

The asymptotic error of chaos expansion approximations for stochastic differential equations

Tony Huschto^a, Mark Podolskij^{b,*}, Sebastian Sager^c

^a*Department of Mathematics, Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany*

^b*Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus, Denmark*

^c*Faculty of Mathematics, Otto-von-Guericke Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*

tony.huschto@roche.com (T. Huschto), mpodolskij@math.au.dk (M. Podolskij), sager@ovgu.de (S. Sager)

Received: 22 November 2018, Revised: 9 March 2019, Accepted: 9 March 2019,
Published online: 23 April 2019

Abstract In this paper we present a numerical scheme for stochastic differential equations based upon the Wiener chaos expansion. The approximation of a square integrable stochastic differential equation is obtained by cutting off the infinite chaos expansion in chaos order and in number of basis elements. We derive an explicit upper bound for the L^2 approximation error associated with our method. The proofs are based upon an application of Malliavin calculus.

Keywords Chaos expansion, Malliavin calculus, numerical approximation, stochastic differential equations

2010 MSC 65C30, 60H10, 60H07

1 Introduction

We consider a one-dimensional continuous stochastic process $(X_t)_{t \in [0, T]}$ that satisfies the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \quad \text{with } X_0 = x_0, \quad (1.1)$$

*Corresponding author.

where $(W_t)_{t \in [0, T]}$ is a Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$. Various numerical approximation schemes for the SDE (1.1) have been proposed and studied in the literature in the past decades. The probably most prominent numerical approximation for the solution of (1.1) is the Euler scheme, which can be described as follows. Let $\varphi_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be the function defined by $\varphi_n(t) = i/n$ when $t \in [i/n, (i+1)/n)$. The continuous Euler approximation scheme is described by

$$dX_t^n = b(\varphi_n(t), X_{\varphi_n(t)}^n)dt + \sigma(\varphi_n(t), X_{\varphi_n(t)}^n)dW_t \quad \text{with } X_0^n = x_0. \quad (1.2)$$

The probabilistic properties of the Euler approximation scheme have been investigated in numerous papers. We refer to the classical works [2, 3, 14, 16, 17, 23] for the studies on weak and strong approximation errors among many others. Asymptotic results in the framework of non-regular coefficients can be found in e.g. [1, 8, 9, 27].

In this paper we take a different route and propose to use the Wiener chaos expansion (also called polynomial chaos in the literature) to approximate the solution of the SDE (1.1). To explain ideas let us fix an orthonormal basis $(e_i)_{i \geq 1}$ of the separable Hilbert space $L^2([0, T])$. It is a well-known statement (cf. [6]) that if $X_t \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \in [0, T]$, where \mathcal{F} is generated by the Brownian motion $(W_t)_{t \in [0, T]}$, it admits the chaotic expansion

$$X_t = \sum_{\alpha \in \mathcal{I}} x_\alpha(t) \Psi^\alpha, \quad (1.3)$$

where x_α are deterministic functions, Ψ^α are mutually orthogonal random projections that are associated to the basis $(e_i)_{i \geq 1}$, and the index set \mathcal{I} is defined via

$$\mathcal{I} := \{\alpha = (\alpha_i)_{i \geq 1} : \alpha_i \in \mathbb{N}_0 \text{ and almost all } \alpha_i \text{'s are } 0\}. \quad (1.4)$$

Such orthogonal expansions have been successfully applied in numerous fields of stochastic and numerical analysis. We refer e.g. to [11–13, 25] for applications of the Wiener chaos expansion in the context of SDEs and to [20–22, 26] for applications of polynomial expansion in modelling, simulation and filtering of stochastic partial differential equations. The aim of this work is to use the chaos expansion (1.3) to numerically approximate the solution of the SDE (1.1). For this purpose we need to truncate the infinite sum in (1.3). An obvious approach is to consider the approximation

$$X_t^{p,k} = \sum_{\alpha \in \mathcal{I}_{p,k}} x_\alpha(t) \Psi^\alpha, \quad (1.5)$$

where the subset $\mathcal{I}_{p,k} \subset \mathcal{I}$ refers to using orthogonal projections Ψ^α only with respect to the first k basis elements $(e_i)_{1 \leq i \leq k}$ and only up to the p th order Wiener chaos. This method is mostly related to the articles [11, 12, 22, 21]. More specifically, the works [21, 22] study the L^2 -error associated with the approximation (1.5), but only for a particular choice of the basis $(e_i)_{i \geq 1}$. In this paper we will study the decay rate of $\mathbb{E}[(X_t - X_t^{p,k})^2]$ when $k, p \rightarrow \infty$ for a general basis $(e_i)_{i \geq 1}$ of $L^2([0, T])$ applying methods from Malliavin calculus.

The paper is structured as follows. In Section 2 we present the elements of Malliavin calculus. The main results of the paper are demonstrated in Section 3. Section 4 is devoted to proofs. In Section 5 we illustrate our approach with exemplary numerical results for the Haar and a trigonometric basis, and propose a heuristic based on sparse indices for computational speedup.

2 Background on Malliavin calculus

In this section we introduce some basic concepts of Malliavin calculus. The interested readers are referred to [24] for more thorough exposition on this subject. Set $\mathbb{H} = L^2([0, T])$ and let $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ denote the scalar product on \mathbb{H} . We note that \mathbb{H} is a separable Hilbert space and denote by $(e_i)_{i \geq 1}$ an orthonormal basis of \mathbb{H} . We consider the *isonormal Gaussian family* $W = \{W(h) : h \in \mathbb{H}\}$ indexed by \mathbb{H} defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. the random variables $W(h)$ are centered Gaussian with a covariance structure determined via

$$\mathbb{E}[W(g)W(h)] = \langle g, h \rangle_{\mathbb{H}}.$$

Here and throughout the paper we assume that $\mathcal{F} = \sigma(W)$. In our setting we consider $W(h) = \int_0^T h_s dW_s$, where W is a standard Brownian motion. We define the normalised Hermite polynomials through the identities

$$H_0(x) := 1, \quad H_n(x) := \frac{(-1)^n}{\sqrt{n!}} \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \left(\exp\left(-\frac{x^2}{2}\right) \right), \quad n \geq 1. \quad (2.1)$$

The n th Wiener chaos \mathcal{H}_n is the closed linear subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ generated by the family of random variables $\{H_n(W(h)) : \|h\|_{\mathbb{H}} = 1\}$. The vector spaces \mathcal{H}_n , $n \geq 0$, are orthogonal and we have the Wiener chaos expansion

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n. \quad (2.2)$$

(See [24, Theorem 1.1.1] for more details.) Next, for $\alpha \in \mathcal{I}$, where \mathcal{I} has been introduced in (1.4), we define the random variable

$$\Psi^\alpha := \prod_{i=1}^{\infty} H_{\alpha_i}(W(e_i)), \quad (2.3)$$

where $(e_i)_{i \geq 1}$ is a fixed orthonormal basis of \mathbb{H} . We define $|\alpha| = \sum_{i=1}^{\infty} \alpha_i$ and $\alpha! = \prod_{i=1}^{\infty} \alpha_i!$ for $\alpha \in \mathcal{I}$. We deduce that the set $\{\Psi^\alpha : \alpha \in \mathcal{I} \text{ with } |\alpha| = n\}$ forms a complete orthonormal basis of the n th Wiener chaos \mathcal{H}_n and consequently $\{\Psi^\alpha : \alpha \in \mathcal{I}\}$ is a complete orthonormal basis of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ (cf. [24, Proposition 1.1.1]).

Now, we introduce multiple stochastic integrals of order $n \in \mathbb{N}$, which are denoted by I_n . For an element $h^{\otimes n} := h \otimes \dots \otimes h$ in $\mathbb{H}^{\otimes n}$ with $\|h\|_{\mathbb{H}} = 1$, we define

$$I_n(h^{\otimes n}) := \sqrt{n!} H_n(W(h)), \quad n \geq 1. \quad (2.4)$$

Assuming that the mapping I_n is linear, the definition (2.4) can be extended to all symmetric elements $h \in \mathbb{H}^{\otimes n}$ by polarisation identity. Finally, for an arbitrary func-

tion $h \in \mathbb{H}^{\otimes n}$ we set

$$I_n(h) := I_n(\tilde{h}),$$

where \tilde{h} denotes the symmetrised version of $h \in \mathbb{H}^{\otimes n}$. By definition I_n maps \mathbb{H} into \mathcal{H}_n , so the multiple integrals of different orders are orthogonal. In particular, they satisfy the isometry property

$$\mathbb{E}[I_m(g)I_n(h)] = n! \langle \tilde{g}, \tilde{h} \rangle_{\mathbb{H}^{\otimes n}} 1_{\{n=m\}}, \quad h \in \mathbb{H}^{\otimes n}, g \in \mathbb{H}^{\otimes m}. \quad (2.5)$$

Furthermore, for any symmetric $h \in \mathbb{H}^{\otimes n}$, $g \in \mathbb{H}^{\otimes m}$, the following multiplication formula holds:

$$I_m(g)I_n(h) = \sum_{r=0}^{\min(m,n)} r! \binom{m}{r} \binom{n}{r} I_{m+n-2r}(g \otimes_r h), \quad (2.6)$$

where the r th contraction $g \otimes_r h$ is defined by

$$\begin{aligned} g \otimes_r h(t_1, \dots, t_{m+n-2r}) &:= \int_{[0,T]^r} g(u_1, \dots, u_r, t_1, \dots, t_{m-r}) \\ &\quad \times h(u_1, \dots, u_r, t_{m-r+1}, \dots, t_{m+n-2r}) du_1 \dots du_r. \end{aligned}$$

(See [24, Proposition 1.1.3].) The Wiener chaos expansion (2.2) transfers to the context of multiple integrals as follows. For each random variable $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ we obtain the orthogonal decomposition

$$F = \sum_{n=0}^{\infty} I_n(g_n), \quad g_n \in \mathbb{H}^{\otimes n}, \quad (2.7)$$

where $I_0 = \mathbb{E}[F]$ and the decomposition is unique when g_n , $n \geq 1$, are assumed to be symmetric (cf. [24, Theorem 1.1.2]).

