Kolmogorov widths and low-rank approximations of parametric elliptic PDEs

Markus Bachmayr, Albert Cohen

Kolmogorov widths and low-rank approximations of parametric elliptic PDEs *

Markus Bachmayr and Albert Cohen

February 10, 2015

Abstract

Kolmogorov n-widths and low-rank approximations are studied for families of elliptic diffusion PDEs parametrized by the diffusion coefficients. The decay of the n-widths can be controlled by that of the error achieved by best n-term approximations using polynomials in the parametric variable. However, we prove that in certain relevant instances where the diffusion coefficients are piecewise constant over a partition of the physical domain, the n-widths exhibit significantly faster decay. This, in turn, yields a theoretical justification of the fast convergence of reduced basis or POD methods when treating such parametric PDEs. Our results are confirmed by numerical experiments, which also reveal the influence of the partition geometry on the decay of the n-widths.

1 Introduction

Solving a parameter-dependent family of partial differential equations for a large number of different parameter values can be computationally demanding, since the solution of the problem for each single set of parameters typically already requires substantial resources. Put in abstract form, one aims to find the solution u of

$$\mathcal{P}(u, y) = 0 \tag{1.1}$$

for many different values of a vector $y = (y_1, \ldots, y_d)$ in a certain range $U \subset \mathbb{R}^d$, where $\mathcal{P}$ is a partial differential operator $\mathcal{P}$ parametrized by y.

We assume that the problem is uniformly well posed in a separable Hilbert space V whose elements depend on a physical variable x ranging in a domain $D \subset \mathbb{R}^m$: for each value of $y \in U$, the solution $u(y)$ belongs to V. We thus regard u either as the function $y \mapsto u(y)$ from U to V, or as the function $(x, y) \mapsto u(x, y) := u(y)(x)$ from $D \times U$ to $\mathbb{R}$.

1

Many methods for reducing the complexity of such a task are explicitly or implicitly based on the existence of efficient approximations u_n of u of the general separable form

$$u_n(x,y) := \sum_{k=1}^{n} v_k(x)\, \phi_k(y), \tag{1.2}$$

for some functions $\{v_1, \ldots, v_n\}$ and $\{\phi_1, \ldots, \phi_n\}$. For instance, in the reduced basis method [17], one constructs $\{v_k\}_{k=1,\ldots,n}$ as particular instances $v_k = u(y^k)$ of the solution for well-chosen values $y^k \in U$, so that the problem can be solved rapidly by applying, for each given y, the Galerkin method in the n-dimensional space spanned these basis functions. Similarly, the POD method [14] constructs $\{v_k\}_{k=1,\ldots,n}$ by applying a principal component analysis on a representative set of instances. Both methods amount to an expansion of u of the above form (1.2), although the latter is not explicitly constructed since the values $(\phi_k(y))_{k=1,\ldots,n}$ are computed for each given y by the Galerkin method. In contrast, polynomial based methods such as in [10, 11] produce an explicit representation (1.2) where the functions $y \mapsto \phi_k(y)$ are multivariate polynomials. More generally, low-rank tensor methods [15] aim at directly building an approximation (1.2) with functions v_k and ϕ_k constructed depending on u.

The efficiency of such approaches depends crucially on the size of $n = n(\varepsilon)$, which should not increase too rapidly with decreasing error tolerance ε in the approximation (1.2). A crucial question is how one can obtain bounds for $n(\varepsilon)$ in terms of ε that are not crude overestimates, but indeed reflect the minimal required number of terms in an expansion of the form (1.2). Such bounds can then give an indication to what extent the considered problem is amenable to such approximations.

First upper bounds can be obtained by placing restrictions on the choice of ϕ_k. A fairly well-studied instance are *sparse polynomial expansions*. In this case, the functions ϕ_k are selected from an a priori chosen basis of tensor product polynomials, for instance

$$u_n(x,y) := \sum_{\nu \in \Lambda_n} u_\nu(x)\, L_\nu(y), \tag{1.3}$$

where each $L_\nu(y) = \prod_{i=1}^{d} L_{\nu_i}(y_i)$ is a tensor product of univariate Legendre polynomials of degrees ν_i in the variable y_i, and Λ_n is a suitable subset of such multi-indices $\nu = (\nu_i)_{i=1,\ldots,d}$ with $\#(\Lambda_n) = n$. In this framework, one can also treat infinite-dimensional U, for instance $y \in [-1,1]^{\mathbb{N}}$, provided that the problem is anisotropic in the sense that the y_i have decreasing influence as $i \to \infty$. We refer to [10, 11] where approximations of the basic form (1.2) with provable convergence rates for the error $\|u - u_n\|_{L^\infty(U,V)}$ as $n \to +\infty$ are derived in the infinite-dimensional framework. These approximations are derived by best n-term truncations of Legendre or Taylor expansions, see also [4, 5] for numerical methods in this line of ideas.

The question that we aim to address in this work is whether one can obtain provably more efficient approximations of the form (1.2) when the ϕ_k are *not* selected from certain fixed basis functions, but are allowed to vary essentially arbitrarily. We do not treat the case of infinite-dimensional y here, but rather consider finite-dimensional parameter domains U in an *isotropic* setting, that is, with all parameters y_i carrying equal weight.

When no limitations are imposed a priori on ϕ_k, the relevant benchmark for reduced models that aim to guarantee uniform accuracy in V for all values y is the Kolmogorov n-width of the so-called *solution manifold* $u(U) \subset V$, which is defined for $n \in \mathbb{N}$ by

$$d_n(u(U))_V := \inf_{\dim(W)=n} \sup_{v \in u(U)} \min_{w \in W} \|v - w\|_V = \inf_{\dim(W)=n} \sup_{y \in U} \min_{w \in W} \|u(y) - w\|_V. \quad (1.4)$$

In general, optimal subspaces W_n corresponding to the infimum in (1.4) are not easily determined, although greedy methods under certain assumptions provide reduced basis spaces for which the approximation errors have comparable rates of decay, see [6, 12].

A more accessible notion of low-rank approximation results from replacing $L^\infty(U, V)$ by $L^2(U, V)$, considering instead the quantities

$$\delta_n(u, \mu)_V^2 := \inf_{\dim(W)=n} \int_U \min_{w \in W} \|u(y) - w\|_V^2 \, d\mu(y), \quad (1.5)$$

where μ is a given probability measure on U. The subspaces for which the infimum is attained can in this case be characterized: $u \in L^2(U, V, \mu)$ induces a Hilbert-Schmidt operator from $L^2(U, \mu)$ to V defined by

$$\mathcal{M}_u : \varphi \mapsto \int_U u(y) \, \varphi(y) \, d\mu(y), \quad (1.6)$$

with singular value decomposition

$$\mathcal{M}_u = \sum_{k=1}^\infty \sigma_k v_k \langle \cdot, \phi_k \rangle_{L^2(U,\mu)}, \quad (1.7)$$

where $\sigma := (\sigma_k)_{k \geq 1} \in \ell^2(\mathbb{N})$ is nonnegative and nonincreasing, and $(v_k)_{k \geq 1}$, $(\phi_k)_{k \geq 1}$ are orthonormal bases of V and $L^2(U, \mu)$, respectively. It is well known that for given n, the infimum in (1.5) is attained by choosing $W_n := \{v_k : k = 1, \ldots, n\}$, and has value

$$\delta_n(u, \mu)_V^2 = \sum_{k>n} \sigma_k^2. \quad (1.8)$$

The operator $\mathcal{M}_u$ also yields a precise notion of rank for $u \in L^2(U, V)$ via

$$\text{rank}(u) := \dim \text{range}(\mathcal{M}_u) = \#\{k : \sigma_k \neq 0\}, \quad (1.9)$$

which, in general, is infinite. Note that μ may in principle be *any* probability measure. However, a natural choice in the case where the y_j are random variables is to take for μ the measure associated to the distribution of y, since the quantity which is minimized in (1.5) may then also be viewed as the quadratic error $\mathbb{E}(\|u(y) - P_W u(y)\|_V^2)$. When the y_i are deterministic, a standard choice is to take for μ the uniform probability measure over U.

Note that since μ is a probability measure, we always have

$$\delta_n(u, \mu)_V \leq d_n(u(U))_V. \quad (1.10)$$

3

On the other hand, polynomial approximation results such as in [10, 11] yield estimates for these quantities since obviously

$$d_n(u(U))_V \leq \|u - u_n\|_{L^\infty(U,V)} \,, \tag{1.11}$$

where u_n is an n-term Legendre expansion of the form (1.3).

The results in this paper construct estimates for both (1.4) and (1.5) which are significantly sharper than those which could be obtained using the above estimate (1.11). The particular class of problems that we focus on are parametric diffusion equations

$$- \operatorname{div}_x\big(a(x,y)\nabla u(x,y)\big) = f(x) \,, \quad x \in D \subset \mathbb{R}^m, \ y \in U := [-1,1]^d, \tag{1.12}$$

with $m, d \in \mathbb{N}$, where the coefficient a has the form

$$a(x,y) = \bar{a}(x) - \sum_{i=1}^{d} y_i \psi_i(x) \,. \tag{1.13}$$

Note that if the initial range of y_i is a more general finite interval I_i, the range $y_i \in [-1, 1]$ can always be ensured up to a renormalization of ψ_i and a modification of $\bar{a}$. While in most references there is a positive sign after $\bar{a}(x)$, we use here a negative sign for later notational convenience.

