

Combined Error Estimates for Local Fluctuations of SPDEs

Christian Kuehn and Patrick Kürschner

August 6, 2018

Abstract

In this work, we study the numerical approximation of local fluctuations of certain classes of parabolic stochastic partial differential equations (SPDEs). Our focus is on effects for small spatially-correlated noise on a time scale before large deviation effects have occurred. In particular, we are interested in the local directions of the noise described by a covariance operator. We introduce a new strategy and prove a Combined ERror ESTimate (CERES) for the four main errors: the spatial discretization error, the local linearization error, the local relaxation error to steady state, and the approximation error via an iterative low-rank matrix algorithm. In summary, we obtain one CERES describing, apart from modelling of the original equations and standard round-off, all the sources of error for a local fluctuation analysis of an SPDE in one estimate. To prove our results, we rely on a combination of methods from optimal Galerkin approximation of SPDEs, covariance moment estimates, analytical techniques for Lyapunov equations, iterative numerical schemes for low-rank solution of Lyapunov equations, and working with related spectral norms for different classes of operators.

Keywords: stochastic partial differential equation, combined error estimates, optimal regularity, Lyapunov equation, low-rank approximation, local fluctuations.

1 Introduction

This work has two main goals. The first - more abstract - goal is to establish a general strategy to find and prove Combined ERror ESTimates (CERES)¹ for dynamical systems involving several sources of error. The second - more specific - goal is to demonstrate CERES for a concrete challenge of an infinite-dimensional stochastic problem. The technically precise formulation of our work starts in Section 2. In this introduction we provide a basic overview of our strategy and our main results. We focus on the evolution equation

$$du = [Au + f(u)] dt + g(u) dW, \quad (1)$$

¹CERES is not only a direct acronym for the main strategy of this work but also contains a nice historical note: The asteroid Ceres was predicted by Gauß using the method of least-squares, which also 'combines' the minimization of several errors to a single function.

where $W = W(x, t)$ is a certain white-noise process, A is a suitable linear operator, f, g are given maps and $u = u(x, t)$ is the unknown function [43]. As a paradigmatic example, one may think of the Laplacian $A = \Delta$, W as a Q -Wiener process with trace-class operator Q and f, g as sufficiently smooth Lipschitz functions- Suppose the deterministic problem, i.e. $g(u) \equiv 0$, has a steady state solution $u = u^*$, which solves $0 = Au^* + f(u^*)$. Suppose the steady state is linearly stable, which just means that the spectrum of the linear operator

$$\tilde{A} := A + D_u f(u^*) \tag{2}$$

is properly contained in the left-half of the complex plane [27, 45]. Assume that the initial condition $u(x, 0)$ is very close to u^* and the noise in (1) is sufficiently small, $0 < \|g(u)\| =: \nu \leq 1$, in comparison to the spectral gap of \tilde{A} to the imaginary axis. Then the probability is extremely large that one observes only local fluctuations of sample paths of the SPDE (1) near u^* on very long time scales. Only when the time scale reaches roughly order $\mathcal{O}(e^{-c/\nu^2})$ for some constant $c > 0$ as $\nu \rightarrow 0$, large deviation effects [53, 43] occur. Here we focus on the initial scale of fluctuations, which we refer to as sub-exponential scale, on which large deviation effects do not play a role. However, the noise does still play an important role near the steady state. Its interplay with the operator \tilde{A} determines the directions, in which we are going to find the process with higher probability locally near u^* . This raises the question, how to numerically compute these directions. A recent practical strategy suggested in the context of a numerical continuation [28] framework for stochastic ordinary differential equations (SODEs) [31] and then extended for SPDEs [32] is to:

- (S1) spatially discretize u with approximation level h and consider the resulting SODEs for a large vector $u_h \in \mathbb{R}^N$;
- (S2) locally linearize the SODEs around u_h^* and consider SODEs for the linear approximation $U_h \in \mathbb{R}^N$, which form an Ornstein-Uhlenbeck (OU) process [16];
- (S3) take the covariance matrix $V_h = V_h(t)$ of the OU process, which satisfies a time-dependent Lyapunov equation [1, 48], and show that it converges quickly to a stationary Lyapunov equation for a matrix $V_* \in \mathbb{R}^{N \times N}$;
- (S4) compute the matrix $V_* \approx ZZ^\top$ using a specialized iterative method involving a low-rank approximation $Z \in \mathbb{R}^{N \times \mathfrak{r}}$ with $\mathfrak{r} \ll N$; j computation steps of an iterative algorithm yield a matrix $V_j \approx V_*$.

Every step (S1)-(S4) produces an error, i.e., the final matrix V_j only provides an approximation to the infinite-dimensional covariance operator $\text{Cov}(u)$ [18]. Hence, one should aim for a result of the form

$$\sup_{t \in [0, T]} \|\text{Cov}(u) - V_j\| \leq \frac{\text{error in}}{\text{Step (S1)}} + \frac{\text{error in}}{\text{Step (S2)}} + \frac{\text{error in}}{\text{Step (S3)}} + \frac{\text{error in}}{\text{Step (S4)}} \tag{3}$$

to really judge the quality from the viewpoint of numerical analysis. Equation (3) is just a prototypical example, i.e., a chain of different error terms does occur in most challenging high-dimensional problems, particularly those involving stochastic aspects.

Remark: We do not consider in (3) the modelling error of SPDEs of the form (1). On the one hand, it is partially included in a stochastic formulation anyway and on the other hand it is always possible to argue in an application, whether other terms or effects matter. The second error we do not include in the CERES is the standard numerical round-off error, which is universal for a given precision.

From a technical viewpoint, each step demands different techniques and then a combination of the different estimates. For this CERES, we decided to consider the spectral or 2-norm $\|\cdot\|_2$ on the infinite- as well as finite-dimensional levels. This simplifies computing a CERES considerably as one may use the standard triangle inequality and suitable embeddings for the expression

$$\| \underbrace{\text{Cov}(u) - \text{Cov}(u_h)}_{\text{Step (S1)}} + \underbrace{\text{Cov}(u_h) - V_h}_{\text{Step (S2)}} + \underbrace{V_h - V_*}_{\text{Step (S3)}} + \underbrace{V_* - V_j}_{\text{Step (S4)}} \|_2. \quad (4)$$

For (S1), we rely on an extension to covariance operators of optimal error estimates for Galerkin finite elements methods for SPDEs [29, 30]. (S2) is treated via small noise approximation in combination with moment equations [49, 33]. (S3) is probably the simplest step as classical results on spectra for Lyapunov equations [5] and decay of the time-dependent problem [23] can be combined. (S4) requires a careful tracing of error estimates for low-rank versions of iterative algorithms, such as alternating direction implicit (ADI) and rational Krylov methods [11].

Our final result for (3) is summarized in Theorem 8.1. It illustrates that many factors can influence the error. For example, the spatial resolution h , the final time T , the Lipschitz constants of f, g , the structure of the operator Q , the spectrum of \tilde{A} , and the low-rank \mathfrak{r} all appear in some form in the final error. Therefore, *balancing* a CERES is the key practical message of our work. Just making a spatial resolution h small or a dynamical error small by taking higher-order terms into account may not be enough in practice, i.e., one has to be aware which error term dominates a CERES in the actual numerical implementation.

The paper is structured as follows: In Section 2 we provide the foundational technical setup for SPDEs (1) using mild solutions. Section 4 develops the error estimates for covariance matrices while Section 5 contains the relevant moment estimates. In Section 6 we transition to the Lyapunov equation and its reduction to steady state. The last technical step is carried out in Section 7 tracing the error results for low-rank iterative schemes for Lyapunov equations. In Section 8, we present the full CERES, which draws upon all the previous results. An outlook to open problems and further applications of our methodology is given in Section 9.

2 SPDE - Mild Solutions

Let $\mathcal{D} \subset \mathbb{R}^d$ be a bounded domain with smooth boundary $\partial\mathcal{D}$ and denote the spatial variable by $x \in \mathcal{D}$. Consider a fixed compact time interval $\mathcal{I} = [0, T]$, $t \in \mathcal{I}$ and a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Furthermore, we fix two Hilbert spaces \mathcal{H} and \mathcal{U} with inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{U}}$ respectively. Let $Q \in \mathcal{L}(\mathcal{U}, \mathcal{U})$ be a symmetric non-negative linear operator and let $W(t)$ denote the associated Q -Wiener process; see [43, Sec.4.1] for the definitions

when Q is of trace class and [43, Sec.4.3] for the case of a cylindrical Wiener process. Let $\mathcal{U}_0 := Q^{1/2}\mathcal{U}$ be a Hilbert space [43, 44] with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{U}_0} := \langle Q^{-1/2} \cdot, Q^{-1/2} \cdot \rangle_{\mathcal{U}}$, where $Q^{-1/2}$ is the pseudoinverse of Q . An operator $M : \mathcal{U}_0 \rightarrow \mathcal{H}$ is a Hilbert-Schmidt operator if the norm

$$\|M\|_{\mathcal{L}_2^0} := \left(\sum_{k=1}^{\infty} \|M\zeta_k\|_{\mathcal{H}}^2 \right)^{1/2}$$

is finite, where the choice of orthonormal basis $\{\zeta_k\}_{k=1}^{\infty}$ for \mathcal{U}_0 turns out to be arbitrary. The space of these Hilbert-Schmidt operators will be denoted accordingly by \mathcal{L}_2^0 . If we consider Hilbert-Schmidt operators on \mathcal{H} , then they will be denoted by $\mathcal{L}_2 = \mathcal{L}_2(\mathcal{H}, \mathcal{H})$.

