ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Mathematics Institute of Computational Science and Engineering

School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report

Nr. 24.2014

Mai 2014

Low-rank tensor approximation for high-order correlation functions of Gaussian random fields

Daniel Kressner, Rajesh Kumar, Fabio Nobile, Christine Tobler

http://mathicse.epfl.ch

Phone: +41 21 69 37648

Address:
EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland

Fax: +41 21 69 32545

Low-rank tensor approximation for high-order correlation functions of Gaussian random fields*

Daniel Kressner† Rajesh Kumar‡* Fabio Nobile‡* Christine Tobler†

May 12, 2014

Abstract Gaussian random fields are widely used as building blocks for modeling stochastic processes. This paper is concerned with the efficient representation of d-point correlations for such fields, which in turn enables the representation of more general stochastic processes that can be expressed as a function of one (or several) Gaussian random fields. Our representation consists of two ingredients. In the first step, we replace the random field by a truncated Karhunen-Loève expansion and analyze the resulting error. The parameters describing the d-point correlation can be arranged in a tensor, but its storage grows exponentially in d. To avoid this, the second step consists of approximating the tensor in a low-rank tensor format, the so called Tensor Train decomposition. By exploiting the particular structure of the tensor, an approximation algorithm is derived that does not need to form this tensor explicitly and allows to process correlations of order as high as $d = 20$. The resulting representation is very compact and its use is illustrated for elliptic partial differential equations with random Gaussian forcing terms.

Keywords: Gaussian random fields, n-points correlations, low-rank approximation, Karhunen-Loève expansion, Tensor Train decomposition.

1 Introduction

This paper is concerned with the efficient representation of d-point correlation functions for a (possibly non-stationary) Gaussian random field f on a bounded domain $D \subset \mathbb{R}^n$. We particularly focus on the case $d > 2$. This gives rise to a number of challenges, as a d-point correlation function is defined on the domain $D \times D \times \ldots \times D \subset \mathbb{R}^{dn}$. Therefore a naive representation would lead to storage requirements growing exponentially in d. In this work, we propose to overcome this issue by combining a truncated Karhunen-Loève (KL) expansion of the random field with low-rank tensor techniques.

Gaussian random fields are widely used as building blocks for modeling stochastic processes. Efficient representations of d-point correlations, as the one proposed in this paper, will then

†ANCHP, EPF Lausanne, station 8, CH-1015 Lausanne, Switzerland. daniel.kressner@epfl.ch. CT is now with Mathworks, Natick, USA. christine.tobler@mathworks.com.

‡CSQI, EPF Lausanne, station 8, CH-1015 Lausanne, Switzerland. fabio.nobile@epfl.ch. RK is now with RICAM, Linz, Austria. rajesh.kumar@oeaw.ac.at.

*Supported by the Italian grant FIRB-IDEAS (Project n. RBID08223Z) "Advanced Numerical Techniques for Uncertainty Quantification in Engineering and Life Science Problems (NUMQUES)".

be useful to derive analogous representations and compute statistics of more general stochastic processes that can be expressed as a function of one (or several) Gaussian random fields.

The target application discussed in this work is the computation of statistics of the solution of a well-posed linear partial differential equation (PDE) with random Gaussian input data. In this case, one can derive exact high-dimensional differential equations, obtained by d-fold tensorization of the original PDE, that relate the d-point correlation of the solution with the d-point correlation of the input data. This problem has already been addressed several times in the literature. We mention, in particular, the works [22, 26, 11] that propose and investigate sparse finite element approximations of the d-moment equation for strongly elliptic PDEs. More recently, extensions of this idea to (non-coercive) problems in mixed form of Hodge-Laplace type have been considered in [4, 13]. Thanks to the fact that the d-point correlation of the solution of the PDEs typically has Sobolev regularity on mixed derivatives, sparse finite element approximations appear to be particularly suited. However, their implementation in a Galerkin framework remains quite cumbersome. An approach using frames has been proposed in [11], which allows to reuse available deterministic codes. However, one still has to construct a hierarchy of nested triangulations and the corresponding discrete deterministic problems.

Our approach is alternative to the ones mentioned above and makes use of a low-rank representation of the tensor associated with the d-point correlation expanded in the KL eigenfunctions. We first focus on the representation of the correlations of a given Gaussian random field in the Tensor Train (TT) format proposed in [17, 19]. Different methods for approximating a given tensor in the TT format exist. In principle, the cross approximation algorithms proposed in [18, 20] are well suited for this purpose, as they only need to access selected entries from the tensor. However, in our preliminary numerical experiments, we have found these methods to be too slow in the context of our application. We therefore propose a new method that takes the particular structure of the tensors at hand into account. We stress that our algorithm heavily relies on the fact that the random field is expanded in an orthonormal basis (like the basis of KL eigenfunctions), and does not directly apply to other types of non-orthogonal expansions such as the one produced by a pivoted Cholesky decomposition [10].

When considering the d-point correlation for the solution of a PDE with the data represented in a low rank tensor format, the tensor structure of the d-moment equation allows to obtain a low rank approximation to the solution in a straightforward manner, without resorting to a Galerkin projection on a sparse approximation space. Our approach has the additional advantages that it can be built on any available deterministic solver as a black-box. Moreover, we observe that the low rank representation for the solution only has a comparably mild dependence on d.

We finally mention that the ideas presented here can be extended to non-linear PDEs or PDEs depending non-linearly on the input random field, as it is the case for instance for PDEs with random coefficients or defined in random domains, by adopting a perturbation approach in the case of small noise, see e.g. [3, 5, 6, 12, 25].

The rest of this paper is organized as follows. In Section 2, we derive the expansion for the d-point correlation of a given Gaussian random field f in terms of tensorized KL eigenfunctions. Moreover, we show that the resulting tensor is sparsifiable in the sense that only M terms need to be retained to achieve an error $O(M^{-\alpha})$ with α independent of d (the involved constant, however, depends strongly on d), under mild assumptions on the decay of the KL expansion. Section 3 is concerned with the approximation of the tensor in the TT format and presents cheaply computable expressions for the error. In Section 4, several numerical examples demonstrate the efficiency of our approach in computing correlations for Gaussian random fields with

covariance function from the Matérn family up to $d = 20$. Section 5 concludes the paper by applying our approach to linear elliptic PDEs with random forcing terms.

2 Series representations of Gaussian random fields and their d-point correlations

Let $D \subset \mathbb{R}^n$ be an open, bounded domain and (Ω, Σ, P) a complete probability space, Ω being the set of outcomes, $\Sigma \subset 2^\Omega$ the σ-algebra of events and $P : \Sigma \to [0,1]$ a probability measure. In this work, we consider a real valued *Gaussian random field* $f : \overline{D} \times \Omega \to \mathbb{R}$ (see e.g. [1]) with continuous covariance function $\mathrm{Cov} : \overline{D} \times \overline{D} \to \mathbb{R}$,

$$\mathrm{Cov}(x,y) = \mathbb{E}[(f(x,\cdot) - \mathbb{E}[f](x))(f(y,\cdot) - \mathbb{E}[f](y))], \qquad x, y \in \overline{D}$$

where $\mathbb{E}[X] = \int_\Omega X(\omega) dP(\omega)$ denotes the expectation of a random variable $X : \Omega \to \mathbb{R}$.

By Mercer's theorem, the random field f can be decomposed into the so called Karhunen-Loève expansion (see e.g. [15, 16])

$$f(x,\omega) = \mathbb{E}[f](x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} Y_i(\omega) \Phi_i(x) \tag{1}$$

where $\lambda_i \geq 0$ are the eigenvalues of the covariance operator $K : L^2(D) \to L^2(D)$ defined by $(K\psi)(x) = \int_{D\times D} \mathrm{Cov}(x,y)\psi(y)\mathrm{d}y$. The corresponding eigenfunctions Φ_i form an orthonormal basis in $L^2(D)$, whereas Y_i are independent standard Gaussian random variables, $Y_i \sim \mathcal{N}(0,1)$. We assume hereafter that the eigenvalues λ_i have been ordered in decreasing order.

It is well known that the series in (1) exhibits mean-square convergence in ω and uniform convergence in x. Let us denote the N-term truncated Karhunen-Loève expansion by $f_N = \mathbb{E}[f] + \sum_{i=1}^N \sqrt{\lambda_i} Y_i \Phi_i$. Since $f - f_N$ is a centered Gaussian random field for any N, we actually have L^p-convergence for any $p > 0$:

$$\lim_{N\to\infty} \sup_{x\in D} \mathbb{E}[|f - f_N|^p] = 0, \qquad \forall p > 0. \tag{2}$$

This result follows from the fact that any centered Gaussian random variable satisfies $\mathbb{E}[|X|^p] = C(p)\mathbb{E}[X^2]^{\frac{p}{2}}$ with $C(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |x|^p e^{-x^2/2} \mathrm{d}x$. In particular, for p even, $C(p) = \gamma_p = \frac{p!}{2^{p/2}(p/2)!}$ is the p-th moment of the standard normal distribution.

We finally recall that

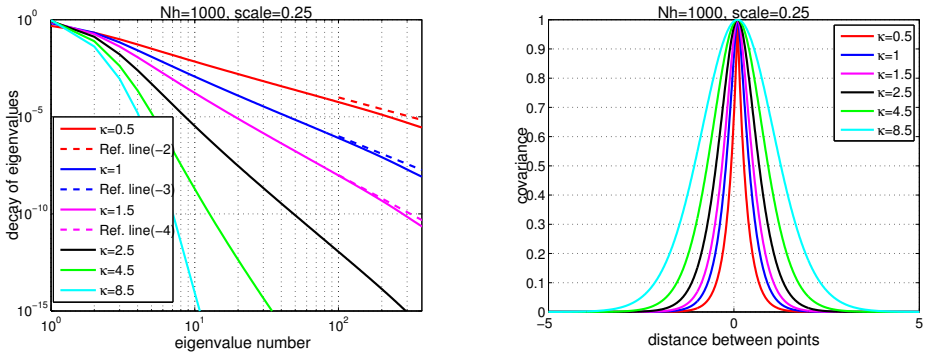
$$\int_D \mathrm{Var}(f) = \sum_{i=1}^{\infty} \lambda_i, \qquad \text{and} \qquad \int_D \mathrm{Var}(f - f_N) = \sum_{i=N+1}^{\infty} \lambda_i.$$

Matérn family. This family of covariance functions, for stationary random fields, has been widely employed for instance in geostatistics applications [8] and will be used in the numerical experiments described in Sections 4 and 5. For all members of this family, the covariance $\mathrm{Cov}(x,y)$ only depends on x, y via $\|x - y\|$:

$$\mathrm{Cov}(x,y) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)} \left(\frac{\|x-y\|}{l_c} \right)^\kappa K_\kappa \left(\frac{\|x-y\|}{l_c} \right), \tag{3}$$

where $K_\kappa(\cdot)$ is the modified Bessel function of the second kind of order $\kappa > 0$, and $l_c > 0$ is a scale parameter that represents a characteristic *correlation length*, see [8]. The corresponding centered and stationary Gaussian random field is κ-times mean square differentiable.

For $\kappa = 0.5$, the Matérn covariance function reduces to the exponential covariance $\mathrm{Cov}(x, y) = e^{-\|x-y\|/l_c}$, while it reduces to the Gaussian covariance $\mathrm{Cov}(x, y) = e^{-\|x-y\|^2/l_c^2}$ for $\kappa \to \infty$. In Figure 1(a) we show the decay of the eigenvalues λ_i of the corresponding covariance operator in one dimension for different values of κ. Notice that the eigenvalues decay asymptotically as $\lambda_i \sim i^{-(\frac{2\kappa}{n}+1)}$ for $i \to \infty$, see e.g. [9].



