Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Tensors-structured Numerical Methods in Scientific Computing: Survey on Recent Advances

(revised version: July 2011)

by

Boris N. Khoromskij

Preprint no.: 21 2010

Tensors-structured Numerical Methods in Scientific Computing: Survey on Recent Advances

Boris N. Khoromskij
Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, D-04103 Leipzig, Germany; {bokh}@mis.mpg.de

Abstract

In the present paper, we give a survey of the recent results and outline future prospects of the tensor-structured numerical methods in applications to multidimensional problems in scientific computing. The guiding principle of the tensor methods is an approximation of multivariate functions and operators relying on certain separation of variables. Along with the traditional canonical and Tucker models, we focus on the recent quantics-TT tensor approximation method that allows to represent N-d tensors with log-volume complexity, $O(d \log N)$. We outline how these methods can be applied in the framework of tensor truncated iteration for the solution of the high-dimensional elliptic/parabolic equations and parametric PDEs. Numerical examples demonstrate that the tensor-structured methods have proved their value in application to various computational problems arising in quantum chemistry and in the multi-dimensional/parametric FEM/BEM modeling—the tool apparently works and gives the promise for future use in challenging high-dimensional applications.

AMS Subject Classification: 65F30, 65F50, 65N35, 65F10
Key words: High-dimensional problems, rank structured tensor approximation, quantics folding of vectors, matrix valued functions, FEM/BEM, computational quantum chemistry, stochastic PDEs.

1 Introduction

In the recent years, the use of tensor-structured data formats was recognized as the basic concept for breaking the "curse of dimensionality" in multidimensional numerical simulations. The guiding principle of the tensor methods is an approximation of multivariate functions and operators relying on a certain separation of variables and keeping the computational process in a low parametric tensor-structured manifold. Modern applications of tensor methods include the challenging high-dimensional problems arising in material sciences, bio-science, stochastic modeling, signal processing, machine learning and data mining, financial mathematics, etc. A survey and a comprehensive bibliography on the traditional tools of multilinear approximation in computer science based on the Tucker and canonical models are presented in [90, 16, 65]. An overview on the modern tensor-structured numerical methods in high-dimensional applications can be found in [56].

In the present survey paper we discuss the prospects of tensor methods in scientific computing, intended for effective solution of certain classes of boundary value, eigenvalue and transient equations posed in the high dimensional physical space.

1.1 Methods of separation of variables

The modern tensor-structured numerical methods are designed for the data-sparse representation of the multivariate functions, related operators and for the construction of tensor-truncated iterative

solvers for the high-dimensional physical equations by their "projection" onto the low separation rank tensor manifold.

In the current discussion, a tensor of order d, or briefly N-d tensor, is thought as a function on a product index set, $\mathbf{A} : I^{\otimes d} \to \mathbb{R}$ with the d-fold product $I^{\otimes d} = I \times \cdots \times I$, and $I = \{1, ..., N\}$. The subsequent exponential scaling in the storage size, N^d, predisposes severe computational difficulties when using the traditional numerical algorithms having the defect of the "curse of dimensionality". Hence, main efforts are focused on the construction of efficient low-complexity representations of the higher order tensors and the related multilinear transforms.

Recent approaches, like the wavelet multiscale methods [12] and the hyperbolic cross (sparse grids) approximation [8, 88, 95, 101, 28, 29, 30] allow to relax the curse of dimensionality, and already enable to handle the moderate dimensional problems, e.g., with $d \leq 10$. The sparse Schur complement methods circumvent the curse of dimension by reduction to the interface or to the wire basket of the boundary [51, 50].

The methods which lead to linear scaling in the dimension are distinctly linked with the idea of separation of variables. The arising discretizations via "formatted" function related tensors typically inherit the separability properties of the initial solutions on the continuous level, usually providing fast exponential convergence in the separation rank. This favorable feature combined with the modern multilinear algebra methods of nonlinear tensor approximation [21, 1, 42, 79, 57, 55] lead to the new concept of numerical schemes in higher dimensions which scale linearly in the dimension parameter d.

In the following, we describe the main ideas behind the tensor numerical methods, outline their theoretical background, and present a number of numerical illustrations on the computational efficiency for typical classes of multidimensional problems.

1.2 Advances of tensor methods in the recent decade

The particularly efficient tensor separation schemes are based on the global-additive type, for example, canonical and Tucker models, and the local-multiplicative type, e.g., matrix product states (MPS) tensor formats. Here we mean that the additive type tensor formats rely on representation by a (global) sum over few separable (rank-1) elements, while the MPS formats count on the contracted product representation correlating each single variable (state) only with a few local neighbors (slightly entangled systems).

The orthogonal rank-$(r_1, ..., r_d)$ Tucker tensor decomposition (cf. [96, 14]) allows to relax the curse of dimension dramatically because of the reduced complexity, $r^d + drN$, where, in practice, the sufficient maximal Tucker rank $r = \max\limits_{1 \leq \ell \leq d} r_\ell$, can be much smaller than N, say, $r = O(\log N)$.

A special case of the (nonorthogonal) rank-$(R, ..., R)$ Tucker representation, usually refereed as the canonical model, implies the linear scaling in d, dRN. In general, the best R-term nonlinear approximation shows rather slow polynomial convergence in the rank parameter R, [92], and it can be calculated by the simple (but non-robust) greedy-type incremental algorithms. For the class of (physically relevant) analytic multivariate functions and Green's kernels the exponential convergence in the separation rank R can be proved [25, 93, 49, 53, 54], that leads to the asymptotically optimal bound on the canonical rank, $R = O(\log N)$, implying the approximation rate $O(N^{-\alpha})$, with $\alpha \geq 1$.

Since the well-known limitations in approximation via the canonical model [17], the mixed (primal-dual) orthogonal Tucker-canonical representation imposing the rank-R canonical core tensor was introduced [49, 57]. It inherits the beneficial features of both models, orthogonality and linear scaling in d. The multigrid accelerated version of the mixed Tucker-canonical format was proved to be efficient in tensor computation of the 3D integral transforms in *ab initio* electronic

structure calculations [57, 58]. Another interesting generalization of the Tucker (and canonical) model, the so-called block Tucker format, was recently introduced [13].

The rank-structured tensor representations that scale linearly in the dimension d and, at the same time, allow the direct SVD-based truncated multilinear operations can be constructed using the idea of selected decoupling of dimensions. Methods based on the so-called matrix product states factorization were introduced in density matrix renormalization group (DMRG) theory in [100] (see also [97] for the algebra of MPS formats). In computational molecular dynamics such methods are known for longer time as hierarchical or binary cascadic multi-configuration methods (MCTDH) used in combination with the Tucker tensor format (see [2, 99, 68] for further details). In the quantum information theory the concept of MPS was also introduced as the state decomposition of slightly entangled systems [98]. The MPS is defined by the explicit one-level factorized representation of an N-d tensor (beneficial feature for the numerical multilinear algebra) complemented by the low complexity SVD/QR rank truncation procedure.

The idea of rank-r dimension splitting in the context of the canonical tensor approximation was addressed in [49, Lemma 2.2] leading to the almost linear scaling in d, $O(dNr^{\log d})$. Recently, the tensor formats based on the MPS-type dimension splitting using the tree-type or hierarchical Tucker (HT) models have attracted much attention (see [38], [79]). The non-hierarchical MPS-type factorized separation of variables specified by the so-called Tensor Train (TT) format was investigated in [79, 72]. The storage is estimated as $O(dr^2N)$, where r is the splitting rank. The DMRG approach to fast linear algebra in the TT-format was addressed in [75]. As the important step, in [80] the TT factorization is extended to the case of incomplete input data by using the TT-Cross approximation via interpolation on the entries of a TT-cross.

1.3 Quantized N-d tensors lead to $d \log N$ complexity: "blessing of dimension"

In high resolution molecular/electronic structure calculations, in FEM applications and in stochastic PDEs the univariate grid size N specifying N-d tensors may be rather large, say, $N \approx 10^4$, thus addressing a question on whether a better than linear asymptotic in N can be attained. For example, it is known for longer time in signal processing, spectroscopy and higher-order statistics that the folding of a sampling vector into a matrix or 3-tensor may reduce the number of essential parameters (samplings) to reconstruct exactly the initial vector in $\mathbb{R}^N$ (cf. [44, 89, 43, 81]). This is due to possible low rank decomposition of arising matrix or tensor data arrays. The relation to a separable representation of polyadics was emphasized in [40, 11].

Clearly, the dyadic folding of a vector can be successively continued up-to tensor-order 3, 4, 5, etc. until the irreducible mode-size of a tensor, $2 \times 2 \times ... \times 2$, is achieved [73, 55]. It was found in [73] by numerical tests that in some cases the dyadic reshaping of $2^L \times 2^L$ matrix leads to a small TT-rank of the resultant quantized matrix of size $(2 \times 2)^{\otimes L}$.

In [55] the diadically quantized representation of vectors (called there *quantics*[1]) was generalized to q-adic folding applied to the general N-d tensors. There the principal question why the folding of a vector to a higher dimensional tensor might lead to an essential data compression was first addressed and rigorously analyses from the approximation aspects. The basic approximation result was proven for a wide class of function generated tensors: the TT-rank of quantized exponential, trigonometric, polynomial etc. N-vectors are shown to be uniformly bounded in N. Moreover, it was shown that the QTT approximation of q-$\log_q N$ tensors obtained by the q-adic multilevel folding, and applied to a class of (analytic) function related N-vectors, $N \times N$-matrices or N-d tensors provides exponential convergence in the TT-rank (see [56] for more detail).

[1]The terminology is borrowed from the methods of separable representation of polyadics, see [11].

This allows to understand why the multifolding of N-d tensors may lead to the $d \log N$-computational complexity (in log-volume size). In particular, the rank structured quantics approximation discovers the similarity patterns in the input data up to the finest resolution level (in contrast to the principles of wavelets multi-resolution schemes), providing the way to the nearly optimal compression rate within the chosen dimensional splitting model posed in the virtual dimension $D = d \log_q N$.

Along with the good approximation properties, the QTT format seems to be crucial for the numerical efficiency of the DMRG type iteration (see [63]) since it reduces the mode size of a tensor from N to 2 (the latter enters at least quadratically in the complexity of DMRG iteration).

In the subsequent paper [27] the uniform rank bound of the quantized polynomial N-vector (renamed as "tensorization of vectors") was proven in the case of hierarchical Tucker format. The more general QTT-rank estimate in terms of the separation rank of a generating function $f(x+y)$ was obtained in [74]. Explicit low-rank representation of the discrete quantized Laplacian and its inverse was derived in [45].

The QTT representation of the Fast Fourier Transform and the Toeplitz matrices discovered in [20] and [46], respectively, leads to logarithmic complexity of the corresponding matrix-vector products. In particular, according to the above results, the convolution of N-d tensors in QTT format can be computed in $O(d \log N)$ operations.

We summarize that the idea of quantized representation in $d \log N$-complexity allows the function/operator calculus on large grids (practically unlimited grid-size) opening the way to accurate quasi-analytic calculations of integrals, and other linear/bilinear mappings in fully discrete form.

1.4 SVD-based numerical tensor approximation

The key point of tensor numerical method based on separation of variables is the efficient computability of formatted tensor decompositions (approximation) of higher order tensors.

The standard linear and multilinear operations on tensors can be realized in rank-structured formats. Due to the increase of separation rank within tensor operations the efficient rank optimization (tensor truncation) plays an important role in the tensor/operator numerical calculus.

The SVD-based approximation in the Tucker format is based on the HOSVD decomposition of the full format tensors [15]. The so-called reduced HOSVD (RHOSVD) approximation applies to the canonical target tensor [57]. The latter scales at most quadratically in the basic model parameters, providing a good initial guess for the (local) ALS-type minimization in the mixed Tucker-canonical format. The efficient multigrid accelerated canonical-to-Tucker rank reduction scheme, which scales linearly in the input parameters, became a powerful tool in 3D electronic structure calculations [48].

A direct method of the Tucker decomposition for $d = 3$, based on the adaptive cross approximation is presented in [76, 3].

Furthermore, the SVD/QR-based algorithms for the rank optimization in the TT/HT dimension splitting schemes have been described in [98, 97, 79, 26].

Quantized high-dimensional tensors can be represented (approximated) in the TT format (the so-called QTT format [55]). In this way the SVD-based rank-truncated multilinear operations on $2 \times ... \times 2$ TT tensors can be applied in the QTT tensor arithmetics.

1.5 On tensor-structured solution of multidimensional equations

Tensor numerical methods are proved to be efficient for data-sparse representation of functions and operators in the Hartree-Fock and Kohn-Sham models in electronic structure calculations [7, 39, 59, 57, 23, 76, 47]. Fully tensorized numerical approach for solving the Hartree-Fock equation,

discretized over $N \times N \times N$ grid, by tensor truncated iteration of complexity $O(N \log N)$, was recently presented in [58]. Other successful applications to high-dimensional eigenvalue problems [37, 54] and to stochastic PDEs [64, 62] are reported. A class of low tensor rank preconditioners for the multidimensional elliptic problems with jumping coefficients in $\mathbb{R}^d$ is proposed in [18]. The employment of separation of variables in machine learning is addressed in [6].

Numerical illustrations presented in Section 4 indicate surprisingly good compression rate, $O(d \log N)$, of the QTT approximation applied to certain function related tensors and matrix-valued functions. Particularly, this includes solutions of model boundary-value and eigenvalue problems in $\mathbb{R}^d$, examples of electron density, Hartree potential and the potential energy surface in electronic structure calculations, as well as some matrix-valued functions of d-Laplacian.

The quantized-TT approximation of log-log complexity scaling, $O(d \log N \log \varepsilon^{-1})$, where $\varepsilon > 0$ is the approximation error, opens the new perspectives for reliable and robust computational methods in higher dimensions that are free from the "curse of dimensionality", and noticeable limitations on the grid-size (blessing of dimension). This approach can be interpreted as a kind of "paradoxically" mesh-less method that is essentially based on the grid representation via reshaping to the virtual higher dimension $D = d \log N$, complemented by the SVD-based rank reduction algorithms.

With impetus to real life applications, we discuss how the QTT approximation method can be employed in the framework of truncated iteration for solving certain classes of elliptic/parabolic equations in higher dimensions with log-scaling in the volume size.

The rest of the paper is organized as follows. Section 2 describes traditional tensor models based on the canonical and Tucked decompositions and then addresses the MPS-type TT and quantized TT (QTT) tensor formats. We discuss the approximation error estimate in terms of the TT/QTT splitting ranks applied to the class of function related tensors. Section 3 analyses the canonical and QTT approximations applied to the solution of certain elliptic boundary-value/eigenvalue problems, as well as to the transient parabolic-type equations. Section 4 presents various numerical examples on the QTT approximation of N-d tensors applied to the elliptic/parabolic equations and to stochastic/multiparametric PDEs. Conclusions summarize the prospects of tensor-structured numerical methods in high-dimensional scientific computing.

2 Introduction to rank-structured tensor formats

This section discusses the most frequently used rank-structured tensor formats. In §2.1 we introduce the basic Tucker and canonical models. In §2.2, the generalized *tensor chain* (TC) representation is addressed characterized by the periodic-type dimensional splitting scheme. §2.3 discusses the two-level Tucker-TT model that is a combination of the orthogonal Tucker decomposition with the TT-representation of the Tucker core. It benefits from almost linear complexity scaling in the rank parameter r, $O(drN + dr^2 \log N)$, due to the fact that usually in the case of function related tensors both the Tucker and TT ranks appear to be much smaller than N, $r \ll N$. The special case of above model is a two-level canonical-TT format that is proved to be useful in the tensor-structured operator calculus.

In our presentation of the Tucker/canonical formats we mainly follow standard notations in [14, 65].

2.1 Basic rank-structured tensor formats

The separable representation of multivariate functions is based on the constructions of a tensor-product Hilbert space [83]. We mainly consider the case of high-order tensors, i.e. the real/complex valued functions on the product index set.

Tensors of order d are defined as the elements of finite dimensional tensor-product Hilbert space $\mathbb{W}_{\mathbf{n}} \equiv \mathbb{W}_{\mathbf{n},d}$ of the d-fold, $\underbrace{N_1 \times ... \times N_d}_{d}$ real/complex-valued arrays, which can be represented componentwise,

$$\mathbf{A} = [A(i_1, ..., i_d)] \quad \text{with} \quad i_\ell \in I_\ell := \{1, ..., N_\ell\}, \quad \text{and} \quad \mathbf{n} = (N_1, ..., N_d).$$

For the ease of presentation, we mainly consider the equal-size tensors, i.e., $I_\ell = I = \{1, ..., N\}$ ($\ell = 1, ..., d$). We call the elements of $\mathbb{W}_{\mathbf{n}} = \mathbb{R}^{\mathcal{I}}$ with $\mathcal{I} = I_1 \times ... \times I_d$, as N-d tensors and use several equivalent notations for the corresponding d-dimensional arrays, $\mathbf{A} \equiv \mathbf{A}_{(N,d)} \equiv \mathbf{A}_{(\mathbf{n},d)} \in \mathbb{W}_{\mathbf{n}}$. The Euclidean scalar product, $\langle \cdot, \cdot \rangle : \mathbb{W}_{\mathbf{n}} \times \mathbb{W}_{\mathbf{n}} \to \mathbb{R}$, is given by

$$\langle \mathbf{A}, \mathbf{B} \rangle := \sum_{\mathbf{i} \in \mathcal{I}} A(\mathbf{i}) B(\mathbf{i}), \quad \mathbf{A}, \mathbf{B} \in \mathbb{W}_{\mathbf{n}}.$$

The storage demand for N-d tensors scales exponentially in d, $\dim \mathbb{W}_{\mathbf{n},d} = N^d$.