Throughout the paper, we assume a to satisfy a *uniform ellipticity assumption*: there exist $0 < r \leq R$ such that

$$0 < r \leq a(x,y) \leq R < \infty \,, \quad x \in D, \ y \in U. \tag{1.14}$$

In view of (1.13), this entails in particular

$$\sum_{i=1}^{d} |\psi_i(x)| \leq \bar{a}(x) - r \,, \quad \sum_{i=1}^{d} |\psi_i(x)| \leq R - \bar{a}(x) \,, \quad x \in D \,. \tag{1.15}$$

We consider the equation (1.12) in its weak form on $V := H_0^1(D)$,

$$\Big(\bar{A} - \sum_{i=1}^{d} y_i A_i\Big) u(y) = f \,, \tag{1.16}$$

where $f \in V'$ and $\bar{A}, A_i \colon V \to V'$ are defined by

$$\langle \bar{A}u, v\rangle := \int_D \bar{a}\nabla u \cdot \nabla v \, dx \,, \quad \langle A_i u, v\rangle := \int_D \psi_i \nabla u \cdot \nabla v \, dx \,, \quad u, v \in V \,. \tag{1.17}$$

A simple, yet relevant, model problem that we study in detail – and where it turns out that general low-rank approximations can indeed perform substantially better than sparse polynomial approximations – is the case of piecewise constant coefficients a, for instance

$$\bar{a} = 1 \quad \text{and} \quad \psi_i = \theta \chi_{D_i}, \ \theta < 1, \tag{1.18}$$
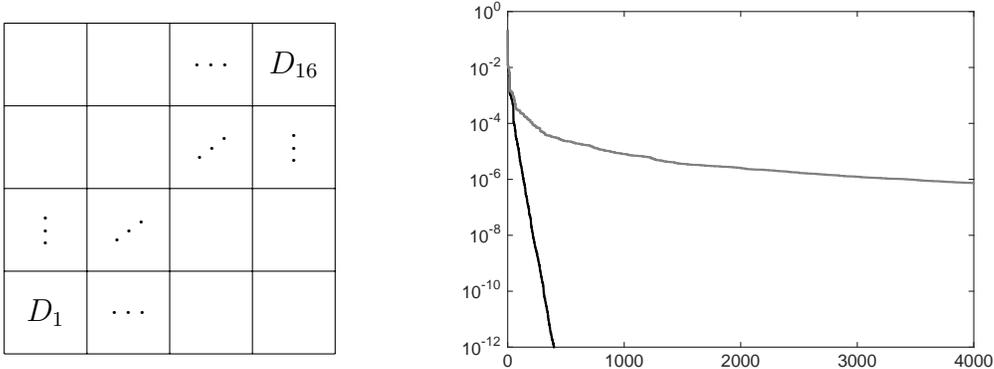
4

Figure 1.1: Example geometry of piecewise constant coefficients ($d = 16$) and decay of singular values (black) and of ordered Legendre coefficients (gray), see also Section 6.

where $\{D_1, \ldots, D_d\}$ is a given partition of D. Such problems arise, for example, in the analysis of heat conductivity properties of components composed of several materials.

As an example of what one may expect in such a case, Figure 1.1 shows a comparison between the numerically observed decay of (i) the singular values σ_k as in (1.7) with μ being the uniform measure in U and (ii) the reordered norms $\|u_\nu\|_V$ of Legendre coefficients in (1.3) where the Legendre polynomials have been normalized in $L^2(U, \mu)$, in the case of a checkerboard partition of the unit square D into 16 smaller squares. Whereas the norms of the Legendre coefficients exhibit a subexponential decay, the singular values are observed to decay exponentially. This indicates that, for this particular case, separable approximations of u using optimally adapted v_k and ϕ_k, for example using reduced basis or POD methods, can perform substantially better than those obtained by best n-term truncation of Legendre series. This phenomenon was already observed in [16] and [3]. Here we provide a first theoretical explanation through a rigorous analysis for various situations. We shall see in particular that this effect is significantly connected to the geometry of the partition.

The rest of our paper is organized as follows: in Section 2, we give two very simple examples for which the solution manifold $u(U)$ happens to be contained in a finite dimensional space, meaning that $d_n(u(U))_V = 0$ for n sufficiently large, while polynomial approximations require that the number of terms n goes to $+\infty$ in order to converge to u. We then consider in Section 3 the general elliptic problem (1.12) with coefficients given by (1.13), and we consider its approximation by its truncated power series

$$u_n(x, y) = \sum_{|\nu| \le k} t_\nu(x) \, y^\nu, \quad n := n(d, k) = \binom{k + d}{d}. \tag{1.19}$$

Our main observation is that under a general condition relating the functions $\bar{a}$ and ψ_i, this approximation has rank $r(n)$ significantly smaller than its number of terms n. Using the estimate

$$d_{r(n)}(u(U))_V \le \|u - u_n\|_{L^\infty(U,V)}, \tag{1.20}$$

5

this allows us to show that the n-widths of $u(U)$ decay significantly faster than the approximation error $\|u - u_n\|_{L^\infty(U,V)}$. The general condition applies in particular to the case where $\bar{a}$ is constant and ψ_i is proportional to χ_{D_i}, which corresponds to the above mentioned case of piecewise constant coefficients. In Section 4, we specialize on the piecewise constant situation and show that low-rank approximations can be obtained through the study of an auxiliary parametric problem posed on the skeleton of the partition and involving various Steklov-Poincaré operators. Using this approach, we prove the exponential decay of the n-widths (which justifies the numerically observed decay of the singular values) in the simple case of a 2×2 checkerboard. We also demonstrate numerically that this behaviour strongly depends on the geometry, in the sense that it is lost when considering a partition of D into 4 quadrilaterals which are not squares. Finally, in Section 5 we discuss the numerical schemes which we use in order to perform the numerical experiments. These schemes combine non-adaptive polynomial approximations and rank reduction by singular value decompositions. The development of numerical schemes which combine rank reduction and adaptive polynomial approximation, along the lines of [2], is the subject of current investigation.

2 Simple examples of exact low-rank representations

We first consider two simple examples where explicit low-rank representations can be given, leading to particularly favorable bounds for the Kolmogorov widths of solution manifolds, when compared to error bounds for the polynomial approximations.

Example 2.1. A trivial instance of a situation where a sparse polynomial expansion in y gives clearly suboptimal approximations is when all ψ_i are proportional to $\bar{a}$, that is, when $\psi_i := c_i \bar{a}$ with values $c_i \in \mathbb{R}$ such that (1.14) holds. Then we may write $u(y)$ as a scalar multiple of $v_1 := u(0) = \bar{A}^{-1} f$ for all $y \in U$, that is

$$u(x,y) = v_1(x)\phi_1(y), \quad \phi_1(y) := \Big(1 - \sum_{i=1}^{d} c_i y_i\Big)^{-1}, \tag{2.1}$$

which shows that $\mathrm{rank}(u) = 1$, or in other words $d_n(u(U))_V = 0$ for $n \geq 1$. However, u has an infinite power series

$$u(x,y) = \sum_{\nu \in \mathbb{N}_0^d} t_\nu(x)\, y^\nu, \quad t_\nu = \Big(\frac{|\nu|!}{\prod_{i=1}^{d} \nu_i!} \prod_{i=1}^{d} c_i^{\nu_i}\Big) v_1, \tag{2.2}$$

and therefore polynomial approximation error bounds are infinitely less favorable than the n-widths.

Example 2.2. We now consider a piecewise constant diffusion coefficient in the case of a one-dimensional interval D. Without loss of generality we set $D =]0,1[$, assume $\{D_1, \ldots, D_d\}$ to be a partition of D into d subintervals, and set $\bar{a} := 1$ and $\psi_i := \theta \chi_{D_i}$ for a $\theta \in]0,1[$.

In the interior of each D_i, we then have $-(1 + \theta y_i)u'' = f$. Hence, with any F such that $F'' = f$, for each y and i we obtain

$$u(y)|_{D_i} \in \text{span}\{\chi_{D_i}, x\, \chi_{D_i}, F\, \chi_{D_i}\}. \tag{2.3}$$

This yields an upper bound of $3d$ degrees of freedom for $u(y)$. Since there are $d+1$ continuity conditions (independent of y) on u at domain and subinterval boundaries, we have

$$\text{rank}(u) \le 2d - 1, \tag{2.4}$$

and therefore $d_n(u(U))_V = 0$ for $n \ge 2d - 1$, independently of the particular choice of f.

The same argument still applies for general $\bar{a}$ that is such that $0 < r \le \bar{a} \le R$ and $\psi_i := \theta \bar{a} \chi_{D_i}$, for some $\theta \in \,]0, 1[$. Since the ψ_i have this particular form, on the interior of each D_i we have the equation

$$-(1 + \theta y_i)(\bar{a}u')' = f, \tag{2.5}$$

which can still be solved by integrating twice as before, and consequently one still has $\text{rank}(u) \le 2d - 1$.

Remark 2.3. One consequence of the previous example is an estimate of the n-widths of the solution manifold in a case where the parameter range is infinite dimensional. More precisely, assume that the diffusion coefficient a ranges in the set

$$\mathcal{A} = \left\{ a \in C^s(\overline{D}) \colon \|a\|_{C^s} \le B,\ 0 < r \le a \le R \right\}, \tag{2.6}$$

for some $B > 0$, where $C^s(\overline{D})$ denotes the usual Hölder space for some $0 < s \le 1$. For any $a \in \mathcal{A}$, there exists a sequence $(a_n)_{n \ge 1}$ of piecewise constant approximations for each uniform partition into n subintervals such that

$$\|a - a_n\|_{L^\infty} \le Bn^{-s}, \quad n \ge 1, \tag{2.7}$$

and such that we also have $0 < r \le a_n \le R$. Note that this rate is optimal since it is well known that $d_n(\mathcal{A})_{L^\infty} \sim n^{-s}$. Denoting by $u(a)$ the corresponding solution to the diffusion equation for this value of a, we know, see e.g. [9], that

$$\|u(a) - u(a_n)\|_V \le r^{-2}\|f\|_{V'}\|a - a_n\|_{L^\infty} \le Br^{-2}\|f\|_{V'}\, n^{-s}. \tag{2.8}$$

We have seen that each $u(a_n)$ is a linear combination of $2n - 1$ basis functions *independent* of a_n. It follows that

$$d_n(u(\mathcal{A}))_V \le Cn^{-s}, \quad n \ge 1. \tag{2.9}$$

In other words, in this particular case, the map $a \mapsto u(a)$ preserves the decay rate n^{-s} between the n-widths of $\mathcal{A}$ in $L^\infty(D)$ and of $u(\mathcal{A})$ in $V = H_0^1(D)$, respectively. A result of this form was proved in [8] for a general holomorphic map $a \mapsto u(a)$, however with a loss of 1 in the value of s.

7

3 Using power expansions

We return to the general formulation (1.16). It has been shown in [11] that under the assumption (1.14), u has a Taylor series expansion

$$u(y) = \sum_{\nu \in \mathbb{N}_0^d} t_\nu y^\nu, \quad t_\nu = \frac{1}{\nu!} \partial_\nu u(0), \tag{3.1}$$

converging unconditionally in $L^\infty(U, V)$. The Taylor coefficients t_ν can be computed by differentiating the equation

$$\Big(\bar{A} + \sum_{i=1}^d y_i A_i\Big) u(y) = f, \tag{3.2}$$

at $y = 0$. For $\nu = 0$ the null multi-index, we have $t_0 = \bar{A}^{-1} f$ and for other values of ν, application of the Leibniz rule gives

$$t_\nu = \sum_{i: \nu_i \neq 0} \bar{A}^{-1} A_i t_{\nu - e_i}, \tag{3.3}$$

where e_i denotes the Kroenecker vector with 1 at position i. Introducing the abbreviations

$$B_i := \bar{A}^{-1} A_i, \quad g := \bar{A}^{-1} f, \tag{3.4}$$

we find that the partial Taylor sums can be re-expressed in the form

$$u_k(y) := \sum_{|\nu| \leq k} t_\nu y^\nu = \sum_{\ell=0}^k \Big(\sum_{i=1}^d y_i B_i\Big)^\ell g. \tag{3.5}$$

Thus the Taylor expansion can also be interpreted as a Neumann series,

$$u(y) = \Big(I - \sum_{i=1}^d y_i B_i\Big)^{-1} \bar{A}^{-1} f = \sum_{\ell=0}^\infty \Big(\sum_{i=1}^d y_i B_i\Big)^\ell g. \tag{3.6}$$

We now introduce the inner product $\langle \cdot, \cdot \rangle_{\bar{A}} := \langle \bar{A} \cdot, \cdot \rangle$, and its corresponding norm $\| \cdot \|_{\bar{A}}$, on $V = H_0^1(D)$, which in the case $\bar{a} = 1$ coincide with the standard inner product and norm of V. More generally, we have the norm equivalence

$$r\|v\|_V^2 \leq \|v\|_{\bar{A}}^2 \leq R\|v\|_V^2. \tag{3.7}$$

The following proposition yields an estimate for the rate of convergence of the series (3.6).

