left[\frac{\langle W_t, \tilde{e}_k\rangle_V \langle W_t, \tilde{e}_j\rangle_V}{\sqrt{\eta_j\eta_k}}\right] = \frac{1}{\sqrt{\eta_j\eta_k}}\langle Q\tilde{e}_k, \tilde{e}_j\rangle_V \\ &= \frac{\sqrt{\eta_k}}{\sqrt{\eta_j}}\langle \tilde{e}_k, \tilde{e}_j\rangle_V = 0 \end{aligned}$$

for $j \neq k$.
Now, we prove that $(\beta_t^j)_{t\in[0,T]}$ is a Brownian motion for any $j \in \mathbb{N}$. For this purpose, we define a partition of $[0,T]$ as $0 = t_0 < t_1 < \ldots < t_m \leq T$, $m \in \mathbb{N}$. Since

$$\beta_{t_k}^j - \beta_{t_{k-1}}^j = \begin{cases} \frac{\langle W_{t_k} - W_{t_{k-1}}, \tilde{e}_j\rangle_V}{\sqrt{\eta_j}}, & \eta_j > 0 \\ 0, & else, \end{cases}$$

the increments $\beta_{t_k}^j - \beta_{t_{k-1}}^j$ and $\beta_{t_i}^j - \beta_{t_{i-1}}^j$ are independent for $i,k \in \{1,\ldots,m\}$, $i \neq k$, $j \in \mathbb{N}$. Furthermore, $\beta_t^j - \beta_s^j \sim N(0, t-s)$ for all $j \in \mathbb{N}$ and $0 \leq s < t$.

It remains to show that $\sigma(\beta_{t_1}^{j_1}, \ldots, \beta_{t_m}^{j_1}), \ldots, \sigma(\beta_{t_1}^{j_n}, \ldots, \beta_{t_m}^{j_n})$ are independent for any $j_1, \ldots, j_n \in \mathbb{N}$, $n \in \mathbb{N}$. This is proved by induction. For $m = 1$, it is obvious. Let us assume that the statement holds for some $m \in \mathbb{N}$ and let $B_{ij} \in \mathcal{B}(\mathbb{R})$, $i \in \{1,\ldots,n\}$, $j \in \{1,\ldots,m+1\}$.
It holds

$$\begin{aligned} &P\Big(\bigcap_{i=1}^n \{\beta_{t_1}^{j_i} \in B_{i1}, \ldots, \beta_{t_m}^{j_i} \in B_{im}, \beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i} \in B_{im+1}\}\Big) \\ &= P\Big(\bigcap_{i=1}^n \bigcap_{k=1}^m \{\beta_{t_k}^{j_i} \in B_{ik}\} \cap \bigcap_{i=1}^n \{\beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i} \in B_{im+1}\}\Big). \end{aligned}$$

Since $\sigma(W_s, s \leq t_m)$ and $\sigma(W_{t_{m+1}} - W_{t_m})$ are independent, $(\beta_t^j)_{j\in\mathbb{N}}$ are independent for fixed $t \in [0,T]$, and by the induction hypothesis, we obtain

$$\begin{aligned} &P\Big(\bigcap_{i=1}^n \{\beta_{t_1}^{j_i} \in B_{i1}, \ldots, \beta_{t_m}^{j_i} \in B_{im}, \beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i} \in B_{im+1}\}\Big) \\ &= P\Big(\bigcap_{i=1}^n \bigcap_{k=1}^m \{\beta_{t_k}^{j_i} \in B_{ik}\}\Big) \cdot P\Big(\bigcap_{i=1}^n \{\beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i} \in B_{im+1}\}\Big) \\ &= \prod_{i=1}^n P\Big(\bigcap_{k=1}^m \{\beta_{t_k}^{j_i} \in B_{ik}\}\Big) \cdot \prod_{i=1}^n P\Big(\{\beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i} \in B_{im+1}\}\Big) \\ &= \prod_{i=1}^n P\Big(\bigcap_{k=1}^m \{\beta_{t_k}^{j_i} \in B_{ik}\} \cap \{\beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i} \in B_{im+1}\}\Big). \end{aligned}$$

This proves the independence of $\sigma(\beta_{t_1}^{j_1}, \ldots, \beta_{t_m}^{j_1}), \ldots, \sigma(\beta_{t_1}^{j_n}, \ldots, \beta_{t_m}^{j_n})$ for any $j_1, \ldots, j_n \in \mathbb{N}$, $n \in \mathbb{N}$ as $\sigma(\beta_{t_1}^{j_i}, \ldots, \beta_{t_m}^{j_i}, \beta_{t_{m+1}}^{j_i}) = \sigma(\beta_{t_1}^{j_i}, \ldots, \beta_{t_m}^{j_i}, \beta_{t_{m+1}}^{j_i} - \beta_{t_m}^{j_i})$ for any $j_i \in \{j_1, \ldots, j_n\}$.

" $\Leftarrow$ "
Let $W_t = \sum_{j\in\mathbb{N}} \sqrt{\eta_j}\beta_t^j \tilde{e}_j$, $t \in [0,T]$ be given; the series is obviously well defined in $L^2(\Omega, \mathcal{F}, P; V)$. We have to show that this process fulfills the properties in Definition 2.6. We only prove that the increments $W_t - W_s$, $0 \le s < t$, are normally distributed with mean 0 and covariance $Q(t-s)$, as the other attributes follow directly from the series representation.
Fix some $n \in \mathbb{N}$; we know that $\langle \sum_{j=1}^n \sqrt{\eta_j}\beta_t^j \tilde{e}_j, v\rangle_V = \sum_{j=1}^n \sqrt{\eta_j}\beta_t^j \langle \tilde{e}_j, v\rangle_V$ is normally distributed for all $v \in V$. The sequence converges in $L^2(\Omega, \mathcal{F}, P; V)$ and with the help of the characteristic function it can be easily shown that the limit is normally distributed as well. Furthermore, we find $\mathrm{E}\big[\langle W_t, v\rangle_V\big] = 0$ and

$$\begin{aligned}\mathrm{E}\big[\langle W_t, u\rangle_V \langle W_t, v\rangle_V\big] &= \lim_{n\to\infty} \mathrm{E}\Big[\Big\langle \sum_{j=1}^n \sqrt{\eta_j}\beta_t^j \tilde{e}_j, u\Big\rangle_V \Big\langle \sum_{j=1}^n \sqrt{\eta_j}\beta_t^j \tilde{e}_j, v\Big\rangle_V\Big] \\ &= \sum_{j\in\mathbb{N}} \eta_j t \langle \tilde{e}_j, u\rangle_V \langle \tilde{e}_j, v\rangle_V = \sum_{j\in\mathbb{N}} t\langle Q\tilde{e}_j, u\rangle_V \langle \tilde{e}_j, v\rangle_V \\ &= t\langle Qu, v\rangle_V\end{aligned}$$

for all $u, v \in V$, $t \in [0,T]$.

It remains to show that the series converges in $L^2(\Omega, \mathcal{F}, P; C([0,T], V))$. For any $n \in \mathbb{N}$, it holds by Doob's maximal inequality, [13, Theorem 3.9],

$$\mathrm{E}\Big[\sup_{t\in[0,T]} \Big\| \sum_{j=1}^n \sqrt{\eta_j}\beta_t^j \tilde{e}_j \Big\|_V^2\Big] \le \sum_{j=1}^n \eta_j \mathrm{E}\Big[\sup_{t\in[0,T]} (\beta_t^j)^2\Big] \le C \sum_{j=1}^n \eta_j.$$

Since Q is a trace class operator, $\sum_{j=1}^\infty \eta_j < \infty$ follows. This proves that the series converges in $L^2(\Omega, \mathcal{F}, P; C([0,T], V))$.

$\square$

As for finite dimensional stochastic differential equations, equation (2.1) is only a formal expression and is to be understood as an integral equation in fact. We refrain from deriving the concept of stochastic integration in Hilbert spaces in depth, as this theory is well studied and not the focus of this work, and simply give the definition of the stochastic integral. For details as well as the theory of integration with respect to a broader class of stochastic processes, we refer to the comprehensive texts [13] and [55], from which the following representation is taken.

Primarily, we introduce the space of Hilbert-Schmidt operators mapping from V to H, denoted as $L_{HS}(V, H)$. We call a bounded linear operator $B : V \to H$ Hilbert-Schmidt if

$$\|B\|_{L_{HS}(V,H)} := \Big(\sum_{j\in\mathbb{N}} \|B\tilde{e}_j\|_H^2\Big)^{\frac{1}{2}} < \infty.$$

Note that this definition holds for any orthonormal basis of V and that the space of Hilbert-

Schmidt operators $(L_{HS}(V,H), \langle\cdot,\cdot\rangle_{L_{HS}(V,H)})$, with inner product given by $\langle B,T\rangle_{L_{HS}(V,H)} := \sum_{j\in\mathbb{N}}\langle B\tilde{e}_j, T\tilde{e}_j\rangle_H$ for all $B,T \in L_{HS}(V,H)$, is a separable Hilbert space, [55, Appendix B].
By the assumptions on Q, there exists a unique decomposition of Q as $Q = Q^{\frac{1}{2}} \circ Q^{\frac{1}{2}}$ such that $Q^{\frac{1}{2}} \in L(V)$ is again symmetric and nonnegative, [55, Proposition 2.3.4]. In the following, it is convenient to work with the Cameron-Martin space V_0 defined by $V_0 := Q^{\frac{1}{2}}V$ as the isometry stated in equation (2.6) below involves this space naturally. The inner product on V_0 is defined as $\langle u,v\rangle_{V_0} = \langle Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}v\rangle_V$ for $u,v\in V_0$. Here, $Q^{-\frac{1}{2}}$ denotes the Pseudo inverse of Q and the space $(V_0, \langle\cdot,\cdot\rangle_{V_0})$ is a Hilbert space, [55, Appendix C].

First, we state the definition of the stochastic integral for some elementary process $(\Psi_t)_{t\in[0,T]}$ which takes finitely many values $\bar{\Psi}_0,\dots,\bar{\Psi}_{m-1} \in L(V,H)$ on $0 = t_0 < t_1 < \ldots < t_m = T$ such that $\bar{\Psi}_n$ is $\mathcal{F}_n$-measurable for all $n \in \{0,\dots,m-1\}$ and $\Psi_t = \sum_{n=0}^{m-1}\bar{\Psi}_n \mathbb{1}_{(t_n,t_{n+1}]}(t)$ for some $m\in\mathbb{N}$. The space of all these processes is denoted as $\mathcal{E}$.
For these elementary processes, the stochastic integral, which is a square-integrable and continuous martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, is defined as

$$\int_0^t \Psi_s \,\mathrm{d}W_s = \sum_{n=0}^{m-1} \bar{\Psi}_n (W_{t_{n+1}\wedge t} - W_{t_n\wedge t})$$

for any $t\in[0,T]$. The following isometry holds

$$\mathrm{E}\Big[\Big\|\int_0^T \Psi_s \,\mathrm{d}W_s\Big\|_H^2\Big] = \mathrm{E}\Big[\int_0^T \|\Psi_s\|^2_{L_{HS}(V_0,H)} \,\mathrm{d}s\Big]. \tag{2.6}$$

Next, we transfer the definition of the stochastic integral to a larger class of processes. Therefore, we introduce the following notation. Let $\Omega_T = [0,T]\times\Omega$, $P_T = \mathrm{d}t\otimes P$, and define the σ-Algebra of all predictable processes $Y : \Omega_T \to \mathbb{R}$ as

$$\mathcal{P}_T := \sigma\big(\{(s,t]\times F_s |\, 0\le s<t\le T, F_s\in\mathcal{F}_s\} \cup \{\{0\}\times F_0 |\, F_0\in\mathcal{F}_0\}\big).$$

Denote the completion of $\mathcal{E}$ by $\mathcal{N}_W^2(0,T;H)$; it can be shown that

$$\begin{aligned}\mathcal{N}_W^2(0,T;H) = \Big\{ Y : \Omega_T \to L_{HS}(V_0,H) \,|\, Y \text{ is } \mathcal{P}_T - \mathcal{B}(L_{HS}(V_0,H))\text{-measurable}\\ \text{and } \mathrm{E}\Big[\int_0^T \|Y_s\|^2_{L_{HS}(V_0,H)}\,\mathrm{d}s\Big] < \infty\Big\}.\end{aligned} \tag{2.7}$$

For processes $(Y_s)_{s\in[0,T]} \in \mathcal{N}_W^2(0,T;H)$, we can extend the definition of the stochastic integral as there exist elementary processes $(\Psi_s^n)_{s\in[0,T]} \in \{T_{|V_0} : T\in L(V,H)\}$, $n\in\mathbb{N}$, such that

$$\lim_{n\to\infty}\Big(\mathrm{E}\Big[\int_0^T \|Y_s - \Psi_s^n\|^2_{L_{HS}(V_0,H)}\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} = 0.$$

Finally, the stochastic integral can be extended to the class of processes

$$\mathcal{N}_W(0,T;H) := \Big\{ Y : \Omega_T \to L_{HS}(V_0,H) \,|\, Y \text{ is } \mathcal{P}_T - \mathcal{B}(L_{HS}(V_0,H))\text{-measurable} \\ P\Big(\int_0^T \|Y_s\|^2_{L_{HS}(V_0,H)} \,\mathrm{d}s < \infty \Big) = 1 \Big\} \tag{2.8}$$

by a localization procedure.

Before we move on to the analysis of the solution to equation (2.1), we state some properties of the stochastic integral that we need later on. The following estimates are essential to prove the existence of a unique solution to SPDE (2.1) as well as in the error analysis of the numerical schemes in Chapter 3 and Chapter 4.

Theorem 2.5

For any process $(Z_s)_{s\in[0,T]}$ that is $\mathcal{P}_T$-$\mathcal{B}(L_{HS}(V_0,H))$-measurable and every $p > 0$, there exists a constant $c_p > 0$ such that

$$\mathrm{E}\Big[\sup_{s\in[0,t]} \Big\| \int_0^s Z_r \,\mathrm{d}W_r \Big\|_H^p \Big] \le c_p \Big(\mathrm{E}\Big[\int_0^t \|Z_r\|^2_{L_{HS}(V_0,H)} \,\mathrm{d}r \Big] \Big)^{\frac{p}{2}}$$

for every $t \in [0,T]$.

[13, Theorem 4.36]

Proof. The proof of this theorem employs Itô's formula and some basic martingale inequalities. It is detailed in [13, Section 4.6]. □

Theorem 2.6

For any process $(Z_s)_{s\in[0,T]}$ that is $\mathcal{P}_T$-$\mathcal{B}(L_{HS}(V_0,H))$-measurable and every $p \ge 2$, there exists a constant $c_p > 0$ such that

$$\mathrm{E}\Big[\sup_{s\in[0,t]} \Big\| \int_0^s Z_r \,\mathrm{d}W_r \Big\|_H^p \Big] \le c_p \Big(\int_0^t \big(\mathrm{E}\big[\|Z_r\|^p_{L_{HS}(V_0,H)}\big]\big)^{\frac{2}{p}} \,\mathrm{d}r \Big)^{\frac{p}{2}}$$

for every $t \in [0,T]$.

[13, Theorem 4.37]

Proof. The main idea of the proof is the same as in the proof of Theorem 2.5. For details, we refer to [13, Section 4.6] again. □

2.3 Existence and Uniqueness of Mild Solutions

With the theory presented in the previous sections, we can finally start to analyze the solvability of SPDE (2.1). Throughout this section, let $Q \in L(V)$ be symmetric and nonnegative with finite trace and denote by $(W_t)_{t\in[0,T]}$ a Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ taking values in V. We employ the notation introduced above.

We prove that SPDE (2.1) possesses a unique solution. Therefore, we need to impose some restrictions on the operators A, F, and B, of course, and most importantly specify what we understand as a solution to this equation. We assume the following.

(A1) $A : D(A) \subset H \to H$ is closed, densely defined, and the infinitesimal generator of an analytic semigroup $S(t) = e^{tA} \in L(H),\ t \in [0,T]$.

(A2) $F : H \to H$ is globally Lipschitz continuous.

(A3) $B : H \to L_{HS}(V_0, H)$ is globally Lipschitz continuous, $B(H_\delta) \subset L_{HS}(V_0, H_\delta)$, and $\|B(y)\|^2_{L_{HS}(V_0,H_\delta)} \leq C^2(1 + \|y\|^2_{H_\delta})$, $y \in H_\delta$, for $\delta \in [0, \frac{1}{2})$, $C > 0$.

(A4) For $\gamma \in [\delta, \delta + \frac{1}{2})$, $p \in [2, \infty)$ the initial condition $\xi : \Omega \to H_\gamma$ is $\mathcal{F}_0$-$\mathcal{B}(H_\gamma)$-measurable and $\mathrm{E}\big[\|\xi\|^p_{H_\gamma}\big] < \infty$.

As for PDEs, there exist different notions of solutions. Based on [13], we define the strong solution.

Definition 2.7 (Strong Solution)
An H-valued predictable process $(X_t)_{t\in[0,T]}$ is called a strong solution of (2.1) if $X_t \in D(A)$ P_T-a.s.,

$$P\Big(\int_0^T (\|X_s\|_H + \|AX_s\|_H)\,\mathrm{d}s < \infty\Big) = 1,$$
$$P\Big(\int_0^T \|B(X_s)\|^2_{L_{HS}(V_0,H)}\,\mathrm{d}s < \infty\Big) = 1,$$

and if it holds, for all $t \in [0,T]$,

$$X_t = X_0 + \int_0^t \big(AX_s + F(X_s)\big)\,\mathrm{d}s + \int_0^t B(X_s)\,\mathrm{d}W_s \quad P\text{-a.s.}$$

[13, Chapter 6,7]

Another concept is the weak solution, where $(X_t)_{t\in[0,T]}$ does not need to take values in $D(A)$.

Definition 2.8 (Weak Solution)
An H-valued predictable process $(X_t)_{t\in[0,T]}$ is called a weak solution of (2.1) if

$$P\Big(\int_0^T \|X_s\|_H\,\mathrm{d}s < \infty\Big) = 1,$$
$$P\Big(\int_0^T \|B(X_s)\|^2_{L_{HS}(V_0,H)}\,\mathrm{d}s < \infty\Big) = 1,$$

and for all $t \in [0,T]$ and $\phi \in D(A^*)$

$$\langle X_t, \phi\rangle_H = \langle X_0, \phi\rangle_H + \int_0^t \big(\langle X_s, A^*\phi\rangle_H + \langle F(X_s), \phi\rangle_H\big)\,\mathrm{d}s + \int_0^t \langle \phi, B(X_s)\,\mathrm{d}W_s\rangle_H \quad P\text{-a.s.}$$

[13, Chapter 6,7]

Finally, we define the mild solution of SPDE (2.1). This is the solution concept that we work with and the numerical schemes in Chapter 3 and Chapter 4 are developed to approximate this process.

Definition 2.9 (Mild Solution)
An H-valued predictable process $(X_t)_{t\in[0,T]}$ is called a mild solution of (2.1) if

$$P\Big(\int_0^T \|X_s\|_H^2 \,\mathrm{d}s < \infty\Big) = 1$$

and if, for all $t \in [0,T]$, it holds

$$X_t = S(t)X_0 + \int_0^t S(t-s)F(X_s)\,\mathrm{d}s + \int_0^t S(t-s)B(X_s)\,\mathrm{d}W_s \quad P\text{-a.s.}$$

[13, Chapter 7]

Obviously, any strong solution is also a weak solution. For further implications, we need to impose restrictions on the integrability, see [47, Appendix G].

Assumptions (A1)–(A4) allow for a unique mild solution of (2.1) as stated in the next theorem; its proof is a combination of the proofs from [13] and [34]. Moreover, some properties of the mild solution are shown.

Theorem 2.7 (Existence and Uniqueness of Mild Solutions)
Assume that (A1)-(A4) are satisfied. Then, there exists a, up to modifications, unique mild solution $X : [0,T]\times\Omega \to H_\gamma$ of SPDE (2.1) with $\sup_{t\in[0,T]} \mathrm{E}[\|X_t\|_{H_\gamma}^p] < \infty$, where p and γ are determined by (A4).
Furthermore, it holds

(i) $(X_t)_{t\in[0,T]}$ has a continuous modification with respect to $\big(\mathrm{E}\big[\|\cdot\|_{H_\gamma}^p\big]\big)^{\frac{1}{p}}$,

(ii) for $r \in [0,\gamma)$, $p \in [2,\infty)$,

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

[34, Theorem 1]

Proof.
The proof of the existence of a mild solution of equation (2.1) is based on a fixed point theorem for contractions. This is a standard technique also used to prove the existence of solutions for other types of differential equations.
For $r \in [0,\infty)$, $p \geq 2$, let $\mathcal{H}_r$ denote the vector space of equivalence classes of predictable processes $Y : [0,T]\times\Omega \to H_r$ such that

$$\sup_{t\in[0,T]} \mathrm{E}\big[\|Y_t\|_{H_r}^p\big] < \infty.$$

Two stochastic processes $X, Y : [0,T] \times \Omega \to H_r$ lie in one equivalence class iff

$$P(X_t = Y_t) = 1 \quad \forall t \in [0,T],$$

that is, if and only if they are modifications of each other.
Define

$$\|Y\|_{\mathcal{H}_r} := \sup_{t\in[0,T]} \big(\mathrm{E}\,\big[\|Y_t\|_{H_r}^p\big]\,\big)^{\frac{1}{p}}$$

for all $Y \in \mathcal{H}_r$, $r \in [0,\infty)$, then the space $(\mathcal{H}_r, \|\cdot\|_{\mathcal{H}_r})$ is a Banach space for all $r \in [0,\infty)$, [13, Chapter 7].
Next, we specify a mapping $\Psi : \mathcal{H}_\delta \to \mathcal{H}_\delta$ for all $t \in [0,T]$ and $(Y_t)_{t\in[0,T]} \in \mathcal{H}_\delta$ as

$$\big(\Psi(Y)\big)_t := e^{At}\xi + \int_0^t e^{A(t-s)} F(Y_s)\,\mathrm{d}s + \int_0^t e^{A(t-s)} B(Y_s)\,\mathrm{d}W_s \quad P\text{-a.s.} \tag{2.9}$$

We show that $\Psi : \mathcal{H}_\delta \to \mathcal{H}_\delta$ is well defined and a contraction.

First, we set $C_A := \sup_{s\in[0,T]} \|e^{As}\|_{L(H)}$ and compute for all $t \in [0,T]$

$$\begin{aligned}\|e^{At}\xi\|_{\mathcal{H}_\gamma} &= \sup_{t\in[0,T]} \big(\mathrm{E}\big[\|e^{At}\xi\|_{H_\gamma}^p\big]\big)^{\frac{1}{p}} \le \sup_{t\in[0,T]} \|e^{At}\|_{L(H)} \big(\mathrm{E}\big[\|\xi\|_{H_\gamma}^p\big]\big)^{\frac{1}{p}} \\ &= C_A\big(\mathrm{E}\big[\|\xi\|_{H_\gamma}^p\big]\big)^{\frac{1}{p}}. \end{aligned} \tag{2.10}$$

With (A1) and (A4) as well as Proposition 3.7 (ii) in [13], we get that $e^{At}\xi$, $t \in [0,T]$, is a predictable stochastic process in $\mathcal{H}_\gamma$.

By Kuratowski's Theorem, see [37], it follows that $H_\delta \in \mathcal{B}(H)$ and $\mathcal{B}(H_\delta) = \mathcal{B}(H) \cap H_\delta$ which shows that $F|_{H_\delta}$ is $\mathcal{B}(H_\delta)$-$\mathcal{B}(H)$-measurable. We use assumption (A1) and (A2) to obtain

$$\begin{aligned}&\mathrm{E}\Big[\Big\|\int_0^t e^{A(t-s)} F(Y_s)\,\mathrm{d}s\Big\|_{H_\gamma}\Big] \le \mathrm{E}\Big[\int_0^t \|(-A)^\gamma e^{A(t-s)}\|_{L(H)} \|F(Y_s)\|_H\,\mathrm{d}s\Big] \\ &\le \mathrm{E}\Big[\int_0^t C_\gamma (t-s)^{-\gamma} \|F(Y_s)\|_H\,\mathrm{d}s\Big] \le C_\gamma \int_0^t (t-s)^{-\gamma} \sup_{u\in[0,T]} \mathrm{E}\big[\|F(Y_u)\|_H\big]\,\mathrm{d}s \\ &\le C_\gamma \sup_{s\in[0,T]} \mathrm{E}\big[1 + \|Y_s\|_H\big] \int_0^t (t-s)^{-\gamma}\,\mathrm{d}s \le C_\gamma \sup_{s\in[0,T]} \big(1 + \mathrm{E}\big[\|Y_s\|_{H_\delta}^p\big]^{\frac{1}{p}}\big) \frac{T^{1-\gamma}}{1-\gamma}\end{aligned}$$

for all $t \in [0,T]$. This implies that the process $\int_0^t e^{A(t-s)} F(Y_s)\,\mathrm{d}s$, $t \in [0,T]$, is adapted and takes values in H_γ for all $(Y_t)_{t\in[0,T]} \in \mathcal{H}_\delta$.
Analogously to the argumentation above, we obtain that $B_\delta : H_\delta \to L_{HS}(V_0, H_\delta)$, defined as $B_\delta(y) = B(y)$ for $y \in H_\delta$, is $\mathcal{B}(H_\delta)$-$\mathcal{B}(L_{HS}(V_0, H_\delta))$-measurable. Then, we estimate the third term with the help of assumption (A3) and Itô's isometry for all $t \in [0,T]$

$$\mathrm{E}\Big[\Big\|\int_0^t e^{A(t-s)} B(Y_s)\,\mathrm{d}W_s\Big\|_{H_\gamma}^2\Big] = \int_0^t \mathrm{E}\big[\|e^{A(t-s)} B(Y_s)\|_{L_{HS}(V_0, H_\gamma)}^2\big]\,\mathrm{d}s$$

$$
\begin{aligned}
&\leq \int_0^t \left\|(-A)^{\gamma-\delta} e^{A(t-s)}\right\|_{L(H)}^2 \mathrm{E}\big[\|B(Y_s)\|_{L_{HS}(V_0,H_\delta)}^2\big]\,\mathrm{d}s \\
&\leq C_{\gamma-\delta} \int_0^t (t-s)^{2(\delta-\gamma)} \mathrm{E}\big[(1+\|Y_s\|_{H_\delta})^2\big]\,\mathrm{d}s \\
&\leq 2C_{\gamma-\delta} \sup_{s\in[0,T]} \big(1+\mathrm{E}\big[\|Y_s\|_{H_\delta}^2\big]\big) \int_0^t (t-s)^{2(\delta-\gamma)}\,\mathrm{d}s \\
&\leq 2C_{\gamma-\delta} \sup_{s\in[0,T]} \big(1+\mathrm{E}\big[\|Y_s\|_{H_\delta}^2\big]\big) \frac{T^{1+2(\delta-\gamma)}}{1+2(\delta-\gamma)}.
\end{aligned}
$$

We obtain that $\int_0^t e^{A(t-s)} B(Y_s)\,\mathrm{d}W_s$, $t\in[0,T]$, is an adapted and H_γ-valued stochastic process for all $(Y_t)_{t\in[0,T]} \in H_\delta$ (see also Remark 1 in [34]).

Let $r\in[0,\gamma]$; now, we prove estimates which are also used to show (i) and (ii). In the following, we employ assumption (A3), Corollary A.1 in [32], and estimates on the semigroup e^{At}, $t\in[0,T]$, as stated in Theorem 2.3. Let $t,s\in[0,T]$ with $s\leq t$, it holds

$$
\begin{aligned}
&\left(\mathrm{E}\left[\left\| \int_0^t e^{A(t-u)} F(Y_u)\,\mathrm{d}u - \int_0^s e^{A(s-u)} F(Y_u)\,\mathrm{d}u \right\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\quad\leq \left(\mathrm{E}\left[\left\| \int_s^t e^{A(t-u)} F(Y_u)\,\mathrm{d}u \right\|_{H_r}^p\right]\right)^{\frac{1}{p}} + \left(\mathrm{E}\left[\left\| \int_0^s \Big(e^{A(t-u)} - e^{A(s-u)}\Big) F(Y_u)\,\mathrm{d}u \right\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\quad\leq \int_s^t (t-u)^{-r} \big(\mathrm{E}\big[\|F(Y_u)\|_H^p\big]\big)^{\frac{1}{p}}\,\mathrm{d}u + \left(\mathrm{E}\left[\left\| \int_0^s \Big(e^{A(t-s)} - I\Big) e^{A(s-u)} F(Y_u)\,\mathrm{d}u \right\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\quad\leq \int_s^t (t-u)^{-r} \big(\mathrm{E}\big[(1+\|Y_u\|_H)^p\big]\big)^{\frac{1}{p}}\,\mathrm{d}u \\
&\qquad + \|(-A)^{r-\gamma-\kappa}(e^{A(t-s)}-I)\|_{L(H)} \int_0^s \big(\mathrm{E}\big[\|e^{A(s-u)}F(Y_u)\|_{H_{\gamma+\kappa}}^p\big]\big)^{\frac{1}{p}}\,\mathrm{d}u \\
&\quad\leq \int_s^t (t-u)^{-r} 2^{1-\frac{1}{p}} \big(1+\big(\mathrm{E}\big[\|Y_u\|_H^p\big]\big)^{\frac{1}{p}}\big)\,\mathrm{d}u \\
&\qquad + (t-s)^{\kappa+\gamma-r} \int_0^s (s-u)^{-\gamma-\kappa} 2^{1-\frac{1}{p}} \big(1+\big(\mathrm{E}\big[\|Y_u\|_H^p\big]\big)^{\frac{1}{p}}\big)\,\mathrm{d}u
\end{aligned}
$$

for all $\kappa\in[0,1-\gamma)$, $r\in[0,\gamma]$, and $(Y_s)_{s\in[0,T]}\in\mathcal{H}_\delta$.
Further computations show

$$
\begin{aligned}
&\left(\mathrm{E}\left[\left\| \int_0^t e^{A(t-u)} F(Y_u)\,\mathrm{d}u - \int_0^s e^{A(s-u)} F(Y_u)\,\mathrm{d}u \right\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\quad\leq C 2^{1-\frac{1}{p}} \Big(1+\sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_H^p\big]\big)^{\frac{1}{p}}\Big) \int_s^t (t-u)^{-r}\,\mathrm{d}u \\
&\qquad + (t-s)^{\kappa+\gamma-r} 2^{1-\frac{1}{p}} C \Big(1+\sup_{u\in[0,T]} \big(\mathrm{E}\,\big[\|Y_u\|_H^p\big]\,\big)^{\frac{1}{p}}\Big) \int_0^s (s-u)^{-\gamma-\kappa}\,\mathrm{d}u \\
&\quad\leq C 2^{1-\frac{1}{p}} \Big(1+\sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_H^p\big]\big)^{\frac{1}{p}}\Big) \left(\frac{(t-s)^{1-r}}{1-r} + (t-s)^{\kappa+\gamma-r} \frac{s^{1-\gamma-\kappa}}{1-\gamma-\kappa}\right) \\
&\quad\leq C 2^{1-\frac{1}{p}} \Big(1+\sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_H^p\big]\big)^{\frac{1}{p}}\Big) \left(\frac{(t-s)^{1-r}}{1-r} \frac{T^{1-\gamma}}{(t-s)^{1-\gamma}} + (t-s)^{\kappa+\gamma-r} \frac{T^{1-\gamma-\kappa}}{1-\gamma-\kappa}\right)
\end{aligned}
$$

for all $s \le t$, $s,t \in [0,T]$, $\kappa \in [0, 1-\gamma)$, and $r \in [0,\gamma]$.
In total, we obtain

$$\begin{aligned}
&\left(\mathrm{E}\left[\Big\| \int_0^t e^{A(t-u)} F(Y_u)\,\mathrm{d}u - \int_0^s e^{A(s-u)} F(Y_u)\,\mathrm{d}u \Big\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\quad \le C 2^{2-\frac{1}{p}} \Big(1 + \sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_H^p\big]\big)^{\frac{1}{p}}\Big) \frac{T^{1-\gamma}}{1-\gamma} (t-s)^{\kappa+\gamma-r}
\end{aligned} \tag{2.11}$$

for all $\kappa \in [0,1-\gamma)$, $r \in [0,\gamma]$, $s \le t$, $s,t \in [0,T]$, and $(Y_t)_{t\in[0,T]} \in \mathcal{H}_\delta$. This shows that $\left(\int_0^t e^{A(t-u)} F(Y_u)\,\mathrm{d}u\right)_{t\in[0,T]}$ is continuous; combining our results on the Bochner integral and employing Proposition 3.7 in [13], implies that $\left(\int_0^t e^{A(t-u)} F(Y_u)\,\mathrm{d}u\right)_{t\in[0,T]}$ has a modification in $\mathcal{H}_\gamma$ for all $(Y_t)_{t\in[0,T]} \in \mathcal{H}_\delta$.

Similarly, we estimate for all $\kappa \in [0, \frac{1}{2} + \delta - \gamma)$, $r \in [0,\gamma]$, $\delta \in [0, \frac{1}{2})$, $s \le t$, $s,t \in [0,T]$, and $(Y_t)_{t\in[0,T]} \in \mathcal{H}_\delta$

$$\begin{aligned}
&\left(\mathrm{E}\left[\Big\| \int_0^t e^{A(t-u)} B(Y_u)\,\mathrm{d}W_u - \int_0^s e^{A(s-u)} B(Y_u)\,\mathrm{d}W_u \Big\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\le \left(\mathrm{E}\left[\Big\| \int_s^t e^{A(t-u)} B(Y_u)\,\mathrm{d}W_u \Big\|_{H_r}^p\right]\right)^{\frac{1}{p}} + \left(\mathrm{E}\left[\Big\| \int_0^s \Big(e^{A(t-u)} - e^{A(s-u)}\Big) B(Y_u)\,\mathrm{d}W_u \Big\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\le \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}} \left(\Big(\int_s^t \big(\mathrm{E}\big[\|(-A)^r e^{A(t-u)} B(Y_u)\|_{L_{HS}(V_0,H)}^p\big]\big)^{\frac{2}{p}}\,\mathrm{d}u\Big)^{\frac{p}{2}\frac{1}{p}}\right. \\
&\quad \left. + \Big(\int_0^s \|(-A)^{r-\gamma-\kappa}(e^{A(t-s)} - I)\|_{L(H)}^2 \big(\mathrm{E}\big[\|e^{A(s-u)} B(Y_u)\|_{L_{HS}(V_0,H_{\gamma+\kappa})}^p\big]\big)^{\frac{2}{p}}\,\mathrm{d}u\Big)^{\frac{p}{2}\frac{1}{p}}\right) \\
&\le p \left(\int_s^t \|(-A)^{r-\delta} e^{A(t-u)}\|_{L(H)}^2 \big(\mathrm{E}\big[\|B(Y_u)\|_{L_{HS}(V_0,H_\delta)}^p\big]\big)^{\frac{2}{p}}\,\mathrm{d}u\right)^{\frac{1}{2}} \\
&\quad + p(t-s)^{\kappa+\gamma-r} \left(\int_0^s (s-u)^{2(\delta-\gamma-\kappa)} \big(\mathrm{E}\big[\|B(Y_u)\|_{L_{HS}(V_0,H_\delta)}^p\big]\big)^{\frac{2}{p}}\,\mathrm{d}u\right)^{\frac{1}{2}} \\
&\le pC\Big(\int_s^t \|(-A)^{r-\delta} e^{A(t-u)}\|_{L(H)}^2 \big(1 + \mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\,\mathrm{d}u\Big)^{\frac{1}{2}} \\
&\quad + pC(t-s)^{\kappa+\gamma-r}\Big(\int_0^s (s-u)^{2(\delta-\gamma-\kappa)} \big(1 + \mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\,\mathrm{d}u\Big)^{\frac{1}{2}}.
\end{aligned}$$

Here, we used Theorem 2.6 and assumption (A3) again.
Further, we obtain with Theorem 2.3 parts b) and c)

$$\begin{aligned}
&\left(\mathrm{E}\left[\Big\| \int_0^t e^{A(t-u)} B(Y_u)\,\mathrm{d}W_u - \int_0^s e^{A(s-u)} B(Y_u)\,\mathrm{d}W_u \Big\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\
&\le p\,C\Big(1 + \sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)\Big(\int_s^t \|(-A)^{r-\delta} e^{A(t-u)}\|_{L(H)}^2\,\mathrm{d}u\Big)^{\frac{1}{2}} \\
&\quad + p\,C\Big(1 + \sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)(t-s)^{\kappa+\gamma-r}\Big(\frac{s^{2(\delta-\gamma-\kappa)+1}}{1+2(\delta-\gamma-\kappa)}\Big)^{\frac{1}{2}} \\
&\le p\,C\Big(1 + \sup_{u\in[0,T]} \big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)\Big(\int_s^t (C_{r-\delta}^2 (t-u)^{2(\delta-r)} + C^2)\,\mathrm{d}u\Big)^{\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
&+ p\,C\Big(1+\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)(t-s)^{\kappa+\gamma-r}\frac{T^{\frac{1+2(\delta-\gamma-\kappa)}{2}}}{(1+2(\delta-\gamma-\kappa))^{\frac{1}{2}}}\\
&\le C_p\Big(1+\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)\left(\frac{(t-s)^{\frac{1}{2}+\delta-r}+(t-s)^{\frac{1}{2}}}{(1+2(\delta-\gamma))^{\frac{1}{2}}}+\frac{T^{\frac{1}{2}+\delta-\gamma-\kappa}(t-s)^{\kappa+\gamma-r}}{(1+2(\delta-\gamma-\kappa))^{\frac{1}{2}}}\right)\\
&\le C_p\Big(1+\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)\left(\frac{T^{\frac{1}{2}+\delta-\gamma-\kappa}(t-s)^{\kappa+\gamma-r}+(t-s)^{\frac{1}{2}}}{(1+2(\delta-\gamma))^{\frac{1}{2}}}\right)\\
&\le C_p\Big(1+\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)\frac{\max(1,T)}{(1+2(\delta-\gamma))^{\frac{1}{2}}}\Big((t-s)^{\kappa+\gamma-r}+(t-s)^{\frac{1}{2}}\Big).
\end{aligned}
$$

This implies

$$
\begin{aligned}
&\left(\mathrm{E}\left[\left\|\int_0^t e^{A(t-u)}B(Y_u)\,\mathrm{d}W_u-\int_0^s e^{A(s-u)}B(Y_u)\,\mathrm{d}W_u\right\|_{H_r}^p\right]\right)^{\frac{1}{p}}\\
&\le C_p\Big(1+\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\Big)\frac{\max(1,T)}{(1+2(\delta-\gamma))^{\frac{1}{2}}}(t-s)^{\min(\kappa+\gamma-r,\frac{1}{2})}<\infty
\end{aligned}
\tag{2.12}
$$

for all $\kappa\in[0,\frac{1}{2}+\delta-\gamma)$, $r\in[0,\gamma]$, $\delta\in[0,\frac{1}{2})$, $s\le t$, $s,t\in[0,T]$, and $(Y_t)_{t\in[0,T]}\in\mathcal{H}_\delta$.
As for the Bochner integral, we obtain that $\left(\int_0^t e^{A(t-u)}B(Y_u)\,\mathrm{d}W_u\right)_{t\in[0,T]}$ has a modification in $\mathcal{H}_\gamma\subset\mathcal{H}_\delta$. Combining the estimates above, shows that $\Psi:\mathcal{H}_\gamma\to\mathcal{H}_\gamma$ is well defined.

It remains to prove that Ψ is a contraction. Let $(Y_t^1)_{t\in[0,T]}$, $(Y_t^2)_{t\in[0,T]}\in\mathcal{H}_\delta$; in the following, we use Corollary A.1 in [32], Theorem 4.37 in [13], and assumptions (A2), (A3). For $t\in[0,T]$, we obtain

$$
\begin{aligned}
&\big(\mathrm{E}\big[\|\Psi(Y^1)_t-\Psi(Y^2)_t\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\\
&\le\left(\mathrm{E}\left[\left\|\int_0^t e^{A(t-s)}(F(Y_s^1)-F(Y_s^2))\,\mathrm{d}s\right\|_{H_\delta}^p\right]\right)^{\frac{1}{p}}\\
&\quad+\left(\mathrm{E}\left[\left\|\int_0^t e^{A(t-s)}(B(Y_s^1)-B(Y_s^2))\,\mathrm{d}W_s\right\|_{H_\delta}^p\right]\right)^{\frac{1}{p}}\\
&\le\int_0^t\big(\mathrm{E}\big[\|e^{A(t-s)}(F(Y_s^1)-F(Y_s^2))\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\,\mathrm{d}s\\
&\quad+C_{\frac{p}{2}}\left(\int_0^t\Big(\mathrm{E}\big[\|(-A)^\delta e^{A(t-s)}(B(Y_s^1)-B(Y_s^2))\|_{L_{HS}(V_0,H)}^p\big]\Big)^{\frac{2}{p}}\,\mathrm{d}s\right)^{\frac{p}{2}\frac{1}{p}}\\
&\le C_\delta\int_0^t(t-s)^{-\delta}\big(\mathrm{E}\big[\|Y_s^1-Y_s^2\|_H^p\big]\big)^{\frac{1}{p}}\,\mathrm{d}s+C_{\frac{p}{2},\delta}\left(\int_0^t(t-s)^{-2\delta}\big(\mathrm{E}\,\big[\|Y_s^1-Y_s^2\|_H^p\big]\,\big)^{\frac{2}{p}}\,\mathrm{d}s\right)^{\frac{1}{2}}\\
&\le C_\delta\int_0^t(t-s)^{-\delta}\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u^1-Y_u^2\|_H^p\big]\big)^{\frac{1}{p}}\,\mathrm{d}s\\
&\quad+C_{\frac{p}{2},\delta}\left(\int_0^t(t-s)^{-2\delta}\Big(\sup_{u\in[0,T]}\big(\mathrm{E}\big[\|Y_u^1-Y_u^2\|_H^p\big]\big)^{\frac{1}{p}}\Big)^2\,\mathrm{d}s\right)^{\frac{1}{2}}\\
&\le\left(C_\delta\frac{T^{1-\delta}}{1-\delta}+C_{\frac{p}{2},\delta}\sqrt{\frac{T^{1-2\delta}}{1-2\delta}}\right)\sup_{s\in[0,T]}\big(\mathrm{E}\,\big[\|Y_s^1-Y_s^2\|_H^p\big]\,\big)^{\frac{1}{p}}
\end{aligned}
$$

$$= \left(C_\delta \frac{T^{1-\delta}}{1-\delta} + C_{\frac{p}{2},\delta} \sqrt{\frac{T^{1-2\delta}}{1-2\delta}} \right) \|Y^1 - Y^2\|_{\mathcal{H}}.$$

This shows

$$\begin{aligned} \|\Psi(Y^1) - \Psi(Y^2)\|_{\mathcal{H}_\delta} &\leq C_{\frac{p}{2},\delta} \left(\frac{T^{1-\delta}}{1-\delta} + \sqrt{\frac{T^{1-2\delta}}{1-2\delta}} \right) \|Y^1 - Y^2\|_{\mathcal{H}} \\ &\leq C_{\frac{p}{2},\delta} \left(\frac{T^{1-\delta}}{1-\delta} + \sqrt{\frac{T^{1-2\delta}}{1-2\delta}} \right) \|Y^1 - Y^2\|_{\mathcal{H}_\delta} \end{aligned}$$

for all $(Y_t^1)_{t\in[0,T]}, (Y_t^2)_{t\in[0,T]} \in \mathcal{H}_\delta$. Therefore, the mapping Ψ is a contraction and allows for a unique fixed point in $\mathcal{H}_\delta$ iff $C_{\frac{p}{2},\delta}\left(\frac{T^{1-\delta}}{1-\delta} + \sqrt{\frac{T^{1-2\delta}}{1-2\delta}}\right) < 1$. If this condition is fulfilled, there exists a predictable and unique (up to modifications) process $(Y_t^*)_{t\in[0,T]} \in \mathcal{H}_\delta$ such that

$$Y_t^* = e^{At}\xi + \int_0^t e^{A(t-s)} F(Y_s^*)\,\mathrm{d}s + \int_0^t e^{A(t-s)} B(Y_s^*)\,\mathrm{d}W_s \quad P\text{-a.s.}, \quad t \in [0,T].$$

We partition the interval $[0,T]$ into subintervals $[0,T^*], [T^*, 2T^*], \ldots, [nT^*, T], n \in \mathbb{N}_0$, such that T^* fulfills $C_{\frac{p}{2},\gamma}\left(\frac{(T^*)^{1-\delta}}{1-\delta} + \sqrt{\frac{(T^*)^{1-2\delta}}{1-2\delta}}\right) < 1$, which completes the proof of the first statement. Finally, the estimates (2.10), (2.11), (2.12), and Proposition 3.7 in [13] show that there exists a predictable modification $X : \Omega_T \to H_\gamma$ of Y^*.
Furthermore, it follows

i) From (2.10), (2.11), and (2.12), we get directly that $(Y_t)_{t\in[0,T]}$ is continuous with respect to $\left(\mathrm{E}\left[\|\cdot\|_{H_\gamma}^p\right]\right)^{\frac{1}{p}}$.

ii) Let $(Y_t^1)_{t\in[0,T]}, (Y_t^2)_{t\in[0,T]} \in H_\delta$; for $t_2 > t_1$, $t_1, t_2 \in [0,T]$, $r \in [0,\gamma)$, $\delta \in [0, \frac{1}{2})$, we obtain by (2.11) and (2.12)

$$\begin{aligned} &\left(\mathrm{E}\left[\|Y_{t_2} - Y_{t_1}\|_{H_r}^p\right]\right)^{\frac{1}{p}} \\ &\leq \left(\mathrm{E}\left[\|e^{At_2}\xi - e^{At_1}\xi\|_{H_r}^p\right]\right)^{\frac{1}{p}} + C2^{2-\frac{1}{p}}\left(1 + \sup_{u\in[0,T]} \mathrm{E}\left[\|Y_u\|_H^p\right]\right)^{\frac{1}{p}} \frac{T^{1-\gamma}}{1-\gamma}(t_2 - t_1)^{\gamma - r} \\ &\quad + C_p\left(1 + \sup_{u\in[0,T]} \mathrm{E}\left[\|Y_u\|_{H_\delta}^p\right]\right)^{\frac{1}{p}} \frac{\max(1,T)}{(1 + 2(\delta-\gamma))^{\frac{1}{2}}}(t_2 - t_1)^{\min(\gamma - r, \frac{1}{2})} \\ &\leq \left(\mathrm{E}\left[\|e^{At_1}\|_{L(H)}^p \|(-A)^{r-\gamma}(e^{A(t_2-t_1)} - I)\|_{L(H)}^p \|\xi\|_{H_\gamma}^p\right]\right)^{\frac{1}{p}} + C_{p,T,\delta,\gamma,r}(t_2 - t_1)^{\min(\gamma - r, \frac{1}{2})} \\ &\leq C_{A,\gamma,T,r}(t_2 - t_1)^{\gamma - r} + C_{p,T,\delta,\gamma,r}(t_2 - t_1)^{\min(\gamma - r, \frac{1}{2})} \\ &\leq C_{p,T,\gamma,\delta,r}(t_2 - t_1)^{\min(\gamma - r, \frac{1}{2})}. \end{aligned}$$

□

We proved that, given (A1)–(A4), SPDE (2.1) allows for a unique solution. In the next chapters we aim at obtaining this solution.

3

Efficient Approximation of Commutative SPDEs

Explicit solutions to SPDE (1.1) are, in general, not computable; it is only possible to obtain the solution process for a few types of SPDEs analytically. In the following, we specify two examples and present different approaches to solve these equations. We emphasize the difficulties that are involved and highlight the features that allow for the computation of an explicit solution.
If the diffusion operator is of the form $B(X_t) = bX_t,\ X_t \in H,\ t \in [0,T],\ b \in \mathbb{R}$, and we assume $V = \mathbb{R}$ additionally, it is possible to obtain an analytical solution - as outlined in the following example.

Example 3.1 (Scalar Brownian Motion)
For some $T \in (0,\infty)$, we compute the strong solution to

$$\begin{aligned} \mathrm{d}X_t &= \Delta X_t \,\mathrm{d}t + bX_t \,\mathrm{d}\beta_t, \quad t \in (0,T],\ b \in \mathbb{R}, \\ X_0(x) &= \sqrt{2}\sum_{n=1}^{\infty} c_n \sin(n\pi x), \quad x \in (0,1),\ c_n \in \mathbb{R},\ n \in \mathbb{N}, \\ X_t(0) &= X_t(1) = 0, \quad t \in (0,T], \end{aligned} \tag{3.1}$$

for some suitable sequence $(c_n)_{n\in\mathbb{N}}$. Here, $(\beta_t)_{t\in[0,T]}$ denotes a scalar Brownian motion and $H = L^2((0,1),\mathbb{R}),\ V = \mathbb{R}$.

If $c_n = \frac{1}{n^2}$, $n \in \mathbb{N}$, for example, one can show that

$$X_t(x) = \sqrt{2} \sum_{n=1}^{\infty} c_n e^{-(n^2\pi^2 + \frac{b^2}{2})t + b\beta_t} \sin(n\pi x) \tag{3.2}$$

is a strong solution to (3.1) for all $t \in [0,T]$, $x \in (0,1)$, see [13] or [21]. This can easily be proved by Itô's formula, which is stated in [7].

Next, we examine an equation with additive noise and show that even for this simple equation, we need to simulate the stochastic integrals involved in the solution process. A similar example is illustrated in [9].

Example 3.2 (Additive Noise)
Let $T \in (0,\infty)$ and $H = V = L^2((0,1),\mathbb{R})$; in the following, we compute the mild solution to

$$\begin{aligned} \mathrm{d}X_t &= (\Delta X_t + 1)\,\mathrm{d}t + b\,\mathrm{d}W_t, \quad b \in \mathbb{R},\ t \in (0,T], \\ X_0(x) &= \sqrt{2}\sum_{n=1}^{\infty} \frac{1}{n^2} \sin(n\pi x), \quad x \in (0,1), \\ X_t(0) &= X_t(1) = 0, \quad t \in (0,T]. \end{aligned} \tag{3.3}$$

Here, $(W_t)_{t\in[0,T]}$ is a Q-Wiener process, see Definition 2.6, and the operator $Q \in L(V)$, with eigenvalues $(\eta_n)_{n\in\mathbb{N}}$, is assumed to have finite trace. This admits the representation

$$W_t = \sum_{j\in\mathbb{N}} \sqrt{\eta_j}\beta_t^j \tilde{e}_j, \quad t \in [0,T],$$

see equation (2.5). Moreover, it holds $-\Delta e_n = \lambda_n e_n$ with $\lambda_n = n^2\pi^2$ and $e_n(x) = \sqrt{2}\sin(n\pi x)$ for all $n \in \mathbb{N}$, $x \in (0,1)$.

