Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Tensor approach to optimal control problems with fractional d-dimensional elliptic operator in constraints

by

Gennadij Heidel, Venera Khoromskaia, Boris N. Khoromskij, and Volker Schulz

Preprint no.: 81 2018

Tensor approach to optimal control problems with fractional d-dimensional elliptic operator in constraints

Gennadij Heidel* Venera Khoromskaia** Boris N. Khoromskij◇
Volker Schulz§

Abstract

We introduce the tensor numerical method for solution of the d-dimensional optimal control problems with fractional Laplacian type operators in constraints discretized on large spacial grids. It is based on the rank-structured approximation of the matrix valued functions of the corresponding fractional elliptic operator. The functions of finite element (finite difference) Laplacian on a tensor grid are diagonalized by using the fast Fourier transform (FFT) matrix and then the low rank tensor approximation to the multi-dimensional core diagonal tensor is computed. The existence of low rank canonical approximation to the class of matrix valued functions of the fractional Laplacian is proved based on the sinc quadrature approximation method applied to the integral transform of the generating function. The equation for the control function is solved by the PCG method with the rank truncation at each iteration step where the low Kronecker rank preconditioner is precomputed by using the canonical decomposition of the core tensor for the inverse of system matrix. The right-hand side, the solution, and the governing operator are maintained in the rank-structured tensor format. Numerical tests for the 2D and 3D control problems confirm the linear complexity scaling of the method in the univariate grid size.

1 Introduction

Optimization problems with partial differential equations (PDEs) as constraints are well known in the mathematical literature. They admit a number of applications in various fields of natural science and have been studied for many years, see [35] for comprehensive presentation. Of particular interest are *tracking-type* problems, where the objective functional is the deviation of the PDE solution from a given target function in some norm.

*University of Trier, FB 4 - Department of Mathematics, D-54286, Trier, Germany (heidel@uni-trier.de)

**Max Planck Institute for Mathematics in the Sciences, Leipzig (vekh@mis.mpg.de); Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg.

◇Max-Planck-Institute for Mathematics in the Sciences, Inselstr. 22-26, D-04103 Leipzig, Germany (bokh@mis.mpg.de).

§University of Trier, FB 4 - Department of Mathematics, D-54286, Trier, Germany (Volker.Schulz@uni-trier.de).

Optimal control problems pose a major challenge from a computational point of view due to the complexity of constraint: evaluating the constraint requires the solution of a partial differential equation. Therefore, to make these problems tractable, specially tailored solvers are required. In the classical sense, partial differential equations are given by local operators, i. e. only local information is required to evaluate the operator at a given point in the domain. When the PDE is discretized, the locality of the operator infers sparsity of the discrete operator—to evaluate the operator on a grid point, only the information at the neighboring grid points is needed. In this case, classical elliptic problem solvers for the forward problem can be modified to apply to the optimization problems, see [16] for an overview. Multigrid methods for elliptic equations are shown to be particularly efficient since their computational complexity is linear in the number of degrees of freedom, i.e. in the number of grid points in the computational domain in $\mathbb{R}^d$, see [4, 5].

In recent literature the study of problems with nonlocal constraints has attracted particular interest, where the operator is not differential but of integral type. The prototypical example of this kind of problems is the *fractional Laplacian* operator, which is attained by taking a fractional power of the classical Laplacian operator, see for example [32, 31, 17, 13]. These problems pose an additional computational challenge: since local information is not sufficient for the evaluation of the operator, the discretized operator will be a dense matrix instead of a sparse one—if implemented in a straightforward way, even the simplest matrix-vector operations would have a quadratic complexity in the number of degrees of freedom in the computational box in $\mathbb{R}^d$. This would make these problems intractable even on moderately fine grids, especially in the three-dimensional case.

A number of approaches have been proposed in the literature to circumvent these difficulties, see papers [19, 17, 37, 11, 18] which consider approximation methods for fractional elliptic operator and [10, 38] related to time dependent problems. An extension method has been proposed [6], which reduces a fractional Laplacian problem to a classical Laplacian problem in a higher-dimensional space, allows to make the problem tractable in some cases. Recently, a proof of concept for an optimal control solver based on the extension approach has been proposed, see [1].

However, note that above methods based on the conventional techniques of numerical analysis are providing a linear scaling in the problem size, thus exhibiting an exponential growth in the storage and computational complexity as $O(n^d)$ in the number of dimensions d, where n is the univariate grid size. One of the promising approaches which exploits the structure of involved quantities for the extraction of essential information in the system of interest is based on the concept of rank-structured tensor representation (approximation) of the target multivariate functions and operators, which allows to avoid the so-called "curse of dimensionality".

In the last decades an extensive research have been focused on different aspects of multilinear algebra, tensor calculus and related issues, see [7, 29, 14, 25] for an overview. Recent tensor numerical methods appeared as a bridging of the well-known tensor decompositions from multilinear algebra with basic results in nonlinear approximation theory by using the low-rank separable representation of the multidimensional functions and operators [12, 13]. In the recent years, the development of low-rank tensor numerical methods has been a prior direction of mathematical research in scientific computing [25, 21, 30, 9]. The main idea of tensor methods is in reducing the numerical solution of the multidimensional integral-

differential equations to one-dimensional tensor product operations.

In this article, we introduce the tensor numerical method for the efficient numerical solution of the d-dimensional optimal control problems with fractional Laplacian type operators in constraints discretized on large spacial grids. We propose and analyze the approximate low-rank structure representations of functions of the fractional Laplacian $(-\Delta)^{\alpha}$, and its inverse $(-\Delta)^{-\alpha}$, $\alpha > 0$, in the bounded domain in $\mathbb{R}^d$, $d = 2, 3$, by using the canonical tensor format. There are many equivalent definitions of fractional elliptic operator based on either integral or spectral formulation, see [32, 31, 17, 13]. Our tensor approach is based on the spectral decomposition of the target operators.

The functions of finite element (finite difference) Laplacian on a tensor grid are diagonalized by using the fast Fourier transform (FFT) matrix and then the low rank tensor approximation to the multi-dimensional core diagonal tensor is computed. In 3D case the multigrid Tucker tensor decomposition [27] and Tucker-to-canonical transform are employed [21]. The theoretical justification on the low-rank canonical and Tucker tensor approximation of functions of the discrete fractional Laplacian is provided. We then show that these low-parametric representations transfer to the solution operator of an optimal control problem governed by the fractional Laplacian. In this way the spacial dimensions can be approximately separated to admit a low-rank tensor structure in the solution vector. Our approach allows to make the solution of the considered optimal control problems computationally tractable in both the two- and three-dimensional cases. We justify that optimal control problems of this class can be solved with a cost that is only slightly larger than linear in the number of grid points in one spacial dimension, independent of the number of spacial dimensions, which is a considerable improvement over classical solvers, whose cost scales in the spacial dimension exponentially. We provide convincing numerical results to support our theoretical reasoning.

Notice that the low-rank tensor methods in the context of fractional time dependent optimal control based on a one-dimensional fractional Laplacian operator have been recently reported in [10].

The rest of the paper is organized as follows. Section 2 describes the considered problem classes. Section 3 discusses the FEM/FDM discretization schemes, formulates the traditional Lagrange multiplies approach and describes the Kronecker product tensor structure in the discrete Laplacain type matrices in many dimensions. In Section 4 we recall the main definition and basic properties of the canonical and Tucker tensor format to be applied for tensor approximation. Section 5 analyzes the tensor approximation of the inverse to fractional Laplace operator and to some related matrix valued functions of fractional Laplacian in $\mathbb{R}^d$ arising in representation of the unknown control and design functions. Finally, in Section 6, we collect the results of numerical tests for 2D and 3D examples which confirm the efficiency of the tensor approximation in the considered class of optimal control problems.

2 Problem setting

Our goal is the construction of fast solution schemes for solving the control problems with d-dimensional fractional elliptic operators in constraints. For this reason we restrict ourself to the case of rectangular domains and to the class of generalized Laplacian type elliptic operators with separable coefficients.

Given the design function $y_\Omega \in L^2(\Omega)$ on $\Omega := (0,1)^d$, $d = 1, 2, 3$, first, we consider the optimization problem for the cost functional

$$\min_{y,u} J(y,u) := \int_\Omega (y(x) - y_\Omega(x))^2 \, dx + \frac{\gamma}{2} \int_\Omega u^2(x) \, dx, \tag{2.1}$$

constrained by the elliptic boundary value problem in Ω for the state variable $y \in H_0^1(\Omega)$,

$$\mathcal{A}y := -\nabla^T \cdot \mathbb{A}(x) \nabla y = \beta u, \quad x \in \Omega, \; u \in L_2(\Omega), \quad \beta > 0, \tag{2.2}$$

endorsed with the homogeneous Dirichlet boundary conditions on $\Gamma = \partial\Omega$, i.e., $y_{|\Gamma} = 0$. The coefficient matrix $\mathbb{A}(x) \in \mathbb{R}^{d\times d}$ is supposed to be symmetric, positive definite and uniformly bounded in Ω with positive constants $c > 0$ and $C > 0$, i.e.,

$$c \, I_{d\times d} \le \mathbb{A}(x) \le C \, I_{d\times d}.$$

Under above assumptions the associated bilinear form

$$A(u,v) = \int_\Omega \mathbb{A}(x) \nabla u(x) \cdot \nabla v(x) \, dx$$

defined on $V \times V$, $V := \{v \in H_0^1(\Omega)\}$ is symmetric, coercive and bounded on V with the same constants c and C.

In what follows, we describe the tensor method for fast numerical solution of the optimization problem with the generalized constraints

$$\mathcal{A}^\alpha y = \beta u(x), \quad x \in \Omega, \tag{2.3}$$

such that for $0 < \alpha \le 1$, the fractional power of the elliptic operator $\mathcal{A}$ is defined by

$$\mathcal{A}^\alpha y(x) = \sum_{i=1}^\infty \lambda_i^\alpha c_i \psi_i(x), \quad y = \sum_{i=1}^\infty c_i \psi_i(x), \tag{2.4}$$

where $\{\psi_i(x)\}_{i=1}^\infty$ is the set of L_2-orthogonal eigenfunctions of the symmetric, positive definite operator $\mathcal{A}$, while $\{\lambda_i\}_{i=1}^\infty$ are the corresponding (real and positive) eigenvalues.

Notice that the elliptic operator inverse $\mathcal{A}^{-1} = \mathcal{T} : L_2(\Omega) \to V$, where $\mathcal{A} = \mathcal{T}^{-1}$, provides the explicit representation for the state variable, $y = \beta \mathcal{T} u = \beta \mathcal{A}^{-1} u$ in case (2.2), while in the general case (2.3) we have

$$y = \beta \mathcal{T}^\alpha u \equiv \beta \mathcal{A}^{-\alpha} u. \tag{2.5}$$

Here $\mathcal{T}$ is a compact, symmetric and positive definite operator on $L_2(\Omega)$ and its eigenpairs $\{\psi_i, \mu_i\}$, $i = 1, \ldots, \infty$, provide an orthonormal basis for $L_2(\Omega)$. Clearly, we have $\lambda_i^{-1} = \mu_i$.

There are several equivalent representations (definitions) for the fractional powers of the symmetric, positive definite operators $\mathcal{A}^\alpha$ and $\mathcal{A}^{-\alpha}$ with $0 < \alpha \le 1$, see for example the survey paper [31]. In particular, the Dunford-Taylor contour integral and the Laplace transform integral representations could be applied. In the presented computational schemes based on low rank tensor decompositions we apply the Laplace transform integral representation. For $\alpha > 0$, the integral representation based on the Laplace transform

$$\mathcal{A}^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t\mathcal{A}} \, dt \tag{2.6}$$

suggests the numerical schemes for low rank canonical tensor representation of the operator (matrix) $\mathcal{A}^{-\alpha}$ by using the sinc quadrature approximations for the integral on the real axis [12]. The efficiency of this approach is based on the exponentially fast convergence of the sinc quadratures on a class of analytic functions. This techniques is to be applied in the present paper for both the theoretical analysis of the rank decomposition schemes and for the description of their constructive representation.