The rank-$(r_1, \ldots, r_d)$ Tucker format contains all tensors in $\mathbb{W}_{\mathbf{n}} = \mathbb{R}^{\mathcal{I}}$, that can be presented in the form of a tensor-by-matrix *contracted product* over the product index set $J := \times_{\ell=1}^{d} J_\ell$, with $J_\ell = \{1, ..., r_\ell\}$, and $\mathbf{r} = (r_1, ..., r_d) \in \mathbb{N}^d$,

$$\mathbf{V} = \boldsymbol{\beta} \times_1 T^{(1)} \times_2 T^{(2)} ... \times_d T^{(d)} \in \mathbb{W}_{\mathbf{n}}, \tag{2.1}$$

and with certain (orthogonal) $N \times r_\ell$ side matrices, $T^{(\ell)} = [t_\ell^1 ... t_\ell^{r_\ell}] \in \mathbb{R}^{I \times J_\ell}$. The coefficients tensor $\boldsymbol{\beta} = [\beta(\nu_1, ..., \nu_d)]$, $\nu_\ell \in J_\ell$, called the *core tensor*, is an element of a (dual) tensor space $\mathbb{B}_{\mathbf{r}} = \mathbb{R}^{J_1 \times ... \times J_d}$. We denote this tensor class by $\boldsymbol{\mathcal{T}}_{\mathbf{r},\mathbf{n}} \subset \mathbb{W}_{\mathbf{n}}$. The storage size is still exponential, $r^d + drN$, since the interdimensional connectivity tensor $\boldsymbol{\beta}$ is fully populated.

The R-term canonical format is defined as the particular case of Tucker model (2.1) specified by rank parameters $r_\ell = R$ ($\ell = 1, ..., d$), and by the diagonal Tucker core $\boldsymbol{\beta} := diag\{\beta_1, ..., \beta_R\}$, such that $\beta(\nu_1, ..., \nu_d) = 0$ except when $\nu_1 = ... = \nu_d$ with $\beta(\nu, ..., \nu) = \beta_\nu$. The orthogonality of $T^{(\ell)}$ is no longer required. We denote by $\boldsymbol{\mathcal{C}}_{R,\mathbf{n}}$ the class of tensors in $\mathbb{W}_{\mathbf{n}}$ whose rank does not exceed R, $rank(\mathbf{V}) \le R$, with the particular parametrization

$$\mathbf{V} = \sum_{k=1}^{R} b_k V_k^{(1)} \times_2 ... \times_d V_k^{(d)} \equiv \sum_{k=1}^{R} b_k \bigotimes_{\ell=1}^{d} V_k^{(\ell)}, \quad b_k \in \mathbb{R}, \tag{2.2}$$

including the normalized canonical vectors $V_k^{(\ell)} \in \mathbb{R}^{I_\ell}$. In spite of linear scaling in d, dRN, the inflexible and rather poor connectivity pattern of this format, parallel with the lack of orthogonality and nonclosedeness of the nonlinear set $\boldsymbol{\mathcal{C}}_{R,\mathbf{n}}$, lead to the well understood computational difficulties with the canonical decomposition. Another possible drawback is the "rigid" constraints due to the equal rank distribution for all dimensions, $r_\ell = R$, $\ell = 1, ..., d$, (though the latter seems to be the intrinsic property in the case of symmetric/antisymmetric tensors). Notice that the orthogonal canonical decomposition is well understood theoretically and allows robust greedy-type computational schemes [103, 65], however, the approximating quality of such representations is rather poor making this tensor format practically inapplicable.

On the other hand, the canonical format can be gainfully combined with the Tucker and some other stable tensor representations (see §2.3). The example of combined tensor models is given by

the rank-R mixed (primal-dual) Tucker-canonical format denoted by $\boldsymbol{\mathcal{T}}[\mathcal{C}_{R,\mathbf{r}}] \subset \boldsymbol{\mathcal{C}}_{R,\mathbf{n}}$. It contains all Tucker tensors in $\boldsymbol{\mathcal{T}}_{\mathbf{r},\mathbf{n}}$ with the Tucker core in $\boldsymbol{\mathcal{C}}_{R,\mathbf{r}}$ (see [49, 57]). This benefits from the linear storage complexity, $drN + dRr$, orthogonality, and the opportunity for *variable directional ranks*, $\mathbf{r} = (r_1, ..., r_d)$. Another example is given by the block Tucker model [13].

In some applications (see e.g. Example 4.2 in §4.1) the so-called weakly (locally) coupled canonical representation is useful. The weakly coupled canonical tensor $\mathbf{V} \in \mathbb{R}^{J \times \mathcal{I}}$, $J = \{1, ..., N_0\}$, is defined componentwise in the index $j \in J$,

$$\mathbf{V}(j) = \sum_{k=1}^{R} b_k V_k^{(1)}(j) \times_2 ... \times_d V_k^{(d)}(j) \equiv \sum_{k=1}^{R} b_k \bigotimes_{\ell=1}^{d} V_k^{(\ell)}(j), \quad b_k \in \mathbb{R}, \tag{2.3}$$

including the normalized canonical matrices $V_k^{(\ell)} \in \mathbb{R}^{I_\ell \times J}$, with the column vectors $V_k^{(\ell)}(j) \in \mathbb{R}^{I_\ell}$, $j \in J$.

2.2 Reduced Tucker sparsity: Tensor chain/train formats

The rank-$\mathbf{r}$ *tensor train (TT)* format belong to the class of MPS representations. It is defined in the spirit of Tucker model, but with essentially reduced "connectivity" constraints (see [79, 72]) so that its storage size scales linearly in both d and N. The generalization of the TT-format to the case of "periodic" index chain is given by the following definition [55].

Definition 2.1 *(Tensor chain/train format). Given the rank parameter $\mathbf{r} = (r_0, ..., r_d)$, and the respective index sets $J_\ell = \{1, ..., r_\ell\}$ ($\ell = 0, 1, ..., d$), with the periodicity constraints $J_0 = J_d$. The rank-$\mathbf{r}$ tensor chain (TC) format contains all elements $\mathbf{V}$ in $\mathbb{W}_{\mathbf{n}} = \mathbb{R}^{\mathcal{I}}$ that can be represented as the chain of contracted products of 3-tensors over the d-fold product index set $J := \times_{\ell=1}^{d} J_\ell$,*

$$\mathbf{V} = \{\times_\ell\}_{\ell=1}^{d} \mathbf{G}^{(\ell)} \quad \textit{with 3rd order core tensors} \quad \mathbf{G}^{(\ell)} \in \mathbb{R}^{J_{\ell-1} \times I_\ell \times J_\ell}, \tag{2.4}$$

or in the index notation,

$$V(i_1, ..., i_d) = \sum_{\alpha_1 \in J_1} \cdots \sum_{\alpha_d \in J_d} G^{(1)}(\alpha_d, i_1, \alpha_1) G^{(2)}(\alpha_1, i_2, \alpha_2) \cdots G^{(d)}(\alpha_{d-1}, i_d, \alpha_d) \equiv G_{i_1}^{(1)} \cdots G_{i_d}^{(d)},$$

where $G_{i_\ell}^{(\ell)}$ is an $r_{\ell-1} \times r_\ell$ matrix. Denote this set of tensors by $TC[\mathbf{r}, d] \equiv TC[\mathbf{r}, \mathbf{n}, d] \subset \mathbb{W}_{\mathbf{n}}$.

In the case $J_0 = J_d = \{1\}$ (open boundary conditions), this construction coincides with the definition of TT format in [72], thus implying $TT[\mathbf{r}, d] \subset TC[\mathbf{r}, d]$.

Storage requirement for the rank-$\mathbf{r}$ TC tensors in (2.4) is bounded by

$$\sum_{\ell=1}^{d} r_{\ell-1} r_\ell \, N \le d r^2 N \quad \text{with} \quad r = \max_\ell r_\ell.$$

Notice that the standard linear and bilinear operations on TT/TC tensors can be performed with the linear complexity scaling in d and n. In particular, for the Hadamard product we have

$$\mathbf{Z} = \mathbf{X} \circ \mathbf{Y}: \quad Z^{(k)}(i_k) = X^{(k)}(i_k) \otimes Y^{(k)}(i_k),$$

implying the formatted representation of the scalar product (in $O(dr^3 N)$ operations)

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \langle \mathbf{X} \circ \mathbf{Y}, \mathbf{1} \rangle.$$

The contracted product with rank-1 tensor is given by

$$\mathbf{Z} = \mathbf{X} \times_1 U^{(1)} \dots \times_d U^{(d)} : \quad Z^{(k)} = \sum_{i_k} X^{(k)}(i_k) U^{(k)}(i_k).$$

For a sum of TT-tensors we have,

$$\mathbf{Z} = \mathbf{X} + \mathbf{Y} : \quad Z^{(k)}(i_k) = \begin{bmatrix} X^{(k)}(i_k) & 0 \\ 0 & Y^{(k)}(i_k) \end{bmatrix}.$$

The above operations can be complemented by the SVD-based rank optimization procedure.

The beneficial properties of the TC/TT formats are collected in the following lemma.

Lemma 2.2 *([55]) The maximal TC rank of $\mathbf{V} \in TC[\mathbf{r}, d]$, does not exceed its canonical rank, $r \le rank(\mathbf{V})$. The rank-R canonical tensor belongs to $TC[\mathbf{r}, d]$ with $\mathbf{r} = (R, ..., R)$, and with diagonal cores.*
$TT[\mathbf{r}]$ is the closed manifold [87], while $TC[\mathbf{r}]$ is closed on the subset $TT[\mathbf{r}]$ [97] (see Remark 2.3).

Recall that the k-th TT rank of a tensor $\mathbf{X} = [X(i_1 ... i_d)]$ is the rank of its matrix unfolding $\mathbf{X}^{(k)}$ $(1 \ge k \ge d-1)$ with the elements

$$\mathbf{X}^{(k)}(i_1 ... i_k; i_{k+1} ... i_d) = X(i_1, ..., i_d).$$

Applicability of the general TC format with $J_0 = J_d \neq \{1\}$, can be motivated, in particular, by the following computational tasks:

- DMRG in FCI electronic structure calculations and quantum information theory [100, 98, 97].
- Rank optimization in the case of highly nonuniform distribution of the TT-rank parameters r_ℓ, $\ell = 1, ..., d$.
- Representation of a periodic MPS.

Remark 2.3 *(On noncloseness of $TC[\mathbf{r},d]$). For given $\mathbf{V} \in TC[\mathbf{r}, d]$, let the core tensor $\mathbf{G}^{(d)}$ be defined by diagonal $r_d \times r_d$-matrices. Then the TC-tensor $\mathbf{V}$ can be represented by a sum of r_d TT tensors. Hence, similar to the case of canonical tensors, it is easy to construct a sequence $\mathbf{V}_k \to 0$ such that the approximant $\mathbf{V}_k$ does not converge as $k \to \infty$.*

Notice that the suboptimal approximation of the full format, canonical or TT-tensors by using the low TT-rank elements can be fulfilled by the algorithm based on SVD/QR decompositions [72]. The best TT-approximation can be computed by ALS type iteration [84] (cf. the Tucker and canonical approximations). In the case of general TC tensors the rank reduction operations require certain modifications based on the use of ALS type iteration applied in the cyclic ordering similar to the case ALS-TT approximation.

The next elegant result explains how to select real and imaginary part of the complex-valued vector in the TT/TC formats [20].

Theorem 2.4 *The complex-valued tensor train $a(j_1, \dots, j_d) = A^{(1)}_{j_1} \dots A^{(d)}_{j_d}$, with ranks $r_0, r_1, ..., r_{d-1}, r_d$ can be represented as tensor train $a(j_1, \dots, j_d) = \hat{A}^{(1)}_{j_1} \dots \hat{A}^{(d)}_{j_d}$, with ranks $r_0, 2r_1, ..., 2r_{d-1}, r_d$, where the cores $\hat{A}^{(p)}_{j_p}$, $p = 1, \dots, d-1$ are real-valued,*

$$\begin{aligned} \hat{A}^{(1)}_{j_1} = \begin{bmatrix} \Re A^{(1)}_{j_1} & \Im A^{(1)}_{j_1} \end{bmatrix}, \quad \hat{A}^{(p)}_{j_p} = \begin{bmatrix} \Re A^{(p)}_{j_p} & \Im A^{(p)}_{j_p} \\ -\Im A^{(p)}_{j_p} & \Re A^{(p)}_{j_p} \end{bmatrix}, \qquad p = 2, \dots, d-1, \\ \hat{A}^{(d)}_{j_d} = \begin{bmatrix} \Re A^{(d)}_{j_d} \\ -\Im A^{(d)}_{j_d} \end{bmatrix} + \imath \begin{bmatrix} \Im A^{(d)}_{j_d} \\ \Re A^{(d)}_{j_d} \end{bmatrix} \end{aligned} \tag{2.5}$$

Example 2.1. Applying Theorem 2.4 to a tensor generated by a function $e^{i(x_1+...+x_d)}$, we derive that TT-rank of the N-d tensor obtained by sampling $f(x_1,...,x_d) = sin(x_1 + ... + x_d)$ or $f(x_1,...,x_d) = cos(x_1 + ... + x_d)$ over $N^{\otimes d}$ tensor grid in $\mathbb{R}^d$, is exactly 2 (see also [79]), while its canonical rank is estimated by $R = d$ [5].

2.3 Quantics representation of N-d tensors

In this section, we address the important statement saying that the class of discrete exponential (resp. trigonometric) N-vectors allows the rank-1 (resp. rank-2) q-folding representation with small $q = 2, 3, ...$, reducing the storage complexity $O(N)$ to the logarithmic bound $O(q \log_q N)$.

Given $q = 2, 3, ...$, we suppose that $N = q^L$ with some $L = 1, 2, ...$ The folding (lifting) and unfolding (reducing) transforms on N-d tensors can be interpreted as the dual reshaping operations specified by the reordering scheme of the respective index sets. Next definition introduces the folding of N-d tensors into the elements of auxiliary D-dimensional tensor space with $D = d \log_q N$.

Definition 2.5 *([55]) Introduce the q-adic folding transform of degree $2 \le p \le L$,*

$$\mathcal{F}_{q,d,p} : \mathbb{W}_{\mathbf{n},d} \to \mathbb{W}_{\mathbf{m},dp}, \quad \mathbf{m} = (\mathbf{m}_1, ..., \mathbf{m}_\ell), \quad \mathbf{m}_\ell = (m_{\ell,1}, ..., m_{\ell,p}),$$

with $m_{\ell,1} = q^{L-p+1}$, and $m_{\ell,\nu} = q$ for $\nu = 2, ..., p$, ($\ell = 1, ..., d$), that reshapes the initial $\mathbf{n}$-d tensor in $\mathbb{W}_{\mathbf{n},d}$ to the quantics space $\mathbb{W}_{\mathbf{m},dp}$ as follows:
(A) For $d = 1$ a vector $\mathbf{X}_{(N,1)} = [X(i)]_{i\in I} \in \mathbb{W}_{N,1}$, is reshaped to the element of $\mathbb{W}_{q^{L-p+1},p}$ by

$$\mathcal{F}_{q,1,p} : \mathbf{X}_{(N,1)} \to \mathbf{Y}_{(\mathbf{m},p)} = [Y(\mathbf{j})] := [X(i)], \quad \mathbf{j} = \{j_1, ..., j_p\},$$

with $j_1 \in \{1, ..., q^{L-p+1}\}$, and $j_\nu \in \{1, ..., q\}$ for $\nu = 2, ..., p$. For fixed i, $j_\nu = j_\nu(i)$ is defined by $j_\nu = 1 + C_{L-p-1+\nu}$, ($\nu = 1, ..., p$), where the $C_{L-p-1+\nu}$ are found from the partial radix-q representation of $i - 1$,

$$i - 1 = C_{L-p} + C_{L-p+1} q^{L-p+1} + \cdots + C_{L-1} q^{L-1}.$$

(B) For $d > 1$ a tensor $\mathbf{A}_{(\mathbf{n},d)} = [A(i_1, ..., i_d)]$, $i_\ell \in I_\ell$, $\ell = 1, ..., d$, is reshaped by

$$\mathcal{F}_{q,d,p} : \mathbf{A}_{(\mathbf{n},d)} \to \mathbf{B}_{(\mathbf{m},dp)} = [B(\mathbf{j}_1, ..., \mathbf{j}_d)] := [A(i_1, ..., i_d)], \quad \mathbf{j}_\ell = \{j_{\ell,1}, ..., j_{\ell,p}\},$$

with $j_{\ell,1} \in \{1, ..., q^{L-p+1}\}$, and $j_{\ell,\nu} \in \{1, ..., q\}$, for $\nu = 2, ..., p$, and for all $\ell = 1, ..., d$. Now the univariate ℓ-mode index i_ℓ is represented by $\mathbf{j}_\ell$ as in the case $d = 1$.

For the maximal degree folding corresponding to $p = L$, the multiindex $\mathbf{j}_\ell - \mathbf{1}$ is the q-adic representation of $i_\ell - 1$ for $i_\ell \in I_\ell$, in radix-q system, such that $j_{\ell,\nu}$ takes values in $\{1, ..., q\}$.