(a) eigenvalues (b) covariance

Figure 1: Decay of eigenvalues (a) and the covariance (b) of the Matérn covariance operator for different values of κ.

2.1 Series expansion of d-point correlations

In the following, we aim to approximate the d-point correlation of f, defined as

$$\mu_f^d : \overline{D}^d \to \mathbb{R}, \qquad \mu_f^d(x_1, \ldots, x_d) := \mathbb{E}\Big[\prod_{\eta=1}^{d} \big(f(x_\eta, \cdot) - \mathbb{E}[f](x_\eta)\big)\Big].$$

Inserting the Karhunen-Loève expansion of f from (1) leads to the following series representation:

$$\mu_f^d(x_1, \ldots, x_d) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_d=1}^{\infty} \mathbb{E}\Big[\prod_{\eta=1}^{d} \sqrt{\lambda_{i_\eta}} Y_{i_\eta}\Big] \bigotimes_{\eta=1}^{d} \Phi_{i_\eta}(x_\eta). \tag{4}$$

Because of (2) this series converges uniformly in $\overline{D}^d$.

We now introduce the multi-index $\mathbf{i} \in \mathbb{N}^d$ and denote by $m_{\mathbf{i}}(\ell)$ the multiplicity of ℓ in the multi-index $\mathbf{i} = (i_1, \ldots, i_d)$, that is,

$$m_{\mathbf{i}}(\ell) := \sharp\{i_j = \ell, j = 1, \ldots, d\} \quad \text{for} \quad \ell = 1, 2, \ldots. \tag{5}$$

Notice that $m_{\mathbf{i}}$ has finite support, having d non-zero entries at most. Thanks to the mutual independence of the random variables Y_i, the series expansion (4) can be equivalently written

as

$$\mu_f^d(x_1,\dots,x_d) = \sum_{\mathbf{i}\in\mathbb{N}^d} \mathcal{C}_{i_1,\dots,i_d} \bigotimes_{\eta=1}^{d} \Phi_{i_\eta}(x_\eta), \qquad \mathcal{C}_{i_1,\dots,i_d} = \prod_{\ell=1}^{\infty} \lambda_\ell^{m_{\mathbf{i}}(\ell)/2} \mathbb{E}\Big[Y_\ell^{m_{\mathbf{i}}(\ell)}\Big]. \tag{6}$$

Recalling that the moments of a standard normal random variable $Y \sim \mathcal{N}(0,1)$ satisfy

$$\gamma_m = \mathbb{E}[Y^m] = \begin{cases} 0, & \text{for } m \text{ odd}, \\ \frac{m!}{2^{m/2}(m/2)!}, & \text{for } m \text{ even}, \end{cases}$$

it follows that $\mathcal{C}_{i_1,\dots,i_d}$ is nonzero only when all multiplicities $m_{\mathbf{i}}(\ell)$ are even. In particular, the d-point correlation μ_f^d is always zero for odd d and we therefore focus on the case of even d from now on.

Assuming that all multiplicities $m_{\mathbf{i}}(\ell)$ are even implies that the index $\mathbf{i} = (i_1,\dots,i_d)$ must be a permutation of $(j_1, j_1, j_2, j_2, \dots, j_{d/2}, j_{d/2})$ for $\mathbf{j} = (j_1,\dots,j_{d/2}) \in \mathbb{N}^{\frac{d}{2}}$. We now define the set $\Sigma \subset \mathbb{N}^{\frac{d}{2}}$ as the set of $d/2$-tuples with increasing entries:

$$\Sigma = \{\mathbf{j} \in \mathbb{N}^{\frac{d}{2}}, \text{ such that } j_1 \le j_2 \le \ldots \le j_{d/2}\}.$$

For any $\mathbf{j} \in \Sigma$, we let $P_{\mathbf{j}}$ denote the set of all unique permutations of $(j_1, j_1, j_2, j_2, \dots, j_{d/2}, j_{d/2})$. For example, when $d = 4$ and $j_1 = 1, j_2 = 2$,

$$P_{\mathbf{j}} = \Big\{(1,1,2,2),(1,2,1,2),(1,2,2,1),(2,1,1,2),(2,1,2,1),(2,2,1,1)\Big\}.$$

This notation allows us to equivalently write the d-point correlation as

$$\mu_f^d(x_1,\dots,x_d) = \sum_{\mathbf{j}\in\Sigma} \sum_{\mathbf{i}\in P_{\mathbf{j}}} \left(\prod_{\ell=1}^{\infty} \lambda_\ell^{m_{\mathbf{j}}(\ell)} \gamma_{2m_{\mathbf{j}}(\ell)}\right) \left(\bigotimes_{\eta=1}^{d} \Phi_{i_\eta}(x_\eta)\right). \tag{7}$$

2.2 Optimal Series truncation and error estimates

In this section, we consider best M-term approximations of the series (7) representing the d-point correlation of f. We derive estimates for the corresponding approximation error, with the main result given by Theorem 2.3 below.

Let $\Lambda_M \subset \Sigma$ be an arbitrary subset of Σ such that $\sum_{\mathbf{j}\in\Lambda_M} \sharp P_{\mathbf{j}} = M$. We consider the approximation of μ_f^d when restricting the first summation in (7) to the subset Λ_M:

$$\mu_{f,\Lambda_M}^d(x_1,\dots,x_d) = \sum_{\mathbf{j}\in\Lambda_M} \sum_{\mathbf{i}\in P_{\mathbf{j}}} \prod_{\ell=1}^{\infty} \lambda_\ell^{m_{\mathbf{j}}(\ell)} \gamma_{2m_{\mathbf{j}}(\ell)} \left(\bigotimes_{\eta=1}^{d} \Phi_{i_\eta}(x_\eta)\right).$$

Notice that this approximation contains exactly M rank-one tensor product terms $\bigotimes_{\eta=1}^{d} \Phi_{i_\eta}(x_\eta)$. It can therefore be viewed as a rank-M approximation in the *Canonical Polyadic* (CP) format [14]. We now aim at estimating the $L^2(D^d)$-error between μ_f^d and μ_{f,Λ_M}^d given by

$$\|\mu_f^d - \mu_{f,\Lambda_M}^d\|_{L^2(D^d)}^2 = \underbrace{\int_D \cdots \int_D}_{d \text{ times}} \left(\sum_{\mathbf{j}\in\Lambda_M^c} \sum_{\mathbf{i}\in P_{\mathbf{j}}} \prod_{\ell=1}^{\infty} \lambda_\ell^{m_{\mathbf{j}}(\ell)} \gamma_{2m_{\mathbf{j}}(\ell)} \left(\Phi_{i_1} \otimes \cdots \otimes \Phi i_d\right)\right)^2,$$

where $\Lambda_M^c = \Sigma \setminus \Lambda_M$. Using the L^2-orthogonality of the basis $\{\Phi_i\}_{i=1}^\infty$, this simplifies to

$$\|\mu_f^d - \mu_{f,\Lambda_M}^d\|_{L^2(D^d)}^2 = \sum_{\mathbf{j}\in\Lambda_M^c} \sum_{\mathbf{i}\in P_\mathbf{j}} \prod_{\ell=1}^\infty \lambda_\ell^{2m_\mathbf{j}(\ell)} \gamma_{2m_\mathbf{j}(\ell)}^2 = \sum_{\mathbf{j}\in\Lambda_M^c} \sharp P_\mathbf{j}\, a_\mathbf{j}, \tag{8}$$

where we have set

$$a_\mathbf{j} = \prod_{\ell=1}^\infty \lambda_\ell^{2m_\mathbf{j}(\ell)} \gamma_{2m_\mathbf{j}(\ell)}^2.$$

To state our results, we need to introduce the following additional notation:

- $\{a_l\}_{l=1}^\infty$ is the sequence of the coefficients $\{a_\mathbf{j},\ \mathbf{j}\in\Sigma\}$ ordered in decreasing order such that $a_l \geq a_{l+1}$. To simplify the notation, we still use the symbol $\{a_l\}$.
- $\mathbf{j}(l)$ is the multi-index $\mathbf{j} \in \Sigma$ corresponding to the l-th coefficient in the ordered sequence $\{a_l\}$.
- $\{\widetilde{a}_l\}_{l=1}^\infty$ is the repeated sequence defined as

$$\{\widetilde{a}_l\}_{l=1}^\infty = \{\underbrace{a_1,\ldots,a_1}_{\sharp P_{\mathbf{j}(1)}}, \underbrace{a_2,\ldots,a_2}_{\sharp P_{\mathbf{j}(2)}}, \ldots\}.$$

We now define the set of M largest coefficients (each one considered with its multiplicity) as

$$\Lambda_M^{\mathsf{opt}} = \{\mathbf{j}\in\Sigma,\ \text{corresponding to the } M \text{ largest } \widetilde{a}_l\}.$$

The following two lemmas will be needed to derive our main result.

Lemma 2.1. *Assuming that the sequence $\{\widetilde{a}_l\}$ is p-summable for some $p \leq 1$, that is,*

$$\sum_{l=1}^\infty \widetilde{a}_l^p = \sum_{l=1}^\infty \sharp P_{\mathbf{j}(l)} a_l^p < \infty,$$

we have

$$\|\mu_f^d - \mu_{f,\Lambda_M^{\mathsf{opt}}}^d\|_{L^2(D^d)} \leq M^{1/2-1/2p} \left(\sum_{l=1}^\infty \widetilde{a}_l^p\right)^{1/2p}.$$

Proof. It can easily be seen from equation (8) that

$$\|\mu_f^d - \mu_{f,\Lambda_M^{\mathsf{opt}}}^d\|_{L^2(D^d)}^2 = \sum_{\mathbf{j}\in(\Lambda_M^{\mathsf{opt}})^c} \sharp P_\mathbf{j}\, a_\mathbf{j} = \sum_{l>M} \widetilde{a}_l.$$

Now, Stechkin's Lemma (see, e.g., [7]) immediately implies

$$\|\mu_f^d - \mu_{f,\Lambda_M^{\mathsf{opt}}}^d\|_{L^2(D^d)}^2 \leq M^{1-1/p} \left(\sum_{l=1}^\infty (\widetilde{a}_l)^p\right)^{1/p}.$$

□

The following lemma quantifies the p-summability of the sequence $\{\widetilde{a}_l\}$.

