Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

Michael Griebel and Guanglian Li

On the decay rate of the singular values of bivariate functions

INS Preprint No. 1702

November 2017

On the decay rate of the singular values of bivariate functions

Michael Griebel[*,†] and Guanglian Li[‡]

November 22, 2017

Abstract

In this work, we establish a new truncation error estimate of the singular value decomposition (SVD) for a class of Sobolev smooth bivariate functions $\kappa \in L^2(\Omega, H^s(D))$, $s \geq 0$ and $\kappa \in L^2(\Omega, \dot{H}^s(D))$ with $D \subset \mathbb{R}^d$, where $H^s(D) := W^{s,2}(D)$ and $\dot{H}^s(D) := \{v \in L^2(D) : (-\Delta)^{s/2} v \in L^2(D)\}$ with $-\Delta$ being the negative Laplacian on D coupled with specific boundary conditions. To be precise, we show the order $\mathcal{O}(M^{-s/d})$ for the truncation error of the SVD series expansion after the M-th term. This is achieved by deriving the sharp decay rate $\mathcal{O}(n^{-1-\frac{2s}{d}})$ for the square of the n-th largest singular value of the associated integral operator, which improves on known results in the literature. We then use this error estimate to analyze an algorithm for solving a class of elliptic PDEs with random coefficient in the multi-query context, which employs the Karhunen-Loève approximation of the stochastic diffusion coefficient to truncate the model.
Keywords: eigenvalue decay, approximation of bivariate functions, Karhunen-Loève approximation, PDEs with random coefficient

1 Introduction

The efficient and accurate approximation of bivariate functions with a certain Sobolev regularity is of central importance in many diverse research areas, which range from functional analysis and machine learning to model reduction. In functional analysis, it is closely related to the so called s-numbers[1] of the kernel operator, especially to its eigenvalues through Weyl's theorem [31]. In machine learning, the decay rate of the eigenvalues or entropy numbers of the covariance integral operator is crucial for estimating the approximation error [14]. Also for many recent model-order reduction algorithms, the eigenvalue decay of an associated compact operator underpins their efficiency in practical applications.

The SVD is a popular tool for high-dimensional problems which aims at obtaining effective low-dimensional approximations. It was developed independently in different disciplines and is known under various names, e.g., as proper orthogonal decomposition (POD), as Karhunen-Loève (KL) expansion, and as principal component analysis (PCA). Its performance relies directly on the decay rate of the singular values of a bivariate kernel function. It is also encountered in multiscale numerical methods for problems in heterogeneous media, e.g., within the partition of unity method [29] or within the generalized multiscale finite element method [9, 19, 27]. The convergence rates of these modern numerical approximation methods again depend on the eigenvalue decay of an associated compact operator. Moreover, in statistical inference, the singular value decay is often used to characterize the smoothing property of the associated integral operator, which directly impacts the optimality of the regularized regression estimator [14, 10].

In this article, we focus on the singular value decomposition (SVD) of a certain class of Sobolev smooth bivariate functions. We specifically consider functions in $L^2(\Omega, H^s(D))$ with $H^s(D) := W^{s,2}(D)$, and we consider functions in $L^2(\Omega, \dot{H}^s(D))$ where $\dot{H}^s(D) := \{v \in L^2(D) : (-\Delta)^{s/2} v \in L^2(D)\}$ with $-\Delta$ being the negative Laplacian on D coupled with specific boundary conditions. Here, $D \subset \mathbb{R}^d$ is a bounded domain with a regular boundary and Ω is a bounded (not necessarily finite-dimensional) domain with dimension d'. We denote $d^* := \min\{d, d'\}$. Throughout this paper, we denote $I := (-1, 1) \subset \mathbb{R}$ and $I^d := (-1, 1)^d$.

[*]Institut für Numerische Simulation, Universität Bonn, Wegelerstraße 6, D-53115 Bonn, Germany; griebel@ins.uni-bonn.de
[†]Fraunhofer SCAI, Schloss Birlinghoven, 53754 Sankt Augustin, Germany
[‡]Institut für Numerische Simulation, Universität Bonn, Wegelerstraße 6, D-53115 Bonn, Germany; li@ins.uni-bonn.de.

[1] s-numbers are a scalar sequence assigned to an operator characterizing its degree of approximability or compactness [31]. They include approximation numbers, eigenvalues or Weyl numbers, among others.

Table 1: State of the art of eigenvalue decay results and their corresponding truncation error estimates.

reference	$\kappa(y,x)$	λ_n	M-term truncation error	
			Condition	Rate
[2]	$L^2(\Omega, H^s(D))$	$\mathcal{O}(n^{-\frac{s}{d}})$	$s > d$	$\mathcal{O}(M^{\frac{1}{2}-\frac{s}{2d}})$
[31]	$L^2(\Omega, H^s(D))$	$\mathcal{O}(n^{-\frac{2s}{d}-\frac{1}{2}})$	$s > \frac{d}{4}$	$\mathcal{O}(M^{\frac{1}{4}-\frac{s}{d}})$
[22]	$H^s(\Omega \times D)$	$\mathcal{O}(n^{-\frac{2s}{d^*}})$	$s > \frac{d^*}{2}$	$\mathcal{O}(M^{\frac{1}{2}-\frac{s}{d^*}})$
[3]	$L^2(\Omega, H^s(I))$	-	$s > 0$	$\mathcal{O}(M^{-s})$
results of this article	$L^2(\Omega, \dot{H}^s(D))$ $L^2(\Omega, H^s(D))$	$\mathcal{O}(n^{-1-\frac{2s}{d}})$	$s > 0$	$\mathcal{O}(M^{-\frac{s}{d}})$

First, let us briefly review related results from functional analysis: For any given bivariate function $\kappa(y,x) \in L^2(\Omega, H^s(D))$, we define the associated covariance function $R(x,x')$ by

$$R(x,x') = \int_\Omega \kappa(y,x)\kappa(y,x')\mathrm{d}y \in H^s(D) \times H^s(D).$$

Furthermore, let $\mathcal{R}$ denote the integral operator on $L^2(D)$ with associated kernel $R(x,x')$. Then, $\mathcal{R}$ is compact and self-adjoint with range in $H^s(D)$. Obviously, the singular values of $\kappa(y,x)$ are equivalent to the square root of the eigenvalues of $\mathcal{R}$. It was already pointed out in [2, Theorem 13.6] that, for $s > d/2$, the n-th largest eigenvalue of $\mathcal{R}$ is at least of the order $\mathcal{O}(n^{-\frac{s}{d}})$. Then, the rate of the M-term truncation error, i.e the truncation error of the SVD series expansion after the M-th term, is at least of the order $\mathcal{O}(M^{\frac{1}{2}-\frac{s}{2d}})$. A related result can be found in e.g. [35]. From a functional analytic point of view, the approximability and compactness of an operator can be characterized by a certain scalar sequence named s-scale, which is unique for the class of operators acting on a Hilbert space [31, Section 2.11.9]. With nonincreasingly ordered eigenvalues $\{\lambda_n\}_{n=1}^\infty$ of $\mathcal{R}$, we have $\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{2d}{4s+d},2}$ [31, Section 6.4.19] and [26, Section 3.c.5]. This means that the n-th largest eigenvalue of $\mathcal{R}$ is at least of the order $\mathcal{O}(n^{-\frac{2s}{d}-\frac{1}{2}})$. Thus, the decay rate of the M-term truncation error is at least of the order $\mathcal{O}(M^{\frac{1}{4}-\frac{s}{d}})$ if $s > d/4$. As a matter of fact, this estimate is even optimal for the special class of integral operators with associated kernels in $H^s(D) \times H^s(D)$.

In the last couple of years, the approximation of high-dimensional stochastic processes based on the KL expansion has gained much popularity [3, 22, 35]. For example, for $\kappa(y,x) \in H^s(\Omega \times D)$, a decay rate of the order $\mathcal{O}(n^{-\frac{2s}{d^*}})$ was established in [22] for the n-th largest eigenvalue of $\mathcal{R}$ using a minmax principle. This results in a rate of the order $\mathcal{O}(M^{\frac{1}{2}-\frac{s}{d^*}})$ for the M-term truncation error if $s > d^*/2$. Furthermore, in [3], a direct estimate of the error rate for the M-term truncation of the order $\mathcal{O}(M^{-\frac{s}{1}})$ was given in the case D being a one-dimensional interval. Moreover, as pointed out in that paper, this truncation estimate can be extended to the higher dimension case as well, which would result in the optimal M-term truncation rate of $\mathcal{O}(M^{-\frac{s}{d}})$. There, however, no decay rate of the eigenvalues was given. For the ease of comparison, we summarize these existing results on the eigenvalue decay rate and on the M-term approximation error in Table 1. Moreover, in [36, 37], the anisotropic Sobolev space $SW^{\mathbf{R}}_{\mathbf{q},\alpha}$ and the anisotropic Nikol'skii space $NH^{\mathbf{R}}_{\mathbf{q}}$ are considered, where $\mathbf{R}$, $\mathbf{q}$ and α are d-dimensional vectors. These spaces consist of periodic functions, which admit an integral representation in form of a convolution of certain L^2 functions with the Bernoulli kernels in each component. In this periodic setting, which is different from ours, an asymptotic decay rate of also the order $\mathcal{O}(M^{-s/d})$ was shown for $NH^{\mathbf{R}}_{\mathbf{q}}$, where $\mathbf{R} = [\mathbf{R}_1, \mathbf{R}_2]$ and $\mathbf{R}_1 = (0, \cdots, 0)$, $\mathbf{R}_2 = (s, \cdots, s)$. Since in this case, $NH^{(0,\cdots,0;s,\cdots,s)}_{(2,\cdots,2)}$ is larger than $SW^{(0,\cdots,0;s,\cdots,s)}_{(2,\cdots,2),\alpha}$, the rate $\mathcal{O}(M^{-s/d})$ also holds for the corresponding anisotropic periodic Sobolev space $SW^{(0,\cdots,0;s,\cdots,s)}_{(2,\cdots,2),\alpha}$. Consequently, an eigenvalue decay of $\mathcal{O}(n^{-2s/d-1})$ can be inferred, which is analogous to our result, however only for the periodic situation and product domains, see [23, 18, 33] for more references on this setting. We now deal with the non-periodic setting and general non-product Lipschitz domains Ω and D.