We compute the coefficients $a_n(t)$, $n \in \mathbb{N}$, $t \in [0,T]$, such that

$$X_t(x) = \sqrt{2}\sum_{n=1}^{\infty} a_n(t) \sin(n\pi x)$$

fulfills equation (3.3).
Obviously, it holds $a_n(0) = \frac{1}{n^2}$ and

$$\mathrm{d}a_n(t) = \big(-n^2\pi^2 a_n(t) + \langle 1, e_n\rangle_H\big)\,\mathrm{d}t + b\sqrt{\eta_n}\,\mathrm{d}\beta_t^n, \quad t \in (0,T],$$

for all $n \in \mathbb{N}$.
First, we compute the Fourier coefficients for all $n \in \mathbb{N}$

$$\langle 1, e_n\rangle_H = \sqrt{2}\int_0^1 \sin(n\pi x)\,\mathrm{d}x = \frac{2\sqrt{2}}{n\pi}\mathbb{1}_{n \text{ odd}}.$$

This yields the system $a_n(0) = \frac{1}{n^2}$ and

$$\mathrm{d}a_n(t) = \Big(- n^2\pi^2 a_n(t) + \frac{2\sqrt{2}}{n\pi}\mathbb{1}_{n\ odd}\Big)\,\mathrm{d}t + b\sqrt{\eta_n}\,\mathrm{d}\beta_t^n, \quad t \in (0,T],$$

for all $n \in \mathbb{N}$.
For each $n \in \mathbb{N}$, $t \in [0,T]$, the solution can easily be obtained and reads as

$$\begin{aligned} a_n(t) &= \frac{1}{n^2} e^{-n^2\pi^2 t} + \frac{2\sqrt{2}}{n\pi}\mathbb{1}_{n\ odd}\int_0^t e^{-n^2\pi^2(t-s)}\,\mathrm{d}s + b\sqrt{\eta_n}\int_0^t e^{-n^2\pi^2(t-s)}\,\mathrm{d}\beta_s^n \\ &= \frac{1}{n^2} e^{-n^2\pi^2 t} + \frac{2\sqrt{2}}{n^3\pi^3}\mathbb{1}_{n\ odd}\big(1 - e^{-n^2\pi^2 t}\big) + b\sqrt{\eta_n}\int_0^t e^{-n^2\pi^2(t-s)}\,\mathrm{d}\beta_s^n. \end{aligned}$$

Then, we get the mild solution of (3.3) *for all $t \in [0,T]$ and $x \in (0,1)$ as*

$$\begin{aligned} X_t(x) =& \sum_{n=1}^{\infty} \Big(\frac{1}{n^2} e^{-n^2\pi^2 t} + b\sqrt{\eta_n}\int_0^t e^{-n^2\pi^2(t-s)}\,\mathrm{d}\beta_s^n \Big)\sqrt{2}\sin(n\pi x) \\ &+ \sum_{n=0}^{\infty} \frac{2\sqrt{2}}{(2n+1)^3\pi^3}\big(1 - e^{-(2n+1)^2\pi^2 t}\big)\sqrt{2}\sin((2n+1)\pi x). \end{aligned}$$

We need to simulate the integrals $\int_0^t e^{-n^2\pi^2(t-s)}\,\mathrm{d}\beta_s^n$ for all $n \in \mathbb{N}$, $t \in [0,T]$, however, which can be done as described in [31].
This example shows that even for additive equations, we are not guaranteed to obtain a closed form solution.

In applications, the model at hand is generally not of such a simple form as in Example 3.1. That is, the diffusion operator is nonlinear in general and the Q-Wiener process belongs to a Hilbert space of infinite dimension such that no analytical solution can be computed. This justifies the extensive research on numerical methods for stochastic partial differential equations.

Stochastic partial differential equations need a distinct numerical treatment. We cannot simply employ the well studied methods developed to solve SODEs. Numerical methods for this class of differential equations are mainly designed for a fixed number of random influences, $K \in \mathbb{N}$, where K is often a factor in the error constant, see [38] or [59], for example. Therefore, they do, in general, not converge when K goes to infinity in the approximation of the Q-Wiener process. Moreover, even if for some $N, K \in \mathbb{N}$, we project the SPDE to a finite dimensional system of SODEs in H_N and obtain an approximation of the Q-Wiener process in V_K (these projections will be described in the following), schemes for SODEs are not necessarily applicable. The projection might distort properties of the original equation. One such example is the commutativity of the SPDE which reads

$$B'(y)\,(B(y)u, v) = B'(y)\,(B(y)v, u)$$

for all $u, v \in V_0$, $y \in H_\beta$, and some $\beta \in [0,1)$ specified below.

Example 3.3 (Projected SPDE is not commutative)
Assume a SPDE of the form

$$\begin{aligned} \mathrm{d}X_t &= \frac{\partial^2}{\partial x^2} X_t \,\mathrm{d}t + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{j^2} \langle X_t, e_i \rangle_H \langle X_t, e_{2j} \rangle_H e_i \,\mathrm{d}\beta_t^j, \quad t \in (0,T], \\ X_0 &= \xi. \end{aligned} \tag{3.4}$$

In this notation, the diffusion operator reads as

$$B(y)u = \sum_{i,j \in \mathbb{N}} \langle y, e_i \rangle_H \langle y, e_{2j} \rangle_H \langle u, \tilde{e}_j \rangle_V e_i$$

for all $y \in H_\beta$, $u \in V_0$. We choose the eigenvalues of Q as $\eta_j = j^{-4}$ for all $j \in \mathbb{N}$. Here, $\{e_i,\ i \in \mathbb{N}\}$ and $\{\tilde{e}_j,\ j \in \mathbb{N}\}$ denote the orthonormal bases of H and V, respectively, introduced in Chapter 2.
In this setting, we have

$$B'(y)\left(B(y)v, u\right) = \sum_{i,k,j,r=1}^{\infty} \Big(\langle y, e_i \rangle_H \mathbb{1}_{k=2j} + \langle y, e_{2j} \rangle_H \mathbb{1}_{k=i} \Big) \langle y, e_k \rangle_H \langle y, e_{2r} \rangle_H \langle v, \tilde{e}_r \rangle_V \langle u, \tilde{e}_j \rangle_V e_i$$

for all $u, v \in V_0$, $y \in H_\beta$. The notation and the derivation of this expression can be found in Section 3.6.

Then, the commutativity condition reads

$$\begin{aligned} &\sum_{k=1}^{\infty} \Big(\langle y, e_i \rangle_H \mathbb{1}_{k=2m} + \langle y, e_{2m} \rangle_H \mathbb{1}_{k=i} \Big) \langle y, e_k \rangle_H \langle y, e_{2n} \rangle_H \\ &= 2 \langle y, e_i \rangle_H \langle y, e_{2m} \rangle_H \langle y, e_{2n} \rangle_H \\ &\stackrel{!}{=} \sum_{k=1}^{\infty} \Big(\langle y, e_i \rangle_H \mathbb{1}_{k=2n} + \langle y, e_{2n} \rangle_H \mathbb{1}_{k=i} \Big) \langle y, e_k \rangle_H \langle y, e_{2m} \rangle_H \\ &= 2 \langle y, e_i \rangle_H \langle y, e_{2m} \rangle_H \langle y, e_{2n} \rangle_H \end{aligned}$$

for all $i \in \mathbb{N}$, $n, m \in \mathcal{J}_K$, $K \in \mathbb{N}$, $y \in H_\beta$. This shows that the equation is commutative. See Section 3.3 for a definition of the set $\mathcal{J}_K$.
Now, we define the projection operator $P_N : H \to H_N$ for $y \in H$ by

$$P_N y := \sum_{n=1}^{N} \langle y, e_n \rangle_H e_n$$

for all $N \in \mathbb{N}$.
We set $X_t^N = P_N X_t$ for $t \in [0,T]$, $N \in \mathbb{N}$, and approximate the Q-Wiener process by some

projection P_K, $K \in \mathbb{N}$, as well; then, we obtain

$$\mathrm{d}X_t^N = P_N \frac{\partial^2}{\partial x^2} X_t^N \,\mathrm{d}t + \sum_{i=1}^{N} \sum_{j=1}^{K} \frac{1}{j^2} \langle X_t^N, e_i \rangle_H \langle X_t^N, e_{2j} \rangle_H e_i \,\mathrm{d}\beta_t^j, \quad t \in (0,T], \tag{3.5}$$
$$X_0^N = P_N \xi.$$

This equation is not commutative anymore, that is, condition (1.3) *does not hold, if we project the spaces such that H_N is identified by $\mathbb{R}^N$, V_K by $\mathbb{R}^K$, for $N, K \in \mathbb{N}$ with $K > N$, and choose indices $i, j \in \{1, \ldots, K\}$ such that $2i > N$ but $2j < N$. In this case, we get*

$$\sum_{r=1}^{N} \langle y, e_r \rangle_H \langle y, e_{2j} \rangle_H \Big(\langle y, e_{2i} \rangle_H \mathbb{1}_{r=l} + \langle y, e_l \rangle_H \mathbb{1}_{r=2i} \Big) = \langle y, e_l \rangle_H \langle y, e_{2j} \rangle_H \langle y, e_{2i} \rangle_H$$

whereas

$$\sum_{r=1}^{N} \langle y, e_r \rangle_H \langle y, e_{2i} \rangle_H \Big(\langle y, e_{2j} \rangle_H \mathbb{1}_{r=l} + \langle y, e_l \rangle_H \mathbb{1}_{r=2j} \Big) = 2 \langle y, e_l \rangle_H \langle y, e_{2j} \rangle_H \langle y, e_{2i} \rangle_H$$

for all $l \in \{1, \ldots, N\}$ and $y \in H_\beta$.

Examples 3.1, 3.2, and 3.3 motivate the development of numerical schemes to approximate the mild solution of SPDE (1.1) in the following sections. First, we detail the framework that we assume throughout this work. We specify the operators and conditions necessary to conduct the analysis of convergence below. As described in the introduction, we develop numerical schemes free of derivatives to approximate the mild solution of SPDE (1.1). We focus on this type of solution in this work as for this process it is known how to construct higher order schemes by means of Taylor approximations, [32], which form the base for the derivative-free methods. We prove the convergence of the schemes and discuss their computational effort, which leads to the notion of the effective order of convergence. This concept combines the theoretical order of convergence with respect to the dimensions of the projection spaces and the time step size with the computational cost necessary to compute one path with a specific method. We are concerned about the effective order of convergence as this is the rate that actually determines the scheme that is favorable with respect to the overall computational effort. The analytical findings are then illustrated and confirmed in numerical simulations. In this chapter, we investigate SPDEs that are commutative.
Approximation schemes for equations of type (1.1) that do not fulfill this assumption are discussed in Chapter 4.

3.1 Setting for SPDEs

Throughout this chapter, let $T \in (0, \infty)$ and let $(\Omega, \mathcal{F}, P)$ denote a probability space endowed with a filtration $(\mathcal{F}_t)_{t \in [0,T]}$ fulfilling the usual conditions. Furthermore, let $(H, \langle \cdot, \cdot \rangle_H)$ and $(V, \langle \cdot, \cdot \rangle_V)$ denote separable real-valued Hilbert spaces. Assume $Q \in L(V)$ to be positive, sym-

metric, and to have finite trace and denote by $(W_t)_{t\in[0,T]}$ a V-valued Q-Wiener process with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, see Definition 2.6. According to Section 2.2, there exists an orthonormal basis of V consisting of eigenfunctions $\tilde{e}_j$, $j \in \mathbb{N}$, of Q. In the following, we consider this basis.
Our aim is to obtain an approximation of the mild solution $(X_t)_{t\in[0,T]}$ to

$$\begin{aligned} \mathrm{d}X_t &= (AX_t + F(X_t))\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t, \quad t \in (0,T], \\ X_0 &= \xi. \end{aligned} \tag{3.6}$$

We denote by A the generator of an analytic semigroup, call F drift operator and B diffusion operator. Let $\mathcal{I}$ and $\mathcal{J}$ denote finite or countable sets such that $\eta_k \neq 0$ for all $k \in \mathcal{J}$. Then, $(W_t)_{t\in[0,T]}$ can be represented as

$$W_t = \sum_{j\in\mathcal{J}} \sqrt{\eta_j}\tilde{e}_j\beta_t^j \quad P\text{-a.s.} \tag{3.7}$$

for all $t \in [0,T]$ with $\eta_j\tilde{e}_j = Q\tilde{e}_j$, $j \in \mathcal{J}$, see Theorem 2.4. The index set $\mathcal{I}$ is specified in condition (C1) below.
We assume that the operators fulfill the following conditions, where we follow the notation in [35] to make our results easily comparable.

(C1) Let $A : D(A) \subseteq H \to H$ be the infinitesimal generator of an analytic semigroup $S(t) = e^{tA} \in L(H)$, $t \in [0,T]$. Denote by e_i, $i \in \mathcal{I}$, its eigenfunctions and by $\lambda_i \in (0,\infty)$, $i \in \mathcal{I}$, its eigenvalues such that $\inf_{i\in\mathcal{I}} \lambda_i > 0$, $-\lambda_i e_i = Ae_i$, $i \in \mathcal{I}$, and A can be written as

$$Au = \sum_{i\in\mathcal{I}} -\lambda_i \langle u, e_i\rangle_H e_i$$

for all $u \in D(A)$ with

$$D(A) := \Big\{u \in H : \sum_{i\in\mathcal{I}} |\lambda_i|^2 |\langle u, e_i\rangle_H|^2 < \infty\Big\}.$$

This implies the existence of an orthonormal basis $\{e_i,\ i \in \mathcal{I}\}$ of H, as described in Section 2.1 or [65]. In the following, we work with this basis. Moreover, we define the interpolation spaces $H_r = D((-A)^r)$ for some $r \in [0,\infty)$, see Section 2.1 for details.

Remark 3.1
In this setting, the analytic semigroup can be expressed as

$$e^{At}y = \sum_{i\in\mathcal{I}} e^{-\lambda_i t}\langle y, e_i\rangle_H e_i$$

for all $y \in H$ and $t \in [0,T]$, [65, p.67]. Therefore, its implementation in a numerical scheme is straightforward.

(C2) Let $F : H_\beta \to H$ be twice continuously Fréchet differentiable with $\sup_{y\in H_\beta} \|F'(y)\|^2_{L(H)} < \infty$ and $\sup_{y\in H_\beta} \|F''(y)\|^2_{L^{(2)}(H_\beta,H)} < \infty$ for some $\beta \in [0,1)$.

We introduce the Cameron-Martin space $V_0 = Q^{\frac{1}{2}}V$ which we use in the following, more details can be found in Section 2.2. Moreover, define $L(V,H)_0 := \{T|_{V_0} \mid T \in L(V,H)\}$, see [55]. Note that $L(V,H)_0$ is a dense subset of $L_{HS}(V_0,H)$, [55, Lemma 2.3.7].

(C3) Assume $B : H_\beta \to L(V,H)_0$, $\beta \in [0,1)$, to be twice continuously Fréchet differentiable with $\sup_{y\in H_\beta} \|B'(y)\|_{L(H,L(V,H))} < \infty$, $\sup_{y\in H_\beta} \|B''(y)\|_{L^{(2)}(H,L(V,H))} < \infty$. Moreover, $B(H_\delta) \subset L(V,H_\delta)$ and for all $u \in H_\delta$, $y,w \in H_\gamma$,

$$\begin{aligned} \|B(u)\|_{L(V,H_\delta)} &\leq C(1+\|u\|_{H_\delta}), \\ \|B'(y)B(y) - B'(w)B(w)\|_{L^{(2)}_{HS}(V_0,H)} &\leq C\|y-w\|_H, \\ \|(-A)^{-\vartheta}B(y)Q^{-\alpha}\|_{L_{HS}(V_0,H)} &\leq C(1+\|y\|_{H_\gamma}) \end{aligned}$$

for some constant $C > 0$ and parameters $\alpha \in (0,\infty)$, $\delta,\vartheta \in (0,\frac{1}{2})$, $\gamma \in [\max(\delta,\beta), \delta+\frac{1}{2})$.

We denote $L^{(2)}(H,L(V,H)) = L(H,L(H,L(V,H))$ and $L^{(2)}_{HS}(V_0,H) = L_{HS}(V_0,L_{HS}(V_0,H))$ here.

(C4) The initial condition $\xi : \Omega \to H_\gamma$ is $\mathcal{F}_0$-$\mathcal{B}(H_\gamma)$-measurable with $E[\|\xi\|^4_{H_\gamma}] < \infty$.

Note that it holds $\sup_{y\in H_\beta} \|B'(y)\|_{L(H,L_{HS}(V_0,H))} \leq \sqrt{\operatorname{tr} Q}\, \sup_{y\in H_\beta} \|B'(y)\|_{L(H,L(V,H))} < \infty$. Since H_β is a dense subset of H, the operator $B\colon H_\beta \to L(V,H)_0$ can be continuously extended to a globally Lipschitz continuous mapping $\tilde{B}\colon H \to L(V,H)_0$. For legibility, we do not distinguish between B and $\tilde{B}$ in the following. For the operator F we proceed analogously.

As we show in the error analysis in Section 3.4, the parameters α, β, and γ determine the maximal order of convergence that can be obtained. This contrasts the results obtained for SODEs, where the rate of convergence can be specified universally for each approximation scheme.
We want to give an idea on the influence of these parameters on the intensity of the assumptions as well as the rate of convergence. For this purpose, we consider the setting underlying the examples in Section 3.6.
The parameter δ is mainly determined by the operator A, more precisely its eigenvalues $(\lambda_i)_{i\in\mathcal{I}}$, the operator B, and its behavior on the boundary. This can be seen in equation (3.37) as well as in [34]. Moreover, if we assume $\eta_j \leq Cj^{-\rho_Q}$ for some $C, \rho_Q > 0$ and all $j \in \mathcal{J}$, it follows from (3.40) and the examples in Section 3.6 that the parameter α increases with decreasing ρ_Q. That is, the faster the sequence of eigenvalues $(\eta_j)_{j\in\mathcal{J}}$ of Q converges to zero, the less restrictive is assumption (C3). Furthermore, α depends on the operator B naturally. The parameter range for ϑ is determined by the connection of the eigenvalues $(\lambda_i)_{i\in\mathcal{I}}$ of A and the operator B, see (3.40). The faster the sequence $(\lambda_i)_{i\in\mathcal{I}}$ increases, the smaller the parameter ϑ can be chosen. The value of $\beta \in [0,1)$ is selected such that the Fréchet derivatives of B and F are bounded on H_β, that is, (C2) and (C3) hold. Finally, γ is determined by the parameters δ and β.

If not stated differently, this is the setting in which the numerical schemes are analyzed in this work. Given (C1)–(C4), assumptions (A1)–(A4) in Section 2.3 hold with $p = 4$ such that there exists a unique mild solution to SPDE (3.6) according to Theorem 2.7. Moreover, the properties (i) and (ii) in Theorem 2.7 hold.

3.2 Effective Order of Convergence

Before we develop approximation schemes for SPDE (3.6), we specify the error criterium we are interested in. The concept that we introduce has also been considered in our article [46]. For some fixed $M, N, K \in \mathbb{N}$, denote by $(Y_m^{M,N,K})_{m\in\{0,\dots,M\}}$ the discrete approximation process of $(X_t)_{t\in[0,T]}$ obtained by some numerical scheme. Here, $Y_m^{M,N,K}$ are $\mathcal{F}_{t_m}$-$\mathcal{B}(H)$-measurable random variables for all $m \in \{0,\dots,M\}$. We examine the strong convergence of the schemes in this work, so we are concerned about

$$\sup_{m\in\{0,\dots,M\}} \Big(\mathrm{E}\big[\|X_{t_m} - Y_m^{M,N,K}\|_H^2\big]\Big)^{\frac{1}{2}} \tag{3.8}$$

for all $M, N, K \in \mathbb{N}$.

In the numerical analysis of SPDEs, we need to discretize the time interval and the infinite dimensional Hilbert space H (referred to as approximation with respect to the space domain) as well as the infinite dimensional stochastic process driving the equation. Therefore, the approximation error (3.8) is determined by diverse sources. In the following sections, we show that the approximation of H by a spectral projection yields an error term depending on $(\lambda_i)_{i\in\mathcal{I}}$, whereas the error resulting from the approximation of the Q-Wiener process involves the eigenvalues $(\eta_j)_{j\in\mathcal{J}}$. We denote these terms by $\mathcal{E}(\lambda, N)$ and $\mathcal{E}(\eta, K)$ for some $N, K \in \mathbb{N}$, respectively. Due to the approximation in the temporal direction, we obtain some error term $\mathcal{E}(M)$ depending on the step size $h > 0$ which, in turn, is determined by $M \in \mathbb{N}$. In total, we get an expression for the strong error subject to the dimensions N and K of the projection spaces and the step size in temporal direction as

$$\sup_{m\in\{0,\dots,M\}} \Big(\mathrm{E}\big[\|X_{t_m} - Y_m^{M,N,K}\|_H^2\big]\Big)^{\frac{1}{2}} \leq \mathcal{E}(\lambda, N) + \mathcal{E}(\eta, K) + \mathcal{E}(M). \tag{3.9}$$

We estimate this expression for the numerical schemes developed in the following sections. In order to obtain an overall order of convergence, we have to balance the error terms according to their respective convergence rates. We are, however, mainly interested in the relation of the error and the computational cost necessary to simulate one sample path with a specific numerical scheme, instead of considering the relation of the error to the dimensions of the approximation spaces only.

The cost model that we consider is based on [74], see also [46]. We employ this model as it is more objective than the comparison of computation time, which may depend on the implementation of the algorithm. In the following, we assume that an arithmetic operation, like the evaluation of the sine function, generates cost of one unit. The cost necessary to obtain information about, for example, some element $v \in H$ by some functional $\varphi : H \to \mathbb{R}$ is taken to be $\mathrm{cost}(\varphi) = c$ for

some $c \gg 1$. Therefore, the cost of arithmetic operations is negligible.
For some fixed $N, K \in \mathbb{N}$, we specify the finite dimensional subsets $\mathcal{I}_N$, $\mathcal{J}_K$ with $\mathcal{I}_N \subset \mathcal{I}$, $\mathcal{J}_K \subset \mathcal{J}$ and $|\mathcal{I}_N| = N$, $|\mathcal{J}_K| = K$, which is the worst case in terms of computational cost necessary to simulate one path. Moreover, we identify H_N by $\mathbb{R}^N$ and V_K by $\mathbb{R}^K$. In the numerical schemes that we describe in the following sections, there are mainly three terms that we need to evaluate. Let $y \in H_\beta$, $u, v \in V_K$, $N, K \in \mathbb{N}$,

- We compute $P_N F(y) = \sum_{i \in \mathcal{I}_N} \langle F(y), e_i \rangle_H e_i$; therefore, we need to obtain the functionals $\langle F(y), e_i \rangle_H$ with $\mathrm{cost}(\langle F(y), e_i \rangle_H) = c$ for all $i \in \mathcal{I}_N$. This implies $\mathrm{cost}(P_N F(y)) = \mathcal{O}(N)$.

- In order to compute $P_N B(y)u = \sum_{i \in \mathcal{I}_N} \sum_{j \in \mathcal{J}_K} \langle B(y)\tilde{e}_j, e_i \rangle_H \langle u, \tilde{e}_j \rangle_V e_i$, we need to evaluate $\langle B(y)\tilde{e}_j, e_i \rangle_H$ for all $i \in \mathcal{I}_N$, $j \in \mathcal{J}_K$ with $\mathrm{cost}(P_N B(y)|_{V_K}) = cNK$. Moreover, given $\langle B(y)\tilde{e}_j, e_i \rangle_H$ and $\langle u, \tilde{e}_j \rangle_V$ for all $i \in \mathcal{I}_N$, $j \in \mathcal{J}_K$, the computation of $P_N B(y)u$ requires K multiplications and $K-1$ summations for each $i \in \mathcal{I}_N$. This yields $\mathrm{cost}(P_N B(\cdot)u) = 2NK - 1$. In total, we obtain $\mathrm{cost}(P_N B(y)u) = \mathcal{O}(NK)$.

- For the Milstein scheme, we additionally compute

$$P_N B'(y)(B(y)u)v = \sum_{i,k \in \mathcal{I}_N} \sum_{j \in \mathcal{J}_K} \langle B'(y)(e_k, \tilde{e}_j), e_i \rangle_H \langle B(y)u, e_k \rangle_H \langle v, \tilde{e}_j \rangle_V e_i,$$

  that is, we need to evaluate $\langle B'(y)(e_k, \tilde{e}_j), e_i \rangle_H$ for all $i, k \in \mathcal{I}_N$, $j \in \mathcal{J}_K$ with $\mathrm{cost}(P_N B'(y)(\cdot,\cdot)|_{H_N, V_K}) = cN^2K$. The total cost necessary to compute $P_N B'(y)(B(y)u)v$ is $\mathrm{cost}(P_N B'(y)(B(y)u)v) = \mathcal{O}(N^2 K)$, which can be obtained similar as for $P_N B(y)u$.

Combining these terms according to the specific scheme, gives the computational cost required to simulate the solution process $(X_t)_{t \in [0,T]}$ at time T. This is detailed in Section 3.4.

Now, we derive the connection of the computational cost and the approximation error (3.8) as our goal is to minimize this term such that the computational cost (CC) does not exceed some specified value $\bar{c} > 0$. That is, we have to solve the optimization problem

$$\min_{N,M,K} \Big(\sup_{m \in \{0,\ldots,M\}} \mathrm{E}\big[\|X_{t_m} - Y_m^{M,N,K}\|_H^2\big] \Big)^{\frac{1}{2}} \quad \text{such that} \quad \mathrm{CC} = \bar{c}$$

for some $\bar{c} > 0$. As a result, we obtain the effective order of convergence, see also [62] for this concept. We compute and compare this value in the following sections for the schemes that we derive.

3.3 Derivation of Efficient Derivative-Free Milstein Scheme

We develop a numerical scheme to approximate the mild solution to SPDE (3.6) in a setting where (C1)-(C4) hold and impose a commutativity condition additionally; this reads as

(C5) $B'(y)\big(B(y)u, v\big) = B'(y)\big(B(y)v, u\big)$ for all $y \in H_\beta$, $u, v \in V_0$.

As explained in the introduction, the goal of this section is to devise an approximation scheme free of derivatives with reduced computational cost but theoretical rate of convergence as high as for the Milstein scheme presented in [35], that is, the scheme obtains a higher effective order of convergence. We call this scheme commutative derivative-free Milstein scheme (cDFM). Our approach to deal with the problem of high dimensionality is not restricted to equations that are pointwise multiplicative in the Q-Wiener process, which is the technique chosen in [35] and [71], but is applicable to a more general type of equation. Equations which do not fulfill this assumption involve an integral or derivative operator for the operator B, for example, see [43, 58, 64]. In the case of pointwise multiplicative operators B, the scheme by Jentzen and Röckner [35] and the derivative-free version introduced by Gan and Wang in [71] are indeed very efficient, and the cDFM obtains the same effective order of convergence, see Section 3.4. This is due to the fact, that in this setting the evaluation of the Jacobian can be done with less computational cost and therefore, the computational effort is of the same order of magnitude for all these schemes. It is, however, not our aim to construct approximation schemes for this class and in a general setting the effective order of convergence of the Milstein scheme in [35] can be improved. This issue is illustrated in the introduction in detail.

The approach that we present is based on an idea to reduce the computational cost by a factor depending on the dimensionality of the equation, originally designed for SODEs by Rößler, see [60, 61, 62], for example. In the error analysis of these numerical schemes for SODEs, the number of Brownian motions $K \in \mathbb{N}$ is a factor in the constant, however. Therefore, we need to design a scheme tailored to SPDEs and prove its convergence by different means. In the setting of SPDEs, the idea is even more powerful as it increases the effective order of convergence.
The operator that we propose to approximate $B'B$ is more flexible than the operator BB defined in [71] as it does not have to fulfill assumptions (1.4) and (1.5). Moreover, we do not have to assume the relation $B(y(x))v(x) = b(x, y(x)) \cdot v(x)$ for all $x \in (0,1)^d$, $y \in H_\beta$, $v \in V_0$, $d = 1, 2, 3$ and the approximations $BB(y, h)(\cdot, \cdot)$, $h > 0$, of the operators $B'(y)B(y)(\cdot, \cdot)$ do not need to be bilinear operators.
By a careful choice of the approximation operator, we maintain the theoretical order of the Milstein scheme in [35] with respect to the spatial and time discretizations, whereas we reduce the large number of function evaluations by one order of magnitude. The commutative derivative-free Milstein scheme that we derive in the following is also described in our article [46].

In the numerical analysis of SPDEs, there are different spaces that need to be approximated. We start with the infinite dimensional solution space H. Concerning the approximation of the space domain, there exist various numerical schemes designed for PDEs which are of interest in the approximation of SPDEs as well. We can make use of methods such as finite differences, finite elements, or spectral Galerkin approximations, as in Example 3.3, which are employed in [1, 35, 41, 49, 70, 78], for example. Here, we decide for a spectral Galerkin projection; this type of approximation has also been chosen in [1, 24, 35], or [71], for instance.
Let P_N denote a projection operator for some $N \in \mathbb{N}$ and define a finite index set $\mathcal{I}_N \subset \mathcal{I}$ with

$|\mathcal{I}_N| = N$. Then $P_N : H \to H_N$ maps H to a finite dimensional subspace $H_N := \operatorname{span}(e_1, \ldots, e_N)$ and is given by

$$P_N y = \sum_{i \in \mathcal{I}_N} \langle y, e_i \rangle_H e_i \qquad \text{for all } y \in H,\ N \in \mathbb{N}.$$

The approximation of the Q-Wiener process is represented similarly. We define $\mathcal{J}_K \subset \mathcal{J}$ with $|\mathcal{J}_K| = K$ and the operator $P_K : V \to V_K$ for all $K \in \mathbb{N}$ as

$$W_t^K := P_K W_t = \sum_{j \in \mathcal{J}_K} \langle W_t, \tilde{e}_j \rangle_V \tilde{e}_j = \sum_{j \in \mathcal{J}_K} \sqrt{\eta_j} \beta_t^j \tilde{e}_j \qquad P\text{-a.s.} \tag{3.10}$$

We use the notation $P_N X_t = X_t^N$, $t \in [0,T]$, and obtain a finite dimensional system of SODEs in H_N for all $N, K \in \mathbb{N}$

$$\begin{aligned} \mathrm{d}X_t^N &= \big(P_N A X_t^N + P_N F(X_t^N)\big)\,\mathrm{d}t + P_N B(X_t^N)\,\mathrm{d}W_t^K, \quad t \in (0,T], \\ X_0^N &= P_N \xi. \end{aligned}$$

Now, merely the temporal discretization is missing. For legibility, we consider an equidistant partition of the time interval $[0,T]$ as $t_m = m \cdot h$, $m \in \{0, \ldots, M\}$, with time step $h = \frac{T}{M}$ for $M \in \mathbb{N}$. We define the increments of the Q-Wiener process as follows

$$\Delta W_m^{K,M}(\omega) := W_{t_{m+1}}^K(\omega) - W_{t_m}^K(\omega) = \sum_{j \in \mathcal{J}_K} \sqrt{\eta_j} \Delta \beta_m^j(\omega) \tilde{e}_j$$

with $\Delta \beta_m^j(\omega) = \beta_{t_{m+1}}^j(\omega) - \beta_{t_m}^j(\omega)$ for all $\omega \in \Omega$, $m \in \{0, \ldots, M-1\}$, $j \in \mathcal{J}_K$, $M, K \in \mathbb{N}$.

Since we assume commutativity as stated in (C5) and by the expression for the finite dimensional stochastic double integral (1.2), we can split the iterated integrals as

$$\begin{aligned} & e^{A(t-s)} \int_s^t B'(X_s)\Big(\int_s^r B(X_s)\,\mathrm{d}W_u^K\Big)\,\mathrm{d}W_r^K \\ &= e^{A(t-s)} \Big(\frac{1}{2} B'(X_s)\big(B(X_s)(W_t^K - W_s^K), (W_t^K - W_s^K)\big) - \frac{t-s}{2} \sum_{j \in \mathcal{J}_K} \eta_j B'(X_s)\big(B(X_s)\tilde{e}_j, \tilde{e}_j\big)\Big) \end{aligned} \tag{3.11}$$

P-a.s. for all $s, t \in [0,T]$, $s \leq t$, $K \in \mathbb{N}$. This expression can easily be simulated and a proof, which mainly employs (1.2), can be found in [35].
Let us fix some arbitrary $M, N, K \in \mathbb{N}$. In this setting the Milstein scheme in [35], denoted as MIL in this work, reads $Y_0^{N,K,M} = P_N \xi$ and

$$\begin{aligned} Y_{m+1}^{N,K,M} = P_N \bigg(e^{Ah} \Big(& Y_m^{N,K,M} + hF(Y_m^{N,K,M}) + B(Y_m^{N,K,M}) \Delta W_m^{K,M} \\ & + \frac{1}{2} B'(Y_m^{N,K,M})\big(B(Y_m^{N,K,M})\Delta W_m^{K,M}, \Delta W_m^{K,M}\big) \\ & - \frac{h}{2} \sum_{j \in \mathcal{J}_K} \eta_j\, B'(Y_m^{N,K,M})\big(B(Y_m^{N,K,M})\tilde{e}_j, \tilde{e}_j\big) \Big) \bigg) \end{aligned}$$

for all $m \in \{0, 1, \ldots, M-1\}$.

Now, we carefully choose the approximation operator of $B'B$ such that we obtain a scheme with computational cost of optimal order. The cDFM reads $Y_0^{N,K,M} = P_N\xi$ and

$$\begin{aligned} Y_{m+1}^{N,K,M} = P_N\Big(e^{Ah}\Big(&Y_m^{N,K,M} + hF(Y_m^{N,K,M}) + B(Y_m^{N,K,M})\Delta W_m^{K,M} \\ &+ \frac{1}{\sqrt{h}}\Big(B\Big(Y_m^{N,K,M} + \frac{1}{2}\sqrt{h}\,P_N B(Y_m^{N,K,M})\Delta W_m^{K,M}\Big) - B(Y_m^{N,K,M})\Big)\Delta W_m^{K,M} \\ &+ \sum_{j\in\mathcal{J}_K} \bar{B}(Y_m^{N,K,M},h,j)\Big)\Big) \end{aligned} \tag{3.12}$$

for all $m \in \{0,1,\ldots,M-1\}$, where, in general, we choose

$$\bar{B}(Y_m^{N,K,M},h,j) = B\Big(Y_m^{N,K,M} - \frac{h}{2}\sqrt{\eta_j}\,P_N B(Y_m^{N,K,M})\tilde{e}_j\Big)\sqrt{\eta_j}\,\tilde{e}_j - B(Y_m^{N,K,M})\sqrt{\eta_j}\,\tilde{e}_j$$

for all $j \in \mathcal{J}_K$, $M,N,K \in \mathbb{N}$, see also [46]. If B is pointwise multiplicative in the Q-Wiener process, we define the approximation operator differently; this operator is specified in equation (3.21) in the next section.

Remark 3.2

In [20], the authors showed that the exponential term e^{At} in the Milstein scheme can be approximated by $(I-At)^{-1}$, $t\in[0,T]$, without a reduction in the theoretical order of convergence. The same should hold true for the cDFM.

3.4 Error Analysis and Computational Cost

For the commutative derivative-free Milstein scheme in (3.12), we obtain the following error estimate, which is the same, apart from constants, as for the Milstein scheme in [35]. We impose assumptions which differ only slightly from the conditions that are necessary to prove the convergence of the Milstein scheme.

Theorem 3.1 (Convergence of cDFM)

Let assumptions (C1)–(C5) be fulfilled. Then, there exists a constant $C_{T,Q}\in(0,\infty)$, independent of N, K, and M, such that for $(Y_m^{N,K,M})_{0\le m\le M}$, defined by the commutative derivative-free Milstein scheme in (3.12)*, it holds*

$$\Big(\mathrm{E}\Big[\big\|X_{t_m} - Y_m^{N,K,M}\big\|_H^2\Big]\Big)^{\frac{1}{2}} \le C_{T,Q}\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-\gamma} + \Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Big)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)}\Big)$$

for all $m\in\{0,1,\ldots,M\}$ and all $N,K,M\in\mathbb{N}$. The parameter values are determined by (C1)–(C4).

Proof. The proof of Theorem 3.1 can be found in Section 3.5, see also our article [46]. □

As we do not only want to compare our scheme to the Milstein scheme but to some type of exponential Euler scheme (EES) as well, analyzed in [49] or in [32] in a slightly different form, we give a short proof of the error estimate adjusted to our setting and notation. Let $M, N, K \in \mathbb{N}$ be arbitrarily fixed; the exponential Euler scheme reads as $Y_0^{EES} = P_N\xi$ and

$$Y_{m+1}^{EES} = P_N\Big(e^{Ah}Y_m^{EES} + A^{-1}(e^{Ah} - I)F(Y_m^{EES}) + e^{Ah}B(Y_m^{EES})\Delta W_m^{K,M}\Big) \tag{3.13}$$

for all $m \in \{0, 1, \ldots, M-1\}$. In this form, the scheme has been introduced in [49] with a finite element discretization. We obtain the following estimate for this scheme.

Theorem 3.2 (Convergence of EES)
Let assumptions (C1)–(C4) be fulfilled. Then, there exists a constant $C_T \in (0, \infty)$, independent of N, K, and M, such that for $(Y_m^{EES})_{0\leq m\leq M}$, defined by the exponential Euler scheme in (3.13), it holds

$$\Big(\mathrm{E}\Big[\big\|X_{t_m} - Y_m^{EES}\big\|_H^2\Big]\Big)^{\frac{1}{2}} \leq C_T\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-\gamma} + \Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Big)^{\alpha} + M^{-\min(\frac{1}{2},\gamma,2(\gamma-\beta))}\Big)$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M, K \in \mathbb{N}$. The parameter values are determined by (C1)–(C4).

Proof. For a proof, we refer to Section 3.5. We present a proof tailored to our setting and employ computations similar to those in the proof of convergence of the Milstein scheme in [35]. The proof differs from the representation in [49] as we do employ a Galerkin approximation instead of a finite element discretization in the numerical scheme. Moreover, in [49], the operators F and B are defined on H which yields a slightly differing result. The proof of the exponential Euler scheme in [32] employs stochastic trees. We do not use this technique. □

Besides, we compare the cDFM to the linear implicit Euler scheme (LIE) in the numerical simulations. This scheme reads as $Y_0^{LIE} = P_N\xi$ and

$$Y_{m+1}^{LIE} = P_N\Big((I - hA)^{-1}\Big(Y_m^{LIE} + hF(Y_m^{LIE}) + B(Y_m^{LIE})\Delta W_m^{K,M}\Big)\Big) \tag{3.14}$$

for all $m \in \{0, 1, \ldots, M-1\}$ and $M, N, K \in \mathbb{N}$, see [40]. Similar as for the EES, we obtain the error estimate for this scheme, which attains the same order. This can be proved by combining estimates from the proof of Theorem 3.2 with the error estimate for the linear implicit Euler scheme in [78] adjusted to our setting.

Our goal is to determine the effective order of convergence for the schemes illustrated above. Therefore, we need to compute their computational costs and combine them with their respective theoretical order of convergence. The following results can also be found in [46].
Let us fix $M, N, K \in \mathbb{N}$; first, we consider the exponential Euler scheme as this is the base for the analysis of the other schemes as well. In order to simulate one path with the EES, we need to evaluate $P_NF(\cdot)$ and $P_NB(\cdot)|_{V_K}$ in each time step $m \in \{0, \ldots, M\}$. Moreover, we compute $P_NB(Y_m^{N,K,M})|_{V_K}$, $m \in \{0, \ldots, M\}$, as described in Section 3.2. We need K realizations of standard normal random variables (rv) $\epsilon \sim N(0, 1)$ in each step, where we assume $\mathrm{cost}(\epsilon) = 1$.

In total, we obtain cost(EES) $= \mathcal{O}(MNK)$.
For the commutative derivative-free Milstein scheme, we have to compute the same terms. Additionally, we need to compute the expression $P_N \bar{B}(Y_m^{N,M,K}, h, j)u$ for $Y_m^{N,M,K} \in H_N$, $m \in \{0, \ldots, M\}$, $j \in \mathcal{J}_K$, $u \in V_K$ with

$$P_N \bar{B}(Y_m^{N,M,K}, h, j)u = \sum_{i \in \mathcal{I}_N} \sum_{j \in \mathcal{J}_K} \langle \bar{B}(Y_m^{N,M,K}, h, j)\tilde{e}_j, e_i \rangle_H \langle u, \tilde{e}_j \rangle_V e_i.$$

The computation of $\langle \bar{B}(Y_m^{N,M,K}, h, j)\tilde{e}_j, e_i \rangle_H$ for all $i \in \mathcal{I}_N$, $j \in \mathcal{J}_K$ has cost of NK in each time step. As described in Section 3.2, we obtain total costs of $\mathcal{O}(NK)$ for this term. Therefore, the computational cost of the cDFM differs by a constant compared to the EES only, and it holds cost(cDFM) $= \mathcal{O}(MNK)$.
For the Milstein scheme, the computational cost is higher by one order of magnitude in N. We need to compute terms $P_N B'(Y_m^{N,M,K})(B(Y_m^{N,M,K})u, v)$ for $Y_m^{N,M,K} \in H_N$, $m \in \{0, \ldots, M\}$, $u, v \in V_K$ along with the terms that have to be computed in the EES. This yields cost(MIL) $= \mathcal{O}(MN^2K)$ to simulate one path with the Milstein scheme.

The main differences between these schemes are summarized in the following table.

	computational cost for evaluation of			
Scheme	$P_N F(x)$	$P_N B(x)$	$P_N B'(x)$	# of $N(0,1)$ rv
MIL	N	KN	KN^2	K
LIE	N	KN	–	K
EES	N	KN	–	K
cDFM	N	$3KN$	–	K

Table 3.1: Number of (nonlinear) function evaluations and random variables for each time step.

In Table 3.1, we observe that the cDFM obtains computational costs which are optimal in some sense. The term $P_N B(y)u$, $y \in H_N$, $u \in V_K$, which is included in all the numerical schemes that we consider, yields costs of $\mathcal{O}(MNK)$. Therefore, we cannot reduce the computational effort below this magnitude.

We solve the following optimization problem to determine the effective order of convergence

$$\min_{N,M,K} \sup_{m \in \{0,\ldots,M\}} \Big(\mathrm{E}\Big[\big\| X_{t_m} - Y_m^{M,N,K} \big\|_H^2 \Big] \Big)^{\frac{1}{2}} \quad \text{such that} \quad \mathrm{CC} = \bar{c}$$

for some fixed $\bar{c} > 0$. To this end, we assume a relationship of the eigenvalues of A and Q with the dimensions of the corresponding projection spaces. Precisely, we assume $\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i = \mathcal{O}(N^{\rho_A})$ and $\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j = \mathcal{O}(K^{-\rho_Q})$ for some $\rho_A, \rho_Q > 0$. Therewith, we obtain an estimate of the error for some numerical scheme as

$$\mathrm{error}(\mathrm{SCHEME}(N,K,M)) = \sup_{m\in\{0,\dots,M\}} \Big(\mathrm{E}\Big[\big\|X_{t_m} - Y_m^{M,N,K}\big\|_H^2\Big]\Big)^{\frac{1}{2}} \leq N^{-\gamma\rho_A} + K^{-\alpha\rho_Q} + M^{-q} \tag{3.15}$$

for some $q > 0$, which is determined by the approximation scheme, and all $N, M, K \in \mathbb{N}$.

General Setting

Let $\bar{c} > 0$ be arbitrarily fixed. As described above, we have $\mathrm{cost}(\mathrm{EES}) = \mathcal{O}(MNK)$ for the exponential Euler scheme. Therefore, we obtain an optimal choice of M, N, K as

$$N = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\rho_Q q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big), \quad K = \mathcal{O}\Big(\bar{c}^{\frac{\gamma\rho_A q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big), \quad M = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\gamma\rho_A\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big).$$

These values balance the error term, which yields the effective order of convergence

$$\mathrm{error}(\mathrm{EES}(N,K,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\alpha\gamma\rho_A\rho_Q q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big). \tag{3.16}$$

For this scheme, it holds $q = q_{EES} = \min(2(\gamma-\beta), \gamma, \frac{1}{2})$, see Theorem 3.2.
We optimize the error for the Milstein scheme in the same fashion. Here, the parameter q equals $q = q_{MIL} = \min(2(\gamma-\beta), \gamma)$, [35]. We obtain

$$N = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\rho_Q q}{(2\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big), \quad K = \mathcal{O}\Big(\bar{c}^{\frac{\gamma\rho_A q}{(2\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big), \quad M = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\gamma\rho_A\rho_Q}{(2\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big),$$

and the optimal order as

$$\mathrm{error}(\mathrm{MIL}(N,K,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\alpha\gamma\rho_A\rho_Q q}{(2\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big). \tag{3.17}$$

Finally, we investigate the effective order of convergence for the commutative derivative-free Milstein scheme. On first sight, we obtain the same result as for the EES, that is

$$N = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\rho_Q q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big), \quad K = \mathcal{O}\Big(\bar{c}^{\frac{\gamma\rho_A q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big), \quad M = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\gamma\rho_A\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big) \tag{3.18}$$

and

$$\mathrm{error}(\mathrm{cDFM}(N,K,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\alpha\gamma\rho_A\rho_Q q}{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}}\Big). \tag{3.19}$$

The parameter q, however, differs. As in the Milstein scheme, we have $q = q_{cDFM} = \min(2(\gamma - \beta), \gamma)$, see Theorem 3.1.
Next, we compare the effective orders of convergence for the different schemes. We consider the reciprocal of the effective order of the cDFM as this term allows to clearly identify the dependence on q

$$\frac{(\alpha\rho_Q+\gamma\rho_A)q+\alpha\gamma\rho_A\rho_Q}{\alpha\gamma\rho_A\rho_Q q} = \frac{\alpha\rho_Q+\gamma\rho_A}{\alpha\gamma\rho_A\rho_Q} + \frac{1}{q}.$$

The reciprocal decreases with increasing q. As $q_{EES} \leq q_{cDFM}$, the effective order of convergence of the cDFM is, in general, higher than for the EES.
A comparison of the cDFM and the Milstein scheme clearly reveals a preference for the cDFM as the order of convergence differs by a factor 2 in the denominator, whereas all parameters take the same values. Thus, the optimal order of the cDFM is higher if $\alpha, \gamma, \rho_Q, \rho_A, q > 0$.

Pointwise Multiplicative Diffusion

If we restrict the operator B to be pointwise multiplicative in the Q-Wiener process, this is the setting considered in [35] and [71], we get differing results. Let $H = V = L^2((0,1), \mathbb{R})$, $(F(y))(x) = f(x, y(x))$, and $(B(y)v)(x) = b(x, y(x)) \cdot v(x)$ for all $y \in H_\beta$, $\beta \in [0,1)$, $v \in V_0$, and $x \in (0,1)$ with $b, f : [0,1] \times \mathbb{R} \to \mathbb{R}$.
As mentioned before, the operator $\bar{B}(y,h,j)$ in the cDFM is chosen differently in this case. Precisely, we compute $Y_0^{N,K,M} = P_N \xi$ and

$$\begin{aligned} Y_{m+1}^{N,K,M} = P_N \Big(e^{Ah} \Big(& Y_m^{N,K,M} + h f(\cdot, Y_m^{N,K,M}) + b(\cdot, Y_m^{N,K,M}) \cdot \Delta W_m^{K,M} \\ & + \frac{1}{\sqrt{h}} \Big(b\Big(\cdot, Y_m^{N,K,M} + \frac{1}{2}\sqrt{h}\, P_N b(\cdot, Y_m^{N,K,M}) \cdot \Delta W_m^{K,M}\Big) - b(\cdot, Y_m^{N,K,M}) \Big) \cdot \Delta W_m^{K,M} \\ & + \sum_{j \in \mathcal{J}_K} \bar{B}(Y_m^{N,K,M}, h, j) \Big) \Big) \end{aligned} \tag{3.20}$$

for all $m \in \{0, \ldots, M-1\}$ with

$$\bar{B}(Y_m^{N,K,M}, h, j) = \Big(b\Big(\cdot, Y_m^{N,K,M} - \frac{h}{2} P_N b(\cdot, Y_m^{N,K,M})\Big) - b(\cdot, Y_m^{N,K,M}) \Big) \eta_j \tilde{e}_j^2 \tag{3.21}$$

for all $j \in \mathcal{J}_K$ and $M, N, K \in \mathbb{N}$. We call this scheme the multiplicative version of the commutative derivative-free Milstein scheme (cDFMM), see also [46]. The computational cost of the multiplicative scheme is lower than for the cDFM as we only need to compute terms such as $P_N b(\cdot, Y_m^{N,M,K}(\cdot))$ here with $\operatorname{cost}(P_N b(\cdot, Y_m^{N,M,K}(\cdot)) = N$ in each time step $m \in \{0, \ldots, M\}$. The same holds true for the other terms involved in this scheme. Moreover, we need to simulate K independent standard normal distributed random variables in each time step. In total, we have $\operatorname{cost}(\text{cDFMM}) = \mathcal{O}(MN + MK)$.
We determine the effective order of convergence as above. Let $\bar{c} > 0$ be arbitrarily fixed; we minimize the error (3.15) such that $\mathcal{O}(MN + MK) = \bar{c}$. This yields a reasonable choice of

$$N = \mathcal{O}\Big(\bar{c}^{\frac{\min(\gamma\rho_A, \alpha\rho_Q) q}{\gamma\rho_A(\min(\gamma\rho_A, \alpha\rho_Q)+q)}}\Big), \qquad K = \mathcal{O}\Big(\bar{c}^{\frac{\min(\gamma\rho_A, \alpha\rho_Q) q}{\alpha\rho_Q(\min(\gamma\rho_A, \alpha\rho_Q)+q)}}\Big), \qquad M = \mathcal{O}\Big(\bar{c}^{\frac{\min(\gamma\rho_A, \alpha\rho_Q)}{\min(\gamma\rho_A, \alpha\rho_Q)+q}}\Big)$$

with $q = q_{cDFMM} = \min(2(\gamma - \beta), \gamma)$. For the effective order of convergence, we get

$$\operatorname{error}(\text{cDFMM}(N,K,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\min(\gamma\rho_A, \alpha\rho_Q) q}{\min(\gamma\rho_A, \alpha\rho_Q)+q}}\Big). \tag{3.22}$$

The Milstein scheme in [35] and the Runge-Kutta type scheme in [71] obtain the same effective order in this setting. The computational cost of the Milstein scheme is significantly reduced, compared to the general setting, as the evaluation of the Jacobian produces cost of N in each time step only. Therefore, it efficiently approximates the solution process $(X_t)_{t\in[0,T]}$ in T. The Runge-Kutta type scheme in [71] does not involve this derivative, which may help to reduce computation time even more.

For the EES, we obtain the same expression as well, but one has to keep in mind that the parameter $q = q_{EES} = \min(2(\gamma-\beta), \gamma, \frac{1}{2})$ differs. We do, however, not restrict our analysis to this special case in the following.

Finite Dimensional Noise

We detail this exception here for completeness and since we illustrate one such example in Section 3.6. Let $K \in \mathbb{N}$ be fixed and assume $|\{\eta_j : \ j \in \mathcal{J}\}| = K$, that is, we choose $\mathcal{J}_K = \mathcal{J}$. Then, there is no error from the approximation of the Q-Wiener process and we analyze

$$\text{error}(\text{SCHEME}(N,M)) = \sup_{m\in\{0,\ldots,M\}} \Big(\mathrm{E}\Big[\big\|X_{t_m} - Y_m^{M,N}\big\|_H^2\Big]\Big)^{\frac{1}{2}} \leq N^{-\gamma\rho_A} + M^{-q} \tag{3.23}$$

for all $M, N \in \mathbb{N}$.