Next, we introduce the notion of Malliavin derivative and its adjoint operator. We define the set of smooth random variables via

$$\mathcal{S} = \{F = f(W(h_1), \dots, W(h_n)) : n \geq 1, h_i \in \mathbb{H}\},$$

where $f \in C_p^\infty(\mathbb{R}^n)$ (i.e. the space of infinitely differentiable functions such that all derivatives exhibit polynomial growth). The k th order Malliavin derivative of $F \in \mathcal{S}$, denoted by $D^k F$, is defined by

$$D^k F = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}} f(W(h_1), \dots, W(h_n)) h_{i_1} \otimes \dots \otimes h_{i_k}. \quad (2.8)$$

Notice that $D^k F$ is a $\mathbb{H}^{\otimes k}$ -valued random variable, and we write $D_x^k F$ for the realisation of the function $D^k F$ at the point $x \in [0, T]^k$. The space $\mathbb{D}^{k,q}$ denotes the completion of the set \mathcal{S} with respect to the norm

$$\|F\|_{k,q} := \left(\mathbb{E}[|F|^q] + \sum_{m=1}^k \mathbb{E}[\|D^m F\|_{\mathbb{H}^{\otimes m}}^q] \right)^{1/q}.$$

We define $\mathbb{D}^{k,\infty} = \bigcap_{q>1} \mathbb{D}^{k,q}$. The Malliavin derivative of the random variable $\Psi^\alpha \in \mathcal{S}$, $\alpha \in \mathcal{I}$, can be easily computed using the definition (2.8) and the formula $H'_n(x) =$

$\sqrt{n}H_{n-1}(x)$:

$$D_s \Psi^\alpha = \sum_{i=1}^{\infty} \sqrt{\alpha_i} e_i(s) \Psi^{\alpha^{-(i)}}, \quad \text{where} \quad (2.9)$$

$$\alpha^{-(i)} := (\alpha_1, \dots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \dots) \quad \text{if } \alpha_i \geq 1. \quad (2.10)$$

Higher order Malliavin derivatives of Ψ^α are computed similarly.

The operator D^k possesses an unbounded adjoint denoted by δ^k , which is often referred to as multiple Skorokhod integral. The following integration by parts formula holds (see [24, Exercise 1.3.7]): if $u \in \text{Dom}(\delta^k)$ and $F \in \mathbb{D}^{k,2}$, where $\text{Dom}(\delta^k)$ consists of all elements $u \in L^2(\Omega; \mathbb{H}^{\otimes k})$ such that the inequality $|\mathbb{E}[\langle D^k F, u \rangle_{\mathbb{H}^{\otimes k}}]| \leq c(E[F^2])^{1/2}$ holds for some $c > 0$ and all $F \in \mathbb{D}^{k,2}$, then we have the identity

$$\mathbb{E}[F \delta^k(u)] = \mathbb{E}[\langle D^k F, u \rangle_{\mathbb{H}^{\otimes k}}]. \quad (2.11)$$

In case the random variable $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ has a chaos decomposition as displayed in (2.7), the statement $F \in \mathbb{D}^{k,2}$ is equivalent to the condition $\sum_{n=1}^{\infty} n^k n! \|g_n\|_{\mathbb{H}^{\otimes n}}^2 < \infty$. In particular, when $F \in \mathbb{D}^{1,2}$ we deduce an explicit chaos representation of the derivative $DF = (D_t F)_{t \in [0, T]}$

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(g_n(\cdot, t)), \quad t \in [0, T], \quad (2.12)$$

where $g_n(\cdot, t) : [0, T]^{n-1} \rightarrow \mathbb{R}$ is obtained from the function g_n by setting the last argument equal to t (see [24, Proposition 1.2.7]). Finally, we present an explicit formula for the Malliavin derivative of a solution of a stochastic differential equation. Assume that $(X_t)_{t \in [0, T]}$ is a solution of a stochastic differential equation (1.1) and $b, \sigma \in C^1(\mathbb{R})$. In this setting $DX_t = (D_s X_t)_{s \in [0, T]}$ is given as the solution of the SDE

$$D_s X_t = \sigma(s, X_s) + \int_s^t \frac{\partial}{\partial x} b(u, X_u) D_s(X_u) du + \int_s^t \frac{\partial}{\partial x} \sigma(u, X_u) D_s(X_u) dW_u, \quad (2.13)$$

for $s \leq t$, and $D_s X_t = 0$ if $s > t$ (see [24, Theorem 2.2.1]). Throughout the paper $\partial/\partial x$ denotes the derivative with respect to the space variable and $\partial/\partial t$ denotes the derivative with respect to the time variable.

3 Main results

We start with the analysis of the Wiener chaos expansion introduced in (1.3). Under square integrability assumption on the solution X of the SDE (1.1) we obtain the Wiener chaos expansion

$$X_t = \sum_{\alpha \in \mathcal{I}} x_\alpha(t) \Psi^\alpha \quad \text{with } \Psi^\alpha = \prod_{i=1}^{\infty} H_{\alpha_i}(W(e_i)).$$

In order to study the strong approximation error we require a good control of the coefficients $x_\alpha(t)$. When X_t is sufficiently smooth in the Malliavin sense, we deduce

the identity

$$x_\alpha(t) = \mathbb{E}[X_t \Psi^\alpha] = \frac{1}{\sqrt{\alpha!}} \mathbb{E} \left[\left\langle D^{|\alpha|} X_t, \bigotimes_{i=1}^{\infty} e_i^{\otimes \alpha_i} \right\rangle_{\mathbb{H}^{\otimes |\alpha|}} \right] \quad (3.1)$$

applying the duality formula (2.11). In fact, the coefficients $x_\alpha(t)$ satisfy a system of ordinary differential equations. The following propagator system has been derived in [12, 21]. We state the proof for completeness.

Theorem 3.1. *Let $(X_t)_{t \in [0, T]}$ be the unique solution of the SDE (1.1) and assume that $X \in L^2(\mathcal{S} \times [0, T])$. Then X_t possesses the chaos expansion (1.3) and the coefficients $x_\alpha(t)$ satisfy the system of ordinary differential equations*

$$x'_\alpha(t) = b_\alpha(t, X_t) + \sum_{j=1}^{\infty} \sqrt{\alpha_j} e_j(t) \sigma_{\alpha^-(j)}(t, X_t), \quad (3.2)$$

$$x_\alpha(0) = 1_{\{\alpha=0\}} x_0.$$

Here $b_\alpha(t, X_t)$ (resp. $\sigma_\alpha(t, X_t)$) denotes the α -coefficient of the Wiener chaos expansion (1.3) associated with the random variable $b(t, X_t)$ (resp. $\sigma(t, X_t)$), and the multi-index $\alpha^-(j)$ is defined by (2.10).

Proof. Using the SDE (1.1) and applying the formula (3.1) we obtain the identity

$$x_\alpha(t) = x_0 1_{\{\alpha=0\}} + \int_0^t \mathbb{E}[\Psi^\alpha b(s, X_s)] ds + \mathbb{E} \left[\Psi^\alpha \int_0^t \sigma(s, X_s) dW_s \right].$$

Applying the formula (3.1) once again for the random variable $b(s, X_s)$ we immediately deduce that

$$b_\alpha(s, X_s) = \mathbb{E}[\Psi^\alpha b(s, X_s)].$$

On the other hand, observing the identity $\delta(1_{[0, t]} \sigma(\cdot, X)) = \int_0^t \sigma(s, X_s) dW_s$, we get by the duality formula (2.11) and (2.9) that

$$\begin{aligned} \mathbb{E} \left[\Psi^\alpha \int_0^t \sigma(s, X_s) dW_s \right] &= \sum_{j=1}^{\infty} \int_0^t \sqrt{\alpha_j} e_j(s) \mathbb{E}[\Psi^{\alpha^-(j)} \sigma(s, X_s)] ds \\ &= \sum_{j=1}^{\infty} \int_0^t \sqrt{\alpha_j} e_j(s) \sigma_{\alpha^-(j)}(s, X_s) ds. \end{aligned}$$

Putting things together we obtain the identity

$$x_\alpha(t) = x_0 1_{\{\alpha=0\}} + \int_0^t b_\alpha(s, X_s) ds + \sum_{j=1}^{\infty} \int_0^t \sqrt{\alpha_j} e_j(s) \sigma_{\alpha^-(j)}(s, X_s) ds.$$

Consequently, the assertion follows after taking the derivative with respect to t . \square

We remark that the propagator system (3.2) is recursive. Let us give some simple examples to illustrate how (3.2) can be solved explicitly.