Proposition 3.1. *Under the assumption* (1.14), *we have*

$$\Big\|\sum_{i=1}^d y_i B_i v\Big\|_{\bar{A}} \leq \rho \|v\|_{\bar{A}}, \quad v \in V, \ y \in U, \tag{3.8}$$

where

$$\rho := 1 - \frac{r}{\|\bar{a}\|_{L^\infty}} < 1. \tag{3.9}$$

Proof. Let $u := \sum_{i=1}^{d} y_i B_i v$, then by definition

$$\int_D \bar{a} \nabla u \cdot \nabla w \, dx = \sum_{i=1}^{d} y_i \int_D \psi_i \nabla v \cdot \nabla w \, dx, \quad w \in V. \tag{3.10}$$

Taking $w = u$, and using the Cauchy-Schwarz inequality, we obtain

$$\|u\|_{\bar{A}}^2 \leq \|u\|_{\bar{A}} \|v\|_{\bar{A}} \sup_{x \in D} \frac{1}{\bar{a}(x)} \sum_{i=1}^{d} |\psi_i(x)|. \tag{3.11}$$

Using (1.15) we arrive at the assertion. $\qquad\square$

Note that the total number of separable terms in u_k is bounded by

$$n = n(d, k) := \#\{\nu: |\nu| \leq k\} = \binom{k + d}{d}, \tag{3.12}$$

and we have $\frac{1}{d!} k^d \leq n(d, k) \leq (k + 1)^d$. Hence in view of Proposition 3.1, u_k provides an approximation with error

$$\|u - u_k\|_{L^\infty(U,V)} \leq \frac{\|g\|_{\bar{A}}}{1 - \rho} \rho^{k+1}, \quad \rho := 1 - \frac{r}{\|\bar{a}\|_\infty}, \tag{3.13}$$

and with rank

$$\mathrm{rank}(u_k) \leq n(d, k) \leq (k + 1)^d. \tag{3.14}$$

Setting $k(d, n) := \max\{k \in \mathbb{N}: n(d, k) \leq n\} \geq n^{\frac{1}{d}} - 1$, we find on the one hand that

$$\mathrm{rank}(u_{k(d,n)}) \leq n, \tag{3.15}$$

and on the other hand that

$$\left\|u - u_{k(d,n)}\right\|_{L^\infty(U,V)} \lesssim \rho^{k(d,n)+1} \leq e^{-|\ln \rho| \, n^{1/d}}. \tag{3.16}$$

In summary, (3.5) thus provides the bound

$$d_n\big(u(U)\big)_V \lesssim e^{-|\ln \rho| \, n^{1/d}} \tag{3.17}$$

for the n-widths of the solution manifold $u(U)$, where the multiplicative constant depends on g and ρ.

As an example of a rather general structural property of the problem that leads to a reduction in the n-widths, we consider the additional assumption

$$\sum_{i=1}^{d} A_i = \theta \bar{A}. \tag{3.18}$$

9

In our setting, this holds if $\psi_i = \bar{a}\varphi_i$ with the φ_i forming a scaled partition of unity, that is, $\sum_{i=1}^{d} \varphi_i = \theta$ for a $\theta \in]0,1[$. The case of piecewise constant a is a particular instance, where $\bar{a}$ is constant and $\psi_i = \bar{a}\theta\chi_{D_i}$ with $\{D_1, \ldots, D_d\}$ a partition of D.

Note that (3.18) is equivalent to $\sum_{i=1}^{d} B_i = \theta I$. Let us consider first the consequences in the case $d = 2$, where this implies

$$u_k(y) = \sum_{j=0}^{k}(y_1 B_1 + y_2 B_2)^j g = \sum_{j=0}^{k}\left(\theta y_2 I + (y_1 - y_2)B_1\right)^j g. \tag{3.19}$$

We can rewrite each term in the sum as

$$(y_1 B_1 + y_2 B_2)^j g = (\theta y_2 I + (y_1 - y_2)B_1)^j g = \sum_{\ell=0}^{j}(\theta y_2)^{j-\ell}\binom{j}{\ell}(y_1 - y_2)^\ell B_1^\ell g. \tag{3.20}$$

Exchanging the order of summations in the sums over ℓ and j, setting

$$v_\ell := B_1^\ell g, \quad \phi_{k,\ell}(y) := (y_1 - y_2)^\ell \sum_{j=\ell}^{k}(\theta y_2)^{j-\ell}\binom{j}{\ell} \tag{3.21}$$

we thus obtain

$$u_k(y) = \sum_{\ell=0}^{k}\phi_{k,\ell}(y)\, v_\ell. \tag{3.22}$$

Consequently, under the assumption (3.18) the rank of u_k has the linear bound

$$\operatorname{rank}(u_k) \leq k + 1, \tag{3.23}$$

which is obviously more favorable than the quadratic bound given by (3.14) in the generic case. The same reasoning can be applied when $d > 2$, leading to the following result.

Proposition 3.2. *If (3.18) holds, for each k there exist $v_\ell \in V$ and d-variate polynomials ϕ_ℓ, with $\ell = 1, \ldots, n(d-1, k)$, such that*

$$u_k(y) = \sum_{\ell=1}^{n(d-1,k)}\phi_{k,\ell}(y)\, v_\ell, \tag{3.24}$$

and therefore

$$\operatorname{rank}(u_k) \leq n(d-1, k) \leq (k+1)^{d-1}. \tag{3.25}$$

Proof. Analogously to the above considerations for $d = 2$, we obtain

$$\sum_{j=0}^{k}\left(\sum_{i=1}^{d} y_i B_i\right)^j = \sum_{\ell=0}^{k}\gamma_{k,\ell}(y_d)\left(\sum_{i=1}^{d-1}(y_i - y_d)B_i\right)^\ell, \quad \gamma_{k,\ell}(y_d) := \sum_{j=\ell}^{k}\binom{j}{\ell}(\theta y_d)^{j-\ell}. \tag{3.26}$$

10

For notational convenience, we define the operators $\tilde{B}_\kappa^{(\ell)}$ for multi-indices $\kappa = (\kappa_1, \ldots, \kappa_{d-1})$ such that

$$(y_1 B_1 + \ldots + y_{d-1} B_{d-1})^\ell = \sum_{|\kappa|=\ell} y^\kappa \tilde{B}_\kappa^{(\ell)}, \quad y^\kappa := y_1^{\kappa_1} \cdots y_{d-1}^{\kappa_{d-1}}. \tag{3.27}$$

Note that the $\tilde{B}_\kappa^{(\ell)}$ are thus precisely the sums of powers of B_i, $i = 1, \ldots, d-1$, that arise in a Taylor expansion of $u(y)$ with respect to $y_1, \ldots, y_{d-1}$. Using this to rewrite the right hand side in (3.26), we arrive at

$$u_k(y) = \sum_{|\kappa| \leq k} \gamma_{k,|\kappa|}(y_d)\, y^\kappa\, \tilde{B}_\kappa^{(|\kappa|)} g. \tag{3.28}$$

Let $\{\kappa_\ell\}_{\ell=1,\ldots,\hat{n}(d-1,k)}$ be an enumeration of $\{|\kappa| \leq k\}$. We then obtain the assertion by choosing

$$v_\ell := \tilde{B}_{\kappa_\ell}^{(|\kappa_\ell|)} g, \quad \phi_{k,\ell}(y) := \gamma_{k,|\kappa_\ell|}(y_d)\, y^{\kappa_\ell}. \qquad \square$$

As can be seen from the proof, the functions v_k in Proposition 3.2 are precisely the Taylor coefficients $t_{(\kappa_1,\ldots,\kappa_{d-1},0)}$ of $u(y)$. As a consequence of the proposition, under the additional assumption (3.18), the generic estimate (3.17) improves to

$$d_n\big(u(U)\big)_V \lesssim e^{-|\ln \rho|\, n^{1/(d-1)}}. \tag{3.29}$$

Note that in view of (1.8) and (1.10), this estimate implies that the singular values in (1.7) satisfy

$$\sum_{k>n} \sigma_k^2 \lesssim e^{-2|\ln \rho|\, n^{1/(d-1)}} \tag{3.30}$$

and in particular

$$\sigma_n \lesssim e^{-|\ln \rho|\, n^{1/(d-1)}}. \tag{3.31}$$

In these estimates, the multiplicative constants depend on g and ρ, similar to (3.17).

4   Using the geometry of the partition

As noted above, the assumption (3.18) can hold under more general conditions than a piecewise constant coefficient a. The reduction in the bound for the n-widths between (3.17) and (3.29), however, weakens rapidly as d increases. In particular, the number of terms required to meet a certain error bound still grows exponentially in d. In order to obtain stronger, but also more specialized results, we now focus on certain parametric forms of a which are described through a partition of D.

11

4.1 Steklov-Poincaré operators

For given $\bar{a} \in L^\infty(D)$ with ess inf $\bar{a} > 0$, we consider $\psi_j := \theta \chi_{D_j} \bar{a}$ with a fixed $\theta \in]0, 1[$, where for the moment, $\{D_i\}$ is a partition of D into general Lipschitz subdomains. We denote the skeleton of this partition by

$$\Gamma := \bigcup_{i=1}^{d} \partial D_i \setminus \partial D \tag{4.1}$$

and define V_Γ as the space of trace values of functions in V on Γ, also denoted by $V_\Gamma := H_{00}^{\frac{1}{2}}(\Gamma)$, which we equip with the norm

$$\|w\|_{V_\Gamma} := \min\{\|v\|_{\bar{A}} \colon v \in V, \ v_{|\Gamma} = w\}. \tag{4.2}$$

We also introduce for each $i = 1, \ldots, d$ the space $V_i := H_0^1(D_i)$ viewed as a closed subspace of $V = H_0^1(D)$ by regarding its elements as extended by zero to all of D. The direct sum $V_1 \oplus \cdots \oplus V_d$ is a closed subspace of V, and we denote by W its $\bar{A}$-orthogonal complement in V.

