Tensor-Train ranks for matrices and their inverses

Ivan Oseledets [1], Eugene Tyrtyshnikov, [1,2] Nickolai Zamarashkin [1]

Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkin Street, 8, Moscow 119333

Abstract

We show that the recent tensor-train (TT) decompositions of matrices come up from its recursive Kronecker-product representations with a systematic use of common bases. The names TTM and QTT used in this case stress the relation with multilevel matrices or quantization that increases artificially the number of levels. Then we investigate how the tensor-train ranks of a matrix can be related with those of its inverse. In the case of a banded Toeplitz matrix, we prove that the tensor-train ranks of its inverse are bounded from above by $1 + (l + u)^2$, where l and u are the bandwidths in the lower and upper parts of the matrix without the main diagonal.

AMS classification: 15A12; 65F10; 65F15

Key words: Tensor ranks; Tensor-Train Decomposition; QTT-ranks; Inverse matrices; Multilevel matrices; Toeplitz matrices; Banded matrices.

1 Introduction

A sound idea of using the separation of variables or low-rank constructions to compactly represent a general matrix *should not* be applied straight-

Email addresses: `ivan@bach.inm.ras.ru` (Ivan Oseledets), `tee@inm.ras.ru` (Eugene Tyrtyshnikov,), `kolya@bach.inm.ras.ru` (Nickolai Zamarashkin).

[1] Supported by the RFBR grants 11-01-00549-a, 09-01-12058, 09-01-91332 (joint with DFG), the Government Contracts П940, П1112, 14.740.11.0345 and Priority Research Grant of the Presidium and of the Department of Mathematical Sciences of the Russian Academy of Sciences.

[2] During this work the author was a visiting professor at the University of Siedlce (Poland) and at the Hausdorff Research Institute for Mathematics in Bonn (Germany).

forwardly, i.e. in the form of a low-rank dyadic (skeleton) decomposition of a given matrix. That cannot help if we deal with a nonsingular matrix. However, this idea works superbly when we consider Kronecker-product representations.

If A is a matrix, then we denote its entries by $A(i,j)$. We write $A = G^1 \otimes G^2$ when A is a block matrix of the form $A = [G^1(i_1,j_1)G^2]$. As a consequence, the sizes of A are the products of the corresponding sizes of G^1 and G^2. In this case A is called the Kronecker product of G^1 and G^2.

The Kronecker product separates *some variables.* In order to see *which variables*, consider the row and column indices of A as multi-indices

$$i = i_1 i_1, \quad j = j_1 j_2,$$

then

$$A(i_1 i_2 \, , \, j_1 j_2) = G^1(i_1,j_1) G^2(i_2,j_2).$$

We prefer to depict multi-indices ignoring brackets and commas. The components of a mutli-index are called *mode-indices.* Thus, in the Kronecker product, the original variables i and j are not separated. Instead, we consider these indices as products of mode-indices, mix them in a special way and create new multi-indices which occur to admit separation. If we *reshape* A into a new matrix $\tilde{A}$ so that (cf. [22])

$$\tilde{A}(i_1 j_1 \, , \, i_2 j_2) = A(i_1 i_2 \, , \, j_1, j_2),$$

then A is a Kronecker product of two matrices if and only if $\tilde{A}$ is a matrix of rank at most 1.

It is seldom, all the same, that a matrix is represented by a single Kronecker product. A much wider option is when A is a sum of a few Kronecker products as follows:

$$A = \sum_{\alpha=1}^{r} G^1_\alpha \otimes G^2_\alpha, \tag{1}$$

r is referred to as a *Kronecker rank of the representation*; therefore, the minimality is not required. If r is minimal possible value for all representations of the form (1) then it is called *Kronecker rank of the matrix A*; note the dependence on the sizes chosen for the mode-indices.

