ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

MATHICSE
Mathematics Institute of Computational Science and Engineering
School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report
Nr. 08.2013
February 2013

A literature survey of low-rank tensor approximation techniques

L. Grasedyck, D. Kressner, C. Tobler

http://mathicse.epfl.ch

Phone: +41 21 69 37648

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

Fax: +41 21 69 32545

A literature survey of low-rank tensor approximation techniques

Lars Grasedyck[1,*], **Daniel Kressner**[2,**], and **Christine Tobler**[2,***]

[1] Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, 52056 Aachen, Germany

[2] ANCHP, MATHICSE, EPF Lausanne, Switzerland

Key words tensor, low rank, multivariate functions, linear systems, eigenvalue problems

MSC (2010) 15A69, 65F10, 65F15

During the last years, low-rank tensor approximation has been established as a new tool in scientific computing to address large-scale linear and multilinear algebra problems, which would be intractable by classical techniques. This survey attempts to give a literature overview of current developments in this area, with an emphasis on function-related tensors.

1 Introduction

This survey is concerned with tensors in the sense of multidimensional arrays. A general tensor of *order* d and size $n_1 \times n_2 \times \cdots \times n_d$ for integers $n_1, n_2, \ldots, n_d$ will be denoted by

$$\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}.$$

An entry of $\mathcal{X}$ is denoted by $\mathcal{X}_{i_1,\ldots,i_d}$ where each index $i_\mu \in \{1, \ldots, n_\mu\}$ refers to the μth mode of the tensor for $\mu = 1, \ldots, d$. For simplicity, we will assume that $\mathcal{X}$ has real entries, but it is of course possible to define complex tensors or, more generally, tensors over arbitrary fields.

A wide variety of applications lead to problems where the data or the desired solution can be represented by a tensor. In this survey, we will focus on tensors that are induced by the discretization of a multivariate function; we refer to the survey [168] and to the books [174, 239] for the treatment of tensors containing observed data. The simplest way a given multivariate function $f(x_1, x_2, \ldots, x_d)$ on a tensor product domain $\Omega = [0, 1]^d$ leads to a tensor is by sampling f on a tensor grid. In this case, each entry of the tensor contains the function value at the corresponding position in the grid. The function f itself may, for example, represent the solution to a high-dimensional partial differential equation (PDE).

As the order d increases, the number of entries in $\mathcal{X}$ increases exponentially for constant $n = n_1 = \cdots = n_d$. This so called *curse of dimensionality* prevents the explicit storage of the entries except for very small values of d. Even for $n = 2$, storing a tensor of order $d = 50$ would require 9 petabyte! It is therefore essential to approximate tensors of higher order in a compressed scheme, for example, a low-rank tensor decomposition. Various such decompositions have been developed, see Section 2. An important difference to tensors containing observed data, a tensor $\mathcal{X}$ induced by a function is usually not given directly but only as the solution of some algebraic equation, e.g., a linear system or eigenvalue problem. This requires the development of solvers for such equations working within the compressed storage scheme. Such algorithms are discussed in Section 3.

The range of applications of low-rank tensor techniques is quickly expanding. For example, they have been used for addressing:

- the approximation of parameter-dependent integrals [15, 164, 190], multi-dimensional integrals [48, 112, 153], and multi-dimensional convolution [152];

* lgr@igpm.rwth-aachen.de

** daniel.kressner@epfl.ch

*** christine.tobler@epfl.ch

- computational tasks in electronic structure calculations, e.g., based on Hartree-Fock or DFT models [23, 24, 31, 44, 45, 47, 87, 141–146, 150, 156–159];
- the solution of stochastic and parametric PDEs [68, 69, 76, 77, 161, 165, 171, 189, 202];
- approximation of Green's functions of high-dimensional PDEs [30, 113];
- the solution of the Boltzmann equation [129, 149], and the chemical master / Fokker-Planck equations [4, 6, 61, 64, 86, 122, 133, 135, 196]
- the solution of high-dimensional Schrödinger equations [46, 160, 172, 184, 192, 193];
- computational tasks in micromagnetics [80, 81];
- rational approximation problems [242];
- computational homogenization [42, 96, 176];
- computational finance [134, 137];
- the approximation of stationary states of stochastic automata networks [38];
- multivariate regression and machine learning [25].

Note that the list above mainly focuses on techniques that involve tensors; low-rank matrix techniques, such as POD and reduced basis methods, have been applied in an even broader setting.

A word of caution is appropriate. Even though the field of low-rank tensor approximation is relatively young, it already appears a daunting task to give proper credit to all developments in this area. This survey is biased towards work related to the TT and hierarchical Tucker decompositions. Surveys with a similar scope are the lecture notes by Grasedyck [102], Khoromskij [147,153,155], and Schneider [233], as well as the monograph by Hackbusch [108]. Less detailed attention may be given to other developments, although we have made an effort to at least touch on all important directions in the area of function-related tensors.

2 Low-rank tensor decompositions

As mentioned in the introduction, it will rarely be possible to store all entries of a higher-order tensor explicitly. Various compression schemes have been developed to reduce storage requirements. For $d = 2$ all these schemes boil down to the well known reduced singular value decomposition (SVD) of matrices [97]; however, they differ significantly for tensors of order $d \geq 3$.

2.1 CP decomposition

The entries of a rank-one tensor $\mathcal{X}$ can be written as

$$\mathcal{X}_{i_1,i_2,\ldots,i_d} = u^{(1)}_{i_1} u^{(2)}_{i_2} \cdots u^{(d)}_{i_d}, \quad 1 \leq i_\mu \leq n_\mu, \quad \mu = 1, \ldots, d. \tag{1}$$

By defining the vectors $u^{(\mu)} := \left(u^{(\mu)}_1, \ldots, u^{(\mu)}_{n_\mu}\right)^T$, a more compact from of this relation is given by

$$\operatorname{vec}(\mathcal{X}) = u^{(d)} \otimes u^{(d-1)} \otimes \cdots \otimes u^{(1)},$$

where $\otimes$ denotes the usual Kronecker product and vec stacks the entries of a tensor into a long column vector, such that the indices are in *reverse* lexicographical order. (Using the vector outer product $\circ$, this relation takes the form $\mathcal{X} = u^{(1)} \circ u^{(2)} \circ \cdots \circ u^{(d)}$.) When $\mathcal{X}$ represents the discretization of a separable function $f(x_1, x_2, \ldots, x_d) = f_1(x_1) f_2(x_2) \ldots f_d(x_d)$ then $\mathcal{X}$ is a rank-one tensor with each vector $u^{(\mu)}$ corresponding to a discretization of f_μ.

The *CP (CANDECOMP/PARAFAC) decomposition* is a sum of rank-one tensors:

$$\operatorname{vec}(\mathcal{X}) = u^{(d)}_1 \otimes u^{(d-1)}_1 \otimes \cdots \otimes u^{(1)}_1 + \cdots + u^{(d)}_R \otimes u^{(d-1)}_R \otimes \cdots \otimes u^{(1)}_R. \tag{2}$$

The *tensor rank* of $\mathcal{X}$ is defined as the minimal R such that $\mathcal{X}$ has a CP decomposition with R terms. Note that, in contrast to matrices, the set CP(R) of tensors of rank at most R is in general not closed, which renders the problem of finding a best low-rank approximation ill-posed [56]. For more properties of the CP decomposition, we refer to the survey paper [168].

The CP decomposition requires the storage of $(n_1 + n_2 + \cdots + n_d)R$ entries, which becomes very attractive for small R. To be able to use the CP decomposition for the approximation of function-related tensors, robust and efficient compression techniques are essential. In particular, the truncation of a rank-R tensor to lower tensor rank is frequently needed. Nearly all existing algorithms are based on carefully adapting existing optimization algorithms, see, once again, [168] for an overview of the literature until around 2009. More recent developments for general tensors include work on increasing the efficiency and robustness of gradient-based and Newton-like methods [2, 72, 74, 140, 223, 224], modifying and improving ALS (alternating least squares) [41, 57, 58, 90, 225], studying the convergence of ALS [51, 194, 253] and reducing the cost of the unfolding operations required during the approximation [222, 269].

One can impose additional structure on the coefficients of the CP decomposition, such as nonnegativity. As this is of primary interest in data analysis applications, a comprehensive discussion is beyond the scope of this survey, see [168].

2.2 Tucker decomposition

A *Tucker decomposition* of a tensor $\mathcal{X}$ takes the form

$$\operatorname{vec}(\mathcal{X}) = (U_d \otimes U_{d-1} \otimes \cdots \otimes U_1) \operatorname{vec}(\mathcal{C}), \tag{3}$$

where $U_1, U_2, \ldots, U_d$, with $U_\mu \in \mathbb{R}^{n_\mu \times r_\mu}$, are called the *factor matrices* or *basis matrices* and $\mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$ is called the *core tensor* of the decomposition.

Like CP, the Tucker decomposition has a long history and we refer to the survey [168] for a more detailed account. In the following, we briefly summarize the basic techniques, which are needed to motivate the TT and HT decompositions discussed below.

The Tucker decomposition is closely related to the matricizations of $\mathcal{X}$. The μth matricization $X^{(\mu)}$ is an $n_\mu \times \big(n_1 \cdots n_{\mu-1} n_{\mu+1} \cdots n_d\big)$ matrix formed in a specific way from the entries of $\mathcal{X}$:

$$X^{(\mu)}_{i_\mu, \ell} := \mathcal{X}_{i_1, \ldots, i_d}, \qquad \ell = 1 + \sum_{\nu<\mu} (i_\nu - 1) \prod_{\eta<\nu} n_\eta + \sum_{\nu>\mu} (i_\nu - 1) \prod_{\substack{\eta<\nu \\ \eta\neq\mu}} n_\eta.$$

In particular, the relation (3) implies

$$X^{(\mu)} = U_\mu \cdot C^{(\mu)} \cdot \big(U_d \otimes \cdots \otimes U_{\mu+1} \otimes U_{\mu-1} \otimes \cdots \otimes U_1\big)^T, \qquad \mu = 1, \ldots, d. \tag{4}$$

It follows that $\operatorname{rank}\big(X^{(\mu)}\big) \leq r_\mu$, as the first factor $U_\mu \in \mathbb{R}^{n_\mu \times r_\mu}$ obviously has rank at most r_μ. This motivates to define the *multilinear rank* (also called μ-rank) of a tensor $\mathcal{X}$ as the tuple

$$(r_1, \ldots, r_d), \quad \text{with} \quad r_\mu = \operatorname{rank}\big(X^{(\mu)}\big).$$

In contrast to the tensor rank related to the CP decomposition, the set T($r_1, \ldots, r_d$) of tensors of μ-rank at most r_μ is closed.

Another consequence of the relation (4) is the higher-order SVD (HOSVD) introduced in [54, 55] for approximating a tensor by a Tucker decomposition (3) of lower multilinear rank. In HOSVD, the columns of each factor matrix U_μ are computed as the k_μ dominant left singular vectors of $X^{(\mu)}$. The core tensor is then obtained by forming $\operatorname{vec}(\mathcal{C}) := \big(U_d \otimes \cdots \otimes U_1\big)^T \operatorname{vec}(\mathcal{X})$. Eventually, this yields

$$\operatorname{vec}\big(\widetilde{\mathcal{X}}\big) := \big(U_d \otimes \cdots \otimes U_1\big) \cdot \operatorname{vec}(\mathcal{C}) \quad \in T(k_1, \ldots, k_d).$$

In contrast to the matrix case, where the SVD yields a best low-rank approximation for all unitarily invariant norms [125, Sec. 7.4.9], the truncated tensor $\widetilde{\mathcal{X}}$ resulting from the HOSVD is usually not optimal. However, we have

$$\|\mathcal{X} - \widetilde{\mathcal{X}}\| \leq \sqrt{d} \min_{\mathcal{Y} \in T(k_1, \ldots, k_d)} \|\mathcal{X} - \mathcal{Y}\|.$$

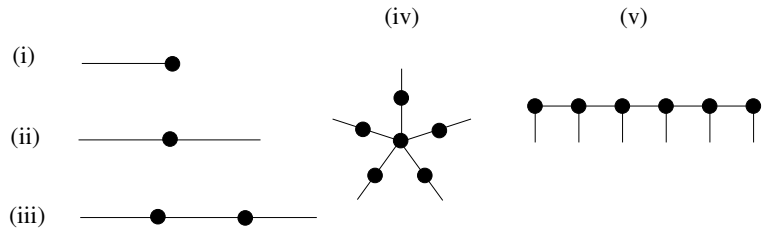

Fig. 1 Tensor network diagrams representing (i) a vector, (ii) a matrix, (iii) a matrix-matrix multiplication, (iv) a tensor in Tucker decomposition, and (v) a tensor in TT decomposition.

