Approximation of Two-Variate Functions: Singular Value Decomposition Versus Regular Sparse Grids

Michael Griebel, Helmut Harbrecht

no. 488

Diese Arbeit ist mit Unterstützung des von der Deutschen Forschungsgemeinschaft getragenen Sonderforschungsbereichs 611 an der Universität Bonn entstanden und als Manuskript vervielfältigt worden.

Bonn, Dezember 2010

APPROXIMATION OF TWO-VARIATE FUNCTIONS: SINGULAR VALUE DECOMPOSITION VERSUS REGULAR SPARSE GRIDS

MICHAEL GRIEBEL AND HELMUT HARBRECHT

ABSTRACT. We compare the cost complexities of two approximation schemes for functions $f \in H^p(\Omega_1 \times \Omega_2)$ which live on the product domain $\Omega_1 \times \Omega_2$ of general domains $\Omega_1 \subset \mathbb{R}^{n_1}$ and $\Omega_2 \subset \mathbb{R}^{n_2}$, namely the singular value / Karhunen-Loeve decomposition and the regular sparse grid representation. Here we assume that suitable finite element methods with associated fixed order r of accuracy are given on the domains Ω_1 and Ω_2 . Then, the sparse grid approximation essentially needs only $\mathcal{O}(\varepsilon^{-\frac{\max\{n_1,n_2\}}{r}})$ unknowns to reach a prescribed accuracy ε provided that the smoothness of fsatisfies $p \geq 2r$, which is an almost optimal rate. The singular value decomposition produces this rate only if f is analytical since otherwise the decay of the singular values is not fast enough. If p < 2r, then the sparse grid approach gives essentially the rate $\mathcal{O}(\varepsilon^{-\frac{2\max\{n_1,n_2\}}{p}})$ while, for the singular value decomposition, we only can prove the rate $\mathcal{O}(\varepsilon^{-\frac{2\min\{r,p\}\min\{n_1,n_2\}+2p\max\{n_1,n_2\}}{(2p-\min\{n_1,n_2\})\min\{r,p\}}})$. This shows that, in our setting and if $n_1 = n_2$, the regular sparse grid approach is superior to the singular value decomposition. We prove the resulting complexities, compare the two approaches and give the results of numerical experiments which demonstrate that these rates are also achieved in practice.

1. INTRODUCTION

The efficient approximate representation of multivariate functions is an important task in numerical analysis and scientific computing. In this paper, we concentrate on functions which live on the product of two domains $\Omega_1 \times \Omega_2$. Already for this simple situation, there exists a large amount of applications: such problems involve for example radiosity models and radiative transfer [30]. Here, Ω_1 denotes the spatial threedimensional domain of a geometric object under consideration and Ω_2 is the sphere \mathbb{S}^2 . Moreover, in the case of space-time discretizations of parabolic problems, Ω_1 is the time interval whereas Ω_2 is the spatial domain [14, 29].

Then, there are various phase space problems where both Ω_1 and Ω_2 are three-dimensional cubes or the full three-dimensional real space. Examples are the Boltzmann equation, kinetic equations or the Langevin equation, see e.g. [1]. Furthermore, non-Newtonian flow can be modeled by a coupled system which consists of the Navier Stokes equation for the flow in a three-dimensional geometry described by Ω_1 and of the Fokker-Planck equation in 3(k-1)-dimensional configuration space on Ω_2 . Here k denotes the number of atoms in a chain-like molecule which constitutes the non-Newtonian behavior of the flow, for details see [3, 20, 23]. Note that the domain of the configuration space is itself again a product of k - 1 spheres.

Another example is two-scale homogenization. After unfolding [5], it gives raise to the product of the macroscopic physical domain and the periodic microscopic domain of the cell problem, see [24]. For multiple scales, a general product appears here which still can be written as the product of two domains, one containing e.g. the macroscopic scale and the other consisting of the product of the domains of the different microscales [19].

Also the two-point correlation functions of linear elliptic boundary value problems with stochastic source terms are known to satisfy a deterministic partial differential equation with the tensor product of the elliptic operator on the product of the physical domain [17, 26, 27]. Higher order moments then involve larger tensor products [28]. This approach extends to stochastic diffusion functions and more general PDEs with stochastic coefficient functions as well as to stochastic domains [16, 18].

Finally we find the product of two domains in quantum mechanics for e.g. the Schrödinger equation for Helium; systems with more than two electrons involve then multiple product domains, of course.

In general, some problems are directly given on the product of two domains, while for other problems the domains themselves are products of lower-dimensional domains. Then, the domain of an n-dimensional problem with e.g. n being some power of two can be split into the tensor product of two domains of dimension n/2 which then can be recursively further split until a terminal situation (a one-dimensional domain or a truly higher dimensional but non-tensor product domain) is reached. Related representation methods have recently been considered in [2, 15, 22].

In this article, we consider the simple case of two domains Ω_1 and Ω_2 only. Here, our analysis covers the situation of the first bisection step in the above mentioned recursion. To this end, for i = 1, 2, let Ω_i denote a domain in \mathbb{R}^{n_i} (or alternatively also an n_i -dimensional manifold in \mathbb{R}^{n_i+1}). We intend to compare the approximation of functions $f(\mathbf{x}, \mathbf{y})$ of two variables on the tensor product domain $\Omega_1 \times \Omega_2$ by either the truncated singular value decomposition

$$f(\mathbf{x}, \mathbf{y}) \approx \sum_{\ell=1}^{M} \sqrt{\lambda_{\ell}} \varphi_{\ell}(\mathbf{x}) \psi_{\ell}(\mathbf{y})$$

or by regular sparse grids

$$f(\mathbf{x}, \mathbf{y}) \approx \sum_{j_1+j_2 \le J} \sum_{k_1 \in \nabla_{j_1}^{(1)}} \sum_{k_2 \in \nabla_{j_2}^{(2)}} \beta_{(j_1, k_1), (j_2, k_2)} \xi_{j_1, k_1}^{(1)}(\mathbf{x}) \xi_{j_2, k_2}^{(2)}(\mathbf{y}).$$

In the first representation, $\{\varphi_\ell\}_{\ell=1}^M$ and $\{\psi_\ell\}_{\ell=1}^M$ are sets of orthonormal functions. They are a-priorily unknown and need thus to be approximated by systematic basis functions of some ansatz spaces $V_J^{(1)} \subset L^2(\Omega_1)$ and $V_J^{(2)} \subset L^2(\Omega_2)$, respectively, which is indeed the case in most practical applications. In the second representation, $\{\xi_{j,k}^{(i)}\}_{k\in\nabla_j^{(i)},j\leq J}$ are in general multilevel or wavelet bases of $V_J^{(i)}$, where the index j refers to the level of resolution and the index k refers to the locality of the basis function (the precise definition will be given in Section 4).

The result of this paper is as follows: we can proof a rate of the order $\mathcal{O}(\varepsilon^{-\frac{2\min\{r,p\}\min\{n_1,n_2\}+2p\max\{n_1,n_2\}}{(2p-\min\{n_1,n_2\})\min\{r,p\}}})$ for the truncated singular value decomposition. For the regular sparse grid method we essentially obtain a rate of the order $\mathcal{O}(\varepsilon^{-2\frac{\max\{n_1,n_2\}}{\min\{p,2r\}}})$ for the degrees of freedom needed to reach a specific prescribed accuracy ε . Here, $p \geq \min\{n_1, n_2\}/2$ denotes the Sobolev smoothness of f, i.e. $f \in H^p(\Omega_1 \times \Omega_2)$, and r denotes the approximation power of the ansatz spaces $V_J^{(1)}$ and $V_J^{(2)}$. This shows

(see Section 5) that, already in the simple case of $n_1 = n_2$, the regular sparse grid approach is superior to the singular value decomposition. We prove the resulting cost complexities, compare the two approaches and give the results of numerical experiments which demonstrate that these rates are also achieved in practice.

We remark that in the case of $n_1 \neq n_2$ the regular sparse grid approach can even be improved by the use of an optimized sparse grid (see Remark 4.4 and compare [4] or [11]). Moreover, adaptive approximation offers the possibility of further improvements. The discussion of such more sophisticated approximation schemes is, however, beyond the scope of the present article and will be considered in more detail in a forthcoming paper.

The remainder of this article is organized as follows: In Section 2 we give a short introduction to multilevel approximation. In Section 3 we describe the singular value decomposition of a two-valued function on $\Omega_1 \times \Omega_2$ and discuss its approximation properties in detail. Section 4 gives the basics of the so-called regular sparse grid approximation of a two-valued function on $\Omega_1 \times \Omega_2$ and presents its error rates and cost complexities. In Section 5, we compare the two approximations. In Section 6, we give the results of numerical experiments which show that the theoretical complexity rates are also achieved in practice. Section 7 concludes with some final remarks.