Lemma 2.2. *Suppose that the sequence $\{\lambda_\ell\}_\ell$ is $2p$-summable, that is, $\sum_{\ell=1}^{\infty}\lambda_\ell^{2p} < +\infty$ for some $p < 1/2$. Then the sequence $\{\widetilde{a}_l\}_l$ is p-summable and*

$$\sum_{l=1}^{\infty}\widetilde{a}_l^p \le \gamma_d\left(\sum_{\ell=1}^{\infty}\lambda_\ell^{2p}\right)^{\frac{d}{2}}.$$

Proof. We have

$$\sum_{l=1}^{\infty}\widetilde{a}_l^p = \sum_{l=1}^{\infty}\sharp P_{\mathbf{j}(l)}a_l^p = \sum_{\mathbf{j}\in\Sigma}\sharp P_{\mathbf{j}}\left(\prod_{\ell=1}^{\infty}\lambda_\ell^{2m_{\mathbf{j}}(\ell)}\gamma_{2m_{\mathbf{j}}(\ell)}^2\right)^p.$$

Let us now introduce sequences $\mathbf{m}\in\mathbb{N}^{\mathbb{N}}$ with $|\mathbf{m}| = \sum_{j=1}^{\infty}m_j < \infty$ and use the notation

$$\mathbf{m}! = \prod_{j=1}^{\infty}m_j! \quad\text{and}\quad 2^{\mathbf{m}} = \prod_{j=1}^{\infty}2^{m_j} = 2^{|\mathbf{m}|}.$$

Substituting the value of $\sharp P_{\mathbf{j}}$ and using the fact that there is a one-to-one correspondence between $\mathbf{j}\in\Sigma$ and $\mathbf{m}\in\mathbb{N}^{\mathbb{N}}$ with $|\mathbf{m}| = d/2$, we obtain

$$\begin{aligned}\sum_{l=1}^{\infty}\widetilde{a}_l^p &= \sum_{|\mathbf{m}|=d/2}\frac{(2|\mathbf{m}|)!}{(2\mathbf{m})!}\left(\prod_{\ell=1}^{\infty}\lambda_\ell^{2\,p\,m_\ell}\gamma_{2m_\ell}^{2p}\right)\\ &= \sum_{|\mathbf{m}|=d/2}\frac{d\,!}{(2\mathbf{m})!}\left(\frac{(2\mathbf{m})!}{2^{\mathbf{m}}\mathbf{m}!}\right)^{2p}\prod_{\ell=1}^{\infty}\lambda_\ell^{2\,p\,m_\ell}\\ &= \frac{d\,!}{2^{dp}(d/2)!}\sum_{|\mathbf{m}|=d/2}\frac{(2\mathbf{m})!^{2p-1}}{\mathbf{m}!^{2p-1}}\frac{(d/2)!}{\mathbf{m}!}\prod_{\ell=1}^{\infty}\lambda_\ell^{2\,p\,m_\ell}.\end{aligned}$$

The last expression can be simplified further to

$$\sum_{l=1}^{\infty}\widetilde{a}_l^p = \frac{d\,!}{2^{dp}(d/2)!}2^{d(2p-1)/2}\sum_{|\mathbf{m}|=d/2}((2\mathbf{m}-1)!!)^{2p-1}\frac{(d/2)!}{\mathbf{m}!}\prod_{\ell=1}^{\infty}\lambda_\ell^{2\,p\,m_\ell}.$$

Since $p < 1/2$, we have $((2m-1)!!)^{2p-1} < 1$ for any $m > 0$ and it therefore follows that

$$\sum_{l=1}^{\infty}\widetilde{a}_l^p \le \frac{d\,!}{2^{d/2}(d/2)!}\sum_{|\mathbf{m}|=d/2}\frac{(d/2)!}{\mathbf{m}!}\prod_{\ell=1}^{\infty}\lambda_\ell^{2\,p\,m_\ell} = \gamma_d\left(\sum_{\ell=1}^{\infty}\lambda_\ell^{2p}\right)^{d/2},$$

where we have used the multinomial theorem in the last step. This yields the desired result. □

The following theorem states the main result of this section.

Theorem 2.3. *Assume there exists $p < 1$ such that*

$$\sum_{\ell=1}^{\infty} \lambda_\ell^p < +\infty, \tag{9}$$

then, for any even $d \geq 2$,

$$\|\mu_f^d - \mu_{f,\Lambda_M^{\mathsf{opt}}}^d\|_{L^2(D^d)} \leq M^{\frac{1}{2}-\frac{1}{p}} \left[\gamma_d \left(\sum_{\ell=1}^{\infty} \lambda_\ell^p\right)^{\frac{d}{2}}\right]^{\frac{1}{p}} \tag{10}$$

with $\gamma_d = \frac{d!}{2^{d/2}(d/2)!}$.

Proof. The result follows directly from combining the results of Lemma 2.1 and Lemma 2.2, with $2p$ replaced by p. □

Remark 2.4 (Exponential decay)**.** *In the case of exponential decay of the eigenvalues of the KL expansion, $\lambda_j \leq Ce^{-sj}$, $s > 0$, the sum in* (9) *is bounded for any $p > 0$:*

$$\sum_{\ell=1}^{\infty} \lambda_\ell^p \leq (Ce^{-s})^p \frac{1}{1-e^{-sp}}.$$

The freedom in choosing the parameter p can be used to optimize the bound (10)*. For this purpose, we follow closely the argument in [2]. Using the asymptotic estimate $(1-e^{-sp}) \sim sp$ for $p \ll 1$ and $\gamma_d \sim \sqrt{2}(d/e)^{d/2}$, obtained from the Stirling approximation, the bound* (10) *can be written as*

$$\|\mu_f^d - \mu_{f,\Lambda_M^{\mathsf{opt}}}^d\|_{L^2(D^d)} \lesssim \sqrt{(Ce^{-s})^d M} \left[\sqrt{2} M^{-1} \left(\frac{d}{esp}\right)^{\frac{d}{2}}\right]^{\frac{1}{p}}.$$

This expression is minimized for $p = \frac{d}{s}\sqrt[d]{2M^{-2}}$, leading to the approximate bound

$$\|\mu_f^d - \mu_{f,\Lambda_M^{\mathsf{opt}}}^d\|_{L^2(D^d)} \lesssim \sqrt{(Ce^{-s})^d M} \exp\{-s2^{-\frac{d+1}{d}} M^{\frac{2}{d}}\}. \tag{11}$$

3 Tensor compression of d-point correlations

In this section, we will discuss a compressed storage scheme for representing the d-point correlation μ_f^d. Starting from the representation (6), we will proceed in two steps.

In the first step, we truncate the infinite sum over $\mathbf{i} \in \mathbb{N}^d$ to a finite sum over $\mathbf{i} \in \{1, \ldots, N\}^d$, which corresponds to considering only the first N terms in the KL expansion (1). The effect of this truncation will be analyzed in Section 3.1.

In the second step, we consider the $N \times \cdots \times N$ tensor $\mathcal{C}$ of order d containing all coefficients

$$\mathcal{C}_{i_1,\ldots,i_d} = \prod_{\ell=1}^{N} \lambda_\ell^{m_{\mathbf{i}}(\ell)/2} \gamma_{m_{\mathbf{i}}(\ell)}, \qquad \mathbf{i} \in \{1, \ldots, N\}^d, \tag{12}$$

where $\gamma_{m_{\mathbf{i}}(\ell)} = \mathbb{E}\big[Y_\ell(\omega)^{m_{\mathbf{i}}(\ell)}\big]$. Even when exploiting the many zero entries of $\mathcal{C}$, constructing or storing this tensor is impossible, except for very small values of d, say $d = 2$ or $d = 4$. It will therefore be necessary to store $\mathcal{C}$ approximately in a data-sparse format. For this purpose, we will make use of the so called tensor train (TT) decomposition [19], which we briefly recall in Section 3.2. Methods for computing exact and approximate TT decompositions of $\mathcal{C}$ are described in Section 3.3 and Section 3.4, respectively.

3.1 Error from truncating the Karhunen-Loève expansion

Assuming that the function $f(x, \omega)$ admits a KL expansion of the form (1), we will consider the truncated KL expansion

$$f_N(x, \omega) = \mathbb{E}(f) + \sum_{i=1}^{N} \sqrt{\lambda_i} Y_i(\omega) \Phi_i(x).$$

This section is concerned with computing the resulting error in the corresponding d-point correlation function.

In practice, the function f is discretized in space and, consequently, the KL expansion only has *finitely* many terms. More specifically, the number $\widetilde{N} \gg N$ of these terms corresponds to the degrees of freedom in the discretization. In the following, we assume such a setting and derive computable formulas for

$$\mathrm{err}_N := \|\mu_f^d - \mu_{f_N}^d\|_{L^2(D^d)}, \tag{13}$$

where μ_f^d and $\mu_{f_N}^d$ are the d-point correlation functions for f and f_N, respectively. We now consider the corresponding $\widetilde{N} \times \cdots \times \widetilde{N}$ tensor $\widetilde{\mathcal{C}}$ and the $N \times \cdots \times N$ tensor $\mathcal{C}$, defined from the (truncated) KL expansions as in (12). Since the functions $\Phi_i(x)$ are L^2 orthonormal, it follows from (6) that

$$\mathrm{err}_N = \|\widetilde{\mathcal{C}} - \mathcal{C}\|_F^2,$$

where we implicitly padded zeros to the different modes of $\mathcal{C}$ to match the size of $\widetilde{\mathcal{C}}$. Note that $\|\cdot\|_F$ denotes the usual Frobenius norm of a tensor.

Since $\widetilde{\mathcal{C}}_{\mathbf{i}}$ and $\mathcal{C}_{\mathbf{i}}$ are identical for all $\mathbf{i} \in \{1, \ldots, N\}^d$, we have

$$\mathrm{err}_N = \sqrt{\|\widetilde{\mathcal{C}}\|_F^2 - \|\mathcal{C}\|_F^2}. \tag{14}$$

Thus, the problem of computing err_N has been reduced to computing norms of tensors. However, even the naive computation of these norms becomes way too expensive for larger d. The following lemma provides a much cheaper way.

Lemma 3.1. *Consider the d-th order tensor $\mathcal{C} \in \mathbb{R}^{N \times \cdots \times N}$ defined by*

$$\mathcal{C}_{i_1, \ldots, i_d} = \prod_{\ell=1}^{N} \lambda_\ell^{m_{\mathbf{i}}(\ell)/2} \gamma_{m_{\mathbf{i}}(\ell)}.$$

for some scalars $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ and $\gamma_0, \ldots, \gamma_d \in \mathbb{R}$. Then

$$\|\mathcal{C}\|_F^2 = d! \left(a^{(1)} * a^{(2)} * \cdots * a^{(N)}\right)_d,$$

where the second factor is the d-th component of the discrete convolution $*$ of the vectors $a^{(j)} \in \ell^2(\mathbb{Z})$, $j = 1, \dots, N$, defined as

$$a_m^{(j)} = \begin{cases} \frac{\lambda_j^m \gamma_m^2}{m!} & \text{if } m \in \{0, \dots, d\}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We aim to calculate

$$\|\mathcal{C}\|_F^2 = \sum_{\mathbf{i} \in \{1,\dots,N\}^d} \mathcal{C}_{\mathbf{i}}^2.$$

Similarly to the discussion in Section 2, the symmetry of $\mathcal{C}_{i_1,\dots,i_d}$ allows us to rewrite this sum in terms of the multi-index $\mathbf{m} = (m_1, \dots, m_d)$:

$$\|\mathcal{C}\|_F^2 = \sum_{\mathbf{i} \in \Sigma} \frac{d!}{\prod_{j=1}^N m_{\mathbf{i}}(j)!} \mathcal{C}_{\mathbf{i}}^2 = d! \sum_{\mathbf{i} \in \Sigma} \prod_{j=1}^N \frac{\lambda_j^{m_{\mathbf{i}}(j)} \gamma_{m_{\mathbf{i}}(j)}^2}{m_{\mathbf{i}}(j)!} = d! \sum_{|\mathbf{m}|=d} \prod_{j=1}^N a_{m_j}^{(j)},$$

where $\Sigma := \big\{(i_1, \dots, i_d) \in \{1, \dots, N\}^d : i_1 \le i_2 \le \cdots \le i_d\big\}$ and $m_{\mathbf{i}}(j)$ is defined as in (5). By recursive application of discrete convolution, defined by

$$(a^{(1)} * a^{(2)})_k = \sum_{j=-\infty}^{\infty} a_j^{(1)} a_{k-j}^{(2)},$$

it follows that $\sum_{|\mathbf{m}|=d} \prod_{j=1}^N a_{m_j}^{(j)} = (a^{(1)} * a^{(2)} * \cdots * a^{(N)})_d$, which concludes the proof. □

Lemma 3.1 shows that $\|\mathcal{C}\|_F$ can be computed within $\mathcal{O}(N^2 d \log(Nd))$ operations, when using the FFT for computing the discrete convolutions. For larger d, say $d \ge 16$, we observed that the FFT leads to accuracy problems, probably because it mixes very small entries with larger entries in the vectors $a^{(j)}$. This problem seems to disappear when implementing the discrete convolution directly according to its definition, which is still comparably cheap.