For the sake of higher compressibility, the *maximal degree folding*, $\mathcal{F}_{q,d,L}$, can be applied.

Example 2.2. For $d = 1$ and $p = 2, 3$, the reshaping map $\mathcal{F}_{q,1,p}$ folds an N-vector to a $N/q \times q$-matrix or to $N/q^2 \times q \times q$, 3-tensor, respectively.

The unfolding transform, e.g., tensor-to-matrix (matricization) or tensor-to-vector (vectorization), may be viewed as the reverse to the folding, $\mathcal{F}^{-1}_{q,d,p}$, see [65] for the conventional definition.

The quantics (folding) transform $\mathcal{F}_{q,d,p}$ exhibits the following main properties:

(F1) $\mathcal{F}_{q,d,p}$ is the linear isometry between $\mathbb{W}_{N,d}$ and $\mathbb{W}_{q^{L-p+1},dp}$.

(F2) The q-folding of a rank-1 tensor $w = x_1 \times ... \times x_d \in \mathbb{W}_{N,d}$, is given by the outer product of componentwise vector reshaping transforms,

$$\mathcal{F}_{q,d,p} w = \mathcal{F}_{q,1,p} x_1 \otimes ... \otimes \mathcal{F}_{q,1,p} x_d.$$

(F3) Let $d = 1$, then for any $p = 2, ..., L$ and $\mathbf{X} = [X(i)] \in \mathbb{C}^N$ we have a bound on the TT rank of a tensor $\mathcal{F}_{q,1,L}\mathbf{X}$,

$$r_{p-1} \le rank(\mathbf{X}_p),$$

where $\mathbf{X}_p$ is the reshaping of $\mathbf{X}$ to a $N/q^{p-1} \times q^{p-1}$ matrix.

Next lemma presents the basic results on the rank-1 (resp. rank-2) q-folding representation of the exponential (resp. trigonometric) vectors [55].

Lemma 2.6 *For given $N = q^L$, with $q = 2, 3, ...$ and $L \in \mathbb{N}$, and for given $c_k, z_k \in \mathbb{C}$ $(k = 1, ..., R)$, the exponential sum N-vector, $\mathbf{X} := \{x_n := \sum_{k=1}^{R} c_k z_k^{n-1}\}_{n=1}^N$, can be reshaped by the q-folding $\mathcal{F}_{q,1,L}$, to the rank-R, q-L tensor in $\mathbb{W}_{\mathbf{q},L}$,*

$$\mathcal{F}_{q,1,L} : \mathbf{X} \to \mathbf{A}_{(q,L)} = \sum_{k=1}^{R} c_k \otimes_{p=1}^{L} [1\ z_k^{q^{p-1}}\ ...\ z_k^{(q-1)q^{p-1}}]^T \in \boldsymbol{\mathcal{C}}_{R,\mathbf{q}}[TT[\mathbf{1}]]. \tag{2.6}$$

The number of representation parameters is reduced from $(N+1)R$ to $(qL+1)R$.

The trigonometric sum N-vector, $\mathbf{X} := \{x_n := \sum_{k=1}^{R} c_k \sin(\alpha_k(n-1))\}_{n=1}^N$, $\alpha_k \in \mathbb{R}$, can be reshaped by the successive q-folding to the rank-R, q-L tensor $\mathbf{A}_{(q,L)}$, where for each of R summands, both the $\mathbb{C}$-rank, and the TT-rank are exactly 2,

$$\mathcal{F}_{q,1,L} : \mathbf{X} \to \mathbf{A}_{(q,L)} = \sum_{k=1}^{R} \mathbf{A}_k \in \mathbb{W}_{q,L}, \quad with \quad \mathbf{A}_k \in TT[\mathbf{2}, L].$$

The number of representation parameters does not exceed $4qLR$.

In the case $q = 2$, the single sin-vector has the explicit rank-2 QTT representation (see Thm. 2.4 and [74]) (with $y_p = 2^{p-L} i_p - 1$, $i_p \in \{0, 1\}$),

$$\mathbf{X} \mapsto [\sin\ y_1 \cos\ y_1] \otimes_{p=2}^{L-1} \begin{bmatrix} \cos\ y_p & -\sin\ y_p \\ \sin\ y_p & \cos\ y_p \end{bmatrix} \otimes \begin{bmatrix} \cos\ y_L \\ \sin\ y_L \end{bmatrix} \in \{0,1\}^{\otimes L},$$

It turns out that the exponential-trigonometric product vector allows the $4 \log_q N$ complexity quantics representation as proven by the following lemma.

Lemma 2.7 *For given $N = q^L$, with $q = 2, 3, ...$, $L \in \mathbb{N}$, and $c, z \in \mathbb{C}, \alpha \in \mathbb{R}$, the single exponential-trigonometric vector $\mathbf{X} := \{x_n := cz^{n-1} \sin(\alpha(n-1))\}_{n=1}^N$, can be reshaped to the q-L tensor, $\mathcal{F}_{q,1,L} : \mathbf{X} \to \mathbf{A}_{(q,L)} \in TT[2]$, whose both $\mathbb{C}$- and TT-ranks do not exceed 2. The number of representation parameters is bounded by $4q \log_q N$.*

Notice that Lemmas 2.6 - 2.7 outline the way to further improvement of the results in [89, 44] on the identifiability of multidimensional harmonic retrieval.

Remark 2.8 *Property (F3) of the quantics folding ensures that the QTT rank of a vector obtained by the equidistant sampling the polynomial of degree m, does not exceed $m+1$. In fact, the column space of the reshaped TT-unfolding matrix is spaned by at most $m+1$ polynomial vectors generated by $1, x, ..., x^m$, respectively.*

The similar result was proven in [27] for the case of quntized (renamed there as tensorization of vectors) hierarchical Tucker representation.

It is worth to note that using *equidistant* sampling points is non mandatory.

Lemma 2.9 *(A) For any $n = 0, 1, ...$, the Chebyshev polynomial $T_n(x) = \cos(n \arccos x)$, $|x| \le 1$, sampled over $N + 1 = 2^L$ Chebyshev nodes $x_j \in [-1, 1]$, can be represented in the quantics space of 2-log N tensors with both $\mathbb{C}$-rank and QTT-rank ≤ 2, uniformly in N.*
(B) The Chebyshev polynomial $T_n(x)$, sampled as a vector $\mathbf{X}$, at Chebyshev nodes, $\theta_j = \arccos x_j$, has the explicit rank-2 QTT representation (with $y_p = 2^{p-L} i_p - 1$, $i_p \in \{0, 1\}$),

$$\mathbf{X} \mapsto [\cos y_1 - \sin y_1] \otimes_{p=2}^{L-1} \begin{bmatrix} \cos y_p & -\sin y_p \\ \sin y_p & \cos y_p \end{bmatrix} \otimes \begin{bmatrix} \cos y_L \\ \sin y_L \end{bmatrix} \in \{0, 1\}^{\otimes L},$$

The general concept on the explicit QTT representation of vectors and matrices can be found in [74], [45], [20], [46], and [56].

Remark 2.10 *The TT-rank of the q-folded discrete Gaussian function $e^{-\alpha t^2}$ sampled over the uniform grid, $\mathbf{G} := \{e^{-\alpha h^2 (n-1)^2}\}_{n=1}^N$, $h > 0$, appears to be greater than 2. Numerical tests show that it remains to be almost uniformly bounded in the vector size N (see Table 3.1 below). Lemma 2.6 implies that the rank-1 quantics representation of $\mathbf{G}$ arises in the case of quadratic mesh grading toward the origin, i.e., by sampling over the points $t_n = \sqrt{h(n-1)}$, $(n = 1, ..., N)$.*

Representation of tensors in low separation rank formats is the key point in the design of fast tensor-structured numerical methods in higher dimension. In fact, it allows the implementation of basic linear and bilinear algebraic operations on tensors such as addition, scalar, Hadamard and convolution products with linear complexity in the univariate tensor size (see [49, 35, 52]). These tensor operations (excepting scalar product) normally increase the separation rank of the resultant tensor. Hence, the complexity control requires further "projection" of such intermediate results to the set of tensors with smaller rank parameter (rank truncation). This leads to nonlinear approximation problem to be addressed in the next section.

2.4 Combination of the Tucker, canonical and TC/TT formats

Following [55], let us denote by $\boldsymbol{\mathcal{T}}_{\mathbf{r}}[TC[\mathbf{r}_1]]$ the *Tucker-TC* format containing all Tucker tensors in $\boldsymbol{\mathcal{T}}_{\mathbf{r},\mathbf{n}}$ with the Tucker core in the rank-$\mathbf{r}_1$ TC format. Now the complexity of representation scales linearly in r, $O(drN + d\, r_1^2\, r)$, while the representation basis is given explicitly by the "optimal" set of orthogonal Tucker vectors. This may be gainfully applied in the framework of the Galerkin scheme. Notice that the rank-R, $\boldsymbol{\mathcal{T}}[\mathcal{C}_{R,\mathbf{r}}]$-format is embedded into the class $\boldsymbol{\mathcal{T}}_{\mathbf{r}}[TT[\mathbf{r}_1]]$ with $\mathbf{r}_1 = (R, ..., R)$ (cf. Lemma 2.2, (C)). Hence, the further TT-rank optimisation of the initial element in $\boldsymbol{\mathcal{T}}[\mathcal{C}_{R,\mathbf{r}}]$ can be accomplished with the SVD-based scheme applied to the small mode-size core tensor in $\boldsymbol{\mathcal{C}}_{r_1,\mathbf{r}}$, $(r \ll N)$.

Another tensor format that might be useful in numerical multilinear algebra is specified as a set of N-d tensors in $\boldsymbol{\mathcal{C}}_{R,\mathbf{n}}$ with $N = q^L$, where each N-vector composing rank-1 terms is represented by the q-L quantized tensor in the $TC[\mathbf{r}, L]$ format. We further denote by $\boldsymbol{\mathcal{C}}_{R,\mathbf{n}}[TC[\mathbf{r}, L]]$ the set of *canonical-TC* tensors represented in form

$$\mathbf{V} = \sum_{k=1}^{R} c_k T_k^{(1)} \times_2 T_k^{(2)} ... \times_d T_k^{(d)} \in \boldsymbol{\mathcal{C}}_{R,\mathbf{n}}[TC[\mathbf{r}, L]], \tag{2.7}$$

where, for $\nu = 1, ..., d$, $T_k^{(\nu)} := \bigotimes_{\ell=1}^{L} \mathbf{G}_{k,\nu}^{(\ell)} \in TC[\mathbf{r}, L]$ with small-size cores $\mathbf{G}_{k,\nu}^{(\ell)} \in \mathbb{R}^{r_{\ell-1} \times q \times r_\ell}$. The complexity of respective representation scales logarithmically in N, $O(Rr^2 d \log N)$, hence, it has advantages for large mode-size N (see §2.3).

3 Multilinear approximation on tensor manifolds

In the following, we choose the manifold $\mathcal{S}$ of rank-structured tensors as one of the above defined tensor classes, and call the elements in $\mathcal{S}$ as $\mathcal{S}$-tensors.

To perform computation over nonlinear manifold (say, the truncated iteration) we need to perform a "projection" of the current iterand onto $\mathcal{S}$. This action is fulfilled by using the tensor truncation operator $T_{\mathcal{S}} : \mathbb{W}_{\mathbf{n},d} \to \mathcal{S}$ defined by

$$\mathbf{A}_0 \in \mathcal{S}_0 \subset \mathbb{W}_{\mathbf{n},d} : \quad T_{\mathcal{S}}\mathbf{A}_0 = \mathrm{argmin}_{\mathbf{T}\in\mathcal{S}} \|\mathbf{A}_0 - \mathbf{T}\|_{\mathcal{S}}, \tag{3.1}$$

that is a challenging nonlinear approximation problem. In practice, the computation of the minimiser $T_{\mathcal{S}}\mathbf{A}_0$ can be performed only approximately. The replacement of $\mathbf{A}_0$ by its approximation in $\mathcal{S}$ is called the *tensor truncation* to $\mathcal{S}$ and denoted by $T_{\mathcal{S}}\mathbf{A}_0$. We discuss the analytic and algebraic methods of approximate solution to the problem (3.1) for different classes of rank-structured tensors $\mathcal{S}$.

3.1 Analytic methods of approximation

Here we discuss the low rank approximation of a special class of higher-order tensors, further called function-related tensors (FRTs), obtained by sampling the multi-variate function over tensor grid in $\mathbb{R}^d$. They directly arise from:

(a) A separable approximation of multi-variate functions;

(b) Nyström/collocation/Galerkin discretisations of integral operators;

(c) The tensor-product approximation of some analytic matrix-valued functions.

The constructive analytic approximation methods are based on *sinc*-quadrature representations [91]. It applies, in particular, to the class of Green kernels (the Poisson, Yukawa, Helmholtz potentials), cf. [69, 32], to certain functions arising in the Boltzmann equation, in electronic structure calculations [39, 4, 57, 58], and to correlation functions in construction the Karhünen-Loéve expansion [88].

3.1.1 Error estimate in terms of analytic generating function

In the following we define FRTs corresponding to collocation-type discretization.

Given the function $g : \Omega := \Pi^d \to \mathbb{R}$, with $\Pi = [a, b]^p$ and $p = 1, 2, 3$, grid-size $n \in \mathbb{N}$, and the mesh-size $h = (b - a)/n$. We denote by $\{x^{(1)}_{\mathbf{i}_1}, ..., x^{(d)}_{\mathbf{i}_d}\}$ with $\mathbf{i}_\ell \in \mathcal{I}_\ell := I_{\ell,1} \times ... \times I_{\ell,p}$ $(\ell = 1, ..., d)$ a set of collocation points living in the midpoints of the tensor grid with mesh-size $h = (b - a)/n$. Here for $\mathbf{i}_\ell = (i_{\ell,1}, ..., i_{\ell,p}) \in \mathcal{I}_\ell$, we have $i_{\ell,m} \in I_n := \{1, ..., n\}$ $(m = 1, ..., p)$.

In our applications we have $d \geq 2$ with some fixed $p \in \{1, 2, 3\}$. In particular, the matrix (operator) decompositions correspond to the choice $p = 2$. In this case we introduce the reordered index set of pairs $\mathcal{M}_\ell := \{\mathbf{m}_\ell : \mathbf{m}_\ell = (i_\ell, j_\ell),\ i_\ell, j_\ell \in I_n\}$ $(\ell = 1, ..., d)$, so that $\mathcal{I} = \mathcal{M}_1 \times ... \times \mathcal{M}_d$ with $\mathcal{M}_\ell = I_n \times I_n$.

The Nyström and Galerkin approximations to function related tensors are discussed in [33, 49]. Here we focus on the collocation-type schemes [32], which are based on tensor-product ansatz functions

$$\psi^{\mathbf{i}}(y_1, ..., y_d) = \prod_{\ell=1}^{d} \psi_\ell^{i_\ell}(y_\ell), \quad \mathbf{i} = (i_1, ..., i_d) \in \mathcal{I}_1 \times ... \times \mathcal{I}_d. \tag{3.2}$$

Definition 3.1 *(Collocation, FRT(C)). Given the tensor-product basis set (3.2), we introduce the variable* $\zeta_{i_\ell}^{(\ell)} := (x_{i_\ell}^{(\ell)}, y_\ell)$ *with the collocation point* $x_{i_\ell}^{(\ell)}$ *and* $y_\ell \in \Pi$*, the pair* $\mathbf{m}_\ell := (i_\ell, j_\ell) \in \mathcal{M}_\ell$ *and define the collocation-type d-th order FRT by* $\mathbf{A} \equiv \mathbf{A}(g) := [a_{\mathbf{m}_1 \dots \mathbf{m}_d}] \in \mathbb{R}^{\mathcal{M}_1 \times \dots \times \mathcal{M}_d}$ *with*

$$a_{\mathbf{m}_1 \dots \mathbf{m}_d} := \int_\Omega g(\zeta_{i_1}^{(1)}, ..., \zeta_{i_d}^{(d)}) \psi^{\mathbf{j}}(y_1, ..., y_d) dy, \quad \mathbf{m}_\ell \in \mathcal{M}_\ell. \tag{3.3}$$

The key observation is that there is a natural duality between separable approximation of the multi-variate generating function and the tensor-product decomposition of the related multi-dimensional array. Hence, the canonical decompositions can be derived by using a corresponding separable expansion of the generating function g (see [33, 35] for more details).