Let $u(x, t; \omega)$ for $\omega \in \Omega$ denote the unknown family of random variables $u : \mathcal{D} \times \mathcal{I} \times \Omega \rightarrow \mathbb{R}$. In the notation we shall always suppress ω from now on and assume that all maps we define are measurable with respect to ω , which will also imply measurability for u below. We write $u(t) = u(\cdot, t) \in \mathcal{H}$ and study the SPDE

$$du(t) = [Au(t) + f(u(t))] dt + g(u(t)) dW(t), \quad u(0) = u_0, \quad (5)$$

as an evolution equation on the Hilbert space \mathcal{H} , which is taken as a suitable function space on \mathcal{D} .

(A0) We assume the operator $A : \text{dom}(A) \subset \mathcal{H} \rightarrow \mathcal{H}$ is linear, self-adjoint positive definite operator with compact inverse and generates an analytic semigroup $t \mapsto e^{tA}$ [22].

f, g are given maps to be specified more precisely below and the initial condition $u_0 \in \mathcal{H}$ is a random variable u_0 . A mild solution $u(t)$ to (5) satisfies

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A} f(u(s)) ds + \int_0^t e^{tA} g(u(s)) dW(s), \quad (6)$$

i.e., the integral equation (6) holds \mathbb{P} -almost surely (\mathbb{P} -a.s.) for $t \in \mathcal{I}$ and

$$\mathbb{P} \left(\int_0^t |u(s)|^2 ds < +\infty \right) = 1, \quad \mathbb{P}\text{-a.s.};$$

see also [43, Ch.4] or [10] for the construction of the stochastic integral. It is well-known that under certain Lipschitz assumptions [43, Sec.7.1] or dissipativity assumptions [43, Sec.7.4.2] on f, g , there exists a unique mild solution. However, since we are interested in numerical error estimates, it is important to have optimal regularity results for mild solutions so we follow [29, 30, 25]. Denote by $\{a_k\}_{k=1}^{\infty}$, $0 > a_1 \geq a_2 \geq \dots$ the eigenvalues of A and by e_k the associated eigenfunctions with $Ae_k = a_k e_k$. For $r \in \mathbb{R}$, define the fractional operator $A^{r/2} : \text{dom}(A^{r/2}) \rightarrow \mathcal{H}$ by

$$A^{r/2}v := -(-A)^{r/2}v = - \sum_{k=1}^{\infty} (-a_k)^{r/2} \langle v, e_k \rangle_{\mathcal{H}} e_k.$$

Set $\dot{\mathcal{H}}^r := \text{dom}(A^{r/2})$ and consider the norm $\|v\|_{\dot{\mathcal{H}}^r} := \|A^{r/2}v\|_{\mathcal{H}}$, which turns $\dot{\mathcal{H}}^r$ into a Hilbert space. Let $\mathcal{L}_{2,r}^0 \subset \mathcal{L}_2^0$ denote the subspace of Hilbert-Schmidt operators with the norm $\|\cdot\|_{\mathcal{L}_{2,r}^0} := \|A^{r/2} \cdot\|_{\mathcal{L}_2^0}$. The following assumptions are assumed to hold from now on (although we are still going to emphasize this several times in statements of theorems below):

(A1) Fix two constants $r \in [0, 1)$ and $p \in [2, \infty)$ for all assumptions. The initial condition $u_0 : \Omega \rightarrow \dot{\mathcal{H}}^{r+1}$ is \mathcal{F}_0 -measurable with

$$\mathbb{E} \left[\|u_0\|_{\dot{\mathcal{H}}^{r+1}}^p \right]^{1/p} \leq C_{\text{ini},r} < +\infty.$$

(A2) $f : \mathcal{H} \rightarrow \dot{\mathcal{H}}^{r-1}$ and there exists a constant $C_f > 0$ such that

$$\|f(u) - f(v)\|_{\dot{\mathcal{H}}^{r-1}} \leq C_f \|u - v\|_{\mathcal{H}}, \quad \text{for all } u, v \in \mathcal{H}. \quad (7)$$

(A3) $g : \mathcal{H} \rightarrow \mathcal{L}_2^0$ and there exists a constant $C_{g,1} > 0$ such that

$$\|g(u) - g(v)\|_{\mathcal{L}_2^0} \leq C_{g,1} \|u - v\|_{\mathcal{H}}, \quad \text{for all } u, v \in \mathcal{H}, \quad (8)$$

and furthermore $g(\dot{\mathcal{H}}^r) \subset \mathcal{L}_{2,r}^0$ holds with the estimate

$$\|g(u)\|_{\mathcal{L}_{2,r}^0} \leq C_{g,2} (1 + \|u\|_{\dot{\mathcal{H}}^r}), \quad \text{for all } u \in \mathcal{H}, \quad (9)$$

and some constant $C_{g,2} > 0$.

Essentially (A2)-(A3) are modifications/extensions of the classical Lipschitz assumptions in [43]. Therefore, one immediately gets:

Theorem 2.1 ([43, Thm.7.4],[24, Thm.1]). *Suppose (A0)-(A3) hold, then there exists a unique mild solution for the SPDE (5).*

In addition, one may provide (optimal) regularity estimates for the mild solution:

Theorem 2.2 ([30, Thm.3.1 & Thm.4.1]). *Suppose (A0)-(A3) hold then the unique mild solution is almost surely in $\dot{\mathcal{H}}^{r+1}$. There exists a constant $C_{\text{spa}} > 0$ such that*

$$\sup_{t \in \mathcal{I}} \left(\mathbb{E} \left[\|u(t)\|_{\dot{\mathcal{H}}^{r+1}}^p \right] \right)^{1/p} \leq C_{\text{ini},r} + C_{\text{spa}} \left(1 + \sup_{t \in \mathcal{I}} \left(\mathbb{E} \left[\|u(t)\|_{\dot{\mathcal{H}}^r}^p \right] \right)^{1/p} \right) \quad (10)$$

holds as a spatial regularity estimate and for temporal regularity we have

$$\sup_{t_1, t_2 \in \mathcal{I}, t_1 \neq t_2} \left(\mathbb{E} \left[\|u(t_1) - u(t_2)\|_{\dot{\mathcal{H}}^s}^p \right] \right)^{1/p} \leq C_{\text{time}} |t_1 - t_2|^{\min(\frac{1}{2}, \frac{1+r-s}{2})} \quad (11)$$

for some constant $C_{\text{time}} > 0$ and for every $s \in [0, r + 1)$.

As for the deterministic counterpart ($g \equiv 0$) of the SPDE (5), one may use regularity estimates to obtain convergence rates and error estimates for associated numerical schemes [50, 15]. In this paper, as already discussed in Section 1, we are interested in the spatial approximation error only. We stated the temporal regularity results only for completeness. It is noted that the constant $C_{\text{spa}} > 0$ in Theorem 2.2 does depend upon p, r, A, f, g and T .

As a typical example for the abstract framework, one should always keep in mind classical one-component reaction-diffusion systems, where

$$\mathcal{H} = L^2(\mathcal{D}) \quad \text{and} \quad A = \Delta := \sum_{k=1}^d \frac{\partial^2}{\partial x^2},$$

but we shall not restrict to this case in this paper and continue to work with the more general setup.

3 SPDE - Spatial Discretization

Let \mathcal{S}_h for $h \in (0, 1]$ denote a continuous family of finite-dimensional subspaces of the Hilbert space $\dot{\mathcal{H}}^1$ with

$$\dim(\mathcal{S}_h) =: N \in \mathbb{N}, \quad (12)$$

which are spanned by N basis elements of an orthonormal basis of $\dot{\mathcal{H}}^1$. The spaces \mathcal{S}_h are going to be used as the spatial discretization spaces; see [50, 15] for a detailed overview. Particularly important examples are the span of a finite set of basis functions of A leading to a spectral Galerkin method, or the span of piecewise polynomial functions on \mathcal{D} , where h is the diameter of the largest element of a suitable mesh on \mathcal{D} , leading to a Galerkin finite element method. Let $R_h : \dot{\mathcal{H}}^1 \rightarrow \mathcal{S}_h$ be the orthogonal projection onto \mathcal{S}_h with respect to the inner product $a(\cdot, \cdot) := \langle A^{1/2} \cdot, A^{1/2} \cdot \rangle_{\mathcal{H}}$ so that we have

$$a(R_h u, v_h) = a(u, v_h), \quad \text{for all } u \in \dot{\mathcal{H}}^1, v_h \in \mathcal{S}_h.$$

The discretized version A_h of A is defined by the requirement that for a given $u_h \in \mathcal{S}_h$, the image $A_h u_h$ is the unique element satisfying

$$a(u_h, v_h) = \langle A_h u_h, v_h \rangle_{\mathcal{H}}, \quad \text{for all } v_h \in \mathcal{S}_h.$$