We consider the commutative derivative-free Milstein scheme first and obtain computational costs of $\text{cost}(\text{cDFM}(N,M)) = \mathcal{O}(MN)$. The minimization of $\text{error}(\text{cDFM}(N,M))$ such that $\mathcal{O}(MNK) = \bar{c}$, for fixed $\bar{c} > 0$, yields an optimal choice of $N = \mathcal{O}\big(\bar{c}^{\frac{q}{\gamma\rho_A+q}}\big)$ and $M = \mathcal{O}\big(\bar{c}^{\frac{\gamma\rho_A}{\gamma\rho_A+q}}\big)$. Then, the effective order of convergence for the cDFM is

$$\text{error}(\text{cDFM}(N,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\gamma\rho_A q}{\gamma\rho_A+q}}\Big). \tag{3.24}$$

As in the previous part, the Milstein scheme has computational costs of order $\text{cost}(\text{MIL}(N,M)) = \mathcal{O}(N^2M)$. The optimization of $\text{error}(\text{MIL}(N,M))$ yields $N = \mathcal{O}\big(\bar{c}^{\frac{q}{\gamma\rho_A+2q}}\big)$ and $M = \mathcal{O}\big(\bar{c}^{\frac{\gamma\rho_A}{\gamma\rho_A+2q}}\big)$. The effective order of convergence for the Milstein scheme is

$$\text{error}(\text{MIL}(N,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\gamma\rho_A q}{\gamma\rho_A+2q}}\Big). \tag{3.25}$$

In both these schemes, it holds $q = \min(2(\gamma-\beta), \gamma)$. Therefore, as in the general case, the commutative derivative-free Milstein scheme obtains a higher effective order of convergence than the Milstein scheme.

If the operators are pointwise multiplicative additionally, the multiplicative version of the cDFM is employed. The computational cost for this scheme is $\text{cost}(\text{cDFMM}(N,M)) = \mathcal{O}(MN+MK)$. This yields a reasonable choice of $N = \mathcal{O}\big(\bar{c}^{\frac{q}{\gamma\rho_A+q}}\big)$ and $M = \mathcal{O}\big(\bar{c}^{\frac{\gamma\rho_A}{\gamma\rho_A+q}}\big)$. The effective order of convergence results in

$$\text{error}(\text{cDFMM}(N,M)) = \mathcal{O}\Big(\bar{c}^{-\frac{\gamma\rho_A q}{\gamma\rho_A+q}}\Big). \tag{3.26}$$

As in the last part the Milstein and the Runge-Kutta type scheme obtain the same effective order of convergence for this particular case. Furthermore, observe that in this special case the cDFM obtains the same order as the cDFMM as well. So, we can choose the cDFM instead.

Again, for the EES, we get the same expressions as for the cDFM, but $q_{EES} = \min(2(\gamma-\beta), \gamma, \frac{1}{2})$.

For the linear implicit Euler scheme, we obtain the same results as for the exponential Euler scheme in all cases as the computational cost as well as the theoretical order of convergence agree, see page 40.

The following table combines the results.

Scheme	Effective Order		
	general	pointwise multiplicative	finite noise
MIL	$\frac{\alpha\gamma\rho_A\rho_Q q_{MIL}}{(2\alpha\rho_Q+\gamma\rho_A)q_{MIL}+\alpha\gamma\rho_A\rho_Q}$	$\frac{\min(\gamma\rho_A,\alpha\rho_Q)q_{MIL}}{\min(\gamma\rho_A,\alpha\rho_Q)+q_{MIL}}$ *	$\frac{\gamma\rho_A q_{MIL}}{\gamma\rho_A+2q_{MIL}}$
LIE	$\frac{\alpha\gamma\rho_A\rho_Q q_{EES}}{(\alpha\rho_Q+\gamma\rho_A)q_{EES}+\alpha\gamma\rho_A\rho_Q}$	$\frac{\min(\gamma\rho_A,\alpha\rho_Q)q_{EES}}{\min(\gamma\rho_A,\alpha\rho_Q)+q_{EES}}$	$\frac{\gamma\rho_A q_{EES}}{\gamma\rho_A+q_{EES}}$
EES	$\frac{\alpha\gamma\rho_A\rho_Q q_{EES}}{(\alpha\rho_Q+\gamma\rho_A)q_{EES}+\alpha\gamma\rho_A\rho_Q}$	$\frac{\min(\gamma\rho_A,\alpha\rho_Q)q_{EES}}{\min(\gamma\rho_A,\alpha\rho_Q)+q_{EES}}$	$\frac{\gamma\rho_A q_{EES}}{\gamma\rho_A+q_{EES}}$
cDFM or cDFMM	$\frac{\alpha\gamma\rho_A\rho_Q q_{cDFM}}{(\alpha\rho_Q+\gamma\rho_A)q_{cDFM}+\alpha\gamma\rho_A\rho_Q}$ *	$\frac{\min(\gamma\rho_A,\alpha\rho_Q)q_{cDFM}}{\min(\gamma\rho_A,\alpha\rho_Q)+q_{cDFM}}$ *	$\frac{\gamma\rho_A q_{cDFM}}{\gamma\rho_A+q_{cDFM}}$ *

Table 3.2: Effective order of convergence for commutative SPDE - by $*$ we indicate the scheme that is superior for each setting, given that $\gamma, \rho_A, \alpha, \rho_Q, q > 0$ and $q_{EES} < q_{cDFM}$. Moreover, it holds $q_{MIL} = q_{cDFM}$.

We compare the effective order for various numerical examples after we present the proofs of convergence of the approximation schemes in the next section.

3.5 Proofs

Proof of Theorem 3.1

Theorem (Convergence of cDFM)
Let assumptions (C1)–(C5) be fulfilled. Then, there exists a constant $C_{T,Q} \in (0,\infty)$, independent of N, K, and M, such that for $(Y_m^{N,K,M})_{0\leq m\leq M}$, defined by the commutative derivative-free Milstein scheme in (3.12)*, it holds*

$$\left(\mathrm{E}\left[\left\|X_{t_m} - Y_m^{N,K,M}\right\|_H^2\right]\right)^{\frac{1}{2}} \leq C_{T,Q}\left(\left(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\right)^{-\gamma} + \left(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j\right)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)}\right)$$

for all $m \in \{0,1,\ldots,M\}$ and $N, K, M \in \mathbb{N}$. The parameter values are determined by (C1)–(C4).

The proof of Theorem 3.1 builds on the proof of convergence in [35] - with an addition which accounts for the approximation of the derivative, see also [46]. We do not incorporate the analysis of the error which may result from the approximation of the coefficients in the spectral projection $P_N y = \sum_{n\in\mathcal{I}_N} \langle y, e_n\rangle_H e_n$, $N \in \mathbb{N}$, $y \in H$, here.

In the proof, we use some generic constant C which may change from line to line.

Proof of Theorem 3.1.
We adopt the representation

$$X_{t_m} = e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}F(X_s)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}B(X_s)\,\mathrm{d}W_s,$$

set $Y_m := Y_m^{N,M,K}$ and $W_m^K := W_m^{K,M}$ for all $m \in \{0,\dots,M\}$, $M,N,K \in \mathbb{N}$, for legibility, and define some auxiliary processes for all $m \in \{0,\dots,M\}$, $M,N,K \in \mathbb{N}$ as

$$\begin{aligned}
\bar{X}_{t_m} :=& P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(X_{t_l})\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(X_{t_l})\,\mathrm{d}W_s^K \\
&+ \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l})\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big),\\
\bar{Y}_{t_m} :=& P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l)\,\mathrm{d}W_s^K \\
&+ \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B'(Y_l)\Big(\int_{t_l}^{s} B(Y_l)\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big)\\
=& P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l)\,\mathrm{d}W_s^K \\
&+ \sum_{l=0}^{m-1} e^{A(t_m-t_l)}\Big(\frac{1}{2}B'(Y_l)\big(B(Y_l)\Delta W_l^K, \Delta W_l^K\big) - \frac{h}{2}\sum_{j\in\mathcal{J}_K}\eta_j B'(Y_l)\big(B(Y_l)\tilde{e}_j,\tilde{e}_j\big)\Big)\Big).
\end{aligned}$$

For all $m \in \{0,\dots,M\}$, $M,N \in \mathbb{N}$, we estimate

$$\mathrm{E}\big[\|X_{t_m} - Y_m\|_H^2\big] = \mathrm{E}\big[\|X_{t_m} - P_NX_{t_m} + P_NX_{t_m} - \bar{X}_{t_m} + \bar{X}_{t_m} - \bar{Y}_{t_m} + \bar{Y}_{t_m} - Y_m\|_H^2\big]$$

in several parts

$$\begin{aligned}
\mathrm{E}\big[\|X_{t_m} - Y_m\|_H^2\big] \leq& 4\Big(\mathrm{E}\big[\|X_{t_m} - P_NX_{t_m}\|_H^2\big] + \mathrm{E}\big[\|P_NX_{t_m} - \bar{X}_{t_m}\|_H^2\big] \\
&+ \mathrm{E}\big[\|\bar{X}_{t_m} - \bar{Y}_{t_m}\|_H^2\big] + \mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big]\Big).
\end{aligned} \tag{3.27}$$

The first term is the error that results from the projection of H to a finite dimensional subspace H_N, $N \in \mathbb{N}$. The second and third term arise due to the approximation of the solution process with the Milstein scheme and the last one is the error that we obtain by approximating the operator $B'B$. The estimates of the first three terms are not specific to our scheme and the ideas originate from [35]. For completeness, we state the whole proof.

For all $m \in \{1,\dots,M\}$, $M,N,K \in \mathbb{N}$, we want to prove the error from the Galerkin projection,

$$\big(\mathrm{E}\big[\|P_NX_{t_m} - \bar{X}_{t_m}\|\big]_H^2\big)^{\frac{1}{2}} \leq \Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}F(X_s) - e^{A(t_m-t_l)}F(X_{t_l})\Big)\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}}$$

$$
\begin{aligned}
&+\Bigg(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}B(X_s)-e^{A(t_m-t_l)}B(X_{t_l})\Big)\,\mathrm{d}W_s^K \\
&\qquad-\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}e^{A(t_m-t_l)}B'(X_{t_l})\Big(\int_{t_l}^{s}B(X_{t_l})\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big\|_H^2\Big]\Bigg)^{\frac{1}{2}} \\
&+\Bigg(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}e^{A(t_m-s)}B(X_s)\,\mathrm{d}(W_s-W_s^K)\Big\|_H^2\Big]\Bigg)^{\frac{1}{2}} \\
&\le C_T\Big(M^{-\min(2(\gamma-\beta),\gamma)}+\Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Big)^{\alpha}\Big),
\end{aligned}
$$

$$
\mathrm{E}\big[\|\bar{X}_{t_m}-\bar{Y}_{t_m}\|\big]_H^2\le\frac{C_T}{M}\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l}-Y_l\|_H^2\big], \tag{3.28}
$$

and

$$
\mathrm{E}\big[\|\bar{Y}_{t_m}-Y_m\|\big]_H^2\le C_{T,Q}M^{-2}
$$

separately.

After estimating these expressions, we obtain

$$
\begin{aligned}
\mathrm{E}\big[\|X_{t_m}-Y_m\|_H^2\big]\le& C_T\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-2\gamma}+\Big(\sup_{j\in\mathcal{I}\setminus\mathcal{I}_K}\eta_j\Big)^{2\alpha}+M^{-2\min(2(\gamma-\beta),\gamma)}\Big) \\
&+\frac{C_T}{M}\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l}-Y_l\|_H^2\big]+C_{T,Q}M^{-2} \\
\le& C_{T,Q}\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-2\gamma}+\Big(\sup_{j\in\mathcal{I}\setminus\mathcal{I}_K}\eta_j\Big)^{2\alpha}+M^{-2\min(2(\gamma-\beta),\gamma)}\Big)
\end{aligned}
$$

for all $m\in\{1,\dots,M\}$, $M,N,K\in\mathbb{N}$, by a discrete version of Gronwall's Lemma.

Spectral Galerkin projection

The error resulting from the spectral Galerkin projection is estimated for all $m\in\{0,\dots,M\}$, $M,N\in\mathbb{N}$ as

$$
\begin{aligned}
\mathrm{E}\big[\|X_{t_m}-P_NX_{t_m}\|_H^2\big]&=\mathrm{E}\big[\|(I-P_N)X_{t_m}\|_H^2\big]\le\mathrm{E}\big[\|(I-P_N)(-A)^{-\gamma}\|_{L(H)}^2\big\|X_{t_m}\big\|_{H_\gamma}^2\big] \\
&=\sup_{\substack{y\in H\\\|y\|_H=1}}\|(I-P_N)(-A)^{-\gamma}y\|_H^2\,\mathrm{E}\big[\|X_{t_m}\|_{H_\gamma}^2\big] \\
&=\sup_{\substack{y\in H\\\|y\|_H=1}}\Big\|(I-P_N)\sum_{k\in\mathcal{I}}\lambda_k^{-\gamma}\langle y,e_k\rangle_He_k\Big\|_H^2\mathrm{E}\big[\|X_{t_m}\|_{H_\gamma}^2\big] \\
&=\sup_{\substack{y\in H\\\|y\|_H=1}}\Big\|\sum_{n\in\mathcal{I}\setminus\mathcal{I}_N}\langle\sum_{k\in\mathcal{I}}\lambda_k^{-\gamma}\langle y,e_k\rangle_He_k,e_n\rangle_He_n\Big\|_H^2\mathrm{E}\big[\|X_{t_m}\|_{H_\gamma}^2\big].
\end{aligned}
$$

Due to (C1)–(C4) and Theorem 2.7, we further obtain

$$\begin{aligned}
\mathrm{E}\big[\|X_{t_m} - P_N X_{t_m}\|_H^2\big] &= \sup_{\substack{y\in H\\ \|y\|_H=1}} \Big\| \sum_{n\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_n^{-\gamma} \langle y, e_n\rangle_H e_n \Big\|_H^2 \mathrm{E}\big[\|X_{t_m}\|_{H_\gamma}^2\big] \\
&\leq C \sup_{\substack{y\in H\\ \|y\|_H=1}} \sum_{n\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_n^{-2\gamma} \langle y, e_n\rangle_H^2 \\
&\leq C\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\Big)^{-2\gamma} \sup_{\substack{y\in H\\ \|y\|_H=1}} \sum_{n\in\mathcal{I}\setminus\mathcal{I}_N} \langle y, e_n\rangle_H^2 \\
&\leq C\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\Big)^{-2\gamma} \sup_{\substack{y\in H\\ \|y\|_H=1}} \|y\|_H \\
&= C\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\Big)^{-2\gamma}
\end{aligned}$$

for all $m \in \{0, \dots, M\}$, $M, N \in \mathbb{N}$. This proves the first part.

In the following, we use

$$\begin{aligned}
\|P_N y\|_H^2 = \Big\| \sum_{n\in\mathcal{I}_N} \langle y, e_n\rangle_H e_n \Big\|_H^2 &= \sum_{n\in\mathcal{I}_N} |\langle y, e_n\rangle_H|^2 \\
&\leq \sum_{n\in\mathcal{I}} |\langle y, e_n\rangle_H|^2 = \|y\|_H^2
\end{aligned}$$

for $y \in H$, $N \in \mathbb{N}$, several times.

Temporal discretization - Nonlinearity F

We prove the error resulting from the temporal discretization of the Bochner integral by partitioning the error into three components, which we again estimate separately. Let $m \in \{1, \dots, M\}$, $M \in \mathbb{N}$, we show

$$\begin{aligned}
&\bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m-s)}F(X_s) - e^{A(t_m-t_l)}F(X_{t_l})\Big)\,\mathrm{d}s \Big\|_H^2\Big]\bigg)^{\frac{1}{2}} \\
&\leq \bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\left(F(X_s) - F(X_{t_l})\right)\mathrm{d}s \Big\|_H^2\Big]\bigg)^{\frac{1}{2}} \\
&\quad + \bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-2} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m-s)} - e^{A(t_m-t_l)}\Big) F(X_{t_l})\,\mathrm{d}s \Big\|_H^2\Big]\bigg)^{\frac{1}{2}} \\
&\quad + \bigg(\mathrm{E}\Big[\Big\| \int_{t_{m-1}}^{t_m} \Big(e^{A(t_m-s)} - e^{A(t_m-t_{m-1})}\Big) F(X_{t_{m-1}})\,\mathrm{d}s \Big\|_H^2\Big]\bigg)^{\frac{1}{2}} \\
&\leq C_T M^{-\min(2(\gamma-\beta),\gamma)}.
\end{aligned}$$

We define $\tilde{X}_{s,l} := X_s - X_{t_l}$ for all $s \in [0, T]$, $l \in \{0, \dots, M-1\}$, $M \in \mathbb{N}$, for legibility. For the first term, we obtain by the triangle inequality and the representation of the mild solution

$(X_t)_{t\in[0,T]}$ for all $m \in \{1,\dots,M\}$, $M \in \mathbb{N}$

$$
\begin{aligned}
&\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\left(F(X_s)-F(X_{t_l})\right)\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq \Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}F'(X_{t_l})(X_s-X_{t_l})\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\quad+\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\Big(\int_0^1\int_0^r \frac{1}{2}F''(X_{t_l}+u\tilde{X}_{s,l})(\tilde{X}_{s,l},\tilde{X}_{s,l})\,\mathrm{d}u\, r\,\mathrm{d}r\Big)\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq \sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}F'(X_{t_l})\left(e^{A(s-t_l)}-I\right)X_{t_l}\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\quad+\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}F'(X_{t_l})\Big(\int_{t_l}^{s} e^{A(s-u)}F(X_u)\,\mathrm{d}u\Big)\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\quad+\Big(\sum_{l=0}^{m-1}\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}F'(X_{t_l})\Big(\int_{t_l}^{s} e^{A(s-u)}B(X_u)\,\mathrm{d}W_u\Big)\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\quad+\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\Big(\int_0^1\int_0^r \frac{1}{2}F''(X_{t_l}+u\tilde{X}_{s,l})(\tilde{X}_{s,l},\tilde{X}_{s,l})\,\mathrm{d}u\, r\,\mathrm{d}r\Big)\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}}.
\end{aligned}
$$

Then, Hölder's inequality implies for all $m \in \{1,\dots,M\}$, $M \in \mathbb{N}$

$$
\begin{aligned}
&\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\left(F(X_s)-F(X_{t_l})\right)\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq \sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\Big\|e^{A(t_m-s)}F'(X_{t_l})\left(e^{A(s-t_l)}-I\right)X_{t_l}\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad+\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\Big\|e^{A(t_m-s)}F'(X_{t_l})\Big(\int_{t_l}^{s} e^{A(s-u)}F(X_u)\,\mathrm{d}u\Big)\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad+\Big(\sum_{l=0}^{m-1}\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\Big\|e^{A(t_m-s)}F'(X_{t_l})\Big(\int_{t_l}^{s} e^{A(s-u)}B(X_u)\,\mathrm{d}W_u\Big)\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad+\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\Big\|e^{A(t_m-s)}\Big(\int_0^1\int_0^r F''(X_{t_l}+u\tilde{X}_{s,l})(X_s-X_{t_l},X_s-X_{t_l})\,\mathrm{d}u\, r\,\mathrm{d}r\Big)\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}}
\end{aligned}
$$

and by (C2), Theorem 2.3, and Theorem 2.7, we get

$$
\begin{aligned}
&\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\left(F(X_s)-F(X_{t_l})\right)\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq C\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\|F'(X_{t_l})\|_{L(H)}^2\|(-A)^{-\gamma}\big(e^{A(s-t_l)}-I\big)\|_{L(H)}^2\|X_{t_l}\|_{H_\gamma}^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad+C\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\|F'(X_{t_l})\|_{L(H)}^2\Big\|\int_{t_l}^{s} e^{A(s-u)}F(X_u)\,\mathrm{d}u\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + C\Big(\sum_{l=0}^{m-1} h\,\mathrm{E}\Big[\int_{t_l}^{t_{l+1}} \|F'(X_{t_l})\|_{L(H)}^2 \Big\|\int_{t_l}^{s} e^{A(s-u)}B(X_u)\,\mathrm{d}W_u\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad + C\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\int_0^1\int_0^r \|F''(X_{t_l}+u\tilde{X}_{s,l})\|_{L^{(2)}(H_\beta,H)}\,\mathrm{d}u\,r\,\mathrm{d}r\|X_s-X_{t_l}\|_{H_\beta}^4\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\leq C\sum_{l=0}^{m-1}\Big(h\int_{t_l}^{t_{l+1}}(s-t_l)^{2\gamma}\,\mathrm{E}\big[\|X_{t_l}\|_{H_\gamma}^2\big]\,\mathrm{d}s\Big)^{\frac{1}{2}} \\
&\quad + C\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[h\int_{t_l}^{t_{l+1}}\Big\|\int_{t_l}^{s} e^{A(s-u)}F(X_u)\,\mathrm{d}u\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad + C\Big(\sum_{l=0}^{m-1} h\,\mathrm{E}\Big[\int_{t_l}^{t_{l+1}}\Big\|\int_{t_l}^{s} e^{A(s-u)}B(X_u)\,\mathrm{d}W_u\Big\|_H^2\,\mathrm{d}s\Big]\Big)^{\frac{1}{2}} \\
&\quad + C\sum_{l=0}^{m-1}\Big(h\int_{t_l}^{t_{l+1}}(s-t_l)^{4\min(\gamma-\beta,\frac{1}{2})}\,\mathrm{d}s\Big)^{\frac{1}{2}}.
\end{aligned}
$$

Next, (C1)–(C4) and Itô's isometry imply

$$
\begin{aligned}
&\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\,(F(X_s)-F(X_{t_l}))\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq CMh^{1+\gamma} + C\sum_{l=0}^{m-1}\Big(h\int_{t_l}^{t_{l+1}}(s-t_l)^2\,\mathrm{d}s\Big)^{\frac{1}{2}} \\
&\quad + \Big(C\sum_{l=0}^{m-1} h\int_{t_l}^{t_{l+1}}\int_{t_l}^{s}\mathrm{E}\big[\|(-A)^{-\delta}\|_{L(H)}^2\,\|B(X_u)\|_{L_{HS}(V_0,H_\delta)}^2\big]\,\mathrm{d}u\,\mathrm{d}s\Big)^{\frac{1}{2}} \\
&\quad + \sum_{l=0}^{m-1}\Big(h^{4\min(\gamma-\beta,\frac{1}{2})+2}\Big)^{\frac{1}{2}} \\
&\leq C_T h^{\gamma} + CMh^2 + \Big(C\sum_{l=0}^{m-1} h\int_{t_l}^{t_{l+1}}(s-t_l)\,\mathrm{d}s\Big)^{\frac{1}{2}} + h^{\min(2(\gamma-\beta),1)} \\
&\leq C_T h^{\gamma} + C_T h + C\big(Mh^3\big)^{\frac{1}{2}} \leq C_T h^{\min(2(\gamma-\beta),\gamma)}
\end{aligned}
$$

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

The estimates of the second and third part follow easily by the triangle inequality, Hölder's inequality, (C1)–(C4), and Theorem 2.3 as well. For all $m\in\{1,\ldots,M\}$, $M\in\mathbb{N}$, we get

$$
\begin{aligned}
&\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-2}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}-e^{A(t_m-t_l)}\Big)F(X_{t_l})\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq \sum_{l=0}^{m-2}\Big(\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}-e^{A(t_m-t_l)}\Big)F(X_{t_l})\,\mathrm{d}s\Big\|_H^2\Big]\Big)^{\frac{1}{2}} \\
&\leq C\sum_{l=0}^{m-2}\Big(h\int_{t_l}^{t_{l+1}}\big\|(-A)e^{A(t_m-s)}\big\|_{L(H)}^2\big\|(-A)^{-1}\big(I-e^{A(s-t_l)}\big)\big\|_{L(H)}^2\,\mathrm{d}s\Big)^{\frac{1}{2}} \\
&\leq C\sum_{l=0}^{m-2}\Big(h\int_{t_l}^{t_{l+1}}\Big(\frac{s-t_l}{t_m-s}\Big)^2\,\mathrm{d}s\Big)^{\frac{1}{2}} \leq C\sum_{l=0}^{m-2}\Big(h\int_{t_l}^{t_{l+1}}\Big(\frac{s-t_l}{(m-l-1)h}\Big)^2\,\mathrm{d}s\Big)^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
&= C \sum_{l=0}^{m-2} \Big(\frac{h^4}{(m-l-1)^2 h^2} \Big)^{\frac{1}{2}} = Ch \sum_{l=0}^{m-2} \frac{1}{m-l-1} = Ch \sum_{l=1}^{m-1} \frac{1}{l} \leq C \frac{1+\ln(M)}{M} \\
&\leq C \frac{M^{1-\gamma}}{M(1-\gamma)} = Ch^{\gamma}.
\end{aligned}
$$

In the last step, we employed some basic computations for $m \in \{1, \dots, M\}$, $M \in \mathbb{N}$

$$
\sum_{l=1}^{m-1} \frac{1}{l} = 1 + \sum_{l=2}^{m-1} \frac{1}{l} \leq 1 + \sum_{l=2}^{M} \frac{1}{l} \leq 1 + \int_1^M \frac{1}{s} \, \mathrm{d}s = 1 + \ln(M)
$$

and for all $r \in [0,1)$ and $x \geq 1$, we get

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

see [33]. Further, we obtain

$$
\begin{aligned}
&\Big(\mathrm{E}\Big[\Big\| \int_{t_{m-1}}^{t_m} \Big(e^{A(t_m - s)} - e^{A(t_m - t_{m-1})} \Big) F(X_{t_{m-1}}) \, \mathrm{d}s \Big\|_H^2 \Big] \Big)^{\frac{1}{2}} \\
&\leq \sqrt{h} \Big(\int_{t_{m-1}}^{t_m} \mathrm{E}\Big[\big\| \big(e^{A(t_m - s)} - e^{A(t_m - t_{m-1})} \big) F(X_{t_{m-1}}) \big\|_H^2 \Big] \, \mathrm{d}s \Big)^{\frac{1}{2}} \\
&\leq \sqrt{h} \Big(\int_{t_{m-1}}^{t_m} C \, \mathrm{d}s \Big)^{\frac{1}{2}} \leq C_T h
\end{aligned}
$$

for all $m \in \{1, \dots, M\}$, $M \in \mathbb{N}$.

Temporal discretization with Milstein scheme - Diffusion B

For the estimation of the error resulting from the discretization of the stochastic integrals, we compute for all $m \in \{1, \dots, M\}$, $M, K \in \mathbb{N}$

$$
\begin{aligned}
&\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m - s)} B(X_s) - e^{A(t_m - t_l)} B(X_{t_l}) \Big) \, \mathrm{d}W_s^K \\
&\qquad - \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} B'(X_{t_l}) \Big(\int_{t_l}^{s} B(X_{t_l}) \, \mathrm{d}W_r^K \Big) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&= \sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} \left(B(X_s) - B(X_{t_l}) \right) \, \mathrm{d}W_s^K \\
&\qquad\quad - \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} B'(X_{t_l}) \Big(\int_{t_l}^{s} B(X_{t_l}) \, \mathrm{d}W_r^K \Big) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&\quad + \mathrm{E}\Big[\Big\| \sum_{l=0}^{m-2} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m - s)} - e^{A(t_m - t_l)} \Big) B(X_s) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&\quad + \mathrm{E}\Big[\Big\| \int_{t_{m-1}}^{t_m} \Big(e^{A(t_m - s)} - e^{A(t_m - t_{m-1})} \Big) B(X_s) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&\leq C_T \Big(M^{-2\gamma} + \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} \Big),
\end{aligned}
\tag{3.29}
$$

where

$$\sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} \left(B(X_s) - B(X_{t_l})\right) \mathrm{d}W_s^K$$
$$- \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big) \,\mathrm{d}W_s^K \Big\|_H^2\Big]$$
$$= \sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} \Big(B'(X_{t_l})(X_s - X_{t_l})$$
$$+ \frac{1}{2}\int_0^1 \Big(\int_0^r B''(X_{t_l} + u(X_s - X_{t_l}))\big(X_s - X_{t_l}, X_s - X_{t_l}\big)\, r \,\mathrm{d}u\Big) \,\mathrm{d}r\Big) \,\mathrm{d}W_s^K$$
$$- \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big) \,\mathrm{d}W_s^K \Big\|_H^2\Big]$$
$$\leq \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| e^{A(t_m-t_l)} B'(X_{t_l})\Big((X_s - X_{t_l}) - \int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big)$$
$$+ e^{A(t_m-t_l)} \frac{1}{2}\int_0^1 \Big(\int_0^r B''(X_{t_l} + u(X_s - X_{t_l}))\big(X_s - X_{t_l}, X_s - X_{t_l}\big) \,\mathrm{d}u\Big)\, r \,\mathrm{d}r \Big\|_{L_{HS}(V_0,H)}^2\Big] \,\mathrm{d}s$$

due to Itô's isometry.

With Theorem 2.3 and Theorem 2.7, we obtain for all $m \in \{1, \ldots, M\}$, $M, K \in \mathbb{N}$

$$\sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} \left(B(X_s) - B(X_{t_l})\right) \mathrm{d}W_s^K$$
$$- \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big) \,\mathrm{d}W_s^K \Big\|_H^2\Big]$$
$$\leq 2\sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| e^{A(t_m-t_l)} B'(X_{t_l})\Big((X_s - X_{t_l}) - \Big(\int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big)\Big) \Big\|_{L_{HS}(V_0,H)}^2\Big] \,\mathrm{d}s$$
$$+ \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\big\| e^{A(t_m-t_l)} \big\|_{L(H)}^2 \, \|X_s - X_{t_l}\|_H^4$$
$$\cdot \int_0^1 \Big(\int_0^r \big\|B''(X_{t_l} + u(X_s - X_{t_l}))\big\|_{L^{(2)}(H, L_{HS}(V_0,H))}^2 \,\mathrm{d}u\Big)\, r \,\mathrm{d}r\Big] \,\mathrm{d}s$$
$$\leq C \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| e^{A(t_m-t_l)} B'(X_{t_l})\Big((X_s - X_{t_l}) - \Big(\int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big)\Big) \Big\|_{L_{HS}(V_0,H)}^2\Big] \,\mathrm{d}s$$
$$+ \frac{h^{1+\min(4\gamma,2)}}{1+\min(4\gamma,2)}\Big).$$

Next, we plug in the expression for the mild solution, use (C3), and get for all $m \in \{1, \ldots, M\}$, $M, K \in \mathbb{N}$

$$\sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} \left(B(X_s) - B(X_{t_l})\right) \mathrm{d}W_s^K$$
$$- \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l}) \,\mathrm{d}W_r^K\Big) \,\mathrm{d}W_s^K \Big\|_H^2\Big]$$

$$
\begin{aligned}
&\leq C \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| e^{A(t_m - t_l)} B'(X_{t_l}) \Big(\big(e^{A(s-t_l)} - I\big) X_{t_l} + \int_{t_l}^{s} e^{A(s-u)} F(X_u)\,\mathrm{d}u \\
&\qquad\qquad + \int_{t_l}^{s} e^{A(s-u)} B(X_u)\,\mathrm{d}(W_u - W_u^K) + \int_{t_l}^{s} \big(e^{A(s-u)} - I\big) B(X_u)\,\mathrm{d}W_u^K \\
&\qquad\qquad + \int_{t_l}^{s} \left(B(X_u) - B(X_{t_l})\right)\,\mathrm{d}W_u^K \Big) \Big\|_{L_{HS}(V_0,H)}^2 \Big]\,\mathrm{d}s + h^{1+\min(4\gamma,2)} \Big) \\
&\leq C \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\big\| \big(e^{A(s-t_l)} - I\big) X_{t_l} \big\|_H^2 \Big]\,\mathrm{d}s + \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| \int_{t_l}^{s} e^{A(s-u)} F(X_u)\,\mathrm{d}u \Big\|_H^2 \Big]\,\mathrm{d}s \\
&\qquad + \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| \int_{t_l}^{s} e^{A(s-u)} B(X_u)\,\mathrm{d}(W_u - W_u^K) \Big\|_H^2 \Big]\,\mathrm{d}s \\
&\qquad + \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| \int_{t_l}^{s} \big(e^{A(s-u)} - I\big) B(X_u)\,\mathrm{d}W_u^K \Big\|_H^2 \Big]\,\mathrm{d}s \\
&\qquad + \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| \int_{t_l}^{s} \left(B(X_u) - B(X_{t_l})\right)\,\mathrm{d}W_u^K \Big\|_H^2 \Big]\,\mathrm{d}s + h^{1+\min(4\gamma,2)} \Big).
\end{aligned}
$$

The proof of

$$
\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\| \int_{t_l}^{s} e^{A(s-u)} B(X_u)\,\mathrm{d}(W_u - W_u^K) \Big\|_H^2 \Big]\,\mathrm{d}s \leq C_T h \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha},
$$

for all $l \in \{0, \ldots, M-1\}$, $M, K \in \mathbb{N}$, can be found in the next part on page 55.

With Theorem 2.3, (C1)–(C4), by Hölder's inequality, and Itô's isometry, we obtain for all $m \in \{1, \ldots, M\}$, $M, K \in \mathbb{N}$

$$
\begin{aligned}
&\sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m - s)} \left(B(X_s) - B(X_{t_l})\right)\,\mathrm{d}W_s^K \\
&\qquad - \int_{t_l}^{t_{l+1}} e^{A(t_m - s)} B'(X_{t_l}) \Big(\int_{t_l}^{s} B(X_{t_l})\,\mathrm{d}W_r^K \Big)\,\mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&\leq \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} \big\| (-A)^{-\gamma} \big(e^{A(s-t_l)} - I\big) \big\|_{L(H)}^2 \mathrm{E}\big[\|(-A)^{\gamma} X_{t_l}\|_H^2 \big]\,\mathrm{d}s \\
&\qquad + \int_{t_l}^{t_{l+1}} (s - t_l) \Big(\int_{t_l}^{s} \mathrm{E}\big[\|e^{A(s-u)} F(X_u)\|_H^2 \big]\,\mathrm{d}u \Big)\,\mathrm{d}s + C_T h \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} \\
&\qquad + \int_{t_l}^{t_{l+1}} \Big(\int_{t_l}^{s} \big\| (-A)^{-\delta} \big(e^{A(s-u)} - I\big) \big\|_{L(H)}^2 \mathrm{E}\big[\|(-A)^{\delta} B(X_u)\|_{L_{HS}(V_0,H)}^2 \big]\,\mathrm{d}u \Big)\,\mathrm{d}s \\
&\qquad + \int_{t_l}^{t_{l+1}} \Big(\int_{t_l}^{s} \mathrm{E}\big[\| \left(B(X_u) - B(X_{t_l})\right) \|_{L_{HS}(V_0,H)}^2 \big]\,\mathrm{d}u \Big)\,\mathrm{d}s + h^{1+\min(4\gamma,2)} \Big) \\
&\leq \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} (s - t_l)^{2\gamma}\, \mathrm{E}\big[\|(-A)^{\gamma} X_{t_l}\|_H^2 \big]\,\mathrm{d}s \\
&\qquad + \int_{t_l}^{t_{l+1}} (s - t_l) \Big(\int_{t_l}^{s} C\, \mathrm{E}\big[\|F(X_u)\|_H^2 \big]\,\mathrm{d}u \Big)\,\mathrm{d}s + C_T h \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha}
\end{aligned}
$$

$$
\begin{aligned}
&+ \int_{t_l}^{t_{l+1}} \Big(\int_{t_l}^{s} (s-u)^{2\delta} \, \mathrm{E}\big[\|B(X_u)\|^2_{L_{HS}(V_0,H_\delta)}\big] \, \mathrm{d}u \Big) \, \mathrm{d}s \\
&+ \int_{t_l}^{t_{l+1}} \Big(\int_{t_l}^{s} \mathrm{E}\big[\| \left(B(X_u) - B(X_{t_l})\right) \|^2_{L_{HS}(V_0,H)}\big] \, \mathrm{d}u \Big) \, \mathrm{d}s + h^{1+\min(4\gamma,2)} \Big)
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{l=0}^{m-1} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} e^{A(t_m-s)} \left(B(X_s) - B(X_{t_l})\right) \mathrm{d}W_s^K \\
&\qquad - \int_{t_l}^{t_{l+1}} e^{A(t_m-s)} B'(X_{t_l}) \Big(\int_{t_l}^{s} B(X_{t_l}) \, \mathrm{d}W_r^K \Big) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&\leq C \sum_{l=0}^{m-1} \Big(h^{2\gamma+1} + h^3 + C_T \, h \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} + h^{2\delta+2} + \int_{t_l}^{t_{l+1}} \Big(\int_{t_l}^{s} (u-t_l)^{\min(2\gamma,1)} \, \mathrm{d}u \Big) \, \mathrm{d}s \\
&\quad + h^{1+\min(4\gamma,2)} \Big) \\
&\leq C \sum_{l=0}^{m-1} \Big(h^{2\gamma+1} + h^3 + h^{2\delta+2} + C_T h \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} + h^{\min(2\gamma,1)+2} + h^{1+\min(4\gamma,2)} \Big) \\
&\leq C_T \Big(\Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j \Big)^{2\alpha} + h^{2\gamma} \Big),
\end{aligned}
$$

where we used $2 + \min(2\gamma, 1) \geq 1 + \min(4\gamma, 2)$.

The second term in (3.29) is estimated for all $m \in \{1, \ldots, M\}$, $M, K \in \mathbb{N}$, using the independence of the increments of the Q-Wiener process in time, the Itô isometry, Theorem 2.7, and (C1)–(C4)

$$
\begin{aligned}
&\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-2} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m-s)} - e^{A(t_m-t_l)} \Big) B(X_s) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&= \sum_{l=0}^{m-2} \mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m-s)} - e^{A(t_m-t_l)} \Big) B(X_s) \, \mathrm{d}W_s^K \Big\|_H^2 \Big] \\
&\leq \sum_{l=0}^{m-2} \int_{t_l}^{t_{l+1}} \big\|(-A)^{-\delta} \big(e^{A(t_m-s)} - e^{A(m-t_l)} \big) \big\|^2_{L(H)} \mathrm{E}\big[\|(-A)^{\delta} B(X_s)\|^2_{L_{HS}(V_0,H)}\big] \, \mathrm{d}s \\
&\leq C \sum_{l=0}^{m-2} \int_{t_l}^{t_{l+1}} \big\|(-A)^{1-\delta} e^{A(t_m-s)} \big\|^2_{L(H)} \big\|(-A)^{-1} \big(I - e^{A(s-t_l)} \big) \big\|^2_{L(H)} \, \mathrm{d}s \\
&\leq C h^2 \sum_{l=0}^{m-2} \int_{t_l}^{t_{l+1}} (t_m - s)^{2(\delta-1)} \, \mathrm{d}s = C h^2 \sum_{l=0}^{m-2} \Big((t_m - t_{l+1})^{2\delta-1} - (t_m - t_l)^{2\delta-1} \Big) \\
&= C h^2 \big((t_m - t_{m-1})^{2\delta-1} - (t_m)^{2\delta-1} \big) \leq C_T h^{2\delta+1} \leq C_T h^{2\gamma}.
\end{aligned}
$$

Finally, we obtain by conditions (C1), (C3), Theorem 2.3, and Theorem 2.7 for all $m \in \{1, \ldots, M\}$,

$M, K \in \mathbb{N}$

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_{t_{m-1}}^{t_m} \left(e^{A(t_m-s)} - e^{A(t_m-t_{m-1})}\right) B(X_s)\,\mathrm{d}W_s^K \Big\|_H^2\Big] \\
&\leq C \int_{t_{m-1}}^{t_m} \|e^{A(t_m-s)}\|_{L(H)}^2 \big\|(-A)^{-\delta}\big(I - e^{A(s-t_{m-1})}\big)\big\|_{L(H)}^2 \mathrm{E}\big[\|(-A)^{\delta} B(X_s)\|_{L_{HS}(V_0,H)}^2\big]\,\mathrm{d}s \\
&\leq C h^{2\delta+1} \leq C h^{2\gamma}.
\end{aligned}$$

Approximation of the Q-Wiener process

Next, we prove the error estimate resulting from the approximation of the Q-Wiener process and employ

$$\mathrm{d}(W_s - W_s^K) = \sum_{j \in \mathcal{J}\setminus\mathcal{J}_K} \sqrt{\eta_j}\tilde{e}_j\,\mathrm{d}\beta_s^j$$

for all $s \in [0,T]$, $K \in \mathbb{N}$.

For all $l \in \{0, \ldots, M-1\}$, $M, K \in \mathbb{N}$, $s \in [0,T]$, it holds

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_{t_l}^{s} e^{A(s-u)} B(X_u)\,\mathrm{d}(W_u - W_u^K) \Big\|_H^2\Big]^{\frac{1}{2}} \\
&= \mathrm{E}\Big[\Big\| \sum_{j \in \mathcal{J}\setminus\mathcal{J}_K} \int_{t_l}^{s} e^{A(s-u)} B(X_u)\sqrt{\eta_j}\,\mathrm{d}\beta_u^j \tilde{e}_j \Big\|_H^2\Big]^{\frac{1}{2}} \\
&= \Big(\sum_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j \int_{t_l}^{s} \mathrm{E}\Big[\big\|e^{A(s-u)} B(X_u) Q^{-\alpha} Q^{\alpha} \tilde{e}_j\big\|_H^2\Big]\,\mathrm{d}u \Big)^{\frac{1}{2}} \\
&= \Big(\sum_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j^{2\alpha+1} \int_{t_l}^{s} \mathrm{E}\Big[\big\|e^{A(s-u)} B(X_u) Q^{-\alpha} \tilde{e}_j\big\|_H^2\Big]\,\mathrm{d}u \Big)^{\frac{1}{2}} \\
&\leq \Big(\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j \Big)^{2\alpha} \int_{t_l}^{s} \mathrm{E}\Big[\sum_{j \in \mathcal{J}} \eta_j \big\|e^{A(s-u)} B(X_u) Q^{-\alpha} \tilde{e}_j\big\|_H^2\Big]\,\mathrm{d}u \Big)^{\frac{1}{2}} \\
&= \Big(\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j \Big)^{2\alpha} \int_{t_l}^{s} \mathrm{E}\Big[\big\|e^{A(s-u)} B(X_u) Q^{-\alpha}\big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}u \Big)^{\frac{1}{2}}.
\end{aligned}$$

By assumptions (C1), (C3), and Theorem 2.3, we get

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_{t_l}^{s} e^{A(s-u)} B(X_u)\,\mathrm{d}(W_u - W_u^K) \Big\|_H^2\Big]^{\frac{1}{2}} \\
&\leq \Big(\Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j \Big)^{2\alpha} \int_{t_l}^{s} \|(-A)^{\vartheta} e^{A(s-u)}\|_{L(H)}^2 \mathrm{E}\Big[\big\|(-A)^{-\vartheta} B(X_u) Q^{-\alpha}\big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}u \Big)^{\frac{1}{2}} \\
&\leq \Big(C \Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j \Big)^{2\alpha} \int_{t_l}^{s} (s-u)^{-2\vartheta}\,\mathrm{d}u \Big)^{\frac{1}{2}} = \Big(C \Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j \Big)^{2\alpha} \frac{(t_l - s)^{-2\vartheta+1}}{2\vartheta - 1} \Big)^{\frac{1}{2}}
\end{aligned}$$

all $s \in [0,T]$, $l \in \{0, \ldots, M-1\}$, $M, K \in \mathbb{N}$.

Proof of (3.28)
Due to the expression for the iterated integral in (3.11), Hölder's inequality, and Itô's isometry, it holds for all $m \in \{1,\dots,M\}$, $M,N,K \in \mathbb{N}$

$$
\begin{aligned}
&\mathrm{E}\big[\|\bar{X}_{t_m} - \bar{Y}_{t_m}\|_H^2\big] \\
&= \mathrm{E}\Big[\Big\|P_N\Big(\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}\left(F(X_{t_l}) - F(Y_l)\right)\mathrm{d}s \\
&\quad + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}\left(B(X_{t_l}) - B(Y_l)\right)\mathrm{d}W_s^K \\
&\quad + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}\Big(B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l})\,\mathrm{d}W_r^K\Big) - B'(Y_l)\Big(\int_{t_l}^{s} B(Y_l)\,\mathrm{d}W_r^K\Big)\Big)\,\mathrm{d}W_s^K\Big)\Big\|_H^2\Big] \\
&\leq 3\Big(Mh\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} \mathrm{E}\big[\|e^{A(t_m-t_l)}\left(F(X_{t_l}) - F(Y_l)\right)\|_H^2\big]\,\mathrm{d}s \\
&\quad + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} \mathrm{E}\big[\|e^{A(t_m-t_l)}\left(B(X_{t_l}) - B(Y_l)\right)\|_{L_{HS}(V_0,H)}^2\big]\,\mathrm{d}s \\
&\quad + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\|B'(X_{t_l})\Big(\int_{t_l}^{s} B(X_{t_l})\,\mathrm{d}W_r^K\Big) - B'(Y_l)\Big(\int_{t_l}^{s} B(Y_l)\,\mathrm{d}W_r^K\Big)\Big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}s\Big) \\
&\leq C_T h\sum_{l=0}^{m-1}\mathrm{E}\big[\|F(X_{t_l}) - F(Y_l)\|_H^2\big] + Ch\sum_{l=0}^{m-1}\mathrm{E}\big[\|B(X_{t_l}) - B(Y_l)\|_{L_{HS}(V_0,H)}^2\big] \\
&\quad + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\|B'(X_{t_l})\Big(\sum_{j\in\mathcal{J}_K}\int_{t_l}^{s} B(X_{t_l})\tilde{e}_j\sqrt{\eta_j}\,\mathrm{d}\beta_r^j\Big) \\
&\qquad\qquad - B'(Y_l)\Big(\sum_{j\in\mathcal{J}_K}\int_{t_l}^{s} B(Y_l)\tilde{e}_j\sqrt{\eta_j}\,\mathrm{d}\beta_r^j\Big)\Big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}s.
\end{aligned}
$$

By assumptions (C1), (C2), (C3), and the properties of the independent Brownian motions $(\beta_t^j)_{t\in[0,T]}$, $j \in \mathcal{J}_K$, we obtain for all $m \in \{1,\dots,M\}$, $M,K \in \mathbb{N}$

$$
\begin{aligned}
\mathrm{E}\big[\|\bar{X}_{t_m} - \bar{Y}_{t_m}\|_H^2\big] &\leq C_T h\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l} - Y_l\|_H^2\big] + Ch\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l} - Y_l\|_H^2\big] \\
&\quad + C\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\Big\|B'(X_{t_l})\Big(\sum_{j\in\mathcal{J}_K} B(X_{t_l})\tilde{e}_j\sqrt{\eta_j}(\beta_s^j - \beta_{t_l}^j)\Big) \\
&\qquad\qquad - B'(Y_l)\Big(\sum_{j\in\mathcal{J}_K} B(Y_l)\tilde{e}_j\sqrt{\eta_j}(\beta_s^j - \beta_{t_l}^j)\Big)\Big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}s
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathrm{E}\big[\|\bar{X}_{t_m} - \bar{Y}_{t_m}\|_H^2\big] \\
&\leq C_T h\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l} - Y_l\|_H^2\big]
\end{aligned}
$$

$$
\begin{aligned}
&+ C \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\| \sum_{j \in \mathcal{J}_K} \sqrt{\eta_j}\big(B'(X_{t_l})\,(B(X_{t_l})\tilde{e}_j) - B'(Y_l)\,(B(Y_l)\tilde{e}_j)\,\big)(\beta_s^j - \beta_{t_l}^j)\|_{L_{HS(V_0,H)}}^2\Big]\,\mathrm{d}s \\
&\leq C_T h \sum_{l=0}^{m-1} \mathrm{E}\big[\|X_{t_l} - Y_l\|_H^2\big] \\
&+ C \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \sum_{j \in \mathcal{J}} \eta_j \mathrm{E}\big[\|B'(X_{t_l})\,(B(X_{t_l})\tilde{e}_j) - B'(Y_l)\,(B(Y_l)\tilde{e}_j)\,\|_{L_{HS(V_0,H)}}^2\big]\mathrm{E}\big[(\beta_s^j - \beta_{t_l}^j)^2\big]\,\mathrm{d}s \\
&\leq C_T h \sum_{l=0}^{m-1} \mathrm{E}\big[\|X_{t_l} - Y_l\|_H^2\big] \\
&+ C \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \mathrm{E}\big[\|B'(X_{t_l})\,(B(X_{t_l})) - B'(Y_l)\,(B(Y_l))\,\|_{L_{HS}^{(2)}(V_0,H)}^2\big](s - t_l)\,\mathrm{d}s \\
&\leq C_T h \sum_{l=0}^{m-1} \mathrm{E}\big[\|X_{t_l} - Y_l\|_H^2\big].
\end{aligned}
$$

Approximation of the derivative

It remains to show that the approximation of the derivative does not distort the convergence order. In the proof, we need the following lemma.

Lemma 3.1

Under assumptions (C1)–(C4), it holds for all $p \in [2, \infty)$, $M, N, K \in \mathbb{N}$, and some constant $C_{p,T,Q} > 0$, independent of M, N, K,

$$
\sup_{m \in \{0,\dots,M\}} \big(\mathrm{E}\big[\|Y_m\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} \leq C_{p,T,Q}\big(1 + \big(\mathrm{E}\big[\|\xi\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\big).
$$

Proof of Lemma 3.1.

We prove this estimate iteratively. Therefore, assume that for fixed $m \in \mathbb{N}$ it holds

$$
\big(\mathrm{E}\big[\|Y_l\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} \leq C_{p,T,Q}\big(1 + \big(\mathrm{E}\big[\|\xi\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\big)
$$

for all $l \in \{0, \dots, m-1\}$, $p \in [2, \infty)$.