This quasi-optimality condition is usually sufficient for the purpose of obtaining an accurate approximation to a function-related tensor.

Various alternatives to improve on the approximation provided by the HOSVD have been developed, see [168] and the references therein. Recent developments include Newton-type methods on manifolds [71, 131, 132, 229], a Jacobi algorithm for symmetric tensors [130], generalizations of Krylov subspace methods [98, 228], and modifications of the HOSVD [259].

2.3 Tensor train decomposition

The need for storing the $r_1 \times \cdots \times r_d$ core tensor $\mathcal{C}$ renders the Tucker decomposition increasingly unattractive as d gets larger. This has motivated the search for decompositions which potentially avoid these exponentially growing memory requirements, while still featuring the two most important advantages of the Tucker decomposition: closedness and SVD-based compression.

One well established candidate for such a decomposition is the so called *TT* (tensor train) decomposition, which takes the form

$$\mathcal{X}_{i_1,\ldots,i_d} = G_1(i_1) \cdot G_2(i_2) \cdots G_d(i_d), \qquad G_\mu(i_\mu) \in \mathbb{R}^{r_{\mu-1} \times r_\mu}, \tag{5}$$

where $r_0 = r_d = 1$. For every mode μ and every index i_μ the coefficients $G_\mu(i_\mu)$ are matrices. In the context of numerical analysis, a decomposition of the form (5) was first proposed in [212, 213, 217]. However, such a decomposition has been proposed earlier in the density-matrix renormalization group method (DMRG) for simulating quantum systems [235, 266]. In this area, the term *matrix product state* (MPS) representation for the decomposition (5) has been established [218]. Suitable conditions that imply a unique MPS representation can be found in [221, 263]. The connection between TT and MPS has been explained in [127].

Similar to the Tucker decomposition, the TT decomposition is closely related to certain matricizations of $\mathcal{X}$. Let $X^{(1,\ldots,\mu)}$ denote the matrix obtained by reshaping the entries of $\mathcal{X}$ into an $(n_1 \cdots n_\mu) \times (n_{\mu+1} \cdots n_d)$ array, such that (5) implies $\operatorname{rank}\big(X^{(1,\ldots,\mu)}\big) \leq r_\mu$ for $\mu = 1, \ldots, d$. Consequently, the tuple containing the ranks of these matricizations is called the *TT-rank* of $\mathcal{X}$. As explained, e.g., in [217] a quasi-best approximation in a TT decomposition for a given TT-rank can be obtained from the SVDs of $X^{(1,\ldots,\mu)}$, similarly to the HOSVD. It is important to avoid the explicit construction of these matrices and the SVDs when truncating a tensor in TT decomposition to lower TT-rank. Such truncation algorithms are described in [217]. On the theoretical side, it turns out that the set $TT(r_1, \ldots, r_{d-1})$ of tensors with TT-ranks bounded by r_μ is closed, and under a full rank condition it actually forms a smooth manifold [123, 255]. The Kähler manifold structure for complex MPS with open and periodic boundary conditions has been studied in [119].

Tensor network diagrams, which have been attributed to Penrose [220], are helpful in visualizing tensor decompositions and their manipulation. Figure 1 gives a few basic examples, see, e.g., [124, 127, 173] for more details. In particular, Figure 1 (v) gives an illustration of the contraction (5) representing a TT decomposition. In view of this diagram, the TT decomposition is also sometimes called linear tensor network [91].

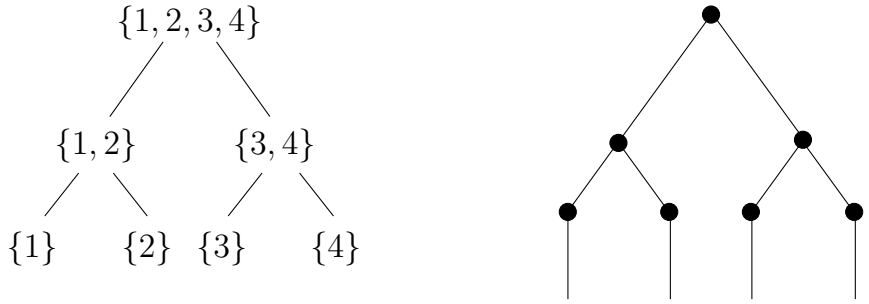

Fig. 2 Left: Binary tree representing mode splitting for HT decomposition. Right: Tensor network diagram representing a tensor in HT decomposition.

In applications related to quantum spin systems, the tensor $\mathcal{X}$ often exhibits symmetries inherited from underlying physical properties. There are variants of MPS/TT that reflect such symmetries in the low-rank decomposition, see [128, 221, 227] and the references therein.

2.4 Hierarchical Tucker decomposition

An alternative way to reduce the complexity of the Tucker decomposition is given by the *hierarchical Tucker (HT) decomposition* [103, 117] (also called *hierarchical tensor representation*). This decomposition is based on the idea of recursively splitting the modes of the tensor, which results in a binary tree $\mathcal{T}$ containing a subset $t \subset \{1, \dots, d\}$ at each node. An example of such a binary tree is given in the left plot of Figure 2. The matricization $X^{(t)}$ of a tensor $\mathcal{X}$ corresponding to such a subset t merges all modes contained in t into row indices of the matrix, and the other modes into column indices. We then consider a hierarchy of matrices U_t whose columns span the image of $X^{(t)}$ for each $t \in \mathcal{T}$. Hence, U_t has exactly $r_t = \operatorname{rank}(X^{(t)})$ columns. The rank tuple $(r_t)_{t\in\mathcal{T}}$ is called the *HT-rank* of $\mathcal{X}$.

The following nestedness property allows for the implicit storage of $(U_t)_{t\in\mathcal{T}}$, and thus of the tensor $\mathcal{X}$: For $t = t_l \cup t_r$, $t_l \cap t_r = \emptyset$, there exists a matrix B_t such that

$$U_t = (U_{t_r} \otimes U_{t_l})B_t, \qquad B_t \in \mathbb{R}^{r_{t_l} r_{t_r} \times r_t}. \tag{6}$$

For simplicity, we have assumed that the ordering of the modes in the tree $\mathcal{T}$ is such that all modes contained in t_l are smaller than the modes contained in t_r. The relation (6) implies that it suffices to store the basis matrices U_t only for the leaf nodes $t = \{1\}, \{2\}, \dots, \{d\}$, and B_t for all other nodes in $\mathcal{T}$. The resulting storage requirements are $O(dnr + dr^3)$, when assuming $r \equiv r_t$ and $n \equiv n_\mu$.

Similarly to the Tucker and TT decompositions, a quasi-best approximation in the HT decomposition for a given HT-rank can be obtained from the SVDs of $X^{(t)}$. Algorithms that avoid the explicit computation of these SVDs when truncating a tensor that is already in HT decomposition are discussed in [103, 117, 173, 175]. As for the TT decomposition, the set of tensors having fixed HT-rank forms a smooth manifold [83, 254, 255].

The tensor network corresponding to the HT decomposition is always a binary tree, see also the right plot of Figure 2. Such tensor tree networks had already been discussed in [238] (without the basis matrices at the leafs). Moreover, the so called multilayer multi-configuration time-dependent Hartree method (ML-MCTDH) introduced in [265] makes use of a decomposition based on general trees instead of binary trees. When allowing for general trees, tensor tree networks include the Tucker decomposition from Section 2.2 as a (quite particular) special case. In the case of a degenerate tree, where at each level, one mode is split from the remaining modes, the HT decomposition becomes equivalent to a *variant* of the TT decomposition discussed in [62, 217]. In contrast to the TT decomposition defined in (5), this variant features additional basis matrices, which may reduce the storage cost for large n_μ. A discussion on the difference between the ranks for the HT and TT decompositions can be found in [105].

2.5 More general tensor network formats

Motivated by an underlying topology describing interactions, tensor networks beyond trees have been considered in the context of renormalization group methods for simulating strongly correlated quantum spin systems. Well-known examples include the so called projected entangled-pair states (PEPS) [260, 261] and the multiscale entanglement renormalization ansatz (MERA) [264]. Both, PEPS and MERA contain cycles in the tensor network. Tree-structured tensor networks, as the hierarchical Tucker and the TT format, are closed [75, 82, 108] in the sense that tensors with ranks at most r_μ form a closed set in $\mathbb{R}^{n_1 \times \cdots \times n_d}$. In general, this statement does not hold for tensor networks containing cycles [177, 178]. Possibly for this reason, more general networks have not yet been considered to a large extend in the numerical analysis community for, e.g., the solution of high-dimensional PDEs, but see [75, 121] for some recent mathematically oriented work.

2.6 Hybrid formats

Adding to the diversity of the formats discussed above, it is possible and sometimes useful to combine different low-rank formats. One popular combination is the Tucker format combined with the CP format for the approximation of the core tensor [148, 156, 157, 170], see [168, Sec. 5.7] for other variations of Tucker and CP. In [93, 110, 111, 115], combinations of low-rank tensor formats with hierarchical matrices are investigated.

2.7 A priori approximation results

As an essential prerequisite for the success of tensor-based computations, it is important to decide whether a tensor generated by a certain multivariate function $f(x_1, \ldots, x_d)$ can be well approximated by a low-rank tensor decomposition. As discussed in Section 2.1, the tensor rank is closely linked to approximating f by a sum of separable functions. Only in exceptional cases, it will be possible to represent f *exactly* by such a sum, see [26, 195, 208].

In general, one is therefore interested in an approximation of the form

$$f(x_1, \ldots, x_d) = \sum_{r=1}^{R} f_r^{(1)}(x_1) \cdot f_r^{(2)}(x_2) \cdots f_r^{(d)}(x_d) + \varepsilon. \tag{7}$$

For a function of the form $f(x_1, \ldots, x_d) = g(x_1 + \cdots + x_d)$, such an approximation can be immediately obtained from approximating g by a sum of exponentials. For this purpose, various approaches have been discussed in [27–29, 36, 37]. Other techniques include applying numerical quadrature to an integral representation of the function, see, e.g., [24, 77, 87, 110, 115], Taylor series expansion [249, 250], and polynomial interpolation [19, 33]. Based on results by Temlyakov [243–245] on bilinear approximation rates, singular value estimates for the hierarchical Tucker decomposition of functions in mixed Sobolev spaces have been obtained in [234], see also [106]. In fact, a sparse grid approximation to f can be turned quite effectively into a low-rank tensor decomposition [108, 117]. General nonlinear best R-term approximation schemes [49, 50, 171] represent another important technique, which we cannot cover in detail.

Even for smooth f, it may not always be possible to attain sufficiently low ranks, especially when the variation of f is too strong across its entire domain of definition. In this case, it can be advantageous to subdivide the domain and approximate f on each subdomain separately with a low-rank tensor decomposition, see [11, 14, 15] for examples. As first discussed in [26], an approximation of the form (7) can also be used to approximate linear operators on tensors, see also Section 2.8 below.