Throughout this paper, the notion "essential" in connection with the complexity estimates means "up to logarithmic terms". Moreover, to avoid the repeated use of generic but unspecified constants, we denote by $C \leq D$ that C is bounded by a multiple of D independently of parameters which C and D may depend on. Obviously, $C \geq D$ is defined as $D \leq C$, and $C \sim D$ as $C \leq D$ and $C \geq D$.

2. Approximation on the subdomains

Let $\Omega \subset \mathbb{R}^n$ be a sufficiently smooth, bounded domain. In general one uses finite elements to approximate functions on $L^2(\Omega)$. In the present paper we focus on the common *h*-method, i.e., on finite elements of fixed approximation order. Then, particularly for applying multiscale techniques, one has a sequence of nested trial spaces

(2.1)
$$V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2(\Omega)$$

such that

$$L^2(\Omega) = \overline{\bigcup_{j \in \mathbb{N}_0} V_j}, \quad V_0 = \bigcap_{j \in \mathbb{N}_0} V_j,$$

called *multiscale analysis*. Each space V_j is defined by a single scale basis $\Phi_j = \{\phi_{j,k}\}$, i.e., $V_j = \operatorname{span}\{\phi_{j,k} : k \in \Delta_j\}$, where Δ_j denotes a suitable index set with cardinality $\#(\Delta_j) \sim 2^{nj}$.

We say that the trial spaces have (approximation) order $r \in \mathbb{N}$ if

(2.2)
$$r = \sup\{s \in \mathbb{R} : \inf_{v_j \in V_j} \|v - v_j\|_{L^2(\Omega)} \lesssim h_j^s \|v\|_s \, \forall v \in H^s(\Omega)\},$$

where the quantity $h_j \sim 2^{-j}$ corresponds to the mesh width of the mesh associated with the subspace V_j on Ω . Note that the integer r > 0 refers in general to the maximal order of polynomials which are locally contained in V_j .

Equation (2.2) implies that a given function $v \in H^s(\Omega)$, $0 \le s \le r$, can be approximated in V_j at a rate h_j^s , that is

(2.3)
$$\inf_{v_j \in V_j} \|v - v_j\|_{L^2(\Omega)} \lesssim h_j^s \|v\|_{H^s(\Omega)}.$$

Thus, when we approximate a function $v \in H^s(\Omega)$ with $0 \le s \le r$ by uniform mesh refinement we obtain the rate h_j^s according to (2.3). Since the meshsize and the number of unknowns in V_j are related by $\dim(V_j) \sim 2^{jn} \sim h_j^{-n}$, we deduce that

$$(2.4) N \sim \varepsilon^{-s/n}$$

unknowns have to be spent to achieve an approximation error ε . The best possible rate $N^{-n/d}$ is achieved if s = r, that is if $v \in H^r(\Omega)$.

Note that we later will employ the definitions, properties and cost complexities individually for each subdomain Ω_i i = 1, 2, that is, we will deal with two multiscale analyses

$$V_0^{(i)} \subset V_1^{(i)} \subset V_2^{(i)} \subset \dots \subset L^2(\Omega_i), \quad i = 1, 2.$$

3. Singular value decomposition

We intend to numerically represent functions $f(\mathbf{x}, \mathbf{y}) \in L^2(\Omega_1 \times \Omega_2)$ on tensor product domains $\Omega_1 \times \Omega_2$ in the most efficient way. One way to solve this approximation problem is to use an ansatz by means of tensor products which separates the variables \mathbf{x} and \mathbf{y} . We consider the approximation

(3.5)
$$f(\mathbf{x}, \mathbf{y}) \approx \sum_{\ell=0}^{M} \alpha_{\ell} \varphi_{\ell}(\mathbf{x}) \psi_{\ell}(\mathbf{y}),$$

with certain coefficients $\alpha_{\ell} \in \mathbb{R}$ and normalized functions $\varphi_{\ell} \in L^2(\Omega_1)$ and $\psi_{\ell} \in L^2(\Omega_2)$. Such an approximation is called *low rank approximation*.

The decay of the α_{ℓ} is important for the fast convergence (in terms of M) of the series (3.5). As we will see, the convergence depends on the smoothness of the function f to be approximated. Moreover, besides determining the coefficients $\{\alpha_{\ell}\}_{\ell \in \mathbb{N}}$, a numerical scheme to compute (3.5) needs to approximate the functions $\{\varphi_{\ell}\}_{\ell \in \mathbb{N}}$ and $\{\psi_{\ell}\}_{\ell \in \mathbb{N}}$ up to an accuracy corresponding to that of (3.5).

It is well known (see e.g. [9, 21, 28] for the proof) that, with respect to the number M of terms, the best possible representation of a function $f \in L^2(\Omega_1 \times \Omega_2)$ in the L^2 -sense is given by the Karhunen-Loeve / singular value decomposition. To this end, we consider the integral operator

$$\mathcal{S}: L^2(\Omega_1) \to L^2(\Omega_2), \quad (\mathcal{S}u)(\mathbf{y}) := \int_{\Omega_1} f(\mathbf{x}, \mathbf{y}) u(\mathbf{x}) d\mathbf{x}.$$

Its adjoint is

$$\mathcal{S}^{\star}: L^2(\Omega_2) \to L^2(\Omega_1), \quad (\mathcal{S}^{\star}u)(\mathbf{x}):= \int_{\Omega_2} f(\mathbf{x}, \mathbf{y}) u(\mathbf{y}) d\mathbf{y}.$$

Then, to obtain the representation (3.5) we need to compute the eigenvalues of the integral operator

(3.6)
$$\mathcal{K} = \mathcal{S}^* \mathcal{S} : L^2(\Omega_1) \to L^2(\Omega_1), \quad (\mathcal{K}u)(\mathbf{x}) := \int_{\Omega_1} k(\mathbf{x}, \mathbf{x}') u(\mathbf{x}') d\mathbf{x}'$$

whose kernel function is given by

(3.7)
$$k(\mathbf{x}, \mathbf{x}') = \int_{\Omega_2} f(\mathbf{x}, \mathbf{y}) f(\mathbf{x}', \mathbf{y}) d\mathbf{y} \in L^2(\Omega_1 \times \Omega_1).$$

This is a Hilbert-Schmidt kernel. Thus, the associated integral operator \mathcal{K} is compact. Moreover, since \mathcal{K} is self-adjoint, there exists a decomposition into eigenpairs $(\lambda_{\ell}, \varphi_{\ell})$ which fulfill

$$\mathcal{K}\varphi_{\ell} = \lambda_{\ell}\varphi_{\ell}, \quad \ell \in \mathbb{N},$$

with non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \to 0$ and eigenfunctions $\{\varphi_\ell\}_{\ell \in \mathbb{N}}$, which constitute an orthonormal basis in $L^2(\Omega_1)$.

We now define for all $\ell \in \mathbb{N}$ with $\lambda_{\ell} > 0$ the function $\psi_{\ell} \in L^2(\Omega_2)$ by

(3.8)
$$\psi_{\ell}(\mathbf{y}) = \frac{1}{\sqrt{\lambda_{\ell}}} (S\varphi_{\ell})(\mathbf{y}) = \frac{1}{\sqrt{\lambda_{\ell}}} \int_{\Omega_1} f(\mathbf{x}, \mathbf{y}) \varphi_{\ell}(\mathbf{x}) d\mathbf{x}.$$

This constitutes a second sequence of orthonormal functions since

$$\begin{aligned} (\psi_k, \psi_\ell)_{L^2(\Omega_2)} &= \frac{1}{\sqrt{\lambda_k \lambda_\ell}} (\mathcal{S}\varphi_k, \mathcal{S}\varphi_\ell)_{L^2(\Omega_2)} = \frac{1}{\sqrt{\lambda_k \lambda_\ell}} (\mathcal{K}\varphi_k, \varphi_\ell)_{L^2(\Omega_1)} \\ &= \frac{\lambda_k}{\sqrt{\lambda_k \lambda_\ell}} (\varphi_k, \varphi_\ell)_{L^2(\Omega_1)} = \delta_{k,\ell}. \end{aligned}$$

If $\lambda_{\ell} = 0$ for some $\ell \in \mathbb{N}$ we can extend this collection of functions properly to obtain an orthonormal basis $\{\psi_{\ell}\}_{\ell \in \mathbb{N}}$ of $L^2(\Omega_2)$. Due to

(3.9)
$$\sqrt{\lambda_{\ell}}\varphi_{\ell}(\mathbf{x}) = \frac{1}{\sqrt{\lambda_{\ell}}} (\mathcal{S}^{\star} \mathcal{S} \varphi_{\ell})(\mathbf{x}) = (\mathcal{S}^{\star} \psi_{\ell})(\mathbf{x}) = \int_{\Omega_2} f(\mathbf{x}, \mathbf{z}) \psi_{\ell}(\mathbf{z}) d\mathbf{z}$$

we finally obtain the representation

(3.10)
$$f(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} \sqrt{\lambda_{\ell}} \varphi_{\ell}(\mathbf{x}) \psi_{\ell}(\mathbf{y}).$$

In view of (3.8) and (3.9), this equation is easily verified by testing with the othonormal basis $\{\varphi_k \otimes \psi_\ell\}_{k,\ell \in \mathbb{N}}$ of $L^2(\Omega_1 \times \Omega_2)$.