In this paper, we shall analyze the eigenvalues of the operator $\mathcal{R}$ for the cases $\kappa(y,x) \in L^2(\Omega, \dot{H}^s(D))$ and $\kappa(y,x) \in L^2(\Omega, H^s(D))$. We establish a decay rate of $\mathcal{O}(n^{-1-\frac{2s}{d}})$ of the n-th largest eigenvalue. This decay estimate is consistent with the numerical findings in [22], and, to the best of our knowledge, it is presently the sharpest one. Such a result for the two different considered cases is obtained in two different ways: For $\kappa(y,x) \in L^2(\Omega, \dot{H}^s(D))$, our proof is based on an alternative representation of the minmax principle [20] and a careful characterization of the operator $\mathcal{R}$. For $\kappa(y,x) \in L^2(\Omega, H^s(D))$, our proof employs Stein's extension theorem and an argument using the result from the first case in the special situation $D = (-1,1)^d$, cf. Proposition 3.1. One distinct feature of our approach is that it can directly give a truncation error estimate of the order $\mathcal{O}(M^{-\frac{s}{d}})$ for bivariate functions with a rather low degree of Sobolev regularity in y-direction.

To illustrate the use of our new decay result, we shall provide a detailed error analysis of an algorithm for solving elliptic PDEs involving high-dimensional stochastic diffusion coefficients $\kappa(y,x)$ in a multi-query context. This task arises in e.g., optimal control, inverse problems, Bayesian inversion and uncertainty quantification (see the references in the surveys [11, 34]). In these algorithms, one usually truncates the stochastic diffusion coefficient, i.e. one truncates its KL series expansion after the M-th term. This leads to an approximative model with finite-dimensional noise that is amenable to practical computations. Our new truncation estimate in Theorem 3.3 then allows, under mild integrability conditions on the source term, to derive an improved error estimate of the stochastic solution due to the KL truncation of the coefficient $\kappa(y,x)$, cf. Theorem 4.2.

The remainder of this paper is organized as follows. In Section 2, we recall preliminaries on the SVD approximation of bivariate functions and basic facts about Sobolev spaces and Lorentz sequence spaces. Then, in Section 3, we establish eigenvalue decay rates for a bivariate function $\kappa(y,x) \in L^2(\Omega, \dot{H}^s(D))$ and $\kappa(y,x) \in L^2(\Omega, H^s(D))$. In Section 4, we discuss an algorithm for elliptic PDEs with random coefficient and give an error analysis of it due to the KL truncation of the stochastic diffusion coefficient. Finally, we give a conclusion in Section 5.

2 Preliminaries

Let us recall some facts on the approximation of bivariate functions. Throughout this paper, we suppose that Ω is a bounded (not necessarily finite-dimensional) domain and $D \subset \mathbb{R}^d$ is equipped with a regular boundary. Now, consider a bivariate function $\kappa(y,x) \in L^2(\Omega \times D) = L^2(\Omega) \times L^2(D)$. The associated integral operator $\mathcal{S} : L^2(D) \to L^2(\Omega)$ is defined by

$$(\mathcal{S}v)(y) = \int_D \kappa(y,x)v(x)\mathrm{d}x, \tag{2.1}$$

with its adjoint operator $\mathcal{S}^* : L^2(\Omega) \to L^2(D)$ defined by

$$(\mathcal{S}^*v)(x) = \int_\Omega \kappa(y,x)v(y)\mathrm{d}y. \tag{2.2}$$

Next, let

$$\mathcal{R} : L^2(D) \to L^2(D), \quad \mathcal{R} = \mathcal{S}^*\mathcal{S}.$$

Then $\mathcal{R}$ is a nonnegative self-adjoint Hilbert-Schmidt operator with its kernel $R \in L^2(D \times D) : D \times D \to \mathbb{R}$ given by[2]

$$R(x,x') = \int_\Omega \kappa(y,x)\kappa(y,x')\mathrm{d}y.$$

Hence, for any $v \in L^2(D)$, we have

$$\mathcal{R}v(x) = \int_D R(x,x')v(x')\mathrm{d}x' = \int_D\int_\Omega \kappa(y,x)\kappa(y,x')v(x')\mathrm{d}y\mathrm{d}x'.$$

According to standard spectral theory for compact operators [42], the operator $\mathcal{R}$ has at most countably many discrete eigenvalues, with zero being the only accumulation point, and each non-zero eigenvalue has only finite multiplicity. Let $\{\lambda_n\}_{n=1}^\infty$ be the sequence of eigenvalues (with multiplicity counted) associated to $\mathcal{R}$, which are ordered nonincreasingly, and let $\{\phi_n\}_{n=1}^\infty$ be the[3] corresponding eigenfunctions. The eigenfunctions $\{\phi_n\}_{n=1}^\infty$ can be chosen to be orthonormal in $L^2(D)$. Furthermore, for any $\lambda_n \neq 0$, we define

$$\psi_n(y) = \frac{1}{\sqrt{\lambda_n}}\int_D \kappa(y,x)\phi_n(x)\mathrm{d}x. \tag{2.3}$$

[2]If the bivariate function $\kappa(y,x)$ represents a stochastic process, $R(x,x')$ is often denoted as the covariance function.
[3]For multiplicity > 1, one can always select an orthonormal basis for the eigenspaces since $\mathcal{R}$ is self-adjoint.

Then it is easy to verify that the sequence $\{\psi_n\}_{n=1}^{\infty}$ is orthonormal in $L^2(\Omega)$. Moreover, the sequence $\{\lambda_n\}_{n=1}^{\infty}$ can be characterized by the so-called approximation numbers (cf. [31, Section 2.3.1]). They are defined by

$$\lambda_n = \inf\{\|\mathcal{R} - L\| : L \in \mathfrak{F}(L^2(D)), \text{rank}(L) < n\} \tag{2.4}$$

where $\mathfrak{F}(L^2(D))$ denotes the set of the finite rank operators on $L^2(D)$. This equivalency is frequently employed to estimate eigenvalues by constructing finite rank approximation operators to $\mathcal{R}$.

The singular value decomposition of the bivariate function $\kappa(y, x)$ then refers to the expansion

$$\kappa(y, x) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\phi_n(x)\psi_n(y),$$

where the series converges in $L^2(\Omega \times D)$. Moreover, for any $M \in \mathbb{N}$, the M-term truncated SVD, denoted by $\kappa_M(y, x)$, is defined by

$$\kappa_M(y, x) = \sum_{n=1}^{M} \sqrt{\lambda_n}\phi_n(x)\psi_n(y). \tag{2.5}$$

The associated M-term KL truncation error is then

$$\|\kappa(y, x) - \kappa_M(y, x)\|_{L^2(\Omega\times D)} = \left\| \sum_{n>M} \sqrt{\lambda_n}\phi_n(x)\psi_n(y) \right\|_{L^2(\Omega\times D)}. \tag{2.6}$$

It is worth emphasizing the optimality of eigenfunctions $\{\phi_n\}_{n=1}^{\infty}$ in the sense that the mean-square error resulting from a finite-rank approximation of $\kappa(y, x)$ is minimized [21]. Thus, the eigenfunctions indeed minimize the truncation error in the L^2-sense, i.e.

$$\min_{\substack{\{c_n(x)\}_{n=1}^{M}\subset L^2(D)\\ \{c_n(x)\}_{n=1}^{M} \text{ orthonormal}}} \left\| \kappa(y, x) - \sum_{n=1}^{M} \left(\int_D \kappa(y, x)c_n(x)\mathrm{d}x \right) c_n(x) \right\|_{L^2(\Omega\times D)} = \sqrt{\sum_{n>M} \lambda_n}. \tag{2.7}$$

There are various articles on the convergence rate of the M-term approximation $\kappa_M(y, x)$ to $\kappa(y, x)$ as $M \to \infty$ [3, 22, 35, 39]. It is well known [31] that, as smoother the kernel $R(x, x')$ is, as faster the decay of the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ is, and thus as faster the decay of the KL truncation error is. Recently, for the heat equation, an exponentially fast decay of the truncation error was shown by exploiting the special structure of the Grammian matrix [3].

In Section 3 below, we will derive a KL truncation error estimate using a new decay rate estimate of the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$, which in turn directly implies the decay rate of the SVD approximation. Our result relies essentially on the following regularity condition on the bivariate function $\kappa(y, x)$:

Assumption 2.1 (Regularity of $\kappa(y, x)$)**.** *There exists some $s \geq 0$ such that*[4] $\kappa(y, x) \in L^2(\Omega, H^s(D))$.

Under Assumption 2.1, by the definition of the kernel $R(x, x')$, we have $R(x, x') \in H^s(D) \times H^s(D)$.

We conclude this section with some notation. Let two Banach spaces V_1 and V_2 be given. Then, $\mathcal{B}(V_1, V_2)$ stands for the Banach space composed of all continuous linear operators from V_1 to V_2 and $\mathcal{B}(V_1)$ stands for $\mathcal{B}(V_1, V_1)$. The set of nonnegative integers is denoted by $\mathbb{N}$. For any index $\alpha \in \mathbb{N}^d$, $|\alpha|$ is the sum of all components. The letter M is reserved for the truncation number of the SVD modes. We write $A \lesssim B$ if $A \leq cB$ for some absolute constant c which depends only on the domain D, and we likewise write $A \gtrsim B$. Finally, $\|\cdot\|$ denotes the Euclidean norm. Moreover, for any $m \in \mathbb{N}$, $1 \leq p \leq \infty$, we follow [1] and define the Sobolev space $W^{m,p}(D)$ by

$$W^{m,p}(D) = \{u \in L^p(D) : D^\alpha u \in L^p(D) \text{ for } 0 \leq |\alpha| \leq m\}.$$

It is equipped with the norm

$$\|u\|_{W^{m,p}(D)} = \begin{cases} \Big(\sum_{0\leq|\alpha|\leq m} \|D^\alpha u\|_{L^p(D)}^p \Big)^{\frac{1}{p}}, \text{ if } 1 \leq p < \infty, \\ \max_{0\leq|\alpha|\leq m} \|D^\alpha u\|_{L^\infty(D)}, \text{ if } p = \infty. \end{cases}$$

[4]Note here that the space $L^2(\Omega, H^s(D))$ is isomorphic to $L^2(\Omega) \times H^s(D)$.