To determine a suitable cutoff parameter N, we can now progressively increase $N = 1, 2, 3, \dots$, until the error (13) is smaller than a given tolerance ϵ. The results of computing the cutoff error for two different decays of λ_j are shown in Figure 2, where we have used $\widetilde{N} = 1000$. These plots correspond quite well with a bound of the form

$$\|\mathcal{C} - \mathcal{C}_N\|_F \lesssim \sqrt{\lambda_{N+1}^2 + \lambda_{N+2}^2 + \cdots + \lambda_{\widetilde{N}}^2}.$$

3.2 Tensor Train decomposition

In the following, we will compress the truncated tensor $\mathcal{C}$ further by making use of the TT (tensor train) decomposition introduced by Oseledets [19].

A general tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is said to be in *TT decomposition* if its entries can be expressed in the form

$$\mathcal{X}_{i_1,\dots,i_d} = \sum_{\alpha_0=1}^{r_0} \sum_{\alpha_1=1}^{r_1} \cdots \sum_{\alpha_d=1}^{r_d} \mathcal{G}_1(\alpha_0, i_1, \alpha_1) \mathcal{G}_2(\alpha_1, i_2, \alpha_2) \cdots \mathcal{G}_d(\alpha_{d-1}, i_d, \alpha_d) \tag{15}$$

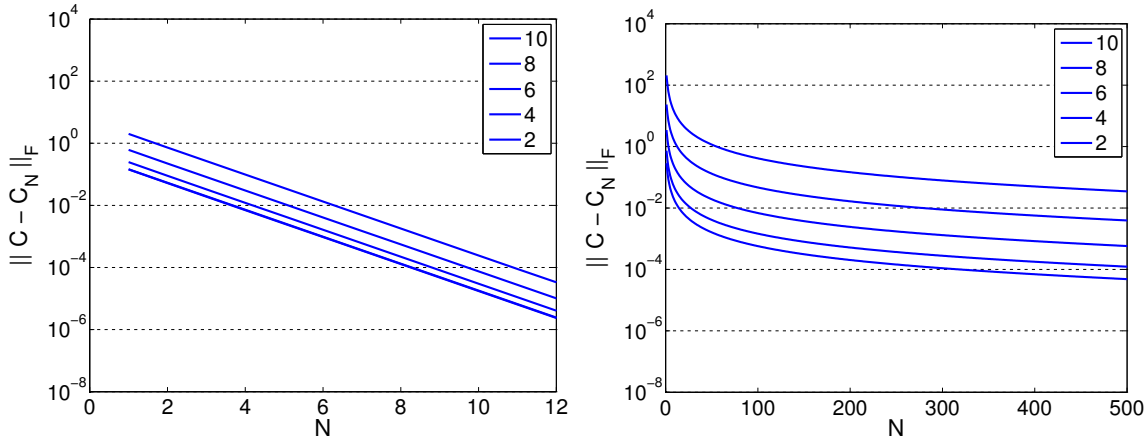


Figure 2: Error $\|\widetilde{\mathcal{C}} - \mathcal{C}\|_F$ for truncating an order d tensor $\widetilde{\mathcal{C}}$, $d = 2, 4, \ldots, 10$. **Left:** KL eigenvalues $\lambda_j = \exp(-j)$. **Right:** KL eigenvalues $\lambda_j = j^{-2}$.

where the tensors $\mathcal{G}_\mu \in \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu}$ for $\mu = 1, \ldots, d$ are the so called *cores* of the TT decomposition. For convenience, we have used MATLAB notation for denoting the entries of the cores.

The minimal $r_0, r_1, \ldots, r_d$ for which (15) holds are called the *TT ranks* of $\mathcal{X}$. Note that we always impose the boundary conditions $r_0 = 1$ and $r_d = 1$ and consequently $\mathcal{G}_1$ and $\mathcal{G}_d$ are actually $n_1 \times r_1$ and $r_{d-1} \times n_d$ matrices, respectively. For moderate TT ranks, the cores of a TT decomposition require much less storage than the full tensor $\mathcal{X}$. For example, if $r_\mu \equiv r$ and $n_\mu \equiv n$ then the storage cost is reduced from $O(n^d)$ down to $O(dnr^2)$.

The TT decomposition is closely connected to the $(1, \ldots, \mu)$-matricization defined as the $n_1 \cdots n_\mu \times n_{\mu+1} \cdots n_d$ matrix with entries

$$X^{(1,\ldots,\mu)}([i_1, \ldots, i_\mu], [i_{\mu+1}, \ldots, i_d]) = \mathcal{X}_{i_1,\ldots,i_d},$$

where $[i_1, \ldots, i_\mu]$ represents the index associated with $(i_1, \ldots, i_\mu)$, see [14] for more details. In particular, the TT rank r_μ is given by the rank of $X^{(1,\ldots,\mu)}$. Moreover, as explained in [19], the singular value decompositions of $X^{(1,\ldots,\mu)}$ for $\mu = 1, \ldots, d-1$ can be used to compute a quasi-optimal approximation of lower TT ranks to $\mathcal{X}$.

3.3 Construction of exact TT decomposition for $\mathcal{C}$

In this section, we will derive a procedure for computing an *exact* TT decomposition (15) of the $N \times \cdots \times N$ tensor $\mathcal{C}$ defined in (12). This will form the basis for the approximate TT decomposition developed in the next section.

As mentioned before, we will only consider the case of even $d = 2k$, for which the representation (7) leads to the following expression for the vectorization of $\mathcal{C}$:

$$\mathrm{vec}(\mathcal{C}) = \sum_{\substack{j_1,\ldots,j_k=1 \\ j_1 \le \cdots \le j_k}}^{N} \prod_{\ell=1}^{N} \lambda_\ell^{m_{\mathbf{j}}(\ell)} \frac{(2m_{\mathbf{j}}(\ell))!}{2^{m_{\mathbf{j}}(\ell)} m_{\mathbf{j}}(\ell)!} \left(\sum_{\mathbf{i} \in P_{\mathbf{j}}} \bigotimes_{\mu=1}^{2k} e_{i_\mu} \right). \tag{16}$$

We will make use of the following fundamental property of TT decompositions [19]. Let $G_\mu \in \mathbb{R}^{r_{\mu-1}N\times r_\mu}$ denote the $(1,2)$-matricizations of the cores $\mathcal{G}_\mu$ for a TT decomposition of $\mathcal{C}$. Then the matrices $U_p \in \mathbb{R}^{N^p\times r_p}$ generated by the recursion

$$U_1 = G_1, \quad U_{\mu+1} = (I_N \otimes U_\mu)G_{\mu+1}, \quad \mu = 1,\dots,d-1 \tag{17}$$

satisfy $\mathcal{R}(U_p) = \mathcal{R}(C^{(1,\dots,p)})$, where $\otimes$ denotes the Kronecker product and $\mathcal{R}$ denotes the range of a matrix.

Our construction of the TT decomposition will go the opposite way. We explicitly construct cores satisfying the relation (17) for some U_μ with $\mathcal{R}(U_\mu) = \mathcal{R}(C^{(1,\dots,\mu)})$. For this purpose, note that (16) immediately implies for $p = 1,\dots,k$ the relations

$$\begin{aligned}\mathcal{R}(C^{(1,\dots,p)}) &= \operatorname{span}\Big\{ \sum_{\mathbf{i}\in P_\mathbf{j}} \bigotimes_{\mu=1}^{p} e_{i_\mu} : \mathbf{j} = (j_1,\dots,j_k) \in \{1,\dots,N\}^k, j_1 \le \cdots \le j_k \Big\} \\ &= \operatorname{span}\Big\{ \sum_{\mathbf{i}\in Q_\mathbf{j}} \bigotimes_{\mu=1}^{p} e_{i_\mu} : \mathbf{j} = (j_1,\dots,j_p) \in \{1,\dots,N\}^p, j_1 \le \cdots \le j_p \Big\}, \end{aligned} \tag{18}$$

where $Q_\mathbf{j}$ is the set of all distinct permutations of $\mathbf{j} = (j_1,\dots,j_p)$. Let us now define the vectors

$$f_\mathbf{j} := \frac{1}{\sqrt{\sharp Q_\mathbf{j}}} \sum_{\mathbf{i}\in Q_\mathbf{j}} \bigotimes_{\mu=1}^{p} e_{i_\mu}. \tag{19}$$

We then define the matrix U_p such that its columns contain the vectors $f_\mathbf{j}$ for all $\mathbf{j} = (j_1,\dots,j_p) \in \{1,\dots,N\}^p$ with $j_1 \le \cdots \le j_p$.

Lemma 3.2. *The matrix U_p defined above is an orthonormal basis of $\mathcal{R}(C^{(1,\dots,p)})$.*

Proof. From (18), it follows that U_p is a basis of $\mathcal{R}(C^{(1,\dots,p)})$. Moreover, the definition (19) implies

$$\langle f_\mathbf{j}, f_{\mathbf{j}'}\rangle = \frac{1}{\sqrt{\sharp Q_\mathbf{j}\cdot\sharp Q_{\mathbf{j}'}}} \sum_{i\in Q_\mathbf{j}} \sum_{i'\in Q_{\mathbf{j}'}} \prod_{\mu=1}^{p} \langle e_{i_\mu}, e_{i'_\mu}\rangle = \begin{cases} 1 & \text{if } \mathbf{j} = \mathbf{j}', \\ 0 & \text{otherwise,} \end{cases}$$

and hence the columns of U_p are orthonormal. □

Example 3.3. *For $d = 4, N = 3$ we have $U_1 = G_1 = (e_1, e_2, e_3)$ and*

$$\begin{aligned} U_2 = \Big(&e_1\otimes e_1, e_2\otimes e_2, e_3\otimes e_3, \dots \\ &(e_1\otimes e_2 + e_2\otimes e_1)/\sqrt{2}, (e_1\otimes e_3 + e_3\otimes e_1)/\sqrt{2}, (e_2\otimes e_3 + e_3\otimes e_2)/\sqrt{2}\Big). \end{aligned}$$

Let the tensor $\mathcal{G}_2 \in \mathbb{R}^{3\times3\times6}$ be defined to have zero entries except for

$$\begin{aligned} &\mathcal{G}_2(1,1,1) = 1, \quad \mathcal{G}_2(2,2,2) = 1, \quad \mathcal{G}_2(3,3,3) = 1, \quad \mathcal{G}_2(1,2,4) = 1/\sqrt{2}, \quad \mathcal{G}_2(2,1,4) = 1/\sqrt{2}, \\ &\mathcal{G}_2(1,3,5) = 1/\sqrt{2}, \quad \mathcal{G}_2(3,1,5) = 1/\sqrt{2}, \quad \mathcal{G}_2(2,3,6) = 1/\sqrt{2}, \quad \mathcal{G}_2(3,2,6) = 1/\sqrt{2}. \end{aligned}$$

Then it can be easily verified that

$$U_2 = (I_3 \otimes U_1)G_2,$$

where $G_2 \in \mathbb{R}^{9\times6}$ is the $(1,2)$-matricization of $\mathcal{G}_2$.