Further more, for $\alpha > 0$ the Dunford-Taylor (or Dunford-Cauchy) contour integral representation reads (see for example [19, 17, 13])

$$\mathcal{A}^{-\alpha} = \frac{1}{2\pi i} \int_{\mathcal{G}} z^{-\alpha} (\mathcal{A} - z\mathcal{I})^{-1} dz, \tag{2.7}$$

where the contour $\mathcal{G}$ in the complex plane includes the spectrum of operator (matrix) $\mathcal{A}$. This representation applies to any $u \in L_2(\Omega)$ and it allows to define the negative fractional powers of elliptic operator as a finite sum of elliptic resolvents $\mathcal{R}_z(\mathcal{L}) = (z\mathcal{I} - \mathcal{L})^{-1}$ by application of certain quadrature approximations, see also [17, 12, 13]. This opens the way for multigrid type or $\mathcal{H}$-matrix (see [13]) schemes approximating the fractional powers of elliptic operator with variable coefficients of rather general form.

It is worth to notice that both integral representations (2.6) and (2.7) can be applied to rather general class of analytic functions of operator $f(\mathcal{A})$, including the case $f(\mathcal{A}) = \mathcal{A}^{-\alpha}$, see [19, 12, 13].

The constraints equation (2.5) allows to derive the Lagrange equation for the control u in the explicit form as follows (see §3 concerning the Lagrange equations)

$$\big(\beta\mathcal{A}^{-\alpha} + \tfrac{\gamma}{\beta}\mathcal{A}^{\alpha}\big)u = y_\Omega, \tag{2.8}$$

for some positive constants $\beta > 0$ and $\gamma > 0$. This equation implies the following representation for the state variable

$$y = \beta\mathcal{A}^{-\alpha}u = \big(\mathcal{I} + \tfrac{\gamma}{\beta^2}\mathcal{A}^{2\alpha}\big)^{-1} y_\Omega. \tag{2.9}$$

The practically interesting range of parameters includes the case $\beta = O(1)$ for the small values of $\gamma > 0$. Out tensor numerical scheme is focused on the solution of equations (2.8) and (2.9) that include the nonlocal operators of "integral-differential" type. The efficiency of the rank-structured tensor approximations presupposes that the design function in the right-hand side of these equations, $y_\Omega(x_1, x_2, x_3)$, allows the low rank separable approximation.

Since we aim for the low-rank (approximate) tensor representation of all functions and operators involved in the above control problem, we further assume that the equation coefficients matrix takes a diagonal form

$$\mathbb{A}(x) = \mathrm{diag}\{a_1(x_1), a_2(x_2)\}, \quad a_\ell(x_\ell) > 0, \quad \ell = 1, 2,$$

in 2D case, and similar for $d = 3$

$$\mathbb{A}(x) = \mathrm{diag}\{a_1(x_1), a_2(x_2), a_3(x_3)\}, \quad a_\ell(x_\ell) > 0, \quad \ell = 1, 2, 3. \tag{2.10}$$

In what follows, we consider the discrete matrix formulation of the optimal control problem (2.1), (2.3) based on the FEM/FDM discretization of d-dimensional elliptic operators

defined on the uniform $n_1 \times n_2 \times \ldots \times n_d$ tensor grid in Ω. We use the collocation grid representation of functions, such that the L_2 scalar product will be substituted by the Euclidean scalar product $(\cdot, \cdot)$ of vectors in $\mathbb{R}^n$.

The fractional elliptic operator $\mathcal{A}^\alpha$ is approximated by its FEM/FDM representation $(\mathcal{A}_h)^\alpha$, subject to the homogeneous Dirichlet boundary condition, where the operator (matrix) $(\mathcal{A}_h)^\alpha$ is defined as in (2.4), where the eigenpairs for the corresponding grid discretization $\mathcal{A}_h$ of the elliptic operator $\mathcal{A}$ are used. Here $h = h_\ell = 1/n_\ell$ is the univariate mesh parameter. The FEM/FDM approximation theory for fractional powers of elliptic operator was presented in [11], see also literature therein.

3 Optimality conditions and representations in a low-rank format

The solution of problem (2.1) with constraint (2.3) requires solving for the necessary optimality conditions. In this section, we will derive these conditions based on a discretize-then-optimize-approach. Then, we will discuss how the involved discretized operators can be applied efficiently in a low-rank format, and how this can be used to design a preconditioned conjugate gradient (PCG) scheme for the necessary optimality conditions.

3.1 Discrete optimality conditions

We consider the discretized version of the control problem (2.1)-(2.3). We assume we have a uniform grid in each space dimension, i. e. we have $N = n_1 n_2$ (for $d = 2$) or $N = n_1 n_2 n_3$ (for $d = 3$) grid points. We will denote the discretized state y, design y_Ω and control u by vectors $\mathbf{y}, \mathbf{y}_\Omega, \mathbf{u} \in \mathbb{R}^N$, respectively. For simplicity, we assume that we use the same approximation for all quantities.

Then, the discrete problem is given as

$$\begin{aligned} \min_{\mathbf{y},\mathbf{u}} &= \frac{1}{2}(\mathbf{y} - \mathbf{y}_\Omega)^T M(\mathbf{y} - \mathbf{y}_\Omega) + \frac{\gamma}{2}\mathbf{u}^T M \mathbf{u} \\ \text{s. t. } A^\alpha \mathbf{y} &= \beta M \mathbf{u}, \end{aligned}$$

where $A = \mathcal{A}_h$ denotes a discretization of the elliptic operator $\mathcal{A}$ by finite elements or finite diferences. The matrix M will be a mass matrix in the finite element case and simply the identity matrix in the finite difference case.

For the discrete adjoint $\mathbf{p}$ define the Lagrangian function

$$L(\mathbf{y}, \mathbf{u}, \mathbf{p}) = \frac{1}{2}(\mathbf{y} - \mathbf{y}_\Omega)^T M(\mathbf{y} - \mathbf{y}_\Omega) + \frac{\gamma}{2}\mathbf{u}^T M \mathbf{u} + \mathbf{p}^T(A^\alpha \mathbf{y} - \beta M \mathbf{u}), \tag{3.1}$$

and compute the necessary first order conditions, given by the Karush-Kuhn-Tucker (KKT) system,

$$\begin{bmatrix} M & O & A^\alpha \\ O & \gamma M & -\beta M \\ A^\alpha & -\beta M & O \end{bmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{y}_\Omega \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \tag{3.2}$$

We can solve the state equation for y, getting

$$\mathbf{y} = \beta A^{-\alpha}\mathbf{u},$$

and the design equation for p getting

$$\mathbf{p} = \tfrac{\gamma}{\beta}\mathbf{u}.$$

Hence the adjoint equation gives us an equation for the control $\mathbf{u}$, namely

$$\left(\beta A^{-\alpha} + \tfrac{\gamma}{\beta}A^{\alpha}\right)\mathbf{u} = \mathbf{y}_{\Omega}. \tag{3.3}$$

3.2 Matrix-vector multiplication in the low-rank format

Now we derive a decomposition of the discrete Laplacian A which is compatible with low-rank data. Let $I_{(i)}$ denote the identity matrix, and $A_{(i)}$ the discretized one-dimensional Laplacian on the given grid in the ith mode, then we have

$$A = A_{(1)} \otimes I_2 \otimes I_3 + I_1 \otimes A_{(2)} \otimes I_3 + I_1 \otimes I_2 \otimes A_{(3)}. \tag{3.4}$$

To obtain the matrices $A_{(\ell)}$, we simply discretize the one-dimensional subproblems

$$\begin{aligned} -y''(x_\ell) &= \beta u(x_\ell) \\ y(0) &= 0 = y(1). \end{aligned}$$

Using a uniform grid with grid size h_i, we obtain the discretizations

$$\underbrace{-\frac{1}{h_\ell}\begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}}_{=A_{(\ell)}} \begin{pmatrix} y_1^{(\ell)} \\ \vdots \\ y_{n_\ell}^{(\ell)} \end{pmatrix} = \beta \begin{pmatrix} u_1^{(\ell)} \\ \vdots \\ u_{n_\ell}^{(\ell)} \end{pmatrix}$$

The one-dimensional operator $A_{(i)}$ has an eigenvalue decomposition in the Fourier basis, i. e.

$$A_{(\ell)} = F_\ell^* \Lambda_{(\ell)} F_\ell.$$

In the case of homogeneous Dirichlet boundary conditions the matrix F_i defines the sin-Fourier transform while $\Lambda_{(\ell)} = \operatorname{diag}\{\lambda_1^{(\ell)}, \ldots, \lambda_n^{(\ell)}\}$, where λ_k denote the eigenvalues of the univariate discrete Laplacian with Dirichlet boundary conditions. These are given by

$$\lambda_k = -\frac{4}{h_\ell^2}\sin^2\left(\frac{\pi k}{2(n_\ell+1)}\right) = -\frac{4}{h_\ell^2}\sin^2\left(\frac{\pi k h_\ell}{2}\right). \tag{3.5}$$

Thus, using the properties of the Kronecker product, we can write the first summand in (3.4) as

$$\begin{aligned} A_1 \otimes I_2 \otimes I_3 &= (F_1^* \Lambda_{(1)} F_1) \otimes (F_2^* I_2 F_2) \otimes (F_3^* I_3 F_3) \\ &= (F_1^* \otimes F_2^* \otimes F_3^*)(\Lambda_{(1)} \otimes I_2 \otimes I_3)(F_1 \otimes F_2 \otimes F_3). \end{aligned}$$

The decomposition of the second and third summand works analogously, thus we can write equation (3.4) as

$$\begin{aligned} A =&(F_1^* \otimes F_2^* \otimes F_3^*)(\Lambda_1 \otimes I_2 \otimes I_3)(F_1 \otimes F_2 \otimes F_3) \\ &+ (F_1^* \otimes F_2^* \otimes F_3^*)(I_1 \otimes \Lambda_2 \otimes I_3)(F_1 \otimes F_2 \otimes F_3) \\ &+ (F_1^* \otimes F_2^* \otimes F_3^*)(I_1 \otimes I_2 \otimes \Lambda_3)(F_1 \otimes F_2 \otimes F_3) \\ =&(F_1^* \otimes F_2^* \otimes F_3^*)\big((\Lambda_1 \otimes I_2 \otimes I_3) + (I_1 \otimes \Lambda_2 \otimes I_3) + (I_1 \otimes I_2 \otimes \Lambda_3)\big)(F_1 \otimes F_2 \otimes F_3). \end{aligned} \tag{3.6}$$

The above expression gives us the eigenvalue decomposition, which can be used to efficiently compute functions of A.

In the case $d = 2$, the expression simplifies to

$$\begin{aligned} A &= (F_1^* \otimes F_2^*)(\Lambda_1 \otimes I_2)(F_1 \otimes F_2) + (F_1^* \otimes F_2^*)(I_1 \otimes \Lambda_2)(F_1 \otimes F_2) \\ &= (F_1^* \otimes F_2^*) \underbrace{\big((\Lambda_1 \otimes I_2) + (I_1 \otimes \Lambda_2)\big)}_{=:\Lambda}(F_1 \otimes F_2). \end{aligned} \tag{3.7}$$

Note that expression (3.7) gives us the eigenvalue decomposition of A. Therefore, for a function $\mathcal{F}$ applied to A, we get

$$\mathcal{F}(A) = (F_1^* \otimes F_2^*)\mathcal{F}(\Lambda)(F_1 \otimes F_2). \tag{3.8}$$

Now assume that $\mathcal{F}(A)$ may be expressed approximately by a linear combination of Kronecker rank 1 operators. Then, to approximate $\mathcal{F}(A)$ it is sufficient to approximate the diagonal matrix $\mathcal{F}(\Lambda)$. Assume we have a decomposition

$$\mathcal{F}(\Lambda) = \sum_{k=1}^{R} \operatorname{diag}\big(\mathbf{u}_1^{(k)} \otimes \mathbf{u}_2^{(k)}\big),$$

with vectors $\mathbf{u}_i^{(k)} \in \mathbb{R}^{n_i}$ and $R \ll \min(n_1, n_2)$. Now let $\mathbf{x} \in \mathbb{R}^N$ be a vector given in a low rank format, i. e.

$$\mathbf{x} = \sum_{j=1}^{S} \mathbf{x}_1^{(j)} \otimes \mathbf{x}_2^{(j)},$$

with vectors $\mathbf{x}_i^{(j)} \in \mathbb{R}^{n_i}$ and $S \ll \min(n_1, n_2)$. Then, we can compute a matrix-vector product

$$\begin{aligned} \mathcal{F}(A)\mathbf{x} &= (F_1^* \otimes F_2^*)\bigg(\sum_{k=1}^{R} \operatorname{diag}\big(\mathbf{u}_1^{(k)} \otimes \mathbf{u}_2^{(k)}\big)\bigg)(F_1 \otimes F_2)\bigg(\sum_{j=1}^{S} \mathbf{x}_1^{(j)} \otimes \mathbf{x}_2^{(j)}\bigg) \\ &= \sum_{k=1}^{R}\sum_{j=1}^{S} F_1^*\big(\mathbf{u}_1^{(k)} \odot F_1\mathbf{x}_1^{(j)}\big) \otimes F_2^*\big(\mathbf{u}_2^{(k)} \odot F_2\mathbf{x}_2^{(j)}\big), \end{aligned} \tag{3.9}$$

where $\odot$ denotes the componentwise (Hadamard) product.

