Adaptive stochastic Galerkin FEM for lognormal coefficients in hierarchical tensor representations

Martin Eigel[1] · Manuel Marschall[1] · Max Pfeffer[2] · Reinhold Schneider[3]

Abstract

Stochastic Galerkin methods for non-affine coefficient representations are known to cause major difficulties from theoretical and numerical points of view. In this work, an adaptive Galerkin FE method for linear parametric PDEs with lognormal coefficients discretized in Hermite chaos polynomials is derived. It employs problem-adapted function spaces to ensure solvability of the variational formulation. The inherently high computational complexity of the parametric operator is made tractable by using hierarchical tensor representations. For this, a new tensor train format of the lognormal coefficient is derived and verified numerically. The central novelty is the derivation of a reliable residual-based a posteriori error estimator. This can be regarded as a unique feature of stochastic Galerkin methods. It allows for an adaptive algorithm to steer the refinements of the physical mesh and the anisotropic Wiener chaos polynomial degrees. For the evaluation of the error estimator to become feasible, a numerically efficient tensor format discretization is developed. Benchmark examples with unbounded lognormal coefficient fields illustrate the performance of the proposed Galerkin discretization and the fully adaptive algorithm.

Mathematics Subject Classification 35R60 · 47B80 · 60H35 · 65C20 · 65N12 · 65N22 · 65J10

✉ Reinhold Schneider
   schneidr@math.tu-berlin.de

   Martin Eigel
   martin.eigel@wias-berlin.de

   Manuel Marschall
   manuel.marschall@wias-berlin.de

   Max Pfeffer
   max.pfeffer@mis.mpg.de

[1] Weierstrass Institute, Mohrenstrasse 39, 10117 Berlin, Germany

[2] Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany

[3] TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

1 Introduction

In the thriving field of uncertainty quantification (UQ), efficient numerical methods for the approximate solution of random PDEs have been a topic of vivid research. As common benchmark problem, one often considers the Darcy equation (as a model for flow through a porous medium) with different types of random coefficients in order to assess the efficiency of a numerical approach. Two important properties are the length of the expansion of random fields, which often directly translates to the number of independent random variables describing the variability in the model, and the type of dependence on these random variables. The affine case with uniform random variables has been studied extensively, since it represents a rather simple model which can easily be treated with standard methods. Opposite to that, the lognormal case with Gaussian random variables is quite challenging, from the analytical as well as numerical point of view. A theory for the solvability of linear elliptic PDEs with respective unbounded coefficients (and hence a lack of uniform ellipticity) in a variational formulation was only developed recently in [25,41,49]. Computationally, the problems quickly become difficult or even intractable with many stochastic dimensions, which might be required to accurately represent the stochasticity in the random field expansion. This paper is concerned with the development of an efficient numerical method for this type of problems.

While popular sample-based Monte Carlo methods obtain dimension-independent convergence rates, these are rather low despite often encountered higher regularity of the parameter dependence. Moreover, such methods can only be used to evaluate functionals of the solution (QoIs = quantities of interest) and an a posteriori error control usually is not feasible reliably. Some recent developments in this field can e.g. be found in [32,33,36,38] for the model problem with a lognormal coefficient. Some ideas on a posteriori adaptivity for Monte Carlo methods can e.g. be found in [11,20].

An alternative are functional (often called spectral) approximations, which for instance are obtained by Stochastic Collocation (SC) [2,43,44], the related Multi-level Quadraure (MLQ) [30,31] and Stochastic Galerkin (SG[1].) methods. The latter in particular is popular in the engeneering sciences since it can be perceived as an extension of classical finite element methods (FEM). These approaches provide a complete parameter to solution map based on which e.g. statistical moments of the stochastic solution can be evaluated. Notably, the regularity of the solution can be exploited in order to obtain quasi-optimal convergence rates. However, the number of random variables and nonlinear parameter representations have a significant impact on the computational feasibility and techniques for a model order reduction are required. Collocation methods with pointwise evaluations in the parameter space are usually constructed either based on some a priori knowledge or by means of an iterative refinement algorithm which takes into account the hierarchical surplus on possible new discretization levels. While these approaches work reasonably well, methods for a reliable error control do not seem immediate since the approximation relies only on interpolation properties. Nevertheless, for the affine case and under certain assumptions, first ideas were recently presented in [28].

---

[1] We usually use SGFEM for Stochastic Galerkin FEM

The computationally more elaborate Stochastic Galerkin methods carry out an orthogonal projection with respect to the energy norm onto a discrete space which is usually spanned by a tensor basis consisting of FE basis functions in physical space and polynomial chaos polynomials orthogonal with respect to the joint parameter measure in (stochastic) parameter space. The use of global polynomials is justified by the high (analytic) regularity of the solution map with respect to the parameters [8,9,34]. However, in particular the large computational cost of Galerkin methods make adaptivity and model reduction techniques a necessity.

In order to achieve this, different paths have been pursued successfully. As a first approach, sparse approximations as in [15,16,19] or [4,5,10] with either a residual based or a hierarchical a posteriori error estimators can be computed. Here, the aim is to restrict an exponentially large discrete basis to the most relevant functions explictly by iteratively constructing a quasi-optimal subset. In [16], convergence of the adaptive algorithm could be shown. Moreover, adjoint based error estimators are considered in [6,48].

As a second approach, an adaptive discretization in hierarchical tensor representations can be derived as described in [21]. These modern compression formats have lately been investigated intensively in the numerical analysis community [3,29,45,46]. It has been examined that with appropriate assumptions the curse of dimensionality can be alleviated, particularly so when employed with typical random PDE problems in UQ, see [12,13] for examples with sample-based reconstruction strategies. Such representations can be understood as an implicit model order reduction technique, closely related to (but more general than e.g.) reduced basis methods.

In the mentioned adaptive approaches, the FE mesh for the physical space and the parametric polynomial chaos space are adapted automatically with respect to the considered problem. In the case of tensor approximations, also the ranks of the representation are adjusted.

However, in all adaptive SGFEM research, so far only the affine case with uniform elliptic coefficient has been considered. In this paper, we extend the ASGFEM approach developed in [21] to the significantly more involved case of lognormal (non-affine) coefficients. This poses several severe complications analytically and numerically. Analytical aspects have recently been tackled in [25,26,34,41]. Numerically, in particular computationally efficient Galerkin methods are quite diffucult to construct for this case and have not been devised. Compression techniques and adaptivity most certainly are required in order to make these problems tractable with SGFEM, as described in this paper. Of particular interest is the construction of a computable a posteriori error estimator, which also greatly benefits from using tensor formats. In order to obtain a well-posed discretization, problem adapted spaces according to the presentation in [49] are used.

Main contributions of this work are a representation of the coefficient in the tensor train (TT) format, the operator discretization in tensor format and the derivation of an reliable residual based error estimatator. This then serves as the basis for an adaptive algorithm which steers all discretization parameters of the SGFEM. The performance of the proposed method is demonstrated with some benchmark problems. Here, the used field models are not artificially bounded or shifted away from zero.

We point out that, to our knowledge, an SGFEM for the lognormal case so far has only been practically computed in the weighted function space setting in [42] for a small 1d model as proof of concept. Moreover, there has not been any adaptive numerical method with reliable a posteriori error estimation as derived in this work. However, we note that our approach relies on the assumption that the coefficient is discretized sufficiently accurately and hence the related discretization error can be neglected. In practice, this can be ensured with high probability by sampling the error of the discrete coefficient. Additionally, since constants in the error bound can become quite large, we interpret the error estimate as a refinement indicator.

It should be noted that a functional adaptive evluation of the forward map allows for the derivation of an explicit adaptive Bayesian inversion with functional tensor representations as in [17]. The results of the present work lay the ground for a similar approach with a Gaussian prior assumption. This will be the topic of future research. Moreover, the described approach enables to construct SGFEM with arbitrary densities (approximated in hierarchical tensor formats). This generalization should also be examined in more detail in further research. Lastly, while sparse discretizations seem infeasible for the lognormal coefficient, a transformation [51] yields a convection-diffusion formulation of the problem with affine parameter dependence, which then again is amenable to an adaptive sparse SGFEM. This direction is examined in [14].

The structure of this paper is as follows: We first introduce the setting of parametric PDEs with our model problem in Sect. 2. The variational formulation in problem dependent weighted function spaces and the finite element (FE) setting are described in Sect. 3. The employed tensor formats and the tensor representations of the coefficient and the differential operator are examined in Sect. 4. As a central part, in Sect. 5 we derive the a posteriori error estimators and define a fully adaptive algorithm in Sect. 6, including efficient ways to compute error indicators for physical and stochastic refinement. Numerical examples in Sect. 7 illustrate the performance of the presented method and conclude the paper.

2 Setting and discretization

In the following, we introduce the considered model problem formally, present its weak formulation and describe the employed discretizations in finite dimensional function spaces. We closely follow the presentations in [26,34,49] regarding the lognormal problem in problem-dependent function spaces. In [21], a related formulation for the solution and evaluation of a posteriori error estimators for parametric PDEs with affine coefficient fields in hierarchical tensor formats is derived.

2.1 Model problem

We assume some bounded domain $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, with Lipschitz boundary ∂D and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $\mathbb{P}$-almost all $\omega \in \Omega$, we consider the random elliptic problem

$$\begin{cases} -\operatorname{div}(a(x,\omega)\nabla u(x,\omega)) = f(x) & \text{in } D, \\ \quad\quad\quad\quad\quad u(x,\omega) = 0 & \text{on } \partial D. \end{cases} \quad (2.1)$$

The coefficient $a : D \times \Omega \mapsto \mathbb{R}$ denotes a lognormal, isotropic diffusion coefficient, i.e., $\log(a)$ is an isotropic Gaussian random field.

Remark 2.1 The source term $f \in L^2(D)$ is assumed deterministic. However, it would not introduce fundamental additional difficulties to also model f and the boundary conditions as stochastic fields not correlated to the coefficient $a(x,\omega)$ as long as appropriate integrability of the data is given.

For the coefficient $a(x,\omega)$ of (2.1), we assume a Karhunen-Loève type expansion of $b := \log(a)$ of the form

$$b(x,\omega) = \sum_{\ell=1}^{\infty} b_\ell(x) Y_\ell(\omega), \quad x \in D, \quad \mathbb{P}\text{- almost all } \omega \in \Omega.$$

Here, the parameter vector $Y = (Y_\ell)_{\ell \in \mathbb{N}}$ consists of independent standard normal Gaussian random variables in $\mathbb{R}$. Then, by passing to the image space $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \gamma)$ with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ of all open sets of $\mathbb{R}^{\mathbb{N}}$ and the Gaussian product measure

$$\gamma := \bigotimes_{\ell \in \mathbb{N}} \gamma_\ell \quad \text{with} \quad \gamma_\ell := \gamma_1 := \mathcal{N}_1 := \mathcal{N}(0,1)$$

$$\text{and} \quad d\gamma_1(y_\ell) = \frac{1}{\sqrt{2\pi}} \exp(-y_\ell^2/2) \, dy_\ell,$$

we can consider the parameter vector $y = (y_\ell)_{\ell \in \mathbb{N}} = (Y_\ell(\omega))_{\ell \in \mathbb{N}}, \omega \in \Omega$.

For any sequence $\boldsymbol{\beta} \in \ell^1(\mathbb{N})$ with

$$\beta_\ell := \|b_\ell\|_{L^\infty(D)} \quad \text{and} \quad \boldsymbol{\beta} = (\beta_\ell)_{\ell \in \mathbb{N}},$$

we define the set

$$\Gamma_{\boldsymbol{\beta}} := \left\{ y \in \mathbb{R}^{\mathbb{N}} : \sum_{\ell=1}^{\infty} \beta_\ell |y_\ell| < \infty \right\}.$$

The set $\Gamma_{\boldsymbol{\beta}}$ of admissible parameter vectors is γ-measurable and of full measure.

Lemma 2.2 ([34, Lemma 2.1]) *For any sequence $\boldsymbol{\beta} \in \ell^1(\mathbb{N})$, there holds*

$$\Gamma_{\boldsymbol{\beta}} \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}) \quad \text{and} \quad \gamma(\Gamma_{\boldsymbol{\beta}}) = 1.$$

For any $y \in \Gamma_{\boldsymbol{\beta}}$, we define the deterministic parametric coefficient

$$a(x, y) = \exp(b(x, y)) = \exp\left(\sum_{\ell=1}^{\infty} b_\ell(x) y_\ell\right), \qquad x \in D. \tag{2.2}$$

This series converges in $L^\infty(D)$ for all $y \in \Gamma_{\boldsymbol{\beta}}$.

Lemma 2.3 ([34, Lemma 2.2]) *For all $y \in \Gamma_{\boldsymbol{\beta}}$, the diffusion coefficient* (2.2) *is well-defined and satisfies*

$$0 < \check{a}(y) := \operatorname*{ess\,inf}_{x \in D} a(x, y) \le a(x, y) \le \operatorname*{ess\,sup}_{x \in D} a(x, y) =: \hat{a}(y) < \infty,$$

with

$$\hat{a}(y) \le \exp\left(\sum_{\ell=1}^{\infty} \beta_\ell |y_\ell|\right) \quad and \quad \check{a}(y) \ge \exp\left(-\sum_{\ell=1}^{\infty} \beta_\ell |y_\ell|\right).$$

Due to Lemmas 2.2 and 2.3 , we consider $\Gamma = \Gamma_{\boldsymbol{\beta}}$ as the parameter space instead of $\mathbb{R}^{\mathbb{N}}$. By Lemma 2.3, the stochastic coefficient $a(x, y)$ is well defined, bounded from above and admits a positive lower bound for almost all $y \in \Gamma$. Thus, the equations (2.1) and (2.2) have a unique solution $u(y) \in \mathcal{X}$ for almost all $y \in \Gamma$.

Let $\mathcal{X} := H_0^1(D)$ denote the closed subspace of functions in the Sobolev space $H^1(D)$ with vanishing boundary values in the sense of trace and define the norm

$$\|v\|_{\mathcal{X}} := \left(\int_D |\nabla v(x)|^2 \, \mathrm{d}x\right)^{1/2}.$$

We denote by $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\mathcal{X}^*, \mathcal{X}}$ the duality pairing of $\mathcal{X}^*$ and $\mathcal{X}$ and consider $f \in L^2(D)$ as an element of the dual $\mathcal{X}^* = H^{-1}(D)$.

