

# TAYLOR EXPANSIONS OF SOLUTIONS OF STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS WITH ADDITIVE NOISE<sup>1</sup>

BY ARNULF JENTZEN AND PETER KLOEDEN

*Johann Wolfgang Goethe University, Frankfurt am Main*

The solution of a parabolic stochastic partial differential equation (SPDE) driven by an infinite-dimensional Brownian motion is in general not a semi-martingale anymore and does in general not satisfy an Itô formula like the solution of a finite-dimensional stochastic ordinary differential equation (SODE). In particular, it is not possible to derive stochastic Taylor expansions as for the solution of a SODE using an iterated application of the Itô formula. Consequently, until recently, only low order numerical approximation results for such a SPDE have been available. Here, the fact that the solution of a SPDE driven by additive noise can be interpreted in the mild sense with integrals involving the exponential of the dominant linear operator in the SPDE provides an alternative approach for deriving stochastic Taylor expansions for the solution of such a SPDE. Essentially, the exponential factor has a mollifying effect and ensures that all integrals take values in the Hilbert space under consideration. The iteration of such integrals allows us to derive stochastic Taylor expansions of arbitrarily high order, which are robust in the sense that they also hold for other types of driving noise processes such as fractional Brownian motion. Combinatorial concepts of trees and woods provide a compact formulation of the Taylor expansions.

**1. Introduction.** Taylor expansions are a fundamental and repeatedly used method of approximation in mathematics, in particular in numerical analysis. Although numerical schemes for ordinary differential equations (ODEs) are often derived in an ad hoc manner, those based on Taylor expansions of the solution of an ODE, the Taylor schemes, provide a class of schemes with known convergence orders against which other schemes can be compared to determine their order. An important component of such Taylor schemes are the iterated total derivatives of the vector field corresponding higher derivatives of the solution, which are obtained via the chain rule; see [8].

An analogous situation holds for Itô stochastic ordinary differential equations (SODEs), except, due to the less robust nature of stochastic calculus, stochastic Taylor schemes instead of classical Taylor schemes are the starting point to obtain

---

Received September 2008; revised July 2009.

<sup>1</sup>Supported by the DFG project “Pathwise numerics and dynamics of stochastic evolution equations.”

*AMS 2000 subject classifications.* 35K90, 41A58, 65C30, 65M99.

*Key words and phrases.* Taylor expansions, stochastic partial differential equations, SPDE, strong convergence, stochastic trees.

consistent higher order numerical schemes, see [29] for the general theory. Another important difference is that SODEs are really just a symbolic representation of stochastic integral equations since their solutions are not differentiable, so an integral version of Taylor expansions based on iterated application of the stochastic chain rule, the Itô formula, is required. Underlying this method is the fact that the solution of a SODE is an Itô-process or, more generally, a semi-martingale and in particular of finite quadratic variation.

This approach fails, however, if a SODE is driven by an additive stochastic process with infinite quadratic variation such as a fractional Brownian motion, because the Itô formula is in general no longer valid. A new method to derive Taylor expansions in such cases was presented in [20, 23, 28]. It uses the smoothness of the coefficients, but only minimal assumptions on the nature of the driving stochastic process. The resulting Taylor expansions there are thus robust with respect to assumptions concerning the driving stochastic process and, in particular, remain valid for other noise processes.

A similar situation holds for stochastic partial differential equations (SPDEs). In this article, we consider parabolic SPDEs with additive noise of the form

$$(1) \quad dU_t = [AU_t + F(U_t)] dt + B dW_t, \quad U_0 = u_0, \quad t \in [0, T],$$

in a separable Hilbert space  $H$ , where  $A$  is an unbounded linear operator,  $F$  is a nonlinear smooth function,  $B$  is a bounded linear operator and  $W_t$ ,  $t \geq 0$ , is an infinite-dimensional Wiener process. (See Section 2 for a precise description of the equation above and the assumptions, we use.) Although the SPDE (1) is driven by a martingale Brownian motion, the solution process is with respect to a reasonable state space in general not a semi-martingale anymore (see [10] for a clear discussion of this problem) and except of special cases a general Itô formula does not exist for its solution (see, e.g., [10, 37]). Hence, stochastic Taylor expansions for the solution of the SPDE (1) cannot be derived as in [29] for the solutions of finite-dimensional SODEs. Consequently, until recently, only low order numerical approximation results for such SPDEs have been available (except for SPDEs with spatially smooth noise; see, e.g., [14]). For example, the stochastic convolution of the semigroup generated by the Laplacian with Dirichlet boundary conditions on the one-dimensional domain  $(0, 1)$  and a cylindrical  $I$ -Wiener process on  $H = L^2((0, 1), \mathbb{R})$  has sample paths which are only  $(\frac{1}{4} - \varepsilon)$ -Hölder continuous (see Section 5.4) and previously considered approximations such as the linear implicit Euler scheme or the linear implicit Crank–Nicolson scheme are not of higher temporal order. The reason is that the infinite-dimensional noise process has only minimal spatial regularity and the convolution of the semigroup and the noise is only as smooth in time as in space. This comparable regularity in time and space is a fundamental property of the dynamics of semigroups; see [41] or also [9], for example. To overcome these problems, we thus need to derive robust Taylor expansions for a SPDE of the form (1) driven by an infinite-dimensional Brownian motion.

An idea used in [24] to derive what was called the exponential Euler scheme for the SPDE (1), that has a better convergence rate than hitherto analyzed schemes, can be exploited here. It is based on the fact that the SPDE (1) can be understood in the mild sense, that is as an integral equation of the form

$$(2) \quad U_t = e^{At} u_0 + \int_0^t e^{A(t-s)} F(U_s) ds + \int_0^t e^{A(t-s)} B dW_s$$

for all  $t \in [0, T]$  rather than as an integral equation obtained by directly integrating the terms of the SPDE (1). (This mild integral equation form of the SPDE is considered in some detail in the monograph [6], (7.1) and (7.3.4), and in the monograph [37], (F.0.2).) The crucial point here is that all integrals in the mild integral equation (2) contain the exponential factor  $e^{A(t-s)}$  of the operator  $A$ , which acts in a sense as a mollifier and ensures that iterated versions of the terms remain in the Hilbert space  $H$ . The main idea of the Taylor expansions presented in this article is to use a classical Taylor expansion for the nonlinearity  $F$  in the mild integral equation above and then to replace the higher order terms recursively by Taylor expansions of lower orders (see Section 3). Hence, this method avoids the need of an Itô formula but nevertheless yields stochastic Taylor expansions of arbitrarily high order for the solution of the SPDE (1) (see Section 5.1 for details). Moreover, these Taylor expansions are robust with respect to the type of noise used and can easily be modified to other types of noise such as fractional Brownian motion.

The paper is organized as follows. In the next section, we describe precisely the SPDE that we are considering and state the assumptions that we require on its terms and coefficients and on the initial value. Then, in the third section, we sketch the idea and the notation for deriving simple Taylor expansions, which we develop in section four in some detail using combinatorial objects, specifically stochastic trees and woods, to derive Taylor expansions of an arbitrarily high order. We also provide an estimate for the remainder terms of the Taylor expansions there. (Proofs are postponed to the final section.) These results are illustrated with some representative examples in the fifth section. Numerical schemes based on these Taylor expansions are presented in the sixth section.

**2. Assumptions.** Fix  $T > 0$  and let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space with a normal filtration  $\mathcal{F}_t$ ,  $t \in [0, T]$ ; see, for example, [6] for details. In addition, let  $(H, \langle \cdot, \cdot \rangle)$  be a separable  $\mathbb{R}$ -Hilbert space with its norm denoted by  $|\cdot|$  and consider the SPDE (1) in the mild integral equation form (2) on  $H$ , where  $W_t$ ,  $t \in [0, T]$ , is a cylindrical  $Q$ -Wiener process with  $Q = I$  on another separable  $\mathbb{R}$ -Hilbert space  $(U, \langle \cdot, \cdot \rangle)$  (space-time white noise),  $B : U \rightarrow H$  is a bounded linear operator and the objects  $A$ ,  $F$  and  $u_0$  are specified through the following assumptions.

**ASSUMPTION 1 (Linear operator  $A$ ).** Let  $I$  be a countable or finite set, let  $(\lambda_i)_{i \in I} \subset \mathbb{R}$  be a family of real numbers with  $\inf_{i \in I} \lambda_i > -\infty$  and let  $(e_i)_{i \in I} \subset H$

be an orthonormal basis of  $H$ . Assume that the linear operator  $A : D(A) \subset H \rightarrow H$  is given by

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

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

**ASSUMPTION 2 (Drift term  $F$ ).** The nonlinearity  $F : H \rightarrow H$  is smooth and regular in the sense that  $F$  is infinitely often Fréchet differentiable and its derivatives satisfy  $\sup_{v \in H} |F^{(i)}(v)| < \infty$  for all  $i \in \mathbb{N} := \{1, 2, \dots\}$ .

Fix  $\kappa \geq 0$  with  $\sup_{i \in \mathcal{I}} (\kappa + \lambda_i) > 0$  and let  $D((\kappa - A)^r)$ ,  $r \in \mathbb{R}$ , denote the domains of powers of the operator  $\kappa - A : D(\kappa - A) = D(A) \subset H \rightarrow H$ , see, for example, [41].

**ASSUMPTION 3 (Stochastic convolution).** There exist two real numbers  $\gamma \in (0, 1)$ ,  $\delta \in (0, \frac{1}{2}]$  and a constant  $C > 0$  such that

$$\int_0^T |(\kappa - A)^\gamma e^{As} B|_{\text{HS}}^2 ds < \infty, \quad \int_0^t |e^{As} B|_{\text{HS}}^2 ds \leq Ct^{2\delta}$$

holds for all  $t \in [0, 1]$ , where  $|\cdot|_{\text{HS}}$  denotes the Hilbert–Schmidt norm for Hilbert–Schmidt operators from  $U$  to  $H$ .

**ASSUMPTION 4 (Initial value  $u_0$ ).** The  $\mathcal{F}_0/\mathcal{B}(D((\kappa - A)^\gamma))$ -measurable mapping  $u_0 : \Omega \rightarrow D((\kappa - A)^\gamma)$  satisfies  $\mathbb{E}|(\kappa - A)^\gamma u_0|^p < \infty$  for every  $p \in [1, \infty)$ , where  $\gamma \in (0, 1)$  is given in Assumption 3.

Similar assumptions are used in the literature on the approximation of this kind of SPDEs (see, e.g., Assumption H1–H3 in [15] or see also [24, 32–34]). This setup also includes trace class noise (see Section 5.5) and finite-dimensional SODEs with additive noise (see Section 5.2).

Henceforth, we fix  $t_0 \in [0, T)$  and denote by  $\mathcal{P}$  the set of all adapted stochastic processes

$$X : \Omega \rightarrow C([t_0, T], H) \quad \text{with} \quad \sup_{t_0 \leq t \leq T} |X_t|_{L^p} < \infty \quad \forall p \geq 1,$$

and with continuous sample paths, where  $|Z|_{L^p} := (\mathbb{E}|Z|^p)^{1/p}$  is the  $L^p$ -norm of a random variable  $Z : \Omega \rightarrow H$ . Under Assumptions 1 and 3, it is well known that the stochastic convolution

$$\int_0^t e^{A(t-s)} B dW_s, \quad t \in [t_0, T],$$

has an (up to indistinguishability) unique version with continuous sample paths (see Lemma 5 in Section 7.3). From now on, we fix such a version of the stochas-

tic convolution. Hence, under Assumptions 1–4 it is well known that there is a pathwise unique adapted stochastic process  $U : \Omega \rightarrow C([0, T], H)$  with continuous sample paths, which satisfies (2) (see Theorems 7.4 and 7.6 in [6]). Even more, this process satisfies

$$(3) \quad \sup_{0 \leq t \leq T} |(\kappa - A)^\gamma U_t|_{L^p} < \infty$$

for all  $p \geq 1$ .

**3. Taylor expansions.** In this section, we present the notation and the basic idea behind the derivation of Taylor expansions. We write

$$\Delta U_t := U_t - U_{t_0}, \quad \Delta t := t - t_0$$

for  $t \in [t_0, T] \subset [0, T]$ , thus  $\Delta U$  denotes the stochastic process  $\Delta U_t, t \in [t_0, T]$ , in  $\mathcal{P}$ . First, we introduce some integral operators and an expression relating them and then we show how they can be used to derive some simple Taylor expansions.

3.1. *Integral operators.* Let  $j \in \{0, 1, 2, 1^*\}$ , where the indices  $\{0, 1, 2\}$  will label expressions containing only a constant value or no value of the SPDE solution, while  $1^*$  will label a certain integral with time dependent values of the SPDE solution in the integrand. Specifically, we define the stochastic processes  $I_j^0 \in \mathcal{P}$  by

$$I_j^0(t) := \begin{cases} (e^{A\Delta t} - I)U_{t_0}, & j = 0, \\ \int_{t_0}^t e^{A(t-s)} F(U_{t_0}) ds, & j = 1, \\ \int_{t_0}^t e^{A(t-s)} B dW_s, & j = 2, \\ \int_{t_0}^t e^{A(t-s)} F(U_s) ds, & j = 1^*, \end{cases}$$

for each  $t \in [t_0, T]$ . Note that the stochastic process  $\int_{t_0}^t e^{A(t-s)} B dW_s$  for  $t \in [t_0, T]$  given by

$$\int_{t_0}^t e^{A(t-s)} B dW_s = \int_0^t e^{A(t-s)} B dW_s - e^{A(t-t_0)} \left( \int_0^{t_0} e^{A(t_0-s)} B dW_s \right)$$

for every  $t \in [t_0, T]$  is indeed in  $\mathcal{P}$ . Given  $i \in \mathbb{N}$  and  $j \in \{1, 1^*\}$ , we then define the  $i$ -multilinear symmetric mapping  $I_j^i : \mathcal{P}^i \rightarrow \mathcal{P}$  by

$$I_j^i[g_1, \dots, g_i](t) := \frac{1}{i!} \int_{t_0}^t e^{A(t-s)} F^{(i)}(U_{t_0})(g_1(s), \dots, g_i(s)) ds,$$

when  $j = 1$  and by

$$I_j^i[g_1, \dots, g_i](t) := \int_{t_0}^t e^{A(t-s)} \left( \int_0^1 F^{(i)}(U_{t_0} + r\Delta U_s)(g_1(s), \dots, g_i(s)) \frac{(1-r)^{(i-1)}}{(i-1)!} dr \right) ds,$$

when  $j = 1^*$  for all  $t \in [t_0, T]$  and all  $g_1, \dots, g_i \in \mathcal{P}$ . One immediately checks that the stochastic processes  $I_j^0 \in \mathcal{P}$ ,  $j \in \{0, 1, 2, 1^*\}$ , and the mappings  $I_j^i: \mathcal{P}^i \rightarrow \mathcal{P}$ ,  $i \in \mathbb{N}$ ,  $j \in \{1, 1^*\}$ , are well defined.