The space $W_0^{m,p}(D)$ is the closure of $C_0^\infty(D)$ in $W^{m,p}(D)$. Its dual space is $W^{-m,q}(D)$, with $1/p+1/q=1$. Also we use $H^m(D)=W^{m,p}(D)$ for $p=2$. The fractional order Sobolev space $W^{s,p}(D)$, $s\geq 0, s\notin\mathbb{N}$, is defined by means of interpolation [1]. Furthermore, we will need the space [40, Section 4.3.2]

$$\tilde{H}^{1/2}(D)=\{v : v\in H^{1/2}(\mathbb{R}^d), \mathrm{supp}(v)\subset\bar{D}\}.$$

Finally, besides Sobolev spaces, we will resort to Lorentz sequence spaces $\ell_{r,w}$. They are useful in the study of s-scales [31], especially for the characterization of the growth of a nonnegative, nonincreasing sequence $\{a_n\}_{n=1}^\infty$. Here, a sequence $\{a_n\}_{n=1}^\infty\in\ell_{r,w}$ if $\{n^{\frac{1}{r}-\frac{1}{w}}a_n\}_{n=1}^\infty\in\ell_w$ [31, Section 2.1.4] with ℓ_w being the classical sequence space which consists of the w-power summable sequences. The following embedding properties hold:

Proposition 2.1. *[31, Section 2.1.11]*

1. $\ell_{r_0,w_0}\subsetneq\ell_{r_1,w_1}$ *for* $0<r_0<r_1<\infty$ *and arbitrary* w_0, $w_1>0$,
2. $\ell_{r,w_0}\subsetneq\ell_{r,w_1}$ *for arbitrary* $r>0$ *and* $0<w_0<w_1\leq\infty$.

3 Eigenvalue decay and KL truncation error

In this section, we establish a sharp eigenvalue decay rate for the Hilbert-Schmidt operator $\mathcal{R}$ and discuss its use in analyzing the KL truncation error. We shall consider the following two cases (a) $\kappa(y,x)\in L^2(\Omega,\dot{H}^s(D))$ and (b) $\kappa(y,x)\in L^2(\Omega,H^s(D))$ separately. Note that $L^2(\Omega,\dot{H}^s(D))\subset L^2(\Omega,H^s(D))$. For case (a) we derive $\{\lambda_n\}_{n=1}^\infty\in\ell_{\frac{d}{2s+d},1}$, i.e., a decay rate of the order $\mathcal{O}(n^{-1-\frac{2s}{d}})$, using a rearrangement argument originating from the minmax principle [20]. For case (b) we show the same decay rate by employing Stein's extension theorem and an argument based on the result for the first case in the special situation $D=(-1,1)^d$, cf. Proposition 3.1. It however involves a constant $C_{\mathrm{ext}}(D,s)$ arising from the extension operator which is difficult to control. Nevertheless, $\mathcal{R}$ belongs to the trace class in either case.

3.1 Case (a): $\kappa(y,x)\in L^2(\Omega,\dot{H}^s(D))$

By definition, the associated operator $\mathcal{R}$ is self-adjoint and nonnegative. The classical Hilbert-Schmidt theorem gives then the series expansion [42]

$$R(x,x')=\sum_{n=1}^\infty\lambda_n\phi_n(x)\phi_n(x'), \tag{3.1}$$

with convergence in the $L^2(D\times D)$-norm.

First, we show a smoothing property of the operator $\mathcal{R}$ under Assumption 2.1.

Lemma 3.1. *Let Assumption 2.1 hold. Then* $\mathcal{R}\in\mathcal{B}(L^2(D),H^s(D))$.

Proof. Consider $s\in\mathbb{N}$. For any $v\in L^2(D)$ and $|\alpha|\leq s$, by taking the α^{th} derivative with respect to x, we have

$$\partial^\alpha(\mathcal{R}v)=\int_\Omega\Big(\int_D\kappa(y,x')v(x')\mathrm{d}x'\Big)\partial^\alpha\kappa(y,x)\mathrm{d}y.$$

With Hölder's inequality, this yields

$$\|\partial^\alpha(\mathcal{R}v)\|_{L^2(D)}^2\leq\|\kappa\|_{L^2(\Omega\times D)}^2\|\partial^\alpha\kappa\|_{L^2(\Omega\times D)}^2\|v\|_{L^2(D)}^2.$$

Summing up all terms for $|\alpha|\leq s$ leads to

$$\begin{aligned}\|\mathcal{R}v\|_{H^s(D)}^2&=\sum_{|\alpha|\leq s}\|\partial^\alpha(\mathcal{R}v)\|_{L^2(D)}^2\leq\sum_{|\alpha|\leq s}\|\kappa\|_{L^2(\Omega\times D)}^2\|\partial^\alpha\kappa\|_{L^2(\Omega\times D)}^2\|v\|_{L^2(D)}^2\\&=\|\kappa\|_{L^2(\Omega\times D)}^2\|\kappa\|_{L^2(\Omega,H^s(D))}^2\|v\|_{L^2(D)}^2.\end{aligned}$$

Consequently,

$$\|\mathcal{R}\|_{\mathcal{B}(L^2(D),H^s(D))}\leq\|\kappa\|_{L^2(\Omega\times D)}\|\kappa\|_{L^2(\Omega,H^s(D))}.$$

This shows the assertion for $s\in\mathbb{N}$. The general case follows from the Riesz-Thorin interpolation theorem. □

To study the eigenvalue decay of the operator $\mathcal{R}$, we need a few more auxiliary tools. Recall the floor function $\lfloor \cdot \rfloor$, defined by $\lfloor r \rfloor = \max\{k \in \mathbb{N} : k \leq r\}$, and the ceiling function $\lceil \cdot \rceil$, defined by $\lceil r \rceil = \min\{k \in \mathbb{N} : k \geq r\}$ for any $r \geq 0$.

Definition 3.1 (Trace condition)**.** *Given $s \geq 0$, let $D \subset \mathbb{R}^d$ be a bounded $\mathcal{C}^{\lfloor s \rfloor, 1}$-domain. Let Δ be the Laplacian with respect to the spatial variable x (on the domain D) and let n be the unit outward normal to the boundary ∂D. We say $v \in H^s(D)$ satisfies the trace condition, if one of the following statement holds.*

(i) *If $s > \frac{1}{2}$: $\Delta^j v = 0$ on ∂D for all $0 \leq j \leq \lfloor \frac{s}{2} - \frac{1}{4} \rfloor$. In the case that $\frac{s}{2} - \frac{1}{4} = \lfloor \frac{s}{2} - \frac{1}{4} \rfloor$, we replace the highest order condition with $\Delta^{\frac{s}{2} - \frac{1}{4}} v \in \tilde{H}^{\frac{1}{2}}(D)$;*

(ii) *If $s > \frac{3}{2}$: $\frac{\partial}{\partial n} \Delta^j v = 0$ on ∂D for all $0 \leq j \leq \lfloor \frac{s}{2} - \frac{3}{4} \rfloor$. In the case that $\frac{s}{2} - \frac{3}{4} = \lfloor \frac{s}{2} - \frac{3}{4} \rfloor$, we replace the highest order condition with $\frac{\partial}{\partial n} \Delta^{\frac{s}{2} - \frac{3}{4}} v \in \tilde{H}^{\frac{1}{2}}(D)$;*

(iii) *If $s > \frac{3}{2}$: $\frac{\partial}{\partial n} \Delta^j v + h \Delta^j v = 0$ for some $h \geq 0$ on ∂D for all $0 \leq j \leq \lfloor \frac{s}{2} - \frac{3}{4} \rfloor$. In the case that $\frac{s}{2} - \frac{3}{4} = \lfloor \frac{s}{2} - \frac{3}{4} \rfloor$, we replace the highest order condition with $\frac{\partial}{\partial n} \Delta^{\frac{s}{2} - \frac{3}{4}} v + h \Delta^{\frac{s}{2} - \frac{3}{4}} v \in \tilde{H}^{\frac{1}{2}}(D)$.*

Next, we introduce the space $\dot{H}^s(D)$ and discuss its properties. Let $A = -\Delta$ represent the negative Laplacian on a subspace of $H^2(D)$ that satisfies Definition 3.1 (for $s = 2$, any bounded convex domain suffices). Then A is nonnegative, invertible and self-adjoint. Furthermore, let $\{\nu_j, \xi_j\}_{j=1}^{\infty}$ be the eigenpairs of A with nondecreasingly ordered eigenvalues. It is well known [13] that[5]

$$\nu_j \geq C_{\text{weyl}}(d) \text{diam}(D)^{-2} j^{2/d}$$

where $C_{\text{weyl}}(d)$ denotes a positive constant depending on d only and $\text{diam}(D)$ represents the diameter of D, and it is clear that $\{\xi_j\}_{j=1}^{\infty}$ forms an orthonormal basis in $L^2(D)$. With $(\cdot, \cdot)$ being the $L^2(D)$ inner product, each $v \in L^2(D)$ admits the expansion $v = \sum_{j=1}^{\infty} (v, \xi_j) \xi_j$. Next, for $s \geq 0$, we define a Hilbert space $\dot{H}^s(D) \subset H^s(D)$ by

$$\dot{H}^s(D) = \left\{ v \in L^2(D) : \sum_{j=1}^{\infty} \nu_j^s \cdot (v, \xi_j)^2 < \infty \right\}. \tag{3.2}$$

This space is endowed with an inner product $(\cdot, \cdot)_s$ defined by

$$(v, w)_s = \sum_{j=1}^{\infty} \nu_j^s (v, \xi_j)(w, \xi_j), \text{ for } v, w \in \dot{H}^s(D).$$

We denote by $|\cdot|_s$ the induced norm. In view that $\nu_j = \mathcal{O}(j^{2/d})$, the norm $|\cdot|_s$ is stronger than the norm $\|\cdot\|_{L^2(D)}$.

Assume that $D \subset \mathbb{R}^d$ is a $\mathcal{C}^{\lfloor s \rfloor, 1}$-bounded domain. Then the space $\dot{H}^s(D)$ can also be characterized by (see [38, Lemma 3.1] and [40, Theorem 4.3.3])

$$\dot{H}^s(D) = \{v \in H^s(D) : v \text{ satisfies Definition 3.1}\}. \tag{3.3}$$

Here, the boundary condition for the operator A is the same as that in (3.3) when $s = 2$.

This fractional-order space $\dot{H}^s(D)$ has many applications in the sparse representation of solutions to elliptic operators (see e.g., [15]). It also shares a certain similarity to the native space associated with a positive Hilbert-Schmidt kernel on $L^2(D)$, cf. [32]. In view of (3.2) and [16, Theorem 4], the orthonormal basis $\{\xi_j\}$ (ONB) is optimal in $\dot{H}^s(D)$. This has to be compared to [41, Theorem 4.22] where a sequence of optimal ONB, i.e., via wavelets, are provided for a smaller space, namely $H_0^s(D)$, with all traces vanishing.