Lemma 3.2 *([32]) Suppose that a multi-variate function* $g : \Omega \subset \mathbb{R}^{\mathbf{d}} \to \mathbb{R}$ *can be approximated by a separable expansion*

$$g_r(\boldsymbol{\zeta}) := \sum_{k=1}^{r} \mu_k \Phi_k^{(1)}(\zeta^{(1)}) \cdots \Phi_k^{(d)}(\zeta^{(d)}) \approx g(\boldsymbol{\zeta}), \quad \boldsymbol{\zeta} = (\zeta^{(1)}, ..., \zeta^{(d)}) \in \mathbb{R}^{\mathbf{d}}, \tag{3.4}$$

where $\mu_k \in \mathbb{R}$ *and* $\Phi_k^\ell : \Pi \subset \mathbb{R}^2 \to \mathbb{R}$*. Define the canonical decomposition (2.2) via* $\mathbf{A}_{(r)} := \mathbf{A}(g_r)$ *(cf. Definition 3.1) with the choice,*

$$V_k^{(\ell)} = \left\{ \int \Phi_k^{(\ell)}(\zeta_i^{(\ell)}) \psi_\ell^j(y_\ell) dy_\ell \right\}_{(i,j) \in \mathcal{M}_\ell} \in \mathbb{R}^{\mathcal{I}_\ell \times \mathcal{J}_\ell}, \quad \ell = 1, ..., d, \quad k = 1, ..., r. \tag{3.5}$$

Then the FRT(C) $\mathbf{A}_{(r)}$ *allows the error estimate*

$$\|\mathbf{A}(g) - \mathbf{A}_{(r)}(g_r)\|_\infty \le C \|g - g_r\|_{L^\infty(\Omega)}.$$

Though in general a decomposition (3.4) with small separation rank r is a complicated numerical task, in many interesting applications efficient approximation methods are available. In particular, for a class of multi-variate functions (say, for certain shift-invariant Green's kernels in $\mathbb{R}^d$) it is possible to obtain a dimensionally-independent Kronecker rank $r = \mathcal{O}(\log n |\log \varepsilon|)$, e.g., based on *sinc*-quadrature methods or an approximation by exponential sums (see examples in [33, 49]).

Next, we discuss the constructive canonical and Tucker decomposition of FRTs applied to a general class of *analytic generating functions* characterized in terms of their Laplace transform.

3.1.2 *sinc*-quadrature approximation in the Hardy space

We use constructive approximation based on the *sinc*-quadrature methods. For the readers convenience we recall the standard approximation results by the *sinc*-methods (cf. [91, 24]). First, we introduce the Hardy space $H^1(D_\delta)$ as the set of all complex-valued functions f, which are analytic in the strip $D_\delta := \{z \in \mathbb{C} : |\Im\mathfrak{m}\, z| < \delta\}$, such that

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

Given $f \in H^1(D_\delta)$, $\mathfrak{h} > 0$, and $M \in \mathbb{N}_0$, the corresponding sinc-quadrature reads as

$$T_M(f, \mathfrak{h}) := \mathfrak{h} \sum_{k=-M}^{M} f(k\mathfrak{h}) \approx \int_{\mathbb{R}} f(\xi) \mathrm{d}\xi. \tag{3.6}$$

Proposition 3.3 *Let $f \in H^1(D_\delta)$, $\mathfrak{h} > 0$, and $M \in \mathbb{N}_0$ be given. If*

$$|f(\xi)| \le C \exp(-b|\xi|) \qquad \text{for all } \xi \in \mathbb{R} \text{ with } b, C > 0, \tag{3.7}$$

then the quadrature error satisfies

$$\left| \int_{\mathbb{R}} f(\xi) \mathrm{d}\xi - T_M(f, \mathfrak{h}) \right| \le C e^{-\sqrt{2\pi\delta b M}} \quad \text{with } \mathfrak{h} = \sqrt{2\pi\delta / bM}$$

and with a positive constant C depending only on f, δ, b (cf. [91]). If f possesses the hyper-exponential decay

$$|f(\xi)| \le C \exp(-b \mathrm{e}^{a|\xi|}) \qquad \text{for all } \xi \in \mathbb{R} \quad \text{with } a, b, C > 0, \tag{3.8}$$

then the choice $\mathfrak{h} = \log(\frac{2\pi a M}{b}) / (aM)$ leads to (cf. [24])

$$\left| \int_{\mathbb{R}} f(\xi) \mathrm{d}\xi - T_M(f, \mathfrak{h}) \right| \le C\, N(f, D_\delta)\, \mathrm{e}^{-2\pi\delta a M / \log(2\pi a M / b)}.$$

Note that $2M + 1$ is the number of quadrature/interpolation points. If f is an even function, the number of quadrature/interpolation points reduces to $M + 1$.

3.1.3 Error bounds for canonical decomposition

We consider a class of multi-variate functions $g : \mathbb{R}^d \to \mathbb{R}$ parametrised by $g(\boldsymbol{\zeta}) = G(\rho(\boldsymbol{\zeta})) \equiv G(\rho)$ with $\rho \equiv \rho(\boldsymbol{\zeta}) = \rho_1(\zeta^{(1)}) + ... + \rho_d(\zeta^{(d)}) > 0$, $\rho_\ell : \mathbb{R}^2 \to \mathbb{R}_+$, where the univariate function $G : \mathbb{R}_+ \to \mathbb{R}$ can be represented via the Laplace transform

$$G(\rho) = \int_{\mathbb{R}_+} \mathcal{G}(\tau) e^{-\rho\tau} d\tau.$$

The FRT(C) approximation corresponds to $p = 2$, $\zeta^{(\ell)} = (x_\ell, y_\ell)$ (cf. Definition 3.1). Without loss of generality, we introduce one and the same scaling function

$$\psi^i(\cdot) = \psi(\cdot + (i-1)h), \quad i \in I_n, \tag{3.9}$$

for all spatial dimensions $\ell = 1, ..., d$, where $h > 0$ is the mesh parameter. We simplify further and set $\rho \equiv \rho(\boldsymbol{\zeta}) = \sum_{\ell=1}^d \rho_0(\zeta^{(\ell)})$, i.e., $\rho_\ell = \rho_0(x_\ell, y_\ell)$ $(\ell = 1, ..., d)$ with $\rho_0 : [a, b]^2 \to \mathbb{R}_+$. For $i \in I_n$, let $\{\bar{x}_i\}$ be the set of cell-centered collocation points on $[a, b]$. For each $i, j \in I_n$, we introduce the parameter dependent integral

$$\Psi_{i,j}(\tau) := \int_{\mathbb{R}^2} e^{-\rho_0(\bar{x}_i, y)\tau} \psi(y + (j-1)h) dy, \quad \tau \ge 0. \tag{3.10}$$

Theorem 3.4 *(FRT(C) approximation [32]). Assume that:*

*(**a**) $\mathcal{G}(\tau)$ has an analytic extension $\mathcal{G}(w)$, $w \in \Omega_{\mathcal{G}}$, into a certain domain $\Omega_{\mathcal{G}} \subset \mathbb{C}$, that can be mapped conformally onto the strip D_δ, such that $w = \varphi(z)$, $z \in D_\delta$ and $\varphi^{-1} : \Omega_{\mathcal{G}} \to D_\delta$;*

*(**b**) for all $(\mathbf{i}, \mathbf{j}) \in \mathcal{I} \times \mathcal{J}$ the transformed integrand*

$$f(z) := \varphi'(z) \mathcal{G}(\varphi(z)) \prod_{\ell=1}^d \Psi_{i_\ell j_\ell}(\varphi(z)) \tag{3.11}$$

belongs to the Hardy space $H^1(D_\delta)$ with $N(f, D_\delta) < \infty$ uniformly in $(\mathbf{i}, \mathbf{j})$;

*(**c**) the function $f(t)$, $t \in \mathbb{R}$, in (3.11) has either exponential (c1) or hyper-exponential (c2) decay as $t \to \pm\infty$ (see Proposition 3.3).*

*Under the assumptions (**a**)-(**c**), we have that, for each $M \in \mathbb{N}$, the FRT(C), $\mathbf{A}(g)$, defined on $[a, b]^d$, allows an exponentially convergent symmetric[2] canonical approximation $\mathbf{A}_{(r)} \in \boldsymbol{\mathcal{C}}_r$ with $V_k^{(\ell)}$*

[2] A d-th order tensor is called symmetric if it is invariant under arbitrary permutations of indices in $\{1, ..., d\}$.

as in (3.5), where the expansion (3.4) is obtained by the substitution of f from (3.11) into the sinc*-quadrature (3.6), such that we have*

$$\|\mathbf{A}(g) - \mathbf{A}_{(r)}\|_\infty \le Ce^{-\alpha M^\nu} \quad \textit{with } r = 2M+1, \tag{3.12}$$

where $\nu = \frac{1}{2}$, $\alpha = \sqrt{2\pi\delta b}$ *in the case (c1) and* $\nu = 1$, $\alpha = \frac{2\pi\delta b}{\log(2\pi aM/b)}$ *in the case (c2).*

Theorem 3.4 proves the existence of the canonical decomposition of the FRT $\mathbf{A}(g)$ with the Kronecker rank $r = \mathcal{O}(|\log\varepsilon|\log 1/h)$ (in the case (c2)) or $r = \mathcal{O}(\log^2\varepsilon)$ (in the case (c1)), which provide an approximation of order $\mathcal{O}(\varepsilon)$. In our applications we usually have $1/h = \mathcal{O}(n)$, where n is the number of grid-points in one spacial direction. Theorem 3.4 applies to translation invariant or spherically symmetric functions, e.g. to the classical Newton, Yukawa and Helmholtz kernels $\frac{1}{\|x-y\|}$, $\frac{e^{-\lambda\|x-y\|}}{\|x-y\|}$ and $\frac{\cos(\lambda\|x-y\|)}{\|x-y\|}$ with $x, y \in \mathbb{R}^3$, $\lambda > 0$.

3.2 SVD-based and ALS-type optimization

The nonlinear approximation problem (3.1) can be solved approximately by algebraic methods that may combine the SVD-based procedures to compute the initial guess, with nonlinear iterations like ALS, gradient or quasi-Newton methods to find the local mimima. We refer to [65, 42, 85] and references therein concerning iterative methods of Tucker-type approximation (computation of local minima in (3.1)).

In the following, we briefly discuss the quasi-optimal direct (non-iterative) approximation based on the robust QR/SVD matrix decompositions. In our context, such methods can be basically considered as a kind of rank truncation procedures applied to the target tensors already given in some tensor-structured format, but with unsatisfactory large rank parameters. Specifically, these may be canonical, Tucker, TT/TC or QTT input tensors.

Since the approximation procedure in the TT-format is performed by finite SVD-based algorithm, we arrive at the robust SVD-based quasi-optimal TT/QTT approximation [79, 72, 55, 61].

Proposition 3.5 *(Tensor truncation). The operator* $T_{\mathcal{S}} : \mathbb{W}_{\mathbf{q},dL} \to \mathcal{S} := TT[\mathbf{r},\mathbf{q},dL]$ *is well defined [55].*

For given $\mathbf{A}_0 \in TT[\mathbf{r}_0,\mathbf{q},dL] \subset \mathbb{W}_{\mathbf{q},dL}$, *and* $\mathbf{r} < \mathbf{r}_0$, *the quasi-optimal approximation to* $T_{\mathcal{S}}\mathbf{A}_0$ *can be computed by QR/SVD based algorithm in* $O(qdLr_0^3)$ *operations [79, 72].*

The quasi-optimal Tucker approximation can be computed by the truncation of singular matrices from higher order SVD (HOSVD), cf. [15].

If $\mathbf{A}_0 \in \mathcal{C}_R$, *then the quasi-optimal rank-*$\mathbf{r}$ *Tucker approximation can be computed by SVD-based RHOSVD algorithm, cf. [57].*

Best orthogonal Tucker approximation can be computed by the ALS iteration provided that the initial guess is given by the HOSVD/RHOSVD methods. Likewise, the best TT/TC tensor approximations can be calculated by ALS iteration starting with the SVD-based representations as above. The ALS iteration in TT format was considered in [84].

3.3 Quantics approximation of function related tensors

The quantics representation of exponential-trigonometric N-vectors can be adapted to the more general class of multidimensional N-d tensors. First, one can apply the previous results to a sum of R exponential-trigonometric terms in d dimensions [55].

Lemma 3.6 *Given $N = q^L$, with some $q = 2, 3, ...$ and $L \in \mathbb{N}$. Let $c_k, z_{k,\ell} \in \mathbb{C}$, ($k = 1, ..., R$, $\ell = 1, ..., d$), be given, then the exponential sum N-d tensor,*

$$\mathbf{A}_{(\mathbf{n},d)} := \{x_{\mathbf{n}} := \sum_{k=1}^{R} c_k \prod_{\ell=1}^{d} z_{k,\ell}^{n_\ell - 1}\}_{\mathbf{n} \in I^{\otimes d}}, \quad I = \{1, ..., N\}, \tag{3.13}$$

can be reshaped to the rank-R, q-dL quantics tensor, $\mathcal{F}_{q,d,L} : \mathbf{A} \to \mathbf{B}_{(q,dL)} \in \boldsymbol{\mathcal{C}}_{R,\mathbf{q}}[TT[\mathbf{1}, dL]]$. The number of representation parameters amounts to $dqR \log_q N$.

Notice that the outer product of d rank-r tensors of order L leads, in general, to the canonical rank-r^{d-1} tensor in the dL-dimensional tensor space. This is the reason why in the case of exponential sum tensor in (3.13), we obtain the better structured rank-R representation in auxiliary dimension, which is not the case for the trigonometric R-term sum of tensors.

To describe the analytic tensor approximation, let us consider the class E_M of function related tensors generated by certain (analytic) functions f that allow the efficient approximation in the set of exponential sums on $[a, b] \subset \mathbb{R}_+$,

$$E_M := \left\{ u(x) = \sum_{k=1}^{M} c_k e^{-t_k x}, \quad c_k, t_k \in \mathbb{R} \right\}, \quad x \in [a, b].$$

In particular, such approximations can be based on the sinc method.

Conjecture 3.7 *([55]) Based on our extensive numerical tests, we assume that the Gaussian-, polynomial- and sinc-vectors obtained via the uniform sampling, allow the quantics approximation by the q-folding, whose TT-rank remains bounded by a small constant (at most 4) uniformly in the vector size N (see Table 3.1).*

Next statement proves the error bound for the *semi-analytic quantics approximation* of function related tensors.

Lemma 3.8 *([55]) Suppose that for given continuous function $f : [a, b] \to \mathbb{R}$, and given $\varepsilon > 0$, there is $u \in E_M$, such that*

$$\max_{x \in [a,b]} \left| f(x) - \sum_{k=1}^{M} c_k e^{-t_k x} \right| \le \varepsilon. \tag{3.14}$$

Then for any $N = q^L$, with some $q = 2, 3, ...$, and $L \in \mathbb{N}$, we have:

(A) The function related N-d tensor $\mathbf{F} = [F_{\mathbf{i}}]$, defined by the entries

$$F_{\mathbf{i}} = f(hi_1 + hi_2 + ... + hi_d), \quad \mathbf{i} \in I^{\otimes d}, \quad a \le dh \le b/N, \quad h > 0,$$

and discretising the multivariate function $g(x_1, ..., x_d) = f(\sum_{\ell=1}^{d} x_\ell)$ over the uniform grid can be approximated by the rank-M, q-dL tensor up to the tolerance ε in the max-norm.

(B) Under condition $a \le dh \le b/N$, the function related N-d tensor $\mathbf{G} = [G_{\mathbf{i}}]$, with the entries

$$G_{\mathbf{i}} = f(\sum_{\ell=1}^{d} x_{\ell,i_\ell}^2), \quad x_{\ell,i_\ell} = \sqrt{hi_\ell}, \quad \mathbf{i} \in I^{\otimes d},$$

discretising the multivariate function $g_1(x_1, ..., x_d) = f(\sum_{\ell=1}^{d} x_\ell^2)$ on the quadratically graded grid $\{x_{\ell,i_\ell}\}$, embedded into the region $a \le \sum_{\ell=1}^{d} x_\ell^2 \le b$, can be approximated by the rank-M, q-dL tensor with the tolerance ε in the max-norm.

In both cases, the number of representation parameters is bounded by $O(dqM \log_q N)$.

Lemma 3.8 allows us to derive the accurate $O(d \log N)$-approximations to the wide class of function related tensors in high dimension. For a class of analytic functions the basic approximability assumption (3.14) can be verified with

$$\varepsilon = O(e^{-\alpha M/\log M}), \quad \alpha > 0,$$

by applying the *sinc*-approximation [32, 49, 53]. The semi-analytic quantics approximations via Lemma 3.8 can be further optimized by applying the rank-r, $r < M$, TT-compression.

Based on Conjuncture 3.7, the quantics representations similar to those in Lemma 3.8 can be derived based on the product-polynomial or product-trigonometric, exponentially convergent approximations, taking into account (3.14).

3.4 Numerics on QTT approximation of N-vectors and N-d-tensors

The following numerics illustrates the behavior of approximation error vs. the TT-rank of the dyadic folding approximation ($q = 2$) applied to different classes of function related vectors/tensors. We apply the MATLAB subroutines in [73] implementing the binary folding of vectors and the low rank TT approximation.

Recall that for any $q = 2, 3, ...$, the QTT-rank (or more precisely, QTT_q-rank) of the exponential and trigonometric vectors equals 1 and 2, respectively. Moreover, it can be proven that the Kronecker-TT rank of the diagonal matrices generated by the exponential and trigonometric vectors equals 1 and 2, respectively (see Lemma 3.12). Furthermore, the product of exponential and sin-functions in arbitrary frequency range has the Q-rank equals to 2, as for the single sin-function.