Furthermore, let $P_h : \dot{\mathcal{H}}^{-1} \rightarrow \mathcal{S}_h$ be the orthogonal projection onto \mathcal{S}_h defined analogously to R_h , i.e., by requiring

$$\langle P_h u, v_h \rangle_{\mathcal{H}} = a(A^{-1} u, v_h), \quad \text{for all } u \in \dot{\mathcal{H}}^{-1}, v_h \in \mathcal{S}_h.$$

Then the spatial discretization of the SPDE can be written as

$$du_h(t) = [A_h u_h(t) + P_h f(u_h(t))] dt + P_h g(u_h(t)) dW(t), \quad u_h(0) = P_h u_0. \quad (13)$$

It is relatively straightforward to check that (13) must also have a unique mild solution; see also the proof of Lemma 3.2. To fix the role of the discretization parameter h , we are going to assume:

(A4) There exists a constant $C_h > 0$ such that

$$\|R_h v - v\|_{\mathcal{H}} \leq C_h h^s \|v\|_{\dot{\mathcal{H}}^s}, \quad \text{for all } v \in \dot{\mathcal{H}}^s, s \in \{1, 2\}, h \in (0, 1]. \quad (14)$$

Recently, the regularity result of Theorem 2.2 has been transferred to yield results on strong/pathwise approximation properties of the finite-dimensional SODE (13) for the infinite-dimensional SPDE (5). Consider the norm

$$\|\cdot\|_{L^p(\Omega; \mathcal{H})} := (\mathbb{E}[\|\cdot\|_{\mathcal{H}}^p])^{1/p} \quad (15)$$

for $p \in [2, +\infty)$ as above. Then one can prove the following error estimate:

Theorem 3.1 ([29, Thm.1.1]). *Suppose (A0)-(A4) hold, then there exists a constant $C_{\text{Gal}} > 0$ such that*

$$\|u(t) - u_h(t)\|_{L^p(\Omega; \mathcal{H})} \leq C_{\text{Gal}} h^{1+r}, \quad \forall t \in \mathcal{I}. \quad (16)$$

In the later development of the error estimates for the covariance operator, we shall need another auxiliary result, which we prove here:

Lemma 3.2. *Suppose (A0)-(A4) hold, then here exists a constant $C_+ > 0$ (independent of h) such that*

$$\sup_{t \in \mathcal{I}} \|u(t) + u_h(t)\|_{L^2(\Omega; \mathcal{H})} \leq C_+.$$

Proof. Indeed, we just have

$$\sup_{t \in \mathcal{I}} \|u(t) + u_h(t)\|_{L^2(\Omega; \mathcal{H})} \leq \sup_{t \in \mathcal{I}} \|u(t)\|_{L^2(\Omega; \mathcal{H})} + \sup_{t \in \mathcal{I}} \|u_h(t)\|_{L^2(\Omega; \mathcal{H})}$$

so the first term is bounded by a direct application of Theorem 2.2 (or in fact, the classical result [43, Thm.7.4(ii)]). For the discretization, note that we may apply the same results since (A0)-(A3) also hold for the discretized SPDE (13). For example, consider (A1) then we have

$$\|P_h f(u) - P_h f(v)\|_{\dot{H}^{r-1}} = \|P_h[f(u) - f(v)]\|_{\dot{H}^{r-1}} \leq \|f(u) - f(v)\|_{\dot{H}^{r-1}} \quad (17)$$

since P_h is an orthogonal projector. The other assumptions are checked similarly. \square

Of course, it should be noted that the constant $C_+ > 0$ will depend on the data of the problem, i.e., on f, g, T, A . However, we will only be interested in the convergence rate in h and we will show later on that we can select T as an order one constant anyhow, while f, g, A are given and satisfy the Lipschitz and semigroup assumptions stated above.

4 SPDEs - Covariance

The next, and first novel, step is to establish numerical error estimates for the covariance. Let $v \in L^2(\Omega; \mathcal{H})$, then one defines the covariance operator [43, 35] of v as

$$\text{Cov}(v) := \mathbb{E}[(v - \mathbb{E}[v]) \otimes (v - \mathbb{E}[v])]. \quad (18)$$

By definition, $\text{Cov}(v) : \mathcal{H} \rightarrow \mathcal{H}$ is a symmetric linear operator. In addition one may check that $\text{Cov}(v)$ is nuclear using the equivalent characterization of nuclear operators M on Hilbert spaces via the condition

$$\text{Tr}(M) := \sum_{k=1}^{\infty} \langle M \xi_k, \xi_k \rangle_{\mathcal{H}} < +\infty$$

for an orthonormal basis $\{\xi_k\}_{k=1}^{\infty}$ of \mathcal{H} . The space of nuclear operators $\mathcal{L}_1(\mathcal{H}, \mathcal{H})$ becomes a Banach space under the norm $\|\cdot\|_{\mathcal{L}_1(\mathcal{H}, \mathcal{H})} := \text{Tr}(\cdot)$ [43, Appendix C]. Note that we have two well-defined covariance operators

$$\text{Cov}(u(t)) \quad \text{and} \quad \text{Cov}(u_h(t)) \quad (19)$$

as the SPDE (5) and the spatially discretized version (13) both have mild solutions in $L^2(\Omega \times \mathcal{I}; \mathcal{H})$ by (11).

Theorem 4.1. *There exists a constant $C_d > 0$ such that*

$$\sup_{t \in \mathcal{I}} \|\text{Cov}(u(t)) - \text{Cov}(u_h(t))\|_{\mathcal{L}_2(\mathcal{H}, \mathcal{H})} \leq C_d h^{1+r}. \quad (20)$$

Proof. The proof is a calculation aiming to use the spatial approximation property of Theorem 3.1. Suppose as a first case that $\mathbb{E}[u] = 0 = \mathbb{E}[u_h]$. Consider an orthonormal basis $\{\xi_k\}_{k=1}^\infty$ of \mathcal{H} . The idea is to estimate the error in the Hilbert-Schmidt norm $\|\cdot\|_{\mathcal{L}_2}$. For the following steps we suppress the argument $u = u(t)$ and $u_h = u_h(t)$ and use the definition of the covariance operator

$$\begin{aligned} \sup_{t \in \mathcal{I}} \|\text{Cov}(u) - \text{Cov}(u_h)\|_{\mathcal{L}_2}^2 &= \sup_{t \in \mathcal{I}} \|\mathbb{E}[u \otimes u] - \mathbb{E}[u_h \otimes u_h]\|_{\mathcal{L}_2}^2, \\ &= \sup_{t \in \mathcal{I}} \|\mathbb{E}[(u - u_h) \otimes (u + u_h)]\|_{\mathcal{L}_2}^2 \\ &= \sup_{t \in \mathcal{I}} \sum_{k=1}^{\infty} \langle \mathbb{E}[(u - u_h) \otimes (u + u_h)] \xi_k, \xi_k \rangle_{\mathcal{H}}^2, \\ &= \sup_{t \in \mathcal{I}} \sum_{k=1}^{\infty} \langle \mathbb{E}[(u - u_h) \otimes (u + u_h)] \xi_k, \xi_k \rangle_{\mathcal{H}}^2, \\ &= \sup_{t \in \mathcal{I}} \sum_{k=1}^{\infty} \mathbb{E}[\langle u - u_h, \xi_k \rangle_{\mathcal{H}} \langle u + u_h, \xi_k \rangle_{\mathcal{H}}]^2, \end{aligned}$$

where we use in the last step the definition of the tensor product $(v_1 \otimes v_2)v_3 := \langle v_2, v_3 \rangle_{\mathcal{H}} v_1$ for $v_k \in \mathcal{H}$ for $k \in \{1, 2, 3\}$. Then a direct application of Cauchy-Schwarz and Parseval's Theorem yield

$$\begin{aligned} \sup_{t \in \mathcal{I}} \|\text{Cov}(u) - \text{Cov}(u_h)\|_{\mathcal{L}_2}^2 &\leq \sup_{t \in \mathcal{I}} \sum_{k=1}^{\infty} \mathbb{E}[\langle u - u_h, \xi_k \rangle_{\mathcal{H}}^2] \mathbb{E}[\langle u + u_h, \xi_k \rangle_{\mathcal{H}}^2], \\ &\leq \sup_{t \in \mathcal{I}} \|u - u_h\|_{L^2(\Omega; \mathcal{H})}^2 \|u + u_h\|_{L^2(\Omega; \mathcal{H})}^2. \end{aligned} \quad (21)$$

In the expression (21), we may estimate the two terms separately by estimating the supremum by the product of suprema. Furthermore, a direct application of Theorem 3.1 to the first term and Lemma 3.2 to the second term give

$$\sup_{t \in \mathcal{I}} \|\text{Cov}(u) - \text{Cov}(u_h)\|_{\mathcal{L}_2}^2 \leq C_{\text{Gal}} C_+ h^{2(r+1)},$$

which yields the result in the basic case of zero means. If $\mathbb{E}[u] \neq 0$ and $\mathbb{E}[u_h] \neq 0$, we observe that

$$\begin{aligned} \sup_{t \in \mathcal{I}} \|u - u_h - \mathbb{E}[u - u_h]\|_{L^2(\Omega; \mathcal{H})}^2 &\leq \sup_{t \in \mathcal{I}} \|u - u_h\|_{L^2(\Omega; \mathcal{H})}^2 + \sup_{t \in \mathcal{I}} \|\mathbb{E}[u - u_h]\|_{L^2(\Omega; \mathcal{H})}^2, \\ &\leq C_{\text{Gal}} h^{2(r+1)} + \sup_{t \in \mathcal{I}} \mathbb{E}[\|u - u_h\|_{L^2(\Omega; \mathcal{H})}]^2, \\ &\leq 2C_{\text{Gal}} h^{2(r+1)}, \end{aligned}$$

using again Theorem 3.1 twice. Furthermore, it is easy to see that

$$\mathbb{E}[\|u + u_h - \mathbb{E}[u + u_h]\|_{L^2(\Omega; \mathcal{H})}]$$

remains bounded for nonzero means. The result now follows repeating the same steps shown for the zero means case also for the nonzero means case. \square

Obviously the constant C_d also depends upon the data of the problem as do C_+ and C_{Gal} but all constants are independent of h , which is the key discretization parameter in the step (S1).

