

Vico-Greengard-Ferrando quadratures in the tensor solver for integral equations

V. Khrulkov¹, M. Rakhuba¹, I. Oseledets^{1,2}

¹Skolkovo Institute of Science and Technology, Russia

²Institute of Numerical Mathematics, Russia

Abstract— Convolution with Green’s function of a differential operator appears in a lot of applications e.g. Lippmann-Schwinger integral equation. Algorithms for computing such are usually non-trivial and require non-uniform mesh. However, recently Vico, Greengard and Ferrando developed method for computing convolution with smooth functions with compact support with spectral accuracy, requiring nothing more than Fast Fourier Transform (FFT). Their approach is very suitable for the low-rank tensor implementation which we develop using Quantized Tensor Train (QTT) decomposition.

1. INTRODUCTION

In this paper we propose an algorithm for computing an approximation to a convolution

$$h(\mathbf{r}') = \int_{\mathbb{R}^3} g(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}) d\mathbf{r},$$

where G is a continuous Green’s function of some PDE. It has been shown recently in [1] that discrete approximation to h yielding spectral accuracy for compactly supported smooth functions ρ can be constructed using relatively simple idea. Idea is based on the fact that continuous Green’s function can be replaced by the truncated one, Fourier transform of which is in $C^\infty(\mathbb{R}^3)$. Applying convolution theorem one gets answer as a Fourier type integral of a rapidly decreasing function from $C^\infty(\mathbb{R}^3)$ which is then discretized on finer grid and evaluated using Discrete Fourier Transform (DFT). This scheme achieves spectral order of accuracy due to the superalgebraic approximation of the continuous Fourier transform by DFT and has been shown to also give second order with small constant for continuous functions. We implemented this scheme using data compression via Quantized Tensor Train (QTT). Approach is straightforward – we compute discrete truncated Green’s function using formulas from [1] and convert it to QTT-format. Then we use standard algorithms in QTT which allow for logarithmic storage with respect to the grid size and accuracy of approximation. Moreover, discrete convolution and solution of integral equations are then performed also with logarithmic complexity by the algorithms described in [2],[3]. Paper is organized as follows:

- We briefly describe the algorithm developed in Section 2
- We present necessary definitions and algorithms from TT theory in Section 3.1
- We describe TT implementation of the algorithm in Section 3.2
- We present numerical results in Section 4

The main reference for Vico-Greengard-Ferrando quadrature is the original paper [1], in the presentation of the algorithm we follow [4].

2. ALGORITHM

We will present the scheme described in [1] in a way which is most suitable for transitioning to TT-format. Let $\rho(\mathbf{r})$ be a smooth function such that $\text{supp } \rho \subset D = [0, L_1] \times [0, L_2] \times [0, L_3]$ and we are interested in computing

$$h(\mathbf{r}) = \int_{\mathbb{R}^3} g(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}', \quad (1)$$

where $g(\mathbf{r})$ is a Green's function of some differential operator. Let us assume that $g(\mathbf{r})$ depends only on $r = |\mathbf{r}|$. If we seek restriction of the solution $h(\mathbf{r})$ to D then $|r - r'|$ in (1) doesn't exceed $L = \sqrt{L_1^2 + L_2^2 + L_3^2}$. Thus if we replace $g(\mathbf{r})$ by $g_L(\mathbf{r}) = g(\mathbf{r}) \operatorname{rect} \frac{r}{2L}$,

$$\int_{\mathbb{R}^3} g(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}' = \int_{\mathbb{R}^3} g_L(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}'.$$

Advantage of using $g_L(\mathbf{r})$ is that since it has compact support its Fourier transform $G_L(\mathbf{s})$ is in $C^\infty(\mathbb{R}^3)$, and is straightforward to compute for many differential operators. We mainly focus on Helmholtz differential operator $\nabla^2 + k^2$ for which the following formula holds:

$$\begin{aligned} g(\mathbf{r}) &= \frac{e^{ikr}}{4\pi r}, \\ G_L(\mathbf{s}) &= \int_{\mathbb{R}^3} g(\mathbf{r}) \operatorname{rect}\left(\frac{r}{2L}\right) e^{i\mathbf{s}\cdot\mathbf{r}} d\mathbf{r} \\ &= \frac{-1 + e^{iLk}(\cos Ls - i\frac{k}{s} \sin Ls)}{(k - s)(k + s)}. \end{aligned} \quad (2)$$