$N \setminus \overline{r}$	$e^{-\alpha x^2}, \alpha = 0.1, 1, 10, 10^2$	$\frac{\sin(\alpha x)}{x}, \alpha = 1, 10, 10^2$	$1/x$	e^{-x}/x	$x, x^{10}, x^{1/10}$
2^{10}	3.2/2.8/2.8/2.2	4.0/4.7/5.5	4	3.5	1.9/2.7/3.9
2^{12}	3.1/2.9/2.9/2.6	3.8/4.8/5.6	4.2	3.8	1.9/2.6/3.9
2^{14}	2.9/2.8/2.8/2.8	3.6/4.7/5.5	4.2	3.8	1.9/2.5/3.9
2^{16}	2.8/2.7/2.8/2.8 (0.03)	3.6/4.5/5.4 (0.048)	4.2 (0.05)	5.3 (0.04)	1.9/2.4/3.9

Table 3.1: QTT_2-ranks of functional N-vectors on large grids, $N = 2^p$.

Tables 3.1 and 3.2, present numerical results on the binary quantics ε-approximation with $\varepsilon = 10^{-6}$, of the function related vectors/tensors corresponding to monomials and fractional power of x, as well as to the functions

$$\frac{1}{x_1 + ... + x_d}, \quad \frac{1}{\|x\|}, \quad \frac{e^{-\alpha\|x\|}}{\|x\|}, \quad e^{-\alpha\|x\|}, \quad \frac{\sin(\alpha\|x\|)}{\|x\|} \quad x \in \mathbb{R}^d.$$

For the following discussion, the average splitting rank of the TT-model, $\overline{r}$, is defined by

$$\overline{r} := \sqrt{\frac{1}{d}\sum_{\ell-1}^{d} r_{\ell=1} r_\ell},$$

providing the complexity bound $\leq 2d\overline{r}^2 \log N$. The CPU time (sec.) corresponding to the finest grid is given in the brackets (see Table 3.1). It scales linearly in the input vector size, N. Notice for comparison that the FFT_N on finest grid requires $t_{FFT}(2^{16}) \approx 0.006$ sec., indicating that the QTT_2 vector-transform to the low rank TT-format in higher dimension $D = \log N$, is almost as fast as the FFT on the same vector size.

Table 3.2 represents the QTT_2-ranks of the functional $N \times N$-arrays generated by sampling over large equidistant grids on $[0,1]^2$, approximated up to the tolerance $\varepsilon = 10^{-6}$. Here Δ_d denotes the d-Laplacian as in (3.16). The matrix notation $diag(f(x))$ means the diagonal $N \times N$-matrix built by the N-vector generated via $f(x)$ sampled on the uniform N-grid. Approximation properties by QTT-format are similar in the case $d \geq 3$.

$N \setminus \overline{r}$	$1/(x_1+x_2)$	$e^{-\|x\|}$	$e^{-\|x\|^2}$	$diag(e^{-x^2})$	$\Delta_2^{-1}\mathbf{1}, \varepsilon = 10^{-6}, 10^{-7}, 10^{-8}$
2^9	5.0	9.4	7.8	3.8	3.6/3.6/3.6
2^{10}	5.1	9.4	7.7	3.9	3.6/3.6/3.6
2^{11}	5.2	9.3	7.5	3.9	3.7/3.7/3.7

Table 3.2: QTT_2-ranks of functional $N \times N$-arrays on large grids, $N = 2^p$.

The above numerical illustrations lead to the following observations.

- The QTT-rank remains almost independent on the vector/matrix size, hence being specified only by analytic properties of generating function.
- The QTT-rank of the discrete Gaussians and polynomials is very small (≤ 3), that means that any M-term exponential or polynomial expansion on large N-grid can be represented with the complexity $O(M \log N)$, and with small constant in front of.
- Vectors generated by singular functions like $1/x, e^{-x}/x$, $\sin(\alpha x)/x$, and x^α, $(\alpha > 0)$ exhibit almost the same QTT-rank as analytic functions, uniformly in N.
- The computation time for TT-approximation scales linearly in the grid size N, hence, being proportional to the cost of FFT_N transform.

3.5 Tensor-structured representation of matrices on N-d tensors

3.5.1 General definitions

The MPS decomposition induces the important concept of matrix product operators (MPO) acting between two tensor-product Hilbert spaces, $\mathbf{A} : \mathbb{X} = \otimes X_\ell \to \mathbb{Y} = \otimes Y_\ell$, each of dimension d.

Definition 3.9 *Introduce the rank-***r** *operator TC (OTC) decomposition symbolized by a set of factorized operators* **A**,

$$\mathbf{A} = \sum_{\alpha \in \mathcal{J}} A^{(1)}(\alpha_d, \alpha_1) A^{(2)}(\alpha_1, \alpha_2) \cdots A^{(d)}(\alpha_{d-1}, \alpha_d),$$

with $A^{(\ell)} = [A^{(\ell)}(\alpha_\ell, \alpha_{\ell+1})]$ *being the operator valued* $r_\ell \times r_{\ell+1}$ *matrix, where* $A^{(\ell)}(\alpha_\ell, \alpha_{\ell+1}) : X_\ell \to Y_\ell$, $(\ell = 1, ..., d)$, *or in the index notation*

$$\begin{aligned} \boldsymbol{A}(i_1, j_1, \ldots, i_d, j_d) &= \sum_{\alpha_1=1}^{r_1} \cdots \sum_{\alpha_{d-1}=1}^{r_{d-1}} A^{(1)}(i_1, j_1, \alpha_1) A^{(2)}(\alpha_1, i_2, j_2, \alpha_2) \cdot \ldots \cdot \\ &\cdot \; A^{(D-1)}(\alpha_{d-2}, i_{d-1}, j_{d-1}, \alpha_{d-1}) A^{(D)}(\alpha_{d-1}, i_d, j_d). \end{aligned} \tag{3.15}$$

The action $\mathbf{A}X$ *on rank-*$\mathbf{r}_X$ *TT-tensor* $X \in \mathbb{X}$ *is defined as the TT/TC element in* $\mathbb{Y}$,

$$(\mathbf{A}X)(i_1, ..., i_d) := \sum_{\alpha \in \mathcal{J}} Y^{(1)}(\alpha_d, i_1, \alpha_1) Y^{(2)}(\alpha_1, i_2, \alpha_2) \cdots Y^{(d)}(\alpha_{d-1}, i_d, \alpha_d),$$

with

$$Y^{(\ell)}(i_\ell) = \sum_{j_\ell=1}^{n_\ell} A^{(\ell)}(i_\ell, j_\ell) \otimes X^{(\ell)}(j_\ell).$$

If a matrix is considered merely as a *vector* in (3.15), and its neither possibility to map vectors to vectors nor related properties are taken into consideration, then we arrive at the same concept of matrix *vector TT ranks*.

Definition 3.10 *A multi-way $m_1 \times n_1 \times \ldots \times m_d \times n_d$-matrix*

$$\boldsymbol{A} : \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_d} \mapsto \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_d}$$

is given, its k-th vector TT rank *is the rank of its unfolding $\boldsymbol{A}^{(k)}$ ($1 \le k \le d-1$) with the elements*

$$\boldsymbol{A}^{(k)}\left(i_1 j_1 \ldots i_k j_k \ ; \ i_{k+1} j_{k+1} \ldots i_d j_d\right) \ = \ \boldsymbol{A}(i_1 j_1 \ldots i_d j_d).$$

In particular this means that the minimal vector ranks of TT decomposition of a certain matrix are somewhat independent from one other, depending on the matrix in the aggregate. So we may consider a minimal rank decomposition, which it holds for that no one of $d-1$ ranks can be reduced without introducing an error in (3.15) even if we allow the others to grow. This makes it reasonable to compare ranks elementwise.

Definition 3.11 *Let us say that a multiway matrix (vector) is of ranks not greater than $r_1 \ldots r_{d-1}$ if and only if for any k: $1 \le k \le d-1$ its k-th vector TT rank is not greater than r_k.*

Their complexity of storage and such basic operations as dot product, multi-dimensional contraction, matrix-by-vector multiplication, rank reduction and orthogonalization of a tensor train is linear with respect to vector TT rank upper bound raised to the power 2 or 3.

Next lemma shows that the diagonal matrices generated by the exponential and trigonometric vectors can be proven to have fixed vector QTT rank [55].

Lemma 3.12 *Quantics Kronecker-TT rank of the diagonal matrices generated by the exponential and trigonometric vectors of size $N = q^L$, equals* 1 *and* 2*, respectively.*

Developing iterative solvers we are likely to be concerned with vector TT ranks of a matrix-by-vector product. Below we introduce the concept of *operator TT rank*.

Definition 3.13 *A multi-way matrix $\boldsymbol{A} : \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_d} \mapsto \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_d}$ given, for any vector $\boldsymbol{X} \in \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_d}$ let us denote vector TT ranks of the matrix-by-vector product $\boldsymbol{AX}$ by $r_1 \ldots r_{d-1}$. Then let us refer to*

$$\max_{\substack{k=1\ldots d-1, \\ \boldsymbol{X} \text{ is of vector TT rank } 1\ldots 1}} r_k$$

as the operator TT rank *of $\boldsymbol{A}$.*

Obviously, the operator TT rank does not exceed the maximum component of vector TT rank.

The QTT rank estimates for a class of Laplacian related operators will be discussed in §3.5.3.

3.5.2 Matrix exponential

We are interested in tensor representation of a matrix exponential $\exp(B)$, where B is a matrix that is already presented in some Kronecker tensor-product format. For the (approximate) computation of the matrix exponential, the traditional *scaling-and-squaring method* can be used, which is the most suitable for rank-structured formats. The matrix B is first scaled by $\frac{1}{2^s}$ with $s \approx \log_2 ||B||$, so that $C := \frac{1}{2^s}||B|| \leq 1$. Then for a scaled matrix C the exponent is computed by truncated Taylor series

$$\exp(C) \approx \sum_{k=0}^{p} \frac{C^k}{k!},$$

in p multiplications using a Horner rule: $C_k = \frac{1}{k} C_{k+1} C + I,\ k = p-1, p-2, \ldots, 1, 0$, and $C_p = I$. All multiplications are performed in a structured format (QTT format here) and the compression is done at each step. After $\exp(C) = C_0$ is computed, the target exponent is obtained by squaring,

$$\exp(B) = \exp(C)^{2^s},$$

in $s \approx \log_2 ||B||$ steps. This iteration requires $p + s$ matrix-by-matrix multiplications with the subsequent rank compression. In the case of rank-r QTT format the complexity is of order $\mathcal{O}(r^6 \log N)$ operations, since after multiplication of two rank-r QTT matrices the result has, in general, the TT ranks r^2, and recompression requires $\mathcal{O}(r^6 \log N)$ operations. It is worth to note that the recompression step usually dominates over the multiplication step, since all mode sizes are equal to 4.

3.5.3 Elliptic operator and its inverse

Our basic example would be the finite difference negative d-Laplacian over uniform tensor grid. It is known to have the Kronecker rank-d representation,

$$\Delta_d = A \otimes I_N \otimes \ldots \otimes I_N + I_N \otimes A \otimes I_N \ldots \otimes I_N + \ldots + I_N \otimes I_N \ldots \otimes A \in \mathbb{R}^{I^{\otimes d} \times I^{\otimes d}}, \tag{3.16}$$

with $A = \Delta_1 = \mathrm{tridiag}\{-1, 2, -1\} \in \mathbb{R}^{N \times N}$, and I_N being the $N \times N$ identity.

Notice that for the canonical rank we have $rank_C(\Delta_d) = d$, while TT-rank of Δ_d is equal to 2 for any dimension due to the explicit representation

$$\Delta_d = \begin{bmatrix} \Delta_1 & I \end{bmatrix} \bowtie \begin{bmatrix} I & 0 \\ \Delta_1 & I \end{bmatrix}^{\otimes (d-2)} \bowtie \begin{bmatrix} I \\ \Delta_1 \end{bmatrix},$$

where the rank product operation "$\bowtie$" is defined as a regular matrix product of the two corresponding core matrices, their blocks being multiplied by means of tensor product (see [45]). The similar bound is true for the Tucker rank $rank_{Tuck}(\Delta_d) = 2$.

The explicit rank-4 QTT representation of Δ_d for $d \geq 2$ is obtained in [45], while for $d = 1$, we have

$$\Delta_1 = \begin{bmatrix} I & J' & J \end{bmatrix} \bowtie \begin{bmatrix} I & J' & J \\ & J & \\ & & J' \end{bmatrix}^{\otimes (d-2)} \bowtie \begin{bmatrix} 2I - J - J' \\ -J \\ -J' \end{bmatrix},$$

with the Pauli matrices

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The analysis of the low QTT-rank approximations of elliptic operator inverse for $d \geq 2$ is based on certain assumptions on the QTT-rank of the matrix exponential family.

Conjecture 3.14 *For any given $\varepsilon > 0$, and for fixed $a, b > 0$, let us assume that the family of matrix exponentials, $\{\exp(-t_k\Delta_1)\}$, $t_k > 0$, $k = -M, ..., M$, allows the QTT ε-approximation with rank-r_Δ being uniformly bounded in the grid size N and in the scaling factors $t_k \in [a, b] \subset \mathbb{R}_{>0}$ (see Table 3.3 for numerical justification).*

With the previous assumption on the QTT-rank of the family of "univariate" matrix exponentials, $\{\exp(-t_k\Delta_1)\}$, $t_k > 0$, $\Delta_1 \in \mathbb{R}^{N\times N}$, one can prove the following Lemma [55].

Lemma 3.15 *Under assumptions of Conjecture 3.14 the matrix*

$$\mathcal{B}_M := \sum_{k=-M}^{M} c_k \bigotimes_{\ell=1}^{d} \exp(-t_k a_\ell \Delta_1), \quad t_k = e^{k\mathfrak{h}}, \quad c_k = \mathfrak{h}t_k, \quad \mathfrak{h} = \pi/\sqrt{M}, \tag{3.17}$$

possesses the rank-$O(\log^2 \varepsilon)$, QTT ε-approximation (or preconditioner if M is small) to the anisotropic d-Laplacian inverse $\Delta_{d,\alpha}^{-1}$, where

$$\Delta_{d,\alpha} := \sum_{\ell=1}^{d} a_\ell \bigotimes_{k=1}^{d} \Delta_1^{\delta_{\ell,k}}, \quad a_\ell > 0, \quad \delta_{\ell,k} \quad \textit{is the Kronecker symbol.} \tag{3.18}$$

The following Lemma summarize the previous discussion (see also [45]).

Proposition 3.16 *The following canonical and TT/QTT rank estimates hold:*
$rank_C(\Delta_d) = d$, $rank_{TT}(\Delta_d) = 2$, $rank_{QTT}(\Delta_d) = 4$, $d \geq 2$.
$rank_{QTT}(\Delta_1) = 3$, $rank_{QTT}(\Delta_1^{-1}) \leq 5$.
Given $a \in \mathbb{R}^N$, then for $d = 1$, $rank_{QTT}\left(\nabla^T diag\{a\}\nabla\right) \leq 7\, rank_{QTT}(a)$.
$rank_{TT}(\Delta_d^{-1}) \leq rank_C(\Delta_d^{-1}) \leq C|\log\varepsilon|\log N$ for ε-rank.
$rank_{QTT}(\Delta_d^{-1}) \leq C|\log\varepsilon|^2 \log N$ for ε-rank.

Table 3.3 represents the average QTT-ranks in approximation of certain function related matrices up to fixed tolerance $\varepsilon = 10^{-5}$. Among others, it includes the important example of matrix exponential (cf. Conjecture 3.14). As above, the matrix $diag(f(\|x\|^2))$, $x = (x_1, x_2)$, is a diagonal matrix with diagonal entries obtained by sampling a function $f(\|x\|^2)$ over uniform grid points situated on the line $x_1 = x_2$. One can observe that rank parameters are small, and depend very mildly on the grid size.

$N \setminus \overline{r}$	$e^{-\alpha\Delta_1}, \alpha = 0.1, 1, 10, 10^2$	Δ_1^{-1}	$diag(1/\|x\|^2)$	$diag(e^{-\|x\|^2})$
2^9	6.2/6.8/9.7/11.2	6.2	5.1	4.0
2^{10}	6.3/6.8/9.5/10.8	6.3	5.3	4.0
2^{11}	6.4/6.8/9.0/10.4	6.2	5.5	4.1

Table 3.3: QTT_2-matrix-ranks of $N \times N$-matrices for large $N = 2^p$.

3.5.4 Fast multi-dimensional FFT and convolution in QTT format

d-dimensional FFT over $N^{\otimes d}$ grid can be realized on the rank-k tensor with the linear-logarithmic cost $O(dkN\log_2 N)$, due to the rank-1 factorized representation

$$F_N^{(d)} = (F_N^{(1)} \otimes I_N ... \otimes I_N)(I_N \otimes F_N^{(2)} ... \otimes I_N)...(I_N \otimes I_N ... \otimes F_N^{(d)}) \equiv F_N^{(1)} \otimes ... \otimes F_N^{(d)},$$

where $F_N^{(\ell)} \in \mathbb{R}^{N\times N}$ represents the univariate FFT matrix along mode ℓ.

Remark 3.17 *Let each canonical vector in the target rank-k tensor contain only $p \le N$ essential Fourier frequencies. Then the total cost of Q-FFT amounts to $O(dpk \log N)$ that is of order $O(dk \log N)$ if $p = O(1)$. The claim is justified by observation that each column (row) of 1D FFT matrix has a rank-1 QTT image. Hence, the QTT representation can be derived first for the columns of 1D FFT matrix and then substituted to the factorised tensor-product decomposition of $F_N^{(d)}$. We conclude the argument by noticing that the complex (resp. real) Fourier harmonics have QTT-rank equals to 1 (resp. 2).*

Fourier and convolution transforms of N-d tensors in log-volume complexity, $O(d \log^2 N)$, can be computed by using the QTT approximation as proposed in [20]. Direct convolution transform of N-d tensors in $O(d \log N)$ operations using the explicit QTT representation of multilevel Toeplitz matrices is developed in [46] (see also the discussion in [31]).