Furthermore, we introduce for each $i = 1, \ldots, d$ the $\bar{A}$-harmonic extension operator E_i from V_Γ to the space

$$H^1(D_i) \cap H_0^1(D) := \{v \in H^1(D_i) \colon v|_{\partial D} = 0\}, \tag{4.3}$$

which maps the trace value $v_\Gamma \in V_\Gamma$ to the extension $E_i v_\Gamma$ defined by

$$\int_{D_i} \bar{a} \nabla E_i v_\Gamma \cdot \nabla w \, dx = 0 \quad \text{for all } w \in V_i, \qquad E_i v|_{\partial D_i \cap \partial D} = 0, \ E_i v|_{\partial D_i \cap \Gamma} = v_\Gamma|_{\partial D_i}. \tag{4.4}$$

In the case where $\bar{a}$ is constant, this is the usual harmonic extension. We define the operator E from V_Γ to V that concatenates these extensions into a function defined on D by

$$E \colon v_\Gamma \mapsto E v_\Gamma, \qquad E v_\Gamma|_{D_i} := E_i v_\Gamma, \ i = 1, \ldots, d. \tag{4.5}$$

Note that the space W coincides with the range of E.

The operators $B_i = \bar{A}^{-1} A_i$ are self-adjoint in the $\bar{A}$-inner product, and for each i they have the properties

$$B_i|_{V_i} = I, \qquad B_i|_{V_j} = 0 \ \text{ for } j \neq i. \tag{4.6}$$

Consequently, these operators also map W to itself.

We now decompose $u(y)$ into the sum of its $\bar{A}$-orthogonal projections onto the V_i and W, according to

$$u(y) = u_W(y) + \sum_{i=1}^{d} u_i(y). \tag{4.7}$$

The component $u_i(y) \in V_i$ solves

$$(1 - \theta y_i) \int_{D_i} \bar{a} \nabla u_i(y) \cdot \nabla v \, dx = \int_{D_i} f v \, dx, \ v \in V_i, \quad \text{and} \quad u_i(y)|_{D \setminus D_i} = 0. \tag{4.8}$$

12

Each u_i is therefore of rank at most 1, depending only on y_i, and $\sum_{i=1}^d u_i$ therefore has rank at most d.

The component u_W can also be written in the form $u_W(y) = E u_\Gamma(y)$ for a $u_\Gamma(y) \in V_\Gamma$, where $u_\Gamma(y)$ is characterized by the variational problem

$$\sum_{i=1}^d (1 - \theta y_i) \int_{D_i} \bar{a} \nabla E u_\Gamma(y) \cdot \nabla E v_\Gamma \, dx = \int_D f \, E v_\Gamma \, dx \,, \quad v_\Gamma \in V_\Gamma \,. \tag{4.9}$$

To obtain a more concise formulation of the equation for u_Γ, we introduce the Steklov-Poincaré operators $S_i \colon V_\Gamma \to V_\Gamma'$, $i = 1, \ldots, d$, and $\bar{S} \colon V_\Gamma \to V_\Gamma'$ defined by

$$\langle S_i v_\Gamma, w_\Gamma \rangle := \int_{D_i} \bar{a} \nabla E v_\Gamma \cdot \nabla E w_\Gamma \, dx, \quad v_\Gamma, w_\Gamma \in V_\Gamma, \tag{4.10}$$

and

$$\bar{S} := \sum_{i=1}^d S_i \,. \tag{4.11}$$

The norm of V_Γ is induced by the inner product

$$(u, v)_\Gamma := \langle \bar{S} u, v \rangle = \int_D \bar{a} \nabla E u \cdot \nabla E v \, dx \,. \tag{4.12}$$

In particular, the operator $\bar{S}$ is bounded, elliptic, and defines an isometry between V_Γ and V_Γ'.

Using the operators S_i and $\bar{S}$, we can now rewrite (4.9) as

$$\left(\bar{S} - \theta \sum_{i=1}^d y_i S_i \right) u_\Gamma(y) = f_\Gamma \,, \tag{4.13}$$

where $f_\Gamma \in V_\Gamma'$ is defined by

$$\langle f_\Gamma, v_\Gamma \rangle := \int_D f \, E v_\Gamma \, dx \,, \quad v_\Gamma \in V_\Gamma \,. \tag{4.14}$$

We thus obtain the Neumann series representation

$$u_\Gamma = \lim_{k \to \infty} u_{k,\Gamma}, \qquad u_{k,\Gamma}(y) := \sum_{\ell=0}^k \left(\sum_{i=1}^d y_i \left(\theta \bar{S}^{-1} S_i \right) \right)^\ell g_\Gamma, \tag{4.15}$$

where $g_\Gamma := \bar{S}^{-1} f_\Gamma$. By a similar argument as in the proof of Proposition 3.1, we obtain that

$$\left\| \sum_{i=1}^d y_i \theta \bar{S}^{-1} S_i v \right\|_{V_\Gamma} \leq \theta \, \|v\|_{V_\Gamma} \,, \quad v \in V_\Gamma, \ y \in U, \tag{4.16}$$

13

which shows that the Neumann series converges in $L^\infty(U, V_\Gamma)$.

Note that the Neumann series for u can be rewritten as

$$u_k(y) = E u_{k,\Gamma}(y) + \sum_{i=1}^{d} u_{k,i}(y), \tag{4.17}$$

where $u_{k,i}(y)$ is the $\bar{A}$-orthogonal projection of $u_k(y)$ onto V_i. Similar to u_i, the rank of $u_{k,i}$ is at most 1. Therefore

$$\operatorname{rank}(u_k) \leq d + \operatorname{rank}(u_{k,\Gamma}). \tag{4.18}$$

The task of obtaining favorable bounds for the rank of u_k is thus in the present setting essentially reduced to obtaining similar such bounds for the rank of $u_{k,\Gamma}$, which is defined on $\Gamma \times U$. We shall see next that what one obtains in this regard can depend strongly on the particular geometry of D and Γ.

4.2 A particular example: four squares

We now turn to the case of piecewise constant coefficients, with each piece varying within the same range of values, where we may assume without loss of generality that $\bar{a} = 1$. We consider the case of the unit square $D =]-\frac{1}{2}, \frac{1}{2}[^2$, partitioned into the four symmetric quadrants,

$$D_1 =]-\tfrac{1}{2}, 0[^2, \quad D_2 =]0, \tfrac{1}{2}[\times]-\tfrac{1}{2}, 0[, \quad D_3 =]-\tfrac{1}{2}, 0[\times]0, \tfrac{1}{2}[, \quad D_4 =]0, \tfrac{1}{2}[^2, \tag{4.19}$$

and with $\psi_i := \theta \chi_{D_i}$, $\theta \in]0, 1[$.

In order to motivate what follows, we first consider a numerical experiment, where we compare the decay of the V-norms of the tensor product Legendre coefficients u_ν as in (1.3) to the decay of singular values σ_k in the decomposition (1.7). We execute these computations under a *fixed* space discretization through a finite element space V_h: the exact solution map $y \mapsto u(y)$ is replaced by $y \mapsto u_h(y)$, where $u_h(y)$ is the Galerkin projection of $u(y)$ onto the space V_h. We compare the results obtained for (a) the symmetric partition (4.19) to those for (b) a non-symmetric partition into four quadrilaterals which may be viewed as a distorted version of (4.19). As explained further in Section 5, in order to obtain result that reflect the behavior of the Legendre coefficients and singular values for the original solution map $y \mapsto u(y)$, it is extremely important to resolve the subdomain interfaces appropriately in the spatial discretization. Figure 4.2 displays the typical meshes that we use for the symmetric partition (4.19) and the distorted partition. We employ piecewise linear finite elements on further refined meshes, which gives 59 365 degrees of freedom, 4 105 of these on subdomain interfaces. For the discretization in the parametric variable, we use Galerkin projection onto tensor product Legendre polynomials of total degree at most 15. We therefore compute $\binom{19}{4} = 3876$ Legendre coefficients $u_{h,\nu}$, each of which is a finite element function from V_h. A detailed discussion of the precise computational procedure is deferred to Section 5. The given results are obtained with the choices $f := 1$ and $\theta := \frac{1}{2}$.
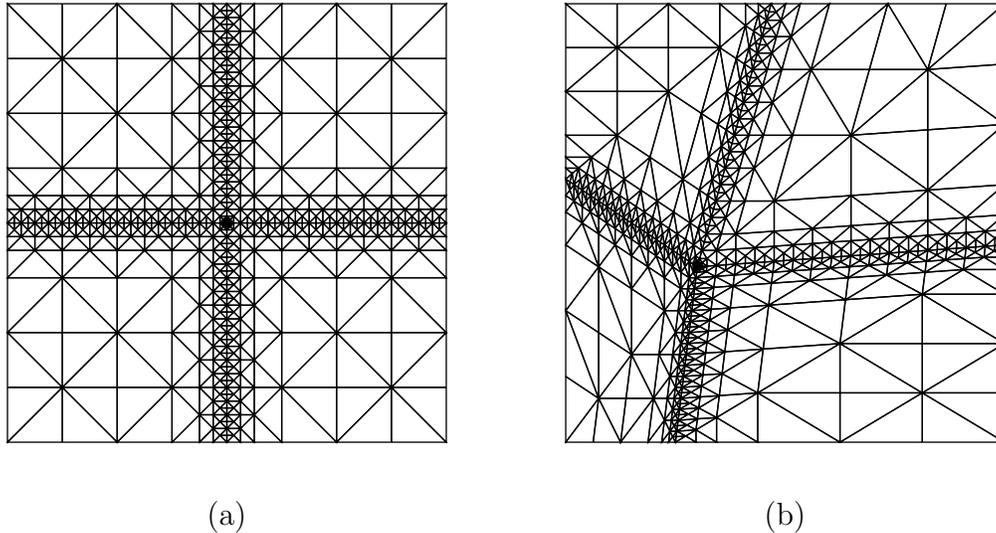
(a)	(b)

Figure 4.2: Example meshes for the two test geometries with 992 triangles and 521 vertices. The meshes used in the actual computation have $119\,272$ triangles and $59\,909$ vertices with a similar gradation towards subdomain boundaries.

In Figure 4.3, we compare the decay of the V-norms of the Legendre coefficients $u_{h,\nu}$ to the decay of singular values σ_k in the decomposition (1.7) where we have replaced V by V_h and $u(y)$ by $\sum_{|\nu|\le 15} u_{h,\nu} L_\nu(y)$ in the definition of $\mathcal{M}_u$. In both cases (a) and (b), the second sequence decays substantially faster. However, although the Legendre coefficients behave almost identically in the two cases, we observe an exponential decrease of the σ_k in case (a), but only a subexponential decrease in case (b). A closer inspection shows that the decay in case (b) is in fact not substantially better than guaranteed by (3.31).