Remark 3.1. The adjoint kernel $\tilde{k}(\cdot, \cdot)$ is just obtained by interchanging the roles of Ω_1 and Ω_2 , i.e.,

$$\widetilde{k}(\mathbf{y},\mathbf{y}') = \int_{\Omega_1} f(\mathbf{x},\mathbf{y}) f(\mathbf{x},\mathbf{y}') d\mathbf{x} \in L^2(\Omega_2 \times \Omega_2).$$

Then, one has the integral operator

$$\widetilde{\mathcal{K}} = \mathcal{SS}^{\star} : L^2(\Omega_2) \to L^2(\Omega_2), \quad (\widetilde{\mathcal{K}}u)(\mathbf{y}) := \int_{\Omega_2} \widetilde{k}(\mathbf{y}, \mathbf{y}') u(\mathbf{y}') d\mathbf{y}'.$$

Again there exists a decomposition into eigenpairs

$$\widetilde{\mathcal{K}}\widetilde{\varphi}_{\ell} = \widetilde{\lambda}_{\ell}\widetilde{\varphi}_{\ell}, \quad \ell \in \mathbb{N},$$

with non-negative eigenvalues $\widetilde{\lambda}_1 \geq \widetilde{\lambda}_2 \geq \cdots \geq \widetilde{\lambda}_m \to 0$ and eigenfunctions $\widetilde{\varphi}_{\ell} \in L^2(\Omega_2)$. We also obtain a second sequence of orthonormal functions $\widetilde{\psi}_{\ell} \in L^2(\Omega_1)$ analogously to (3.8). The functions $\{\widetilde{\varphi}_{\ell}\}_{\ell \in \mathbb{N}}$ and $\{\widetilde{\psi}_{\ell}\}_{\ell \in \mathbb{N}}$ will be the same as before but now their roles are exchanged. Moreover, the eigenvalues λ_{ℓ} and $\widetilde{\lambda}_{\ell}$ of \mathcal{K} and $\widetilde{\mathcal{K}}$ coincide.

We shall prove the following auxiliary result concerning the mapping properties of the integral operators S and S^* . To this end, for s > 0we mean by $H^{-s}(\Omega) := (H^s(\Omega))'$ the dual of the Sobolev space $H^s(\Omega)$ (which is usually denoted by $\widetilde{H}^{-s}(\Omega)$).

Lemma 3.2. Assume that $f \in H^p(\Omega_1 \times \Omega_2)$. Then, the operators

$$\mathcal{S}: H^{-s}(\Omega_1) \to H^{p-s}(\Omega_2), \quad \mathcal{S}^{\star}: H^{-s}(\Omega_2) \to H^{p-s}(\Omega_1)$$

are continuous for all $s \in [0, p]$.

Proof. From $H^p(\Omega_1 \times \Omega_2) \subset H^{0,p}_{mix}(\Omega_1 \times \Omega_2)$ it follows that $f \in H^{0,p}_{mix}(\Omega_1 \times \Omega_2)$. Therefore, the operator $\mathcal{S}: L^2(\Omega_1) \to H^p(\Omega_2)$ is continuous since

$$\begin{split} \|\mathcal{S}u\|_{H^{p}(\Omega_{2})} &= \sup_{\|v\|_{H^{-p}(\Omega_{2})=1}} (\mathcal{S}u, v)_{L_{2}(\Omega_{2})} \\ &= \sup_{\|v\|_{H^{-p}(\Omega_{2})=1}} (f, u \otimes v)_{L^{2}(\Omega_{1} \times \Omega_{2})} \\ &\leq \sup_{\|v\|_{H^{-p}(\Omega_{2})=1}} \|f\|_{H^{0,p}_{mix}(\Omega_{1} \times \Omega_{2})} \|u \otimes v\|_{H^{0,-p}_{mix}(\Omega_{1} \times \Omega_{2})} \\ &\sim \|f\|_{H^{0,p}_{mix}(\Omega_{1} \times \Omega_{2})} \|u\|_{L^{2}(\Omega_{1})}. \end{split}$$

Note that we have used here that $H^{0,-p}_{mix}(\Omega_1 \times \Omega_2) \cong L^2(\Omega_1) \otimes H^{-p}(\Omega_2).$

In complete analogy one shows that $S^* : L^2(\Omega_2) \to H^p(\Omega_1)$ is continuous which proves the desired assertion for s = 0. By duality one also infers the assertion for s = p. The assertion for $s \in (0, p)$ is finally obtained by interpolation. With the above lemma we are able to determine the decay rate of the eigenvalues of the integral operator $\mathcal{K} = \mathcal{S}^* \mathcal{S}$ with kernel (3.7). Note here that our result improves the one from [28] by a factor of two since we exploit the regularity in both variables of the kernel function which satisfies only $k \in H^{p,p}_{mix}(\Omega_1 \times \Omega_1)$.

Theorem 3.3. Consider $f \in H^p(\Omega_1 \times \Omega_2)$ with associated kernel k from (3.7) and associated integral operator \mathcal{K} from (3.6). Then, the eigenvalues $\{\lambda_\ell\}_{\ell \in \mathbb{N}}$ of \mathcal{K} decay like

(3.11)
$$\lambda_{\ell} \lesssim \ell^{-\frac{2p}{\min\{n_1, n_2\}}} \quad as \ \ell \to \infty.$$

Proof. We shall investigate the dependance of the decay of the eigenvalues of the integral operator $\mathcal{K} = \mathcal{S}^* \mathcal{S}$ on the smoothness p of f. To this end, we introduce *new* finite element spaces $U_N \subset L^2(\Omega_2)$, which consist of N discontinuous, piecewise polynomial functions of total degree p on a quasi-uniform triangulation of Ω_2 with mesh width $h_N \sim N^{-1/n_2}$. Then, there holds the following approximation result for the L^2 -orthogonal projection $P_N : L^2(\Omega_2) \to U_N$:

$$\|(I - P_N)w\|_{L^2(\Omega_2)} \le c_p N^{-p/n_2} \|w\|_{H^p(\Omega_2)},$$

provided that $w \in H^p(\Omega_2)$. Then, since $\mathcal{S}^* P_N \mathcal{S} : L^2(\Omega_1) \to L^2(\Omega_1)$ is an operator of finite rank N, the min-max principle of Courant-Fischer implies

$$\begin{split} \lambda_{N+1} &= \min_{\substack{V \subset L^{2}(\Omega_{1}) \\ \dim V^{\perp} \leq N}} \max_{\substack{u \in V \\ \|u\|_{L^{2}(\Omega_{1})} = 1}} (\mathcal{K}u, u)_{L^{2}(\Omega_{1})} \\ &\leq \max_{\substack{u \perp \operatorname{img}(\mathcal{S}^{*}P_{N}\mathcal{S}) \\ \|u\|_{L^{2}(\Omega_{1})} = 1}} (\mathcal{S}^{*}\mathcal{S}u, u)_{L^{2}(\Omega_{1})} \\ &= \max_{\substack{u \perp \operatorname{img}(\mathcal{S}^{*}P_{N}\mathcal{S}) \\ \|u\|_{L^{2}(\Omega_{1})} = 1}} (\mathcal{S}^{*}(I - P_{N})\mathcal{S}u, u)_{L^{2}(\Omega_{1})} \\ &= \max_{\substack{u \perp \operatorname{img}(\mathcal{S}^{*}P_{N}\mathcal{S}) \\ \|u\|_{L^{2}(\Omega_{1})} = 1}} ((I - P_{N})\mathcal{S}u, (I - P_{N})\mathcal{S}u)_{L^{2}(\Omega_{2})} \\ &\leq \sup_{\|u\|_{L^{2}(\Omega_{1})} = 1} \|(I - P_{N})\mathcal{S}u\|_{L^{2}(\Omega_{2})}^{2} \\ &\leq N^{-2p/n_{2}} \sup_{\|u\|_{L^{2}(\Omega_{1})} = 1}} \|\mathcal{S}u\|_{H^{p}(\Omega_{2})}^{2}. \end{split}$$

Since $\mathcal{S} : L^2(\Omega_1) \to H^p(\Omega_2)$ is continuous according to Lemma 3.2, we arrive at $\lambda_{N+1} \lesssim N^{-2p/n_2}$. Applying the same arguments to the operator $\widetilde{\mathcal{K}}$, one gets the decay rate $\lambda_{N+1} \lesssim N^{-2p/n_1}$. Substituting finally N+1 by ℓ yields the desired result. \Box

Altogether, if $f \in H^p(\Omega_1 \times \Omega_2)$, then Theorem 3.3 implies that the associated coefficients $\{\sqrt{\lambda_\ell}\}$ in the expansion (3.10) of f decay like $\sqrt{\lambda_\ell} \lesssim \ell^{-p/\min\{n_1,n_2\}}$. This leads to the following theorem.