Example 3.2. (i) (*Scaled Brownian motion with drift*) We start with the toy example of a scaled Brownian motion with drift:

$$X_t = bt + \sigma W_t.$$

In this case we obviously have that $x_\alpha(t) = 0$ for any $\alpha \in \mathcal{I}$ with $|\alpha| \geq 2$. Applying formula (3.2) we obtain the representation

$$X_t = bt + \sigma \sum_{j=1}^{\infty} \left(\int_0^t e_j(s) ds \right) \int_0^T e_j(s) dW_s,$$

which is a well known Karhunen–Loève expansion of the scaled Brownian motion.

(ii) (*Geometric Brownian motion*) Let us consider the geometric Brownian motion defined via the SDE

$$dX_t = bX_t dt + \sigma X_t dW_t, \quad X_0 = x_0 > 0.$$

In this setting the propagator system (3.2) translates to

$$x'_\alpha(t) = bx_\alpha(t) + \sigma \sum_{j=1}^{\infty} \sqrt{\alpha_j} e_j(t) x_{\alpha - (j)}(t),$$

$$x_\alpha(0) = 1_{\{\alpha=0\}} x_0.$$

This system of ordinary differential equations can be solved recursively. For $\alpha = 0$ we have $x'_0(t) = bx_0(t)$ and hence $x_0(t) = x_0 \exp(bt)$. If α is the j th canonical unit vector in \mathcal{I} (and hence $|\alpha| = 1$) we obtain the differential equation

$$x'_\alpha(t) = bx_\alpha(t) + \sigma e_j(t) x_0(t),$$

$$x_\alpha(0) = 1_{\{\alpha=0\}} x_0.$$

Hence, $x_\alpha(t) = x_0 \sigma \exp(bt) \int_0^t e_j(s) ds$. Following this recursion we obtain the general formula

$$x_\alpha(t) = \frac{1}{\sqrt{\alpha!}} x_0 \sigma^p \exp(bt) \prod_{j=1}^{\infty} \left(\int_0^t e_j(s) ds \right)^{\alpha_j}$$

for any $\alpha \in \mathcal{I}$ with $|\alpha| = p$.

For a general specification of the drift coefficient b and diffusion coefficient σ in model (1.1) the propagator system (3.2) cannot be solved explicitly. Thus, precise infinite dimensional Wiener chaos expansion (1.3) is out of reach. For simulation purposes it is an obvious idea to consider a finite subset of \mathcal{I} in the expansion (1.3). We introduce the index set

$$\mathcal{I}_{p,k} := \{ \alpha \in \mathcal{I} : |\alpha| \leq p \text{ and } \alpha_i = 0 \text{ for all } i > k \}. \quad (3.3)$$

The approximation of X_t is now defined via (1.5):

$$X_t^{p,k} = \sum_{\alpha \in \mathcal{I}_{p,k}} x_\alpha(t) \Psi^\alpha. \quad (3.4)$$

We remark that the quantity $X_t^{p,k}$ is more useful than the Euler approximation X_t^n introduced in (1.2) if we are interested in the approximation of the first two moments of X_t . Indeed, the first two moments of $X_t^{p,k}$ are given explicitly by

$$\mathbb{E}[X_t^{p,k}] = x_0(t) \quad \text{and} \quad \mathbb{E}[(X_t^{p,k})^2] = \sum_{\alpha \in \mathcal{I}_{p,k}} x_\alpha^2(t),$$

while higher order moments can be computed via an application of the multiplication formula (2.6).

The strong approximation error associated with the truncation (3.4) has been previously studied in [21, 22] in a slightly different context. In particular, in both papers the authors consider one specific basis $(e_i)_{i \geq 1}$ of $L^2([0, T])$ whereas we are interested in the asymptotic analysis for a general basis $(e_i)_{i \geq 1}$. While [22] mostly uses methods from analysis, our approach is based upon Malliavin calculus and is close in spirit to [21]. The main result of our paper gives an upper bound on the L^2 -error $\mathbb{E}[(X_t^{p,k} - X_t)^2]$.

Theorem 3.3. *Let $(X_t)_{t \in [0, T]}$ be the solution of the SDE (1.1). Suppose that the coefficient functions b and σ satisfy the Lipschitz and linear growth conditions*

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|, \quad t \in [0, T], \quad (3.5)$$

$$|b(t, x)|^2 + |\sigma(t, x)|^2 \leq K^2(1 + |x|^2), \quad t \in [0, T].$$

Moreover, assume that $b, \sigma \in C^{1, \infty}([0, T] \times \mathbb{R})$, where $C^{1, \infty}([0, T] \times \mathbb{R})$ denotes the space of functions $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ that are once continuously differentiable in the first component and infinitely differentiable in the second component, such that

$$\left| \frac{\partial^{l+m}}{\partial t^l \partial x^m} b(t, x) - \frac{\partial^{l+m}}{\partial t^l \partial x^m} b(t, y) \right| + \left| \frac{\partial^{l+m}}{\partial t^l \partial x^m} \sigma(t, x) - \frac{\partial^{l+m}}{\partial t^l \partial x^m} \sigma(t, y) \right| \leq K|x - y|,$$

for $t \in [0, T]$, any $l = 0, 1$, and $m \geq 0$. Then it holds that

$$\mathbb{E}[(X_t^{p,k} - X_t)^2] \leq C(1 + x_0^2) \left(\frac{1}{(p+1)!} + \sum_{l=k+1}^{\infty} \left(E_l^2(t) + \int_0^t E_l^2(\tau) d\tau \right) \right), \quad (3.6)$$

where $C = C(t, K)$ is a positive constant and the function $E_l(t)$ is defined by

$$E_l(t) := \int_0^t e_l(s) ds. \quad (3.7)$$

Let us give some remarks about the statement (3.6). First of all, recalling the Karhunen–Loève expansion $W_t = \sum_{l=1}^{\infty} E_l(t) \int_0^T e_l(s) dW_s$, we readily deduce that

$$\mathbb{E}[W_t^2] = \sum_{l=1}^{\infty} E_l^2(t) < \infty \quad \text{and}$$

$$\int_0^t \mathbb{E}[W_s^2] ds = \sum_{l=1}^{\infty} \int_0^t E_l^2(\tau) d\tau < \infty.$$

Hence, the upper bound on the right-hand side of (3.6) indeed converges to 0 when $p, k \rightarrow \infty$. We also note that the error associated with the truncation of the chaos order does not depend on the particular choice of the basis $(e_i)_{i \geq 1}$, while the error associated with truncation of basis strongly depends on $(e_i)_{i \geq 1}$ (which is not really surprising). Furthermore, we remark that it is extremely computationally costly to

compute all coefficients $x_\alpha(t)$, $\alpha \in \mathcal{I}_{p,k}$, for a large chaos order p . Thanks to the first bound $((p+1)!)^{-1}$ in (3.6) it is sufficient to use small values of p in practical situations (usually $p \leq 4$).

Example 3.4. Let us explicitly compute the last terms of the upper bound (3.6) for two prominent bases of $L^2([0, 1])$.

(i) (*Trigonometric basis*) Consider the orthonormal basis $(e_i)_{i \geq 1}$ given by

$$e_1(t) = 1, \quad e_{2j}(t) = \sqrt{2} \sin(2\pi jt), \quad e_{2j+1}(t) = \sqrt{2} \cos(2\pi jt), \quad j \geq 1.$$

In this setting we obtain that

$$E_{2j}^2(t) = \frac{1}{2\pi^2 j^2} (1 - \cos(2\pi jt))^2, \quad E_{2j+1}^2(t) = \frac{1}{2\pi^2 j^2} (1 - \sin(2\pi jt))^2,$$

for any $j \geq 1$. Consequently, we deduce that

$$\sum_{l=k+1}^{\infty} \left(E_l^2(t) + \int_0^t E_l^2(\tau) d\tau \right) \leq \frac{C}{k} \quad (3.8)$$

for some $C > 0$.