In the other direction, SVD-based approximations of a function-related tensor yield an approximation of the underlying function, where the L^2-norm approximation error can be directly controlled by the truncated singular values. For smooth functions, best approximations in tensor formats are known to inherit the regularity of the approximated function [252, 254]. This also holds for SVD based quasi-best approximations, for which even the smoothness of the error can be controlled [109]. This can be used, e.g., for deriving L^∞ error estimates [109] or approximation results for the basis matrices U_μ [234].

For quantum many-body systems, the approximability of the ground state by a low-rank TT decomposition is closely linked to the concepts of entropy and entanglement, see [221] for an introduction.

2.8 Low-rank decomposition of linear operators

The matrix representation of a linear operator

$$\mathcal{A} : \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \to \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_d}, \quad \mathcal{X} \mapsto \mathcal{A}(\mathcal{X}),$$

can be viewed as an $m_1 n_1 \times m_2 n_2 \times \cdots \times m_d n_d$ tensor after pairing up row and column indices:

$$\mathcal{A}_{i_1,i_2,\ldots,i_d;j_1,j_2,\ldots,j_d} \qquad \Rightarrow \qquad \mathcal{A}_{(i_1,j_1),(i_2,j_2),\ldots,(i_d,j_d)}.$$

This view allows to apply any of the low-rank tensor decompositions discussed above to $\mathcal{A}$. Such a low-rank decomposition of $\mathcal{A}$ is useful, e.g., for performing the matrix-vector product $\mathcal{A}(\mathcal{X})$ efficiently when $\mathcal{X}$ itself admits a low-rank decomposition. This idea appears to be ubiquitous in the literature on low-rank tensor decomposition, see [26] for an early reference. For example, in the study of strongly correlated quantum systems, matrix product operators (MPO) were introduced in [262, 270], which corresponds to representing $\mathcal{A}$ in the TT decomposition. As pointed out in [116], having $\mathcal{A}$ in a low-rank decomposition also allows to compute an approximate inverse by combining the Newton-Hotelling-Schulz algorithm with truncation, see Section 3.1.

2.9 Tensorization

Reverting the process of vectorization, the entries x_j of a vector $x \in \mathbb{R}^N$, $N = 2^d$, can be rearranged into an $2 \times 2 \times \cdots \times 2$ tensor $\mathcal{X}$ of order d. (The binary representation $j - 1 = \sum_{\mu=1}^{d} 2^{\mu-1} i_\mu$ yields a simple mapping to the corresponding multi-index $(i_1, \ldots, i_d)$ of $\mathcal{X}$.) In turn, a low-rank approximation of $\mathcal{X}$ yields a compression of the original vector x. In combination with the TT decomposition, this idea of *tensorization* or *quantization* is usually called *Quantics-TT* or *Quantized-TT* (QTT); it was first used as a compression scheme for matrices in [205], and introduced for a more general setting in [154].

Quantization is particularly interesting when the vector x represents a function $f : I \to \mathbb{R}$ evaluated at 2^d points, usually uniformly distributed in the interval. The exact and approximate ranks of $\mathcal{X}$ for various functions f have been discussed for the TT and HT decompositions in [104, 108, 154].

Applying quantization to each mode of a tensor $\mathcal{X} \in \mathbb{R}^{N \times \cdots \times N}$ of order D, $N = 2^d$, yields a $2 \times 2 \times \cdots \times 2$ tensor $\mathcal{Y}$ of order $d \cdot D$. This gives rise to a variety of mixed low-rank tensor decompositions, as discussed in [62, 154].

QTT has been applied to the solution of PDEs and eigenvalue problems [144, 154, 162], evaluation of boundary integrals in BEM [164], convolution [107] and the FFT [63, 108, 230]. A connection between QTT and the wavelet transform is discussed in [215].

One important ingredient of QTT is that the involved matrices can be represented in a way that conforms to the format [162]. For the following matrices, QTT representations have been discussed: Toeplitz matrices [136, 216], (inverse) Laplace operators [139, 205], linear diffusion operators [62, 138].

2.10 Software

There are several MATLAB toolboxes available for dealing with tensors in CP and Tucker decomposition, including the Tensor Toolbox [12, 13], the N-way toolbox [7], the PLS_Toolbox [267], and the Tensorlab [240]. The TT-Toolbox [207] provides MATLAB classes covering tensors in TT and QTT decomposition, as well as linear operators. There is also a Python implementation of the TT-toolbox called `ttpy` [204]. The `htucker` toolbox [173] provides a MATLAB class representing a tensor in HT decomposition.

The *TensorCalculus library* [79] is a mathematically oriented C++ library, allowing for computations with general tensor networks. The Heidelberg MCTDH Package [268] is a set of Fortran programs for multi-dimensional quantum dynamics. ALPS [18] provides C++ libraries for simulating strongly correlated quantum mechanical systems, including DMRG. Block [237] is a C++ implementation of the DMRG algorithms discussed in [236]. The tensor contraction engine [10] automatically generates near-optimal code for tensor contractions in many-body electronic structure methods.

3 Algorithms

In applications for function-related tensors, $\mathcal{X}$ is often given implicitly, e.g., as the solution to a linear system or eigenvalue problem. There are mainly two different types of approaches to obtain an approximation to $\mathcal{X}$. A first class of methods is based on combining classical iterative algorithm with repeated low-rank truncation. A second class is based on reformulating the problem at hand as an optimization problem, constraining the admissible set to low-rank tensors, and applying various optimization techniques.

3.1 Iterative methods combined with truncation

In principle, any vector iteration for solving a linear algebra problem involving a tensor can be combined with truncation in any of the low-rank decompositions discussed above. To illustrate the basic principle, let us consider the (preconditioned) Richardson iteration for solving a linear system $\mathcal{A}(\mathcal{X}) = \mathcal{B}$:

$$\mathcal{X}_{k+1} = \mathcal{X}_k + \omega\mathcal{P}\big(\mathcal{B} - \mathcal{A}(\mathcal{X}_k)\big),$$

where $\mathcal{P}$ is a preconditioner and ω is a suitably chosen scalar. Letting $\mathcal{T}$ denote truncation in any of the low-rank tensor decompositions discussed above, we obtain the truncated Richardson iteration

$$\mathcal{X}_{k+1} = \mathcal{T}\Big(\mathcal{X}_k + \omega\mathcal{P}\big(\mathcal{B} - \mathcal{A}(\mathcal{X}_k)\big)\Big),$$

which has been proposed [165] in combination with the CP decomposition.

Other examples of combining iterative methods with low-rank truncation include:

- the (shift-and-invert) power method combined with CP [26, 27];
- the (shift-and-invert) power and Lanczos methods combined with CP and Tucker [11, 114];
- a restarted Lanczos method combined with TT [126];
- conjugate gradient type-methods for symmetric eigenvalue problems combined with truncation for low-rank matrices [180], as well as for TT [187], HT [172] and QTT [181].
- the Richardson method combined with QTT [161] and low-rank matrix decompositions [189];
- a projection method combined with HT [16];
- the conjugate gradient method, BiCGStab and other Krylov subspace methods combined with HT [169, 171, 246] and low-rank matrix decompositions [32];
- GMRES combined with TT [59].

When applying an iterative method in combination with a low-rank decompositions, the ranks will inevitably grow quite quickly in the course of the iteration. For example, the sum of k tensors can multiply the ranks by k, while the pointwise (Hadamard) product between two tensors can even square the ranks. To gain efficiency, it is therefore advisable to not let this rank growth happen explicitly and combine such an operation directly with truncation. This has been discussed for sums in [173, 246] and for the Hadamard product (and other bilinear operations) in [232], see also [206].

Preconditioners not only help accelerate convergence but numerical evidence suggests that an effective preconditioner leads to a limited intermediate rank growth. At the same time, it is mandatory that the preconditioner can be applied efficiently to a low-rank tensor decomposition. In particular, this is the case when the preconditioner itself admits a low-rank decomposition in the sense of Section 2.8. There are various techniques to construct such preconditioners. An early technique is based on best approximation in the Frobenius norm by a Kronecker product [251, 256], by a short sum of Kronecker products [203], or by a more general low-rank tensor decomposition. Other techniques include the use of approximate inverses for high-dimensional Laplace operators [151, 165, 172, 210], low-rank manipulation of the PDE coefficients [60], low-rank tensor approximation of multilevel preconditioners [8], and low-rank tensor diagonal preconditioners for wavelet discretizations [11].

3.2 Optimization-based algorithms

In many cases, a linear algebra problems involving a tensor can be posed as an optimization problem. For example, it is well known that a symmetric positive definite linear system $\mathcal{A}(\mathcal{X}) = \mathcal{B}$ can be turned into

$$\min_{\mathcal{X}\in\mathbb{R}^{n_1\times\cdots\times n_d}} \frac{1}{2}\langle\mathcal{X},\mathcal{A}(\mathcal{X})\rangle - \langle\mathcal{X},\mathcal{B}\rangle, \tag{8}$$

where $\langle\cdot,\cdot\rangle$ corresponds to the standard inner product for the vectorization of the tensors. For nonsymmetric linear systems, an optimization problem can be obtained by minimizing the norm of the residual. For symmetric eigenvalue problems, the Rayleigh-quotient minimization or, more generally, the trace minimization principle can be used.

Once the optimization problem is set up, the set of admissible tensors $\mathcal{X}$ is then constrained to a low-rank decomposition, for example, to all tensors with fixed tensor rank or fixed multilinear rank. Even when the original optimization problem, such as (8), is convex and in principle simple, the resulting constrained optimization problem is highly nonlinear and non-convex in general. A number of heuristic approaches to the solution of such constrained optimization problems are available, including ALS (alternating linear scheme). The basic principle of ALS is to optimize every factor of the low-rank decomposition separately and to sweep over all factors repeatedly. It is probably most natural to combine ALS with the CP decomposition, see [27,68] for an application of this idea to linear systems. The combination of ALS with the TT decomposition has been considered in [65–67, 124]. The convergence of ALS for the TT decomposition has been studied in [226].

The ALS scheme can be improved in various ways. One quite successful improvement for decompositions described by tensor networks is to join two neighboring factors, optimize the resulting supernode, and split the result into separate factors by a low-rank factorization. Originally, this so called DMRG method had been developed for the simulation of strongly correlated quantum lattice systems, see [235] for an overview. Later on, the ideas of DMRG have been picked up and extended to other applications in the numerical analysis community in a series of papers [65, 124, 127, 160, 171, 206].

There is a growing interest in applying so called Riemannian optimization techniques [1] to (8). Examples include nonlinear conjugate gradient or Newton-like methods on manifolds of low-rank matrices [35, 191, 198, 257, 258] or low-rank tensors [52, 53, 131, 132, 255], see also Section 3.4. For tensors in CP decomposition, the approximate solution of (8) by gradient techniques has been discussed in [78].

3.3 Successive rank-1 approximation

A tempting and surprisingly successful approach to the solution of high-dimensional problems is to build up a low-rank approximation from successive rank-1 approximations. This idea has been suggested in the context of various applications, including Fokker-Planck equations [4] and stochastic partial differential equations [202].

We will illustrate the basic idea with a simple example. Consider a linear system $\mathcal{A}(\mathcal{X}) = \mathcal{B}$ with the solution tensor $\mathcal{X} \in \mathbb{R}^{n_1\times\cdots\times n_d}$. Assume we already have a CP approximation $\mathcal{X}_r$ of tensor rank r:

$$\mathrm{vec}(\mathcal{X}_r) = u_1^{(d)} \otimes u_1^{(d-1)} \otimes \cdots \otimes u_1^{(1)} + \cdots + u_r^{(d)} \otimes u_r^{(d-1)} \otimes \cdots \otimes u_r^{(1)}. \tag{9}$$

We then search for a rank-1 correction

$$\mathcal{W} = u_{r+1}^{(d)} \otimes u_{r+1}^{(d-1)} \otimes \cdots \otimes u_{r+1}^{(1)},$$

such that $\mathcal{X}_{r+1} = \mathcal{X}_r + \mathcal{W}$ is an improved approximation, that is,

$$\mathcal{A}(\mathcal{X}_{r+1}) \approx \mathcal{B} \qquad \Leftrightarrow \qquad \mathcal{A}(\mathcal{W}) \approx \mathcal{B} - \mathcal{A}(\mathcal{X}_r) \tag{10}$$

Analogous to Section 3.2, the unknown vectors $w^{(1)}, \ldots, w^{(d)}$ can be determined by turning (10) into a nonlinear (optimization) problem and applying standard methods, such as the alternating direction method. This procedure is repeated until the residual $\mathcal{A}(\mathcal{X}_r) - \mathcal{B}$ is sufficiently small. Of course, such a greedy approach will not yield the best rank-R approximation after R steps [241]. However, it is important to

remember that these methods aim at a more moderate goal, to obtain a reasonable approximation after R steps, with R not too large. Convergence results in this direction can be found in [3, 39, 40, 84–86, 179].