Lemma 1.1 *Given a decomposition (1), suppose that A is of size $m \times n$ with $m = m_1 m_2$ and $n = n_1 n_2$, and G^1_α and G^2_α are $m_1 \times n_1$ and $m_2 \times n_2$, respectively. Then A can be viewed as a block matrix*

$$A = [A_{i_1 j_1}], \quad 1 \le i_1 \le m_1, \quad 1 \le j_1 \le n_1,$$

where the dimension of linear span of the blocks $A_{i_1 j_1}$ does not exceed r, and is equal to r if and only if r is the Kronecker rank of A.

Proof. It is sufficient to note that each block $A_{i_1 j_1}$ is an r-term linear combination of the blocks G^2_α. $\square$

All the matrices in the right-hand side of (1) contain $r(m_1 n_1 + m_2 n_2)$ entries. This can be notably less than n^2, the total number of entries in the matrix A, and thus, the ansatz of (1) can be useful as a compact (compressed) representation of A. Further compression is envisaged when we use Kronecker products with more than two factors:

$$A = \sum_{\alpha=1}^{r} \bigotimes_{s=1}^{d} G^s_\alpha. \tag{2}$$

In the pursuit of better compression we are interested to take the sizes of G^s_α as small as possible. Hence, d has to be maximized. If $m = n = 2^d$ and all G^s_α are of size 2×2, then the right-hand side of (2) is determined by $4r \log_2 n$ parameters, and in case r grows slowly as a functon of n, it is dramatically less than n^2.

An actual advantage depends, of course, on r. Anyway, we hope to enjoy modest values of r at least in some applications, and hence, we might be very interested in increasing the number of factors in the Kronecker products. Formula (2) defines the so called *canonical polyadic* (CP) decomposition of A. Some standard tensor decompositions and their applications are lately surveyed in the work [1], numerical calculus in higher dimensions is discussed in [2,4].

A big problem with exploiting CP in practice is the lack of reliable and fast algorithms in the case $d \geq 3$. It is opposed to the case $d = 2$ which is easily treated, e.g. by application of the SVD. Nevertheless, one may naturally employ an idea of reducing a d-factor case to those with two factors. For example, this was behind numerical examples in [21].

Consider the case $d = 3$ and assume that A is $m \times n$ with $m = m_1 m_2 m_3$ and $n = n_1 n_2 n_3$. The whole enterprise consists in three steps. First, we treat A as a block matrix with $m_1 \times n_1$ block entries and find a two-factor decomposition

$$A = \sum_{\alpha_1=1}^{r_1} G^1_{\alpha_1} \otimes A^1_{\alpha_1}.$$

It can be found even with minimal possible number of summands. Second, we find two-factor decompositions for the smaller matrices $A^1_{\alpha_1}$ which are viewed as block matrices with $m_2 \times n_2$ block entries:

$$A^1_{\alpha_1} = \sum_{\alpha_2=1}^{r_2} G^2_{\alpha_1\alpha_2} \otimes G^3_{\alpha_1\alpha_2}.$$

Note that r_2 may depend on α_1, here we simply take the maximal value for different α_1. Third, by putting the former and latter together we obtain

$$A = \sum_{\alpha_1=1}^{r_1} \sum_{\alpha_2=1}^{r_2} G^1_{\alpha_1} \otimes G^2_{\alpha_1\alpha_2} \otimes G^3_{\alpha_1\alpha_2}.$$

In the general case we assume that

$$m = \prod_{s=1}^{d} m_s, \quad n = \prod_{s=1}^{d} n_s, \tag{3}$$

then we need d steps and finish with a decomposition

$$A = \sum_{\alpha_1=1}^{r_1} \dots \sum_{\alpha_{d-1}=1}^{r_{d-1}} G^1_{\alpha_1} \otimes G^2_{\alpha_1\alpha_2} \otimes \dots \otimes G^{d-1}_{\alpha_1\dots\alpha_{d-1}} \otimes G^d_{\alpha_1\dots\alpha_{d-1}}.$$

The result, however, does not bring us permanent satisfaction: although we have obtained some CP decomposition, it contains too much a reduntant total number of the involved parameters. And with increasing the number of factors this redundancy tends to grow even significantly, exponentially in d.