By the triangle inequality, we get for all $m \in \{1, \dots, M\}$, $M, K \in \mathbb{N}$, and $p \in [1, \infty)$

$$
\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \\
&\leq C\Bigg(\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} + \sum_{l=0}^{m-1}\Bigg(\mathrm{E}\Bigg[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} F(Y_l)\,\mathrm{d}s\Big\|_{H_\delta}^p\Bigg]\Bigg)^{\frac{1}{p}} \\
&\quad + \Bigg(\mathrm{E}\Bigg[\Big\|\int_{t_0}^{t_m} \sum_{l=0}^{m-1} e^{A(t_m - t_l)} B(Y_l)\mathbb{1}_{[t_l,t_{l+1})}(s)\,\mathrm{d}W_s^K\Big\|_{H_\delta}^p\Bigg]\Bigg)^{\frac{1}{p}} \\
&\quad + \sum_{l=0}^{m-1}\Bigg(\mathrm{E}\Bigg[\Big\|e^{A(t_m - t_l)}\frac{1}{\sqrt{h}}\Big(B\Big(Y_l + \frac{\sqrt{h}}{2}P_N B(Y_l)\Delta W_l^K\Big) - B(Y_l)\Big)\Delta W_l^K\Big\|_{H_\delta}^p\Bigg]\Bigg)^{\frac{1}{p}}
\end{aligned}
$$

$$+\sum_{l=0}^{m-1}\Big(\mathrm{E}\Big[\Big\|e^{A(t_m-t_l)}\sum_{j\in\mathcal{J}_K}\bar{B}(Y_l,h,j)\Big\|_{H_\delta}^p\Big]\Big)^{\frac{1}{p}}\Big)^2.$$

By Hölder's inequality and Theorem 2.6, we obtain for all $p\in[2,\infty)$ and $m\in\{1,\dots,M\}$, $M,K\in\mathbb{N}$

$$\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\\
&\leq C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}+C\Big(\sum_{l=0}^{m-1}\Big(\int_{t_l}^{t_{l+1}}\mathrm{E}\Big[\big\|(-A)^\delta e^{A(t_m-t_l)}F(Y_l)\big\|_H^p\Big]\,\mathrm{d}s\Big)^{\frac{1}{p}}h^{1-\frac{1}{p}}\Big)^2\\
&\quad+C\int_{t_0}^{t_m}\Big(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}e^{A(t_m-t_l)}B(Y_l)\mathbb{1}_{[t_l,t_{l+1})}(s)\Big\|_{L_{HS}(V_0,H_\delta)}^p\Big]\Big)^{\frac{2}{p}}\,\mathrm{d}s\\
&\quad+C\Big(\sum_{l=0}^{m-1}(t_m-t_l)^{-\delta}\Big(\mathrm{E}\Big[\Big\|\frac{1}{\sqrt{h}}\Big(B\Big(Y_l+\frac{\sqrt{h}}{2}P_NB(Y_l)\Delta W_l^K\Big)-B(Y_l)\Big)\Delta W_l^K\Big\|_H^p\Big]\Big)^{\frac{1}{p}}\Big)^2\\
&\quad+C\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\big\|(-A)^\delta e^{A(t_m-t_l)}\bar{B}(Y_l,h,j)\big\|_H^p\Big]\Big)^{\frac{1}{p}}\Big)^2.
\end{aligned}$$

Now, we concentrate on

$$\bar{B}(Y_l,h,j)=\Big(B\big(Y_l-\frac{h}{2}P_NB(Y_l)\sqrt{\eta_j}\tilde{e}_j\big)-B(Y_l)\Big)\sqrt{\eta_j}\tilde{e}_j$$

and use the following Taylor expansions for all $l\in\{0,\dots,M\}$, $j\in\mathcal{J}_K$, $M,N,K\in\mathbb{N}$

$$\begin{aligned}
&B\Big(Y_l+\frac{\sqrt{h}}{2}P_NB(Y_l)\Delta W_l^K\Big)\Delta W_l^K\\
&\quad=B(Y_l)\Delta W_l^K+B'(\xi(Y_l,\Delta W_l^K))\Big(\frac{\sqrt{h}}{2}P_NB(Y_l)\Delta W_l^K,\Delta W_l^K\Big)\\
&B\Big(Y_l-\frac{h}{2}P_NB(Y_l)\sqrt{\eta_j}\tilde{e}_j\Big)\sqrt{\eta_j}\tilde{e}_j\\
&\quad=B(Y_l)\sqrt{\eta_j}\tilde{e}_j+B'(\xi(Y_l,j))\Big(-\frac{h}{2}P_NB(Y_l)\sqrt{\eta_j}\tilde{e}_j,\sqrt{\eta_j}\tilde{e}_j\Big).
\end{aligned}\tag{3.30}$$

Precisely, we have

$$\xi(Y_l,\Delta W_l^K)=Y_l+\theta\frac{\sqrt{h}}{2}P_NB(Y_l)\Delta W_l^K$$

and

$$\xi(Y_l,j)=Y_l-\theta\frac{h}{2}P_NB(Y_l)\sqrt{\eta_j}\tilde{e}_j$$

for some $\theta\in(0,1)$ and all $l\in\{0,\dots,M\}$, $j\in\mathcal{J}_K$, $M,K\in\mathbb{N}$.

As $\xi(Y_l,\Delta W_l^K)\in H_N$, $\xi(Y_l,j)\in H_N$ for all $l\in\{0,\dots,M\}$, $j\in\mathcal{J}_K$, $M,K,N\in\mathbb{N}$, it holds $\xi(Y_l,\Delta W_l^K),\xi(Y_l,j)\in H_\beta$ for arbitrary $l\in\{0,\dots,M\}$, $j\in\mathcal{J}_K$, $M,K,N\in\mathbb{N}$.

Due to (C1)–(C3), Theorem 2.3, and Theorem 2.6, we get

$$
\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \\
&\leq C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + Ch^{2-\frac{2}{p}} M \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} (t_m - t_l)^{-\delta p} \,\mathrm{d}s\Big)^{\frac{2}{p}} \big(\mathrm{E}\big[\|F(Y_l)\|_H^p\big]\big)^{\frac{2}{p}} \\
&\quad + C_p \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \bigg(\mathrm{E}\bigg[\Big\| \sum_{l=0}^{m-1} e^{A(t_m-t_l)} B(Y_l) \mathbb{1}_{[t_l,t_{l+1})}(s)\Big\|_{L_{HS}(V_0,H_\delta)}^p\bigg]\bigg)^{\frac{2}{p}} \mathrm{d}s \\
&\quad + C\frac{M}{h} \sum_{l=0}^{m-1} (t_m - t_l)^{-2\delta} \bigg(\mathrm{E}\bigg[\Big\| B'(\xi(Y_l, \Delta W_l^K)) \frac{\sqrt{h}}{2} P_N B(Y_l) \Delta W_l^K \Big\|_{L(V,H)}^p \|\Delta W_l^K\|_V^p\bigg]\bigg)^{\frac{2}{p}} \\
&\quad + CM \sum_{l=0}^{m-1} \bigg(\sum_{j\in\mathcal{J}_K} (t_m - t_l)^{-\delta} \bigg(\mathrm{E}\bigg[\Big\| - B'(\xi(Y_l, j)) \frac{h}{2} P_N B(Y_l) \sqrt{\eta_j}\tilde{e}_j \Big\|_{L(V,H)}^p \|\sqrt{\eta_j}\tilde{e}_j\|_V^p\bigg]\bigg)^{\frac{1}{p}} \bigg)^2 \\
&\leq C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{p,T} h^{1-\frac{2}{p}} \sum_{l=0}^{m-1} \big(h(t_m - t_l)^{-\delta p}\big)^{\frac{2}{p}} \big(1 + \big(\mathrm{E}\big[\|Y_l\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C \sum_{l=0}^{m-1} \big(\mathrm{E}\big[\|B(Y_l)\|_{L_{HS}(V_0,H_\delta)}^p\big]\big)^{\frac{2}{p}} \int_{t_l}^{t_{l+1}} \|(-A)^{-\delta}\|_{L(H)}^2 \|(-A)^{\delta} e^{A(t_m-t_l)}\|_{L(H)}^2 \,\mathrm{d}s \\
&\quad + CM \sum_{l=0}^{m-1} (t_m - t_l)^{-2\delta} \Big(\mathrm{E}\big[\|B'(\xi(Y_l, \Delta W_l^K))\|_{L(H,L(V,H))}^p \|B(Y_l)\Delta W_l^K\|_{H_\delta}^p \|\Delta W_l^K\|_V^p\big]\Big)^{\frac{2}{p}} \\
&\quad + CM \sum_{l=0}^{m-1} (t_m - t_l)^{-2\delta} \Big(\sum_{j\in\mathcal{J}_K} \sqrt{\eta_j} h \big(\mathrm{E}\big[\|B'(\xi(Y_l, j))\|_{L(H,L(V,H))}^p \|B(Y_l)\sqrt{\eta_j}\tilde{e}_j\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} \Big)^2 \\
&\leq C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + h^{1-2\delta} C_{p,T} \sum_{l=0}^{m-1} (m-l)^{-2\delta} \big(1 + \big(\mathrm{E}\big[\|Y_l\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p \sum_{l=0}^{m-1} h\,(t_m - t_l)^{-2\delta} \,\big(1 + \big(\mathrm{E}\big[\|Y_l\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p M h^{-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta} \big(\mathrm{E}\big[(1 + \|Y_l\|_{H_\delta}^p) \|\Delta W_l^K\|_V^{2p}\big]\big)^{\frac{2}{p}} \\
&\quad + C_{p,T} h^{1-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta} \bigg(\sum_{j\in\mathcal{J}_K} \sqrt{\eta_j} \Big(\mathrm{E}\Big[(1 + \|Y_l\|_{H_\delta}^p) \|\sqrt{\eta_j}\tilde{e}_j\|_V^p\Big]\Big)^{\frac{1}{p}} \bigg)^2.
\end{aligned}
$$

for all $m \in \{1, \ldots, M\}$, $M, K, N \in \mathbb{N}$, $p \in [2, \infty)$.

Similar as in [31], we obtain for $\delta \in (0, \frac{1}{2})$ and all $m \in \{1, \ldots, M\}$, $M \in \mathbb{N}$

$$
\sum_{l=0}^{m-1} (m-l)^{-2\delta} = \sum_{l=1}^{m} \frac{1}{l^{2\delta}} \leq 1 + \int_1^M \frac{1}{r^{2\delta}} \,\mathrm{d}r = 1 + \frac{M^{1-2\delta} - 1}{1 - 2\delta} \leq \frac{M^{1-2\delta}}{1 - 2\delta}. \tag{3.31}
$$

Therefore, we get for all $m \in \{1, \ldots, M\}$, $M \in \mathbb{N}$, $p \in [2, \infty)$

$$
\big(\mathrm{E}\big[\|Y_m\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \leq C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + h^{1-2\delta} C_{T,p,Q} \sum_{l=0}^{m-1} (m-l)^{-2\delta} \big(1 + \big(\mathrm{E}\big[\|Y_l\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big)
$$

$$\leq C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{T,p,Q} + h^{1-2\delta} C_{T,p,Q} \sum_{l=0}^{m-1} (m-l)^{-2\delta} \big(\mathrm{E}\big[\|Y_l\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}.$$

By the discrete Gronwall Lemma, we finally obtain for all $m \in \{1, \ldots, M\}$, $M \in \mathbb{N}$, $p \in [2, \infty)$

$$\begin{aligned}\big(\mathrm{E}\big[\|Y_m\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \leq& \Big(C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{T,p,Q}\Big)\, e^{C_{T,p,Q} h^{1-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}} \\ \leq& C_{T,p,Q} \left(1 + \big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\right).\end{aligned}$$

For the cDFMM scheme with

$$\bar{B}(Y_l, h, j) = \Big(b\Big(\cdot, Y_l - \frac{h}{2} P_N b(\cdot, Y_l)\Big) - b(\cdot, Y_l)\Big) \eta_j \tilde{e}_j^2$$

for all $l \in \{0, \ldots, M\}$, $j \in \mathcal{J}_K$, $M, N, K \in \mathbb{N}$, the result follows analogously. □

With this at hand, we can prove the last estimate in (3.27) for all $m \in \{1, \ldots, M\}$, $M, N, K \in \mathbb{N}$

$$\begin{aligned}&\mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big] \\ &= \mathrm{E}\Big[\Big\| P_N\Big(e^{At_m} X_0 + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} F(Y_l)\, \mathrm{d}s + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} B(Y_l)\, \mathrm{d}W_s^K \\ &\quad + \sum_{l=0}^{m-1} \Big(\frac{1}{2} e^{A(t_m - t_l)} B'(Y_l) \left(B(Y_l) \Delta W_l^K, \Delta W_l^K\right) - \frac{h}{2} e^{A(t_m - t_l)} \sum_{j \in \mathcal{J}_K} \eta_j B'(Y_l) \left(B(Y_l) \tilde{e}_j, \tilde{e}_j\right)\Big)\Big) \\ &\quad - P_N\Big(e^{At_m} X_0 + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} F(Y_l)\, \mathrm{d}s + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m - t_l)} B(Y_l)\, \mathrm{d}W_s^K \\ &\quad + \sum_{l=0}^{m-1} e^{A(t_m - t_l)} \frac{1}{\sqrt{h}} \Big(B\Big(Y_l + \frac{1}{2}\sqrt{h} P_N B(Y_l) \Delta W_l^K\Big) - B(Y_l)\Big) \Delta W_l^K \\ &\quad + \sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} e^{A(t_m - t_l)} \bar{B}(Y_l, h, j)\Big)\Big\|_H^2\Big].\end{aligned}$$

We are left with

$$\begin{aligned}&\mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big] \\ &= \mathrm{E}\Big[\Big\| P_N\Big(\sum_{l=0}^{m-1} e^{A(t_m - t_l)} \Big(\frac{1}{2} B'(Y_l) \left(B(Y_l) \Delta W_l^K, \Delta W_l^K\right) - \frac{h}{2} \sum_{j \in \mathcal{J}_K} \eta_j B'(Y_l) \left(B(Y_l) \tilde{e}_j, \tilde{e}_j\right)\Big)\Big) \\ &\quad - P_N\Big(\sum_{l=0}^{m-1} e^{A(t_m - t_l)} \frac{1}{\sqrt{h}} \Big(B\Big(Y_l + \frac{1}{2}\sqrt{h} P_N B(Y_l) \Delta W_l^K\Big) - B(Y_l)\Big) \Delta W_l^K\Big) \\ &\quad - P_N\Big(\sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} e^{A(t_m - t_l)} \bar{B}(Y_l, h, j)\Big)\Big\|_H^2\Big]\end{aligned}$$

for all $m \in \{1, \ldots, M\}$, $M, N, K \in \mathbb{N}$. Again, we consider

$$\bar{B}(Y_l, h, j) = \Big(B\big(Y_l - \frac{h}{2} P_N B(Y_l) \sqrt{\eta_j} \tilde{e}_j\big) - B(Y_l)\Big) \sqrt{\eta_j} \tilde{e}_j$$

for all $l \in \{0, \ldots, M\}$, $j \in \mathcal{J}_K$, $M, N, K \in \mathbb{N}$ first and use Taylor approximations similar to (3.30).

For legibility, we define $S_l := \frac{\sqrt{h}}{2} P_N B(Y_l) \Delta W_l^K$, $\tilde{S}_{lj} := \frac{h}{2} P_N B(Y_l) \sqrt{\eta_j} \tilde{e}_j$ for all $l \in \{0, \ldots, M\}$, $j \in \mathcal{J}_K$, $M, N, K \in \mathbb{N}$, and insert these expressions. This implies

$$\begin{aligned}
&\mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big] \\
&\leq \mathrm{E}\Big[\Big\| P_N \Big(\sum_{l=0}^{m-1} e^{A(t_m - t_l)} \Big(\frac{1}{2} B'(Y_l)\big(B(Y_l)\Delta W_l^K, \Delta W_l^K\big) - \frac{h}{2} \sum_{j \in \mathcal{J}_K} \eta_j B'(Y_l)\big(B(Y_l)\tilde{e}_j, \tilde{e}_j\big)\Big)\Big) \\
&\quad - P_N \Big(\sum_{l=0}^{m-1} e^{A(t_m - t_l)} \Big(\frac{1}{\sqrt{h}} B'(Y_l) \Big(\frac{\sqrt{h}}{2} P_N B(Y_l) \Delta W_l^K, \Delta W_l^K \Big) \\
&\qquad + \frac{1}{2\sqrt{h}} B''(\xi(Y_l, \Delta W_l^K)) \Big(\frac{\sqrt{h}}{2} P_N B(Y_l) \Delta W_l^K, \frac{\sqrt{h}}{2} P_N B(Y_l) \Delta W_l^K \Big) \Delta W_l^K \Big)\Big) \\
&\quad - P_N \Big(\sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} e^{A(t_m - t_l)} \Big(B'(Y_l) \Big(-\frac{h}{2} P_N B(Y_l) \sqrt{\eta_j} \tilde{e}_j, \sqrt{\eta_j} \tilde{e}_j \Big) \\
&\qquad + \frac{1}{2} B''(\xi(Y_l, j)) \Big(-\frac{h}{2} P_N B(Y_l) \sqrt{\eta_j} \tilde{e}_j, -\frac{h}{2} P_N B(Y_l) \sqrt{\eta_j} \tilde{e}_j \Big) \sqrt{\eta_j} \tilde{e}_j \Big)\Big) \Big\|_H^2 \Big] \\
&\leq C \Bigg(\Bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} e^{A(t_m - t_l)} \frac{1}{2\sqrt{h}} B''(\xi(Y_l, \Delta W_l^K)) \Big(S_l, S_l \Big) \Delta W_l^K \Big\|_H^2 \Big] \Bigg)^{\frac{1}{2}} \Bigg)^2 \\
&\quad + C \Bigg(\Bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} e^{A(t_m - t_l)} \frac{1}{2} B''(\xi(Y_l, j)) \Big(\tilde{S}_{lj}, \tilde{S}_{lj} \Big) \sqrt{\eta_j} \tilde{e}_j \Big\|_H^2 \Big] \Bigg)^{\frac{1}{2}} \Bigg)^2
\end{aligned}$$

for all $m \in \{1, \ldots, M\}$, $M, N, K \in \mathbb{N}$.
Then, we obtain by assumptions (C1) and (C3) for all $m \in \{1, \ldots, M\}$, $M, N, K \in \mathbb{N}$

$$\begin{aligned}
&\mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big] \\
&\leq \Bigg(\sum_{l=0}^{m-1} \frac{C}{\sqrt{h}} \Big(\mathrm{E}\Big[\| e^{A(t_m - t_l)} B''(\xi(Y_l, \Delta W_l^K)) \big(S_l, S_l\big) \Delta W_l^K \|_H^2 \Big] \Big)^{\frac{1}{2}} \Bigg)^2 \\
&\quad + C \Bigg(\sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} \Big(\mathrm{E}\Big[\| e^{A(t_m - t_l)} \frac{1}{2} B''(\xi(Y_l, j)) \Big(\tilde{S}_{lj}, \tilde{S}_{lj} \Big) \sqrt{\eta_j} \tilde{e}_j \|_H^2 \Big] \Big)^{\frac{1}{2}} \Bigg)^2 \\
&\leq \Bigg(C \sum_{l=0}^{m-1} \frac{1}{\sqrt{h}} \Bigg(\mathrm{E}\Big[\|B''(\xi(Y_l, \Delta W_l^K)\|_{L^{(2)}(H, L(V,H))}^2 \Big\| \frac{\sqrt{h}}{2} P_N B(Y_l) \Delta W_l^K \Big\|_H^4 \|\Delta W_l^K\|_V^2 \Big] \Bigg)^{\frac{1}{2}} \Bigg)^2 \\
&\quad + \Bigg(C \sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} \Big(\mathrm{E}\Big[\|B''(\xi(Y_l, j))\|_{L^{(2)}(H, L(V,H))}^2 \Big\| \frac{h}{2} P_N B(Y_l) \sqrt{\eta_j} \tilde{e}_j \Big\|_H^4 \|\sqrt{\eta_j} \tilde{e}\|_V^2 \Big] \Big)^{\frac{1}{2}} \Bigg)^2.
\end{aligned}$$

With (C1)–(C4) again and the fact that Q is trace class, we get for all $m \in \{1, \dots, M\}$, $M, K \in \mathbb{N}$

$$
\begin{aligned}
&\mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big] \\
&\leq \Big(C \sum_{l=0}^{m-1} \frac{\sqrt{h}}{4} \Big(\big(\mathrm{E}\big[\|B(Y_l)\|^4_{L(V,H_\delta)} \|\Delta W_l^K\|_V^6\big]\big)^{\frac{1}{2}}\Big)^2 + \Big(C \sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} \frac{h^2}{4} \eta_j^{\frac{3}{2}} \big(\mathrm{E}\big[\|B(Y_l)\|^4_{L(V,H_\delta)}\big]\big)^{\frac{1}{2}}\Big)^2 \\
&\leq \Big(C \sum_{l=0}^{m-1} \sqrt{h} \big(1 + \mathrm{E}\big[\|Y_l\|^4_{H_\delta}\big]\big)^{\frac{1}{2}} \big(\mathrm{E}\big[\|\Delta W_l^K\|_V^6\big]\big)^{\frac{1}{2}}\Big)^2 + \Big(C \sum_{l=0}^{m-1} \sum_{j \in \mathcal{J}_K} h^2 \eta_j^{\frac{3}{2}} \big(1 + \mathrm{E}\big[\|Y_l\|^4_{H_\delta}\big]\big)^{\frac{1}{2}}\Big)^2 \\
&\leq \Big(C \sum_{l=0}^{m-1} h^2 \big(C\big(1 + \mathrm{E}\big[\|Y_l\|^4_{H_\delta}\big]\big)\big)^{\frac{1}{2}}\Big)^2 + \Big(C \sum_{l=0}^{m-1} \Big(\sup_{j \in \mathcal{J}_K} \sqrt{\eta_j}\Big) \operatorname{tr} Q \, h^2 \big(1 + \mathrm{E}\big[\|Y_l\|^4_{H_\delta}\big]\big)^{\frac{1}{2}}\Big)^2 \\
&\leq \Big(C \sum_{l=0}^{m-1} h^2\Big)^2 + \Big(C \sum_{l=0}^{m-1} \Big(\sup_{j \in \mathcal{J}} \sqrt{\eta_j}\Big) \operatorname{tr} Q \, h^2\Big)^2 \leq C_{T,Q} h^2.
\end{aligned}
$$

This proves the error estimate for the general case.

Now, let $l \in \{0, \dots, M\}$, $j \in \mathcal{J}_K$, $M, N, K \in \mathbb{N}$; for

$$
\bar{B}(Y_l, h, j) = \Big(b\Big(\cdot, Y_l - \frac{h}{2} P_N b(\cdot, Y_l)\Big) - b(\cdot, Y_l)\Big) \eta_j \tilde{e}_j^2,
$$

we use

$$
\begin{aligned}
b\Big(\cdot, Y_l - \frac{h}{2} P_N b(\cdot, Y_l)\Big) \eta_j \tilde{e}_j^2 =& b(\cdot, Y_l) \eta_j \tilde{e}_j^2 + b'(\cdot, Y_l) \cdot \Big(-\frac{h}{2} P_N b(\cdot, Y_l)\Big) \eta_j \tilde{e}_j^2 \\
&+ \frac{1}{2} b''(\cdot, \xi(Y_l, j)) \cdot \Big(-\frac{h}{2} P_N b(\cdot, Y_l)\Big) \cdot \Big(-\frac{h}{2} P_N b(\cdot, Y_l)\Big) \eta_j \tilde{e}_j^2
\end{aligned}
$$

and the estimate

$$
\mathrm{E}\big[\|\bar{Y}_{t_m} - Y_m\|_H^2\big] \leq C_{T,Q} h^2
$$

follows for all $m \in \{0, \dots, M\}$, $M \in \mathbb{N}$, as above. □

Proof of Theorem 3.2

Theorem (Convergence of EES)
Let assumptions (C1)–(C4) be fulfilled. Then, there exists a constant $C_T \in (0, \infty)$, independent of N, K, and M, such that for $(Y_m^{EES})_{0 \leq m \leq M}$, defined by the exponential Euler scheme in (3.13), *it holds*

$$
\Big(\mathrm{E}\Big[\big\|X_{t_m} - Y_m^{EES}\big\|_H^2\Big]\Big)^{\frac{1}{2}} \leq C_T \Big(\Big(\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i\Big)^{-\gamma} + \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j\Big)^{\alpha} + M^{-\min(\frac{1}{2}, \gamma, 2(\gamma - \beta))}\Big)
$$

for all $m \in \{0, 1, \dots, M\}$ and all $N, M, K \in \mathbb{N}$. The parameter values are determined by (C1)–(C4).

Proof of Theorem 3.2.
We use the following notation throughout the proof

$$
\begin{aligned}
X_{t_m} =& e^{At_m} X_0 + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-s)} F(X_s)\,\mathrm{d}s + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-s)} B(X_s)\,\mathrm{d}W_s, \\
\bar{X}_{t_m} =& P_N\Big(e^{At_m} X_0 + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-s)} F(X_{t_l})\,\mathrm{d}s + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} B(X_{t_l})\,\mathrm{d}W_s^K \Big),
\end{aligned}
$$

and

$$
\begin{aligned}
Y_m &= Y_m^{EES} \\
&= P_N\Big(e^{At_m} X_0 + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-s)} F(Y_l^{EES})\,\mathrm{d}s + \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)} B(Y_l^{EES})\,\mathrm{d}W_s^K \Big)
\end{aligned}
$$

for all $m \in \{1, \ldots, M\}$ and $M,\, N,\, K \in \mathbb{N}$.
In the following, we compute the error estimate in several parts according to

$$
\big(\mathrm{E}\big[\|X_{t_m} - Y_m\|_H^2\big]\big)^{\frac{1}{2}} = \Big(\mathrm{E}\big[\|X_{t_m} - P_N X_{t_m} + P_N X_{t_m} - \bar{X}_{t_m} + \bar{X}_{t_m} - Y_m\|_H^2\big]\Big)^{\frac{1}{2}} \tag{3.32}
$$

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

Throughout the proof we use some generic constant C which may change in each step.
The first term accounts for the projection $P_N : H \to H_N$, $N \in \mathbb{N}$, and can be estimated independently of the numerical scheme. We obtain this estimate as in the proof of Theorem 3.1 on page 47, see also [35],

$$
\big(\mathrm{E}\big[\|X_{t_m} - P_N X_{t_m}\|_H^2\big]\big)^{\frac{1}{2}} \leq C\Big(\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i\Big)^{-\gamma} \tag{3.33}
$$

for all $m \in \{0, 1, \ldots, M\}$ and all $N, M \in \mathbb{N}$.
The remaining terms arise due to the approximation in time and the approximation of the Q-Wiener process.

Similar as in the proof of Theorem 3.1 and [35], we show for all $m \in \{1, \ldots, M\}$ and $M, N, K \in \mathbb{N}$,

$$
\begin{aligned}
&\big(\mathrm{E}\big[\|P_N X_{t_m} - \bar{X}_{t_m}\|_H^2\big]\big)^{\frac{1}{2}} \\
&= \Bigg(\mathrm{E}\Bigg[\Big\|P_N\Big(\sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m-s)}(F(X_s) - F(X_{t_l}))\Big)\,\mathrm{d}s \\
&\qquad + P_N\Big(\sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \Big(e^{A(t_m-s)} B(X_s) - e^{A(t_m-t_l)} B(X_{t_l})\Big)\,\mathrm{d}W_s^K\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\quad + P_N\Big(\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}B(X_s)\,\mathrm{d}(W_s-W_s^K)\Big)\Big\|_H^2\Big]\Big)^{\frac{1}{2}}\\
&\leq C\bigg(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}(F(X_s)-F(X_{t_l}))\Big)\,\mathrm{d}s\Big\|_H^2\Big]\bigg)^{\frac{1}{2}}\\
&\quad+\bigg(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}B(X_s)-e^{A(t_m-t_l)}B(X_{t_l})\Big)\,\mathrm{d}W_s^K\Big\|_H^2\Big]\bigg)^{\frac{1}{2}}\\
&\quad+\bigg(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}B(X_s)\,\mathrm{d}(W_s-W_s^K)\Big\|_H^2\Big]\bigg)^{\frac{1}{2}}\\
&\leq C_T\Big(M^{-\min(\frac{1}{2},\gamma,2(\gamma-\beta))}+\Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Big)^{\alpha}\Big)
\end{aligned}\tag{3.34}
$$

and

$$
\mathrm{E}\big[\|\bar{X}_{t_m}-Y_m\|_H^2\big]\leq\frac{C_T}{M}\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l}-Y_l\|_H^2\big]
$$

separately.

The estimate for the first term in (3.34) can be obtained analogously as in the proof of the derivative free Milstein scheme, see page 48 or [35],

$$
\bigg(\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\,(F(X_s)-F(X_{t_l}))\,\mathrm{d}s\Big\|_H^2\Big]\bigg)^{\frac{1}{2}}\leq C_T M^{-\min(2(\gamma-\beta),\gamma)}\tag{3.35}
$$

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

Next, we analyze the second term in (3.34) in several steps. This yields

$$
\begin{aligned}
&\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}B(X_s)-e^{A(t_m-t_l)}B(X_{t_l})\Big)\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&\leq C\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\,(B(X_s)-B(X_{t_l}))\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&\quad+C\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-2}\int_{t_l}^{t_{l+1}}\Big(e^{A(t_m-s)}-e^{A(t_m-t_l)}\Big)B(X_{t_l})\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&\quad+C\mathrm{E}\Big[\Big\|\int_{t_{m-1}}^{t_m}\Big(e^{A(t_m-s)}-e^{A(t_m-t_{m-1})}\Big)B(X_{t_{m-1}})\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&\leq C_T\Big(M^{-\min(1,2\gamma)}+\Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Big)^{2\alpha}\Big)
\end{aligned}
$$

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

For the first part, we obtain by Theorem 2.7, (C1), (C3), and Jensen's inequality

$$
\begin{aligned}
&\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\,(B(X_s)-B(X_{t_l}))\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&=\sum_{l=0}^{m-1}\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\,(B(X_s)-B(X_{t_l}))\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&=\sum_{l=0}^{m-1}\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\Big(\int_0^1 B'(X_{t_l}+r(X_s-X_{t_l}))(X_s-X_{t_l})\,\mathrm{d}r\Big)\,\mathrm{d}W_s^K\Big\|_H^2\Big]\\
&\le\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\mathrm{E}\Big[\Big\|e^{A(t_m-s)}\int_0^1 B'(X_{t_l}+r(X_s-X_{t_l}))(X_s-X_{t_l})\,\mathrm{d}r\Big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}s\\
&\le\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} C\mathrm{E}\Big[\Big(\int_0^1\big\|B'(X_{t_l}+r(X_s-X_{t_l}))\big\|_{L(H,L_{HS}(V_0,H))}\,\|X_s-X_{t_l}\|_H\,\mathrm{d}r\Big)^2\Big]\,\mathrm{d}s\\
&\le C\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\mathrm{E}\big[\|X_s-X_{t_l}\|_H^2\big]\,\mathrm{d}s\le\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}(s-t_l)^{2\min(\gamma,\frac12)}\,\mathrm{d}s\\
&=\sum_{l=0}^{m-1}h^{2\min(\gamma,\frac12)+1}\le C_T h^{\min(2\gamma,1)}
\end{aligned}
$$

for all $m\in\{1,\dots,M\}$, $M,K\in\mathbb{N}$.
The proof of the second and third part as well as the last term in (3.34) are the same as for the cDFM, see the estimates on pages 54 to 55.

Finally, we estimate the last term in (3.32) by Hölder's inequality and Itô's isometry. We get

$$
\begin{aligned}
\mathrm{E}\big[\|\bar{X}_{t_m}-Y_m\|_H^2\big]=&\mathrm{E}\Big[\Big\|P_N\Big(\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}\,(F(X_{t_l})-F(Y_l))\,\mathrm{d}s\\
&+\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}\,(B(X_{t_l})-B(Y_l))\,\mathrm{d}W_s^K\Big)\Big\|_H^2\Big]\\
\le&2Mh\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\mathrm{E}\Big[\big\|e^{A(t_m-s)}\,(F(X_{t_l})-F(Y_l))\,\big\|_H^2\Big]\,\mathrm{d}s\\
&+2\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}}\mathrm{E}\Big[\big\|e^{A(t_m-t_l)}\,(B(X_{t_l})-B(Y_l))\,\big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}s\\
\le&C_T h\sum_{l=0}^{m-1}\mathrm{E}\big[\|F(X_{t_l})-F(Y_l)\|_H^2\big]+Ch\sum_{l=0}^{m-1}\mathrm{E}\big[\|B(X_{t_l})-B(Y_l)\|_{L_{HS}(V_0,H)}^2\big]\\
\le&C_T h\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l}-Y_l\|_H^2\big]
\end{aligned}
\tag{3.36}
$$

for all $m\in\{1,\dots,M\}$, $M,N,K\in\mathbb{N}$.

Combining (3.33), (3.34), and (3.36), we get

$$\begin{aligned}\mathrm{E}\big[\|X_{t_m}-Y_m\|_H^2\big] \leq& C_T\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-2\gamma}+\Big(\sup_{j\in\mathcal{I}\setminus\mathcal{I}_K}\eta_j\Big)^{2\alpha}+M^{-2\min(\frac{1}{2},\gamma,2(\gamma-\beta))}\Big)\\ &+\frac{C_T}{M}\sum_{l=0}^{m-1}\mathrm{E}\big[\|X_{t_l}-Y_l\|_H^2\big]\\ \leq& C_T\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-2\gamma}+\Big(\sup_{j\in\mathcal{I}\setminus\mathcal{I}_K}\eta_j\Big)^{2\alpha}+M^{-2\min(\frac{1}{2},\gamma,2(\gamma-\beta))}\Big)\end{aligned}$$

for all $m\in\{0,\ldots,M\}$, $M,N,K\in\mathbb{N}$ by Gronwall's Lemma. □

3.6 Numerical Analysis

In this section, we give some examples to illustrate and confirm the theoretical results derived in the previous sections. We compare the commutative derivative-free Milstein scheme to the linear implicit and exponential Euler scheme as well as the Milstein scheme in the following and illustrate its superiority in a general setting. We begin with an example that allows for an exact solution and show that all the schemes approximate this process with the expected orders of convergence. Then, we present an example where the operator B is pointwise multiplicative in the Q-Wiener process. Furthermore, we investigate various equations in a more general setting to compare the effective order of convergence of the cDFM with the benchmark schemes.

Explicit Solution

We begin with the equation introduced in Example 3.1. That is, we solve

$$\begin{aligned}\mathrm{d}X_t &= \Delta X_t\,\mathrm{d}t + X_t\,\mathrm{d}\beta_t, \quad t\in(0,1],\\ X_0(x) &= \sqrt{2}\sum_{n\in\mathbb{N}}\frac{1}{n^2}\sin(n\pi x), \quad x\in(0,1),\\ X_t(0) &= X_t(1)=0, \quad t\in(0,1],\end{aligned}$$

where $(\beta_t)_{t\in[0,T]}$ is a scalar Brownian motion and $H=L^2((0,1),\mathbb{R})$, $V=\mathbb{R}$.

We need to identify the parameters introduced in (C1)–(C4) as these determine the order of convergence. Since $A=\Delta$, it holds $e_i(x)=\sqrt{2}\sin(i\pi x)$, $x\in(0,1)$, and $\lambda_i=i^2\pi^2$ for all $i\in\mathbb{N}$. Thus, we can read off ρ_A as $\rho_A=2$. Moreover, we get $\delta\in(0,\frac{1}{2})$ by [34] as $b(x,y(x))=y(x)$ for all $x\in[0,1]$, $y\in H$. We select the maximal value for δ and obtain by Theorem 3.1 that $q_{cDFM}=q_{MIL}=\gamma\in[\frac{1}{2},1)$; from Theorem 3.2, we get $q_{EES}=\frac{1}{2}$ for the exponential Euler scheme. We choose $\gamma=1-\varepsilon$ for some $\varepsilon>0$ in the following. The parameters ρ_Q and α are not relevant in this setting as there is no error due to an approximation of the Q-Wiener process.
Here, the diffusion operator B is pointwise multiplicative in the stochastic process. Therefore, the multiplicative version of the cDFM is applicable and we expect the cDFMM and the Milstein scheme to converge with the same order error(cDFMM) = error(MIL) $=\mathcal{O}\big(\bar{c}^{-\frac{2}{3}+\varepsilon}\big)$, see

equation (3.26). For this special case, the Runge-Kutta type scheme in [71] is applicable and is simulated here as well. The effective order of convergence is expected to be the same as for the other two schemes. Moreover, we choose $N = 2, 2^2, \ldots, 2^6$ and $M = N^2$. For the linear implicit and the exponential Euler scheme, we expect a lower effective order of convergence of $\text{error(LIE)} = \text{error(EES)} = \mathcal{O}\big(\bar{c}^{-\frac{2}{5}+\varepsilon}\big)$ and set $M = N^4$.

We simulate 300 paths to determine the mean-square error. In the following Figure 3.1, the dashed and dotted lines represent the expected effective orders of convergence - these are attained for all the schemes.

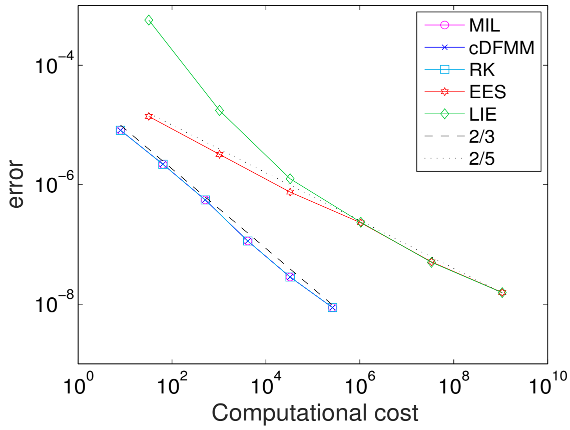

Figure 3.1: Error for $N = 2, 4, 8, 16, 32, 64$ and 300 paths for multiplicative SPDE with exact solution in log-log scale.

For the equations considered below, no explicit solutions exist to the best of our knowledge. Therefore, we compare the numerical schemes with an approximation computed with small step size $h = \frac{T}{M}$, $M \in \mathbb{N}$, and large N, K in the following.

Pointwise Multiplicative Operator

We consider an example that has been analyzed in [35] and [71]. Again, this equation involves a diffusion operator that is pointwise multiplicative. In the following analysis, we show that the cDFMM obtains the same order of convergence as the Milstein scheme and the derivative-free version in [71].
For $T = 1$, we investigate the equation

$$\mathrm{d}X_t = \left(\frac{1}{100}\Delta X_t + 1 - X_t\right)\,\mathrm{d}t + \frac{1 - X_t}{1 + X_t^2}\,\mathrm{d}W_t, \quad t \in (0, T],$$

on $H = V = L^2((0,1), \mathbb{R})$ with $X_0(x) = 0$ and $X_t(0) = X_t(1) = 0$ for all $t \in (0,1]$, $x \in (0,1)$. As A is the Laplacian, we get $\lambda_i = \frac{1}{100}\pi^2 i^2$, $e_i(x) = \sqrt{2}\sin(i\pi x)$ for all $x \in (0,1)$, $i \in \mathbb{N}$. Furthermore, we set $\eta_j = j^{-2}$ and $\tilde{e}_j = e_j$ for all $j \in \mathbb{N}$. We do not analyze the equation at this point and refer the reader to [35] for more details. The analysis in [35] yields $\beta = \frac{1}{5}$, $\alpha \in (0, \frac{3}{4})$,

$\gamma \in (\frac{1}{2}, \frac{3}{4})$, and an expected error of

$$\bigg(\mathrm{E}\bigg[\int_0^1 \big|X_T(x) - Y_M^{N,M,K}(x)\big|^2 \,\mathrm{d}x\bigg]\bigg)^{\frac{1}{2}} \leq C \cdot N^{r-\frac{3}{2}},$$

for $r \in (0,\infty)$. Here, $M = N^3$, $K = N$ are chosen for the linear implicit and exponential Euler and $M = N^2$, $K = N$ for the other schemes according to Section 3.4. As a substitute for the exact solution, we use an approximation computed with an implicit version of the Milstein scheme, introduced in [12], with $N_X = K_X = 2^8$ and $M_X = 2^{21}$.
As this equation is pointwise multiplicative in the Q-Wiener process, the cDFMM is expected to converge with the same order as the Milstein scheme and the Runge-Kutta type scheme by [71]. In this setting, both schemes are very efficient and the cDFMM does not attain an improved effective order of convergence. All higher order schemes are expected to outperform the linear implicit and exponential Euler schemes, however. This is confirmed in Figure 3.2.

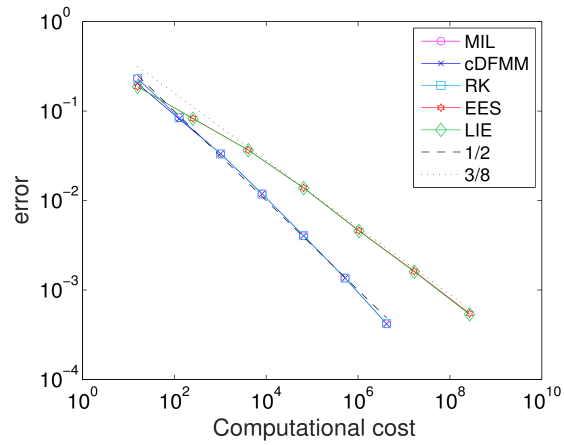

Figure 3.2: Error of Milstein, Runge-Kutta type (RK), multiplicative commutative derivative-free Milstein scheme, linear implicit Euler and exponential Euler scheme for $N = 2, 4, 8, 16, 32, 64, 128$ and 200 paths in log-log scale.

Numerical Analysis of Effective Order of Convergence

Let us now concentrate on equations that are not pointwise multiplicative in the Q-Wiener process, that is, we do not assume $B(y(x))v = b(x, y(x)) \cdot v(x)$, $b : (0,1)^d \times \mathbb{R} \to \mathbb{R}$, $y \in H_\beta$, $v \in V_0$, $x \in (0,1)^d$, $d = 1,2,3$. If, for example, the diffusion operator B is an integral or derivative operator this condition is not fulfilled. For these SPDEs, the commutative derivative-free Milstein scheme is superior in terms of the effective order of convergence. Here, we focus on integral operators B to illustrate this advantage.
First, we introduce the notation and define the operators that we are concerned about in this section. Let $\mu_{ij} : H_\beta \to \mathbb{R}$, $\phi_{ij}^k : H_\beta \to \mathbb{R}$ for all $i, k \in \mathcal{I}$, $j \in \mathcal{J}$ and define

$$B(y)u = \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \mu_{ij}(y)\langle u, \tilde{e}_j\rangle_V e_i,$$

$$
\begin{aligned}
B'(y)\left(B(y)v,u\right) &= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} D\mu_{ij}(y)(B(y)v)\langle u,\tilde{e}_j\rangle_V e_i \\
&= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{I}} \phi_{ij}^k(y)\langle B(y)v,e_k\rangle_H\langle u,\tilde{e}_j\rangle_V e_i \\
&= \sum_{i,k\in\mathcal{I}}\sum_{j,r\in\mathcal{J}} \phi_{ij}^k(y)\mu_{kr}(y)\langle v,\tilde{e}_r\rangle_V\langle u,\tilde{e}_j\rangle_V e_i
\end{aligned}
$$

for $y\in H_\beta$, $u,v\in V_0$, where $D\mu_{ij}: H_\beta\to L(H,\mathbb{R})$ denotes the Fréchet derivative of μ_{ij} for all $i\in\mathcal{J}$, $j\in\mathcal{J}$. Note, the functionals μ_{ij}, ϕ_{ij}^k, $i,k\in\mathcal{I}$, $j\in\mathcal{J}$ have to be chosen such that $B(y)u\in H$ and $B'(y)\left(B(y)v,u\right)\in H$ for all $y\in H_\beta$, $u,v\in V_0$.

Before we consider some sample SPDEs, we rewrite the numerical schemes such that they fit this notation. Let us fix some $N,M,K\in\mathbb{N}$ throughout this section and let $m\in\{0,\dots,M-1\}$. For the Milstein scheme, we get

$$
\begin{aligned}
Y_{m+1}^{N,M,K} =& P_N\bigg(e^{Ah}\Big(Y_m^{N,M,K}+hF(Y_m^{N,M,K})+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\mu_{ij}(Y_m^{N,M,K})\sqrt{\eta_j}\Delta\beta_m^j e_i \\
&+\frac{1}{2}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\sum_{k\in\mathcal{I}}\phi_{ij}^k(Y_m^{N,M,K})\sum_{r\in\mathcal{J}_K}\mu_{kr}(Y_m^{N,M,K})\sqrt{\eta_r}\Delta\beta_m^r\sqrt{\eta_j}\Delta\beta_m^j e_i \\
&-\frac{h}{2}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\sum_{k\in\mathcal{I}}\eta_j\phi_{ij}^k(Y_m^{N,M,K})\mu_{kj}(Y_m^{N,M,K})e_i\Big)\bigg)
\end{aligned}
$$

and for the commutative derivative-free Milstein scheme, we obtain

$$
\begin{aligned}
&Y_{m+1}^{N,M,K} \\
&= P_N\bigg(e^{Ah}\Big(Y_m^{N,M,K}+hF(Y_m^{N,M,K})+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\mu_{ij}(Y_m^{N,M,K})\langle\Delta W_m^K,\tilde{e}_j\rangle_V e_i \\
&\quad+\frac{1}{\sqrt{h}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\Big(\mu_{ij}\Big(Y_m^{N,M,K}+\frac{\sqrt{h}}{2}P_N\Big(\sum_{k\in\mathcal{I}}\sum_{l\in\mathcal{J}}\mu_{kl}(Y_m^{N,M,K})\langle\Delta W_m^K,\tilde{e}_l\rangle_V e_k\Big)\Big) \\
&\qquad\qquad-\mu_{ij}(Y_m^{N,M,K})\Big)\langle\Delta W_m^K,\tilde{e}_j\rangle_V e_i \\
&\quad+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\Big(\mu_{ij}\Big(Y_m^{N,M,K}-\frac{h}{2}P_N\Big(\sqrt{\eta_j}\sum_{k\in\mathcal{I}}\mu_{kj}(Y_m^{N,M,K})e_k\Big)\Big)-\mu_{ij}(Y_m^{N,M,K})\Big)\sqrt{\eta_j}e_i\Big)\bigg) \\
&= P_N\bigg(e^{Ah}\Big(Y_m^{N,M,K}+hF(Y_m^{N,M,K})+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\mu_{ij}(Y_m^{N,M,K})\sqrt{\eta_j}\Delta\beta_m^j e_i \\
&\quad+\frac{1}{\sqrt{h}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\Big(\mu_{ij}\Big(Y_m^{N,M,K}+\frac{\sqrt{h}}{2}P_N\Big(\sum_{k\in\mathcal{I}}\sum_{l\in\mathcal{J}_K}\mu_{kl}(Y_m^{N,M,K})\sqrt{\eta_l}\Delta\beta_m^l e_k\Big)\Big) \\
&\qquad\qquad-\mu_{ij}(Y_m^{N,M,K})\Big)\sqrt{\eta_j}\Delta\beta_m^j e_i \\
&\quad+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\Big(\mu_{ij}\Big(Y_m^{N,M,K}-\frac{h}{2}P_N\Big(\sqrt{\eta_j}\sum_{k\in\mathcal{I}}\mu_{kj}(Y_m^{N,M,K})e_k\Big)\Big)-\mu_{ij}(Y_m^{N,M,K})\Big)\sqrt{\eta_j}e_i\Big)\bigg).
\end{aligned}
$$

Finally, we get for the linear implicit Euler scheme

$$Y_{m+1}^{LIE} = P_N\Big(\Big(I - hA\Big)^{-1}\Big(Y_m^{LIE} + hF(Y_m^{LIE}) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\mu_{ij}(Y_m^{LIE})\sqrt{\eta_j}\Delta\beta_m^j e_i\Big)\Big)$$

and the exponential Euler scheme reads

$$Y_{m+1}^{EES} = P_N\Big(e^{Ah}Y_m^{EES} + A^{-1}\Big(e^{Ah} - I\Big)F(Y_m^{EES}) + e^{Ah}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_K}\mu_{ij}(Y_m^{EES})\sqrt{\eta_j}\Delta\beta_m^j e_i\Big).$$

The Runge-Kutta type scheme by [71] is not applicable in this setting.

We fix the framework for the sample equations and choose $\mathcal{I} = \mathcal{J} = \mathbb{N}$. Let $T = 1$ and $H = V = L^2((0,1),\mathbb{R})$. Assume $A = \frac{\Delta}{100}$ with $\lambda_i = \frac{1}{100}\pi^2 i^2$, $e_i(x) = \sqrt{2}\sin(i\pi x)$ for all $x \in (0,1)$, $i \in \mathbb{N}$, and Dirichlet boundary conditions $X_t(0) = X_t(1) = 0$ for all $t \in (0,T]$. We consider $\{e_i,\ i \in \mathbb{N}\}$ as orthonormal basis of H. Moreover, we set $F(y) = 1 - y$ for $y \in H_\beta$ and $X_0(x) = 0$ for all $x \in (0,1)$. Let Q be a trace class operator and let $(W_t)_{t\in[0,T]}$ denote a Q-Wiener process in V. The eigenvalues of Q are denoted by η_j and the corresponding eigenfunctions by $\tilde{e}_j$ for all $j \in \mathbb{N}$; we choose $\eta_j = j^{-3}$ and $\tilde{e}_j = e_j$ for all $j \in \mathbb{N}$. We approximate the mild solution to various SPDEs in the following and, if not stated differently, assume this setting.

It is easily verified that conditions (C1), (C2), and (C4) hold in this context, see also [35].
We have to confirm that assumptions (C3) and (C5) are fulfilled. Therefore, we transfer these conditions to our framework and derive their dependence on the functionals $\mu_{ij}(y)$ and $\phi_{ij}^k(y)$ for $i,k \in \mathcal{I}$, $j \in \mathcal{J}$, $y \in H_\beta$.
We start with (C3) and rewrite $\|B(y)\|_{L(V,H_\delta)}$ for $y \in H_\delta$. We assume that μ_{ij}, $i \in \mathcal{I}$, $j \in \mathcal{J}$ and $\delta \in (0,\frac{1}{2})$ are chosen such that $B(H_\delta) \subset L(V,H_\delta)$. For $y \in H_\delta$, we obtain

$$\begin{aligned}
\|B(y)\|_{L(V,H_\delta)} &= \sup_{\substack{v\in V\\ \|v\|_V=1}} \|B(y)v\|_{H_\delta} = \sup_{\substack{v\in V\\ \|v\|_V=1}} \Big\|\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\mu_{ij}(y)\langle v,\tilde{e}_j\rangle_V e_i\Big\|_{H_\delta} \\
&= \sup_{\substack{v\in V\\ \|v\|_V=1}} \Big\|(-A)^\delta\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\mu_{ij}(y)\langle v,\tilde{e}_j\rangle_V e_i\Big\|_H \\
&= \sup_{\substack{v\in V\\ \|v\|_V=1}} \Big\|\sum_{k\in\mathcal{I}}\lambda_k^\delta\langle\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\mu_{ij}(y)\langle v,\tilde{e}_j\rangle_V e_i, e_k\rangle_H e_k\Big\|_H \\
&= \sup_{\substack{v\in V\\ \|v\|_V=1}} \Big\|\sum_{k\in\mathcal{I}}\lambda_k^\delta\sum_{j\in\mathcal{J}}\mu_{kj}(y)\langle v,\tilde{e}_j\rangle_V e_k\Big\|_H \\
&\le \sup_{\substack{v\in V\\ \|v\|_V=1}} \Big(\sum_{k\in\mathcal{I}}\sum_{j\in\mathcal{J}}\lambda_k^\delta|\mu_{kj}(y)||\langle v,\tilde{e}_j\rangle_V|\Big) \le \sum_{k\in\mathcal{I}}\sum_{j\in\mathcal{J}}\lambda_k^\delta|\mu_{kj}(y)|.
\end{aligned} \tag{3.37}$$

Condition (C3) requires $\|B(y)\|_{L(V,H_\delta)} \le C(1+\|y\|_{H_\delta})$. This is analyzed for the specific examples below.

Moreover, we rewrite

$$\begin{aligned}
&\|B'(z)B(z) - B'(w)B(w)\|^2_{L^{(2)}_{HS}(V_0,H)} \\
&= \sum_{k,l\in\mathcal{J}} \left\|\sqrt{\eta_k}\sqrt{\eta_l}\left(B'(z)(B(z)\tilde{e}_k,\tilde{e}_l) - B'(w)(B(w)\tilde{e}_k,\tilde{e}_l)\right)\right\|^2_H \\
&= \sum_{k,l\in\mathcal{J}} \eta_k\eta_l \Big\| \sum_{i,r\in\mathcal{I}} \phi^r_{il}(z)\langle \sum_{j_1\in\mathcal{I}}\sum_{j_2\in\mathcal{J}} \mu_{j_1j_2}(z)\langle\tilde{e}_k,\tilde{e}_{j_2}\rangle_V e_{j_1}, e_r\rangle_H e_i \\
&\qquad\qquad - \sum_{i,r\in\mathcal{I}} \phi^r_{il}(w)\langle \sum_{j_1\in\mathcal{I}}\sum_{j_2\in\mathcal{J}} \mu_{j_1j_2}(w)\langle\tilde{e}_k,\tilde{e}_{j_2}\rangle_V e_{j_1}, e_r\rangle_H e_i \Big\|^2_H \\
&= \sum_{k,l\in\mathcal{J}} \eta_k\eta_l \sum_{i,r_1,r_2\in\mathcal{I}} \left(\phi^{r_1}_{il}(z)\mu_{r_1k}(z) - \phi^{r_1}_{il}(w)\mu_{r_1k}(w)\right)\left(\phi^{r_2}_{il}(z)\mu_{r_2k}(z) - \phi^{r_2}_{il}(w)\mu_{r_2k}(w)\right)
\end{aligned}$$

for $w, z \in H_\gamma$.