It is easy to check that $G_L(\mathbf{s})$ is indeed smooth and nonsingular. If we denote Fourier transform of $\rho(\mathbf{r})$ by $\widehat{\rho}(\mathbf{s})$ we obtain the final formula:

$$h(\mathbf{r}) = \left(\frac{1}{2\pi}\right)^3 \int_{\mathbb{R}^3} e^{i\mathbf{s}\cdot\mathbf{r}} \widehat{\rho}(\mathbf{s}) G_L(\mathbf{s}) d\mathbf{s}. \quad (3)$$

This integral is then discretized using trapezoidal rule and computed using DFT. However to cancel slightly oscillatory behavior of $\widehat{G}_L(s)$ zero padding by a factor of at least 3 is required (we will use factor 4 to keep grid size being a power of 2 - for the analysis see [4]). Suppose that domain D is discretized using uniform grid with N_i nodes in corresponding dimensions and function ρ is sampled on this grid yielding an array ρ_{ijk} i.e.

$$\rho_{ijk} := \rho(ih_1, jh_2, kh_3), \quad h_i = \frac{L_i}{N_i}.$$

Algorithm then is summarized as following

Algorithm 1 Basic Vico-Greengard-Ferrando quadrature

- 1: Zero pad ρ_{ijk} by a factor of 4 and compute 3d FFT defining $\widehat{\rho}(\mathbf{s})$ for $\mathbf{s} = \frac{\pi}{2}(\frac{s_1}{L_1}, \frac{s_2}{L_2}, \frac{s_3}{L_3})$, $s_i \in \{-2N_i, \dots, 2N_i - 1\}$.
 - 2: Evaluate $G_L(\mathbf{s})$ for $L = \sqrt{L_1^2 + L_2^2 + L_3^2}$ for \mathbf{s} defined above, and multiply elementwise by $\widehat{\rho}(\mathbf{s})$.
 - 3: Perform 3d IFFT on the array defined above and truncate the result keeping first $N_1 \times N_2 \times N_3$ entries, obtaining approximation to $h(\mathbf{r})$ on the grid.
-

2.1. Convolution form of the algorithm

Let us write Algorithm 1 more explicitly. All steps together can be represented as follows

$$h_{ijk} = \frac{1}{4N_1} \frac{1}{4N_2} \frac{1}{4N_3} \sum_{s_1=-2N_1}^{2N_1-1} \sum_{s_2=-2N_2}^{2N_2-1} \sum_{s_3=-2N_3}^{2N_3-1} G_L(\mathbf{s}) \widehat{\rho}(\mathbf{s}) e^{2\pi i \frac{s_1 i}{4N_1}} e^{2\pi i \frac{s_2 j}{4N_2}} e^{2\pi i \frac{s_3 k}{4N_3}}, \quad (4)$$

where

$$\widehat{\rho}(\mathbf{s}) = \sum_{i'=0}^{4N_1-1} \sum_{j'=0}^{4N_2-1} \sum_{k'=0}^{4N_3-1} \rho_{i'j'k'} e^{-2\pi i \frac{s_1 i'}{4N_1}} e^{-2\pi i \frac{s_2 j'}{4N_2}} e^{-2\pi i \frac{s_3 k'}{4N_3}}. \quad (5)$$

By plugging (5) into (4) and by changing the order of summation it is easy to see that

$$h_{ijk} = \sum_{i'=0}^{4N_1-1} \sum_{j'=0}^{4N_2-1} \sum_{k'=0}^{4N_3-1} G_{i-i', j-j', k-k'}^M \rho_{i'j'k'}, \quad (6)$$

where

$$G_{i-i',j-j',k-k'}^M = \frac{1}{4N_1} \frac{1}{4N_2} \frac{1}{4N_3} \sum_{s_1=-2N_1}^{2N_1-1} \sum_{s_2=-2N_2}^{2N_2-1} \sum_{s_3=-2N_3}^{2N_3-1} G_L(\mathbf{s}) e^{2\pi i \frac{s_1(i-i')}{4N_1}} e^{2\pi i \frac{s_2(j-j')}{4N_2}} e^{2\pi i \frac{s_3(k-k')}{4N_3}}. \quad (7)$$

