



Universität
Basel

Universitätsbibliothek

edoc

Institutional Repository of the University of Basel

University Library

Schoenbeinstrasse 18-20

CH-4056 Basel, Switzerland

<http://edoc.unibas.ch/>

UB | Für alle offen

Year: 2015

Efficient approximation of random fields for numerical applications

Harbrecht, Helmut and Peters, Michael and Siebenmorgen, Markus

Posted at edoc, University of Basel

Official URL: <http://edoc.unibas.ch/dok/A6419797>

This is the peer reviewed version of the following article:

Harbrecht, Helmut and Peters, Michael and Siebenmorgen, Markus. (2015) *Efficient approximation of random fields for numerical applications*. Numerical linear algebra with applications, Vol. 22, H. 4. S. 596-617.

which has been published in final form at <http://dx.doi.org/10.1002/nla.1976>. This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

Efficient approximation of random fields for numerical applications[†]

Helmut Harbrecht, Michael Peters*, Markus Siebenmorgen

Mathematisches Institut, Rheinsprung 21, 4051 Basel, Switzerland

SUMMARY

This article is dedicated to the rapid computation of separable expansions for the approximation of random fields. We consider approaches based on techniques from the approximation of non-local operators on the one hand and based on the pivoted Cholesky decomposition on the other hand. Especially, we provide an a-posteriori error estimate for the pivoted Cholesky decomposition in terms of the trace. Numerical examples are provided to validate and quantify the presented methods. Copyright © 0000 John Wiley & Sons, Ltd.

Received ...

1. INTRODUCTION

In this article, we present and compare two different approaches for the approximation of random fields in $L^2_{\mathbb{P}}(\Omega, H^p(D))$ for a spatial domain $D \subset \mathbb{R}^d$ and a separable, complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Stochastic fields appear for example in the modeling of diffusion problems with random data, see e.g. [1], and in machine learning, see e.g. [2]. To make a stochastic field $a(\mathbf{x}, \omega)$ feasible for numerical computations in a stochastic Galerkin or stochastic collocation method, see e.g. [1, 3, 4, 5, 6, 7] and the references therein, one has to separate the spatial variable \mathbf{x} and the stochastic variable ω . Since $L^2_{\mathbb{P}}(\Omega, H^p(D)) \cong L^2_{\mathbb{P}}(\Omega) \otimes H^p(D)$, see e.g. [8], this can be accomplished by computing a basis representation of a in $L^2_{\mathbb{P}}(\Omega) \otimes H^p(D)$. A very common approach to obtain such a representation is the *Karhunen-Loève expansion*, cf. [1, 9], which can be regarded as the linear operator analogue of the singular value decomposition of matrices.

The main task in the computation of a Karhunen-Loève expansion is the solution of a symmetric and positive semidefinite eigen-problem. In this context, approaches to efficiently compute the Karhunen-Loève expansion have been made by means of the *Fast Multipole Method* (FMM) based on interpolation, cf. [10], in [11] and with the aid of \mathcal{H} -matrices, cf. [12], in [13]. The idea in these articles is to provide a data-sparse representation of the covariance operator which is then used to solve the related eigen-problem numerically by a Krylov subspace method, cf. [14]. Of course, another algorithm for the efficient approximation of non-local operators, like the *Adaptive Cross Approximation* (ACA), cf. [15, 16], or the *Wavelet Galerkin Scheme* (WGS), cf. [17, 18], can be considered as well for the representation of the covariance operator. Nevertheless, the major drawback of these approaches is that the number of eigenvalues to be computed has to be known in advance which might be a strong assumption in practice.

To overcome this obstruction, we present here an alternative approach based on the *Pivoted Cholesky Decomposition* (PCD). The PCD is an established tool in the simulation of Gaussian

*Correspondence to: E-mail: michael.peters@unibas.ch

[†]This research has been supported by the Swiss National Science Foundation (SNSF) through the project “Rapid Solution of Boundary Value Problems on Stochastic Domains”.

processes and the computation of low-rank approximations to covariance matrices, see e.g. [2, 19, 20]. It can be interpreted as a single-block ACA with applicable total pivoting, cf. [21]. Hence, only the main diagonal of the discretized operator has to be precomputed, which can be performed in essentially, i.e. up to possible poly-logarithmic terms, linear complexity, if the quadrature proposed in [22] is applied to discretize the operator. Then, in each step of the algorithm, the quality of the approximation with respect to the stochastic field is controllable by means of the trace. If the desired accuracy is achieved, the algorithm stops with an M -term approximation to the operator. If M is substantially smaller than the dimension of the ansatz space, we end up with a remarkable computational speed-up. The related Karhunen-Loève expansion might then be computed in a post-processing step. In this case, the PCD yields a full but relatively small eigenproblem if the operator under consideration exhibits a certain smoothness. This eigenproblem might be solved numerically by e.g. the *QR-algorithm*, cf. [23].

Now the following question arises: which approach is more efficient? We will try to answer this question numerically by comparing the PCD with methods lend from the approximation of non-local operators. We employ here ACA for the data-sparse approximation of the covariance operator which results in a fast matrix-vector product. Thus, a Krylov subspace method – we use the *Implicit Restarted Arnoldi Method* (IRAM), cf. [24, 25, 26] – is feasible to compute the desired eigenvalues of largest magnitude.

Finally, we would like to emphasize that, although we focus here on the application to random fields, the presented methods are also applicable in the more general case of approximating bivariate functions in $L^2(D_1) \otimes L^2(D_2)$ for two domains $D_1 \subset \mathbb{R}^{d_1}$ and $D_2 \subset \mathbb{R}^{d_2}$.

The rest of this article is structured as follows. Section 2 is devoted to the approximation of random fields. We start by introducing the Karhunen-Loève expansion. After this, we introduce error bounds for the approximation of random fields in terms of operator traces. Especially, we discuss here error estimates including discretization and truncation error. We also provide bounds for the decay of the covariance operator's eigenvalues. In Section 3, we provide the theoretical background for the pivoted Cholesky decomposition. To that end, we consider separable representations of random fields which are more general than the Karhunen-Loève expansion. After this, we establish error estimates for the approximation of random fields by the pivoted Cholesky decomposition. These estimates are essential for the a-posteriori control of the approximation error. Section 4 introduces a special class of covariance functions based on the Matérn kernel functions. We choose this class of covariance functions for our numerical tests, since we a-priori know the decay rate of the respective eigenvalues. In particular, we are also able to analytically compute the eigenfunctions and eigenvalues in the case of the unit sphere \mathbb{S}^2 . Thus, these kernels provide an excellent benchmark to compare both approaches. Section 5 is dedicated to testing the numerical performance of the methods under consideration. We start here by giving some information on the implementation of the considered methods and especially introduce in brief the ACA. Then, we will solve the eigenvalue problem for covariance operators related to some of the Matérn kernels from Section 4 on different geometries. Finally, we sum up the results presented within this article in Section 6.

In the following, in order to avoid the repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that C can be bounded by a multiple of D , independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \approx D$ as $C \lesssim D$ and $C \gtrsim D$.

2. APPROXIMATION OF RANDOM FIELDS

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with σ -field $\mathcal{F} \subset 2^\Omega$ and probability measure \mathbb{P} . In order to ensure that $L^2_{\mathbb{P}}(\Omega)$ is separable, we also assume that Ω is a separable set. Furthermore, let $D \subset \mathbb{R}^d$ for $d = 2, 3$ be a sufficiently smooth and bounded domain.

For $p \geq 0$, the *Lebesgue-Bochner space* $L^2_{\mathbb{P}}(\Omega; H^p(D))$ consists of all maps $a: \Omega \rightarrow H^p(D)$ that satisfy

$$\|v\|_{L^2_{\mathbb{P}}(\Omega; H^p(D))} := \left(\int_{\Omega} \|v(\cdot, \omega)\|_{H^p(D)}^2 d\mathbb{P}(\omega) \right)^{1/2} < \infty. \quad (1)$$

In the sequel, it is convenient to identify $L_{\mathbb{P}}^2(\Omega; H^p(D))$ according to

$$L_{\mathbb{P}}^2(\Omega; H^p(D)) \cong H^p(D) \otimes L_{\mathbb{P}}^2(\Omega).$$

For further details on Lebesgue-Bochner spaces see e.g. [27].

2.1. Karhunen-Loève expansion

Let $a \in L_{\mathbb{P}}^2(\Omega; H^p(D))$ be a random field. We define the related *centered* random field $a_0(\mathbf{x}, \omega)$ via

$$a_0(\mathbf{x}, \omega) := a(\mathbf{x}, \omega) - \bar{a}(\mathbf{x}) := a(\mathbf{x}, \omega) - \int_{\Omega} a(\mathbf{x}, \omega) d\mathbb{P}(\omega) \quad (2)$$

and the corresponding Hilbert-Schmidt operators, i.e. $\mathcal{S}: L_{\mathbb{P}}^2(\Omega) \rightarrow H^p(D)$ with

$$(\mathcal{S}u)(\mathbf{x}) = \int_{\Omega} a_0(\mathbf{x}, \omega)u(\omega) d\mathbb{P}(\omega) \quad \text{for } u \in L_{\mathbb{P}}^2(\Omega)$$

and its adjoint $\mathcal{S}^*: \tilde{H}^{-p}(D) \rightarrow L_{\mathbb{P}}^2(\Omega)$ with

$$(\mathcal{S}^*u)(\omega) = \int_D a_0(\mathbf{y}, \omega)u(\mathbf{y}) d\mathbf{y} \quad \text{for } u \in \tilde{H}^{-p}(D).$$

Then, we especially find that $\mathcal{S}\mathcal{S}^*: \tilde{H}^{-p}(D) \rightarrow H^p(D)$ is given by

$$(\mathcal{S}\mathcal{S}^*u)(\mathbf{x}) = \int_D \int_{\Omega} a_0(\mathbf{x}, \omega)a_0(\mathbf{y}, \omega) d\mathbb{P}(\omega)u(\mathbf{y}) d\mathbf{y} =: (\mathcal{C}u)(\mathbf{x}) \quad (3)$$

which is the *covariance operator* related to the stochastic field a .

Obviously, the Hilbert-Schmidt norms of \mathcal{S} and \mathcal{S}^* correspond to the $L_{\mathbb{P}}^2(\Omega; L^2(D))$ -norm of a . The products of Hilbert-Schmidt operators form the *trace-class* of operators, cf. [28].

Definition 2.1. A bounded linear operator $\mathcal{A}: L^2(D) \rightarrow L^2(D)$ is of trace-class if $\text{Tr } \mathcal{A} := \sum_{m=1}^{\infty} ((\mathcal{A}^* \mathcal{A})^{\frac{1}{2}} \varphi_m, \varphi_m)_{L^2(D)} < \infty$ holds for an arbitrary orthonormal basis $\{\varphi_m\}_m$ in $L^2(D)$.

