Tensor Approach to Optimal Control problems with Fractional Elliptic Operator

ALGORITHMIC OPTIMIZATION

www.alop.uni-trier.de

Universität Trier

Volker Schulz

Gennadij Heidel
Britta Schmitt

Boris Khoromskij / Venera Khoromskaia (MPI Leipzig)

Applications causing recent interest

Application fields of fractional operators:

- viscoelastics
- biophysics
- nonlocal electrostatics
- anomalous diffusion
- heat equation in plasmonic nanostructure networks/composite materials
- ...

An optimal control problem

Given a function $y_\Omega \in L^2(\Omega)$ on $\Omega := (0,1)^d$, we consider the optimization problem

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

$$\text{s.t.} \; -\Delta y = \beta u$$

$$y, u \in H_0^1(\Omega)$$

An optimal control problem

Given a function $y_\Omega \in L^2(\Omega)$ on $\Omega := (0,1)^d$, we consider the optimization problem

$$\min_{y,u} J(y,u) := \int_\Omega (y(x) - y_\Omega(x))^2 \, dx + \frac{\gamma}{2} \int_\Omega u^2(x) \, dx$$
$$\text{s.t. } \mathcal{A}^\alpha y = \beta u$$

where $\mathcal{A}^\alpha$ is the *spectral* fractional Laplacian operator for some $\alpha \in (0,1)$.

The KKT system

$$\begin{bmatrix} id & 0 & \mathcal{A}^\alpha \\ 0 & \gamma id & -\beta id \\ \mathcal{A}^\alpha & -\beta id & 0 \end{bmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = \begin{pmatrix} y_\Omega \\ 0 \\ 0 \end{pmatrix}$$

The KKT system

$$\begin{bmatrix} id & 0 & \mathcal{A}^\alpha \\ 0 & \gamma id & -\beta id \\ \mathcal{A}^\alpha & -\beta id & 0 \end{bmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = \begin{pmatrix} y_\Omega \\ 0 \\ 0 \end{pmatrix} \begin{array}{l} \Rightarrow (\beta \mathcal{A}^{-\alpha} + \frac{\gamma}{\beta}\mathcal{A}^\alpha)u = y_\Omega \\ \Rightarrow p = \frac{\gamma}{\beta} u \\ \Rightarrow y = \beta \mathcal{A}^{-\alpha} u \end{array}$$

The KKT system

$$\begin{bmatrix} id & 0 & \mathcal{A}^\alpha \\ 0 & \gamma id & -\beta id \\ \mathcal{A}^\alpha & -\beta id & 0 \end{bmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = \begin{pmatrix} y_\Omega \\ 0 \\ 0 \end{pmatrix} \quad \begin{array}{l} \Rightarrow (\beta\mathcal{A}^{-\alpha} + \frac{\gamma}{\beta}\mathcal{A}^\alpha)u = y_\Omega \\ \Rightarrow p = \frac{\gamma}{\beta} u \\ \Rightarrow y = \beta \mathcal{A}^{-\alpha} u \end{array}$$

Thus, we find the following necessary optimality conditions:

$$u = \left(\beta\mathcal{A}^{-\alpha} + \tfrac{\gamma}{\beta}\mathcal{A}^\alpha\right)^{-1} y_\Omega$$

for the control u, and

$$y = \beta\mathcal{A}^{-\alpha} u = \left(\mathcal{I} + \tfrac{\gamma}{\beta^2}\mathcal{A}^{2\alpha}\right)^{-1} y_\Omega$$

for the state y.

The KKT system

$$\begin{bmatrix} id & 0 & \mathcal{A}^\alpha \\ 0 & \gamma id & -\beta id \\ \mathcal{A}^\alpha & -\beta id & 0 \end{bmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = \begin{pmatrix} y_\Omega \\ 0 \\ 0 \end{pmatrix} \quad \begin{array}{l} \Rightarrow (\beta\mathcal{A}^{-\alpha} + \frac{\gamma}{\beta}\mathcal{A}^\alpha)u = y_\Omega \\ \Rightarrow p = \frac{\gamma}{\beta}u \\ \Rightarrow y = \beta\mathcal{A}^{-\alpha}u \end{array}$$

Thus, we find the following necessary optimality conditions:

$$u = \underbrace{\left(\beta\mathcal{A}^{-\alpha} + \frac{\gamma}{\beta}\mathcal{A}^\alpha\right)^{-1}}_{G_1} y_\Omega$$

for the control u, and

$$y = \beta \underbrace{\mathcal{A}^{-\alpha}}_{G_3} u = \underbrace{\left(\mathcal{I} + \frac{\gamma}{\beta^2}\mathcal{A}^{2\alpha}\right)^{-1}}_{G_2} y_\Omega$$

for the state y.