The solution process  $U$  of (2) obviously satisfies

$$(4) \quad \Delta U_t = (e^{A\Delta t} - I)U_{t_0} + \int_{t_0}^t e^{A(t-s)} F(U_s) ds + \int_{t_0}^t e^{A(t-s)} B dW_s$$

or, in terms of the above integral operators,

$$\Delta U_t = I_0^0(t) + I_{1^*}^0(t) + I_2^0(t)$$

for every  $t \in [t_0, T]$ , which we can write symbolically in the space  $\mathcal{P}$  as

$$(5) \quad \Delta U = I_0^0 + I_{1^*}^0 + I_2^0.$$

The stochastic processes  $I_j^i[g_1, \dots, g_i]$  for  $g_1, \dots, g_i \in \mathcal{P}$ ,  $j \in \{0, 1, 2\}$  and  $i \in \{0, 1, 2, \dots\}$  only depend on the solution at time  $t = t_0$ . These terms are therefore useful approximations for the solution process  $U_t$ ,  $t \in [t_0, T]$ . However, the stochastic processes  $I_{1^*}^i[g_1, \dots, g_i]$  for  $g_1, \dots, g_i \in \mathcal{P}$  and  $i \in \{0, 1, 2, \dots\}$  depend on the solution path  $U_t$  with  $t \in [t_0, T]$ . Therefore, we need a further expansion for these processes. For this, we will use the important formula

$$(6) \quad \begin{aligned} I_{1^*}^0 &= I_1^0 + I_{1^*}^1[\Delta U] \\ &= I_1^0 + I_{1^*}^1[I_0^0] + I_{1^*}^1[I_{1^*}^0] + I_{1^*}^1[I_2^0], \end{aligned}$$

which is an immediate consequence of integration by parts and (5), and the iterated formula

$$(7) \quad \begin{aligned} I_{1^*}^i[g_1, \dots, g_i] &= I_1^i[g_1, \dots, g_i] + I_{1^*}^{(i+1)}[\Delta U, g_1, \dots, g_i] \\ &= I_1^i[g_1, \dots, g_i] + I_{1^*}^{(i+1)}[I_0^0, g_1, \dots, g_i] \\ &\quad + I_{1^*}^{(i+1)}[I_{1^*}^0, g_1, \dots, g_i] + I_{1^*}^{(i+1)}[I_2^0, g_1, \dots, g_i] \end{aligned}$$

for every  $g_1, \dots, g_i \in \mathcal{P}$  and every  $i \in \mathbb{N}$  (see Lemma 1 for a proof of the equations above).

3.2. *Derivation of simple Taylor expansions.* To derive a further expansion of (5), we insert formula (6) to the stochastic process  $I_{1^*}^0$ , that is,

$$I_{1^*}^0 = I_1^0 + I_{1^*}^1[I_0^0] + I_{1^*}^1[I_{1^*}^0] + I_{1^*}^1[I_2^0]$$

into (5) to obtain

$$\Delta U = I_0^0 + (I_1^0 + I_{1^*}^1[I_0^0] + I_{1^*}^1[I_{1^*}^0] + I_{1^*}^1[I_2^0]) + I_2^0,$$

which can also be written as

$$(8) \quad \Delta U = I_0^0 + I_1^0 + I_2^0 + I_{1^*}^1[I_0^0] + I_{1^*}^1[I_{1^*}^0] + I_{1^*}^1[I_2^0].$$

If we can show that the double integral terms  $I_{1*}^1[I_0^0]$ ,  $I_{1*}^1[I_{1*}^0]$  and  $I_{1*}^1[I_2^0]$  are sufficiently small (indeed, this will be done in the next section), then we obtain the approximation

$$(9) \quad \Delta U \approx I_0^0 + I_1^0 + I_2^0,$$

or, using the definition of the stochastic processes  $I_0^0$ ,  $I_1^0$  and  $I_2^0$ ,

$$\Delta U_t \approx (e^{A\Delta t} - I)U_{t_0} + \int_{t_0}^t e^{A(t-s)} F(U_{t_0}) ds + \int_{t_0}^t e^{A(t-s)} B dW_s$$

for  $t \in [t_0, T]$ . Hence,

$$(10) \quad U_t \approx e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s$$

for  $t \in [t_0, T]$  is an approximation for the solution of SPDE (1). Since the right-hand side of (10) depends on the solution only at time  $t_0$ , it is the first nontrivial Taylor expansion of the solution of the SPDE (1). The remainder terms  $I_{1*}^1[I_0^0]$ ,  $I_{1*}^1[I_{1*}^0]$  and  $I_{1*}^1[I_2^0]$  of this approximation can be estimated by

$$|I_{1*}^1[I_0^0](t) + I_{1*}^1[I_{1*}^0](t) + I_{1*}^1[I_2^0](t)|_{L^2} \leq C(\Delta t)^{(1+\min(\gamma,\delta))}$$

for every  $t \in [t_0, T]$  with a constant  $C > 0$  and where  $\gamma \in (0, 1)$  and  $\delta \in (0, \frac{1}{2}]$  are given in Assumption 3 (see Theorem 1 in the next section for details).

We write  $Y_t = O((\Delta t)^r)$  with  $r > 0$  for a stochastic process  $Y \in \mathcal{P}$ , if  $|Y_t|_{L^2} \leq C(\Delta t)^r$  holds for all  $t \in [t_0, T]$  with a constant  $C > 0$ . Therefore, we have

$$\begin{aligned} U_t - \left( e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s \right) \\ = O((\Delta t)^{(1+\min(\gamma,\delta))}) \end{aligned}$$

or

$$(11) \quad \begin{aligned} U_t = e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s \\ + O((\Delta t)^{(1+\min(\gamma,\delta))}). \end{aligned}$$

The approximation (10) thus has order  $1 + \min(\gamma, \delta)$  in the above strong sense. It plays an analogous role to the simplest strong Taylor expansion in [29] on which the Euler–Maruyama scheme for finite-dimensional SODEs is based and was in fact used in [24] to derive the exponential Euler scheme for the SPDE (1). Note that the Euler–Maruyama scheme in [29] approximates the solution of an SODE with additive noise locally with order  $\frac{3}{2}$ . Here, in the case of trace class noise, we will have  $\gamma = \frac{1}{2} - \varepsilon$ ,  $\delta = \frac{1}{2}$  (see Section 5.5), and therefore the exponential Euler scheme for the SPDE (1) in [24] also approximates the solution locally with order  $\frac{3}{2} - \varepsilon$  (see Section 5.5.2), while other schemes in use, in particular the linear

implicit Euler scheme or the Crank–Nicolson scheme, approximate the solution with order  $\frac{1}{2}$  instead of order  $\frac{3}{2}$  as in the finite-dimensional case. Therefore, the Taylor approximation (11) attains the classical order of the Euler approximation for finite-dimensional SODEs and in general the Taylor expansion introduced above lead to numerical schemes for SPDEs, which converge with a higher order than other schemes in use (see Section 6).

3.3. *Higher order Taylor expansions.* Further expansions of the remainder terms in a Taylor expansion give a Taylor expansion of higher order. To illustrate this, we will expand the terms  $I_{1*}^1[I_0^0]$  and  $I_{1*}^1[I_2^0]$  in (8). From (7), we have

$$I_{1*}^1[I_0^0] = I_1^1[I_0^0] + I_{1*}^2[I_0^0, I_0^0] + I_{1*}^2[I_{1*}^0, I_0^0] + I_{1*}^2[I_2^0, I_0^0]$$

and

$$I_{1*}^1[I_2^0] = I_1^1[I_2^0] + I_{1*}^2[I_0^0, I_2^0] + I_{1*}^2[I_{1*}^0, I_2^0] + I_{1*}^2[I_2^0, I_2^0],$$

which we insert into (8) to obtain

$$\Delta U = I_0^0 + I_1^0 + I_2^0 + I_1^1[I_0^0] + I_1^1[I_2^0] + R,$$

where the remainder term  $R$  is given by

$$R = I_{1*}^1[I_{1*}^0] + I_{1*}^2[I_0^0, I_0^0] + I_{1*}^2[I_{1*}^0, I_0^0] + I_{1*}^2[I_2^0, I_0^0] + I_{1*}^2[I_0^0, I_2^0] + I_{1*}^2[I_{1*}^0, I_2^0] + I_{1*}^2[I_2^0, I_2^0].$$

From Theorem 1 in the next section, we will see  $R = O((\Delta t)^{(1+2\min(\gamma,\delta))})$ . Thus, we have

$$\Delta U = I_0^0 + I_1^0 + I_2^0 + I_2^1[I_0^0] + I_2^1[I_2^0] + O((\Delta t)^{(1+2\min(\gamma,\delta))}),$$

which can also be written as

$$\begin{aligned} U_t &= e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s \\ &+ \int_{t_0}^t e^{A(t-s)} F'(U_{t_0})(e^{A\Delta s} - I) U_{t_0} ds \\ &+ \int_{t_0}^t e^{A(t-s)} F'(U_{t_0}) \int_{t_0}^s e^{A(s-r)} B dW_r ds + O((\Delta t)^{(1+2\min(\gamma,\delta))}). \end{aligned}$$

This approximation is of order  $1 + 2\min(\gamma, \delta)$ .

By iterating this idea, we can construct further Taylor expansions. In particular, we will show in the next section how a Taylor expansion of arbitrarily high order can be achieved.



**4. Systematic derivation of Taylor expansions of arbitrarily high order.**

The basic mechanism for deriving a Taylor expansion for the SPDE (1) was explained in the previous section. We illustrate now how Taylor expansions of arbitrarily high order can be derived and will also estimate their remainder terms. For this, we will identify the terms occurring in a Taylor expansions by combinatorial objects, that is, trees. It is a standard tool in numerical analysis to describe higher order terms in a Taylor expansion via rooted trees (see, e.g., [2] for ODEs and [1, 38–40] for SODEs). In particular, we introduce a class of trees which is appropriate for our situation and show how the trees relate to the desired Taylor expansions.

4.1. *Stochastic trees and woods.* We begin with the definition of the trees that we need, adapting the standard notation of the trees used in the Taylor expansion of SODEs (see, e.g., Definition 2.3.1 in [38] as well as [1, 39, 40]).

Let  $N \in \mathbb{N}$  be a natural number and let

$$\mathbf{t}' : \{2, \dots, N\} \rightarrow \{1, \dots, N - 1\}, \quad \mathbf{t}'' : \{1, \dots, N\} \rightarrow \{0, 1, 2, 1^*\}$$

be two mappings with the property that  $\mathbf{t}'(j) < j$  for all  $j \in \{2, \dots, N\}$ . The pair of mappings  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'')$  is a S-tree (stochastic tree) of length  $N = l(\mathbf{t})$  nodes.

Every S-tree can be represented as a graph, whose nodes are given by the set  $\text{nd}(\mathbf{t}) := \{1, \dots, N\}$  and whose arcs are described by the mapping  $\mathbf{t}'$  in the sense that there is an edge from  $j$  to  $\mathbf{t}'(j)$  for every node  $j \in \{2, \dots, N\}$ . In view of a rooted tree,  $\tau'$  also codifies the parent and child pairings and is therefore often referred as son-father mapping (see, e.g., Definition 2.1.5 in [38]). The mapping  $\mathbf{t}''$  is an additional labeling of the nodes with  $\mathbf{t}''(j) \in \{0, 1, 2, 1^*\}$  indicating the type of node  $j$  for every  $j \in \text{nd}(\mathbf{t})$ . The left picture in Figure 1 corresponds to the tree  $\mathbf{t}_1 = (\mathbf{t}'_1, \mathbf{t}''_1)$  with  $\text{nd}(\mathbf{t}_1) = \{1, 2, 3, 4\}$  given by

$$\mathbf{t}'_1(4) = 1, \quad \mathbf{t}'_1(3) = 2, \quad \mathbf{t}'_1(2) = 1$$

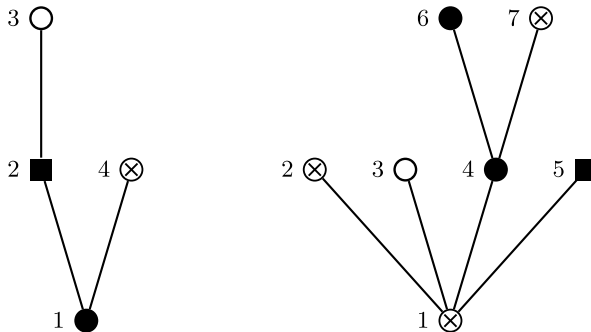


FIG. 1. Two examples of stochastic trees.



FIG. 2. The stochastic wood  $w_0$  in  $\mathbf{SW}$ .

and

$$t''_1(1) = 1, \quad t''_1(2) = 1^*, \quad t''_1(3) = 2, \quad t''_1(4) = 0.$$

The root is always presented as the lowest node. The number on the left of a node in Figure 1 is the number of the node of the corresponding tree. The type of the nodes in Figure 1 depends on the additional labeling of the nodes given by  $t''_1$ . More precisely, we represent a node  $j \in \text{nd}(t_1)$  by  $\otimes$  if  $t''_1(j) = 0$ , by  $\bullet$  if  $t''_1(j) = 1$ , by  $\circ$  if  $t''_1(j) = 2$ , and finally by  $\blacksquare$  if  $t''_1(j) = 1^*$ . The right picture in Figure 1 corresponds to the tree  $t_2 = (t'_2, t''_2)$  with  $\text{nd}(t_2) = \{1, \dots, 7\}$  given by

$$\begin{aligned} t'_2(7) &= 4, & t'_2(6) &= 4, & t'_2(5) &= 1, \\ t'_2(4) &= 1, & t'_2(3) &= 1, & t'_2(2) &= 1 \end{aligned}$$

and

$$\begin{aligned} t''_2(1) &= 0, & t''_2(2) &= 0, & t''_2(3) &= 2, & t''_2(4) &= 1, \\ t''_2(5) &= 1^*, & t''_2(6) &= 1, & t''_2(7) &= 0. \end{aligned}$$

We denote the set of all stochastic trees by  $\mathbf{ST}$  and will also consider a tuple of trees, that is, a wood. The set of S-woods (stochastic woods) is defined by

$$\mathbf{SW} := \bigcup_{n=1}^{\infty} (\mathbf{ST})^n.$$

Of course, we have the embedding  $\mathbf{ST} \subset \mathbf{SW}$ . A simple example of an S-wood which will be required later is  $w_0 = (t_1, t_2, t_3)$  with  $t_1, t_2$  and  $t_3$  given by  $l(t_1) = l(t_2) = l(t_3) = 1$  and  $t''_1(1) = 0, t'_2(1) = 1^*, t''_3(1) = 2$ . This is shown in Figure 2 where the left tree corresponds to  $t_1$ , the middle one to  $t_2$  and the right tree corresponds to  $t_3$ .

4.2. *Construction of stochastic trees and woods.* We define an operator on the set  $\mathbf{SW}$ , that will enable us to construct an appropriate stochastic wood step by step. Let  $w = (t_1, \dots, t_n)$  with  $n \in \mathbb{N}$  be a S-wood with  $t_i = (t'_i, t''_i) \in \mathbf{ST}$  for every  $i \in \{1, 2, \dots, n\}$ . Moreover, let  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, l(t_i)\}$  be given and suppose that  $t''_i(j) = 1^*$ , in which case we call the pair  $(i, j)$  an active node of  $w$ . We denote the set of all active nodes of  $w$  by  $\text{acn}(w)$ .