For the positive and symmetric operator \mathcal{C} , i.e. $(\mathcal{C}u, u)_{L^2(D)} = (\mathcal{S}^*u, \mathcal{S}^*u)_{L^2(\Omega)} \geq 0$, it holds $\text{Tr } \mathcal{C} = \sum_{m=1}^{\infty} (\mathcal{C}\varphi_m, \varphi_m)_{L^2(D)}$. Moreover, we derive

$$\begin{aligned} \text{Tr } \mathcal{C} &= \sum_{m=1}^{\infty} (\mathcal{C}\varphi_m, \varphi_m)_{L^2(D)} = \sum_{m=1}^{\infty} \|\mathcal{S}^* \varphi_m\|_{L_{\mathbb{P}}^2(\Omega)}^2 = \int_{\Omega} \sum_{m=1}^{\infty} (a_0(\cdot, \omega), \varphi_m)_{L^2(D)}^2 d\mathbb{P}(\omega) \\ &= \int_{\Omega} \|a_0(\cdot, \omega)\|_{L^2(D)}^2 d\mathbb{P}(\omega) = \|a_0\|_{L_{\mathbb{P}}^2(\Omega; L^2(D))}^2. \end{aligned} \quad (4)$$

Notice that for continuous *correlation kernels*, we have in particular

$$\text{Tr } \mathcal{C} = \int_D k(\mathbf{x}, \mathbf{x}) d\mathbf{x} \quad \text{with } k(\mathbf{x}, \mathbf{y}) := \int_{\Omega} a_0(\mathbf{x}, \omega)a_0(\mathbf{y}, \omega) d\mathbb{P}(\omega) \quad (5)$$

which is a consequence of Mercer's theorem, cf. [29]. For more details on trace-class operators, we refer the reader to [28, 30].

A very common representation of random fields in terms of the covariance operator's eigen-pairs $\{(\lambda_m, \varphi_m)\}_m$ is given by the *Karhunen-Loève expansion*.

Definition 2.2. Let $a \in L_{\mathbb{P}}^2(\Omega; L^2(D))$ be a random field. The representation

$$a(\mathbf{x}, \omega) = \bar{a}(\mathbf{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m} \varphi_m(\mathbf{x}) X_m(\omega) \quad (6)$$

is called Karhunen-Loève expansion with respect to a . Here, $\bar{a}(\mathbf{x})$ denotes the mean of a as defined in (2). The random variables $\{X_m\}_m$ are given according to

$$X_m(\omega) := \frac{1}{\sqrt{\lambda_m}} (\mathcal{S}^* \varphi_m)(\omega) = \frac{1}{\sqrt{\lambda_m}} \int_D a_0(\mathbf{x}, \omega) \varphi_m(\mathbf{x}) \, d\mathbf{x}$$

and satisfy $(X_m, X_n)_{L^2_{\mathbb{P}}(\Omega)} = \delta_{m,n}$ as well as $(X_m, 1)_{L^2_{\mathbb{P}}(\Omega)} = 0$.

The Karhunen-Loève expansion can be regarded as the continuous analogue to the singular value decomposition of matrices. For numerical issues, we have to truncate the series in (6) appropriately after $M \in \mathbb{N}$ terms. The question how small M can actually be chosen in order to achieve a certain precision is closely related to the decay of the eigenvalues of \mathcal{C} , which depends on the smoothness index p . Results on the decay of the eigenvalues have been established for periodic functions already in [31]. Nevertheless, since we do not want to restrict ourselves to this situation, we refer here to the more general results in [32, Theorem 3.3, Theorem 3.5].

Theorem 2.3. *Let $a \in L^2_{\mathbb{P}}(\Omega; H^p(D))$ with $p > d/2$. Then, the eigenvalues of the covariance operator $\mathcal{C}: \tilde{H}^{-p}(D) \rightarrow H^p(D)$ decay like $\lambda_m \lesssim m^{-2p/d}$ as $m \rightarrow \infty$ and it holds $\varepsilon(M) := \sum_{m=M+1}^{\infty} \lambda_m \lesssim M^{\frac{1}{2} - \frac{p}{d}}$.*

2.2. Finite element approximation

For the approximation of spatial functions in $L^2(D)$, we employ piece-wise polynomial (discontinuous) finite elements of order s . To that end, we introduce a family of quasi-uniform triangulations \mathcal{T}_h for D with mesh width h and define the spaces

$$V_h^s := \{v_h : D \rightarrow \mathbb{R} : v|_T \text{ is a polynomial of order } s \text{ for all } T \in \mathcal{T}_h\} \subset L^2(D). \quad (7)$$

Then, given a function $v \in H^t(D)$ with $0 \leq t \leq s$, it holds due to the Bramble-Hilbert lemma the approximation estimate

$$\|v - Q_h v\|_{L^2(D)} := \inf_{v_h \in V_h^s} \|v - v_h\|_{L^2(D)} \lesssim h^t \|v\|_{H^t(D)} \quad (8)$$

uniformly in h , see e.g. [33, 34]. For the spatial approximation of $a(\mathbf{x}, \omega)$, i.e.

$$a_h(\mathbf{x}, \omega) := (Q_h a)(\mathbf{x}, \omega) = (Q_h \bar{a})(\mathbf{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m} (Q_h \varphi_m)(\mathbf{x}) X_m(\omega),$$

we obtain in terms of the trace the following approximation result in V_h^s . Even though this result has already been derived in [11, Theorem 2.10], we shall present here a proof which employs another technique required for our considerations later on.

Theorem 2.4. *Let $N = \dim V_h^s$, $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ be the eigenvalues of the covariance operator \mathcal{C} and $\lambda_{1,h} \geq \lambda_{2,h} \geq \dots \geq \lambda_{N,h} \geq 0$ those of $\mathcal{C}_h := Q_h \mathcal{C} Q_h$. Then, it holds*

$$\|a_0 - Q_h a_0\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))}^2 = \text{Tr } \mathcal{C} - \text{Tr } \mathcal{C}_h$$

and therefore

$$\|a_0 - Q_h a_0\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))}^2 = \sum_{m=1}^N (\lambda_m - \lambda_{m,h}) + \sum_{m=N+1}^{\infty} \lambda_m.$$

Proof

Let $\{\varphi_m\}_m$ be an orthonormal basis of $L^2(D)$ such that either $\varphi_m \in \text{Im } Q_h$ or $\varphi_m \in \text{Im}(I - Q_h)$ holds. Therefore, we obtain $(\mathcal{S}^*(I - Q_h)\varphi_m, \mathcal{S}^*Q_h\varphi_m)_{L^2_{\mathbb{P}}(\Omega)} = 0$. Thus, we infer by (4) that

$$\begin{aligned} \|a_0 - Q_h a_0\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))}^2 &= \sum_{m=1}^{\infty} (\mathcal{S}^* \varphi_m, \mathcal{S}^* \varphi_m)_{L^2_{\mathbb{P}}(\Omega)} - (\mathcal{S}^* Q_h \varphi_m, \mathcal{S}^* Q_h \varphi_m)_{L^2_{\mathbb{P}}(\Omega)} \\ &= \text{Tr } \mathcal{C} - \text{Tr } \mathcal{C}_h. \end{aligned}$$

□

For the rest of this article, we refer to $\{(\lambda_m, \varphi_m)\}_m$ as the eigen-pairs of \mathcal{C} (in decreasing order) and to $\{(\lambda_{m,h}, \varphi_{m,h})\}_{m=1}^N$ as the eigen-pairs of \mathcal{C}_h . Moreover, it is convenient to set $\bar{a}_h(\mathbf{x}) := (Q_h \bar{a})(\mathbf{x})$ and $a_{0,h} := a_h - \bar{a}_h$. By the application of the theorem and the approximation estimate (8) it is straightforward to show the following

Corollary 2.5. *The trace error satisfies $0 \leq \text{Tr } \mathcal{C} - \text{Tr } \mathcal{C}_h \lesssim h^{2 \min\{s,p\}}$ if $a \in L_{\mathbb{P}}^2(\Omega; H^p(D))$, where the hidden constant involves the $L_{\mathbb{P}}^2(\Omega; H^p(D))$ -norm of a .*

Theorem 2.4 remains valid if we introduce the additional orthogonal projection $P_h: V_h^s \rightarrow U$ onto an M -dimensional subspace $U \subset V_h^s$. The related projected stochastic field is given by

$$a_{h,M} := \bar{a}_h + P_h a_{0,h} = Q_h \bar{a} + P_h Q_h a_0$$

and its covariance according to $\mathcal{C}_{h,M} := P_h \mathcal{C}_h P_h$. We arrive at the subsequent approximation result.

Theorem 2.6. *Let $\mathcal{C}_h = Q_h \mathcal{C} Q_h$, $\mathcal{C}_{h,M} = P_h \mathcal{C}_h P_h$ and $a_{h,M} = \bar{a}_h + P_h a_{0,h}$. Then, there holds*

$$\|a - a_{h,M}\|_{L_{\mathbb{P}}^2(\Omega; L^2(D))}^2 \lesssim h^{2 \min\{s,p\}} + (\text{Tr } \mathcal{C}_h - \text{Tr } \mathcal{C}_{h,M}), \quad (9)$$

where the hidden constant involves the $L_{\mathbb{P}}^2(\Omega; H^p(D))$ -norm of a .

Proof

By Theorem 2.4 and Corollary 2.5, it holds

$$\begin{aligned} \|a - a_{h,M}\|_{L_{\mathbb{P}}^2(\Omega; L^2(D))}^2 &\leq \|(I - Q_h)\bar{a}\|_{L^2(D)}^2 \\ &\quad + \|a_0 - a_{0,h}\|_{L_{\mathbb{P}}^2(\Omega; L^2(D))}^2 + \|a_{0,h} - P_h a_{0,h}\|_{L_{\mathbb{P}}^2(\Omega; L^2(D))}^2 \\ &\lesssim h^{2 \min\{s,p\}} + (\text{Tr } \mathcal{C} - \text{Tr } \mathcal{C}_h) + (\text{Tr } \mathcal{C}_h - \text{Tr } \mathcal{C}_{h,M}) \\ &\lesssim h^{2 \min\{s,p\}} + (\text{Tr } \mathcal{C}_h - \text{Tr } \mathcal{C}_{h,M}). \end{aligned}$$

□

The theorem indicates that, after fixing the ansatz space V_h^s , the approximation error of the stochastic field is controllable in terms of the discretized operators \mathcal{C}_h and $\mathcal{C}_{h,M}$. The optimal choice of P_h in terms of minimizing the trace error is the orthogonal projection onto the dominant invariant subspace of \mathcal{C}_h , i.e. $U_{M,h} := \text{span}\{\varphi_{1,h}, \dots, \varphi_{M,h}\}$ corresponding to the M dominant eigenvalues of \mathcal{C}_h . The related Karhunen-Loève expansion then reads

$$a_{h,M}(\mathbf{x}, \omega) = \bar{a}_h(\mathbf{x}) + \sum_{m=1}^M \sqrt{\lambda_{m,h}} \varphi_{m,h}(\mathbf{x}) X_m(\omega), \quad (10)$$

where the random variables are given according to

$$X_m(\omega) := \frac{1}{\sqrt{\lambda_{m,h}}} \int_D (P_h Q_h a_0)(\mathbf{x}, \omega) \varphi_{m,h}(\mathbf{x}) \, d\mathbf{x}. \quad (11)$$

Notice that, in this setting, the discretization of the stochastic field implies a change of the stochastic model induced by (11).

If, however, $U_{M,h}$ and thus P_h are not known exactly, they have to be approximated appropriately. This induces an additional error and we have to assume that $\lambda_{M+1,h}$ is distinct from $\lambda_{M,h}$, cf. [35, 36]. Nevertheless, any subspace is feasible as long as the difference $\text{Tr } \mathcal{C}_h - \text{Tr } \mathcal{C}_{h,M}$ becomes small for moderate sizes of M .