Next, we examine the conditions on the derivatives of B. We denote the Fréchet derivative of ϕ^k_{ij} in direction of e_r, $i,k,r\in\mathcal{I}$, $j\in\mathcal{J}$ as $\hat{\phi}^r_{ijk}$ and obtain for all $y\in H_\beta$, $z,w\in H$

$$\begin{aligned}
\|B'(y)\|_{L(H,L(V,H))} &= \sup_{\substack{w\in H\\ \|w\|_H=1}} \|B'(y)w\|_{L(V,H)} = \sup_{\substack{w\in H, v\in V\\ \|w\|_H=1,\|v\|_V=1}} \|B'(y)(w,v)\|_H \\
&= \sup_{\substack{w\in H, v\in V\\ \|w\|_H=1,\|v\|_V=1}} \Big\| \sum_{i,k\in\mathcal{I}}\sum_{j\in\mathcal{J}} \phi^k_{ij}(y)\langle w,e_k\rangle_H\langle v,\tilde{e}_j\rangle_V e_i\Big\|_H \\
&\le \sup_{\substack{w\in H, v\in V\\ \|w\|_H=1,\|v\|_V=1}} \Big(\sum_{i,k\in\mathcal{I}}\sum_{j\in\mathcal{J}} |\phi^k_{ij}(y)||\langle w,e_k\rangle_H||\langle v,\tilde{e}_j\rangle_V|\Big) \\
&\le \sum_{i,k\in\mathcal{I}}\sum_{j\in\mathcal{J}} |\phi^k_{ij}(y)|
\end{aligned} \tag{3.38}$$

and for the second derivative, we get

$$\begin{aligned}
\|B''(y)\|_{L^{(2)}(H,L(V,H))} &= \sup_{\substack{z\in H\\ \|z\|_H=1}} \|B''(y)z\|_{L(H,L(V,H))} = \sup_{\substack{z,w\in H\\ \|z\|_H=\|w\|_H=1}} \|B''(y)(z,w)\|_{L(V,H)} \\
&= \sup_{\substack{z,w\in H, v\in V\\ \|z\|_H=\|w\|_H=\|v\|_V=1}} \Big\| \sum_{i,r,l\in\mathcal{I}}\sum_{j\in\mathcal{J}} \hat{\phi}^l_{ijr}(y)\langle z,e_l\rangle_H\langle w,e_r\rangle_H\langle v,\tilde{e}_j\rangle_V e_i\Big\|_H \\
&\le \sup_{\substack{z,w\in H, v\in V\\ \|z\|_H=\|w\|_H=\|v\|_V=1}} \Big(\sum_{i,r,l\in\mathcal{I}}\sum_{j\in\mathcal{J}} |\hat{\phi}^l_{ijr}(y)||\langle z,e_l\rangle_H||\langle w,e_r\rangle_H||\langle v,\tilde{e}_j\rangle_V|\Big) \\
&\le \sum_{i,r,l\in\mathcal{I}}\sum_{j\in\mathcal{J}} |\hat{\phi}^l_{ijr}(y)|.
\end{aligned} \tag{3.39}$$

The last condition in (C3) reads

$$\begin{aligned}
\|(-A)^{-\vartheta}B(z)Q^{-\alpha}\|_{L_{HS}(V_0,H)} &= \Big(\sum_{k\in\mathcal{J}} \|(-A)^{-\vartheta}B(z)Q^{-\alpha+\frac{1}{2}}\tilde{e}_k\|^2_H\Big)^{\frac{1}{2}} \\
&= \Big(\sum_{k\in\mathcal{J}} \eta_k^{1-2\alpha}\Big\|\sum_{i\in\mathcal{I}} \lambda_i^{-\vartheta}\langle B(z)\tilde{e}_k,e_i\rangle_H e_i\Big\|^2_H\Big)^{\frac{1}{2}}
\end{aligned}$$

$$= \Big(\sum_{k\in\mathcal{J}} \eta_k^{1-2\alpha} \Big\| \sum_{i\in\mathcal{I}} \lambda_i^{-\vartheta} \mu_{ik}(z) e_i \Big\|_H^2 \Big)^{\frac{1}{2}}$$
$$= \Big(\sum_{k\in\mathcal{J}} \eta_k^{1-2\alpha} \sum_{i\in\mathcal{I}} \lambda_i^{-2\vartheta} \mu_{ik}^2(z) \Big)^{\frac{1}{2}} \tag{3.40}$$

for $z \in H_\gamma$.

We need assumption (C5) to rewrite the iterated stochastic integral as in equation (3.11). Therefore, condition (C5) reads

$$\sum_{k\in\mathcal{I}} \phi_{im}^k(y)\mu_{kn}(y) = \sum_{k\in\mathcal{I}} \phi_{in}^k(y)\mu_{km}(y) \tag{3.41}$$

for all $y \in H_\beta$, $i \in \mathcal{I}$, $m, n \in \mathcal{J}_K$, $K \in \mathbb{N}$ in this framework.
Now, we analyze some specific examples and investigate the effective order of convergence numerically. We prove (C3) and (C5) based on the preparatory work for each equation.

Example 1 - Linear Equation

Let $\mu : H_\beta \to \mathbb{R}$ be given by $\mu_{ij}(y) = \frac{\langle y, e_i\rangle_H}{i^4+j^4}$; we obtain $\phi_{ij}^k(y) = \begin{cases} 0, & k \neq i \\ \frac{1}{i^4+j^4}, & k = i \end{cases}$ for all $i, k \in \mathcal{I}$, $j \in \mathcal{J}$, $y \in H_\beta$.

Before we investigate the results of the numerical simulations, we show step by step that (C3) is fulfilled. By (3.37), we get for $y \in H_\beta$

$$\begin{aligned}\|B(y)\|_{L(V,H_\delta)} &\leq \sum_{k\in\mathcal{I}}\sum_{j\in\mathcal{J}} \lambda_k^\delta |\mu_{kj}(y)| = \sum_{k\in\mathcal{I}}\sum_{j\in\mathcal{J}} \Big(\frac{1}{100}\pi^2 k^2\Big)^\delta \frac{|\langle y, e_k\rangle_H|}{k^4+j^4} \\ &\leq C \sum_{k\in\mathcal{I}}\sum_{j\in\mathcal{J}} \frac{1}{k^{2-2\delta}} \frac{1}{j^2} \|(-A)^{-\delta}\|_{L(H)} \|y\|_{H_\delta}.\end{aligned}$$

We obtain $\|B(y)\|_{L(V,H_\delta)} \leq C(1 + \|y\|_{H_\delta})$ for all $y \in H_\delta$ and $\delta \in (0, \frac{1}{2})$ due to Theorem 2.3.

For

$$\begin{aligned}&\|B'(z)B(z) - B'(w)B(w)\|^2_{L_{HS}^{(2)}(V_0,H)} \\ &= \sum_{k,l\in\mathcal{J}} \eta_k\eta_l \sum_{i,r_1,r_2\in\mathcal{I}} \big(\phi_{il}^{r_1}(z)\mu_{r_1k}(z) - \phi_{il}^{r_1}(w)\mu_{r_1k}(w)\big)\big(\phi_{il}^{r_2}(z)\mu_{r_2k}(z) - \phi_{il}^{r_2}(w)\mu_{r_2k}(w)\big)\end{aligned}$$

we compute

$$\begin{aligned}\|B'(z)B(z) - B'(w)B(w)\|^2_{L_{HS}^{(2)}(V_0,H)} &= \sum_{k,l\in\mathcal{J}} \frac{1}{k^3}\frac{1}{l^3} \sum_{i,r_1\in\mathcal{I}} \frac{1}{(i^4+l^4)^2} \mathbb{1}_{i=r_1} \Big(\frac{\langle z, e_{r_1}\rangle_H}{r_1^4+k^4} - \frac{\langle w, e_{r_1}\rangle_H}{r_1^4+k^4}\Big)^2 \\ &= \sum_{k,l\in\mathcal{J}} \frac{1}{k^3}\frac{1}{l^3} \sum_{i\in\mathcal{I}} \frac{1}{(i^4+l^4)^2} \Big(\frac{\langle z, e_i\rangle_H}{i^4+k^4} - \frac{\langle w, e_i\rangle_H}{i^4+k^4}\Big)^2\end{aligned}$$

$$= \sum_{k,l\in\mathcal{J}} \frac{1}{k^3}\frac{1}{l^3} \sum_{i\in\mathcal{I}} \frac{1}{(i^4+l^4)^2}\Big(\frac{\langle z-w, e_i\rangle_H}{i^4+k^4}\Big)^2$$
$$\le C \sum_{k,l\in\mathcal{J}} \frac{1}{k^7}\frac{1}{l^7} \sum_{i\in\mathcal{I}} \frac{1}{i^8}\|z-w\|_H^2$$

for $v, w \in H_\gamma$.
Next, we consider the Fréchet derivatives of B. It holds,

$$\|B'(y)\|_{L(H,L(V,H))} \le \sum_{i,k\in\mathcal{I}}\sum_{j\in\mathcal{J}} \frac{1}{i^4+j^4}\mathbb{1}_{k=i} \le \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \frac{1}{i^2}\frac{1}{j^2},$$

that is,

$$\|B'(y)\|_{L(H,L(V,H))} < \infty$$

for all $y \in H_\beta$.
As $\hat{\phi}^r_{ijk}(y) = 0$ for all $i, k, r \in \mathcal{I}$, $j \in \mathcal{J}$, and $y \in H_\beta$, it follows

$$\sup_{y\in H_\beta} \|B''(y)\|_{L^{(2)}(H,L(V,H))} < \infty.$$

Finally, we determine α and ϑ such that

$$\|(-A)^{-\vartheta}B(z)Q^{-\alpha}\|_{L_{HS}(V_0,H)} \le C(1+\|z\|_{H_\gamma})$$

is fulfilled for $z \in H_\gamma$

$$\|(-A)^{-\vartheta}B(z)Q^{-\alpha}\|_{L_{HS}(V_0,H)} = \Big(\sum_{k\in\mathcal{J}} \eta_k^{1-2\alpha} \sum_{i\in\mathcal{I}} \lambda_i^{-2\vartheta}\mu_{ik}^2(z)\Big)^{\frac{1}{2}}$$
$$= \Big(\sum_{k\in\mathcal{J}} \frac{1}{k^{3(1-2\alpha)}} \sum_{i\in\mathcal{I}} \lambda_i^{-2\vartheta}\frac{\langle z, e_i\rangle_H^2}{(i^4+k^4)^2}\Big)^{\frac{1}{2}}$$
$$\le C\Big(\sum_{k\in\mathcal{J}} \frac{1}{k^{3(1-2\alpha)+4}} \sum_{i\in\mathcal{I}} \frac{1}{i^{4+4\vartheta}}\Big)^{\frac{1}{2}}\|z\|_{H_\gamma}.$$

We find

$$\|(-A)^{-\vartheta}B(z)Q^{-\alpha}\|_{L_{HS}(V_0,H)} \le C(1+\|z\|_{H_\gamma})$$

for $\alpha \in (0,1)$, $\vartheta \in (0,\frac{1}{2})$, and all $z \in H_\gamma$.

The parameters in assumption (C3) are therefore given by: $\delta, \vartheta \in (0, \frac{1}{2})$, $\alpha \in (0,1)$. We choose the maximal value for δ and $\beta = 0$, which implies $\gamma \in [\frac{1}{2}, 1)$ and $q_{cDFM} = \gamma$. Then, we select $q_{cDFM} = \gamma = \alpha = 1 - \varepsilon$ for some arbitrary $\varepsilon > 0$.

It remains to verify (C5). We get

$$\begin{aligned}\sum_{k\in\mathcal{I}}\phi_{im}^k(y)\mu_{kn}(y) &= \sum_{k\in\mathcal{I}}\frac{1}{i^4+m^4}\mathbb{1}_{k=i}\frac{\langle y,e_k\rangle_H}{k^4+n^4} = \frac{1}{i^4+m^4}\frac{\langle y,e_i\rangle_H}{i^4+n^4}\\ &= \sum_{k\in\mathcal{I}}\frac{1}{i^4+n^4}\mathbb{1}_{k=i}\frac{\langle y,e_k\rangle_H}{k^4+m^4} = \sum_{k\in\mathcal{I}}\phi_{in}^k(y)\mu_{km}(y)\end{aligned}$$

for all $i\in\mathcal{I}$, $n,m\in\mathcal{J}_K$, $K\in\mathbb{N}$, $y\in H_\beta$.

The remaining parameters in the error estimate are $\rho_Q=3$ and $\rho_A=2$; this yields $K=N^{\frac{2}{3}}$ and $M=N^2$ for the cDFM and the Milstein scheme according to (3.18). In the exponential and linear implicit Euler scheme, we need to take $M=N^4$, however. Then we compute the effective order of convergence as $\operatorname{error}(\mathrm{MIL}(N,K,M))=\mathcal{O}\Big(\bar{c}^{-\frac{3}{7}+\varepsilon}\Big)$ and $\operatorname{error}(\mathrm{LIE}(N,K,M))=\operatorname{error}(\mathrm{EES}(N,K,M))=\mathcal{O}\Big(\bar{c}^{-\frac{6}{17}+\varepsilon}\Big)$, whereas for the commutative derivative-free Milstein scheme, we obtain $\operatorname{error}(\mathrm{cDFM}(N,K,M))=\mathcal{O}\Big(\bar{c}^{-\frac{6}{11}+\varepsilon}\Big)$ for some $\varepsilon>0$.

The following logarithmic plot of the error for $N=2,2^2,\ldots,2^5$ and 700 paths confirms the theoretical results in Section 3.4. In order to compute the error, we compare the simulation results to an approximation obtained with the linear implicit Euler scheme with $N_X=2^6$, $K_X=2^4$, and $M_X=2^{20}$ as no explicit solution is available. As suggested by the analysis in Section 3.4, we observe a higher effective order of convergence for the cDFM than for the Milstein or Euler schemes.

		Milstein			cDFM		
N	M	CC	Error	Std	CC	Error	Std
2	4	$2^{4+\frac{2}{3}}$	$3.0\cdot10^{-2}$	$1.5\cdot10^{-3}$	$3\cdot2^{3+\frac{2}{3}}$	$3.0\cdot10^{-2}$	$1.5\cdot10^{-3}$
4	2^4	$2^{8+\frac{4}{3}}$	$2.5\cdot10^{-2}$	$3.0\cdot10^{-4}$	$3\cdot2^{6+\frac{4}{3}}$	$2.5\cdot10^{-2}$	$3.0\cdot10^{-4}$
8	2^6	2^{14}	$1.7\cdot10^{-2}$	$6.0\cdot10^{-5}$	$3\cdot2^{10}$	$1.7\cdot10^{-2}$	$6.0\cdot10^{-5}$
16	2^8	$2^{16+\frac{8}{3}}$	$6.3\cdot10^{-3}$	$1.1\cdot10^{-5}$	$3\cdot2^{12+\frac{8}{3}}$	$6.3\cdot10^{-3}$	$1.1\cdot10^{-5}$
32	2^{10}	$2^{20+\frac{10}{3}}$	$1.6\cdot10^{-3}$	$2.0\cdot10^{-6}$	$3\cdot2^{15+\frac{10}{3}}$	$1.6\cdot10^{-3}$	$2.0\cdot10^{-6}$

		Linear Implicit Euler			Exponential Euler		
N	M	CC	Error	Std	CC	Error	Std
2	2^4	$2^{5+\frac{2}{3}}$	$2.2\cdot10^{-2}$	$4.0\cdot10^{-3}$	$2^{5+\frac{2}{3}}$	$2.3\cdot10^{-2}$	$4.0\cdot10^{-3}$
4	2^8	$2^{10+\frac{4}{3}}$	$2.7\cdot10^{-2}$	$6.5\cdot10^{-4}$	$2^{10+\frac{4}{3}}$	$2.7\cdot10^{-2}$	$6.5\cdot10^{-4}$
8	2^{12}	2^{17}	$1.7\cdot10^{-2}$	$1.2\cdot10^{-4}$	2^{17}	$1.7\cdot10^{-2}$	$1.1\cdot10^{-4}$
16	2^{16}	$2^{20+\frac{8}{3}}$	$6.1\cdot10^{-3}$	$2.3\cdot10^{-5}$	$2^{20+\frac{8}{3}}$	$6.1\cdot10^{-3}$	$2.3\cdot10^{-5}$
32	2^{20}	$2^{25+\frac{10}{3}}$	$1.5\cdot10^{-3}$	$3.9\cdot10^{-6}$	$2^{25+\frac{10}{3}}$	$1.5\cdot10^{-3}$	$3.9\cdot10^{-6}$

Table 3.3: Error and standard deviation for 700 paths for Example 1, computed with batches of size 50 ([38, p.312]). CC denotes the computational cost needed to evaluate the term that dominates the computational effort according to Table 3.1 for all time steps M.

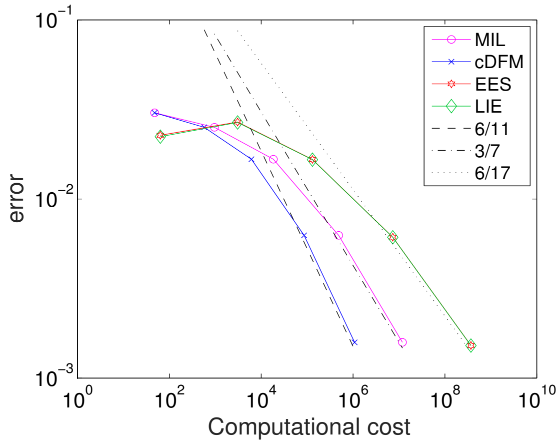

Figure 3.3: Error against computational cost for Example 1 for 700 paths and $N = 2, 4, 8, 16, 32$ in log-log scale.

Example 2 - Differing Bases

In contrast to the standard setting, we choose $\tilde{e}_j = \sqrt{2}\cos(j\pi x)$, $j \in \mathcal{J}$, $x \in (0,1)$, here. This example shows that we can choose different basis functions for H and V. Let $\mu_{ij}(y) = \frac{1}{j^2}\sum_{p\in\mathcal{I}} \frac{\langle y, e_p\rangle_H}{i^3+p^4}$ for all $i \in \mathcal{I}$, $j \in \mathcal{J}$, $y \in H_\beta$. This yields $\phi_{ij}^k(y) = \frac{1}{j^2}\frac{1}{i^3+k^4}$ for all $i, k \in \mathcal{I}$, $j \in \mathcal{J}$, and $y \in H_\beta$.

As in the previous example, we need to check that (C3) and (C5) are fulfilled. We do not detail these computations here as they can be conducted analogously to the above. We simply state the ranges for the parameters, which are $\delta \in (0, \frac{1}{4})$, $\vartheta \in (0, \frac{1}{2})$, $\alpha \in (0, 1)$. We select the maximal value for δ. For $\beta = 0$, this implies $\gamma \in [\frac{1}{4}, \frac{3}{4})$ and, for the cDFM, we get $q_{cDFM} = \gamma$. Then, the optimal parameter choice is $q_{cDFM} = \gamma = \frac{3}{4} - \varepsilon$ and $\alpha = 1 - \varepsilon$ for some $\varepsilon > 0$.
We verify the commutativity condition, however,

$$\sum_{k\in\mathcal{I}} \phi_{im}^k(y)\mu_{kn}(y) = \sum_{k\in\mathcal{I}} \frac{1}{m^2}\frac{1}{i^3+k^4}\frac{1}{n^2}\sum_{p\in\mathcal{I}} \frac{\langle y, e_p\rangle_H}{k^3+p^4} = \sum_{k\in\mathcal{I}} \phi_{in}^k(y)\mu_{km}(y)$$

for all $i \in \mathcal{I}$, $n, m \in \mathcal{J}_K$, $y \in H_\beta$, $K \in \mathbb{N}$.
For this parameter setting, we derive the following relation between N, M, K from Section 3.4: $K = \sqrt{N}$, $M = N^2$ for the Milstein and commutative derivative-free Milstein scheme and $K = \sqrt{N}$, $M = N^3$ for the linear implicit and the exponential Euler. The exact solution is replaced by an approximation obtained with the linear implicit Euler scheme for $N_X = 2^7$, $K_X = \sqrt{N_X}$, and $M_X = 2^{18}$ as no explicit solution is available.
First, we compute the effective order of convergence of the cDFM as $\text{error}(\text{cDFM}(N, K, M)) = \mathcal{O}\Big(\bar{c}^{-\frac{3}{7}+\varepsilon}\Big)$. This rate is higher than for the other numerical schemes, which all achieve the same rate $\text{error}(\text{MIL}(N, K, M)) = \text{error}(\text{LIE}(N, K, M)) = \text{error}(\text{EES}(N, K, M)) = \mathcal{O}\Big(\bar{c}^{-\frac{1}{3}+\varepsilon}\Big)$. Again, the expected difference in the effective order of convergence is confirmed by the following table and plot.

		Milstein			cDFM		
N	M	CC	Error	Std	CC	Error	Std
2	4	$2^{4+\frac{1}{2}}$	$3.2 \cdot 10^{-2}$	$3.0 \cdot 10^{-3}$	$3 \cdot 2^{3+\frac{1}{2}}$	$3.2 \cdot 10^{-2}$	$3.0 \cdot 10^{-3}$
4	2^4	2^9	$2.5 \cdot 10^{-2}$	$5.0 \cdot 10^{-4}$	$3 \cdot 2^7$	$2.5 \cdot 10^{-2}$	$5.0 \cdot 10^{-4}$
8	2^6	$2^{12+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$6.2 \cdot 10^{-5}$	$3 \cdot 2^{9+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$6.2 \cdot 10^{-5}$
16	2^8	2^{18}	$6.6 \cdot 10^{-3}$	$2.0 \cdot 10^{-5}$	$3 \cdot 2^{14}$	$6.6 \cdot 10^{-3}$	$2.0 \cdot 10^{-5}$
32	2^{10}	$2^{20+\frac{5}{2}}$	$2.0 \cdot 10^{-3}$	$7.0 \cdot 10^{-6}$	$3 \cdot 2^{15+\frac{5}{2}}$	$2.0 \cdot 10^{-3}$	$7.0 \cdot 10^{-6}$
64	2^{12}	2^{27}	$4.6 \cdot 10^{-4}$	$5.4 \cdot 10^{-6}$	$3 \cdot 2^{21}$	$4.6 \cdot 10^{-4}$	$5.4 \cdot 10^{-6}$

		Linear Implicit Euler			Exponential Euler		
N	M	CC	Error	Std	CC	Error	Std
2	2^3	$2^{4+\frac{1}{2}}$	$2.4 \cdot 10^{-2}$	$4.6 \cdot 10^{-3}$	$2^{4+\frac{1}{2}}$	$2.4 \cdot 10^{-2}$	$4.7 \cdot 10^{-3}$
4	2^6	2^9	$2.7 \cdot 10^{-2}$	$6.2 \cdot 10^{-4}$	2^9	$2.7 \cdot 10^{-2}$	$6.7 \cdot 10^{-4}$
8	2^9	$2^{12+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-4}$	$2^{12+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$1.8 \cdot 10^{-4}$
16	2^{12}	2^{18}	$6.5 \cdot 10^{-3}$	$5.3 \cdot 10^{-5}$	2^{18}	$6.7 \cdot 10^{-3}$	$5.9 \cdot 10^{-5}$
32	2^{15}	$2^{20+\frac{5}{2}}$	$1.9 \cdot 10^{-3}$	$7.9 \cdot 10^{-6}$	$2^{20+\frac{5}{2}}$	$2.0 \cdot 10^{-3}$	$9.6 \cdot 10^{-6}$
64	2^{18}	2^{27}	$4.3 \cdot 10^{-4}$	$4.5 \cdot 10^{-6}$	2^{27}	$4.8 \cdot 10^{-4}$	$5.8 \cdot 10^{-6}$

Table 3.4: Error and standard deviation for Example 2 obtained from 500 paths. CC denotes the computational cost needed to evaluate the term that dominates the computational effort according to Table 3.1 for all time steps M.

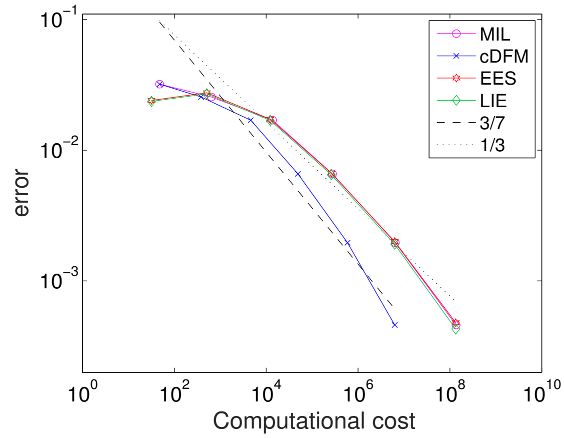

Figure 3.4: Error against computational cost for Example 2 for 500 paths and $N = 2, 4, 8, 16, 32, 64$ in log-log scale.

Example 3 - Nonlinear Equation

Equations that are nonlinear can also be treated with the approximation schemes presented above; we consider functionals $\mu_{ij} : H_\beta \to \mathbb{R}$, $i \in \mathcal{I}$, $j \in \mathcal{J}$ which are nonlinear in H_β, here. Let this functional be given as

$$\mu_{ij}(y) = \sum_{p\in\mathcal{I}} \Big(\frac{1}{i^{\frac{3}{2}} j^2} \frac{1}{i+j+p^2} \mathbb{1}_{(j-1)^2+1 \leq i,p \leq j^2}$$

$$+\frac{1}{j^2}\sum_{r=0}^{j-2}\frac{1}{i^{\frac{3}{2}}}\frac{1}{(r+1)}\frac{1}{i+(r+1)+p^2}\mathbb{1}_{r^2+1\le i,p\le (r+1)^2}\Big)e^{-\langle y,e_p\rangle_H^2},$$

for $\phi_{ij}^k(y)$, we obtain

$$\begin{aligned}\phi_{ij}^k(y) =&\Big(\frac{1}{i^{\frac{3}{2}}j^2}\frac{1}{i+j+k^2}\mathbb{1}_{(j-1)^2+1\le i,k\le j^2}+\frac{1}{j^2}\sum_{r=0}^{j-2}\frac{1}{i^{\frac{3}{2}}}\frac{1}{(r+1)}\frac{1}{i+(r+1)+k^2}\mathbb{1}_{r^2+1\le i,k\le (r+1)^2}\Big)\\ &\cdot e^{-\langle y,e_k\rangle_H^2}(-2\langle y,e_k\rangle_H)\end{aligned}$$

for all $i,k\in\mathcal{I}$, $j\in\mathcal{J}$, and $y\in H_\beta$. Here, we return to $\tilde{e}_j=e_j$ for all $j\in\mathcal{J}$.

We only elaborate on assumption (C5) here, as this condition is not verified at first glance. It holds

$$\begin{aligned}&\sum_{k\in\mathcal{I}}\phi_{im}^k(y)\mu_{kn}(y)\\ &=\sum_{k\in\mathcal{I}}\Big(\frac{1}{i^{\frac{3}{2}}m^2}\frac{1}{i+m+k^2}\mathbb{1}_{(m-1)^2+1\le i,k\le m^2}\\ &\qquad+\frac{1}{m^2}\sum_{r=0}^{m-2}\frac{1}{i^{\frac{3}{2}}(r+1)}\frac{1}{i+(r+1)+k^2}\mathbb{1}_{r^2+1\le i,k\le (r+1)^2}\Big)e^{-\langle y,e_k\rangle_H^2}(-2\langle y,e_k\rangle_H)\\ &\qquad\cdot\Big(\sum_{p\in\mathcal{I}}\frac{1}{k^{\frac{3}{2}}n^2}\frac{1}{k+n+p^2}\mathbb{1}_{(n-1)^2+1\le k,p\le n^2}\\ &\qquad\quad+\sum_{p\in\mathcal{I}}\frac{1}{n^2}\sum_{r=0}^{n-2}\frac{1}{k^{\frac{3}{2}}(r+1)}\frac{1}{k+(r+1)+p^2}\mathbb{1}_{r^2+1\le k,p\le (r+1)^2}\mathbb{1}_{k=p}\Big)e^{-\langle y,e_p\rangle_H^2}\\ &=\sum_{k=\max((m-1)^2,(n-1)^2)+1}^{\min(m^2,n^2)}\sum_{p=(n-1)^2+1}^{n^2}\frac{1}{i^{\frac{3}{2}}m^2}\frac{1}{i+m+k^2}\frac{1}{k^{\frac{3}{2}}n^2}\frac{1}{k+n+p^2}\mathbb{1}_{(m-1)^2+1\le i\le m^2}\\ &\quad\cdot e^{-\langle y,e_k\rangle_H^2-\langle y,e_p\rangle_H^2}(-2\langle y,e_k\rangle_H)\\ &\quad+\sum_{r=0}^{n-2}\sum_{k=\max((m-1)^2,r^2)+1}^{\min(m^2,(r+1)^2)}\sum_{p=(r^2+1}^{(r+1)^2}\frac{1}{m^2n^2}\frac{1}{i^{\frac{3}{2}}}\frac{1}{i+m+k^2}\frac{1}{k^{\frac{3}{2}}(r+1)}\frac{1}{k+(r+1)+p^2}\\ &\quad\cdot\mathbb{1}_{(m-1)^2+1\le i\le m^2}e^{-\langle y,e_k\rangle_H^2-\langle y,e_p\rangle_H^2}(-2\langle y,e_k\rangle_H)\\ &\quad+\sum_{r=0}^{m-2}\sum_{k=\max((n-1)^2,r^2)+1}^{\min(n^2,(r+1)^2)}\sum_{p=(n-1)^2+1}^{n^2}\frac{1}{m^2}\frac{1}{i^{\frac{3}{2}}(r+1)}\frac{1}{i+(r+1)+k^2}\frac{1}{k^{\frac{3}{2}}n^2}\frac{1}{k+n+p^2}\\ &\quad\cdot\mathbb{1}_{r^2+1\le i\le (r+1)^2}e^{-\langle y,e_k\rangle_H^2-\langle y,e_p\rangle_H^2}(-2\langle y,e_k\rangle_H)\\ &\quad+\sum_{r=0}^{n-2}\sum_{\tilde{r}=0}^{m-2}\sum_{k=\max(\tilde{r}^2,r^2)+1}^{\min((\tilde{r}+1)^2,(r+1)^2)}\sum_{p=r^2+1}^{(r+1)^2}\frac{1}{m^2n^2}\frac{1}{i^{\frac{3}{2}}(\tilde{r}+1)}\frac{1}{i+(\tilde{r}+1)+k^2}\frac{1}{k^{\frac{3}{2}}(r+1)}\\ &\quad\cdot\mathbb{1}_{\tilde{r}^2+1\le i\le (\tilde{r}+1)^2}\frac{1}{k+(r+1)+p^2}e_ie^{-\langle y,e_k\rangle_H^2-\langle y,e_p\rangle_H^2}(-2\langle y,e_k\rangle_H)\\ &=\sum_{k,p=(m-1)^2+1}^{m^2}\frac{1}{i^{\frac{3}{2}}k^{\frac{3}{2}}m^4}\frac{1}{i+m+k^2}\frac{(-2\langle y,e_k\rangle_H)}{k+m+p^2}e^{-\langle y,e_k\rangle_H^2}e^{-\langle y,e_p\rangle_H^2}\mathbb{1}_{(m-1)^2+1\le i\le m^2}\mathbb{1}_{m=n}\end{aligned}$$

$$+ \sum_{k,p=(m-1)^2+1}^{m^2} \frac{1}{n^2m^2} \frac{1}{i^{\frac{3}{2}}} \frac{1}{i+m+k^2} \frac{1}{k^{\frac{3}{2}}m} \frac{(-2\langle y,e_k\rangle_H)}{k+m+p^2} e^{-\langle y,e_k\rangle_H^2} e^{-\langle y,e_p\rangle_H^2} \mathbb{1}_{(m-1)^2+1\leq i\leq m^2} \mathbb{1}_{m<n}$$

$$+ \sum_{k,p=(n-1)^2+1}^{n^2} \frac{1}{m^2} \frac{1}{i^{\frac{3}{2}}n} \frac{1}{i+n+k^2} \frac{1}{k^{\frac{3}{2}}n^2} \frac{(-2\langle y,e_k\rangle_H)}{k+n+p^2} e^{-\langle y,e_k\rangle_H^2} e^{-\langle y,e_p\rangle_H^2} \mathbb{1}_{(n-1)^2+1\leq i\leq n^2} \mathbb{1}_{n<m}$$

$$+ \sum_{r=0}^{\min(n-2,m-2)} \sum_{k,p=r^2+1}^{(r+1)^2} \frac{1}{m^2n^2} \frac{1}{i^{\frac{3}{2}}(r+1)} \frac{1}{i+(r+1)+k^2} \frac{1}{k^{\frac{3}{2}}(r+1)} \frac{1}{k+(r+1)+p^2}$$

$$\cdot \mathbb{1}_{r^2+1\leq i\leq (r+1)^2} e^{-\langle y,e_k\rangle_H^2 - \langle y,e_p\rangle_H^2} (-2\langle y,e_k\rangle_H)$$

$$= \sum_{k\in\mathcal{I}} \phi_{in}^k(y)\mu_{km}(y)$$

for all $i \in \mathcal{I}$, $m, n \in \mathcal{J}_K$, $K \in \mathbb{N}$, and $y \in H_\beta$. So, the equation is commutative.

Assumption (C3) holds true with $\delta \in (0, \frac{1}{4})$, $\vartheta \in (0, \frac{1}{2})$, $\alpha \in (0,1)$, and $\beta = 0$, which yields $q_{cDFM} = \gamma \in [\frac{1}{4}, \frac{3}{4})$ for a maximal δ. These are the same parameters as in the preceding example. Therefore, we obtain $K = \sqrt{N}$, $M = N^2$ for the Milstein and commutative derivative-free Milstein scheme and $K = \sqrt{N}$, $M = N^3$ for the other schemes. Naturally, for the effective orders of convergence we expect error(MIL(N,K,M)) = error(LIE(N,K,M)) = error(EES(N,K,M)) = $\mathcal{O}\Big(\bar{c}^{-\frac{1}{3}+\varepsilon}\Big)$ and error(cDFM(N,K,M)) = $\mathcal{O}\Big(\bar{c}^{-\frac{3}{7}+\varepsilon}\Big)$ again.
In the computation of the mean square-error, we employ an approximation obtained with the linear implicit Euler scheme for $N_X = 2^7$, $K_X = \sqrt{N_X}$, and $M_X = 2^{18}$.

		Milstein			cDFM		
N	M	CC	Error	Std	CC	Error	Std
2	4	$2^{4+\frac{1}{2}}$	$2.9 \cdot 10^{-2}$	$2.3 \cdot 10^{-3}$	$3 \cdot 2^{3+\frac{1}{2}}$	$2.8 \cdot 10^{-2}$	$2.1 \cdot 10^{-3}$
4	2^4	2^9	$2.5 \cdot 10^{-2}$	$3.8 \cdot 10^{-4}$	$3 \cdot 2^7$	$2.5 \cdot 10^{-2}$	$3.9 \cdot 10^{-4}$
8	2^6	$2^{12+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$6.3 \cdot 10^{-5}$	$3 \cdot 2^{9+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$6.4 \cdot 10^{-5}$
16	2^8	2^{18}	$6.6 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$	$3 \cdot 2^{14}$	$6.6 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$
32	2^{10}	$2^{20+\frac{5}{2}}$	$1.9 \cdot 10^{-3}$	$3.5 \cdot 10^{-6}$	$3 \cdot 2^{15+\frac{5}{2}}$	$1.9 \cdot 10^{-3}$	$3.5 \cdot 10^{-6}$
64	2^{12}	2^{27}	$4.4 \cdot 10^{-4}$	$1.2 \cdot 10^{-6}$	$3 \cdot 2^{21}$	$4.4 \cdot 10^{-4}$	$1.2 \cdot 10^{-6}$

		Linear Implicit Euler			Exponential Euler		
N	M	CC	Error	Std	CC	Error	Std
2	2^3	$2^{4+\frac{1}{2}}$	$1.8 \cdot 10^{-2}$	$1.7 \cdot 10^{-3}$	$2^{4+\frac{1}{2}}$	$1.9 \cdot 10^{-2}$	$1.7 \cdot 10^{-3}$
4	2^6	2^9	$2.6 \cdot 10^{-2}$	$3.8 \cdot 10^{-4}$	2^9	$2.6 \cdot 10^{-2}$	$4.5 \cdot 10^{-4}$
8	2^9	$2^{12+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$8.2 \cdot 10^{-4}$	$2^{12+\frac{3}{2}}$	$1.7 \cdot 10^{-2}$	$1.0 \cdot 10^{-4}$
16	2^{12}	2^{18}	$6.4 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$	2^{18}	$6.6 \cdot 10^{-3}$	$1.8 \cdot 10^{-5}$
32	2^{15}	$2^{20+\frac{5}{2}}$	$1.9 \cdot 10^{-3}$	$4.9 \cdot 10^{-6}$	$2^{20+\frac{5}{2}}$	$2.0 \cdot 10^{-3}$	$5.1 \cdot 10^{-6}$
64	2^{18}	2^{27}	$4.2 \cdot 10^{-4}$	$2.6 \cdot 10^{-7}$	2^{27}	$4.9 \cdot 10^{-4}$	$2.2 \cdot 10^{-6}$

Table 3.5: Error and standard deviation for Example 3 obtained from 500 paths. CC denotes the computational cost needed to evaluate the term that dominates the computational effort according to Table 3.1 for all time steps M.

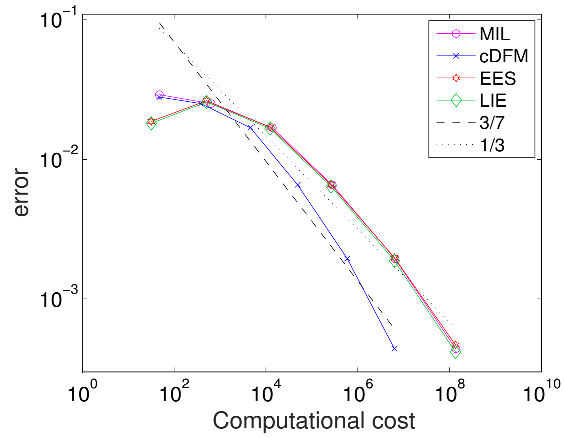

Figure 3.5: Error against computational cost for Example 3 for 500 paths and $N = 2, 4, 8, 16, 32, 64$ in log-log scale.

The various examples that we investigated confirm the theory developed and described in Section 3.2 to Section 3.5. For equations that are pointwise multiplicative, the cDFMM and the Milstein scheme converge with the same order. For general SPDEs of type (1.1), however, we observe the predicted increase in the effective order of convergence for the commutative derivative-free Milstein scheme.

3.7 Some Notes on Implementation

As there are various spaces that have to be approximated or discretized in the numerical schemes that we introduced in the last sections, we outline their implementation for clarification. Let us fix $M, N, K \in \mathbb{N}$; additionally to the projections P_N, P_K, described in the derivation of the algorithm above, we need to discretize the underlying space. If $H = L^2((0,1), \mathbb{R})$, for example, we specify a grid on the interval $(0,1)$ by the grid points $x_j = \frac{j}{N_x+1}$, $j \in \{1, \ldots, N_x\}$, $N_x \in \mathbb{N}$.

We define the numerical method by some operator $\Phi : H_N \times V_K \times \mathbb{R} \to H$ as

$$Y_{m+1}^{K,M} = e^{Ah} Y_m^{N,K,M} + \Phi(Y_m^{N,K,M}, \Delta W_m^{K,M}, h)$$

for all $m \in \{0, \ldots, M-1\}$. For the exponential Euler scheme, for example, we have

$$\Phi(Y_m^{N,K,M}, \Delta W_m^{K,M}, h) = A^{-1}(e^{Ah} - I) F(Y_m^{N,K,M}) + e^{Ah} B(Y_m^{N,K,M}) \Delta W_m^{K,M}$$

for all $m \in \{0, \ldots, M-1\}$.

Now, let an initial condition $X_0 = \sum_{i \in \mathcal{I}} c_i e_i$ with $c_i \in \mathbb{R}$ for all $i \in \mathcal{I}$ be given, that is, we know the coefficients $c_i = \langle X_0, e_i \rangle_H$, for all $i \in \mathcal{I}$. Furthermore, assume that we have computed an approximation $Y_m^{K,M}$ of the solution at time $t_m = m \cdot h$ for some $m \in \{0, \ldots, M-1\}$ and have calculated the corresponding Fourier coefficients $\langle Y_m^{K,M}, e_i \rangle_H$ for all $i \in \mathcal{I}_N$.

For all $m \in \{1, \dots, M\}$, $M, N, K \in \mathbb{N}$, the solution in the next step is then obtained as

1. Compute the projection of $Y_m^{K,M}$ at grid points x_j, $j \in \{1, \dots, N_x\}$, as

$$Y_m^{N,K,M}(x_j) = P_N Y_m^{K,M}(x_j) = \sum_{i \in \mathcal{I}_N} \langle Y_m^{K,M}, e_i \rangle_H e_i(x_j).$$

2. Draw K random variables $\epsilon_k \sim N(0,1)$, $k \in \{1, \dots, K\}$, and project the Q-Wiener process as

$$\Delta W_m^{K,M}(x_j) = \sum_{k \in \mathcal{J}_K} \sqrt{\eta_k} \sqrt{h} \epsilon_k \tilde{e}_k(x_j)$$

at all grid points x_j, $j \in \{1, \dots, N_x\}$.

3. Compute the solution in the next time step at all grid points x_j, $j \in \{1, \dots, N_x\}$, with the numerical scheme

$$Y_{m+1}^{K,M}(x_j) = e^{Ah} Y_m^{N,K,M}(x_j) + \Phi(Y_m^{N,K,M}(x_j), \Delta W_m^{K,M}(x_j), h).$$

4. Calculate the Fourier coefficients $\langle Y_{m+1}^{K,M}, e_i \rangle_H$, $i \in \mathcal{I}_N$. If again, for example, $H = L^2((0,1))$, we employ a quadrature rule such as

$$\langle Y_{m+1}^{K,M}, e_i \rangle_H \approx \frac{1}{N_x + 1} \sum_{j=1}^{N_x} e_i(x_j) Y_{m+1}^{K,M}(x_j)$$

for all $i \in \mathcal{I}_N$.

This sequence is repeated until we reach $\langle Y_M^{K,M}, e_i \rangle_H$, $i \in \mathcal{I}_N$; then, we conduct step 1. once more to obtain $Y_M^{N,K,M}$.

In the error analysis, we are interested in the term $\big(\mathrm{E}\big[\|X_T - Y_M^{N,M,K}\|_H^2\big]\big)^{\frac{1}{2}}$ for all $M, N, K \in \mathbb{N}$. We stay with the example from above and fix $H = L^2((0,1), \mathbb{R})$. Denote by $P \in \mathbb{N}$ the number of paths that we simulate and by Y_M^p and X_T^p the discrete approximation, obtained as described above, and the solution process at time T computed in simulation p, $p \in \{1, \dots, P\}$. Then, we approximate the error as

$$\left(\mathrm{E}\left[\int_0^1 \left|X_T(x) - Y_M^{N,M,K}(x)\right|^2 \mathrm{d}x\right]\right)^{\frac{1}{2}} \approx \left(\frac{1}{P} \frac{1}{(N_x+1)} \sum_{p=1}^{P} \sum_{j=1}^{N_x} |X_T^p(x_j) - Y_M^p(x_j)|^2\right)^{\frac{1}{2}}.$$

This is just a basic composite trapezoidal rule, [17, Chapter 9]. We refrain from analyzing or improving the approximation at this point by choosing, for example, another quadrature rule as this part is not specific for SPDEs and therefore not the focus of our work. In the numerical analyses in Section 3.6 and Section 4.4, we chose $N_x = N$. These examples show that the approximation of the inner product with the composite trapezoidal rule does not reduce the order of convergence below the expected value.

4

Non-Commutative SPDEs: Numerical Schemes with Higher Orders

If the commutativity condition (C5) is not fulfilled for a SPDE of type (1.1), it is not possible to rewrite the iterated stochastic integrals

$$\int_s^t B'(X_s)\Big(\int_s^r B(X_s)\,\mathrm{d}W_u^K\Big)\,\mathrm{d}W_r^K, \tag{4.1}$$

$s,t \in [0,T]$, $s \leq t$, $K \in \mathbb{N}$, as in equation (3.11); thus, the commutative derivative-free Milstein scheme is not applicable. Therefore, we need to analyze and approximate these iterated integrals to obtain a numerical scheme for such equations. If not stated differently, we assume the setting presented in Section 3.1 throughout this chapter.

Below, we introduce two algorithms to approximate stochastic integrals of type (4.1). In a second step, these methods are incorporated in a derivative-free Milstein scheme to approximate equation (1.1). We analyze the strong convergence for both combinations and compare the methods analytically and numerically to the exponential Euler scheme.

4.1 Approximation of Iterated Stochastic Integrals

The schemes that we propose to approximate the stochastic double integrals derive from the methods developed by Kloeden, Platen, and Wright in [39] and Wiktorsson in [75] for finite dimensional SODEs; they are based on the truncation of a series representation of the Brownian bridge process. We transfer these schemes to our setting and adjust the proofs. For Algorithm 1, based on [39], we obtain an estimate for the sum of squared errors instead of the maximal error only, whereas the error estimate for Algorithm 2, adapted from [75], differs in the setting of SPDEs driven by a Q-Wiener process of trace class.

In order to devise algorithms to approximate the iterated stochastic integral (4.1) in the setting of SPDEs, we define the operators $B \in L(V,H)_0$, $G \in L(H, L(V,H)_0)$. Furthermore, we assume $(W_t)_{t\in[0,T]}$ to be a Q-Wiener process with finite trace. These assumptions are imposed on SPDE (1.1) to derive the numerical schemes in Sections 3.3 and 4.2 as well and are therefore not limiting.

In the following, we denote the iterated stochastic integrals by

$$I_{(i,j)}(h) := I_{(i,j)}(t,t+h) = \int_t^{t+h} \int_t^s \mathrm{d}\beta_r^i \,\mathrm{d}\beta_s^j, \tag{4.2}$$

for all $i,j \in \mathcal{J}$, $h > 0$, $t, t+h \in [0,T]$.
As $(W_t)_{t\in[0,T]}$ is an infinite dimensional stochastic process, represented by an infinite sum (2.5), we truncate this series such that it can be simulated with a numerical scheme. We denote this approximation by $(W_t^K)_{t\in[0,T]}$, $K \in \mathbb{N}$, which is defined in (3.10).
Thereby, we can express the iterated stochastic integral for any $h > 0$, $t, t+h \in [0,T]$, $K \in \mathbb{N}$, as

$$\begin{aligned}
\int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big) \,\mathrm{d}W_s^K &= \int_t^{t+h} G\Big(\int_t^s B \sum_{i\in\mathcal{J}_K} \sqrt{\eta_i}\tilde{e}_i \,\mathrm{d}\beta_r^i\Big) \sum_{j\in\mathcal{J}_K} \sqrt{\eta_j}\tilde{e}_j \,\mathrm{d}\beta_s^j \\
&= \int_t^{t+h} \int_t^s \sum_{i\in\mathcal{J}_K} \sqrt{\eta_i} \,\mathrm{d}\beta_r^i \sum_{j\in\mathcal{J}_K} \sqrt{\eta_j} \,\mathrm{d}\beta_s^j \; G\big(B\tilde{e}_i, \tilde{e}_j\big) \\
&= \sum_{i\in\mathcal{J}_K} \sum_{j\in\mathcal{J}_K} \sqrt{\eta_i}\sqrt{\eta_j} I_{(i,j)}(h) \; G\big(B\tilde{e}_i, \tilde{e}_j\big).
\end{aligned}$$

For simplicity of notation, we assume $\mathcal{J}_K = \{1,\ldots,K\}$ and $\eta_1 > \eta_2 > \ldots > \eta_K$, $K \in \mathbb{N}$, in this chapter. Consequently, we aim at devising a scheme to approximate

$$I_{(i,j)}^Q(h) := \sqrt{\eta_i}\sqrt{\eta_j} \int_t^{t+h} \int_t^s \mathrm{d}\beta_r^i \,\mathrm{d}\beta_s^j \tag{4.3}$$

for all $i,j \in \{1,\ldots,K\}$, $K \in \mathbb{N}$, $h > 0$, $t, t+h \in [0,T]$.

Algorithm 1
In order to define Algorithm 1, we mainly adjust the scheme of [39] to our setting. Moreover, we obtain an estimate for the sum of squared errors, similar to (4.4) below, which remains independent of the number of Brownian motions driving the equation.

In the setting of SODEs, we denote the approximation of the stochastic double integral (4.2) with the algorithm derived in [39] by $\tilde{I}_{(i,j)}(h)$ for all $i,j \in \{1,\ldots,K\}$, $K \in \mathbb{N}$, $h > 0$; here, the integer $K \in \mathbb{N}$ represents the number of independent Brownian motions $(\beta_t^j)_{t\geq 0}$, $j \in \{1,\ldots,K\}$, in the SODE that is to be approximated. In [38] and [39], the authors showed that for any $h > 0$ the simulation error can be estimated as

$$\max_{i,j\in\{1,\dots,K\}} \mathrm{E}\big[|\tilde{I}_{i,j}(h) - I_{i,j}(h)|^2\big] \leq C\frac{h^2}{D}, \tag{4.4}$$

where $D \in \mathbb{N}$ is the integer at which the series representation of $I_{i,j}(h)$, $i,j \in \{1,\dots,K\}$, $K \in \mathbb{N}$, $h > 0$, is truncated to obtain the approximation.