The spectral fractional Laplacian

Let $\Omega \in \mathbb{R}^d$ be a bounded Lipschitz domain, and let λ_k and e_k be the eigenvalues and the corresponding eigenfunctions of the Laplacian, i. e.

$$\begin{aligned} -\Delta e_k &= \lambda_k e_k && \text{in } \Omega, \\ e_k &= 0 && \text{on } \partial\Omega, \end{aligned}$$

and the functions e_k are an orthonormal basis of $L^2(\Omega)$. Then, for $\alpha \in [0,1]$ and a function $g \in H_0^1(\Omega)$

$$g = \sum_{k=1}^{\infty} a_k e_k,$$

we consider the operator

$$\mathcal{A}^\alpha g = \sum_{k=1}^{\infty} a_k \lambda_k^\alpha e_k.$$

The Riesz fractional Laplacian

For $\alpha \in (0,1)$, the fractional Laplacian $(-\Delta)^\alpha$ of a function $g : \mathbb{R}^d \to \mathbb{R}$ at a point $x \in \mathbb{R}^d$ is defined by

$$(-\Delta)^\alpha g(x) := C_{d,\alpha} \int_{\mathbb{R}^d} \frac{g(x) - g(y)}{\|x - y\|^{d+2\alpha}} \mathrm{d}y.$$

The Riesz fractional Laplacian

For $\alpha \in (0,1)$, the fractional Laplacian $(-\Delta)^\alpha$ of a function $g : \mathbb{R}^d \to \mathbb{R}$ at a point $x \in \mathbb{R}^d$ is defined by

$$(-\Delta)^\alpha g(x) := C_{d,\alpha} \int_{\mathbb{R}^d} \frac{g(x) - g(y)}{\|x - y\|^{d+2\alpha}} \mathrm{d}y.$$

- coincides with $\mathcal{A}^\alpha$ on $\mathbb{R}^d$, cf. details in: Lischke et al. (2018, arXiv:1801.09767)
- leads to multilevel Toeplitz structures on tensor grids (Ch. Vollmann, V. Schulz, CVS 2019)

General nonlocal operator

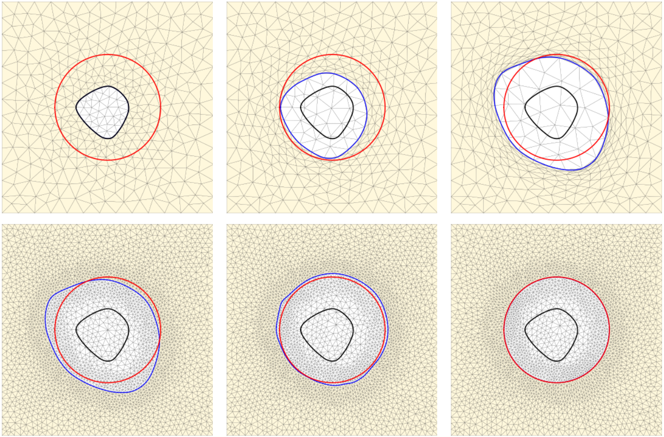
$$L_\phi g(x) = \int_\Omega g(x)\phi(x,y) - g(y)\phi(y,x)dy$$

- nonlocal calculus developed by Max Gunzburger et. al.
- unstructured discretization and shape optimization discussed in
  - Ch. Vollmann: *Nonlocal Models with Truncated Interaction Kernels– Analysis, Finite Element Methods and Shape Optimization*, PhD dissertation Trier University, 2019
  - V. Schulz, Ch. Vollmann: *Shape optimization for interface identification in nonlocal models*, arXiv:1909.08884, 2019

$\rightarrow$ more details in 2nd RICAM workshop in two weeks...

Example of nonlocal shape numerics

ALGORITHMIC OPTIMIZATION



A tale of two fractional Laplacians

On a bounded domain, the operators are different.

Theorem, Servadei/Valdinoci (2014)

The operators $\mathcal{A}^\alpha$ and $(-\Delta)^\alpha$ are not the same, since they have different eigenvalues and eigenfunctions (with respect to Dirichlet boundary conditions). In particular,

- the first eigenvalues of $(-\Delta)^\alpha$ is strictly less than that of $\mathcal{A}^\alpha$
- the eigenfunctions of $(-\Delta)^\alpha$ are only Hölder continuous up to the boundary, in contrast with those of $\mathcal{A}^\alpha$, which are as smooth up to the boundary as the boundary allows.

A tale of two fractional Laplacians

On a bounded domain, the operators are different.

Theorem, Servadei/Valdinoci (2014)

The operators $\mathcal{A}^\alpha$ and $(-\Delta)^\alpha$ are not the same, since they have different eigenvalues and eigenfunctions (with respect to Dirichlet boundary conditions). In particular,

- the first eigenvalues of $(-\Delta)^\alpha$ is strictly less than that of $\mathcal{A}^\alpha$
- the eigenfunctions of $(-\Delta)^\alpha$ are only Hölder continuous up to the boundary, in contrast with those of $\mathcal{A}^\alpha$, which are as smooth up to the boundary as the boundary allows.

Lischke et al. (2018, arXiv:1801.09767): Numerical tests for the error between $\mathcal{A}^\alpha$ and $(-\Delta)^\alpha$.