Using the sin-FFT, expression (3.9) can be computed in factored form in $\mathcal{O}(RSn \log n)$ flops, where $n = \max(n_1, n_2)$.

In the case $d = 3$, with completely analogous reasoning, equation (3.9) becomes

$$\mathcal{F}(A)\mathbf{x} = \sum_{k=1}^{R}\sum_{j=1}^{S} F_1^*\big(\mathbf{u}_1^{(k)} \odot F_1\mathbf{x}_1^{(j)}\big) \otimes F_2^*\big(\mathbf{u}_2^{(k)} \odot F_2\mathbf{x}_2^{(j)}\big) \otimes F_3^*\big(\mathbf{u}_3^{(k)} \odot F_3\mathbf{x}_3^{(j)}\big), \tag{3.10}$$

and similar in the case of $d > 3$.

3.3 The low-rank PCG scheme

For operators `func` and `precond` given in a low-rank format, such as (3.9) (for $d = 2$) or (3.10) (for $d = 3$), Krylov subspace methods can be applied very efficiently, since they only require matrix-vector products. The formulation of the PCG method in Algorithm 1 is independent of d, as long as an appropriate rank truncation procedure `trunc` is chosen.

Algorithm 1 Preconditioned CG method in low-rank format

Input: Rank truncation procedure `trunc`, rank tolerance parameter ε, linear function in low rank format `fun`, preconditioner in low rank format `precond`, right-hand side tensor $\mathbf{B}$, initial guess $\mathbf{X}^{(0)}$, maximal iteration number $k_{\max}$

$\mathbf{R}^{(0)} \leftarrow \mathbf{B} - \texttt{fun}(\mathbf{X}^{(0)})$
$\mathbf{Z}^{(0)} \leftarrow \texttt{precond}(\mathbf{R}^{(0)})$
$\mathbf{Z}^{(0)} \leftarrow \texttt{trunc}(\mathbf{Z}^{(0)}, \varepsilon)$
$\mathbf{P}^{(0)} \leftarrow (\mathbf{Z}^{(0)})$
$k \leftarrow 0$
repeat
  $\mathbf{S}^{(k)} \leftarrow \texttt{fun}(\mathbf{P}^{(k)})$
  $\mathbf{S}^{(k)} \leftarrow \texttt{trunc}(\mathbf{S}^{(k)}, \varepsilon)$
  $\alpha_k \leftarrow \frac{\langle \mathbf{R}^{(k)}, \mathbf{Z}^{(k)} \rangle}{\langle \mathbf{P}^{(k)}, \mathbf{S}^{(k)} \rangle}$
  $\mathbf{X}^{(k+1)} \leftarrow \mathbf{X}^{(k)} + \alpha_k \mathbf{P}^{(k)}$
  $\mathbf{X}^{(k+1)} \leftarrow \texttt{trunc}(\mathbf{X}^{(k+1)}, \varepsilon)$
  $\mathbf{R}^{(k+1)} \leftarrow \mathbf{R}^{(k)} - \alpha_k \mathbf{S}^{(k)}$
  $\mathbf{R}^{(k+1)} \leftarrow \texttt{trunc}(\mathbf{R}^{(k+1)}, \varepsilon)$
  if $\mathbf{R}^{(k+1)}$ is sufficiently small **then**
    return $\mathbf{X}^{(k+1)}$
    break
  end if
  $\mathbf{Z}^{(k+1)} \leftarrow \texttt{precond}(\mathbf{R}^{(k+1)})$
  $\mathbf{Z}^{(k+1)} \leftarrow \texttt{trunc}(\mathbf{Z}^{(k+1)}, \varepsilon)$
  $\beta_k \leftarrow \frac{\langle \mathbf{R}^{(k+1)}, \mathbf{Z}^{(k+1)} \rangle}{\langle \mathbf{Z}^{(k)}, \mathbf{R}^{(k)} \rangle}$
  $\mathbf{P}^{(k+1)} \leftarrow \mathbf{Z}^{(k+1)} + \beta_k \mathbf{P}^{(k)}$
  $\mathbf{P}^{(k+1)} \leftarrow \texttt{trunc}(\mathbf{P}^{(k+1)}, \varepsilon)$
  $k \leftarrow k + 1$
until $k = k_{\max}$

Output: Solution $\mathbf{X}$ of $\texttt{fun}(\mathbf{X}) = \mathbf{B}$

As the rank truncation procedure, in our implementation we apply the reduced SVD algorithm in 2D case and the RHOSVD based canonical-to-Tucker-to-canonical algorithm (see [27]) as described in Section 4.

4 Rank-structured decomposition of function related tensors

The basic tensor formats used in multilinear algebra for the low-rank representation of the multidimensional tensors are the canonical [20] and Tucker [36] tensor formats. These rank-structured tensor representations have been commonly used for the quantitative analysis of correlations in multi-dimensional experimental data in chemometrics, psychometrics and signal processing [7, 29]. Application of these and of the new tensor formats in scientific computing and in computational chemistry are described in [25, 21], where it is shown that due to logarithmically low Tucker tensor ranks for function related tensors [22] solution of the multidimensional integral-differential equations is reduced to principally one-dimensional calculations. For the readers convenience, in this section we recall the main tensor techniques applied in this paper.

A tensor of order d in a full format, is defined as a multidimensional array over a d-tuple index set:

$$\mathbf{T} = [t_{i_1,\ldots,i_d}] \equiv [t(i_1,\ldots,i_d)] \in \mathbb{R}^{n_1\times\ldots\times n_d} \text{ with } \quad i_\ell \in I_\ell := \{1,\ldots,n_\ell\}. \tag{4.1}$$

For a tensor in a full size format (assuming $n_\ell = n$, $\ell = 1,\ldots d$) the required storage, as well as operations with tensors scale exponentially[1] in the dimension size d, as n^d.

To avoid exponential scaling in the dimension, the rank-structured separable representations (approximations) of the multidimensional tensors can be used. By using a tensor product of vectors $\mathbf{u}^{(\ell)}$ in each dimension $\ell = 1,\ldots d$, a rank-1 tensor is constructed as

$$\mathbf{T} = \mathbf{u}^{(1)} \otimes \ldots \otimes \mathbf{u}^{(d)} \in \mathbb{R}^{n_1\times\ldots\times n_d},$$

with entries $t_{i_1,\ldots,i_d} = u^{(1)}_{i_1}\cdots u^{(d)}_{i_d}$. Such tensor requires only dn numbers for storage (taking for simplicity $n_\ell = n$). The rank-1 canonical tensor may be considered as a discrete counterpart of a separable d-variate function represented as the product of univariate functions,

$$f(x_1, x_2, \ldots, x_d) = f_1(x_1)f_2(x_2)\ldots f_d(x_d).$$

A tensor in the R-term canonical format is defined by a sum of rank-1 tensors

$$\mathbf{T} = \sum\nolimits_{k=1}^{R} \xi_k \mathbf{u}^{(1)}_k \otimes \ldots \otimes \mathbf{u}^{(d)}_k, \quad \xi_k \in \mathbb{R}, \tag{4.2}$$

where $\mathbf{u}^{(\ell)}_k \in \mathbb{R}^{n_\ell}$ are normalized vectors, and R is the canonical rank. The storage cost of this parametrization is bounded by dRn. An alternative (contracted product) notation for a canonical tensor can be used (cf. the Tucker tensor format, (4.4))

$$\mathbf{T} = \mathbf{C} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 \cdots \times_d U^{(d)}, \tag{4.3}$$

[1]This exponential growth in dimension size is the so-called "curse of dimensionality" [2].

where $\mathbf{C} = diag\{c_1, ..., c_d\} \in \mathbb{R}^{R\otimes d}$ is a super-diagonal tensor, $U^{(\ell)} = [\mathbf{u}_1^{(\ell)} \dots \mathbf{u}_R^{(\ell)}] \in \mathbb{R}^{n_\ell \times R}$ are the side matrices, and $\times_\ell$ denotes the contacted product in mode ℓ.

For $d \geq 3$, there is lack of stable algorithms to compute the canonical rank representation of a general tensor $\mathbf{T}$, that is, with the minimal number R in representation (4.2), and the respective decomposition with the polynomial cost in d, i.e., the computation of the canonical decomposition is in general an N-P hard problem.

The Tucker tensor format is suitable for stable numerical decompositions with a fixed truncation threshold. We say that the tensor $\mathbf{T}$ is represented in the rank-$\mathbf{r}$ orthogonal Tucker format with the rank parameter $\mathbf{r} = (r_1, \dots, r_d)$ if

$$\mathbf{T} = \sum_{\nu_1=1}^{r_1} \dots \sum_{\nu_d=1}^{r_d} \beta_{\nu_1,\dots,\nu_d} \mathbf{v}_{\nu_1}^{(1)} \otimes \mathbf{v}_{\nu_2}^{(2)} \dots \otimes \mathbf{v}_{\nu_d}^{(d)}, \tag{4.4}$$

where $\{\mathbf{v}_{\nu_\ell}^{(\ell)}\}_{\nu_\ell=1}^{r_\ell} \in \mathbb{R}^{n_\ell}$, $\ell = 1, \dots, d$ represents a set of orthonormal vectors and $\boldsymbol{\beta} = [\beta_{\nu_1,\dots,\nu_d}] \in \mathbb{R}^{r_1\times\dots\times r_d}$ is the Tucker core tensor. The storage cost for the Tucker tensor format is bounded by $drn + r^d$, with $r = |\mathbf{r}| := \max_\ell r_\ell$. Using the orthogonal side matrices $V^{(\ell)} = [\mathbf{v}_1^{(\ell)} \dots \mathbf{v}_{r_\ell}^{(\ell)}]$, the Tucker tensor decomposition can be presented by using contracted products,

$$\mathbf{T}_{(\mathbf{r})} = \boldsymbol{\beta} \times_1 V^{(1)} \times_2 V^{(2)} \times_3 \dots \times_d V^{(d)}. \tag{4.5}$$

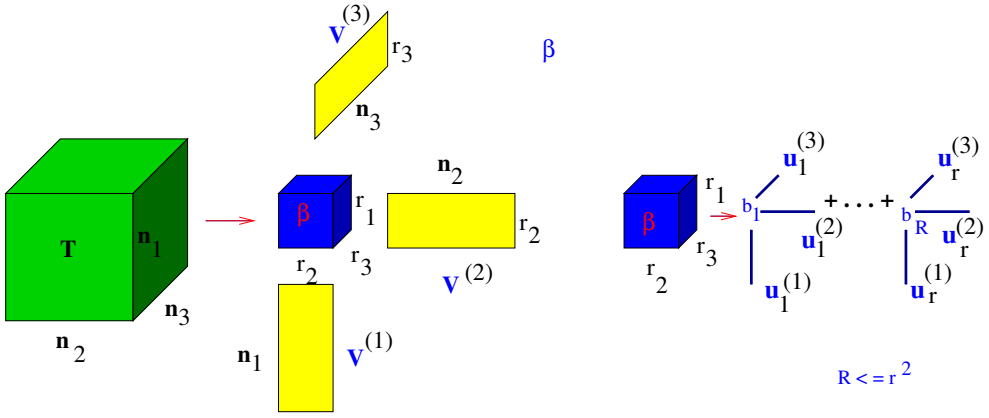


Figure 4.1: Tucker tensor decomposition of the full size tensor, and canonical tensor decomposition of the Tucker core.