5 SODEs - Linearization

Having estimated the error of the discretization, we are now dealing with a finite-dimensional system of stochastic ordinary differential equations (SODEs) given by

$$du_h(t) = [A_h u_h(t) + P_h f(u_h(t))] dt + P_h g(u_h(t)) dW(t), \quad u_h(0) = P_h u_0, \quad (22)$$

where $u_h \in \mathbb{R}^N$. We assume that the original SPDE (5) and its projection have a locally asymptotically stable homogeneous steady state for zero noise and that we only study the small noise regime.

(A5) Assume u^* satisfies $Au^* + f(u^*) = 0$. Furthermore, suppose f is Fréchet differentiable, $D_u f(u^*) = f'(u^*)\text{Id}$ and

$$\text{spec}(A + f'(u^*)\text{Id}) \subset \{\rho \in \mathbb{R} : \text{Re}(\rho) < 0\}. \quad (23)$$

Furthermore, assume $u_h^* := P_h u^*$ satisfies $A_h u_h^* + P_h f(u_h^*) = 0$ for all $h \in (0, 1]$.

(A6) Suppose there exists a constant $v > 0$ such that

$$\|P_h g(u_h(t))\|_2 \leq v \quad (24)$$

for all $h \in (0, 1]$.

Below we are also going to assume that v is chosen sufficiently small to get a good approximation of the linearized system. The goal is to provide a finite-time estimate for the difference between the covariance matrix $\text{Cov}(u_h(t))$ of (22) and covariance matrix $\text{Cov}(U_h(t))$ of the linearized OU process

$$dU_h(t) = [A_h + P_h [D_u f](u^*)] U_h(t) dt + P_h g(u^*) dW(t), \quad U_h(0) = P_h u_0. \quad (25)$$

Obviously, for general nonlinearities f , there is no infinite-time asymptotic estimate. Furthermore, without additional assumptions on the nonlinearity f , it is usually not possible to estimate the error between a linearized and a nonlinear system. To simplify the notation we let

$$u_h =: z, \quad A_h u_h(t) + P_h f(u_h(t)) =: F(z), \quad P_h g(u_h(t)) =: G(z),$$

as well as

$$U_h =: Z, \quad [A_h + P_h[D_{u^*}f](u^*)] =: \mathcal{A}, \quad P_h g(u^*) =: B.$$

Hence, we have to compare the SODE

$$dz = F(z) dt + G(z) dW^R, \quad z_0 = z(0), \quad (26)$$

near a steady state $z^* := u_h^*$ to the SODE

$$dZ = \mathcal{A}Z dt + B dW^R, \quad Z_0 = Z(0), \quad (27)$$

where $W^R = (W_1(t), \dots, W_R(t))^\top \in \mathbb{R}^R$ is a vector of independent Brownian motions. Without loss of generality we may assume that $z^* \equiv (0, \dots, 0)^\top$ since we can always translate the steady if necessary. Let $\mathbf{p} = (p_1, p_2, \dots, p_N) \in (\mathbb{N}_0)^N$ be a multi-index, define the mean values of z by $\mu := \mathbb{E}[z] \in \mathbb{R}^N$ and the centered moments as

$$\mathbb{E}[(z - \mu)^{\mathbf{p}}] := \mathbb{E}[(z_1 - \mu_1)^{p_1} (z_2 - \mu_2)^{p_2} \cdots (z_N - \mu_N)^{p_N}]. \quad (28)$$

To make the notation more compact, we also introduce 'altered' multi-indices as follows:

$$\mathbf{p}(k : \zeta) = (p_1, p_2, \dots, p_{k-1}, p_k + \zeta, p_{k+1}, \dots, p_N)$$

for $\zeta \in \mathbb{Z}$ and multiple arguments pertain to changes in the respective components. Furthermore, we consider the diffusion operators

$$\mathcal{G}(z) := G(z)G(z)^\top, \quad \mathcal{B} := BB^\top. \quad (29)$$

Lemma 5.1. *The evolution equations for the centered moments $\mathbb{E}[(z - \mu)^{\mathbf{p}}]$ of (26) are given by*

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[(z - \mu)^{\mathbf{p}}] &= \sum_{k=1}^N p_k \mathbb{E}[F_k(z)(z - \mu)^{\mathbf{p}(k:-1)}] + \frac{1}{2} \sum_{k=1}^N p_k(p_k - 1) \mathbb{E}[\mathcal{G}_{kk}(z)(z - \mu)^{\mathbf{p}(k:-2)}] \\ &\quad + \sum_{l=2}^N \sum_{k=1}^{l-1} p_l p_k \mathbb{E}[\mathcal{G}_{kl}(z)(z - \mu)^{\mathbf{p}(k:-1, l:-1)}] \end{aligned}$$

Proof. The calculation of the SODEs for the moments follows from first applying Itô's formula to monomials of the form $z^{\mathbf{p}}$ as shown in [49, Sec. 4.1]. However, we then need to average via $\mathbb{E}[\cdot]$, and terms of the form $\int(\cdot) dW^R$ average to zero as they satisfy the martingale property [26] by the Lipschitz assumption on G . \square

Consider the linear approximation of (26) given by (27). We use the notation $\nu := \mathbb{E}[Z]$ and we may assume without loss of generality that \mathcal{A} is already diagonal; indeed, by assumptions (A0) and (A5) we already have that \mathcal{A} is symmetric with real spectrum so we can apply a coordinate change to make \mathcal{A} diagonal. Now we have two processes and we would like to compare $\text{Cov}(z)$ with $\text{Cov}(Z)$. From Lemma 5.1 it follows that

$$\frac{d}{dt} \mu = \mathbb{E}[F(z)], \quad \frac{d}{dt} \nu = \mathcal{A}\nu. \quad (30)$$

It is relatively easy to write down the formal evolution equations for the covariances. For the diagonal entries we have

$$\frac{d}{dt}\text{Cov}(z)_{ii} = 2\mathbb{E}[F_i(z)(z_i - \mu_i)] + \mathbb{E}[\mathcal{G}_{ii}(z)] \quad (31)$$

$$\frac{d}{dt}\text{Cov}(Z)_{ii} = 2\mathbb{E}[(\mathcal{A}Z)_i(Z_i - \nu_i)] + \mathcal{B}_{ii} \quad (32)$$

Consider the difference $\text{Cov}(z) - \text{Cov}(Z) =: \text{Cov}^\Delta$ and also defined the remainders

$$R^F(z) := F(z) - \mathcal{A}z, \quad R^G(z) := G(z) - \mathcal{B}.$$

Observe that the remainder R^F is Lipschitz

$$\|R^F(z_1) - R^F(z_2)\| \leq (C_F + \|\mathcal{A}\|)\|z_1 - z_2\|, \quad (33)$$

by assumptions (A2) and (A5) for some constant $C_F > 0$, where $\|\cdot\|$ always denotes the 2-norm. Regarding the remainder R^G , recall that we assumed a uniform noise bound in (A6) so that

$$\|R^G(z_1) - R^G(z_2)\| \leq C_G v^2 \quad (34)$$

for some constant $C_G > 0$. Then one finds, using that \mathcal{A} is diagonal and via assumption (A5), that

$$\begin{aligned} \frac{d}{dt}\text{Cov}_{ii}^\Delta &= 2\mathbb{E}[F_i(z)(z_i - \mu_i) - (\mathcal{A}Z)_i(Z_i - \nu_i)] + \mathbb{E}[G_{ii}(z) - \mathcal{B}_{ii}], \\ &= 2\mathcal{A}_{ii}\mathbb{E}[(z_i + \mu_i - \mu_i - R_i^F(z)/\mathcal{A}_{ii})(z_i - \mu_i) - (Z_i + \nu_i - \nu_i)(Z_i - \nu_i)] + \mathbb{E}[R_{ii}^G(z)], \\ &= 2\mathcal{A}_{ii}(\text{Cov}_{ii}^\Delta + \mathbb{E}[(\mu_i - R_i^F(z)/\mathcal{A}_{ii})(z_i - \mu_i) - \nu_i(Z_i - \nu_i)]) + \mathbb{E}[R_{ii}^G(z)], \\ &= 2\mathcal{A}_{ii}(\text{Cov}_{ii}^\Delta + \mathbb{E}[R_i^F(z)/\mathcal{A}_{ii}(\mu_i - z_i)]) + \mathbb{E}[R_{ii}^G(z)]. \end{aligned}$$