The following derivation is taken from [39] to a large extent. We derive an approximation of the Brownian bridge process $(\beta_s^j - \frac{s}{t}\beta_t^j)_{0\leq s\leq t}$ for all $j \in \{1,\dots,K\}$, $K \in \mathbb{N}$, $t \in [0,T]$ by its Fourier series first. Therewith, we obtain an approximation of $I_{(i,j)}^Q(t,t+h)$ for all $i,j \in \{1,\dots,K\}$, $K \in \mathbb{N}$, $h > 0$, $t, t+h \in [0,T]$, depending on realizations of the increments of the Brownian motion $\Delta\beta_h^j = \beta_{t+h}^j - \beta_t^j$, $j \in \{1,\dots,K\}$, $K \in \mathbb{N}$, $h > 0$.
The series expansion, which converges in $L^2(\Omega)$, reads

$$\beta_s^j - \frac{s}{t}\beta_t^j = \frac{1}{2}a_0^j + \sum_{r=1}^{\infty}\Big(a_r^j \cos\Big(\frac{2r\pi s}{t}\Big) + b_r^j \sin\Big(\frac{2r\pi s}{t}\Big)\Big) \tag{4.5}$$

with

$$\begin{aligned} a_r^j &= \frac{2}{t}\int_0^t (\beta_u^j - \frac{u}{t}\beta_t^j)\cos\Big(\frac{2r\pi u}{t}\Big)\,\mathrm{d}u, \\ b_r^j &= \frac{2}{t}\int_0^t (\beta_u^j - \frac{u}{t}\beta_t^j)\sin\Big(\frac{2r\pi u}{t}\Big)\,\mathrm{d}u \end{aligned}$$

for all $j \in \{1,\dots,K\}$, $K \in \mathbb{N}$, $r \in \mathbb{N}_0$, and all $0 \leq s \leq t \leq T$.
By simple computations, we obtain

$$\mathrm{E}\big[a_r^j\big] = \mathrm{E}\big[b_r^j\big] = 0, \tag{4.6}$$

$$\mathrm{E}\big[a_r^i a_r^j\big] = \mathrm{E}\big[b_r^i b_r^j\big] = \mathrm{E}\big[a_r^i b_k^i\big] = \mathrm{E}\big[a_r^i b_r^i\big] = 0, \tag{4.7}$$

for all $i,j \in \{1,\dots,K\}$, $i \neq j$, $K \in \mathbb{N}$, $r,k \in \mathbb{N}$, and

$$\mathrm{E}\big[a_r^j a_k^j\big] = \frac{1}{r^2\pi^2}\mathrm{E}\Big[\int_0^t \cos\Big(\frac{2r\pi s}{t}\Big)\cos\Big(\frac{2k\pi s}{t}\Big)\,\mathrm{d}s\Big] = \begin{cases} 0, & k \neq r \\ \frac{t}{2\pi^2 r^2}, & k = r, \end{cases} \tag{4.8}$$

$$\mathrm{E}\big[b_r^j b_k^j\big] = \begin{cases} 0, & k \neq r \\ \frac{t}{2\pi^2 r^2}, & k = r, \end{cases} \tag{4.9}$$

for all $j \in \{1,\dots,K\}$, $K \in \mathbb{N}$ $r,k \in \mathbb{N}$.
We rearrange expression (4.5) and truncate the series at some index $D \in \mathbb{N}$. This yields the following approximation of the Brownian motions $(\beta_s^j)_{0\leq s\leq t}$ for all $j \in \{1,\dots,K\}$ and all $s,t \in [0,T]$, $s \leq t$

$$\beta_s^j(D) = \frac{s}{t}\beta_t^j + \frac{1}{2}a_0^j + \sum_{r=1}^{D}\Big(a_r^j \cos\Big(\frac{2r\pi s}{t}\Big) + b_r^j \sin\Big(\frac{2r\pi s}{t}\Big)\Big). \tag{4.10}$$

In fact, we are interested in computing an integral with respect to this process; it converges to the Stratonovich integral $J(h)$ for $D \to \infty$ according to Wong and Zakai [76, 77] and [38, Chapter 6.1]. Therefore, we obtain an approximation $\tilde{J}^Q_{(i,j)}(h)$, $i, j \in \{1, \dots, K\}$, $K \in \mathbb{N}$, $h > 0$ for the stochastic double integral in the Stratonovich sense from which we obtain the Itô integral $\tilde{I}^Q_{(i,j)}(h)$, $i, j \in \{1, \dots, K\}$, $h > 0$, as

$$I^Q_{(i,j)}(h) = J^Q_{(i,j)}(h) - \frac{1}{2} h \eta_i \mathbb{1}_{i=j},$$

see [38, p.174].
From [38, p.171], we know

$$I^Q_{(i,i)}(h) = \frac{\eta_i \big(\Delta\beta^i_h\big)^2 - \eta_i h}{2}$$

for all $i \in \{1, \dots, K\}$, $k \in \mathbb{N}$, $h > 0$; therefore, we have to approximate $\tilde{J}^Q_{(i,j)}(h)$ or $\tilde{I}^Q_{(i,j)}(h)$, respectively, for $i, j \in \{1, \dots, K\}$, $i \neq j$, $K \in \mathbb{N}$, only. As $\tilde{I}^Q_{(i,j)}(h) = \tilde{J}^Q_{(i,j)}(h)$ for all $i, j \in \{1, \dots, K\}$, $i \neq j$, $K \in \mathbb{N}$, $h > 0$, we obtain an approximation of the Itô integral directly by integrating with respect to the process given by (4.10).

Since $\int_0^t f(u)\, \mathrm{d}\beta^j_u = f(t)\beta^j_t - \int_0^t f'(u)\beta^j_u \, \mathrm{d}u$ for a continuously differentiable function $f : [0, t] \to \mathbb{R}$, see [38, p.89], and $a^j_0 = \frac{2}{t} \int_0^t \beta^j_u \, \mathrm{d}u - \beta^j_t$ for all $j \in \{1, \dots, K\}$, $t \in [0, T]$, the iterated stochastic integral can be computed as

$$\begin{aligned}
I^Q_{(i,j)}(h) =& I^Q_{(i,j)}(0,h) = \sqrt{\eta_i}\sqrt{\eta_j} \int_0^h \int_0^u \mathrm{d}\beta^i_r \, \mathrm{d}\beta^j_u = \sqrt{\eta_i}\sqrt{\eta_j} \int_0^h \beta^i_u \, \mathrm{d}\beta^j_u \\
=& \sqrt{\eta_i}\sqrt{\eta_j} \int_0^h \Big(\frac{u}{h}\beta^i_h + \frac{1}{2} a^i_0 + \sum_{r=1}^{\infty} \Big(a^i_r \cos\Big(\frac{2r\pi u}{h}\Big) + b^i_r \sin\Big(\frac{2r\pi u}{h}\Big)\Big)\Big) \, \mathrm{d}\beta^j_u \\
=& \sqrt{\eta_i}\sqrt{\eta_j} \Big(\frac{\beta^i_h}{h} \int_0^h u \, \mathrm{d}\beta^j_u + \frac{1}{2} a^i_0 \beta^j_h + \int_0^h \Big(\sum_{r=1}^{\infty} a^i_r \cos\Big(\frac{2r\pi u}{h}\Big) + b^i_r \sin\Big(\frac{2r\pi u}{h}\Big)\Big) \, \mathrm{d}\beta^j_u \Big) \\
=& \sqrt{\eta_i}\sqrt{\eta_j} \Big(\frac{\beta^i_h}{h} \int_0^h u \, \mathrm{d}\beta^j_u + \frac{1}{2} a^i_0 \beta^j_h \\
&+ \sum_{r=1}^{\infty} \Big(a^i_r \Big(\beta^j_h + \int_0^h \frac{2r\pi}{h} \sin\Big(\frac{2r\pi u}{h}\Big) \beta^j_u \, \mathrm{d}u \Big) - b^i_r \int_0^h \frac{2r\pi}{h} \cos\Big(\frac{2r\pi u}{h}\Big) \beta^j_u \, \mathrm{d}u \Big)\Big) \\
=& \sqrt{\eta_i}\sqrt{\eta_j} \Big(\frac{1}{2} \beta^i_h \beta^j_h - \frac{1}{2} (a^j_0 \beta^i_h - a^i_0 \beta^j_h) + \sum_{r=1}^{\infty} \Big(a^i_r \Big(\beta^j_h + r\pi\big(b^j_r - \frac{\beta^j_h}{r\pi}\big)\Big) - \frac{2r\pi}{h} b^i_r \frac{h}{2} a^j_r \Big)\Big) \\
=& \sqrt{\eta_i}\sqrt{\eta_j} \Big(\frac{1}{2} \beta^i_h \beta^j_h - \frac{1}{2} (a^j_0 \beta^i_h - a^i_0 \beta^j_h) + \pi \sum_{r=1}^{\infty} r (a^i_r b^j_r - b^i_r a^j_r) \Big)
\end{aligned}$$

for all $i, j \in \{1, \dots, K\}$, $i \neq j$, $K \in \mathbb{N}$, $h > 0$.
Following [39] and [75], we do not approximate the stochastic double integral but the Lévy stochastic area integrals

$$A^Q_{(i,j)}(h) = \sqrt{\eta_i}\sqrt{\eta_j} \Big(-\frac{1}{2} (a^j_0 \beta^i_h - a^i_0 \beta^j_h) + h \frac{\pi}{h} \sum_{r=1}^{\infty} r (a^i_r b^j_r - b^i_r a^j_r) \Big) \tag{4.11}$$

for all $i, j \in \{1, \ldots, K\},\ i \neq j,\ h > 0$, instead. Furthermore, we only need to simulate $A^Q_{(i,j)}(h)$ for $i, j \in \{1, \ldots, K\}$ with $i < j$. This is sufficient due to the following relations, see [75],

$$
\begin{aligned}
I^Q_{(i,j)}(h) &= \frac{\sqrt{\eta_i}\sqrt{\eta_j}\Delta\beta^i_h\ \Delta\beta^j_h - h\eta_i\delta_{ij}}{2} + A^Q_{(i,j)}(h) \quad P\text{-a.s.} && (4.12)\\
A^Q_{(j,i)}(h) &= -A^Q_{(i,j)}(h) && (4.13)\\
A^Q_{(i,i)}(h) &= 0 && (4.14)
\end{aligned}
$$

for all $i, j \in \{1, \ldots, K\}$, $h > 0$.
In the following, we use the descriptive notation from [75], which follows directly from (4.11) and (4.6)-(4.9) and was introduced in [39], to simulate

$$
A^Q_{(i,j)}(h) = \frac{h}{2\pi}\sum_{r=1}^{\infty}\frac{1}{r}\Big(U_{ri}\Big(Z_{rj} + \sqrt{\frac{2}{h}}\sqrt{\eta_j}\Delta\beta^j_h\Big) - U_{rj}\Big(Z_{ri} + \sqrt{\frac{2}{h}}\sqrt{\eta_i}\Delta\beta^i_h\Big)\Big) \tag{4.15}
$$

with random variables $U_{ri} \sim N(0, \eta_i)$, $Z_{ri} \sim N(0, \eta_i)$, and $\Delta\beta^i_h \sim N(0, h)$ that are all independent for $i, j \in \{1, \ldots, K\}$, $r \in \mathbb{N}$, $h > 0$.

We denote

$$
a^r_{(i,j)}(h) := U_{ri}\Big(Z_{rj} + \sqrt{\frac{2}{h}}\sqrt{\eta_j}\Delta\beta^j_h\Big) - U_{rj}\Big(Z_{ri} + \sqrt{\frac{2}{h}}\sqrt{\eta_i}\Delta\beta^i_h\Big) \tag{4.16}
$$

for all $i, j \in \{1, \ldots, K\}$, $r \in \mathbb{N}$, $h > 0$ and define the index set

$$
\begin{aligned}
\mathcal{I}_A &= ((1,2), \ldots, (1,K), \ldots, (l, l+1), \ldots, (l, K), \ldots, (K-1, K))\\
&= (I_1, \ldots, I_L) && (4.17)
\end{aligned}
$$

for $L = \frac{K(K-1)}{2}$. As mentioned above, we only need to approximate $A^Q_I(h)$, $h > 0$ for $I \in \mathcal{I}_A$. We truncate the series in (4.15) at some index $D \in \mathbb{N}$ to obtain the approximation. For $I_k \in \mathcal{I}_A$ with $I_k = (i, j)$, $i, j \in \{1, \ldots, K\}$, $K \in \mathbb{N}$, $k \in \{1, \ldots, L\}$, we define

$$
a^{(D)}_{I_k}(h) := \sum_{r=1}^{D}\frac{1}{r}\Big(U_{ri}\Big(Z_{rj} + \sqrt{\frac{2}{h}}\sqrt{\eta_j}\Delta\beta^j_h\Big) - U_{rj}\Big(Z_{ri} + \sqrt{\frac{2}{h}}\sqrt{\eta_i}\Delta\beta^i_h\Big)\Big) = \sum_{r=1}^{D}\frac{1}{r}a^r_{I_k}(h)
$$

for all $h > 0$ and get an approximation of the vector of area integrals as

$$
\begin{aligned}
\tilde{A}(h) &:= \frac{h}{2\pi}\Big(a^{(D)}_{(1,2)}(h), \ldots, a^{(D)}_{(1,K)}(h), \ldots, a^{(D)}_{(l,l+1)}(h), \ldots, a^{(D)}_{(l,K)}(h), \ldots, a^{(D)}_{(K-1,K)}(h)\Big)^T\\
&= \frac{h}{2\pi}\Big(\big(a^{(D)}_I(h)\big)_{I\in\mathcal{I}_A}\Big)^T. && (4.18)
\end{aligned}
$$

For all $h > 0$, we define the vector of truncation errors as

$$
R_D(h) := \frac{h}{2\pi}\Big(\Big(\sum_{r=D+1}^{\infty}\frac{1}{r}a^r_I(h)\Big)_{I\in\mathcal{I}_A}\Big)^T. \tag{4.19}
$$

Now, we formulate Algorithm 1. The notation in the following instruction is taken from [75] to some extent, as the scheme can be efficiently implemented this way. For some $h > 0$, $D, K \in \mathbb{N}$

1. Simulate
$$\Delta W_h^Q = \left(\sqrt{\eta_1}\Delta\beta_h^1, \ldots, \sqrt{\eta_K}\Delta\beta_h^K\right)^T \sim N(0_K, h\mathcal{C})$$
with $\mathcal{C} = \mathrm{diag}(\eta_1, \ldots, \eta_K)$

2. Approximate $\tilde{A}(h)$ as
$$\begin{aligned}\tilde{A}(h) &= \frac{h}{2\pi}\left(\left(a_I^{(D)}(h)\right)_{I\in\mathcal{I}_A}\right)^T \\ &= \frac{h}{2\pi}\sum_{r=1}^{D}\frac{1}{r}H_K\left(U_r \otimes \left(Z_r + \sqrt{\frac{2}{h}}\Delta W_h^Q\right) - \left(Z_r + \sqrt{\frac{2}{h}}\Delta W_h^Q\right) \otimes U_r\right)\end{aligned}$$
where $U_k = (U_{k1}, \ldots, U_{kK})^T \sim N(0_K, \mathcal{C})$, $Z_k = (Z_{k1}, \ldots, Z_{kK})^T \sim N(0_K, \mathcal{C})$ for all $k \in \{1, \ldots, D\}$ and
$$H_K = \begin{pmatrix} 0_{K-1\times 1} & I_{K-1} & 0_{K-1\times K(K-1)} \\ 0_{K-2\times K+2} & I_{K-2} & 0_{K-2\times K(K-2)} \\ \vdots & \vdots & \vdots \\ 0_{K-l\times(l-1)K+l} & I_{K-l} & 0_{K-l\times K(K-l)} \\ \vdots & \vdots & \vdots \\ 0_{1\times(K-2)K+K-1} & 1 & 0_{1\times K} \end{pmatrix}$$

3. Define $\mathbf{A}^Q = 0_{K\times K}$ and set $\mathbf{A}_{I_k}^Q = \tilde{A}(h)_k$ for $I_k \in \mathcal{I}_A$, $k \in \{1, \ldots, L\}$. Compute
$$\tilde{I}^Q = (\tilde{I}_{(i,j)}^Q)_{i,j\in\{1,\ldots,K\}} = \mathbf{A}^Q - \left(\mathbf{A}^Q\right)^T + \frac{\Delta W_h^Q(\Delta W_h^Q)^T}{2} - \frac{h}{2}\mathcal{C}$$

Now, we determine the truncation error for this approximation method.

Theorem 4.1 (Convergence of Algorithm 1)
Assume that Q is a trace class operator and $(W_t)_{t\in[0,T]}$ a Q-Wiener process. Furthermore, let $B \in L(V, H)_0$ and $G \in L(H, L(V, H)_0)$. Then, it holds
$$\mathrm{E}\left[\left\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K}\tilde{I}_{(i,j)}^Q(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big)\right\|_H^2\right] \leq C_Q\frac{h^2}{\pi^2 D}$$
for some $C_Q > 0$ and some arbitrary $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$.

Proof. The proof is given in Section 4.3. □

Note, that this estimate is independent of K.
For completeness, we state the following lemma. We want to emphasize that in the numerical scheme we compare the approximation of the iterated stochastic integral to the integral with

respect to $(W_t^K)_{t\in[0,T]}$, $K \in \mathbb{N}$, that is, we employ Theorem 4.1. The following estimate is not part of the proof of convergence of the derivative-free Milstein scheme in Section 4.3 and does, therefore, not imply a reduction in the order of convergence.

Lemma 4.1
Let $B \in L(V,H)_0$, $G \in L(H, L(V,H)_0)$, and let $(W_t)_{t\in[0,T]}$ be a Q-Wiener process of trace class. Then, it holds

$$\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r\Big)\,\mathrm{d}W_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \leq C_Q h^2 + C_Q \frac{h^2}{\pi^2 D}$$

for some $C_Q > 0$ and some arbitrary $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$.

Proof. For a proof we refer to Section 4.3. □

Algorithm 2
First, we consider SODEs again and denote the approximation of the iterated stochastic integral (4.2) with the scheme in [75] by $\hat{I}_{(i,j)}(h)$ for all $i,j \in \{1,\ldots,K\}$, $K \in \mathbb{N}$, $h > 0$. In the finite dimensional case, the error resulting from the approximation of $I_{(i,j)}(h)$ by $\hat{I}_{(i,j)}(h)$ for all $i,j \in \{1,\ldots,K\}$, $K \in \mathbb{N}$, $h > 0$ is

$$\sum_{\substack{i,j=1\\ i<j}}^{K} \mathrm{E}\Big[\big|I_{(i,j)}(h) - \hat{I}_{(i,j)}(h)\big|^2\Big] \leq \frac{5h^2}{24\pi^2 D^2} K^2(K-1), \tag{4.20}$$

where $D \in \mathbb{N}$ is again the index at which the sum is truncated such that we obtain the approximation of the iterated stochastic integral and K is the number of independent Brownian motions, see [75] for details and a proof.
In the numerical approximation of SODEs, D is chosen such that the overall order of convergence is not distorted; for example, $D \geq \frac{\sqrt{5K^2(K-1)}}{\sqrt{24\pi^2 h}}$ is selected in the Milstein scheme, [50, 75]. Therefore, the simulation of the iterated integrals is costly; precisely, we have to compute DM terms to obtain $\hat{I}_{(i,j)}(h)$, $h = \frac{T}{M}$, for all $i,j \in \{1,\ldots,K\}$, $K \in \mathbb{N}$, and all time steps, where $M \in \mathbb{N}$ is the number of time steps used to approximate the solution of the SODE. This adds computational cost of order $M^{\frac{3}{2}}$ to the total effort, consult [38] or [75] for more details on this issue.
The error estimate (4.20) depends on the number of Brownian motions $K \in \mathbb{N}$ as well. This suggests that the computational cost increases much faster in the setting of SPDEs as the number of independent Brownian motions is, in general, not fixed. The eigenvalues of the Q-Wiener process are, however, not incorporated in the error estimate yet. The fact that we integrate with respect to a Q-Wiener process in the setting of SPDEs, where Q as a trace class operator, yields an error estimate which depends on the rate of decay of the eigenvalues η_j, $j \in \mathcal{J}$, of Q.

The following derivation is based on [75]. As before, the series (4.18) is truncated at some index $D \in \mathbb{N}$. The tail sum R_D, however, is approximated by a multivariate normal distributed random vector additionally, as described in [75].

Now, we determine the distribution of the tail sum; we compute the covariance matrix of $\big(a_I^r(h)\big)_{I\in\mathcal{I}_A}$ conditional on $Z^r = (Z_{r1},\dots,Z_{rK})$ and $\Delta W_h^Q = (\sqrt{\eta_1}\Delta\beta_h^1,\dots,\sqrt{\eta_K}\Delta\beta_h^K)^T$ for each $r\in\mathbb{N}$, $h>0$, $K\in\mathbb{N}$

$$\Sigma^r_{|Z^r,\Delta W_h^Q} := \begin{pmatrix} \mathrm{Var}(a^r_{I_1}) & \mathrm{Cov}(a^r_{I_1}a^r_{I_2}) & \dots & \mathrm{Cov}(a^r_{I_1}a^r_{I_L}) \\ \mathrm{Cov}(a^r_{I_1}a^r_{I_2}) & \mathrm{Var}(a^r_{I_2}) & & \\ \vdots & & \ddots & \mathrm{Cov}(a^r_{I_{L-1}}a^r_{I_L}) \\ \mathrm{Cov}(a^r_{I_1}a^r_{I_L}) & & & \mathrm{Var}(a^r_{I_L}) \end{pmatrix} \tag{4.21}$$

with

$$\mathrm{Var}(a^r_{I_k}) = \eta_i\Big(Z_{rj} + \sqrt{\eta_j}\sqrt{\frac{2}{h}}\Delta\beta_h^j\Big)^2 + \eta_j\Big(Z_{ri} + \sqrt{\eta_i}\sqrt{\frac{2}{h}}\Delta\beta_h^i\Big)^2$$

for all $I_k = (i,j)\in\mathcal{I}_A$, $k\in\{1,\dots,L\}$, $i,j\in\{1,\dots,K\}$, $K\in\mathbb{N}$, $r\in\mathbb{N}$.
We define $d_i^r := Z_{ri} + \sqrt{\eta_i}\sqrt{\frac{2}{h}}\Delta\beta_h^i$, $i\in\{1,\dots,K\}$, $K\in\mathbb{N}$, $r\in\mathbb{N}$ and get for all $I_k=(i,j)\in\mathcal{I}_A$, $I_l=(m,n)\in\mathcal{I}_A$, $k,l\in\{1,\dots,L\}$, $l\neq k$, $i,j,m,n\in\{1,\dots,K\}$, $K\in\mathbb{N}$,

$$\begin{aligned}\mathrm{Cov}(a^r_{I_k}a^r_{I_l}) &= \mathrm{E}\Big[\big(U_{ri}(Z_{rj} + \frac{2}{h}\sqrt{\eta_j}\Delta\beta_h^j) - (Z_{ri} + \frac{2}{h}\sqrt{\eta_i}\Delta\beta_h^i)U_{rj}\big) \\ &\qquad \cdot \big(U_{rn}(Z_{rm} + \frac{2}{h}\sqrt{\eta_m}\Delta\beta_h^m) - (Z_{rn} + \frac{2}{h}\sqrt{\eta_n}\Delta\beta_h^n)U_{rm}\big)\Big|Z^r,\Delta W_h^Q\Big] \\ &= \begin{cases} 0, & i\neq n,m,\ j\neq n,m \\ \eta_i d_j^r d_m^r, & i=n \\ -\eta_i d_j^r d_n^r, & i=m \\ -\eta_j d_i^r d_m^r, & j=n \\ \eta_j d_i^r d_n^r, & j=m. \end{cases}\end{aligned}$$

This implies

$$R_D(h)_{|Z,\Delta W_h^Q} \sim N\Big(0_L, \Big(\frac{h}{2\pi}\Big)^2 \sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)$$

for $Z=(Z^r)_{r\in\mathbb{N}}$ and $D\in\mathbb{N}$. Hence, we can approximate the tail sum by simulating a random vector

$$V_D(h) = \frac{2\pi}{h}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)^{-\frac{1}{2}} R_D(h)$$

with $V_D(h)_{|Z,\Delta W_h^Q} \sim N(0_L, I_{L\times L})$ and

$$R_D(h) = \frac{h}{2\pi}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)^{\frac{1}{2}} V_D(h) \tag{4.22}$$

for arbitrary $D\in\mathbb{N}$.
It remains to examine, how the covariance matrix evolves with $D\to\infty$. For some $D\in\mathbb{N}$, we

define

$$\Sigma^{(D+)} := \Big(\sum_{r=D+1}^{\infty} \frac{1}{r^2} \Big)^{-1} \sum_{r=D+1}^{\infty} \frac{1}{r^2} \Sigma^r_{|Z^r, \Delta W_h^Q} \tag{4.23}$$

and

$$\Sigma^{(\infty)} := \mathrm{E}\Big[\Sigma^1_{|Z^1, \Delta W_h^Q} \Big| \Delta W_h^Q \Big] \tag{4.24}$$

with diagonal elements

$$\Sigma^{(\infty)}_{kk} = 2\eta_i\eta_j + \frac{2}{h}\eta_i\eta_j(\Delta\beta_h^j)^2 + \frac{2}{h}\eta_i\eta_j(\Delta\beta_h^i)^2$$

and

$$\Sigma^{(\infty)}_{kl} = \begin{cases} 0, & i \neq n, m, \ j \neq n, m \\ \frac{2}{h}\eta_i\sqrt{\eta_j}\sqrt{\eta_m}\Delta\beta_h^j\Delta\beta_h^m, & i = n, \\ -\frac{2}{h}\eta_i\sqrt{\eta_j}\sqrt{\eta_n}\Delta\beta_h^j\Delta\beta_h^n, & i = m, \\ -\frac{2}{h}\eta_j\sqrt{\eta_i}\sqrt{\eta_m}\Delta\beta_h^i\Delta\beta_h^m, & j = n, \\ \frac{2}{h}\eta_j\sqrt{\eta_i}\sqrt{\eta_n}\Delta\beta_h^i\Delta\beta_h^n, & j = m, \end{cases}$$

for all $k, l \in \{1, \ldots, L\}$, $l \neq k$ with $I_k = (i, j)$, $I_l = (m, n)$, $I_k, I_l \in \mathcal{I}_A$, $i, j, m, n \in \{1, \ldots, K\}$, $K \in \mathbb{N}$.
From the proof of Theorem 4.2 below, we have

$$\lim_{D\to\infty} \mathrm{E}\big[\|\Sigma^{(D+)} - \Sigma^{(\infty)}\|_F^2 \big] = 0,$$

where F denotes the Frobenius norm. Thus it follows

$$\frac{2\pi}{h} \Big(\sum_{r=D+1}^{\infty} \frac{1}{r^2} \Big)^{-\frac{1}{2}} R_D(h) \xrightarrow{d} \zeta \sim N\big(0_L, \Sigma^{(\infty)}\big)$$

for $D \longrightarrow \infty$.

Combining the above, we obtain an algorithm very similar to the one in [75], where steps $1, 2$, and 4 equal Algorithm 1. For some $h > 0$, $D, K \in \mathbb{N}$

1. Simulate
$$\Delta W_h^Q = \big(\sqrt{\eta_1}\Delta\beta_h^1, \ldots, \sqrt{\eta_K}\Delta\beta_h^K\big)^T \sim N(0_K, h\mathcal{C})$$
with $\mathcal{C} = \mathrm{diag}(\eta_1, \ldots, \eta_K)$

2. Approximate $\tilde{A}(h)$ as

$$\begin{aligned} \tilde{A}(h) &= \frac{h}{2\pi}\Big(\Big(a_I^{(D)}(h) \Big)_{I \in \mathcal{I}_A} \Big)^T \\ &= \frac{h}{2\pi} \sum_{r=1}^{D} \frac{1}{r} H_K \Big(U_r \otimes \Big(Z_r + \sqrt{\frac{2}{h}}\Delta W_h^Q \Big) - \Big(Z_r + \sqrt{\frac{2}{h}}\Delta W_h^Q \Big) \otimes U_r \Big) \end{aligned}$$

with $U_k \sim N(0_K, \mathcal{C})$, $Z_k \sim N(0_K, \mathcal{C})$, $k \in \{1, \ldots, D\}$ and

$$H_K = \begin{pmatrix} 0_{K-1\times 1} & I_{K-1} & 0_{K-1\times K(K-1)} \\ 0_{K-2\times K+2} & I_{K-2} & 0_{K-2\times K(K-2)} \\ \vdots & \vdots & \vdots \\ 0_{K-l\times (l-1)K+l} & I_{K-l} & 0_{K-l\times K(K-l)} \\ \vdots & \vdots & \vdots \\ 0_{1\times (K-2)K+K-1} & 1 & 0_{1\times K} \end{pmatrix}$$

3. Simulate $V_D(h) \sim N(0_L, I_{L\times L})$ and compute

$$\hat{A}(h) = \tilde{A}(h) + \frac{h}{2\pi}\Big(\sum_{r=D+1}^{\infty} \frac{1}{r^2}\Big)^{\frac{1}{2}} \sqrt{\Sigma^{(\infty)}} V_D(h) \tag{4.25}$$

4. Define $\mathbf{A}^Q = 0_{K\times K}$ and set $\mathbf{A}^Q_{I_k} = \hat{A}(h)_k$ for $I_k \in \mathcal{I}_A$, $k \in \{1, \ldots, L\}$. Compute

$$\hat{I}^Q = (\hat{I}^Q_{(i,j)})_{i,j\in\{1,\ldots,K\}} = \mathbf{A}^Q - \big(\mathbf{A}^Q\big)^T + \frac{\Delta W_h^Q (\Delta W_h^Q)^T}{2} - \frac{h}{2}\mathcal{C}$$

Note, that the matrix $\sqrt{\Sigma^{(\infty)}}$ has to be computed by a Cholesky decomposition. Next, we analyze the error resulting from this algorithm.

Theorem 4.2 (Convergence of Algorithm 2)
Assume that Q is a trace class operator and $(W_t)_{t\in[0,T]}$ a Q-Wiener process. Furthermore, let $B \in L(V,H)_0$ and $G \in L(H, L(V,H)_0)$. Then, it holds

$$\mathrm{E}\bigg[\Big\| \int_t^{t+h} G\Big(\int_t^s B \, \mathrm{d}W_r^K\Big) \, \mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}^Q_{(i,j)}(h) \; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\bigg] \leq C_Q \frac{h^2}{D^2} \eta_K^{-1}$$

for some $C_Q > 0$ and all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$.

Proof. Again, the proof is detailed in Section 4.3 □

Remark 4.1
Note, that if $(W_t)_{t\in[0,T]}$ is a cylindrical Wiener process, we obtain the same estimate (4.20) as in the finite dimensional case.

In general, we obtain convergence for $K, M \to \infty$ if we choose $h = \frac{T}{M}$ and $D > \eta_K^{-\frac{1}{2}} h^{1-\theta}$ for some $\theta > 0$. For Algorithm 1, we require $D > h^{2-\theta}$, instead. However, we need a more careful choice of D to maintain the rate of convergence of the numerical scheme developed in the next section. This issue is detailed in the following.

4.2 Error Analysis and Computational Cost

Consider a numerical scheme with strong order of convergence $q > 0$ in the time step $h = \frac{T}{M}$, $T > 0$, $M \in \mathbb{N}$ to approximate the mild solution to equation (1.1). We aim at simulating the iterated stochastic integrals such that the order of convergence is not reduced. Therefore, we need to choose $D \geq M^{2q-1}$ for Algorithm 1, whereas Algorithm 2 requires $D \geq M^{q-\frac{1}{2}}\eta_K^{-\frac{1}{2}}$, see Theorem 4.1 and Theorem 4.2. The error estimates for the iterated stochastic integrals in the last section show that it depends on the relation of $N, M, K \in \mathbb{N}$ and therewith the parameters $\gamma, \alpha, \rho_A, \rho_Q$ which of the two approximation algorithms is superior in the infinite dimensional setting - Table 4.1 summarizes these results. This contrasts the approximation of SODEs in finite dimensions, where the number of independent Brownian motions is fixed and therefore the algorithm developed in [75] involves less computational effort.

In the following, we develop a derivative-free Milstein scheme for the case that the operator $B'B$ is not commutative, that is, (C5) does not hold. We show that the scheme that we propose below is more efficient for most parameter constellations than the exponential Euler scheme if $q_{EES} < q_{DFM}$, which holds in general.
In contrast to the commutative equation, we approximate the derivative by one term only, compare Section 3.3. The discrete approximation process that is obtained with this scheme is denoted by $(Y_m^Q)_{0\leq m\leq M}$, $M \in \mathbb{N}$. The derivative-free Milstein scheme (DFM) reads $Y_0^Q = P_N\xi$ and

$$\begin{aligned} Y_{m+1}^Q =& P_N\bigg(e^{Ah}Y_m^Q + he^{Ah}F(Y_m^Q) + e^{Ah}B(Y_m^Q)\Delta W_m^{K,M} \\ &+ e^{Ah}\sum_{j\in\mathcal{J}_K}\Big(B\Big(Y_m^Q + \sum_{i\in\mathcal{J}_K} P_N B(Y_m^Q)\tilde{e}_i \bar{I}_{(i,j)}^Q(h)\Big) - B\big(Y_m^Q\big)\Big)\tilde{e}_j\bigg) \end{aligned} \tag{4.26}$$

for all $m \in \{0,\ldots,M-1\}$, $N, M, K \in \mathbb{N}$. Here, $\bar{I}_{i,j}^Q(h)$ denotes the approximation of $I_{i,j}^Q(h)$ obtained with Algorithm 1 or Algorithm 2 for all $i, j \in \mathcal{J}_K$, $h > 0$, $K \in \mathbb{N}$. If the iterated integral is approximated by Algorithm 1, we call the scheme DFM1, whereas if we employ Algorithm 2, it is denoted as DFM2.
In order to prove the error estimate below, we need to impose one of the following assumptions additionally.

(C6a) $Q^{\frac{1}{2}}$ is a trace class operator.

(C6b) Q is a trace class operator and $\sup_{y\in H_\beta}\|B(y)\|_{L(V,H)} < C$ for some $C > 0$.

Therewith, we can show the convergence of the derivative-free Milstein scheme.

Theorem 4.3 (Convergence of DFM)
Let assumptions (C1)–(C4) and (C6a) or (C6b) be fulfilled. Then, there exists a constant $C_{T,Q} \in (0,\infty)$, independent of N, K, M, and D, such that for $(Y_m^Q)_{0\leq m\leq M}$, defined by the DFM in

(4.26), *it holds*

$$\Big(\mathrm{E}\Big[\big\|X_{t_m} - Y_m^Q\big\|_H^2\Big]\Big)^{\frac{1}{2}} \leq C_{T,Q}\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N}\lambda_i\Big)^{-\gamma} + \Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K}\eta_j\Big)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)} + \mathcal{E}(D)\Big)$$

for all $m \in \{0,1,\ldots,M\}$ and all $N,K,M,D \in \mathbb{N}$. The error term $\mathcal{E}(D)$ is determined by Algorithm 1 or 2 and given in Theorem 4.1 or Theorem 4.2, respectively. The parameters are determined by (C1)–(C4) and (C6).

Proof. The proof of Theorem 4.3 is stated in Section 4.3. □

This is the same estimate as for the commutative derivative-free Milstein scheme, see Theorem 3.1, if $D \in \mathbb{N}$ is chosen carefully. The computational cost, however, increases as we have to simulate the iterated integrals.

Next, we determine the effective order of convergence for these schemes. Below, we assume $\eta_j \leq Cj^{-\rho_Q}$ for some $\rho_Q, C > 0$ and all $j \in \mathcal{J}_K$, $K \in \mathbb{N}$. As for the commutative scheme, we minimize the error (3.8) such that the computational cost does not exceed some specified value $\bar{c} > 0$. For Algorithm 1, the condition on the computational cost reads $3MNK + M^{2q}\frac{K(K-1)}{2} = \bar{c}$ with $q = \min(2(\gamma-\beta),\gamma)$. We obtain different results depending on the parameters; as a reasonable choice for N, M, K, we get

$$M = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\rho_Q}{2q(\alpha\rho_Q+1)}}\Big), \quad N = \mathcal{O}\Big(\bar{c}^{\frac{\alpha\rho_Q}{2\gamma\rho_A(\alpha\rho_Q+1)}}\Big), \quad K = \mathcal{O}\Big(\bar{c}^{\frac{1}{2(\alpha\rho_Q+1)}}\Big)$$

if $2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q > 0$ and $q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$. In this case, we obtain for the effective order of convergence

$$\mathrm{err(DFM1)} = \mathcal{O}\Big(\bar{c}^{-\frac{\alpha\rho_Q}{2(\alpha\rho_Q+1)}}\Big). \tag{4.27}$$

For any other parameter set, we obtain as for the commutative scheme

$$M = \mathcal{O}\Big(\bar{c}^{\frac{\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}}\Big), \quad N = \mathcal{O}\Big(\bar{c}^{\frac{q\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}}\Big), \quad K = \mathcal{O}\Big(\bar{c}^{\frac{q\gamma\rho_A}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}}\Big)$$

and the effective order of convergence equals

$$\mathrm{err(DFM1)} = \mathcal{O}\Big(\bar{c}^{-\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}}\Big). \tag{4.28}$$

Concerning Algorithm 2, the condition on the computational cost is $3MNK + M^{q+\frac{1}{2}}\eta_K^{-\frac{1}{2}}\frac{K(K-1)}{2} = \bar{c}$. If $\sqrt{\Sigma^{(\infty)}}$ is not explicitly computable but obtained by a Cholesky decomposition instead, we get an additional term in the computational cost. We want to keep this analysis independent of this factor, however, and do not include the effort to calculate $\sqrt{\Sigma^{(\infty)}}$ in this analysis.
We get different results subject to the parameter constellation. Precisely, if $2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0$ and $q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$, we obtain

$$M = \mathcal{O}\Big(\bar{c}^{\frac{2\alpha\rho_Q}{(2\alpha\rho_Q+\rho_Q+4)q+\alpha\rho_Q}}\Big), \quad N = \mathcal{O}\Big(\bar{c}^{\frac{2\alpha\rho_Q q}{\gamma\rho_A((2\alpha\rho_Q+\rho_Q+4)q+\alpha\rho_Q)}}\Big), \quad K = \mathcal{O}\Big(\bar{c}^{\frac{2q}{(2\alpha\rho_Q+\rho_Q+4)q+\alpha\rho_Q}}\Big)$$

with effective order of convergence

$$\mathrm{err}(\mathrm{DFM2}) = \mathcal{O}\Big(\bar{c}^{-\frac{2\alpha\rho_Q q}{(2\alpha\rho_Q+\rho_Q+4)q+\alpha\rho_Q}}\Big). \tag{4.29}$$

Any other parameter constellation yields the same effective order as the commutative scheme, see (4.28).

We compare the effective order of convergence across the parameter sets to determine which scheme to use for a given equation. The benchmark scheme is the exponential Euler scheme whose effective order of convergence is

$$\mathrm{err}(\mathrm{EES}) = \mathcal{O}\Big(\bar{c}^{-\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}}\Big),$$

as described in Section 3.4. Here the parameter q equals $q = q_{EES} = \min(\gamma, \frac{1}{2}, 2(\gamma-\beta))$. We are mainly interested in treating cases where $q_{EES} < q_{DFM}$, which we assume from now on. Table 4.1 clearly illustrates how the decision on the preferred scheme depends on the underlying parameters $q, \alpha, \gamma, \rho_A, \rho_Q$ and that it cannot be universally identified. The EES is outperformed by the DFM for most parameter constellations, however. We merely compare the effective order of convergence for the case that $2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q > 0$ and $q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$ or $2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0$ and $q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$ does not hold. In this case, the superiority of the DFM1 or DFM2 follows as in Section 3.4 for the commutative derivative-free Milstein scheme. The results for other parameter constellations are summarized in Table 4.1.

We do want to emphasize again the contrast to the approximation of finite dimensional SODEs with numerical schemes that involve an approximation of stochastic double integral; for SPDEs, we do not have a preference for one of the schemes DFM1 or DFM2 independently of the equation to be solved. The overview in Table 4.1 clearly illustrates the dependence on the parameters $q, \alpha, \gamma, \rho_A, \rho_Q$. Note, for completeness, that for the Exponential Euler scheme, assumptions (C6a) and (C6b) as well as parts of (C3) are not required to hold. Therefore, for those equations that do not fulfill these conditions, this scheme might be beneficial.

Conditions			Scheme	EOC
$2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) \leq 0, \quad 2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q \leq 0$				
$q \geq \alpha(2q-1)$			DFM1*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$q < \alpha(2q-1)$			DFM2*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) \leq 0, \quad 2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q > 0$				
$q \leq \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$	$q \geq \alpha(2q-1)$		DFM1*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
	$q < \alpha(2q-1)$		DFM2*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$			DFM2	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0, \quad 2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q \leq 0$				
$q \leq \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$	$q \geq \alpha(2q-1)$		DFM1*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
	$q < \alpha(2q-1)$		DFM2*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$			DFM1	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0,\ 2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q > 0,\ \gamma\rho_A > \rho_Q(\alpha - \gamma\rho_A)$				
$q \leq \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$	$q \geq \alpha(2q-1)$		DFM1*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
	$q < \alpha(2q-1)$		DFM2*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$\frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)} < q \leq \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$			DFM1	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$	$q \geq \alpha(2q-1)$	$\alpha\rho_Q > \gamma\rho_A$	DFM1	$\frac{\alpha\rho_Q}{2(\alpha\rho_Q+1)}$
		$\alpha\rho_Q \leq \gamma\rho_A$	EES	$\frac{\alpha\rho_Q\gamma\rho_A q_{EES}}{(\alpha\rho_Q+\gamma\rho_A)q_{EES}+\alpha\rho_Q\gamma\rho_A}$
	$q < \alpha(2q-1)$	$q > \frac{\alpha\rho_Q\gamma\rho_A}{d_1}$	DFM2	$\frac{2\alpha\rho_Q q}{(2\alpha\rho_Q+\rho_Q+4)q+\alpha\rho_Q}$
		$q \leq \frac{\alpha\rho_Q\gamma\rho_A}{d_1}$	EES	$\frac{\alpha\rho_Q\gamma\rho_A q_{EES}}{(\alpha\rho_Q+\gamma\rho_A)q_{EES}+\alpha\rho_Q\gamma\rho_A}$
$2\alpha\rho_Q\gamma\rho_A + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0,\ 2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q > 0,\ \gamma\rho_A \leq \rho_Q(\alpha - \gamma\rho_A)$				
$q \leq \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q}$	$q \geq \alpha(2q-1)$		DFM1*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
	$q < \alpha(2q-1)$		DFM2*	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$\frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\gamma\rho_A-\alpha\rho_Q} < q \leq \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$			DFM2	$\frac{q\gamma\rho_A\alpha\rho_Q}{(\alpha\rho_Q+\gamma\rho_A)q+\gamma\rho_A\alpha\rho_Q}$
$q > \frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A+\rho_Q\gamma\rho_A+2(\gamma\rho_A-\alpha\rho_Q)}$	$q \geq \alpha(2q-1)$	$\alpha\rho_Q > \gamma\rho_A$	DFM1	$\frac{\alpha\rho_Q}{2(\alpha\rho_Q+1)}$
		$\alpha\rho_Q \leq \gamma\rho_A$	EES	$\frac{\alpha\rho_Q\gamma\rho_A q_{EES}}{(\alpha\rho_Q+\gamma\rho_A)q_{EES}+\alpha\rho_Q\gamma\rho_A}$
	$q < \alpha(2q-1)$	$q > \frac{\alpha\rho_Q\gamma\rho_A}{d_1}$	DFM2	$\frac{2\alpha\rho_Q q}{(2\alpha\rho_Q+\rho_Q+4)q+\alpha\rho_Q}$
		$q \leq \frac{\alpha\rho_Q\gamma\rho_A}{d_1}$	EES	$\frac{\alpha\rho_Q\gamma\rho_A q_{EES}}{(\alpha\rho_Q+\gamma\rho_A)q_{EES}+\alpha\rho_Q\gamma\rho_A}$

Table 4.1: This overview identifies the numerical scheme which is selected to approximate a given non-commutative SPDE of type (1.1). The scheme that is indicated is chosen by its effective order of convergence (EOC). For some parameter sets, there is no difference in the order of convergence; in this case, the computational cost is consulted as an additional factor (indicated by $*$). We define $d_1 = 2\alpha\rho_Q\gamma\rho_A - \rho_Q\gamma\rho_A - 2(\gamma\rho_A - \alpha\rho_Q)$ and assume $q_{EES} < q = q_{DFM}$.

4.3 Proofs

In the following proofs we use some generic constant C which may change from line to line.

Proof of Theorem 4.1

Theorem (Convergence of Algorithm 1)
Assume that Q is a trace class operator and $(W_t)_{t\in[0,T]}$ a Q-Wiener process. Furthermore, let $B \in L(V,H)_0$ and $G \in L(H, L(V,H)_0)$. Then, it holds

$$\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \leq C_Q \frac{h^2}{\pi^2 D}$$

for some $C_Q > 0$ and some arbitrary $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$.

Proof of Theorem 4.1.
For all $h > 0$, $t, t+h \in [0,T]$, $K \in \mathbb{N}$, we get

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&= \mathrm{E}\Big[\Big\| \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} I^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big) - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&= \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \mathrm{E}\big[(I^Q_{(i,j)}(h) - \tilde{I}^Q_{(i,j)}(h))^2\big] \|G\big(B\tilde{e}_i, \tilde{e}_j\big)\|_H^2
\end{aligned}$$

as $\mathrm{E}\big[I^Q_{(i,j)}(h) I^Q_{(m,n)}(h)\big] = 0$ for all $i,j,m,n \in \mathcal{J}_K$ with $(i,j) \neq (m,n)$, $K \in \mathbb{N}$, see [38]. The same holds for the approximations $\tilde{I}^Q_{(i,j)}(h)$, $i,j \in \mathcal{J}_K$, $K \in \mathbb{N}$, $h > 0$.
By the assumptions on B and G, we obtain

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \mathrm{E}\big[(I^Q_{(i,j)}(h) - \tilde{I}^Q_{(i,j)}(h))^2\big] \|G\|^2_{L(H,L(V,H))} \|B\|^2_{L(V,H)} \\
&\leq C \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \mathrm{E}\big[(I^Q_{(i,j)}(h) - \tilde{I}^Q_{(i,j)}(h))^2\big]
\end{aligned}$$

for all $h > 0$, $t, t+h \in [0,T]$, $K \in \mathbb{N}$.
Due to (4.12)-(4.14), it is enough to examine $A^Q_{(i,j)}(h)$, $\tilde{A}^Q_{(i,j)}(h)$, $h > 0$, for $i,j \in \mathcal{J}_K$, $K \in \mathbb{N}$, with $i < j$. This implies

$$\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big]$$

$$\leq C \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \mathrm{E}\big[(A^Q_{(i,j)}(h)-\tilde{A}^Q_{(i,j)}(h))^2\big]$$
$$= 2C \sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}} \mathrm{E}\big[(A^Q_{(i,j)}(h)-\tilde{A}^Q_{(i,j)}(h))^2\big] \tag{4.30}$$

for all $h>0$, $t,t+h\in[0,T]$, $K\in\mathbb{N}$.
By (4.11) and the properties of a^j_r, b^j_r, $r\in\mathbb{N}$, $j\in\mathcal{J}_K$, $K\in\mathbb{N}$, given in (4.6)-(4.9), we obtain

$$\mathrm{E}\Big[\Big\|\int_t^{t+h} G\Big(\int_t^s B\,\mathrm{d}W^K_r\Big)\,\mathrm{d}W^K_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big]$$
$$\leq 2Ch^2 \sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}} \mathrm{E}\Big[\Big(\frac{\pi}{h}\sum_{r=D+1}^{\infty} r\sqrt{\eta_i}\sqrt{\eta_j}(a^i_r b^j_r - b^i_r a^j_r)\Big)^2\Big]$$
$$= 2Ch^2\frac{\pi^2}{h^2}\sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}}\sum_{r=D+1}^{\infty} r^2\eta_i\eta_j\mathrm{E}\big[(a^i_r b^j_r)^2 + (b^i_r a^j_r)^2\big]$$
$$= 4Ch^2\frac{\pi^2}{h^2}\sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}}\sum_{r=D+1}^{\infty} r^2\eta_i\eta_j\Big(\frac{h}{2\pi^2r^2}\Big)^2$$
$$\leq C\,(\operatorname{tr} Q)^2\sum_{r=D+1}^{\infty}\frac{h^2}{\pi^2r^2}$$

for all $h>0$, $t,t+h\in[0,T]$, $D,K\in\mathbb{N}$.
As in [39], we finally estimate

$$\sum_{r=D+1}^{\infty}\frac{1}{r^2}\leq\int_D^{\infty}\frac{1}{s^2}\,\mathrm{d}s=\frac{1}{D}$$

and in total, we obtain

$$\mathrm{E}\Big[\Big\|\int_t^{t+h} G\Big(\int_t^s B\,\mathrm{d}W^K_r\Big)\,\mathrm{d}W^K_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big]$$
$$\leq C\,(\operatorname{tr} Q)^2\frac{h^2}{D\pi^2} \tag{4.31}$$

for all $h>0$, $t,t+h\in[0,T]$, $D,K\in\mathbb{N}$. □

Proof of Lemma 4.1

Lemma

Let $B\in L(V,H)_0$, $G\in L(H,L(V,H)_0)$, and let $(W_t)_{t\in[0,T]}$ be a Q-Wiener process of trace class. Then, it holds

$$\mathrm{E}\Big[\Big\|\int_t^{t+h} G\Big(\int_t^s B\,\mathrm{d}W_r\Big)\,\mathrm{d}W_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big] \leq C_Q h^2 + C_Q\frac{h^2}{\pi^2 D}$$

for some $C_Q > 0$ and some arbitrary $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$.

Proof of Lemma 4.1.
We determine the error resulting from the projection of the Q-Wiener process in (3.10) combined with the approximation of the iterated integrals. For all $h > 0$, $t, t+h \in [0,T]$, $K \in \mathbb{N}$, we decompose the error such that

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r\Big)\,\mathrm{d}W_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq \mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r\Big)\,\mathrm{d}W_s - \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big\|_H^2\Big] \\
&\quad + \mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big]. \qquad (4.32)
\end{aligned}$$

The second term can be estimated as stated in the proof of Theorem 4.1. By Itô's isometry and the properties of the operators B and G, we get for all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r\Big)\,\mathrm{d}W_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq 2\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r\Big)\,\mathrm{d}W_s\Big\|_H^2\Big] + 2\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big\|_H^2\Big] + \operatorname{tr} Q \frac{h^2}{\pi^2 D} \\
&\leq 2\int_t^{t+h} \mathrm{E}\Big[\Big\| G\int_t^s B \,\mathrm{d}W_r\Big\|_{L_{HS}(V_0,H)}^2\Big]\,\mathrm{d}s + 2\sum_{j\in\mathcal{J}_K} \eta_j \int_t^{t+h} \mathrm{E}\Big[\Big\| G\Big(\int_t^s B \,\mathrm{d}W_r^K, \tilde{e}_j\Big)\Big\|_H^2\Big]\,\mathrm{d}s \\
&\quad + \operatorname{tr} Q \frac{h^2}{\pi^2 D} \\
&\leq 2C\int_t^{t+h}\int_t^s \mathrm{E}\big[\|B\|_{L_{HS}(V_0,H)}^2\big]\,\mathrm{d}r\,\mathrm{d}s + 2C\sum_{i,j\in\mathcal{J}_K} \eta_i\eta_j \int_t^{t+h}\int_t^s \mathrm{E}\big[\|B\tilde{e}_i\|_H^2\big]\,\mathrm{d}r\,\mathrm{d}s + \operatorname{tr} Q \frac{h^2}{\pi^2 D}.
\end{aligned}$$

By the assumptions on B, we further obtain

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r\Big)\,\mathrm{d}W_s - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \tilde{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq Ch^2 + C(\operatorname{tr} Q)^2 h^2 + \operatorname{tr} Q \frac{h^2}{\pi^2 D} \leq C_Q h^2 + \operatorname{tr} Q \frac{h^2}{\pi^2 D}.
\end{aligned}$$

for all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$. □

Proof of Theorem 4.2

Theorem (Convergence of Algorithm 2)
Assume that Q is a trace class operator and $(W_t)_{t\in[0,T]}$ a Q-Wiener process. Furthermore, let $B \in L(V,H)_0$ and $G \in L(H, L(V,H)_0)$. Then, it holds

$$\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(\int_t^s B \,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \leq C_Q \frac{h^2}{D^2}\eta_K^{-1}$$

for some $C_Q > 0$ and all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$.

Proof of Theorem 4.2.
As in the proof of Theorem 4.1, particularly equation (4.30), we obtain

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(B\int_t^s \mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big] \\
&\le C\sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \mathrm{E}\big[(I^Q_{(i,j)}(h)-\hat{I}^Q_{(i,j)}(h))^2\big] = 2C\sum_{I=1}^{L}\mathrm{E}\big[(A_I^Q(h)-\hat{A}_I^Q(h))^2\big]
\end{aligned}$$

for all $h > 0$, $t, t+h \in [0,T]$, $K \in \mathbb{N}$.