Moreover, since $\rho_{i'j'k'}$ is 0 for $i' \geq N_1, j' \geq N_2, k' \geq N_3$ (see step 1 of Algorithm 1) and we truncate the result, formula (6) simplifies and finally:

$$h_{ijk} = \sum_{i'=0}^{N_1-1} \sum_{j'=0}^{N_2-1} \sum_{k'=0}^{N_3-1} G_{i-i',j-j',k-k'}^M \rho_{i'j'k'}. \quad (8)$$

We see that (8) takes the form of a discrete aperiodic convolution with discrete Green's function G^M (which we will call *mollified Green's function*). One can notice that to fully determine G^M it is sufficient to run Algorithm 1 once for a special right hand side $\rho_{ijk} = \delta_{i0}\delta_{j0}\delta_{k0}$. Formula (8) plays essential role in the further analysis. Multiplication by multilevel Toeplitz matrix generated by G^M can be performed with logarithmic complexity in QTT format as described in [2], and we discuss necessary definitions and algorithms in the next section.

3. LOW-RANK TENSOR APPROACH

3.1. TT and QTT formats

To understand the QTT format let us start with describing the TT-format, which is a nonlinear low-parametric representation of multidimensional arrays, called *tensors*. Tensor $\mathcal{X} \in \mathbb{C}^{n_1 \times n_2 \times \dots \times n_d}$ is said to be in the TT-format if it represents as

$$\mathcal{X}_{i_1 i_2 \dots i_d} = X^{(1)}(i_1) X^{(2)}(i_2) \dots X^{(d)}(i_d), \quad (9)$$

where $X^{(k)}(i_k) \in \mathbb{C}^{r_{k-1} \times r_k}$, $r_0 = r_d = 1$, $i_k = 1, \dots, n_k$. Matrices $X^{(k)}$ are called *TT-cores* and r_k are called *TT-ranks*. Notice that if $r = \max_k r_k$ is small, then there is a significant compression to store \mathcal{X} . Indeed, initial tensor requires storing n^d parameters, while to store its TT-representation only $\mathcal{O}(dn^2)$ parameters are needed.

In fact, one could use TT representation to store and to work with arising in Algorithm 1 3-dimensional arrays. However we will use a more sophisticated approach called QTT format, which allows for additional storage reduction compared to TT. QTT format is the following modification of the TT format. First we assume that $d = 3$, $n_i = 2^{l_i}$, $i = 1, 2, 3$. Then each “physical” index i, j, k is represented in the binary format, i.e.

$$i = i_1 + 2^1 i_2 + \dots + 2^{l_i-1} i_d, \quad i_m = 0, 1, \quad m = 1, \dots, d$$

and we have initial tensor ρ_{ijk} encoded as a $(l_1 + l_2 + l_3)$ -dimensional array $\tilde{\rho}$:

$$\rho_{ijk} \equiv \tilde{\rho}_{i_1 \dots i_{l_1} j_1 \dots j_{l_2} k_1 \dots k_{l_3}}.$$

TT decomposition of $\tilde{\rho}$ is called the QTT decomposition. The storage of the QTT decomposition is $\mathcal{O}(r^2(l_1 + l_2 + l_3)) = \mathcal{O}(r^2 \log n)$, so if $r = \max_i r_i$ is bounded, the total storage scales logarithmically. In practice tensors of exact low-rank rarely occur. Typically one fixes accuracy ϵ and tries to find best approximation with this accuracy. It has been shown that in some applications ranks grow as $r = \mathcal{O}(\log^\alpha \epsilon^{-1})$, $\alpha > 1$ [5, 6].