Using the Lipschitz conditions (33) and the bound (34) one has

$$\begin{aligned} |\mathbb{E}[R_i^F(z)/\mathcal{A}_{ii}(\mu_i - z_i) + R_{ii}^G(z)]| &\leq \frac{(\|\mathcal{A}\| + C_F)|\mu_i|}{|\mathcal{A}_{ii}|}\mathbb{E}[\|z\|] \\ &\quad + \frac{(\|\mathcal{A}\| + C_F)}{|\mathcal{A}_{ii}|}\mathbb{E}[\|z\|^2] + C_G v^2 \end{aligned} \quad (35)$$

we denote the right-hand side of the last inequality by $\eta_{ii}(t)$, which implies the final estimate

$$\text{Cov}_{ii}^\Delta(t) \leq \text{Cov}_{ii}^\Delta(0) + \underbrace{\int_0^t \eta_{ii}(s) ds}_{=: \eta^*(t)} + 2\mathcal{A}_{ii} \int_0^t \text{Cov}_{ii}^\Delta(s) ds, \quad (36)$$

where we note that $\eta^*(t)$ is non-decreasing. Applying Gronwall's inequality to (36) yields the following result:

Lemma 5.2. *Suppose the assumptions (A0)-(A6) hold for all $t \in \mathcal{I}$, then*

$$\text{Cov}_{ii}^\Delta(t) \leq [\text{Cov}_{ii}^\Delta(0) + \eta^*(t)] e^{2\mathcal{A}_{ii}t}. \quad (37)$$

holds for all $t \in \mathcal{I}$.

Note that an estimate of the form (37) is fully expected to hold since it states that the growth of the difference between the covariances in the case of Lipschitz F and sufficiently bounded noise is controlled by the first- and second-moments of the nonlinear process. In particular, if we are close to a linear SODE or the spectral gap is very large, then we have an excellent finite-time approximation on \mathcal{I} , while large noise and a strong nonlinearity make the approximation worse. The next step is to look at the off-diagonal terms and consider the case $i > j$ (the case $i < j$ is similar). The evolution equations are

$$\begin{aligned}\frac{d}{dt}\text{Cov}(z)_{ij} &= \mathbb{E}[F_i(z)(z_j - \mu_j)] + \mathbb{E}[F_j(z)(z_i - \mu_i)] + \mathbb{E}[G_{ij}(z)] \\ \frac{d}{dt}\text{Cov}(Z)_{ij} &= \mathbb{E}[(\mathcal{A}Z)_i(Z_j - \nu_j)] + \mathbb{E}[(\mathcal{A}Z)_j(Z_i - \nu_i)] + \mathcal{B}_{ij}\end{aligned}$$

A similar calculation as above leads one to define

$$\begin{aligned}\eta_{ij}(t) &= \left(\frac{(\|\mathcal{A}\| + C_F)|\mu_i|}{|\mathcal{A}_{ii}|} + \frac{(\|\mathcal{A}\| + C_F)|\mu_j|}{|\mathcal{A}_{jj}|} C_G v^2 \right) \mathbb{E}[\|z\|] \\ &\quad + \left[\frac{(\|\mathcal{A}\| + C_F)}{|\mathcal{A}_{ii}|} + \frac{(\|\mathcal{A}\| + C_F)}{|\mathcal{A}_{jj}|} \right] \mathbb{E}[\|z\|^2].\end{aligned}$$

As before, we use the notation $\eta_{ij}^*(t) := \int_0^t \eta_{ij}(s) ds$ and use Gronwall's inequality to obtain the result:

Lemma 5.3. *Suppose the assumptions (A0)-(A6) hold for all $t \in \mathcal{I}$, then*

$$\text{Cov}_{ij}^\Delta(t) \leq [\text{Cov}_{ij}^\Delta(0) + \eta_{ij}^*(t)] e^{(\mathcal{A}_{ii} + \mathcal{A}_{jj})t}. \quad (38)$$

holds for all $t \in \mathcal{I}$.

Of course, the estimates (37)-(38) may blow-up as $t \rightarrow +\infty$ if the stationary distribution cannot be approximated well by an OU process. Indeed, if the noise is small, this is exactly the effect of large deviations [53, 43], which are going to occur on an asymptotic time scale $\mathcal{O}(e^{c/v^2})$ as $v \rightarrow 0$. However, the estimate is rather explicit, i.e., if we know the Lipschitz constant, the spectral gap, the noise level and/or have some a-priori knowledge of the norms $\|z\|$ and/or $\|z\|^2$, then we can select $\text{Cov}_{ij}^\Delta(0) = 0$ and $\mathcal{A}_{ij} < 0$ is going to give initially decay so that the linear approximation (27) will be a good approximation for a certain initial time-scale. The worst-case bound is the following:

Theorem 5.4. *Suppose the assumptions (A0)-(A6) hold for all $t \in \mathcal{I}$, then there exists a constant $C_1 > 0$ and a function $\eta^*(t) := \max_{ij} \eta_{ij}(t)$ such that*

$$\|\text{Cov}(z(t)) - \text{Cov}(Z(t))\|_2 \leq C_1 [\|\text{Cov}(z_0) - \text{Cov}(Z_0)\|_2 + \eta^*(t)] e^{-t \min_i |\mathcal{A}_{ii}|}. \quad (39)$$

Proof. Using Lemma 5.2 and Lemma 5.3 we have

$$\begin{aligned}\|\text{Cov}(z(t)) - \text{Cov}(Z(t))\|_2 &\leq C \max_{ij} [(\text{Cov}(z_0) - \text{Cov}(Z_0))_{ij} + \eta_{ij}^*(t)] e^{(\mathcal{A}_{ii} + \mathcal{A}_{jj})t} \\ &\leq C [\|\text{Cov}(z_0) - \text{Cov}(Z_0)\|_2 + \eta^*(t)] e^{-t \min_i |\mathcal{A}_{ii}|}\end{aligned}$$

so that the result follows. \square

In summary, Theorem 5.4 just states that one has to be extremely careful to trust a local linear approximation in a numerical context for problems with large noise and/or a very strong nonlinearity, which both lead to a very quick pathwise sampling of a non-Gaussian stationary distribution. However, for small noise, sub-exponential time scales and/or a weak nonlinearity, the local approximation via linearization and covariance operators for of an OU process is going to correctly expose the relevant directions of fluctuations.

6 Lyapunov Equation - Algebraic Reduction

Recall that we have now considered a spatial discretization of the initial SPDE and we have localized the problem near a locally deterministically stable steady state via linearization. We calculated an upper bound on the discretization error and on the linearization error on a finite time scale. However, although we now can work with the linear SODE problem

$$dU = \mathcal{A}U dt + B dW^R, \quad U_0 = U(0), \quad (40)$$

we are still surprisingly *far from a practical computable problem for many applications!* It is well-known [8, 31] that $V(t) := \text{Cov}(U(t))$ satisfies the matrix ordinary differential equation (ODE) given by

$$\frac{d}{dt}V = \mathcal{A}V + V\mathcal{A}^\top + BB^\top =: L_{\mathcal{A}}V + \mathcal{B}. \quad (41)$$

The stationary problem is given by

$$0 = \mathcal{A}V + V\mathcal{A}^\top + \mathcal{B}. \quad (42)$$

It is well-known [1], that under the assumption (A5), there exists a unique stationary solution V_* to (42), which is stable for the time-dependent problem (41). However, we work on a finite-time interval \mathcal{I} so we need a convergence rate.

Theorem 6.1. *Suppose (A5) holds so that \mathcal{A} also has a spectral gap then there exists a constant $C_\tau > 0$ such that*

$$\|V(t) - V_*\|_2 \leq C_\tau (\|V(0) - V_*\|_2) e^{-t \min_i (|\text{Re}(\lambda_i)|)} \quad (43)$$

or, alternatively

$$\|V(t) - V_*\|_2 \leq C_\tau(H) (\|V(0) - V_*\|_2) e^{-2t/\|H\|_2}, \quad (44)$$

where H solves $\mathcal{A}H + H\mathcal{A}^\top + 2\text{Id} = 0$.

Proof. For (43), one first uses that the eigenvalues of the linear operator $L_{\mathcal{A}}$ are given by $\lambda_i + \lambda_j$ where $\{\lambda_k\}_{k=1}^N$ are the eigenvalues of \mathcal{A} [5]. For (44), we use that V_* is a steady state of (41) to obtain

$$\begin{aligned} V(t) &= V_* + \exp(t\mathcal{A})(V(0) - V_*)\exp(t\mathcal{A}^\top) \\ \Rightarrow \|V(t) - V_*\|_2 &\leq \|\exp(\mathcal{A}t)\|_2^2 \|V(0) - V_*\|_2 \end{aligned}$$

Then by [23, Theorem 3.1]

$$\|\exp(t\mathcal{A})\|_2^2 \leq C_\tau(H) \exp\left(\frac{-2t}{\|H\|_2}\right),$$

which leads to (44). □

Basically, Theorem 6.1 gives an estimate, which allows us to reduce the computation to the Lyapunov equation (42) to the stationary case as long we do not start far away from the locally linearized approximate solution. Unfortunately, direct solution methods, such as the Bartels-Stewart algorithm [4] are unlikely to work as the space dimension N grows drastically in practice as we decrease h . Furthermore, N increases due to the curse of dimensionality as d increases. Direct solution methods come with an complexity $\mathcal{O}(N^3)$ and with storage requirements of $\mathcal{O}(N^2)$, which limits their applicability to problems of moderate sizes. In the following we, therefore, propose to use special methods for large-scale Lyapunov equation that work with complexities and storage requirements linear in N .