Theorem 3.4. Let $f \in H^p(\Omega_1 \times \Omega_2)$ and $p > \min\{n_1, n_2\}/2$. Then it holds

(3.12)
$$\left\| f - \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})} \lesssim M^{\frac{1}{2} - \frac{p}{\min\{n_{1}, n_{2}\}}}.$$

Proof. Assume without loss of generality that $n_1 \leq n_2$. Then, due to $\lambda_{\ell} \sim \ell^{-2p/n_1}$ by Theorem 3.3, we obtain by the orthonormality of the function sets $\{\phi_{\ell}\}$ and $\{\psi_{\ell}\}$ that

$$\left\| f - \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2} = \left\| \sum_{\ell=M+1}^{\infty} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$
$$= \sum_{\ell=M+1}^{\infty} \lambda_{\ell} \lesssim \sum_{\ell=M+1}^{\infty} \ell^{-\frac{2p}{n_{1}}}.$$

Since $2p/n_1 > 1$ we can estimate the sum by an integral as

$$\sum_{\ell=M+1}^{\infty} \ell^{-\frac{2p}{n_1}} \le \int_M^\infty x^{-\frac{2p}{n_1}} dx = \frac{M^{1-\frac{2p}{n_1}}}{\frac{2p}{n_1}-1}$$

which leads to the desired result (3.12).

Consequently, to ensure the error bound

(3.13)
$$\left\| f - \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})} \lesssim \varepsilon$$

we need to choose the expansion degree M as

(3.14)
$$M \sim \varepsilon^{\frac{2\min\{n_1, n_2\}}{\min\{n_1, n_2\} - 2p}}.$$

10

- **Remark 3.5.** (1) Regularity in terms of mixed derivatives does not further improve the results. The property $f \in H^{p,p}_{mix}(\Omega_1 \times \Omega_2)$ again yields the estimate $\lambda_{\ell} \leq \ell^{-\frac{2p}{\max\{n_1,n_2\}}}$ for the eigenvalues of \mathcal{K} .
 - (2) The use of the Sobolev regularity might give a too low decay rate (3.11). For example, for the exponential kernel k(x, y) = exp(-|x y|) on the unit square, we have k ∈ H^{3/2-δ}((0,1) × (0,1)) for all δ > 0 but we observe √λ_ℓ ~ ℓ⁻² instead of √λ_ℓ ~ ℓ^{-3/2}, see Fig. 6.3 in Section 6. Here, we expect that the use of Besov regularity would give the correct decay.
 - (3) Our considerations and thus estimate (3.12) do not apply if $0 \le p \le \min\{n_1, n_2\}/2$. However, it still holds $\sum_{\ell=0}^{\infty} \lambda_{\ell} < \infty$ since \mathcal{K} is a Hilbert-Schmidt operator.

Depending on the smoothness of f, we are able to prove the following result on the regularity of the functons in the collections $\{\varphi_{\ell}\}$ and $\{\psi_{\ell}\}$. This result will be essential for any numerical computation of the truncated singular value decomposition.

Lemma 3.6. Let $f \in H^p(\Omega_1 \times \Omega_2)$. Then, the eigenfunctions $\{\varphi_\ell\}$ and $\{\psi_\ell\}$ are in $H^p(\Omega_1)$ and $H^p(\Omega_2)$, respectively, and satisfy

(3.15)
$$\|\varphi_{\ell}\|_{H^{p}(\Omega_{1})} \lesssim \frac{1}{\sqrt{\lambda_{\ell}}}, \quad \|\psi_{\ell}\|_{H^{p}(\Omega_{2})} \lesssim \frac{1}{\sqrt{\lambda_{\ell}}}, \quad \ell \in \mathbb{N}.$$

Proof. In view of (3.9) and Lemma 3.2, we deduce

$$\|\varphi_{\ell}\|_{H^{p}(\Omega_{1})} = \frac{1}{\sqrt{\lambda_{\ell}}} \|\mathcal{S}^{\star}\psi_{\ell}\|_{H^{p}(\Omega_{1})} \lesssim \frac{1}{\sqrt{\lambda_{\ell}}} \|\psi_{\ell}\|_{L^{2}(\Omega_{2})} = \frac{1}{\sqrt{\lambda_{\ell}}}$$

for all $\ell \in \mathbb{N}$. The second estimate is proven in complete analogy. \Box

So far, we used an *exact* description of the eigenfunctions. However this does not hold in practice. Instead, the eigenvalues $\{\lambda_\ell\}_{\ell=1}^M$ and eigenfunctions $\{\varphi_\ell\}_{\ell=1}^M$ and $\{\psi_\ell\}_{\ell=1}^M$ need to be approximately computed in the finite element spaces $V_j^{(i)} \subset L^2(\Omega_i)$, introduced in Section 2. According to (2.3) and (3.15), when we spend N degrees of freedom, we have

(3.16)

$$\begin{aligned} \|\varphi_{\ell} - \varphi_{\ell,N}\|_{L^{2}(\Omega_{1})} &\lesssim N^{-\frac{\min\{p,r\}}{n_{1}}} \|\varphi_{\ell}\|_{H^{\min\{p,r\}}(\Omega)} \lesssim \frac{1}{\sqrt{\lambda_{\ell}}} N^{-\frac{\min\{p,r\}}{n_{1}}}, \\ \|\psi_{\ell} - \psi_{\ell,N}\|_{L^{2}(\Omega_{2})} &\lesssim N^{-\frac{\min\{p,r\}}{n_{2}}} \|\psi_{\ell}\|_{H^{\min\{p,r\}}(\Omega)} \lesssim \frac{1}{\sqrt{\lambda_{\ell}}} N^{-\frac{\min\{p,r\}}{n_{2}}}. \end{aligned}$$

Here, $\varphi_{\ell,N}$ and $\psi_{\ell,N}$ denote the numerical approximations to φ_{ℓ} and ψ_{ℓ} , respectively.

According to (3.16), to ensure $\|\varphi_{\ell} - \varphi_{\ell,N}\|_{L^2(\Omega_1)} \lesssim \varepsilon/\sqrt{\lambda_{\ell}M}$ (we will later see that this is the proper accuracy), we have to spent (cf. (3.14))

$$N_{\varphi} \sim \left(\frac{\varepsilon}{\sqrt{M}}\right)^{-\frac{n_1}{\min\{r,p\}}} \sim \varepsilon^{\frac{2pn_1}{(\min\{n_1,n_2\}-2p)\min\{r,p\}}}$$

unknowns for the representation of $\varphi_{\ell,N}$ and, to ensure $\|\psi_{\ell} - \psi_{\ell,N}\|_{L^2(\Omega_2)} \lesssim \varepsilon/\sqrt{\lambda_{\ell}M}$, we have to spent

$$N_{\psi} \sim \left(\frac{\varepsilon}{\sqrt{M}}\right)^{-\frac{n_2}{\min\{r,p\}}} \sim \varepsilon^{\frac{2pn_2}{(\min\{n_1,n_2\}-2p)\min\{r,p\}}}$$

unknowns for the representation of $\psi_{\ell,N}$, respectively. In the sequel, we will spend always $N := \max\{N_{\varphi}, N_{\psi}\}$ degrees of freedom which does not deteriorate the cost complexity. In particular, N does not depend on ℓ , i.e., all eigenfunctions $\{\varphi_{N,\ell}\}$ and $\{\psi_{N,\ell}\}$ are approximated in the same ansatz spaces.