A number of improvements to the simple scheme outlined above have been proposed to increase its convergence speed, see, e.g., [4, 95, 200, 202]. A connection between best rank-1 or, more generally, best rank-m approximations to a nonlinear eigenvalue problem is explained in [201]. This connection also motivates the use of the term *generalized spectral decomposition*. Many further developments, improvements, and extensions of successive low-rank approximation techniques have taken place during the last years; we refer to [43] for an overview.

3.4 Low-rank methods for dynamical problems

Let us consider a dynamical system on $\mathbb{R}^{n_1 \times \cdots \times n_d}$:

$$\dot{\mathcal{X}}(t) = F(\mathcal{X}(t)), \quad \mathcal{X}(t_0) = \mathcal{X}_0, \tag{11}$$

for which a typical application is the spatial discretization of a time-dependent d-dimensional PDE. *Dynamical low-rank methods* aim to determine an approximation $\mathcal{Y}(t)$ in a manifold $\mathcal{M}$ of low-rank tensors by restricting the dynamics of (11) to the tangent space $T_{\mathcal{Y}(t)}\mathcal{M}$:

$$\dot{\mathcal{Y}}(t) \in T_{\mathcal{Y}(t)}\mathcal{M} \quad \text{such that} \quad \|\mathcal{Y}(t) - F(\mathcal{Y}(t))\| = \min! \tag{12}$$

As explained in [183], this approximation is closely related to the can Dirac-Frenkel-McLachlan variational principle in quantum molecular dynamics.

Initially proposed for low-rank matrix manifolds in [166], dynamical low-rank methods have been extended to low-rank tensors in Tucker [167, 199], TT/MPS [118, 123, 163, 186], and HT [9, 186, 255] decomposition. The efficient and robust numerical integration of (12) is crucial to the success of dynamical low-rank methods; apart from the references above, this aspect has been discussed in [184, 185].

A more immediate approach to (11) is to combine a standard time stepping method, such as the explicit and implicit Euler methods, with low-rank truncation in every time step [64]. An alternative, which allows to control the error global-in-time, is to apply iterative solvers to a space-time formulation [5, 8, 61, 64, 94].

3.5 Black box approximation

Suppose that a matrix A or a tensor $\mathcal{X}$ is defined through a function that returns entries at arbitrary positions. Then the goal of *black box approximation* is to find a good low-rank approximation based only on relatively few entries. It is important to emphasize that the selection of the entries can be controlled by the user. In this respect, this situation is quite different from the growing area of tensor completion, see, e.g., [92, 182], where the selection of the entries is usually prescribed by the application.

For an $m \times n$ matrix A, the so called *cross approximation* method [19, 34, 100, 248] produces an approximation of the form

$$A(:, J)A(I, J)^{-1}A(I, :), \tag{13}$$

where MATLAB notation is used to denote the submatrices of A corresponding to the index sets $I \subset \{1, \ldots, m\}$ and $J \subset \{1, \ldots, n\}$. In the pth step of cross approximation as described in [21, 209], the entry of largest magnitude in the column j_p of $A - A(:, J)A(I, J)^{-1}A(I, :)$ is calculated and its position is denoted by i_p. Then the entry of largest magnitude in the row i_p of that matrix is calculated and its position is denoted by j_{p+1}. Moreover, both index sets are updated: $I \leftarrow I \cup \{i_p\}$ and $J \leftarrow J \cup \{j_p\}$. Volume maximization is an alternative entry selection strategy that attempts to maximize $\left|\det\left(A(I, J)\right)\right|$, see [99, 248].

A first extension of cross approximation to tensors was proposed in [209] for approximating third-order tensors by a Tucker decomposition. Essentially, this extension consists of applying the algorithm above to an arbitrary matricization of the tensor. However, the rows of this matricization (corresponding to slices of the tensor) are further approximated by a low rank matrix, again using cross approximation. In [88, 211], this method has been combined with multilevel ideas and applied to quantum chemistry. The adaptive cross approximation for the Tucker decomposition has been analyzed in [20]. Another variant for third and

fourth order tensors, focussing on interpolation properties, is discussed in [197]. A cross approximation method for the TT decomposition has been proposed in [214, 231].

Based on fiber crosses, a quite different extension for tensors of arbitrary order can be found in [73]. In this method a set of multi-indices $I^1, \dots, I^p \in [1, n_1] \times \cdots [1, n_d]$ is computed successively. Subspaces $\mathcal{U}_\mu$ are constructed approximately containing all μ-mode fibers passing through at least one of the multi-indices. The tensor is then approximated by a CP decomposition (2) under the constraint $u_j^{(\mu)} \in \mathcal{U}_\mu$, using general optimization methods. The multi-indices are selected by performing an alternating direction search along the fibers of the error tensor. Also based on fiber crosses, a black box approximation for a tensor in the HT decomposition is given in [14, 17].

Randomized algorithms represent an alternative way to quickly extract a low-rank approximation from partial information on the entries of the matrix, see [120] and the references therein. These ideas have been extended to low-rank tensor decompositions, for the Tucker decomposition in [70, 89, 247] and for the TT decomposition in [219].

3.6 Other algorithms

For linear systems and eigenvalue problems with a very particular structure, it is sometime possible to design specialized algorithms that can be more efficient and easier to analyze. This applies, in particular, to discretizations of the multi-dimensional Poisson equation [22, 101, 170, 188].

4 Acknowledgments

We thank Jonas Ballani, Wolfgang Hackbusch, Thomas Huckle, Boris Khoromskij, Anthony Nouy, Ivan Oseledets, André Uschmajew, and Bart Vandereycken for many helpful suggestions on a preliminary draft of this survey.

References

[1] P. A. Absil, R. Mahony, and R. Sepulchre, Optimization algorithms on matrix manifolds (Princeton University Press, Princeton, NJ, 2008).

[2] E. Acar, D. M. Dunlavy, and T. G. Kolda, A scalable optimization approach for fitting canonical tensor decompositions, J. Chemometrics **25**(2), 67–86 (2011).

[3] A. Ammar, F. Chinesta, and A. Falcó, On the convergence of a greedy rank-one update algorithm for a class of linear systems, Arch. Comput. Methods Eng. **17**(4), 473–486 (2010).

[4] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids, J. Non-Newton. Fluid Mech. **139**(3), 153–176 (2006).

[5] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modelling of complex fluids: Part II: Transient simulation using space-time separated representations, J. Non-Newton. Fluid Mech. **144**(2–3), 98–121 (2007).

[6] A. Ammar, M. Normandin, F. Daim, D. González, E. Cueto, and F. Chinesta, Non incremental strategies based on separated representations: applications in computational rheology, Commun. Math. Sci. **8**(3), 671–695 (2010).

[7] C. A. Andersson and R. Bro, The N-way toolbox for MATLAB, Chemometrics and Intelligent Laboratory Systems **52**(1), 1–4 (2000), Available at `http://www.models.life.ku.dk/nwaytoolbox`.

[8] R. Andreev and C. Tobler, Multilevel preconditioning and low rank tensor iteration for space-time simultaneous discretizations of parabolic PDEs, Tech. Rep. 2012–16, Seminar for Applied Mathematics, ETH Zürich, 2012, Available at `http://www.sam.math.ethz.ch/reports/2012/16`.

[9] A. Arnold and T. Jahnke, On the approximation of high-dimensional differential equations in the hierarchical Tucker format, Tech. rep., KIT, Karlsruhe, Germany, 2012.

[10] A. A. Auer et al., Automatic code generation for many-body electronic structure methods: the tensor contraction engine, Molecular Physics **104**(2), 211–228 (2006).

[11] M. Bachmayr, Adaptive low-rank wavelet methods and applications to two-electron Schrödinger equations, PhD thesis, RWTH Aachen, Germany, 2012.

[12] B. W. Bader and T. G. Kolda, Algorithm 862: MATLAB tensor classes for fast algorithm prototyping, ACM Trans. Math. Software **32**(4), 635–653 (2006), Available at `http://csmr.ca.sandia.gov/~tgkolda/TensorToolbox/`.
[13] B. W. Bader and T. G. Kolda, MATLAB Tensor Toolbox Version 2.4, March 2010, Available at `http://csmr.ca.sandia.gov/~tgkolda/TensorToolbox/`.
[14] J. Ballani, Fast evaluation of near-field boundary integrals using tensor approximations, Dissertation, Universität Leipzig, 2012.
[15] J. Ballani, Fast evaluation of singular BEM integrals based on tensor approximations, Numer. Math. **121**(3), 433–460 (2012).
[16] J. Ballani and L. Grasedyck, A projection method to solve linear systems in tensor format, Numer. Linear Algebra Appl. **20**(1), 27–43 (2013).
[17] J. Ballani, L. Grasedyck, and M. Kluge, Black box approximation of tensors in hierarchical Tucker format, Linear Algebra Appl. **438**(2), 639–657 (2013).
[18] B. Bauer et al., The ALPS project release 2.0: open source software for strongly correlated systems, J. Stat. Mech. **5** (2011), Available at `http://alps.comp-phys.org`.
[19] M. Bebendorf, Approximation of boundary element matrices, Numer. Math. **86**(4), 565–589 (2000).
[20] M. Bebendorf, Adaptive cross approximation of multivariate functions, Constr. Approx. **34**, 149–179 (2011).
[21] M. Bebendorf and S. Rjasanow, Adaptive low-rank approximation of collocation matrices, Computing **70**(1), 1–24 (2003).
[22] B. Beckermann, D. Kressner, and C. Tobler, An error analysis of Galerkin projection methods for linear systems with tensor product structure, Tech. rep., MATHICSE, EPF Lausanne, Switzerland, 2012.
[23] U. Benedikt, A. A. Auer, M. Espig, and W. Hackbusch, Tensor decomposition in post-Hartree-Fock methods. I. two-electron integrals and MP2, J. Chem. Phys **134**(5), 054118–1–054118–12 (2011).
[24] C. Bertoglio and B. N. Khoromskij, Low-rank quadrature-based tensor approximation of the Galerkin projected Newton/Yukawa kernels, Computer Physics Communications **183**(4), 904–912 (2012).
[25] G. Beylkin, J. Garcke, and M. J. Mohlenkamp, Multivariate regression and machine learning with sums of separable functions, SIAM J. Sci. Comput. **31**(3), 1840–1857 (2009).
[26] G. Beylkin and M. J. Mohlenkamp, Numerical operator calculus in higher dimensions, Proc. Natl. Acad. Sci. USA **99**(16), 10246–10251 (2002).
[27] G. Beylkin and M. J. Mohlenkamp, Algorithms for numerical analysis in high dimensions, SIAM J. Sci. Comput. **26**(6), 2133–2159 (2005).
[28] G. Beylkin and L. Monzón, On approximation of functions by exponential sums, Appl. Comput. Harmon. Anal. **19**(1), 17–48 (2005).
[29] G. Beylkin and L. Monzón, Approximation by exponential sums revisited, Appl. Comput. Harmon. Anal. **28**(2), 131–149 (2010).
[30] D. J. Biagioni, Numerical construction of Green's functions in high dimensional elliptic problems with variable coefficients and analysis of renewable energy data via sparse and separable approximations, PhD thesis, University of Colorado, 2012.
[31] T. Blesgen, V. Gavini, and V. Khoromskaia, Approximation of the electron density of aluminium clusters in tensor-product format, J. Comput. Phys. **231**(6), 2551–2564 (2012).
[32] M. Bollhöfer and A. K. Eppler, Low-rank Cholesky-factor Krylov subspace methods for generalized projected Lyapunov equations, Tech. rep., To appear in "System Reduction for Nanoscale IC Design" (Peter Benner, ed.), Mathematics in Industry (Springer), 2012.
[33] S. Börm, $\mathcal{H}_2$-matrices – an efficient tool for the treatment of dense matrices, Habilitationsschrift, Christian-Albrechts-Universität zu Kiel, 2006.
[34] S. Börm and L. Grasedyck, Hybrid cross approximation of integral operators, Numer. Math. **101**(2), 221–249 (2005).
[35] N. Boumal and P. A. Absil, RTRMC: A Riemannian trust-region method for low-rank matrix completion, in: Proceedings of the Neural Information Processing Systems Conference (NIPS), (2011).
[36] D. Braess and W. Hackbusch, Approximation of $1/x$ by exponential sums in $[1, \infty)$, IMA J. Numer. Anal. **25**(4), 685–697 (2005).
[37] D. Braess and W. Hackbusch, On the efficient computation of high-dimensional integrals and the approximation by exponential sums, in: Multiscale, Nonlinear and Adaptive Approximation, (Springer, 2009), pp. 39–74.
[38] P. Buchholz, Product form approximations for communicating Markov processes, Performance Evaluation **67**(9), 797–815 (2010).
[39] E. Cancès, V. Ehrlacher, and T. Lelièvre, Convergence of a greedy algorithm for high-dimensional convex nonlinear problems, Mathematical Models and Methods in Applied Sciences **21**(12), 2433–2467 (2011).
[40] E. Cancès, V. Ehrlacher, and T. Lelièvre, Greedy algorithms for high-dimensional non-symmetric linear problems, arxiv:1210.6688, 2012.
[41] Y. Chen, D. Han, and L. Qi, New ALS methods with extrapolating search directions and optimal step size for complex-valued tensor decompositions, Signal Processing, IEEE Transactions on **59**(12), 5888–5898 (2011).