(ii) (*Haar basis*) The Haar basis is a collection of functions

$$\{e_0, e_{j,n} : j = 1, \dots, 2^{n-1}, n \geq 1\}$$

defined as follows:

$$e_0(t) = 1, \quad e_{j,n}(t) = \begin{cases} 2^{(n-1)/2} : & t \in [2^{-n+1}(j-1), 2^{-n}(2j-1)], \\ -2^{(n-1)/2} : & t \in [2^{-n}(2j-1), 2^{-n+1}j], \\ 0 : & \text{else.} \end{cases}$$

In this case we deduce the following representation for $E_{j,n}(t)$:

$$E_{j,n}(t) = \begin{cases} 2^{(n-1)/2}(t - 2^{-n+1}(j-1)) : & t \in [2^{-n+1}(j-1), 2^{-n}(2j-1)], \\ 2^{(n-1)/2}(2^{-n+1}j - t) : & t \in [2^{-n}(2j-1), 2^{-n+1}j], \\ 0 : & \text{else.} \end{cases}$$

Since $\max_{t \in [0,1]} E_{j,n}^2(t) = 2^{-(n+1)}$ and $E_{j,n}(t) \neq 0$ only for $t \in [2^{-n+1}(j-1), 2^{-n+1}j]$, we finally obtain that

$$\sum_{l=n+1}^{\infty} \sum_{j=1}^{2^n} \left(E_{j,l}^2(t) + \int_0^t E_{j,l}^2(\tau) d\tau \right) \leq C 2^{-n}. \quad (3.9)$$

Note that the exponential asymptotic rate for a fixed p th order Wiener chaos is equivalent to $\mathcal{O}(k^{-1})$, if we choose $k = 2^n$ basis elements (e_i) . We will come back to these two bases in Section 5 where we provide exemplary numerical results that highlight this identical asymptotic rate, but also the different error distributions over time.

4 Proofs

Throughout the proofs $C = C(t, K)$ denotes a generic positive constant, which might change from line to line. We start with a proposition that gives an upper bound for the L^2 -norm of the Malliavin derivatives of X_t .

Proposition 4.1. *Under conditions of Theorem 3.3 we obtain the estimate*

$$\mathbb{E}[(D_{s_1, \dots, s_n}^n X_t)^2] \leq C^n (1 + x_0^2). \quad (4.1)$$

Proof. We show the assertion of Proposition 4.1 by induction over n . For $n = 0$ the result is a well-known consequence of the Lipschitz and linear growth conditions (3.5) on the functions b and σ (see e.g. [15, Theorem 2.9, page 289] which relates the second moment of X_t to the second moment of the initial condition $X_0 = x_0$). For $n = 1$ we know from the formula (2.13) that the process $(D_s X_t)_{s \in [0, T]}$ satisfies the SDE

$$D_s X_t = \sigma(s, X_s) + \int_s^t \frac{\partial}{\partial x} b(u, X_u) D_s(X_u) du + \int_s^t \frac{\partial}{\partial x} \sigma(u, X_u) D_s(X_u) dW_u,$$

for $s \leq t$, and $D_s X_t = 0$ for $s > t$. Next, we introduce the two-dimensional stochastic process $(Y_{s;t}^{(1)})_{t \geq s}$ via

$$Y_{s;t}^{(1)} := (X_t, D_s X_t).$$

Note that the process $(Y_{s;t}^{(1)})_{t \geq s}$ satisfies a two-dimensional SDE with initial condition given by $Y_{s;s}^{(1)}$. In particular, setting $Y_{s;t}^{(1)} = (Y_{s;t}^{(1,1)}, Y_{s;t}^{(1,2)})$, we have the representation

$$D_s X_t = \sigma(s, Y_{s;s}^{(1,1)}) + \int_s^t \bar{b}(u, Y_{s;u}^{(1)}) du + \int_s^t \bar{\sigma}(u, Y_{s;u}^{(1)}) dW_u,$$

where $\bar{b}(u, y) = \frac{\partial}{\partial x} b(u, y_1) y_2$ and $\bar{\sigma}(u, y) = \frac{\partial}{\partial x} \sigma(u, y_1) y_2$ for $y = (y_1, y_2)$. Applying again the estimates of [15, Theorem 2.9, page 289] to the diffusion process $(Y_{s;t}^{(1)})_{t \geq s}$, we conclude the estimate

$$\begin{aligned} \mathbb{E}[(D_s X_t)^2] &\leq \mathbb{E}[\|Y_{s,t}^{(1)}\|^2] \leq C(1 + \mathbb{E}[\|Y_{s,s}^{(1)}\|^2]) \\ &\leq C(1 + \mathbb{E}[X_s^2] + \mathbb{E}[\sigma(s, X_s)^2]) \\ &\leq C(1 + x_0^2). \end{aligned} \quad (4.2)$$

Now, we will perform the induction step. Notice that $D_s X_t \in \mathbb{D}^{1,\infty}$, so its Malliavin derivative satisfies again an SDE according to [24, Theorem 2.2.2]. We define the stochastic process

$$Y_{s_1, \dots, s_n; t}^{(n)} := (X_t, D_{s_1} X_t, D_{s_1, s_2}^2 X_t, \dots, D_{s_1, \dots, s_n}^n X_t).$$

Assume that the assertion of Theorem 3.3 holds for all $k = 1, \dots, n-1$. In order to compute the estimate, we have to consider the initial values of the SDE system satisfied by $Y_{s_1, \dots, s_n; t}^{(n)}$. The formula below can be found in [24, Theorem 2.2.2]. Before we state it, we need some additional notation. The stochastic process $D^m X_t = \{D_{s_1, \dots, s_m}^m X_t \mid (s_1, \dots, s_m) \in [0, T]^m\}$ depends on the m time points s_1, \dots, s_m . For any subset $J = \{j_1 < \dots < j_\eta\}$ of $\{1, \dots, m\}$ with $|J| = \eta \leq m$ elements, denote $s(J) = (s_{j_1}, \dots, s_{j_\eta})$. Further on, we define

$$\mathfrak{z}(t, s_1, \dots, s_m) = \sum_{\mathfrak{Q}^m} \frac{\partial^m}{\partial x^m} \sigma(t, X_t) D_{s(J_1)}^{|s(J_1)|} X_t \cdots D_{s(J_\nu)}^{|s(J_\nu)|} X_t, \quad (4.3)$$

and

$$\eta(t, s_1, \dots, s_m) = \sum_{\mathfrak{P}^m} \frac{\partial^m}{\partial x^m} b(t, X_t) D_{s(J_1)}^{|s(J_1)|} X_t \cdots D_{s(J_v)}^{|s(J_v)|} X_t, \quad (4.4)$$

where the sums run over the set \mathfrak{P}^m of all partitions $J_1 \cup \dots \cup J_v$ of $\{1, \dots, m\}$, where J_1, \dots, J_v are disjoint sets. We determine $\mathfrak{z}(t) = \sigma(t, X_t)$ as well. With these notations at hand, we find by (2.13) and induction that the n th order Malliavin derivative $D_{s_1, \dots, s_n}^n X_t$ satisfies the linear SDE

$$\begin{aligned} D_{s_1, \dots, s_n}^n X_t &= \sum_{i=1}^n \mathfrak{z}(s_i, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \\ &\quad + \int_{\hat{s}}^t \eta(u, s_1, \dots, s_n) du + \int_{\hat{s}}^t \mathfrak{z}(u, s_1, \dots, s_n) dW_u \end{aligned} \quad (4.5)$$

for $\hat{s} := \max\{s_1, \dots, s_n\} \leq t$ and $D_{s_1, \dots, s_n}^n X_t = 0$ else. Hence, its initial value is given by

$$\sum_{i=1}^n \mathfrak{z}(s_i, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n),$$

where

$$\begin{aligned} \mathfrak{z}(s_1, \dots, s_n) &= \frac{\partial^n}{\partial x^n} \sigma(s_1, X_{s_1}) \\ &\quad \times (D_{s_2} X_{s_1} \cdots D_{s_n} X_{s_1} + D_{s_2, s_3}^2 X_{s_1} \cdot D_{s_4} X_{s_1} \cdots D_{s_n} X_{s_1} + \cdots + D_{s_2, \dots, s_n}^{n-1} X_{s_1}). \end{aligned}$$

Finally, we apply [15, Theorem 2.9, page 289] as for the case $n = 1$ and taking into account that the assertion holds for $k = 1, \dots, n - 1$:

$$\begin{aligned} \mathbb{E}[|D_{s_1, \dots, s_n}^n X_t|^2] &\leq \mathbb{E}[\|Y_{s_1, \dots, s_n; t}^{(n)}\|^2] \\ &\leq C \left(1 + \mathbb{E}[|X_{\hat{s}}|^2] + \cdots + \mathbb{E} \left[\left| \sum_{i=1}^k \mathfrak{z}(s_i, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \right|^2 \right] \right) \\ &\leq C^n (1 + x_0^2). \end{aligned}$$