For any $s > 0$, one can now define the fractional power operator $T = A^{s/2}$ on $\dot{H}^s(D)$ [8, 25] by

$$Tv = \sum_{j=1}^{\infty} \nu_j^{\frac{s}{2}} \cdot (v, \xi_j) \cdot \xi_j.$$

[5] More precisely, under a zero Dirichlet and Neumann boundary condition, there holds $\nu_j \geq \frac{C_{\text{weyl}}(d)(j+1)^{\frac{2}{d}}}{\text{diam}(D)^2}$ and $\frac{C_{\text{weyl}}(d) j^{\frac{2}{d}}}{\text{diam}(D)^2}$, respectively (cf. [28, Theorem 5.3]).

Equivalently, it can be written as a Dunford-Schwartz integral in the complex plane $\mathbb{C}$. By construction, T is nonnegative and self-adjoint and gives an isomorphism between $\dot{H}^s(D)$ and $L^2(D)$. It possesses the same eigenfunctions as $A = -\Delta$.

Let $\{\mu_j\}_{j=1}^{\infty}$ be the eigenvalues of T in nondecreasing order. Then the relations $T = A^{s/2}$ and $\nu_j \geq C_{\text{weyl}}(d)\text{diam}(D)^{-2}j^{2/d}$ yield

$$\mu_j \geq C_{\text{weyl}}(d)^{\frac{s}{2}}\text{diam}(D)^{-s}j^{s/d}. \tag{3.4}$$

One readily verifies that $T \in \mathcal{B}(\dot{H}^s(D), L^2(D))$ and that

$$\|Tv\|_{L^2(D)}^2 = \sum_{j=1}^{\infty} \nu_j^s (v, \xi_j)^2 = |v|_s^2 \quad \text{for all } v \in \dot{H}^s(D). \tag{3.5}$$

Now we are ready to state our regularity assumption on $\kappa(y, x)$.

Assumption 3.1 (Regularity of $\kappa(y, x)$)**.** *Let $D \subset \mathbb{R}^d$ be a $\mathcal{C}^{\lfloor s \rfloor,1}$-bounded domain for some $s \geq 0$ and let $\kappa(y, x) \in L^2(\Omega, \dot{H}^s(D))$.*

Note here that, by Lemma 3.1, the eigenfunctions ϕ_n satisfy $\phi_n \in H^s(D)$, and there holds $\|\phi_n\|_{H^s(D)} = \mathcal{O}(\lambda_n^{-1})$. Indeed, by definition, we can obtain

$$\phi_n = \lambda_n^{-1}\mathcal{R}\phi_n.$$

Then after taking the $\|\cdot\|_{H^s(D)}$ norm on both sides, the desired result follows with Lemma 3.1 and the fact that $\|\phi_n\|_{L^2(D)} = 1$. Under Assumption 3.1, $\phi_n(x)$ satisfies the same trace condition as $\kappa(y, \cdot)$. This property is of critical importance in the proof later on.

In view of the characterization (3.3), Assumption 3.1 may seem restrictive. Nevertheless, it is natural when s is small. Moreover, in the context of Proper Orthogonal Decomposition (POD) methods for parameterized elliptic problems, the (bivariate) function $\kappa(y, \cdot) \in H_0^1(D)$ represents the solution for each fixed $y \in \Omega$, i.e. it is the solution map from the parameter space Ω to the solution space $H_0^1(D)$. This implies directly Definition 3.1. Coincidentally, such a trace condition is also needed to ensure the fast convergence of the modified Fourier expansion method [24].

Now, let Assumption 3.1 hold. Then $T\kappa(y, x) \in L^2(\Omega \times D)$ and the eigenfunctions of the operator $\mathcal{R}$ satisfy $\phi_n(x) \in \dot{H}^s(D)$. Moreover,

$$\|T\kappa\|_{L^2(\Omega\times D)}^2 = \int_\Omega \|T\kappa(y, \cdot)\|_{L^2(D)}^2 \,\mathrm{d}y,$$

which is equivalent to $\|\kappa\|_{L^2(\Omega, \dot{H}^s(D))}^2$ by (3.5). Next we define

$$R_T(x, x') = \int_\Omega T\kappa(y, x')T\kappa(y, x)\mathrm{d}y \in L^2(D \times D),$$

and denote by $\mathcal{R}_T : L^2(D) \to L^2(D)$ the Hilbert-Schmidt operator associated with the kernel function $R_T(x, x')$. Obviously, $\mathcal{R}_T$ is compact and self-adjoint on $L^2(D)$. Let

$$\mathcal{R}_1 = T\mathcal{R}T, \tag{3.6}$$

with its domain $\mathscr{D}(\mathcal{R}_1) = \dot{H}^s(D)$. Then, its adjoint operator $\mathcal{R}_1^*$ is given by $\mathcal{R}_1^* = \mathcal{R}_T$. In fact, for any $v \in \dot{H}^s(D)$ and $w \in L^2(D)$, by Fubini's theorem and the symmetry of T, we have

$$\begin{aligned}(\mathcal{R}_1 v, w) = (T\mathcal{R}Tv, w) &= \int_\Omega (\int_D \kappa(y, x')Tv(x')\mathrm{d}x')(\int_D T\kappa(y, x)w(x)\mathrm{d}x)\mathrm{d}y \\ &= \int_\Omega (\int_D T\kappa(y, x')v(x')\mathrm{d}x')(\int_D T\kappa(y, x)w(x)\mathrm{d}x)\mathrm{d}y = (v, \mathcal{R}_T w).\end{aligned}$$

The following result shows the boundedness and the symmetry of the operator $\mathcal{R}_1$ restricted to $\dot{H}^s(D)$.

Lemma 3.2. *Let Assumption 3.1 hold. Then the following statements are valid:*

(i) $\mathcal{R}_1 \in \mathcal{B}(\dot{H}^s(D), L^2(D))$.

(ii) $\mathcal{R}_T \mid_{\dot{H}^s(D)} = \mathcal{R}_1$. *Hence, $\mathcal{R}_1$ is a symmetric operator with $\mathcal{R}_T$ as its self-adjoint extension operator.*

(iii) *There holds the identity*

$$\|T\kappa\|^2_{L^2(\Omega\times D)} = \sum_{n=1}^{\infty} \lambda_n \|T\phi_n\|^2_{L^2(D)} . \tag{3.7}$$

Proof. Since T is self-adjoint, we have for any $v \in \dot{H}^s(D)$

$$\begin{aligned}\|\mathcal{R}_1 v\|_{L^2(D)} &= \left\| \int_\Omega \int_D \kappa(y,x') T v(x') \mathrm{d}x' T\kappa(y,x) \mathrm{d}y \right\|_{L^2(D)} = \left\| \int_\Omega \int_D T\kappa(y,x') v(x') \mathrm{d}x' T\kappa(y,x) \mathrm{d}y \right\|_{L^2(D)} \\ &\leq \|v\|_{L^2(D)} \|T\kappa\|^2_{L^2(\Omega\times D)} \leq |v|_s \|T\kappa\|^2_{L^2(\Omega\times D)} \leq |v|_s \|\kappa\|^2_{L^2(\Omega,\dot{H}^s(D))} .\end{aligned}$$

Hence, $\mathcal{R}_1 \in \mathcal{B}(\dot{H}^s(D), L^2(D))$, which shows assertion (i).

To show assertion (ii), we proceed as follows: For any $g_1 \in \dot{H}^s(D)$ and $g_2 \in L^2(D)$, by the definition of $\mathcal{R}_1$, cf. (3.6), and by Fubini's theorem, we have

$$\begin{aligned}\int_D (\mathcal{R}_1 g_1)(x) g_2(x) \mathrm{d}x &= \int_\Omega (\int_D (T\kappa)(y,x') g_1(x') \mathrm{d}x')(\int_D (T\kappa)(y,x) g_2(x) \mathrm{d}x) \mathrm{d}y \\ &= \int_D (\int_D R_T(x',x) g_1(x') \mathrm{d}x') g_2(x) \mathrm{d}x = \int_D (\mathcal{R}_T g_1)(x) g_2(x) \mathrm{d}x.\end{aligned}$$

Thus, there holds

$$\mathcal{R}_1 = \mathcal{R}_T \mid_{\dot{H}^s(D)} . \tag{3.8}$$

Observe that $\mathcal{R}_1 = \mathcal{R}_1^*|_{\dot{H}^s(D)}$ and that $\mathcal{R}_T = \mathcal{R}_T^*$. Hence, $\mathcal{R}_1$ is a symmetric operator with $\mathcal{R}_T$ as its self-adjoint extension [42, pp. 197].

Finally, we show (3.7). By the definition of the trace of an operator and an application of Parseval's identity together with the self-adjointness of $\mathcal{R}_T$, we obtain

$$\mathrm{Tr}(\mathcal{R}_T) = \int_D R_T(x,x) \mathrm{d}x = \|T\kappa\|^2_{L^2(\Omega\times D)} . \tag{3.9}$$

Since $\{\xi_i\}_{i=1}^\infty \subset \dot{H}^s(D)$ is an orthonormal basis in $L^2(D)$ and since the trace is independent of the specific choice of the orthonormal basis [42, pp. 281] and relation (3.8), we deduce

$$\begin{aligned}\mathrm{Tr}(\mathcal{R}_T) &= \sum_{i=1}^\infty (\mathcal{R}_1 \xi_i, \xi_i) = \sum_{i=1}^\infty (T\mathcal{R}T\xi_i, \xi_i) = \sum_{i=1}^\infty (\mathcal{R}T\xi_i, T\xi_i) \\ &= \sum_{n=1}^\infty \lambda_n \sum_{i=1}^\infty (\phi_n, T\xi_i)^2 = \sum_{n=1}^\infty \lambda_n \sum_{i=1}^\infty (T\phi_n, \xi_i)^2 \\ &= \sum_{n=1}^\infty \lambda_n \|T\phi_n\|^2_{L^2(D)} ,\end{aligned}$$

where we have employed Parseval's identity and relation (3.5) in the last two identities. Then (3.7) follows from (3.9). □

Note at this point that $\mathcal{R}_1$ is not defined on the whole space $L^2(D)$ but only on a dense subspace $\dot{H}^s(D)$. Thus its adjoint with respect to $L^2(D)$ is bounded on $L^2(D)$ while $\mathcal{R}_1$ itself is unbounded on $L^2(D)$. We, however, consider $\mathcal{R}_1$ only on its restriction to $\dot{H}^s(D)$, which resolves this issue.

Now we are ready to derive the decay estimate of the eigenvalues of the operator $\mathcal{R}$.