Now, we introduce the trees  $t_{n+1} = (t'_{n+1}, t''_{n+1}), t_{n+2} = (t'_{n+2}, t''_{n+2})$  and  $t_{n+3} = (t'_{n+3}, t''_{n+3})$  in  $\mathbf{ST}$  by  $\text{nd}(t_{n+m}) = \{1, \dots, l(t_i), l(t_i) + 1\}$  and

$$\begin{aligned} t'_{n+m}(k) &= t'_i(k), & k &= 2, \dots, l(t_i), \\ t''_{n+m}(k) &= t''_i(k), & k &= 1, \dots, l(t_i), \end{aligned}$$

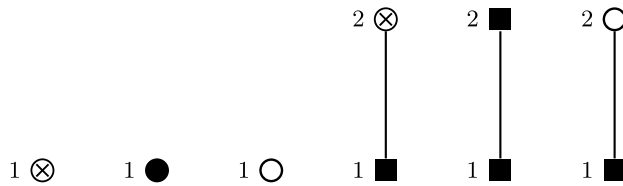


FIG. 3. The stochastic wood  $\mathbf{w}_1$  in SW.

$$\mathbf{t}'_{n+m}(l(\mathbf{t}_i) + 1) = j, \quad \mathbf{t}''_{n+m}(l(\mathbf{t}_i) + 1) = \begin{cases} 1^*, & m = 2, \\ (m - 1), & \text{else,} \end{cases}$$

for  $m = 1, 2, 3$ . Finally, we consider the S-tree  $\tilde{\mathbf{t}}^{(i,j)} = (\tilde{\mathbf{t}}', \tilde{\mathbf{t}}'')$  given by  $\tilde{\mathbf{t}}' = \mathbf{t}'_i$ , but with  $\tilde{\mathbf{t}}''(k) = \mathbf{t}''_i(k)$  for  $k \neq j$  and  $\tilde{\mathbf{t}}''(j) = 1$ . Then, we define

$$E_{(i,j)}\mathbf{w} = E_{(i,j)}(\mathbf{t}_1, \dots, \mathbf{t}_n) := (\mathbf{t}_1, \dots, \mathbf{t}_{i-1}, \tilde{\mathbf{t}}^{(i,j)}, \mathbf{t}_{i+1}, \dots, \mathbf{t}_{n+3})$$

and consider the set of all woods that can be constructed by iteratively applying the  $E_{(i,j)}$  operations, that is, we define

$$\mathbf{SW}' := \{\mathbf{w}_0\} \cup \left\{ \mathbf{w} \in \mathbf{SW} \left| \begin{array}{l} \exists n \in \mathbb{N}, i_1, j_1, \dots, i_n, j_n \in \mathbb{N}: \\ \forall l = 1, \dots, n: (i_l, j_l) \in \text{acn}(E_{(i_{l-1}, j_{l-1})} \cdots E_{(i_1, j_1)} \mathbf{w}_0), \\ \mathbf{w} = E_{(i_n, j_n)} \cdots E_{(i_1, j_1)} \mathbf{w}_0 \end{array} \right. \right\}$$

for the  $\mathbf{w}_0$  introduced above. To illustrate these definitions, we present some examples using the initial stochastic wood  $\mathbf{w}_0$  given in Figure 2. We present these examples here in a brief way, and, later in Section 5.1, we describe more detailed the main advantages of the particular examples considered here. First, the active nodes of  $\mathbf{w}_0$  are  $\text{acn}(\mathbf{w}_0) = \{(2, 1)\}$ , since the first node in the second tree in  $\mathbf{w}_0$  is the only node of type  $1^*$ . Hence,  $E_{(2,1)}\mathbf{w}_0$  is well defined and the resulting stochastic wood  $\mathbf{w}_1 = E_{(2,1)}\mathbf{w}_0$ , which has six trees, is presented in Figure 3. Writing  $\mathbf{w}_1 = (\mathbf{t}_1, \dots, \mathbf{t}_6)$ , the left tree in Figure 3 corresponds to  $\mathbf{t}_1$ , the second tree in Figure 3 corresponds to  $\mathbf{t}_2$ , and so on. Moreover, we have

$$(12) \quad \text{acn}(\mathbf{w}_1) = \{(4, 1), (5, 1), (5, 2), (6, 1)\}$$

for the active nodes of  $\mathbf{w}_1$ , so  $\mathbf{w}_2 = E_{(4,1)}\mathbf{w}_1$  is also well defined. It is presented in Figure 4. In Figure 5, we present the stochastic wood  $\mathbf{w}_3 = E_{(6,1)}\mathbf{w}_2$ , which is

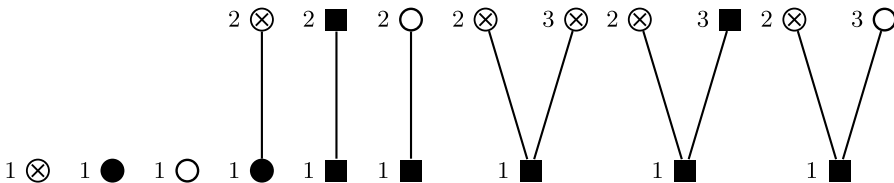


FIG. 4. The stochastic wood  $\mathbf{w}_2$  in SW.

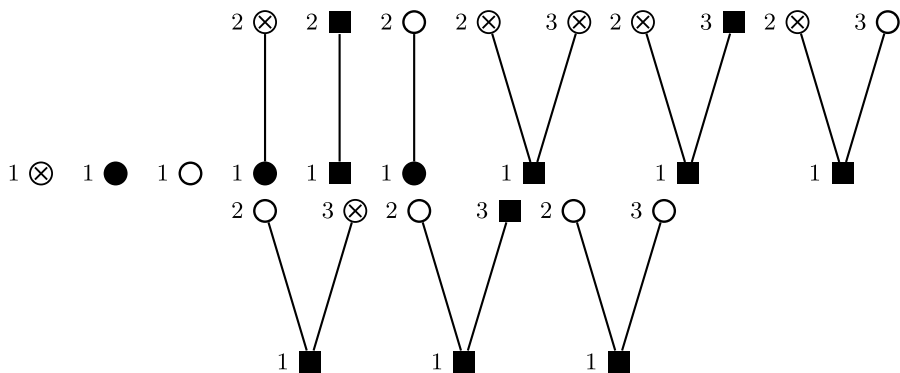


FIG. 5. The stochastic wood  $w_3$  in SW.

well defined since

$$(13) \quad \text{acn}(w_2) = \{(5, 1), (5, 2), (6, 1), (7, 1), (8, 1), (8, 3), (9, 1)\}.$$

For the S-wood  $w_3$ , we have

$$(14) \quad \text{acn}(w_3) = \left\{ \begin{array}{l} (5, 1), (5, 2), (7, 1), (8, 1), (8, 3), (9, 1), \\ (10, 1), (11, 1), (11, 3), (12, 1) \end{array} \right\}.$$

Since  $(7, 1) \in \text{acn}(w_3)$ , the stochastic wood  $w_4 = E_{(7,1)}w_3$  is well defined and presented in Figure 6. For the active nodes, we obtain

$$(15) \quad \text{acn}(w_4) = \left\{ \begin{array}{l} (5, 1), (5, 2), (8, 1), (8, 3), (9, 1), (10, 1), (11, 1), \\ (11, 3), (12, 1), (13, 1), (14, 1), (14, 4), (15, 1) \end{array} \right\}.$$

Finally, we present the stochastic wood  $w_5 = E_{(12,1)}E_{(10,1)}E_{(9,1)}w_4$  with

$$(16) \quad \text{acn}(w_5) = \left\{ \begin{array}{l} (5, 1), (5, 2), (8, 1), (8, 3), (11, 1), (11, 3), (13, 1), (14, 1), \\ (14, 4), (15, 1), (16, 1), (17, 1), (17, 4), (18, 1), (19, 1), \\ (20, 1), (20, 4), (21, 1), (22, 1), (23, 1), (23, 4), (24, 1) \end{array} \right\}$$

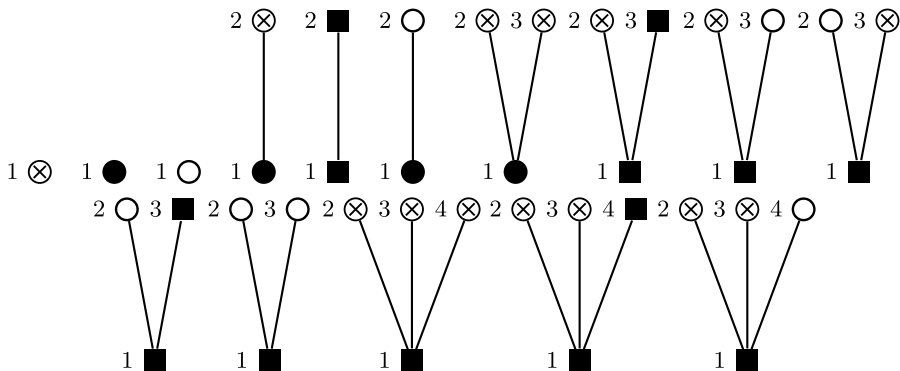


FIG. 6. The stochastic wood  $w_4$  in SW.

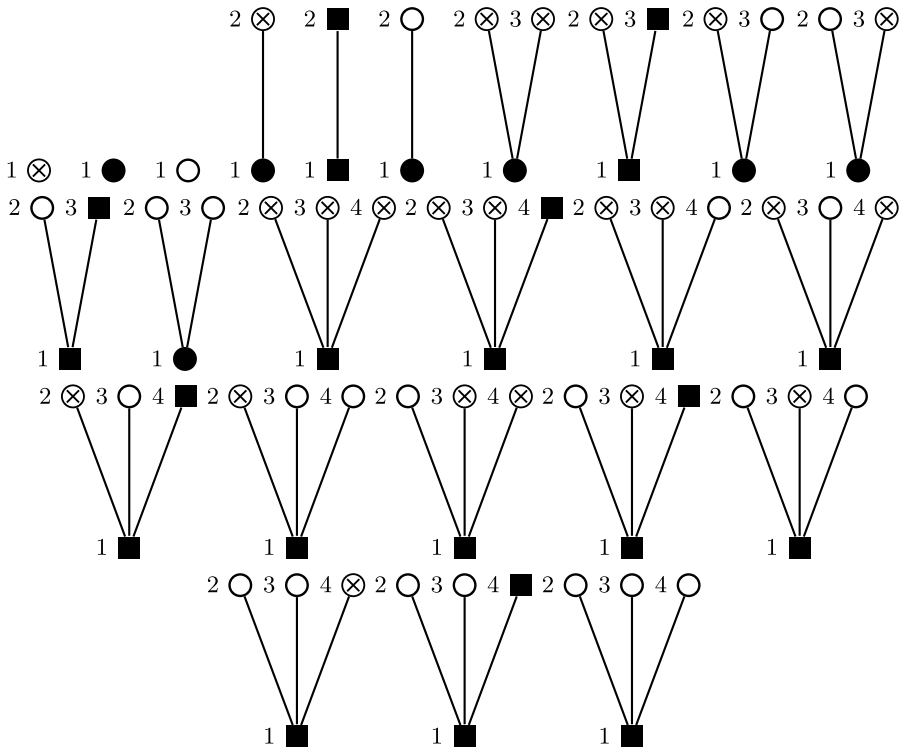


FIG. 7. The stochastic wood  $\mathbf{w}_5$  in  $\mathbf{SW}$ .

in Figure 7. By definition, the S-woods  $\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_5$  are in  $\mathbf{SW}'$ , but the stochastic wood given in Figure 1 is not in  $\mathbf{SW}'$ .

4.3. *Subtrees.* Let  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'')$  be a given S-tree with  $l(\mathbf{t}) \geq 2$ . For two nodes  $k, l \in \text{nd}(\mathbf{t})$  with  $k \leq l$ , we say that  $l$  is a grandchild of  $k$  if there exists a sequence  $k_1 = k < k_2 < \dots < k_n = l$  of nodes with  $n \in \mathbb{N}$  such that  $\mathbf{t}'(k_{v+1}) = k_v$  for every  $v \in \{1, \dots, n-1\}$ . Suppose now that  $j_1, \dots, j_n \in \text{nd}(\mathbf{t})$  with  $n \in \mathbb{N}$  and  $j_1 < \dots < j_n$  are the nodes of  $\mathbf{t}$  such that  $\mathbf{t}'(j_i) = 1$  for every  $i = 1, \dots, n$ . Moreover, for a given  $i \in \{1, \dots, n\}$  suppose that  $j_{i,1}, j_{i,2}, \dots, j_{i,l_i} \in \text{nd}(\mathbf{t})$  with  $j_i = j_{i,1} < j_{i,2} < \dots < j_{i,l_i} \leq l(\mathbf{t})$  and  $l_i \in \mathbb{N}$  are the grandchildren of  $j_i$ . Then, we define a tree  $\mathbf{t}_i = (\mathbf{t}'_i, \mathbf{t}''_i) \in \mathbf{ST}$  with  $l(\mathbf{t}_i) := l_i$  and

$$j_{i,\mathbf{t}'_i(k)} = \mathbf{t}'(j_{i,k}), \quad \mathbf{t}''_i(k) = \mathbf{t}''(j_{i,k})$$

for all  $k \in \{2, \dots, l_i\}$  and  $\mathbf{t}''_i(1) = \mathbf{t}''(j_i)$ . We call the trees  $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbf{ST}$  defined in this way the *subtrees* of  $\mathbf{t}$ . For example, the subtrees of the right tree in Figure 1 are presented in Figure 8.

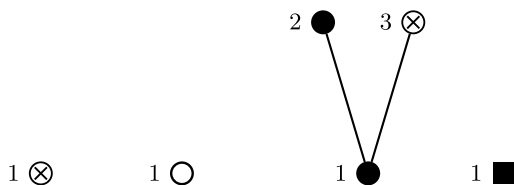


FIG. 8. Subtrees of the right tree in Figure 1.

4.4. *Order of a tree.* Later stochastic woods in  $\mathbf{SW}'$  will represent Taylor expansions and Taylor approximations of the solution process  $U$  of the SPDE (1). Additionally, we will estimate the approximation orders of these Taylor approximations. To this end, we introduce the order of a stochastic tree and of a stochastic wood, which is motivated by Lemma 4 below. More precisely, let  $\text{ord} : \mathbf{ST} \rightarrow [0, \infty)$  be given by

$$\begin{aligned} \text{ord}(\mathbf{t}) := & l(\mathbf{t}) + (\gamma - 1)|\{j \in \text{nd}(\mathbf{t}) | \mathbf{t}''(j) = 0\}| \\ & + (\delta - 1)|\{j \in \text{nd}(\mathbf{t}) | \mathbf{t}''(j) = 2\}| \end{aligned}$$

for every S-tree  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{ST}$ . For example, the order of the left tree in Figure 1 is  $2 + \gamma + \delta$  and the order of the right tree in Figure 1 is  $3 + 3\gamma + \delta$  (since the right tree has three nodes of type 0, three nodes of type 1, respectively,  $1^*$ , and one node of type 2).

In addition, we say that a tree  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'')$  in  $\mathbf{ST}$  is *active* if there is a  $j \in \text{nd}(\mathbf{t})$  such that  $\mathbf{t}''(j) = 1^*$ . In that sense a S-tree is active if it has an active node. Moreover, we define the order of an S-wood  $\mathbf{w} = (\mathbf{t}_1, \dots, \mathbf{t}_n) \in \mathbf{SW}$  with  $n \in \mathbb{N}$  as

$$\text{ord}(\mathbf{w}) := \min\{\text{ord}(\mathbf{t}_i), 1 \leq i \leq n | \mathbf{t}_i \text{ is active}\}.$$

To illustrate this definition, we calculate the order of some stochastic woods. First of all, the stochastic wood in Figure 2 has order 1, since only the middle tree in Figure 2 is active. More precisely, the node  $(2, 1)$  of the S-wood  $\mathbf{w}_0$  is an active node and therefore the second tree is active. The second tree in Figure 2 has order 1 (since it only consists of one node of type  $1^*$ ). Hence, the S-wood  $\mathbf{w}_0$  has order 1. Since the last three trees are active in the stochastic wood  $\mathbf{w}_1$  in Figure 3 [see (12) for the active nodes of  $\mathbf{w}_1$ ], we obtain that the stochastic wood in Figure 3 has order  $1 + \min(\gamma, \delta)$ . The last three trees in the S-wood  $\mathbf{w}_1$  have order  $1 + \gamma$ , 2 and  $1 + \delta$ , respectively. As a third example, we consider the S-wood  $\mathbf{w}_2$  in Figure 4. The active nodes of  $\mathbf{w}_2$  are presented in (13). Hence, the last five S-trees are active. They have the orders 2,  $1 + \delta$ ,  $1 + 2\gamma$ ,  $2 + \gamma$  and  $1 + \gamma + \delta$ . The minimum of the five real numbers is  $1 + \min(2\gamma, \delta)$ . Therefore, the order of the S-wood  $\mathbf{w}_2$  in Figure 4 is  $1 + \min(2\gamma, \delta)$ . A similar calculation shows that the order of the stochastic wood  $\mathbf{w}_3$  in Figure 5 is  $1 + 2 \min(\gamma, \delta)$  and that the order of the stochastic wood  $\mathbf{w}_4$  in Figure 6 is  $1 + \min(3\gamma, \gamma + \delta, 2\delta)$ . Finally, we obtain that the stochastic wood  $\mathbf{w}_5$  in Figure 7 is of order  $1 + 3 \min(\gamma, \delta, \frac{1}{3})$ .