For any $y \in \Gamma$, the variational formulation of (2.1) reads: find $u(y) \in \mathcal{X}$ such that

$$B(u(y), v; y) = \langle f, v \rangle, \qquad \text{for all } v \in \mathcal{X}, \tag{2.3}$$

where $B : \mathcal{X} \times \mathcal{X} \times \Gamma \to \mathbb{R}$ is defined by

$$B(u(y), v; y) := \int_D a(x, y) \nabla u(y) \cdot \nabla v \, \mathrm{d}x.$$

Hence, pathwise existence and uniqueness of the solution $u(y)$ is obtained by the Lax-Milgram lemma due to uniform ellipticity for any fixed $y \in \Gamma$. In particular, for all $y \in \Gamma$, (2.3), it holds

$$\|u(y)\| \le \frac{1}{\check{a}(y)} \|f\|_{\mathcal{X}^*},$$

with some $0 < \check{a}(y) \le a(x, y)$ on D. The integration of (2.3) over Γ with respect to the standard normal Gaussian measure γ does not lead to a well-defined problem since the coefficient $a(x, y)$ is not uniformly bounded in $y \in \Gamma$ and not bounded away from zero. Hence, a more involved approach has to be pursued, which is elaborated in Sect. 3.2. Alternative results for this equation were presented in [25,26,49].

The formulation of (2.1) as a parametric deterministic elliptic problem with solution $u(y) \in \mathcal{X}$ for each parameter $y \in \Gamma$ reads

$$- \operatorname{div}(a(x, y)\nabla u(x, y)) = f(x) \quad \text{for } x \in D, \quad u(x, y) = 0 \quad \text{for } x \in \partial D.$$

3 Variational formulation and discretization

This section is concerned with the introduction of appropriate function spaces required for the discretization of the model problem. In particular, a problem-adapted probability measure is introduced which allows for a well-defined formulation of the weak problem rescaled polynomial chaos basis.

3.1 Problem-adapted function spaces

Let $\mathcal{F}$ be the set of finitely supported multi-indices

$$\mathcal{F} := \{\mu \in \mathbb{N}_0^\infty \,;\, |\operatorname{supp} \mu| < \infty\} \quad \text{where} \quad \operatorname{supp} \mu := \{m \in \mathbb{N} \,;\, \mu_m \ne 0\}.$$

A full tensor index set of order $M \in \mathbb{N}$ is defined by

$$\begin{aligned} \Lambda &:= \{(\mu_1, \ldots, \mu_M, 0, \ldots) \in \mathcal{F} : \mu_m = 0, \ldots, d_m - 1, \ m = 1, \ldots, M\} \\ &\simeq \Lambda_1 \times \ldots \times \Lambda_M \times \{0\} \ldots \subset \mathcal{F}, \end{aligned}$$

with complete index sets of size d_m given by

$$\Lambda_m := \{0, \ldots, d_m - 1\}, \quad m = 1, \ldots, M.$$

For any such subset $\Lambda \subset \mathcal{F}$, we define $\operatorname{supp} \Lambda := \bigcup_{\mu \in \Lambda} \operatorname{supp} \mu \subset \mathbb{N}$.

We denote by $(H_n)_{n=0}^\infty$ the orthonormal basis of $L^2(\mathbb{R}, \gamma_m) = L^2(\mathbb{R}, \gamma_1)$ with respect to the standard Gaussian measure consisting of Hermite polynomials H_n of degree $n \in \mathbb{N}_0$ on $\mathbb{R}$. An orthogonal basis of $L^2(\Gamma, \gamma)$ is obtained by tensorization of the univariate polynomials, see [49,50]. To reduce notation, we drop the explicit dependency on the sigma-algebra which is always assumed to be rich enough. For any multi-index $\mu \in \mathcal{F}$, the tensor product polynomial $H_\mu := \bigotimes_{m=1}^\infty H_{\mu_m}$ in $y \in \Gamma$ is expressed as the finite product

$$H_\mu(y) = \prod_{m=1}^\infty H_{\mu_m}(y_m) = \prod_{m \in \operatorname{supp} \mu} H_{\mu_m}(y_m).$$

For practical computations, an analytic expression for the triple product of Hermite polynomials can be used.

Lemma 3.1 ([40,50]) *For $\mu, \nu, \xi \in \mathcal{F}, m \in \mathbb{N}$, it holds*

$$H_{\nu_m} H_{\mu_m} = \sum_{\xi_m=0}^{\min(\nu_m, \mu_m)} \kappa_{\nu_m, \mu_m, \nu_m+\mu_m-2\xi_m} H_{\nu_m+\mu_m-2\xi_m},$$

where for $\eta_m = \nu_m + \mu_m - 2\xi_m$

$$\kappa_{\nu_m, \mu_m, \eta_m} := \int_{\mathbb{R}} H_{\nu_m}(y_m) H_{\mu_m}(y_m) H_{\eta_m}(y_m) \, d\gamma_m(y_m)$$

$$= \begin{cases} \frac{\sqrt{\nu_m! \mu_m! \eta_m!}}{\xi_m!(\nu_m-\xi_m)!(\mu_m-\xi_m)!}, & \nu_m + \mu_m - \eta_m \text{ is even and} \\ & |\nu_m - \mu_m| \leq \eta_m \leq \nu_m + \mu_m, \\ 0, & \text{otherwise.} \end{cases}$$

Lemma 3.2 ([22,41]) *Let $Y \sim \mathcal{N}_1$, $t \in \mathbb{R}$ and $X = \exp(tY) \in L^2(\mathbb{R}, \gamma)$. The expansion of $X = \exp(tY)$ in Hermite polynomials is given by*

$$X = \exp(tY) = \sum_{n \in \mathbb{N}_0} c_n H_n \quad \text{with} \quad c_n = \frac{t^n}{\sqrt{n!}} \exp(t^2/2).$$

We recall some results from [49] required in our setting. Let $\boldsymbol{\sigma} = (\sigma_m)_{m \in \mathbb{N}} \in \exp(\ell^1(\mathbb{N}))$ and define

$$\gamma_{\boldsymbol{\sigma}} := \bigotimes_{m=1}^{\infty} \gamma_{\sigma_m} := \bigotimes_{m=1}^{\infty} \mathcal{N}_{\sigma_m^2} := \bigotimes_{m=1}^{\infty} \mathcal{N}(0, \sigma_m^2).$$

Then, $d\mathcal{N}_{\sigma_m^2} = \zeta_{\boldsymbol{\sigma}, m} \, d\mathcal{N}_1$ where

$$\zeta_{\boldsymbol{\sigma}, m}(y_m) := \frac{1}{\sigma_m} \exp\left(-\frac{1}{2}(\sigma_m^{-2} - 1)y_m^2\right)$$

is the one-dimensional Radon–Nikodym derivative of γ_{σ_m} with respect to γ_m, i.e., the respective probability density. We assume that the sequence $\boldsymbol{\sigma}$ depends exponentially on $\boldsymbol{\beta} = (\beta_m)_{m \in \mathbb{N}}$ and some $\varrho \in \mathbb{R}$, namely

$$\sigma_m(\varrho) := \exp(\varrho \beta_m), \quad m \in \mathbb{N},$$

and define

$$\gamma_\varrho := \gamma_{\boldsymbol{\sigma}(\varrho)} \quad \text{and} \quad \zeta_{\varrho, m} := \zeta_{\boldsymbol{\sigma}(\varrho), m}.$$

By multiplication, this yields the multivariate identity

$$\mathrm{d}\gamma_\varrho(y) = \zeta_\varrho(y)\,\mathrm{d}\gamma(y) \quad \text{with} \quad \zeta_\varrho(y) = \prod_{m=1}^{\infty} \zeta_{\varrho,m}(y_m).$$

A basis of orthonormal polynomials with respect to the weighted measure γ_ϱ can be defined by the transformation

$$\tau_\varrho : \mathbb{R}^\infty \to \mathbb{R}^\infty, \quad (y_m)_{m\in\mathbb{N}} \mapsto \left(e^{-\varrho\beta_m} y_m\right)_{m\in\mathbb{N}}.$$

Then, for all $v \in L^2(\Gamma, \gamma)$,

$$\int_\Gamma v(y)\mathrm{d}\gamma(y) = \int_\Gamma v(\tau_\varrho(y))\,\mathrm{d}\gamma_\varrho(y).$$

We define the scaled Hermite polynomials $H_\mu^{\tau_\varrho} := H_\mu \circ \tau_\varrho$.

Remark 3.3 Lemmas 3.1 and 3.2 are also valid with the transformed multivariate Hermite polynomials H^{τ_ϱ}. In particular, $\kappa_{\xi_m, v_m, \mu_m}$ does not change under transformation and the expansion in Lemma 3.2 holds by substituting $t \in \mathbb{R}$ with $\sigma_m t$ in the corresponding dimension $m \in \mathbb{N}$.

3.2 Weak formulation in problem-dependent spaces

In order to obtain a well-posed variational formulation of (2.1) on $L^2(\Gamma, \gamma; \mathcal{X})$, we follow the approach in [49] and introduce a measure γ_ϱ which is stronger than γ and assume integrability of f with respect to this measure. For $\varrho > 0$ and $0 \le \theta < 1$, let the bilinear form $B_{\theta\varrho} : \mathcal{V}_{\theta\varrho} \times \mathcal{V}_{\theta\varrho} \to \mathbb{R}$ be given by

$$B_{\theta\varrho}(w, v) := \int_\Gamma \int_D a(x, y)\nabla w(x, y) \cdot \nabla v(x, y)\,\mathrm{d}x\mathrm{d}\gamma_{\theta\varrho}(y). \qquad (3.1)$$

The solution space is then defined as the Hilbert space

$$\mathcal{V}_{\theta\varrho} := \{v : \Gamma \to \mathcal{X} \quad \mathcal{B}(\Gamma)\text{-measurable}\,;\ B_{\theta\varrho}(v, v) < \infty\},$$

endowed with the inner product $B_{\theta\varrho}(\cdot, \cdot)$, the induced energy norm $\|v\|_{\theta\varrho} := B_{\theta\varrho}(v, v)^{1/2}$ for $v \in \mathcal{V}_{\theta\varrho}$ and the respective dual pairing $\langle \cdot, \cdot \rangle_{\theta\varrho}$ between $\mathcal{V}_{\theta\varrho}^*$ and $\mathcal{V}_{\theta\varrho}$. The different employed spaces are related as follows.

Lemma 3.4 ([49, Proposition 2.43]) *For $0 < \theta < 1$,*

$$L^2(\Gamma, \gamma_\varrho; \mathcal{X}) \subset \mathcal{V}_{\theta\varrho} \subset L^2(\Gamma, \gamma; \mathcal{X})$$

are continuous embeddings.

It can be shown that the bilinear form $B_{\theta\varrho}(\cdot, \cdot)$ is $\mathcal{V}_{\theta\varrho}$-elliptic in the sense of the following Lemma.

Lemma 3.5 ([49, Lemma 2.41, 2.42]) *For $w, v \in L^2(\Gamma, \gamma_\varrho; \mathcal{X})$,*

$$|B_{\theta\varrho}(w, v)| \leq \hat{c}_{\vartheta\varrho} \|w\|_{L^2(\Gamma, \gamma_\varrho; \mathcal{X})} \|v\|_{L^2(\Gamma, \gamma_\varrho; \mathcal{X})} \tag{3.2}$$

and for $v \in L^2(\Gamma, \gamma; \mathcal{X})$,

$$B_{\theta\varrho}(v, v) \geq \check{c}_{\vartheta\varrho} \|v\|_{L^2(\Gamma, \gamma; \mathcal{X})}^2. \tag{3.3}$$

Moreover, for $f \in \mathcal{X}^*$ we define the continuous linear form

$$F_{\theta\varrho}(v) := \int_\Gamma \int_D f(x) v(x, y) \mathrm{d}x \mathrm{d}\gamma_{\theta\varrho}(y).$$

For $F_{\theta\varrho} \in \mathcal{V}_{\theta\varrho}^*$, (3.2) and (3.3) in particular lead to the unique solvability of the variational problem in $\mathcal{V}_{\theta\varrho}$,

$$B_{\theta\varrho}(u, v) = F_{\theta\varrho}(v) \qquad \text{for all } v \in \mathcal{V}_{\theta\varrho},$$

and $u \in \mathcal{V}_{\theta\varrho}$ is the unique solution of (2.1).

3.3 Deterministic discretization

We discretise the deterministic space $\mathcal{X}$ by a conforming finite element space $\mathcal{X}_p(\mathcal{T}) := \mathrm{span}\{\varphi_i\}_{i=1}^N \subset \mathcal{X}$ of degree p on some simplicial regular mesh $\mathcal{T}$ of domain D with the set of faces $\mathcal{S}$ (i.e., edges for $d = 2$) and basis functions φ_i.

In order to circumvent complications due to an inexact approximation of boundary values, we assume that D is a polygon. We denote by $P_p(\mathcal{T})$ the space of piecewise polynomials of degree p on the triangulation $\mathcal{T}$. The assumed FE discretization with Lagrange elements of order p then satisfies $\mathcal{X}_p(\mathcal{T}) \subset P_p(\mathcal{T}) \cap C(\overline{\mathcal{T}})$. For any element $T \in \mathcal{T}$ and face $F \in \mathcal{S}$, we set the entity sizes $h_T := \mathrm{diam}\, T$ and $h_F := \mathrm{diam}\, F$. Let n_F denote the exterior unit normal on any face F. The jump of some $\chi \in H^1(D; \mathbb{R}^d)$ on $F = \overline{T_1} \cap \overline{T_2}$ in normal direction $[\![\chi]\!]_F$ is then defined by

$$[\![\chi]\!]_F := \chi|_{T_1} \cdot n_F - \chi|_{T_2} \cdot n_F.$$

By ω_T and ω_F we denote the element and facet patches defined by the union of all elements which share at least a vertex with T or F, respectively. Consequently, the Clément interpolation operator $I : \mathcal{X} \to \mathcal{X}_p(\mathcal{T})$ satisfies

$$\|v - Iv\|_{L^2(T)} \leq c_{\mathcal{T}} h_T |v|_{\mathcal{X}, \omega_T} \quad \text{for } T \in \mathcal{T},$$
$$\|v - Iv\|_{L^2(F)} \leq c_{\mathcal{S}} h_F^{1/2} |v|_{\mathcal{X}, \omega_F} \quad \text{for } F \in \mathcal{S},$$

where the seminorms $|\cdot|_{\mathcal{X},\omega_T}$ and $|\cdot|_{\mathcal{X},\omega_F}$ are the restrictions of $\|\cdot\|_{\mathcal{X}}$ to ω_T and ω_F, respectively.