It is straightforward to generalize Example 3.3 and construct core tensors $\mathcal{G}_2, \ldots, \mathcal{G}_k$ that satisfy the recurrence (17). As the tensor $\mathcal{C}$ is supersymmetric, the remaining core tensors $\mathcal{G}_d, \mathcal{G}_{d-1}, \ldots, \mathcal{G}_{k+1}$ are computed by simply permuting the tensors $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_k$, as follows:

$$\mathcal{G}_{d-\mu+1}(j, i, \ell) = \mathcal{G}_\mu(\ell, i, j), \quad \mu = 1, \ldots, k.$$

With this choice of core tensors, the recurrence (17) is satisfied. We will now modify the middle core tensor $\mathcal{G}_k$ such that also the TT decomposition (15) itself is satisfied. The corresponding matricization $C^{(1,\ldots,k)}$ is symmetric and hence the matrix U_k provides an orthonormal basis for the row and column spaces, implying

$$C^{(1,\ldots,k)} = U_k M U_k^T \quad \text{with} \quad M := U_k^T C^{(1,\ldots,k)} U_k.$$

Note that the entries of M can be explicitly computed as follows:

$$M(\widehat{j}, \widehat{t}) = \sqrt{\sharp Q_{\mathbf{j}} \cdot \sharp Q_{\mathbf{t}}}\, \mathcal{C}_{j_1,\ldots,j_k,t_1,\ldots,t_k},$$

where $\widehat{j}$ and $\widehat{t}$ represent the columns of U_k associated with $\mathbf{j} = (j_1, \ldots, j_k)$ and $\mathbf{t} = (t_1, \ldots, t_k)$, respectively. By the recurrence (17), we have $U_k = (I_N \otimes U_{k-1})G_k$. Hence, setting $\overline{G}_k := G_k M$ or, equivalently,

$$\overline{\mathcal{G}}_k(j_{k-1}, i_k, j_k) := \sum_{\ell=1}^{r_k} \mathcal{G}_k(j_{k-1}, i_k, \ell) M(\ell, j_k),$$

yields

$$C^{(1,\ldots,k)} = \overline{U}_k U_k^T.$$

for $\overline{U}_k := (I_N \otimes U_{k-1})\overline{G_k}$. Together with the recurrence (17), this shows that $\mathcal{C}$ admits a TT decomposition with the cores

$$\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_{k-1}, \overline{\mathcal{G}}_k, \mathcal{G}_{k+1}, \ldots, \mathcal{G}_d.$$

A MATLAB implementation of the described procedure for constructing an exact TT decomposition $\mathcal{C}$ is available [24]. The tensor is constructed by the function call `C = constr_tt(d, lambda, 0)`, where `lambda` represents the vector $(\lambda_1, \ldots, \lambda_N)$ and `d` is the order d of $\mathcal{C}$.

Note that the TT rank r_p of $\mathcal{C}$ equals the number of columns of U_p and hence $r_p = \binom{N+p-1}{p}$. This exponential growth of the ranks limits the practicality of the exact TT decomposition and therefore an approximation is required.

3.4 Construction of approximate TT decomposition for $\mathcal{C}$

In the following, we propose a simple modification of the exact construction from the previous section to limit the growth of the TT ranks. As before, we aim to successively construct implicit representations of (approximate) orthonormal bases $\widehat{U}_1, \widehat{U}_2, \ldots, \widehat{U}_k$ for the matricizations of $\mathcal{C}$. However, in each step we will prune columns of $\widehat{U}_\mu$ that only have a negligible impact on the overall approximation quality of the TT decomposition. This will yield a smaller orthonormal basis $\widetilde{U}_\mu$, from which we continue the construction. This idea leads to Algorithm 1, which should be considered as a top level description only. As we will see below, none of the intermediate tensors $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_k$ needs to be constructed explicitly.

Algorithm 1 Construct approximation of $\mathcal{C}$ in TT format

Input: Tensor $\mathcal{C}$ defined as in (12) by $d = 2k$ and $\lambda_1, \ldots, \lambda_N$. Truncation tolerance ϵ.
Output: Tensor $\widehat{\mathcal{C}}$ in TT decomposition, such that $\|\mathcal{C} - \widehat{\mathcal{C}}\|_F \leq \epsilon$.
1: Set $\mathcal{C}_0 = \mathcal{C}$.
2: Set $\widehat{U}_1 = I_N$.
3: **for** $p = 1, \ldots, k$ **do**
4: Choose indices of $\widehat{U}_\mu$ to form $\widetilde{U}_\mu$, s.t. $\|(I - \widetilde{U}_\mu \widetilde{U}_\mu^T) C_{\mu-1}^{(1,\ldots,\mu)}\|_F^2 \leq \epsilon^2/d$.
5: Set $\mathrm{vec}(\mathcal{C}_\mu) := (\widetilde{U}_\mu \widetilde{U}_\mu^T \otimes I_{N^{d-2\mu}} \otimes \widetilde{U}_\mu \widetilde{U}_\mu^T)\mathrm{vec}(\mathcal{C}_{\mu-1})$.
6: Construct $\widehat{U}_{\mu+1}$ by joining columns of $I_N \otimes \widetilde{U}_\mu$ that represent permutations of the same multi-index.
7: **end for**
8: Set $\widehat{\mathcal{C}} = \mathcal{C}_k$.

For the tensor returned by Algorithm 1, it holds that

$$\|\mathcal{C} - \widehat{\mathcal{C}}\|_F^2 = \|\mathcal{C}_0 - \mathcal{C}_k\|_F^2 = \|\mathcal{C}_0 - \mathcal{C}_1 + \mathcal{C}_1 - \mathcal{C}_2 + \cdots + \mathcal{C}_{k-1} - \mathcal{C}_k\|_F^2 = \sum_{\mu=1}^{k} \|\mathcal{C}_{\mu-1} - \mathcal{C}_\mu\|_F^2,$$

where the last equality follows from the fact that, for all $1 \leq \mu < \nu \leq k$, the entries of the tensor $\mathcal{C}_{\nu-1} - \mathcal{C}_\nu$ are zero at the positions of the nonzero entries of $\mathcal{C}_{\mu-1} - \mathcal{C}_\mu$ and hence these differences are orthogonal to each other. Moreover, the following bound holds for $\mu = 1, \ldots, k$:

$$\begin{aligned}
\|\mathcal{C}_{\mu-1} - \mathcal{C}_\mu\|_F^2 &= \|(I_{N^d} - \widetilde{U}_\mu \widetilde{U}_\mu^T \otimes I_{N^{d-2\mu}} \otimes \widetilde{U}_\mu \widetilde{U}_\mu^T)\mathrm{vec}(\mathcal{C}_{\mu-1})\|_2^2 \\
&\leq 2\|(I_{N^d} - I_{N^{d-\mu}} \otimes \widetilde{U}_\mu \widetilde{U}_\mu^T)\mathrm{vec}(\mathcal{C}_{\mu-1})\|_2^2 \\
&= 2\|(I_{N^\mu} - \widetilde{U}_\mu \widetilde{U}_\mu^T) C_{\mu-1}^{(1,\ldots,\mu)}\|_F^2
\end{aligned}$$

Thus, to ensure $\|\mathcal{C} - \widehat{\mathcal{C}}\|_F \leq \epsilon$, it is sufficient to require that

$$\left\|(I_{N^\mu} - \widetilde{U}_\mu \widetilde{U}_\mu^T) C_{\mu-1}^{(1,\ldots,\mu)}\right\|_F^2 \leq \frac{\epsilon^2}{d}, \tag{20}$$

which coincides with the criterion in Line 4 of Algorithm 1.

We now discuss the efficient evaluation of the criterion (20). Let $\Omega_\mu \subset \{\mathbf{j} : j_1 < \cdots < j_\mu\}$ contain the multi-indices corresponding to the columns that have been pruned from $\widehat{U}_\mu$ to yield the reduced basis $\widetilde{U}_\mu$. Then

$$\left\|(I_{N^\mu} - \widetilde{U}_\mu \widetilde{U}_\mu^T) C_{\mu-1}^{(1,\ldots,\mu)}\right\|_F^2 = \sum_{\mathbf{j} \in \Omega_\mu} \sharp Q_{\mathbf{j}} \cdot \|\widehat{U}_\mu(:, \widehat{j})^T C_{\mu-1}^{(1,\ldots,\mu)}\|_F^2 = \sum_{\mathbf{j} \in \Omega_\mu} \sharp Q_{\mathbf{j}} \cdot \|C_{j_1,\ldots,j_p,:,\ldots,:}\|_F^2,$$

where $\widehat{U}_\mu(:, \widehat{j})$ represents the column of $\widehat{U}_\mu$ associated with $\mathbf{j} = (j_1, \ldots, j_p)$. Based on this formula, we compute each error $\sharp Q_{\mathbf{j}} \cdot \|C_{j_1,\ldots,j_p,:,\ldots,:}\|_F^2$ entailed by pruning the corresponding column of $\widehat{U}_\mu$, and remove as many columns as possible such that the sum of the errors does not exceed ϵ^2/d.

A variation of Lemma 3.1 can be used to perform the evaluation of $\|C_{j_1,\ldots,j_\mu,:,\ldots,:}\|_F$ very efficiently.

Lemma 3.4. *Consider the d-th order tensor* $\mathcal{C} \in \mathbb{R}^{N\times\cdots\times N}$ *defined in* (12) *and a multi-index* $\mathbf{t} = (t_1, \ldots, t_p) \in \{1, \ldots, N\}^p$. *Then*

$$\|\mathcal{C}_{t_1,\ldots,t_p,:,\ldots,:}\|_F^2 = (d-p)!(\widetilde{a}^{(1)} * \widetilde{a}^{(2)} * \cdots * \widetilde{a}^{(N)})_d,$$

where the second factor is the d-th component of the discrete convolution $*$ *of the vectors* $\widetilde{a}^{(j)} \in \ell^2(\mathbb{Z})$, $j = 1, \ldots, N$, *defined as*

$$\widetilde{a}_m^{(j)} = \begin{cases} \frac{\lambda_j^m \gamma_m^2}{(m-m_{\mathbf{t}}(j))!} & \text{if } m \in \{m_{\mathbf{t}}(j), m_{\mathbf{t}}(j)+1, \ldots, d\}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Analogously to the proof of Lemma 3.1, it can be shown that

$$\|\mathcal{C}_{t_1,\ldots,t_p,:,\ldots,:}\|_F^2 = (d-p)! \sum_{\mathbf{i}\in\Sigma} \prod_{j=1}^{N} \frac{\lambda_j^{m_{\mathbf{i}}(j)} \gamma_{m_{\mathbf{i}}(j)}^2}{(m_{\mathbf{i}}(j) - m_{\mathbf{t}}(j))!},$$

where

$$\Sigma := \big\{(t_1, \ldots, t_p, s_{p+1}, \ldots, s_d) : s_j \in \{1, \ldots, N\}\ \forall j = p+1, \ldots, d \text{ and } s_{p+1} \le \cdots \le s_d\big\}.$$

Note that only the multiplicities $m_{\mathbf{i}}(j)$, $j = 1, \ldots, N$ are used inside the sum, and that every d-tuple in Σ is uniquely associated with an N-tuple $(m_1, \ldots, m_N) \in \widetilde{\Sigma}$:

$$\widetilde{\Sigma} := \big\{(m_1, \ldots, m_N) \in \mathbb{N}_0^d : m_1 + m_2 + \cdots + m_N = d \text{ and } m_j \ge m_{\mathbf{t}}(j)\ \forall j = 1, \ldots, N\big\}. \quad (21)$$

Replacing the sum over Σ by a sum over $\widetilde{\Sigma}$ leads to

$$\|\mathcal{C}_{t_1,\ldots,t_p,:,\ldots,:}\|_F^2 = (d-p)! \sum_{\mathbf{m}\in\widetilde{\Sigma}} \prod_{j=1}^{N} a_{m_j}^{(j)},$$

with $a_m^{(j)} = \frac{\lambda_j^m \gamma_m^2}{(m-m_{\mathbf{t}}(j))!}$. The condition $m_j \ge m_{\mathbf{t}}(j)$ in $\widetilde{\Sigma}$ can be enforced by replacing $a_{m_j}^{(j)}$ by $\widetilde{a}_{m_j}^{(j)}$, with $\widetilde{a}_{m_j}^{(j)} = 0$ for $m_j < m_{\mathbf{t}}(j)$. As described in the proof of Lemma 3.1, the resulting summation can be expressed by repeated application of convolutions. □

Once Algorithm 1 has computed the reduced bases $\widehat{U}_1, \ldots, \widehat{U}_k$, the core tensors of the TT decomposition for $\widehat{\mathcal{C}}$ can be computed analogous to the procedure discussed in Section 3.3. In fact, none of these bases needs to be formed explicitly; it is sufficient to keep track of the multi-indices corresponding to the pruned column indices. For the technical details, we refer to the MATLAB implementation [24]. The approximate tensor $\widehat{\mathcal{C}}$ is constructed by the function call `C_hat = constr_tt(d, lambda, epsilon)`, where `lambda` is the vector $(\lambda_1, \ldots, \lambda_N)$, `d` is the order d of $\mathcal{C}$, and `epsilon` is the maximally allowed error ϵ.