[42] F. Chinesta, A. Ammar, F. Lemarchand, P. Beauchene, and F. Boust, Alleviating mesh constraints: Model reduction, parallel time integration and high resolution homogenization, Comput. Meth. Appl. Mech. Engrg. **197**(5), 400–413 (2008).
[43] F. Chinesta, P. Ladeveze, and E. Cueto, A short review on model order reduction based on proper generalized decomposition, Arch. Comput. Methods Eng. **18**, 395–404 (2011).
[44] S. R. Chinnamsetty, M. Espig, H. J. Flad, and W. Hackbusch, Canonical tensor products as a generalization of Gaussian-type orbitals, Z. Phys. Chem. **224**(3–4), 681–694 (2010).
[45] S. R. Chinnamsetty, M. Espig, B. N. Khoromskij, W. Hackbusch, and H. J. Flad, Tensor product approximation with optimal rank in quantum chemistry, J. Chem. Phys **127**(8), 084110 (2007).
[46] S. R. Chinnamsetty, W. Hackbusch, and H. J. Flad, The tensor product approximation to single-electron systems, Tech. Rep. 23, MPI MIS Leipzig, 2009.
[47] S. R. Chinnamsetty, W. Hackbusch, and H. J. Flad, Efficient multi-scale computation of products of orbitals in electronic structure calculations, Comput. Vis. Sci. **13**(8), 397–408 (2010).
[48] S. R. Chinnamsetty, H. Luo, W. Hackbusch, H. J. Flad, and A. Uschmajew, Bridging the gap between quantum Monte Carlo and F12-methods, J. Chem. Phys **401**, 36–44 (2012).
[49] A. Cohen, R. DeVore, and C. Schwab, Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs, Found. Comput. Math. **10**(6), 615–646 (2010).
[50] A. Cohen, R. DeVore, and C. Schwab, Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's, Anal. Appl. (Singap.) **9**(1), 11–47 (2011).
[51] P. Comon, X. Luciani, and A. L. F. de Almeida, Tensor decompositions, alternating least squares and other tales, J. Chemometrics **23**(7–8), 393–405 (2009).
[52] O. Curtef, G. Dirr, and U. Helmke, Conjugate gradient algorithms for best rank-1 approximation of tensors, PAMM **7**(1), 1062201–1062202 (2007).
[53] O. Curtef, G. Dirr, and U. Helmke, Riemannian optimization on tensor products of Grassmann manifolds: Applications to generalized Rayleigh-quotients, SIAM J. Matrix Anal. Appl. **33**(1), 210–234 (2012).
[54] L. De Lathauwer, Signal Processing Based on Multilinear Algebra, PhD thesis, Katholike Universiteit Leuven, 1997.
[55] L. De Lathauwer, B. De Moor, and J. Vandewalle, A multilinear singular value decomposition, SIAM J. Matrix Anal. Appl. **21**(4), 1253–1278 (2000).
[56] 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. **20**(3), 1084–1127 (2008).
[57] H. de Sterck, A nonlinear GMRES optimization algorithm for canonical tensor decomposition, SIAM J. Sci. Comput. **34**(3), A1351–A1379 (2012).
[58] H. De Sterck and K. Miller, An adaptive algebraic multigrid algorithm for low-rank canonical tensor decomposition, Tech. Rep. arXiv:1111.6091, Nov 2011.
[59] S. V. Dolgov, TT-GMRES: on solution to a linear system in the structured tensor format, arXiv preprint 1206.5512 (To appear in: Rus. J. of Num. An. and Math. Model.), 2012.
[60] S. V. Dolgov, V. A. Kazeev, and B. N. Khoromskij, The tensor-structured solution of one-dimensional elliptic differential equations with high-dimensional parameters, Tech. Rep. 51, MPI MIS Leipzig, 2012.
[61] S. V. Dolgov and B. N. Khoromskij, Tensor-product approach to global time-space-parametric discretization of chemical master equation, Tech. Rep. 68, MPI MIS Leipzig, 2012.
[62] S. V. Dolgov and B. N. Khoromskij, Two-level Tucker-TT-QTT format for optimized tensor calculus, Tech. Rep. 19, MPI MIS Leipzig, 2012.
[63] S. V. Dolgov, B. N. Khoromskij, and D. V. Savostyanov, Superfast Fourier transform using QTT approximation, J. Fourier Anal. Appl. **18**(5), 519–953 (2012).
[64] S. V. Dolgov, B. N. Khoromskij, and I. V. Oseledets, Fast solution of parabolic problems in the tensor train/quantized tensor train format with initial application to the Fokker-Planck equation, SIAM J. Sci. Comput. **34**(6), A3016–A3038 (2012).
[65] S. V. Dolgov and I. V. Oseledets, Solution of linear systems and matrix inversion in the TT-format, SIAM J. Sci. Comput. **34**(5), A2718–A2739 (2012).
[66] S. V. Dolgov and D. V. Savostyanov, Alternating minimal energy methods for linear systems in higher dimensions. Part I: SPD systems, arXiv preprint 1301.6068, 2013.
[67] S. V. Dolgov and D. V. Savostyanov, Alternating minimal residual methods for the solution of high-dimensional linear systems in the tensor train format, in preparation, 2013.
[68] A. Doostan and G. Iaccarino, A least-squares approximation of partial differential equations with high-dimensional random inputs, J. Comput. Phys. **228**(12), 4332–4345 (2009).
[69] A. Doostan, A. Validi, and G. Iaccarino, Non-intrusive low-rank separated approximation of high-dimensional stochastic models, arXiv preprint 1210.1532, 2012.
[70] P. Drineas and M. W. Mahoney, A randomized algorithm for a tensor-based generalization of the singular value decomposition, Linear Algebra Appl. **420**(2–3), 553–571 (2007).

[71] L. Eldén and B. Savas, A Newton-Grassmann method for computing the best multilinear rank-(r_1, r_2, r_3) approximation of a tensor, SIAM J. Matrix Anal. Appl. **31**(2), 248–271 (2009).
[72] M. Espig, Effiziente Bestapproximation mittels Summen von Elementartensoren in hohen Dimensionen, PhD thesis, Fakultät für Mathematik und Informatik, Universität Leipzig, 2008.
[73] M. Espig, L. Grasedyck, and W. Hackbusch, Black box low tensor-rank approximation using fiber-crosses, Constr. Appr. **30**(3), 557–597 (2009).
[74] M. Espig and W. Hackbusch, A regularized Newton method for the efficient approximation of tensors represented in the canonical tensor format, Numer. Math. **122**(3), 489–525 (2012).
[75] M. Espig, W. Hackbusch, S. Handschuh, and R. Schneider, Optimization problems in contracted tensor networks, Comput. Vis. Sci. **14**, 271–285 (2011).
[76] M. Espig, W. Hackbusch, A. Litvinenko, H. Matthies, and E. Zander, Efficient analysis of high dimensional data in tensor formats, in: Sparse Grids and Applications, edited by J. Garcke and M. Griebel, Lecture Notes in Computational Science and Engineering Vol. 88 (Springer, 2013), pp. 31–56.
[77] M. Espig, W. Hackbusch, A. Litvinenko, H. G. Matthies, and P. Wähnert, Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats, Comput. Math. Appl. (2013), To appear. Available at `http://dx.doi.org/10.1016/j.camwa.2012.10.008`.
[78] M. Espig, W. Hackbusch, T. Rohwedder, and R. Schneider, Variational calculus with sums of elementary tensors of fixed rank, Numer. Math. **122**(52), 469–488 (2012).
[79] M. Espig, M. Schuster, A. Killaitis, N. Waldren, P. Wähnert, S. Handschuh, and H. Auer, TensorCalculus library, 2012, Available at `http://gitorious.org/tensorcalculus`.
[80] L. Exl, C. Abert, N. J. Mauser, T. Schrefl, H. P. Stimming, and D. Suess, FFT-based Kronecker product approximation to micromagnetic long-range interactions, Tech. Rep. arXiv:1212.3509, 2012.
[81] L. Exl, W. Auzinger, S. Bance, M. Gusenbauer, F. Reichel, and T. Schrefl, Fast stray field computation on tensor grids, J. Comput. Phys. **231**(7), 2840–2850 (2012).
[82] A. Falcó and W. Hackbusch, On minimal subspaces in tensor representations, Found. Comput. Math. **12**, 765–803 (2012).
[83] A. Falcó, W. Hackbusch, and A. Nouy, Geometric structures in tensor representations, Tech. Rep. 9, MPI MIS Leipzig, 2013.
[84] A. Falcó and A. Nouy, A proper generalized decomposition for the solution of elliptic problems in abstract form by using a functional Eckart-Young approach, J. Math. Anal. Appl. **376**(2), 469–480 (2011).
[85] A. Falcó and A. Nouy, Proper generalized decomposition for nonlinear convex problems in tensor Banach spaces, Numer. Math. **121**, 503–530 (2012).
[86] L. E. Figueroa and E. Süli, Greedy approximation of high-dimensional Ornstein-Uhlenbeck operators, Found. Comput. Math. **12**(5), 573–623 (2012).
[87] H. J. Flad, W. Hackbusch, B. N. Khoromskij, and R. Schneider, Concepts of data-sparse tensor-product approximation in many-particle modelling, in: Matrix methods: theory, algorithms and applications, (World Sci. Publ., Hackensack, NJ, 2010), pp. 313–347.
[88] H. J. Flad, B. N. Khoromskij, D. V. Savostyanov, and E. E. Tyrtyshnikov, Verification of the cross 3D algorithm on quantum chemistry data, Rus. J. Numer. Anal. Math. Model. **23**(4), 329–344 (2008).
[89] S. Friedland, V. Mehrmann, A. Miedlar, and M. Nkengla, Fast low rank approximations of matrices and tensors, Electron. J. Linear Algebra **22**, 1031–1048 (2011).
[90] S. Friedland, V. Mehrmann, R. Pajarola, and S. K. Suter, On best rank one approximation of tensors, arXiv preprint 1112.5914, 2012.
[91] F. Fröwis, V. Nebendahl, and W. Dür, Tensor operators: Constructions and applications for long-range interaction systems, Phys. Rev. A **81**(Jun), 062337 (2010).
[92] S. Gandy, B. Recht, and I. Yamada, Tensor completion and low-n-rank tensor recovery via convex optimization, Inverse Problems **27**(2), 025010 (2011).
[93] I. P. Gavrilyuk, W. Hackbusch, and B. N. Khoromskij, Hierarchical tensor-product approximation to the inverse and related operators for high-dimensional elliptic problems, Computing **74**, 131–157 (2005).
[94] I. P. Gavrilyuk and B. N. Khoromskij, Quantized-TT-Cayley transform to compute dynamics and spectrum of high-dimensional Hamiltonians, Comput. Meth. Appl. Math. **1**(3), 273–290 (2011).
[95] L. Giraldi, Contributions aux Méthodes de Calcul Basées sur l'Approximation de Tenseurs et Applications en Mécanique Numérique, PhD thesis, École Centrale de Nantes, 2012, Prelimary version.
[96] L. Giraldi, A. Nouy, G. Legrain, and P. Cartraud, Tensor-based methods for numerical homogenization from high-resolution images, Comput. Meth. Appl. Mech. Engrg. **254**, 154–169 (2013).
[97] G. H. Golub and C. F. Van Loan, Matrix Computations, third edition (Johns Hopkins University Press, Baltimore, MD, 1996).
[98] S. A. Goreinov, I. V. Oseledets, and D. V. Savostyanov, Wedderburn rank reduction and Krylov subspace method for tensor approximation. Part 1: Tucker case, SIAM J. Sci. Comput. **34**(1), A1–A27 (2012).