4 $O(d \log N)$-solvers for elliptic/parabolic problems

In the present section, we show how the Tucker and QTT approximation method can be applied in the framework of truncated iteration for solving certain elliptic/parabolic equations in higher dimensions, and providing log-complexity scaling in the volume size N^d. The construction of such methods is based on the fact that solutions of some classes of boundary-value/eigenvalue problems are well approximated in the canonical, Tucker and quantics-TT formats.

We follow the concept of approximate (truncated) iteration based on use of the rank-structured tensor representation of matrix-vector operations in the framework of *preconditioned iterative solvers* (cf. [53, 37, 54]).

4.1 Tensor-truncated iteration for linear elliptic systems

Consider the model discrete elliptic problem of stationary type,

$$\mathcal{L}U \equiv (\mathcal{D} + \mathcal{V})U = F, \quad U \in \mathbb{R}^{\mathcal{I}}, \tag{4.1}$$

where $\mathcal{D} \in \mathbb{R}^{\mathcal{I}\times\mathcal{I}}$ represents the elliptic diffusion operator, $-\nabla a(x)\nabla$, defined on tensor-product domain in $\mathbb{R}^d$, and a matrix $\mathcal{V} \in \mathbb{R}^{\mathcal{I}\times\mathcal{I}}$, represents some physically relevant potential. We assume that $a\Delta_d \le \mathcal{D} + \mathcal{V} \le b\Delta_d$, for some $a, b > 0$. Moreover, matrices $\mathcal{D}$ and $\mathcal{V}$ are supposed to have a low-rank Kronecker $\mathcal{S}$-tensor representation.

Example 4.1. The solution of the discrete Poisson equation in $\mathbb{R}^d$,

$$\Delta_d U = F \quad \text{with rank-1 r.h.s.} \quad F = \otimes_{\ell=1}^{d} f_\ell, \quad f_\ell \in \mathbb{R}^N, \tag{4.2}$$

is (approximately) represented in the rank-$(2M+1)$, canonical tensor format,

$$U = \Delta_d^{-1} F \simeq U_M := \sum_{k=-M}^{M} c_k \bigotimes_{\ell=1}^{d} \exp(-t_k \Delta_1)\, f_\ell, \tag{4.3}$$

providing the exponential convergence rate in the rank parameter (see *sinc*-approximation in [25]),

$$\left\| \Delta_d^{-1} F - U_M \right\| \le C e^{-\pi\sqrt{M}} \|F\|.$$

Hence, the low-rank QTT-representation of the univariate vectors f_ℓ implies the existence of the low QTT-rank solution U_M.

Lemma 4.1 *Under the choice $M = C\log^2\varepsilon$, representation (4.3) approximates the exact solution U in (4.1) up to the relative tolerance $\varepsilon > 0$. This representation has the storage complexity $O(d\ \mathrm{rank}_{QTT}(\Delta_1)\log^2\varepsilon\log N)$.*

For the linear system (4.1) the simple truncated preconditioned iteration takes the form

$$U_0 \in \mathcal{S}: \quad \widetilde{U}_{m+1} = U_m - \mathcal{B}(\mathcal{L}U_m - F), \quad U_{m+1} := T_{\mathcal{S}}(\widetilde{U}_{m+1}), \quad m = 0, 1, ...$$

The preconditioner $\mathcal{B} = \mathcal{B}_M$ can be chosen as inverse of the shifted anisotropic Laplacian. The truncated preconditioned iteration as above can be applied (with the respective modification of the matrix $\mathcal{L}$) to the numerical problems arising in FEM/BEM applications, in quantum chemistry (calculation of the Hartree potential), as well as in the implicit time-stepping schemes in financial mathematics (the Black-Schole equation), and in many-particles dynamics (the master and Fokker-Planck equations).

Table 4.1 presents the average QTT_2-ranks of the finite difference solution to the Poisson equation in the unite cube in $\mathbb{R}^d$, up to the tolerance $\varepsilon = 10^{-5}$,

$$\Delta_d U = F \in \mathbb{R}^{N\times...\times N}, \quad \text{considered on large} \quad \underbrace{N \times ... \times N}_{d} \quad \text{uniform grids}$$

with $N = 2^p$, and for $d = 25, 50, 100, 200$. We study the case $F = 1$, that corresponds to the non-separable solution with weak singularities at the edges and corner points (non-analytic solutions), requiring large spacial grids to provide higher resolution. The CPU time (sec.) is presented, that does not count the (problem independent) preprocessing cost required to compute the QTT_2 representation of $1D$ matrix exponentials, $\{\exp(-t_k\Delta_1)\}$, $t_k > 0$, $k = -M, ..., M$, of size $N \times N$ (the latter can be precomputed once and stored).

$N \setminus d$	25		50		100		200	
	r	*time*	r	time	r	time	r	time
2^7	10.0	0.0124	8.0	0.023	7.8	0.047	5.2	0.089
2^8	10.1	0.015	8.2	0.026	6.4	0.05	5.1	0.1
2^9	10.2	0.016	8.2	0.03	6.4	0.06	5.1	0.12
2^{10}	10.2	0.0177	8.4	0.03	6.4	0.061	5.0	0.127

Table 4.1: QTT_2-ranks for solution of the Poisson equation on large $N^{\otimes d}$ grids, $N = 2^p$.

One observes the systematic decay of the average rank parameter $\overline{r}$, in the growing dimension d, with stabilization to the small value about several ones. It is worth to note that the CPU computing time increases only logarithmically in the grid size N and linearly in d, as predicted by the theory: The total numerical cost is estimated by $O(d\log\varepsilon^{-1}\log N)$.

For the loading vector F, corresponding to the R-term sum of trigonometric functions, the QTT-rank is estimated by $r \le R$, again leading to the log-log computational cost $O(dR\log\varepsilon^{-1}\log N)$.

Example 4.2. (*Tensor-truncated solvers in the case of parameter-dependent coefficients*). Consider a class of high dimensional parametric elliptic problems, arising, for example, in stochastic PDEs (parameter-dependent coefficients in $\mathcal{L}$, [88, 64]). The governing equation is formulated as follows: Given an elliptic operator

$$\mathcal{A} := -\mathrm{div}_x\left(a_M(y,x)\mathrm{grad}_x\right) \quad \text{and} \quad f \in L^2(D), \quad D \in \mathbb{R}^d, \quad d = 1, 2, 3,$$

where the coefficient $a_M(y,x)$ is smooth in $x \in D$, $y = (y_1, ..., y_M) \in \Gamma := [-1,1]^M$, $M \le \infty$. Find $u_M \in L^2(\Gamma) \times H_0^1(D)$, such that

$$\begin{aligned} \mathcal{A}u_M(y,x) &= f(x) \quad \text{in } D, \quad \forall y \in \Gamma, \\ u_M(y,x) &= 0 \qquad \text{on } \partial D, \quad \forall y \in \Gamma. \end{aligned}$$

In the simplest case of random field that is linear in the stochastic variable (*additive case*), we have

$$a_M(y,x) := a_0(x) + \sum_{m=1}^{M} a_m(x) y_m, \quad M \to \infty, \tag{4.4}$$

$a_m \in L^\infty(D)$, $m = 0, ..., M$, obtained via the truncated Karhunen-Loéve expansion [88]. In the so-called *log-additive case* the coefficient is given by

$$a_M(y,x) := e^{a_0(x) + \sum\limits_{m=1}^{M} a_m(x) y_m}. \tag{4.5}$$

In discretizations of diffusion problems with random input, the dimension M of the parameter space could become arbitrarily large. Concerning the coefficient function $a_M(y,x)$, we assume that there exists $a_{min} > 0$, such that (see [64]),

(A) $a_{min} \le a_0(x) < \infty$,

(B) $\left| \sum\limits_{m=1}^{M} a_m(x) y_m \right| \le \gamma a_{min}$ with $\gamma < 1$, and for $|y_m| < 1$ $(m = 1, ..., M)$.

In the case $d = 1$, under some technical assumptions, we are able to derive an estimate on the separation rank in variables $y_1, ..., y_M$, in the weakly coupled format,

$$\operatorname{grad}_x u_M(y,x) \approx \sum_{k=-K}^{K} f_k^{(1)}(y_1, x) \cdots f_k^{(M)}(y_M, x), \tag{4.6}$$

that separates terms in coupled variables $z_m = (y_m, x) \in [-1,1] \times \mathbb{R}^d$ $(m = 1, ..., M)$.

Denote by $v = (-\Delta_x^{-1}) f$ the solution of the associated Poisson equation in D, introduce the waiting coefficients $\sigma_m = \|a_m\| / (\sum_{m=1}^{M} \|a_m\|) > 0$, and define the reassembled coefficients $b_m(y_m, x) = \sigma_m a_0(x) + a_m(x) y_m$, $(m = 1, ..., M)$.

Proposition 4.2 *Assume that $d = 1$, $\operatorname{grad}_x u_M(y,x) \in C(D)$ for all $y \in \Gamma$, $\operatorname{grad}_x v(x) \in C(D)$, $b_m(y_m, x) \ge 0$, and let conditions (A), (B) above be valid. Then, in the additive case, there exist positive constants $c_k, t_k \in \mathbb{R}$, such that*

$$\left\| \operatorname{grad}_x u_M(y,x) - \sum_{k=-K}^{K} c_k \prod_{m=1}^{M} e^{-t_k b_m(y_m,x)} (\operatorname{grad}_x v(x) + C_0) \right\|_{L^\infty} \le C e^{-\beta K / \log K},$$

where $C, \beta > 0$ do not depend on M and K. In log-additive case, the above expansion is exact with $K = 0$.

Proof. Our assumptions imply the pointwise equality

$$\operatorname{grad}_x u_M(y,x) = \frac{1}{a_M(y,x)} (\operatorname{grad}_x v(x) + C_0), \quad x \in D, \quad y \in \Gamma,$$

where the reciprocal coefficient $\frac{1}{a_M(y,x)} = (\sum_{m=1}^{M} b_m(y_m,x))^{-1}$ can be separated by the $(2K+1)$-term *sinc*-quadrature, with the exponential convergence in the number of terms (similar to [33, 64]). ■

Proposition 4.2 leads to the constructive rank-$(2K+1)$ approximation of $\mathrm{grad}_x u_M(y,x)$ in the weakly coupled format (2.3), and it opens the way to the design of direct QTT-truncated solver for the class of elliptic problems with parameter-dependent coefficients for $d=1$. In fact, the solution $u_M(y,x)$ can be obtained at any collocation point by N_0-term quadrature including the weakly separable representation of $\mathrm{grad}_x u_M(y,x)$ that leads to the complexity estimate $O(M(2K+1)N_0N)$, where N is the disretization parameter in the variables y_m (linear scaling in M). In the log-additive case (4.5), the previous result remains valid with $K=0$ since the reciprocal coefficient is separable with rank one. This lead to the complexity bound $O(MN_0N)$.

In the case $d \geq 2$ this construction can be used as the rank-structured preconditioner (see [18] on the efficient preconditioning in the canonical format).

Notice that the rank estimate for the Tucker-type approximation of $u_M(y,x)$ in $(M+1)$-dimensional space can be derived based on the analyticity results in [10].

We end up with numerical illustrations. For this application a fully separable low-rank canonical approximations are constructed in $(M+1)$-dimensional tensor space [64], where the d-dimensional physical variable enters the tensor space as a single argument. $\mathcal{S}$-truncated preconditioned iteration applies to solving discrete sPDE in the form

$$\mathcal{L}(J)U(J) = F, \quad J \in \mathbb{R}^{J_1 \times \ldots \times J_M},$$

leaving on the huge index set of size $N^{\otimes(M+1)}$. Both the solution vector $U = U(J)$, and the stochastic-FEM system matrix $\mathcal{L} = \mathcal{L}(J)$ depend on $J \in \mathbb{R}^{J_1 \times \ldots \times J_M}$ discretizing the continuous multi-parameter $y \in \Gamma$.

In the following example, presented in [64], we choose $d=1$, $M=20$, $\mathcal{S} = \mathcal{C}_R$, and $\mathcal{B} := \mathcal{L}(0)^{-1}$. Figure 4.1 illustrates numerics in the case of variable coefficients with exponential decay ($\alpha = 1$, $N = 63$, $R \leq 5$),

$$a_m(x) = 0.5\, e^{-\alpha m} sin(mx), \quad m = 1, 2, ..., M, \quad x \in (0, \pi).$$

Numerical results for sPDEs based on QTT and hierarchical formats are presented in [62, 66].

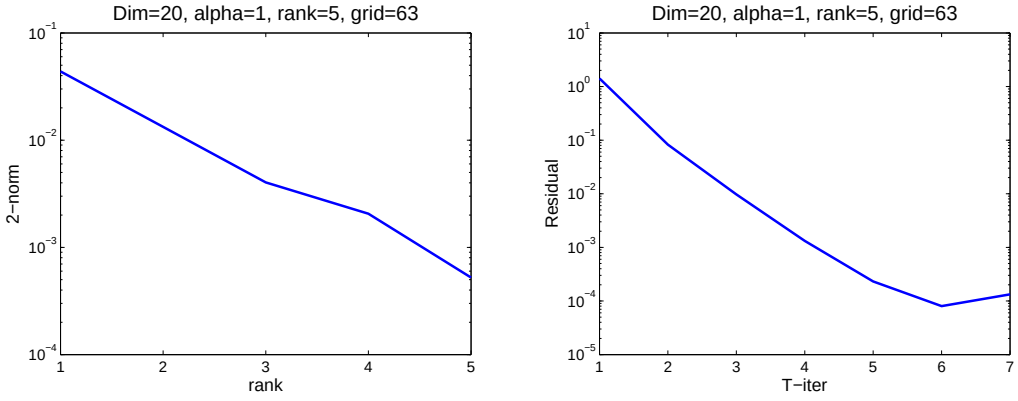

Figure 4.1: Rank approximation (left) and convergence of $\mathcal{S}$-truncated iteration (right).

4.2 Spectral problems

We proceed with the spectral problem as follows,

$$\mathcal{L}U \equiv (\Delta_d + \mathcal{V})U = \lambda U, \quad U \in \mathbb{R}^{\mathcal{I}}. \tag{4.7}$$

The example on existence of low QTT-rank solutions of eigenvalue problems can be observed in the case of d-Laplacian.

Example 4.3. The eigenvectors $U_{\mathbf{i}}$, $\mathbf{i} \in I^{\otimes d}$, of the algebraic eigenvalue problem

$$\Delta_d U = \lambda U, \quad U \in \mathbb{R}^{\mathcal{I}},$$

are exactly in the rank-2 quantics tensor format with the oscillating trigonometric canonical vectors, $U_{\mathbf{i}} = \bigotimes_{\ell=1}^{d} \sin(i_\ell x_\ell)$. Lemma 3.6,(B) allows us to verify this representation in the rank-2, q-quantics format for any $q = 2, 3, ...$

Solving the spectral problem (4.7) in the rank structured format $\mathcal{S}$ can be realized via the tensor-truncated preconditioned inverse iteration as follows,

$$U_0 \in \mathcal{S}: \quad \widetilde{U}_{m+1} = U_m - \mathcal{B}(\mathcal{L}U_m - \mu_m U_m), \quad U_{m+1} := T_{\mathcal{S}}(\widetilde{U}_{m+1}),$$

$$U_{m+1} : U_{m+1}/\|U_{m+1}\|, \quad \mu_{m+1} = (\mathcal{L}U_{m+1}, U_{m+1}), \quad m = 0, 1, ...$$

The preconditioner $\mathcal{B} = \mathcal{B}_M$ can be chosen as inverse of the shifted Laplacian (see [37, 54] for more details).

Alternatively, one can apply the tensor-truncated Green function iteration that takes the form (cf. [53] and references therein)

$$\widetilde{U}_{m+1} = (\Delta_d - E_m I)^{-1}\mathcal{V}U_m, \quad U_{m+1} := T_{\mathcal{S}}(\widetilde{U}_{m+1}), \quad U_{m+1} := \frac{U_{m+1}}{||U_{m+1}||},$$

and E_{m+1} is updated at each step as a Rayleigh quotient, $E_{m+1} = \langle \mathcal{L}U_{m+1}, U_{m+1}\rangle$.

Example 4.4. In the case of Schrödinger equation for hydrogen atom,

$$(-\frac{1}{2}\Delta - \frac{1}{\|x\|})u = \lambda u, \quad x \in \mathbb{R}^3, \quad u \in H^1(\mathbb{R}^3), \tag{4.8}$$

the physically relevant eigenpair with minimal eigenvalue is given by $u_1(x) = e^{-\|x\|}$, $\lambda_1 = -0.5$, where both $e^{-\|x\|}$ and $\frac{1}{\|x\|}$, can be proven to provide accurate approximation in the low rank binary folding format, due to Lemma 3.8, and applying the *sinc*-quadrature approximation in [49, 53, 4] (see Tables 3.2 and 4.3 for numerical examples). The particular numerical illustrations on the truncated Green function iteration for the equation (4.8) are presented in Table 4.2, see [61] for more details. We use $N \times N \times N$-grid, and demonstrate the $O(\log N)$ CPU-time scaling.