2.3. Approximation of eigenvalues

In the remainder of this section, we will briefly consider how well the decay of the eigenvalues $\{\lambda_m\}_m$ of \mathcal{C} is reflected by *Ritz-values* $\lambda_{1,h}, \dots, \lambda_{N,h}$ related to \mathcal{C}_h . The approximation error of the Ritz-values is measured in terms of the *gap* between the invariant subspace $U_m = \text{span}\{\varphi_1, \dots, \varphi_m\} \subset L^2(D)$ of \mathcal{C} and the approximation space V_h^s , i.e.

$$\theta(U_m) := \sup_{u \in U_m, \|u\|_{L^2(D)}=1} \|(I - Q_h)u\|_{L^2(D)}.$$

In order to control the gap, we have to take the eigenfunctions' regularity into account. By [32, Remark 3.8], it holds for $a \in H^p(D) \otimes L^2_{\mathbb{P}}(\Omega)$ that the eigenfunctions $\{\varphi_m\}_m$ of the covariance operator \mathcal{C} satisfy $\|\varphi_m\|_{H^t(D)} \lesssim \lambda_m^{-\frac{t}{2p}}$ for $0 \leq t \leq \min\{s, p\}$ and $\lambda_m \neq 0$. Thus, the orthogonal projection of φ_m onto V_h^s yields the error estimate

$$\|(I - Q_h)\varphi_m\|_{L^2(D)} \lesssim h^t \|\varphi_m\|_{H^t(D)} \lesssim \lambda_m^{-\frac{t}{2p}} h^t. \quad (12)$$

Now, we can estimate the gap as follows.

Lemma 2.7. *Let $U_m = \text{span}\{\varphi_1, \dots, \varphi_m\} \subset L^2(D)$ be an invariant subspace of \mathcal{C} . Then, it holds for $0 \leq t \leq \min\{s, p\}$ and $\lambda_1, \dots, \lambda_m \neq 0$ that*

$$\theta(U_m) \lesssim \sqrt{m} \lambda_m^{-\frac{t}{2p}} h^t. \quad (13)$$

Proof

Let $u = \sum_{i=1}^m \alpha_i \varphi_i$ with $\|u\|_{L^2(D)} = 1$. Thus, it holds $\sum_{i=1}^m \alpha_i^2 = 1$. Then, with $\alpha = (\alpha_1, \dots, \alpha_m)$, we have

$$\begin{aligned} \theta(U_m) &= \sup_{u \in U_m, \|u\|_{L^2(D)}=1} \|(I - Q_h)u\|_{L^2(D)} = \sup_{\|\alpha\|_2=1} \left\| \sum_{i=1}^m \alpha_i (I - Q_h)\varphi_i \right\|_{L^2(D)} \\ &\leq \sup_{\|\alpha\|_2=1} \sum_{i=1}^m |\alpha_i| \|(I - Q_h)\varphi_i\|_{L^2(D)} \lesssim \sup_{\|\alpha\|_2=1} \sum_{i=1}^m |\alpha_i| \lambda_i^{-\frac{t}{2p}} h^t \lesssim \sqrt{m} \lambda_m^{-\frac{t}{2p}} h^t, \end{aligned}$$

where we used (12) in the second to last step. \square

Remark 2.8. *In order to achieve convergence for the m -th eigenvalue, we have to guarantee $\theta(U_m) < 1$ which imposes a restriction to the mesh width h of the discretization. Moreover, we have $m \leq N = \dim V_h^s$.*

From [37, Theorem 9.2.2.2] we obtain finally a convergence result, which relates the eigenvalues' rate of approximation to the gap.

Theorem 2.9. *Let $U_m = \text{span}\{\varphi_1, \dots, \varphi_m\}$ be an invariant subspace of \mathcal{C} such that $\dim(Q_h U_m) = m$. Then, the approximation $\lambda_{i,h}$ to the i -th eigenvalue by the Rayleigh-Ritz method, i.e. $\mathcal{C}_h \varphi_{i,h} = \lambda_{i,h} \varphi_{i,h}$, satisfies the estimate*

$$0 \leq \lambda_i - \lambda_{i,h} \leq \lambda_i (\theta(U_i))^2 \quad \text{for all } 1 \leq i \leq m. \quad (14)$$

Thus, we can approximate the eigenvalues in V_h^s according to

$$0 \leq \lambda_i - \lambda_{i,h} \lesssim i \lambda_i^{\frac{p-t}{p}} h^{2t} \quad \text{for all } 1 \leq i \leq m \text{ and } 0 \leq t \leq \min\{s, p\}.$$

Especially for $s \ll p$, the eigenvalues of \mathcal{C}_h exhibit a similar rate of decay as the eigenvalues of \mathcal{C} up to a relative error of ih^{2s} .

3. THE PIVOTED CHOLESKY DECOMPOSITION

Based on the observation in Theorem 2.6 and the subsequent discussion, we consider in this section a more general approach for the representation of a random field. For this purpose, we rather refer here to the separable decomposition

$$a_{h,M}(\mathbf{x}, \omega) = \bar{a}_h(\mathbf{x}) + \sum_{m=1}^M \psi_{m,h}(\mathbf{x}) Y_m(\omega) \quad (15)$$

than to the orthogonal decomposition (10). In the expansion (15), we assume that $\{\psi_{m,h}\}_{m=1}^M \subset V_h^s$ with appropriately modified random variables $\{Y_m\}_{m=1}^M$.

Algorithm 1: Pivoted Cholesky decomposition ([21])

Data: matrix $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{N \times N}$ and error tolerance $\varepsilon > 0$

Result: low-rank approximation $\mathbf{A}_M = \sum_{i=1}^M \ell_i \ell_i^T$ such that $\text{trace}(\mathbf{A} - \mathbf{A}_M) \leq \varepsilon$

begin

 set $M := 1$;

 set $\mathbf{d} := \text{diag}(\mathbf{A})$ and $\text{error} := \|\mathbf{d}\|_{\ell^1}$;

 initialize $\boldsymbol{\pi} := [1, 2, \dots, N]$;

while $\text{error} > \varepsilon$ **do**

 set $i := \arg \max\{d_{\pi_j} : j = M, M+1, \dots, N\}$;

 swap π_M and π_i ;

 set $\ell_{M,\pi_M} := \sqrt{d_{\pi_M}}$;

for $M+1 \leq i \leq N$ **do**

 compute $\ell_{M,\pi_i} := \left(a_{\pi_M, \pi_i} - \sum_{j=1}^{M-1} \ell_{j,\pi_M} \ell_{j,\pi_i} \right) / \ell_{M,\pi_M}$;

 update $d_{\pi_i} := d_{\pi_i} - \ell_{M,\pi_M} \ell_{M,\pi_i}$;

 compute $\text{error} := \sum_{i=M+1}^n d_{\pi_i}$;

 increase $M := M + 1$;

One possibility to obtain a separable expansion (15) is to compute the pivoted Cholesky decomposition of the coefficient matrix of \mathcal{C}_h with respect to a basis in V_h^s . To that end, let $\boldsymbol{\Phi}(\mathbf{x}) := [\phi_1(\mathbf{x}), \dots, \phi_N(\mathbf{x})]$ denote an orthonormal basis of V_h^s . Then, the coefficient matrix of \mathcal{C}_h with respect to $\boldsymbol{\Phi}$ is given by

$$\mathbf{C} = [(\mathcal{C}_h \phi_j, \phi_i)_{L^2(D)}]_{i,j=1}^N \in \mathbb{R}^{N \times N}. \quad (16)$$

For each finite dimensional ansatz space, the matrix \mathbf{C} is symmetric and positive semidefinite. Thus, \mathbf{C} exhibits a (possibly pivoted) Cholesky decomposition. By pivoting the Cholesky decomposition (see Algorithm 1), we achieve numerical stability on the one hand, cf. [38, 39], and, if the eigenvalues of \mathbf{C} decay sufficiently fast, a low-rank approximation on the other hand, cf. [21]. Especially, the approximation error of the (truncated) pivoted Cholesky decomposition is a-posteriori controllable in terms of the (discrete) trace, i.e. $\text{trace}(\mathbf{A}) := \sum_{i=1}^N \langle \mathbf{A} \mathbf{e}_i, \mathbf{e}_i \rangle = \sum_{i=1}^N a_{i,i}$, where $\mathbf{e}_1, \dots, \mathbf{e}_N$ denotes the canonical basis in \mathbb{R}^N .

3.1. Separable expansions of random fields

In the following, we establish the connection between the approximation to the random field obtained by the pivoted Cholesky decomposition and the Karhunen-Loève expansion of $Q_h a(\mathbf{x}, \omega)$.

We denote the spectral decompositions related to \mathbf{C} by $\mathbf{C} = \sum_{i=1}^N \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ with $\lambda_i \in [0, \infty)$ and $\mathbf{v}_i \in \mathbb{R}^N$. Therefore, the Karhunen-Loève expansion of $Q_h a(\mathbf{x}, \omega)$ is given by

$$a_h(\mathbf{x}, \omega) = \Phi(\mathbf{x}) \bar{\mathbf{a}}_h + \sum_{i=1}^N \sqrt{\lambda_i} \Phi(\mathbf{x}) \mathbf{v}_i X_i(\omega)$$

with respect to the orthonormal basis Φ of V_h^s and $\bar{\mathbf{a}}_h := [(\bar{a}, \phi_i)_{L^2(D)}]_{i=1}^N \in \mathbb{R}^n$. Moreover, the random variables $\{X_i\}_{i=1}^N$ are given by (11), where $P_h = I$ is chosen as the identity mapping on V_h^s . This representation can be rewritten in matrix notation as

$$a_h(\mathbf{x}, \omega) - \Phi(\mathbf{x}) \bar{\mathbf{a}}_h =: \Phi(\mathbf{x}) \mathbf{V} \Sigma \mathbf{X}(\omega) \quad (17)$$

with $\mathbf{V} := [\mathbf{v}_1, \dots, \mathbf{v}_N]$, $\Sigma := \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_N})$ and $\mathbf{X}(\omega) := [X_1(\omega), \dots, X_N(\omega)]^\top$. The matrix $(\mathbf{V} \Sigma)^\top \in \mathbb{R}^{N \times N}$ from (17) exhibits a QR-decomposition: $\mathbf{Q} \mathbf{L}^\top = (\mathbf{V} \Sigma)^\top$ or $\mathbf{L} \mathbf{Q}^\top = \mathbf{V} \Sigma$, respectively. Here, \mathbf{Q} denotes an orthogonal matrix, i.e. $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I} \in \mathbb{R}^{N \times N}$, and $\mathbf{L} \in \mathbb{R}^{N \times N}$ is a lower triangular matrix. We shall next define the transformed random vector $\mathbf{Y}(\omega) := \mathbf{Q}^\top \mathbf{X}(\omega)$. Then, $\mathbf{Y}(\omega)$ also consists of N uncorrelated and centered random variables, since it holds

$$\int_{\Omega} \mathbf{Y}(\omega) \mathbf{Y}^\top(\omega) d\mathbb{P}(\omega) = \mathbf{Q}^\top \int_{\Omega} \mathbf{X}(\omega) \mathbf{X}^\top(\omega) d\mathbb{P}(\omega) \mathbf{Q} = \mathbf{Q}^\top \mathbf{I} \mathbf{Q} = \mathbf{I}.$$

That the random variables $Y_i(\omega)$ are also centered, follows from the fact that they are weighted sums of centered random variables. Thus, we obtain a representation equivalent to (17) according to

$$\Phi(\mathbf{x}) \mathbf{V} \Sigma \mathbf{X}(\omega) = \Phi(\mathbf{x}) \mathbf{L} \mathbf{Y}(\omega)$$

where the change of basis \mathbf{Q} only acts on the basis of $L_{\mathbb{P}}^2(\Omega)$. Moreover, we observe

$$\mathbf{C} = \mathbf{V} \Sigma (\mathbf{V} \Sigma)^\top = \mathbf{L} \mathbf{Q}^\top \mathbf{Q} \mathbf{L}^\top = \mathbf{L} \mathbf{L}^\top.$$

Since \mathbf{L} is a lower triangular matrix, we thus end up with the Cholesky decomposition of \mathbf{C} . In the following, without loss of generality, we will especially assume, that $\mathbf{L} \mathbf{L}^\top$ corresponds to the *pivoted* Cholesky decomposition of \mathbf{C} .