A kind of remedy is still easily available. If $d = 2^p$, then A can be viewed as a block matrix with $(m_1 \dots m_{d/2}) \times (n_1 \dots n_{d/2})$ block entries. Then, we begin with

$$A = \sum_{\alpha_1=1}^{r_1} A^1_{\alpha_1} \otimes A^2_{\alpha_1},$$

where, in the next step, $A^1_{\alpha_1}$ are treated as block matrices with $(m_1 \dots m_{d/4}) \times (n_1 \dots n_{d/4})$ block entries and $A^2_{\alpha_1}$ are processed similarly as block matrices with $(m_{d/2+1} \dots m_{d/2+d/4}) \times (n_{d/2+1} \dots n_{d/2+d/4})$ block entries. Eventually we obtain some CP decomposition, where the number of summation parameters in the factors does not exceed $\log_2 d$. If $r = \max r_s$, then the total number of involved parameters amounts to $r^{\log_2 d}$; previously it was r^{d-1}. Note, all the same, that the values of r_s are not the same as before. But, if r is sufficiently small then we may think of this new strategy as of substantial improvement, because the total number of representation parameters now grows in d polynomially. Note that $r^{\log_2 d} = d^{\log_2 r}$.

In the next section we show how the above constructions can be modified to yield a representation with the number of parameters depending on d just linearly. Strictly speaking, this is true only when the values of r_s arising in the construction do not depend on d. Whatever important, the study of how these quantities behave in various situations should be postponed until after the search for a potentially good ansatz is completed. In the result, we obtain the recent tensor-train representations in the case of matrices [14,15,16,17,18,19]. A somewhat new thing is the approach: with Kronecker products we do not encounter an issue of how should we mix the indices and the cases of vectors and matrices do not differ.

Note that recursive approaches as a way to compact representations of multi-index arrays were first proposed in [6,14,15]. Some algorithms with hierarchical structure are then derived in [3]. However, a recursive reduction of dimensionality naturally leads to tensor trains proving to be most convenient for the design of efficient algorithms; see [16,19].

As soon as the tensor trains for matrices are expounded, we proceed to the study of relations between the tensor-train ranks of matrices and their inverses. In the case of a banded Toeplitz matrix, we prove a result announced in [23]: the tensor-train ranks of the inverse are bounded from above by $1 + (l+u)^2$, where l and u are the bandwidths in the lower and upper parts of the matrix without the main diagonal.

2 Tensor trains for matrices

In the first place, we should realize that CP is not a must if we are interested just in better compression schemes for A. Nevertheless, let us elaborate on the above naive approach with some CP as a purpose (maybe a wrong one) and do not refuse of it too hastily.

Let us begin with $d = 3$. We do not change the first step, as before it yields

$$A = \sum_{\alpha_1=1}^{r_1} G^1_{\alpha_1} \otimes A^1_{\alpha_1}.$$

The second step, however, should be revisited. Previously the two-factor decompositions for each of smaller matrices $A^1_{\alpha_1}$ were looked for *independently*:

$$A^1_{\alpha_1} = \sum_{\alpha_2=1}^{r_2} G^2_{\alpha_1\alpha_2} \otimes G^3_{\alpha_1\alpha_2}.$$