In general, the problem of the best Tucker tensor approximation to a given tensor is the following minimization problem,

$$\mathbf{T}_0 \in \mathbb{V}_\mathbf{n} : \quad f(\mathbf{T}) := \|\mathbf{T} - \mathbf{T}_0\|^2 \to min \quad \text{over} \quad \mathbf{T} \in \boldsymbol{\mathcal{T}}_\mathbf{r}, \tag{4.6}$$

which is equivalent to the maximization problem [8]

$$g(V^{(1)}, ..., V^{(d)}) := \left\| \mathbf{T}_0 \times_1 V^{(1)^T} \times ... \times_d V^{(d)^T} \right\|^2 \to max \tag{4.7}$$

over the set $\mathcal{M}_\ell$ of orthogonal matrices $V^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell}$, $\ell = 1, 2, \ldots, d$. For given maximizers $V^{(\ell)}$, the core $\boldsymbol{\beta}$ that minimizes (4.6) is computed as

$$\boldsymbol{\beta} = \mathbf{T}_0 \times_1 V^{(1)^T} \times_2 \ldots \times_d V^{(d)^T} \in \mathbb{R}^{r_1 \times \ldots \times r_d}, \tag{4.8}$$

yielding contracted product representation (4.5). The Tucker tensor format provides a stable decomposition algorithm [8], based on following steps (consider for simplicity $d = 3$).

- SVD-type higher order singular value decomposition (HOSVD) of matrix unfolding of a tensor $T_{(\ell}$, $\ell = 1, 2, 3$ for finding the so-called initial guess.
- Tucker ALS iteration for finding the best orthogonal approximation to side matrices $V^{(\ell)}$, $\ell = 1, 2, \ldots d$, by fixing all dimensions except one;
- Calculation of the core $\boldsymbol{\beta}$ by simple contractions (4.8).

Figure 4.2 shows the construction of the unfolding matrix of a 3D tensor for mode $\ell = 1$. The complexity of HOSVD is $O(n^{d+1})$ which makes this techniques intractable for higher dimensions and in 3D in case of larger mode sizes n_ℓ, $\ell = 1, 2, 3$.

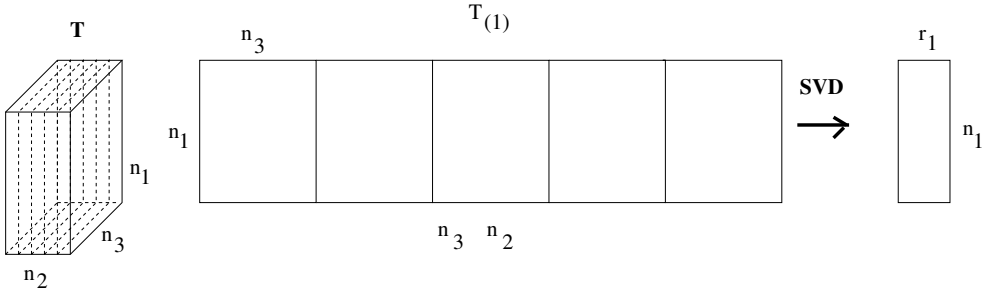


Figure 4.2: HOSVD of the full size $n_1 \times n_2 \times n_3$ tensor (SVD of its matrix unfolding).

It appeared that the Tucker tensor decomposition for function related tensors exhibits exceptional approximation properties. It was proven [22] and demonstrated numerically that for a class of higher order tensors arising from the discretization of linear operators and functions in $\mathbb{R}^d$ the approximation error of the Tucker decomposition decays exponentially in the Tucker rank [22, 26]. Previous papers on the low-rank approximation of the multidimensional functions and operators, in particular, based on the sinc-quadratures [12] described the constructive way for the analytical low-rank canonical representations.

The multigrid Tucker tensor decomposition for function related tensors introduced in [27] leads to solution of the minimization problem (4.6) over a sequence of nested subspaces

$$\mathbb{T}_{\mathbf{r},0} \subset \ldots \subset \mathbb{T}_{\mathbf{r},m} \subset \ldots \subset \mathbb{T}_{\mathbf{r},M},$$

using the sequence of dyadic refined grids of size $n = n_0\, 2^{m-1}$ with $m = 1, \ldots, M$. The Tucker decomposition problem for a tensor $\mathbf{T}_{0,M} \in \mathbb{V}_{\mathbf{n}_M}$ obtained as discretization of a function over the fine grid of size n_M is based on the successive reiteration of the ALS Tucker approximation on a sequence of refined grids. The initial guess is computed by HOSVD only at the coarsest grid with $n_0 \ll n_M$, at the moderate cost $O(n_0^{d+1})$. For finer grids the initial guess is obtained

by interpolation of the orthogonal Tucker vectors in $\{V_{m-1}^{(\ell)}\}_{\ell=1}^{d}$ from the previous grid level. The multigrid approach applies the HOSVD requirements only to the coarsest grid, and thus reduces complexity of the 3D Tucker tensor decomposition from $O(n_M^4)$ to $O(n_M^3)$, where n_M is the largest univariate grid size.

Figure 4.3 displays the exponentially fast convergence in $\mathbf{r}$ of the Tucker tensor approximation error in Frobenius norm,

$$E_F = \frac{\|\mathbf{T}_0 - \mathbf{T}_{\mathbf{r},n_M}\|}{\|\mathbf{T}_0\|}$$

for the 3D tensor $\mathbf{G}_2 = [g_2(i,j,k)]$ corresponding to a function of three variables, see (5.13), with $\alpha = 1/2$ (left panel) and $\alpha = 1/10$ (right panel). The function g_2 is discretized on a sequence of $n \times n \times n$ 3D Cartesian grids. Figure 4.3 shows that the separable representation for this function with accuracy of the order of $\approx 10^{-7}$, can be constructed using ten Tucker vectors in each space dimension, nearly independently on the size of discretization grid.

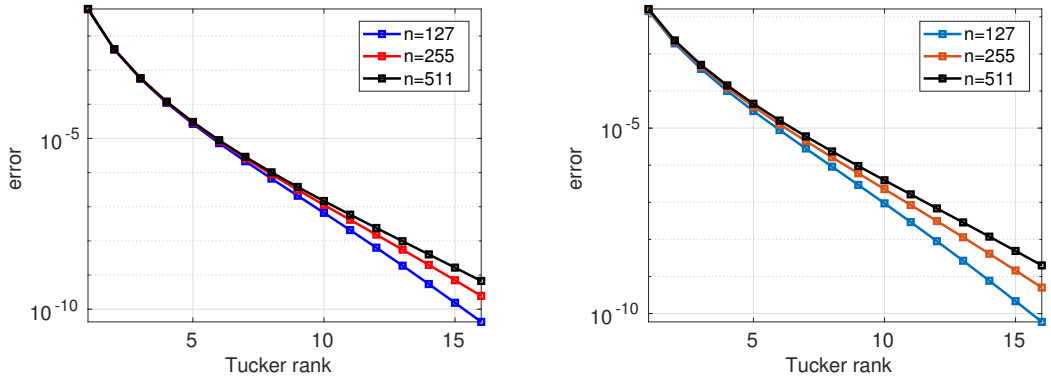


Figure 4.3: Multigrid Tucker approximation of $\mathbf{G}_2$ for $\alpha = 1/2$ (left) and $\alpha = 1/10$ (right) vs. Tucker rank r for $n \times n \times n$ 3D Cartesian grids with univariate grid size $n = 127, 255, 511$.

For tensors given in the rank-R canonical format, $\mathbf{T}_C \in \boldsymbol{\mathcal{C}}_{R,\mathbf{n}}$, with large ranks and with large univariate size n, both construction of the full size tensor representation and HOSVD become intractable. The remedy is the reduced higher order singular value decomposition (RHOSVD) introduced in [27] as a part of the canonical-to-Tucker (C2T) decompositions for function-related tensors, which does not require the construction of a full size tensor. Then minimization problem in $\mathcal{T}_{\mathbf{r}}$ for tensors given in R-term canonical tensor format

$$\mathbf{T}_C \in \boldsymbol{\mathcal{C}}_{R,\mathbf{n}} \subset \mathbb{V}_{\mathbf{n}}: \quad \mathbf{T}_{(\mathbf{r})} = \mathrm{argmin}_{V \in \mathcal{T}_{\mathbf{r},\mathbf{n}}} \|\mathbf{T}_C - \mathbf{T}\|_{\mathbb{V}_{\mathbf{r}}}, \tag{4.9}$$

is equivalent to [27]

$$[Z^{(1)}, ..., Z^{(d)}] = \mathrm{argmax}_{V^{(\ell)} \in \mathcal{M}_\ell} \left\| \sum_{\nu=1}^{R} \xi_\nu \left(V^{(1)^T} \mathbf{u}_\nu^{(1)} \right) \otimes ... \otimes \left(V^{(d)^T} \mathbf{u}_\nu^{(d)} \right) \right\|_{\mathbb{V}_{\mathbf{r}}}^2, \tag{4.10}$$

$V^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell}$ – orthogonal. Instead of HOSVD, the initial guess $Z_0^{(\ell)}$ is computed by RHOSVD, i.e., by the truncated SVD of side matrices $U^{(\ell)} = [\mathbf{u}_1^{(\ell)} ... \mathbf{u}_R^{(\ell)}]$ in (4.3),

$$U^{(\ell)} \approx Z_0^{(\ell)} \Sigma_0 W_0^{(\ell)}, \tag{4.11}$$

under the compatibility condition $r_\ell \le rank(U^{(\ell)})$. The RHOSVD exhibits complexity $O(nR^2)$ (or $O(n^2R)$ if $n \le R$) (does not depend on the dimension of a canonical tensor) and thus it allows to avoid the curse of dimensionality for the Tucker-type decompositions, see [25, 21] for details. The Tucker tensor is built by the contractions of the tensors of type (4.3) with matrices of type (4.11) in the course of ALS iteration and final calculation of a Tucker core $\boldsymbol{\beta}$, see details in [27, 21]. Figure 4.4 shows the RHOSVD of one of the canonical matrices in (4.3).

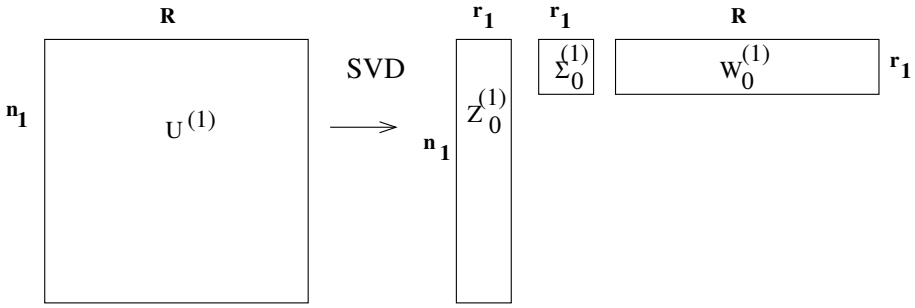


Figure 4.4: RHOSVD for the side matrix $U^{(\ell)}$, $\ell = 1$ of the rank-R canonical tensor.

In general, C2T with RHOSVD and Tucker-to-canonical (T2C) decomposition[2] are main tools for reducing the ranks of a canonical tensor representation of the multivariate functions in the process of the numerical solution of the multidimensional problem. In this paper, they are used for the low rank approximation of the target matrix valued functions as well as for reducing the tensor ranks in the course of iteration process for solution of the considered 3D control problems.

5 Low-rank tensor approximation of functions of fractional Laplacian

In this section we analyze the rank structured tensor decompositions of various matrix (tensor) valued functions on the discrete Laplacian arising in the solution of problems (2.8) and (2.9) including different combinations of fractional Laplacian in $\mathbb{R}^d$.

The elements of the core diagonal matrix Λ in (3.6) can be represented as a three-tensor

$$\mathbf{G} = [g(i_1, i_2, i_3)] \in \mathbb{R}^{n_1 \times n_2 \times n_3}, \quad i_\ell \in \{1, \dots, n_\ell\},$$

where

$$g(i_1, i_2, i_3) = \lambda_{i_1} + \lambda_{i_2} + \lambda_{i_3},$$

implying that $\mathbf{G}$ has the exact rank-3 decomposition. In the case $d = 2$ we have similar two-term representation, $g(i_1, i_2) = \lambda_{i_1} + \lambda_{i_2}$.