Using the Cholesky decomposition of \mathbf{C} , we obtain the separable representation

$$a_h(\mathbf{x}, \omega) = \Phi(\mathbf{x}) \bar{\mathbf{a}}_h + \sum_{i=1}^N \Phi(\mathbf{x}) \ell_i Y_i(\omega) \quad (18)$$

for $a_h(\mathbf{x}, \omega)$ with $\mathbf{L} = [\ell_1, \dots, \ell_N]$. Whereas, the related truncated Cholesky decomposition leads to the truncated expansion

$$a_{h,M}(\mathbf{x}, \omega) = \Phi(\mathbf{x}) \bar{\mathbf{a}}_h + \sum_{i=1}^M \Phi(\mathbf{x}) \ell_i Y_i(\omega).$$

Notice that this is exactly the representation (15) with $\psi_i(\mathbf{x}) = \Phi(\mathbf{x}) \ell_i$.

Remark 3.1. *The separable representation (18) of the stochastic field is based on the knowledge of an appropriate matrix $\mathbf{R} \in \mathbb{R}^{N \times N}$, a square root of the coefficient matrix, such that $\mathbf{C} = \mathbf{R} \mathbf{R}^\top$. It is known that for two different square roots, i.e. $\mathbf{C} = \mathbf{R} \mathbf{R}^\top = \tilde{\mathbf{R}} \tilde{\mathbf{R}}^\top$, there exists an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{N \times N}$ such that $\tilde{\mathbf{R}} = \mathbf{R} \mathbf{Q}^\top$. The change of the representation (17) due to the application of \mathbf{Q} is then performed by the change of the basis functions in $L_{\mathbb{P}}^2(\Omega)$, i.e. $\mathbf{Y}(\omega) := \mathbf{Q}^\top \mathbf{X}(\omega)$. Thus, any square root of \mathbf{C} yields a separable representation of $a_h(\mathbf{x}, \omega)$. Nevertheless, we focus on the pivoted Cholesky decomposition here.*

3.2. Approximation by the pivoted Cholesky decomposition

The approximation error of a given stochastic field induced by truncating of the pivoted Cholesky decomposition is controllable in accordance with the following theorem.

Theorem 3.2. *For the coefficient matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ given by (16), let $\mathbf{C}_M = \mathbf{L}_M \mathbf{L}_M^\top \in \mathbb{R}^{N \times N}$ denote its (truncated) pivoted Cholesky decomposition computed by Algorithm 1 such that $\text{trace}(\mathbf{C} - \mathbf{C}_M) < \varepsilon$ holds for some $\varepsilon > 0$. Then, for the related stochastic fields, we have the error estimate $\|a_h - a_{h,M}\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))} < \sqrt{\varepsilon}$.*

Proof

Let $\mathbf{C} = \mathbf{L}\mathbf{L}^\top$ denote the pivoted Cholesky decomposition of \mathbf{C} . We define

$$(\mathcal{S}_h^* u)(\omega) := \int_D (\Phi(\mathbf{x})\mathbf{L}\mathbf{Y}(\omega))^\top u(\mathbf{x}) \, d\mathbf{x}$$

and $P_h: V_h^s \rightarrow \text{span}\{\phi_1(\mathbf{x}), \dots, \phi_M(\mathbf{x})\}$ the orthogonal projection onto the space spanned by the first M basis functions. Then, it holds in complete analogy to the proof of Theorem 2.4 that

$$\begin{aligned} \|a_h - a_{h,M}\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))}^2 &= \sum_{i=1}^N (\mathcal{S}_h^* \phi_i, \mathcal{S}_h^* \phi_i)_{L^2_{\mathbb{P}}(\Omega)} - (\mathcal{S}_h^* P_h \phi_i, \mathcal{S}_h^* P_h \phi_i)_{L^2_{\mathbb{P}}(\Omega)} \\ &= \sum_{i=1}^N \langle \mathbf{L}^\top \mathbf{e}_i, \mathbf{L}^\top \mathbf{e}_i \rangle - \sum_{i=1}^M \langle \mathbf{L}^\top \mathbf{e}_i, \mathbf{L}^\top \mathbf{e}_i \rangle = \text{trace}(\mathbf{C} - \mathbf{C}_M) < \varepsilon. \end{aligned}$$

□

The theorem states that the choice $\varepsilon \approx h^{\min\{p,s\}}$ in the pivoted Cholesky decomposition guarantees, together with inequality (9), the (optimal) error estimate

$$\|a - a_{h,M}\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))} \lesssim h^{\min\{p,s\}}.$$

The major advantage of this approach is, that at no time the coefficient matrix \mathbf{C} has to be fully assembled. It is sufficient to provide access to single entries of this matrix while processing the pivoted Cholesky decomposition. The error in the approximation of the random field a is then a-posteriori controllable by the trace.

Given that the pivoted Cholesky decomposition for \mathbf{C} truncates with $M \ll N$ terms and $\mathbf{C}_M = \mathbf{L}_M \mathbf{L}_M^\top \in \mathbb{R}^{N \times N}$, where $\mathbf{L}_M \in \mathbb{R}^{N \times M}$, the computation of the related Karhunen-Loève expansion is performed with complexity $\mathcal{O}(M^2 N)$, cf. [21]. This can be achieved by computing the eigenvalues of $\mathbf{L}_M^\top \mathbf{L}_M \in \mathbb{R}^{M \times M}$ which coincide with those of \mathbf{C}_M . Then, if $\mathbf{v}_1, \dots, \mathbf{v}_M$ denote the orthonormal vectors of the small eigen-problem, the eigenvectors of \mathbf{C}_M are given by $\mathbf{L}\mathbf{v}_1, \dots, \mathbf{L}\mathbf{v}_M$ and we have

$$(\mathbf{L}\mathbf{v}_i)^\top (\mathbf{L}\mathbf{v}_j) = \mathbf{v}_i^\top \mathbf{L}^\top \mathbf{L} \mathbf{v}_j = \lambda_i \delta_{i,j} \quad \text{for all } i, j = 1, \dots, M. \quad (19)$$

Thus, the related Karhunen-Loève decomposition is given by

$$a_{h,M}(\mathbf{x}, \omega) = \bar{a}_h(\mathbf{x}) + \sum_{i=1}^M \Phi(\mathbf{x}) \mathbf{L}\mathbf{v}_i \tilde{X}_i(\omega). \quad (20)$$

If the laws of the random variables $X_i(\omega)$ are known, we obtain the relation

$$\tilde{\mathbf{X}}(\omega) = [\mathbf{v}_1, \dots, \mathbf{v}_M]^\top \mathbf{Y}(\omega)$$

with \mathbf{v}_i from (20). Otherwise, the related random variables $\tilde{X}_i(\omega)$ can be determined by a *maximum likelihood* estimate, cf. [11]. Notice that for the important Gaussian case $\mathbf{X}(\omega) \sim [\mathcal{N}(0, 1)]^N$, we have for any orthogonal transform $\mathbf{Q}\mathbf{X}(\omega) \sim [\mathcal{N}(0, 1)]^N$ and thus $\tilde{\mathbf{X}}$ has the same law as \mathbf{X} .

4. THE MATÉRN CLASS OF KERNELS

For our numerical tests, we consider a special subset of Hilbert-Schmidt kernels, namely the *Matérn class* of kernel functions, cf. [40]. They are very often used as covariance kernels for the definition of stochastic fields. In accordance with [2], they are defined as follows.

Definition 4.1. Let $r := \|\mathbf{x} - \mathbf{y}\|_2$ and $\ell \in (0, \infty)$. Then, the Matérn covariance function of order $\nu > 0$ is given by

$$k_\nu(r) := \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}r}{\ell} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu}r}{\ell} \right). \quad (21)$$

Here, Γ denotes the gamma function and K_ν denotes the modified Bessel function of the second kind of order ν , cf. [41].

The expression (21) simplifies if $\nu = p + 1/2$ with $p \in \mathbb{N}$. In this case, [2] provides

$$k_{p+1/2}(r) = \exp \left(-\frac{\sqrt{2\nu}r}{\ell} \right) \frac{p!}{(2p)!} \sum_{i=0}^p \frac{(p+i)!}{i!(p-i)!} \left(\frac{\sqrt{8\nu}r}{\ell} \right)^{p-i}.$$

Especially, we have

$$\begin{aligned} \nu = \frac{1}{2}, \quad k_{1/2}(r) &= \exp \left(-\frac{r}{\ell} \right), \\ \nu = \frac{3}{2}, \quad k_{3/2}(r) &= \left(1 + \frac{\sqrt{3}r}{\ell} \right) \exp \left(-\frac{\sqrt{3}r}{\ell} \right), \\ \nu = \frac{5}{2}, \quad k_{5/2}(r) &= \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2} \right) \exp \left(-\frac{\sqrt{5}r}{\ell} \right), \\ \nu = \frac{7}{2}, \quad k_{7/2}(r) &= \left(1 + \frac{\sqrt{7}r}{\ell} + \frac{14r^2}{5\ell^2} + \frac{49\sqrt{7}r^3}{15\ell^3} \right) \exp \left(-\frac{\sqrt{7}r}{\ell} \right), \\ \nu = \frac{9}{2}, \quad k_{9/2}(r) &= \left(1 + \frac{3r}{\ell} + \frac{27r^2}{7\ell^2} + \frac{18r^3}{7\ell^3} + \frac{27r^4}{35\ell^3} \right) \exp \left(-\frac{3r}{\ell} \right), \\ \nu = \infty, \quad k_\infty(r) &= \exp \left(-\frac{r^2}{2\ell^2} \right). \end{aligned} \quad (22)$$

A visualization of this kernels for different values of ν is given in Figure 1. Obviously, the Sobolev smoothness of the kernel k_ν is controlled by the *smoothness parameter* ν .

For increasing values of ν , the respective kernel function k_ν exhibits successively more regularity. Especially, the eigenvalues of the Matérn correlation kernels decay like

$$\lambda_m \leq Cm^{-(1+\frac{2\nu}{d})} \quad (23)$$

for some $C > 0$, cf. [42]. Thus, since the decay of the covariance operator's eigenvalues is known in advance, they are very well suited for numerical examples.

Obviously, the Matérn kernels provide rotational symmetry, i.e. they are invariant under isometries of D , since they are only dependent on the particular distance of the points \mathbf{x} and \mathbf{y} . Thus, we obtain analytic expressions for the eigenvalues of the underlying Hilbert-Schmidt operators, if we choose $D = \mathbb{S}^{d-1}$ to be the unit sphere in \mathbb{R}^d . More precisely, we may apply the Funk-Hecke formula, cf. [43], which reads as follows.