Yet another fractional operator

$$ {}^{R}L^{\beta} := -\,{}^{R}D_{x_1}^{\beta_1} - {}^{R}D_{x_2}^{\beta_2}\,, \quad \beta_1, \beta_2 \in (1,2) $$

${}^{R}D_{x_i}^{\beta_i}$: 1D Riemann-Liouville derivative

Yet another fractional operator

ALGORITHMIC OPTIMIZATION

$$^{R}L^{\beta} := -\,^{R}D_{x_1}^{\beta_1} - \,^{R}D_{x_2}^{\beta_2}\,, \quad \beta_1, \beta_2 \in (1,2)$$

$$^{R}D_{x_i}^{\beta_i} : \text{1D Riemann-Liouville derivative}$$

This operator is considered in the related publications:

- S. Dolgov, J. W. Pearson, D. V. Savostyanov, M. Stoll:*Fast tensor product solvers for optimization problems with fractional differential equations as constraints*, Applied Mathematics and Computation, 2016
- T. Breiten, V. Simoncini, M. Stoll: *Low-rank solvers for fractional differential equations*, ETNA 2016
- S. Pougkakiotis, J. W. Pearson, S. Leveque, J. Gondzio: *Fast Solution Methods for Convex Fractional Differential Equation Optimization*, arXiv:1907.13428, 2019

Note:

$$(-\Delta)^{\alpha} \neq \mathcal{A}^{\alpha} \neq \,^{R}L^{(2\alpha,2\alpha)}$$

Separation of variables and the Laplacian

For a function with separated variables, the Laplacian can be applied in one dimension: Let

$$g : (0,1)^2 \to \mathbb{R},\; g(x_1, x_2) = g_1(x_1)g_2(x_2).$$

Then

$$-\Delta g(x_1, x_2) = -g_1''(x_1)g_2(x_2) - g_1(x_1)g_2''(x_2).$$

Separation of variables and the Laplacian

For a function with separated variables, the Laplacian can be applied in one dimension: Let

$$g : (0,1)^2 \to \mathbb{R},\ \ g(x_1, x_2) = g_1(x_1)g_2(x_2).$$

Then

$$-\Delta g(x_1, x_2) = -g_1''(x_1)g_2(x_2) - g_1(x_1)g_2''(x_2).$$

The case

$$g(x_1, x_2) = \sum_{j=1}^{S} g_1^{(j)}(x_1)g_2^{(j)}(x_2)$$

follows immediately.

Low rank and the discretized Laplacian

Now consider (FEM/FDM) discretizations $A_{(1)}$, $A_{(2)}$ of the one-dimensional Laplacian in the first and variables, respectively. The a discretization of $\mathcal{A} = -\Delta$ on $(0,1)^2$ is given by

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

Let $\mathbf{g} = \mathbf{g}_1 \otimes \mathbf{g}_2$ be the function g evaluated on the grid. It follows that

$$A\mathbf{g} = A_{(1)}\mathbf{g}_1 \otimes \mathbf{g}_2 + \mathbf{g}_1 \otimes A_{(2)}\mathbf{g}_2.$$

$\Rightarrow$ The Laplacian $\mathcal{A}$ admits a tensor product representation with *Kronecker rank* 2.

Low rank and the discretized Laplacian

Therefore, the discrete Laplacian A can be applied efficiently to function g, given by

$$g(x_1, x_2) = \sum_{j=1}^{S} g_1^{(j)}(x_1) g_2^{(j)}(x_2).$$

We get

$$A\mathbf{g} = \sum_{j=1}^{S} \left(A_{(1)} \mathbf{g}_1^{(j)} \otimes \mathbf{g}_2^{(j)} + \mathbf{g}_1^{(j)} \otimes A_{(2)} \mathbf{g}_2^{(j)} \right),$$

for discretizations $\mathbf{g}_i^{(j)}$ of $g_i^{(j)}$.

From now on: Let A^α be the discretization of $\mathcal{A}^\alpha$.

Low rank for solution operators?

Gavrilyuk/Hackbusch/Khoromskij (2005): using the integral representation

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

(based on the Laplace transform), it can be shown that (discretized) $A^{-\alpha}$ has exponentially decaying singular values, and thus admits a low Kronecker rank approximation.

Proof: Sinc quadrature $\rightarrow$ exponentially decaying coefficients.

Low rank for solution operators?

Gavrilyuk/Hackbusch/Khoromskij (2005): using the integral representation

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

(based on the Laplace transform), it can be shown that (discretized) $A^{-\alpha}$ has exponentially decaying singular values, and thus admits a low Kronecker rank approximation.

Proof: Sinc quadrature $\rightarrow$ exponentially decaying coefficients.