7 Lyapunov Equation - Low-Rank and Computation

In this section, we want to consider and numerically approximate a low-rank approximation of V_* . This involves two steps, which can be accomplished by carefully tracing the literature: (1) understanding error estimates for the low-rank approximation and (2) finding an error estimate for the computation in a low-rank iterative algorithm for solving Lyapunov equations.

7.1 Low-rank Approximations and Singular Value Decay

Consider the singular value decomposition of V_*

$$\begin{aligned} V_* &= Y \Sigma X^\top, \quad Y^\top Y = X^\top X = \text{Id}, \\ \Sigma &= \text{diag}(\sigma_1(V_*), \dots, \sigma_N(V_*)), \\ \sigma_1(V_*) &\geq \dots \geq \sigma_\ell(V_*) > \sigma_{\ell+1}(V_*) = \sigma_n(V_*) = 0, \end{aligned}$$

where $\ell = \text{rank}(V_*)$. The best approximation of V_* of rank \mathfrak{r} is, by the Eckhart-Young theorem [19, Theorem 2.4.8.], obtained by

$$V_* \approx V_*^{\text{lr}, \mathfrak{r}} := \sum_{i=1}^{\mathfrak{r}} \sigma_i u_i x_i^\top = Y_{\mathfrak{r}} \Sigma_{\mathfrak{r}} X_{\mathfrak{r}}, \quad (45)$$

where $Y_{\mathfrak{r}}$, $X_{\mathfrak{r}}$ contain the first \mathfrak{r} columns (left and right singular vectors) of Y , R , and $\Sigma_{\mathfrak{r}}$ the \mathfrak{r} largest singular values. Note that since $V_* = V_*^\top$, $X_{\mathfrak{r}} = Y_{\mathfrak{r}}$ can be chosen. The approximation error is given by

$$\|V_* - V_*^{\text{lr}, \mathfrak{r}}\|_2 \leq \sigma_{\mathfrak{r}+1}. \quad (46)$$

If the singular values of V_* decay rapidly towards zero, a small error can be achieved with small values of \mathfrak{r} . In fact, solutions of large-scale Lyapunov equations with low-rank inhomogeneities often show a fast singular value decay. This phenomenon has been theoretically investigated, e.g., in [42, 20, 2, 47]. By assumption (A0) we restrict to the symmetric case $\mathcal{A} = \mathcal{A}^\top$. The following basic estimate on the singular value decay can be found in [42]

$$\sigma_{R\mathfrak{r}+1}(V_*) \leq \sigma_1(V_*) \left(\prod_{i=0}^{p-1} \frac{\kappa(\mathcal{A})^{(2i+1)/(2p)} - 1}{\kappa(\mathcal{A})^{(2i+1)/(2p)} + 1} \right)^2, \quad 1 \leq R\mathfrak{r} \leq N, \quad (47)$$

where $\kappa(\mathcal{A}) = \|\mathcal{A}\|_2 \|\mathcal{A}^{-1}\|_2$. A more precise bound is developed in [47] using $\text{spec}(\mathcal{A}) \subset [a, b]$ and quantities related to elliptic functions and integrals, which we briefly recall in the following definition.

Definition 7.1. *The elliptic integral of the first kind defined on $[0, 1]$ with respect to the modulus $0 < k < 1$ is*

$$s_k(x) := \int_0^x \frac{1}{\sqrt{(1-t^2)(1-k^2t^2)}} dt.$$

Define also the elliptic functions sn, dn by

$$\text{sn}(s_k) = x(s_k), \quad \text{dn}(s_k) = \sqrt{1 - k^2 \text{sn}^2(s_k)}.$$

The associated complete elliptic integral (w.r.t. the modulus k) is the value $K := s_k(1)$. The complementary modulus is $k' = \sqrt{1 - k^2}$ and the associated complementary complete elliptic integral by $K' = s_{k'}(1)$. The nome q is defined by $q := \exp(-\pi K'/K)$.

Lemma 7.2. *The nome q and complementary modulus k' also satisfy the identity [36]*

$$\sqrt{k'} = \frac{1 - 2q + 2q^4 - 2q^9 + \dots}{1 + 2q + 2q^4 + 2q^9 + \dots}. \quad (48)$$

Using these quantities, the singular values of V_* can be bounded by the following result from [47, 14].

Theorem 7.3. ([47, Theorem 2.1.1.], [14]) *Let $\mathcal{A} = \mathcal{A}^\top$ and $\text{spec}(\mathcal{A}) \in [a, b]$ with $a := \min \lambda_i < b := \max \lambda_i < 0$. Set the complementary modulus to $k' = b/a = 1/\kappa(\mathcal{A})$ and set k , the complete elliptic integrals K' and K , as well as the nome q via the relations in Definition 7.1. Then it holds for the singular values of the solution of (42)*

$$\sigma_{Rr+1}(V_*) \leq \sigma_1(V_*) \left(\frac{1 - \sqrt{k'_t}}{1 + \sqrt{k'_t}} \right)^2, \quad 1 \leq Rr \leq N, \quad (49)$$

where k'_t relates to q^r in the same way k' is build from q via (48):

$$\sqrt{k'_t} = \frac{1 - 2q^r + 2q^{4r} - 2q^{9r} + \dots}{1 + 2q^r + 2q^{4r} + 2q^{9r} + \dots}.$$

We stress that, while better than (47), the bound (49) might not be very sharp in practice, where one often observes an even faster singular value decay. One reason is that Theorem 7.3 only uses $\kappa(\mathcal{A})$, i.e., the extremal eigenvalue a, b of \mathcal{A} . More realistic, but also more difficult to compute, bounds can be obtained by using more than these two eigenvalue of \mathcal{A} [47].

Remark: *In case of non-symmetric \mathcal{A} , the above bounds are not applicable. In this case much less is known about the singular value decay of V_* , see, e.g., the discussion in [3].*

Judging by (49), the decay of the singular values depends mainly on $\kappa(\mathcal{A})$ and the value $R = \text{coldim}(B)$. For instance for fixed a , the closer b to the origin, i.e. the smaller the spectral

gap, the larger $\kappa(\mathcal{A})$ and, consequently, the closer k, k' will be to one and, respectively, zero. Hence, K' and K will tend towards $\frac{\pi}{2}$ and ∞ , respectively, leading to q close to one. In this extreme situation, q^τ will also be close to one, leading, in the end, to

$$\frac{1 - \sqrt{k'_\tau}}{1 + \sqrt{k'_\tau}} \approx 1$$

such that no singular value decay might be observed. In particular, the main insight is that the spectral gap also plays a *numerical analysis* role at the error in step (S4). Furthermore, the column dimension R of B plays a significant role in the bound (49). Recall the B is the result from evaluating g at u^* and approximating the stochastic process $W(t)$. Suppose we truncate the Q -Wiener process [43] to obtain a numerical approximation [39]

$$W(t) \approx \sum_{i=1}^R \sqrt{\lambda_{Q,i}} \beta_i(t) \zeta_i$$

for a basis $\{\zeta_i\}_{i=1}^\infty$ of \mathcal{H} , eigenvalues $\lambda_{Q,i}$ of Q and independent Brownian motions $\beta_i(t)$ is used. The scalars $\lambda_{Q,i}$ form a non-increasing sequence. Hence, the slower the $\lambda_{Q,i}$ decrease, the higher the value of R should be chosen and, consequently, the slower the decay of the singular values of V_* . This implies the key practical guideline that space-time white noise is more difficult to treat numerically than a Q -trace-class Wiener process.

7.2 Error Bounds for Numerically Computed Low-Rank Solutions

Using the low-rank approximation (45) is not practical as it requires to first obtain V_* and then compute its singular valued decomposition. Numerical methods for large-scale Lyapunov equations [48, 7] typically directly compute low-rank factors $\mathcal{Z} \in \mathbb{R}^{N \times \tau}$, $\tau \ll N$ that form a low-rank approximate solution $V_*^{\text{lr}, \tau} = \mathcal{Z}\mathcal{Z}^\top$ which will, however, not be optimal in the sense of the SVD based approximation (45) but close if the method is properly executed. The advantage of these methods is that they are able to provide accurate low-rank solutions in a very efficient manner by utilizing tools from large-scale numerical linear algebra. By exploiting, e.g., the sparsity of \mathcal{A} and the low-rank R of the inhomogeneity, state-of-the-art methods are able to compute low-rank solution factors at complexities and memory requirements of $\mathcal{O}(N)$.