[2]This decomposition applies to a small size core tensor see Figure 4.1 for the mixed Tucker-canonical format of type

$$\mathbf{T}_{(\mathbf{r})} = \left(\sum_{k=1}^{R} b_k \mathbf{u}_k^{(1)} \otimes \ldots \otimes \mathbf{u}_k^{(d)} \right) \times_1 V^{(1)} \times_2 V^{(2)} \times_3 \ldots \times_d V^{(d)},$$

providing the canonical tensor rank of the order of $R \le r^2$ [21].

We consider the matrices A_1, A_2 and A_3 defined as the matrix valued functions of the discrete Laplacain $\mathcal{A}_h$ by the equations

$$A_1 = \mathcal{A}_h^{-\alpha}, \tag{5.1}$$

$$A_2 = \mathcal{A}_h^{-\alpha} + \mathcal{A}_h^{\alpha}, \tag{5.2}$$

and

$$A_3 = \left(\mathcal{A}_h^{-\alpha} + \mathcal{A}_h^{\alpha}\right)^{-1} = A_2^{-1}, \tag{5.3}$$

respectively. It is worth to notice that the matrix A_3 defines the solution operator in equation (2.8), which allows to calculate the optimal control in terms of the design function $\mathbf{y}_\Omega$ on the right-hand side of (2.8) as follows

$$\mathbf{u}^* = A_3 \mathbf{y}_\Omega. \tag{5.4}$$

Finally, the state variable is calculated by

$$\mathbf{y} = A_4 \mathbf{y}_\Omega, \quad \text{and} \quad A_4 = (I + \mathcal{A}_h^{2\alpha})^{-1} = \mathcal{A}_h^{\alpha} A_2^{-1}. \tag{5.5}$$

In the following numerics we consider the behavior of the SVD or Tucker decompositions for the corresponding multi-indexed core tensors/matrices G_1, G_2, G_3 and G_4 representing the matrix valued functions A_1, A_2, A_3 and A_4 in the Fourier basis. It is well suited for the rank-structured algebraic operations since the d-dimensional Fourier transform matrix has Kronecker rank equals to one (see §3)

$$F = F_1 \otimes F_2 \otimes F_3.$$

Let $\{\lambda_i\}_{i=1}^n$ be the set of eigenvalues for the 1D finite difference Laplacian in $H_0^1(0,1)$ discretized on the uniform grid with the mesh size $h = 1/(n+1)$.

In the 2D case we analyze the singular value decomposition of the $n \times n$ core matrices

$$G_p = [g_p(i,j)], \quad p = 1,\, 2,\, 3,\, 4, \tag{5.6}$$

with entries defined by

$$g_1(i,j) = \frac{1}{(\lambda_i + \lambda_j)^\alpha}, \tag{5.7}$$

$$g_2(i,j) = \frac{1}{(\lambda_i + \lambda_j)^\alpha} + (\lambda_i + \lambda_j)^\alpha, \tag{5.8}$$

$$g_3(i,j) = \left(\frac{1}{(\lambda_i + \lambda_j)^\alpha} + (\lambda_i + \lambda_j)^\alpha\right)^{-1}, \tag{5.9}$$

$$g_4(i,j) = \left(1 + (\lambda_i + \lambda_j)^{2\alpha}\right)^{-1}, \tag{5.10}$$

In the 3D case we consider the Tucker decomposition of the $n \times n \times n$ core tensor

$$\mathbf{G}_p = [g_p(i,j,k)], \quad p = 1,2,3,4 \tag{5.11}$$

with entries defined by

$$g_1(i,j,k) = \frac{1}{(\lambda_i + \lambda_j + \lambda_k)^\alpha}, \tag{5.12}$$

$$g_2(i,j,k) = \frac{1}{(\lambda_i+\lambda_j+\lambda_k)^\alpha} + (\lambda_i+\lambda_j+\lambda_k)^\alpha, \tag{5.13}$$

$$g_3(i,j,k) = \left(\frac{1}{(\lambda_i+\lambda_j+\lambda_k)^\alpha} + (\lambda_i+\lambda_j+\lambda_k)^\alpha\right)^{-1}, \tag{5.14}$$

$$g_4(i,j,k) = \left(1+(\lambda_i+\lambda_j+\lambda_k)^{2\alpha}\right)^{-1}. \tag{5.15}$$

The error estimate for the rank decomposition of the matrices G_p and the respective 3D tensors $\mathbf{G}_p$ can be derived based on the sinc-approximation theory. We consider the class of matrix valued functions of the discrete Laplacian, $A_1(\mathcal{A}_h), \ldots, A_4(\mathcal{A}_h)$, given by (5.1) – (5.3) and (5.5), respectively. In view of the FFT diagonalization, the tensor approximation problem is reduced to the analysis of the corresponding function related tensors $\mathbf{G}_1, \ldots, \mathbf{G}_4$ specified by multivariate functions of the discrete argument, $g_1, \ldots, g_4$, given by (5.12) – (5.15).

Our approach applies to the class of simplified "directionally fractional" Laplacian-type operators $\mathcal{A}_\alpha$ (with the diagonal coefficient as in (2.10)) obtained by the modification of (2.2) as follows

$$\mathcal{A}_\alpha := -\sum_{\ell=1}^{d} \left(\nabla_\ell^T \cdot a_\ell(x_\ell)\nabla_\ell\right)^\alpha, \quad \alpha > 0. \tag{5.16}$$

In this case the core tensors in (5.12) – (5.15) representing the operator $\mathcal{A}_\alpha$ in the Fourier basis are simplified to (for $a_\ell(x_\ell) = const$, $d = 3$)

$$\widetilde{g}_1(i,j,k) = \frac{1}{\lambda_i^\alpha+\lambda_j^\alpha+\lambda_k^\alpha}, \tag{5.17}$$

$$\widetilde{g}_2(i,j,k) = \frac{1}{\lambda_i^\alpha+\lambda_j^\alpha+\lambda_k^\alpha} + \lambda_i^\alpha+\lambda_j^\alpha+\lambda_k^\alpha, \tag{5.18}$$

$$\widetilde{g}_3(i,j,k) = \left(\frac{1}{\lambda_i^\alpha+\lambda_j^\alpha+\lambda_k^\alpha} + \lambda_i^\alpha+\lambda_j^\alpha+\lambda_k^\alpha\right)^{-1}, \tag{5.19}$$

$$\widetilde{g}_4(i,j,k) = \left(1+(\lambda_i^\alpha+\lambda_j^\alpha+\lambda_k^\alpha)^2\right)^{-1}, \tag{5.20}$$

correspondingly. Hence the rank behavior in the tensor decomposition of these discrete functions is completely similar to the case of Laplacian like operator in (5.12) – (5.15) corresponding to the fractional power $\alpha = 1$. This special case was considered in [10].

In the general case of variable coefficients $a_\ell(x_\ell) > 0$ in (5.16) the orthogonal matrices F_ℓ, $\ell = 1, \ldots, d$, of the univariate Fourier transforms in (3.6) should be substituted by the orthogonal matrices of the eigenvalue decomposition for the discretized elliptic operators $A_{h,\ell}$ (stiffness matrices) corresponding to $A_\ell = -\nabla_\ell^T \cdot a_\ell(x_\ell)\nabla_\ell$. The eigenvalues in (5.17) – (5.20) are obtained from the solution of the eigenvalue problem $A_{h,\ell}\mathbf{u}_i = \lambda_{i,\ell}\mathbf{u}_i$, $\ell = 1, \ldots, d$.

In the rest of this section, we sketch the proof of the existence of the low rank canonical/Tucker decomposition of the core tensors $\mathbf{G}_p$, $p = 1, \ldots, 4$. Following [34], we define the Hardy space $H^1(D_\delta)$ of functions which are analytic in the strip

$$D_\delta := \{z \in \mathbb{C} : |\Im\, z| \le \delta\}, \quad 0 < \delta < \frac{\pi}{2},$$

and satisfy

$$N(f, D_\delta) := \int_{\mathbb{R}} (|f(x + i\delta)| + |f(x - i\delta)|)\, dx < \infty.$$

Recall that for $f \in H^1(D_\delta)$, the integral

$$I(f) = \int_\Omega f(x)dx \quad (\Omega = \mathbb{R} \text{ or } \Omega = \mathbb{R}_+)$$

can be approximated by the Sinc-quadrature (trapezoidal rule)

$$T(f, \mathfrak{h}) := \mathfrak{h} \sum_{k=-\infty}^{\infty} f(k\mathfrak{h}), \qquad |I(f) - T(f, \mathfrak{h})| = O(e^{-\pi\delta/\mathfrak{h}}), \quad \mathfrak{h} \to 0,$$

that converges exponentially fast in $\mathfrak{h} \to 0$. Similar estimates hold for computable truncated sums, see [34]

$$T_M(f, \mathfrak{h}) := \mathfrak{h} \sum_{k=-M}^{M} f(k\mathfrak{h}). \tag{5.21}$$

Indeed, if $f \in H^1(D_\delta)$ and $|f(x)| \le C \exp(-b|x|)$ for all $x \in \mathbb{R}$ $b, C > 0$, then

$$|I(f) - T_M(f, \mathfrak{h})| \le C \left[\frac{e^{-2\pi\delta/\mathfrak{h}}}{1 - e^{-2\pi\delta/\mathfrak{h}}} N(f, D_\delta) + \frac{1}{b} e^{-b\mathfrak{h}M} \right]. \tag{5.22}$$

In our context, the Sinc-quadrature approximation applies to multivariate functions $F : \mathbb{R}^d \to \mathbb{R}$ of a sum of single variables, say, $F(x_1, \ldots, x_d) = f(\rho)$ with $\rho = \sum_{\ell=1}^d f_\ell(x_\ell) > 0$, where $f_\ell : \mathbb{R} \to \mathbb{R}_+$, by using the integral representation of the analytic univariate function $f : \mathbb{R}_+ \to \mathbb{R}$,

$$f(\rho) = \int_\Omega G(t) e^{-\rho E(t)} \mathrm{d}t, \quad \Omega \in \{\mathbb{R}, \mathbb{R}_+, (a, b)\}.$$

In the cases (5.1) – (5.3) and (5.5) the related functions $f(\rho)$ take the particular form

$$f(\rho) = \rho^{-\alpha}, \quad f(\rho) = (\rho^\alpha + \rho^{-\alpha})^{-1}, \quad f(\rho) = (1 + \rho^{2\alpha})^{-1}.$$

The univariate function f may have point singularities or cusps at $\rho = 0$, say, $f(\rho) = \rho^{\pm\alpha}$. Applying the Sinc-quadrature (5.21) to the Laplace-type transform leads to rank-R separable approximation of the function F,

$$F(x) = f(\rho) = f(f_1(x_1) + \ldots + f_d(x_d)) \approx \sum_{\nu=1}^{R} \omega_\nu G(t_\nu) e^{-\rho E(t_\nu)} = \sum_{\nu=1}^{R} c_\nu \prod_{\ell=1}^{d} e^{-f_\ell(x_\ell) E(t_\nu)}, \tag{5.23}$$

with $c_\nu = \omega_\nu G(t_\nu)$ and $R = 2M + 1$.

Notice that the generating function f can be defined on the discrete argument, i.e., on the multivariate index $\mathbf{i} = (i_1, \ldots, i_d)$, $i_\ell = 1, \ldots, n$, such that each univariate function f_ℓ in (5.23) is defined on the index set $\{i_\ell\}$, $\ell = 1, \ldots, d$. In our particular applications to functions of the discrete fractional Laplacian $\mathcal{A}_h^\alpha$ we have

$$f_\ell(i_\ell) = \lambda_{i_\ell} = -\frac{4}{h_\ell^2} \sin^2\left(\frac{\pi i_\ell h_\ell}{2}\right), \tag{5.24}$$

where λ_{i_ℓ} denote the eigenvalues of the univariate discrete Laplacian with the Dirichlet boundary conditions, see (3.5) and (5.12) – (5.15).