Remark 3.7. If p > r, then we even may estimate

$$\|\varphi_{\ell}\|_{H^{r}(\Omega)} \lesssim \left(\frac{1}{\sqrt{\lambda_{\ell}}}\right)^{\frac{r}{p}}, \quad \|\psi_{\ell}\|_{H^{r}(\Omega)} \lesssim \left(\frac{1}{\sqrt{\lambda_{\ell}}}\right)^{\frac{r}{p}}$$

by using interpolation arguments. Hence, (3.16) can be improved by

$$\|\varphi_{\ell} - \varphi_{\ell,N}\|_{L^{2}(\Omega_{1})} \lesssim \lambda_{\ell}^{-\frac{\min\{r,p\}}{2p}} N^{-\frac{\min\{p,r\}}{n_{1}}},$$
$$\|\psi_{\ell} - \psi_{\ell,N}\|_{L^{2}(\Omega_{2})} \lesssim \lambda_{\ell}^{-\frac{\min\{r,p\}}{2p}} N^{-\frac{\min\{p,r\}}{n_{2}}}.$$

As a consequence, if p > r, the number of unknowns for approximating the eigenfunctions can be reduced when ℓ increases. However, to exploit this fact for the computation of the truncated singular value decomposition, one needs the specific information on the smoothness index p.

12

We assume that the approximate eigenfunctions are normalized and pairwise orthogonal, i.e., for $1 \leq \ell, \ell' \leq M$ we have

$$\int_{\Omega_1} \varphi_{\ell,N}(\mathbf{x}) \varphi_{\ell',N}(\mathbf{x}) d\mathbf{x} = \delta_{\ell,\ell'}, \quad \int_{\Omega_2} \psi_{\ell,N}(\mathbf{y}) \psi_{\ell',N}(\mathbf{y}) d\mathbf{y} = \delta_{\ell,\ell'}.$$

Concerning the approximation of the M largest eigenvalues in the space V_N , defined via their Rayleigh quotients

$$\lambda_{\ell,N} = \int_{\Omega_1} \int_{\Omega_1} k(\mathbf{x}, \mathbf{x}') \varphi_{\ell,N}(\mathbf{x}) \varphi_{\ell,N}(\mathbf{x}') d\mathbf{x} d\mathbf{x}', \quad \ell = 1, 2, \dots, M.$$

we assume the following estimate

(3.17)
$$0 \le \lambda_{\ell} - \lambda_{\ell,N} \lesssim \lambda_{\ell} \|\varphi_{\ell} - \varphi_{\ell,N}\|_{L^{2}(\Omega_{1})}^{2}, \quad \ell = 1, 2, \dots, M.$$

We emphasize that these assumptions, in particular (3.17), are all satisfied if one computes the approximation $\{(\lambda_{N,\ell}, \varphi_{N,\ell})\}_{\ell=1}^{M}$ to the eigenpairs $\{(\lambda_{\ell}, \varphi_{\ell})\}_{\ell=1}^{M}$ with a Ritz-Galerkin method in the space $V_{j}^{(1)}$ with dim $V_{j}^{(1)} \sim N$. For the proof we refer the reader to [8].

Based on (3.12), (3.16) and (3.17), we can estimate the error of the discretized truncated singular value decomposition.

Theorem 3.8. Let $f \in H^p(\Omega_1 \times \Omega_2)$ and choose

$$M \sim \varepsilon^{\frac{2\min\{n_1,n_2\}}{\min\{n_1,n_2\}-2p}}, \quad N \sim \varepsilon^{\frac{2p\max\{n_1,n_2\}}{(\min\{n_1,n_2\}-2p)\min\{r,p\}}}.$$

Then, the truncated singular value decomposition satisfies the error estimate

$$\left\| f - \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell,N}} (\varphi_{\ell,N} \otimes \psi_{\ell,N}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})} \lesssim \varepsilon$$

uniformly in $\varepsilon > 0$.

Proof. It holds

$$E := \left\| \sum_{\ell=0}^{\infty} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) - \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell,N}} (\varphi_{\ell,N} \otimes \psi_{\ell,N}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$

$$\leq \left\| \sum_{\ell=M+1}^{\infty} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$

$$+ \left\| \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell}} (\varphi_{\ell} \otimes \psi_{\ell}) - \sqrt{\lambda_{\ell,N}} (\varphi_{\ell,N} \otimes \psi_{\ell,N}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$

According to (3.12) and (3.13) the squared truncation error is bounded by $c\varepsilon^2$, and thus

$$E \lesssim \varepsilon^{2} + \left\| \sum_{\ell=0}^{M} \left(\sqrt{\lambda_{\ell}} - \sqrt{\lambda_{\ell,N}} \right) (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2} \\ + \left\| \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell,N}} (\varphi_{\ell} \otimes (\psi_{\ell} - \psi_{\ell,N})) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2} \\ + \left\| \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell,N}} ((\varphi_{\ell} - \varphi_{\ell,N}) \otimes \psi_{\ell,N}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$

We now estimate the three terms on the right hand side separately. In view of (3.17), we have $\sqrt{\lambda_{\ell}} \geq \sqrt{\lambda_{\ell,N}}$ and, with (3.16), we obtain

.

$$0 \le \left|\sqrt{\lambda_{\ell}} - \sqrt{\lambda_{\ell,N}}\right|^2 \le \lambda_{\ell} - \lambda_{\ell,N} \lesssim \lambda_{\ell} \|\varphi_{\ell} - \varphi_{\ell,N}\|_{L^2(\Omega_1)}^2 \lesssim \frac{\varepsilon^2}{M}.$$

This yields

(3.18)
$$\left\|\sum_{\ell=0}^{M} \left(\sqrt{\lambda_{\ell}} - \sqrt{\lambda_{\ell,N}}\right) (\varphi_{\ell} \otimes \psi_{\ell}) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$
$$= \sum_{\ell=0}^{M} \left|\sqrt{\lambda_{\ell}} - \sqrt{\lambda_{\ell,N}}\right|^{2} \lesssim \sum_{\ell=0}^{M} \frac{\varepsilon^{2}}{M} \lesssim \varepsilon^{2}.$$

Next, with $\lambda_{\ell,N} \leq \lambda_{\ell}$, we have

$$(3.19) \qquad \left\| \sum_{\ell=0}^{M} \sqrt{\lambda_{\ell,N}} \left(\varphi_{\ell} \otimes \left(\psi_{\ell} - \psi_{\ell,N} \right) \right) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$
$$= \sum_{\ell=0}^{M} \lambda_{\ell,N} \| \varphi_{\ell} \otimes \left(\psi_{\ell} - \psi_{\ell,N} \right) \|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2}$$
$$\leq \sum_{\ell=0}^{M} \lambda_{\ell} \| \varphi_{\ell} \|_{L^{2}(\Omega_{1})}^{2} \| \psi_{\ell} - \psi_{\ell,N} \|_{L^{2}(\Omega_{2})}^{2}$$
$$\lesssim \sum_{\ell=0}^{M} \frac{\varepsilon^{2}}{M} \lesssim \varepsilon^{2},$$

14

and likewise

(3.20)
$$\left\|\sum_{\ell=0}^{M} \sqrt{\lambda_{\ell,N}} \left(\left(\varphi_{\ell} - \varphi_{\ell,N}\right) \otimes \psi_{\ell,N} \right) \right\|_{L^{2}(\Omega_{1} \times \Omega_{2})}^{2} \lesssim \varepsilon^{2}.$$

Plugging (3.18)–(3.20) into the above estimate of E finally leads to the desired estimate $E \leq \varepsilon^2$.

Altogether, since we have to deal with M eigenfunctions with N coefficients each, we arrive at the following theorem.

Theorem 3.9. The number of degrees of freedom needed to approximate a function $f \in H^p(\Omega_1 \times \Omega_2)$ by the singular value decomposition approach (3.5) with (3.16) and (3.17) to a prescribed accuracy ε is

(3.21)
$$\operatorname{dof}_{svd}(\varepsilon) = M \cdot N \sim \varepsilon^{-\frac{2\min\{r,p\}\min\{n_1,n_2\}+2p\max\{n_1,n_2\}}{(2p-\min\{n_1,n_2\})\min\{r,p\}}}$$

We emphasize that the estimate (3.21) does not include the work to be spent for *computing* the singular values nor the eigenfunctions. Here, a naive approach would result in a cost complexity of order $M \cdot N^2$, the use of fast methods for nonlocal operators would result in an almost linear or even linear complexity per eigenpair. Note that in any case at least linear complexity $\mathcal{O}(M \cdot N)$ is required, see e.g. [6, 7, 28].

4. Regular sparse grids

In general one uses finite elements to approximate the functions on $L^2(\Omega_i)$, i = 1, 2. Then, particularly for applying multiscale techniques, one has a sequence of nested trial spaces

(4.22)
$$V_0^{(i)} \subset V_1^{(i)} \subset V_2^{(i)} \subset \dots \subset L^2(\Omega_i), \quad i = 1, 2,$$

called *multiscale analysis*.