This completes the proof of Proposition 4.1. \square

Now, we proceed with the proof of the main result of Theorem 3.3, which follows the ideas of [21, Proof of Theorem 2.2]. Observe the decomposition

$$\mathbb{E}[(X_t^{p,k} - X_t)^2] \leq 2 \underbrace{\sum_{n=p+1}^{\infty} \sum_{|\alpha|=n} x_{\alpha}^2(t)}_{=: A_1(p)} + 2 \underbrace{\sum_{l=k+1}^{\infty} \sum_{n=0}^p \sum_{\substack{|\alpha|=n \\ d(\alpha)=l}} x_{\alpha}^2(t)}_{=: A_2(p,k)}, \quad (4.6)$$

where $d(\alpha) := \max\{i \geq 1 : \alpha_i > 0\}$. To determine an upper bound for $A_1(p)$, we use the chaos expansions (1.3) and (2.7) of X_t , i.e.,

$$X_t = \sum_{n=0}^{\infty} \sum_{|\alpha|=n} x_\alpha(t) \Psi^\alpha = \sum_{n=0}^{\infty} I_n(\xi_n(\mathbf{t}^n; t)),$$

with $\mathbf{t}^n = (t_1, \dots, t_n)$ and symmetric kernel functions $\xi_n(\mathbf{t}^n; t)$ being defined by

$$\xi_n(\cdot; t) = \frac{1}{n!} \mathbb{E}[D^n X_t]$$

for all $n \in \mathbb{N}_0$. We conclude that

$$\begin{aligned} \sum_{|\alpha|=n} x_\alpha^2(t) &= \mathbb{E}\left[\left(\sum_{|\alpha|=n} x_\alpha(t) \Psi^\alpha\right)^2\right] = \mathbb{E}[(I_n(\xi_n(\mathbf{t}^n; t)))]^2 \\ &= n! \langle \xi_n(\mathbf{t}^n; t), \xi_n(\mathbf{t}^n; t) \rangle_{\mathfrak{H}} \\ &= (n!)^2 \int^{(n);t} (\xi_n(\mathbf{t}^n; t))^2 d\mathbf{t}^n, \end{aligned} \quad (4.7)$$

where we abbreviate

$$\int^{(n);t} f(\cdot) d\mathbf{t}^n := \int_0^t \dots \int_0^{t_2} f(\cdot) dt_1 \dots dt_n.$$

Therefore, we deduce via (4.7) that

$$A_1(p) = \sum_{n=p+1}^{\infty} (n!)^2 \int^{(n);t} (\xi_n(\mathbf{t}^n; t))^2 d\mathbf{t}^n = \sum_{n=p+1}^{\infty} \int^{(n);t} \mathbb{E}[(D_{t_1, \dots, t_n}^n X_t)^2] d\mathbf{t}^n. \quad (4.8)$$

Finally, using (4.8) and applying Proposition 4.1 we obtain

$$\begin{aligned} A_1(p) &= \sum_{n=p+1}^{\infty} \int^{(n);t} \mathbb{E}[(D_{t_1, \dots, t_n}^n X_t)^2] d\mathbf{t}^n \\ &\leq (1 + x_0^2) \sum_{n=p+1}^{\infty} C^n \int^{(n);t} d\mathbf{t}^n \\ &= (1 + x_0^2) \sum_{n=p+1}^{\infty} \frac{(Ct)^n}{n!} \\ &\leq C(1 + x_0^2) \frac{1}{(p+1)!}. \end{aligned} \quad (4.9)$$

The treatment of the term $A_2(p, k)$ is more involved. Recalling the definition of $d(\alpha) = \max\{i \geq 1 : \alpha_i > 0\}$, we introduce the characteristic set (i_1, \dots, i_n) of $\alpha \in \mathcal{I}$ for $i_1 \leq i_2 \leq \dots \leq i_n$ and $|\alpha| = n$. It is defined as follows: i_1 is the index of the first non-zero component of α . If $\alpha_{i_1} = 1$ then i_2 is the index of the second non-zero component of α ; otherwise $i_1 = i_2 = \dots = i_{\alpha_{i_1}}$ and $i_{\alpha_{i_1}+1}$ is the second non-zero component of α . The same operation is repeated for the index $i_{\alpha_{i_1}+1}$ and further non-zero components of α . In this fashion, the characteristic set is constructed, resulting in the observation that $d(\alpha) = i_n$. To give a simple example, consider the multiindex $\alpha = (2, 0, 1, 4, 0, 0, \dots) \in \mathcal{I}$. Here the characteristic set of α is given by

$$(1, 1, 3, 4, 4, 4, 4).$$

For any $\alpha \in \mathcal{I}$ with $|\alpha| = n$ and a basis $(e_i)_{i \geq 1}$ of $\mathbb{H} = L^2([0, T])$, we denote by $\tilde{e}_\alpha(\mathbf{t}^n)$ the (scaled) symmetrised form of $\bigotimes e_i^{\otimes \alpha_i}$ defined via

$$\tilde{e}_\alpha(\mathbf{t}^n) := \sum_{\pi \in \mathfrak{P}^n} e_{i_1}(t_{\pi(1)}) \cdots e_{i_n}(t_{\pi(n)}), \quad (4.10)$$

where (i_1, \dots, i_n) is the characteristic set of α and the sum runs over all permutations π within the permutation group \mathfrak{P}^n of $\{1, \dots, n\}$. From (3.1) we know that $x_\alpha(t) = \mathbb{E}[X_t \Psi^\alpha]$. On the other hand, we have the representation

$$\Psi^\alpha = \frac{1}{\sqrt{\alpha!}} I_{|\alpha|} \left(\bigotimes_{i=1}^{\infty} e_i^{\otimes \alpha_i} \right) = \frac{1}{\sqrt{\alpha!}} I_{|\alpha|} \left(\frac{1}{|\alpha|!} \tilde{e}_\alpha \right).$$

Now, for any $\alpha \in \mathcal{I}$ with $|\alpha| = n$, we obtain the identity

$$x_\alpha(t) = \frac{n!}{\sqrt{\alpha!}} \int^{(n);t} \xi_n(\mathbf{t}^n; t) \tilde{e}_\alpha(\mathbf{t}^n) d\mathbf{t}^n, \quad (4.11)$$

by (2.11). Since

$$\tilde{e}_\alpha(\mathbf{t}^n) = \sum_{j=1}^n e_{i_n}(t_j) \cdot \tilde{e}_{\alpha^-(i_n)}(\mathbf{t}_j^n),$$

where \mathbf{t}_j^n is obtained from \mathbf{t}^n by omitting t_j and $\alpha^-(\cdot)$ denotes the diminished multi-index as defined in (2.10), we deduce that

$$x_\alpha(t) = \frac{n!}{\sqrt{\alpha!}} \sum_{j=1}^n \int^{(n-1);t} \left(\int_{t_{j-1}}^{t_{j+1}} \xi_n(\mathbf{t}^n; t) e_{i_n}(t_j) dt_j \right) \tilde{e}_{\alpha^-(i_n)}(\mathbf{t}_j^n) d\mathbf{t}_j^n, \quad (4.12)$$

with $t_0 := 0$ and $t_{n+1} := t$, by changing the order of integration. Then for any $i_n = l \geq 1$ we integrate by parts to deduce

$$\int_{t_{j-1}}^{t_{j+1}} \xi_n(\mathbf{t}^n; t) e_l(t_j) dt_j = [\xi_n(\mathbf{t}^n; t) E_l(t_j)]_{t_j=t_{j-1}}^{t_j=t_{j+1}} - \int_{t_{j-1}}^{t_{j+1}} \frac{\partial}{\partial t_j} \xi_n(\mathbf{t}^n; t) E_l(t_j) dt_j, \quad (4.13)$$

where

$$E_i(s) = \int_0^s e_i(u) du.$$

Now, we use substitution in (4.12) by renaming \mathbf{t}_j^n in the following way for each j : With $s_i = t_i$ for all $i \leq j - 1$ and $s_i = t_{i+1}$ for all $i > j - 1$, we have $\mathbf{s}^{n-1} := \mathbf{t}_j^n$ by setting $s_0 = 0$ and $s_n = t$. Moreover, we denote with $\mathbf{s}^{n-1,r}$, $r = 1, \dots, n - 1$, the set that is generated from \mathbf{s}^{n-1} by taking s_r twice. To finalize this notation, we set $\mathbf{s}^{n-1,0} = (s_0, s_1, \dots, s_{n-1})$ and $\mathbf{s}^{n-1,n} = (s_1, \dots, s_{n-1}, s_n)$. Then

$$[\xi_n(\mathbf{t}^n; t) E_l(t_j)]_{t_j=t_{j-1}}^{t_j=t_{j+1}} = \xi_n(\mathbf{s}^{n-1,j}) E_l(s_j) - \xi_n(\mathbf{s}^{n-1,j-1}) E_l(s_{j-1}), \quad j = 1, \dots, n.$$