Theorem 3.1. *Let Assumption 3.1 hold. Then $\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{d}{d+2s},1}$. In particular,*

$$\lambda_n \leq C_{\mathrm{em}}(d,s) C_{weyl}(d)^{-s} \mathrm{diam}(D)^{2s} \|\kappa\|^2_{L^2(\Omega,\dot{H}^s(D))} n^{-1-\frac{2s}{d}},$$

where $C_{\mathrm{em}}(d,s)$ denotes the embedding constant for $\ell_{\frac{d}{d+2s},1} \hookrightarrow \ell_{\frac{d}{d+2s},\infty}$.

Proof. By Lemma 3.2 (iii), it holds

$$\int_\Omega \|T\kappa(y,\cdot)\|^2_{L^2(D)}\,\mathrm{d}y = \sum_{n=1}^\infty \lambda_n \|T\phi_n\|^2_{L^2(D)}.$$

Since $\kappa(y,x) \in L^2(\Omega, \dot{H}^s(D))$, we have $T\kappa \in L^2(\Omega \times D)$ and

$$\sum_{n=1}^\infty \lambda_n \|T\phi_n\|^2_{L^2(D)} = \|\kappa\|^2_{L^2(\Omega,\dot{H}^s(D))} < \infty. \tag{3.10}$$

Next we apply a rearrangement trick which originates from the minmax principle [20, equation (11)] and obtain

$$\begin{aligned}\sum_{n=1}^\infty \lambda_n \|T\phi_n\|^2_{L^2(D)} &= (\lambda_1-\lambda_2)\|T\phi_1\|^2_{L^2(D)} + (\lambda_2-\lambda_3)\left(\|T\phi_1\|^2_{L^2(D)} + \|T\phi_2\|^2_{L^2(D)}\right)\\ &\quad + (\lambda_3-\lambda_4)\left(\|T\phi_1\|^2_{L^2(D)} + \|T\phi_2\|^2_{L^2(D)} + \|T\phi_3\|^2_{L^2(D)}\right) + \cdots\\ &\geq (\lambda_1-\lambda_2)\mu_1^2 + (\lambda_2-\lambda_3)(\mu_1^2+\mu_2^2) + (\lambda_3-\lambda_4)(\mu_1^2+\mu_2^2+\mu_3^2) + \cdots\\ &= \sum_{n=1}^\infty \lambda_n\mu_n^2 \geq \sum_{n=1}^\infty \lambda_n C_{\text{weyl}}(d)^s \mathrm{diam}(D)^{-2s} n^{\frac{2s}{d}}.\end{aligned} \tag{3.11}$$

Here, the last inequality is due to (3.4). The third line follows from the fact that for any $L^2(D)$-orthonormal system $\{e_n\}_{n=1}^m \subset \dot{H}^s(D)$ of m elements, the sum $\sum_{n=1}^m \|Te_n\|^2_{L^2(D)}$ achieves its minimum only if $\{e_n\}_{n=1}^m$ are the eigenfunctions corresponding to the first m smallest eigenvalues of the operator T, i.e.

$$\sum_{n=1}^m \|Te_n\|^2_{L^2(D)} \geq \sum_{n=1}^m \mu_n^2.$$

Thus, by the definition of the Lorentz sequence space, it follows from (3.10) and (3.11) that

$$\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{d}{d+2s},1} \quad \text{and} \quad \|\{\lambda_n\}_{n=1}^\infty\|_{\ell_{\frac{d}{d+2s},1}} \leq C_{\text{weyl}}(d)^{-s}\mathrm{diam}(D)^{2s}\|\kappa\|^2_{L^2(\Omega,\dot{H}^s(D))}.$$

Now Proposition 2.1 (ii) implies that $\ell_{\frac{d}{d+2s},1} \subsetneq \ell_{\frac{d}{d+2s},\infty}$ and, as a consequence,

$$\sup_n \left\{ n^{1+\frac{2s}{d}}\lambda_n :\ n \in \mathbb{N}\right\} := \|\{\lambda_n\}_{n=1}^\infty\|_{\ell_{\frac{d}{d+2s},\infty}} < \infty.$$

Due to the embedding $\ell_{\frac{d}{d+2s},1} \hookrightarrow \ell_{\frac{d}{d+2s},\infty}$, we obtain

$$\|\{\lambda_n\}_{n=1}^\infty\|_{\ell_{\frac{d}{d+2s},\infty}} \leq C_{\text{em}}(d,s)\|\{\lambda_n\}_{n=1}^\infty\|_{\ell_{\frac{d}{d+2s},1}}$$

with an embedding constant $C_{\text{em}}(d,s)$. Then $\lambda_n \leq \|\{\lambda_n\}_{n=1}^\infty\|_{\ell_{\frac{d}{d+2s},\infty}} n^{-1-\frac{2s}{d}}$, which gives the desired assertion. □

Next we show the eigenvalue decay rate $\mathcal{O}(n^{-1-\frac{2s}{d}})$ for the case $\kappa \in L^2(\Omega, H^s(I^d))$ via a similar argument. Here, due to the special structure of the domain $I^d = (-1,1)^d$, the boundary regularity is relaxed, and so is the trace condition on the function $\kappa(y,x)$, cf. Assumption 3.1. This result will later be needed in Section 3.2 as a stepping stone to obtain the decay rate for the case of the general domain $L^2(\Omega, H^s(D))$.

Proposition 3.1. *Let Assumption 2.1 hold and let $D = I^d$. Then $\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{d}{d+2s},1}$. In particular,*

$$\lambda_n \leq C_{\text{em}}(d,s)\|\kappa\|^2_{L^2(\Omega,H^s(I^d))} n^{-1-\frac{2s}{d}}.$$

Proof. The proof is analogous to that of Theorem 3.1, which essentially relies on Lemma 3.2. To this end, we define $A_j := -\partial_j((1-x_j^2)\partial_j)$, i.e. the one-dimensional singular Sturm-Liouville operator in the variable x_j, $1 \leq j \leq d$. Its n-th smallest nonzero eigenvalue is $n(n+1)$ for $n = 1,2,\cdots$. Now, let T be a tensor product of certain fractional powers of A_j, i.e., let $T = \Pi_{j=1}^d A_j^{\alpha_j}$, with $\alpha = \{\alpha_j\}_{j=1}^d \subset \mathbb{R}_+^d$ and $\alpha_j = \frac{s}{2d}$ for all $j = 1,\cdots,d$. One can readily check that T is a nonnegative and self-adjoint operator on $H^s(I^d)$ and its n-th smallest nonzero eigenvalue is $n^{\frac{s}{2d}}(n+1)^{\frac{s}{2d}}$ for $n = 1,2,\cdots$. Now, we can define the operators $\mathcal{R}_T$ and $\mathcal{R}_1$ as previously (actually, in this case, we now have $\mathcal{R}_T = \mathcal{R}_1$). Then, Lemma 3.2 and the desired assertion follow. □

Remark 3.1. *A close inspection of the proof indicates that the regularity requirement $\kappa \in L^2(\Omega, H^s(I^d))$ in Proposition 3.1 can be relaxed. We now need just the existence of a multi-index $\alpha = \{\alpha_j\}_{j=1}^d \subset \mathbb{R}_+^d$ with $\alpha_j = \frac{s}{2d}$ for all $j = 1, \cdots, d$ such that, with associated operator T, there holds $T\kappa \in L^2(\Omega, L^2(I^d))$.*

3.2 Case (b): $L^2(\Omega, H^s(D))$

Now we provide a singular value decay rate for the general case (b), i.e., $\kappa(y,x) \in L^2(\Omega, H^s(D))$. Our proof employs Stein's extension theorem and Proposition 3.1.

To this end, we will introduce the definition of approximation numbers. Given two Banach spaces E and F, the n-th approximation number $a_n(W)$ of an operator $W \in \mathcal{B}(E,F)$ is defined by

$$a_n(W) := \inf\{\|W - L\| : L \in \mathfrak{F}(E,F),\ \text{rank}(L) < n\}, \tag{3.12}$$

where $\mathfrak{F}(E,F)$ denotes the set of the finite rank operators.

Theorem 3.2. *(Eigenvalue estimate for the general space $L^2(\Omega, H^s(D))$.) Assume that D satisfies the strong local Lipschitz condition. Let $\kappa(y,x) \in L^2(\Omega, H^s(D))$. Then*

$$\lambda_n \le \text{diam}(D)^{2s} C_{\text{em}}(d,s) C_{\text{ext}}(D,s) \|\kappa\|^2_{L^2(\Omega, H^s(D))} n^{-1-\frac{2s}{d}},$$

where $C_{\text{ext}}(D,s)$ is a constant depending only on D and s.

Proof. Let $K \supset D$ be a d-dimensional cube with $\text{diam}(K) = \text{diam}(D)$. Then, by the strong local Lipschitz property of the domain D, Stein's extension theorem implies the existence of a bounded linear operator $\mathcal{E} : H^s(D) \to H^s(K)$ satisfying

$$\mathcal{E}\phi = \phi \text{ in } D \text{ and } \|\mathcal{E}\phi\|_{H^s(K)} \le \sqrt{C_{\text{ext}}(D,s)}\, \|\phi\|_{H^s(D)} \text{ for all } \phi \in H^s(D) \tag{3.13}$$

with $C_{\text{ext}}(D,s)$ being the extension constant that depends on D and s only.

This extension operator $\mathcal{E}$ allows for defining a bivariate function $\tilde{\kappa}(y,\cdot) \in H^s(K)$ for all $y \in \Omega$, s.t.

$$\tilde{\kappa}(y,\cdot) = \kappa(y,\cdot) \text{ in } D \quad \text{and} \quad \|\tilde{\kappa}\|_{L^2(\Omega, H^s(K))} \le \sqrt{C_{\text{ext}}(D,s)}\, \|\kappa\|_{L^2(\Omega, H^s(D))}. \tag{3.14}$$

We will denote $\mathcal{R}_K \in \mathcal{B}(L^2(K))$ as the corresponding Hilbert-Schmidt operator with the covariance function of $\tilde{\kappa}$ as its kernel. Its eigenvalues in a nonincreasing order are $\{\tilde{\lambda}_n\}_{n=1}^\infty$.