Let $\|\cdot\|_F$ denote the Frobenius norm; with the expressions for $R_D(h)$ in (4.22), $\Sigma^{(D+)}$ in (4.23), and the definition of the algorithm (4.25), we get

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(B\int_t^s \mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}^Q_{(i,j)}(h)\; G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big] \\
&\le 2C\sum_{I=1}^{L}\mathrm{E}\Big[\Big(\Big(R_D(h)-\frac{h}{2\pi}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{\frac12}\sqrt{\Sigma^{(\infty)}}V_D(h)\Big)_I\Big)^2\Big] \\
&= 2C\sum_{I=1}^{L}\mathrm{E}\Big[\Big(\Big(\frac{h}{2\pi}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)^{\frac12}V_D(h)-\frac{h}{2\pi}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{\frac12}\sqrt{\Sigma^{(\infty)}}V_D(h)\Big)_I\Big)^2\Big] \\
&= \frac{Ch^2}{2\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\sum_{I=1}^{L}\mathrm{E}\Big[\Big(\Big(\Big(\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-\frac12}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)^{\frac12}-\sqrt{\Sigma^{(\infty)}}\Big)V_D(h)\Big)_I\Big)^2\Big] \\
&= C\frac{h^2}{2\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\sum_{I=1}^{L}\mathrm{E}\Big[\Big(\Big(\sqrt{\Sigma^{(D+)}}-\sqrt{\Sigma^{(\infty)}}\Big)V_D(h))_I\Big)^2\Big] \\
&= C\frac{h^2}{2\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\sum_{I=1}^{L}\mathrm{E}\Big[\mathrm{E}\Big[\Big(\Big(\Big(\sqrt{\Sigma^{(D+)}}-\sqrt{\Sigma^{(\infty)}}\Big)V_D(h)\Big)_I\Big)^2\Big|Z,\Delta W_h^Q\Big]\Big] \\
&= C\frac{h^2}{2\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\sum_{i_1,i_2=1}^{L}\mathrm{E}\Big[\Big(\Big(\sqrt{\Sigma^{(D+)}}-\sqrt{\Sigma^{(\infty)}}\Big)_{(i_1,i_2)}\Big)^2\Big] \\
&= C\frac{h^2}{2\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\mathrm{E}\Big[\big\|\sqrt{\Sigma^{(D+)}}-\sqrt{\Sigma^{(\infty)}}\big\|_F^2\Big]
\end{aligned}$$

for all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$. Here, we used that $V_D(h)_{|Z,\Delta W_h^Q} \sim N(0_L, I_{L\times L})$ for all $h > 0$, $D, L \in \mathbb{N}$.

Next, we employ the following lemma from [75].

Lemma 4.2

Let A and C be symmetric positive definite matrices and denote the smallest eigenvalue of matrix C by λ_{min}. Then, it holds

$$\|A^{\frac12}-C^{\frac12}\|_F^2 \le \frac{1}{\sqrt{\lambda_{min}}}\|A-C\|_F^2.$$

Proof of Lemma 4.2. A proof can be found in [75, Lemma 4.1]. □

As stated above, we assume $\eta_1 > \eta_2 > \ldots > \eta_K$ for all $K \in \mathbb{N}$. We decompose $\Sigma^{(\infty)}$ for some $K \in \mathbb{N}$ as

$$\Sigma^{(\infty)} = 2\eta_{K-1}\eta_K I_{L\times L} + \hat{\Sigma}^{(\infty)}$$

to determine its smallest eigenvalue.
The matrix $\hat{\Sigma}^{(\infty)}$ is defined as follows; for the diagonal elements, we get

$$\hat{\Sigma}^{(\infty)}_{kk} = \Sigma^{(\infty)}_{kk} - 2\eta_{K-1}\eta_K = 2(\eta_i\eta_j - \eta_{K-1}\eta_K) + \frac{2}{h}\eta_i\eta_j(\Delta\beta^j_h)^2 + \frac{2}{h}\eta_i\eta_j(\Delta\beta^i_h)^2$$

for all $k \in \{1,\ldots,L\}$ with $I_k = (i,j)$, $i,j \in \mathcal{J}_K$, $I_k \in \mathcal{I}_A$, $K \in \mathbb{N}$, and $h > 0$.
For entries other than the diagonal elements, we get $\hat{\Sigma}^{(\infty)}_{kl} = \Sigma^{(\infty)}_{kl}$ for all $I_k = (i,j)$, $I_l = (m,n)$, $i,j,m,n \in \mathcal{J}_K$, $i_k, i_l \in \mathcal{I}_A$, $k,l \in \{1,\ldots,L\}$, $K \in \mathbb{N}$.
As the matrix $\hat{\Sigma}^{(\infty)}$ is positive semi-definite, the smallest eigenvalue $\lambda_{\min}$ of $\Sigma^{(\infty)}$ fulfills $\lambda_{\min} \geq 2\eta_K^2$ for all $K \in \mathbb{N}$.
The covariance matrices $\Sigma^{(D+)}$ and $\Sigma^{(\infty)}$ are symmetric positive definite and we get by Lemma 4.2 and the definitions of $\Sigma^{(D+)}$, $\Sigma^{(\infty)}$ in (4.23) and (4.24), respectively,

$$\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(B\int_t^s \mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq C\frac{h^2}{2\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\mathrm{E}\Big[\big\|\sqrt{\Sigma^{(D+)}} - \sqrt{\Sigma^{(\infty)}}\big\|_F^2\Big] \\
&\leq C\frac{h^2}{2\sqrt{2}\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\mathrm{E}\big[\|\Sigma^{(D+)} - \Sigma^{(\infty)}\|_F^2\big] \\
&= C\frac{h^2}{2\sqrt{2}\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)\mathrm{E}\Big[\Big\|\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-1}\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q} - \mathrm{E}\Big[\Sigma^1_{|Z^1,\Delta W_h^Q}\Big|\Delta W_h^Q\Big]\Big\|_F^2\Big] \\
&= C\frac{h^2}{\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big) \\
&\qquad\cdot\mathrm{E}\Big[\Big\|\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-1}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q} - \sum_{r=D+1}^{\infty}\frac{1}{r^2}\mathrm{E}\Big[\Sigma^1_{|Z^1,\Delta W_h^Q}\Big|\Delta W_h^Q\Big]\Big)\Big\|_F^2\Big] \\
&= C\frac{h^2}{\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-1} \\
&\quad\cdot\sum_{i_1,i_2=1}^{L}\mathrm{E}\Big[\mathrm{E}\Big[\Big(\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q} - \sum_{r=D+1}^{\infty}\frac{1}{r^2}\mathrm{E}\Big[\Sigma^r_{|Z^r,\Delta W_h^Q}\Big|\Delta W_h^Q\Big]\Big)_{(i_1,i_2)}\Big)^2\Big|\Delta W_h^Q\Big]\Big]
\end{aligned}$$

for all $h > 0$, $t, t+h \in [0,T]$, $D,K,L \in \mathbb{N}$.
Moreover, we obtain for all $h > 0$, $t, t+h \in [0,T]$, $D,K,L \in \mathbb{N}$

$$\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(B\int_t^s \mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}^Q_{(i,j)}(h)\, G\big(B\tilde{e}_i,\tilde{e}_j\big)\Big\|_H^2\Big]$$

$$
\begin{aligned}
&\le C\frac{h^2}{\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-1}\sum_{i_1,i_2=1}^{L}\mathrm{E}\Big[\mathrm{Var}\Big(\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)_{(i_1,i_2)}\Big|\Delta W_h^Q\Big)\Big]\\
&= C\frac{h^2}{\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-1}\sum_{i_1,i_2=1}^{L}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\mathrm{E}\Big[\mathrm{Var}\Big(\Big(\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)_{(i_1,i_2)}\Big|\Delta W_h^Q\Big)\Big].
\end{aligned}
$$

Next, we compute the expectation; we insert the expressions for $\Sigma^r_{|Z^r,\Delta W_h^Q}$, $r\in\mathbb{N}$, and $\Sigma^{(\infty)}$, see (4.21) and (4.24), and split the sum into diagonal entries and the remaining part of the matrix. This yields for all $h>0$, $t,t+h\in[0,T]$, $D,K,L\in\mathbb{N}$

$$
\begin{aligned}
&\sum_{i_1,i_2=1}^{L}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\mathrm{E}\Big[\mathrm{Var}\Big(\Big(\Sigma^r_{|Z^r,\Delta W_h^Q}\Big)_{(i_1,i_2)}\Big|\Delta W_h^Q\Big)\Big]\\
&=\sum_{i_1=1}^{L}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\mathrm{E}\Big[\mathrm{E}\Big[\Big(\Sigma^r_{|Z^r,\Delta W_h^Q}-\Sigma^{(\infty)}\Big)^2_{(i_1,i_1)}\Big|\Delta W_h^Q\Big]\Big]\\
&\quad+\sum_{\substack{i_1,i_2=1\\ i_1\neq i_2}}^{L}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\mathrm{E}\Big[\mathrm{E}\Big[\Big(\Sigma^r_{|Z^r,\Delta W_h^Q}-\Sigma^{(\infty)}\Big)^2_{(i_1,i_2)}\Big|\Delta W_h^Q\Big]\Big]\\
&=\sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\Bigg(\mathrm{E}\Big[\mathrm{E}\Big[\Big(\eta_i\Big(Z_{rj}^2+2Z_{rj}\sqrt{\eta_j}\sqrt{\frac{2}{h}}\Delta\beta_h^j\Big)\\
&\qquad\qquad+\eta_j\Big(Z_{ri}^2+2Z_{ri}\sqrt{\eta_i}\sqrt{\frac{2}{h}}\Delta\beta_h^i\Big)-2\eta_i\eta_j\Big)^2\Big|\Delta W_h^Q\Big]\Big]\Bigg)\\
&\quad+\sum_{\substack{i,j,m,n\in\mathcal{J}_K\\ i<j,m<n}}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\Bigg(\mathrm{E}\Big[\mathrm{E}\Big[\eta_i^2\Big(Z_{rj}Z_{rm}+Z_{rj}\sqrt{\eta_m}\sqrt{\frac{2}{h}}\Delta\beta_h^m+Z_{rm}\sqrt{\eta_j}\sqrt{\frac{2}{h}}\Delta\beta_h^j\Big)^2\mathbb{1}_{i=n}\mathbb{1}_{j\neq m}\\
&\qquad\qquad+\eta_i^2\Big(Z_{rj}Z_{rn}+Z_{rj}\sqrt{\eta_n}\sqrt{\frac{2}{h}}\Delta\beta_h^n+Z_{rn}\sqrt{\eta_j}\sqrt{\frac{2}{h}}\Delta\beta_h^j\Big)^2\mathbb{1}_{i=m}\mathbb{1}_{j\neq n}\\
&\qquad\qquad+\eta_j^2\Big(Z_{ri}Z_{rm}+Z_{ri}\sqrt{\eta_m}\sqrt{\frac{2}{h}}\Delta\beta_h^m+Z_{rm}\sqrt{\eta_i}\sqrt{\frac{2}{h}}\Delta\beta_h^i\Big)^2\mathbb{1}_{j=n}\mathbb{1}_{i\neq m}\\
&\qquad\qquad+\eta_j^2\Big(Z_{ri}Z_{rn}+Z_{ri}\sqrt{\eta_n}\sqrt{\frac{2}{h}}\Delta\beta_h^n+Z_{rn}\sqrt{\eta_i}\sqrt{\frac{2}{h}}\Delta\beta_h^i\Big)^2\mathbb{1}_{j=m}\mathbb{1}_{i\neq n}\Big|\Delta W_h^Q\Big]\Big]\Bigg).
\end{aligned}
\tag{4.33}
$$

We compute the terms in (4.33) separately and obtain

$$
\begin{aligned}
&\mathrm{E}\Big[\mathrm{E}\Big[\Big(\eta_i\Big(Z_{rj}^2+2Z_{rj}\sqrt{\eta_j}\sqrt{\frac{2}{h}}\Delta\beta_h^j\Big)+\eta_j\Big(Z_{ri}^2+2Z_{ri}\sqrt{\eta_i}\sqrt{\frac{2}{h}}\Delta\beta_h^i\Big)-2\eta_i\eta_j\Big)^2\Big|\Delta W_h^Q\Big]\Big]\\
&=\mathrm{E}\Big[3\eta_i^2\eta_j^2+2\eta_i^2\eta_j^2-4\eta_i^2\eta_j^2+\frac{8}{h}\eta_i^2\eta_j^2(\Delta\beta_h^j)^2+3\eta_i^2\eta_j^2-4\eta_i^2\eta_j^2+\frac{8}{h}\eta_i^2\eta_j^2(\Delta\beta_h^i)^2+4\eta_i^2\eta_j^2\Big]\\
&=20\eta_i^2\eta_j^2
\end{aligned}
$$

and

$$
\mathrm{E}\Big[\mathrm{E}\Big[\eta_i^2\Big(Z_{rj}Z_{rm}+Z_{rj}\sqrt{\eta_m}\sqrt{\frac{2}{h}}\Delta\beta_h^m+Z_{rm}\sqrt{\eta_j}\sqrt{\frac{2}{h}}\Delta\beta_h^j\Big)^2\mathbb{1}_{i=n}\mathbb{1}_{j\neq m}\Big|\Delta W_h^Q\Big]\Big]
$$

$$
\begin{aligned}
&= \mathrm{E}\Big[\Big(\eta_i^2\Big(\eta_j\eta_m + \frac{2}{h}\eta_j\eta_m(\Delta\beta_h^m)^2 + \frac{2}{h}\eta_m\eta_j(\Delta\beta_h^j)^2\Big)\Big)\mathbb{1}_{i=n}\mathbb{1}_{j\neq m}\Big] \\
&= 5\eta_i^2\eta_j\eta_m\mathbb{1}_{i=n}\mathbb{1}_{j\neq m}
\end{aligned}
$$

for all $i, j, m, n \in \mathcal{J}_K$, $K \in \mathbb{N}$ with $i < j, m < n$. For the other terms of this type, we get similar results.

Moreover, we obtain for all $D \in \mathbb{N}$

$$
\sum_{r=D+1}^{\infty} \frac{1}{r^4} \leq \int_D^\infty \frac{1}{s^4}\,\mathrm{d}s = \frac{1}{3D^3}
$$

and

$$
\sum_{r=D+1}^{\infty} \frac{1}{r^2} \geq \int_{D+1}^\infty \frac{1}{s^2}\,\mathrm{d}s = \frac{1}{D+1}.
$$

So, we get

$$
\Big(\sum_{r=D+1}^{\infty} \frac{1}{r^4}\Big)\Big(\sum_{r=D+1}^{\infty} \frac{1}{r^2}\Big)^{-1} \leq \frac{D+1}{3D^3} \leq \frac{2}{3D^2}
$$

for all $D \in \mathbb{N}$.
In total, we get for all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$

$$
\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(B\int_t^s \mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}_{(i,j)}^Q(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq C\frac{h^2}{\eta_K\pi^2}\Big(\sum_{r=D+1}^{\infty}\frac{1}{r^2}\Big)^{-1}\sum_{r=D+1}^{\infty}\frac{1}{r^4}\Bigg(\sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}} 20\eta_i^2\eta_j^2 \\
&\quad + \sum_{\substack{i,j,m,n\in\mathcal{J}_K\\ i<j,m<n}} 5\Big(\eta_i^2\eta_j\eta_m\mathbb{1}_{i=n}\mathbb{1}_{j\neq m} + \eta_i^2\eta_j\eta_n\mathbb{1}_{i=m}\mathbb{1}_{j\neq n} + \eta_j^2\eta_i\eta_m\mathbb{1}_{j=n}\mathbb{1}_{i\neq m} + \eta_j^2\eta_i\eta_n\mathbb{1}_{j=m}\mathbb{1}_{i\neq n}\Big)\Bigg) \\
&\leq C\frac{h^2}{\eta_K\pi^2}\frac{2}{3D^2}\sum_{\substack{i,j\in\mathcal{J}_K\\ i<j}}\Big(20\eta_i^2\eta_j^2 + 5\eta_i^2\eta_j\sum_{\substack{m\in\mathcal{J}_K\\ m\neq j,m\neq i}}\eta_m + 5\eta_j^2\eta_i\sum_{\substack{m\in\mathcal{J}_K\\ m\neq j,m\neq i}}\eta_m\Big).
\end{aligned}
$$

Finally, this implies

$$
\begin{aligned}
&\mathrm{E}\Big[\Big\| \int_t^{t+h} G\Big(B\int_t^s \mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i\in\mathcal{J}_K}\sum_{j\in\mathcal{J}_K} \hat{I}_{(i,j)}^Q(h)\; G\big(B\tilde{e}_i, \tilde{e}_j\big)\Big\|_H^2\Big] \\
&\leq C\frac{h^2}{\eta_K\pi^2}\frac{2}{3D^2}\Big(20\Big(\sup_{j\in\mathcal{J}_K}\eta_j\Big)^2\big(\operatorname{tr}Q\big)^2 + 10\Big(\sup_{j\in\mathcal{J}_K}\eta_j\Big)\big(\operatorname{tr}Q\big)^3\Big) \\
&\leq C_Q\frac{h^2}{\eta_K D^2} \qquad (4.34)
\end{aligned}
$$

for all $h > 0$, $t, t+h \in [0,T]$, $D, K \in \mathbb{N}$. □

Proof of Theorem 4.3

Theorem (Convergence of DFM)
Let assumptions (C1)–(C4) and (C6a) or (C6b) be fulfilled. Then, there exists a constant $C_{T,Q} \in (0,\infty)$*, independent of* N*,* K*,* M*, and* D*, such that for* $(Y_m^Q)_{0\leq m\leq M}$*, defined by the DFM in* (4.26)*, it holds*

$$\Big(\mathrm{E}\Big[\|X_{t_m} - Y_m^Q\|_H^2\Big]\Big)^{\frac{1}{2}} \leq C_{T,Q}\Big(\Big(\inf_{i\in\mathcal{I}\setminus\mathcal{I}_N} \lambda_i\Big)^{-\gamma} + \Big(\sup_{j\in\mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)} + \mathcal{E}(D)\Big)$$

for all $m \in \{0,1,\ldots,M\}$ *and all* $N,K,M,D \in \mathbb{N}$*. The error term* $\mathcal{E}(D)$ *is determined by Algorithm 1 or 2 and given in Theorem 4.1 or Theorem 4.2, respectively. The parameters are determined by (C1)–(C4) and (C6).*

Proof of Theorem 4.3.
Throughout the proof, we use the following notation for all $m \in \{0,\ldots,M\}$, $M,N,K \in \mathbb{N}$,

$$\begin{aligned}
X_{t_m} =& e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}F(X_s)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-s)}B(X_s)\,\mathrm{d}W_s,\\
Y_m^{MIL} =& P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^{MIL})\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^{MIL})\,\mathrm{d}W_s^K\\
&+ \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B'(Y_l^{MIL})\Big(\int_{t_l}^{s} B(Y_l^{MIL})\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big),\\
\bar{Y}_m =& P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^Q)\,\mathrm{d}W_s^K\\
&+ \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B'(Y_l^Q)\Big(\int_{t_l}^{s} B(Y_l^Q)\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\Big),
\end{aligned}$$

and

$$\begin{aligned}
Y_m =& P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^Q)\,\mathrm{d}W_s^K\\
&+ \sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K} e^{A(t_m-t_l)}\Big(B\Big(Y_l^Q + \sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}^Q(h_l)\Big)\tilde{e}_j - B(Y_l)\tilde{e}_j\Big)\Big)
\end{aligned}$$

with $h_l := t_{l+1} - t_l$ for all $l \in \{0,\ldots,m-1\}$.

We estimate the error in several parts according to

$$\begin{aligned}
&\Big(\mathrm{E}\big[\|X_{t_m} - Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}}\\
&= \Big(\mathrm{E}\big[\|X_{t_m} - Y_m^{MIL} + Y_m^{MIL} - Y_m + Y_m - Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}}\\
&\leq \Big(\mathrm{E}\big[\|X_{t_m} - Y_m^{MIL}\|_H^2\big]\Big)^{\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
&+\Big(\mathrm{E}\big[\|Y_m^{MIL}-\bar{Y}_m+\bar{Y}_m-Y_m+Y_m-Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}}\\
&\leq \Big(\mathrm{E}\big[\|X_{t_m}-Y_m^{MIL}\|_H^2\big]\Big)^{\frac{1}{2}}\\
&\quad+\Big(\mathrm{E}\big[\|Y_m^{MIL}-\bar{Y}_m\|_H^2\big]\Big)^{\frac{1}{2}}+\Big(\mathrm{E}\big[\|\bar{Y}_m-Y_m\|_H^2\big]\Big)^{\frac{1}{2}}+\Big(\mathrm{E}\big[\|Y_m-Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}}
\end{aligned}
$$

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

The first part, which is the error resulting from the approximation with the Milstein scheme, is estimated in the same way as for the commutative scheme. For details, read the proof of Theorem 3.1 in Section 3.5.

We prove the following lemma on the moments of the approximation process $(Y_m^Q)_{m\in\{0,\ldots,M\}}$, $M \in \mathbb{N}$, which will be used throughout the proof if (C6b) does not hold. For now, we denote the approximation of $I_{(i,j)}^Q(h_l)$ by $\bar{I}_{(i,j)}^Q(h_l)$ for all $l \in \{0,\ldots,M\}$, $i,j \in \mathcal{J}_K$, $M,K \in \mathbb{N}$, independently of the algorithm that is employed to approximate it. We distinguish Algorithm 1 and Algorithm 2 during the proof.

Lemma 4.3

Let conditions (C1)–(C4) and (C6a) be fulfilled; then, it holds for some arbitrary $M,N,K \in \mathbb{N}$ and some constant $C_{p,T,Q} > 0$, independent of M,N,K,

$$
\sup_{m\in\{0,\ldots,M\}} \big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} \leq C_{p,T,Q}\big(1+\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}\big)
$$

for all $p \in [2,\infty)$.

Proof of Lemma 4.3.

For $m \in \{1,\ldots,M\}$, $M \in \mathbb{N}$, we assume that the statement has been proved for all Y_l^Q with $l \in \{0,\ldots,m-1\}$.

The triangle inequality implies for all $m \in \{1,\ldots,M\}$, $M,N,K \in \mathbb{N}$, and $p \in [1,\infty)$

$$
\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\\
&\leq \Bigg(C\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{1}{p}}+\sum_{l=0}^{m-1}\Bigg(\mathrm{E}\Bigg[\Big\|\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s\Big\|_{H_\delta}^p\Bigg]\Bigg)^{\frac{1}{p}}\\
&\quad+\Bigg(\mathrm{E}\Bigg[\Big\|\int_{t_0}^{t_m}\sum_{l=0}^{m-1} e^{A(t_m-t_l)}B(Y_l^Q)\mathbb{1}_{[t_l,t_{l+1})}(s)\,\mathrm{d}W_s^K\Big\|_{H_\delta}^p\Bigg]\Bigg)^{\frac{1}{p}}\\
&\quad+\Bigg(\mathrm{E}\Bigg[\Big\|\sum_{l=0}^{m-1} e^{A(t_m-t_l)}\sum_{j\in\mathcal{J}_K}\Big(B(Y_l^Q+\sum_{i\in\mathcal{J}_K}P_N B(Y_l^Q)\tilde{e}_i\bar{I}_{(i,j)}^Q(h_l))\tilde{e}_j-B(Y_l^Q)\tilde{e}_j\Big)\Big\|_{H_\delta}^p\Bigg]\Bigg)^{\frac{1}{p}}\Bigg)^2.
\end{aligned}
$$

By Hölders inequality, Theorem 2.6, and a Taylor expansion of the difference approximation, we obtain for all $p \in [2, \infty)$ and all $m \in \{1, \dots, M\}$, $M, K \in \mathbb{N}$

$$\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \\
&\leq C_p \Bigg(\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + \bigg(\sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} \mathrm{E}\Big[\big\|(-A)^\delta e^{A(t_m - t_l)} F(Y_l^Q)\big\|_H^p\Big] \,\mathrm{d}s\Big)^{\frac{1}{p}} h^{1-\frac{1}{p}} \bigg)^2 \\
&\quad + \int_{t_0}^{t_m} \bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} e^{A(t_m - t_l)} B(Y_l^Q) \mathbb{1}_{[t_l, t_{l+1})}(s) \Big\|_{L_{HS}(V_0, H_\delta)}^p\Big]\bigg)^{\frac{2}{p}} \mathrm{d}s \\
&\quad + \bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} (-A)^\delta e^{A(t_m - t_l)} \sum_{j \in \mathcal{J}_K} B'(\xi(Y_l^Q, j)) \Big(\sum_{i \in \mathcal{J}_K} P_N B(Y_l^Q) \tilde{e}_i \bar{I}_{(i,j)}^Q(h_l), \tilde{e}_j \Big) \Big\|_H^p\Big]\bigg)^{\frac{2}{p}} \Bigg).
\end{aligned}$$

Due to (C1)–(C3), Theorem 2.3, and since $\xi(Y_l^Q, j) \in H_\beta$ for all $l \in \{0, \dots, m-1\}$, $j \in \mathcal{J}_K$, and $M, N, K \in \mathbb{N}$, we get

$$\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big])\big)^{\frac{2}{p}} \\
&\leq C_p \big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_p h^{2-\frac{2}{p}} M \sum_{l=0}^{m-1} \Big(\int_{t_l}^{t_{l+1}} (t_m - t_l)^{-\delta p} \,\mathrm{d}s\Big)^{\frac{2}{p}} \big(\mathrm{E}\big[\|F(Y_l^Q)\|_H^p\big]\big)^{\frac{2}{p}} \\
&\quad + C_p \sum_{l=0}^{m-1} \int_{t_l}^{t_{l+1}} \bigg(\mathrm{E}\Big[\Big\| \sum_{l=0}^{m-1} e^{A(t_m - t_l)} B(Y_l^Q) \mathbb{1}_{[t_l, t_{l+1})}(s) \Big\|_{L_{HS}(V_0, H_\delta)}^p\Big]\bigg)^{\frac{2}{p}} \mathrm{d}s \\
&\quad + C_p M \sum_{l=0}^{m-1} (t_m - t_l)^{-2\delta} \bigg(\mathrm{E}\Big[\Big\| \sum_{j \in \mathcal{J}_K} B'(\xi(Y_l^Q, j)) \Big(\sum_{i \in \mathcal{I}_K} P_N B(Y_l^Q) \tilde{e}_i \bar{I}_{(i,j)}^Q(h_l), \tilde{e}_j \Big) \Big\|_H^p\Big]\bigg)^{\frac{2}{p}} \\
&\leq C_p \big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{p,T} h^{1-\frac{2}{p}} \sum_{l=0}^{m-1} \Big(h (t_m - t_l)^{-\delta p}\Big)^{\frac{2}{p}} \big(1 + \big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p \sum_{l=0}^{m-1} \Big(\mathrm{E}\big[\|B(Y_l^Q)\|_{L_{HS}(V_0, H_\delta)}^p\big]\Big)^{\frac{2}{p}} \int_{t_l}^{t_{l+1}} \big\|(-A)^{-\delta}\big\|_{L(H)}^2 \big\|(-A)^\delta e^{A(t_m - t_l)}\big\|_{L(H)}^2 \,\mathrm{d}s \\
&\quad + C_p M \sum_{l=0}^{m-1} (t_m - t_l)^{-2\delta} \bigg(\sum_{j \in \mathcal{J}_K} \Big(\mathrm{E}\Big[\|B'(\xi(Y_l^Q, j))\|_{L(H, L(V, H))}^p \Big\| B(Y_l^Q) \sum_{i \in \mathcal{J}_K} \bar{I}_{(i,j)}^Q(h_l) \tilde{e}_i \Big\|_H^p\Big]\Big)^{\frac{1}{p}} \bigg)^2 \\
&\leq C_p \big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + h^{1-2\delta} C_{p,T} \sum_{l=0}^{m-1} (m - l)^{-2\delta} \big(1 + \big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p \sum_{l=0}^{m-1} h (t_m - t_l)^{-2\delta} \Big(\mathrm{E}\big[\|B(Y_l^Q)\|_{L_{HS}(V_0, H_\delta)}^p\big]\Big)^{\frac{2}{p}} \\
&\quad + C_p M h^{-2\delta} \sum_{l=0}^{m-1} (m - l)^{-2\delta} \bigg(\sum_{j \in \mathcal{J}_K} \Big(\mathrm{E}\big[\|B(Y_l^Q)\|_{L(V, H_\delta)}^p\big]\Big)^{\frac{1}{p}} \sum_{i \in \mathcal{J}_K} \big(\mathrm{E}\big[(\bar{I}_{(i,j)}^Q(h_l))^p\big]\big)^{\frac{1}{p}} \bigg)^2
\end{aligned}$$

for all $m \in \{1, \dots, M\}$, $M, K \in \mathbb{N}$ and $p \in [2, \infty)$.

Further, we obtain by (C6a) and the distribution properties of $\bar{I}_{(i,j)}$, $i, j \in \mathcal{J}_K$, see Section 4.1, for all $m \in \{1, \ldots, M\}$, $M, K \in \mathbb{N}$, $p \in [2, \infty)$

$$\begin{aligned}
&\big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \\
&\leq C_p\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{p,T} h^{1-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}\big(1 + \big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p h \sum_{l=0}^{m-1} (t_m - t_l)^{-2\delta}\big(1 + \big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p M h^{-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}\Big(\big(1 + \mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} \sum_{i,j \in \mathcal{J}_K} \big(\mathrm{E}\big[|\bar{I}_{(i,j)}(h_l)\sqrt{\eta_i}\sqrt{\eta_j}|^{2p}\big]\big)^{\frac{1}{2p}}\Big)^2 \\
&\leq C_p\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{p,T} h^{1-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}\big(1 + \big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big) \\
&\quad + C_p M h^{-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}\Big(\big(1 + \mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{1}{p}} \sum_{i,j \in \mathcal{J}_K} \sqrt{\eta_i}\sqrt{\eta_j}\, h\Big)^2 \\
&\leq C_p\big(\mathrm{E}\big[\,\|X_0\|_{H_\delta}^p\,\big]\big)^{\frac{2}{p}} + C_{p,T} h^{1-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}(1 + \big(\mathrm{E}[\|Y_l^Q\|_{H_\delta}^p]\big)^{\frac{2}{p}}) \\
&\quad + C_{p,Q} h^{1-2\delta} \sum_{l=0}^{m-1} (m-l)^{-2\delta}\big(1 + \big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big).
\end{aligned}$$

As in equation (3.31) in the proof of Theorem 3.1, we obtain for $\delta \in (0, \frac{1}{2})$ and all $m \in \{1, \ldots, M\}$, $M \in \mathbb{N}$

$$\sum_{l=0}^{m-1} (m-l)^{-2\delta} \leq \frac{M^{1-2\delta}}{1-2\delta}.$$

Therefore, we get for all $m \in \{1, \ldots, M\}$, $M \in \mathbb{N}$, $p \in [2, \infty)$ the estimate

$$\big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} \leq C_p\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{p,T,Q} + h^{1-2\delta} C_{p,T,Q} \sum_{l=0}^{m-1} (m-l)^{-2\delta}\big(\mathrm{E}\big[\|Y_l^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}$$

and the discrete Gronwall Lemma implies

$$\begin{aligned}
\big(\mathrm{E}\big[\|Y_m^Q\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} &\leq \Big(C_p\big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}} + C_{p,T,Q}\Big) e^{C_{p,T,Q} \sum_{l=0}^{m-1} (m-l)^{-2\delta} h^{1-2\delta}} \\
&\leq C_{p,T,Q}\big(1 + \big(\mathrm{E}\big[\|X_0\|_{H_\delta}^p\big]\big)^{\frac{2}{p}}\big)
\end{aligned}$$

for all $m \in \{1, \ldots, M\}$, $M \in \mathbb{N}$, $p \in [2, \infty)$. □

Now, we continue with the proof of Theorem 4.3. The estimate of $\mathrm{E}\big[\|Y_m^{MIL} - \bar{Y}_m\|_H^2\big]$ can be obtained as in Theorem 3.1 on page 56 for all $m \in \{0, \ldots, M\}$, $M, K \in \mathbb{N}$. We get

$$\mathrm{E}\big[\|Y_m^{MIL} - \bar{Y}_m\|_H^2\big] \leq C_T h \sum_{l=0}^{m-1} \mathrm{E}\big[\|Y_l^{MIL} - Y_l^Q\|_H^2\big] \tag{4.35}$$

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

For the next term, we obtain

$$
\begin{aligned}
&\mathrm{E}\big[\|\bar{Y}_m - Y_m\|_H^2\big] = \\
&\mathrm{E}\Big[\Big\| P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^Q)\,\mathrm{d}W_s^K\Big) \\
&\quad + P_N\Big(\sum_{i,j\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}\sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B'(Y_l^Q)\Big(\int_{t_l}^{s} B(Y_l^Q)\tilde{e}_i\,\mathrm{d}\beta_r^i\Big)\tilde{e}_j\,\mathrm{d}\beta_s^j\Big) \\
&\quad - P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^Q)\,\mathrm{d}W_s^K\Big) \\
&\quad - P_N\Big(\sum_{l=0}^{m-1} e^{A(t_m-t_l)}\sum_{j\in\mathcal{J}_K}\Big(B\Big(Y_l^Q + \sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big)\tilde{e}_j - B(Y_l^Q)\tilde{e}_j\Big)\Big)\Big\|_H^2\Big] \\
=&\mathrm{E}\Big[\Big\| P_N\Big(\sum_{l=0}^{m-1}\sum_{i,j\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}e^{A(t_m-t_l)}B'(Y_l^Q)\big(B(Y_l^Q)\tilde{e}_i,\tilde{e}_j\big)I_{(i,j)}(h_l)\Big) \\
&\quad - P_N\Big(\sum_{l=0}^{m-1} e^{A(t_m-t_l)}\sum_{j\in\mathcal{J}_K}\Big(B\Big(Y_l^Q + \sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big)\tilde{e}_j - B(Y_l^Q)\tilde{e}_j\Big)\Big)\Big\|_H^2\Big]
\end{aligned}
$$

for all $m \in \{1, \ldots, M\}$, $M, N, K \in \mathbb{N}$.
A Taylor approximation of second order of the second term, similar to (3.30), assumption (C3), and the triangle inequality imply

$$
\begin{aligned}
&\mathrm{E}\big[\|\bar{Y}_m - Y_m\|_H^2\big] \\
&\leq \mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1} e^{A(t_m-t_l)}\sum_{j\in\mathcal{J}_K}\frac{1}{2}\int_0^1\int_0^r B''\Big(Y_l^Q + u\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big) \\
&\qquad\Big(\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l), \sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big)\tilde{e}_j\, r\,\mathrm{d}u\,\mathrm{d}r\Big\|_H^2\Big] \\
&\leq \mathrm{E}\Big[\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\frac{1}{2}\int_0^1\int_0^r \Big\| e^{A(t_m-t_l)}B''\Big(Y_l^Q + u\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big) \\
&\qquad\Big(\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l), \sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big)\tilde{e}_j\Big\|_H r\,\mathrm{d}u\,\mathrm{d}r\Big)^2\Big] \\
&\leq C\mathrm{E}\Big[\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\int_0^1\int_0^r \Big\| B''\Big(Y_l^Q + u\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big)\Big\|_{L^{(2)}(H,L(V,H))} \\
&\qquad\qquad\Big\|\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}(h_l)\Big\|_H^2 r\,\mathrm{d}u\,\mathrm{d}r\Big)^2\Big] \\
&\leq C\mathrm{E}\Big[\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\Big\| B(Y_l^Q)\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}\tilde{e}_i I_{(i,j)}(h_l)\Big\|_H^2\Big)^2\Big]
\end{aligned}
$$

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

Assume that (C6a) holds; by condition (C3) and due to Lemma 4.3, we obtain

$$\begin{aligned}
&\mathrm{E}\big[\|\bar{Y}_m - Y_m\|_H^2\big] \\
&\leq C\mathrm{E}\Big[\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\|B(Y_l^Q)\|_{L(V,H)}^2\Big\|\sum_{i\in\mathcal{J}_K}\sqrt{\eta_j}\sqrt{\eta_i}\tilde{e}_i I_{(i,j)}(h_l)\Big\|_V^2\Big)^2\Big] \\
&\leq C\mathrm{E}\Big[\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\|B(Y_l^Q)\|_{L(V,H_\delta)}^2\sum_{i_1,i_2\in\mathcal{J}_K}\eta_j\sqrt{\eta_{i_1}}\sqrt{\eta_{i_2}}I_{(i_1,j)}(h_l)I_{(i_2,j)}(h_l)\langle\tilde{e}_{i_1},\tilde{e}_{i_2}\rangle_V\Big)^2\Big] \\
&= C\mathrm{E}\Big[\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\|B(Y_l^Q)\|_{L(V,H_\delta)}^2\sum_{i\in\mathcal{J}_K}\eta_j\eta_i I_{(i,j)}^2(h_l)\Big)^2\Big] \\
&\leq C\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\sum_{i\in\mathcal{J}_K}\Big(\mathrm{E}\big[\|B(Y_l^Q)\|_{L(V,H_\delta)}^4\big]^{\frac{1}{2}}\Big)\Big(\mathrm{E}\Big[\big(\eta_j\eta_i I_{(i,j)}^2(h_l)\big)^2\Big]\Big)^{\frac{1}{2}}\Big)^2 \\
&\leq C\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\sum_{i\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\eta_j^2\eta_i^2 I_{(i,j)}^4(h_l)\Big]\Big)^{\frac{1}{2}}\Big)^2
\end{aligned}$$

for all $m \in \{1,\ldots,M\}$, $M, K \in \mathbb{N}$. If (C6b) holds instead, the estimate follows similarly.

Eventually, we obtain

$$\begin{aligned}
\mathrm{E}\big[\|\bar{Y}_m - Y_m\|_H^2\big] &\leq C\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K}\eta_j\sum_{i\in\mathcal{J}_K}\eta_i\Big(\mathrm{E}\Big[I_{(i,j)}^4(h_l)\Big]\Big)^{\frac{1}{2}}\Big)^2 \\
&\leq C\Big(\sum_{l=0}^{m-1}(\operatorname{tr}Q)^2h^2\Big)^2 \leq C_T h^2(\operatorname{tr}Q)^4 \qquad (4.36)
\end{aligned}$$

by the distribution properties of $I_{(i,j)}(h_l)$, $l \in \{0,\ldots,m-1\}$, $i,j \in \mathcal{J}_K$ for all $m \in \{1,\ldots,M\}$, $M, K \in \mathbb{N}$.

Next, we determine the error that results from the approximation of the iterated stochastic integral, that is, we estimate the term $\Big(\mathrm{E}\big[\|Y_m - Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}}$, for all $m \in \{0,\ldots,M\}$, $M \in \mathbb{N}$. We compute Taylor approximations of first order of the approximation operators of the derivative and obtain for all $m \in \{1,\ldots,M\}$, $M, N, K \in \mathbb{N}$

$$\begin{aligned}
&\mathrm{E}\big[\|Y_m - Y_m^Q\|_H^2\big] \\
&= \mathrm{E}\Big[\Big\|P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^Q)\,\mathrm{d}W_s^K \\
&\qquad + \sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K} e^{A(t_m-t_l)}\Big(B\Big(Y_l^Q + \sum_{i\in\mathcal{J}_K}P_N B(Y_l^Q)\tilde{e}_i I_{(i,j)}^Q(h_l)\Big)\tilde{e}_j - B(Y_l^Q)\tilde{e}_j\Big)\Big) \\
&\qquad - P_N\Big(e^{At_m}X_0 + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}F(Y_l^Q)\,\mathrm{d}s + \sum_{l=0}^{m-1}\int_{t_l}^{t_{l+1}} e^{A(t_m-t_l)}B(Y_l^Q)\,\mathrm{d}W_s^K
\end{aligned}$$

$$+\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K} e^{A(t_m-t_l)}\Big(B\Big(Y_l^Q+\sum_{i\in\mathcal{J}_K} P_N B(Y_l)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l)\Big)\tilde{e}_j - B(Y_l^Q)\tilde{e}_j\Big)\Big)\Big\|_H^2\Big]$$

$$\begin{aligned}
&= \mathrm{E}\Big[\Big\|P_N\Big(\sum_{l=0}^{m-1}\sum_{j\in\mathcal{J}_K} e^{A(t_m-t_l)}\Big(B'(Y_l^Q)\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l),\tilde{e}_j\Big)\\
&\quad+\frac{1}{2}e^{A(t_m-t_l)}B''(\xi(Y_l^Q,j))\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l),\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l)\Big)\tilde{e}_j\\
&\quad- B'(Y_l^Q)\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l),\tilde{e}_j\Big)\\
&\quad-\frac{1}{2}e^{A(t_m-t_l)}B''(\bar{\xi}(Y_l,j))\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l),\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l)\Big)\tilde{e}_j\Big)\Big)\Big\|_H^2\Big].
\end{aligned}$$

By rearranging the expression, we obtain

$$\begin{aligned}
&\mathrm{E}\big[\|Y_m-Y_m^Q\|_H^2\big]\\
&\leq C\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1} e^{A(t_m-t_l)}\Big(\int_{t_l}^{t_{l+1}} B'(Y_l^Q)\Big(\int_{t_l}^{s} B(Y_l^Q)\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K\\
&\qquad\qquad - \sum_{i,j\in\mathcal{J}_K}\bar{I}^Q_{(i,j)}B'(Y_l^Q)(B(Y_l^Q)\tilde{e}_i,\tilde{e}_j)\Big)\Big\|_H^2\Big]\\
&\quad+C\mathrm{E}\Big[\Big\|\sum_{l=0}^{m-1} e^{A(t_m-t_l)}\sum_{j\in\mathcal{J}_K}\frac{1}{2}\\
&\qquad\Big(B''(\xi(Y_l^Q,j))\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l),\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l)\Big)\tilde{e}_j\\
&\qquad- B''(\bar{\xi}(Y_l^Q,j))\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l),\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l)\Big)\tilde{e}_j\Big)\Big\|_h^2\Big]\\
&\leq C\sum_{l=0}^{m-1}\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} B'(Y_l^Q)\Big(\int_{t_l}^{s} B(Y_l^Q)\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K-\sum_{i,j\in\mathcal{J}_K}\bar{I}^Q_{(i,j)}(h_l)B'(Y_l^Q)(B(Y_l^Q)\tilde{e}_i,\tilde{e}_j)\Big\|_H^2\Big]\\
&\quad+CM\sum_{l=0}^{m-1}\Bigg(\sum_{j\in\mathcal{J}_K}\Bigg(\mathrm{E}\Big[\Big\|\frac{1}{2}\Big(B''(\xi(Y_l^Q,j))\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l),\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i I^Q_{(i,j)}(h_l)\Big)\tilde{e}_j\\
&\qquad- B''(\bar{\xi}(Y_l^Q,j))\Big(\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l),\sum_{i\in\mathcal{J}_K} P_N B(Y_l^Q)\tilde{e}_i \bar{I}^Q_{(i,j)}(h_l)\Big)\tilde{e}_j\Big)\Big\|_H^2\Big]\Bigg)^{\frac{1}{2}}\Bigg)^2
\end{aligned}$$

for all $m\in\{1,\dots,M\}$, $M,N,K\in\mathbb{N}$.

The first term can be estimated as in Theorem 4.1 or Theorem 4.2, respectively. We get for Algorithm 1

$$\begin{aligned}
&\mathrm{E}\Big[\Big\|\int_{t_l}^{t_{l+1}} B'(Y_l^Q)\Big(\int_{t_l}^{s} B(Y_l^Q)\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K-\sum_{i,j\in\mathcal{J}_K}\tilde{I}^Q_{(i,j)}(h_l)B'(Y_l^Q)(B(Y_l^Q)\tilde{e}_i,\tilde{e}_j)\Big\|_H^2\Big]\\
&\leq C_Q\frac{h^2}{\pi^2 D}
\end{aligned}$$

and for Algorithm 2

$$\mathrm{E}\Big[\Big\| \int_{t_l}^{t_{l+1}} B'(Y_l^Q)\Big(\int_{t_l}^{s} B(Y_l^Q)\,\mathrm{d}W_r^K\Big)\,\mathrm{d}W_s^K - \sum_{i,j\in\mathcal{J}_K} \hat{I}_{(i,j)}^Q(h_l) B'(Y_l^Q)(B(Y_l^Q)\tilde{e}_i,\tilde{e}_j)\Big\|_H^2\Big]$$
$$\leq C_Q \frac{h^2}{D^2}\eta_K^{-1}$$

for all $l \in \{0,\dots,m-1\}$, $m \in \{1,\dots,M\}$, $h>0$, and $M,K,D \in \mathbb{N}$. In the following, we keep the proof independent of the algorithm to approximate the iterated stochastic integral and denote this error by $\mathcal{E}(h,D)$ for all $h>0$, $D\in\mathbb{N}$.
Then, we get by assumption (C3)

$$\begin{aligned}
&\mathrm{E}\big[\|Y_m - Y_m^Q\|_H^2\big]\\
\leq& C\sum_{l=0}^{m-1}\mathcal{E}(h,D)\\
&+CM\sum_{l=0}^{m-1}\Big(\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\|B''(\xi(Y_l^Q,j))\|_{L^{(2)}(H,L(V,H))}^2\|\sum_{i\in\mathcal{J}_K}B(Y_l^Q)\tilde{e}_i I_{(i,j)}^Q(h_l)\|_H^4\Big]\Big)^{\frac{1}{2}}\Big)^2\\
&+CM\sum_{l=0}^{m-1}\Big(\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\|B''(\bar{\xi}(Y_l^Q,j))\|_{L^{(2)}(H,L(V,H))}^2\|\sum_{i\in\mathcal{J}_K}B(Y_l^Q)\tilde{e}_i \bar{I}_{(i,j)}^Q(h_l)\|_H^4\Big]\Big)^{\frac{1}{2}}\Big)^2\\
\leq& C\sum_{l=0}^{m-1}\mathcal{E}(h,D)\\
&+CM\sum_{l=0}^{m-1}\Big(\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\Big(\sum_{i_1,i_2\in\mathcal{J}_K}I_{(i_1,j)}^Q(h_l)I_{(i_2,j)}^Q(h_l)\langle\tilde{e}_{i_1},\tilde{e}_{i_2}\rangle_V\Big)^2\|B(Y_l^Q)\|_{L(V,H)}^4\Big]\Big)^{\frac{1}{2}}\Big)^2\\
&+CM\sum_{l=0}^{m-1}\Big(\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\Big(\sum_{i_1,i_2\in\mathcal{J}_K}\bar{I}_{(i_1,j)}^Q(h_l)\bar{I}_{(i_2,j)}^Q(h_l)\langle\tilde{e}_{i_1},\tilde{e}_{i_2}\rangle_V\Big)^2\|B(Y_l^Q)\|_{L(V,H)}^4\Big]\Big)^{\frac{1}{2}}\Big)^2\\
\leq& C\sum_{l=0}^{m-1}\mathcal{E}(h,D)+CM\sum_{l=0}^{m-1}\Big(\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\Big(\sum_{i\in\mathcal{J}_K}(I_{(i,j)}^Q(h_l))^2\Big)^2\|B(Y_l^Q)\|_{L(V,H)}^4\Big]\Big)^{\frac{1}{2}}\Big)^2\\
&+CM\sum_{l=0}^{m-1}\Big(\sum_{j\in\mathcal{J}_K}\Big(\mathrm{E}\Big[\Big(\sum_{i\in\mathcal{J}_K}(\bar{I}_{(i,j)}^Q(h_l))^2\Big)^2\|B(Y_l^Q)\|_{L(V,H)}^4\Big]\Big)^{\frac{1}{2}}\Big)^2.
\end{aligned}$$

Due to the properties of $I_{(i,j)}^Q(h_l)$, $\bar{I}_{(i,j)}^Q(h_l)$ for $l\in\{0,\dots,M-1\}$, $i,j\in\mathcal{J}_K$, $M,K\in\mathbb{N}$, assumptions (C3), (C6b) or (C6a), and Lemma 4.3, we obtain for all $m\in\{1,\dots,M\}$, $M\in\mathbb{N}$

$$\begin{aligned}
\mathrm{E}\big[\|Y_m - Y_m^Q\|_H^2\big] \leq& C\sum_{l=0}^{m-1}\mathcal{E}(h,D)+C_QM\sum_{l=0}^{m-1}\Big(\Big(h^4\,\mathrm{E}\big[\|B(Y_l^Q)\|_{L(V,H)}^4\big]\Big)^{\frac{1}{2}}\Big)^2\\
\leq& C\sum_{l=0}^{m-1}\mathcal{E}(h,D)+C_{T,Q}\sum_{l=0}^{m-1}h^3\\
\leq& C\sum_{l=0}^{m-1}\mathcal{E}(h,D)+C_{T,Q}h^2. \qquad (4.37)
\end{aligned}$$

Now, let $\bar{I}^Q_{(i,j)}(h_l) = \tilde{I}^Q_{(i,j)}(h_l)$ for $l \in \{0, \ldots, M\}$, $i, j \in \mathcal{J}_K$, $M, K \in \mathbb{N}$, that is, we approximate the iterated integrals by Algorithm 1. Then, we get for all $m \in \{1, \ldots, M\}$, $M, D \in \mathbb{N}$

$$\mathrm{E}\big[\|Y_m - Y_m^Q\|_H^2\big] \leq C \sum_{l=0}^{m-1} C_Q \frac{h^2}{\pi^2 D} + C_{T,Q} h^2 \leq C_{T,Q}\Big(\frac{h}{D} + h^2\Big).$$

For Algorithm 2, on the other hand, we obtain

$$\mathrm{E}\big[\|Y_m - Y_m^Q\|_H^2\big] \leq C_Q \sum_{l=0}^{m-1} \frac{h^2}{D^2}\eta_K^{-1} + C_{T,Q} h^2 \leq C_{T,Q}\Big(\frac{h}{D^2}\eta_K^{-1} + h^2\Big).$$

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

We combine estimates (4.35), (4.36), and (4.37) in order to obtain

$$\begin{aligned}\mathrm{E}\big[\|Y_m^{MIL} - Y_m^Q\|_H^2\big] &\leq C_T h \sum_{l=0}^{m-1} \mathrm{E}\big[\|Y_l^{MIL} - Y_l^Q\|_H^2\big] + C_T h^2 (\operatorname{tr} Q)^4 + CM\mathcal{E}(h, D) + C_{T,Q} h^2 \\ &\leq CM\mathcal{E}(h, D) + C_{T,Q} h^2\end{aligned}$$

for all $m \in \{1, \ldots, M\}$, $M, D \in \mathbb{N}$ by Gronwall's Lemma.

In total, this yields for the DFM1

$$\Big(\mathrm{E}\big[\|X_{t_m} - Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}} \leq C_{T,Q}\Big(\Big(\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i\Big)^{-\gamma} + \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j\Big)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)} + \frac{\sqrt{h}}{\sqrt{D}}\Big),$$

and for the DFM2 it holds

$$\Big(\mathrm{E}\big[\|X_{t_m} - Y_m^Q\|_H^2\big]\Big)^{\frac{1}{2}} \leq C_{T,Q}\Big(\Big(\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i\Big)^{-\gamma} + \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j\Big)^{\alpha} + M^{-\min(2(\gamma-\beta),\gamma)} + \frac{\sqrt{h}}{D}\eta_K^{-\frac{1}{2}}\Big)$$

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

4.4 Numerical Analysis

We illustrate the theoretical results obtained in the previous sections with some numerical simulations now. Therefore, we compare the derivative-free Milstein scheme combined with Algorithm 1 or Algorithm 2 to the exponential Euler scheme in this section. We refrain from implementing a combination of the Milstein scheme with Algorithm 1 or Algorithm 2, respectively, as the effective order of convergence of this scheme is lower than for the DFM1 and DFM2; this follows in the same way as in Section 3.4.