As explained in Section 5, the Galerkin discretization in the Legendre basis up to some total order ℓ allows us to exactly compute the truncated power series $u_{h,\ell}(y) = \sum_{|\nu|\le\ell} t_{h,\nu} y^\nu$ associated to the map $y \mapsto u_h(y)$. Figure 4.4 displays the ranks of the partial sums $u_{h,k}$ for $k = 1,\dots,10$ together with the corresponding singular values. In case (a), these ranks are observed to increase approximately linearly in k before the smallest non-zero singular values drop below machine precision, whereas the ranks grow more rapidly (faster than quadratically) in case (b). Further experiments reveal that there is no qualitative change in the results if one prescribes a less regular right hand side f, which shows that the decay of singular values is not tied to the spatial smoothness of u.

Our aim is now to explain the substantially better low-rank approximability observed in the case of a symmetric geometry. The remainder of this section is devoted to the proof of the following theorem, which requires several preparatory steps. Note that under our present assumptions, $\|\cdot\|_{\bar{A}}$ equals the standard norm on $H_0^1(D)$.

Theorem 4.1. *Let $d = 4$, and let D_i, $i = 1,\dots,4$, be defined as in (4.19). For each $k \in \mathbb{N}$, and for $n(k) := 8k + 5$, there exist $v_1,\dots,v_{n(k)} \in V$ and d-variate polynomials $\phi_1,\dots,\phi_{n(k)}$,*
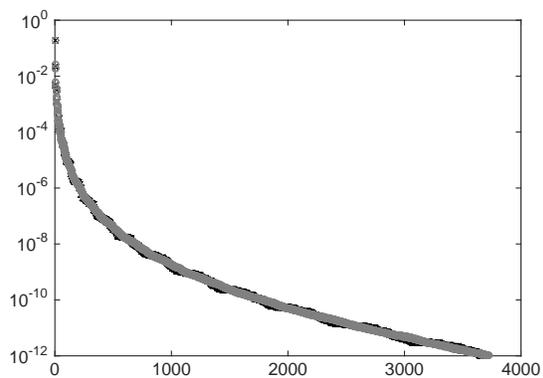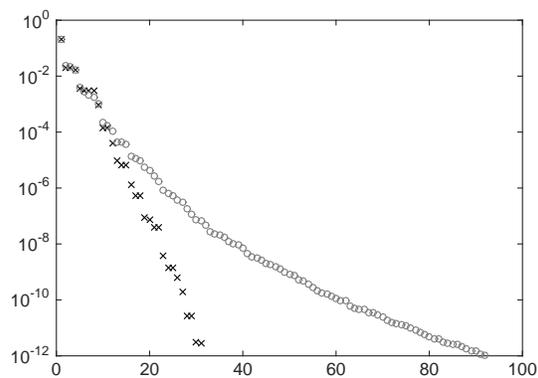
Sorted H^1-norms of Legendre coefficients　　　　　　　Singular values

Figure 4.3: Decay of sorted H^1-norms of Legendre coefficients and of singular values, both for geometries (a) and (b) in Figure 4.2 (marked by $\times$ and $\circ$, respectively).
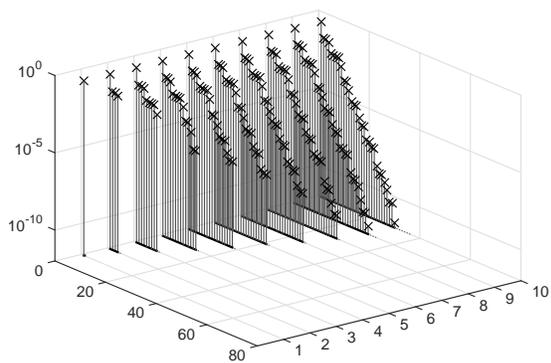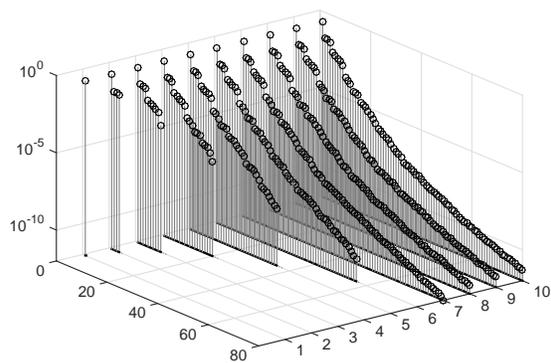
(a)　　　　　　　　　　　　　　　　　　　　(b)

Figure 4.4: Ranks and first 80 singular values of the partial sums of the Neumann series for orders $k = 1, \ldots, 10$, and for geometries (a) and (b) in Figure 4.2 (marked by $\times$ and $\circ$, respectively).

each of total degree k, such that the truncated power series $u_k(y) = \sum_{|\nu| \leq k} t_\nu y^\nu$ has the exact expression

$$u_k(y) = \sum_{\ell=1}^{n(k)} \phi_\ell(y)\, v_\ell. \tag{4.20}$$

Therefore $\operatorname{rank}(u_k) \leq 8k + 5$.

Combining this result with the approximation estimate (3.13) between u and u_k immediately yields the following corollary.

Corollary 4.2. *With the same notations as in Theorem 4.1, we have*

$$\sup_{y \in U} \left\| u(y) - \sum_{\ell=1}^{n(k)} \phi_\ell(y)\, v_\ell \right\|_V \leq \frac{\left\| \bar{A}^{-1} f \right\|_V}{1 - \theta}\, \theta^{k+1}, \tag{4.21}$$

and consequently $d_n\big(u(U)\big)_V \lesssim \exp\big(-\frac{|\ln\theta|}{8} n\big)$.

We begin by introducing the notation

$$\Gamma_1 := [-\tfrac{1}{2}, \tfrac{1}{2}] \times \{0\}, \quad \Gamma_2 := \{0\} \times [-\tfrac{1}{2}, \tfrac{1}{2}], \tag{4.22}$$

for the horizontal and vertical components of the skeleton Γ of the decomposition (4.19). The space V_Γ introduced at the beginning of this section can be decomposed into the three closed subspaces

$$\begin{aligned}
\hat{V}_1 &:= \{v \in V_\Gamma : v|_{\Gamma_1} \text{ and } v|_{\Gamma_2} \text{ are even}\}, \\
\hat{V}_2 &:= \{v \in V_\Gamma : v|_{\Gamma_1} \text{ odd}, v|_{\Gamma_2} = 0\}, \\
\hat{V}_3 &:= \{v \in V_\Gamma : v|_{\Gamma_2} \text{ odd}, v|_{\Gamma_1} = 0\}.
\end{aligned} \tag{4.23}$$

The following result decribes the action of E on these spaces. Its proof is immediate from symmetry arguments and is therefore omitted.

Proposition 4.3. *For the harmonic extension Ew of $w \in V_\Gamma$, we have the following symmetry and antisymmetry properties depending on the particular subspaces: if $w \in \hat{V}_1$, then $Ew(-x_1, x_2) = Ew(x_1, x_2)$ and $Ew(x_1, -x_2) = Ew(x_1, x_2)$ for a.e. $x \in D$; if $w \in \hat{V}_2$, then $Ew(-x_1, x_2) = -Ew(x_1, x_2)$ and $Ew(x_1, -x_2) = Ew(x_1, x_2)$; and if $w \in \hat{V}_3$, then $Ew(x_1, -x_2) = -Ew(x_1, x_2)$ and $Ew(-x_1, x_2) = Ew(x_1, x_2)$.*

For the proofs of the following results, it will be convenient to introduce the notation

$$(u, v)_{\Gamma, i} := \int_{D_i} \nabla Eu \cdot \nabla Ev\, dx, \tag{4.24}$$

so that $(u, v)_\Gamma = \sum_{i=1}^{4} (u, v)_{\Gamma, i}$.

Proposition 4.4. *The spaces $\hat{V}_i$, $i = 1, 2, 3$, form an orthogonal decomposition of V_Γ with respect to the inner product $(\cdot, \cdot)_\Gamma$.*

Proof. Note first that a unique decomposition into even and odd parts of this form always exists. To show orthogonality, for symmetry reasons, it suffices to consider the pairings $\hat{V}_1$, $\hat{V}_2$ and $\hat{V}_2$, $\hat{V}_3$. If either $u \in \hat{V}_3$, $v \in \hat{V}_2$ or $u \in \hat{V}_1$, $v \in \hat{V}_2$, Proposition 4.3 gives

$$(u, v)_{\Gamma,2} = -(u, v)_{\Gamma,1}, \quad (u, v)_{\Gamma,3} = -(u, v)_{\Gamma,4}, \tag{4.25}$$

and consequently $(u, v)_\Gamma = 0$. □

We now consider the following operators:

$$H_0 := \bar{S}, \qquad\qquad H_2 := S_1 + S_2 - S_3 - S_4,$$
$$H_1 := S_1 - S_2 - S_3 + S_4, \qquad H_3 := S_1 - S_2 + S_3 - S_4.$$

For convenience, we set $G_i := \theta \bar{S}^{-1} H_i$, $i = 0, \ldots, 3$. For each k, the partial sums $u_{k,\Gamma}$ introduced in (4.15) can be expressed in terms of the operators G_i. We introduce the new variables

$$z_0 = z_0(y) := \frac{1}{4}(y_1 + y_2 + y_3 + y_4), \qquad z_2 = z_2(y) := \frac{1}{4}(y_1 + y_2 - y_3 - y_4),$$

$$z_1 = z_1(y) := \frac{1}{4}(y_1 - y_2 - y_3 + y_4), \qquad z_3 = z_3(y) := \frac{1}{4}(y_1 - y_2 + y_3 - y_4),$$

so that

$$\sum_{i=0}^{3} z_i G_i = \sum_{i=1}^{4} y_i (\theta \bar{S}^{-1} S_i). \tag{4.26}$$

In view of (4.15), we thus obtain the representation

$$u_{k,\Gamma}(z) = \sum_{\ell=0}^{k} \left(\sum_{i=0}^{3} z_i G_i \right)^{\ell} g_\Gamma. \tag{4.27}$$

As we shall now show, the operators G_i act on the spaces (4.23) in a very particular way, which will eventually allow us to obtain the desired rank estimate.