The scaling argument in the proof of Proposition 3.1 together with (3.14) then leads to

$$\tilde{\lambda}_n \le \text{diam}(D)^{2s} C_{\text{em}}(d,s) C_{\text{ext}}(D,s) \|\kappa\|^2_{L^2(\Omega, H^s(K))} n^{-1-\frac{2s}{d}}. \tag{3.15}$$

Given $\epsilon > 0$, the equivalence of approximation numbers and eigenvalues in a Hilbert space [31, Section 2.11.15] combined with (3.12) implies the existence of a self-adjoint operator $L_K \in \mathcal{B}(L^2(K))$ with $\text{rank}(L_K) < n$, satisfying

$$\|\mathcal{R}_K - L_K\|_{\mathcal{B}(L^2(K))} \le \tilde{\lambda}_n + \epsilon. \tag{3.16}$$

Note that L_K can be regarded as a rank $< n$ operator on $L^2(D)$. To prove the desired result, we only need to show

$$\|\mathcal{R} - L_K\|_{\mathcal{B}(L^2(D))} \le \|\mathcal{R}_K - L_K\|_{\mathcal{B}(L^2(K))}. \tag{3.17}$$

Then an application of (3.12) together with (3.16) and (3.15) leads to the assertion, after letting $\epsilon \to 0$.

To derive (3.17), we obtain by definition that

$$\begin{aligned}\|\mathcal{R} - L_K\|_{\mathcal{B}(L^2(D))} &= \sup_{0 \ne v \in L^2(D)} \frac{((\mathcal{R} - L_K)v, v)}{(v,v)} = \sup_{\substack{0 \ne \tilde{v} \in L^2(K) \\ \tilde{v}|_{K\setminus D} = 0}} \frac{((\mathcal{R}_K - L_K)\tilde{v}, \tilde{v})_K}{(\tilde{v},\tilde{v})_K} \\ &\le \sup_{0 \ne \tilde{v} \in L^2(K)} \frac{((\mathcal{R}_K - L_K)\tilde{v}, \tilde{v})_K}{(\tilde{v},\tilde{v})_K} = \|\mathcal{R}_K - L_K\|_{\mathcal{B}(L^2(K))}.\end{aligned}$$

Here, $(\cdot,\cdot)_K$ is the inner product on $L^2(K)$. This proves (3.17), and thus completes the proof. □

Note that this proof of (3.16) is in spirit similar to the proof of [7, Theorem 3.1]. Furthermore, note that we made the assumption of $C^{\lfloor s\rfloor,1}$-boundedness for the domain D of $\dot{H}^s(D)$ in Theorem 3.1 to guarantee the higher regularity of that space. Now, for the more general case of Theorem 3.2 involving $H^s(D)$, we do not need such higher regularity on the domain any more but just assume D to satisfy the strong local Lipschitz condition.

The results in Theorems 3.1 and 3.2 essentially improve the known eigenvalue estimates in [35] and [22]. There, decay rates of $\mathcal{O}(n^{-\frac{s}{d}})$ and $\mathcal{O}(n^{-\frac{2s}{d^*}})$ were shown, respectively, but both under somewhat higher regularity conditions on the bivariate function $\kappa(y,x)$ and using a finite element approximation (and not an orthogonal basis). Formally, our results are in the spirit of the estimate in [31, Section 6.4.31] and [26, Section 3.c.5], where it was established that $\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{2d}{4s+d},2}$. Nevertheless, that result is slightly weaker than our $\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{d}{d+2s},1}$ from Theorem 3.2 and Theorem 3.1, due to the strict inclusion $\ell_{\frac{d}{d+2s},1} \subsetneq \ell_{\frac{2d}{4s+d},2}$, cf. Proposition 2.1. Our higher decay rate stems from the nonnegativity and the symmetry of the covariance function $R(x,x')$, which was not exploited in the previous results ([31, Section 6.4.31] and [26, Section 3.c.5]).

3.3 Examples and KL truncation error estimates

Now, we provide two examples to illustrate the optimality of our estimate.

Example 3.1. *For $s=0$ we have $\kappa(y,x)\in L^2(\Omega\times D)$. Since*

$$\sum_{n=1}^\infty \lambda_n = \|\kappa\|^2_{L^2(\Omega\times D)} < \infty,$$

we immediately obtain $\{\lambda_n\}_{n=1}^\infty \in \ell_1$, which is the best possible estimate for this type of operator. Theorem 3.1 also implies $\{\lambda_n\}_{n=1}^\infty \in \ell_1$. But in contrast, we can have by [31, Section 6.2.15 or Section 6.4.31] only $\{\lambda_n\}_{n=1}^\infty \in \ell_2 \supsetneq \ell_1$. Thus our estimate is clearly superior.

Example 3.2. *The isotropic Matérn kernel is defined by*

$$G_\nu(|x-y|) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right)^\nu K_\nu\left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right) \text{ in } D\times D,$$

where ν is the smoothing parameter, σ^2 is the variance, ρ is a length scale parameter, Γ is the Gamma function, and K_ν denotes the modified Bessel function of the second kind. Consider the one-dimensional case with $D=(0,1)$, $\sigma=1$, $\rho=1$ and take $\nu=1/2$ and $3/2$, i.e.,

$$G_{\frac{1}{2}}(|x-y|) = e^{-|x-y|} \in H^{\frac{3}{2}-\delta}(D\times D)$$

and

$$G_{\frac{3}{2}}(|x-y|) = (1+\sqrt{3}|x-y|)e^{-\sqrt{3}|x-y|} \in H^{\frac{7}{2}-\delta}(D\times D),$$

respectively, where $\delta\in(0,1/2)$ is arbitrary. Such kernels are popular in machine learning [14]. The decay rates of their singular values have been numerically computed in [22, Section 6.3]. The results therein show that the square of the n-th largest singular value of $G_{\frac{1}{2}}(|x-y|)$ and $G_{\frac{3}{2}}(|x-y|)$ decays like $\mathcal{O}(n^{-4})$ and $\mathcal{O}(n^{-8})$, respectively, which is in excellent agreement with the theoretical predictions of our Theorem 3.2. In addition, we can infer that $G_{\frac{1}{2}}\in L^2(D,\dot{H}^{\frac{3}{2}-\delta}(D))$. Thus, the same decay rate can also be deduced from our Theorem 3.1 in that case.

Now we will consider the convergence rate of the KL truncation error (2.6). The following error estimate is an immediate consequence of the Theorems 3.1 and 3.2.

Theorem 3.3. *Let Assumption 2.1 hold. Then, for any $1\le M\in\mathbb{N}$, there holds*

$$\left\|\kappa(y,x) - \sum_{n=1}^M \sqrt{\lambda_n}\phi_n(x)\psi_n(y)\right\|_{L^2(\Omega\times D)} \le C(M+1)^{-\frac{s}{d}}$$

with the constant

$$C := \operatorname{diam}(D)^s C_{\mathrm{em}}(d,s)^{\frac{1}{2}} C_{\mathrm{ext}}(D,s)^{\frac{1}{2}} \|\kappa\|_{L^2(\Omega,H^s(D))}\sqrt{\frac{d}{2s}}.$$

Proof. Under the given assumptions, Theorem 3.2 can be applied to obtain $\{\lambda_n\}_{n=1}^\infty \in \ell_{\frac{d}{2s+d},1}$ and

$$\lambda_n \le \operatorname{diam}(D)^{2s} C_{\mathrm{em}}(d,s) C_{\mathrm{ext}}(D,s) \|\kappa\|^2_{L^2(\Omega,H^s(D))} n^{-1-\frac{2s}{d}}. \tag{3.18}$$

For any $1 \le M \in \mathbb{N}$, the $L^2(D)$-orthonormality of the sequence $\{\phi_n(x)\}_{n=1}^\infty$ then yields

$$\left\| \kappa(y,x) - \sum_{n=1}^{M} \sqrt{\lambda_n}\phi_n(x)\psi_n(y) \right\|^2_{L^2(\Omega\times D)} = \sum_{n=M+1}^{\infty} \lambda_n.$$

In view of the eigenvalue estimate (3.18), we obtain immediately

$$\begin{aligned} &\left\| \kappa(y,x) - \sum_{n=1}^{M} \sqrt{\lambda_n}\phi_n(x)\psi_n(y) \right\|^2_{L^2(\Omega\times D)} \\ &\le \operatorname{diam}(D)^{2s} C_{\mathrm{em}}(d,s) C_{\mathrm{ext}}(D,s) \|\kappa\|^2_{L^2(\Omega,H^s(D))} \sum_{n=M+1}^{\infty} n^{-\frac{2s}{d}-1} \\ &\le \operatorname{diam}(D)^{2s} C_{\mathrm{em}}(d,s) C_{\mathrm{ext}}(D,s) \|\kappa\|^2_{L^2(\Omega,H^s(D))} \frac{d}{2s}(M+1)^{-\frac{2s}{d}}, \end{aligned}$$

which shows the desired assertion after taking the square root on both sides. □

Note here that the error estimate of our Theorem 3.3 for the KL truncation of a bivariate function $\kappa(y,x) \in L^2(\Omega, H^s(D))$ for the case $D = I$ is identical to that in [3, Proposition 3.1], which was derived for the special situation $\kappa(y,x) \in L^2(\Omega, H^s(I))$ with $I = (-1,1)$ only. The authors of [3] however mention that their proof can be generalized to the situation when $D = I^d := [-1,1]^d$, which would result in a rate of $\mathcal{O}(M^{-s/d})$ for the M-term truncation error.

Last we show the sharpness of Theorem 3.3.

Theorem 3.4. *Let $M > 0$ be given. Then there exists a bivariate function $g(y,x) \in H^s(I^d \times I^d)$ satisfying Assumption 3.1, such that*

$$\inf_{\substack{\{u_n\}_{n=1}^M \subset L^2(I^d) \\ \{v_n\}_{n=1}^M \subset L^2(I^d)}} \left\| g(y,x) - \sum_{n=1}^{M} u_n(x)v_n(y) \right\|_{L^2(I^d\times I^d)} \gtrsim M^{-\frac{s}{d}}.$$

Proof. The proof follows as in [4], where $g(y,x) \in \dot{H}^s(I^d) \times \dot{H}^s(I^d)$ was constructed to show the lower bound in the truncation estimate. □

Note that Theorem 3.4 implies the sharpness of our eigenvalue decay estimates of Theorems 3.1 and 3.2 by a simple contradiction argument.

4 Application to elliptic PDEs with random coefficient

In this section, we use the results of Section 3 to analyze a model order reduction algorithm for a class of elliptic PDEs with random coefficient in the multi-query context. In the algorithm, we apply the Karhunen-Loève approximation to the stochastic diffusion coefficient $\kappa(y,x)$ to arrive at a truncated model with finite-dimensional noise. We shall provide an error analysis below. Throughout this section, we assume that the conditions of Theorem 3.3 are satisfied.