3.2. Translation of the algorithm to the QTT format

Computation of kernel G^M . To use (8) we first need to find G^M (7). For this purpose we run Algorithm 1 for $\rho_{ijk} = \delta_{i0}\delta_{j0}\delta_{k0}$. Precomputations are done in the full format, in other words we form the whole dense tensor G^M and utilize TT-SVD algorithm [7] to find its QTT representation. TT-SVD algorithm is based on the computation of SVD decompositions of tensors reshaped into full 2D matrices and therefore is quite expensive. In principle one could use DFT in the QTT format [8] to avoid forming full tensors. Unfortunately, we found that intermediate tensors arising in Algorithm 1 are of large rank. We will address this problem in our future work.

Computation of ρ . Tensor ρ can be already given in the QTT representation. This can happen, e.g. if we are running a certain iterative process involving computation of convolution (1) and all operations in this process are done within the QTT format. Otherwise, ρ_{ijk} can be approximated with logarithmic complexity by using the cross approximation method [9], which adaptively samples elements of a tensor. In this case we just need ρ_{ijk} be given as a function which returns value by given 3 indices i, j, k .

Computation of convolution $G^M * \rho$. Next goal is to find convolution of tensors G^M and ρ (8). The convolution can be considered as a multiplication of multilevel Toeplitz matrix generated by G^M and vector ρ . Matrices can also be represented in the TT and by analogy in the QTT format. The definition is similar to that of TT-tensor: given matrix (operator) $\mathcal{A}_{i_1 \dots i_d j_1 \dots j_d}$, which acts on vector $\mathcal{X}_{j_1 \dots j_d}$ such that

$$\mathcal{Y}_{i_1 \dots i_d} = \sum_{j_1 \dots j_d} \mathcal{A}_{i_1 \dots i_d j_1 \dots j_d} \mathcal{X}_{j_1 \dots j_d}.$$

its TT-decomposition is defined as

$$\mathcal{A}_{i_1 \dots i_d j_1 \dots j_d} = A^{(1)}(i_1, j_1) A^{(2)}(i_2, j_2) \dots A^{(d)}(i_d, j_d),$$

where $A^{(k)}(i_k, j_k) \in \mathbb{C}^{R_{k-1} \times R_k}$, $R_0 = R_d = 1$, $i_k = 1, \dots, n_k$. QTT decomposition of 3D operator $G_{i-i', j-j', k-k'}^M$ is defined by analogy with the QTT decomposition of a tensor – we quantize indices i, j, k and i', j', k' , group them pairwise and then compute TT decomposition:

$$\begin{aligned} G_{i-i', j-j', k-k'}^M &= G^{(1)}(i_1, i'_1) \dots G^{(l_1+2)}(i_{l_1+2}, i'_{l_1+2}) \\ &\quad G^{(l_1+3)}(j_1, j'_1) \dots G^{(l_1+l_2+4)}(j_{l_2+2}, j'_{l_2+2}) \\ &\quad G^{(l_1+l_2+5)}(k_1, k'_1) \dots G^{(l_1+l_2+l_3+6)}(k_{l_3+2}, k'_{l_3+2}). \end{aligned}$$

We use approach from [2] and analytically construct QTT representation of the induced multilevel Toeplitz matrix $G_{i-i', j-j', k-k'}^M$ given QTT representation of $G_{i,j,k}^M$. Then matrix-vector product (8) can be done in different ways. We used optimization procedure AMEn (alternating minimal energy method) [10, 11] which allows for rank adaptation compared to standard ALS (alternating least squares) [12] optimization which works with the representation of a given size.

4. NUMERICAL EXPERIMENTS

Approximating G^M using QTT. Firstly we show that using QTT representation greatly reduces number of degrees of freedom (DOF) of G^M . Suppose that tensor \mathcal{X} is given in the QTT format with ranks $\{r_1, \dots, r_d\}$. Then it is easy to count total number of DOF of \mathcal{X} :

$$\text{DOF}(\mathcal{X}) = 2(r_1 + r_d) + \sum_{i=2}^d 2r_{i-1}r_i.$$

By applying this formula to G^M computed as described in 3.2 for various values of k (while keeping $L = 1$) we obtained the following results (see Figure 1). This shows advantages of using QTT.