[99] S. A. Goreinov, I. V. Oseledets, D. V. Savostyanov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, How to find a good submatrix, in: Matrix Methods: Theory, Algorithms, Applications, (World Scientific, Hackensack, NY, 2010), pp. 247–256.
[100] S. A. Goreinov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, A theory of pseudoskeleton approximations, Linear Algebra Appl. **261**, 1–21 (1997).
[101] L. Grasedyck, Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure, Computing **72**(3–4), 247–265 (2004).
[102] L. Grasedyck, Hierarchical low rank approximation of tensors and multivariate functions, 2010, Lecture notes of Zurich summer school on "Sparse Tensor Discretizations of High-Dimensional Problems".
[103] L. Grasedyck, Hierarchical singular value decomposition of tensors, SIAM J. Matrix Anal. Appl. **31**(4), 2029–2054 (2010).
[104] L. Grasedyck, Polynomial approximation in hierarchical Tucker format by vector-tensorization, Tech. rep., RWTH Aachen, Germany, 2010.
[105] L. Grasedyck and W. Hackbusch, An introduction to hierarchical (H-) rank and TT-rank of tensors with examples, Comput. Meth. Appl. Math. **11**(3), 291–304 (2011).
[106] M. Griebel and H. Harbrecht, Approximation of two-variate functions: singular value decomposition versus regular sparse grids, Preprint 1109, INS, Universität Bonn, 2011.
[107] W. Hackbusch, Tensorisation of vectors and their efficient convolution, Numer. Math. **119**, 465–488 (2011).
[108] W. Hackbusch, Tensor Spaces and Numerical Tensor Calculus (Springer, 2012).
[109] W. Hackbusch, L^∞ estimation of tensor truncations, Numer. Math. (2013), To appear. Available at `http://www.mis.mpg.de/publications/preprints/2012/prepr2012-17.html`.
[110] W. Hackbusch and B. N. Khoromskij, Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators. I. Separable approximation of multi-variate functions, Computing **76**(3-4), 177–202 (2006).
[111] W. Hackbusch and B. N. Khoromskij, Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators. II. HKT representation of certain operators, Computing **76**(3-4), 203–225 (2006).
[112] W. Hackbusch and B. N. Khoromskij, Tensor-product approximation to operators and functions in high dimensions, J. Complexity **23**(4-6), 697–714 (2007).
[113] W. Hackbusch and B. N. Khoromskij, Tensor-product approximation to multidimensional integral operators and Green's functions, SIAM J. Matrix Anal. Appl. **30**(3), 1233–1253 (2008).
[114] W. Hackbusch, B. N. Khoromskij, S. A. Sauter, and E. E. Tyrtyshnikov, Use of tensor formats in elliptic eigenvalue problems, Numer. Linear Algebra Appl. **19**(1), 133–151 (2012).
[115] W. Hackbusch, B. N. Khoromskij, and E. E. Tyrtyshnikov, Hierarchical Kronecker tensor-product approximations, J. Numer. Math. **13**(2), 119–156 (2005).
[116] W. Hackbusch, B. N. Khoromskij, and E. E. Tyrtyshnikov, Approximate iterations for structured matrices, Numer. Math. **109**(3), 365–383 (2008).
[117] W. Hackbusch and S. Kühn, A new scheme for the tensor representation, J. Fourier Anal. Appl. **15**(5), 706–722 (2009).
[118] J. Haegeman, I. Cirac, T. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Time-dependent variational principle for quantum lattices, Phys. Rev. Lett. **107**(7), 070601 (2011).
[119] J. Haegeman, M. Mariën, T. J. Osborne, and F. Verstraete, Geometry of matrix product states: metric, parallel transport and curvature, arXiv preprint 1210.7710, 2012.
[120] N. Halko, P. G. Martinsson, and J. A. Tropp, Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions, SIAM Rev. **53**(2), 217–288 (2011).
[121] S. Handschuh, Changing the topology of tensor networks, Tech. Rep. 15, MPI MIS Leipzig, 2012.
[122] M. Hegland and J. Garcke, On the numerical solution of the chemical master equation with sums of rank one tensors, in: Proceedings of the 15th Biennial Computational Techniques and Applications Conference, CTAC-2010, edited by W. McLean and A. J. Roberts, ANZIAM J. Vol. 52 (August 2011), pp. C628–C643.
[123] S. Holtz, T. Rohwedder, and R. Schneider, On manifolds of tensors of fixed TT-rank, Numer. Math. **120**(4), 701–731 (2010).
[124] S. Holtz, T. Rohwedder, and R. Schneider, The alternating linear scheme for tensor optimization in the tensor train format, SIAM J. Sci. Comput. **34**(2), A683–A713 (2012).
[125] R. A. Horn and C. R. Johnson, Matrix analysis, second edition (Cambridge University Press, 2013).
[126] T. Huckle and K. Waldherr, Subspace iteration methods in terms of matrix product states, PAMM **12**(1), 641–642 (2012).
[127] T. Huckle, K. Waldherr, and T. Schulte-Herbrüggen, Computations in quantum tensor networks, Linear Algebra Appl. **438**(2), 750–781 (2013).
[128] T. Huckle, K. Waldherr, and T. Schulte-Herbrüggen, Exploiting matrix symmetries and physical symmetries in matrix product states and tensor trains, Linear and Multilinear Algebra **61**(1), 91–122 (2013).
[129] I. Ibragimov and S. Rjasanow, Three way decomposition for the Boltzmann equation, J. Comput. Math. **27**(2-3), 184–195 (2009).

[130] M. Ishteva, P. A. Absil, and P. Van Dooren, Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors, Tech. rep. UCL-INMA-2011.011, UC Louvain, Belgium, 2011.
[131] M. Ishteva, P. A. Absil, S. Van Huffel, and L. De Lathauwer, Best low multilinear rank approximation of higher-order tensors, based on the Riemannian trust-region scheme, SIAM J. Matrix Anal. Appl. **32**(1), 115–135 (2011).
[132] M. Ishteva, L. De Lathauwer, P. A. Absil, and S. Van Huffel, Differential-geometric Newton method for the best rank-(R_1, R_2, R_3) approximation of tensors, Numer. Algorithms **51**(2), 179–194 (2009).
[133] T. Jahnke and W. Huisinga, A dynamical low-rank approach to the chemical master equation, Bulletin of Mathematical Biology **70**, 2283–2302 (2008).
[134] E. Jondeau, E. Jurczenko, and M. Rockinger, Moment component analysis: An illustration with international stock markets, Swiss Finance Institute Research Paper No. 10-43, 2010, Available at `http://dx.doi.org/10.2139/ssrn.1694643`.
[135] V. Kazeev, M. Khammash, M. Nip, and C. Schwab, Direct solution of the chemical master equation using quantized tensor trains, Tech. Rep. 2013-04, Seminar for Applied Mathematics, ETH Zürich, 2013.
[136] V. Kazeev, B. N. Khoromskij, and E. E. Tyrtyshnikov, Multilevel Toeplitz matrices generated by QTT tensor-structured vectors and convolution with logarithmic complexity, Tech. Rep. 36, MPI MIS Leipzig, 2011.
[137] V. Kazeev, O. Reichmann, and C. Schwab, hp-DG-QTT solution of high-dimensional degenerate diffusion equations, Tech. Rep. 2012-11, Seminar for Applied Mathematics, ETH Zurich, 2012.
[138] V. Kazeev, O. Reichmann, and C. Schwab, Low-rank tensor structure of linear diffusion operators in the TT and QTT formats, Tech. Rep. 2012-13, Seminar for Applied Mathematics, ETH Zurich, 2012.
[139] V. A. Kazeev and B. N. Khoromskij, Low-rank explicit QTT representation of Laplace operator and its inverse, SIAM J. Matrix Anal. Appl. **33**(3), 742–758 (2012).
[140] V. A. Kazeev and E. E. Tyrtyshnikov, Structure of the Hessian matrix and an economical implementation of Newton's method in the problem of canonical approximation of tensors, Comput. Math. Math. Phys. **50**(6), 927–945 (2010).
[141] V. Khoromskaia, Computation of the Hartree-Fock exchange in the tensor-structured format., Comput. Meth. Appl. Math. **10**(2), 204–218 (2010).
[142] V. Khoromskaia, Numerical solution of the Hartree-Fock equation by multilevel tensor-structured methods, PhD thesis, Technische Universität Berlin, 2010.
[143] V. Khoromskaia, D. Andrae, and B. N. Khoromskij, Fast and accurate 3D tensor calculation of the Fock operator in a general basis, Comput. Phys. Commun. **183**(11), 2392–2404 (2012).
[144] V. Khoromskaia, B. N. Khoromskij, and R. Schneider, QTT representation of the Hartree and exchange operators in electronic structure calculations, Comput. Meth. Appl. Math. **11**(3), 327–341 (2011).
[145] V. Khoromskaia and B. Khoromskij, Møller-Plesset (MP2) energy correction using tensor factorizations of the grid-based two-electron integrals, Tech. Rep. 26, MPI MIS Leipzig, 2013.
[146] V. Khoromskaia, B. Khoromskij, and R. Schneider, Tensor-structured factorized calculation of two-electron integrals in a general basis, Tech. Rep. 29, MPI MIS Leipzig, 2012.
[147] B. N. Khoromskij, An introduction to structured tensor-product approximation of discrete nonlocal operators, 2005, Lecture notes No. 27, Leipzig.
[148] B. N. Khoromskij, Structured rank-$(r_1, \ldots, r_d)$ decomposition of function-related tensors in $\mathbb{R}^d$, Comput. Meth. Appl. Math. **6**(2), 194–220 (2006).
[149] B. N. Khoromskij, Structured data-sparse approximation to high order tensors arising from the deterministic Boltzmann equation, Math. Comp. **76**(259), 1291–1315 (2007).
[150] B. N. Khoromskij, On tensor approximation of Green iterations for Kohn-Sham equations, Comput. Vis. Sci. **11**, 259–271 (2008).
[151] B. N. Khoromskij, Tensor-structured preconditioners and approximate inverse of elliptic operators in $\mathbb{R}^d$, Constr. Approx. **30**, 599–620 (2009).
[152] B. N. Khoromskij, Fast and accurate tensor approximation of a multivariate convolution with linear scaling in dimension, J. Comput. Appl. Math. **234**(11), 3122–3139 (2010).
[153] B. N. Khoromskij, Introduction to tensor numerical methods in scientific computing, 2011, Lecture notes, available at `http://www.math.uzh.ch/fileadmin/math/preprints/06_11.pdf`.
[154] B. N. Khoromskij, $O(d \log N)$-quantics approximation of $N - d$ tensors in high-dimensional numerical modeling, Constr. Approx. **34**, 257–280 (2011).
[155] B. N. Khoromskij, Tensors-structured numerical methods in scientific computing: Survey on recent advances, Chemometrics and Intelligent Laboratory Systems **110**(1), 1–19 (2012).
[156] B. N. Khoromskij and V. Khoromskaia, Low rank Tucker-type tensor approximation to classical potentials, Central European Journal of Mathematics **5**, 523–550 (2007).
[157] B. N. Khoromskij and V. Khoromskaia, Multigrid accelerated tensor approximation of function related multidimensional arrays, SIAM J. Sci. Comput. **31**(4), 3002–3026 (2009).
[158] B. N. Khoromskij, V. Khoromskaia, S. R. Chinnamsetty, and H. J. Flad, Tensor decomposition in electronic structure calculations on 3D cartesian grids, J. Comput. Phys. **228**(16), 5749–5762 (2009).