This gives a second method to approximate functions in tensor product spaces: By choosing complementary spaces

$$W_{j}^{(i)} = \operatorname{span} \left\{ \xi_{j,k}^{(i)} : k \in \nabla_{j}^{(i)} := \Delta_{j}^{(i)} \setminus \Delta_{j-1}^{(i)} \right\}, \quad i = 1, 2,$$

such that

$$V_j^{(i)} = W_j^{(i)} \oplus V_{j-1}^{(i)}, \quad V_0^{(i)} = W_0^{(i)},$$

we can approximate a function $f \in L^2(\Omega_1 \times \Omega_2)$ in the so called *regular* sparse grid space, see [4],

(4.23)
$$V_J^{(1)} \otimes V_J^{(2)} := \bigoplus_{j_1+j_2 \le J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}.$$

Then, a function $\widehat{f}_J \in V_J^{(1)} \otimes V_J^{(2)}$ is represented as

(4.24)
$$\widehat{f}_J(\mathbf{x}, \mathbf{y}) = \sum_{j_1+j_2 \le J} \sum_{k_1 \in \nabla_{j_1}^{(1)}} \sum_{k_2 \in \nabla_{j_2}^{(2)}} \beta_{(j_1, k_1), (j_2, k_2)} \xi_{j_1, k_1}^{(1)}(\mathbf{x}) \xi_{j_2, k_2}^{(2)}(\mathbf{y}).$$

Note here that the regular sparse grid space has substantially less unknowns than the full tensor product space

$$V_J^{(1)} \otimes V_J^{(2)} := \bigoplus_{j_1, j_2 \le J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}.$$

Lemma 4.1. There holds

$$\dim V_J^{(1)} \otimes V_J^{(2)} = \begin{cases} J 2^{Jn_1}, & \text{if } n_1 = n_2, \\ 2^{J \max\{n_1, n_2\}}, & \text{if } n_1 \neq n_2. \end{cases}$$

Proof. The first assertion is already well known, see e.g. [4]. Here, we shall assume without loss of generality that $n_1 \ge n_2$. Then, it follows

$$\sum_{j_1+j_2 \leq J} 2^{j_1n_1+j_2n_2} = 2^{Jn_1} \sum_{j_1=0}^J 2^{(j_1-J)n_1} \sum_{j_2=0}^{J-j_1} 2^{j_2n_2}$$
$$= 2^{Jn_1} \sum_{j_1=0}^J 2^{(j_1-J)(n_1-n_2)}$$
$$\lesssim \begin{cases} J2^{Jn_1}, & \text{if } n_1 = n_2, \\ 2^{Jn_1}, & \text{if } n_1 > n_2. \end{cases}$$

Sparse grids can be constructed via hierarchical bases, interpolets and wavelet-like bases (see e.g. [11, 26, 27, 28, 31]) or even directly by finite elements in terms of frames (see e.g. [10, 12, 18]). For a survey on sparse grids we refer the reader to [4] and the references therein.

From the approximation property (2.3) it follows that the order of approximation for a function $f \in H^s(\Omega_1 \times \Omega_2)$ $(0 \le s \le r)$ in the full tensor product space $V_j \otimes V_j$ is s. Nonetheless, according to [4, 11, 25], by spending additional smoothness in terms of *mixed Sobolev spaces* $H_{mix}^{s,s}(\Omega_1 \times \Omega_2)$, essentially the same approximation order is provided in the regular sparse grid space (4.23). The following result has been proven in [25], see also [13].

Lemma 4.2. Let $f \in H^{s,s}_{mix}(\Omega_1 \times \Omega_2)$, $0 \leq s \leq r$. Then, the L^2 -orthogonal projection \widehat{P}_J onto the sparse tensor product space $\widehat{V_J \otimes V_J}$ satisfies

$$(4.25) \ \left\| f - \widehat{P}_J f \right\|_{L^2(\Omega_1 \times \Omega_2)} \lesssim \begin{cases} 2^{-Js} \sqrt{J} \| f \|_{H^{s,s}(\Omega_1 \times \Omega_2)}, & \text{if } s = r, \\ 2^{-Js} \| f \|_{H^{s,s}(\Omega_1 \times \Omega_2)}, & \text{if } 0 \le s < r. \end{cases}$$

Thus, in view of Lemma 4.1, for a given function $f \in H^{s,s}_{mix}(\Omega_1 \times \Omega_2)$ with $s \geq r$ and a desired accuracy $\varepsilon > 0$, we essentially, i.e., up to logarithmic terms, have to spent

(4.26)
$$N \sim \varepsilon^{-\frac{\max\{n_1, n_2\}}{r}}$$

unknowns to guarantee this accuracy. Note that this rate is essentially of optimal order in the present situation, i.e., for the case $s \ge r$. If fprovides less smoothness, say $f \in H^{s,s}_{mix}(\Omega_1 \times \Omega_2)$ and s < r, then (4.25) induces the reduced approximation rate 2^{-Js} . Since the the sparse grid space owns at most $J2^{J\max\{n_1,n_2\}}$ degrees of freedom, we conclude that essentially

$$N \sim \varepsilon^{-\frac{\max\{n_1, n_2\}}{s}}$$

degrees of freedom are needed to guarantee an accuracy of ε . Thus, altogether, we essentially obtain for $f \in H^{s,s}_{mix}(\Omega \times \Omega)$, $s \ge 0$, the rate

$$N \sim \varepsilon^{-\frac{\max\{n_1, n_2\}}{\min\{s, r\}}}$$

If we now assume as above that f is an isotropic function such that $f \in H^p(\Omega_1 \times \Omega_2)$, we conclude, due to $H^p(\Omega_1 \times \Omega_2) \subset H^{\frac{p}{2},\frac{p}{2}}_{mix}(\Omega_1 \times \Omega_2)$, that $f \in H^{\frac{p}{2},\frac{p}{2}}_{mix}(\Omega_1 \times \Omega_2)$. This leads to the following theorem.

Theorem 4.3. The number of degrees of freedom needed to approximate a function $f \in H^p(\Omega_1 \times \Omega_2)$ to a prescribed accuracy ε by the regular sparse grid approach (4.24) is essentially

(4.27)
$$\operatorname{dof}_{sg}(\varepsilon) = N \sim \varepsilon^{-\frac{2\max\{n_1, n_2\}}{\min\{p, 2r\}}}.$$

Remark 4.4. Following the idea of an equilibrated cost-benefit rate (see [4]), one arrives at the optimized sparse grid

(4.28)
$$V_J^{(1)} \otimes V_J^{(2)} := \bigoplus_{j_1 \sqrt{\frac{n_1 + r}{n_2 + r}} + j_2 \sqrt{\frac{n_2 + r}{n_1 + r}} \le J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}.$$

As we will show in a forthcoming paper such a problem-adapted sparse grid is superior to the standard regular sparse grid (4.23) if the spatial dimensions of Ω_1 and Ω_2 differ, i.e., if $n_1 \neq n_2$.

5. Comparison of the two approximations

Now, we will compare the two approximation schemes. We are interested for which values of p, r, n_1 and n_2 the regular sparse grid approach is asymptotically superior to the singular value decomposition. To this end, we shall distinguish three regimes of the smoothness parameter pwhere one should have in mind the condition $p \ge \min\{n_1, n_2\}$. Nevertheless, the subsequent discussion has to be carefully interpreted, cf. Remark 3.5.

5.1. The case $p \geq 2r$. In this case it holds $f \in H^{r,r}_{mix}(\Omega_1 \times \Omega_2)$. Then, according to Theorem 4.3, it follows that the regular sparse grid approach produces essentially optimal approximation rates, whereas the truncated singular value decomposition is only of essentially optimal complexity if f is analytical. This follows from the fact that $M \sim |\log(\varepsilon)|^{\min\{n_1,n_2\}}$ dominating singular values are needed, see [28].