N	Time for 1 iter.	Iter.	Eigenvalue error
2^7	8.5	8	6.1e-03
2^8	13	8	1.5e-03
2^9	18	8	4.0e-04
2^{10}	25	8	1.0e-04

Table 4.2: Schrödinger equation for hydrogen atom, $M = 20$.

The tensor truncated iteration can be applied, with the respective modification to the nonlinear case, $\mathcal{L} = \mathcal{L}(U)$, arising in the spectral problems in electronic structure calculations (the Hartree-Fock and Kohn-Sham equations).

Example 4.5. The Hartree-Fock (HF) equation for determination of the ground state of a molecular system consisting of M nuclei and N electrons is given by the following nonlinear eigenvalue problem in $L^2(\mathbb{R}^3)$,

$$(\mathcal{F}_\Phi \phi_i)(x) = \lambda_i \, \phi_i(x), \quad \int_{\mathbb{R}^3} \phi_i(x)\phi_j(x)\,\mathrm{d}x = \delta_{ij}, \; i,j = 1,...,N \; , \tag{4.9}$$

with $\mathcal{F}_\Phi$ being the non-linear Fock operator

$$\mathcal{F}_\Phi(\cdot) := -\frac{1}{2}\Delta(\cdot) - \sum_{\nu=1}^{M} \frac{Z_\nu}{\|x - A_\nu\|}(\cdot) + V_H(x)(\cdot) + V_E(\cdot) \; ,$$

where the Hartree potential is defined by $V_H(x) := \int_{\mathbb{R}^3} \frac{\tau(y,y)}{\|x-y\|}\,\mathrm{d}\mathbf{y}$, and the nonlocal exchange operator is given as $V_E\phi := -\frac{1}{2}\int_{\mathbb{R}^3} \frac{\tau(x,y)}{\|x-y\|}\phi(y)\,\mathrm{d}y$. Here, $1/\|\cdot\| : \mathbb{R}^3 \to \mathbb{R}$ corresponds to the Newton potential, and $Z_\nu \in \mathbb{R}_+$, $A_\nu \in \mathbb{R}^3$ $(\nu = 1,...,M)$ specify charges and positions of M nuclei. The electron density matrix $\tau : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, is given by $\tau(x,y) = 2\sum_{i=1}^{N} \phi_i(x)\phi_i^*(y)$.

Note that both exchange and the Hartree potentials include the 3D convolution transform with the Newton convolving kernel, that has to be computed at each step of the iterations on nonlinearity. Hence these terms represent the most complicated part in the numerical treatment of the Hartree-Fock equation. Efficient tensor-structured methods to compute these operators are presented in [52, 57, 58].

In this case, instead of truncated preconditioned inverse iteration, one can employ the $\mathcal{S}$-truncated Krylov subspace-type iteration (the so-called DIIS iteration [9, 82]) applied to the Galerkin system in the problem adapted basis, (see [58]), leading to the numerical cost $O(N \log N)$. Figure 4.2 (left) shows convergence of the SCF iteration for all electron case of H_2O. This challenging problem is solved efficiently by the multigrid-accelerated tensor method [57, 58] that made possible calculations on large 3D Cartesian grids up to the volume size $N = 8192^3$. Figure 4.2 (right) shows convergence of the HF energy for the corresponding grid levels.

Numerical illustrations on the QTT approximation of functions and operators arising in the solution of Hartree-Fock equation are given in Table 4.3. It illustrates the low QTT-rank approximability of the Newton potential $1/\|x\|$, in $\mathbb{R}^3$, the electron density $\rho(x) = \tau(x,x)$, $x \in \mathbb{R}^3$, and the Hartree potential, $V_H := 1/\|x\| * \rho$, of CH_4 molecule discretized over large $N \times N \times N$ spatial grid, and computed in [57] by the multigrid Tucker-canonical decomposition. In all cases the approximation accuracy $\varepsilon = 10^{-6}$ in the QTT-format is achieved. In Table 4.3, Hartree(S) and Hartree(F) correspond to the middle $N \times N$ matrix slice and to the full $N \times N \times N$-array representing the discrete Hartree potential, respectively.

N	128	256	512	1024
$1/\|x\|$	13.8	16.0	17.5	18.0
$\rho(x)$	32.0	40.0	45.8	48.6
Hartree(S)	13.7	14.2	14.2	13.9
Hartree(F)	32.1	34.9	20.2	28.2

Table 4.3: QTT_2-ranks of $1/\|x\|$, the Hartree potential, and electron density of CH_4.

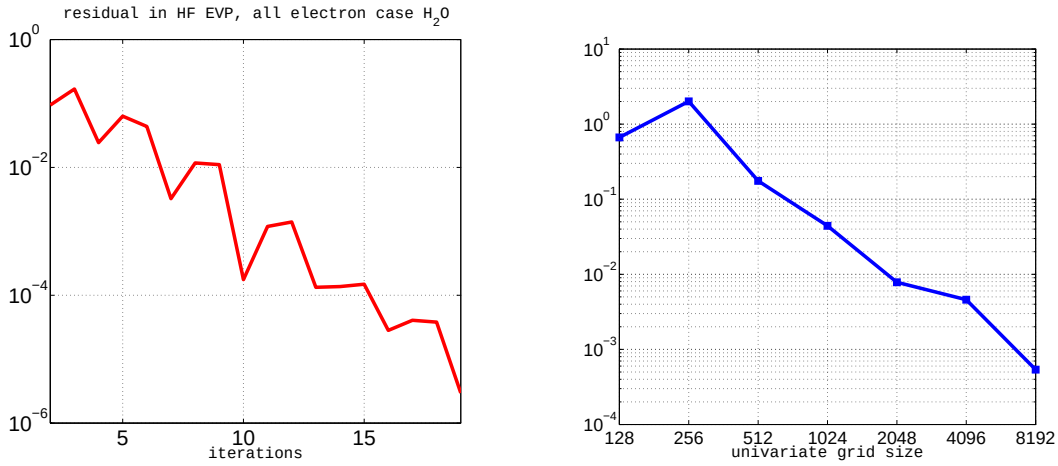

Figure 4.2: Multilevel convergence of the tensor-truncated SCF iteration applied to the all electron case of H_2O (left), and the energy error vs. N (right).

We conclude that QTT-ranks of the Newton and Hartree potentials are rather small and increase merely logarithmically in N, ensuring $O(\log N)$ scaling. In turn, the QTT-rank of the electron density ρ, has the tendency to approach the canonical rank of the respective density tensor.

This numerics indicates the way toward solving the nonlinear Hartree-Fock equation in quantics tensor format with the almost grid-independent computational cost $O(\log \varepsilon^{-1} \log N)$, that may provide the tool to efficient *ab initio* numerical grid-based simulation of large molecules.

4.3 Parabolic equations

Consider the transient semi-discrete equation of parabolic type,

$$U(0) \in \mathbb{R}^{\mathcal{I}}: \quad \frac{\partial U}{\partial t} + \mathcal{L}U = 0, \quad U(t) \in \mathbb{R}^{\mathcal{I}},\ t > 0, \tag{4.10}$$

with $\mathcal{L} = \Delta_d + \mathcal{V}(t)$, where the matrix $\mathcal{V} \in \mathbb{R}^{\mathcal{I}\times\mathcal{I}}$ may represent certain interaction potential. Commonly used solvers are based on spectral decomposition of $\mathcal{L}$ and then treating dynamics explicitly. We discuss some alternative approaches.

Example 4.6 (*Iterative solvers with time-stepping*). The simple truncated implicit integrator (say, the Euler scheme) takes the form (the scheme of better choice, say Crank-Nicolsol scheme, can be used),

$$U_0 = T_{\mathcal{S}}(U(0)): \quad \widetilde{U}_{m+1} = (\mathcal{I} + \tau_m \mathcal{L})^{-1} U_m, \quad U_{m+1} := T_{\mathcal{S}}(\widetilde{U}_{m+1}), \quad m = 0, 1, ... \tag{4.11}$$

with time stepping parameter $\tau_m > 0$, where the action of inverse matrix can be implemented as the solution of the stationary equation (4.1) with $F = U_m$, and with singularly perturbed elliptic operator $\mathcal{L}_m = \tau_m^{-1}\mathcal{I} + \mathcal{L}$, and with the initial guess taken from the previous time step. At this point also the ALS and DMRG type preconditioned iteration can be applied [19].

In the case of unbounded domains, the low tensor rank preconditioner can be chosen as inverse of the translated Laplacian with positive shift (separable approximation to the Yukawa type Green kernel), that is proven to have the low canonical rank approximation [53, 54]. In the case of finite

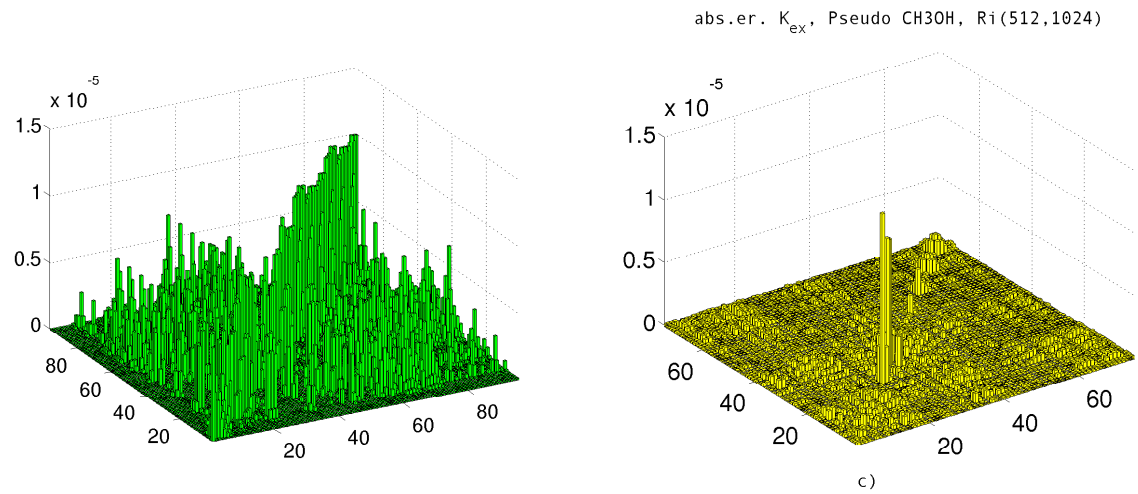

Figure 4.3: Absolute error in the Coulomb matrix of C_2H_6, 8192^3-grid (left); Absolute error of the exchange matrix for the *pseudopotential* case of CH_3OH on the $n \times n \times n$ grid, $n = 1024$.

computational domain, one can apply the rank-$(2M+1)$ preconditioner like (3.17) in the form,

$$\mathcal{B}_M := \sum_{k=-M}^{M} c_k \bigotimes_{\ell=1}^{d} \exp(-t_k(\frac{1}{d\tau_m} I_n + \Delta_\ell)), \tag{4.12}$$

providing exponentially convergent in M approximation of $\mathcal{L}_m$.

The *tensor truncated implicit integrator* described above can be applicable to the transient Hartree-Fock model (cf. [2, 68]), the heat transfer equation, to the time-dependent Stocks and Navier-Stockes equations in fluid dynamics, to the multidimensional Black-Schole equation of financial mathematics, as well as to the deterministic/stochastic equations of chemical kinetics (the chemical master and Fokker-Planck equations).

Example 4.7 *(Direct tensor approximation of the solution operator).* In the case of heat equation the formal representation of the matrix exponential family

$$U(t) = e^{-\mathcal{L}t} U_0 \approx T_{\mathcal{S}}(e^{-\mathcal{L}t}) U_0, \quad t \geq 0,$$

provides mean for application of the tensor-structured (say in QTT format) matrix exponential to each fixed $t > 0$, described in Section 3.5.2 (scaling-and-squaring method with tensor truncation). This approach allows a considerable coarsening in the time stepping, reducing the number of grid points in the time domain to $O(\log T)$, to compute the solution at $t = T$. In the case of moderate T the time stepping can be avoided completely.

Example 4.8 (*Molecular dynamics without time stepping*). Important example in molecular dynamics is given by the Schrödinger equation for the motion of d nuclei obtained from the Born-Oppenheimer approximation (see [68] for more detail),

$$i\frac{\partial \psi}{\partial t} = \mathcal{H}\psi, \quad \mathcal{H} = T + V, \quad \psi(0) = \psi_0, \tag{4.13}$$

with kinetic energy $T = -\sum_{n=1}^{d} \frac{\hbar^2}{2M_n} \Delta_{x_n}$ and a potential $V = V(x_1, ..., x_d)$, $x_\ell \in \mathbb{R}^3$ $(\ell = 1, ..., d)$ being an approximation to an electronic potential energy surface $E(x_1, ..., x_d)$.

The explicit solution operator $e^{-i\mathcal{H}t}$ could not be approximated by QTT-matrix exponential with uniform bound on the TT-ranks.

To get rid of this problem we introduce the operator regularization scheme $\Sigma_p = \mathcal{H}^{-p} e^{-i\mathcal{H}t}$, $U_p = \mathcal{H}^p U(x,0)$, $p \geq 1$, leading to stable QTT approximation,

$$U(x,t) = e^{-i\mathcal{H}t} U(x,0) \approx (T_{\mathcal{S}} \Sigma_p) T_{\mathcal{S}} U_p, \quad t \geq 0.$$

For the numerical example, we consider 1D quantum harmonic oscillator, $V(x) = \frac{1}{2}\|x\|^2$, propagating the wave packet (exact eigenfunction) by $U_n(x)e^{-i(n+1/2)t}$ on level $n = 0$, with $d = 1$, $T = 1.0$, $\varepsilon = 10^{-6}$, $p = 2$, and with spatial grid-size N. Denote the average rank by $rank_{QTT}$. The next table indicates that QTT-ranks of both initial wave packet and the resultant solutions are about several ones.

N	$rank(\mathcal{H}^{-p}cos(\mathcal{H}T))$	$rank(\mathcal{H}^{-p}cos(\mathcal{H}t)U_p)$	$rank(U_p)$
2^8	33.8	3.7	4.7
2^9	33.2	3.7	4.7
2^{10}	32.5	3.6	4.9

The computational cost is bounded by $= O(\log T \log N)$.

Results on spectral calculations for high-dimensional Hamiltonians in (4.13) by using the implicit integrators via GMRES and DMRG solvers will be presented in forthcoming paper.

5 Conclusions

In the present survey, we discuss the prospects of tensor-structured data formats as the basic numerical tools for applications in high-dimensional scientific computing.

We address the traditional canonical, Tucker and mixed models as well as the matrix product states formats in the form of tensor train expansion. Then we present the modern $O(d \log N)$-complexity quantics-TT (QTT) approximation method applied to high-order N-d tensors. In our applications these tensors arise as the grid representation of physically relevant functions and operators in $\mathbb{R}^d$. The idea of QTT method is based on the folding of initial N-d tensor to the auxiliary higher dimensional space of q-D tensors with $D = d \log_q N$ and $q = 2, 3$. The rigorous analysis on the low-rank quantics-TT tensor approximation to the class of function related N-d arrays, indicates the log-log complexity scaling, $O(d \log \varepsilon^{-1} \log N)$, in both the grid size and numerical precision $\varepsilon > 0$. The nonlinear QTT tensor approximation can be implemented on the base of stable, QR/SVD algebraic decompositions with controlled accuracy $\varepsilon > 0$.

It is described how the mixed Tucker-canonical and QTT approximation methods can be applied in the framework of truncated iteration for solving certain classes of elliptic/parabolic equations in higher dimensions with log-scaling in the basic discretisation parameters. In particular, the QTT method can be applied to electronic structure calculations, molecular dynamics, stochastic PDEs and in the traditional FEM/BEM modeling in $\mathbb{R}^d$.

Various numerical tests illustrate the high compression rate provided by the quantics-TT method applied to multidimensional data arrays arising in the traditional FEM calculations, in numerical quantum chemistry and in stochastic PDEs.

We hope that the mixed Tucker-canonical model combined with the QTT numerical methods will open the new prospects for developing reliable and robust computational schemes in higher dimensions that are free from the "curse of dimensionality" and noticeable limitations on the univariate grid-size.

Acknowledgments. I gratefully acknowledge Wolfgang Hackbusch (MPI MiS, Leipzig), Eugene Tyrtyshnikov (INM RAS, Moscow), Ivan Gavrilyuk (BA Eisenach), Venera Khoromskaia (MPI MiS, Leipzig), Heinz-Jurgen Flad (TU Berlin), Christoph Schwab (ETH Zuerich), Reinhold

Schneider (TU Berlin), Stefan Sauter (Uni. Zuerich), and Ivan Oseledets (INM RAS, Moscow) for the productive collaboration on tensor numerical methods. I am also appreciative to Christian Lubich (Uni. Tuebingen), Lars Grasedyck (TU Aachen), Lieven De Lathauwer (Uni. Leuven), Thomas Huckle (TU Muenchen), Dmitry Savostianov (INM RAS, Moscow), Sergey Dolgov (MPI MiS, Leipzig), Vladimir Kazeev (INM RAS, Moscow), Christine Tobler (ETH Zuerich), and Nikos D. Sidiropoulos (TU of Crete) for valuable discussions and cooperation on different topics addressed in this survey.

References

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds.* Princeton University Press, Princeton, NJ, 2008.