Theorem 4.2. Let $\mathbf{x} \in \mathbb{S}^{d-1}$ and $f \in C([-1, 1])$, then it holds

$$\int_{\mathbb{S}^{d-1}} f(\mathbf{x}^\top \mathbf{y}) Y_m(\mathbf{y}) d\sigma_{\mathbf{y}} = \lambda_m Y_m(\mathbf{x})$$

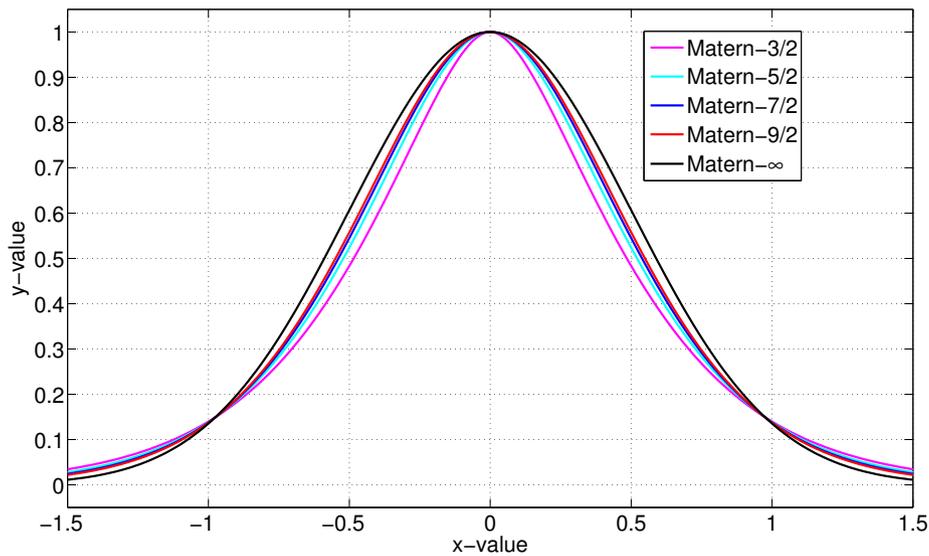


Figure 1. Different values for the smoothness parameter ν .

with

$$\lambda_m = |\mathbb{S}^{d-2}| \int_{-1}^1 P_m(d; t) f(t) (1-t^2)^{\frac{d-3}{2}} dt.$$

Here, Y_m corresponds to a spherical harmonic of order m and $P_m(d; t)$ denotes the polynomial

$$P_m(d; t) := m! \Gamma\left(\frac{d-1}{2}\right) \sum_{i=0}^{\lfloor m/2 \rfloor} \left(\frac{-1}{4}\right)^i \frac{(1-t^2)^i t^{m-2i}}{i!(m-2i)! \Gamma\left(i + \frac{d-1}{2}\right)}.$$

A proof of this theorem can be found in [43]. Especially, for the case $d = 3$, the polynomials $P_m(3; t)$, correspond to the Legendre polynomials, cf. [43].

Notice, that the Funk-Hecke formula applies to all kernel functions on \mathbb{S}^{d-1} , which depend only on the Euclidean distance $r(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$. This is easily seen due to $r(\mathbf{x}, \mathbf{y}) = r(\mathbf{x}^\top \mathbf{y}) = \sqrt{2 - 2\mathbf{x}^\top \mathbf{y}}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^{d-1}$. Figure 2 shows the distribution of the Matérn-kernels' eigenvalues for $\nu = 3/2, 5/2, 7/2, 9/2$ on \mathbb{S}^2 up to an order of magnitude of 10^{-10} for the correlation length $\ell = 1$. The constant C is estimated by a least-square fit for the ratio of the rate given by formula (23) for $C = 1$ and the exact eigenvalues given by Theorem 4.2. The obtained values of C for each kernel under consideration are denoted in the legend of Figure 2. The plot indicates, that the fitted rates perfectly match the asymptotic behavior of the eigenvalues.

5. NUMERICAL RESULTS

The numerical tests in this section are performed on parametric surfaces $\Gamma \subset \mathbb{R}^3$. These surfaces have recently been considered in the context of solving boundary integral equations, cf. [44] and the references therein. The following subsection provides some details on the numerical implementation of the considered methods.

5.1. Implementation

For the sake of completeness, we start by briefly recalling the idea of ACA. For more details, we refer to [44] and the references therein.

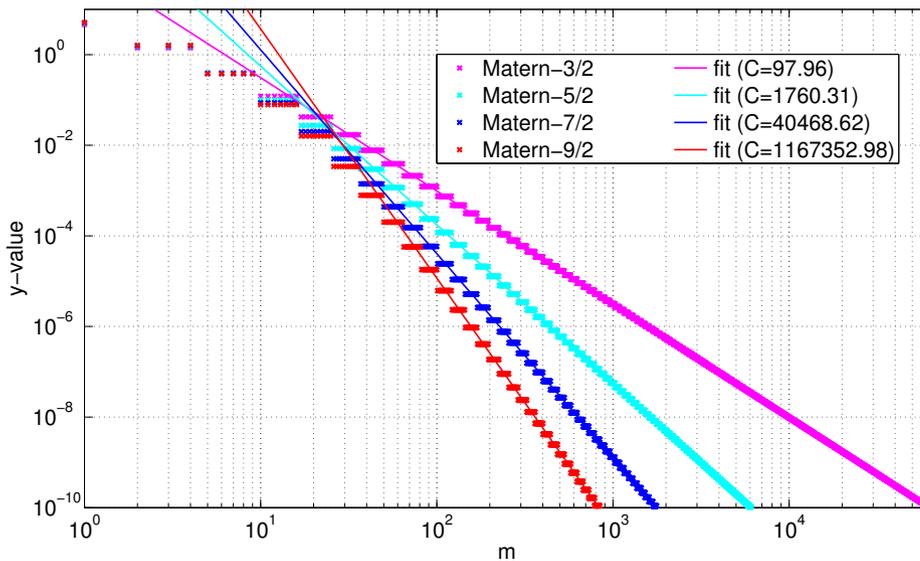


Figure 2. Decay of the eigenvalues with related fits.

For the application of ACA, we exploit that the correlation kernel $k(\mathbf{x}, \mathbf{y})$ of the covariance operator $C: \dot{H}^{-p}(D) \rightarrow H^p(D)$ is *asymptotically smooth*, i.e. the kernel is smooth away from the diagonal $\{(\mathbf{x}, \mathbf{y}) \in D \times D : \mathbf{x} = \mathbf{y}\}$. Then, the coefficient matrix C from (16) can be partitioned into *admissible*, i.e. compressible, matrix blocks corresponding to the *far-field* of C and *non-admissible*, i.e. non-compressible, matrix blocks corresponding to the *near-field* of C . ACA provides a means to compress the admissible matrix blocks of C . A visualization of the block partitioning and the related ranks of C in case of the Matérn-9/2 kernel on \mathbb{S}^2 is depicted for level 4 in Figure 3. In each admissible matrix block, we approximate C by a truncated, partially pivoted Gaussian

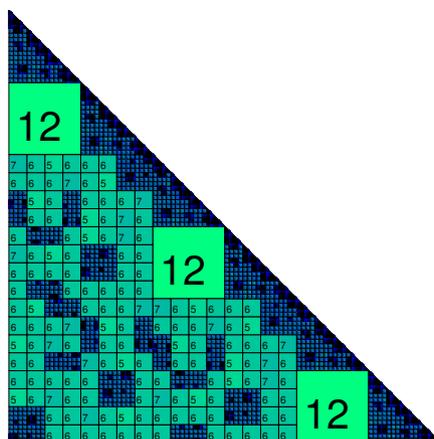


Figure 3. Partition of the matrix C for the Matérn-9/2 kernel with inscribed ranks.

elimination, cf. [16]. To this end, we define the vectors $\ell_m, \mathbf{u}_m \in \mathbb{R}^n$ by the following iterative

scheme, where $\mathbf{C}_{\text{adm}} = [c_{i,j}]_{i,j=1}^n$ is the admissible matrix block under consideration:

$$\text{for } m = 1, 2, \dots \text{ set } \mathbf{u}_m = \hat{\mathbf{u}}_m / [\hat{\mathbf{u}}_m]_{j_m} \text{ with}$$

$$\hat{\mathbf{u}}_m = [c_{i_m,j}]_{j=1}^n - \sum_{j=1}^{m-1} [\ell_j]_{i_m} \mathbf{u}_j \text{ and } \ell_m = [c_{i,j_m}]_{i=1}^n - \sum_{i=1}^{m-1} [u_i]_{j_m} \ell_i. \quad (24)$$

A criterion to guarantee the convergence of the algorithm is to choose the pivot element located in (i_m, j_m) -position as the maximum element in modulus of the remainder $\mathbf{C}_{\text{adm}} - \mathbf{L}_{m-1} \mathbf{U}_{m-1}$, where we define the matrices $\mathbf{L}_{m-1} := [\ell_1, \dots, \ell_{m-1}]$ and $\mathbf{U}_{m-1} := [\mathbf{u}_1 \dots, \mathbf{u}_{m-1}]^\top$. This would require the assembly of the whole matrix block \mathbf{C}_{adm} , which is not feasible in practice. Therefore, we employ another pivoting strategy: We choose j_m such that $[\hat{\mathbf{u}}_m]_{j_m}$ is the largest element in modulus of the row $\hat{\mathbf{u}}_m$.

We finally stop the iteration if the criterion $\|\ell_{m+1}\|_2 \|\mathbf{u}_{m+1}\|_2 \leq \varepsilon \|\mathbf{L}_m \mathbf{U}_m\|_F$ is met for some desired accuracy $\varepsilon > 0$. Here and in the following, $\|\cdot\|_F$ denotes the Frobenius norm. Notice that the stopping criterion supposes a *saturation assumption*, i.e. the assumption that the error is reduced in each step by a constant factor. Compressing each block with the prescribed accuracy ε yields the overall error estimate $\|\mathbf{C} - \tilde{\mathbf{C}}\|_F \lesssim \varepsilon \|\mathbf{C}\|_F$, where $\tilde{\mathbf{C}}$ denotes the compressed matrix.

Obviously, the complexity for the computation of the rank- m -approximation $\mathbf{L}_m \mathbf{U}_m$ to the block \mathbf{C}_{adm} is $\mathcal{O}(2m^2n)$ and the storage cost is $\mathcal{O}(2mn)$. The latter one can be further reduced by the application of a singular value decomposition and neglecting non-relevant singular values.

Notice that (24) combined with total pivoting would result for symmetric matrix blocks in an algorithm quite similar to the pivoted Cholesky decomposition. Nevertheless, for PCD, we do not have to partition the system matrix into far- and near-field, but directly employ Algorithm 1 to \mathbf{C} . In this sense, we may think of PCD as a single-block ACA with total pivoting. Here, the total pivoting is not prohibitive expensive since it is a-priori known that the pivots are located along the main diagonal of \mathbf{C} . Furthermore, we have in contrast to ACA a rigorous stopping criterion based on the quantity $\text{trace}(\mathbf{C} - \mathbf{L}_m \mathbf{L}_m^\top)$.