One iterative method for solving (42) is based on the fact that for any $\alpha \in \mathbb{C}_-$, (42) is equivalent to the matrix equation

$$X = \mathcal{C}(\alpha)X\mathcal{C}(\alpha) + \mathcal{B}(\alpha)\mathcal{B}(\alpha)^H$$

$$\text{with } \mathcal{C}(\alpha) := (\mathcal{A} + \alpha\text{Id})^{-1}(\mathcal{A} - \bar{\alpha}\text{Id}), \quad \mathcal{B}(\alpha) := \sqrt{-2\text{Re}(\alpha)}(\mathcal{A} + \alpha\text{Id})^{-1}B,$$

where $(\cdot)^H$ is the Hermitian conjugate. This motivates the self-evident iteration scheme

$$X_j = \mathcal{C}(\alpha_j)X_{j-1}\mathcal{C}(\alpha_j) + \mathcal{B}(\alpha_j)\mathcal{B}(\alpha_j)^H$$

for varying $\alpha_j \in \mathbb{C}_-$, which is the alternating directions implicit (ADI) iteration for Lyapunov equations [52]. In order to be applicable to large-scale equations, one uses $X_0 = 0$ and, after

a series of basic algebraic manipulation [38], one arrives at the low-rank ADI (LR-ADI) iteration [38, 6, 34]

$$\begin{aligned} H_j &= (\mathcal{A} + \alpha_j \text{Id})^{-1} \mathcal{W}_{j-1}, \quad \mathcal{W}_j = \mathcal{W}_{j-1} - 2\text{Re}(\alpha_j) H_j, \\ \mathcal{Z}_j &= [\mathcal{Z}_{j-1}, \sqrt{-2\text{Re}(\alpha_j) H_j}] \in \mathbb{C}^{N \times R_j} \end{aligned}$$

for $\mathcal{W}_0 := B$, $j \geq 1$. It produces low-rank approximations of the solution of (42) of the form $V_* \approx V_j := \mathcal{Z}_j \mathcal{Z}_j^H$. The numbers $\alpha_j \in \mathbb{C}_-$ are referred to as shift parameters and are crucial for a fast convergence of the LR-ADI iteration. Obviously, the main numerical effort of the LR-ADI iteration comes from the solution of the linear systems of equations $(\mathcal{A} + \alpha_j \text{Id}) H_j = \mathcal{W}_{j-1}$ for H_j which can for sparse \mathcal{A} be done efficiently by sparse-direct [12] or iterative solvers [46]. The error of the constructed low-rank approximation V_j is given by

$$V_j - V_* = \mathcal{J}_j V_* \mathcal{J}_j^H, \quad \mathcal{J}_j := \prod_{i=1}^j \mathcal{C}_i, \quad \mathcal{C}_i := \mathcal{C}(\alpha_i) \quad (50)$$

such that

$$\begin{aligned} \|V_j - V_*\|_2 &\leq \kappa(S)^2 \|V_*\|_2 R_j^2, \quad \Theta_j := \prod_{i=1}^j \rho_i, \\ \rho_i &= \rho(\mathcal{C}_i) = \max_{z \in \text{spec}(\mathcal{A})} \left| \frac{z - \bar{\alpha}_i}{z + \alpha_i} \right|, \end{aligned}$$

and S is the matrix containing the eigenvectors of \mathcal{A} . Since $\text{spec}(\mathcal{A}) \in \mathbb{C}_-$ and $\alpha_i \in \mathbb{C}_-$, it holds $\rho_i < 1$, $\forall i \geq 1$ and, thus, $\Theta_j = \rho_j \Theta_{j-1} < \Theta_{j-1}$, indicating that the sequence of the spectral radii Θ_j is monotonically decreasing and, in the limit, will approach the value zero. The shifts α_i should therefore be chosen such that Θ_j is as small as possible leading to the ADI parameter problem

$$\{\alpha_1^*, \dots, \alpha_j^*\} = \arg \min \Theta_j = \arg \min_{\alpha_i \in \mathbb{C}_-} \max_{z \in \text{spec}(\mathcal{A})} \left| \prod_{i=1}^j \frac{z - \bar{\alpha}_i}{z + \alpha_i} \right|. \quad (51)$$

This problem is in general a formidable task which has been addressed in numerous works, e.g., in [14, 52, 47, 6, 34]. Again, the situation simplifies for the important case $\mathcal{A} = \mathcal{A}^\top$, where real shifts are usually sufficient (and \mathcal{Z}_j will also be real). Note that in this case $\kappa(S) = 1$. In summary, we have the following relevant result for our purposes:

Theorem 7.4. ([51, 47, 52]) *With the same assumptions and settings for k , k' , K , K' , q as in Theorem 7.3, construct real shifts parameters $\alpha_1, \dots, \alpha_j \in \mathbb{R}_-$ by*

$$\alpha_i = \text{adn}((2i - 1)K/2j), \quad 1 \leq i \leq j \quad (52)$$

with the elliptic function dn from Definition 7.1. Using these shifts, the smallest value of the spectral radius $\Theta_j = \Theta_j(\alpha_1, \dots, \alpha_j)$ in (51) is $\frac{1 - \sqrt{k'_j}}{1 + \sqrt{k'_j}}$, where the modulus k'_j is associated to the nome q^j via (48). Hence, carrying out j steps of the LR-ADI iteration using (52) yields

$$\|V_j - V_*\|_2 \leq \|V_*\|_2 \left(\frac{1 - \sqrt{k'_j}}{1 + \sqrt{k'_j}} \right)^2.$$

We can, thus, expect to approximate the solution V_* by the LR-ADI low-rank approximation V_j at a speed similar to the predicted singular value decay by Theorem 7.3. For other low-rank algorithms for solving (42), e.g., rational Krylov subspace method [11], there exists similar error bounds [11].

8 Summary and Main Result

Although we stated and proved all our main error estimates, it is helpful to summarize the results to provide a CERES. Recall from the introduction that our setup considered four steps. To combine the four steps, we have to link operators on $\dot{\mathcal{H}}^1$ with the finite-dimensional approximation spaces \mathcal{S}_h . If $L_h : \mathcal{S}_h \rightarrow \mathcal{S}_h$ is a finite-dimensional linear operator, then we can always view it as an infinite-dimensional linear operator L on $\dot{\mathcal{H}}^1$ by declaring basis vectors not in $P_h \dot{\mathcal{H}}^1$ to lie in $\text{nullspace}(L)$.

Theorem 8.1. *Suppose (A0)-(A6) hold. Let $\text{Cov}(u)$ denote the covariance operator of the SPDE (5). Then a low-rank solution V_j of the locally linearized discretized problem on \mathcal{S}_h near the steady state u^* computed after j ADI steps satisfies the CERES*

$$\sup_{t \in [0, T]} \|\text{Cov}(u) - V_j\|_{\mathcal{L}_2} \leq \sup_{t \in \mathcal{I}} [\text{err}_{(S1)} + \text{err}_{(S2)} + \text{err}_{(S3)} + \text{err}_{(S4)}] \quad (53)$$

and the individual error terms are given by

$$\text{err}_{(S1)} = C_d h^{1+r}, \quad (54)$$

$$\text{err}_{(S2)} = C_1 [\|\text{Cov}(u_h(0)) - \text{Cov}(U_h(0))\|_2 + \eta^*(t)] e^{-t \min_i |\mathcal{A}_{ii}|}, \quad (55)$$

$$\text{err}_{(S3)} = C_\tau (\|\text{Cov}(U_h(0)) - V_*\|_2) e^{t \max(\text{spec}(\mathcal{A}))} \quad (56)$$

$$\text{err}_{(S4)} = \|V_*\|_2 \left(1 - \sqrt{k'_j}\right)^2 / \left(1 + \sqrt{k'_j}\right)^2, \quad (57)$$

where $C_d, C_1, C_\tau, k'_j > 0$ are constant depending upon the data of the problem (i.e. on A, f, g, Q), $u_h(0), U_h(0)$ are initial conditions for the discretized full and linearized problems, $V_* = \lim_{t \rightarrow +\infty} U_h(t)$ is the finite asymptotic limit for the stationary problem, $\eta^*(t)$ depends upon the data and first two moments of the SPDE, and \mathcal{A} is the leading-order discretized linear operator part near u^* .

Proof. Just applying a triangle inequality to the left-hand side of (53), the proof follows from a direct application of Theorems 4.1, 5.4, 6.1, and 7.4. \square

Theorem 8.1 illustrates very well that it would be short-sighted to just look at one source of error. For example, even if the spatial discretization h is extremely small, the actual error could be very large if the spectral gap is small, i.e., the deterministic steady state is only weakly locally stable. We may also summarize the behaviour of the different constants and terms in the CERES into more practical observations, which effects lead to *smaller error*:

- A large gap in the spectrum of the local linearization of the deterministic part exists.
- A small spatial discretization level h is chosen.

- One starts with initial data close to the local approximating OU stationary distribution.
- The nonlinear part of f does not have a strong effect on sub-exponential time scales.
- The noise is small to stay for a long time in the sub-exponential regime.
- The Q -trace-class operator has fast decaying eigenvalues.
- The iteration number j in the low-rank Lyapunov solver is chosen large.

For a detailed discussion of these effects, we refer to the individual proofs of the different parts of the CERES in previous sections. However, it is very interesting to note that certain effects, which decrease the error occur in *multiple steps*. Obviously this is true for the spectral gap of $\mathcal{A} = A_h + P_h D_u f(u^*)$ in (S2) and (S3) but also for the type of noise, which crucially influences the, usually growing, function $\eta^*(t)$ in (S2) as well as the convergence rate of a low-rank approximation in (S4).