4.5. *Trees and stochastic processes.* To identify each tree in  $\mathbf{ST}$  with a predictable stochastic process in  $\mathcal{P}$ , we define two functions  $\Phi: \mathbf{ST} \rightarrow \mathcal{P}$  and  $\Psi: \mathbf{ST} \rightarrow \mathcal{P}$ , recursively. For a given S-tree  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{ST}$ , we define  $\Phi(\mathbf{t}) := I_{\mathbf{t}''(1)}^0$  when  $\mathbf{t}''(1) \in \{0, 2\}$  or  $l(\mathbf{t}) = 1$  and, when  $l(\mathbf{t}) \geq 2$  and  $\mathbf{t}''(1) \in \{1, 1^*\}$ , we define

$$\Phi(\mathbf{t}) := I_{\mathbf{t}''(1)}^n[\Phi(\mathbf{t}_1), \dots, \Phi(\mathbf{t}_n)],$$

where  $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbf{ST}$  with  $n \in \mathbb{N}$  are the subtrees of  $\mathbf{t}$ . In addition, for an arbitrary  $\mathbf{t} \in \mathbf{ST}$ , we define  $\Psi(\mathbf{t}) := 0$  if  $\mathbf{t}$  is an active tree and  $\Psi(\mathbf{t}) = \Phi(\mathbf{t})$  otherwise. Finally, for a S-wood  $\mathbf{w} = (\mathbf{t}_1, \dots, \mathbf{t}_n)$  with  $n \in \mathbb{N}$  we define  $\Phi(\mathbf{w})$  and  $\Psi(\mathbf{w})$  by

$$\Phi(\mathbf{w}) = \Phi(\mathbf{t}_1) + \dots + \Phi(\mathbf{t}_n), \quad \Psi(\mathbf{w}) = \Psi(\mathbf{t}_1) + \dots + \Psi(\mathbf{t}_n).$$

As an example, we have

$$(17) \quad \Phi(\mathbf{w}_0) = I_0^0 + I_{1^*}^0 + I_2^0 \quad \text{and} \quad \Psi(\mathbf{w}_0) = I_0^0 + I_2^0$$

for the elementary stochastic wood  $\mathbf{w}_0$  (see Figure 2). Hence, we obtain

$$(18) \quad \Phi(\mathbf{w}_0) = \Delta U$$

from (5) and (17). Since (2, 1) is an active node of  $\mathbf{w}_0$ , we obtain

$$(19) \quad \Phi(\mathbf{w}_1) = I_0^0 + I_1^0 + I_2^0 + I_{1^*}^1[I_0^0] + I_{1^*}^1[I_{1^*}^0] + I_{1^*}^1[I_2^0]$$

and

$$(20) \quad \Psi(\mathbf{w}_1) = I_0^0 + I_1^0 + I_2^0$$

for the S-wood  $\mathbf{w}_1 = E_{(2,1)}\mathbf{w}_0$  presented in Figure 3. Moreover, in view of (6) and (7), we have

$$(21) \quad \Phi(\mathbf{w}) = \Phi(E_{(i,j)}\mathbf{w})$$

for every active node  $(i, j) \in \text{acn}(\mathbf{w})$  and every stochastic wood  $\mathbf{w} \in \mathbf{SW}'$ .

Hence, we obtain

$$\Phi(\mathbf{w}_1) = \Phi(E_{(2,1)}\mathbf{w}_0) = \Phi(\mathbf{w}_0)$$

due to the equation above and the definition of  $\mathbf{w}_1$ . Hence, we obtain  $\Phi(\mathbf{w}_1) = \Delta U$ , which can also be seen from (19), since the right-hand side of (19) is nothing other than (8). We also note that the right-hand side of (20) is just the exponential Euler approximation in (9), so we obtain

$$\Delta U = \Phi(\mathbf{w}_1) \approx \Psi(\mathbf{w}_1).$$

With the above notation and definitions we are now able to present the main result of this article, which is a representation formula for the solution of the SPDE (1) via Taylor expansions and an estimate of the remainder terms occurring in the Taylor expansions.

**THEOREM 1.** *Let Assumptions 1–4 be fulfilled and let  $\mathbf{w} \in \mathbf{SW}'$  be an arbitrary stochastic wood. Then, for each  $p \in [1, \infty)$ , there is a constant  $C_p > 0$  such that*

$$(22) \quad \begin{aligned} U_t &= U_{t_0} + \Phi(\mathbf{w})(t), \\ (\mathbb{E}[|U_t - U_{t_0} - \Psi(\mathbf{w})(t)|^p])^{1/p} &\leq C_p \cdot (t - t_0)^{\text{ord}(\mathbf{w})} \end{aligned}$$

holds for every  $t \in [t_0, T]$ , where  $U_t, t \in [0, T]$ , is the solution of the SPDE (1). Here the constant  $C_p > 0$  is independent of  $t$  and  $t_0$  but depends on  $p$  as well as  $\mathbf{w}, T$  and the coefficients of the SPDE (1).

The representation of the solution here is a direct consequence of (18) and (21). The proof for the estimate in (22) will be given in Section 7. Here,  $\Phi(\mathbf{w}) = \Delta U$  is the increment of the solution of the SPDE (1), while  $\Psi(\mathbf{w})$  is the Taylor approximation of the increment of the solution and  $\Phi(\mathbf{w}) - \Psi(\mathbf{w})$  is its remainder for every arbitrary  $\mathbf{w} \in \mathbf{SW}'$ . Since there are woods in  $\mathbf{SW}'$  with arbitrarily high orders, Taylor expansions of arbitrarily high order can be constructed by successively applying the operator  $E_{(i,j)}$  to the initial S-wood  $\mathbf{w}_0$ . Finally, the approximation result of Theorem 1 can also be written as

$$U_t = U_{t_0} + \Psi(\mathbf{w}) + O((\Delta t)^{\text{ord}(\mathbf{w})})$$

for every stochastic wood  $\mathbf{w} \in \mathbf{SW}'$ . Here, we also remark that Assumptions 1–4 can be weakened. In particular, instead of Assumption 3, one can assume that the nonlinearity  $F : V \rightarrow V$  is only  $i$ -times Fréchet differentiable with  $i \in \mathbb{N}$  sufficiently high and that the derivatives of  $F$  satisfy only local estimates, where  $V \subset H$  is a continuously embedded Banach space. Nevertheless, it is usual to present Taylor expansions for stochastic differential equations under such restrictive assumptions here (see [29]) and then after considering a particular numerical scheme one reduces these assumptions by pathwise localization techniques (see [11, 26] for SDEs and [22] for SPDEs).

**5. Examples.** We present some examples here to illustrate the Taylor expansions introduced above.

5.1. *Abstract examples of the Taylor expansions.* We begin with some abstract examples of the Taylor expansions.

5.1.1. *Taylor expansion of order 1.* The first Taylor expansion of the solution is given by the initial stochastic wood  $\mathbf{w}_0$  (see Figure 2), that is, we have  $\Phi(\mathbf{w}_0) = \Delta U$  approximated by  $\Psi(\mathbf{w}_0)$  with order  $\text{ord}(\mathbf{w}_0)$ . Precisely, we have

$$\Psi(\mathbf{w}_0)(t) = (e^{A\Delta t} - I)U_{t_0} + \int_{t_0}^t e^{A(t-s)} B dW_s$$



and

$$\Phi(\mathbf{w}_0)(t) = (e^{A\Delta t} - I)U_{t_0} + \int_{t_0}^t e^{A(t-s)} F(U_s) ds + \int_{t_0}^t e^{A(t-s)} B dW_s$$

for every  $t \in [t_0, T]$  due to (17). Since  $\text{ord}(\mathbf{w}_0) = 1$  (see Section 4.4), we finally obtain

$$U_t = e^{A\Delta t} U_{t_0} + \int_{t_0}^t e^{A(t-s)} B dW_s + O(\Delta t)$$

for the Taylor expansion corresponding to the S-wood  $\mathbf{w}_0$ .

5.1.2. *Taylor expansion of order  $1 + \min(\gamma, \delta)$ .* In order to derive a higher order Taylor expansion, we expand the stochastic wood  $\mathbf{w}_0$ . To this end, we consider the Taylor expansion given by the S-wood  $\mathbf{w}_1 = E_{(2,1)} \mathbf{w}_0$  (see Figure 3). Here,  $\Phi(\mathbf{w}_1)$  and  $\Psi(\mathbf{w}_1)$  are presented in (19) and (20). Since  $\text{ord}(\mathbf{w}_1) = 1 + \min(\gamma, \delta)$  (see Section 4.4), we obtain

$$U_t = e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s + O((\Delta t)^{(1+\min(\gamma, \delta))})$$

for the Taylor expansion corresponding to the S-wood  $\mathbf{w}_1$ . This example corresponds to the exponential Euler scheme introduced in [24], which was already discussed in Section 3.2 [see (11)].

5.1.3. *Taylor expansion of order  $1 + \min(2\gamma, \delta)$ .* For a higher order Taylor expansion, we now have several possibilities to further expand the stochastic wood  $\mathbf{w}_1$ . For instance, we could consider the Taylor expansion given by the stochastic wood  $E_{(5,1)} E_{(5,2)} \mathbf{w}_1$  [see (12) for the active nodes of  $\mathbf{w}_1$ ]. Since our main goal is always to obtain higher order approximations with the least possible terms and since the fifth tree of  $\mathbf{w}_1$  [given by the nodes (5, 1) and (5, 2)] is of order 2 (see Section 4.4 for details), we concentrate on expanding the lower order trees of  $\mathbf{w}_1$ . Finally, since oftentimes  $\gamma \leq \delta$  in examples (see Section 5.4 and also Section 5.5), we consider the stochastic wood  $\mathbf{w}_2 = E_{(4,1)} \mathbf{w}_1$  (see Figure 4). It is of order  $1 + \min(2\gamma, \delta)$  (see Section 4.4) and the corresponding Taylor approximation  $\Psi(\mathbf{w}_2)$  of  $\Phi(\mathbf{w}_2) = \Delta U$  is given by  $\Psi(\mathbf{w}_2) = I_0^0 + I_1^0 + I_2^0 + I_1^1 [I_0^0]$ . This yields

$$U_t = e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s + \int_{t_0}^t e^{A(t-s)} F'(U_{t_0})(e^{A\Delta s} - I)U_{t_0} ds + O((\Delta t)^{(1+\min(2\gamma, \delta))})$$

for the Taylor expansion corresponding to the S-wood  $\mathbf{w}_2$ .

5.1.4. *Taylor expansion of order  $1 + 2 \min(\gamma, \delta)$ .* In examples, we oftentimes have  $\gamma = \frac{1}{4} - \varepsilon, \delta = \frac{1}{4}$  for an arbitrarily small  $\varepsilon \in (0, \frac{1}{4})$  (space-time white noise, see Section 5.4) or  $\gamma = \frac{1}{2} - \varepsilon, \delta = \frac{1}{2}$  for an arbitrarily small  $\varepsilon \in (0, \frac{1}{2})$  (trace-class noise). In these cases, the sixth stochastic tree in  $\mathbf{w}_2$  turns out to be the active tree of the lowest order. Therefore, we consider the stochastic wood  $\mathbf{w}_3 = E_{(6,1)}\mathbf{w}_2$  (see Figure 5) here. It has order  $1 + 2 \min(\gamma, \delta)$  (see Section 4.4) and we have

$$\Psi(\mathbf{w}_3) = I_0^0 + I_1^0 + I_2^0 + I_1^1[I_0^0] + I_1^1[I_2^0],$$

which implies

$$\begin{aligned} U_t &= e^{A\Delta t}U_{t_0} + \left(\int_0^{\Delta t} e^{As} ds\right)F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)}B dW_s \\ &+ \int_{t_0}^t e^{A(t-s)}F'(U_{t_0})(e^{A\Delta s} - I)U_{t_0} ds \\ &+ \int_{t_0}^t e^{A(t-s)}F'(U_{t_0}) \int_{t_0}^s e^{A(s-r)}B dW_r ds + O((\Delta t)^{1+2\min(\gamma,\delta)}). \end{aligned}$$

This example corresponds to the Taylor expansion introduced in the beginning in Section 3.3.

5.1.5. *Taylor expansion of order  $1 + \min(3\gamma, \gamma + \delta, 2\delta)$ .* Since we often have  $\gamma \leq \rho$  and  $\gamma < \frac{1}{2}$  in the examples below, the seventh stochastic tree in  $\mathbf{w}_3$  has the lowest order in these examples. Therefore, we consider the Taylor approximation corresponding to the S-wood  $\mathbf{w}_4 = E_{(7,1)}\mathbf{w}_3$  (see Figure 6), which is given by

$$\begin{aligned} U_t &= e^{A\Delta t}U_{t_0} + \left(\int_0^{\Delta t} e^{As} ds\right)F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)}B dW_s \\ &+ \int_{t_0}^t e^{A(t-s)}F'(U_{t_0})(e^{A\Delta s} - I)U_{t_0} ds \\ &+ \frac{1}{2} \int_{t_0}^t e^{A(t-s)}F''(U_{t_0})((e^{A\Delta s} - I)U_{t_0}, (e^{A\Delta s} - I)U_{t_0}) ds \\ &+ \int_{t_0}^t e^{A(t-s)}F'(U_{t_0}) \int_{t_0}^s e^{A(s-r)}B dW_r ds \\ &+ O((\Delta t)^{(1+\min(3\gamma,\gamma+\delta,2\delta))}). \end{aligned}$$

It is of order  $1 + \min(3\gamma, \gamma + \delta, 2\delta)$ , which can be seen in Section 4.4.