[159] B. N. Khoromskij, V. Khoromskaia, and H. J. Flad, Numerical solution of the Hartree–Fock equation in multilevel tensor-structured format, SIAM J. Sci. Comput. **33**(1), 45–65 (2011).
[160] B. N. Khoromskij and I. V. Oseledets, DMRG+QTT approach to computation of the ground state for the molecular Schrödinger operator, Tech. Rep. 69, MPI MIS Leipzig, 2010.
[161] B. N. Khoromskij and I. V. Oseledets, Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDEs, Comput. Meth. Appl. Math. **10**(4), 376–394 (2010).
[162] B. N. Khoromskij and I. V. Oseledets, QTT approximation of elliptic solution operators in higher dimensions, Rus. J. Numer. Anal. Math. Model. **26**(3), 303–322 (2011).
[163] B. N. Khoromskij, I. V. Oseledets, and R. Schneider, Efficient time-stepping scheme for dynamics on TT-manifolds, Tech. Rep. 24, MPI MIS Leipzig, 2012.
[164] B. N. Khoromskij, S. A. Sauter, and A. Veit, Fast quadrature techniques for retarded potentials based on TT/QTT tensor approximation, Comput. Meth. Appl. Math. **11**(3), 342–362 (2011).
[165] B. N. Khoromskij and C. Schwab, Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs, SIAM J. Sci. Comput. **33**(1), 364–385 (2011).
[166] O. Koch and C. Lubich, Dynamical low-rank approximation, SIAM J. Matrix Anal. Appl. **29**(2), 434–454 (2007).
[167] O. Koch and C. Lubich, Dynamical tensor approximation, SIAM J. Matrix Anal. Appl. **31**(5), 2360–2375 (2010).
[168] T. G. Kolda and B. W. Bader, Tensor decompositions and applications, SIAM Review **51**(3), 455–500 (2009).
[169] D. Kressner, M. Plešinger, and C. Tobler, A preconditioned low-rank CG method for parameter-dependent Lyapunov matrix equations, Tech. rep., MATHICSE, EPF Lausanne, Switzerland, 2012, Available at `http://anchp.epfl.ch`.
[170] D. Kressner and C. Tobler, Krylov subspace methods for linear systems with tensor product structure, SIAM J. Matrix Anal. Appl. **31**(4), 1688–1714 (2010).
[171] D. Kressner and C. Tobler, Low-rank tensor Krylov subspace methods for parametrized linear systems, SIAM J. Matrix Anal. Appl. **32**(4), 1288–1316 (2011).
[172] D. Kressner and C. Tobler, Preconditioned low-rank methods for high-dimensional elliptic PDE eigenvalue problems, Comput. Meth. Appl. Math. **11**(3), 363–381 (2011).
[173] D. Kressner and C. Tobler, `htucker` – a MATLAB toolbox for tensors in hierarchical Tucker format, Tech. rep., MATHICSE, EPF Lausanne, Switzerland, 2012, Available at `http://anchp.epfl.ch/htucker`.
[174] P. M. Kroonenberg, Applied Multiway Data Analysis (Wiley, 2008).
[175] S. Kühn, Hierarchische Tensordarstellung, Dissertation, Universität Leipzig, 2012.
[176] H. Lamari, A. Ammar, P. Cartraud, G. Legrain, F. Chinesta, and F. Jacquemin, Routes for efficient computational homogenization of nonlinear materials using the proper generalized decompositions, Arch. Comput. Methods Eng. **17**, 373–391 (2010).
[177] J. M. Landsberg, Tensors: geometry and applications, Graduate Studies in Mathematics, Vol. 128 (American Mathematical Society, Providence, RI, 2012).
[178] J. M. Landsberg, Y. Qi, and K. Ye, On the geometry of tensor network states, Quantum Inf. Comput. **12**(3-4), 346–354 (2012).
[179] C. Le Bris, T. Lelièvre, and Y. Maday, Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations, Constr. Approx. **30**(3), 621–651 (2009).
[180] O. S. Lebedeva, Block tensor conjugate gradient-type method for Rayleigh quotient minimization in two-dimensional case, Comput. Math. Math. Phys. **50**(5), 749–765 (2010).
[181] O. S. Lebedeva, Tensor conjugate-gradient-type method for Rayleigh quotient minimization in block QTT-format, Russian J. Numer. Anal. Math. Modelling **26**(5), 465–489 (2011).
[182] J. Liu, P. Musialski, P. Wonka, and J. Ye, Tensor completion for estimating missing values in visual data, in: Computer Vision, 2009 IEEE 12th International Conference on, (29 2009-oct. 2 2009), pp. 2114–2121.
[183] C. Lubich, On variational approximations in quantum molecular dynamics, Math. Comp. **74**(250), 765–779 (2005).
[184] C. Lubich, From quantum to classical molecular dynamics: reduced models and numerical analysis, Zurich Lectures in Advanced Mathematics (European Mathematical Society (EMS), Zürich, 2008).
[185] C. Lubich and I. V. Oseledets, A projector-splitting integrator for dynamical low-rank approximation, arXiv preprint 1301.5512, 2013.
[186] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken, Dynamical approximation of hierarchical Tucker and tensor-train tensors, Tech. rep., July 2012.
[187] T. Mach, Computing inner eigenvalues of matrices in tensor train matrix format, Tech. Rep. MPIMD/11-09, MPI Magdeburg, 2011, Available at `http://www.mpi-magdeburg.mpg.de/preprints/2011/09/`.
[188] T. Mach and J. Saak, Towards an ADI iteration for tensor structured equations, Tech. rep., MPI Magdeburg, 2012, Available at `www.mpi-magdeburg.mpg.de/preprints/2011/12/`.

[189] H. G. Matthies and E. Zander, Solving stochastic systems with low-rank tensor compression, Linear Algebra Appl. **436**(10), 3819–3838 (2012).
[190] P. Meszmer and J. Ballani, Tensor structured evaluation of singular volume integrals, Tech. Rep. 70, MPI MIS Leipzig, 2012.
[191] G. Meyer, S. Bonnabel, and R. Sepulchre, Regression on fixed-rank positive semidefinite matrices: a Riemannian approach, J. Mach. Learn. Res. **12**, 593–625 (2011).
[192] H. D. Meyer, F. Gatti, and G. A. Worth (eds.), Multidimensional Quantum Dynamics: MCTDH Theory and Applications (Wiley-VCH Verlag GmbH & Co. KGaA, 2009).
[193] H. D. Meyer, U. Manthe, and L. Cederbaum, The multi-configurational time-dependent hartree approach, Chemical Physics Letters **165**(1), 73 – 78 (1990).
[194] M. J. Mohlenkamp, Musings on multilinear fitting, Linear Algebra Appl. **438**(2), 834–852 (2013).
[195] M. J. Mohlenkamp and L. Monzón, Trigonometric identities and sums of separable functions, Math. Intelligencer **27**(2), 65–69 (2005).
[196] B. Mokdad, A. Ammar, M. Normandin, F. Chinesta, and J. R. Clermont, A fully deterministic micro-macro simulation of complex flows involving reversible network fluid models, Math. Comput. Simulation **80**(9), 1936–1961 (2010).
[197] K. K. Naraparaju and J. Schneider, Generalized cross approximation for 3d-tensors, Comput. Vis. Sci. **14**(3), 105–115 (2011).
[198] T. Ngo and Y. Saad, Scaled gradients on Grassmann manifolds for matrix completion, Preprint ys-2012-5, Dept. Computer Science and Engineering, University of Minnesota, Minneapolis, MN, 2012.
[199] A. Nonnenmacher and C. Lubich, Dynamical low-rank approximation: applications and numerical experiments, Mathematics and Computers in Simulation **79**(4), 1346–1357 (2008).
[200] A. Nouy, A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, Comput. Meth. Appl. Mech. Engrg. **196**(45–48), 4521–4537 (2007).
[201] A. Nouy, Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms, Comput. Meth. Appl. Mech. Engrg. **197**(51–52), 4718–4736 (2008).
[202] A. Nouy, Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems, Arch. Comput. Methods Eng. **17**, 403–434 (2010).
[203] V. Olshevsky, I. V. Oseledets, and E. E. Tyrtyshnikov, Tensor properties of multilevel Toeplitz and related matrices, Linear Algebra Appl. **412**(1), 1–21 (2006).
[204] I. Oseledets, S. Dolgov, and D. Savostyanov, ttpy, 2013, Available at `https://github.com/oseledets/ttpy/`.
[205] I. V. Oseledets, Approximation of $2^d \times 2^d$ matrices using tensor decomposition, SIAM J. Matrix Anal. Appl. **31**(4), 2130–2145 (2010).
[206] I. V. Oseledets, DMRG approach to fast linear algebra in the TT–format, Comput. Meth. Appl. Math **11**(3), 382–393 (2011).
[207] I. V. Oseledets, MATLAB TT-Toolbox Version 2.2, May 2011, Available at `http://spring.inm.ras.ru/osel/?page_id=24`.
[208] I. V. Oseledets, Constructive representation of functions in low-rank tensor formats, Constr. Appr. **37**(1), 1–18 (2013).
[209] I. V. Oseledets, D. V. Savostianov, and E. E. Tyrtyshnikov, Tucker dimensionality reduction of three-dimensional arrays in linear time, SIAM J. Matrix Anal. Appl. **30**(3), 939–956 (2008).
[210] I. V. Oseledets, D. V. Savostyanov, and E. E. Tyrtyshnikov, Linear algebra for tensor problems, Computing **85**(3), 169–188 (2009).
[211] I. V. Oseledets, D. V. Savostyanov, and E. E. Tyrtyshnikov, Cross approximation in tensor electron density computations, Numer. Linear Algebra Appl. **17**(6), 935–952 (2010).
[212] I. V. Oseledets and E. E. Tyrtyshnikov, Breaking the curse of dimensionality, or how to use SVD in many dimensions, SIAM J. Sci. Comput. **31**(5), 3744–3759 (2009).
[213] I. V. Oseledets and E. E. Tyrtyshnikov, Tensor tree decomposition does not need a tree, Preprint (Submitted to Linear Algebra Appl) 2009-04, INM RAS, Moscow, 2009.
[214] I. V. Oseledets and E. E. Tyrtyshnikov, TT-cross approximation for multidimensional arrays, Linear Algebra Appl. **432**(1), 70–88 (2010).
[215] I. V. Oseledets and E. E. Tyrtyshnikov, Algebraic wavelet transform via quantics tensor train decomposition, SIAM J. Sci. Comput. **33**(3), 1315–1328 (2011).
[216] I. V. Oseledets, E. E. Tyrtyshnikov, and N. L. Zamarashkin, Tensor-train ranks of matrices and their inverses, Comput. Meth. Appl. Math **11**(3), 394–403 (2011).
[217] I. Oseledets, Tensor-train decomposition, SIAM J. Sci. Comput. **33**(5), 2295–2317 (2011).
[218] S. Östlund and S. Rommer, Thermodynamic limit of density matrix renormalization, Phys. Rev. Lett. **75**(Nov), 3537–3540 (1995).
[219] V. Y. Pan, G. Qian, and A. L. Zheng, Randomized matrix computations, arXiv preprint 1210.7476, 2012.