Finally, we remark that the tensor $\widehat{\mathcal{C}}$ returned by Algorithm 1 is not necessarily supersymmetric, as the pruning of the columns does, in general, not preserve all permutations of a given multi-index.

4 Numerical Experiments

In this section, we assess the performance of the TT decomposition for approximating d-point correlation functions. For this purpose, we apply Algorithm 1 with a prescribed tolerance ϵ. This is followed by the higher-order SVD, in order to compress the tensor further within the same tolerance. We investigate numerically how the obtained TT ranks grow as a function of ϵ, depending on the dimension d and the spatial smoothness of the random field.

In all examples, we consider a one-dimensional centered Gaussian random field with unit variance on the interval $[0, 1]$. The random field has been discretized on a uniform grid of grid-size N_h, where N_h has always been chosen sufficiently large to not affect the obtained ranks. In all cases, the middle TT rank $r_{d/2}$ of the TT tensor turned out to be the largest one. Since the CP rank of a tensor represents an upper bound for each TT rank, it makes sense to compare $r_{d/2}$ with the CP rank M predicted by the bounds (10) and (11) to attain the same accuracy ϵ. For convenience, we recall the asymptotics of these bounds:

$$\begin{aligned} &\text{finite regularity } \textstyle\sum_i \lambda_i^p < +\infty : && \epsilon \sim C M^{\frac{1}{2}-\frac{1}{p}}, \\ &\text{exponential decay } \lambda_j \sim e^{-sj} : && \epsilon \sim C \exp\{-s 2^{-\frac{d+1}{d}} M^{\frac{2}{d}}\}. \end{aligned}$$

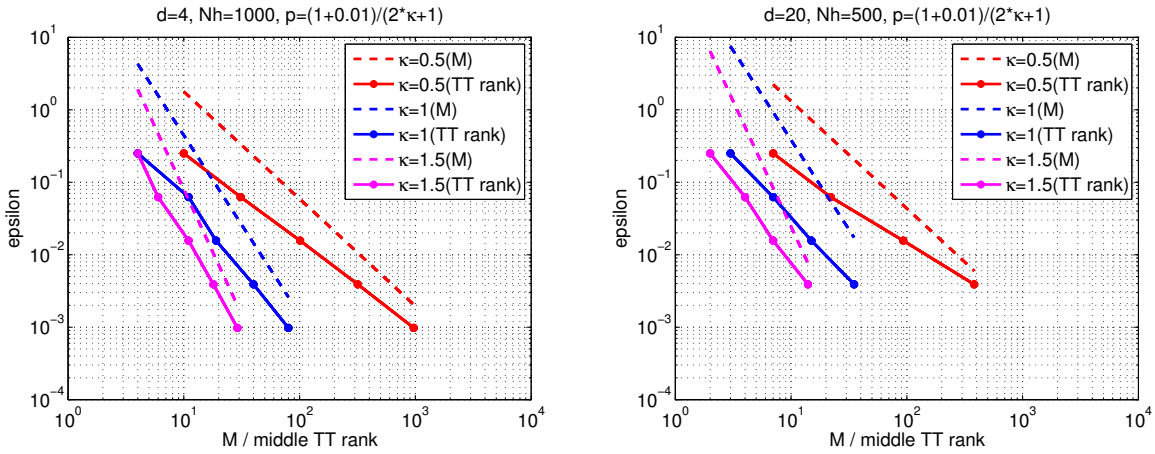
It is important to remark that these bounds feature constants C that are potentially very large and have not been taken into account in our plots.

4.1 Matérn covariance

We use the Matérn covariance (3) with the parameters $\kappa \in \{0.5, 1, 1.5\}$ and $l_c = 0.25$. The KL eigenvalues are computed using linear finite elements on a grid with $N_h + 1$ equidistant points. These eigenvalues are then used to construct the approximation of d-point correlation functions of f, using the KL expansion and the approximate TT decomposition with given accuracy ϵ. We analyze the d-point correlations for various d by taking a sequence of decreasing tolerances $\epsilon_1 > \epsilon_2 > \cdots$.

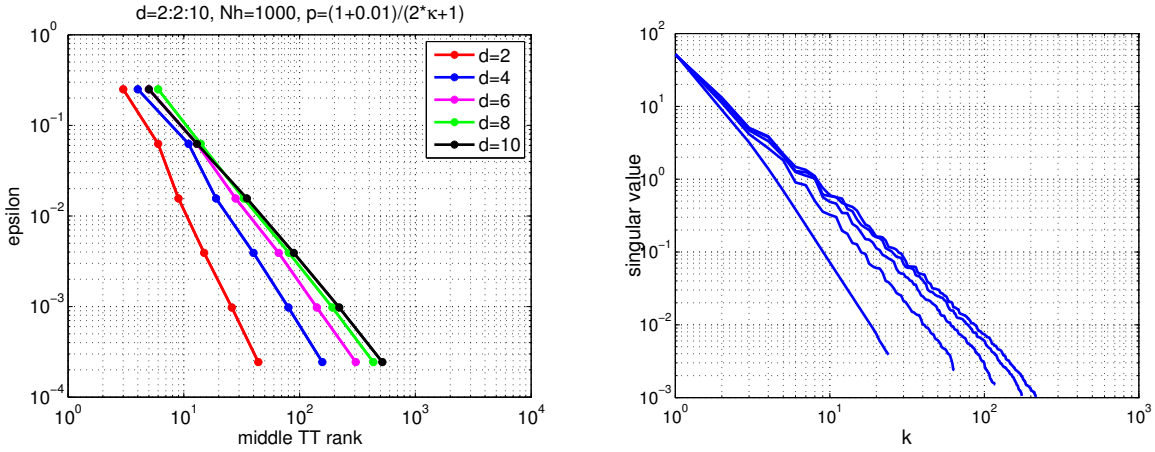
Figure 3 shows the resulting middle TT ranks $r_{d/2}$ (continuous line) and the theoretical bound M (dashed line) as a function of ϵ for $d = 4$ and $d = 20$. Since the eigenvalues of the covariance operator decay as $\lambda_j \sim j^{-(2\kappa+1)}$, we use $p = 1.01/(2\kappa + 1)$ in the theoretical bound, which makes the sum in (9) bounded. For larger κ, the slopes of the observed accuracy do not match the predicted asymptotic rate, due to the large constants involved in the theoretical bounds. This becomes even more apparent in the left plot of Figure 4, which compares the results for $d = 2, 4, 6, 8, 10$. Although the theoretically predicted rate does not depend on d, the observed maximum TT-rank shows a slight deterioration with respect to the dimension d.

The right plot of Figure 4 displays the singular values of all matricizations of the 10-point correlation of f. Note that – due to symmetry – there are only 5 matricizations with different singular values. The (1)-matricization shows the most favorable behavior; its singular values decay with the rate $2\kappa + 1$, consistent with the decay of the KL eigenvalues. The $(1, \ldots, d/2)$-matricization, which determines the middle TT rank, shows the worst behavior. We refer to [21] for a theoretical explanation of this phenomenon.



(a) $d = 4$ (b) $d = 20$

Figure 3: Matérn covariance with $\kappa \in \{0.5, 1, 1.5\}$ and $l_c = 0.25$. Prescribed tolerance ϵ vs. middle TT rank of the TT approximation for the d point correlation, compared to the asymptotic rate (M) predicted by (10) with $p = 1.01/(2\kappa + 1)$.

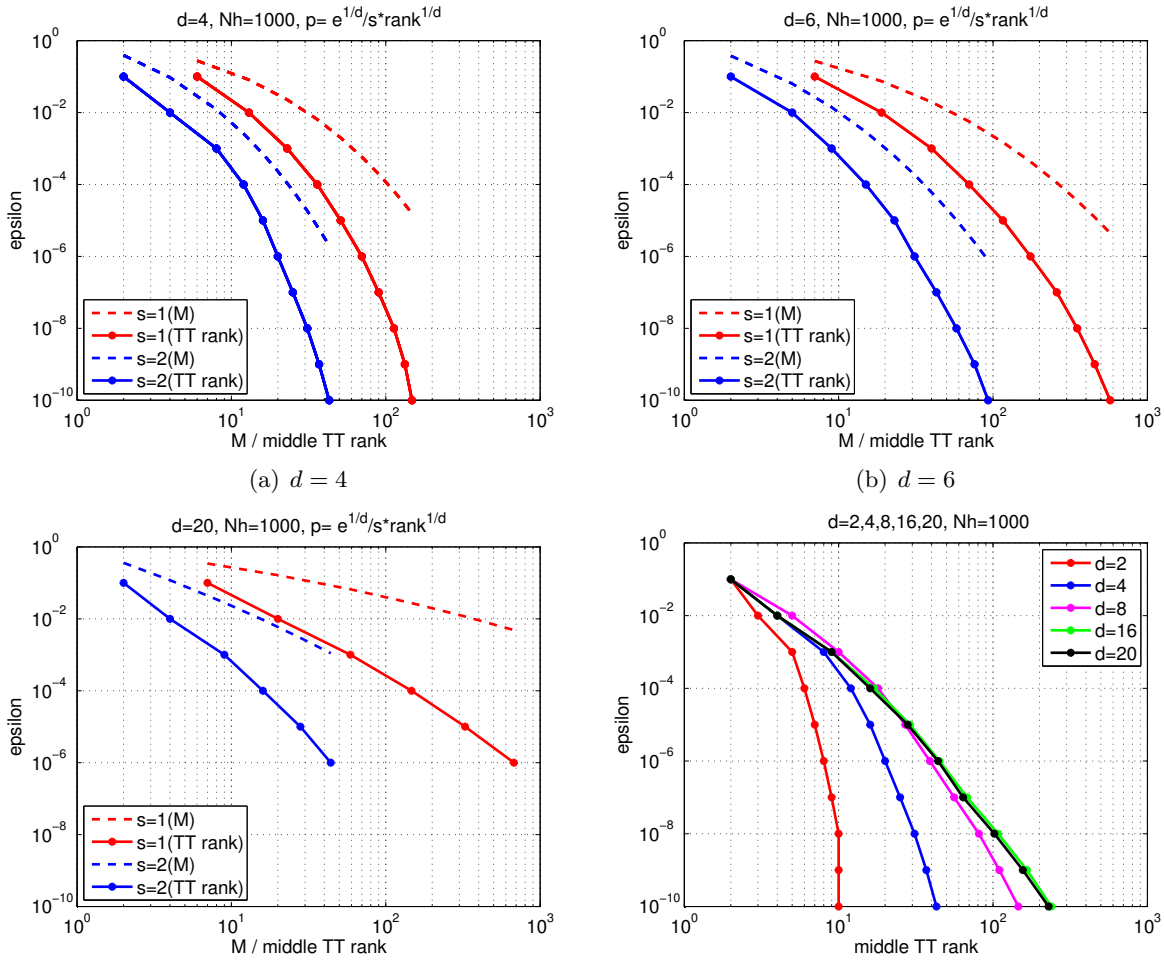


(a) d=2,4,6,8,10 (b) d=10

Figure 4: Matérn covariance with $\kappa = 1$ and $l_c = 0.25$. (a) Prescribed tolerance ϵ vs. middle TT rank of the TT approximation for the d point correlation. (b) Singular values for each matricization of the $d = 10$ point correlation.