Because $E_l(s_0) = E_l(0) = 0$ and $E_l(s_n) = E_l(t)$, from (4.13) we see that by summing over j all terms except one cancel out. Hence, setting

$$\begin{aligned} \psi_l(\mathbf{s}^{n-1}; t) &:= \xi_n(\mathbf{s}^{n-1,n}) E_l(t) \\ &- \int_0^{s_1} \frac{\partial}{\partial s_1} \xi_n(\tau, \mathbf{s}^{n-1}; t) E_l(\tau) d\tau \\ &- \sum_{j=2}^{n-1} \int_{s_{j-1}}^{s_j} \frac{\partial}{\partial s_j} \xi_n(\dots, s_{j-1}, \tau, s_{j+1}, \dots; t) E_l(\tau) d\tau \\ &- \int_{s_{n-1}}^t \frac{\partial}{\partial s_n} \xi_n(\mathbf{s}^{n-1}, \tau; t) E_l(\tau) d\tau, \end{aligned}$$

we obtain from (4.12)

$$\begin{aligned} \sum_{\substack{|\alpha|=n \\ i_n=d(\alpha)=l}} x_\alpha^2(t) &= \sum_{\substack{|\alpha|=n \\ i_n=d(\alpha)=l}} \left(\frac{n!}{\sqrt{\alpha}!} \int^{(n-1);t} \psi_l(\mathbf{s}^{n-1}; t) \tilde{e}_{\alpha-(l)}(\mathbf{s}^{n-1}) d\mathbf{s}^{n-1} \right)^2 \\ &\leq n^2 \sum_{|\beta|=n-1} \left(\frac{(n-1)!}{\sqrt{\beta}!} \int^{(n-1);t} \psi_l(\mathbf{s}^{n-1}; t) \tilde{e}_\beta(\mathbf{s}^{n-1}) d\mathbf{s}^{n-1} \right)^2, \end{aligned}$$

since $|\alpha^-(i_{|\alpha|})| = |\alpha| - 1$ and $\alpha! \geq \alpha^-(i_{|\alpha|})!$. In order to interpret the last sum, consider the random variable $I_{n-1}(\psi_l(\mathbf{s}^{n-1}; t))$. Then, following (4.7) and the identity (4.11), we conclude that

$$\begin{aligned} \mathbb{E}[(I_{n-1}(\psi_l(\mathbf{s}^{n-1}; t)))^2] &= (n-1)!^2 \int^{(n-1);t} \psi_l^2(\mathbf{s}^{n-1}; t) d\mathbf{s}^{n-1} \\ &= \sum_{|\beta|=n-1} \left(\frac{(n-1)!}{\sqrt{\beta}!} \int^{(n-1);t} \psi_l(\mathbf{s}^{n-1}; t) \tilde{e}_\beta(\mathbf{s}^{n-1}) d\mathbf{s}^{n-1} \right)^2. \end{aligned}$$

Thus, we get the estimate

$$\sum_{\substack{|\alpha|=n \\ i_n=l}} x_\alpha^2(t) \leq (n!)^2 \int^{(n-1);t} \psi_l^2(\mathbf{s}^{n-1}; t) d\mathbf{s}^{n-1}.$$

Now, using the Cauchy–Schwarz inequality we deduce that

$$\begin{aligned} \psi_t^2(\mathbf{s}^{n-1}; t) &\leq (n+1) \left(\xi_n^2(\mathbf{s}^{n-1, n}) E_t^2(t) \right. \\ &\quad \left. + \int_0^t E_t^2(\tau) d\tau \cdot \sum_{j=1}^n \int_{s_{j-1}}^{s_j} \left| \frac{\partial}{\partial s_j} \xi_n(\dots s_{j-1}, \tau, s_{j+1}, \dots; t) \right|^2 d\tau \right). \end{aligned}$$

Recall the identity $\xi_n(\cdot; t) = (n!)^{-1} \mathbb{E}[D^n X_t]$. Using similar arguments as in the proof of Proposition 4.1, we obtain the inequality

$$\left| \frac{\partial}{\partial t_i} \xi_n(t_1, \dots, t_n; t) \right|^2 \leq (1 + x_0^2) \frac{C^n}{(n!)^2}, \quad i = 1, \dots, n.$$

Since $|s_j - s_{j-1}| \leq t$, we finally conclude that

$$\psi_t^2(\mathbf{s}^{n-1}; t) \leq (1 + x_0^2) \frac{C^n(n+1)}{(n!)^2} \left(E_t^2(t) + \int_0^t E_t^2(\tau) d\tau \right). \quad (4.14)$$

Consequently, we deduce

$$\sum_{\substack{|\alpha|=n \\ i_n=l}} x_\alpha^2(t) \leq (1 + x_0^2) \frac{C^n(n+1)}{(n-1)!} \left(E_t^2(t) + \int_0^t E_t^2(\tau) d\tau \right).$$

Putting things together we conclude

$$\begin{aligned} A_2(p, k) &= \sum_{l=k+1}^{\infty} \sum_{n=1}^p \sum_{\substack{|\alpha|=n \\ i_n=l}} x_\alpha^2(t) \\ &\leq (1 + x_0^2) \sum_{n=1}^{\infty} \frac{C^n(n+1)}{(n-1)!} \cdot \sum_{l=k+1}^{\infty} \left(E_t^2(t) + \int_0^t E_t^2(\tau) d\tau \right) \\ &\leq C(1 + x_0^2) \sum_{l=k+1}^{\infty} \left(E_t^2(t) + \int_0^t E_t^2(\tau) d\tau \right). \end{aligned} \quad (4.15)$$

Combining (4.9) and (4.15) we obtain the result of Theorem 3.3.

5 Numerical results and sparse truncation

To get an idea of the performance of the polynomial chaos expansion for different bases and truncations, we apply the trigonometric basis from Example 3.4 (i) and the Haar basis from Example 3.4 (ii) to the geometric Brownian motion from Example 3.2,

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \quad X_0 = x_0 \quad (5.1)$$

on the horizon $t \in [0, 1]$ with $\mu = \sigma = x_0 = 1$. In this setting we have the identity

$$\text{Var}(X_t) = x_0^2 \exp(2\mu t) (\exp(\sigma^2 t) - 1),$$

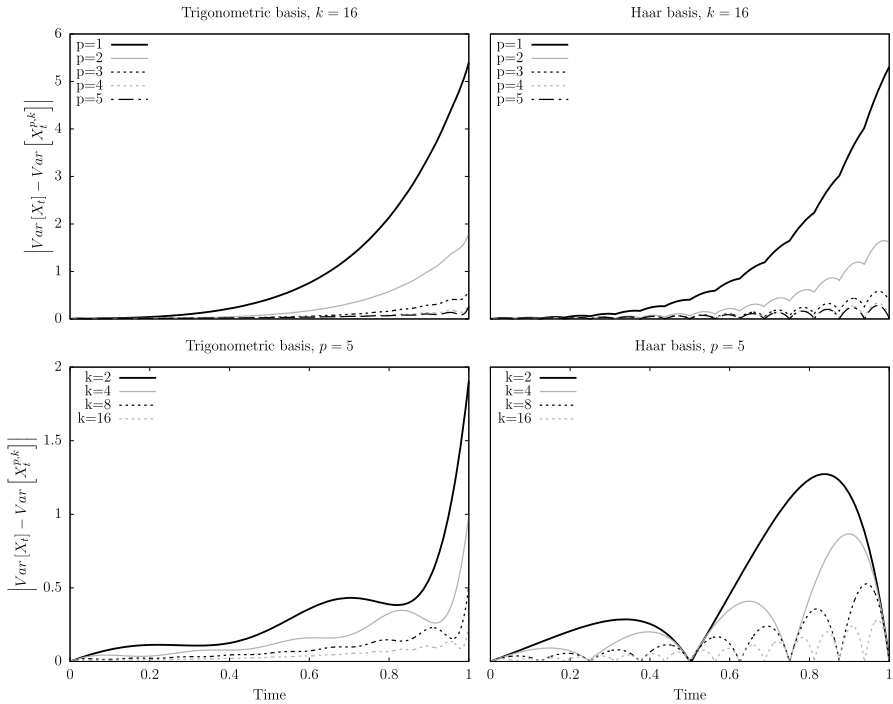


Fig. 1. Absolute errors of the variance as calculated with different polynomial chaos expansions for the Geometric Brownian Motion, $|\text{Var}[X_t] - \text{Var}[X_t^{p,k}]|$. The trigonometric and Haar bases (with $k = 2^n$) from Example 3.4 are shown

which can be used to compare the absolute errors of approximated variances using different bases and values for k and p . We used the `ode45` solver with standard settings within `Matlab 2017b` (64bit, single core) on an Intel Core i5-6300U (2.4GHz, 8GB RAM).

5.1 Absolute errors of variances

Figure 1 shows the absolute errors between the variances of the analytical solution and different polynomial chaos expansions. One observes the fast decay of the approximation error for fixed k and increasing p , and vice versa. The main difference between the observed errors is that the maximum of the trigonometric expansion is always at the end of the time horizon at $t = 1$, while the Haar basis leads to almost zero error on k equidistant time points, including $t = 0$ and $t = 1$.