In this case the integral representation of the function $f(\rho) = \rho^{-\alpha}$, $\rho = \sum_{\ell=1}^{d} f_\ell(x_\ell) > 0$, with f_ℓ given by (5.24) $\alpha > 0$, takes a form

$$\rho^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-\rho t}\, dt, \quad \rho \in [1, B],\ B > 1. \tag{5.25}$$

Several efficient sinc-approximation schemes for classes of multivariate functions and operators have been developed, see [12, 15, 22, 25]. In the case (5.25) the simple modification of Lemma 2.51 in [25] can be applied, see also [15]. To that end, the substitution $t = \phi(u) := \log(1 + e^u)$, that maps $\phi : \mathbb{R} \to \mathbb{R}_+$, leads to the integral

$$\rho^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_{\mathbb{R}} f_1(u) du, \quad f_1(u) = \frac{[\log(1 + e^u)]^{\alpha-1} e^{-\rho \log(1+e^{-u})}}{1 + e^{-u}},$$

which can be approximated by the sinc-quadrature with the choice $\mathfrak{h} = \frac{\gamma\pi}{\sqrt{M}}$ with some $0 < \gamma < 1$. This argument justifies the accurate low-rank representation of functions in (5.7) and (5.12) representing the fractional Laplacian inverse.

The numerical results presented in Section 6 clearly illustrate the high accuracy of of the low-rank approximations to the other matrix valued functions of the fractional Laplacian defined in (5.7) – (5.10) and in (5.12) – (5.15). In our numerical tests the rank decomposition of the 2D and 3D tensors under consideration was performed by the multigrid Tucker-to-canonical scheme as described in Section 4.

Remark 5.1 *The presented approach is applicable with minor modifications to the case of more general elliptic operators of the form (for $d = 3$)*

$$\mathcal{A} = F(-\Delta_1) \otimes I_2 \otimes I_3 + I_1 \otimes F(-\Delta_1) \otimes I_3 + I_1 \otimes I_2 \otimes F(-\Delta_1), \tag{5.26}$$

where F is the rather general analytic function of the univariate Laplacian. In this case the fractional operators $\mathcal{A}^\alpha$ and $\mathcal{A}^{-\alpha}$ can be introduced by the similar way, where the values $\{\lambda_i\}$ in the representation of the core tensors $\mathbf{G}_p$, $p = 1, 2, 3, 4$, should be substituted by $\{F(\lambda_i)\}$. Given the symmetric positive definite operator (matrix) X, then the particular choice $F(X) = X^{\pm\alpha}$, $\alpha > 0$, suites well for the presented approach.

We notice that the rank structured approximation of the control problem with the Laplacian type operator in the form (5.26) with the particular choice $F(-\Delta_1) = (-\Delta_1)^\alpha$ was considered in [10]. Since the operator in (5.26) only includes the univariate fractional Laplacain, this case can be treated as for the standard Laplacian type control problems.

6 Numerics on rank-structured approximations

In this section we analyze the rank decomposition of all matrix entities involved in the solution operator (2.8). For the ease of exposition, in what follows, we set the model constants as $\beta = \gamma = 1$ and assume that $n_1 = n_2 = n_3$. Recall that $A = F^* G F$ with the notation

$A = \mathcal{A}_h$, where $\mathcal{A}_h$ is the FDM approximation to the elliptic operator $\mathcal{A}$ and G is the diagonal core matrix represented in terms of eigenvalues of the discrete Laplacian $A = \mathcal{A}_h$. All numerical simulations are performed in MATLAB on a laptop.

First, we illustrate the smoothing properties of the elliptic operator $\mathcal{A}_h^{-\alpha}$ in 2D (or by the other words, the localization properties of the fractional operator $\mathcal{A}_h^{\alpha}$) in the equation for control depending on the fractional power $\alpha > 0$. Figures 6.1, 6.1, 6.3 and 6.4 represent the shape of the design function y_Ω and the corresponding optimal control $\mathbf{u}^*$ in the equation (5.4) computed for different values of the parameter α and for fixed grid size $n = 255$.

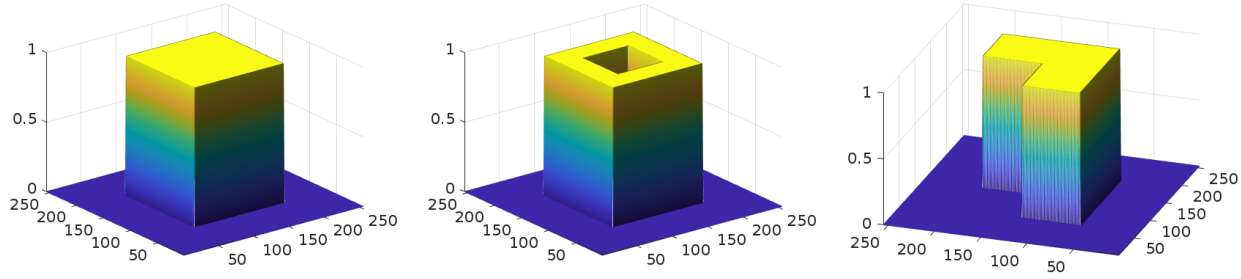


Figure 6.1: Shapes of the right-hand side y_Ω used in the 2D equation (5.4) computed with n=255.

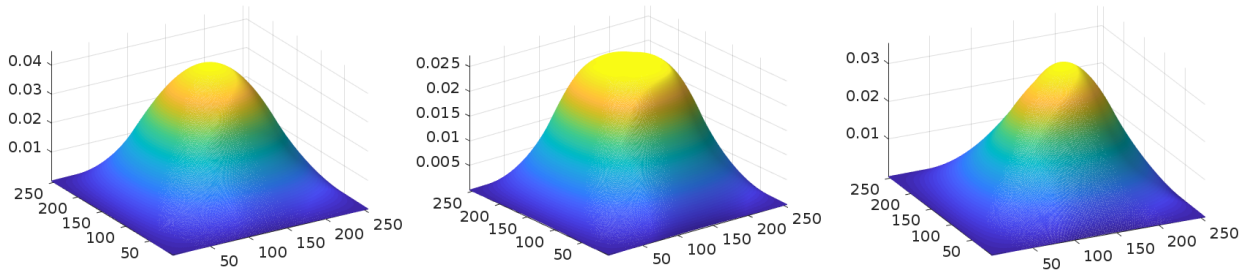


Figure 6.2: Solutions $\mathbf{u}^*$ for above right-hand sides y_Ω with $\alpha = 1$ for $n = 255$.

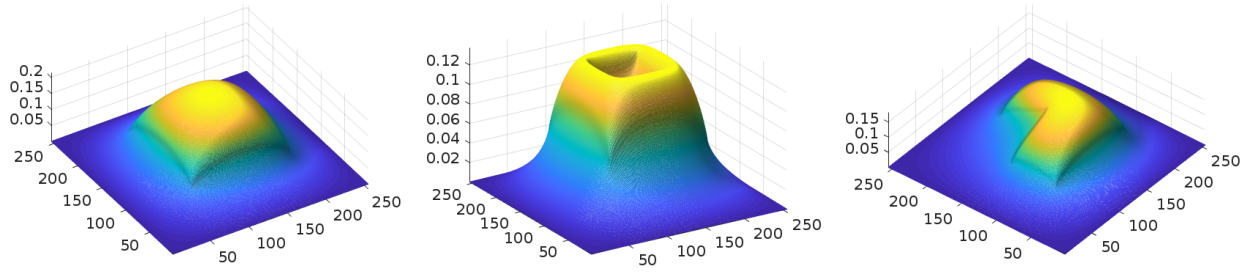


Figure 6.3: Solutions $\mathbf{u}^*$ for above right-hand sides with $\alpha = 1/2$ for $n = 255$.

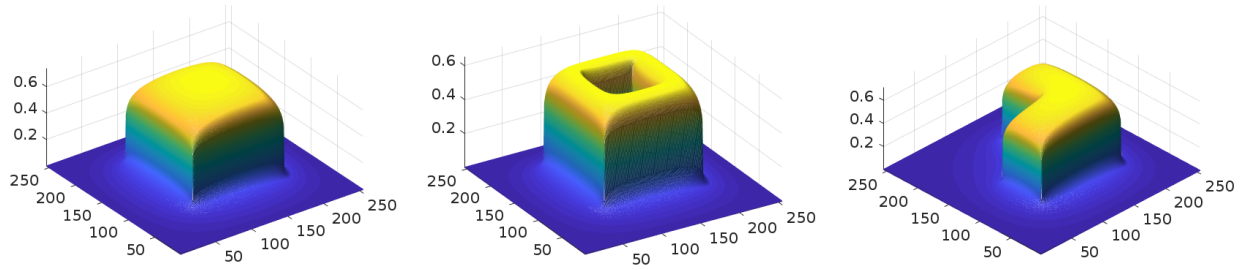


Figure 6.4: Solutions $\mathbf{u}^*$ for above right-hand sides with $\alpha = 1/10$ for $n = 255$.

One observes the nonlocal features of the elliptic resolvent $\mathcal{A}_h^{-1}$ and highly localized action of the operator $\mathcal{A}_h^{-\frac{1}{10}}$.

6.1 Numerical tests for 2D case

Figure 6.5, left, represents the singular values of the matrix G_1, with entries given by (5.7) for different univariate grid size $n = 255, 511$, and 1023 and fixed $\alpha = 1$ (Laplacian inverse). Figure 6.5, right, shows the decay of respective singular values for G_1 with fixed univariate grid size $n = 511$ and for different $\alpha = 1, 1/2, 1/10$.

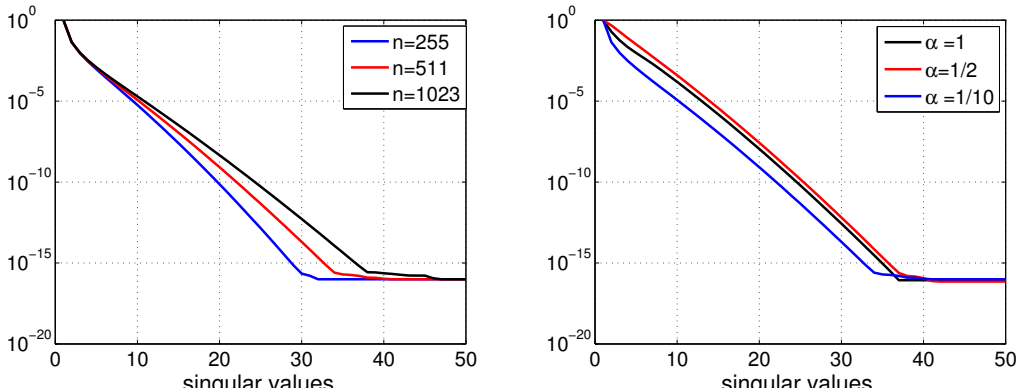


Figure 6.5: Decay of singular values for G_1 with $\alpha = 1$ vs. n (left); singular values for G_1 vs. $\alpha > 0$ with fixed $n = 511$ (right).

Figure 6.6 demonstrates the behavior of singular values for matrices G_2 and G_3, with the entries corresponding to (5.8) and (5.9), respectively, vs. $\alpha = 1, 1/2, 1/10$ with fixed univariate grids size $n = 511$. In all cases we observe exponentially fast decay of the singular values which means there exists the accurate low Kronecker rank approximation of the matrix functions A_1, A_2 and A_3 (see equations (5.1), (5.2) and (5.3)) including fractional powers of the elliptic operator.

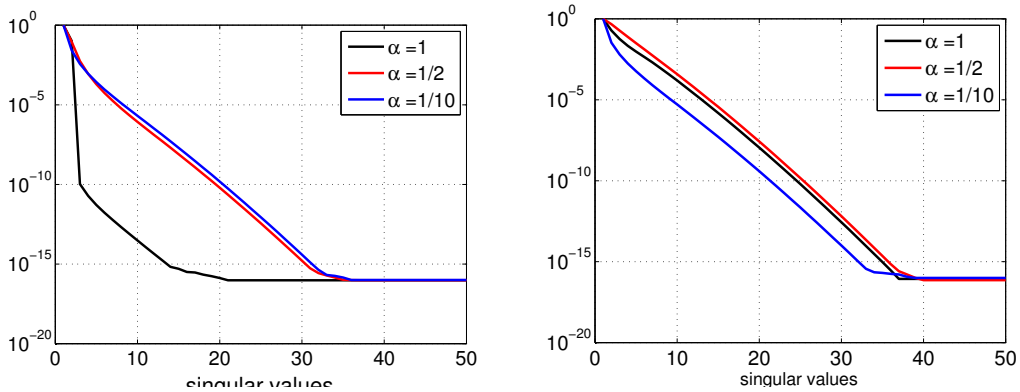


Figure 6.6: Decay of singular values of G_2 (left) and G_3 (right) vs. $\alpha = 1, 1/2, 1/10$ for $n = 511$.