5.2. The case $2r > p \ge r$. According to Theorems 3.9 and 4.3, the truncated singular value decomposition has the complexity

$$dof_{svd}(\varepsilon) \sim \varepsilon^{-\frac{2r\min\{n_1,n_2\}+2p\max\{n_1,n_2\}}{(2p-\min\{n_1,n_2\})r}}$$

and the regular sparse grid approach has the complexity

$$\operatorname{dof}_{sg}(\varepsilon) \sim \varepsilon^{-\frac{2\max\{n_1, n_2\}}{p}}.$$

The sparse grid approximation is asymptotically superior to the truncated singular value decomposition if $dof_{sg}(\varepsilon) \lesssim dof_{svd}(\varepsilon)$ which holds if

$$-\frac{2\max\{n_1, n_2\}}{p} \le \frac{2r\min\{n_1, n_2\} + 2p\max\{n_1, n_2\}}{(2p - \min\{n_1, n_2\})r}$$

One readily infers that this inequality is equivalent to

$$0 \le p^2 - rp\left(1 + \frac{|n_1 - n_2|}{\max\{n_1, n_2\}}\right) + r\min\{n_1, n_2\} =: g(p).$$

The polynomial g(p) cannot be easily discussed for general values of p, r, n_1 and n_2 . But at least for the simple case $n = n_1 = n_2$, the polynomial g(p) might be bounded from below by

$$g(p) = p^2 - rp + rn \ge rp - rp + rn \ge 0$$

due to $p \ge r$. Thus, in this case the sparse grid approach exhibits higher rates of convergence.

5.3. The case $r > p \ge 0$. Theorems 3.9 and 4.3 lead to the complexities

$$\operatorname{dof}_{svd}(\varepsilon) \sim \varepsilon^{-\frac{2(n_1+n_2)}{2p-\min\{n_1,n_2\}}}$$

for the singular value decomposition and

$$\mathrm{dof}_{sg}(\varepsilon) \sim \varepsilon^{-\frac{2\max\{n_1,n_2\}}{p}}$$

for the regular sparse grid approach. Now, it holds $dof_{sg}(\varepsilon) \lesssim dof_{svd}(\varepsilon)$ if the inequality

$$\frac{2\max\{n_1, n_2\}}{p} \le \frac{2(n_1 + n_2)}{2p - \min\{n_1, n_2\}}$$

is satisfied. Straightforward manipulation yields the condition

$$n_1 n_2 > p |n_1 - n_2|$$

for the superiority of the sparse grid approximation. Especially this condition is satisfied if the dimensions n_1 and n_2 coincide.

Let us remark here that in the situation r > p an approximation in the *full* tensor product space is of the same complexity as the approximation by regular sparse grids.

6. Numerical experiments

In our numerical experiments we consider the special situation that f is a symmetric function (i.e., f(x, y) = f(y, x)) and that $\Omega_1 = \Omega_2 = (0, 1)$ (i.e., $n_1 = n_2 = 1$). Hence, the singular value decomposition of f is just its spectral decomposition.



FIGURE 6.1. The sparse grid approximation of the Gauss kernel converges essentially like N^{-2} .

On level j, we subdivide (0, 1) into 2^j intervals of length 2^{-j} which leads to $2^j + 1$ ansatz functions. The approximation spaces V_j under consideration are then generated by continuous piecewise linear polynomials on an equidistant subdivision of the interval (0, 1) (i.e., r = 2).

6.1. Gauss kernel. In our first example we discuss the approximation of the Gauss kernel

$$f(x,y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{|x-y|^2}{\sigma^2}\right), \quad \sigma > 0.$$

Since f is known to be analytical and thus arbitrarily smooth, particularly $f \in H^{2,2}_{mix}((0,1) \times (0,1))$, the regular sparse grid approach converges with the optimal rate $h_j^2 \sqrt{j} \sim 4^{-j} \sqrt{j}$. This is also observed by our numerical results, see Fig. 6.1. Of course, the smaller the value of σ the larger the constant which appears in front of the complexity estimate.

According to [28], the singular values of an analytical kernel decay exponentially (if $n_1 = n_2 = 1$). In fact, in case of the Gauss kernel the eigenvalues decay even double exponentially, cf. Fig. 6.2, where the decay is the faster the larger σ is. Since the eigenfunctions are also analytical, it holds

$$\|\varphi_{\ell}\|_{H^r(0,1)} \lesssim \lambda_{\ell}^{-s}$$
 for any $s > 0$.



FIGURE 6.2. The eigenvalues of the Gauss kernel decay double exponentionally.

Therefore, it suffices to compute the eigenfunctions with accuracy $\mathcal{O}(h_j^2)$ (see [28] for the details) which leads to a quadratic rate of convergence within $\mathcal{O}(N \log^{1/2}(N))$ cost. Altogether, this shows that both approaches converge in our setting for the Gauss kernel with essentially the same rate.

6.2. Exponential kernel. Our second example is concerned with the exponential kernel

$$k(x,y) = \exp(-|x-y|)$$

Since k is only Lipschitz continuous at the diagonal x = y, it follows that $k \in H^{3/2-\delta}((0,1) \times (0,1))$ for any $\delta > 0$. According to Theorem 3.9 we therefore can essentially guarantee the rate $dof_{svd}(\varepsilon) \sim \varepsilon^{-2}$. Nonetheless, as already mentioned in Remark 3.5, the singular values decay like $\sqrt{\lambda_{\ell}} \sim \ell^{-2}$ which is faster than predicted. In addition, it turned out in our numerical tests that also the eigenfunctions $\{\varphi_{\ell}\}$ are more regular than expected, satisfying $\|\varphi_{\ell}\|_{H^2(0,1)} \sim \sqrt{\lambda_{\ell}} \sim \ell^2$. The first nine eigenfunctions are depicted in Fig. 6.4. By repeating our analysis with these settings (that is, p = r = 2 and $n_1 = n_2 = 1$) one can shows that

$$\operatorname{dof}_{svd}(\varepsilon) \sim \varepsilon^{-4/3}$$

which is much better than predicted.



FIGURE 6.3. The eigenvalues of the exponential kernel (blue line) decay quadratically (indicated by the grey shadow).



FIGURE 6.4. The first nine eigenfunctions of the exponential kernel.



FIGURE 6.5. The sparse grid approximation of the exponential kernel converges essentially like $N^{-0.75}$.

On the other hand, in accordance with Theorem 4.3, the regular sparse grid approach realizes (essentially) the same rate: $dof_{sg}(\varepsilon) \sim \varepsilon^{-4/3}$. Indeed, this is validated by our computations (see Fig. 6.5).

Note here that only a regular sparse grid is involved and no locally adapted sparse grid is used, see e.g. [4] and the references therein. There, we expect a doubling of the convergence rate, i.e., essentially the rate $N^{-1.5}$. The analysis of such a nonlinear approximation scheme is however beyond of the scope of this paper.

7. Concluding Remarks

In the present paper we compared the cost complexities of the truncated singular value decomposition and the regular sparse grid approach. We have shown that the regular sparse grid provides an efficient tool to approximate two-variate functions. Its cots complexity is at least equal to the truncated singular value decomposition. In certain situations it is even superior.

In case of the sparse grid approach we envision further improvements by the use of an optimized sparse grid and local adaptivity. Both techniques would increase the performance of the sparse grid approach.

In case of the singular value decomposition the truncation length is determined by the smoothness of the function under consideration and is thus fixed. Therefore, improvements for the truncated singular value decomposition can only be achieved by a more efficient representation of the eigenfunctions.

References

- R. Balescu. Statistical Dynamics, Matter out of Equilibrium. Imperial College Press, Imperial College, London, 1997.
- M. Bebendorf. Adaptive cross approximation of multivariate functions. *Preprint 453*, Berichtsreihe des SFB 611, Universität Bonn, Germany, 2008 (to appear in Constructive Approximation).
- [3] J.W. Barrett, D. Knezevic, and E. Süli. Kinetic Models of Dilute Polymers: Analysis, Approximation and Computation. 11th School on Mathematical Theory in Fluid Mechanics 22-29 May 2009, Kacov, Czech Republic Necas Center for Mathematical Modeling, Prague, 2009.
- [4] H.J. Bungartz and M. Griebel. Sparse grids. Acta Numerica 13 (2004) 147–269.
- [5] D. Cioranescu, A. Damlamian, and G. Griso. The periodic unfolding method in homogenization. SIAM J. Appl. Math. 40 (2008) 1585–1620.
- [6] W. Dahmen, T. Rohwedder, R. Schneider, and A. Zeiser. Adaptive eigenvalue computation: complexity estimates. *Numer. Math.* 110 (2008) 277–312.
- [7] X. Dai, J. Xu, and A. Zhou. Convergence and optimal complexity of adaptive finite element eigenvalue computations. *Numer. Math.* 110 (2008) 313–355.
- [8] E. Dyakonov. Optimization in Solving Elliptic Problems. CRC Press, Boca Raton, 1996.
- [9] R.G. Ghanem and P.D. Spanos. Stochastic Finite Elements: A Spectral Approach. Springer-Verlag, 1991.
- [10] M. Griebel. Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen. Teubner Skripten zur Numerik. B.G. Teubner, Stuttgart, 1994.
- [11] M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Math. Comput.* 78 (2009) 2223–2257.
- [12] M. Griebel and P. Oswald. On additive Schwarz preconditioners for sparse grid discretizations. *Numer. Math.* 66 (1994) 449–463.