Solving scattering problems. To further test our approach we solve the Lippmann-Schwinger equation which is used for solving scattering problems:

$$\sigma(\mathbf{r}) + k^2 q(\mathbf{r}) \int_{\mathbb{R}^3} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \sigma(\mathbf{r}') d\mathbf{r}' = -k^2 q(\mathbf{r}) \phi^{\text{inc}}, \quad (10)$$

and then we find

$$\phi^{\text{scat}} = \int_{\mathbb{R}^3} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \sigma(\mathbf{r}') d\mathbf{r}'.$$

We used rounding by $\epsilon = 10^{-7}$ in our computations. Firstly we fixed $k = 1$ and $L = 32\pi$ and took $q(\mathbf{r})$ to be a 3D gaussian:

$$q(\mathbf{r}) = e^{\frac{-|\mathbf{r}-\mathbf{r}'|^2}{2a^2}}, \quad (11)$$

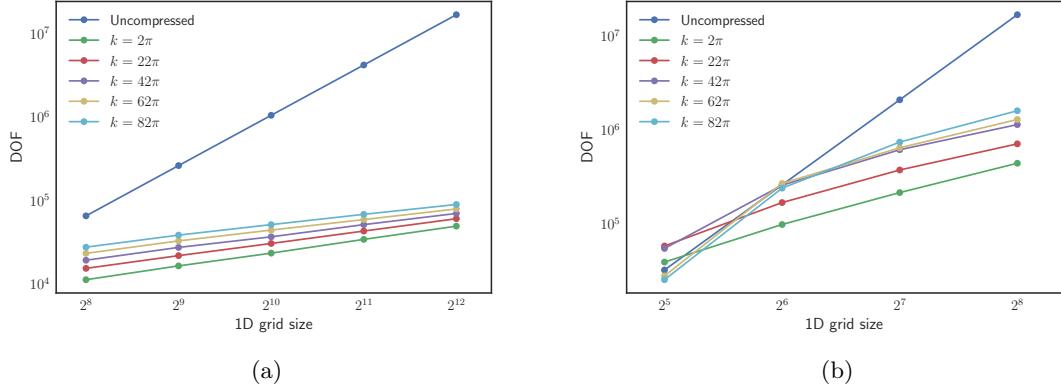


Figure 1: DOF of G^M in 2D (1a) and 3D (1b) for $\epsilon = 10^{-7}$ and $L = 1$

Table 1: Relative error (err) and effective rank (erank) for different grid sizes and rounding errors ϵ . Results are presented for two types of function q : Gaussian (11) and smoothed cube (12).

	3D grid size	32^3		64^3		128^3	
		err	erank	err	erank	err	erank
Gaussian	$\epsilon = 10^{-3}$	8e-1	14	8e-2	15	7e-2	14
	$\epsilon = 10^{-5}$	8e-1	24	4e-4	25	4e-4	26
	$\epsilon = 10^{-7}$	8e-1	30	5e-6	34	4e-6	38
Smoothed cube	$\epsilon = 10^{-3}$	8e-1	22	8e-2	35	6e-2	34
	$\epsilon = 10^{-5}$	8e-1	29	5e-4	57	5e-4	59
	$\epsilon = 10^{-7}$	8e-1	29	2e-5	73	6e-6	80

with $\mathbf{r}' = \left(\frac{L}{2}, \frac{L}{2}, \frac{L}{2}\right)$, and $a = \frac{L}{10}$, and

$$\phi^{\text{inc}}(x, y, z) = e^{ix}.$$

To solve arising systems in the QTT format we used AMEn [10, 11], which allows for rank adaptation. Taking the solution computed on a grid with size 256^3 as a reference we measured