The fully discrete approximation space with $|\Lambda| < \infty$ is given by

$$
\mathcal{V}_N := \mathcal{V}_N(\Lambda; \mathcal{T}, p) := \left\{ v_N(x, y) = \sum_{\mu \in \Lambda} v_{N,\mu}(x) H_\mu^{\tau_{\theta\varrho}}(y); \ v_{N,\mu} \in \mathcal{X}_p(\mathcal{T}) \right\},
$$

and it holds $\mathcal{V}_N(\Lambda; \mathcal{T}, p) \subset \mathcal{V}_{\theta\varrho}$. The Galerkin projection of u is the unique $u_N \in \mathcal{V}_N(\Lambda; \mathcal{T}, p)$ which satisfies

$$
B_{\theta\varrho}(u_N, v_N) = F_{\theta\varrho}(v_N) \qquad \text{for all } v_N \in \mathcal{V}_N(\Lambda; \mathcal{T}, p).
$$

We define a tensor product interpolation operator $\mathcal{I} \colon L^2(\Gamma, \gamma; \mathcal{X}) \to \mathcal{V}_N(\Lambda; \mathcal{T}, p)$ for $v = \sum_{\mu \in \mathcal{F}} v_\mu H_\mu \in L^2(\Gamma, \gamma; \mathcal{X})$ by setting

$$
\mathcal{I}v := \sum_{\mu \in \Lambda} (I v_\mu) H_\mu.
$$

For $v \in \mathcal{V}_{\vartheta\varrho}(\Lambda) := \left\{ v = \sum_{\mu \in \Lambda} v_\mu H_\mu; \ v_\mu \in \mathcal{X} \right\}$ and all $T \in \mathcal{T}$, this yields the interpolation estimate

$$
\begin{aligned}
\|(\mathrm{id} - \mathcal{I})v\|_{L^2(\Gamma,\gamma;L^2(T))} &= \left(\int_\Gamma \left\| \sum_{\mu \in \Lambda} (\mathrm{id} - I) v_\mu H_\mu(y) \right\|_{L^2(T)}^2 \, d\gamma(y) \right)^{1/2} \\
&= \left(\sum_{\mu \in \Lambda} \|(\mathrm{id} - I) v_\mu\|_{L^2(T)}^2 \right)^{1/2} \\
&\leq c_T h_T \left(\int_\Gamma \left| \sum_{\mu \in \Lambda} v_\mu H_\mu(y) \right|_{\mathcal{X},\omega_T}^2 \, d\gamma(y) \right)^{1/2} \\
&= c_T h_T |v|_{\mathcal{V}_{\vartheta\varrho},\omega_T}.
\end{aligned}
\tag{3.4}
$$

Likewise, on the edges $F \in \mathcal{S}$ we derive

$$
\|v - \mathcal{I}v\|_{L^2(\Gamma,\gamma;L^2(F))} \leq c_{\mathcal{S}} h_F^{1/2} |v|_{\mathcal{V}_{\vartheta\varrho},\omega_F}.
$$

Here,

$$
\begin{aligned}
|v|_{\mathcal{V}_{\vartheta\varrho},\omega_T}^2 &:= \int_\Gamma |v(y)|_{\mathcal{X},\omega_T}^2 \, d\gamma(y), \\
|v|_{\mathcal{V}_{\vartheta\varrho},\omega_F}^2 &:= \int_\Gamma |v(y)|_{\mathcal{X},\omega_F}^2 \, d\gamma(y).
\end{aligned}
$$

Theorem 3.6 ([49], Theorem 2.45). *If* $f \in L^p(\Gamma, \gamma_\varrho; \mathcal{X}^*)$ *for* $p > 2$, *the Galerkin projection* $u_N \in \mathcal{V}_N$ *satisfies*

$$\|u - u_N\|_{L^2(\Gamma, \gamma; \mathcal{X})} \leq \frac{\hat{c}_{\vartheta_\varrho}}{\check{c}_{\vartheta_\varrho}} \inf_{v_N \in \mathcal{V}_N(\Lambda; \mathcal{T}, p)} \|u - v_N\|_{L^2(\Gamma, \gamma_\varrho; \mathcal{X})}.$$

Remark 3.7 It should be noted that the constant $\frac{\hat{c}_{\vartheta_\varrho}}{\check{c}_{\vartheta_\varrho}}$ tends to ∞ as $\varrho \to \{0, \infty\}$ and for $\theta \to \{0, 1\}$, see Remark 2.46 in [49]. This is expected, as the problem is ill-posed in these limits. In order to obtain reasonable upper error bounds, the parameters have hence to be chosen judiciously. A more detailed investigation of an optimal parameter choice is postponed to future research.

4 Decomposition of the operator

In this section, we introduce the discretization of the operator in an appropriate tensor format. For this, an efficient representation of the non-linear coefficient is derived. We first introduce basic aspects of the employed Tensor Train (TT) format.

4.1 The tensor train format

We only provide a brief overview of the notation used regarding the tensor train representation. For further details, we refer the reader to [3,21] and the references therein.

Any function $w_N \in \mathcal{V}_N(\Lambda; \mathcal{T}, p)$ can be written as

$$w_N = \sum_{i=0}^{N-1} \sum_{\mu \in \Lambda} W(i, \mu) \varphi_i H_\mu^{\tau_{\theta_\varrho}}.$$

Thus, the discretization space is isomorphic to the tensor space, namely

$$\mathcal{V}_N(\Lambda; \mathcal{T}, p) \simeq \mathbb{R}^{N \times d_1 \times \cdots \times d_M}.$$

The tensor W grows exponentially with the order M, which constitutes the so called *curse of dimensionality*. We employ a low-rank decomposition of the tensor for a dimension reduction. In this paper, we adhere to the *Tensor Train (TT)* format for tensor decomposition [46]. This seems reasonable, as the components (of the operator and hence the solution) are of decreasing importance due to the decay of the coefficient functions b_m and therefore we can expect decreasing ranks in the tensor train format. Nevertheless, other tensor formats are also feasible in principle.

The TT representation of a tensor $W \in \mathbb{R}^{N \times d_1 \times \cdots \times d_M}$ is given as

$$W(i, \mu_1, \ldots, \mu_M) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_M=1}^{r_M} W_0(i, k_1) \prod_{m=1}^{M} W_m(k_m, \mu_m, k_{m+1}).$$

For simplicity of notation, we set $r_0 = r_{M+1} = 1$. If all dimensions r_m are minimal, then this is called the *TT decomposition* of W and $\text{rank}_{\text{TT}}(W) := r = (1, r_1, \ldots, r_M, 1)$ is called the *TT rank* of W. The TT decomposition always exists and it can be computed in polynomial time using the *hierarchical SVD* (HSVD) [35]. A truncated HSVD yields a *quasi-optimal* approximation in the Frobenius norm [27,29,39,46]. Most algebraic operations can be performed efficiently in the TT format [3].

Once the function w_N has a low-rank TT decomposition, it is advisable to obtain a similar representation for the Galerkin operator on $\mathcal{V}_N(\Lambda; \mathcal{T}, p)$ in order to allow for efficient tensor solvers. For the lognormal coefficient $a(x, y)$, this can only be done approximately.

Later, it will be useful to express the storage complexity of a tensor train. We distinguish the degrees of freedom given by the tensor train representation and the full (uncompressed) degrees of freedom. For a tensor $U \in \mathbb{R}^{q_0 \times \cdots \times q_L}$ of TT-rank $r = (1, r_1, \ldots, r_L, 1)$, the dimension of the low-rank tensor manifold is given by

$$\text{tt-dofs}(U) := \sum_{\ell=1}^{L-1} (r_\ell q_\ell r_{\ell+1} - r_{\ell+1}^2) + r_L q_L, \tag{4.1}$$

while the dimension of the full tensor space and hence its representation is

$$\text{full-dofs}(U) := \prod_{\ell=0}^{L} q_\ell.$$

One can conclude from (4.1) that the complexity of tensor trains depend only linearly on the dimension, i.e., we have to store

$$\mathcal{O}(L \hat{q} \hat{r}^2), \quad \hat{r} = \max\{r\} \quad \hat{q} = \max\{q_0, \ldots, q_L\}$$

entries instead of $\mathcal{O}(\hat{q}^L)$, which is much smaller for moderate TT-ranks r.

4.2 TT representation of the non-linear coefficient

s

We approximate the coefficient

$$a(x, y) = \exp\left(\sum_{\ell=1}^{\infty} b_\ell(x) y_\ell\right) = \prod_{\ell=1}^{\infty} e^{b_\ell(x) y_\ell} \tag{4.2}$$

using the coefficient splitting algorithm described in [47]. This results in a discretized coefficient on a tensor set

$$\Delta = \{(v_1, \ldots, v_L, 0, \ldots) \in \mathcal{F} : v_\ell = 0, \ldots, q_\ell - 1, \ \ell = 1, \ldots, L\}$$

with TT-rank $s = (1, s_1, \ldots, s_L, 1)$. Here, we exploit Lemma 3.2, i.e., the fact that every factor of (4.2) has a Hermite expansion of the form

$$\exp(b_\ell(x)y_\ell) = \sum_{v_\ell=0}^{\infty} c_{v_\ell}^{(\ell)}(x) H_{v_\ell}^{\tau_{\theta\varrho}}(y_\ell) \tag{4.3}$$

with

$$c_{v_\ell}^{(\ell)}(x) = \frac{(b_\ell(x)\sigma_\ell(\vartheta\varrho))^{v_\ell}}{\sqrt{v_\ell!}} \exp((b_\ell(x)\sigma_\ell(\vartheta\varrho))^2/2).$$

The procedure is as follows: First, we fix an adequate quadrature rule for solving the involved integrals by choosing quadrature points $\chi_q \in D$ and weights $w_q \in \mathbb{R}$ for $q = 1, \ldots, P_{\text{quad}}$. We begin the discretization at the right most side and define the correlation matrix

$$C_L(v_L, v'_L) := \sum_{q=1}^{P_{\text{quad}}} c_{v_L}^{(L)}(\chi_q) c_{v'_L}^{(L)}(\chi_q) w_q$$

$$\approx \int_D c_{v_L}^{(L)}(x) c_{v'_L}^{(L)}(x) \, dx$$

for $v_L, v'_L = 0, \ldots, q_L - 1$. This means that we have truncated the expansion (4.3) according to the tensor set Δ, which yields an approximation of the factors. This matrix is symmetric and positive semidefinite and it therefore admits an eigenvalue decomposition

$$C_L(v_L, v'_L) = \sum_{k_L=1}^{q_L} \lambda_{k_L} A_L(k_L, v_L) A_L(k_L, v'_L).$$

This yields reduced basis functions

$$\tilde{c}_{k_L}^{(L)}(\chi_q) := \sum_{v_L=0}^{q_L-1} A_L(k_L, v_L) c_{v_L}^{(L)}(\chi_q)$$

for $k_L = 1, \ldots, s_L$ that we can store explicitly at the quadrature points of the integral. If we choose $s_L = q_L$ then this is just a transformation without any reduction.

We proceed successively for $\ell = L - 1, \ldots, 1$ by defining correlation matrices

$$C_\ell(v_\ell, k_{\ell+1}, v'_L, k'_{\ell+1}) := \sum_{q=1}^{P_{\text{quad}}} c_{v_\ell}^{(\ell)}(\chi_q) \tilde{c}_{k_{\ell+1}}^{(\ell+1)}(\chi_q) c_{v'_\ell}^{(\ell)}(\chi_q) \tilde{c}_{k'_{\ell+1}}^{(\ell+1)}(\chi_q) w_q$$

Algorithm 1 Algorithm for coefficient splitting.

Input: Coefficients $c_{\nu_1}^{(1)}, \ldots, c_{\nu_L}^{(L)}$, $\nu_\ell = 0, \ldots, q_\ell - 1$, $\ell = 1, \ldots, L$;
                   ranks $s_1, \ldots, s_L$;
                   quadrature rule (χ_q, w_q), $q = 1, \ldots, P_{\text{quad}}$;
Output: TT Tensor components $A_1, \ldots, A_L$;

Init $s_{L+1} = 1$; $\tilde{c}_1^{(L+1)} \equiv 1$
for $\ell = L, \ldots, 1$ **do**
    Arrange correlation matrix C_ℓ:

$$C_\ell(\nu_\ell, k_{\ell+1}, \nu'_\ell, k'_{\ell+1}) := \sum_{q=1}^{P_{\text{quad}}} c_{\nu_\ell}^{(\ell)}(\chi_q) \tilde{c}_{k_{\ell+1}}^{(\ell+1)}(\chi_q) c_{\nu'_\ell}^{(\ell)}(\chi_q) \tilde{c}_{k'_{\ell+1}}^{(\ell+1)}(\chi_q) w_q;$$

Compute eigenvalue decomposition:

$$C_\ell(\nu_\ell, k_{\ell+1}, \nu'_\ell, k'_{\ell+1}) = \sum_{k_\ell=1}^{s_\ell} \lambda_{k_\ell} A_\ell(k_\ell, \nu_\ell, k_{\ell+1}) A_\ell(k_\ell, \nu'_\ell, k'_{\ell+1});$$

Set reduced basis functions:

$$\tilde{c}_{k_\ell}^\ell(\chi_q) = \sum_{\nu_\ell=0}^{q_\ell-1} \sum_{k_{\ell+1}=1}^{s_{\ell+1}} A_\ell(k_\ell, \nu_\ell, k_{\ell+1}) c_{\nu_\ell}^\ell(\chi_q) \tilde{c}_{k_{\ell+1}}^{\ell+1}(\chi_q);$$

end for
return $A_1, \ldots, A_L$;

---

with eigenvalue decompositions

$$C_\ell(\nu_\ell, k_{\ell+1}, \nu'_L, k'_{\ell+1}) = \sum_{k_\ell=1}^{q_\ell} \lambda_{k_\ell} A_\ell(k_\ell, \nu_\ell, k_{\ell+1}) A_\ell(k_\ell, \nu'_\ell, k'_{\ell+1})$$

and the resulting reduced basis functions at the quadrature points

$$\tilde{c}_{k_\ell}(\ell)(\chi_p) = \sum_{\nu_\ell=0}^{q_\ell-1} \sum_{k_{\ell+1}=1}^{s_\ell} A_\ell(k_\ell, \nu_\ell, k_{\ell+1}) c_{\nu_\ell}^{(\ell)}(\chi_q) \tilde{c}_{k_{\ell+1}}^{(\ell+1)}(\chi_q),$$

see Algorithm 1.