Table 1 in the Appendix lists approximation errors, numbers of coefficients, and computational times for several choices of p and k . The exponential increase in the number of coefficients with increasing p and k leads to an exponential increase in runtime, as expected. To overcome this issue, we propose a heuristic to reduce the computational burden.

5.2 Sparse truncation

The information contained in the coefficient functions $x_\alpha(\cdot)$ decays with increasing order p of the basis polynomials Ψ^α and the decaying rate of the Gaussian expansion, i.e., the index k of the basis polynomials e_k used for constructing Ψ^α [22].

Hence, we can define *sparse index sets* for reducing the number of multi-indices α from the full truncated index set $\mathcal{I}_{p,k}$ without losing too much information, i.e., accuracy in a numerical simulation.

Definition 1 (Sparse truncation of first order). Let p be the maximum order of the index $\alpha \in \mathcal{I}_{p,k}$. Then the *first order sparse index* $\mathfrak{r} = (r_1, \dots, r_k)$ satisfies $p = r_1 \geq r_2 \geq \dots \geq r_k$ and we define the *first order sparse index set*

$$\mathcal{I}_{p,k}^{\mathfrak{r}} = \{\alpha \in \mathcal{I}_{p,k} \mid \alpha_i \leq r_i \forall i \leq k\}. \quad (5.2)$$

Example 5.1. Let $k = 5$ and $p = 3$. Then a possible choice of the sparse index is $\mathfrak{r} = (3, 2, 2, 1, 1)$. For constructing the first order polynomials all five first order basis polynomials can be used. The second order polynomials are comprised by all possible combinations of first order basis polynomials depending on e_1, \dots, e_5 and the second order basis polynomials of e_1, e_2, e_3 . Analogously, the third order polynomials are constructed. By using this first order sparse index set $\mathcal{I}_{p,k}^{\mathfrak{r}}$ the number of coefficient functions $x_\alpha(\cdot)$ appearing within the propagator system (3.2) can be reduced drastically without impairing the solution much. The full index set $\mathcal{I}_{p,k}$ consists of $\frac{(k+p)!}{k!p!} = 56$ terms, whereas this first order sparse index set includes 42 terms.

An even more rigorous reduction of the number of coefficients included in the propagator system can be achieved by using a *second order sparse index*, i.e., a series of first order sparse indices $(\mathfrak{r}^j)_{j=0,\dots,p}$ that depend on the actual order of the polynomials Ψ^α with $j = |\alpha|$, [22]. This approach allows to exclude crossing products of random variables $W(e_i)$ from the construction of higher order basis polynomials Ψ^α that add only negligible information to the system.

Definition 2 (Sparse truncation of second order). Let p be the maximum order of the index $\alpha \in \mathcal{I}_{p,k}$. Then the *second order sparse index* $(\mathfrak{r}) = (\mathfrak{r}^j)_{j \leq p}$ is a series of first order sparse indices $\mathfrak{r}^j = (r_1^j, \dots, r_k^j)$ satisfying $j = r_1^j \geq r_2^j \geq \dots \geq r_k^j$ for all $j = |\alpha| \leq p$ and we define the *second order sparse index set*

$$\mathcal{I}_{p,k}^{(\mathfrak{r})} = \{\alpha \in \mathcal{I}_{p,k} \mid \alpha_i \leq r_i^j \forall i \leq k, \forall j \leq p\}. \quad (5.3)$$

Example 5.2. Considering again the setting of the previous example with $k = 5$ and $p = 3$, one possible choice of a second order sparse index is given via $\mathfrak{r}^1 = (1, 1, 1, 1, 1)$, $\mathfrak{r}^2 = (2, 2, 2, 1, 0)$, and $\mathfrak{r}^3 = (3, 2, 0, 0, 0)$. In constructing basis polynomials of order $|\alpha| = 3$ we can use all combinations of basis polynomials depending on the first two random variables e_1 and e_2 up to orders 3 and 2, respectively. Thus, these are $\sqrt{6}H_3(e_1)$, $\sqrt{2}H_2(e_1)H_1(e_2)$, and $\sqrt{2}H_1(e_1)H_2(e_2)$, compare [22].

Table 2 in the Appendix lists the first and second order sparse indices that were used for the numerical study. The results in Table 1 show that at the price of a slightly higher error, and of course the loss of a guaranteed upper bound as in (3.6), the computational times could be reduced by several orders of magnitude.

Appendix

Table 1. Computational times in seconds and error estimates for trigonometric and Haar bases from Example 3.4 and different values of p and k . The number of coefficients n_{coeff} depends also on the type of truncation, compare Table 2 for details. $\text{error}_{t=1}$ is the absolute error of the expansion's variance $|\text{Var}[X_t] - \text{Var}[X_t^{p,k}]|$ at $t = 1$ while $\text{error}_{\text{max}}$ is the maximum value over $t \in [0, 1]$

k	p	n_{coeff}	type	Trigonometric basis			Haar basis		
				time	$\text{error}_{t=1}$	$\text{error}_{\text{max}}$	time	$\text{error}_{t=1}$	$\text{error}_{\text{max}}$
2	1	3	full	0.04	6.04	6.04	0.03	5.31	5.31
4	1	5	full	0.02	5.68	5.68	0.02	5.31	5.31
8	1	9	full	0.22	5.49	5.49	0.04	5.30	5.30
16	1	17	full	0.05	5.40	5.40	0.19	5.31	5.31
32	1	33	full	0.17	5.35	5.35	1.31	5.30	5.30
64	1	65	full	1.28	5.33	5.33	10.12	5.31	5.31
2	2	6	full	0.02	3.04	3.04	0.02	1.61	1.83
4	2	15	full	0.03	2.35	2.35	0.05	1.61	1.76
8	2	45	full	0.20	1.98	1.98	0.82	1.61	1.69
8	2	41	sp ¹	0.31	1.99	1.99	0.24	1.61	1.69
8	2	19	sp ²	0.04	2.16	2.16	0.07	1.61	1.72
16	2	153	full	3.85	1.80	1.80	18.61	1.61	1.65
16	2	141	sp ³	1.00	1.80	1.80	4.60	1.61	1.65
16	2	27	sp ⁴	0.05	2.07	2.07	0.19	1.61	1.67
32	2	561	full	86.83	1.71	1.71	554.61	1.61	1.63
32	2	537	sp ⁵	22.80	1.71	1.71	143.90	1.61	1.63
32	2	69	sp ⁶	0.31	1.84	1.84	3.22	1.61	1.64
64	2	2145	full	2189.70	1.66	1.66	17234.51	1.61	1.62
2	3	10	full	0.02	2.15	2.15	0.02	0.38	1.35
4	3	35	full	0.10	1.29	1.29	0.19	0.38	1.01
8	3	165	full	2.76	0.84	0.84	10.58	0.38	0.76
8	3	127	sp ⁷	0.53	0.85	0.85	1.77	0.37	0.76
8	3	37	sp ⁸	0.06	1.11	1.11	0.13	0.38	0.86
16	3	969	full	200.34	0.61	0.61	1070.17	0.38	0.58
16	3	763	sp ⁹	45.61	0.62	0.62	169.16	0.38	0.59
16	3	45	sp ¹⁰	0.14	1.02	1.02	0.39	0.38	0.76
2	4	15	full	0.02	1.94	1.94	0.02	0.07	1.32
4	4	70	full	0.30	1.04	1.04	0.79	0.07	0.87
8	4	495	full	26.99	0.57	0.57	117.55	0.07	0.56
8	4	303	sp ¹¹	4.34	0.57	0.57	12.30	0.07	0.52
8	4	32	sp ¹²	0.05	0.96	0.96	0.12	0.07	0.75
16	4	4845	full	5986.44	0.32	0.32	30591.94	0.07	0.33
16	4	40	sp ¹³	0.08	0.87	0.87	0.36	0.07	0.67
32	4	92	sp ¹⁴	0.50	0.59	0.59	3.56	0.07	0.45
2	5	21	full	0.03	1.91	1.91	0.03	0.01	1.27
4	5	126	full	0.92	1.00	1.00	2.04	0.01	0.87
8	5	1287	full	192.98	0.51	0.51	855.07	0.01	0.53
8	5	599	sp ¹⁵	10.04	0.51	0.51	50.03	0.01	0.55
8	5	36	sp ¹⁶	0.03	0.92	0.92	0.11	0.01	0.74
16	5	20349	full	120469.96	0.26	0.26	600591.17	0.01	0.28
16	5	44	sp ¹⁷	0.06	0.83	0.83	0.28	0.01	0.65
32	5	98	sp ¹⁸	0.51	0.55	0.55	3.60	0.01	0.42