[2] M.H. Beck, A. Jäckle, G.A. Worth, and H.-D. Meyer, *The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets.* Phys. Reports 324 (2000), 1-105.

[3] M. Bebebndorf, *Adaptive Cross Approximation of Multivariate Functions.* J. Constructive Approximation, to appear (2010).

[4] C. Bertoglio, and B.N. Khoromskij, *Low rank tensor-product approximation of projected Green kernels via sinc-quadratures.* Preprint 79/2008, MPI MiS 2008 (submitted).

[5] G. Beylkin and M.J. Mohlenkamp, *Numerical operator calculus in higher dimensions.* Proc. Natl. Acad. Sci. USA **99** (2002) 10246–10251.

[6] G. Beylkin, J. Garcke, and M.J. Mohlenkamp, *Multivariate regression and machine learnig with sums of separable functions.* SIAM. J. of Sci. Comp., **31**(3): 1840-1857(2009).

[7] G. Beylkin and M.J. Mohlenkamp, *Algorithms for numerical analysis in high dimension.* SIAM J. Sci. Comput. **26** (6) (2005) 2133-2159.

[8] H.J. Bungartz, and M. Griebel, *Sparse grids.* Acta Numerica, 1-123, 2004.

[9] E. Cancés and Le Bris, *On the convergence of SCF algorithms for the Hartree-Fock equations.* ESAIM: M2AN, vol. 34/4, 2000, 749-774.

[10] A. Cohen, R. DeVore, and Ch. Schwab: *Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs*, Report 2009-02, Seminar for Applied Mathematics, ETH Zürich, CH-8092 Zürich.

[11] P. Comon, and B. Mourrain, *Decomposition of quantics in sums of powers of linear forms.* Signal Processing, **53** (2), 1996, 96-107.

[12] W. Dahmen, S. Proessdorf, and R. Schneider, *Multiscale methods for pseudo-differential equations on manifolds.* In: Wavelet analysis and its applications, 5, Academic Press, 1995.

[13] L. De Lathauwer, *Decomposition of higher-order tensor in blockterms — Part I: Lemmas for partitioned matrices.* SIAM J. Matr. Anal. Appl., vol. 30, no. 3, 2008, 1022-1032.

[14] L. De Lathauwer, B. De Moor, J. Vandewalle, *On the best rank-*1 *and rank-*$(R_1, ..., R_N)$ *approximation of higher-order tensors.* SIAM J. Matrix Anal. Appl., 21 (2000) 1324-1342.

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

[16] *L. de Lathauwer*, A survey of tensor methods, in IEEE International Symposium on Circuits and Systems, May 2009, pp. 2773–2776. doi10.1109/iscas.2009.5118377.

[17] V. De Silva and L.-H. Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem.* SIAM J. Matrix Anal. Appl., 30 (2008), no. 3, 1084-1127.

[18] S.V. Dolgov, B.N. Khoromskij, I. Oseledets, and E.E. Tyrtyshnikov. *Tensor Structured Solution of Elliptic Problems with Jumping Coefficients in* $\mathbb{R}^d$. Preprint 55/2010, MPI MiS, Leipzig 2010 (submitted).

[19] S.V. Dolgov, B.N. Khoromskij, I. Oseledets, *On fast solution of multi-dimensional parabolic equations in the TT/QTT format.* MPI Leipzig 2011. In preparation.

[20] S.V. Dolgov, B.N. Khoromskij, and D. Savostianov. *Fourier transform of d-tensors in* log-*volume complexity using QTT approximation.* Preprint 18/2011, MPI MiS, Leipzig 2011 (submitted).

[21] A. Edelman, T.A. Arias, and S.T. Steven, *The geometry of algorithms with orthogonality constraints.* SIAM J. Matr. Anal. Appl., 20 (1999), no. 2, 303-353.

[22] H.-J. Flad, W. Hackbusch, B.N. Khoromskij, and R. Schneider, *Concept of data-sparse tensor-product approximation in many-particle modeling.* In: "Matrix Methods: Theory, Algorithms, Applications", E. Tyrtyshnikov, et.al. eds., World Scientific Publishing, Singapoure, 2010 (313-347).

[23] H.-J. Flad, B.N. Khoromskij, D.V. Savostyanov, and E.E. Tyrtyshnikov, *Verification of the cross 3d algorithm on quantum chemistry data.* J. Numer. Anal. and Math. Modelling, 4 (2008), 1-16.

[24] I.P. Gavrilyuk, W. Hackbusch, and B.N. Khoromskij, *Data-sparse approximation to a class of operator-valued functions.* Math. Comp. **74** (2005), 681–708.

[25] 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.

[26] L. Grasedyck, *Hierarchical Singular Value Decomposition of Tensors.* Preprint 27/2009, MPI MIS Leipzig, 2009.

[27] L. Grasedyck, *Polynomial Approximation in Hierarchical Tucker Format by Vector Tensorization.* Preprint 43, DFG/SPP1324, RWTH Aachen, 2010.

[28] M. Griebel and J. Hamaekers, *Sparse grids for the Schroedinger equation.* M2AN, 41 (2007), pp. 215-247.

[29] M. Griebel, D. Oeltz, and P. Vassilevski, *Space-time approximation with sparse grids.* SIAM J. Sci. Comput., 28(2):701-727, 2005.

[30] M. Griebel, and D. Oeltz, *A Sparse Grid Space-Time Discretization Scheme for Parabolic Problems.* Computing, 81(1):1-34, 2007.

[31] W. Hackbusch, *Tensorization of vectors and their efficient convolution.* DFG-SPP1324 Preprint 65, University Marburg, 2011.

[32] W. Hackbusch and B. N. Khoromskij, *Tensor-product Approximation to Operators and Functions in High dimension.* Journal of Complexity **23** (2007), 697-714.

[33] 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.

[34] W. Hackbusch, B.N. Khoromskij and S. Sauter. *Adaptive Galerkin Boundary Element Methods with Panel Clustering.* Numer. Math. **105** (2007), 603-631.

[35] W. Hackbusch, B.N. Khoromskij and E.E. Tyrtyshnikov, *Hierarchical Kronecker tensor-product approximations.* J. Numer. Math. **13** (2005), 119–156.

[36] W. Hackbusch, B.N. Khoromskij and E. Tyrtyshnikov, *Approximate iteration for structured matrices.* Numer. Math., **109** (2008), 365-383.

[37] W. Hackbusch, B.N. Khoromskij, S. Sauter, and E. Tyrtyshnikov, *Use of Tensor Formats in Elliptic Eigenvalue Problems.* Preprint 78/2008, MPI MiS Leipzig 2008 (Numer. Lin Alg. Appl., to appear).

[38] W. Hackbusch, and S. Kühn, *A new scheme for the tensor representation.* J. of Fourier Anal. and Appl., 15:706-722, 2009.

[39] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, and G. Beylkin, *Multiresolution quantum chemistry: Basic theory and initial applications.* J. of Chemical Physics, 121 (23): 11587-11598, 2004.

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

[41] S. Holtz, Th. Rohwedder, and R. Schneider, *On manifold of tensors of fixed TT-rank.* Tehn. Rep. 61, TU Berlin, 2010.

[42] M. Ishteva, L. De Lathauwer, P.-A. Absil, and S. Van Huffel, *Differential-geometric Newton algorithm for the best rank-(r_1, r_2, r_3) approximation of tensors.* Numerical Algorithms, **51**(2):179-194, 2009.

[43] A. Jaravine, I. Ibraghimov, and V. Orekhov, *Removal of time barrier for high-resolution multidimensional NMR spectroscopy.* Nature Methods, **3**(8), 2006, 605-607.

[44] T. Jiang, and N.D. Sidiropoulos, *Almost-Sure Identifiability of Multidimensional Harmonic Retrieval.* IEEE Transactions on Signal Proceedings, **49** (9), 2001, 1849-1859.

[45] V. Kazeev, and B.N. Khoromskij. *Explicit low-rank QTT representation of Laplace operator and its inverse.* Preprint 75/2010, MPI MiS, Leipzig 2010 (SIMAX, submitted).

[46] V. Kazeev, B.N. Khoromskij, and E.E. Tyrtyshnikov. *Tensor structure of multilevel Toeplitz matrices leads to convolution with logarithmic complexity.* Preprint 2011, MPI MiS, Leipzig 2011.

[47] V. Khoromskaia, *Computation of the Hartree-Fock Exchange in the Tensor-structured Format.* Comput. Methods in Applied Math., Vol. 10(2010), No 2, pp.204-218.

[48] V. Khoromskaia, *Numerical Solution of the Hartree-Fock Equation by Multilevel Tensor-structured Methods.* PhD thesis, TU Berlin, 2010.
http://opus.kobv.de/tuberlin/volltexte/2011/2948/.

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

[50] B.N. Khoromskij and J.M. Melenk. *Boundary Concentrated Finite Element Methods. SIAM J. Numer. Anal.*, **41**, No. 1 (2003), 1–36.

[51] 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.

[52] B.N. Khoromskij. *Fast and Accurate Tensor Approximation of a Multivariate Convolution with Linear Scaling in Dimension.* J. of Comp. Appl. Math., **234** (2010) 3122-3139.

[53] B. N. Khoromskij, *On tensor approximation of Green iterations for Kohn-Sham equations.* Comp. and Visualization in Sci., **11** (2008) 259-271.

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

[55] B.N. Khoromskij, *$O(d \log N)$-Quantics Approximation of N-d Tensors in High-Dimensional Numerical Modeling.* J. Constr. Approx. 2011, DOI: 10.1007/s00365-011-9131-1. Preprint 55/2009 MPI MiS, Leipzig 2009.

[56] B.N. Khoromskij, *Introduction to Tensor Numerical Methods in Scientific Computing.* Lecture Notes, Preprint 06-2011, University of Zuerich, Institute of Mathematics, 2011, pp 1 - 238. http://www.math.uzh.ch/fileadmin/math/preprints/06_11.pdf.

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

[58] B.N. Khoromskij, V. Khoromskaia, and H.-J. Flad. *Numerical solution of the Hartree-Fock equation in multilevel tensor-structured format.* SIAM J. Sci. Comp., **33**(1), 2011, 45-65.

[59] B. N. Khoromskij, V. Khoromskaia, S. R. Chinnamsetty, H.-J. Flad, *Tensor Decomposition in Electronic Structure Calculations on 3D Cartesian Grids.* J. of Comput. Phys. **228** (2009) 5749-5762.

[60] B. N. Khoromskij and V. Khoromskaia, *Low Rank Tucker Tensor Approximation to the Classical Potentials.* Central European J. of Math., **5**(3) 2007, 1-28.

[61] B.N. Khoromskij, and I. Oseledets. *Quantics-TT approximation of elliptic solution operators in higher dimensions.* Preprint MPI MiS 79/2009, Leipzig 2009. To appear in J. of Numerical Math., 2011.

[62] B.N. Khoromskij, and I. Oseledets. *Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDEs.* Comp. Meth. in Applied Math., 10(4):34-365, 2010.

[63] B.N. Khoromskij, and I. Oseledets. *DMRG+QTT approach to high-dimensional quantum molecular dynamics.* Preprint 68/2010, MPI MiS, Leipzig 2010 (Numer. Math., submitted).

[64] B.N. Khoromskij, and Ch. Schwab, *Tensor-Structured Galerkin Approximation of Parametric and Stochastic Elliptic PDEs.* SIAM J. Sci. Comp., **33**(1), 2011, 1-25.

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

[66] D. Kressner, and C. Tobler, *Iterative solution of parameter-dependent linear systems.* Preprint SAM, ETH Zuerich, 2010.

[67] C. Le Bris, *Computational chemistry from the perspective of numerical analysis.* Acta Numerica (2005), 363 - 444.

[68] Ch. Lubich, *From quantum to classical molecular dynamics: reduced models and numerical analysis.* Zurich Lectures in Advanced Mathematics, EMS, 2008.

[69] V. Mazyja, G. Schmidt, *Approximate Approximations,* Math. Surveys and Monographs, vol. 141, AMS 2007.

[70] H.-D. Meyer, F. Gatti, and G.A. Worth, *Multidimensional Quantum Dynamics: MCTDH Theory and Applications.* Willey-VCH, Wienheim, 2009.

[71] V. Olshevsky, I. Oseledets, E. Tyrtyshnikov, *Tensor properties of multilevel Toeplitz and related matrices*, *Linear Algebra Appl.* 412 (2006), 1–21.

[72] I.V. Oseledets, *Compact matrix form of the d-dimensional tensor decomposition.* Preprint 2009-01, INM RAS, Moscow 2009.

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

[74] I.V. Oseledets, *Constructive representation of functions in tensor formats.* Preprint 2010-04, INM, Moscow, 2010.

[75] I.V. Oseledets, *DMRG approach to fast linear algebra in the TT-format.* Preprint 2011, INM Moscow.

[76] I.V. Oseledets, D.V. Savostyanov, and E.E. Tyrtyshnikov, *Tucker dimensionality reduction of three-dimensional arrays in linear time.* SIMAX, 30(3), 939-956 (2008).

[77] I.Oseledets, D. Savostyanov, E.Tyrtyshnikov, *Linear algebra for tensor problems*, Computing, 85 (2009), 169–188.

[78] I.Oseledets, E.Tyrtyshnikov, *Recursive decomposition of multidimensional tensors*, Doklady Mathematics, vol. 80, no. 1 (2009), pp. 460–462.

[79] I.V. Oseledets, and E.E. Tyrtyshnikov, *Breaking the Curse of Dimensionality, or How to Use SVD in Many Dimensions.* SIAM J. Sci. Comp., v. 31(5) (2009), 3744-3759.

[80] I.V. Oseledets, and E.E. Tyrtyshnikov, *TT-Cross Approximation for Multidimensional arrays.* Lin. Alg. Appl., 432 (2010), 70–88.

[81] J-M. Papy, L. De Lathauwer, and S. Van Huffel, *Exponential Data Fitting Using Multilinear Algebra: the Decimative Case.* Chemometrics, (www.interscience.willey.com) DOI:10.1002/cem.1212.

[82] P. Pulay, *Improved SCF convergence acceleration.* J. Comput. Chem. **3**, 556-560 (1982).

[83] M. Reed, and B. Simon, *Methods of Modern Mathematical Physics, I. Functional Analysis*, AP, NY, 1972.

[84] Th. Rohwedder, S. Holtz, and R. Schneider, *The alternation least square scheme for tensor optimisation in the TT-format.* Preprint DGF-Schwerpunktprogramm 1234 71, 2010.

[85] B. Savas, and L.-H. Lim, *Quisi-Newton methods on Grassmanians and multilinear approximation of tensors.* SIAM J. Sci. Comp., 2010, to appear.

[86] D.Savostyanov, E.Tyrtyshnikov, *Approximate multiplication of tensor matrices based on the individual filtering of factors*, Comput. Math. and Math. Physics, vol.49, no.10 (2009), pp. 1662–1677.

[87] R. Schneider, Th. Rohwedder, J. Blauert, A. Neelov, *Direct minimization for calculating invariant subspaces in density functional computations of the electronic structure*, J. of Comp. Math., 27 (2009), no. 2-3, 360-387.

[88] C. Schwab and R. Todor, *Sparse finite elements for elliptic problems with stochastic loading.* Numer. Math., 95(4), 707-713, 2003.

[89] N.D. Sidiropoulos, *Generalized Carathéodory's Uniqueness of Harmonic Parametrization to N Dimensions.* IEEE Trans. Inform. Theory, **47** (2001) 1687-1690.

[90] A. Smilde, R. Bro and P.Geladi. *Multiway analysis.* Wiley, 2004.

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

[92] V.N. Temlyakov, *Nonlinear methods of approximation.* Foundation of Comp. Math., Springer, 2008.

[93] E.E. Tyrtyshnikov: *Kronecker-product approximations for some function-related matrices.* Linear Algebra Appl. **379** (2004), 423–437.

[94] E.E. Tyrtyshnikov, Tensor approximations of matrices generated by asymptotically smooth functions, *Sbornik: Mathematics* **194**, No. 5-6 (2003), 941–954.

[95] R.-A. Todor, and Ch. Schwab: *Convergence rate for sparse approximation of elliptic problems with stochastic coefficients.* IMA J. of Numer. Anal. (2007) **27**, 232-261.

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

[97] F. Verstraete, D. Porras, and J.I. Cirac, *DMRG and periodic boundary conditions: A quantum information perspective.* Phys. Rev. Lett., 93(22): 227205, Nov. 2004.

[98] G. Vidal, *Efficient classical simulation of slightly entangled quantum computations.* Phys. Rev. Lett. 91(14), 2003, 147902-1 147902-4.

[99] H. Wang, and M. Thoss, *Multilayer formulation of the multiconfiguration time-dependent Hartree theory.* J. Chem. Phys. 119 (2003), 1289-1299.

[100] S.R. White, *Density-matrix algorithms for quantum renormalization groups.* Phys. Rev. B, v. 48(14), 1993, 10345-10356.

[101] H. Yserentant, *The hyperbolic cross space approximation of electronic wavefunctions.* Numer. Math., 105 (2007) 659-690.

[102] N.Zamarashkin, I.Oseledets, E.Tyrtyshnikov, *The tensor structure of the inverse of a banded Toeplitz matrix*, Doklady Mathematics, vol. 80, no. 2 (2009), pp. 669–670.

[103] T. Zhang, and G.H. Golub, *Rank-one approximation to high order tensors.* SIAM J. Matrix Anal. Appl. **23** (2001), 534-550.