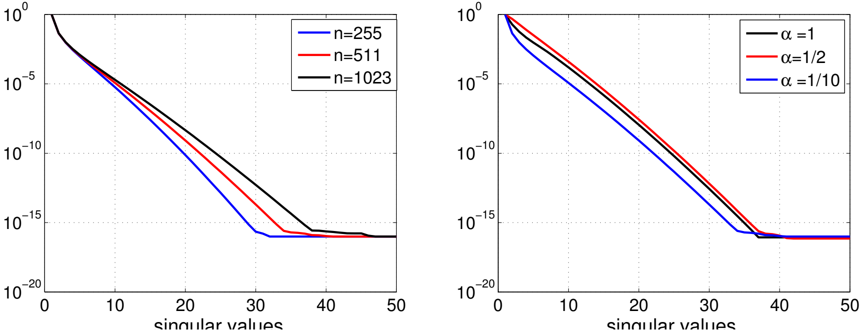
Open question: Similar results for

$$G_1 := \left(\beta A^{-\alpha} + \tfrac{\gamma}{\beta} A^{\alpha}\right)^{-1}, \quad G_2 := \left(I + \tfrac{\gamma}{\beta^2} A^{2\alpha}\right)^{-1}.$$

But: Numerical experiments (later) show similar behaviour for all three operators.

Singular value decay for $A^{-\alpha}$

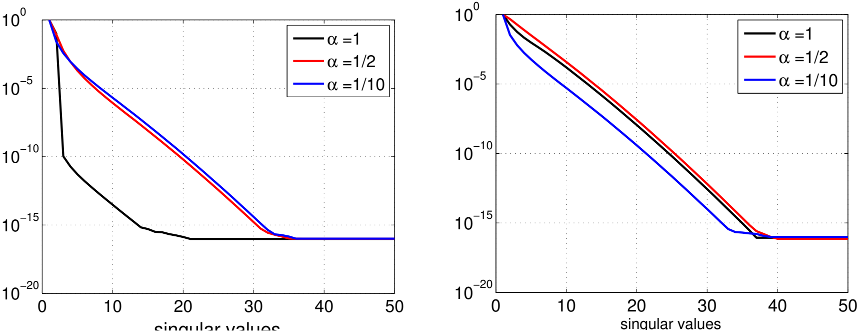
ALGORITHMIC OPTIMIZATION



Decay of singular values with $\alpha = 1$ in vs. n (left);
singular values vs. $\alpha > 0$ with fixed $n = 511$ (right).

Decay for discretized solution operators

ALGORITHMIC OPTIMIZATION



Decay of singular values of G_1 (left) and G_2 (right) vs. $\alpha = 1, 1/2, 1/10$ for $n = 511$.

Laplacian eigenvalue decomposition

Let $A_{(i)}$ be the discretized one-dimensional Laplacian on a uniform grid. Then $A_{(i)}$ is diagonalized in the sine basis, i. e.

$$A_{(i)} = F_i^* \Lambda_{(i)} F_i,$$

where $\Lambda_{(i)} = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$, and the action of F_i (F_i^*) is given by the (inverse) sine transform.

Thus, we can write

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

and, for a function f applied to A, we get

$$f(A) = (F_1^* \otimes F_2^*) f(\Lambda)(F_1 \otimes F_2).$$

Data in low rank format

Now assume that $f(A)$ may be approximated by a linear combination of Kronecker rank 1 matrices. Then, to approximate $f(A)$, it is sufficient to approximate $f(\Lambda)$ (e. g. by a truncated SVD).

Assume we have a decomposition

$$f(\Lambda) = \sum_{k=1}^{R} \operatorname{diag}\left(\mathbf{u}_1^{(k)} \otimes \mathbf{u}_2^{(k)}\right),$$

and let

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

Fast application of the fractional Laplacian via FFT

Now we can compute a matrix-vector product

$$f(A)\mathbf{x} = (F_1^* \otimes F_2^*)\left(\sum_{k=1}^{R} \operatorname{diag}\left(\mathbf{u}_1^{(k)} \otimes \mathbf{u}_2^{(k)}\right)\right)(F_1 \otimes F_2)\left(\sum_{j=1}^{S} \mathbf{x}_1^{(j)} \otimes \mathbf{x}_2^{(j)}\right)$$

$$= \sum_{k=1}^{R}\sum_{j=1}^{S} F_1^*\left(\mathbf{u}_1^{(k)} \odot F_1\mathbf{x}_1^{(j)}\right) \otimes F_2^*\left(\mathbf{u}_2^{(k)} \odot F_2\mathbf{x}_2^{(j)}\right),$$

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

Fast application of the fractional Laplacian via FFT

Now we can compute a matrix-vector product

$$f(A)\mathbf{x} = (F_1^* \otimes F_2^*)\left(\sum_{k=1}^{R} \operatorname{diag}\left(\mathbf{u}_1^{(k)} \otimes \mathbf{u}_2^{(k)}\right)\right)(F_1 \otimes F_2)\left(\sum_{j=1}^{S} \mathbf{x}_1^{(j)} \otimes \mathbf{x}_2^{(j)}\right)$$

$$= \sum_{k=1}^{R}\sum_{j=1}^{S} F_1^*\left(\mathbf{u}_1^{(k)} \odot F_1\mathbf{x}_1^{(j)}\right) \otimes F_2^*\left(\mathbf{u}_2^{(k)} \odot F_2\mathbf{x}_2^{(j)}\right),$$