- [13] M. Griebel, P. Oswald, and T. Schiekhofer. Sparse grids for boundary integral equations. *Numer. Math.* 83 (1999), 279–312.
- [14] M. Griebel and D. Oeltz. A sparse grid space-time discretization scheme for parabolic problems. *Computing* 81 (2007) 1–34.
- [15] W. Hackbusch and S. Kühn. A new scheme for the tensor representation. The Journal of Fourier Analysis and Applications. 15(5) (2009) 706–722
- [16] H. Harbrecht. A finite element method for elliptic problems with stochastic input data. Appl. Numer. Math. 60 (2010) 227–244.
- [17] H. Harbrecht, R. Schneider, and C. Schwab. Sparse second moment analysis for elliptic problems in stochastic domains. *Numer. Math.* 109 (2008) 167–188.
- [18] H. Harbrecht, R. Schneider, and C. Schwab. Multilevel frames for sparse tensor product spaces. *Numer. Math.* **110** (2008) 199–220.
- [19] V.H. Hoang and C. Schwab. High-dimensional finite elements for elliptic problems with multiple scales. SIAM Multiscale Model. Simul. 3 (2005) 168–194.
- [20] C. Le Bris and T. Lelièvre. Multiscale modelling of complex fuids: A mathematical initiation. In *Multiscale Modeling and Simulation in Science*, Lecture Notes in Computational Science and Engineering, vol. 66, Springer, Berlin, 2009, 49–138.
- [21] M. Lòeve. Probability Theory, Vol. I+II, Springer, New York, 1978.
- [22] I.V. Oseledets and E.E. Tyrtyshnikov. Breaking the curse of dimensionality, or how to use SVD in many dimensions. SIAM J. Sci. Comput. 31(5) (2009) 3744-3759.
- [23] A. Lozinski, R.G. Owens, and T.N. Phillips. The Langevin and Fokker-Planck equations in polymer rheology. In R. Glowinski (ed.) *Handbook* of Numerical Analysis XVI/XVII, Elsevier North-Holland, 2010.
- [24] A. Matache. Sparse two-scale FEM for homogenization problems. SIAM J. Sci. Comput. 17 (2002) 709-720.
- [25] T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. M2AN Math. Model. Numer. Anal. 38 (2004) 93–127.
- [26] T. von Petersdorff and C. Schwab. Sparse wavelet methods for operator equations with stochastic data. Applications of Mathematics 51(2) (2006) 145–180.

- [27] C. Schwab and R.A. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.* 95 (2003) 707–734.
- [28] C. Schwab and R.A. Todor. Karhunen-Loéve approximation of random fields by generalized fast multipole methods. J. Comput. Phys. 217 (2006) 100–122.
- [29] R. Stevenson and C. Schwab. Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comput.* 78 (2009) 1293–1318.
- [30] G. Widmer, R. Hiptmair and C. Schwab. Sparse adaptive finite elements for radiative transfer. J. Comp. Phys. 227(12) (2008) 6071-6105.
- [31] C. Zenger. Sparse grids. In Parallel algorithms for partial differential equations, Proceedings of the 6th GAMM-Seminar, Kiel/Germany, 1990, Notes Numer. Fluid Mech. 31, 241–251, Vieweg, Braunschweig (1991).

MICHAEL GRIEBEL, INSTITUT FÜR NUMERISCHE SIMULATION, UNIVER-SITÄT BONN, WEGELERSTR. 6, 53115 BONN, GERMANY *E-mail address*: griebel@ins.uni-bonn.de

HELMUT HARBRECHT, INSTITUT FÜR ANGEWANDTE ANALYSIS UND NU-MERISCHE SIMULATION, PFAFFENWALDRING 57, 70569 STUTTGART, GER-MANY

E-mail address: harbrecht@ians.uni-stuttgart.de

26

Bestellungen nimmt entgegen:

Sonderforschungsbereich 611 der Universität Bonn Endenicher Allee 60 D - 53115 Bonn

 Telefon:
 0228/73 4882

 Telefax:
 0228/73 7864

 E-Mail:
 astrid.link@ins.uni-bonn.de

http://www.sfb611.iam.uni-bonn.de/

Verzeichnis der erschienenen Preprints ab No. 460

- 460. Brenier, Yann; Otto, Felix; Seis, Christian: Upper Bounds on Coarsening Rates in Demixing Binary Viscous Liquids
- 461. Bianchi, Alessandra; Bovier, Anton; Ioffe, Dmitry: Pointwise Estimates and Exponential Laws in Metastable Systems Via Coupling Methods
- 462. Basile, Giada; Bovier, Anton: Convergence of a Kinetic Equation to a Fractional Diffusion Equation; erscheint in: Review Markov Processes and Related Fields
- 463. Bartels, Sören; Roubíček, Tomáš: Thermo-Visco-Elasticity with Rate-Independent Plasticity in Isotropic Materials Undergoing Thermal Expansion
- 464. Albeverio, Sergio; Torbin, Grygoriy: The Ostrogradsky-Pierce Expansion: Probability Theory, Dynamical Systems and Fractal Geometry Points of View
- 465. Capella Kort, Antonio; Otto, Felix: A Quantitative Rigidity Result for the Cubic to Tetragonal Phase Transition in the Geometrically Linear Theory with Interfacial Energy
- 466. Philipowski, Robert: Stochastic Particle Approximations for the Ricci Flow on Surfaces and the Yamabe Flow
- 467. Kuwada, Kazumasa; Philipowski, Robert: Non-explosion of Diffusion Processes on Manifolds with Time-dependent Metric; erscheint in: Mathematische Zeitschrift
- 468. Bacher, Kathrin; Sturm, Karl-Theodor: Ricci Bounds for Euclidean and Spherical Cones
- 469. Bacher, Kathrin; Sturm, Karl-Theodor: Localization and Tensorization Properties of the Curvature-Dimension Condition for Metric Measure Spaces
- 470. Le Peutrec, Dorian: Small Eigenvalues of the Witten Laplacian Acting on *p*-Forms on a Surface
- 471. Wirth, Benedikt; Bar, Leah; Rumpf, Martin; Sapiro, Guillermo: A Continuum Mechanical Approach to Geodesics in Shape Space
- 472. Berkels, Benjamin; Linkmann, Gina; Rumpf, Martin: An SL (2) Invariant Shape Median
- 473. Bartels, Sören; Schreier, Patrick: Local Coarsening of Triangulations Created by Bisections

- 474. Bartels, Sören: A Lower Bound for the Spectrum of the Linearized Allen-Cahn Operator Near a Singularity
- 475. Frehse, Jens; Löbach, Dominique: Improved Lp-Estimates for the Strain Velocities in Hardening Problems
- 476. Kurzke, Matthias; Melcher, Christof; Moser, Roger: Vortex Motion for the Landau-Lifshitz-Gilbert Equation with Spin Transfer Torque
- 477. Arguin, Louis-Pierre; Bovier, Anton; Kistler, Nicola: The Genealogy of Extremal Particles of Branching Brownian Motion
- 478. Bovier, Anton; Gayrard, Véronique: Convergence of Clock Processes in Random Environments and Ageing in the p-Spin SK Model
- 479. Bartels, Sören; Müller, Rüdiger: Error Control for the Approximation of Allen-Cahn and Cahn-Hilliard Equations with a Logarithmic Potential
- 480. Albeverio, Sergio; Kusuoka, Seiichiro: Diffusion Processes in Thin Tubes and their Limits on Graphs
- 481. Arguin, Louis-Pierre; Bovier, Anton; Kistler, Nicola: Poissonian Statistics in the Extremal Process of Branching Brownian Motion
- 482. Albeverio, Sergio; Pratsiovyta, Iryna; Torbin, Grygoriy: On the Probabilistic, Metric and Dimensional Theories of the Second Ostrogradsky Expansion
- 483. Bulíček, Miroslav; Frehse, Jens: C^α-Regularity for a Class of Non-Diagonal Elliptic Systems with p-Growth
- 484. Ferrari, Partik L.: From Interacting Particle Systems to Random Matrices
- 485. Ferrari, Partik L.; Frings, René: On the Partial Connection Between Random Matrices and Interacting Particle Systems
- 486. Scardia, Lucia; Zeppieri, Caterina Ida: Line-Tension Model as the Γ–Limit of a Nonlinear Dislocation Energy
- 487. Bolthausen, Erwin; Kistler, Nicola: A Quenched Large Deviation Principle and a Parisi Formula for a Perceptron Version of the Grem
- 488. Griebel, Michael; Harbrecht, Helmut: Approximation of Two-Variate Functions: Singular Value Decomposition Versus Regular Sparse Grids