Table 2. List of first and second order sparse indices used in Section 5.2 together with the number of resulting coefficient functions. The reference numbers coincide with those in Table 1

symbol	p	k	order	index $\mathfrak{r}/(\mathfrak{r})$	n_{coeff}
sp ¹	2	8	1	$\mathfrak{r} = (2, 2, 2, 2, 1, 1, 1, 1)$	41
sp ²	2	8	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, 0, 0, 0)$	19
sp ³	2	16	1	$\mathfrak{r} = (2, 2, 2, 2, 1, \dots, 1)$	141
sp ⁴	2	16	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, \dots, 0)$	27
sp ⁵	2	32	1	$\mathfrak{r} = (2, 2, 2, 2, 2, 2, 2, 2, 1, \dots, 1)$	537
sp ⁶	2	32	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 2, 2, 2, 2, 0, \dots, 0)$	69
sp ⁷	3	8	1	$\mathfrak{r} = (3, 3, 2, 2, 1, 1, 1, 1)$	127
sp ⁸	3	8	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, 0, 0, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 2, 0, 0, 0, 0)$	37
sp ⁹	3	16	1	$\mathfrak{r} = (3, 3, 2, 2, 1, \dots, 1)$	763
sp ¹⁰	3	16	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 2, 0, \dots, 0)$	45
sp ¹¹	4	8	1	$\mathfrak{r} = (4, 4, 2, 2, 1, 1, 1, 1)$	303
sp ¹²	4	8	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^4 = (4, 3, 0, \dots, 0)$	32
sp ¹³	4	16	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 0, \dots, 0)$ $\mathfrak{r}^4 = (4, 3, 0, \dots, 0)$	40
sp ¹⁴	4	32	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^4 = (4, 4, 0, \dots, 0)$	92
sp ¹⁵	5	8	1	$\mathfrak{r} = (5, 5, 2, 2, 1, 1, 1, 1)$	599
sp ¹⁶	5	8	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^4 = (4, 3, 0, \dots, 0)$ $\mathfrak{r}^5 = (5, 3, 0, \dots, 0)$	36
sp ¹⁷	5	16	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 0, \dots, 0)$ $\mathfrak{r}^4 = (4, 3, 0, \dots, 0)$ $\mathfrak{r}^5 = (5, 3, 0, \dots, 0)$	44
sp ¹⁸	5	32	2	$\mathfrak{r}^1 = (1, \dots, 1)$ $\mathfrak{r}^2 = (2, 2, 2, 2, 2, 2, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^3 = (3, 3, 2, 2, 0, \dots, 0)$ $\mathfrak{r}^4 = (4, 4, 0, \dots, 0)$ $\mathfrak{r}^5 = (5, 5, 0, \dots, 0)$	98

Funding

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 647573), from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 314838170, GRK 2297 MathCoRe, from the project “Ambit fields: probabilistic properties and statistical inference” funded by Villum Fonden, and from CREATES funded by the Danish National Research Foundation.

References

- [1] Arnold, S.: Approximation schemes for SDEs with discontinuous coefficients. PhD thesis, ETH Zürich (2006)
- [2] Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Relat. Fields* **104**(1), 43–60 (1996). [MR1367666](#). <https://doi.org/10.1007/BF01303802>
- [3] Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.* **2**(2), 93–128 (1996). [MR1401964](#). <https://doi.org/10.1515/mcma.1996.2.2.93>
- [4] Benth, F.E., Gjerde, J.: Convergence rates for finite element approximations of stochastic partial differential equations. *Stoch. Stoch. Rep.* **63**, 313–326 (1998). [MR1658087](#). <https://doi.org/10.1080/17442509808834153>
- [5] Benth, F.E., Theting, T.G.: Some regularity results for the stochastic pressure equation of Wick-type. *Stoch. Anal. Appl.* **20**, 1191–1223 (2002). [MR1943795](#). <https://doi.org/10.1081/SAP-120015830>
- [6] Cameron, R.H., Martin, W.T.: The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Ann. Math.* **48**, 385–392 (1947). [MR0020230](#). <https://doi.org/10.2307/1969178>
- [7] Cao, Y.: On convergence rate of Wiener-Itô expansion for generalized random variables. *Stochastics* **78**, 179–187 (2006). [MR2241915](#). <https://doi.org/10.1080/17442500600768641>
- [8] Chan, K.S., Stramer, O.: Weak consistency of the Euler method for numerically solving stochastic differential equations with discontinuous coefficients. *Stoch. Process. Appl.* **76**(1), 33–44 (1998). [MR1638015](#). [https://doi.org/10.1016/S0304-4149\(98\)00020-9](https://doi.org/10.1016/S0304-4149(98)00020-9)
- [9] Halidias, N., Kloeden, P.E.: A note on strong solution of stochastic differential equations with a discontinuous drift coefficient. *J. Appl. Math. Stoch. Anal.* (2006). [MR2221001](#). <https://doi.org/10.1155/JAMSA/2006/73257>
- [10] Holden, H., Øksendal, B., Ubøe, J., Zhang, T.: *Stochastic partial differential equations*. Springer (2010). [MR2571742](#). <https://doi.org/10.1007/978-0-387-89488-1>
- [11] Huschto, T., Sager, S.: Stochastic optimal control in the perspective of the Wiener chaos. In: *Proceedings of the 12th European Control Conference*, pp. 3059–3064 (2013)
- [12] Huschto, T., Sager, S.: Solving stochastic optimal control problems by a Wiener chaos approach. *Vietnam J. Math.* **42**, 83–113 (2014). [MR3167032](#). <https://doi.org/10.1007/s10013-014-0060-8>
- [13] Isobe, E., Sato, S.: Wiener-Hermite expansion of a process generated by an Itô stochastic differential equation. *J. Appl. Probab.* **20**, 754–765 (1983). [MR0720467](#). <https://doi.org/10.2307/3213587>

- [14] Jacod, J., Protter, P.: Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26**, 267–307 (1998). [MR1617049](#). <https://doi.org/10.1214/aop/1022855419>
- [15] Karatzas, I., Shreve, S.E.: *Brownian motion and stochastic calculus*. Springer (2007). [MR1121940](#). <https://doi.org/10.1007/978-1-4612-0949-2>
- [16] Kloeden, P.E., Platen, E.: *Numerical solution of stochastic differential equations. Applications of Mathematics (New York)*, vol. 23. Springer, Berlin (1992). [MR1214374](#). <https://doi.org/10.1007/978-3-662-12616-5>
- [17] Kurtz, T.G., Protter, P.: Wong-Zakai corrections, random evolutions and numerical schemes for SDEs. In: *Stochastic Analysis*, pp. 331–346. Academic Press, New York (1991). [MR1119837](#). <https://doi.org/10.1016/B978-0-12-481005-1.50023-5>
- [18] Levajković, T., Seleši, D.: Chaos expansion methods for stochastic differential equations involving the Malliavin derivative-Part I. *Publ. Inst. Math.* **90**, 65–84 (2011). [MR2907094](#). <https://doi.org/10.2298/PIM1104065L>
- [19] Levajković, T., Seleši, D.: Chaos expansion methods for stochastic differential equations involving the Malliavin derivative-Part II. *Publ. Inst. Math.* **90**, 85–98 (2011). [MR2907095](#). <https://doi.org/10.2298/PIM1104085L>
- [20] Lototsky, S.V., Rozovskii, B.L.: Stochastic differential equations: A Wiener chaos approach. In: *From stochastic calculus to mathematical finance*, pp. 433–506. Springer, Berlin, Heidelberg (2006). [MR2234286](#). https://doi.org/10.1007/978-3-540-30788-4_23
- [21] Lototsky, S.V., Mikulevičius, R., Rozovskii, B.L.: Nonlinear filtering revisited: A spectral approach. *SIAM J. Control Optim.* **35**, 435–446 (1997). [MR1436632](#). <https://doi.org/10.1137/S0363012993248918>
- [22] Luo, W.: *Wiener chaos expansion and numerical solutions of stochastic partial differential equations*. PhD thesis, Caltech (2006). [MR3078549](#)
- [23] Mikulevičius, R., Platen, E.: Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.* **15**, 233–239 (1991). [MR1121206](#). <https://doi.org/10.1002/mana.19911510114>
- [24] Nualart, D.: *The Malliavin calculus and related topics*, 2nd edn. Springer, Berlin (2006). [MR2200233](#)
- [25] Veretennikov, A.Y., Krylov, N.V.: On explicit formulas for solutions of stochastic equations. *Math. USSR Sb.* **29**, 239–256 (1976). <https://doi.org/10.1070/SM1976v029n02ABEH003666>
- [26] Wan, X., Rozovskii, B.L., Karniadakis, G.E.: A stochastic modeling methodology based on weighted Wiener chaos and Malliavin calculus. In: *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 14189–14194 (2009). [MR2539729](#). <https://doi.org/10.1073/pnas.0902348106>
- [27] Yan, L.: The Euler scheme with irregular coefficients. *Ann. Probab.* **30**(3), 1172–1194 (2002). [MR1920104](#). <https://doi.org/10.1214/aop/1029867124>