The implementations of both ACA and PCD rely on the same single-scale code, which means, they use the same quadrature routines for the assembly of the entries of the matrix \mathbf{C} . In case of ACA, we end up with a data-sparse representation $\tilde{\mathbf{C}}$ of \mathbf{C} . Thus, to obtain a representation of the approximate stochastic field $a_{h,M}$ similar to (10), we have still to compute the dominant eigen-pairs of $\tilde{\mathbf{C}}$. The representation of $\tilde{\mathbf{C}}$ yields a fast matrix-vector product. Therefore, we employ ARPACK, cf. [25], to solve the eigen-problem for the compressed matrix $\tilde{\mathbf{C}}$. The size of the Krylov subspace in ARPACK is chosen twice the number of desired eigenvalues, which is a reasonable choice according to [25]. Both methods have been implemented in the C-programming language, cf. [45]. Furthermore, in the implementation of ACA, we have employed level 1 and 2 BLAS[†] routines in the assembly of the matrix $\tilde{\mathbf{C}}$ and in the matrix-vector product, whenever possible. Additionally, exploiting the symmetry of \mathbf{C} , we only assembled the lower triangular part of the matrix $\tilde{\mathbf{C}}$ for ACA. Notice that a LAPACK[‡]-style implementation of the PCD, cf. [46], is not applicable since it relies on the assembly of the entire matrix \mathbf{C} , which is not feasible for large values of N [§].

5.2. Numerical examples

All computations are carried out single threaded on a computing server with two Intel(R) Xeon(R) X5550 CPUs with a clock rate of 2.67GHz and 48GB of main memory. Furthermore, we set the correlation length of the Matérn kernels to $\ell = 1$ in each example. For the spatial discretization we choose piece-wise constant finite elements, i.e. the ansatz space is V_h^1 , cf. (7). For ACA, we set the truncation error $\varepsilon \approx h^2$. Thus, we expect for both approaches a linear rate of convergence for the

[†]<http://www.netlib.org/blas/>

[‡]<http://www.netlib.org/lapack/>

[§]For example for $N = 10^5$, the storage of the matrix \mathbf{C} would require about 80GB of memory in 8-Byte double precision.

stochastic field in terms of the (continuous) traces, i.e.

$$\|a - a_{h,M}\|_{L^2_{\mathbb{P}}(\Omega; L^2(D))} \lesssim h + \sqrt{\text{Tr } \mathcal{C}_h - \text{Tr } \mathcal{C}_{h,M}}.$$

Therefore, we will measure the error by the quantity $\sqrt{\text{Tr } \mathcal{C}_h - \text{Tr } \mathcal{C}_{h,M}} / \sqrt{\text{Tr } \mathcal{C}_h}$. Notice that we consider here the relative error in order to make the error independent of the scaling of the eigenvalues which depends on the size of the domain, cf. (5).

First example

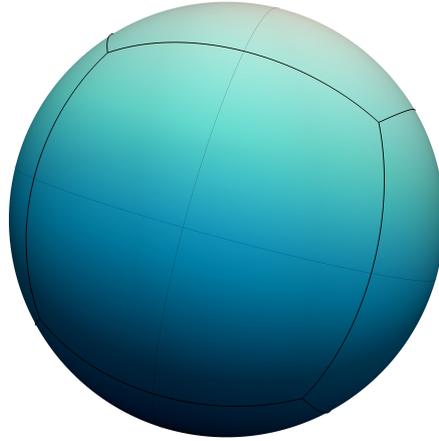


Figure 4. The unit sphere \mathbb{S}^2 represented by 6 patches.

As a benchmark, we consider the three-dimensional unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ represented by 6 congruent patches, see Figure 4. Thus, with the knowledge from the preceding Section 4, we can compute the exact eigenvalues and eigenfunctions of the Matérn covariance functions as reference. Furthermore, we can estimate the truncation error due to (23). For the truncation error related to the Matérn covariance with smoothness parameter ν , it holds that

$$\sqrt{\sum_{m=M}^{\infty} \lambda_m} \lesssim \sqrt{\int_M^{\infty} Cx^{-1-\nu} dx} = \sqrt{\frac{1}{\nu} CM^{-\nu}}. \quad (25)$$

Notice that the dimension is essentially $d = 2$ here, since we restrict the Matérn kernels to the unit sphere \mathbb{S}^2 . Thus, to bound the truncation error of the Karhunen-Loève expansion by h , we have to ensure that

$$\sqrt{\frac{1}{\nu} CM^{-\nu}} \leq h \quad \Rightarrow \quad M \geq \left(\frac{C}{\nu h^2} \right)^{\frac{1}{\nu}}.$$

With the estimation of the constant C at hand, cf. Figure 2, we could now compute the related length of the Karhunen-Loève expansion. Unfortunately, this approach yields very large numbers of eigen-pairs to be approximated by ACA. Therefore, we choose another approach. We consider for each respective kernel the sum of those eigenvalues with magnitude larger than 10^{-10} as an approximation to the actual trace of the kernel, i.e.

$$\int_{\mathbb{S}^2} k_{\nu}(\mathbf{x}, \mathbf{x}) ds_{\mathbf{x}} = \sum_{m=1}^{M_{\max}} \lambda_m + \varepsilon,$$

cf. (5), with $M_{\max} = \arg \min_m \{\lambda_m > 10^{-10}\}$. The resulting truncation error is computable due to the knowledge of the exact traces which are equal to 4π for every ν . We have $\varepsilon = 4.18 \cdot 10^{-6}$ for

j	$\nu = 3/2$	$\nu = 5/2$	$\nu = 7/2$	$\nu = 9/2$
1	6 (9)	4 (4)	4 (4)	4 (4)
2	18 (25)	13 (16)	11 (16)	9 (9)
3	48 (49)	25 (25)	20 (25)	17 (25)
4	120 (121)	45 (49)	33 (36)	26 (36)
5	305 (324)	79 (81)	49 (49)	40 (49)
6	768 (789)	139 (144)	76 (81)	57 (64)
7	1928 (1936)	243 (256)	113 (121)	78 (81)
8	4807 (4900)	423 (441)	166 (169)	107 (121)

Table I. Different values for the cut-off parameter M_j on the unit sphere \mathbb{S}^2 .

j	$\nu = 3/2$	$\nu = 5/2$	$\nu = 7/2$	$\nu = 9/2$
1	5 (6)	5 (6)	4 (5)	4 (5)
2	19 (21)	14 (14)	12 (13)	11 (12)
3	49 (56)	29 (32)	23 (24)	21 (22)
4	137 (158)	53 (58)	38 (41)	32 (35)
5	359 (414)	97 (107)	58 (62)	46 (49)
6	935 (1082)	167 (185)	89 (96)	64 (69)
7	2415 (2812)	295 (327)	132 (143)	90 (96)
8	— (7158)	513 (569)	197 (214)	122 (130)

Table II. Ranks determined by PCD on the unit sphere \mathbb{S}^2 .

$\nu = 3/2, \varepsilon = 2.43 \cdot 10^{-7}$ for $\nu = 5/2, \varepsilon = 4.93 \cdot 10^{-8}$ for $\nu = 7/2$, and $\varepsilon = 1.70 \cdot 10^{-8}$ for $\nu = 9/2$. The rank on each level j with mesh width $h \approx 2^{-j}$ is then determined according to

$$M_j = \arg \min_{k \in \{1, \dots, M_{\max}\}} \left\{ \sum_{m=1}^{M_{\max}} \lambda_m - \sum_{m=1}^k \lambda_m < h^2 \sum_{m=1}^{M_{\max}} \lambda_m \right\}, \quad (26)$$

that is the trace error relative to the scaling of the eigenvalues. The finest level j which we consider here is 8, resulting in 393216 finite elements. For the levels $j = 1, \dots, 8$ and $\nu = 3/2, 5/2, 7/2, 9/2$, the related cut-off parameters M_j are found in Table I. The number in the brackets denotes the size necessary to resolve clusters of eigenvalues by approximating only complete subspaces related to the multiplicity of the respective eigenvalue. This is proposed in [25] in order to achieve the optimal performance of ARPACK.

Table II shows the ranks determined by PCD. The numbers in front of the brackets correspond to the recompressed ranks, the numbers within the brackets denote the original rank. As it turns out, the ranks computed by PCD are rather optimal in the sense that they reflect the estimated length of the Karhunen-Loève expansion determined by formula (23). Especially for increasing smoothness of the kernel function, the determined rank gets successively better.

Remark 5.1. *We end up with the spectral decomposition of the approximate covariance $\mathcal{C}_{h,M}$ when we solve the eigen-problem (19) for PCD. By truncating the obtained decomposition (20) with the prescribed relative accuracy h^2 , we achieve an a-posteriori recompression of the PCD. This procedure may at most double the approximation error but reduces the rank by up to 10% on average in our computations for this article.*

The error plots and related computational times for the numerical experiments on the unit sphere are depicted in Figure 5 and in Figure 6, respectively. Unfortunately, the computations of ACA as well as PCD with recompression for $\nu = 3/2$ and level 8, i.e. for 393216 finite elements, could not be carried out since the available main memory has been insufficient.

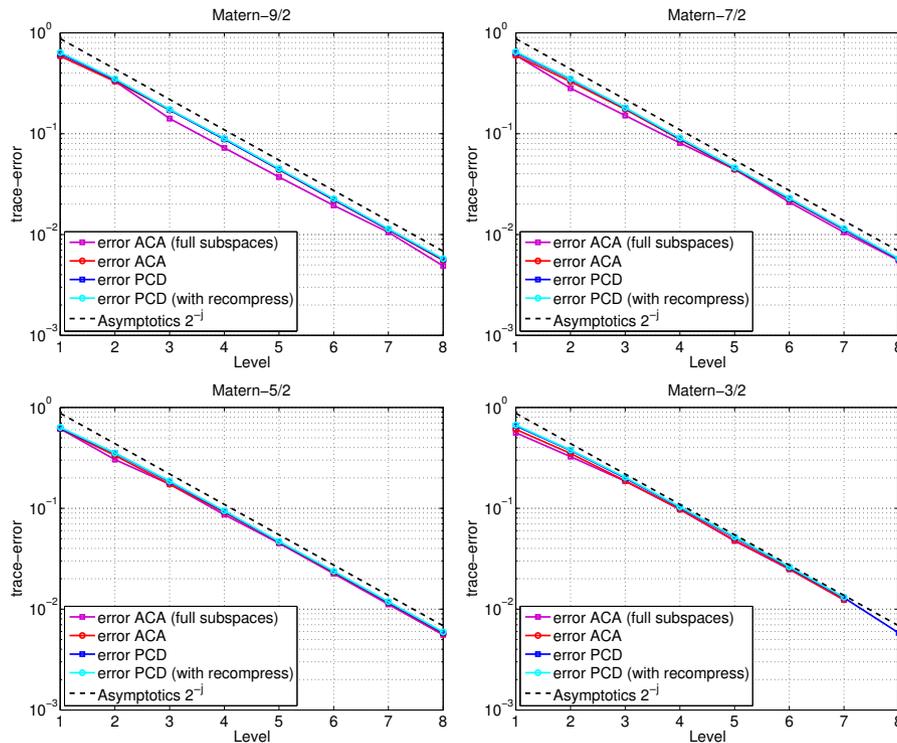


Figure 5. Numerical results (errors) on the unit sphere \mathbb{S}^2 .

Figure 5 shows the trace error for each particular kernel. The expected rate $h \approx 2^{-j}$ is indicated in the plots by the dashed black line. The magenta colored line with boxes shows the error for ACA with clusters of eigenvalues resolved (full subspaces), whereas the red line with circles shows the error for ACA with the exact number of eigenvalues computed by (26). The error of the PCD is indicated by the blue lines with squares and finally the error of PCD with recompression is indicated by the cyan colored line with circles. It turns out that all four methods provide the expected rate of convergence in this example. For overview purposes, we have chosen the same colors and markers for each particular method in the subsequent visualizations.