Lemma 4.5. *One has $G_i(\hat{V}_i) = \{0\}$ for $i = 1, 2, 3$.*

Proof. Since $\bar{S}$ is an isomorphism from V_Γ to V_Γ', we prove the equivalent statement that $H_i(\hat{V}_i) = \{0\}$ for $i = 1, 2, 3$. Let $u_1 \in \hat{V}_1$ and $v \in V_\Gamma$. Then by definition,

$$\langle H_1 u_1, v \rangle = (u_1, v)_{\Gamma,1} - (u_1, v)_{\Gamma,2} - (u_1, v)_{\Gamma,3} + (u_1, v)_{\Gamma,4}. \tag{4.28}$$

If $v \in \hat{V}_1$, the four integrals on the right hand side all give the same value, and hence $\langle H_1 u_1, v \rangle = 0$; if $v \in \hat{V}_2$ or $v \in \hat{V}_3$, then using a change of variables and the above antisymmetry properties, we obtain

$$(u_1, v)_{\Gamma,4} = -(u_1, v)_{\Gamma,1}, \quad (u_1, v)_{\Gamma,3} = -(u_1, v)_{\Gamma,2}. \tag{4.29}$$

18

We have thus shown $\langle H_1 u_1, v \rangle = 0$ for any $v \in V_\Gamma$, or equivalently that $H_1(\hat{V}_1) = \{0\}$. Next, let $u_2 \in \hat{V}_2$. Then

$$\langle H_2 u_2, v \rangle = (u_2, v)_{\Gamma,1} + (u_2, v)_{\Gamma,2} - (u_2, v)_{\Gamma,3} - (u_2, v)_{\Gamma,4} . \tag{4.30}$$

If $v \in \hat{V}_1$ or $v \in \hat{V}_3$, we have

$$(u_2, v)_{\Gamma,2} = -(u_2, v)_{\Gamma,1} , \quad (u_2, v)_{\Gamma,4} = -(u_2, v)_{\Gamma,3} . \tag{4.31}$$

If $v \in \hat{V}_2$, again all four integrals give the same value. Thus for all $v \in V_\Gamma$, $\langle H_2 u_2, v \rangle = 0$, or equivalently $H_2(\hat{V}_2) = \{0\}$. The statement $H_3(\hat{V}_3) = \{0\}$ follows by exchanging the role of the coordinates (x_1, x_2). $\qquad \square$

Lemma 4.6. *One has*

$$G_2(\hat{V}_1), G_1(\hat{V}_2) \subset \hat{V}_3 , \quad G_3(\hat{V}_1), G_1(\hat{V}_3) \subset \hat{V}_2 , \quad \text{and} \quad G_3(\hat{V}_2), G_2(\hat{V}_3) \subset \hat{V}_1 . \tag{4.32}$$

Proof. By the symmetries in the problem, it suffices to consider $G_2(\hat{V}_1)$, $G_1(\hat{V}_2)$, and $G_3(\hat{V}_2)$. We begin with the following observation: if $u_j, v_j \in \hat{V}_j$, then

$$\langle H_i u_j, v_j \rangle = 0 , \quad \text{for all } i, j \in \{1, 2, 3\}. \tag{4.33}$$

To see this, we argue as in the proof of Lemma 4.5, noting that

$$(u_j, v_j)_{\Gamma,n} = (u_j, v_j)_{\Gamma,m} , \quad n, m \in \{1, \ldots, 4\} , \tag{4.34}$$

since $E u_j$ and $E v_j$ have the same symmetry or antisymmetry properties. We now treat the three above cases.

(i) $G_2(\hat{V}_1) \subset \hat{V}_3$: let $v_1 \in \hat{V}_1$. In view of Proposition 4.4, we need to show that $H_2 v_1$ annihilates the elements in $\hat{V}_1$ and $\hat{V}_2$. With $w_1 \in \hat{V}_1$ and $w_2 \in \hat{V}_2$, we have $\langle H_2 v_1, w_2 \rangle = \langle v, H_2 w_2 \rangle = 0$ from Lemma 4.5, and $\langle H_2 v_1, w_1 \rangle = 0$ follows from (4.33). Altogether, this gives $G_2(\hat{V}_1) \subset \hat{V}_3$.

(ii) $G_1(\hat{V}_2) \subset \hat{V}_3$: let $v_2 \in \hat{V}_2$. With $w_1 \in \hat{V}_1$ and $w_2 \in \hat{V}_2$, we again have $\langle H_1 v_2, w_1 \rangle = \langle v_2, H_1 w_1 \rangle = 0$, and $\langle H_1 v_2, w_2 \rangle = 0$ by (4.33), showing $G_1(\hat{V}_2) \subset \hat{V}_3$.

(iii) $G_3(\hat{V}_2) \subset \hat{V}_1$: let $v_2 \in \hat{V}_2$. With $w_2 \in \hat{V}_2$ and $w_3 \in \hat{V}_3$, we have $\langle H_3 v_2, w_2 \rangle = 0$ by (4.33) and $\langle H_3 v_2, w_3 \rangle = \langle v_2, H_3 w_3 \rangle = 0$, which shows $G_3(\hat{V}_2) \subset \hat{V}_1$. $\qquad \square$

Lemma 4.7. *For any $v_2 \in \hat{V}_2$, $v_3 \in \hat{V}_3$, one has*

$$G_2 G_3 v_2 = G_1 v_2 , \quad G_3 G_2 v_3 = G_1 v_3 , \quad G_3^2 v_2 = v_2 , \quad G_2^2 v_3 = v_3 . \tag{4.35}$$

Proof. By symmetry arguments, it suffices to show the first and the third statement. Let $p_3 := G_2 G_3 v_2 \in \hat{V}_3$, $p_1 := G_3 v_2 \in \hat{V}_1$, and $q_3 := G_1 v_2 \in \hat{V}_3$. The first statement now reads $p_3 = q_3$. By definition,

$$\langle \bar{S} p_3, u \rangle = \langle H_2 p_1, u \rangle , \quad u \in V_\Gamma, \tag{4.36}$$

19

$$\langle \bar{S}p_1, v \rangle = \langle H_3 v_2, v \rangle \,, \quad v \in V_\Gamma \,, \tag{4.37}$$

as well as

$$\langle \bar{S}q_3, w \rangle = \langle H_1 v_2, w \rangle \,, \quad w \in V_\Gamma \,. \tag{4.38}$$

For given $u \in \hat{V}_3$, we now set

$$v := (\chi_{\{x \in D \colon x_2 \leq 0\}} - \chi_{\{x \in D \colon x_2 > 0\}}) u \in \hat{V}_1 \,. \tag{4.39}$$

For this choice of v, the definition of $\{H_1, H_2, H_3\}$ shows that we have $\langle H_2 p_1, u \rangle = \langle \bar{S}p_1, v \rangle$ as well as $\langle H_3 v_2, v \rangle = \langle H_1 v_2, u \rangle$. Using (4.36) and (4.37) this implies $\langle \bar{S}p_3, u \rangle = \langle H_1 v_2, u \rangle$ for any $u \in \hat{V}_3$. By (4.38), we thus reach

$$\langle \bar{S}p_3, u \rangle = \langle \bar{S}q_3, u \rangle, \quad u \in \hat{V}_3. \tag{4.40}$$

By Lemma 4.5, both sides in the above equality vanish in the the case where $u \in \hat{V}_1 \oplus \hat{V}_2$. Since $\bar{S}$ is an isomorphism, it follows that $p_3 = q_3$.

Let $p_2 := G_3^2 v_2$ and $q_1 := G_3 v_2$. The third statement reads $p_2 = v_2$. By definition of p_2 and q_1,

$$\langle \bar{S}p_2, u \rangle = \langle H_3 q_1, u \rangle, \quad u \in V_\Gamma \,, \tag{4.41}$$

$$\langle \bar{S}q_1, v \rangle = \langle H_3 v_2, v \rangle \,, \quad v \in V_\Gamma \,. \tag{4.42}$$

For $u \in \hat{V}_2$, we set

$$v := (\chi_{\{x \in D \colon x_1 \leq 0\}} - \chi_{\{x \in D \colon x_1 > 0\}}) u \in \hat{V}_1. \tag{4.43}$$

For this choice of v, the definition of H_3 shows that we have $\langle H_3 q_1, u \rangle = \langle \bar{S}q_1, v \rangle$ and $\langle H_3 v_2, v \rangle = \langle \bar{S}v_2, u \rangle$. Using (4.41) and (4.42), we reach

$$\langle \bar{S}p_2, u \rangle = \langle \bar{S}v_2, u \rangle, \quad u \in \hat{V}_2. \tag{4.44}$$

By Lemma 4.5, both sides in the above equality vanish in the the case where $u \in \hat{V}_1 \oplus \hat{V}_3$. Since $\bar{S}$ is an isomorphism, it follows that $p_2 = v_2$. $\qquad\square$

Recall that our aim is to bound the rank of the representation (4.27) for each order k, which is bounded by the number of linearly independent elements of V_Γ that are created applying compositions of the operators G_i, $i = 1, 2, 3$, up to order k, to g_Γ. In other words, defining

$$F_k(v) := \{v\} \cup \{G_{i_1} \cdots G_{i_j} v \colon 1 \leq j \leq k, \ i_1, \ldots, i_j \in \{1, 2, 3\}\} \tag{4.45}$$

for any given $v \in V_\Gamma$, we have $\mathrm{rank}(u_{k,\Gamma}) \leq n(k) := \dim \mathrm{span}\, F_k(g_\Gamma)$. Note that this strategy for obtaining rank bounds can be regarded as the opposite of the one used in Section 3, which was based on instead grouping terms in the partial sums u_k according to common polynomial degrees.

One has the trivial bound

$$n(k) \leq \#F_k(g_\Gamma) \leq 3^k + 1. \tag{4.46}$$

$$
\begin{array}{ccc}
g_1 & g_2 \xleftarrow{G_3} & g_3 \xleftarrow{G_2} \\
\swarrow \qquad \searrow & \swarrow \qquad & \swarrow \\
G_2 g_1 \qquad\qquad G_3 g_1 & G_1 g_2 \xleftarrow{G_2} G_3 g_2 & G_1 g_3 \xleftarrow{G_3} G_2 g_3 \\
\end{array}
$$

g_1

$G_2 g_1 \xrightarrow{\ G_2\ }$ $G_3 g_1 \xrightarrow{\ G_3\ }$

$G_1 G_2 g_1 \xleftarrow{G_3} G_2^2 g_1$ $G_1 G_3 g_1 \xleftarrow{G_2} G_3^2 g_1$

$G_1^2 G_2 g_1 \xleftarrow{G_2} G_3 G_1 G_2 g_1$ $G_1^2 G_3 g_1 \xleftarrow{G_3} G_2 G_1 G_3 g_1$