This results in a first component

$$a_0[k_1](\chi_q) := \tilde{c}_{k_1}^{(1)}(\chi_q)$$

for $k_1 = 1, \ldots, s_1$. Note that on the one hand, it is possible to evaluate this component at any point $x \in D$ by converting the reduced basis functions back into their original form by means of the tensor components A_ℓ and the coefficient functions $c_{\nu_\ell}^{(\ell)}$. More
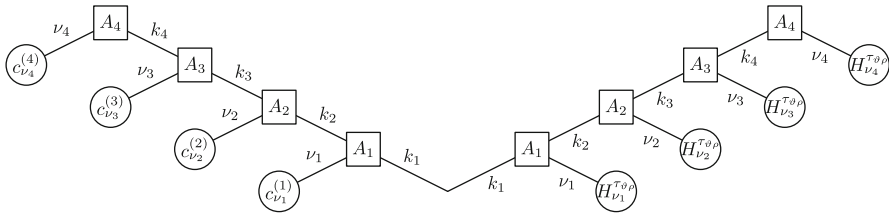
Fig. 1 A coefficient splitting for $L = 4$

specifically,

$$a_0[k_1](x) = \sum_{v_1=0}^{q_1} \sum_{k_2=1}^{s_2} A_1(k_1, v_1, k_2) c_{v_1}^{(1)}(x) \tilde{c}_{k_2}^{(2)}(x)$$

$$= \dots$$

$$= \sum_{k_2=1}^{s_2} \cdots \sum_{k_L=1}^{s_L} \prod_{\ell=1}^{L} \left(\sum_{v_\ell=0}^{q_\ell-1} A_\ell(k_\ell, v_\ell, k_\ell) c_{v_\ell}^{(\ell)}(x) \right).$$

On the other hand, each original coefficient function is approximated by the reduced basis representation

$$c_{v_L}^{(L)}(x) \approx \sum_{k_L=1}^{s_L} A_L(k_L, v_L) \tilde{c}_{k_L}^{(L)}(x),$$

$$c_{v_\ell}^{(\ell)}(x) \tilde{c}_{k_{\ell+1}}^{(\ell+1)} \approx \sum_{k_\ell=1}^{s_\ell} A_\ell(k_\ell, v_\ell, k_{\ell+1}) \tilde{c}_{k_\ell}^{(\ell)} \qquad \text{for all } \ell = L - 1, \dots, 1.$$

This approximation is exact if the ranks $s = (s_1, \dots, s_L)$ are full.

By the described procedure we obtain an approximate discretization $a_{\Delta,s} \approx a$ in a TT-like format that is continuous in the first component and that has the decomposition

$$a_{\Delta,s}(x, y) = \sum_{k_1=1}^{s_1} \cdots \sum_{k_L=1}^{s_L} a_0[k_1](x) \left(\sum_{v \in \Delta} \prod_{\ell=1}^{L} A_\ell(k_\ell, v_\ell, k_{\ell+1}) H_{v_\ell}^{\tau_{\theta\varrho}}(y_\ell) \right), \quad (4.4)$$

see Fig. 1.

On $\mathcal{V}_N(\Lambda; \mathcal{T}, p)$ the linear Galerkin operator $\mathbf{A}$ is in the TT matrix or matrix product operator (MPO) format:

$$\mathbf{A}(i, \mu, j, \mu') = \sum_{k_1=1}^{s_1} \cdots \sum_{k_M=1}^{s_M} \mathbf{A}_0(i, j, k_1) \prod_{m=1}^{M} \mathbf{A}_m(k_m, \mu_m, \mu'_m, k_{m+1}), \quad (4.5)$$

where

$$\mathbf{A}_0(i, j, k_1) = \int_D a_0[k_1](x) \nabla \varphi_i \cdot \nabla \varphi_j \, dx$$

and for all $m = 1, \ldots, M - 1$

$$\mathbf{A}_m(k_m, \mu_m, \mu_m', k_{m+1}) = \sum_{v_m=0}^{q_m-1} A_m(k_m, v_m, k_{m+1})$$

$$\times \int_{\mathbb{R}} H_{v_m}^{\tau_{\theta\varrho}} H_{\mu_m}^{\tau_{\theta\varrho}} H_{\mu_m'}^{\tau_{\theta\varrho}} \, d\gamma_{\theta\varrho,m}(y_m)$$

$$= \sum_{v_m=0}^{q_m-1} A_m(k_m, v_m, k_{m+1}) \kappa_{\mu_m,\mu_m',v_m}$$

and

$$\mathbf{A}_M(k_m, \mu_M, \mu_M') = \sum_{v_M=0}^{q_M-1} \sum_{k_{m+1}=1}^{s_{m+1}} \cdots \sum_{k_L=1}^{s_L} A_M(k_m, v_M, k_{m+1}) \kappa_{\mu_M,\mu_M',v_M}$$

$$\times \prod_{\ell=m+1}^{L} A_\ell(k_\ell, 0, k_{\ell+1}).$$

Since the integral over the triple product $\kappa_{\mu_m,\mu_m',v_m} = 0$ for all $v_m > 2 \max(\mu_m, \mu_m')$, it is sufficient to set $q_\ell = 2d_\ell - 1$ for all $\ell = 1, \ldots, L$. If the rank s of the decomposition of the coefficient is full, then the discretised coefficient $a_{\Delta,s}$ is exact *on the discrete space* $\mathcal{V}_N$ (up to quadrature errors).

However, this is generally infeasible as the rank would grow exponentially with M. Therefore, a truncation of the rank becomes necessary and the coefficient is only an approximation. We assume in the following that the error that is due to this approximation of the coefficient is small. A thorough estimation of this error is subject to future research.

Remark 4.1 A similar approach to decomposing the coefficient has been chosen in [23] where the knowledge of the eigenfunctions of the covariance operator was assumed a priori. This means that one has an orthogonal basis also for the deterministic part in x and all that remains to do is to decompose the coefficient tensor for this basis representation. This is also done using some quadrature and the L^2-error of this approximation can be estimated.

According to the discretization of the coefficient, we introduce the splitting of the operator,

$$\mathcal{A}(v) = \mathcal{A}_+(v) + \mathcal{A}_-(v),$$

with the active and inactive operators $\mathcal{A}_+ : \mathcal{V}_N(\Lambda; \mathcal{T}, p) \to \mathcal{V}_N(\Lambda; \mathcal{T}, p)^*$ and $\mathcal{A}_- : \mathcal{V}_N(\Lambda; \mathcal{T}, p) \to \mathcal{V}_N(\Lambda; \mathcal{T}, p)^*$ defined by

$$\begin{aligned} \mathcal{A}_+(v) &:= -\operatorname{div}(a_{\Delta,s}\nabla v), \\ \mathcal{A}_-(v) &:= -\operatorname{div}\left((a - a_{\Delta,s})\nabla v\right), \end{aligned} \tag{4.6}$$

for $v \in \mathcal{V}_N(\Lambda; \mathcal{T}, p)$.

The above considerations yield that the approximate operator $\mathcal{A}_+$ can be represented by the following tensor product structure

$$\mathcal{A}_+ = \sum_{k_1=1}^{s_1} \cdots \sum_{k_\ell=1}^{s_L} \mathcal{A}_0[k_1] \otimes \mathcal{A}_1[k_1, k_2] \otimes \cdots \otimes \mathcal{A}_L[k_L],$$

with the operator components

$$\mathcal{A}_0[k_1] : \mathcal{X} \to \mathcal{X}^*, \qquad \mathcal{A}_0[k_1](v_x) = -\operatorname{div}(a_0[k_1]\nabla v_x),$$

and for all $\ell = 1, \ldots, L$,

$$\mathcal{A}_\ell[k_\ell, k_{\ell+1}] : L^2(\mathbb{R}, \gamma_{\vartheta\varrho,m}) \to L^2(\mathbb{R}, \gamma_{\vartheta\varrho,m})$$

$$\mathcal{A}_\ell[k_\ell, k_{\ell+1}](v_y) = \sum_{v_\ell=0}^{q_\ell-1} A_\ell(k_\ell, v_\ell, k_{\ell+1}) H_{v_\ell}^{\tau_{\theta\varrho}} v_y.$$

5 Error estimates

This section is concerned with the derivation of a reliable a posteriori error estimator based on the stochastic residual. In comparison to the derivation in [15,16,21], the log-normal coefficient requires a more involved approach directly related to the employed weighted function spaces introduced in Sect. 3. In theory, an additional error occurs because of the discretization of the coefficient which we assume to be negligible. The developed adaptive algorithm makes possible a computable a posteriori steering of the error components by a refinement of the FE mesh and the anisotropic Hermite polynomial chaos of the solution. The efficient implementation is due to the formulation in the TT format the ranks of which are also set adaptively.

The definition of the operators as in (4.6) leads to a decomposition of the residual

$$\mathcal{R}(v) = \mathcal{R}_+(v) + \mathcal{R}_-(v),$$

with

$$\mathcal{R}_+(v) := f - \mathcal{A}_+(v), \qquad \mathcal{R}_-(v) := -\mathcal{A}_-(v).$$

The discrete solution $w_N \in \mathcal{V}_N$ reads

$$w_N = \sum_{i=0}^{N-1} \sum_{\mu \in \Lambda} W(i, \mu) \varphi_i H_\mu^{\tau_{\theta\varrho}}.$$

We assume that the operator is given in its approximate semi-discrete form $\mathcal{A}_+$ and aim to estimate the energy error

$$\|u - w_N\|_{\mathcal{A}_+}^2 = \int_\Gamma \int_D a_{\Delta,s} |\nabla(u - w_N)|^2 \, dx d\gamma_{\vartheta_\varrho}(y).$$

Remark 5.1 As stated before, we assume that the error that results from approximating the coefficient is small. Estimation of this error is subject to future research. Work in this direction has e.g. been carried out in [7,23]. Additionally, we require that the bounds (3.2) and (3.3) still hold, possibly with different constants $\hat{c}_{\vartheta_\varrho}^+$ and $\check{c}_{\vartheta_\varrho}^+$. This is for example guaranteed if $a_{\Delta,s}$ is positive, i.e., if

$$a_{\Delta,s}(x, y) > 0 \quad \forall x \in D, y \in \Gamma.$$

Then, since the approximated coefficient is polynomial in y, the arguments in Lemma 3.5 yield the same constants

$$\hat{c}_{\vartheta_\varrho}^+ = \hat{c}_{\vartheta_\varrho}, \qquad \check{c}_{\vartheta_\varrho}^+ = \check{c}_{\vartheta_\varrho}.$$

We recall Theorem 5.1 from [15] and also provide the proof for the sake of a complete presentation. Note that the result allows for non-orthogonal approximations $w_N \in \mathcal{V}_N$.

Theorem 5.2 *Let $\mathcal{V}_N \subset \mathcal{V}_{\vartheta_\varrho}$ a closed subspace and $w_N \in \mathcal{V}_N$, and let u_N denote the $\mathcal{A}_+$ Galerkin projection of u onto $\mathcal{V}_N$. Then it holds*

$$\|u - w_N\|_{\mathcal{A}_+}^2 \leq \left(\sup_{v \in \mathcal{V}_{\theta_\varrho} \setminus \{0\}} \frac{|\langle \mathcal{R}_+(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta_\varrho}|}{\check{c}_{\theta_\varrho}^+ \|v\|_{L^2(\Gamma, \gamma; \mathcal{X})}} + c_\mathcal{I} \|u_N - w_N\|_{\mathcal{A}_+} \right)^2$$
$$+ \|u_N - w_N\|_{\mathcal{A}_+}^2.$$

Here, $\mathcal{I}$ denotes the Clément interpolation operator in (3.4) and $c_\mathcal{I}$ is the operator norm of $\mathrm{id} - \mathcal{I}$ with respect to the energy norm $\|\cdot\|_{\mathcal{A}_+}$. The constant $\check{c}_{\theta_\varrho}^+$ is derived from the assumed coercivity of the bilinear form induced by $\mathcal{A}_+$ similar to (3.2) and (3.3).

Proof Due to Galerkin orthogonality of u_N, it holds

$$\|u - w_N\|_{\mathcal{A}_+}^2 = \|u - u_N\|_{\mathcal{A}_+}^2 + \|u_N - w_N\|_{\mathcal{A}_+}^2.$$

By the Riesz representation theorem, the first part is

$$\|u - u_N\|_{\mathcal{A}_+} = \sup_{v \in \mathcal{V}_{\theta\varrho} \backslash \{0\}} \frac{|\langle \mathcal{R}_+(u_N), v \rangle_{\theta\varrho}|}{\|v\|_{\mathcal{A}_+}}.$$

We now utilise the Galerkin orthogonality and introduce the bounded linear map $\mathcal{I} : \mathcal{V}_{\theta\varrho} \to \mathcal{V}_N$ to obtain

$$\|u - u_N\|_{\mathcal{A}_+} = \sup_{v \in \mathcal{V}_{\theta\varrho} \backslash \{0\}} \frac{|\langle \mathcal{R}_+(u_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}|}{\|v\|_{\mathcal{A}_+}}.$$

Since we do not have access to the Galerkin solution u_N, we reintroduce w_N

$$
\begin{aligned}
\|u - u_N\|_{\mathcal{A}_+} &\leq \sup_{v \in \mathcal{V}_{\theta\varrho} \backslash \{0\}} \frac{|\langle \mathcal{R}_+(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}|}{\|v\|_{\mathcal{A}_+}} \\
&\quad + \frac{|\langle \mathcal{R}_+(u_N) - \mathcal{R}_+(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}|}{\|v\|_{\mathcal{A}_+}} \\
&\leq \sup_{v \in \mathcal{V}_{\theta\varrho} \backslash \{0\}} \frac{|\langle \mathcal{R}_+(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}|}{\|v\|_{\mathcal{A}_+}} \\
&\quad + \frac{\|u_N - w_N\|_{\mathcal{A}_+} \|(\mathrm{id} - \mathcal{I})v\|_{\mathcal{A}_+}}{\|v\|_{\mathcal{A}_+}} \\
&\leq \sup_{v \in \mathcal{V}_{\theta\varrho} \backslash \{0\}} \frac{|\langle \mathcal{R}_+(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}|}{\|v\|_{\mathcal{A}_+}} + c_{\mathcal{I}} \|w_N - u_N\|_{\mathcal{A}_+}.
\end{aligned}
$$

We apply the coercivity of the operator $\mathcal{A}_+$ to the denominator, which yields the desired result. For the last inequality, we used the boundedness of $\mathcal{I}$ in the energy norm by defining the constant as the operator norm

$$c_{\mathcal{I}} := \sup_{v \in \mathcal{V}_{\theta\varrho} \backslash \{0\}} \frac{\|(\mathrm{id} - \mathcal{I})v\|_{\mathcal{A}_+}}{\|v\|_{\mathcal{A}_+}}.$$

$\square$

Since the product of the Hermite polynomials for each $m = 1, \ldots, M$ has degree at most $q_m + d_m - 2$, it is useful to define the index set

$$
\begin{aligned}
\varXi := \varDelta + \varLambda := \big\{ \eta = (\eta_1, \ldots, \eta_L, 0, \ldots) : \\
\eta_m = 0, \ldots, q_m + d_m - 2, \; m = 1, \ldots, M; \\
\eta_\ell = 0, \ldots, q_\ell - 1, \; \ell = M + 1, \ldots, L \big\}.
\end{aligned}
$$

Then, the residual can be split into an active and an inactive part by using the tensor sets $\varXi$ and $\varLambda$,

$$\mathcal{R}_+(w_N) = f - \mathcal{A}_+(w_N)$$

$$= f + \sum_{\eta \in \Xi} \operatorname{div} \operatorname{res}(\cdot, \eta) H_\eta^{\tau_{\theta\varrho}}$$

$$= \mathcal{R}_{+,\Lambda}(w_N) + \mathcal{R}_{+,\Xi\setminus\Lambda}(w_N),$$

with

$$\mathcal{R}_{+,\Lambda}(w_N) = f + \sum_{\eta \in \Lambda} \operatorname{div} \operatorname{res}(\cdot, \eta) H_\eta^{\tau_{\theta\varrho}},$$

$$\mathcal{R}_{+,\Xi\setminus\Lambda}(w_N) = \sum_{\eta \in \Xi\setminus\Lambda} \operatorname{div} \operatorname{res}(\cdot, \eta) H_\eta^{\tau_{\theta\varrho}},$$

where $\operatorname{div} \operatorname{res}(\cdot, \eta) \in \mathcal{X}^*$ for all $\eta \in \Xi$.