5.1.6. *Taylor expansion of order  $1 + 3 \min(\gamma, \delta, \frac{1}{3})$ .* In the case  $\gamma < \frac{1}{2}$ , the 9th, 10th and 12th stochastic tree in  $\mathbf{w}_4$  all have lower or equal orders than the fifth stochastic tree in  $\mathbf{w}_4$ . Therefore, we consider the S-wood  $\mathbf{w}_5 = E_{(12,1)}E_{(10,1)}E_{(9,1)}\mathbf{w}_4$

(see Figure 7) with the Taylor approximation

$$\begin{aligned}
 U_t &= e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s \\
 &+ \int_{t_0}^t e^{A(t-s)} F'(U_{t_0})(e^{A\Delta s} - I)U_{t_0} ds \\
 &+ \frac{1}{2} \int_{t_0}^t e^{A(t-s)} F''(U_{t_0})((e^{A\Delta s} - I)U_{t_0}, (e^{A\Delta s} - I)U_{t_0}) ds \\
 &+ \int_{t_0}^t e^{A(t-s)} F''(U_{t_0}) \left( (e^{A\Delta s} - I)U_{t_0}, \int_{t_0}^s e^{A(s-r)} B dW_r \right) ds \\
 &+ \frac{1}{2} \int_{t_0}^t e^{A(t-s)} F''(U_{t_0}) \left( \int_{t_0}^s e^{A(s-r)} B dW_r, \int_{t_0}^s e^{A(s-r)} B dW_r \right) ds \\
 &+ \int_{t_0}^t e^{A(t-s)} F'(U_{t_0}) \int_{t_0}^s e^{A(s-r)} B dW_r ds + O((\Delta t)^{(1+3 \min(\gamma, \delta, 1/3))}).
 \end{aligned}$$

REMARK 1. Not all Taylor expansions for general finite-dimensional SODEs in [29] are used in practice due to cost and difficulty of computing the higher iterated integrals in the expansions. For SODEs with additive noise, however, the Wagner–Platen scheme is often used since the iterated integrals appearing in it are linear functionals of the Brownian motion process, thus Gaussian distributed and hence easy to simulate. A similar situation holds for the above Taylor expansions of SPDEs. In particular, the conditional distribution [with respect to  $F'(U_{t_0})$ ] of the expression

$$\int_{t_0}^t e^{A(t-s)} F'(U_{t_0}) \int_{t_0}^s e^{A(s-r)} B dW_r ds$$

for  $t \in [t_0, T]$  in Section 5.1.4 is Gaussian distributed and, in principle, easy to simulate (see also Section 6).

5.2. *Taylor expansions for finite-dimensional SODEs.* Of course, the abstract setting for stochastic partial differential equations of evolutionary type in Section 2 in particular covers the case of finite-dimensional SODEs with additive noise. The main purpose of the Taylor expansions in this article is to overcome the need of an Itô formula in the infinite-dimensional setting. In contrast, in the finite-dimensional case, Itô’s formula is available and the whole machinery developed here is not needed. Nevertheless, we apply in this subsection the Taylor expansions introduced above to stochastic ordinary differential equations with additive noise to compare them with the well-known stochastic Taylor expansions for SODEs in the monograph [29]. These considerations are not so relevant in the view of applications, since the finite-dimensional case is well studied in the literature (see, e.g., [35] and the above named monograph), but more for a theoretical understanding of

the new Taylor expansions introduced here. More precisely, only in this subsection let  $H = \mathbb{R}^d$  with  $d \in \mathbb{N}$  be the  $d$ -dimensional  $\mathbb{R}$ -Hilbert space of real  $d$ -tuples with the scalar product

$$\langle v, w \rangle = v_1 \cdot w_1 + \dots + v_d \cdot w_d$$

for every  $v = (v_1, \dots, v_d) \in H$  and every  $w = (w_1, \dots, w_d) \in H$ . Let also  $U = \mathbb{R}^m$  with  $m \in \mathbb{N}$  and suppose that  $(W_t)_{t \in [0, T]}$  is a standard  $m$ -dimensional Brownian motion. Furthermore, we suppose that the eigenfunctions and the eigenvalues of the linear operator  $-A$  in Assumption 1 are given by  $e_1 = (1, 0, \dots, 0) \in H$ ,  $e_2 = (0, 1, 0, \dots, 0) \in H, \dots, e_m = (0, \dots, 0, 1) \in H$  and  $\lambda_1 = \lambda_2 = \dots = \lambda_d = 0$  with  $\mathcal{I} = \{1, \dots, d\}$ . So, in this case  $A$  is of course a boring bounded linear operator with  $D(A) = H = \mathbb{R}^d$  and  $Av = 0$  for every  $v \in D(A)$ . Furthermore, note that  $D((\kappa - A)^r) = H = \mathbb{R}^d$  for every  $r \in \mathbb{R}$  with an arbitrary  $\kappa > 0$ . The bounded linear mapping  $B : \mathbb{R}^m \rightarrow \mathbb{R}^d$  is then a  $d \times m$ -matrix. Due to Assumption 2, the drift term  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is then a smooth function with globally bounded derivatives as it is assumed in [29]. The initial value  $x_0 : \Omega \rightarrow \mathbb{R}^d$  is then simply a  $\mathcal{F}_0/\mathcal{B}(\mathbb{R}^d)$ -measurable mapping, which satisfies  $\mathbb{E}|x_0|^p < \infty$  for every  $p \in [1, \infty)$ . So, the SPDE (1) is in that case in fact a SODE and is given by

$$(23) \quad dU_t = F(U_t) dt + B dW_t, \quad U_0 = u_0,$$

for  $t \in [0, T]$ . Now, we apply the abstract Taylor expansions introduced above to that simple example. Therefore, note that the parameters in Assumption 3 are given by  $\gamma = 1 - \varepsilon$  and  $\delta = \frac{1}{2}$  for every arbitrarily small  $\varepsilon \in (0, 1)$ . First of all, we have

$$U_t = U_{t_0} + B \cdot (W_t - W_{t_0}) + O(\Delta t)$$

(see Section 5.1.1). Thus, this Taylor approximation corresponds in the case of finite-dimensional SODEs to the Taylor approximation for SODEs with the multi-index set

$$\mathcal{A} = \{v, (1), (2), \dots, (m)\}$$

in Theorem 5.5.1 in [29]. Here, we only mention the multi-index set, which uniquely determines the stochastic Taylor expansion in [29] and refer to the above named monograph for a detailed description of the stochastic Taylor expansions for SODEs there.

The exponential Euler approximation in Section 5.1.2 yields

$$(24) \quad U_t = U_{t_0} + F(U_{t_0}) \cdot (t - t_0) + B \cdot (W_t - W_{t_0}) + O((\Delta t)^{3/2}).$$

This is nothing else than the corresponding one-step approximation of the classical Euler–Maruyama scheme (see Section 10.2 in [29]) and is in the setting of [29] given by the multi-index set

$$\mathcal{A} = \{v, (0), (1), (2), \dots, (m)\}$$

in Theorem 5.5.1 there. In that sense, the name of the exponential Euler scheme is indeed justified. While in this article, the Taylor approximation (24) is obtained via an expansion of the  $I_j^i$ -operators (see Lemma 1 and Section 3.1), in [29] the stochastic Taylor approximation (24) is achieved by applying Itô’s formula to the integrand  $F(U_t)$  in the SODE (23). Finally, the Taylor approximation in Section 5.1.4 reduces to

$$U_t = U_{t_0} + F(U_{t_0}) \cdot (t - t_0) + B \cdot (W_t - W_{t_0}) + F'(U_{t_0}) \cdot B \cdot \left( \int_{t_0}^t \int_{t_0}^s dW_r ds \right) + O((\Delta t)^2).$$

The approximation above is nothing else than the one-step approximation of the stochastic Taylor approximation given by the multi-index set

$$\mathcal{A} = \left\{ \begin{array}{l} v, (0), (1), (2), \dots, (m), \\ (1, 0), (2, 0), \dots, (m, 0), \\ (1, 1), (2, 1), \dots, (m, 1), \\ \vdots \\ (1, m), (2, 1), \dots, (m, m) \end{array} \right\}$$

in Theorem 5.5.1 in [29]. In [29], it is obtained via again applying Itô’s formula.

To sum up, although the method for deriving Taylor expansions in this article ( $I_j^i$ -operators) is different to the method in [29] (Itô’s formula), the resulting Taylor approximations above coincide.

5.3. *Simultaneous diagonalizable case.* We illustrate Assumption 3 with the case where  $A$  and  $B$  are simultaneous diagonalizable (see, for example, Section 5.5.1 in [6]). This assumption is commonly considered in the literature for approximations of SPDEs (see, e.g., Section 2 in [33] or see also [24, 32, 34]). Suppose that  $U = H$  and that  $B : H \rightarrow H$  is given by

$$Bv = \sum_{i \in \mathcal{I}} b_i \langle e_i, v \rangle e_i \quad \forall v \in H,$$

where  $b_i, i \in \mathcal{I}$ , is a bounded family of real numbers and  $e_i, i \in \mathcal{I}$ , is the family of eigenfunctions of the operator  $A$  (see Assumption 1). Concerning Assumption 3, note that

$$\begin{aligned} & \int_0^T |(\kappa - A)^\gamma e^{As} B|_{\text{HS}}^2 ds \\ &= \sum_{i \in \mathcal{I}} (\kappa + \lambda_i)^{2\gamma} b_i^2 \left( \int_0^T e^{-2\lambda_i s} ds \right) \\ &\leq \sum_{i \in \mathcal{I}} (\kappa + \lambda_i)^{2\gamma} b_i^2 \left( \int_0^T e^{-2\lambda_i s} e^{-2\kappa s} ds \right) e^{2\kappa T} \end{aligned}$$

$$\begin{aligned} &= \sum_{i \in \mathcal{I}} (\kappa + \lambda_i)^{2\gamma} b_i^2 \left( \int_0^T e^{-2(\kappa + \lambda_i)s} ds \right) e^{2\kappa T} \\ &= \frac{e^{2\kappa T}}{2} \left( \sum_{i \in \mathcal{I}} b_i^2 (\kappa + \lambda_i)^{(2\gamma-1)} (1 - e^{-2(\lambda_i + \kappa)T}) \right) \\ &\leq \frac{e^{2\kappa T}}{2} \left( \sum_{i \in \mathcal{I}} b_i^2 (\kappa + \lambda_i)^{(2\gamma-1)} \right) \end{aligned}$$

for a given  $\gamma > 0$ , so

$$\int_0^T |(\kappa - A)^\gamma e^{As} B|_{\text{HS}}^2 < \infty$$

follows from

$$(25) \quad \sum_{i \in \mathcal{I}} b_i^2 (\kappa + \lambda_i)^{(2\gamma-1)} < \infty$$

for a given  $\gamma > 0$ . In this case, we also have

$$(26) \quad \int_0^t |e^{As} B|_{\text{HS}}^2 ds \leq Ct^{2\delta}$$

for every  $t \in [0, 1]$  with  $\delta := \min(\gamma, \frac{1}{2})$  and a constant  $C > 0$ .

5.4. *Space-time white noise.* This example will be a special case of the previous one. Let  $H = L^2((0, 1), \mathbb{R})$  be the space of equivalence classes of square integrable measurable functions from the interval  $(0, 1)$  to  $\mathbb{R}$  with the scalar product and the norm

$$\langle u, v \rangle = \int_0^1 u(x)v(x) dx, \quad |u| = \sqrt{\int_0^1 |u(x)|^2 dx}, \quad u, v \in H.$$

Let  $U = H$  and let  $B = I : H \rightarrow H$  be the identity operator. In addition, assume that  $\alpha : (0, 1) \rightarrow \mathbb{R}$  is a bounded measurable function and that the operator  $F : H \rightarrow H$  is given by

$$F(v)(x) := (Fv)(x) := \alpha(x) \cdot v(x), \quad x \in (0, 1),$$

for all  $v \in H$ , which clearly satisfies Assumption 2. Also note that

$$F'(v)w = F(w) \quad \text{and} \quad F^{(i)}(v)(w_1, \dots, w_i) = 0$$

for all  $v, w, w_1, \dots, w_i \in H$  and all  $i \in \{2, 3, \dots\}$ . Furthermore, let  $A = \frac{\partial^2}{\partial x^2} : D(A) \subset H \rightarrow H$  be the Laplace operator with Dirichlet boundary condition, that is,

$$Au = \sum_{n=1}^\infty -\lambda_n \langle e_n, u \rangle e_n, \quad u \in H,$$

where

$$\lambda_n = \pi^2 n^2, \quad e_n(x) = \sqrt{2} \sin(n\pi x), \quad x \in (0, 1),$$

for each  $n \in \mathcal{I} := \mathbb{N}$ . Of course, the  $e_n$ ,  $n \in \mathbb{N}$ , form an orthonormal basis of  $H$  (Assumption 1). Additionally, we choose  $\kappa = 0$ .

Let  $t_0 = 0$  and  $T = 1$ . In view of (25), Assumption 3 requires  $\gamma = \frac{1}{4} - \varepsilon$  for every arbitrarily small  $\varepsilon > 0$ . However, instead of (26), we obtain here the stronger result  $\delta = \frac{1}{4}$ , since

$$\|e^{As}\|_{\text{HS}} \leq C \left(\frac{1}{s}\right)^{1/4}$$

for every  $s \in (0, 1]$  and a constant  $C > 0$  (see Remark 2 in [32]). Finally, let  $u_0 \in H$  be an arbitrary (deterministic) function in  $H$ , which satisfies Assumption 4. The SPDE (1) is then given by

$$dU_t(x) = \left[ \frac{\partial^2}{\partial x^2} U_t(x) + \alpha(x)U_t(x) \right] dt + dW_t, \quad U_t(0) = U_t(1) = 0,$$

with  $U_0(x) = u_0(x)$  for  $x \in (0, 1)$  and  $t \in [0, 1]$ . After considering Assumptions 1–4 for this example, we now present the Taylor approximations in this case. Here,  $\varepsilon \in (0, \frac{1}{4})$  is always an arbitrarily small real number in  $(0, \frac{1}{4})$ .

5.4.1. *Taylor expansion of order 1.* For an approximation of  $U_t$  of order one for small  $t > 0$ , we obtain

$$U_t = e^{At}u_0 + \int_0^t e^{A(t-s)} dW_s + O(\Delta t)$$

(see Section 5.1.1).

5.4.2. *Taylor expansion of order  $\frac{5}{4} - \varepsilon$ .* Here, we have

$$U_t = e^{At}u_0 + A^{-1}(e^{At} - I)Fu_0 + \int_0^t e^{A(t-s)} dW_s + O((\Delta t)^{(5/4-\varepsilon)})$$

for an approximation of order  $\frac{5}{4} - \varepsilon$  (see Section 5.1.2).

5.4.3. *Taylor expansion of order  $\frac{5}{4}$ .* In the next step, we obtain

$$U_t = e^{At}u_0 + A^{-1}(e^{At} - I)Fu_0 + \int_0^t e^{A(t-s)} dW_s + \left( \int_0^t e^{A(t-s)} F(e^{As} - I) ds \right) u_0 + O((\Delta t)^{5/4})$$

for an approximation of order  $\frac{5}{4}$  (see Section 5.1.3).

5.4.4. *Taylor expansion of order  $\frac{3}{2} - \varepsilon$ .* Here, we have

$$U_t = e^{At}u_0 + A^{-1}(e^{At} - I)Fu_0 + \int_0^t e^{A(t-s)}F\left(\int_0^s e^{A(s-r)}dW_r\right)ds + \left(\int_0^t e^{A(t-s)}F(e^{As} - I)ds\right)u_0 + \int_0^t e^{A(t-s)}dW_s + O((\Delta t)^{(3/2-\varepsilon)})$$

for an approximation of order  $\frac{3}{2} - \varepsilon$  (see Section 5.1.4).

5.4.5. *Taylor expansion of order  $\frac{7}{4} - \varepsilon$ .* Since  $F$  is linear here with  $F'(v) \equiv F$  and  $F''(v) \equiv 0$  for all  $v \in H$ , the approximation above is even more of order  $\frac{7}{4} - \varepsilon$ , that is,

$$U_t = e^{At}u_0 + A^{-1}(e^{At} - I)Fu_0 + \int_0^t e^{A(t-s)}F\left(\int_0^s e^{A(s-r)}dW_r\right)ds + \left(\int_0^t e^{A(t-s)}F(e^{As} - I)ds\right)u_0 + \int_0^t e^{A(t-s)}dW_s + O((\Delta t)^{(7/4-\varepsilon)})$$

(see Section 5.1.6).