$g_2 \xleftarrow{G_3}$

$G_1 g_2 \xleftarrow{G_2} G_3 g_2 \quad\xrightarrow{G_2}$

$G_1^2 g_2 \xleftarrow{G_3} G_2 G_1 g_2 \quad\xrightarrow{G_3}$

$G_1^3 g_2 \xleftarrow{G_2} G_3 G_1^2 g_2$

$g_3 \xleftarrow{G_2}$

$G_1 g_3 \xleftarrow{G_3} G_2 g_3 \quad\xrightarrow{G_3}$

$G_1^2 g_3 \xleftarrow{G_2} G_3 G_1 g_3 \quad\xrightarrow{G_2}$

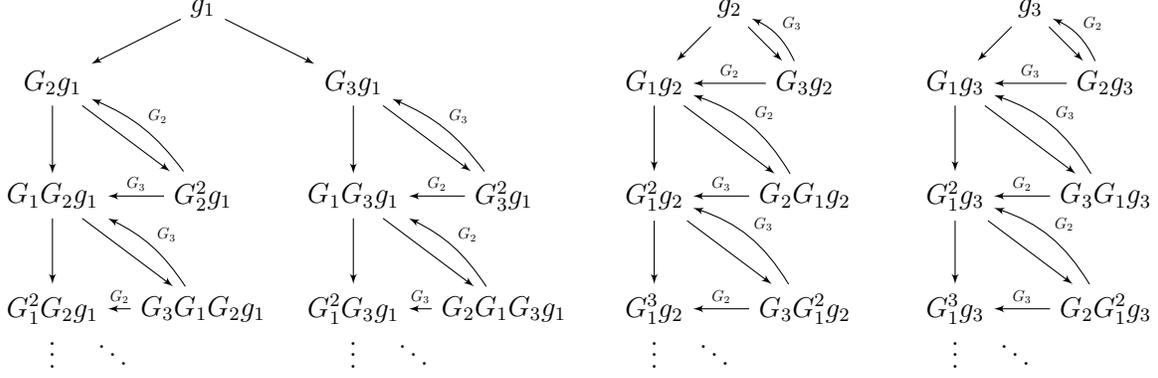$G_1^3 g_3 \xleftarrow{G_3} G_2 G_1^2 g_3$

$\vdots \quad \ddots$

Figure 4.5: Illustration of the proof of Lemma 4.8, where $g_\Gamma = g_1 + g_2 + g_3$.

Using the decomposition

$$g_\Gamma = g_1 + g_2 + g_3, \tag{4.47}$$

with $g_i \in \hat{V}_i$, we obviously have

$$\operatorname{span} F_k(g_\Gamma) \subset \operatorname{span}\{F_k(g_1) \cup F_k(g_2) \cup F_k(g_3)\}, \tag{4.48}$$

and therefore

$$n(k) \leq n_1(k) + n_2(k) + n_3(k), \quad n_i(k) := \dim \operatorname{span} F_k(g_i). \tag{4.49}$$

Lemmas 4.5 and 4.6 allow us to bound the number of non-zero elements in $F_k(g_i)$, and therefore $n_i(k)$, by $2^{k+1} - 1$. We thus obtain the slightly better estimate

$$n(k) \leq 3(2^{k+1} - 1). \tag{4.50}$$

Note that these estimates are both much weaker than the bound of order k^3 that we have for $\operatorname{rank}(u_k)$ as a consequence of the results in Section 3. However, with the help of Lemma 4.7, we now show that the number of non-zero elements in each $F_k(g_i)$ in fact grows only *linearly* in k, which explains our initial numerical observations and leads us to the proof of Theorem 4.1.

Lemma 4.8. *Let $g_2 \in \hat{V}_2$. Then for each $k \in \mathbb{N}$, the only non-zero elements in $F_k(g_2)$ are*

$$\left\{G_1^i g_2 : 0 \leq i \leq k\right\} \cup \left\{G_2 G_1^i g_2 : i \text{ odd}, 1 \leq i < k\right\} \cup \left\{G_3 G_1^i g_2 : i \text{ even}, 0 \leq i < k\right\}. \tag{4.51}$$

Let $g_3 \in \hat{V}_3$. Then for each $k \in \mathbb{N}$, the only non-zero elements in $F_k(g_3)$ are

$$\left\{G_1^i g_3 : 0 \leq i \leq k\right\} \cup \left\{G_3 G_1^i g_3 : i \text{ odd}, 1 \leq i < k\right\} \cup \left\{G_2 G_1^i g_3 : i \text{ even}, 0 \leq i < k\right\}. \tag{4.52}$$

The idea behind the proof of Lemma 4.8, which combines Lemmas 4.5, 4.6, and 4.7, is illustrated on the two graphs on the right of Figure 4.5. The nodes in these graphs represent

the only possible non-zero functions that can be created by repeated application of the operators G_i, $i = 1, 2, 3$, to a given $g_2 \in \hat{V}_2$ or $g_3 \in \hat{V}_3$. The first graph on the left represents the only possible non-zero functions that can be created by repeated application of the G_i to a given $g_1 \in \hat{V}_1$. This first graph is actually obtained by applying the third and second graphs on its first two nodes $G_2 g_1 \in \hat{V}_3$ and $G_3 g_1 \in \hat{V}_2$.

Proof. It suffices to prove the result for $F_k(g_2)$, since the result for $F_k(g_3)$ follows by symmetry. For each k, we use the abbreviation $F_k := F_k(g_2)$ and denote the set in (4.51) by

$$R_k = R_{k,1} \cup R_{k,2} \cup R_{k,3}. \tag{4.53}$$

We show that $F_k = R_k \cup \{0\}$ for all $k \geq 1$. In the case $k = 1$, we have

$$F_k(g_2) = \{g_2, G_1 g_2, G_2 g_2, G_3 g_2\} = \{g_2, G_1 g_2, 0, G_3 g_2\}, \tag{4.54}$$

which proves the claim in this case. We assume the claim holds for a $k \geq 1$, and we thus need to show that $F_{k+1} = R_{k+1} \cup \{0\}$. The inclusion $R_{k+1} \cup \{0\} \subset F_{k+1}$ follows trivially from the definitions. To show that $F_{k+1} \subset R_{k+1} \cup \{0\}$, note first that from the definition of F_k, we have

$$F_{k+1} = \{g_2\} \cup \bigcup_{i=1}^{3} G_i(F_k) \tag{4.55}$$

and hence, by our induction hypothesis,

$$F_{k+1} = \{g_2\} \cup \bigcup_{i=1}^{3} G_i(R_k \cup \{0\}). \tag{4.56}$$

By Lemma 4.6, $G_1^i g_2 \in V_2$ for i even and $G_1^i g_2 \in V_3$ for i odd. Thus

$$G_1(R_{k,1}) \subset R_{k,1} \cup \{G_1^{k+1} g_2\} = R_{k+1,1},$$
$$G_2(R_{k,1}) = \{0\} \cup \{G_2 G_1^i g_2 \colon i \text{ odd}, 1 \leq i \leq k\} = \{0\} \cup R_{k+1,2},$$
$$G_3(R_{k,1}) = \{0\} \cup \{G_3 G_1^i g_2 \colon i \text{ even}, 0 \leq i \leq k\} = \{0\} \cup R_{k+1,3}.$$

By Lemma 4.6, $G_1(R_{k,2}) = G_1(R_{k,3}) = \{0\}$ and by Lemma 4.7, $G_i(R_{k,j}) \subset \{0\} \cup R_{k+1,1}$ for $i, j \in \{2, 3\}$. Since furthermore $\{g_2\} \subset R_{k+1,1}$ for each k, we arrive at $F_{k+1} = R_{k+1} \cup \{0\}$. $\square$

With Lemma 4.8 at hand, we can now prove Theorem 4.1.

Proof of Theorem 4.1. Invoking (4.18), we know that

$$\mathrm{rank}(u_k) \leq 4 + \mathrm{rank}(u_{k,\Gamma}), \tag{4.57}$$

and by the representation (4.27) the rank of $u_{k,\Gamma}$ is bounded by the number of linearly independent vectors in $F_k(g_\Gamma)$.

After an orthogonal decomposition of g_Γ in the form $g_\Gamma = g_1 + g_2 + g_3$ with $g_i \in \hat{V}_i$, Lemma 4.8 can be applied separately to $G_3 g_1$, $G_2 g_1$, g_2, g_3. This shows that the number of non-zero elements in $F_k(g_i)$ is at most $(2k + 1)$ for $i = 2$ and $i = 3$, and at most $1 + 2(2k - 1)$ for $i = 1$. In conclusion, $F_k(g_\Gamma)$ contains at most $1 + 2(2k - 1) + 2(2k + 1) = 8k + 1$ linearly independent elements, which confirms the estimate $8k + 5$ for $\mathrm{rank}(u_k)$. $\square$

5   Numerical construction of low-rank approximations

The numerical scheme used for the computational examples in this work is based on the iteration

$$u_0 := \bar{A}^{-1} f = g\,, \qquad u_k := \bar{A}^{-1}\Big(f + \sum_{i=1}^{d} y_i A_i u_{k-1}\Big) = g + \sum_{i=1}^{d} y_i B_i u_{k-1}\,, \ k \in \mathbb{N}\,. \tag{5.1}$$

As a consequence of Proposition 3.1, this is a fixed point iteration with linear convergence in $L^\infty(U, V)$. The iterates u_k are precisely the partial sums (3.5), that is,

$$u_k(y) = \sum_{|\nu| \le k} t_\nu y^\nu = \sum_{\ell=0}^{k}\Big(\sum_{i=1}^{d} y_i B_i\Big)^\ell g\,. \tag{5.2}$$

Since we shall use the singular value decomposition to obtain low-rank approximations, we regard u and the u_k as elements of $L^2(U, V) \simeq V \otimes L^2(U)$ with the uniform probability measure μ on U, and consider convergence in this norm.

We first introduce a discretization in the parametric variable y, based on a choice of basis for $L^2(U)$. We use here the orthonormal tensor product Legendre polynomials L_ν, $\nu \in \mathbb{N}_0^d$. The set of basis indices used in the computation is of the form

$$\Lambda_J := \{\nu \in \mathbb{N}_0^d \colon |\nu| \le J\}\,, \tag{5.3}$$

so that $\mathrm{span}\{L_\nu\}_{\nu \in \Lambda_J} = \mathbb{P}_J$, the space of polynomials of total degree at most J. For each i, the operator corresponding to multiplication by y_i on $L^2(U)$ is replaced by its Galerkin discretization $\mathbf{M}_i := \big(\langle y_i L_\nu, L_\mu \rangle_{L^2(U)}\big)_{\mu,\nu \in \Lambda_J}$, which is a bidiagonal matrix (we refer to [13] for further details). Note that in this semidiscrete setting, as long as $k < J$, all operations in (5.1) are represented exactly, and consequently we *exactly recover* u_k for $k < J$. As k increases further, u_k converges to the Galerkin projection $G_J u$ of u onto the space $V \otimes \mathbb{P}_J$. This projection is defined by

$$\int_U \int_D a(x,y) \nabla_x G_J u(x,y) \nabla_x v(x,y)\, dx\, d\mu(y) = \int_U \int_D f(x) v(x,y)\, dx\, d\mu(y)\,, \ \ v \in V \otimes \mathbb{P}_J\,. \tag{5.4}$$

The discretization is completed by replacing V by a fixed finite element subspace V_h of dimension M_h. The corresponding discretizations of $\bar{A}$, A_i, and f are denoted by $\bar{\mathbf{A}}$, $\mathbf{A}_i$ and $\mathbf{f}$, respectively. The iterates in the discretised version of (5.1) can then be regarded as matrices $\mathbf{u}_k \in \mathbb{R}^{M_h \times N_J}$ with $N_J := \#\Lambda_J = n(d, J)$. As $k \to +\infty$, these iterates converge to the representation coefficients of the Galerkin projection of u on the space $V_h \otimes \mathbb{P}_J$.