For all $\eta \in \Xi$, the function res is given as

$$\operatorname{res}(x, \eta) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_M=1}^{r_M} \sum_{k_1'=1}^{s_1} \cdots \sum_{k_L'=1}^{s_L} \operatorname{res}_0[k_1, k_1'](x)$$

$$\times \left(\prod_{m=1}^{M} R_m(k_m, k_m', \eta_m, k_{m+1}, k_{m+1}') \prod_{\ell=M+1}^{L} A_\ell(k_\ell', \eta_\ell, k_{\ell+1}') \right)$$

with continuous first component

$$\operatorname{res}_0[k_1, k_1'](x) = \sum_{i=0}^{N-1} a_0[k_1'](x) W_0(i, k_1) \nabla \varphi_i(x)$$

and stochastic components for $m = 1, \ldots, M$,

$$R_m(k_m, k_m', \eta_m, k_{m+1}, k_{m+1}')$$
$$= \sum_{v_m=0}^{q_m-1} \sum_{\mu_m=0}^{d_m-1} A(k_m', v_m, k_{m+1}') W_m(k_m, \mu_m, k_{m+1}) \kappa_{\mu_m, v_m, \eta_m}.$$

The function res is again a TT tensor with continuous first component with TT ranks $r_m s_m$ for $m = 1, \ldots, M$ and s_ℓ for $\ell = M + 1, \ldots, L$. The physical dimensions are $d_m + q_m - 2$ for all $m = 1, \ldots, M$ and $d_\ell - 1$ for $\ell = M + 1, \ldots, L$.

The above considerations suggest that the error can be decomposed into errors that derive from the respective approximations in the deterministic domain, the parametric domain and in the ranks. This is indeed the case, as we will see in the following. In a nutshell, if u_N is the Galerkin solution in $\mathcal{V}_N$ and u_Λ is the Galerkin solution in the semi-discrete space $\mathcal{V}(\Lambda)$, then the *deterministic error* $\operatorname{err}_{\mathrm{det}} = \|u_\Lambda - u_N\|_{\mathcal{A}_+}$ corresponds to the error of the *active residual* $\mathcal{R}_{+,\Lambda}$, the *parametric error* $\operatorname{err}_{\mathrm{param}} = \|u - u_\Lambda\|_{\mathcal{A}_+}$ corresponds to the *inactive residual* $\mathcal{R}_{+,\Xi\setminus\Lambda}$ and the error made by restricting the ranks is the error in the discrete space $\operatorname{err}_{\mathrm{disc}}(w_N) = \|u_N - w_N\|_{\mathcal{A}_+}^2$, see Fig. 2 for an illustration.
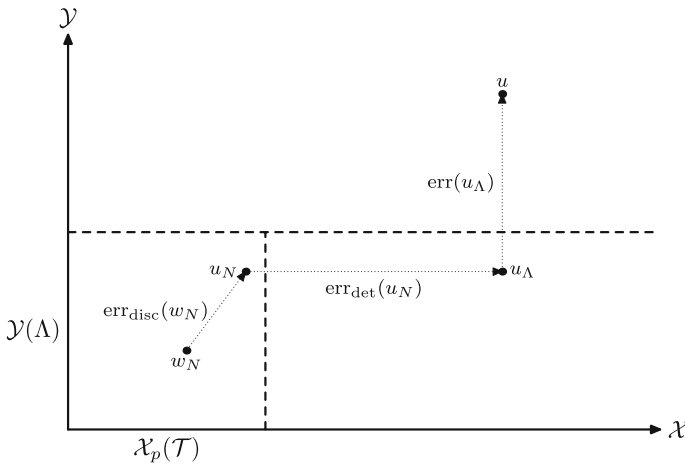
Fig. 2 An illustration of the different errors

5.1 Deterministic error estimation

We define the *deterministic error estimator*

$$\mathrm{est}_{\mathrm{det}}(w_N) := \left(\sum_{T \in \mathcal{T}} \mathrm{est}^2_{\mathrm{det},T}(w_N) + \sum_{F \in \mathcal{S}} \mathrm{est}^2_{\mathrm{det},F}(w_N) \right)^{1/2},$$

$$\mathrm{est}_{\mathrm{det},T}(w_N) := h_T \| f + \mathrm{div}\, \sigma^{\theta\varrho}_{\Lambda}(w_N) \zeta_{\vartheta\varrho} \|_{L^2(\Gamma,\gamma;L^2(T))},$$

$$\mathrm{est}_{\mathrm{det},F}(w_N) := h_F^{1/2} \| [\![\sigma^{\theta\varrho}_{\Lambda}(w_N)]\!]_F \zeta_{\vartheta\varrho} \|_{L^2(\Gamma,\gamma;L^2(F))},$$

where the flow is given by the residual contributions

$$\sigma^{\theta\varrho}_{\Lambda} := \sum_{\eta \in \Lambda} \mathrm{res}(\cdot, \eta) H_{\eta}^{\tau_{\theta\varrho}}.$$

This estimates the active residual as follows.

Proposition 5.3 *For any $v \in \mathcal{V}_{\vartheta\varrho}$ and any $w_N \in \mathcal{V}_N$, it holds*

$$\frac{|\langle \mathcal{R}_{+,\Lambda}(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}|}{\|v\|_{L^2(\Gamma,\gamma;\mathcal{X})}} \le c_{\mathrm{det}}\, \mathrm{est}_{\mathrm{det}}(w_N).$$

Proof By localization to the elements of the triangulation $\mathcal{T}$ and integration by parts,

$$\langle \mathcal{R}_{+,\Lambda}(w_N), (\mathrm{id} - \mathcal{I})v \rangle_{\theta\varrho}$$
$$= \int_{\Gamma} \sum_{T \in \mathcal{T}} \int_T f\big((\mathrm{id} - \mathcal{I})v\big) - a_{\Delta,s} \nabla w_N \cdot \nabla\big((\mathrm{id} - \mathcal{I})v\big) \, \mathrm{d}x \mathrm{d}\gamma_{\vartheta\varrho}(y)$$

$$= \sum_{T \in \mathcal{T}} \int_\Gamma \int_T \left(f + \operatorname{div} \sigma_\Lambda^{\theta\varrho}(w_N)\right)\left((\operatorname{id} - \mathcal{I})v\right) \zeta_{\vartheta\varrho}(y) \, \mathrm{d}x \mathrm{d}\gamma(y)$$

$$+ \sum_{F \in \mathcal{S}} \int_\Gamma \int_F [\![\sigma_\Lambda^{\theta\varrho}(w_N)]\!]_F \left((\operatorname{id} - \mathcal{I})v\right) \zeta_{\vartheta\varrho}(y) \mathrm{d}(x) \mathrm{d}\gamma(y).$$

The Cauchy-Schwarz inequality yields

$$|\langle \mathcal{R}_{+,\Lambda}(w_N), (\operatorname{id} - \mathcal{I})v \rangle_{\theta\varrho}|$$
$$\leq \sum_{T \in \mathcal{T}} \|(f + \operatorname{div} \sigma_\Lambda^{\theta\varrho}(w_N))\zeta_{\vartheta\varrho}\|_{L^2(\Gamma,\gamma;L^2(T))} \|(\operatorname{id} - \mathcal{I})v\|_{L^2(\Gamma,\gamma;L^2(T))}$$
$$+ \sum_{F \in \mathcal{S}} \|[\![\sigma_\Lambda^{\theta\varrho}(w_N)]\!]_F \zeta_{\vartheta\varrho}\|_{L^2(\Gamma,\gamma;L^2(F))} \|(\operatorname{id} - \mathcal{I})v\|_{L^2(\Gamma,\gamma;L^2(F))}.$$

With the interpolation properties (3.4) we obtain

$$|\langle \mathcal{R}_{+,\Lambda}(w_N), (\operatorname{id} - \mathcal{I})v \rangle_{\theta\varrho}|$$
$$\leq \sum_{T \in \mathcal{T}} h_T c_\mathcal{T} \|(f + \operatorname{div} \sigma_\Lambda^{\theta\varrho}(w_N))\zeta_{\vartheta\varrho}\|_{L^2(\Gamma,\gamma;L^2(T))} |v|_{\mathcal{V}_{\vartheta\varrho},\omega_T}$$
$$+ \sum_{F \in \mathcal{S}} h_F^{1/2} c_\mathcal{S} \|[\![\sigma_\Lambda^{\theta\varrho}(w_N)]\!]_F \zeta_{\vartheta\varrho}\|_{L^2(\Gamma,\gamma;L^2(F))} |v|_{\mathcal{V}_{\vartheta\varrho},\omega_F}.$$

Since the overlaps of the patches ω_T and ω_F are bounded uniformly, a Cauchy-Schwarz estimate leads to

$$|\langle \mathcal{R}_{+,\Lambda}(w_N), (\operatorname{id} - \mathcal{I})v \rangle_{\theta\varrho}| \leq c_{\det} \operatorname{est}_{\det}(w_N) \|v\|_{L^2(\Gamma,\gamma;\mathcal{X})}.$$

Here, the constant $c_{\det}$ depends on the properties of the interpolation operator (3.4).

□

Remark 5.4 Note that an L^2-integration of the residual, which is an element of the dual space $\mathcal{V}_{\vartheta\varrho}^*$, is possible since the solution consists of finite element functions. These are piecewise polynomial and thus smooth on each element $T \in \mathcal{T}$.

5.2 Tail error estimation

The *parametric or tail estimator* is given by

$$\operatorname{est}_{\operatorname{param}}(w_N) := \left(\int_\Gamma \int_D \left(\sum_{\eta \in \mathcal{E} \setminus \Lambda} \operatorname{res}(x,\eta) H_\eta^{\tau_{\theta\varrho}}(y) \zeta_{\vartheta\varrho}(y)\right)^2 \mathrm{d}x \mathrm{d}\gamma(y)\right)^{1/2}$$

and bounds the parametric error as follows.

Proposition 5.5 *For any $v \in \mathcal{V}_{\vartheta\varrho}$ and any $w_N \in \mathcal{V}_N$, it holds*

$$\frac{|\langle \mathcal{R}_{+,\Xi\setminus\Lambda}(w_N), (\mathrm{id}-\mathcal{I})v\rangle_{\theta\varrho}|}{\|v\|_{L^2(\Gamma,\gamma;\mathcal{X})}} \le \mathrm{est}_{\mathrm{param}}(w_N).$$

Proof. Recall that $\langle \mathcal{R}_{+,\Xi\setminus\Lambda}(w_N), \mathcal{I}v\rangle_{\theta\varrho} = 0$ since $\mathcal{I}v \in \mathcal{V}_N$.

Instead of factorizing out the L^∞-norm of the diffusion coefficient as in [15,16,21], we use the Cauchy-Schwarz inequality to obtain

$$\langle \mathcal{R}_{+,\Xi\setminus\Lambda}(w_N), v\rangle_{\theta\varrho}$$
$$= \int_\Gamma \int_D \Big(\sum_{\eta\in\Xi\setminus\Lambda} \mathrm{res}(x,\eta) H_\eta^{\tau_{\theta\varrho}}(y) \Big) \cdot \nabla v(x,y)\, \zeta_{\vartheta\varrho}(y)\, \mathrm{d}x\mathrm{d}\gamma(y)$$
$$\le \int_\Gamma \int_D \Big(\sum_{\eta\in\Xi\setminus\Lambda} \mathrm{res}(x,\eta) H_\eta^{\tau_{\theta\varrho}}(y)\, \zeta_{\vartheta\varrho}(y) \Big)^2 \mathrm{d}x\mathrm{d}\gamma(y) \|v\|_{L^2(\Gamma,\gamma;\mathcal{X})}$$
$$= \mathrm{est}_{\mathrm{param}}(w_N) \|v\|_{L^2(\Gamma,\gamma;\mathcal{X})}. \qquad \square$$

5.3 Algebraic error estimation

In order to define the *algebraic error estimator*, we need to state the linear basis change operator that translates integrals over two Hermite polynomials in the measure $\gamma_{\vartheta\varrho}$ to the measure γ:

$$\mathbf{H}_{\vartheta\varrho\to 0} : \mathbb{R}^{N\times d_1\times\cdots\times d_M} \to \mathbb{R}^{N\times d_1\times\cdots\times d_M},$$
$$\mathbf{H}_{\vartheta\varrho\to 0} := Z_0 \otimes Z_1 \otimes \cdots \otimes Z_M,$$
$$Z_0(i,j) := \int_D \nabla\varphi_i \cdot \nabla\varphi_j\, \mathrm{d}x,$$
$$Z_m(\mu_m,\mu'_m) := \int_{\mathbb{R}} H_{\mu_m}^{\tau_{\theta\varrho}}(y_m) H_{\mu'_m}^{\tau_{\theta\varrho}}(y_m)\, \mathrm{d}\gamma_m(y_m) \quad \text{for all } m = 1,\ldots,M.$$

This yields the estimator

$$\mathrm{est}_{\mathrm{disc}}(w_N) := \|(\mathbf{A}(W)-F)\mathbf{H}_{\vartheta\varrho\to 0}^{-1/2}\|_{\ell^2(\mathbb{R}^{N\times d_1\times\cdots\times d_M})}. \qquad (5.1)$$

Proposition 5.6 *For any $w_N \in \mathcal{V}_N$ and the Galerkin solution $u_N \in \mathcal{V}_N$, it holds*

$$\|u_N - w_N\|_{\mathcal{A}_+} \le (\check{c}_{\vartheta\varrho}^+)^{-1}\, \mathrm{est}_{\mathrm{disc}}(w_N).$$

Proof For $v_N = \sum_{i=0}^{N-1} \sum_{\mu\in\Lambda} V(i,\mu)\varphi_i H_\mu^{\tau_{\theta\varrho}} \in \mathcal{V}_N$, it holds

$$\int_\Gamma \int_D \nabla v_N \cdot \nabla v_N\, \mathrm{d}x\mathrm{d}\gamma(y) = \langle V\mathbf{H}_{\vartheta\varrho\to 0}, V\rangle = \|V\mathbf{H}_{\vartheta\varrho\to 0}^{1/2}\|_{\ell^2(\mathbb{R}^{N\times d_1\times\cdots\times d_M})}^2.$$

With this and using the coercivity of $\mathcal{A}_+$, we can see that

$$
\begin{aligned}
\|w_N - u_N\|_{\mathcal{A}_+}^2 &= \langle \mathcal{A}_+(w_N - u_N), (w_N - u_N)\rangle_{\theta\varrho} \\
&= \langle \mathbf{A}W - F, W - U\rangle \\
&= \langle (\mathbf{A}W - F)\mathbf{H}_{\vartheta\varrho\to 0}^{-1/2}, (W - U)\mathbf{H}_{\vartheta\varrho\to 0}^{1/2}\rangle \\
&\leq \|(\mathbf{A}W - F)\mathbf{H}_{\vartheta\varrho\to 0}^{-1/2}\|_{\ell^2(\mathbb{R}^{N\times d_1\times\cdots\times d_M})} \|w_N - u_N\|_{L^2(\Gamma,\gamma;\mathcal{X})} \\
&\leq (\check{c}_{\vartheta\varrho}^+)^{-1} \|(\mathbf{A}W - F)\mathbf{H}_{\vartheta\varrho\to 0}^{-1/2}\|_{\ell^2(\mathbb{R}^{N\times d_1\times\cdots\times d_M})} \|w_N - u_N\|_{\mathcal{A}_+}
\end{aligned}
$$

and thus

$$
\|w_N - u_N\|_{\mathcal{A}_+} \leq (\check{c}_{\vartheta\varrho}^+)^{-1} \operatorname{est_{disc}}(w_N).
$$

$\square$

5.4 Overall error estimation

A combination of the above estimates yields an overall error estimator.