9 Outlook

We stress again that our results presented here should be viewed as a first key step to introduce a general framework of CERES for high-dimensional stochastic problems, where many different sources of error naturally occur. In this regard, a multitude of problems can, and should, now be tackled from a similar viewpoint. For example, uncertainty quantification of random partial differential equations (RPDE) [17, 54, 40, 41] contains an entire chain of error sources such as truncation error for polynomial chaos, dynamical error if the RPDE is just an approximation, error from the large-scale numerical linear algebra, and/or error due to reduced bases, just to name a few. Hence, to compute a CERES in a single norm, such as the spectral norm we used here, for all the steps would be very worthwhile. Similar issues also appear for problems involving large deviations and transition paths in high-dimensional energy landscapes [9, 13, 37, 21], where developing a CERES would definitely be very helpful.

On the very concrete level of the SPDE (5) studied here, several interesting directions could be pursued. Firstly, we do not claim that all our estimates are sharp and/or the assumptions are the theoretically weakest possible. Already the CERES presented in this work is interesting and difficult to chain together properly from its different components. Nevertheless, improvements might be possible, e.g., if f is a strongly dissipative non-Lipschitz nonlinearity, we expect that the results still hold from a dynamical perspective but essentially the steps (S1)-(S2) would then require a major extension or even a completely new approach.

Furthermore, it could be desirable to specify the constant C_d precisely as the techniques in [29, 30] are more explicit than the final statements on optimal regularity. Unfortunately this would entail re-writing the entire optimal regularity proof for the spatial discretization so we refrain from attempting to carry this program in this work. Similar remarks also apply to the constant C_1 but in this case it is important to make the potentially growing function $\eta^*(t)$ explicit, which we have achieved in Section 5.

Acknowledgments: CK would like to thank the VolkswagenStiftung for support via a Lichtenberg Professorship. Furthermore, CK would like to thank Daniele Castellano for interesting discussions regarding moment equations.

References

- [1] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Springer, 2005.
- [2] A.C. Antoulas, D.C. Sorensen, and Y. Zhou. On the decay rate of Hankel singular values and related issues. *Syst. Control Lett.*, 46(5):323–342, 2002.
- [3] J. Baker, M. Embree, and J. Sabino. Fast singular value decay for Lyapunov solutions with nonnormal coefficients. *SIAM Matrix Anal. Appl.*, 36(2):656–668, 2015.
- [4] R.H. Bartels and G.W. Stewart. A solution of the equation $AX + XB = C$. *Commun. ACM*, 15:820–826, 1972.
- [5] R. Bellman. *Introduction to Matrix Analysis*. McGraw-Hill, 1960.
- [6] P. Benner, P. Kürschner, and J. Saak. Self-generating and efficient shift parameters in ADI methods for large Lyapunov and Sylvester equations. *Electron. Trans. Numer. Anal.*, 43:142–162, 2014.
- [7] P. Benner and J. Saak. Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey. *GAMM Mitteilungen*, 36(1):32–52, 2013.
- [8] N. Berglund and B. Gentz. *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*. Springer, 2006.
- [9] P.G. Bolhuis, D. Chandler, C. Dellago, and P.L. Geissler. Transition path sampling: throwing ropes over rough mountain passes. *Ann. Rev. Phys. Chem.*, 53:291–318, 2002.
- [10] R.C. Dalang and L. Quer-Sardanyons. Stochastic integrals for SPDEs: A comparison. *Expositiones Math.*, 29(1):67–109, 2011.
- [11] V. Druskin, L.A. Knizhnerman, and V. Simoncini. Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation. *SIAM J. Numer. Anal.*, 49:1875–1898, 2011.
- [12] I.S. Duff, A.M. Erisman, and J.K. Reid. *Direct Methods for Sparse Matrices*. Clarendon Press, Oxford, UK, 1989.
- [13] W. E., W. Ren, and E. Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66(5):052301, 2002.
- [14] N.S. Ellner and E.L. Wachspress. Alternating direction implicit iteration for systems with complex spectra. *SIAM J. Numer. Anal.*, 28(3):859–870, 1991.
- [15] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer, 2004.
- [16] C. Gardiner. *Stochastic Methods*. Springer, Berlin Heidelberg, Germany, 4th edition, 2009.
- [17] R.G. Ghanem and P.D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Courier Corporation, 2003.
- [18] B. Goldys and J.M.A.M. Van Neerven. Transition semigroups of Banach space-valued Ornstein-Uhlenbeck processes. *Acta Applicandae Mathematica*, 76:283–330, 2003.

- [19] G.H. Golub and C. van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.
- [20] L. Grasedyck. Existence of a low rank or H -matrix approximant to the solution of a Sylvester equation. *Numer. Lin. Alg. Appl.*, 11:371–389, 2004.
- [21] G. Henkelman, B.P. Uberuaga, and H. Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.*, 113:9901–9904, 2000.
- [22] D. Henry. *Geometric Theory of Semilinear Parabolic Equations*. Springer, Berlin Heidelberg, Germany, 1981.
- [23] G.-D. Hu and M. Liu. The weighted logarithmic matrix norm and bounds of the matrix exponential. *Linear Algebra and its Applications*, 390(1):145–154, 2004.
- [24] A. Jentzen and M. Röckner. A Milstein scheme for SPDEs. *arXiv:1001.2751v4*, pages 1–37, 2012.
- [25] A. Jentzen and M. Röckner. Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise. *J. Differential Equat.*, 252:114–136, 2012.
- [26] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, NY, 2nd edition, 2002.
- [27] H. Kielhoefer. *Bifurcation Theory: An Introduction with Applications to PDEs*. Springer, 2004.
- [28] B. Krauskopf, H.M. Osinga, and J. Galán-Vique, editors. *Numerical Continuation Methods for Dynamical Systems: Path following and boundary value problems*. Springer, 2007.
- [29] R. Kruse. Optimal error estimates for Galerkin finite element methods for stochastic partial differential equations with multiplicative noise. *IMA. J. Numer. Anal.*, 34:217–251, 2014.
- [30] R. Kruse and S. Larsson. Optimal regularity for semilinear stochastic partial differential equations with nonlinear multiplicative trace-class noise. *Electron. J. Probab.*, 17:1–19, 2012.
- [31] C. Kuehn. Deterministic continuation of stochastic metastable equilibria via Lyapunov equations and ellipsoids. *SIAM J. Sci. Comp.*, 34(3):A1635–A1658, 2012.
- [32] C. Kuehn. Numerical continuation and SPDE stability for the 2d cubic-quintic Allen-Cahn equation. *SIAM/ASA J. Uncertain. Quantif.*, 3(1):762–789, 2015.
- [33] C. Kuehn. Moment closure - a brief review. In E. Schöll, S. Klapp, and P. Hövel, editors, *Control of Self-Organizing Nonlinear Systems*, pages 253–271. Springer, 2016.
- [34] P. Kürschner. *Efficient Low-Rank Solution of Large-Scale Matrix Equations*. Dissertation, OVGU Magdeburg, 2016.

- [35] A. Lang, S. Larsson, and C. Schwab. Covariance structure of parabolic stochastic partial differential equations. *Stoch. PDE: Anal. Comp.*, 1(2):351–364, 2013.
- [36] D.F. Lawden. *Elliptic Functions and Applications*. Springer New York, 1989.
- [37] T. Lelièvre, G. Stoltz, and M. Rousset. *Free Energy Computations: A Mathematical Perspective*. World Scientific, 2010.
- [38] J.-R. Li and J. White. Low rank solution of Lyapunov equations. *SIAM J. Matrix Anal. Appl.*, 24(1):260–280, 2002.
- [39] G.J. Lord, C.E. Powell, and T. Shardlow. *An Introduction to Computational Stochastic PDEs*. CUP, 2014.
- [40] O. Le Maître and O.M. Knio. *Spectral Methods for Uncertainty Quantification: with applications to computational fluid dynamics*. Springer, 2010.
- [41] F. Nobile, R. Tempone, and C.G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2309–2345, 2008.
- [42] T. Penzl. Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. *Syst. Control Lett.*, 40:139–144, 2000.
- [43] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992.
- [44] C. Prévot and M. Röckner. *A Concise Course on Stochastic Partial Differential Equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, 2008.
- [45] J.C. Robinson. *Infinite-Dimensional Dynamical Systems*. CUP, 2001.
- [46] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, PA, 2003.
- [47] John Sabino. *Solution of Large-Scale Lyapunov Equations via the Block Modified Smith Method*. PhD thesis, Rice University, Houston, Texas, June 2007. Available from: http://www.caam.rice.edu/tech_reports/2006/TR06-08.pdf.
- [48] V. Simoncini. Computational methods for linear matrix equations. *SIAM Rev.*, 58(3):377–441, 2016.
- [49] L. Socha. *Linearization Methods for Stochastic Dynamic Systems*. Springer, 2008.
- [50] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer, 2006.
- [51] E.L. Wachspress. Extended application of alternating direction implicit iteration model problem theory. *Journal of the Society for Industrial and Applied Mathematics*, 11(4):994–1016, 1963.
- [52] E.L. Wachspress. *The ADI model problem*. Springer New York, 2013.
- [53] A.D. Wentzell and M.I. Freidlin. On small random perturbations of dynamical systems. *Russ. Math. Surv.*, 25:1–55, 1970.
- [54] D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.