4.2 Exponentially decaying KL eigenvalues

We repeat the experiments from the previous section for exponentially decaying KL eigenvalues. More specifically, $\lambda_k = \exp(-ks)$ with $s = 1$ and $s = 2$ is considered. Again, the theoretical bound is compared with the maximal rank of the TT approximation for 4, 6, 8, and 20 point correlation functions of f using a varying tolerance ϵ in the TT algorithm. The obtained results are displayed in Figure 5. Again, there is a slight mismatch with the theoretically predicted asymptotic rate, but this time the observed convergence of the TT approximation is actually better.



(a) $d = 4$ (b) $d = 6$

(c) $d = 20$ (d) $d = 2, 4, 8, 16, 20$

Figure 5: Exponential decay $\lambda_j = e^{-js}$ of KL eigenvalues for $s \in \{1, 2\}$. Prescribed tolerance ϵ vs. middle TT rank of the TT approximation for the d point correlation. (a)–(c) Comparison to the asymptotic rate (M) predicted by (11). (d) Comparison for different d using $s = 2$.

5 Application to linear elliptic PDEs with random forcing term

In this section, we illustrate how the techniques from this paper can be applied to study stochastic linear elliptic equations of the form

$$-\mathrm{div}(a(\cdot)\nabla u(\cdot,\omega)) = f(\cdot,\omega) \quad \text{on } D, \quad \text{for a.e. } \omega \in \Omega \tag{22}$$

with boundary condition $u(\cdot,\omega) = 0$ on ∂D.

In the following, we give a very brief description of the discretization of (22) and the resulting equation for the (discretized) correlations of u. We refer to, e.g., [23, 26] for more details. A standard finite element (FE) discretization of (22) leads to a linear system

$$Au(\omega) = f(\omega),$$

where A is an $\widetilde{N} \times \widetilde{N}$ symmetric positive definite matrix and $\widetilde{N}$ equals the dimension of the FE space. To simplify the notation, we again use u, f for expressing the discretizations of the corresponding functions.

Using the fact that the expectation $\mathbb{E}(u)$ satisfies the mean field equation $A\mathbb{E}(u) = \mathbb{E}(f)$, the d-point correlation $\mu_u^d \in \mathbb{R}^{\widetilde{N}\times\cdots\times\cdots\widetilde{N}}$ of u is related to the d-point correlation $\mu_f^d \in \mathbb{R}^{\widetilde{N}\times\cdots\times\cdots\widetilde{N}}$ of f via the linear system

$$(\underbrace{A \otimes A \otimes \ldots \otimes A}_{d-\text{times}})\mathrm{vec}(\mu_u^d) = \mathrm{vec}(\mu_f^d). \tag{23}$$

where $\otimes$ again denotes the Kronecker product.

5.1 Tensor techniques

As explained in Section 2, a truncated KL expansion allows us to (approximately) represent the d-point correlation of f as

$$\mathrm{vec}(\mu_f^d) = (\Phi \otimes \cdots \otimes \Phi)\mathrm{vec}(\mathcal{C}_f),$$

where $\Phi \in \mathbb{R}^{\widetilde{N}\times N}$ contains the orthonormal basis of retained KL eigenfunctions and $\mathcal{C}_f$ is the tensor defined as in (16). By (23), this implies that the d-point correlation of u is given by

$$\mathrm{vec}(\mu_u^d) = (A^{-1}\Phi \otimes \cdots \otimes A^{-1}\Phi)\mathrm{vec}(\mathcal{C}_f).$$

Let $A^{-1}\Phi = \Phi_u R_u$ be a QR decomposition, that is, Φ_u is again an orthonormal basis and R_u is an upper triangular matrix. Then

$$\mathrm{vec}(\mu_u^d) = (\Phi_u \otimes \cdots \otimes \Phi_u)\underbrace{(R_u \otimes \cdots \otimes R_u)\mathrm{vec}(\mathcal{C}_f)}_{=:\mathrm{vec}(\mathcal{C}_u)}. \tag{24}$$

The tensor $\mathcal{C}_u$ is obtained from the tensor $\mathcal{C}_f$ by multiplying each of its modes with R_u.

Suppose that $\widehat{\mathcal{C}}_f$ is the tensor in TT decomposition obtained from Algorithm 1, such that $\|\mathcal{C}_f - \widehat{\mathcal{C}}_f\|_F \leq \epsilon$. Then an approximation $\widehat{\mathcal{C}}_u$ for $\mathcal{C}_u$ is obtained by multiplying each mode of $\widehat{\mathcal{C}}_f$ with R_u, which can be performed very cheaply for a tensor in TT decomposition. The orthonormality of Φ and Φ_u then imply the error bound

$$\|\widehat{\mathcal{C}}_u - \mathcal{C}_u\|_F = \big\|(R_u \otimes \cdots \otimes R_u)\big(\mathrm{vec}(\widehat{\mathcal{C}}_f) - \mathrm{vec}(\mathcal{C}_f)\big)\big\|_2 \leq \|A^{-1}\|_2^d \cdot \epsilon.$$

Note that $\|A^{-1}\|_2$ is bounded from above uniformly with respect to the dimension $\widetilde{N}$ of the FE space. Note also that $\widehat{\mathcal{C}}_u$ and $\widehat{\mathcal{C}}_f$ have identical TT ranks but, as we will see below, it is often possible to recompress $\widehat{\mathcal{C}}_u$ to lower rank without significantly increasing the error.

5.2 Example: One-dimensional PDE

First, we consider an example that admits a direct approximation of μ_u^d. For $D = (0, \pi)$ we obtain the boundary value problem

$$-u''(x,\omega) = f(x,\omega), \quad x \in (0,\pi), \quad \omega \in \Omega, \tag{25}$$

with zero Dirichlet boundary conditions. Assuming that $f(x,\omega) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} Y_j(\omega) \sin(jx)$, it is natural to use a spectral discretization of (25) with the basis $\sin(x), \sin(2x), \ldots, \sin(\widetilde{N}x)$. This yields the linear system

$$A\mathbf{u}(\omega) = \mathbf{f}(\omega), \qquad A = \operatorname{diag}(-1^2, -2^2, \ldots, -\widetilde{N}^2).$$

Since the matrix A is diagonal, the KL expansion of $u(x,\omega)$ is known a priori for this case:

$$u(x,\omega) = -\sum_{j=1}^{\infty} j^{-2} \sqrt{\lambda_j} Y_j(\omega) \sin(jx).$$

Using the method described in Section 3.2, this gives us the possibility to directly compute low-rank approximations $\widehat{\mathcal{C}}_f$ and $\widehat{\mathcal{C}}_u$ corresponding to the correlations μ_f^d and μ_u^d, respectively.

To study the approximability of the correlations, let us consider the dominant singular values of the different matricizations for $\widehat{\mathcal{C}}_f$ and $\widehat{\mathcal{C}}_u$ for the cases $\lambda_j = \exp(-j)$ for $d = 14$ (Figure 6) and $\lambda_j = j^{-3}$ for $d = 8$ (Figure 7). In both cases, the singular values are observed to decay much faster for $\widehat{\mathcal{C}}_u$, giving the possibility to approximate d-point correlations of u with TT decompositions of comparably low rank.

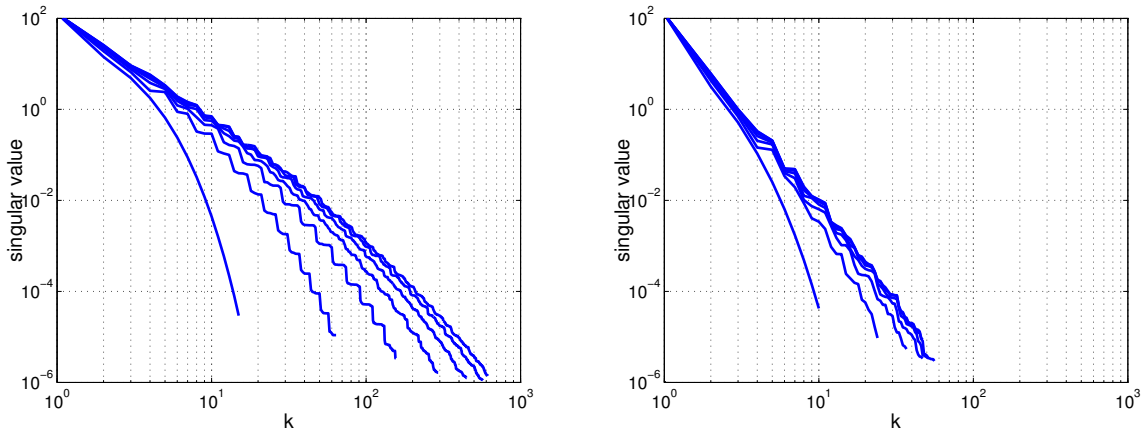
5.3 Example: Two-dimensional PDE

Finally, we consider the PDE (22) with a two-dimensional domain of the form $D = (0,\pi) \times (0,\pi)$. The equation is discretized using piecewise linear FEs on an unstructured mesh with 1932 vertices. We consider a tensorized random field $f(x,y,\omega) = f_1(x,\omega) f_2(y,\omega)$ with f_1,f_2 independend Gaussian random fields with Matérn covariance function, for two different values of $\kappa = 0.5, \kappa = 1$ together with $l_c = 0.35$ and $l_c = 0.25$, respectively.

Using the method described in Section 5.1, we compute a low-rank approximation $\widehat{\mathcal{C}}_u$ to the d-point correlation for the solution u of the discretized PDE (22). Figure 8 shows the middle rank of the TT decomposition of $\widehat{\mathcal{C}}_u$ for different prescribed tolerances ϵ. Once again, the growth of the middle rank is rather mild as d increases. As can be seen in Figure 9 for $\kappa = 0.5$, the singular values decay much faster for u than for f, confirming our observation for the 1D-problem. The plots for $\kappa = 1$ look similar and yield the same conclusion.