[220] R. Penrose, Applications of negative dimensional tensors, in: Combinatorial Mathematics and its Applications (Proc. Conf., Oxford, 1969), (Academic Press, London, 1971), pp. 221–244.
[221] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations, Quantum Info. Comput. **7**(5), 401–430 (2007).
[222] A. H. Phan, P. Tichavský, and A. Cichocki, CANDECOMP/PARAFAC decomposition of high-order tensors through tensor reshaping, arXiv preprint 1211.3796, 2012.
[223] A. H. Phan, P. Tichavský, and A. Cichocki, On fast computation of gradients for CANDECOMP/PARAFAC algorithms, arXiv preprint 1204.1586, 2012.
[224] A. H. Phan, P. Tichavský, and A. Cichocki, Low complexity damped gauss–newton algorithms for CANDECOMP/PARAFAC, SIAM Journal on Matrix Analysis and Applications **34**(1), 126–147 (2013).
[225] M. Rajih, P. Comon, and R. A. Harshman, Enhanced line search: a novel method to accelerate PARAFAC, SIAM J. Matrix Anal. Appl. **30**(3), 1128–1147 (2008).
[226] T. Rohwedder and A. Uschmajew, Local convergence of alternating schemes for optimization of convex problems in the TT format, Tech. rep., 2011.
[227] M. Sanz, M. M. Wolf, D. Pérez-García, and J. I. Cirac, Matrix product states: Symmetries and two-body Hamiltonians, Phys. Rev. A **79**(Apr), 042308 (2009).
[228] B. Savas and L. Eldén, Krylov-type methods for tensor computations I, Linear Algebra Appl. **438**(2), 891–918 (2013).
[229] B. Savas and L. H. Lim, Quasi-Newton methods on Grassmannians and multilinear approximations of tensors, SIAM J. Sci. Comput. **32**(6), 3352–3393 (2010).
[230] D. V. Savostyanov, QTT-rank-one vectors with QTT-rank-one and full-rank Fourier images, Linear Algebra Appl. **436**(9), 3215–3224 (2012).
[231] D. V. Savostyanov and I. V. Oseledets, Fast adaptive interpolation of multi-dimensional arrays in tensor train format, in: Proceedings of 7th International Workshop on Multidimensional Systems (nDS), (IEEE, 2011).
[232] D. V. Savostyanov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, Fast truncation of mode ranks for bilinear tensor operations, Numer. Linear Algebra Appl. **19**(1), 103–111 (2012).
[233] R. Schneider, Tensor product approximation and numerical solution of the electronic Schrödinger equation, 2012, John von Neumann Gastvorlesung, available at `http://www-m2.ma.tum.de/bin/view/Allgemeines/JohnVonNeumannLectureSS12`.
[234] R. Schneider and A. Uschmajew, Approximation rates for the hierarchical tensor format in certain periodic Sobolev spaces, Tech. rep., MATHICSE, EPFL, 2013.
[235] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Annals of Physics **326** (2011).
[236] S. Sharma and G. K. L. Chan, Spin-adapted density matrix renormalization group algorithms for quantum chemistry, J. Chem. Phys. **136**(12), 124121 (2012).
[237] S. Sharma and G. K. L. Chan, Block v 0.9.6, 2013, Available at `http://www.princeton.edu/chemistry/chan/software/dmrg/`.
[238] Y. Y. Shi, L. M. Duan, and G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network, Physical Review A **74**(2), 022320 (2006).
[239] A. Smilde, R. Bro, and P. Geladi, Multi-way Analysis: Applications in the Chemical Sciences (Wiley, 2004).
[240] L. Sorber, M. Van Barel, and L. De Lathauwer, Tensorlab v1.0, 2013, Available online, `http://esat.kuleuven.be/sista/tensorlab/`.
[241] A. Stegeman and P. Comon, Subtracting a best rank-1 approximation may increase tensor rank, Linear Algebra Appl. **433**(7), 1276–1300 (2010).
[242] N. Takaaki, Algebraic reconstruction of the general-order poles of a meromorphic function, Inverse Problems **28**(2), 025008 (2012).
[243] V. N. Temlyakov, Approximations of functions with bounded mixed derivative, Trudy Mat. Inst. Steklov. **178**, 113 (1986), In Russian; English translation in Proc. Steklov Inst. Math. 1989, no. 1 (178).
[244] V. N. Temlyakov, Bilinear approximation and related questions, Trudy Mat. Inst. Steklov pp. 229–248 (1992), In Russian; English translation: Proc. Steklov Inst. Math. 1993, no. 4 (194), 245–265.
[245] V. N. Temlyakov, Estimates for the best bilinear approximations of functions and approximation numbers of integral operators, Mat. Zametki **51**(5), 125–134, 159 (1992), In Russian; English translation in Math. Notes 1992, 51(5), 510–517.
[246] C. Tobler, Low-rank Tensor Methods for Linear Systems and Eigenvalue Problems, PhD thesis, ETH Zurich, Switzerland, 2012.
[247] C. E. Tsourakakis, MACH: fast randomized tensor decompositions, in: SIAM International Conference on Data Mining, (2010).
[248] E. E. Tyrtyshnikov, Incomplete cross approximation in the mosaic-skeleton method, Computing **64**(4), 367–380 (2000), International GAMM-Workshop on Multigrid Methods (Bonn, 1998).
[249] E. E. Tyrtyshnikov, Tensor approximations of matrices generated by asymptotically smooth functions, Mat. Sb. **194**(6), 147–160 (2003).

[250] E. E. Tyrtyshnikov, Kronecker-product approximations for some function-related matrices, Linear Algebra Appl. **379**, 423–437 (2004), Tenth Conference of the International Linear Algebra Society.
[251] E. Ullmann, A Kronecker product preconditioner for stochastic Galerkin finite element discretizations, SIAM J. Sci. Comput. **32**(2), 923–946 (2010).
[252] A. Uschmajew, Regularity of tensor product approximations tosquare integrable functions, Constr. Approx. **34**, 371–391 (2011).
[253] A. Uschmajew, Local convergence of the alternating least squares algorithm for canonical tensor approximation, SIAM J. Matrix Anal. Appl. **33**(2), 639–652 (2012).
[254] A. Uschmajew, Zur Theorie der Niedrigrangapproximation in Tensorprodukten von Hilberträumen, PhD thesis, Technische Universität Berlin, 2013.
[255] A. Uschmajew and B. Vandereycken, The geometry of algorithms using hierarchical tensors, Tech. rep., January 2012.
[256] C. F. Van Loan and N. Pitsianis, Approximation with Kronecker products, in: Linear algebra for large scale and real-time applications (Leuven, 1992), , NATO Adv. Sci. Inst. Ser. E Appl. Sci., Vol. 232 (Kluwer Acad. Publ., Dordrecht, 1993), pp. 293–314.
[257] B. Vandereycken, Low-rank matrix completion by Riemannian optimization, Tech. rep., MATHICSE, EPF Lausanne, Switzerland, 2012.
[258] B. Vandereycken and S. Vandewalle, A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations, SIAM J. Matrix Anal. Appl. **31**(5), 2553–2579 (2010).
[259] N. Vannieuwenhoven, R. Vandebril, and K. Meerbergen, A new truncation strategy for the higher-order singular value decomposition, SIAM J. Sci. Comput. **34**(2), A1027–A1052 (2012).
[260] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, Tech. Rep. arXiv:cond-mat/0407066, 2004.
[261] F. Verstraete and J. I. Cirac, Valence-bond states for quantum computation, Phys. Rev. A **70**(Dec), 060302 (2004).
[262] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix product density operators: Simulation of finite-temperature and dissipative systems, Phys. Rev. Lett. **93**(Nov), 207204 (2004).
[263] G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, Phys. Rev. Lett. **91**, 147902 (2003).
[264] G. Vidal, Entanglement renormalization, Phys. Rev. Lett. **99**(Nov), 220405 (2007).
[265] H. Wang and M. Thoss, Multilayer formulation of the multiconfiguration time-dependent Hartree theory, J. Chem. Phys **119**(3), 1289–1299 (2003).
[266] S. R. White, Density matrix formulation for quantum renormalization groups, Phys. Rev. Lett. **69**, 2863–2866 (1992).
[267] B. M. Wise and N. B. Gallagher, PLS Toolbox 7.0.3, 2013, Available at `http://www.eigenvector.com`.
[268] G. A. Worth, M. H. Beck, A. Jäckle, and H. D. Meyer, The MCTDH package, version 8.3 (2003), See `http://www.pci.uni-heidelberg.de/tc/usr/mctdh/`.
[269] G. Zhou, A. Cichocki, and S. Xie, Accelerated canonical polyadic decomposition by using mode reduction, arXiv preprint 1211.3500, 2012.
[270] M. Zwolak and G. Vidal, Mixed-state dynamics in one-dimensional quantum lattice systems: A time-dependent superoperator renormalization algorithm, Phys. Rev. Lett. **93**(Nov), 207205 (2004).

Recent publications :

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

43.2012 M. ASTORINO, F. CHOULY, A. QUARTERONI:
Multiscale coupling of finite element and lattice Boltzmann methods for time dependent problems

44.2012 D. KRESSNER, X. LIU:
Structured canonical forms for products of (skew-)symmetric matrices and the matrix equation XAX = B

45.2012 B. BECKERMANN, D. KRESSNER, C. TOBLER:
An error analysis of Galerkin projection methods for linear systems with tensor product structure

46.2012 J. BECK, F. NOBILE, L. TAMELLINI, R. TEMPONE:
A quasi-optimal sparse grids procedure for groundwater flows

47.2012 A. ABDULLE, Y. BAI:
Fully discrete analysis of the heterogeneous multiscale method for elliptic problems with multiple scales

48.2012 G. MIGLORATI, F. NOBILE, E. VON SCHWERIN, R. TEMPONE:
Approximation of quantities of interest in stochastic PDES by the random discrete L^2 projection on polynomial spaces

01.2013 A. ABDULLE, A. BLUMENTHAL:
Stabilized multilevel Monte Carlo method for stiff stochastic differential equations

02.2013 D. N. ARNOLD, D. BOFFI, F. BONIZZONI:
Tensor product finite element differential forms and their approximation properties

03.2013 N. GUGLIELMI, D. KRESSNER, C. LUBICH:
Low-rank differential equations for Hamiltonian matrix nearness problems

04.2013 P. CHEN, A. QUARTERONI, G. ROZZA:
A weighted reduced basis method for elliptic partial differential equations with random input data

05.2013 P. CHEN, A. QUARTERONI, G. ROZZA:
A weighted empirical interpolation method: a priori convergence analysis and applications

06.2013 R. SCHNEIDER, A. USCHMAJEW:
Approximation rates for the hierarchical tensor format in periodic Sobolev spaces

07.2013 C. BAYER, H. HOEL, E. VON SCHWERIN, R. TEMPONE:
On non-asymptotic optimal stopping criteria in Monte Carlo simulation.

08.2013 L. GRASEDYCK, D. KRESSNER, C. TOBLER:
A literature survey of low-rank tensor approximation techniques