Let D be an open bounded domain in $\mathbb{R}^d$ with a strong local Lipchitz boundary, and let $(\Omega, \Sigma, \mathcal{P})$ be a given probability space. Consider the elliptic PDE with random coefficient

$$\begin{aligned} \mathcal{L}u(y,\cdot) &= f, \quad x \in D, \\ u(y,\cdot) &= 0, \quad x \in \partial D, \end{aligned} \tag{4.1}$$

for a.e. $y \in \Omega$, where the elliptic operator $\mathcal{L}$ is defined by

$$\mathcal{L}u(y,\cdot) = -\nabla\cdot(\kappa(y,x)\nabla u(y,x)),$$

and ∇ denotes taking the derivative with respect to the spatial variable x. We assume the diffusion coefficient $\kappa(y,x)$ to be $\kappa(y,\cdot) \in L^\infty(D)$ almost surely and the force term $f(x)$ to be $f \in H^{-1}(D)$. In model (4.1), the dependence of the diffusion coefficient $\kappa(y,x)$ on a stochastic variable $y \in \Omega$ reflects imprecise knowledge or lack of information.

The extra-coordinate y poses significant computational challenges. One popular approach is the stochastic Galerkin method [5]. There, one often approximates the stochastic diffusion coefficient $\kappa(y,x)$ by a finite sum of products of deterministic and stochastic orthogonal basis (with respect to a certain probability measure). This gives a computationally more tractable finite-dimensional noise model. Then, the choice of the employed orthogonal basis[6] is crucial for the accurate and efficient approximation to $\kappa(y,x)$.

In this article, we just consider the KL approximation $\kappa_M(y,x)$ of the random field $\kappa(y,x)$, cf. (2.5). First, we specify the functional analytic setting. Let $V = H_0^1(D)$ with the inner product $\langle v_1, v_2\rangle = (\nabla v_1, \nabla v_2)$ and the induced norm $|v|_{H^1(D)} = \sqrt{\langle v, v\rangle}$, and let $H^{-1}(D)$ be its dual space. Then, for any given $y \in \Omega$, the weak formulation of problem (4.1) is to find $u(y,x) \in V$ such that

$$\int_D \kappa(y,x)\nabla u(y,x)\cdot\nabla v(x)\mathrm{d}x = \int_D f(x)v(x)\mathrm{d}x \quad \forall v \in V. \tag{4.2}$$

To analyze its well-posedness, we make some conventional assumptions [17].

Assumption 4.1 (Uniform ellipticity assumption on κ)**.** *There exist some constants α and β, $0 < \alpha < \beta$ such that*

$$\alpha \le \kappa(y,x) \le \beta, \ \forall (y,x) \in \Omega \times D.$$

Under Assumption 4.1, the weak formulation (4.2) is well-posed due to the Lax-Milgram theorem, and

$$|u(y,\cdot)|_{H^1(D)} \le \alpha^{-1}\|f\|_{H^{-1}(D)} \quad \forall y \in \Omega. \tag{4.3}$$

Thus $\mathcal{L} : V \to H^{-1}(D)$ is an invertible linear operator with inverse $\mathcal{S} = \mathcal{L}^{-1} : H^{-1}(D) \to V$ that both depend on the stochastic diffusion coefficient $\kappa(y,x)$. Clearly, $\mathcal{S}$ is a self-adjoint operator for all $y \in \Omega$.

To analyze the truncated model with the KL truncation $\kappa_M(y,x)$, we further assume the following two conditions on the $L^2(\Omega)$-orthonormal bases $\psi_n(y)$ from (2.3), and on the truncated series $\kappa_M(y,x)$.

Assumption 4.2. *There exists some $\theta > 0$ such that $|\psi_n(y)| \le \theta < \infty$, $\forall\, n \in \mathbb{N}$ and $y \in \Omega$.*

Assumption 4.3 (Uniform ellipticity assumption on κ_M)**.** *There exist some constants[7] α and β, $0 < \alpha < \beta$ such that*

$$\alpha \le \kappa_M(y,x) \le \beta, \ \forall (y,x) \in \Omega \times D.$$

Note here that Assumption 4.2 allows to derive a KL truncation error in $L^2(D)$ which is uniform in y. Indeed, under Assumption 4.2, and for $\psi_n(y)$ as defined in (2.3), Theorem 3.3 implies

$$\|\kappa(y,\cdot) - \kappa_M(y,\cdot)\|^2_{L^2(D)} = \sum_{n>M}\lambda_n|\psi_n(y)|^2 \le \theta^2\sum_{n>M}\lambda_n \lesssim \theta^2\frac{d}{2s}(M+1)^{-\frac{2s}{d}} \quad \forall y \in \Omega. \tag{4.4}$$

Furthermore, Assumptions 4.2 and 4.3 together enable the control of the KL truncation error in $L^p(D)$ which is uniform in y. Indeed, for $p \ge 2$ and for given $y \in \Omega$, we obtain the bound

$$\begin{aligned}\|\kappa(y,\cdot) - \kappa_M(y,\cdot)\|_{L^p(D)} &\le \|\kappa(y,\cdot) - \kappa_M(y,\cdot)\|^{\frac{2}{p}}_{L^2(D)}\|\kappa(y,\cdot) - \kappa_M(y,\cdot)\|^{\frac{p-2}{p}}_{L^\infty(D)}\\ &\lesssim \theta^{\frac{2}{p}}\left(\frac{d}{2s}\right)^{\frac{1}{p}}(M+1)^{-\frac{2s}{dp}}\beta^{\frac{p-2}{p}}.\end{aligned} \tag{4.5}$$

This estimate will now be used to bound the error of the solution to problem (4.1) due to the KL truncation. After substituting the KL approximation $\kappa_M(y,x)$ of the diffusion coefficient $\kappa(y,x)$ into problem (4.1), we arrive at a truncated problem with finite-dimensional noise: For a.e. $y \in \Omega$

$$\begin{aligned}\mathcal{L}_M u_M(y,\cdot) &= f, \quad x \in D,\\ u_M(y,\cdot) &= 0, \quad x \in \partial D,\end{aligned} \tag{4.6}$$

[6] Note that instead of an expansion in the eigenbasis, there are other choices, like a polynomial chaos expansion [35, 12]. Moreover, there is the expansion with respect to the hierarchical Faber basis or some wavelet type basis, i.e. to a *local* basis. In certain situation this allows to further improve on the approximation rate of u, for details, see [12, 6].

[7] Here, for notational simplicity, we have assumed the same constants as in Assumption 4.1.

where $\mathcal{L}_M$ is the elliptic differential operator with the diffusion coefficient κ_M. The corresponding weak formulation is then to find $u_M(y,x) \in V$ such that

$$\int_D \kappa_M(y,x)\nabla u_M(y,x)\cdot\nabla v(x)\mathrm{d}x = \int_D f(x)v(x)\mathrm{d}x \quad \forall v \in V, \tag{4.7}$$

for any given $y \in \Omega$. Under Assumption 4.3 on the KL truncation κ_M, we get the well-posedness of problem (4.7) by the Lax-Milgram theorem. As before, we set

$$\mathcal{S}_M = \mathcal{L}_M^{-1}.$$

Then $\mathcal{S}_M$ is a self-adjoint operator for all $y \in \Omega$. Clearly, the solution $u_M(y,x) = \mathcal{S}_M(y)f$ corresponds to the perturbed coefficient $\kappa_M(y,x)$ (relative to the unperturbed coefficient $\kappa(y,x)$).

The next lemma quantifies the effect of the perturbation of the coefficient $\kappa(y,x)$ on the solution $u(y,x)$.

Lemma 4.1. *Let Assumptions 4.1, 4.2 and 4.3 hold. Then, for given $p_1 \geq 2$ and $\frac{1}{p_1}+\frac{1}{p_2}=\frac{1}{2}$, we have*

$$|u(y,\cdot)-u_M(y,\cdot)|_{H^1(D)} \lesssim \frac{1}{\alpha}\theta^{\frac{2}{p_1}}(\frac{d}{2s})^{\frac{1}{p_1}}(M+1)^{-\frac{2s}{dp_1}}\beta^{\frac{p_1-2}{p_1}}\|\nabla u_M(y,\cdot)\|_{L^{p_2}(D)}.$$

Proof. From the weak formulations for $u(y,x)$ and $u_M(y,x)$, cf. (4.2) and (4.7), we obtain for any $y \in \Omega$

$$\begin{aligned}&\int_D \kappa(y,x)\nabla(u(y,x)-u_M(y,x))\cdot\nabla v(x)\mathrm{d}x \\ &\qquad= \int_D(\kappa_M(y,x)-\kappa(y,x))\nabla u_M(y,x)\cdot\nabla v(x)\mathrm{d}x \quad \forall v \in V.\end{aligned} \tag{4.8}$$

By setting $v = u - u_M \in V$ in the weak formulation (4.8), using Assumption 4.1 and the generalized Hölder inequality, we have

$$\begin{aligned}\alpha\,|u(y,\cdot)-u_M(y,\cdot)|^2_{H^1(D)} &\leq \int_D \kappa(y,x)|\nabla(u(y,x)-u_M(y,x))|^2\mathrm{d}x \\ &= \int_D(\kappa_M(y,x)-\kappa(y,x))\nabla u_M(y,x)\cdot\nabla(u(y,x)-u_M(y,x))\mathrm{d}x \\ &\leq \|\kappa_M(y,\cdot)-\kappa(y,\cdot)\|_{L^{p_1}(D)}\,|u(y,\cdot)-u_M(y,\cdot)|_{H^1(D)}\,\|\nabla u_M(y,\cdot)\|_{L^{p_2}(D)}.\end{aligned}$$

Consequently, by (4.5), we get

$$|u(y,\cdot)-u_M(y,\cdot)|_{H^1(D)} \lesssim \frac{1}{\alpha}\theta^{\frac{2}{p_1}}(\frac{d}{2s})^{\frac{1}{p_1}}(M+1)^{-\frac{2s}{dp_1}}\beta^{\frac{p_1-2}{p_1}}\|\nabla u_M(y,\cdot)\|_{L^{p_2}(D)}.$$

□

The estimate in Lemma 4.1 depends on the bound $\|\nabla u_M(y,\cdot)\|_{L^{p_2}(D)}$. Using Meyers' Theorem [30], this term can be directly controlled by the force term $f(x)$, provided that it possesses higher integrability.