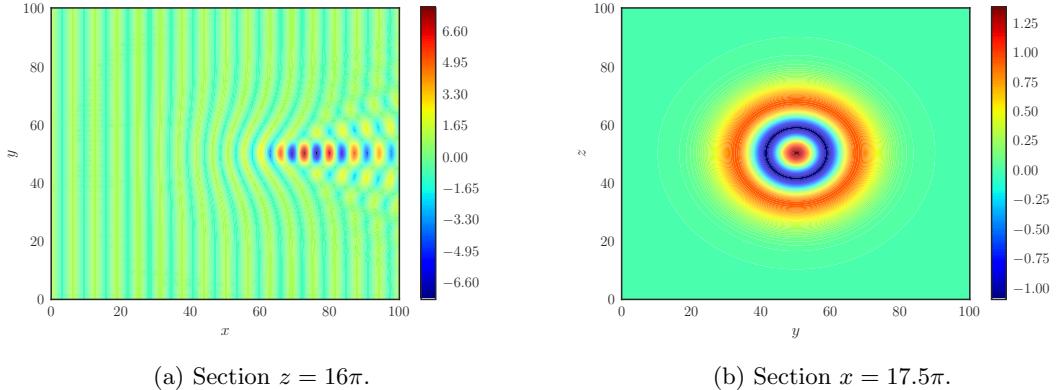


Figure 2: Scattering on the Gaussian with standard deviation 3.2π in $D = [0, 32\pi] \times [0, 32\pi] \times [0, 32\pi]$ for $k = 1$. Effective rank of the solution is equal to 39.

relative error of the solutions computed on smaller grids for various rounding parameters ϵ . Results are given in the Table 1. We see that error is roughly equal to the ϵ even for modest number of grid nodes per wavelength.

As a next experiment we performed the same computations for $q(r)$ representing smoothed cube:

$$q(\mathbf{r}) = e^{-0.5\left(\frac{|\mathbf{r}-\mathbf{r}'|}{a}\right)^8}, \quad (12)$$

for $\mathbf{r}' = (\frac{L}{2}, \frac{L}{2}, \frac{L}{2})$ and $a = \frac{L}{4}$. For the results see Figure 3.

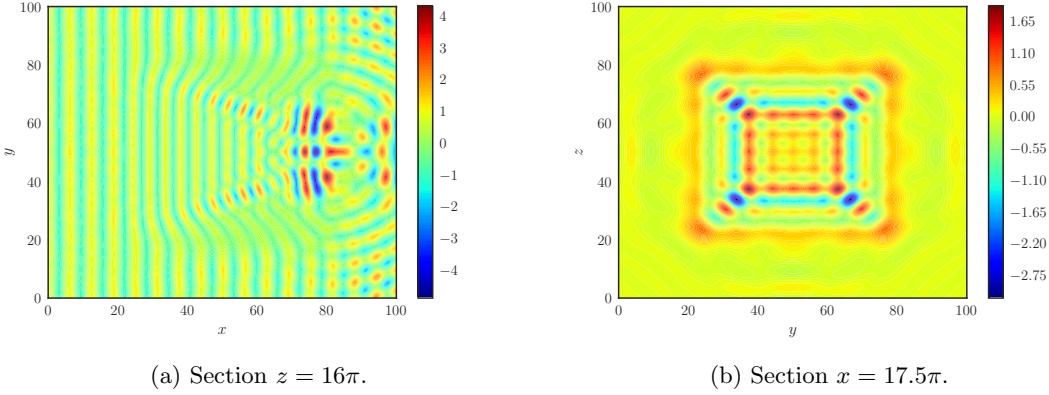


Figure 3: Scattering on the smoothed cube of size 8π in $D = [0, 32\pi] \times [0, 32\pi] \times [0, 32\pi]$ for $k = 1$. Effective rank of the solution is equal to 81.

Scattering problems on quasiperiodic structures. As a final test we took $q(\mathbf{r})$ to be a periodic grid of smoothed cubes (structures like this are extremely suitable for QTT computations). Namely, denoting $q(\mathbf{r})$ defined by formula (12) by $q_{\mathbf{r}',a}(\mathbf{r})$ we solve the equation (10) in the domain $[0, 2] \times [0, 1] \times [0, 10]$ on the grid $64 \times 64 \times 1024$ for $q(\mathbf{r})$ defined as

$$q(\mathbf{r}) = \sum_{i=1}^{19} \sum_{j=0}^3 \sum_{k=0}^1 q_{(0.25+0.5j, 0.25+0.5k, 0.5i), 0.1}(\mathbf{r})$$

and for $k = 4\pi$. ϕ^{inc} in this case is a plane wave propagating in z -direction

$$\phi^{\text{inc}}(x, y, z) = e^{4\pi iz}.$$

Figure 4 demonstrates our results.