Corollary 5.7 *For any* $w_N \in \mathcal{V}_N$, *the energy error can be bounded by*

$$
\|u - w_N\|_{\mathcal{A}_+}^2 \leq (\check{c}_{\vartheta\varrho}^+)^{-2} \operatorname{est_{all}}(w_N)^2
$$

with the error estimator given by

$$
\begin{aligned}
\operatorname{est_{all}}(w_N)^2 &:= \Big(c_{\det} \operatorname{est_{det}}(w_N) + \operatorname{est_{param}}(w_N) + c_{\mathcal{I}} \operatorname{est_{disc}}(w_N)\Big)^2 \\
&\quad + \operatorname{est_{disc}}(w_N)^2.
\end{aligned}
$$

Remark 5.8 In order to get suitable measures for the estimators, the squared density $\zeta_{\vartheta\varrho}^2$ appears, which upon scaling with

$$
c_\sigma := \prod_{\ell=1}^{L} \frac{1}{\sigma_\ell \sqrt{2 - \sigma_\ell^2}}
$$

again is a Gaussian measure with standard deviation $\sigma' = (\sigma_\ell')_{1\leq\ell\leq L}$ for

$$
\sigma_\ell' := \frac{\sigma_\ell}{\sqrt{2 - \sigma_\ell^2}}.
$$

First, this adds a restriction on ϑ such that the argument in the square root is positive. This is fulfilled if $\exp(2\vartheta\varrho\alpha_1) < 2$, since $(\alpha_m)_{1\leq m\leq M}$ is a decreasing series, which can be ensured for some ϑ small enough. Second, it is important to check whether the

new measure is *weaker* or *stronger* than $\gamma_{\vartheta\varrho}$ [49], i.e., which space contains the other. Since

$$\sigma'_\ell = \frac{\sigma_\ell}{\sqrt{2 - \sigma_\ell^2}} = \frac{\exp(\vartheta\varrho\alpha_\ell)}{\sqrt{2 - \exp(2\vartheta\varrho\alpha_\ell)}} \geq \exp(\vartheta\varrho\alpha_\ell) = \sigma_\ell,$$

functions that are integrable with respect to the measure $\gamma_{\vartheta\varrho}$ are not necessarily integrable with respect to the squared measure. However, since f is independent of the parameters and $\mathcal{A}_+(w_N) \in \mathcal{X}^* \otimes \mathcal{Y}(\Xi)$ has a polynomial chaos expansion of finite degree, the residual $\mathcal{R}_+(w_N)$ is integrable over the parameters for any Gaussian measure and therefore it is also integrable with respect to the squared measure.

5.5 Efficient computation of the different estimators

The error estimators can be calculated efficiently in the TT format. For each element $T \in \mathcal{T}$ of the triangulation, the residual estimator is given by

$$\begin{aligned}
\text{est}_{\text{det},T}(w_N)^2 &= h_T^2 \|(f + \text{div}\,\sigma_\Lambda^{\theta\varrho}(w_N))\zeta_{\vartheta\varrho}\|_{L^2(\Gamma,\gamma;L^2(T))}^2 \\
&= h_T^2 \int_\Gamma \int_T \left(f + \sum_{\eta \in \Lambda} \text{div}\,\text{res}(x,\eta) H_\eta^{\tau_{\theta\varrho}} \right)^2 \zeta_{\vartheta\varrho}^2 \, dx \, d\gamma(y) \\
&= h_T^2 (f,f)_{L^2(T)} \int_\Gamma \zeta_{\vartheta\varrho}^2 \, d\gamma(y) \\
&\quad + 2h_T^2 \sum_{\eta \in \Lambda} (f, \text{div}\,\text{res}(x,\eta))_{L^2(T)} \int_\Gamma H_\eta^{\tau_{\theta\varrho}} \zeta_{\vartheta\varrho}^2 d\gamma(y) \\
&\quad + \sum_{\eta \in \Lambda} \sum_{\eta' \in \Lambda} (\text{div}\,\text{res}(x,\eta), \text{div}\,\text{res}(x,\eta'))_{L^2(T)} \\
&\quad \times \int_\Gamma H_\eta^{\tau_{\theta\varrho}} H_{\eta'}^{\tau_{\theta\varrho}} \zeta_{\vartheta\varrho}^2 d\gamma(y).
\end{aligned}$$

A complication of the change of the measure to γ and the involved weight $\zeta_{\vartheta\varrho}^2$ is the fact that the shifted Hermite polynomials $H^{\tau_{\theta\varrho}}$ are not orthogonal with respect to this measure. However, this property can be restored easily by calculating the basis change integrals beforehand. This results in another tensor product operator that is defined element-wise for $\eta, \eta' \in \Xi$ by

$$\begin{aligned}
\tilde{\mathbf{H}}(\eta,\eta') &:= \tilde{Z}_1(\eta_1,\eta_1') \cdots \tilde{Z}_L(\eta_L,\eta_L'), \\
\tilde{Z}_\ell(\eta_\ell,\eta_\ell') &:= \int_\Gamma H_{\eta_\ell}^{\tau_{\theta\varrho}} H_{\eta_\ell'}^{\tau_{\theta\varrho}} \zeta_{\vartheta\varrho,\ell}^2 d\gamma(y_\ell).
\end{aligned}$$

This operator encodes the basis change to the squared measure and can be inserted in order to calculate the scalar product. With this, the estimator takes the form

$$\text{est}_{\text{det},T}(w_N)^2 = h_T^2(f, f)_{L^2(T)} \int_\Gamma \zeta_{\vartheta\varrho}^2 \, d\gamma(y)$$
$$+ 2h_T^2 \sum_{\eta \in \Lambda} \tilde{\mathbf{H}}(\eta, 0)(f, \text{div res}(x, \eta))_{L^2(T)}$$
$$+ \sum_{\eta \in \Lambda} \sum_{\eta' \in \Lambda} \tilde{\mathbf{H}}(\eta, \eta')(\text{div res}(x, \eta), \text{div res}(x, \eta'))_{L^2(T)}.$$

Since $\tilde{\mathbf{H}}$ is a tensor product operator, this summation can be done component-wise, i.e., performing a matrix-vector multiplication of every component of the operator $\tilde{\mathbf{H}}$ with the corresponding component of the tensor function r.

Similarly, for the jump over the edge F we obtain the estimator

$$\text{est}_{\text{det},F}(w_N)^2 = h_F \sum_{\eta \in \Lambda} \sum_{\eta' \in \Lambda} \tilde{\mathbf{H}}(\eta, \eta')([\![\text{res}(x, \eta)]\!]_F, [\![\text{res}(x, \eta')]\!]_F)_{L^2(F)}.$$

Analogously to the affine case dealt with in [21], both of these estimators can then be computed efficiently in the TT format. The parametric error estimator $\text{est}_{\text{param}}(w_N)$ can be estimated in a similar way.

To gain additional information about the residual influence of certain stochastic dimensions, we sum over specific index sets. Let

$$\Xi_m := \{(\eta_1, \ldots, d_m, \ldots, \eta_M, 0, \ldots) \in \mathcal{F} : x\eta_\ell = 0, \ldots, d_\ell - 1,$$
$$\ell = 1, \ldots, \not{m}, \ldots, M\},$$

where the strike through means that ℓ takes all values but m. For every $m = 1, 2, \ldots$, and $w_N \in \mathcal{V}_N$ we define

$$\text{est}_{\text{param},m}(w_N)^2 := \int_\Gamma \int_D \zeta_{\vartheta\varrho}^2 \left| \sum_{\eta \in \Xi_m} \text{res}(x, \eta) H_\eta^{\tau_{\theta\varrho}} \right|^2 \, dx \, d\gamma(y).$$

Using the same arguments and notation as above, we can simplify

$$\text{est}_{\text{param},m}(w_N)^2 = \int_D \sum_{\eta \in \Xi_n} \left(\text{res}(x, \eta) \cdot \sum_{\eta' \in \Xi_m} \tilde{\mathbf{H}}(\eta, \eta') \text{res}(x, \eta') \right) \, dx.$$

These operations, including the calculation of the discrete error estimator (5.1), can be executed efficiently in the TT format.

6 Fully adaptive algorithm

With the derived error estimators of the preceding sections, it is possible to refine all discretization parameters accordingly. As discussed before, the deterministic estimator assesses the error that arises from the finite element method. The discrete error estimator evaluates the error made by a low rank approximation. The rest of the error is estimated by the parametric error estimator.

The adaptive algorithm described in this section is similar to the algorithms presented in [15,16,21]. Given some mesh $\mathcal{T}$, a fixed polynomial degree p, a finite tensor set $\Lambda \subset \mathcal{F}$, and a start tensor W with TT rank r, we assume that a numerical approximation $w_N \in \mathcal{V}_N$ is obtained by a function

$$w_N^+ \leftarrow \mathrm{Solve}[\Lambda, \mathcal{T}, r, W]$$

where the superscript $+$ always denotes the updated object. In our implementation, we used the preconditioned ALS algorithm but other solution algorithms are feasible as well. The lognormal error indicators from Sect. 5 and thus the overall upper bound $\mathrm{est}_{\mathrm{all}}(w_N)$ in Corollary 5.7 are computed by the methods

$$
\begin{aligned}
(\mathrm{est}_{\det,T}(w_N))_{T\in\mathcal{T}}, \mathrm{est}_{\det}(w_N) &\leftarrow \mathrm{Estimate}_x[w_N, \Lambda, \mathcal{T}, p],\\
(\mathrm{est}_{\mathrm{param},m}(w_N))_{m\in\mathbb{N}}, \mathrm{est}_{\mathrm{param}}(w_N) &\leftarrow \mathrm{Estimate}_y[w_N, \Lambda],\\
\mathrm{est}_{\mathrm{disc}}(w_N) &\leftarrow \mathrm{Estimate}_{\mathrm{ALS}}[w_N],\\
\mathrm{est}_{\mathrm{all}}(w_N) &\leftarrow \mathrm{Estimate}_{\mathrm{all}}[\mathrm{est}_{\det}, \mathrm{est}_{\mathrm{param}}, \mathrm{est}_{\mathrm{disc}}].
\end{aligned}
$$

Depending on which error is largest, either the mesh is refined, or the index set Λ is enlarged, or the rank r of the solution is increased. This is done as follows:

- If the deterministic error $\mathrm{est}_{\det}(w_N)$ outweighs the others, we mark the elements $T \in \mathcal{T}$ that have the largest error $\mathrm{est}_{\det,T}(w_N)$ until the sum of errors on the elements in this marked subset $\mathcal{T}_{\mathrm{mark}} \subset \mathcal{T}$ exceeds a certain ratio $0 < \vartheta < 1$. This is called the *Dörfler marking strategy*

$$\sum_{T\in\mathcal{T}_{\mathrm{mark}}} \mathrm{est}_{\det,T}(w_N) \geq \vartheta\, \mathrm{est}_{\det}(w_N).$$

We denote this procedure by

$$\mathcal{T}_{\mathrm{mark}} \leftarrow \mathrm{Mark}_x[\vartheta, (\mathrm{est}_{\det,T}(w_N, \Lambda))_{T\in\mathcal{T}}, \mathrm{est}_{\det}(w_N, \Lambda, \mathcal{T})].$$

The elements in this subset are subsequently refined by

$$\mathcal{T}^+ \leftarrow \mathrm{Refine}_x[\mathcal{T}, \mathcal{T}_{\mathrm{mark}}].$$

- In case the parametric error $\mathrm{est}_{\mathrm{param}}(w_N)$ dominates, we use estimators $\mathrm{est}_{\mathrm{param},m}(w_N)$ in order to determine which components need to be refined. Here,

Algorithm 2 The TTASGFEM algorithm.