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

$\Rightarrow$ Can be computed in $\mathcal{O}(RSn\log n)$ flops

Fast application of the fractional Laplacian via FFT

Now we can compute a matrix-vector product

$$f(A)\mathbf{x} = (F_1^* \otimes F_2^*)\left(\sum_{k=1}^{R} \operatorname{diag}\left(\mathbf{u}_1^{(k)} \otimes \mathbf{u}_2^{(k)}\right)\right)(F_1 \otimes F_2)\left(\sum_{j=1}^{S} \mathbf{x}_1^{(j)} \otimes \mathbf{x}_2^{(j)}\right)$$

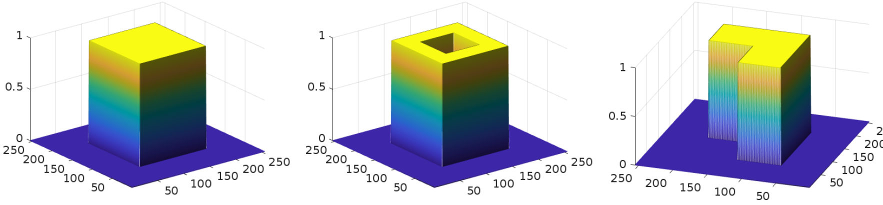
$$= \sum_{k=1}^{R}\sum_{j=1}^{S} F_1^*\left(\mathbf{u}_1^{(k)} \odot F_1\mathbf{x}_1^{(j)}\right) \otimes F_2^*\left(\mathbf{u}_2^{(k)} \odot F_2\mathbf{x}_2^{(j)}\right),$$

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

$\Rightarrow$ Can be computed in $\mathcal{O}(RSn\log n)$ flops

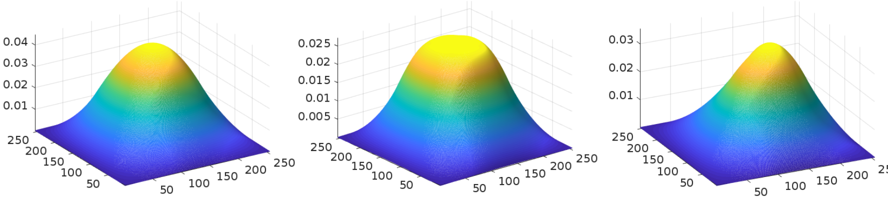
Computational time on a 2^{15}-by-2^{15} grid ($> 10^9$): under 1 s
Preprocessing time: 300 s

Shapes of the right hand side y_Ω

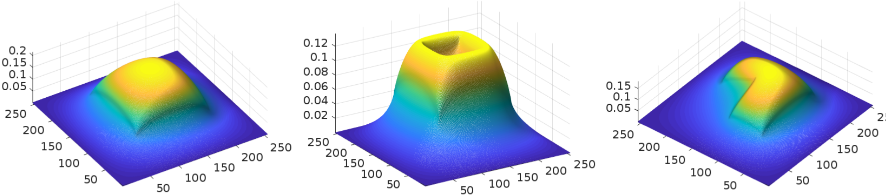


Shapes of the right-hand side y_Ω computed with n=255.

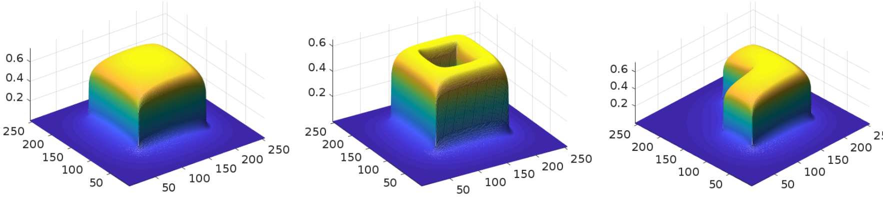
Computed u with $\alpha = 1$



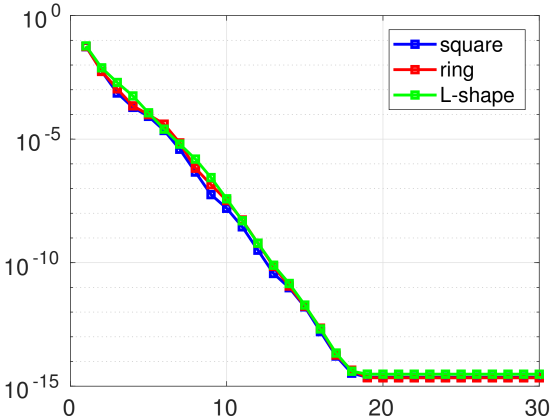
Computed u with $\alpha = 1/2$



Computed u with $\alpha = 1/10$



Decay of the error with respect to rank



Error norm of optimal solution in full rank vs low rank format with $n = 255$ and $\alpha = 1/2$.