Decay of the error for the optimal control obtained as the solution of equation (5.4) with rank-R approximation of the solution operator A_3 is shown in Figure 6.7.

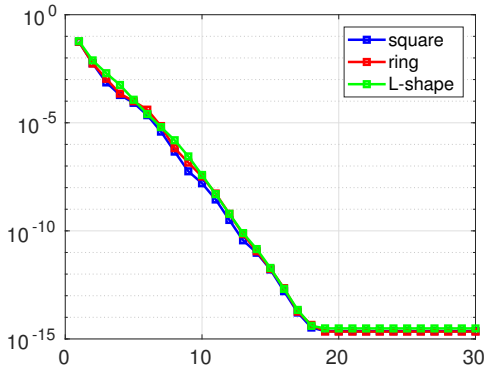


Figure 6.7: Decay of the error for the optimal control vs. truncation rank parameter.

As we have shown theoretically in Section 3, a single PCG iteration has a complexity, which is slightly higher than linear in the univariate grid size n. Figure 6.8 shows that the CPU times show the expected behavior. Thus, with Figure 6.8 and Tables 6.1 and 6.2, the overall cost of the algorithm is almost linear in the univariate grid size n for the problem discretized on $n \times n$ 2D Cartesian grid.

We also test the properties of the low-rank discrete operator as a preconditioner. This means, we solve the equations in $\mathbb{R}^d$, $d = 2, 3$,

$$A^{\alpha}\mathbf{x} = \mathbf{b}, \tag{6.1}$$

$$\left(I + A^{2\alpha}\right)\mathbf{x} = \mathbf{b}, \tag{6.2}$$

$$\left(A^{\alpha} + A^{-\alpha}\right)\mathbf{x} = \mathbf{b}, \tag{6.3}$$

with a preconditioned conjugate gradient scheme, using a low-rank direct solver as a preconditioner discussed above. We simplify the notation by $A = \mathcal{A}_h$.

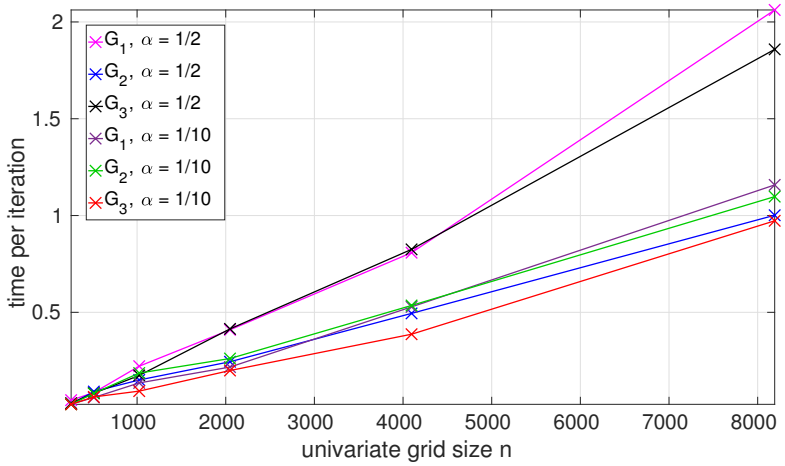


Figure 6.8: CPU times (sec) vs. univariate grid size n for a single iteration of Algorithm 1 for a 2D problem, for different fractional operators and fixed preconditioner rank $r = 5$.

In numerical tests we solve the equations (6.1) - (6.3) on a grid of size n, using a rank-r preconditioner. Tables 6.1 and 6.2 show the number of CG iteration counts for convergence to a relative residual of 10^{-6} of (6.1)-(6.3) with $\alpha = 1/2$ and $\alpha = 1/10$, respectively. The dash '—' indicates failure to converge to converge in 100 iterations.

	g_1				g_4				g_3			
r \ n	256	512	1024	2048	256	512	1024	2048	256	512	1024	2048
1	20	24	24	29	—	—	83	80	20	24	24	19
2	—	—	3	3	73	—	38	36	—	—	3	3
3	7	9	10	14	99	—	17	16	7	9	10	14
4	5	6	6	9	31	—	3	3	5	6	6	9
5	4	4	4	5	11	–	5	5	4	4	4	5
6	3	3	3	4	6	13	2	2	3	3	3	4
7	3	3	3	3	4	7	6	4	3	3	3	3
8	2	2	2	2	3	5	4	2	2	2	2	2
9	2	2	2	2	3	4	3	4	2	2	2	2
10	2	2	2	2	3	3	2	3	2	2	2	2

Table 6.1: PCG iteration counts for convergence to a relative residual of 10^{-6} for the equations (6.1) - (6.3) for a 2D fractional Laplacian with $\alpha = 1/2$ vs. the univariate grid size n and separation rank r.

As can be seen in Tables 6.1 and 6.2, we achieve almost grid-independent preconditioning; the iteration counts only grow logarithmically with the number of grid points, as can be expected from the theoretical reasoning. As can be seen in Table 6.1, the ranks of the preconditioner should be chosen sufficiently large to ensure reliability. In the cases tested here, $r = 6$ is sufficient to achieve reliable preconditioning even in the most difficult case of equation (6.2) with $\alpha = 1/2$.

	g_1				g_4				g_3			
r \ n	256	512	1024	2048	256	512	1024	2048	256	512	1024	2048
1	9	9	10	11	11	13	14	16	7	7	8	9
2	6	4	7	8	7	8	8	9	5	5	6	6
3	4	5	5	6	5	5	6	7	4	4	5	5
4	4	4	4	5	4	4	4	5	3	4	4	4
5	3	3	4	4	3	4	4	4	3	3	3	4
6	3	3	3	4	3	3	3	4	2	3	3	3
7	2	3	3	3	2	3	3	3	2	2	3	3
8	2	2	2	3	2	2	2	3	2	2	2	3
9	2	2	2	2	2	2	2	2	2	2	2	3
10	2	2	2	2	2	2	2	2	2	2	2	2

Table 6.2: PCG iteration counts for convergence to a relative residual of 10^{-6} for the equations (6.1) - (6.3) for a 2D fractional Laplacian with $\alpha = 1/10$ vs. the univariate grid size n and separation rank r.

6.2 Numerical tests for 3D case

In the following examples we solve the problems governed by the 3D operators in (5.1) – (5.3), with a 3D fractional Laplacian with $\alpha = 1, 1/2$ and $\alpha = 1/10$. The rank-structured approximation to the above fractional operators is performed by using the multigrid Tucker decomposition of the 3D tensors $\mathbf{G}_k$, $k = 1, 2, 3, 4$, described by (5.12) – (5.14), and the consequent Tucker-to-canonical decomposition of the Tucker core tensor thus obtaining a canonical tensor with a smaller rank. The rank truncation procedure in the PCG Algorithm 1 is performed by using the RHOSVD tensor approximation and its consequent transform to the canonical format, see Section 4.

Figures 6.9 – 6.11 demonstrate the exponential convergence of the approximation error with respect to the Tucker rank for operators given by (6.1) – (6.3).

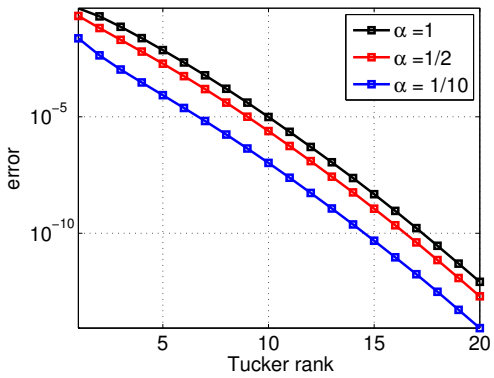
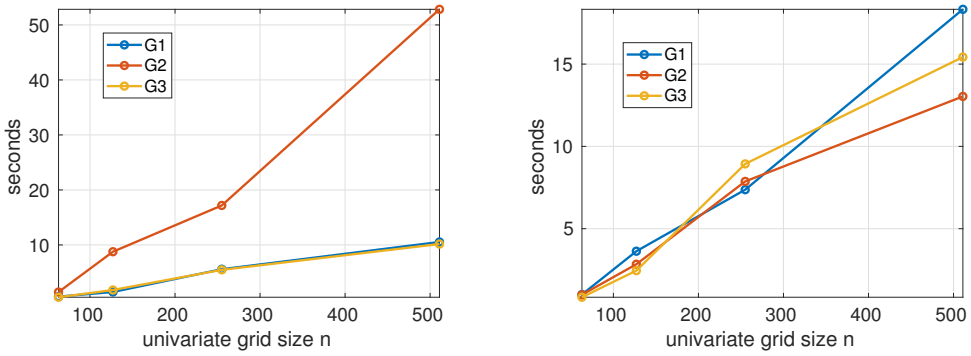


Figure 6.9: Tucker tensor approximation of $\mathbf{G}_1$ vs. rank parameter for $\alpha = 1, 1/2, 1/10$.

We solve the equations (6.1) - (6.3) using $n \times n \times n$ 3D Cartesian grids with the univariate grid size n, using a rank-r preconditioner. Tables 6.3 and 6.4 show the number of CG



$\alpha = 1/2$, $r = 4$ $\alpha = 1/10$, $r = 7$

Figure 6.10: CPU times (in seconds) vs. univariate grid size n for a single iteration of Algorithm 1 for a 3D problem, for different fractional operators and fixed preconditioner rank r.

iteration counts for convergence to a relative residual of 10^{-6} of (6.1) - (6.3) with $\alpha = 1/2$ and $\alpha = 1/10$, respectively.

Similarly to the previous subsection, we see that the low-rank approximation gives us an approximately grid-independent preconditioner. In the cases tested here, $r = 6$ is sufficient to achieve reliable preconditioning even in the most difficult case of equation (6.2) with $\alpha = 1/2$.

	g_1				g_4				g_3			
r \ n	64	128	256	512	64	128	256	512	64	128	256	512
4	1	2	1	1	1	6	1	2	1	2	1	1
5	1	1	1	2	1	1	8	4	1	1	1	2
6	1	1	1	1	2	2	1	1	1	1	1	1
7	1	3	1	2	1	1	5	4	1	2	1	2
8	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	2	1	6	5	4	1	1	1	2
10	1	1	1	1	1	6	1	1	1	1	1	1

Table 6.3: PCG iteration counts for convergence to a relative residual of 10^{-6} for the equations (6.1) - (6.3) for a 3D fractional Laplacian with $\alpha = 1/2$. Here n is the univariate grid size, r is the separation rank.

Our numerical test indicates that all three matrices A_1, A_2 and A_3, as well as the corresponding three-tensors have ε-rank approximation such that the rank parameter depends logarithmically on ε, i.e., $r = O(|\log \varepsilon|)$, that means that the low rank representation of the design function y_Ω ensures the low rank representation of both optimal control and optimal state variable.

We show as well that, using rank-structured tensor methods for the numerical solution of this optimization problem using the operators of type A_1, A_2 and A_3 can be implemented at low cost that scales linearly in the univariate grid size, $O(n \log n)$, see Figure 6.10.

	g_1				g_4				g_3			
r \ n	64	128	256	512	64	128	256	512	64	128	256	512
4	2	1	9	20	2	1	10	17	1	1	9	18
5	1	1	1	1	1	1	1	1	1	2	1	13
6	1	1	1	2	1	1	1	2	1	1	1	7
7	1	1	1	2	1	1	1	2	1	1	2	1
8	1	1	1	1	1	1	1	1	1	1	1	2
9	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	2	1	1	1	2	1	1	1	1

Table 6.4: CG iteration counts for convergence to a relative residual of 10^{-6} for the equations (6.1)-(6.3) for a 3D fractional Laplacian with $\alpha = 1/10$. Here n is the univariate grid size, r is the separation rank.