5.4.6. *Taylor of order  $2 - \varepsilon$ .* We also consider the Taylor expansion given by the stochastic wood

$$E_{(24,1)}E_{(22,1)}E_{(21,1)}E_{(19,1)}E_{(18,1)}E_{(16,1)}E_{(15,1)}E_{(13,1)}\mathbf{w}_5,$$

where  $\mathbf{w}_5$  is presented in Figure 7. Since  $F$  is linear here, we see that the corresponding Taylor approximation is the same as in the both examples above, so we obtain

$$U_t = e^{At}u_0 + A^{-1}(e^{At} - I)Fu_0 + \int_0^t e^{A(t-s)}F\left(\int_0^s e^{A(s-r)}dW_r\right)ds + \left(\int_0^t e^{A(t-s)}F(e^{As} - I)ds\right)u_0 + \int_0^t e^{A(t-s)}dW_s + O((\Delta t)^{(2-\varepsilon)}).$$

By further expansions, one can show that this approximation is in fact of order 2.

5.5. *Trace class noise.* In this subsection, we compute the smoothness parameters  $\gamma$  and  $\delta$  in Assumption 3 for the case of trace class noise (see, e.g., Sections 4.1 and 5.4.1 in [6]). This assumption is also commonly considered in the literature for approximations of SPDEs (see, e.g., [15, 33]). Precisely, we suppose that  $B : U \rightarrow H$  is a Hilbert–Schmidt operator, that is,  $|B|_{\text{HS}} < \infty$ . Hence, we obtain

$$\int_0^t |e^{As}B|_{\text{HS}}^2 ds \leq \int_0^t (|e^{As}|^2 \cdot |B|_{\text{HS}}^2) ds \leq e^{2\kappa} |B|_{\text{HS}}^2 t$$



and therefore

$$\sqrt{\int_0^t |e^{As} B|_{\text{HS}}^2 ds} \leq e^\kappa |B|_{\text{HS}} \sqrt{t}$$

for all  $t \in [0, 1]$ . Moreover, we have

$$\begin{aligned} \int_0^T |(\kappa - A)^r e^{As} B|_{\text{HS}}^2 ds &\leq \int_0^T |(\kappa - A)^r e^{As}|^2 |B|_{\text{HS}}^2 ds \\ &= \left( \int_0^T |(\kappa - A)^r e^{(A-\kappa)s} e^{\kappa s}|^2 ds \right) |B|_{\text{HS}}^2 \\ &\leq \left( \int_0^T |(\kappa - A)^r e^{(A-\kappa)s}|^2 ds \right) e^{2\kappa T} |B|_{\text{HS}}^2 \\ &\leq \left( \int_0^T s^{-2r} ds \right) e^{2\kappa T} |B|_{\text{HS}}^2 < \infty \end{aligned}$$

for all  $r \in [0, \frac{1}{2})$ . Hence, we obtain  $\gamma = \frac{1}{2} - \varepsilon$  and  $\delta = \frac{1}{2}$  for every arbitrarily small  $\varepsilon \in (0, \frac{1}{2})$  in this situation. Now, we present the Taylor expansions from Section 5.1 again in this special situation.

5.5.1. *Taylor expansion of order 1.* Here, we have

$$U_t = e^{A\Delta t} U_{t_0} + \int_{t_0}^t e^{A(t-s)} B dW_s + O(\Delta t)$$

for a Taylor approximation of order 1 (see Section 5.1.1).

5.5.2. *Taylor expansion of order  $\frac{3}{2} - \varepsilon$ .* For a Taylor approximation of order  $\frac{3}{2} - \varepsilon$  (see Section 5.1.2), we obtain

$$U_t = e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s + O((\Delta t)^{(3/2-\varepsilon)}).$$

Here and below,  $\varepsilon \in (0, \frac{1}{2})$  is an arbitrarily small real number in  $(0, \frac{1}{2})$ .

5.5.3. *Taylor expansion of order  $\frac{3}{2}$ .* The Taylor approximation in Section 5.1.3 reduces to

$$\begin{aligned} U_t &= e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s \\ &\quad + \int_{t_0}^t e^{A(t-s)} F'(U_{t_0})(e^{A\Delta s} - I)U_{t_0} ds + O((\Delta t)^{3/2}). \end{aligned}$$

5.5.4. *Taylor expansion of order 2 - ε.* Here, we obtain

$$\begin{aligned}
 U_t &= e^{A\Delta t} U_{t_0} + \left( \int_0^{\Delta t} e^{As} ds \right) F(U_{t_0}) + \int_{t_0}^t e^{A(t-s)} B dW_s \\
 &\quad + \int_{t_0}^t e^{A(t-s)} F'(U_{t_0})(e^{A\Delta s} - I)U_{t_0} ds \\
 &\quad + \int_{t_0}^t e^{A(t-s)} F'(U_{t_0}) \int_{t_0}^s e^{A(s-r)} B dW_r ds + O((\Delta t)^{(2-\varepsilon)})
 \end{aligned}$$

for a Taylor expansion of order 2 - ε. This example corresponds to the Taylor expansion introduced in Section 5.1.4.

5.6. *A special example of trace class noise.* Let  $H = U = L^2((0, 1)^3, \mathbb{R})$  be the space of equivalence classes of square integrable measurable functions from  $(0, 1)^3$  to  $\mathbb{R}$  and consider two distinct Hilbert bases  $e_i, i \in \mathcal{I} := \mathbb{N}^3$ , and  $f_i, i \in \mathcal{I}$ , in  $H$  given by

$$e_i(x_1, x_2, x_3) = 2^{3/2} \sin(i_1\pi x_1) \sin(i_2\pi x_2) \sin(i_3\pi x_3)$$

and

$$\begin{aligned}
 f_i(x_1, x_2, x_3) \\
 &= c_{(i_1-1)} c_{(i_2-1)} c_{(i_3-1)} \cos((i_1 - 1)\pi x_1) \cos((i_2 - 1)\pi x_2) \cos((i_3 - 1)\pi x_3)
 \end{aligned}$$

for every  $i = (i_1, i_2, i_3) \in \mathcal{I} = \mathbb{N}^3$  and every  $x = (x_1, x_2, x_3) \in (0, 1)^3$ , where  $c_n := \sqrt{2}$  for every  $n \in \mathbb{N}$  and  $c_0 = 1$ . Then, consider the Hilbert-Schmidt operator  $B : U \rightarrow H$  given by

$$Bu = \sum_{i \in \mathbb{N}^3} \frac{\langle f_i, u \rangle}{(i_1 \cdot i_2 \cdot i_3)} e_i$$

for all  $u \in U = H$ . Moreover, let  $\lambda_i, i \in \mathbb{N}^3$ , be a family of real numbers given by  $\lambda_i = \pi^2(i_1^2 + i_2^2 + i_3^2)$  for all  $i = (i_1, i_2, i_3) \in \mathbb{N}^3$ . Finally, consider  $A = (\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}) : D(A) \subset H \rightarrow H$  (with Dirichlet boundary conditions) given by

$$Av = \sum_{i \in \mathbb{N}^3} -\lambda_i \langle e_i, v \rangle e_i$$

for all  $v \in D(A)$ , where  $D(A)$  is given by

$$D(A) = \left\{ v \in H \mid \sum_{i \in \mathbb{N}^3} (i_1^2 + i_2^2 + i_3^2) |\langle e_i, v \rangle|^2 < \infty \right\}.$$

Then, the SPDE (1) reduces to

$$dU_t(x) = \left[ \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) U_t(x) + F(U_t(x)) \right] dt + \sqrt{Q} dW_t(x)$$

with  $U|_{\partial(0,1)^3} = 0$  for  $x \in (0, 1)^3$  and  $t \in [0, T]$ . Assumptions 1–4 are fulfilled with  $\delta = \frac{1}{2}$  and  $\gamma = \frac{1}{2} - \varepsilon$  for every arbitrarily small  $\varepsilon \in (0, \frac{1}{2})$  (see Section 5.5). The Taylor approximations in that situation are presented in Section 5.5.

**6. Numerical schemes based on the Taylor expansions.** In this section, some numerical schemes based on the Taylor expansions in this article are presented. We refer to [21, 22, 24, 25] for estimations of the convergence orders of these schemes and also for numerical simulations for these schemes.

For numerical approximations of SPDEs, one has to discretize both the time interval  $[0, T]$  and the  $\mathbb{R}$ -Hilbert space  $H$ . For the discretization of the space  $H$ , we use a spectral Galerkin approximation based on the eigenfunctions of the linear operator  $A : D(A) \subset H \rightarrow H$ . More precisely, let  $(\mathcal{I}_N)_{N \in \mathbb{N}}$  be a sequence of increasing finite nonempty subsets of  $\mathcal{I}$ , that is,  $\emptyset \neq \mathcal{I}_N \subset \mathcal{I}_M \subset \mathcal{I}$  for all  $N, M \in \mathbb{N}$  with  $N \leq M$  and let  $H_N := \text{span}\langle e_i, i \in \mathcal{I}_N \rangle$  be the finite-dimensional span of  $|\mathcal{I}_N|$ -eigenfunctions for  $N \in \mathbb{N}$ . The bounded linear mappings  $P_N : H \rightarrow H_N$  are then given by  $P_N(v) = \sum_{i \in \mathcal{I}_N} \langle e_i, v \rangle e_i$  for every  $v \in H$  and every  $N \in \mathbb{N}$ .

6.1. *The exponential Euler scheme.* Based on the Taylor approximation in Sections 3.2 and 5.1.2, we consider the family of random variables  $Y_k^{N,M} : \Omega \rightarrow H_N, k = 0, 1, \dots, M, N, M \in \mathbb{N}$ , given by  $Y_0^{N,M} = P_N(u_0)$  and

$$(27) \quad \begin{aligned} Y_{k+1}^{N,M} &= e^{AT/M} Y_k^{N,M} + \left( \int_0^{T/M} e^{As} ds \right) (P_N F)(Y_k^{N,M}) \\ &\quad + P_N \left( \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} B dW_s \right) \end{aligned}$$

for every  $k = 0, 1, \dots, M - 1$  and every  $N, M \in \mathbb{N}$ . This scheme is introduced and analyzed in [24]. As already mentioned, it is called the *exponential Euler scheme* there.

In the setting of deterministic PDEs, that is, in the case  $B = 0$ , this scheme reduces to

$$Y_{k+1}^{N,M} = e^{AT/M} Y_k^{N,M} + \left( \int_0^{T/M} e^{As} ds \right) (P_N F)(Y_k^{N,M})$$

for every  $k = 0, 1, \dots, M - 1$  and every  $N, M \in \mathbb{N}$ . This scheme and similar schemes, usually referred as exponential integrators, have for deterministic PDEs been intensively studied in the literature (see, e.g., [3–5, 16–19, 27, 30, 31, 36]). Such schemes are easier to simulate than may seem on the first sight (see [4]). In the stochastic setting, we refer to Sections 3 and 4 in [24] for a detailed description for the simulation of the scheme (27), in particular, for the generation of the random variables used there.

6.2. *The Taylor scheme indicated by Section 5.1.3.* In view of Section 5.1.3, we obtain the Taylor scheme  $Y_k^{N,M} : \Omega \rightarrow H_N, k = 0, 1, \dots, M, N, M \in \mathbb{N}$ , given by  $Y_0^{N,M} = P_N(u_0)$  and

$$\begin{aligned} Y_{k+1}^{N,M} &= e^{AT/M} Y_k^{N,M} + \left( \int_0^{T/M} e^{As} ds \right) (P_N F)(Y_k^{N,M}) \\ &\quad + \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} (P_N F')(Y_k^{N,M}) \\ &\quad \quad \quad \times ((e^{A(s-kT/M)} - I) Y_k^{N,M}) ds \\ &\quad + P_N \left( \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} B dW_s \right) \end{aligned}$$

for every  $k = 0, 1, \dots, M - 1$  and every  $N, M \in \mathbb{N}$ .

6.3. *The Taylor scheme indicated by Section 5.1.4.* The Taylor approximation in Section 5.1.4 yields the Taylor scheme  $Y_k^{N,M} : \Omega \rightarrow H_N, k = 0, 1, \dots, M, N, M \in \mathbb{N}$ , given by  $Y_0^{N,M} = P_N(u_0)$  and

$$\begin{aligned} Y_{k+1}^{N,M} &= e^{AT/M} Y_k^{N,M} + \left( \int_0^{T/M} e^{As} ds \right) (P_N F)(Y_k^{N,M}) \\ &\quad + \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} (P_N F')(Y_k^{N,M}) \\ &\quad \quad \quad \times ((e^{A(s-kT/M)} - I) Y_k^{N,M}) ds \\ &\quad + \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} (P_N F')(Y_k^{N,M}) \\ &\quad \quad \quad \times \left( P_N \left( \int_{kT/M}^s e^{A(s-u)} B dW_u \right) \right) ds \\ &\quad + P_N \left( \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} B dW_s \right) \end{aligned}$$

for every  $k = 0, 1, \dots, M - 1$  and every  $N, M \in \mathbb{N}$ .

6.4. *A Runge–Kutta scheme for SPDEs.* In principle, we can proceed with the next Taylor approximations and obtain numerical schemes of higher order. These schemes would however be of limited practical use due to cost and difficulty of computing the higher iterated integrals as well as the higher order derivatives in the Taylor approximations. Therefore, we follow a different approach and derive a derivative free numerical scheme with simple integrals—a so called *Runge–Kutta scheme for SPDEs*. We would like to mention that this way is the usual procedure

for numerical schemes for differential equations: Taylor expansions and their corresponding Taylor schemes provide the underlying theory for deriving numerical schemes, but are rarely implemented in practice. Instead of these Taylor schemes other numerical schemes, which are easier to compute but still depend on the Taylor expansions such as Runge–Kutta schemes or multi-step schemes (see, e.g., [8] for details) are used.