In order to exploit the low-rank approximability, we modify the discretized version of (5.1) by introducing additional truncations based on the sizes of the singular values. This requires that we work using low-rank representations of the iterates. If $\mathbf{u}_k$ is given in low-rank form $\mathbf{u}_k = \mathbf{V}_k \mathbf{\Phi}_k^T$ with $\mathbf{V}_k \in \mathbb{R}^{M_h \times r_k}$, $\mathbf{\Phi}_k \in \mathbb{R}^{N_J \times r_k}$ with $r_k \le \min\{M_h, N_J\}$, one step

of the discretized iteration can be done by computing

$$\tilde{\mathbf{u}}_{k+1} := (\bar{\mathbf{A}}^{-1}\mathbf{f})\mathbf{e}_0^T + \sum_{i=1}^{d}(\bar{\mathbf{A}}^{-1}\mathbf{A}_i\mathbf{V}_k)(\mathbf{M}_i\boldsymbol{\Phi}_k)^T . \qquad (5.5)$$

Here $\mathbf{e}_0 := (\delta_{0\nu})_{\nu\in\Lambda_J}$, regarded as a column vector. Thus (5.5) yields a representation of the form $\tilde{\mathbf{u}}_{k+1} = \tilde{\mathbf{V}}_{k+1}\tilde{\boldsymbol{\Phi}}_{k+1}^T$ with formal *representation rank* $\tilde{r}_{k+1} = nr_k + 1$, which may be larger than rank$(\tilde{\mathbf{u}}_{k+1})$. To detect and computationally exploit further low-rank structure in $\tilde{\mathbf{u}}_{k+1}$, we truncate its singular value decomposition (SVD) up to a prescribed tolerance ε_k in the Frobenius norm. Note that in doing so, we need to use the norm induced by $\bar{A}$ on V_h. To achieve this, we compute a sparse Cholesky decomposition $\bar{\mathbf{A}} = \mathbf{L}\mathbf{L}^T$, perform the truncated SVD on $\mathbf{L}^T\tilde{\mathbf{u}}_{k+1}$, and pre-multiply the result by $\mathbf{L}^{-T}$. The actual SVD can performed at a total cost of order $\max\{M_h, N_J\}\,\tilde{r}_{k+1}^2$ by orthogonalizing the columns of $\mathbf{L}^T\tilde{\mathbf{V}}_{k+1}$ and $\tilde{\boldsymbol{\Phi}}_{k+1}$ and then performing an SVD of a matrix of size $\tilde{r}_{k+1} \times \tilde{r}_{k+1}$. We define $\mathbf{u}_{k+1} = \mathbf{V}_{k+1}\boldsymbol{\Phi}_{k+1}^T$, with rank $r_{k+1} \le \tilde{r}_{k+1}$, as the result of this rank truncation.

Note that these new iterates $\mathbf{u}_k$ differ from the initial iterates defined by only performing the discretised version of (5.1). With appropriately chosen truncation parameters ε_k, as $k \to \infty$ the iterates $\mathbf{u}_k$ converge linearly to the representation coefficients of the Galerkin projection of u on $V_h \otimes \mathbb{P}_J$, similar to the initial iterates. The precise choice of the ε_k is crucial not only for ensuring convergence, but also for the efficiency of the resulting scheme. A choice that can be shown to achieve this balance is given in [2]. In our present setting, our aim is less to obtain an efficient method than to closely reproduce the exact u_k, and we thus use $\varepsilon_k = 10^{-15}$, i.e., a truncation tolerance close to machine precision, for all tests in this work.

Concerning the choice of V_h, no triangles in the mesh should be intersected by subdomain boundaries, since this has a negative effect on the approximation error. Note further that if this is taken into account, as a consequence of (4.18), the maximum rank of the $\mathbf{u}_k$ that can be observed numerically is bounded by d plus the number of degrees of freedom on the skeleton Γ. This means in particular that in order to obtain results that reflect the actual decay of the singular values of u, when using only uniformly refined triangulations one needs very fine meshes to produce sufficiently many degrees of freedom on Γ. For this reason, in our numerical experiments we use meshes with a strong gradation towards Γ.

6 Conclusions and outlook

We have constructed low-rank approximations of solutions of certain parametric elliptic problems, based on assumptions on additional structures in these problems. Our analysis reveals particular mechanisms which, for specific problems, may ensure significant improvements between low-rank approximations based on truncated polynomial expansions and optimal low-rank approximations.

Whereas the rather general assumption (3.18) yields only a rather modest improvement in the n-widths as compared to estimates obtained by a direct polynomial expansion, we
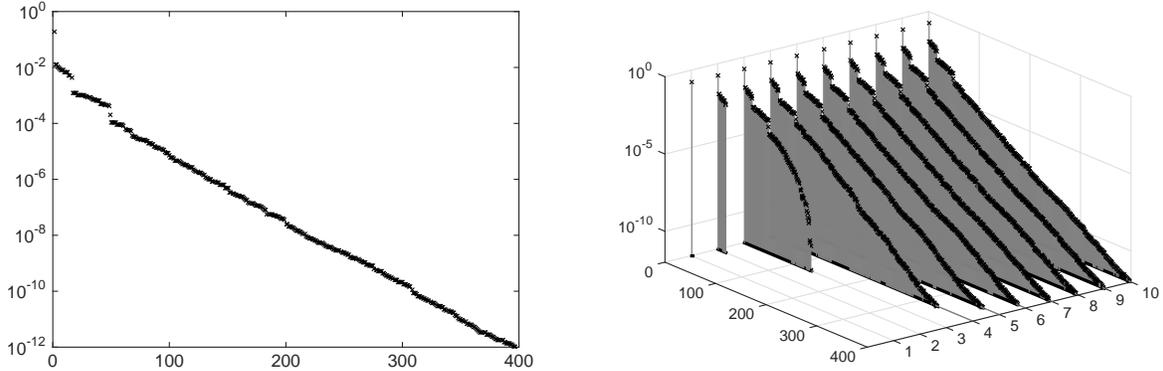
Figure 6.6: Decay of singular values and ranks of partial sums for 16 parameters.

arrive at a much stronger estimate in the particular setting considered in Section 4.2. As demonstrated there, this improvement depends rather strongly on the symmetries in the geometry of the problem. Numerical experiments indicate that the exponential decay of the singular values is maintained in the case of a similar $d = m \times m$ checkerboard structure on the unit square, however with no general proof available at the present stage.

As an illustration we reconsider in more detail the problem with 16 parameters of Figure 1.1, with a partition of D analogous to (4.19), but with a 4×4 checkerboard pattern of squares of side length $\frac{1}{4}$. The numerical realization was done as in the numerical experiments in Section 4, here using total degree 5, resulting in 20 349 Legendre coefficients. The spatial discretization uses piecewise linear finite elements with 112 961 degrees of freedom, with 12 345 of these on subdomain interfaces.

A more detailed view of the ranks of partial sums u_k and of the decay of the corresponding singular values is shown in Figure 6.6. The observed decay of singular values is slower than in the example of Section 4, but clearly still exponential. However, here the ranks of partial sums increase *faster* than linearly. This indicates that analyzing the ranks of partial sums alone, as in the 2×2 checkerboard case, might not be sufficient to explain the exponential decay in this case, and that a finer analysis of the singular values associated with these partial sums may be required.

References

[1] R. Andreev and C. Schwab, *Sparse Tensor Approximation of Parametric Eigenvalue Problems*, in Numerical Analysis of Multiscale Problems (I.G. Graham, T.Y. Hou, O. Lakkis and R. Scheichl (Eds)), Springer Lecture Notes in Comp. Sci. Eng., **83**, 203-242, Springer Verlag, 2011.

[2] M. Bachmayr and W. Dahmen, *Adaptive near-optimal rank tensor approximation for high-dimensional operator equations*, to appear in J. Foundations of Computational

Mathematics, 2014, DOI 10.1007/s10208-013-9187-3.

[3] J. Ballani and L. Grasedyck, *Hierarchical tensor approximation of output quantities of parameter-dependent PDEs*, ANCHP Preprint, EPF Lausanne, 2014.

[4] J. Beck, F. Nobile, L. Tamellini, and R. Tempone, *On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods*, Mathematical Models and Methods in Applied Sciences, **22**, 1-33, 2012.

[5] J. Beck, F. Nobile, L. Tamellini, and R.Tempone, *Convergence of quasi-optimal stochastic Galerkin methods for a class of PDEs with random coefficients*, Computers and Mathematics with Applications, **67**, 732-751, 2014.

[6] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *Convergence rates for greedy algorithms in reduced basis methods*, SIAM Journal of Mathematical Analysis, **43**, 1457-1472, 2011.

[7] A. Chkifa, A. Cohen, R. DeVore and C. Schwab, *Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs*, M2AN, **47**, 253-283, 2013.

[8] A. Cohen and R. DeVore, *Kolmogorov widths under holomorphic mappings*, to appear in IMA Journal of Applied Mathematics, 2015.

[9] A. Cohen and R. DeVore, *High dimensional approximation of parametric PDEs*, to appear in Acta Numerica, 2015.

[10] A. Cohen, R. DeVore and C. Schwab, *Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs*, J. Foundations of Computational Mathematics, **10**, 615-646, 2010.

[11] A. Cohen, R. DeVore and C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic PDE's*, Analysis and Applications, **9**, 1-37, 2011.

[12] R. DeVore, G. Petrova, and P. Wojtaszczyk, *Greedy algorithms for reduced bases in Banach spaces*, Constructive Approximation, **37**, 455-466, 2013.

[13] C. J. Gittelson and C. Schwab, *Sparse tensor discretization of high-dimensional parametric and stochastic PDEs*, Acta Numerica, **20**, 2011.

[14] M. Kahlbacher and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems*, Discussiones Mathematicae: Differential Inclusions, Control and Optimization, **27**, 95-117, 2007.

[15] B. Khoromskij and C. Schwab, *Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs*, SIAM Journal of Scientific Computing, **33**, 364-385, 2011.

[16] D. Kressner and C. Tobler, *Low-rank tensor Krylov subspace methods for parameterized linear systems*, SIAM J. Matrix Anal. Appl., **32**, 1288-1316, 2011.

[17] G. Rozza, D.B.P. Huynh, and A.T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Archives of Computational Methods in Engineering, **15**, 229-275, 2008.

Markus Bachmayr
UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
bachmayr@ann.jussieu.fr

Albert Cohen
UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
cohen@ann.jussieu.fr