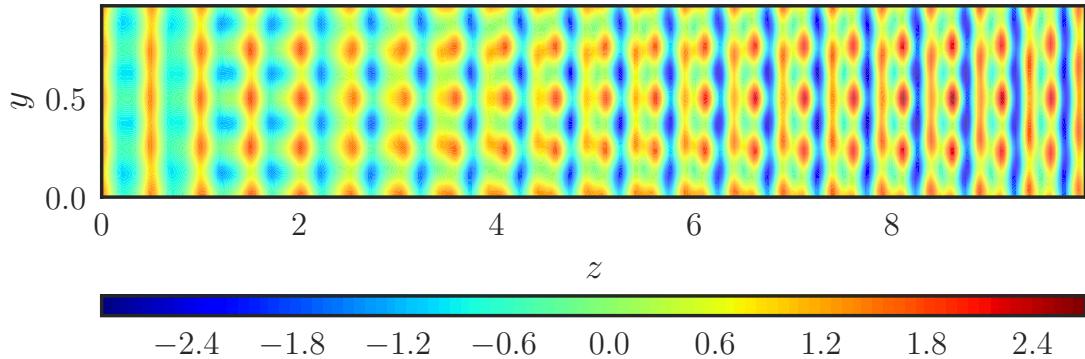


Figure 4: Scattering on the periodic cube structure in $D = [0, 2] \times [0, 1] \times [0, 10]$ for $k = 2\pi$. Effective rank of the solution is equal to 67. Section $x = 1$.

Acknowledgements

This study was supported by the Ministry of Education and Science of the Russian Federation (grant 14.756.31.0001), by RFBR grants 16-31-60095-mol-a-dk, 16-31-00372-mol-a and by Skoltech NGP program.

REFERENCES

1. F. Vico, L. Greengard, and M. Ferrando, “Fast convolution with free-space Green’s functions,” *Journal of Computational Physics*, vol. 323, pp. 191–203, 2016.
2. V. Kazeev, B. Khoromskij, and E. Tyrtyshnikov, “Multilevel Toeplitz matrices generated by tensor-structured vectors and convolution with logarithmic complexity,” *SIAM J. Sci. Comput.*, vol. 35, no. 3, pp. A1511–A1536, 2013.
3. I. V. Oseledets, “Approximation of $2^d \times 2^d$ matrices using tensor decomposition,” *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 4, pp. 2130–2145, 2010.
4. L. a. Klinteberg, D. S. Shamshirgar, and A.-K. Tornberg, “Fast Ewald summation for free-space Stokes potentials,” *arXiv preprint arXiv:1607.04808*, 2016.
5. V. Kazeev and C. Schwab, “Quantized tensor-structured finite elements for second-order elliptic pdes in two dimensions,” tech. rep., SAM research report 2015-24, ETH Zürich, 2015.
6. V. Kazeev, I. Oseledets, M. Rakhuba, and C. Schwab, “QTT-finite-element approximation for multiscale problems I: model problems in one dimension,” *Adv. Comp. Math.*, 2016.
7. I. V. Oseledets, “Tensor-train decomposition,” *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011.
8. S. V. Dolgov, B. N. Khoromskij, and D. V. Savostyanov, “Superfast Fourier transform using QTT approximation,” *J. Fourier Anal. Appl.*, vol. 18, no. 5, pp. 915–953, 2012.
9. I. V. Oseledets and E. E. Tyrtyshnikov, “TT-cross approximation for multidimensional arrays,” *Linear Algebra Appl.*, vol. 432, no. 1, pp. 70–88, 2010.
10. S. V. Dolgov and D. V. Savostyanov, “Alternating minimal energy methods for linear systems in higher dimensions. Part I: SPD systems,” *arXiv preprint 1301.6068*, 2013.
11. S. V. Dolgov and D. V. Savostyanov, “Alternating minimal energy methods for linear systems in higher dimensions. Part II: Faster algorithm and application to nonsymmetric systems,” *arXiv preprint 1304.1222*, 2013.
12. S. Holtz, T. Rohwedder, and R. Schneider, “The alternating linear scheme for tensor optimization in the tensor train format,” *SIAM J. Sci. Comput.*, vol. 34, no. 2, pp. A683–A713, 2012.