Figure 6 shows the computational times for every method and each particular kernel. There seems to be no significant difference in the times for ACA with clusters of eigenvalues resolved and ACA with the exact number of eigenvalues from (26) for all kernels under consideration. Nevertheless, we observe that PCD is about a factor of 10 times faster than ACA for the smoother kernels, i.e. $\nu = 5/2, 7/2, 9/2$ and about a factor of 2 times faster for $\nu = 3/2$. Furthermore, we observe that the computation time consumed by ACA for $\nu = 5/2, 7/2, 9/2$ is mostly caused by the assembly of the matrix \tilde{C} , indicated by the green colored line with diamonds, whereas the computation time is governed by the eigenvalue computation for $\nu = 3/2$.

Second example

In our second example, we consider the plate geometry shown in Figure 7. It is a rectangle with 30 inscribed, equi-spaced circular holes, which is represented by 120 patches and scaled to a size of 2×2.4 . Here, the computations are carried out on levels $j = 1, \dots, 6$, where level 6 corresponds to 491520 finite elements. Figure 8 contains a visualization of the four orthonormal eigenfunctions corresponding to the four largest eigenvalues of the Matérn kernel with $\nu = 3/2$.

In this example, we do not know the number of eigenvalues necessary to achieve the desired precision with ACA and ARPACK. Therefore, we use here the ranks provided by PCD with recompression as reference. The respective values are found in Table III. Again, the numbers in

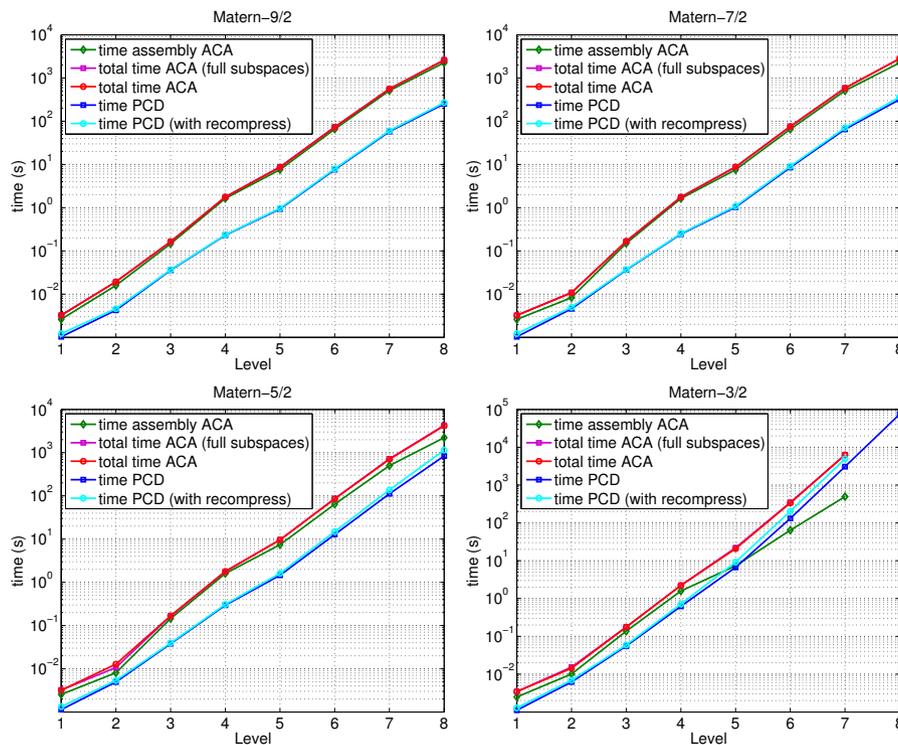


Figure 6. Numerical results (computational times) on the unit sphere \mathbb{S}^2 .

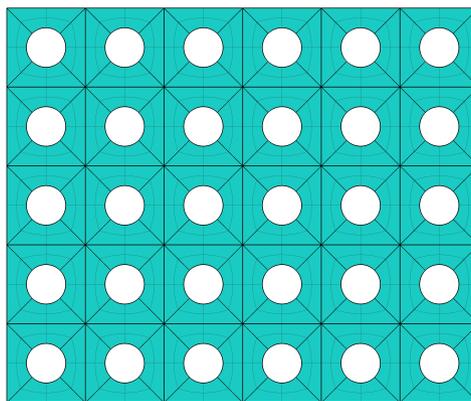


Figure 7. Plate geometry represented by 120 patches.

front of the brackets correspond to the recompressed ranks of PCD and the numbers within the brackets denote the original ranks.

The error plots and related computational times for the numerical experiments on the plate geometry are presented in Figure 9 and in Figure 10, respectively.

The trace error for each particular kernel, i.e. $\nu = 3/2, 5/2, 7/2, 9/2, \infty$, and the different methods is found in Figure 9. Again, PCD provides exactly the expected rate of convergence. The behavior of ACA is not that monotone as in the previous example. In case of the smoother kernels, i.e. $\nu = 7/2, 9/2, \infty$, the rate of convergence deteriorates in the last step. For $\nu = 3/2$ we have a contrary behavior. The rate of convergence is increased from level 2 to 3 and in the last step. Finally, we observe for $\nu = 5/2$ an increased rate of convergence from level 4 to 5 on the one hand and an increase of the error in the last step on the other hand. Possibly, these effects are caused by a lack of resolution of the faster oscillating eigenfunctions which are involved in the deflation process

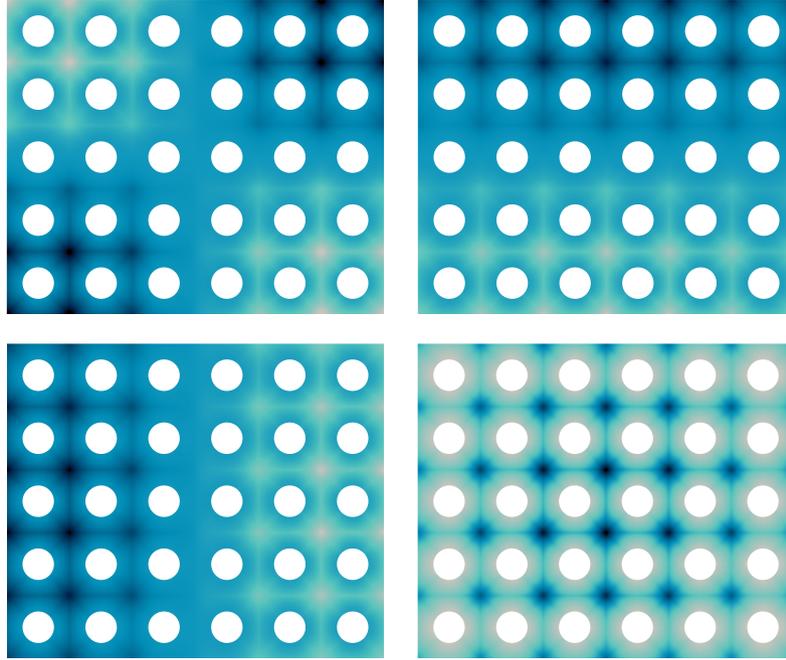


Figure 8. First four orthonormal eigenfunctions on the plate geometry and Matérn kernel for $\nu = 3/2$.

j	$\nu = 3/2$	$\nu = 5/2$	$\nu = 7/2$	$\nu = 9/2$	$\nu = \infty$
1	13 (14)	11 (11)	9 (9)	9 (9)	7 (7)
2	34 (36)	19 (20)	15 (15)	14 (14)	10 (10)
3	78 (86)	31 (33)	25 (26)	20 (21)	13 (13)
4	178 (196)	52 (56)	35 (37)	29 (30)	17 (17)
5	416 (459)	87 (93)	49 (52)	38 (39)	21 (21)
6	983(1085)	141 (151)	71 (75)	53 (55)	25 (26)

Table III. Ranks determined by PCD on the plate geometry.

of the implicit restarted Arnoldi method and the resulting impact on the computation of the sought eigenvalues.

Figure 10 shows the computational times for every method and each particular kernel. Here, the times for the recompression of PCD are rather moderate due to the low ranks. Nevertheless, the benefit of the recompression is relatively small here, especially for the smoother kernels, cf. Table III. Again, PCD outperforms ACA by about a factor of 10, even in the case of $\nu = 3/2$. This is due to the coupling of the number of eigenvalues to be computed to the ranks provided by PCD. Furthermore, as in the previous example, the computation time for ACA is governed by the assembly of the matrix for $\nu = 5/2, 7/2, 9/2, \infty$ and by the eigenvalue computation for $\nu = 3/2$.

6. CONCLUDING REMARKS

The present article is devoted to the efficient approximation of random fields for numerical applications. It is state of the art to compute a separable representation of the random field under consideration. A very common approach to determine such a representation is the (truncated) Karhunen-Loève expansion. Here, one has to solve the eigen-problem for the related covariance operator. We have tackled this task by combining the Adaptive Cross Approximation and ARPACK. Nevertheless, a major drawback of this approach is that the number of eigen-pairs to

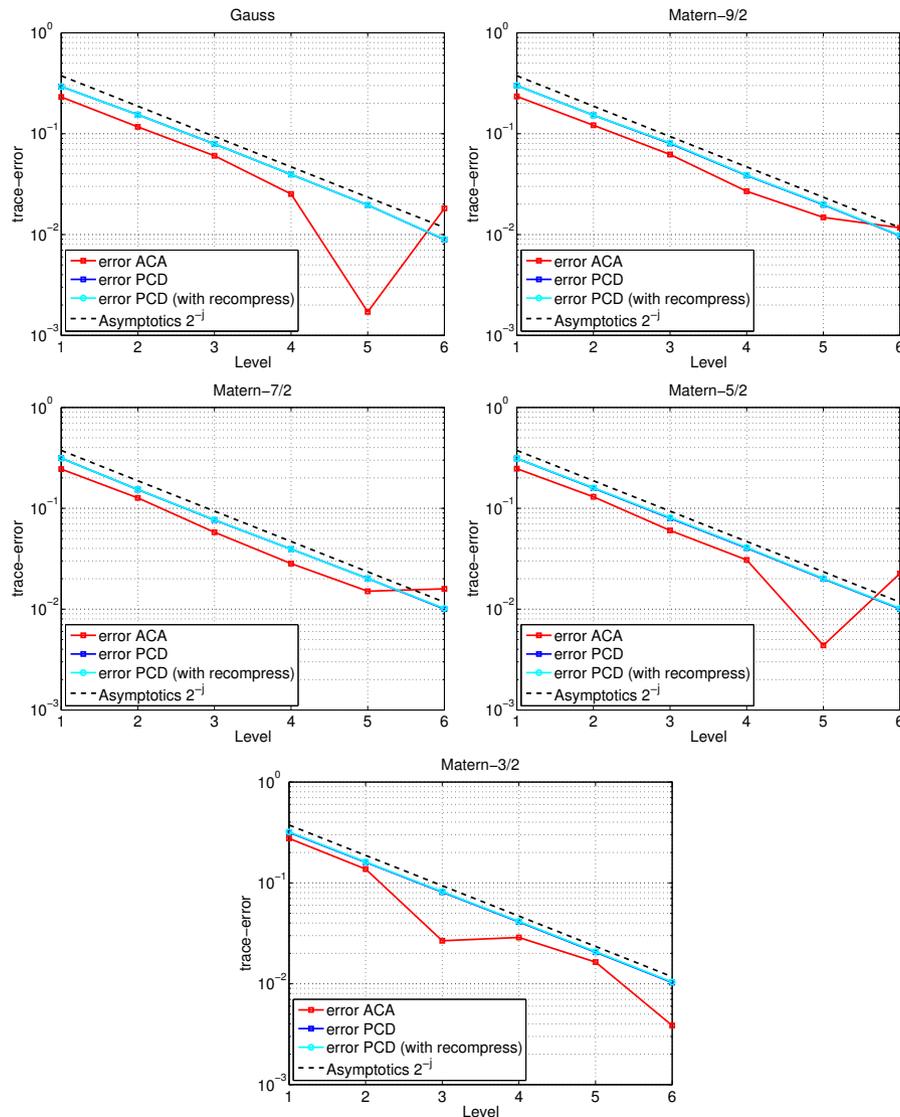


Figure 9. Numerical results (errors) on the plate geometry.

be approximated has to be known in advance. This might be a problem in practice since the correct number is not feasible in many applications. With the pivoted Cholesky decomposition, we provide a method which overcomes this obstruction. Due to the knowledge of the discretized covariance operator's main diagonal, we are able to a-posteriori control the approximation error in terms of the trace. If, for the application at hand, an orthogonal decomposition of the stochastic field is required, this can be realized relatively cheap by the pivoted Cholesky decomposition in a post-processing step. The numerical experiments suggest that both approaches provide the optimal rate of convergence. In the comparison of the computational times, we observe however that the pivoted Cholesky decomposition is the superior method.