PCG iterations for 2D and $\alpha = 1/10$

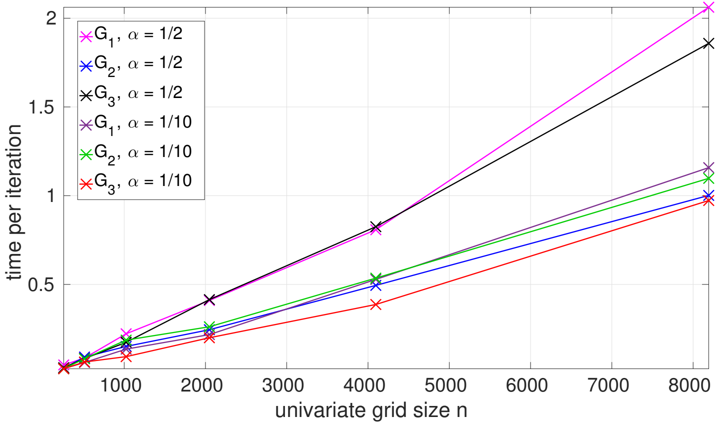
ALGORITHMIC OPTIMIZATION

	G_1				G_2				G_3			
r \ n	256	512	1024	2048	256	512	1024	2048	256	512	1024	2048
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

- univariate grid size n
- preconditioner rank r
- convergence to relative residual of 10^{-6}

CPU times for 2D

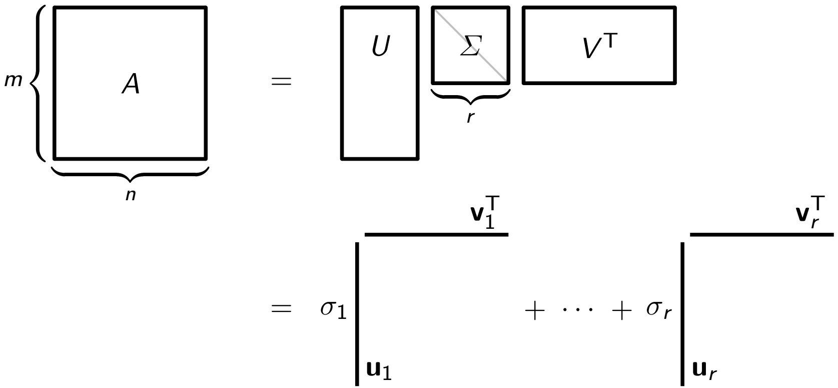
ALGORITHMIC OPTIMIZATION



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

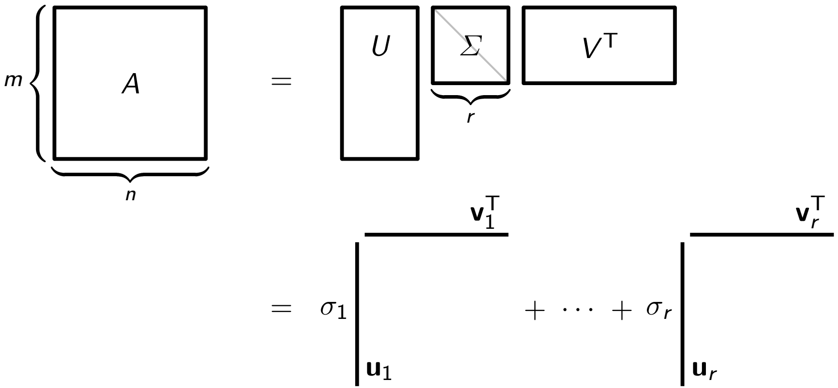
3D: from SVD ...

Low-rank $m \times n$ matrix given by the *singular value decomposition*:



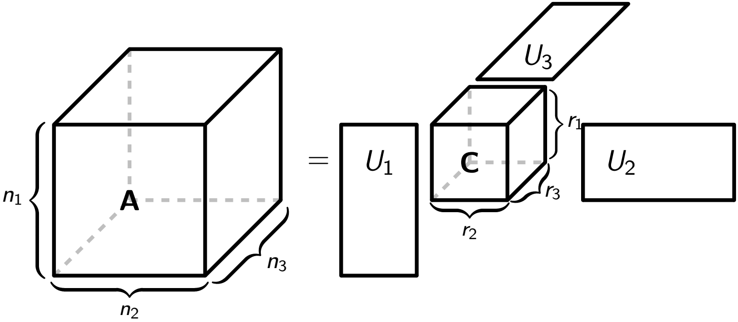
3D: from SVD ...

Low-rank $m \times n$ matrix given by the *singular value decomposition*:

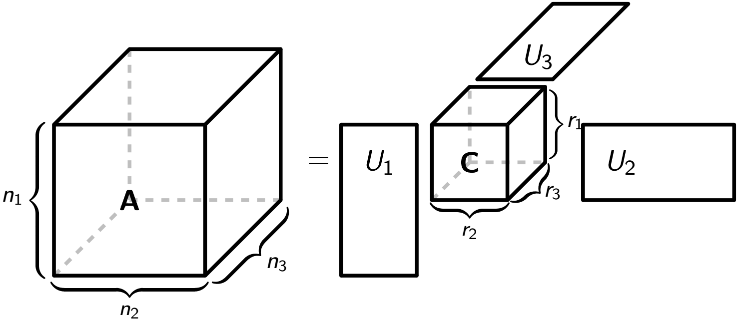


Formally: $A = U\Sigma V^{\mathsf{T}} = \sum_{k=1}^{r} \sigma_k \, \mathbf{u}_k \, \mathbf{v}_k^{\mathsf{T}} =: \sum_{k=1}^{r} \sigma_k \, \mathbf{u}_k \circ \mathbf{v}_k$

...over the Tucker format...