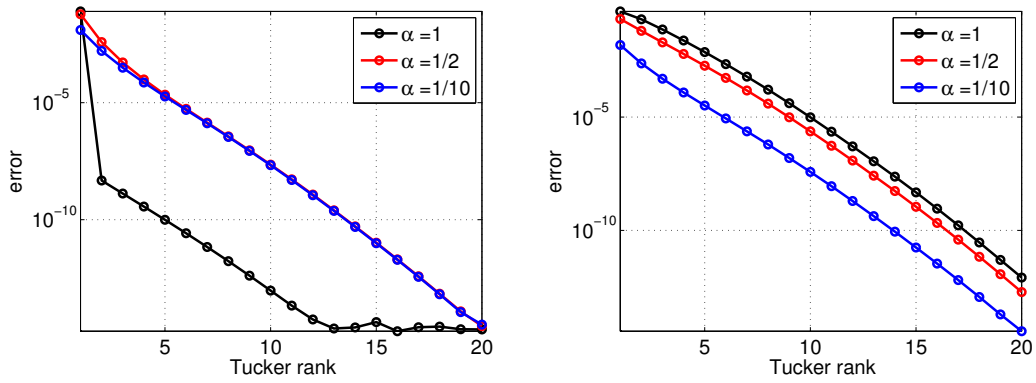


Figure 6.11: Tucker tensor approximation of $\mathbf{G}_2$ and $\mathbf{G}_3$ vs. rank parameter for $\alpha = 1, 1/2, 1/10$.

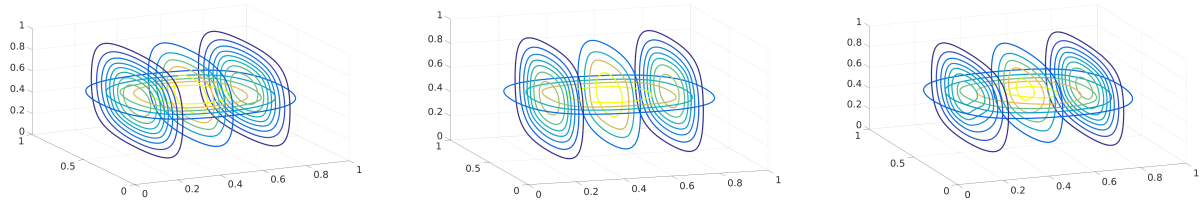


Figure 6.12: Solutions of the equation with 3D right-hand sides (analogous to Figure 6.1) with $\alpha = 1$ for $n = 255$.

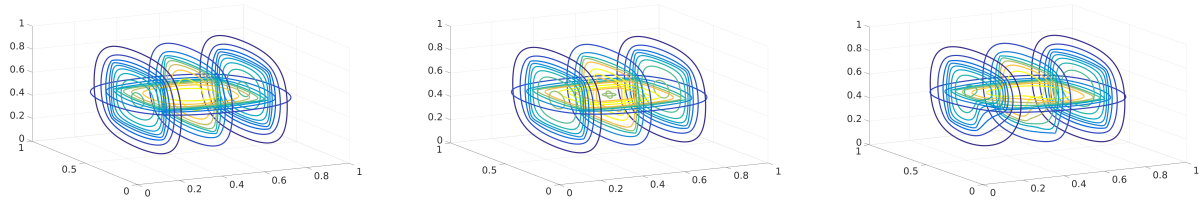


Figure 6.13: Solutions of the equation with 3D right-hand sides (analogous to Figure 6.1) with $\alpha = 1/2$ for $n = 255$.

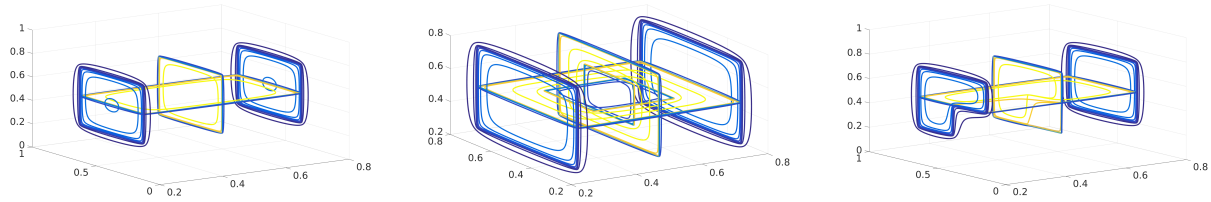


Figure 6.14: Solutions of the equation with 3D right-hand sides (analogous to Figure 6.1) with $\alpha = 1/10$ for $n = 255$.

7 Conclusions

We have introduced and analyzed a new approach for the optimal control of a fractional Laplacian equation using tensor numerical methods. The fractional Laplacian is diagonalized in the FFT basis on a tensor grid and a low-rank approximation to the core diagonal tensor is computed. We present the novel rank-structured representation of functions of the fractional elliptic operator based on sinc-approximation method applied to the core tensor. This representation exhibits the exponential decay of the approximation error in the rank parameter.

These results apply to the fractional Laplacian itself, as well as to the solution operators of a fractional control problem, resulting from first-order necessary conditions. Due to the separation of the spatial variables, the application of the arising matrix-valued functions of a fractional Laplacian to a given rank-structured vector has a complexity which is nearly linear (linear-logarithmic) in the univariate grid size, independently of the spatial dimension of the problem.

The PCG algorithm for solving the equation for control function with adaptive rank truncation is implemented. In 3D case the rank truncation is based on the RHOSVD-Tucker approximation and its transform to the low-rank canonical form. The numerical study illustrates the exponential decay of the approximation error of the canonical tensor decompositions of the target tensors in the rank parameter, and indicates the almost linear complexity scaling of the rank-truncated PCG solver in the univariate grid size n for 3D problems discretized on $n \times n \times n$ Cartesian grid. The low-rank preconditioner provides the uniform convergence rate in the grid size n and other model parameters. All numerical simulations are performed in Matlab on a laptop.

The approach can be applied to generalized Laplacian-type control operator and to the case of fractional elliptic operator with variable coefficients. On the other hand, the further reduction of the numerical complexity to the logarithmic scale can be achieved by using the quantized-TT (QTT) representation of all discrete functions and operators involved (see [24, 33, 25]).

References

[1] H. Antil and E. Otárola. *A FEM for an Optimal Control Problem of Fractional Powers of Elliptic Operators.* SIAM J. Control Optim., 53(6), 2015, 3433-3456.

[2] R. E. Bellman. *Dynamic programming.* Princeton University Press, 1957.

[3] P. Benner. *Solving large-scale control problems.* IEEE control systems, 24 (1), 2004, pp. 44-59.

[4] A. Borzi and V. Schulz. *Multigrid methods for PDE optimization.* SIAM Review 51(2), 2009, 361-395.

[5] A. Borzi and V. Schulz. *Computational optimization of systems governed by partial differential equations.* Soc. for Ind. and Appl. Math., Philadelphia, 2012.

[6] L. Caffarelli and L. Silvestre. *An Extension Problem Related to the Fractional Laplacian.* Communications in Partial Differential Equations, 32 (8), 2007, pp. 1245-1260.

[7] A. Cichocki and Sh. Amari. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications.* Wiley, 2002.

[8] L. De Lathauwer, B. De Moor, J. Vandewalle. *A multilinear singular value decomposition.* SIAM J. Matrix Anal. Appl., 21 (2000) 1253-1278.

[9] S. Dolgov and I.V. Oseledets. *Solution of linear systems and matrix inversion in the TT-format.* SIAM J. Sci. Comput. 34 (5), 2011, A2718-A2739.

[10] S. Dolgov, J. Pearson, D. Savostyanov and M. Stoll. *Fast tensor product solvers for optimization problems with fractional differential equations as constraints.* Appl. Math. Comp., 273, 2016, 604-623.

[11] B. Duan, R. Lazarov and J. Pasciak. *Numerical approximation of fractional powers of elliptic operators.* arXiv:1803.10055v1, 2018.

[12] I.P. Gavrilyuk, W. Hackbusch, and B. N. Khoromskij. *Tensor-product approximation to elliptic and parabolic solution operators in higher dimensions.* Computing **74** (2005), 131-157.

[13] I. P. Gavrilyuk, W. Hackbusch and B. N. Khoromskij. *Data-Sparse Approximation to Operator-Valued Functions of Elliptic Operator.* Math. Comp. 73, (2003), 1297-1324.

[14] W. Hackbusch. *Tensor spaces and numerical tensor calculus.* Springer, Berlin, 2012.

[15] W. Hackbusch and B.N. Khoromskij. *Low-rank Kronecker product approximation to multi-dimensional nonlocal operators. Part I. Separable approximation of multi-variate functions.* Computing **76** (2006), 177-202.

[16] R. Herzog and K. Kunisch. *Algorithms for PDE constrained optimization.* GAMM, 33 (2010), 163-176.

[17] Nicholas Hale, Nicholas J Higham, and Lloyd N Trefethen. *Computing A^α, $\log(A)$, and related matrix functions by contour integrals.* SIAM J. on Numerical Analysis, 46 (2), 2008, 2505-2523.

[18] S. Harizanov, R. Lazarov, P. Marinov, S. Margenov and Ya. Vutov. *Optimal solvers for linear systems with fractional powers of sparse spd matrices.* Preprint arXiv:1612.04846v3, 2018.

[19] Nicholas J Higham. *Functions of Matrices.* SIAM, Philadelphia, 2008.

[20] F.L. Hitchcock. *The expression of a tensor or a polyadic as a sum of products.* J. Math. Physics, 6 (1927), 164-189.

[21] Venera Khoromskaia and Boris N. Khoromskij. *Tensor Numerical Methods in Quantum Chemistry.* Research monograph, De Gruyter Verlag, Berlin, 2018.

[22] B. N. Khoromskij. *Structured Rank-$(r_1, ..., r_d)$ Decomposition of Function-related Tensors in $\mathbb{R}^d$.* Comp. Meth. Applied Math., 6, (2006), 2, 194-220.

[23] B. N. Khoromskij. *Tensor-Structured Preconditioners and Approximate Inverse of Elliptic Operators in $\mathbb{R}^d$.* Constructive Approximation, 30:599-620 (2009).

[24] B. N. Khoromskij. *$O(d \log N)$-Quantics Approximation of N-d Tensors in High-Dimensional Numerical Modeling.* Constr. Approx., v.34(2), 2011, 257-289.

[25] Boris N. Khoromskij. *Tensor Numerical Methods in Scientific Computing.* Research monograph, De Gruyter Verlag, Berlin, 2018.

[26] B. N. Khoromskij and V. Khoromskaia. *Low Rank Tucker-Type Tensor Approximation to Classical Potentials.* Central European J. of Math., 5(3), pp.523-550, 2007.

[27] B. N. Khoromskij and V. Khoromskaia. *Multigrid Tensor Approximation of Function Related Arrays.* SIAM J. Sci. Comp., 31(4), 3002-3026 (2009).

[28] B.N. Khoromskij and G. Wittum. *Numerical Solution of Elliptic Differential Equations by Reduction to the Interface.* Research monograph, LNCSE, No. 36, Springer-Verlag 2004.

[29] T. G. Kolda and B. W. Bader. *Tensor Decompositions and Applications.* SIAM Rev, 51(2009), no. 3, pp.455-500.

[30] D. Kressner and C Tobler. *Krylov subspace methods for linear systems with tensor product structure.* SIAM J Matr. Anal. Appl., 31 (4), 2011, 1688-1714, 2011.

[31] M. Kwaśnicki. *Ten equivalent definitions of the fractional Laplace operator.* Functional Calculus and Applied Analysis, 20(1):7-51, 2017.

[32] A. Lischke, G. Pang, M. Gulian, F. Song, Ch. Glusa, X. Zheng, Z. Mao, W. Cei, M. M. Meerschaert, M. Ainsworth, G. E. Karniadakis. *What is the fractional Laplacian?* arXiv:1801.09767v1, 2018.

[33] I.V. Oseledets. *Approximation of $2^d \times 2^d$ matrices using tensor decomposition.* SIAM J. Matrix Anal. Appl., 31(4):2130-2145, 2010.

[34] F. Stenger. *Numerical methods based on Sinc and analytic functions.* Springer-Verlag, 1993.

[35] F. Troeltzsch. *Optimal control of partial differential equations: theory, methods and applications.* AMS, Providence, Rhode Island, 2010.

[36] L. R. Tucker. *Some mathematical notes on three-mode factor analysis.* Psychometrika, 31 (1966) 279-311.

[37] P. N. Vabishchevich. *Numerically solving an equation for fractional powers of elliptic operators.* J. Comput. Phys., 282, 2015, pp. 289-302.

[38] P. N. Vabishchevich. *Numerical solution of time-dependent problems with fractional power elliptic operator.* Comput. Meth. Applied Math., 18 (1), 111-128, 2018.