To derive a Runge–Kutta scheme for SPDEs, we consider the Taylor approximation in Section 5.1.4 (see also the Taylor scheme above) from  $\frac{kT}{M}$  to  $\frac{(k+1)T}{M}$  and obtain

$$\begin{aligned} U_{(k+1)h} &\approx e^{Ah} U_{kh} + \left( \int_0^h e^{As} ds \right) F(U_{kh}) + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} B dW_s \\ &\quad + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} F'(U_{kh}) ((e^{A(s-kh)} - I) U_{kh}) ds \\ &\quad + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} F'(U_{kh}) \left( \int_{kh}^s e^{A(s-r)} B dW_r \right) ds \end{aligned}$$

and hence

$$\begin{aligned} U_{(k+1)h} &\approx e^{Ah} U_{kh} + he^{Ah} F(U_{kh}) + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} B dW_s \\ &\quad + he^{Ah} F'(U_{kh}) \left( \frac{1}{h} \int_{kh}^{(k+1)h} (e^{A(s-kh)} - I) U_{kh} ds \right) \\ &\quad + he^{Ah} F'(U_{kh}) \left( \frac{1}{h} \int_{kh}^{(k+1)h} \int_{kh}^s e^{A(s-r)} B dW_r ds \right) \\ &\approx e^{Ah} U_{kh} + he^{Ah} F(U_{kh}) + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} B dW_s \\ &\quad + he^{Ah} F'(U_{kh}) \left[ \frac{1}{h} \int_{kh}^{(k+1)h} (e^{A(s-kh)} - I) \int_0^{kh} e^{A(kh-r)} B dW_r ds \right] \\ &\quad + he^{Ah} F'(U_{kh}) \left( \frac{1}{h} \int_{kh}^{(k+1)h} \int_{kh}^s e^{A(s-r)} B dW_r ds \right) \end{aligned}$$

with  $h := \frac{T}{M}$  for  $k = 0, 1, \dots, M - 1$  and  $M \in \mathbb{N}$ . This yields

$$U_{(k+1)h} \approx e^{Ah} U_{kh} + he^{Ah} F(U_{kh} + Z_k^M) + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} B dW_s$$

with the random variables

$$\begin{aligned} Z_k^M &= \frac{1}{h} \int_{kh}^{(k+1)h} (e^{A(s-kh)} - I) \int_0^{kh} e^{A(kh-r)} B dW_r ds \\ &\quad + \frac{1}{h} \int_{kh}^{(k+1)h} \int_{kh}^s e^{A(s-r)} B dW_r ds \end{aligned}$$

for  $k = 0, 1, \dots, M - 1$  and  $M \in \mathbb{N}$ . The corresponding numerical scheme  $Y_k^{N,M} : \Omega \rightarrow H_N, k = 0, 1, \dots, M, N, M \in \mathbb{N}$ , is then given by  $Y_0^{N,M} = P_N(u_0)$  and

$$Y_{k+1}^{N,M} = e^{AT/M} Y_k^{N,M} + \frac{T}{M} e^{AT/M} (P_N F)(Y_k^{N,M} + P_N(Z_k^M)) + P_N \left( \int_{kT/M}^{(k+1)T/M} e^{A((k+1)T/M-s)} B dW_s \right)$$

for every  $k = 0, 1, \dots, M - 1$  and every  $N, M \in \mathbb{N}$ . This Runge–Kutta scheme for SPDEs is introduced and analyzed in [21]. Under non-global Lipschitz coefficients of the SPDE, it is analyzed in [22]. Note that the random variables occurring in the scheme above are Gaussian distributed and therefore easy to simulate (see also Remark 1 and [21, 22] for details). More precisely, in the case of one-dimensional stochastic reaction diffusion equations with space–time white noise it is shown in the articles cited above that this scheme converges with the overall order  $\frac{1}{4}$ —with respect to the number of independent standard normal distributed random variables and the number of arithmetical operations used to compute the scheme instead of the overall order  $\frac{1}{6}$  of classical numerical schemes (see, for instance, [7, 12, 13, 42]) such as the linear implicit Euler scheme.

**7. Proofs.**

7.1. *Proofs of (6) and (7).*

LEMMA 1. *Let Assumptions 1–4 be fulfilled and let  $i \in \mathbb{N}$  be given. Then, we have*

$$I_{1*}^0 = I_1^0 + I_{1*}^1[I_0^0] + I_{1*}^1[I_{1*}^0] + I_{1*}^1[I_2^0]$$

and

$$I_{1*}^i[g_1, \dots, g_i] = I_1^i[g_1, \dots, g_i] + I_{1*}^{(i+1)}[I_0^0, g_1, \dots, g_i] + I_{1*}^{(i+1)}[I_{1*}^0, g_1, \dots, g_i] + I_{1*}^{(i+1)}[I_2^0, g_1, \dots, g_i]$$

for all  $g_1, \dots, g_i \in \mathcal{P}$ .

PROOF. We begin with the first equation. Since we have

$$\begin{aligned} F(U_s) &= F(U_{t_0}) + \int_0^1 F'(U_{t_0} + r(U_s - U_{t_0}))(U_s - U_{t_0}) dr \\ &= F(U_{t_0}) + \int_0^1 F'(U_{t_0} + r\Delta U_s)(\Delta U_s) dr \\ &= F(U_{t_0}) + \int_0^1 F'(U_{t_0} + r\Delta U_s)(I_0^0(s)) dr \\ &\quad + \int_0^1 F'(U_{t_0} + r\Delta U_s)(I_{1*}^0(s)) dr + \int_0^1 F'(U_{t_0} + r\Delta U_s)(I_2^0(s)) dr \end{aligned}$$

for every  $s \in [t_0, T]$  due to the fundamental theorem of calculus and (5), we obtain

$$\begin{aligned} I_{1*}^0(t) &= \int_{t_0}^t e^{A(t-s)} F(U_s) ds \\ &= \int_{t_0}^t e^{A(t-s)} F(U_{t_0}) ds + \int_{t_0}^t e^{A(t-s)} \left( \int_0^1 F'(U_{t_0} + r \Delta U_s)(I_0^0(s)) dr \right) ds \\ &\quad + \int_{t_0}^t e^{A(t-s)} \left( \int_0^1 F'(U_{t_0} + r \Delta U_s)(I_{1*}^0(s)) dr \right) ds \\ &\quad + \int_{t_0}^t e^{A(t-s)} \left( \int_0^1 F'(U_{t_0} + r \Delta U_s)(I_2^0(s)) dr \right) ds, \end{aligned}$$

which implies

$$I_{1*}^0(t) = I_1^0(t) + I_{1*}^1[I_0^0](t) + I_{1*}^1[I_{1*}^0](t) + I_{1*}^1[I_2^0](t)$$

for all  $t \in [t_0, T]$ . Moreover, we have

$$\begin{aligned} &\int_0^1 F^{(i)}(U_{t_0} + r \Delta U_s)(g_1(s), \dots, g_i(s)) \frac{(1-r)^{(i-1)}}{(i-1)!} dr \\ &= \left[ -F^{(i)}(U_{t_0} + r \Delta U_s)(g_1(s), \dots, g_i(s)) \frac{(1-r)^i}{i!} \right]_{r=0}^{r=1} \\ &\quad + \int_0^1 F^{(i+1)}(U_{t_0} + r \Delta U_s)(\Delta U_s, g_1(s), \dots, g_i(s)) \frac{(1-r)^i}{i!} dr \\ &= \frac{1}{i!} F^{(i)}(U_{t_0})(g_1(s), \dots, g_i(s)) \\ &\quad + \int_0^1 F^{(i+1)}(U_{t_0} + r \Delta U_s)(\Delta U_s, g_1(s), \dots, g_i(s)) \frac{(1-r)^i}{i!} dr \end{aligned}$$

for all  $s \in [t_0, T]$  and all  $g_1, \dots, g_i \in \mathcal{P}$  due to integration by parts and therefore, we also obtain the second equation.  $\square$

7.2. *Proof of Theorem 1.* For the proof of Theorem 1, we need the following lemma.

LEMMA 2. *Let  $X : [t_0, T] \times \Omega \rightarrow [0, \infty)$  be a predictable stochastic process. Then, we obtain*

$$\left| \int_{t_0}^t X_s ds \right|_{L^r} \leq \int_{t_0}^t |X_s|_{L^r} ds$$

for every  $t \in [t_0, T]$  and every  $r \in [1, \infty)$ , where both sides could be infinite.

PROOF. First, we consider the case, where  $X_t \leq C$  is bounded by a constant  $C > 0$  for all  $t \in [0, T]$ . Here, we have

$$\begin{aligned} \mathbb{E} \left[ \left( \int_{t_0}^t X_s ds \right)^r \right] &= \int_{t_0}^t \mathbb{E} \left[ \left( \int_{t_0}^t X_u du \right)^{(r-1)} X_s \right] ds \\ &\leq \int_{t_0}^t |X_s|_{L^r} ds \left( \mathbb{E} \left[ \left( \int_{t_0}^t X_u du \right)^r \right] \right)^{(r-1)/r} \end{aligned}$$

for every  $t \in [t_0, T]$  and every  $r \in [1, \infty)$  due to Hölder's inequality. Since

$$\mathbb{E} \left[ \left( \int_{t_0}^t X_u du \right)^r \right] < \infty$$

is finite for every  $t \in [t_0, T]$  and every  $r \in [1, \infty)$  due to the boundedness of  $X : [t_0, T] \times \Omega \rightarrow [0, \infty)$ , we obtain the assertion. In the general case, we can approximate the stochastic process  $(X_t)_{t \in [0, T]}$  by bounded processes  $(X_t^N)_{t \in [0, T]}$  for  $N \in \mathbb{N}$  given by

$$X_t^N := \min(N, X_t)$$

for all  $t \in [0, T]$  and all  $N \in \mathbb{N}$ . This shows the assertion.  $\square$

We also need the Burkholder–Davis–Gundy inequality in infinite dimensions (see Lemma 7.7 in [6]).

LEMMA 3. *Let  $X : [t_0, T] \times \Omega \rightarrow \text{HS}(U, H)$  be a predictable stochastic process with  $\mathbb{E} \int_{t_0}^T |X_s|_{\text{HS}}^2 ds < \infty$ . Then, we obtain*

$$\left| \int_{t_0}^t X_s dW_s \right|_{L^p} \leq p \left( \int_{t_0}^t \|X_s\|_{\text{HS}}^2 ds \right)^{1/2}$$

for every  $t \in [t_0, T]$  and every  $p \in [1, \infty)$ , where both sides could be infinite.

In view of the definitions of the mappings  $\Phi$  and  $\Psi$ , Theorem 1 immediately follows from the next lemma. For this, the subset  $\mathbf{ST}' \subset \mathbf{ST}$  of stochastic trees given by

$$\begin{aligned} \mathbf{ST}' := \left\{ \mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{ST} \mid \forall k \in \text{nd}(\mathbf{t}) : ((\exists l \in \text{nd}(\mathbf{t}) : \mathbf{t}'(l) = k) \right. \\ \left. \implies (\mathbf{t}''(k) \in \{1, 1^*\})) \right\} \end{aligned}$$

is used.

LEMMA 4. *Let  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{ST}'$  be an arbitrary stochastic tree in  $\mathbf{ST}'$ . Then, for each  $p \geq 1$ , there exists a constant  $C_p > 0$  such that*

$$(\mathbb{E}[|\Phi(\mathbf{t})(t)|^p])^{1/p} \leq C_p \cdot (t - t_0)^{\text{ord}(\mathbf{t})}$$

holds for all  $t \in [t_0, T]$ , where  $C_p$  is independent of  $t$  and  $t_0$  but depends on  $p, \mathbf{t}, T$  and the SPDE (1).



PROOF. Due to Jensen’s inequality, we can assume without loss of generality that  $p \in [2, \infty)$  holds. We will prove now the assertion by induction with respect to the number of nodes  $l(\mathbf{t}) \in \mathbb{N}$ .

In the base case  $l(\mathbf{t}) = 1$ , we have  $\Phi(\mathbf{t}) = I_{\mathbf{t}''(1)}^0$  by definition. Hence, we obtain

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &= |I_{\mathbf{t}''(1)}^0(t)|_{L^p} = |I_0^0(t)|_{L^p} = |(e^{A\Delta t} - I)U_{t_0}|_{L^p} \\ &\leq |(\kappa - A)^{-\gamma}(e^{A\Delta t} - I)| \cdot |(\kappa - A)^\gamma U_{t_0}|_{L^p} \\ &= |(\kappa - A)^{-\gamma}(e^{(A-\kappa)\Delta t} - e^{-\kappa\Delta t})| \cdot e^{\kappa\Delta t} \cdot |(\kappa - A)^\gamma U_{t_0}|_{L^p} \\ &\leq |(\kappa - A)^{-\gamma}(e^{(A-\kappa)\Delta t} - e^{-\kappa\Delta t})| \cdot e^{\kappa T} \cdot \left( \sup_{0 \leq s \leq T} |(\kappa - A)^\gamma U_s|_{L^p} \right) \\ &\leq C_p \cdot |(\kappa - A)^{-\gamma}(e^{(A-\kappa)\Delta t} - e^{-\kappa\Delta t})| \end{aligned}$$

and therefore

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &\leq C_p \cdot (|(\kappa - A)^{-\gamma}(e^{(A-\kappa)\Delta t} - I)| + |(\kappa - A)^{-\gamma}(I - e^{-\kappa\Delta t})|) \\ &\leq C_p \cdot (|(\kappa - A)^{-\gamma}(e^{(A-\kappa)\Delta t} - I)| + |(I - e^{-\kappa\Delta t})|) \\ &\leq C_p \cdot (\Delta t)^\gamma + C_p \cdot (\Delta t) \leq C_p \cdot (\Delta t)^\gamma \end{aligned}$$

for every  $t \in [t_0, T]$  in the case  $\mathbf{t}''(1) = 0$ . Here and below,  $C_p > 0$  is a constant, which changes from line to line but is independent of  $t$  and  $t_0$ . Moreover, by Lemma 2, we obtain

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &= |I_{\mathbf{t}''(1)}^0(t)|_{L^p} = |I_{1^*}^0(t)|_{L^p} = \left| \int_{t_0}^t e^{A(t-s)} F(U_s) ds \right|_{L^p} \\ &\leq \int_{t_0}^t (|e^{A(t-s)} F(U_s)|_{L^p}) ds \leq C_p \cdot \left( \int_{t_0}^t |F(U_s)|_{L^p} ds \right) \\ &\leq C_p \cdot \left( \int_{t_0}^t (1 + |U_s|_{L^p}) ds \right) \leq C_p \cdot (\Delta t) \end{aligned}$$

for every  $t \in [t_0, T]$  in the case  $\mathbf{t}''(1) = 1^*$  and

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &= |I_{\mathbf{t}''(1)}^0(t)|_{L^p} = |I_1^0(t)|_{L^p} \\ &= \left| \int_{t_0}^t e^{A(t-s)} F(U_{t_0}) ds \right|_{L^p} \\ &\leq \int_{t_0}^t (|e^{A(t-s)} F(U_{t_0})|_{L^p}) ds \leq C_p \cdot \left( \int_{t_0}^t |F(U_{t_0})|_{L^p} ds \right) \\ &\leq C_p \cdot \left( \int_{t_0}^t (1 + |U_{t_0}|_{L^p}) ds \right) \leq C_p \cdot (\Delta t) \end{aligned}$$

for every  $t \in [t_0, T]$  in the case  $\mathbf{t}''(1) = 1$ . Finally, due to Lemma 3, we obtain

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &= |I_{\mathbf{t}''(1)}^0(t)|_{L^p} = |I_2^0(t)|_{L^p} = \left| \int_{t_0}^t e^{A(t-s)} B dW_s \right|_{L^p} \\ &\leq C_p \cdot \left( \int_{t_0}^t \|e^{A(t-s)} B\|_{\text{HS}}^2_{L^p} ds \right)^{1/2} \\ &\leq C_p \cdot \left( \int_0^{(\Delta t)} |e^{As} B|_{\text{HS}}^2 ds \right)^{1/2} \\ &\leq C_p \cdot (\Delta t)^\delta \end{aligned}$$

for every  $t \in [t_0, T]$  in the case  $\mathbf{t}''(1) = 2$ . This shows  $|\Phi(\mathbf{t})(t)|_{L^p} \leq C_p \cdot (\Delta t)^{\text{ord}(\mathbf{t})}$  for every  $t \in [t_0, T]$  in the base case  $l(\mathbf{t}) = 1$ .