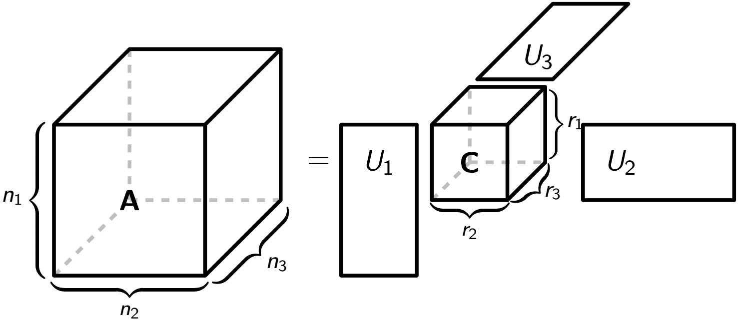


...over the Tucker format...



Formally: $A = \mathbf{C} \times_1 U_1 \cdots \times_d U_d = \mathbf{C} \bigtimes_{i=1}^{d} U_i$

...over the Tucker format...



Formally: $A = \mathbf{C} \times_1 U_1 \cdots \times_d U_d = \mathbf{C} \bigtimes_{i=1}^{d} U_i$

Storage cost: $\mathcal{O}(r^d + dnr)$

... to the truncated HOSVD

A Tucker decomposition of a rank-**r** tensor can be computed by computing an SVD for each tensor mode.

... to the truncated HOSVD

ALGORITHMIC OPTIMIZATION

A Tucker decomposition of a rank-$\mathbf{r}$ tensor can be computed by computing an SVD for each tensor mode.

We are interested in computing rank-$\mathbf{r}$ approximations.

... to the truncated HOSVD

A Tucker decomposition of a rank-$\mathbf{r}$ tensor can be computed by computing an SVD for each tensor mode.

We are interested in computing rank-$\mathbf{r}$ approximations. To this end, let $\mathrm{P}^i_{r_i}$ be the best rank-r_i approximation operator in the ith mode. Then the rank-$\mathbf{r}$ truncated HOSVD operator $\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}$ is given by

$$\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}\,\mathbf{A} := \mathrm{P}^1_{r_1}\cdots\mathrm{P}^d_{r_d}\,\mathbf{A}.$$

... to the truncated HOSVD

A Tucker decomposition of a rank-$\mathbf{r}$ tensor can be computed by computing an SVD for each tensor mode.

We are interested in computing rank-$\mathbf{r}$ approximations. To this end, let $\mathrm{P}^{i}_{r_i}$ be the best rank-r_i approximation operator in the ith mode. Then the rank-$\mathbf{r}$ truncated HOSVD operator $\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}$ is given by

$$\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}\,\mathbf{A} := \mathrm{P}^{1}_{r_1}\cdots\mathrm{P}^{d}_{r_d}\,\mathbf{A}.$$

Remark:

... to the truncated HOSVD

A Tucker decomposition of a rank-$\mathbf{r}$ tensor can be computed by computing an SVD for each tensor mode.

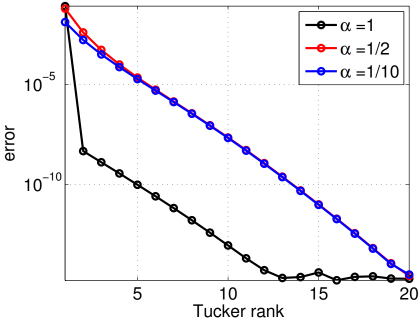
We are interested in computing rank-$\mathbf{r}$ approximations. To this end, let $\mathrm{P}^i_{r_i}$ be the best rank-r_i approximation operator in the ith mode. Then the rank-$\mathbf{r}$ truncated HOSVD operator $\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}$ is given by

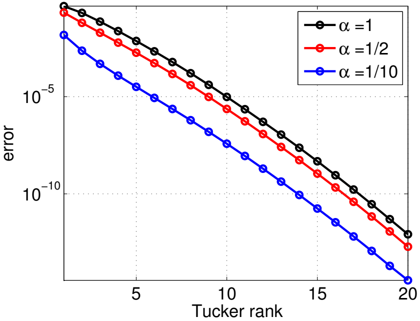
$$\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}\,\mathbf{A} := \mathrm{P}^1_{r_1} \cdots \mathrm{P}^d_{r_d}\,\mathbf{A}.$$

Remark:

- $\mathrm{P}^{\mathrm{HO}}_{\mathbf{r}}$ only gives a *quasi-best* rank-$\mathbf{r}$ approximation

Tucker-ALS approximation in 3D





Tucker-ALS approximation error of G_1 (left) and G_2 (right) vs. $\alpha = 1, 1/2, 1/10$ for $d = 3$.