Theorem 4.1 (Meyers' theorem)**.** *There exist a number $p_2 > 2$ and a positive constant $C(\alpha,\beta,D,p_2) > 0$, which both depend only on α,β, D and p_2, such that if $f \in W^{-1,p_2'}(D)$, with ${p_2}^{-1}+{p_2'}^{-1}=1$, then the solution $u_M(y,\cdot) \in W_0^{1,p_2}(D)$ and satisfies*

$$\|u_M(y,\cdot)\|_{W_0^{1,p_2}(D)} \leq C(\alpha,\beta,D,p_2)\,\|f\|_{W^{-1,p_2'}(D)}.$$

The largest possible number p_2 in Theorem 4.1 is called Meyer's exponent and is denoted by P.

Assumption 4.4. *$f \in W^{-1,p_2'}(D)$, for some $2 < p_2 < P$.*

Finally, under Assumptions 4.1, 4.2, 4.3 and 4.4, and by combining the preceding results, we obtain the following error estimate of the solution u_M due to KL truncation:

Theorem 4.2. *Let Assumptions 4.1, 4.2, 4.3, and 4.4 hold. Then, for $p_1 = \frac{2p_2'}{2-p_2'}$, we have*

$$|u(y,\cdot)-u_M(y,\cdot)|_{H^1(D)} \lesssim C(\alpha,\beta,D,p_2)\frac{1}{\alpha}\theta^{\frac{2}{p_1}}(\frac{d}{2s})^{\frac{1}{p_1}}(M+1)^{-\frac{2s}{dp_1}}\beta^{\frac{p_1-2}{p_1}}\|f\|_{W^{-1,p_2'}(D)}.$$

Theorem 4.2 provides an error estimate of $\mathcal{S}_M$ to $\mathcal{S}$ in the operator norm. Indeed, we have

$$|(\mathcal{S}-\mathcal{S}_M)f|_{H^1(D)} \lesssim C(\alpha,\beta,D,p_2)\frac{1}{\alpha}\theta^{\frac{2}{p_1}}(\frac{d}{2s})^{\frac{1}{p_1}}(M+1)^{-\frac{2s}{dp_1}}\beta^{\frac{p_1-2}{p_1}}\|f\|_{W^{-1,p_2'}(D)}.$$

This in particular implies the convergence in operator norm as $M \to \infty$.

5 Concluding remarks

In this paper, we have derived a new estimate for the sigular value decay rate of order $\mathcal{O}(n^{-1/2-\frac{s}{d}})$ for bivariate functions in $L^2(\Omega, \dot{H}^s(D))$ and $L^2(\Omega, H^s(D))$. This result improves on known results in the literature. Our new estimate was established by analyzing the eigenvalue of the kernel operator $\mathcal{R}$ using two different techniques, i.e., a rearrangement trick originating from the minmax principle, and Stein's extension together with an operator theoretic argument, respectively. We demonstrated its usefulness in the analysis of an algorithm for solving stochastic elliptic PDEs, which employs the Karhunen-Loève truncation of the stochastic diffusion coefficient and provided an error estimate for the truncation approximation. Our improved decay rate and the resulting error estimate can be applied to many other problems as well.

Note furthermore that our approach allows to also deal with negative values s of isotropic smoothness on D. A simple consideration shows the validity of our result also for the case $s \in (-d/2, 0)$. This may be helpful for integral operators with weakly singular kernels, which have applications in, e.g., image and video processing.

Note finally that we have only analyzed the KL truncation error for approximating Sobolev smooth bivariate functions at the continuous level. This is of course only the first step in the analysis of a numerical method which is, after discretization, based on such a truncated series expansion. Any efficient overall numerical algorithm still needs a proper sampling or discretization method to approximate the integrals on Ω (which is a challenging task when it comes to high-dimensional problems), and a suitable discretization algorithm on D to approximate the continuous eigenfunctions (which we here assumed to have at our disposal). Beyond the KL truncation error, these two additional types of approximation errors surely need also be taken into account. The further balancing of all these errors and their corresponding numerical costs will be future work.

Acknowledgements

The authors were supported by the Hausdorff Center for Mathematics in Bonn and the Sonderforschungsbereich 1060 The Mathematics of Emergent Effects funded by the Deutsche Forschungsgemeinschaft. We thank Priv.-Doz. Dr. Christian Rieger for carefully reading our paper and pointing out certain aspects of kernels and bivariate functions, which helped to improve our result. We also thank Dr. Tino Ullrich for his help and fruitful discussion.

References

[1] R. Adams and J. Fournier. *Sobolev Spaces*. Elsevier/Academic Press, Amsterdam, 2003.

[2] S. Agmon. *Lectures on Elliptic Boundary Value Problems*. Princeton, N.J.-Toronto-London, 1965.

[3] M. Azaïez and F. Belgacem. Karhunen-Loève's truncation error for bivariate functions. *Comput. Methods Appl. Mech. Engrg.*, 290:57–72, 2015.

[4] M. Babaev. Approximation to the Sobolev classes W_q^r of functions of several variables by bilinear forms in L_p for $2 \leq q \leq p \leq \infty$. *Mat. Zametki*, 62(1):18–34, 1997.

[5] I. Babuška, R. Tempone, and G. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825, 2004.

[6] M. Bachmayr, A. Cohen, D. D$\tilde{u}$ng, and C. Schwab. Fully discrete approximation of parametric and stochastic elliptic PDEs. *SIAM J. Numer. Anal.*, 55(5):2151–2186, 2017.

[7] M. Bachmayr, A. Cohen, and G. Migliorati. Representations of Gaussian random fields and approximation of elliptic PDEs with lognormal coefficients. *J. Fourier Anal. Appl.*, 2017. To appear.

[8] A. Balakrishnan. Fractional powers of closed operators and the semigroups generated by them. *Pacific J. Math.*, 10:419–437, 1960.

[9] V. Calo, Y. Efendiev, J. Galvis, and G. Li. Randomized oversampling for generalized multiscale finite element methods. *Multiscale Model. Simul.*, 14(1):482–501, 2016.

[10] A. Caponnetto and E. De Vito. Optimal rates for the regularized least-squares algorithm. *Found. Comput. Math.*, 7(3):331–368, 2007.

[11] A. Cohen and R. DeVore. Approximation of high-dimensional parametric PDEs. *Acta Numer.*, 24:1–159, 2015.

[12] A. Cohen, R. DeVore, and C. Schwab. Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. *Found. Comput. Math.*, 10(6):615–646, 2010.

[13] R. Courant and D. Hilbert. *Methods of Mathematical Physics. Vol. I.* Interscience Publishers, Inc., New York, N.Y., 1953.

[14] F. Cucker and S. Smale. On the mathematical foundations of learning. *Bull. Amer. Math. Soc. (N.S.)*, 39(1):1–49, 2002.

[15] W. Dahmen, R. DeVore, L. Grasedyck, and E. Süli. Tensor-sparsity of solutions to high-dimensional elliptic partial differential equations. *Found. Comput. Math.*, 16(4):813–874, 2016.

[16] R. DeVore. Nonlinear approximation. *Acta Numer.*, 7:51–150, 1998.

[17] R. DeVore. The theoretical foundation of reduced basis methods. In *Model Reduction and Approximation: Theory and Algorithms*. SIAM, Philadelphia, PA, 2016.

[18] D. Dũng, V. Temlyakov, and T. Ullrich. Hyperbolic Cross Approximation. *Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser/Springer*, 2017. To appear.

[19] Y. Efendiev, J. Galvis, and T. Hou. Generalized multiscale finite element methods (GMsFEM). *J. Comput. Phys.*, 251:116–135, 2013.

[20] C. Foias, O. Manley, and L. Sirovich. Empirical and Stokes eigenfunctions and the far-dissipative turbulent spectrum. *Phys. Fluids A*, 2(3):464–467, 1990.

[21] R. Ghanem and P. Spanos. *Stochastic Finite Elements: a Spectral Approach*. Courier Corporation, 2003.

[22] M. Griebel and H. Harbrecht. Approximation of bi-variate functions: singular value decomposition versus sparse grids. *IMA J. Numer. Anal.*, 34(1):28–54, 2014.

[23] S. Heinrich. On the optimal error of degenerate kernel methods. *J. Integral Equations*, 9(3):251–266, 1985.

[24] A. Iserles and S. Nørsett. From high oscillation to rapid approximation. I. Modified Fourier expansions. *IMA J. Numer. Anal.*, 28(4):862–887, 2008.

[25] H. Komatsu. Fractional powers of operators. *Pacific J. Math.*, 19:285–346, 1966.

[26] H. König. *Eigenvalue Distribution of Compact Operators*. Birkhäuser Verlag, Basel, 1986.

[27] G. Li. *Multiscale Model Reduction for High-contrast Flow Problems*. PhD thesis, Texas A&M University, 2015.

[28] P. Li and S. Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156(3-4):153–201, 1986.

[29] J. Melenk and I. Babuška. The partition of unity finite element method: basic theory and applications. *Comput. Methods Appl. Mech. Engrg.*, 139(1-4):289–314, 1996.

[30] N. Meyers. An L^p-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 17:189–206, 1963.

[31] A. Pietsch. *Eigenvalues and s-numbers*. Cambridge University Press, Cambridge, 1987.

[32] R. Schaback and H. Wendland. Approximation by positive definite kernels. In *Advanced Problems in Constructive Approximation*, pages 203–222. Springer, 2002.

[33] R. Schneider and A. Uschmajew. Approximation rates for the hierarchical tensor format in periodic Sobolev spaces. *J. Complexity*, 30(2):56–71, 2014.

[34] C. Schwab and C. Gittelson. Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numer.*, 20:291–467, 2011.

[35] C. Schwab and R. Todor. Karhunen-Loève approximation of random fields by generalized fast multipole methods. *J. Comput. Phys.*, 217(1):100–122, 2006.

[36] V. Temlyakov. Bilinear approximation and applications. *Trudy Mat. Inst. Steklov.*, 187:191–215, 1989. Translated in Proc. Steklov Inst. Math. **1**990, no. 3, 221–248, Studies in the theory of differentiable functions of several variables and its applications, 13 (Russian).

[37] V. Temlyakov. Estimates for the best bilinear approximations of functions and approximation numbers of integral operators. *Mat. Zametki*, 51(5):125–134, 159, 1992.

[38] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin, 2006.

[39] R. Todor. Robust eigenvalue computation for smoothing operators. *SIAM J. Numer. Anal.*, 44(2):865–878, 2006.

[40] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. Johann Ambrosius Barth Verlag, Heidelberg, second edition, 1995.

[41] H. Triebel. *Theory of Function Spaces III*. Monographs in Mathematics. Birkhäuser Basel, 2006.

[42] K. Yosida. *Functional Analysis*. Springer-Verlag, Berlin-New York, 1980.