Input:	solution w_N with solution tensor W and rank r;
	mesh $\mathcal{T}$ with degrees p;
	index set Λ;
	Dörfler marking ratio ϑ;
	accuracy $\epsilon_{\text{TTASGFEM}}$;
Output:	New solution w_N^+ with new solution tensor W^+;
	updated mesh $\mathcal{T}^+$;
	updated index set Λ^+;

Init $w_N^+ = w_N$; $W^+ = W$; $\mathcal{T}^+ = \mathcal{T}$; $\Lambda^+ = \Lambda$; $r^+ = r$; $\text{est}_{\text{all}} = \infty$;
repeat

$$
\begin{aligned}
w_N^+ &\leftarrow \text{Solve}[\Lambda^+, \mathcal{T}^+, r^+, W^+] \\
(\text{est}_{\text{det},T})_{T \in \mathcal{T}^+}, \text{est}_{\text{det}} &\leftarrow \text{Estimate}_x[w_N^+, \Lambda^+, \mathcal{T}^+, p] \\
(\text{est}_{\text{param},m})_{m \in \mathbb{N}}, \text{est}_{\text{param}} &\leftarrow \text{Estimate}_y[w_N^+, \Lambda^+] \\
\text{est}_{\text{disc}} &\leftarrow \text{Estimate}_{\text{ALS}}[w_N^+] \\
\text{est}_{\text{all}} &\leftarrow \text{Estimate}_{\text{all}}[\text{est}_{\text{det}}, \text{est}_{\text{param}}, \text{est}_{\text{disc}}]
\end{aligned}
$$

if $\text{est}_{\text{det}} = \max\{\text{est}_{\text{det}}, \text{est}_{\text{param}}, \text{est}_{\text{disc}}\}$ **then**

$$
\begin{aligned}
\mathcal{I}_{\text{mark}} &\leftarrow \text{Mark}_x[\vartheta, (\text{est}_{\text{det},T})_{T \in \mathcal{T}^+}, \text{est}_{\text{det}}] \\
\mathcal{T}^+ &\leftarrow \text{Refine}_x[\mathcal{T}, \mathcal{I}_{\text{mark}}]
\end{aligned}
$$

else if $\text{est}_{\text{param}} = \max\{\text{est}_{\text{det}}, \text{est}_{\text{param}}, \text{est}_{\text{disc}}\}$ **then**

$$
\begin{aligned}
\mathcal{I}_{\text{mark}} &\leftarrow \text{Mark}_y[(\text{est}_{\text{param},m})_{m \in \mathbb{N}}, \text{est}_{\text{param}}] \\
\Lambda^+ &\leftarrow \text{Refine}_y[\Lambda, \mathcal{I}_{\text{mark}}]
\end{aligned}
$$

else

$$
W^+ \leftarrow \text{Refine}_{\text{TT}}[W^+]
$$

end if
until $\text{est}_{\text{all}} < \epsilon_{\text{TTASGFEM}}$

we also mark until the Dörfler property is satisfied, that is, we obtain a subset $\mathcal{I}_{\text{mark}} \subset \mathbb{N}$ such that

$$
\sum_{m \in \mathcal{I}_{\text{mark}}} \text{est}_{\text{param},m}(w_N) \geq \vartheta \, \text{est}_{\text{param}}(w_N).
$$

This is the marking

$$
\mathcal{I}_{\text{mark}} \leftarrow \text{Mark}_y[\vartheta, (\text{est}_{\text{param},m}(w_N))_{m \in \mathbb{N}}, \text{est}_{\text{param}}(w_N)],
$$

and we refine by increasing $d_m^+ \leftarrow d_m + 1$ for $m \in \mathcal{I}_{\text{mark}}$

$$
\Lambda^+ \leftarrow \text{Refine}_y[\Lambda, \mathcal{I}_{\text{mark}}].
$$

- Finally, if $\text{est}_{\text{disc}}(w_N)$ exceeds the other errors, we simply add a random tensor of rank 1 to the solution tensor W. This increases all TT ranks of W by 1 almost surely. It would also be possible to add an approximation of the discrete residual as this is also in TT format. However, since the ALS algorithm will be performed after the refinement, the advantage of this rather costly improvement has shown to be negligible [52]. Thus we get

$$
W^+ \leftarrow \text{Refine}_{\text{TT}}[W].
$$

A single iteration step of the adaptive algorithm returns either a refined $\mathcal{T}^+$ or Λ^+ or the tensor format solution with increased rank W^+ and then solves the problem with these properties. This is done repeatedly until the overall error estimator $\mathrm{err}_{\mathrm{all}}(w_N)$ is sufficiently small, i.e., defined error bound $\epsilon_{\mathrm{TTASGFEM}}$ or a maximum problem size is reached. This procedure is given by the function TTASGFEM, displayed in Algorithm 2. Note that we omit the dependence of w_N for the estimators in the algorithm. The upper error bounds directly follow from Corollary 5.7.

7 Numerical experiments

In this section we examine the performance of the proposed adaptive algorithm with some benchmark problems. The computations are carried out with the open source framework ALEA [18]. The FE discretization is based on the open source package FEniCS [24]. The shown experiments are similar to the ones of the predecessor paper [21] in Section 7. As before, the model equation (2.1) is computed for different lognormal coefficients on the square domain. The derived error estimator is used as a refinement indicator. Of particular interest is the observed convergence of the true (sampled) expected energy error and the behaviour of the error indicator. Moreover, we comment on the complexity of the coefficient discretization.

7.1 Evaluation of the error

The accuracy of the computed approximation is determined by a Monte Carlo estimation. For this, a set of N_{MC} independent samples $\left(y^{(i)}\right)_{i=1}^{N_{\mathrm{MC}}}$ of the stochastic parameter vector is considered. By the tensor structure of the probability measure $\gamma = \bigotimes_{m \in \mathbb{N}} \gamma_1$, each entry of the vector valued sample $y^{(i)}$ is drawn with respect to $\mathcal{N}_1$. The parametric solution $w_N \in \mathcal{V}_N(\Lambda; \mathcal{T}, p)$ obtained by the adaptive algorithm at a sample $y^{(i)}$, is compared to the corresponding (deterministic) sampled solution $u(y^{(i)}) \in \mathcal{X}_{p'}(\widetilde{\mathcal{T}})$ on a much finer finite element space subject to a reference triangulation $\widetilde{\mathcal{T}}$, obtained from uniform refinements of the finest adaptively created mesh, up to the point where we obtain at least 10^5 degrees of freedom with Lagrange elements of order $p' = 3$. For computations, the truncation parameter applied to the infinite sum in (2.2), controlling the length of the affine field representation, is set to $M = 100$. The mean squared error is determined by a Monte Carlo quadrature,

$$\int_{\Gamma} \|u(y) - w_N(y)\|_{\mathcal{X}}^2 \mathrm{d}\gamma(y) \approx \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} \|u(y^{(i)}) - w_N(y^{(i)})\|_{\mathcal{X}_p(\widetilde{\mathcal{T}})}^2.$$

Note that the sample reference quantity $u(y)$ exhibits approximation solely by a finite element method and does not involve the computation of a unified reference solution. A choice of $N_{\mathrm{MC}} = 250$ proved to be sufficient to obtain a reliable estimate in our experiments.

7.2 The stochastic model problem

In the numerical experiments, we consider the stationary diffusion problem (2.1) on the square domain $D = (0, 1)^2$. As in [15,16,19,21], the expansion coefficients of the stochastic field (2.2) are given by

$$b_m(x) := \gamma m^{-\sigma} \cos(2\pi \varrho_1(m)x_1) \cos(2\pi \varrho_2(m)x_2), \text{ for } m \geq 1. \qquad (7.1)$$

We chose the scaling $\gamma = 0.9$.

Moreover,

$$\varrho_1(m) = m - k(m)(k(m) + 1)/2 \quad \text{and} \quad \varrho_2(m) = k(m) - \varrho_1(m),$$

with $k(m) = \lfloor -1/2 + \sqrt{1/4 + 2m} \rfloor$, i.e., the coefficient functions b_m enumerate all planar Fourier sine modes in increasing total order. To illustrate the influence which the stochastic coefficient plays in the adaptive algorithm, we examine the expansion with varying decay, setting σ in (7.1) to different values. The computations are carried out with conforming FE spaces of polynomial degree 1 and 3.

The fully discretized problem is solved in the TT format using the *Alternating Least Squares (ALS)* algorithm as introduced in [21]. Other algorithms like Riemannian optimization are also feasible [1,37]. The ALS has a termination threshold of 10^{-12} that needs to be reached in the Frobenius norm of the difference of two successive iteration results.

For our choice of the coefficient field, the introduced weights in (3.1) are set to $\varrho = 1$ and $\theta = 0.1$.

7.3 Tensor train representation of the coefficient

Since the tensor approximation of the coefficient is the starting point for the discretization of the operator, we begin with an examination of the rank dependence of the coefficient approximation scheme given in Algorithm 1. For this, we fix the multi-index set such that the incorporated Hermite polynomials are of degree 15 in each stochastic dimension, i.e.,

$$\Delta = \{v \in \mathcal{F} : v_\ell = 0, \ldots, 16, \ell = 1, \ldots, L\}.$$

As a benchmark, we chose a slow decay with $\sigma = 2$ and set the quadrature rule[2] to exactly integrate polynomials of degree 7 with a grid such that the number of quadrature nodes is at least $P_{quad} = 10^4$. In the following, the relative root mean squared (RRMS) error

$$\mathcal{E}(a, a_{\Delta,s}) := \left(\mathbb{E}[\|a - a_{\Delta,s}\|^2_{L^2(D)} / \|a\|^2_{L^2(D)}] \right)^{1/2}$$

---

[2] The statements are with regard to the respective FE discretization in FEniCS.

Table 1 Comparison of different coefficient field approximations with fixed stochastic parameter set of order $L = \{10, 50, 100\}$, fixed stochastic polynomial degree 15 and different rank configurations having maximal rank $s_{max} = \{10, 20, 50, 100\}$

s_{max}	$\mathcal{E}(a, a_{\Delta,s})$	tt-dofs(A)	full-dofs(A)	CPU-time (s)
$L = 10$				
10	1.21×10^{-3}	2.04×10^{6}	3.02×10^{17}	1.7
20	6.25×10^{-4}	4.12×10^{6}		4.56
50	6.18×10^{-4}	1.05×10^{7}		16.52
100	6.18×10^{-4}	2.15×10^{7}		70.29
$L = 50$				
10	1.17×10^{-3}	2.11×10^{6}	3.26×10^{65}	9.14
20	3.26×10^{-4}	4.36×10^{6}		13.56
50	5.98×10^{-5}	1.20×10^{7}		50.60
100	5.16×10^{-5}	2.75×10^{7}		247.63
$L = 100$				
10	1.16×10^{-3}	2.18×10^{6}	5.25×10^{125}	19.09
20	3.20×10^{-4}	4.66×10^{6}		27.16
50	6.33×10^{-5}	1.38×10^{7}		100.53
100	7.47×10^{-6}	3.50×10^{7}		498.47

is compared to the growth in degrees of freedom with respect to various tensor ranks. Numerically, this expression is computed by a Monte Carlo approximation as described in Sect. 7.1 with the reference mesh $\widetilde{\mathcal{T}}$.

By denoting $A = (A_0, A_1, \ldots, A_L)$ the component tensors of $a_{\Delta,s}$ in (4.4), where $A_0[x_k, k_1] = a_0[k_1](x_k)$ corresponds to the evaluation of $a_0[k_1]$ at every node x_k of $\widetilde{\mathcal{T}}$, we can apply the notion of tt-dofs (4.1) to the discrete tensor train coefficient. To highlight the crucial choice of the rank parameter $s = (1, s_1, \ldots, s_L, 1)$ in $a_{\Delta,s}$, we let the maximal attainable rank s_{max} vary.

It can be seen in Table 1 that a higher rank yields a more accurate approximation of the coefficient, as long as the stochastic expansion L is long enough. For small numbers of L, the RRMS does not decrease any further than 6×10^{-4}, even when increasing the maximal attainable rank. Otherwise, for up to $L = 100$ terms in the affine coefficient field parametrisation, a small rank parameter of $s_{max} = 10$ gives a reasonable RRMS error of 1×10^{-3} which can be improved by increasing s_{max}.

It should be pointed out that the used approximation only becomes feasible computationally by the employed tensor format since without low-rank representation the amount of degrees of freedom grows exponentially in der number of expansion dimensions. Furthermore, since the tensor complexity and arithmetic operations depend only linearly on the number of expansion dimensions, the computational time scales linearly as illustrated in the last column of Table 1 showing the average run time of 10 runs.[3]

---

[3] Run on an 8GB Intel Core i7-6500 laptop.

Fig. 3 Relative sampled mean squared $H^1(D)$ error for fixed stochastic dimension $L = 5$ and adaptive refinement of the physical mesh. Each setting is shown along its respective overall error estimator as defined in Corollary 5.7 and plotted against the total number of degrees of freedom in the TT representation. Considered are FE approximations of order $p = 1$ and $p = 3$
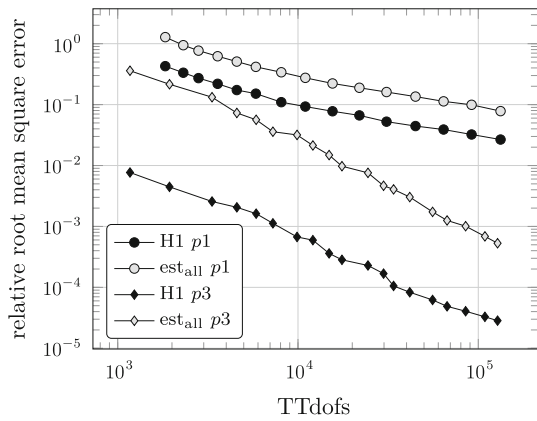
7.4 Adaptivity in physical space

As a first example for an adaptive discretization, we examine the automatic refinement of the physical FE space only. For this, the stochastic expansion of the coefficient is fixed to some small value $L = 5$ in (4.4), which also is assumed for the corresponding reference solution. The considered polynomial (Hermite chaos) degree of the approximation in each dimension is fixed to $d_1 = \ldots = d_5 = 10$, which can be considered sufficiently large for the respective problem, given an algebraic decay rate of $\sigma = 2$ in the coefficient.

Although this experiment does not illustrate the performance of the complete algorithm, it nevertheless depicts the varying convergence rates for different polynomial orders in the FE approximation. The rank of the tensor train approximation is fixed to $r_{\max} = 10$, which is sufficient due to the low stochastic dimensions.

It can be observed in Fig. 3 that the error estimator follows the rate of convergence of the sampled error. Moreover, the higher-order FE method exhibits a higher convergence rate as expected. This could also be observed in the (much simpler) affine case scrutinized in the preceding works [15,16,19,21].

7.5 Fully adaptive algorithm

The fully adaptive algorithm described in Algorithm 2 is instantiated with a small initial tensor with full tensor rank $r_1 = 2$ consisting of a single $M = 1$ stochastic component discretized with a linear polynomial $d_1 = 2$ and a physical mesh with $|\mathcal{T}| = 32$ elements for linear ansatz functions $p = 1$ and $|\mathcal{T}| = 8$ for $p = 3$. The marking parameter is set to $\vartheta = 0.5$.