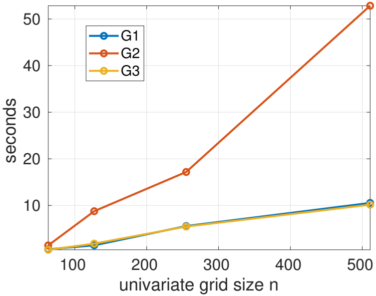
PCG iterations for 3D and $\alpha = 1/2$

ALGORITHMIC OPTIMIZATION

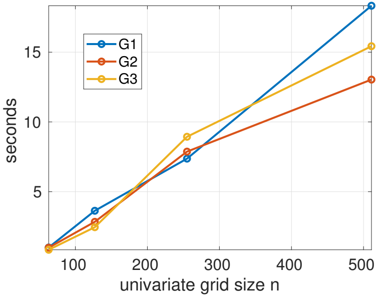
	G_1				G_2				G_3			
r \ n	64	128	256	512	64	128	256	512	64	128	256	512
4	1	2	1	1	1	6	1	2	1	2	1	1
5	1	1	1	2	1	1	8	4	1	1	1	2
6	1	1	1	1	2	2	1	1	1	1	1	1
7	1	3	1	2	1	1	5	4	1	2	1	2
8	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	2	1	6	5	4	1	1	1	2
10	1	1	1	1	1	6	1	1	1	1	1	1

- univariate grid size n
- preconditioner rank r
- convergence to relative residual of 10^{-6}

CPU times for 3D



$\alpha = 1/2, \; r = 4$



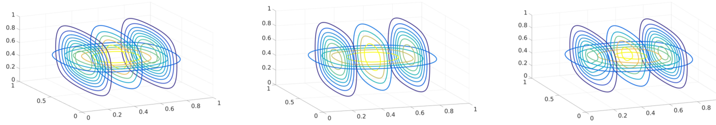
$\alpha = 1/10, \; r = 7$

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

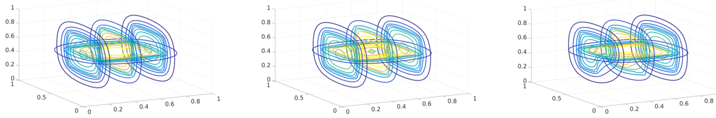
3D level curves

ALGORITHMIC OPTIMIZATION

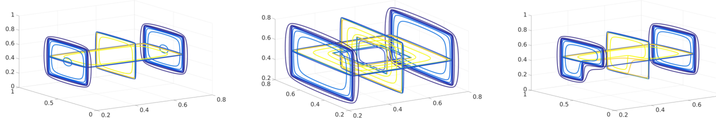
$\alpha = 1$:



$\alpha = \frac{1}{2}$:



$\alpha = \frac{1}{10}$:



Solutions u for analogous right-hand sides ($n = 255$).

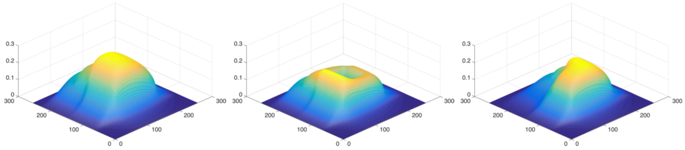
Generalization to varying coefficient

Generalization to different elliptic equation [Schmitt 2019]:

$$(-\mathrm{div}(\mathbb{A}\,\mathrm{grad}))^\alpha y = \beta u\,, \qquad \mathbb{A}(x_1, x_2) = \begin{bmatrix} a_1(x_1) & 0 \\ 0 & a_2(x_2) \end{bmatrix}$$

- needs affordable preparation step for 1D eigenvalues/vectors
- similarly good numerical complexity
- more general $\mathbb{A}$ in preparation

$\alpha = \frac{1}{2}$:



Publications

- Gennadij Heidel: *Optimization in Tensor Spaces for Data Science and Scientific Computing*, PhD Dissertation, Trier University 2019
- Gennadij Heidel, Venera Khoromskaia, Boris N. Khoromskij, Volker Schulz: *Tensor approach to optimal control problems with fractional d-dimensional elliptic operator in constraints*, 2018, arXiv:1809.01971 (revised version submitted to SICON)
- Britta Schmitt: *Niedrigrang-Tensorapproximation bei heterogen verteilten Optimalsteuerungsproblemen*, Masters's Thesis, Trier University, 2019
- Britta Schmitt, Venera Khoromskaia, Boris N. Khoromskij, Volker Schulz: *Low-Rank Tensor Approximation of Heterogenously Distributed Optimal Control Problems* (in preparation)

Conclusions

- Solution of a nonlocal PDE on a tensor grid in $\mathcal{O}(RSn \log n) \ll \mathcal{O}(n^d)$
- Numerical complexity independent of d
- Even for $\alpha = 1$ the methodology is significantly faster than multigrid
- Numerical results convincing; theoretical justification w.r.t. low rank still open