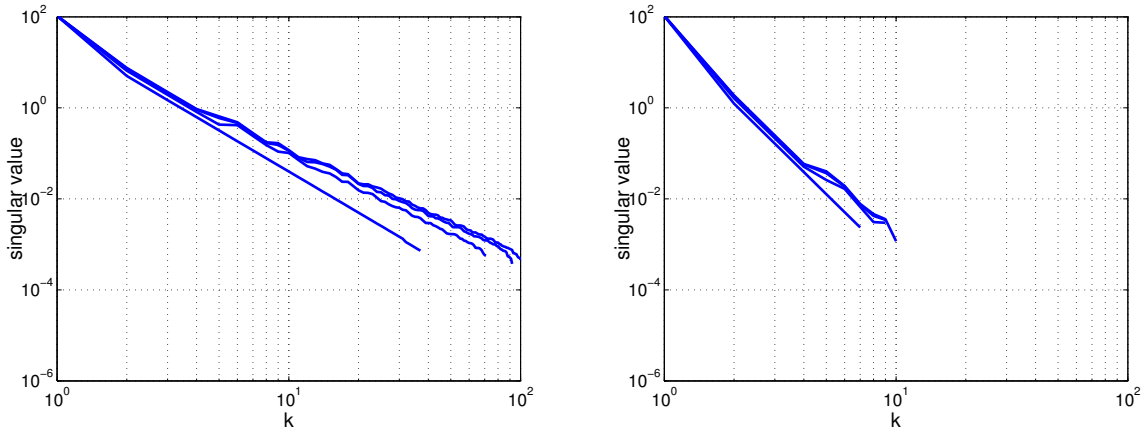
6 Conclusions

The combination of truncated Karhunen-Loève expansion with low-rank tensor techniques is an effective mean to represent d-point correlation functions of Gaussian random fields. The error



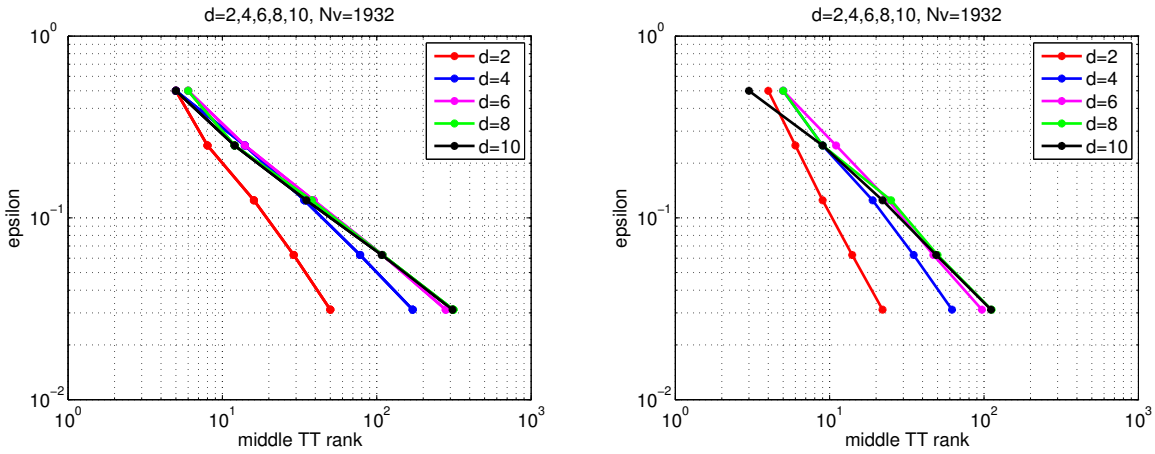
(a) Correlation of f (b) Correlation of u

Figure 6: 1-D problem (25) with exponential KL eigenvalue decay $\lambda_j = \exp(-j)$: Singular values for all matricizations of the tensors $\widehat{\mathcal{C}}_f$ and $\widehat{\mathcal{C}}_u$ corresponding to $d = 14$ point correlations of f and u, respectively.



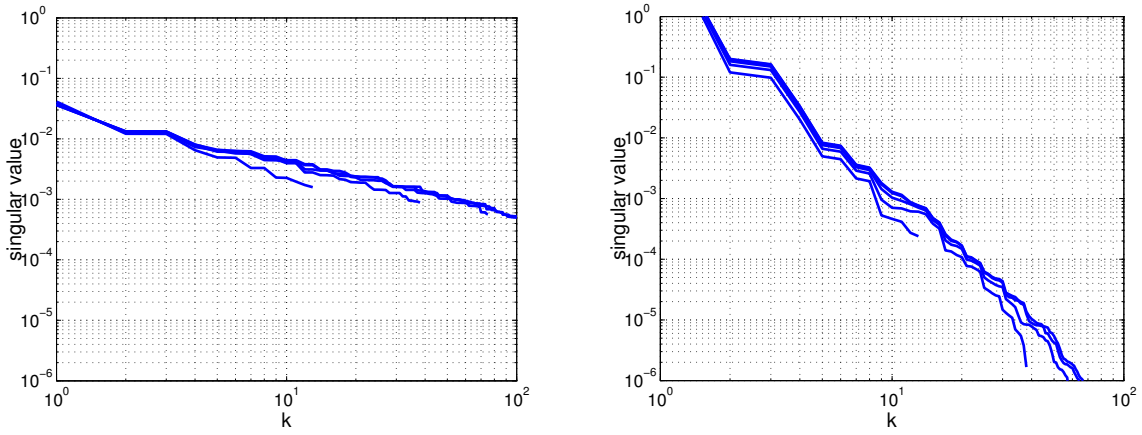
(a) Correlation of f (b) Correlation of u

Figure 7: 1-D problem (25) with polynomial KL eigenvalue decay $\lambda_j = j^{-3}$: Singular values for all matricizations of tensors $\widehat{\mathcal{C}}_f$ and $\widehat{\mathcal{C}}_u$ corresponding to $d = 8$ point correlations of f and u, respectively.



(a) $\kappa = 0.5, l_c = 0.50$ (b) $\kappa = 1, l_c = 0.25$

Figure 8: 2D-Problem with Matérn covariance for two different sets of parameters: Prescribed tolerance ϵ vs. middle TT rank of the TT approximation for the d point correlation of the solution u.



(a) Correlation of f (b) Correlation of u

Figure 9: 2D-Problem with Matérn covariance for $\kappa = 0.5, l_c = 0.35$: Singular values for all matricizations of the tensors $\widehat{\mathcal{C}}_f$ and $\widehat{\mathcal{C}}_u$ corresponding to $d = 10$ point correlations of f and u, respectively.

resulting from the truncation of the expansion admits an analysis that matches the numerical observations quite well. In contrast, it is surprising and not predicted by existing bounds [21] that the ranks of the involved TT decompositions depend only rather mildly on d. Together with Algorithm 1, this allows us to conveniently handle orders as high as $d = 20$.

To illustrate how our construction can be used in the solution of stochastic PDEs, we have restricted ourselves to the comparably simple case of random forcing terms. However, our work has already been used to cover more general situations [3, 5].

References

[1] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007.

[2] J. Beck, F. Nobile, L. Tamellini, and R. Tempone. Convergence of quasi-optimal stochastic Galerkin methods for a class of PDEs with random coefficients. *Comput. Math. Appl.*, 67(4):732–751, 2014.

[3] F. Bonizzoni. *Analysis and approximation of moment equations for PDEs with stochastic data*. PhD thesis, Department of Mathematics, Politecnico di Milano, Italy, 2013.

[4] F. Bonizzoni, A. Buffa, and F. Nobile. Moment equations for the mixed formulation of the Hodge Laplacian with stochastic data. *IMA Journal of Numerical Analysis*, 2013. in press.

[5] F. Bonizzoni, F. Nobile, and D. Kressner. Tensor Train approximation of moment equations for the log-normal Darcy problem. Technical report, MATHICSE, EPFL, 2014.

[6] A. Chernov and Ch. Schwab. First order k-th moment finite element analysis of nonlinear operator equations with stochastic data. *Math. Comp.*, 82(284):1859–1888, 2013.

[7] A. Cohen, R. DeVore, and Ch. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. *Anal. Appl.*, 9(1):11–47, 2011.

[8] P. J. Diggle and P. J. Ribeiro. *Model-based geostatistics*. Springer Series in Statistics. Springer, New York, 2007.

[9] I. Graham, F. Kuo, J. Nichols, R. Scheichl, C. Schwab, and I. Sloan. Quasi-Monte Carlo finite element methods for elliptic PDEs with log-normal random coefficient. SAM Research Report 2013–14, ETH Zurich, 2013.

[10] H. Harbrecht, M. Peters, and R. Schneider. On the low-rank approximation by the pivoted Cholesky decomposition. *Appl. Numer. Math.*, 62(4):428–440, 2012.

[11] H. Harbrecht, R. Schneider, and Ch. Schwab. Multilevel frames for sparse tensor product spaces. *Numer. Math.*, 110(2):199–220, 2008.

[12] H. Harbrecht, R. Schneider, and Ch. Schwab. Sparse second moment analysis for elliptic problems in stochastic domains. *Numer. Math.*, 109(3):385–414, 2008.

[13] R. Hiptmair, C. Jerez-Hanckes, and Ch. Schwab. Sparse tensor edge elements. *BIT*, 53(4):925–939, 2013.

[14] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[15] M. Loève. *Probability theory. I.* Springer-Verlag New York, USA, 4th ed. edition, 1977.

[16] M. Loève. *Probability theory. II.* Springer-Verlag New York, USA, 4th ed. edition, 1978.

[17] I. V. Oseledets and S. V. Dolgov. Solution of linear systems and matrix inversion in the TT-format. *SIAM J. Sci. Comput.*, 34(5):A2718–A2739, 2012.

[18] I. V. Oseledets and E. E. Tyrtyshnikov. TT-cross approximation for multidimensional arrays. *Linear Algebra Appl.*, 432(1):70–88, 2010.

[19] I.V. Oseledets. Tensor train decomposition. *SIAM J. Sc. Comp.*, 33:2295–2317, 2011.

[20] D. V. Savostyanov and I. V. Oseledets. Fast adaptive interpolation of multi-dimensional arrays in tensor train format. In *Proceedings of 7th International Workshop on Multidimensional Systems (nDS)*. IEEE, 2011.

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

[22] C. Schwab and R.A. Todor. Sparse finite elements for stochastic elliptic problems—higher order moments. *Computing*, 71(1):43–63, 2003.

[23] Ch. Schwab and R. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numerische Mathematik*, 95:707–734, 2003.

[24] C. Tobler. MATLAB software package `constr_tt`, 2013. Available from `http://anchp.epfl.ch/software/misc`.

[25] R. A. Todor. *Sparse Perturbation Algorithms for Elliptic PDEs with Stochastic Data*. PhD thesis, ETH Zurich, Switzerland, 2005.

[26] T. von Petersdorff and C. Schwab. Sparse wavelet methods for operator equations with stochastic data. *Appl. Math.*, 51:145–180, 2006.

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

10.2014 NATHAN COLLIER, ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, ERIK VON SCHWERIN, RAÚL TEMPONE:
A continuation multilevel Monte Carlo algorithm

11.2014 LUKA GRUBISIC, DANIEL KRESSNER:
On the eigenvalue decay of solutions to operator Lyapunov equations

12.2014 FABIO NOBILE, LORENZO TAMELLINI, RAÚL TEMPONE:
Convergence of quasi-optimal sparse grid approximation of Hilbert-valued functions: application to random elliptic PDEs

13.2014 REINHOLD SCHNEIDER, ANDRÉ USCHMAJEW:
Convergence results for projected line-search methods on varieties of low-rank matrices via Lojasiewicz inequality

14.2014 DANIEL KRESSNER, PETAR SIRKOVIC:
Greedy law-rank methods for solving general linear matrix equations

15.2014 WISSAM HASSAN, MARCO PICASSO:
An anisotropic adaptive finite element algorithm for transonic viscous flows around a wing

16.2014 ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, ERIK VON SCHWERIN, RAÚL TEMPONE:
Optimization of mesh hierarchies in multilevel Monte Carlo samplers

17.2014 DANIEL KRESSNER, FRANCISCO MACEDO:
Low-rank tensor methods for communicating Markov processes

18.2014 ASSYR ABDULLE, GILLES VILMART, KONSTANTINOS C. ZYGALAKIS:
Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics

19.2014 ELEONORA MUSHARBASH, FABIO NOBILE, TAO ZHOU:
Con the dynamically orthogonal approximation of time dependent random PDEs

20.2014 FROILÁN M. DOPICO, JAVIER GONZÁLEZ, DANIEL KRESSNER, VALERIA SIMONCINI:
Projection methods for large T-Sylvester equations

21.2014 LUCA DEDÈ, CHRISTOPH JÄGGLI, ALFIO QUARTERONI:
Isogeometric numerical dispersion analysis for elastic wave propagation

22.2014 MARCO DISCACCIATI, PAOLA GERVASIO, ALFIO QUARTERONI:
Interface control domain decomposition (ICDD) method for Stokes-Darcy coupling

23.2014 FRANCESCO BALLLARIN, ANDREA MANZONI, ALFIO QUARTERONI, GIANLUIGI ROZZA:
Supremizer stabilization of POD-Galerkin approximation of parametrized Navier-Stokes equations

24.2014 DANIEL KRESSNER, RAJESH KUMAR, FABIO NOBILE, CHRISTINE TOBLER:
Low-rank tensor approximation for high-order correlation functions of Gaussian random fields