Figure 4 depicts the real (sampled) mean squared energy error and the corresponding overall error estimator for FE discretizations of degrees $p = 1, 3$. On the left-hand side of the figure, a lower decay rate ($\sigma = 2$) in the coefficient representation is used, resulting in more stochastic modes to be relevant for an accurate approximation. The right-hand side shows results for a faster decay rate with $\sigma = 4$. As expected, the convergence rate for the $p = 3$ FE discretizations is faster than with $p = 1$ FE.
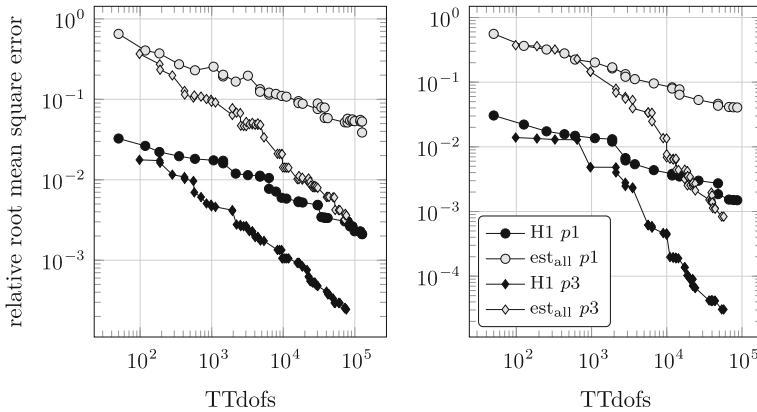
Fig. 4 Relative, sampled, mean squared $H^1(D)$ error for the fully adaptive refinement. Each setting is shown along its respective overall error estimator as defined in Corollary 5.7 and plotted against the total number of degrees of freedom in the representation. Considered are finite elements approximation of order $p = 1$ and $p = 3$ and slow decay ($\sigma = 2$, left) and fast decay ($\sigma = 4$, right)

Table 2 Stochastic expansion length, maximal polynomial chaos degrees, tensor ranks and degrees of freedom for the physical and (compressed) stochastic discretizations as well as for the operator in TT format for selected iteration steps in the fully adaptive algorithm

Iteration	M	$d^{\max}$	$r^{\max}$	m-dofs(W_N)	tt-dofs(W_N)	op-dofs(W_N)
Finite element order $p = 1$ and slow coefficient decay $\sigma = 2$
5/37	1	1	2	292	584	873
15/37	2	2	5	1577	7900	73,965
20/37	3	4	8	2330	18,656	170,247
30/37	5	4	13	6586	85,,847	678,382
37/37	6	5	19	6586	126,330	734,880
Finite element order $p = 1$ and fast coefficient decay $\sigma = 4$
5/22	1	1	2	302	604	1200
10/22	1	3	3	941	2826	11,196
15/22	1	4	5	2951	14,755	44,115
22/22	2	5	9	9608	86,499	962,022

Finite elements of order $p = 1$ are used to solve the problem with coefficient decay rates of $\sigma \in \{2, 4\}$

Moreover, for the simpler problem with fast decay, we achieve a smaller error with a comparable number of degrees of freedom than in the harder slow decay setting.

Remark 7.1 Note that we neglect the (large) factor $\check{c}^{+}_{\theta\varrho}$ in Corollary 5.7, which depends on the choice of weights θ and ϱ of the discrete space. A detailed analysis of how to optimally select these weights is outside the scope of this article.

We conclude the numerical observations with Table 2 and 3 to display the behaviour of the fully adaptive algorithm. To allow for more insights, we redefine the physical mesh resolution as m-dofs(W_N), which is the number of FE degrees of freedom for the respective parametric solution W_N. Moreover, we define the number of degrees

Table 3 Stochastic expansion length, maximal polynomial chaos degrees, tensor ranks and degrees of freedom for the physical and (compressed) stochastic discretizations as well as for the operator in TT format for selected iteration steps in the fully adaptive algorithm

Iteration	M	$d^{\max}$	$r^{\max}$	m-dofs(W_N)	tt-dofs(W_N)	op-dofs(W_N)
Finite element order $p = 3$ and slow coefficient decay $\sigma = 2$						
5/65	1	2	2	139	417	816
20/65	5	3	12	241	3040	37,537
35/65	7	6	21	403	9856	130,799
50/65	9	11	33	568	27,662	239,273
65/65	16	11	46	1078	76,173	353,153
Finite element order $p = 3$ and fast coefficient decay $\sigma = 4$						
5/67	1	2	2	322	966	2226
20/67	2	7	7	991	9974	101,110
35/67	4	13	15	1207	20,423	250,913
50/67	6	21	19	1909	39,998	434,669
67/67	6	34	22	2404	61,920	607,076

Finite elements of order $p = 3$ are used to solve the problem with coefficient decay rates of $\sigma \in \{2, 4\}$

of freedom incorporated in the operator (4.5) generated to obtain the current solution. For a tensor train operator of dimension $\mathbf{A} \in \mathbb{R}^{(N \times N) \times (q_1 \times q_1) \times \dots \times (q_L \times q_L)}$ and rank $s = (s_1, \dots, s_L)$, we define

$$\text{op-dofs}(\mathbf{A}) := N^2 s_1 - s_1^2 + \sum_{\ell=1}^{L-1} (s_\ell q_\ell^2 s_{\ell+1} - s_{\ell+1}^2) + s_L q_L^2.$$

Note that, using a sparse representation of the operator in the first (physical) component, it is possible to reduce number of op-dofs.

Tables 2 and 3 highlight some iteration steps and the employed stochastic expansion length M, the maximal polynomial degree $d^{\max}$ (which was usually naturally attained in the first stochastic component), the maximal tensor train rank $r^{\max}$ (again most of the time this rank is the separation rank of the physical space and the first stochastic dimension) and the corresponding degrees of freedom. Notably, for the fast coefficient decay $\sigma = 4$, the stochastic expansion is very short since most of the randomness is due to the first few stochastic parameters. It is reasonable that the respective parameter dimensions require higher polynomial degrees in the approximation. For the more involved slow decay $\sigma = 2$ setting we observe a larger stochastic expansion of the solution and larger tensor ranks.

References

1. Absil, P.-A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton (2008)
2. Babuška, I., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. **45**(3), 1005–1034 (2007)
3. Bachmayr, M., Schneider, R., Uschmajew, A.: Tensor networks and hierarchical tensors for the solution of high-dimensional partial differential equations. Found. Comput. Math. **16**(6), 1423–1472 (2016)
4. Bespalov, A., Powell, C.E., Silvester, D.: Energy norm a posteriori error estimation for parametric operator equations. SIAM J. Sci. Comput. **36**(2), A339–A363 (2014)
5. Bespalov, A., Silvester, D.: Efficient adaptive stochastic Galerkin methods for parametric operator equations. SIAM J. Sci. Comput. **38**(4), A2118–A2140 (2016)
6. Bryant, C.M., Prudhomme, S., Wildey, T.: Error decomposition and adaptivity for response surface approximations from pdes with parametric uncertainty. SIAM/ASA J. Uncertain. Quant. **3**(1), 1020–1045 (2015)
7. Charrier, J.: Strong and weak error estimates for elliptic partial differential equations with random coefficients. SIAM J. Numer. Anal. **50**(1), 216–246 (2012)
8. Cohen, A., DeVore, R., Schwab, C.: Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. Found. Comput. Math. **10**(6), 615–646 (2010)
9. Cohen, A., Devore, R., Schwab, C.: Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. Anal. Appl. (Singapore) **9**(1), 11–47 (2011)
10. Crowder, A.J., Powell, C.E., Bespalov, A.: Efficient adaptive multilevel stochastic Galerkin approximation using implicit a posteriori error estimation. arXiv preprint arXiv:1806.05987 (2018)
11. Detommaso, G., Dodwell, T., Scheichl, R.: Continuous level Monte Carlo and sample-adaptive model hierarchies. arXiv preprint arXiv:1802.07539 (2018)
12. Dolgov, S., Khoromskij, B.N., Litvinenko, A., Matthies, H.G.: Polynomial chaos expansion of random coefficients and the solution of stochastic partial differential equations in the tensor train format. SIAM/ASA J. Uncertain. Quant. **3**(1), 1109–1135 (2015)
13. Dolgov, S., Scheichl, R.: A hybrid alternating least squares–tt cross algorithm for parametric pdes. arXiv preprint arXiv:1707.04562 (2017)
14. Eigel, M., Martin, G., Claude, J.: Adaptive stochastic Galerkin FEM for a log-transformed PDE with non-affine unbounded random coefficients (2018). (in preparation)
15. Eigel, M., Gittelson, C.J., Schwab, C., Zander, E.: Adaptive stochastic Galerkin FEM. Comput. Methods Appl. Mech. Eng. **270**, 247–269 (2014)
16. Eigel, M., Gittelson, C.J., Schwab, C., Zander, E.: A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes. ESAIM Math. Model. Numer. Anal. **49**(5), 1367–1398 (2015)
17. Eigel, M., Marschall, M., Schneider, R.: Sampling-free bayesian inversion with adaptive hierarchical tensor representations. Inverse Prob. **34**(3), 035010 (2018)
18. Eigel, M., Marschall, M., Zander, E.: alea—A python framework for spectral methods and low-rank approximations in uncertainty quantification. https://bitbucket.org/aleadev/alea
19. Eigel, M., Merdon, C.: Local equilibration error estimators for guaranteed error control in adaptive stochastic higher-order Galerkin finite element methods. SIAM/ASA J. Uncertain. Quant. **4**(1), 1372–1397 (2016)
20. Eigel, M., Merdon, C., Neumann, J.: An adaptive multilevel monte carlo method with stochastic bounds for quantities of interest with uncertain data. SIAM/ASA J. Uncertain. Quant. **4**(1), 1219–1245 (2016)
21. Eigel, M., Pfeffer, M., Schneider, R.: Adaptive stochastic Galerkin FEM with hierarchical tensor representations. Numerische Mathematik **136**(3), 765–803 (2017)

22. Ernst, O.G., Mugler, A., Starkloff, H.-J., Ullmann, E.: On the convergence of generalized polynomial chaos expansions. ESAIM Math. Model. Numer. Anal. **46**(2), 317–339 (2012)
23. Espig, M., Hackbusch, W., Litvinenko, A., Matthies, H.G., Wähnert, P.: Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats. Comput. Math. Appl. **67**(4), 818–829 (2014)
24. FEniCS project—automated solution of differential equations by the finite element method. http://fenicsproject.org
25. Galvis, J., Sarkis, M.: Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity. SIAM J. Numer. Anal. **47**(5), 3624–3651 (2009)
26. Gittelson, C.J.: Stochastic Galerkin discretization of the log-normal isotropic diffusion problem. Math. Models Methods Appl. Sci. **20**(2), 237–263 (2010)
27. Grasedyck, L.: Hierarchical singular value decomposition of tensors. SIAM J. Matrix Anal. Appl. **31**(4):2029–2054 (2009/10)
28. Guignard, D., Nobile, F.: A posteriori error estimation for the stochastic collocation finite element method. Technical report (2017)
29. Hackbusch, W., Schneider, R.: Tensor spaces and hierarchical tensor representations. In: Extraction of Quantifiable Information from Complex Systems, volume 102 of Lecture Notes in Computer Science Engineering, pp. 237–261. Springer, Cham (2014)
30. Harbrecht, H., Peters, M., Siebenmorgen, M.: Multilevel accelerated quadrature for pdes with lognormally distributed diffusion coefficient. SIAM/ASA J. Uncertain. Quant. **4**(1), 520–551 (2016)
31. Harbrecht, H., Peters, M., Siebenmorgen, M.: On the quasi-monte carlo method with halton points for elliptic pdes with log-normal diffusion. Math. Comput. **86**(304), 771–797 (2017)
32. Herrmann, L., Schwab, C.: QMC algorithms with product weights for lognormal-parametric, elliptic pdes (2017)
33. Herrmann, L., Schwab, C.: Multilevel quasi-monte carlo integration with product weights for elliptic pdes with lognormal coefficients. Techncial Report 2017–2019, Seminar for Applied Mathematics. ETH Zürich, Switzerland (2017)
34. Hoang, V.H., Schwab, C.: *N*-term Wiener chaos approximation rate for elliptic PDEs with lognormal Gaussian random inputs. Math. Models Methods Appl. Sci. **24**(4), 797–826 (2014)
35. Holtz, S., Rohwedder, T., Schneider, R.: On manifolds of tensors of fixed TT-rank. Numer. Math. **120**(4), 701–731 (2012)
36. Kazashi, Y.: Quasi-monte carlo integration with product weights for elliptic pdes with log-normal coefficients. arXiv preprint arXiv:1701.05974 (2017)
37. Kressner, D., Steinlechner, M., Vandereycken, B.: Low-rank tensor completion by Riemannian optimization. BIT **54**(2), 447–468 (2014)
38. Kuo, F., Scheichl, R., Schwab, C., Sloan, I., Ullmann, E.: Multilevel quasi-monte carlo methods for lognormal diffusion problems. Math. Comput. **86**(308), 2827–2860 (2017)
39. Loève, M.: Probability theory. II. Graduate Texts in Mathematics, vol. 46, 4th edn. Springer, New York (1978)
40. Malliavin, P.: Stochastic Analysis, vol. 313. Springer, Berlin (2015)
41. Mugler, A.: Verallgemeinertes polynomielles Chaos zur Lösung stationärer Diffusionsprobleme mit zufälligen Koeffizienten. PhD thesis, BTU Cottbus (2013)
42. Mugler, A., Starkloff, H.-J.: On the convergence of the stochastic Galerkin method for random elliptic partial differential equations. ESAIM Math. Model. Numer. Anal. **47**(5), 1237–1263 (2013)
43. Nobile, F., Tempone, R., Webster, C.G.: An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. SIAM J. Numer. Anal. **46**(5), 2411–2442 (2008)
44. Nobile, F., Tempone, R., Webster, C.G.: A sparse grid stochastic collocation method for partial differential equations with random input data. SIAM J. Numer. Anal. **46**(5), 2309–2345 (2008)
45. Nouy, A.: Low-rank methods for high-dimensional approximation and model order reduction. In: Benner, P., Cohen, A., Ohlberger, M., Willcox, K. (eds.) Model reduction and approximation, pp. 171–226. SIAM, Philadelphia, PA (2017)
46. Oseledets, I.V., Tyrtyshnikov, E.E.: Breaking the curse of dimensionality, or how to use SVD in many dimensions. SIAM J. Sci. Comput. **31**(5), 3744–3759 (2009)
47. Pfeffer, M.: Tensor methods for the numerical solution of high-dimensional parametric partial differential equations. PhD thesis, TU Berlin Dissertation No. 19533, Technische Universität Berlin (2018)
48. Prudhomme, S., Bryant, C.M.: Adaptive surrogate modeling for response surface approximations with application to bayesian inference. Adv. Model. Simul. Eng. Sci. **2**(1), 22 (2015)

49. Schwab, C., Gittelson, C.J.: Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. Acta Numer. **20**, 291–467 (2011)
50. Ullmann, E.: Solution strategies for stochastic finite element discretizations. PhD thesis, TU Bergakademie Freiberg (2008)
51. Ullmann, E., Elman, H.C., Ernst, O.G.: Efficient iterative solvers for stochastic Galerkin discretizations of log-transformed random diffusion problems. SIAM J. Sci. Comput. **34**(2), A659–A682 (2012)
52. Uschmajew, A., Vandereycken, B.: Line-search methods and rank increase on low-rank matrix varieties. In: NOLTA2014, Luzern, Switzerland, September 14–18, 2014, pp. 52–55. IEICE, Tokyo (2014)