REFERENCES

1. Ghanem R, Spanos P. *Stochastic finite elements. A spectral approach*. Springer: New York, 1991.
2. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press: Cambridge, 2005.

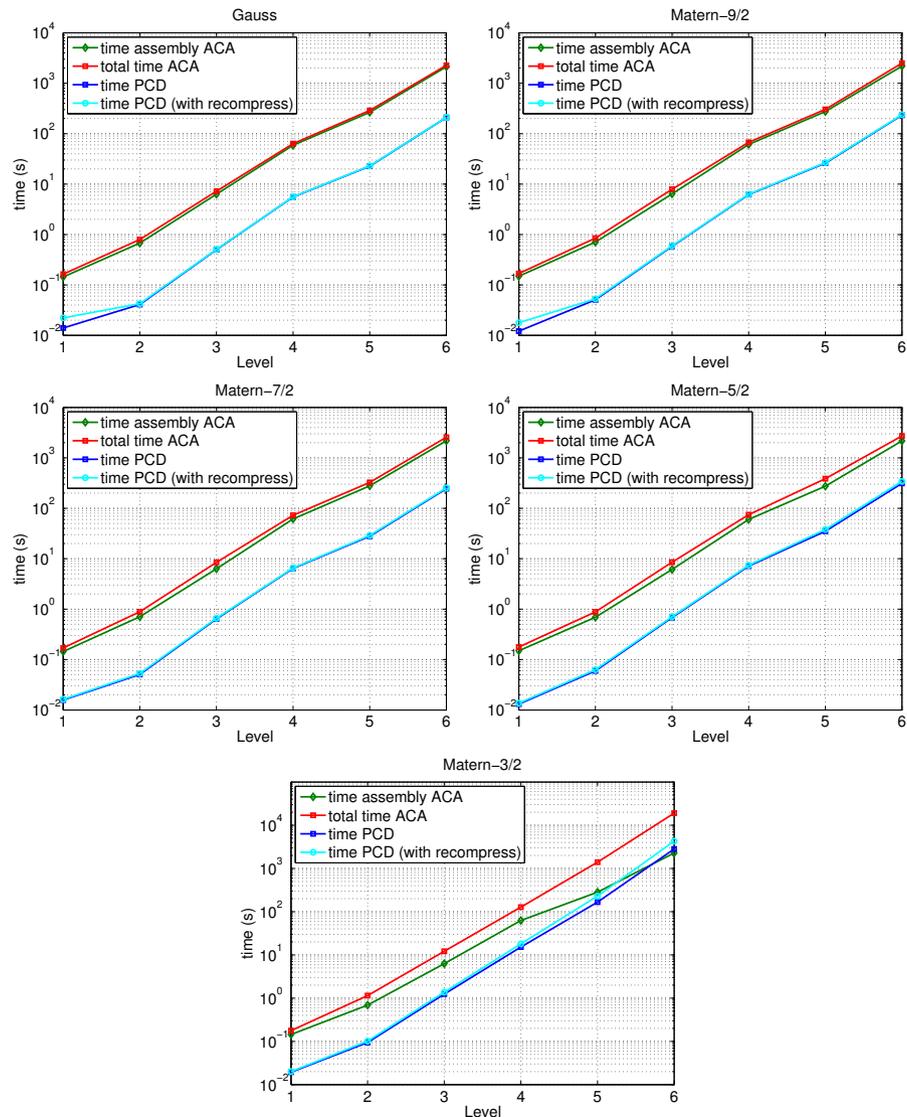


Figure 10. Numerical results (computational times) on the plate geometry.

3. Babuška I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis* 2007; **45**(3):1005–1034.
4. Babuška I, Tempone R, Zouraris G. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis* 2004; **42**(2):800–825.
5. Frauenfelder P, Schwab C, Todor R. Finite elements for elliptic problems with stochastic coefficients. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**(2-5):205–228.
6. Matthies H, Keese A. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**(12-16):1295–1331.
7. Schwab C, Gittelson CJ. Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numerica* 2011; **20**:291–467.
8. Light WA, Cheney EW. *Approximation theory in tensor product spaces*. Lecture notes in mathematics, Springer: Berlin-Heidelberg, 1985.
9. Loève M. *Probability theory. I+II*. Fourth edn., no. 45 in Graduate Texts in Mathematics, Springer: New York, 1977.
10. Giebermann K. Multilevel approximation of boundary integral operators. *Computing* 2001; **67**(3):183–207.
11. Schwab C, Todor R. Karhunen-Loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics* 2006; **217**:100–122.
12. Hackbusch W. *Hierarchische Matrizen: Algorithmen und Analysis*. Springer: Heidelberg, 2009.
13. Eiermann M, Ernst OG, Ullmann E. Computational aspects of the stochastic finite element method. *Computing and Visualization in Science* 2007; **10**(1):3–15.

14. Saad Y. *Numerical Methods for Large Eigenvalue Problems*. Manchester University Press: Manchester, 1992.
15. Bebendorf M, Rjasanow S. Adaptive low-rank approximation of collocation matrices. *Computing* 2003; **70**(1):1–24.
16. Bebendorf M. Approximation of boundary element matrices. *Numerische Mathematik* 2000; **86**(4):565–589.
17. Dahmen W, Harbrecht H, Schneider R. Compression techniques for boundary integral equations. Asymptotically optimal complexity estimates. *SIAM Journal on Numerical Analysis* 2006; **43**(6):2251–2271.
18. Harbrecht H, Schneider R. Wavelet Galerkin schemes for boundary integral equations. Implementation and quadrature. *SIAM Journal on Scientific Computing* 2006; **27**(4):1347–1370.
19. Beebe NHF, Linderberg J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *International Journal of Quantum Chemistry* 1977; **12**(4):683–705.
20. Foster L, Waagen A, Aijaz N, Hurley M, Luis A, Rinsky J, Satyavolu C, Way MJ, Gazis P, Srivastava A. Stable and efficient Gaussian process calculations. *The Journal of Machine Learning Research* 2009; **10**:857–882.
21. Harbrecht H, Peters M, Schneider R. On the low-rank approximation by the pivoted Cholesky decomposition. *Applied Numerical Mathematics* 2012; **62**:28–440.
22. Sauter SA, Schwab C. Quadrature for hp-Galerkin BEM in \mathbb{R}^3 . *Numerische Mathematik* 1997; **78**(2):211–258.
23. Golub GH, Van Loan CF. *Matrix Computations*. Fourth edn., Johns Hopkins University Press: Baltimore, 2012.
24. Lehoucq RB, Sorensen DC. Deflation techniques for an implicitly restarted Arnoldi iteration. *SIAM Journal on Matrix Analysis and Applications* 1996; **17**(4):789–821.
25. Lehoucq RB, Sorensen DC, Yang C. *Arpack User's Guide: Solution of Large-Scale Eigenvalue Problems With Implicitly Restarted Arnoldi Methods (Software, Environments, Tools)*. SIAM: Philadelphia, 1998.
26. Sorensen DC. Implicit application of polynomial filters in a k-step Arnoldi method. *SIAM Journal on Matrix Analysis and Applications* 1992; **13**(1):357–385.
27. Diestel J, Uhl JJ. *Vector Measures*. Mathematical surveys and monographs, American Mathematical Society, 1977.
28. Schatten R. *Norm Ideals of Completely Continuous Operators*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer: Berlin-Göttingen-Heidelberg, 1960.
29. Mercer J. Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 1909; :415–446.
30. Schatten R. *A Theory of Cross-Spaces*. Princeton University Press: Princeton, 1950.
31. Temlyakov VN. Approximations of functions with bounded mixed derivative. *Trudy Matematicheskogo Instituta im. VA Steklova* 1986; **178**:3–113.
32. Griebel M, Harbrecht H. Approximation of bi-variate functions: singular value decomposition versus sparse grids. *IMA Journal of Numerical Analysis* 2014; **34**(1):28–54.
33. Braess D. *Finite Elements. Theory, fast solvers, and applications in solid mechanics*. Second edn., Cambridge University Press: Cambridge, 2001.
34. Brenner S, Scott L. *The mathematical theory of finite element methods*. Third edn., Springer: Berlin, 2008.
35. Knyazev A. New estimates for ritz vectors. *Mathematics of Computation of the American Mathematical Society* 1997; **66**(219):985–995.
36. Ovtchinnikov E. Cluster robust error estimates for the rayleigh–ritz approximation i: Estimates for invariant subspaces. *Linear algebra and its applications* 2006; **415**(1):167–187.
37. D'yakonov EG, McCormick SF. *Optimization in Solving Elliptic Problems*. CRC Press: Boca Raton, 1996.
38. Higham NJ. Analysis of the cholesky decomposition of a semi-definite matrix. in *Reliable Numerical Computation*, Oxford University Press: Oxford, 1990; 161–185.
39. Higham NJ. *Accuracy and Stability of Numerical Algorithms*. Second edn., Society for Industrial and Applied Mathematics, 2002.
40. Matern B. *Spatial Variation*. Springer Lecture Notes in Statistics, Springer: New York, 1986.
41. Abramowitz M, Stegun IA. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series, Dover Publications, 1964.
42. Graham I, Kuo F, Nichols J, Scheichl R, Schwab C, Sloan I. Quasi-Monte Carlo finite element methods for elliptic PDEs with log-normal random coefficient. *SAM-Report* 2013; **2013-14**.
43. Müller C. *Analysis of Spherical Symmetries in Euclidean Spaces*. No. 129 in Applied Mathematical Sciences Series, Springer: Berlin-Heidelberg, 1998.
44. Harbrecht H, Peters M. Comparison of fast boundary element methods on parametric surfaces. *Computer Methods in Applied Mechanics and Engineering* 2013; **261-262**:39–55.
45. Kernighan BW, Ritchie DM. *The C programming language*. Prentice-Hall: Upper Saddle River, NJ, 1988.
46. Lucas C. LAPACK-style codes for level 2 and 3 pivoted Cholesky factorizations. *LAPACK Working* 2004; .