Suppose now that  $l(\mathbf{t}) \in \{2, 3, \dots\}$ . Since  $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{ST}'$ , we must have  $\mathbf{t}''(1) \in \{1, 1^*\}$ . Let  $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbf{ST}'$  with  $n \in \mathbb{N}$  be the subtrees of  $\mathbf{t}$ . Note that  $\mathbf{t}_1, \dots, \mathbf{t}_n$  are indeed in  $\mathbf{ST}'$ . Then, by definition, we have

$$\Phi(\mathbf{t})(t) = I_{\mathbf{t}''(1)}^n[\Phi(\mathbf{t}_1), \dots, \Phi(\mathbf{t}_n)](t)$$

for every  $t \in [t_0, T]$ . Therefore, by Lemma 2, we obtain

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &= |I_{\mathbf{t}''(1)}^n[\Phi(\mathbf{t}_1), \dots, \Phi(\mathbf{t}_n)](t)| = |I_{1^*}^n[\Phi(\mathbf{t}_1), \dots, \Phi(\mathbf{t}_n)](t)|_{L^p} \\ &= \left| \int_{t_0}^t e^{A(t-s)} \left( \int_0^1 F^{(n)}(U_{t_0} + r \Delta U_s) \right. \right. \\ &\quad \left. \left. \times (\Phi(\mathbf{t}_1)(s), \dots, \Phi(\mathbf{t}_n)(s)) \frac{(1-r)^{n-1}}{(n-1)!} dr \right) ds \right|_{L^p} \\ &\leq C_p \cdot \int_{t_0}^t \left| \int_0^1 F^{(n)}(U_{t_0} + r \Delta U_s) \right. \\ &\quad \left. \times (\Phi(\mathbf{t}_1)(s), \dots, \Phi(\mathbf{t}_n)(s)) \frac{(1-r)^{n-1}}{(n-1)!} dr \right|_{L^p} ds \\ &\leq C_p \cdot \int_{t_0}^t \int_0^1 \left| F^{(n)}(U_{t_0} + r \Delta U_s) \right. \\ &\quad \left. \times (\Phi(\mathbf{t}_1)(s), \dots, \Phi(\mathbf{t}_n)(s)) \frac{(1-r)^{n-1}}{(n-1)!} \right|_{L^p} dr ds \\ &\leq C_p \cdot \int_{t_0}^t \int_0^1 |F^{(n)}(U_{t_0} + r \Delta U_s)(\Phi(\mathbf{t}_1)(s), \dots, \Phi(\mathbf{t}_n)(s))|_{L^p} dr ds \\ &\leq C_p \cdot \int_{t_0}^t \int_0^1 \|\Phi(\mathbf{t}_1)(s)\| \cdots \|\Phi(\mathbf{t}_n)(s)\|_{L^p} dr ds \end{aligned}$$

and hence

$$\begin{aligned} |\Phi(\mathbf{t})(t)|_{L^p} &\leq C_p \cdot \int_{t_0}^t \left| |\Phi(\mathbf{t}_1)(s)| \cdots |\Phi(\mathbf{t}_n)(s)| \right|_{L^p} ds \\ &\leq C_p \cdot \left( \int_{t_0}^t |\Phi(\mathbf{t}_1)(s)|_{L^{pn}} \cdots |\Phi(\mathbf{t}_n)(s)|_{L^{pn}} ds \right) \\ &\leq C_p \cdot \left( \int_{t_0}^t ((\Delta s)^{\text{ord}(\mathbf{t}_1)} \cdots (\Delta s)^{\text{ord}(\mathbf{t}_n)}) ds \right) \\ &\leq C_p \cdot (\Delta t)^{(1+\text{ord}(\mathbf{t}_1)+\cdots+\text{ord}(\mathbf{t}_n))} = C_p \cdot (\Delta t)^{\text{ord}(\mathbf{t})} \end{aligned}$$

for every  $t \in [t_0, T]$  in the case  $\mathbf{t}''(1) = 1^*$ , since  $l(\mathbf{t}_1), \dots, l(\mathbf{t}_n) \leq l(\mathbf{t}) - 1$  and we can apply the induction hypothesis to the subtrees. A similar calculation shows the result when  $\mathbf{t}''(1) = 1$ .  $\square$

### 7.3. Properties of the stochastic convolution.

LEMMA 5. *Let Assumptions 1 and 3 be fulfilled. Then, there exists an adapted stochastic process  $O : \Omega \rightarrow C([0, T], H)$  with continuous sample paths, which is a modification of the stochastic convolution  $\int_0^t e^{A(t-s)} B dW_s$ ,  $t \in [0, T]$ , that is, we have*

$$\mathbb{P} \left[ \int_0^t e^{A(t-s)} B dW_s = O_t \right] = 1$$

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

PROOF. Let  $Z : [0, T] \times \Omega \rightarrow H$  be an arbitrary adapted stochastic process with

$$Z_t = \int_0^t e^{A(t-s)} B dW_s, \quad \mathbb{P}\text{-a.s.},$$

for all  $t \in [0, T]$ . Due to Assumption 3, such an adapted stochastic process exists. Moreover,  $Z : [0, T] \times \Omega \rightarrow H$  is centered and square integrable with

$$\mathbb{E}|Z_t|^2 = \int_0^t |e^{A(t-s)} B|_{\text{HS}}^2 ds = \int_0^t |e^{As} B|_{\text{HS}}^2 ds$$

for all  $t \in [0, T]$ . Let  $\theta := \min(\delta, \gamma)$ ,  $p \geq 2$  and  $0 \leq t_1 \leq t_2 \leq T$  be given. Then, we have

$$\begin{aligned} Z_{t_2} - Z_{t_1} &= \int_{t_1}^{t_2} e^{A(t_2-s)} B dW_s \\ &\quad + \int_0^{t_1} (e^{A(t_2-s)} - e^{A(t_1-s)}) B dW_s, \quad \mathbb{P}\text{-a.s.}, \end{aligned}$$

and

$$\begin{aligned} |Z_{t_2} - Z_{t_1}|_{L^p} &\leq p \left( \int_{t_1}^{t_2} \|e^{A(t_2-s)} B\|_{\text{HS}}^2 |L^p|^2 ds \right)^{1/2} \\ &\quad + p \left( \int_0^{t_1} \|(e^{A(t_2-s)} - e^{A(t_1-s)}) B\|_{\text{HS}}^2 |L^p|^2 ds \right)^{1/2} \\ &= p \left( \int_0^{(t_2-t_1)} |e^{As} B|_{\text{HS}}^2 ds \right)^{1/2} \\ &\quad + p \left( \int_0^{t_1} |(e^{A(t_2-t_1)} - I) e^{As} B|_{\text{HS}}^2 ds \right)^{1/2} \end{aligned}$$

due to Lemma 3. Hence, due to Assumption 3, we obtain

$$\begin{aligned} |Z_{t_2} - Z_{t_1}|_{L^p} &\leq C(t_2 - t_1)^\delta \\ &\quad + C|(\kappa - A)^{-\gamma} (e^{A(t_2-t_1)} - I)| \left( \int_0^{t_1} |(\kappa - A)^\gamma e^{As} B|_{\text{HS}}^2 ds \right)^{1/2} \\ &\leq C(t_2 - t_1)^\delta + C(t_2 - t_1)^\gamma \left( \int_0^T |(\kappa - A)^\gamma e^{As} B|_{\text{HS}}^2 ds \right)^{1/2} \\ &\leq C(t_2 - t_1)^\delta + C(t_2 - t_1)^\gamma \leq C(t_2 - t_1)^\theta, \end{aligned}$$

where  $C > 0$  is a constant changing from line to line. Since  $\theta > 0$  is greater than zero and since  $p \geq 2$  was arbitrary, there exists a version of  $(Z_t)_{t \in [0, T]}$  with continuous sample paths due to Kolmogorov's theorem (see, e.g., Chapter 3 in [6]).

□

**Acknowledgments.** We strongly thank the anonymous referee for his careful reading and his very valuable advice.

## REFERENCES

- [1] BURRAGE, K. and BURRAGE, P. M. (2000). Order conditions of stochastic Runge–Kutta methods by  $B$ -series. *SIAM J. Numer. Anal.* **38** 1626–1646. [MR1813248](#)
- [2] BUTCHER, J. C. (1987). *The Numerical Analysis of Ordinary Differential Equations: Runge–Kutta and General Linear methods*. Wiley, Chichester. [MR878564](#)
- [3] BERLAND, H., OWREN, B. and SKAFLESTAD, B. (2005).  $B$ -series and order conditions for exponential integrators. *SIAM J. Numer. Anal.* **43** 1715–1727. [MR2182146](#)
- [4] BERLAND, H., SKAFLESTAD, B. and WRIGHT, W. (2007). Expint—A Matlab package for exponential integrators. *ACM Trans. Math. Software* **33** 1–17.
- [5] COX, S. M. and MATTHEWS, P. C. (2002). Exponential time differencing for stiff systems. *J. Comput. Phys.* **176** 430–455. [MR1894772](#)
- [6] DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and Its Applications* **44**. Cambridge Univ. Press, Cambridge. [MR1207136](#)

- [7] DAVIE, A. M. and GAINES, J. G. (2001). Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. *Math. Comp.* **70** 121–134. [MR1803132](#)
- [8] DEUFLHARD, P. and BORNEMANN, F. (2002). *Scientific Computing with Ordinary Differential Equations. Texts in Applied Mathematics* **42**. Springer, New York. [MR1912409](#)
- [9] ENGEL, K. and NÄGEL, R. (1999). *One-Parameter Semigroups for Linear Evolution Equations*. Springer, Berlin.
- [10] GRADINARU, M., NOURDIN, I. and TINDEL, S. (2005). Itô's- and Tanaka's-type formulae for the stochastic heat equation: The linear case. *J. Funct. Anal.* **228** 114–143. [MR2170986](#)
- [11] GYÖNGY, I. (1998). A note on Euler's approximations. *Potential Anal.* **8** 205–216. [MR1625576](#)
- [12] GYÖNGY, I. (1998). Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I. *Potential Anal.* **9** 1–25. [MR1644183](#)
- [13] GYÖNGY, I. (1999). Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II. *Potential Anal.* **11** 1–37. [MR1699161](#)
- [14] GYÖNGY, I. and KRYLOV, N. (2003). On the splitting-up method and stochastic partial differential equations. *Ann. Probab.* **31** 564–591. [MR1964941](#)
- [15] HAUSENBLAS, E. (2003). Approximation for semilinear stochastic evolution equations. *Potential Anal.* **18** 141–186. [MR1953619](#)
- [16] HOCHBRUCK, M., LUBICH, C. and SELHOFER, H. (1998). Exponential integrators for large systems of differential equations. *SIAM J. Sci. Comput.* **19** 1552–1574. [MR1618808](#)
- [17] HOCHBRUCK, M. and OSTERMANN, A. (2005). Explicit exponential Runge–Kutta methods for semilinear parabolic problems. *SIAM J. Numer. Anal.* **43** 1069–1090. [MR2177796](#)
- [18] HOCHBRUCK, M. and OSTERMANN, A. (2005). Exponential Runge–Kutta methods for parabolic problems. *Appl. Numer. Math.* **53** 323–339. [MR2128529](#)
- [19] HOCHBRUCK, M., OSTERMANN, A. and SCHWEITZER, J. (2008/09). Exponential Rosenbrock-type methods. *SIAM J. Numer. Anal.* **47** 786–803. [MR2475962](#)
- [20] JENTZEN, A. (2007). *Numerische Verfahren hoher Ordnung für zufällige Differentialgleichungen, Diplomarbeit*. Johann Wolfgang Goethe Univ. Frankfurt am Main.
- [21] JENTZEN, A. (2009). Higher order pathwise numerical approximations of SPDEs with additive noise. Submitted.
- [22] JENTZEN, A. (2009). Pathwise numerical approximations of SPDEs with additive noise under non-global Lipschitz coefficients. *Potential Anal.* **4** 375–404.
- [23] JENTZEN, A. and KLOEDEN, P. E. (2009). Pathwise Taylor schemes for random ordinary differential equations. *BIT* **49** 113–140. [MR2486131](#)
- [24] JENTZEN, A. and KLOEDEN, P. E. (2009). Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **465** 649–667. [MR2471778](#)
- [25] JENTZEN, A. and KLOEDEN, P. E. (2009). The numerical approximation of stochastic partial differential equations. *Milan J. Math.* **77** 205–244.
- [26] JENTZEN, A., KLOEDEN, P. E. and NEUENKIRCH, A. (2009). Pathwise approximation of stochastic differential equations on domains: Higher order convergence rates without global Lipschitz coefficients. *Numer. Math.* **112** 41–64. [MR2481529](#)
- [27] KASSAM, A.-K. and TREFETHEN, L. N. (2005). Fourth-order time-stepping for stiff PDEs. *SIAM J. Sci. Comput.* **26** 1214–1233. [MR2143482](#)
- [28] KLOEDEN, P. E. and JENTZEN, A. (2007). Pathwise convergent higher order numerical schemes for random ordinary differential equations. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **463** 2929–2944. [MR2360184](#)
- [29] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations. Applications of Mathematics (New York)* **23**. Springer, Berlin. [MR1214374](#)

- [30] KROGSTAD, S. (2005). Generalized integrating factor methods for stiff PDEs. *J. Comput. Phys.* **203** 72–88. [MR2104391](#)
- [31] LAWSON, J. D. (1967). Generalized Runge–Kutta processes for stable systems with large Lipschitz constants. *SIAM J. Numer. Anal.* **4** 372–380. [MR0221759](#)
- [32] MÜLLER-GRONBACH, T. and RITTER, K. (2007). Lower bounds and nonuniform time discretization for approximation of stochastic heat equations. *Found. Comput. Math.* **7** 135–181. [MR2324415](#)
- [33] MÜLLER-GRONBACH, T. and RITTER, K. (2007). An implicit Euler scheme with nonuniform time discretization for heat equations with multiplicative noise. *BIT* **47** 393–418. [MR2334046](#)
- [34] MÜLLER-GRONBACH, T., RITTER, K. and WAGNER, T. (2008). Optimal pointwise approximation of a linear stochastic heat equation with additive space-time white noise. In *Monte Carlo and Quasi-Monte Carlo Methods 2006* 577–589. Springer, Berlin. [MR2479247](#)
- [35] MILSTEIN, G. N. (1995). *Numerical Integration of Stochastic Differential Equations. Mathematics and Its Applications* **313**. Kluwer, Dordrecht. [MR1335454](#)
- [36] OSTERMANN, A., THALHAMMER, M. and WRIGHT, W. M. (2006). A class of explicit exponential general linear methods. *BIT* **46** 409–431. [MR2238680](#)
- [37] PRÉVÔT, C. and RÖCKNER, M. (2007). *A Concise Course on Stochastic Partial Differential Equations. Lecture Notes in Math.* **1905**. Springer, Berlin. [MR2329435](#)
- [38] RÖSSLER, A. (2009). *Runge–Kutta Methods for the Numerical Solution of Stochastic Differential Equations*. Shaker, Aachen.
- [39] RÖSSLER, A. (2004). Stochastic Taylor expansions for the expectation of functionals of diffusion processes. *Stoch. Anal. Appl.* **22** 1553–1576. [MR2095070](#)
- [40] RÖSSLER, A. (2006). Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations. *Stoch. Anal. Appl.* **24** 97–134. [MR2198539](#)
- [41] SELL, G. R. and YOU, Y. (2002). *Dynamics of Evolutionary Equations. Applied Mathematical Sciences* **143**. Springer, New York. [MR1873467](#)
- [42] SHARDLOW, T. (1999). Numerical methods for stochastic parabolic PDEs. *Numer. Funct. Anal. Optim.* **20** 121–145. [MR1683281](#)

INSTITUTE OF MATHEMATICS  
JOHANN WOLFGANG GOETHE UNIVERSITY  
ROBERT-MAYER-STR. 6-10  
D-60325 FRANKFURT AM MAIN  
GERMANY  
E-MAIL: [jentzen@math.uni-frankfurt.de](mailto:jentzen@math.uni-frankfurt.de)  
[kloeden@math.uni-frankfurt.de](mailto:kloeden@math.uni-frankfurt.de)