Consider a setting similar to the one outlined in Section 3.6 and fix $H = V = L^2((0,1), \mathbb{R})$. Moreover, we choose $F(y) = 1 - y$ for all $y \in H_\beta$, $e_i = \tilde{e}_i = \sqrt{2}\sin(i\pi x)$ for all $i \in \mathbb{N}$, $x \in (0,1)$, and assume $X_t(0) = X_t(1) = 0$, $X_0(x) = 0$ for all $t \in (0, T]$, $x \in (0,1)$.

Let $\mu_{ij}(y) = \frac{\langle y, e_j \rangle_H}{i^p + j^4}$ for all $i \in \mathcal{I}$, $j \in \mathcal{J}$, $y \in H_\beta$ and some $p > 2$.

This implies $\phi^k_{ij}(y) = \begin{cases} 0, & k \neq j \\ \frac{1}{i^p + j^4}, & k = j \end{cases}$ for all $i, k \in \mathcal{I}$, $j \in \mathcal{J}$, $y \in H_\beta$.

The equation does not fulfill (C5) as

$$\sum_{k \in \mathcal{I}} \phi^k_{im}(y) \mu_{kn}(y) = \frac{1}{i^p + m^4} \frac{\langle y, e_n \rangle_H}{m^p + n^4}$$

but

$$\sum_{k \in \mathcal{I}} \phi^k_{in}(y) \mu_{km}(y) = \frac{1}{i^p + n^4} \frac{\langle y, e_m \rangle_H}{n^p + m^4}$$

holds for all $y \in H_\beta$ and all $i \in \mathcal{I}$, $n, m \in \mathcal{J}_K$, $K \in \mathbb{N}$.
Conditions (C2) and (C4) are, however, obviously fulfilled.
In the following examples, we choose $A = \frac{\Delta}{100}$, that is, $\lambda_i = \frac{\pi^2 i^2}{100}$, $e_i = \sqrt{2}\sin(i\pi x)$ for all $i \in \mathbb{N}$, $x \in (0,1)$ such that (C1) holds. Next, we verify (C3). From Section 3.6, we get

$$\|B(y)\|_{L(V,H_\delta)} \leq \sum_{k \in \mathcal{I}} \sum_{j \in \mathcal{J}} \lambda_k^\delta |\mu_{kj}(y)| \leq \sum_{k \in \mathcal{I}} \sum_{j \in \mathcal{J}} \frac{1}{j^2} \pi^{2\delta} \frac{1}{k^{\frac{p}{2} - 2\delta}} \|y\|_H$$

for all $y \in H_\delta$. It holds, $\|B(y)\|_{L(V,H_\delta)} \leq C(1 + \|y\|_{H_\delta})$ for all $y \in H_\delta$ if $\delta < \frac{p-2}{4}$. We select the maximal value for δ in the examples below.
Moreover, we get

$$\begin{aligned} \|(-A)^{-\vartheta} B(z) Q^{-\alpha}\|_{L_{HS}(V_0,H)} =& \Big(\sum_{k \in \mathcal{J}} \eta_k^{1-2\alpha} \sum_{i \in \mathcal{I}} \lambda_i^{-2\vartheta} \mu^2_{ik}(z) \Big)^{\frac{1}{2}} \\ \leq& \Big(\sum_{k \in \mathcal{J}} \frac{1}{k^{\rho_Q(1-2\alpha)+4}} \sum_{i \in \mathcal{I}} \frac{1}{i^{p+4\vartheta}} \|z\|_H^2 \Big)^{\frac{1}{2}} \end{aligned}$$

for all $z \in H_\gamma$. So, $\|(-A)^{-\vartheta} B(z) Q^{-\alpha}\|_{L_{HS}(V_0,H)} \leq C(1 + \|z\|_{H_\gamma})$ holds true for all $z \in H_\gamma$ if $\alpha < \frac{\rho_Q + 3}{2\rho_Q}$.

Example 1 - Similar Effective Orders for DFM1 and DFM2

First, let $p = 4$ and $\rho_Q = 3$. In this case, (C3) holds with $\delta \in (0, \frac{1}{2})$ and $\alpha \in (0,1)$. Moreover, (C6a) is fulfilled as $\rho_Q > 2$. These parameters yield $\gamma \in [\frac{1}{2}, 1)$ for $\beta = 0$. We select $q_{DFM} = \gamma = 1 - \varepsilon$, $\alpha = 1 - \varepsilon$ for some $\varepsilon \in (0, \frac{2}{5})$.
According to Table 4.1, we determine which scheme to use in this setting. It holds

$$2\gamma\rho_A\alpha\rho_Q + \gamma\rho_A - \alpha\rho_Q > 0 \quad \text{and} \quad 2\gamma\rho_A\alpha\rho_Q + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0.$$

Next, we compute $\gamma\rho_A > \rho_Q(\alpha - \gamma\rho_A)$ and

$$\frac{\alpha\rho_Q\gamma\rho_A}{2\alpha\rho_Q\gamma\rho_A + \gamma\rho_A - \alpha\rho_Q} < \frac{3}{5} < q.$$

Moreover, we get $q > \alpha(2q-1)$, $\alpha\rho_Q > \gamma\rho_A$ and, finally, we show

$$\frac{\alpha\rho_Q\gamma\rho_A}{2\gamma\rho_A\alpha\rho_Q + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q)} < \frac{3}{8} < q.$$

Therefore, in this setting, we expect the scheme DFM1 to achieve a higher effective order of convergence.
Let us fix some $N \in \mathbb{N}$, we obtain the relation of N, M, K as $M = N^2$ and $K = N^{\frac{2}{3}}$. For the expected order of convergence, we get error(DFM1) = error(DFM2) $= \mathcal{O}(\bar{c}^{-\frac{3}{8}+\varepsilon})$ for some $\varepsilon > 0$. In the EES, however, we take $M = N^4$, $K = N^{\frac{2}{3}}$ and obtain error(EES) $= \mathcal{O}(\bar{c}^{-\frac{6}{17}+\varepsilon})$.

The following table and plot illustrate the results of the numerical analysis. We simulate 300 paths with the approximation schemes and compare the results to a substitute for the exact solution. We choose an approximation computed with the linear implicit Euler scheme with $N_X = 2^6$, $K_X = 2^4$, and $M_X = 2^{20}$ for comparison. In Figure 4.1, one observes that all the schemes converge with the expected order of convergence. Moreover, from Table 4.2 we get an impression on the difference in the computational cost of the numerical schemes.

			DFM1			DFM2		
N	M	K	CC	Error	Std	CC	Error	Std
2	4	$2^{\frac{2}{3}}$	64	$3.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-3}$	71	$3.4 \cdot 10^{-2}$	$3.6 \cdot 10^{-3}$
4	2^4	$2^{\frac{4}{3}}$	1344	$2.5 \cdot 10^{-2}$	$2.1 \cdot 10^{-4}$	1574	$2.5 \cdot 10^{-2}$	$2.1 \cdot 10^{-4}$
8	2^6	2^2	30720	$1.7 \cdot 10^{-2}$	$2.7 \cdot 10^{-5}$	30720	$1.7 \cdot 10^{-2}$	$2.7 \cdot 10^{-5}$
16	2^8	$2^{\frac{8}{3}}$	1462272	$6.3 \cdot 10^{-3}$	$7.0 \cdot 10^{-6}$	1679055	$6.3 \cdot 10^{-3}$	$7.0 \cdot 10^{-6}$
32	2^{10}	$2^{\frac{10}{3}}$	58753024	$1.6 \cdot 10^{-3}$	$8.4 \cdot 10^{-6}$	66832237	$1.6 \cdot 10^{-3}$	$8.3 \cdot 10^{-6}$

			Linear Implicit Euler			Exponential Euler		
N	M	K	CC	Error	Std	CC	Error	Std
2	2^4	$2^{\frac{2}{3}}$	64	$2.2 \cdot 10^{-2}$	$5.6 \cdot 10^{-3}$	64	$2.2 \cdot 10^{-2}$	$5.6 \cdot 10^{-3}$
4	2^8	$2^{\frac{4}{3}}$	3072	$2.6 \cdot 10^{-2}$	$5.1 \cdot 10^{-4}$	3072	$2.7 \cdot 10^{-2}$	$5.2 \cdot 10^{-4}$
8	2^{12}	2^2	131072	$1.7 \cdot 10^{-2}$	$3.9 \cdot 10^{-5}$	131072	$1.7 \cdot 10^{-2}$	$4.0 \cdot 10^{-5}$
16	2^{16}	$2^{\frac{8}{3}}$	7340032	$6.1 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$	7340032	$6.1 \cdot 10^{-3}$	$1.2 \cdot 10^{-6}$
32	2^{20}	$2^{\frac{10}{3}}$	369098752	$1.5 \cdot 10^{-3}$	$2.6 \cdot 10^{-6}$	369098752	$1.5 \cdot 10^{-3}$	$2.7 \cdot 10^{-6}$

Table 4.2: Error and standard deviation for Example 1 obtained from 300 paths. CC denotes the computational cost computed as $CC(\text{DFM1}) = 3MNK + M^{2q}\frac{K(K-1)}{2}$, $CC(\text{DFM2}) = 3MNK + M^{q+\frac{1}{2}}K^{\frac{\rho_Q}{2}}\frac{K(K-1)}{2}$, and $CC(\text{EES}) = CC(\text{LIE}) = MNK$.

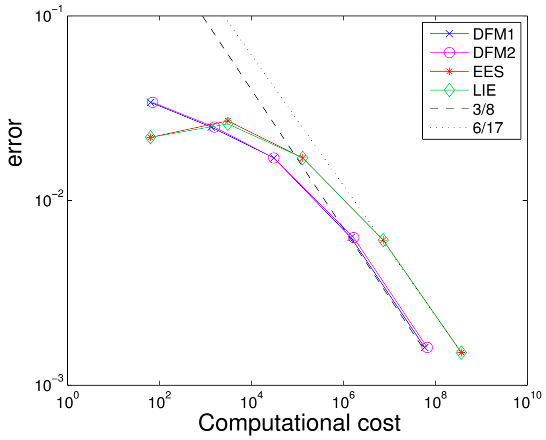

Figure 4.1: Error against computational cost for Example 1 for 300 paths and $N = 2, 4, 8, 16, 32$ in log-log scale.

Example 2 - Lower Computational Cost for DFM1

Now, we choose $p = 3$ and $\rho_Q = 3$. Therewith, we get $\delta \in (0, \frac{1}{4})$ and $\alpha \in (0, 1)$. As before, (C6a) holds due to $\rho_Q > 2$. For $\beta = 0$, we get $q_{DFM} = \gamma \in [\frac{1}{4}, \frac{3}{4})$. We set $\gamma = \frac{3}{4} - \varepsilon$ and $\alpha = 1 - \varepsilon$ with $\varepsilon \in (0, \frac{1}{10})$. Again, we compute

$$2\gamma\rho_A\alpha\rho_Q + \gamma\rho_A - \alpha\rho_Q > 0 \quad \text{ and } \quad 2\gamma\rho_A\alpha\rho_Q + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0.$$

Moreover, we get $\gamma\rho_A > \rho_Q(\alpha - \gamma\rho_A)$ and $\frac{\alpha\rho_Q\gamma\rho_A}{2\gamma\rho_A\alpha\rho_Q + \gamma\rho_A - \alpha\rho_Q} < \frac{9}{14} < q$. Finally, we compute $q > \alpha(2q - 1)$ and $\alpha\rho_Q > \gamma\rho_A$. This parameter constellation suggests, see Table 4.1, that the DFM1 converges with the highest effective order of convergence, given in (4.27).

For both DFM1 and DFM2, we get $M = N^2$, $K = N^{\frac{1}{2}}$, and error(DFM1) $= \mathcal{O}(\bar{c}^{-\frac{3}{8}+\varepsilon})$, error(DFM2) $= \mathcal{O}(\bar{c}^{-\frac{6}{17}+\varepsilon})$ for some $\varepsilon > 0$. For the exponential Euler scheme, we have $q_{EES} = \frac{1}{2}$ and obtain $M = N^3$, $K = N^{\frac{1}{2}}$. The order of convergence equals error(EES) $= \mathcal{O}(\bar{c}^{-\frac{1}{3}+\varepsilon})$.

For this example, we employ an approximation computed by the linear implicit Euler with $N_X = 2^6$, $K_X = 2^4$, and $M_X = 2^{18}$ instead of the exact solution. In Table 4.3 and Figure 4.2, we observe that the predicted effective orders of convergence of the derivative-free Milstein schemes and the exponential Euler scheme are outperformed. A reason might be that the estimates in Section 3.6, leading to the choice of the parameters, are not sharp. Moreover, the DFM2 involves a higher computational effort than the DFM1 which leads to the lower effective order of convergence.

		DFM1			DFM2		
N	M	CC	Error	Std	CC	Error	Std
2	4	56	$3.0 \cdot 10^{-2}$	$2.7 \cdot 10^{-3}$	64	$3.0 \cdot 10^{-2}$	$2.7 \cdot 10^{-3}$
4	2^4	448	$2.5 \cdot 10^{-2}$	$5.3 \cdot 10^{-4}$	475	$2.5 \cdot 10^{-2}$	$5.3 \cdot 10^{-4}$
8	2^6	6144	$1.7 \cdot 10^{-2}$	$6.6 \cdot 10^{-5}$	7430	$1.7 \cdot 10^{-2}$	$6.6 \cdot 10^{-5}$
16	2^8	73728	$6.3 \cdot 10^{-3}$	$3.0 \cdot 10^{-5}$	98304	$6.3 \cdot 10^{-3}$	$3.0 \cdot 10^{-5}$
32	2^{10}	1081344	$1.7 \cdot 10^{-3}$	$2.6 \cdot 10^{-5}$	1866831	$1.7 \cdot 10^{-3}$	$2.6 \cdot 10^{-5}$

		Linear Implicit Euler			Exponential Euler		
N	M	CC	Error	Std	CC	Error	Std
2	2^3	32	$2.1 \cdot 10^{-2}$	$2.8 \cdot 10^{-3}$	32	$2.1 \cdot 10^{-2}$	$2.9 \cdot 10^{-3}$
4	2^6	512	$2.6 \cdot 10^{-2}$	$4.2 \cdot 10^{-4}$	512	$2.7 \cdot 10^{-2}$	$4.5 \cdot 10^{-4}$
8	2^9	12288	$1.7 \cdot 10^{-2}$	$1.5 \cdot 10^{-4}$	12288	$2.0 \cdot 10^{-2}$	$1.7 \cdot 10^{-4}$
16	2^{12}	262144	$6.2 \cdot 10^{-3}$	$3.5 \cdot 10^{-5}$	262144	$6.4 \cdot 10^{-3}$	$3.8 \cdot 10^{-5}$
32	2^{15}	6291456	$1.6 \cdot 10^{-3}$	$3.2 \cdot 10^{-5}$	6291456	$1.7 \cdot 10^{-3}$	$3.6 \cdot 10^{-5}$

Table 4.3: Error and standard deviation for Example 2 obtained from 500 paths. CC denotes the computational cost computed as $CC(\text{DFM1}) = 3MNK + M^{2q}\frac{K(K-1)}{2}$, $CC(\text{DFM2}) = 3MNK + M^{q+\frac{1}{2}} K^{\frac{\rho_Q}{2}} \frac{K(K-1)}{2}$, and $CC(\text{EES}) = CC(\text{LIE}) = MNK$.

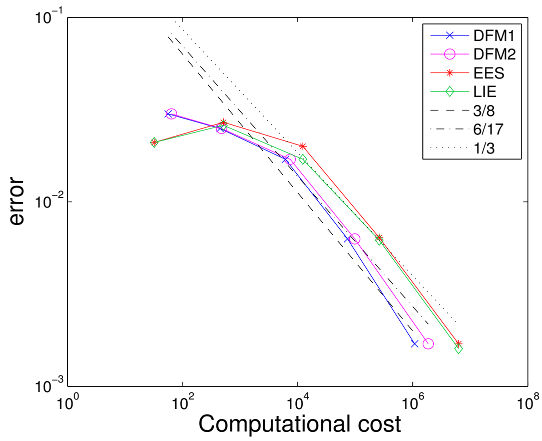

Figure 4.2: Error against computational cost for Example 2 for 500 paths and $N = 2, 4, 8, 16, 32$ in log-log scale.

Example 3 - Differing Effective Orders

Finally, we choose $\rho_Q = 4$ and $p = 4$. This yields $\alpha \in (0, \frac{7}{8})$, $\delta \in (0, \frac{1}{2})$, and $q_{DFM} = \gamma \in [\frac{1}{2}, 1)$ for $\beta = 0$. Here, we set $\gamma = 1 - \varepsilon$ and $\alpha = \frac{7}{8} - \varepsilon$ for some $\varepsilon \in (0, \frac{1}{4})$.
We compute the following expressions to determine the optimal scheme in this setting

$$2\gamma\rho_A\alpha\rho_Q + \gamma\rho_A - \alpha\rho_Q > 0, \quad 2\gamma\rho_A\alpha\rho_Q + \rho_Q\gamma\rho_A + 2(\gamma\rho_A - \alpha\rho_Q) > 0,$$

and

$$\frac{\alpha\rho_Q\gamma\rho_A}{2\gamma\rho_A\alpha\rho_Q + \gamma\rho_A - \alpha\rho_Q} < \frac{14}{27} < q.$$

Moreover, we compute $\gamma\rho_A > \rho_Q(\alpha - \gamma\rho_A)$, $\alpha(2q-1) < q$, and $\alpha\rho_Q > \gamma\rho_A$. By Table 4.1, we identify the DFM1 as the optimal scheme.

For this parameter constellation, the effective order of convergence of the schemes DFM1 and DFM2 is given in (4.27) and (4.29), respectively. This yields $M = N^2$, $K = N^{\frac{4}{7}}$ for the DFM2 and error(DFM2) $= \mathcal{O}(\bar{c}^{-\frac{14}{37}+\varepsilon})$ for some $\varepsilon > 0$. For the DFM1, however, we get $M = N^2$, $K = N^{\frac{4}{7}}$ and error(DFM1) $= \mathcal{O}(\bar{c}^{-\frac{7}{18}+\varepsilon})$. Finally, for the exponential Euler scheme, we choose $M = N^4$, $K = N^{\frac{4}{7}}$ and obtain error(EES) $= \mathcal{O}(\bar{c}^{-\frac{14}{39}+\varepsilon})$.
Here, the exact solution is replaced by an approximation computed with the linear implicit Euler scheme for $N_X = 2^6$, $K_X = 2^{\frac{24}{7}}$, and $M_X = 2^{20}$.

		DFM1			DFM2		
N	M	CC	Error	Std	CC	Error	Std
2	4	64	$3.0 \cdot 10^{-2}$	$2.3 \cdot 10^{-3}$	80	$3.0 \cdot 10^{-2}$	$2.3 \cdot 10^{-3}$
4	2^4	1344	$2.5 \cdot 10^{-2}$	$3.1 \cdot 10^{-4}$	2304	$2.5 \cdot 10^{-2}$	$3.1 \cdot 10^{-4}$
8	2^6	30720	$1.7 \cdot 10^{-2}$	$4.6 \cdot 10^{-5}$	55296	$1.7 \cdot 10^{-2}$	$4.6 \cdot 10^{-5}$
16	2^8	716800	$6.3 \cdot 10^{-3}$	$9.3 \cdot 10^{-6}$	1085440	$6.3 \cdot 10^{-3}$	$9.3 \cdot 10^{-6}$
32	2^{10}	30146560	$1.6 \cdot 10^{-3}$	$9.0 \cdot 10^{-6}$	59506688	$1.6 \cdot 10^{-3}$	$9.0 \cdot 10^{-6}$

		Linear Implicit Euler			Exponential Euler		
N	M	CC	Error	Std	CC	Error	Std
2	2^4	64	$2.2 \cdot 10^{-2}$	$4.7 \cdot 10^{-3}$	64	$2.3 \cdot 10^{-2}$	$4.9 \cdot 10^{-3}$
4	2^8	3072	$2.7 \cdot 10^{-2}$	$4.8 \cdot 10^{-4}$	3072	$2.7 \cdot 10^{-2}$	$5.0 \cdot 10^{-4}$
8	2^{12}	131072	$1.7 \cdot 10^{-2}$	$8.3 \cdot 10^{-5}$	131072	$1.7 \cdot 10^{-2}$	$8.5 \cdot 10^{-5}$
16	2^{16}	5242880	$6.1 \cdot 10^{-3}$	$1.8 \cdot 10^{-5}$	5242880	$6.1 \cdot 10^{-3}$	$1.8 \cdot 10^{-5}$
32	2^{20}	268435456	$1.5 \cdot 10^{-3}$	$1.6 \cdot 10^{-6}$	268435456	$1.5 \cdot 10^{-3}$	$1.7 \cdot 10^{-6}$

Table 4.4: Error and standard deviation for Example 3 obtained from 500 paths. CC denotes the computational cost computed as $CC(\text{DFM1}) = 3MNK + M^{2q}\frac{K(K-1)}{2}$, $CC(\text{DFM2}) = 3MNK + M^{q+\frac{1}{2}}K^{\frac{\rho_Q}{2}}\frac{K(K-1)}{2}$, and $CC(\text{EES}) = CC(\text{LIE}) = MNK$.

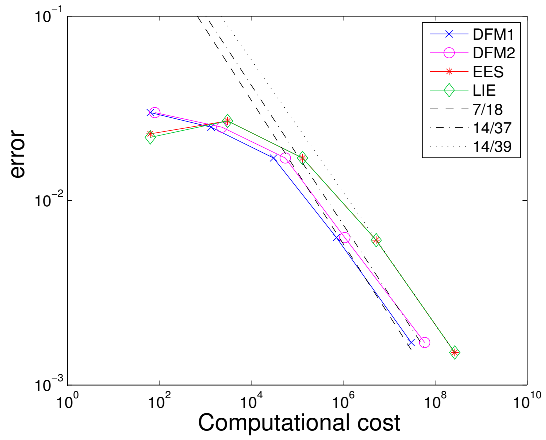

Figure 4.3: Error against computational cost for Example 3 for 500 paths and $N = 2, 4, 8, 16, 32$ in log-log scale.

In Examples 1, 2, and 3, the derivative-free Milstein scheme combined with Algorithm 1 or Algorithm 2, respectively, obtains a higher effective order of convergence than the exponential Euler scheme. This confirms the theoretical analysis that we conducted in Section 4.1 to Section 4.3. In our examples, we selected the maximal value for α and γ, which yields $\alpha\rho_Q = \frac{\rho_Q+3}{2} - \varepsilon\rho_Q \approx \frac{\rho_Q+3}{2} > \frac{5}{2}$. On the other hand, $\gamma\rho_A < \rho_A = 2$. Therefore, in this kind of setting with A denoting the Laplacian, the DFM1 always outperforms the exponential Euler scheme, see Table 4.1. Concerning the differences between the schemes DFM1 and DFM2, one has to keep in mind that we did not include the additional computational effort that arises if the matrix $\sqrt{\Sigma^{(\infty)}}$ is computed by a Cholesky decomposition.

5
From Local to Global Error Estimates

In the error analyses in the last chapters, we computed global error estimates for the approximation schemes directly, see Section 3.5 and Section 4.3. A different approach to obtain these estimates is the estimation of the local error, from which the global error can be inferred.

Let $T \in (0, \infty)$, denote by $(X_t)_{t\in[0,T]}$ the real-valued exact solution of some differential equation and by $(Y_m)_{m\in\{0,\dots,M\}}$, $M \in \mathbb{N}$, the approximation obtained with some numerical scheme.
For ODEs and SODEs, there exist well known results on the connection between local errors, that is, the error that results by conducting one step with a numerical scheme, and the global error of $X_T - Y_M$ for some fixed $M \in \mathbb{N}$. The definitions of the local errors e_m, $m \in \{1, \dots, M\}$, and the global error depend on the type of differential equation that is considered. For SODEs, we compute the mean-square error, for example.

For ODEs, the order of convergence is reduced by one if the global instead of the local error is considered, that is, for some $p > 1$, $h := \max_{m\in\{1,\dots,M\}} h_m$, and $m \in \{1, \dots, M\}$, $M \in \mathbb{N}$, we have

$$e_m \leq C h_m^{p+1} \qquad \Rightarrow \qquad |X_T - Y_M| \leq \tilde{C} h^p,$$

see [23, Chapter II, Theorem 3.4, Theorem 3.6] for the assumptions that have to be fulfilled.

For SODEs, however, it has been shown by Milstein, [50, Theorem 1.1], that the strong order of convergence $p > \frac{1}{2}$ is reduced by $\frac{1}{2}$ only when passing from the local to the global error, that is,

$$e_m \leq C h^{p+1} \qquad \Rightarrow \qquad \big(\mathrm{E}\big[|X_T - Y_M|^2\big]\big)^{\frac{1}{2}} \leq \tilde{C} h^{p+\frac{1}{2}}.$$

In [8], Chen and Hong analyzed semi-discrete schemes for SPDEs of type

$$Y_{m+1} = \hat{S}(h)Y_m + \Phi(Y_m, \Delta W_m^M, h)$$

for all $m \in \{0, \dots, M-1\}$, $M \in \mathbb{N}$, $h > 0$. They showed that the difference in the order of convergence is influenced by the approximation of the semigroup e^{Ah} by some operator $\hat{S}(h)$ with $\|e^{Ah} - \hat{S}(h)\|_{L(H)} \leq Ch^r$ for $r > 0$, $h \geq 0$, as well. For a local error of order $p \geq \frac{1}{2}$, computed in $\| \cdot \|_{H_\alpha}$ for some $\alpha \geq 0$ determined by their setting, they proved

$$\mathrm{E}\big[\|X_T - Y_M\|_{H_\alpha}^2\big] \leq Ch^{\min(r, p-\frac{1}{2})}$$

in [8, Theorem 4.2].

In [14] and [32, p.157,158], however, the authors observed the same order of convergence for the local and the global error in numerical simulations. In the following, we show that this discovery can be proved for numerical schemes of general type

$$Y_{m+1}^{N,M,K} = P_N S(h) Y_m^{N,M,K} + \Phi(Y_m^{N,M,K}, \Delta W_m^{K,M}, h)$$

for all $m \in \{0, \dots, M-1\}$, $M, N, K \in \mathbb{N}$ and all $h \in [h_{\min}, 1)$ with $h_{\min} > 0$ if we assume the setting outlined in Section 3.1. The result can be obtained for semigroups $S(t)$, $t \geq 0$, with $\|S(t)\|_{L(H)} < 1$ for all $t > h_{\min}$. In this case, the semigroup inhibits the accumulation of local errors.

Theorem 5.1 (Local and Global Error Estimates)
Let $M, N, K \in \mathbb{N}$ be arbitrarily fixed. Assume that the local error fulfills

$$e_m \leq Ch^p + \Big(\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i\Big)^{-\gamma} + \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j\Big)^{\alpha} \tag{5.1}$$

for some $p > 0$, all $m \in \{1, \dots, M\}$, and $\|S(t)\|_{L(H)} < 1$ for all $t \in (0, T]$. Furthermore, we require the numerical scheme to fulfill

$$\mathrm{E}\big[\|\Phi(w, u, h) - \Phi(y, u, h)\|_H^2\big] \leq Ch\mathrm{E}\big[\|w - y\|_H^2\big] \tag{5.2}$$

for all $h > 0$, $w, y \in H_N$, $u \in V_K$. Then, there exists a constant $C_{T,S} > 0$, independent of M, N, K, such that we obtain the global error for all $h \in [h_{\min}, 1)$, $h_{\min} > 0$, as

$$\Big(\mathrm{E}\big[\|X_T - Y_M^{N,M,K}\|_H^2\big]\Big)^{\frac{1}{2}} \leq C_{T,S}\Big(Ch^p + \Big(\inf_{i \in \mathcal{I} \setminus \mathcal{I}_N} \lambda_i\Big)^{-\gamma} + \Big(\sup_{j \in \mathcal{J} \setminus \mathcal{J}_K} \eta_j\Big)^{\alpha}\Big).$$

Proof.
We fix $M_S := \|S(h_{\min})\|_{L(H)} < 1$. Denote by $X_{t,(s,x)}$ the solution process at time $t \in [0,T]$ starting in $X_s = x$ for $0 \leq s < t$, $x \in H_\gamma$. $Y_{t,(s,x)}$, $s, t \in [0, T]$, $s < t$, is to be understood

analogously.
First, we rewrite the global error such that we obtain a dependence on the local error $e_m = \big(\mathrm{E}\big[\|X_{t_m} - Y_{m,(m-1,X_{t_{m-1}})}\|_H^2\big]\big)^{\frac{1}{2}}$, $m \in \{1,\dots,M\}$, $M \in \mathbb{N}$. Let $a > 0$, then we obtain by Young's inequality

$$
\begin{aligned}
&\mathrm{E}\big[\|X_T - Y_M\|_H^2\big] \\
&= \mathrm{E}\big[\|X_T - Y_{M,(M-1,X_{t_{M-1}})} + Y_{M,(M-1,X_{t_{M-1}})} - Y_M\|_H^2\big] \\
&\le (1+a)\mathrm{E}\big[\|X_T - Y_{M,(M-1,X_{t_{M-1}})}\|_H^2\big] + \big(1+\frac{1}{a}\big)\mathrm{E}\big[\|Y_{M,(M-1,X_{t_{M-1}})} - Y_M\|_H^2\big] \\
&\le (1+a)e_M^2 \\
&\quad + \big(1+\frac{1}{a}\big)\mathrm{E}\big[\|e^{Ah}(X_{t_{M-1}} - Y_{M-1}) + \Phi(X_{t_{M-1}}, \Delta W_{M-1}^{K,M}, h) - \Phi(Y_{M-1}, \Delta W_{M-1}^{K,M}, h)\|_H^2\big]
\end{aligned}
$$

for all $M, K \in \mathbb{N}$ such that $\frac{T}{M} \ge h_{\min}$.
Further, we obtain with Young's inequality for $a_1 > 0$ and assumption (5.2)

$$
\begin{aligned}
\mathrm{E}\big[\|X_T - Y_M\|_H^2\big] \le &(1+a)e_M^2 + \big(1+\frac{1}{a}\big)\big(1+\frac{1}{a_1}\big)\mathrm{E}\big[\|e^{Ah}(X_{t_{M-1}} - Y_{M-1})\|_H^2\big] \\
&+ \big(1+\frac{1}{a}\big)(1+a_1)\mathrm{E}\Big[\big\|\Phi(X_{t_{M-1}}, \Delta W_{M-1}^{K,M}, h) - \Phi(Y_{M-1}, \Delta W_{M-1}^{K,M}, h)\big\|_H^2\Big] \\
\le &(1+a)e_M^2 + \big(1+\frac{1}{a}\big)\big(1+\frac{1}{a_1}\big)\mathrm{E}\big[\|e^{Ah}(X_{t_{M-1}} - Y_{M-1})\|_H^2\big] \\
&+ \big(1+\frac{1}{a}\big)(1+a_1)Ch\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big]
\end{aligned}
$$

for all $M, K \in \mathbb{N}$ such that $\frac{T}{M} \ge h_{\min}$.
For legibility, let $c_1 := \big(1+\frac{1}{a}\big)\big(1+\frac{1}{a_1}\big)$ and $c_2 := \big(1+\frac{1}{a}\big)(1+a_1)$. Therewith, we get for all $M, K \in \mathbb{N}$ with $\frac{T}{M} \ge h_{\min}$ the estimate

$$
\mathrm{E}\big[\|X_T - Y_M\|_H^2\big] \le (1+a)e_M^2 + c_1 M_S^2 \mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big] + c_2 Ch\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big].
$$

Now, we replace $\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big]$ in the second term inductively. This yields

$$
\begin{aligned}
\mathrm{E}\big[\|X_T - Y_M\|_H^2\big] &\le (1+a)e_M^2 + c_1 M_S^2\Big((1+a)e_{M-1}^2 + c_1 M_S^2\mathrm{E}\big[\|X_{t_{M-2}} - Y_{M-2}\|_H^2\big] \\
&\quad + c_2 Ch\mathrm{E}\big[\|X_{t_{M-2}} - Y_{M-2}\|_H^2\big]\Big) + c_2 Ch\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big] \\
&= (1+a)e_M^2 + c_1 M_S^2(1+a)e_{M-1}^2 + c_1^2 M_S^4\mathrm{E}\big[\|X_{t_{M-2}} - Y_{M-2}\|_H^2\big] \\
&\quad + c_1 M_S^2 c_2 Ch\mathrm{E}\big[\|X_{t_{M-2}} - Y_{M-2}\|_H^2\big] + c_2 Ch\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big] \\
&\le (1+a)e_M^2 + c_1 M_S^2(1+a)e_{M-1}^2 + c_1^2 M_S^4\Big((1+a)e_{M-2}^2 \\
&\quad + c_1 M_S^2\mathrm{E}\big[\|X_{t_{M-3}} - Y_{M-3}\|_H^2\big] + c_2 Ch\mathrm{E}\big[\|X_{t_{M-3}} - Y_{M-3}\|_H^2\big]\Big) \\
&\quad + c_1 M_S^2 c_2 Ch\mathrm{E}\big[\|X_{t_{M-2}} - Y_{M-2}\|_H^2\big] + c_2 Ch\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big] \\
&= (1+a)e_M^2 + c_1 M_S^2(1+a)e_{M-1}^2 + c_1^2 M_S^4(1+a)e_{M-2}^2 \\
&\quad + c_1^3 M_S^6\mathrm{E}\big[\|X_{t_{M-3}} - Y_{M-3}\|_H^2\big] + c_1^2 M_S^4 c_2 Ch\mathrm{E}\big[\|X_{t_{M-3}} - Y_{M-3}\|_H^2\big] \\
&\quad + c_1 M_S^2 c_2 Ch\mathrm{E}\big[\|X_{t_{M-2}} - Y_{M-2}\|_H^2\big] + c_2 Ch\mathrm{E}\big[\|X_{t_{M-1}} - Y_{M-1}\|_H^2\big]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \ldots \leq (1+a)e_M^2 + (1+a)\sum_{i=0}^{M-1} c_1^{i+1} M_S^{2(i+1)} e_{M-1-i}^2 \\
&\qquad + c_2 Ch \sum_{i=0}^{M-1} c_1^i M_S^{2i} \mathrm{E}\big[\|X_{t_{M-1-i}} - Y_{M-1-i}\|_H^2\big]
\end{aligned}
$$

for all $M, K \in \mathbb{N}$ such that $\frac{T}{M} \geq h_{\min}$.

We need $c_1 M_S^2 < 1$ to hold in the following; this can be obtained by choosing $c_1 = (1+\frac{1}{a})(1+\frac{1}{a_1}) < \frac{1}{M_S^2}$ which is possible for all $h \in [h_{\min}, 1)$ as $M_S < 1$. Furthermore, as the local error e_m^2 does not depend on m for all $m \in \{0, \ldots, M\}$, $M \in \mathbb{N}$, see (5.1), we denote this term by $\mathcal{E}$. Therewith, we get by Gronwall's Lemma

$$
\begin{aligned}
\mathrm{E}\big[\|X_T - Y_M\|_H^2\big] &\leq (1+a)\mathcal{E} + (1+a)\mathcal{E}\sum_{i=0}^{M-1} c_1^{i+1} M_S^{2(i+1)} + c_2 Ch \sum_{i=0}^{M-1} \mathrm{E}\big[\|X_{t_{M-1-i}} - Y_{M-1-i}\|_H^2\big] \\
&\leq (1+a)\mathcal{E} + (1+a)\mathcal{E}\frac{1}{1 - c_1 M_S^2} + c_2 Ch \sum_{m=0}^{M-1} \mathrm{E}\big[\|X_{t_m} - Y_m\|_H^2\big] \\
&\leq C_{c_1, M_S}(1+a)\mathcal{E} e^{c_2 CT} \\
&\leq C_{T, M_S}\Big(Ch^{2p} + \Big(\inf_{i \in \mathcal{I}\setminus\mathcal{I}_N} \lambda_i\Big)^{-2\gamma} + \Big(\sup_{j \in \mathcal{J}\setminus\mathcal{J}_K} \eta_j\Big)^{2\alpha}\Big)
\end{aligned}
$$

for all $M, N, K \in \mathbb{N}$ such that $\frac{T}{M} \geq h_{\min}$. $\square$

Remark 5.1

The condition $\|S(t)\|_{L(H)} < 1$ is crucial for this estimate as the smoothing effect of the semigroup is the main difference compared to SODEs. However, we have to bound $h \geq h_{\min} > 0$ away from zero as $\|S(0)\|_{L(H)} = 1$.

6

Conclusion and Remarks

As indicated by their effective order of convergence, derivative-free numerical schemes of higher orders efficiently approximate the mild solution $(X_t)_{t\in[0,T]}$ to SPDEs of type (1.1). This rate is determined by combining the theoretical order with the computational cost (CC) according to

$$\min_{N,M,K}\Big(\sup_{m\in\{0,\dots,M\}}\mathrm{E}\Big[\big\|X_{t_m}-Y_m^{M,N,K}\big\|_H^2\Big]\Big)^{\frac{1}{2}}\quad\text{such that}\quad \mathrm{CC}=\bar{c}$$

for some $\bar{c}>0$. The optimization problem yields a convergence rate in terms of the computational cost $\bar{c}$; it is this parameter that actually determines the scheme that is superior with respect to the overall computational cost. We derived this value for various approximation schemes and showed that for general equations the effective order of convergence is higher for the derivative-free Milstein schemes than for the Milstein or the exponential Euler scheme. We developed differing schemes for equations that are commutative and equations that do not fulfill this assumption. For commutative equations, the simulation of the scheme is straightforward due to expression (3.11), and the results are summarized in Table 3.2. If the SPDE is not commutative, we cannot rewrite the iterated stochastic integrals in terms of increments of the Q-Wiener process. Therefore, the approximation of these SPDEs requires the simulation of these integrals. We presented two algorithms to approximate iterated stochastic integrals of type

$$\int_s^t B'(X_s)\Big(\int_s^r B(X_s)\,\mathrm{d}W_u^K\Big)\,\mathrm{d}W_r^K$$

for $s,t\in[0,T]$, $s\leq t$, $K\in\mathbb{N}$. For details on Algorithm 1 or Algorithm 2, we refer to page 86 and page 89. The approximation schemes that we derived obtain higher theoretical as well as, for most parameter constellations, effective orders of convergence than the exponential Euler scheme, see Theorem 4.3 and Table 4.1. Algorithm 2, however, could be improved in terms of the computational effort if one obtained a closed form for the matrix $\sqrt{\Sigma^{(\infty)}}$, see equation (4.24). In this case, there would be no need for a numerical algorithm to compute the Cholesky decomposition of $\Sigma^{(\infty)}$.

We proved estimates on the global error of the approximation methods directly. A different approach to show the convergence of the numerical schemes is the estimation of the local error. In a second step, the global error can then be inferred from this estimate. For ODEs and SODEs, there exist universal results on the connection of these two error terms, see [23] and [50]. We stated some ideas on the relation of local and global error estimates for SPDEs; these are in line with empirical results from [32]. However, we are restricted to a step size $h > h_{\min}$, $h_{\min} > 0$. So far, there exists no general result on this relationship which is consistent with the numerical findings.

As the focus of this work was on the strong convergence of the numerical methods, it remains to analyze the weak convergence of derivative-free approximation schemes. There exists literature on the weak approximation of SPDEs, for example, [10, 15, 16, 29, 72], but there are no results on derivative-free schemes so far.
Further research could focus on the systematic derivation of higher order schemes that are free of derivatives as well. That is, one could investigate how to transfer the general idea of Runge-Kutta schemes from other types of differential equations to SPDEs. In view of this, [32] is to be mentioned. In this work, Jentzen and Kloeden derived Taylor expansions of arbitrary order for the mild solution of SPDEs. Moreover, Hochbruck und Ostermann developed exponential Runge-Kutta schemes for parabolic PDEs in [27]. These works constitute a promising basis for research in this direction.

Notation

$(H, \langle \cdot, \cdot \rangle_H),\ (V, \langle \cdot, \cdot \rangle_V)$ are real separable Hilbert spaces

$L(V, H) := \big\{ B : V \to H \,|\, B$ is linear and bounded$\big\}$

$L(V) := L(V, V)$

$H_r := D((-A)^r),\ r \in [0, \infty)$ *p.*14

$L_{HS}(V, H) := \Big\{ B \in L(V) : \Big(\sum_{j \in \mathbb{N}} \|B\tilde{e}_j\|_H^2 \Big)^{\frac{1}{2}} < \infty,\ \tilde{e}_j,\ j \in \mathbb{N},$ basis of $V \Big\}$ *p.*18

$V_0 := Q^{\frac{1}{2}} V$ *p.*19

$\mathcal{E} := \big\{ (\Psi_t)_{t \in [0,T]} \in L(V, H) : (\Psi_t)_{t \in [0,T]}$ elementary$\big\}$ *p.*19

$$\mathcal{N}_W^2(0, T; H) = \Big\{ Y : [0, T] \times \Omega \to L_{HS}(V_0, H) \mid Y \text{ is } \mathcal{P}_T - \mathcal{B}(L_{HS}(V_0, H)) - \text{measurable and } \mathrm{E}\Big[\int_0^T \|Y_s\|_{L_{HS}(V_0,H)}^2 \, \mathrm{d}s \Big] < \infty \Big\}$$ *p.*19

$$\mathcal{N}_W(0, T; H) = \Big\{ Y : [0, T] \times \Omega \to L_{HS}(V_0, H) \mid Y \text{ is } \mathcal{P}_T - \mathcal{B}(L_{HS}(V_0, H)) - \text{measurable and } P\Big(\int_0^T \|Y_s\|_{L_{HS}(V_0,H)}^2 \, \mathrm{d}s < \infty \Big) = 1 \Big\}$$ *p.*20

$L(V, H)_0 := \big\{ \left. T \right|_{V_0} \mid T \in L(V, H) \big\}$ *p.*34

$L^{(2)}(H, L(V, H)) = L(H, L(H, L(V, H))$ *p.*34

$L_{HS}^{(2)}(V_0, H) = L_{HS}(V_0, L_{HS}(V_0, H))$ *p.*34

List of Figures

List of Tables

Bibliography

[1] I. A. Adamu and G. J. Lord. Numerical approximation of multiplicative SPDEs. *Int. J. Comput. Math.*, 89(18):2603–2621, 2012.

[2] R. Anton, D. Cohen, S. Larsson, and X. Wang. Full discretization of semilinear stochastic wave equations driven by multiplicative noise. *SIAM J. Numer. Anal.*, 54(2):1093–1119, 2016.

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

[4] A. Barth and A. Lang. L^p and almost sure convergence of a Milstein scheme for stochastic partial differential equations. *Stochastic Process. Appl.*, 123(5):1563–1587, 2013.

[5] S. Becker, A. Jentzen, and P. E. Kloeden. An exponential Wagner-Platen type scheme for SPDEs. *ArXiv e-prints*, September 2013.

[6] D. Blömker, M. Kamrani, and S. M. Hosseini. Full discretization of the stochastic Burgers equation with correlated noise. *IMA J. Numer. Anal.*, 33(3):825–848, 2013.

[7] Z. Brzeźniak, J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Itô's formula in UMD Banach spaces and regularity of solutions of the Zakai equation. *J. Differential Equations*, 245(1):30–58, 2008.

[8] C. Chen and J. Hong. Symplectic Runge-Kutta semi-discretization for stochastic Schrödinger equation. *ArXiv e-prints*, October 2014.

[9] P.-L. Chow. *Stochastic partial differential equations*. Advances in Applied Mathematics. CRC Press, Boca Raton, FL, second edition, 2015.

[10] D. Conus, A. Jentzen, and R. Kurniawan. Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients. *ArXiv e-prints*, August 2014.

[11] E. J. Coutts and G. J. Lord. Effects of noise on models of spiny dendrites. *J. Comput. Neurosci.*, 34(2):245–257, 2013.

[12] G. Da Prato, A. Jentzen, and M. Roeckner. A mild Itô formula for SPDEs. *ArXiv e-prints*, September 2010.

[13] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[14] A. M. Davie and J. G. Gaines. Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. *Math. Comp.*, 70(233):121–134, 2001.

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

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

[17] P. Deuflhard and A. Hohmann. *Numerical analysis in modern scientific computing: an introduction*, volume 43 of *Texts in Applied Mathematics*. Springer, New York, second edition, 2003.

[18] R. Fitzhugh. Thresholds and plateaus in the Hodgkin-Huxley nerve equations. *J. Gen. Physiol.*, 43(5):867–896, 1960.

[19] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, 1(6):445–466, 1961.

[20] B. Ghayebi and S. M. Hosseini. A simplified Milstein scheme for SPDEs with multiplicative noise. *Abstr. Appl. Anal.*, 2014.

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

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

[23] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I: nonstiff problems*, volume 8 of *Springer Series in Computational Mathematics*. Springer, Berlin, second edition, 1993.

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

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

[26] E. Hausenblas. Finite element approximation of stochastic partial differential equations driven by Poisson random measures of jump type. *SIAM J. Numer. Anal.*, 46(1):437–471, 2008.

[27] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numer.*, 19:209–286, 2010.

[28] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117(4):500–544, 1952.

[29] L. Jacobe de Naurois, A. Jentzen, and T. Welti. Weak convergence rates for spatial spectral Galerkin approximations of semilinear stochastic wave equations with multiplicative noise. *ArXiv e-prints*, August 2015.

[30] A. Jentzen, P. Kloeden, and G. Winkel. Efficient simulation of nonlinear parabolic SPDEs with additive noise. *Ann. Appl. Probab.*, 21(3):908–950, 2011.

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

[32] A. Jentzen and P. E. Kloeden. *Taylor approximations for stochastic partial differential equations*, volume 83 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, PA, 2011.

[33] A. Jentzen and M. Röckner. A Milstein scheme for SPDEs. *ArXiv e-prints*, January 2012.

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

[35] A. Jentzen and M. Röckner. A Milstein scheme for SPDEs. *Found. Comput. Math.*, 15(2):313–362, 2015.

[36] G. Kallianpur and J. Xiong. *Stochastic differential equations in infinite-dimensional spaces*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 26. Institute of Mathematical Statistics, Hayward, CA, 1995. Expanded version of the lectures delivered as part of the 1993 Barrett Lectures at the University of Tennessee, Knoxville, TN.

[37] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer, New York, 1995.

[38] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer, Berlin, second corrected printing edition, 1995.

[39] P. E. Kloeden, E. Platen, and I. W. Wright. The approximation of multiple stochastic integrals. *Stoch. Anal. Appl.*, 10(4):431–441, 1992.

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

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

[42] R. Kruse and S. Larsson. Optimal regularity for semilinear stochastic partial differential equations with multiplicative noise. *Electron. J. Probab.*, 17(65):1–19, 2012.

[43] M. C. Kunze. Stochastic reaction-diffusion systems with Hölder continuous multiplicative noise. *Stoch. Anal. Appl.*, 33(2):331–355, 2015.

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

[45] A. Lang, P.-L. Chow, and J. Potthoff. Erratum: Almost sure convergence of a semi-discrete Milstein scheme for SPDEs of Zakai type. *Stochastics*, 84(4):561, 2012.

[46] C. Leonhard and A. Rößler. An efficient derivative-free Milstein scheme for stochastic partial differential equations with commutative noise. *ArXiv e-prints*, September 2015.

[47] W. Liu and M. Röckner. *Stochastic partial differential equations: an introduction.* Universitext. Springer, Cham, 2015.

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

[49] G. J. Lord and A. Tambue. Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise. *IMA J. Numer. Anal.*, 33(2):515–543, 2013.

[50] G. N. Milstein. *Numerical integration of stochastic differential equations*, volume 313 of *Mathematics and its Applications.* Kluwer Academic Publishers, Dordrecht, 1995. Translated and revised from the 1988 Russian original.

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

[52] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.

[53] É. Pardoux. Sur des équations aux dérivées partielles stochastiques monotones. *C. R. Acad. Sci. Paris Sér. A-B*, 275:A101–A103, 1972.

[54] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences.* Springer, New York, 1983.

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

[56] M. Reed and B. Simon. *Methods of modern mathematical physics. I: Functional analysis.* Academic Press, Inc., New York, second edition, 1980.

[57] M. Renardy and R. C. Rogers. *An introduction to partial differential equations*, volume 13 of *Texts in Applied Mathematics.* Springer, New York, first edition, 1993.

[58] M. Röckner and X. Zhang. Stochastic tamed 3D Navier-Stokes equations: existence, uniqueness and ergodicity. *Probab. Theory Related Fields*, 145(1-2):211–267, 2009.

[59] A. Rößler. Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Stoch. Anal. Appl.*, 24(1):97–134, 2006.

[60] A. Rößler. Second order Runge-Kutta methods for Stratonovich stochastic differential equations. *BIT*, 47(3):657–680, 2007.

[61] A. Rößler. Second order Runge-Kutta methods for Itô stochastic differential equations. *SIAM J. Numer. Anal.*, 47(3):1713–1738, 2009.

[62] A. Rößler. Runge-Kutta methods for the strong approximation of solutions of stochastic differential equations. *SIAM J. Numer. Anal.*, 48(3):922–952, 2010.

[63] B. L. Rozovskii. *Stochastic evolution systems: linear theory and applications to non-linear filtering*, volume 35 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990.

[64] M. Sauer and W. Stannat. Analysis and approximation of stochastic nerve axon equations. *Math. Comp.*, 2016.

[65] G. R. Sell and Y. You. *Dynamics of evolutionary equations*, volume 143 of *Applied Mathematical Sciences*. Springer, New York, 2002.

[66] W. Stannat. Stability of travelling waves in stochastic Nagumo equations. *ArXiv e-prints*, January 2013.

[67] H. C. Tuckwell. Analytical and simulation results for the stochastic spatial FitzHugh-Nagumo model neuron. *Neural Comput.*, 20(12):3003–3033, 2008.

[68] H. C. Tuckwell. Stochastic partial differential equations in neurobiology: linear and nonlinear models for spiking neurons. In *Stochastic biomathematical models*, volume 2058 of *Lecture Notes in Math.*, pages 149–173. Springer, Berlin, 2013.

[69] J. B. Walsh. An introduction to stochastic partial differential equations. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 265–439. Springer, Berlin, 1986.

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

[71] X. Wang and S. Gan. A Runge-Kutta type scheme for nonlinear stochastic partial differential equations with multiplicative trace class noise. *Numer. Algorithms*, 62(2):193–223, 2013.

[72] X. Wang and S. Gan. Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise. *J. Math. Anal. Appl.*, 398(1):151–169, 2013.

[73] D. Werner. *Funktionalanalysis.* Springer, Berlin, third edition, 2000.

[74] A. G. Werschulz. *The computational complexity of differential and integral equations: An information-based approach.* Oxford Mathematical Monographs. Oxford University Press, New York, 1991.

[75] M. Wiktorsson. Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions. *Ann. Appl. Probab.*, 11(2):470–487, 2001.

[76] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Stat.*, 36(5):1560–1564, 1965.

[77] E. Wong and M. Zakai. On the relation between ordinary and stochastic differential equations. *Internat. J. Engrg. Sci.*, 3:213–229, 1965.

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

[79] X. Zhang. Regularities for semilinear stochastic partial differential equations. *J. Funct. Anal.*, 249(2):454–476, 2007.