Now let us try to do the same *simultaneously.*

The matrices $A^1_{\alpha_1}$ are block matrices with $m_2 \times n_2$ block entries, denote them by $A^1_{\alpha_1, i_2 j_2}$. According to Lemma 1.1, for any fixed value of α_1 each of these blocks is a linear combination of the blocks of $G^3_{\alpha_1\alpha_2}$. Consider the span

$$\mathcal{L} = \mathrm{span}\{A^1_{\alpha_1, i_2 j_2}, \quad 1 \le \alpha_1 \le r_1, \quad 1 \le i_2 \le m_2, \quad 1 \le j_2 \le n_2\}$$

of *all blocks* of $A^1_{\alpha_1}$ for all values of α_1 and construct a basis of $\mathcal{L}$. Let it define the matrices $G^3_{\alpha_2}$, where

$$1 \le \alpha_2 \le r_2 = \dim \mathcal{L}.$$

Consequently, every block $A^1_{\alpha_1, i_2 j_2}$ is a linear combination of the common basis matrices $G^3_{\alpha_2}$. Denoting the coefficients by $G^2_{\alpha_1\alpha_2}(i_2, j_2)$, we get

$$A^1_{\alpha_1, i_2 j_2} = \sum_{\alpha_2=1}^{r_2} G^2_{\alpha_1\alpha_2}(i_2, j_2)\, G^3_{\alpha_2}.$$

Hence, by the definition of Kronecker product,

$$A^1_{\alpha_1} = \sum_{\alpha_2=1}^{r_2} G^2_{\alpha_1\alpha_2} \otimes G^3_{\alpha_2}, \quad G^2_{\alpha_1\alpha_2} = [G^2_{\alpha_1\alpha_2}(i_2 j_2)].$$

Finally,

$$A = \sum_{\alpha_1=1}^{r_1} \sum_{\alpha_2=1}^{r_2} G^1_{\alpha_1} \otimes G^2_{\alpha_1\alpha_2} \otimes G^3_{\alpha_2}. \tag{4}$$

We call (4) a *tensor-train decomposition of matrix* A, in the case of three levels of its block partitionings. For the definition and machinery of multilevel matrices see [11,12]. Recursive application of this approach in the case of d levels gives

$$A = \sum_{\alpha_1=1}^{r_1} \dots \sum_{\alpha_{d-1}=1}^{r_{d-1}} G^1_{\alpha_1} \otimes G^2_{\alpha_1\alpha_2} \otimes \dots \otimes G^{d-1}_{\alpha_{d-2}\alpha_{d-1}} \otimes G^d_{\alpha_{d-1}}. \tag{5}$$

For brevity of notation it is convenient to introduce additional indices

$$\alpha_0 = \alpha_d = 1$$

and present the same decomposition in the form

$$A = \sum_{\substack{1 \le \alpha_s \le r_s \\ s=1,\dots,d-1}} \bigotimes_{k=1}^{d} G^k_{\alpha_{k-1}\alpha_k}. \tag{6}$$

The right-hand side of (5) or (6) is a *tensor train for a d-level matrix.* The quantities r_s are called *tensor-train (TT) ranks.*

Note that *any* matrix A can be viewed as a d-level matrix with sizes being products of smaller sizes according to (3): we can set $m_s = 1$ or $n_s = 1$ for some values of s. In particular, if $n = 1$ then A is a column matrix (vector), and its tensor-train decomposition is one of a vector. In such a case, a d-level structure allows one to consider A as a d-index array with the entries $A(i_1, \dots, i_d)$. The TT-decomposition originally defined for this case [16] looks as follows:

$$A(i_1, \ldots, i_d) = \sum_{\alpha_1, \ldots, \alpha_{d-1}} \prod_{k=1}^{d} g_k(\alpha_{k-1}, i_k, \alpha_k).$$

Introducing the vectors $G^k_{\alpha_{k-1}\alpha_k}$ as ones with the entries $g_k(\alpha_{k-1}, i_k, \alpha_k)$, we arrive at the Kronecker-product expression

$$A = \sum_{\alpha_1, \ldots, \alpha_{d-1}} \bigotimes_{k=1}^{d} G^k_{\alpha_{k-1}\alpha_k},$$

which exactly coincides with (6). A possible alternative is a *matrix-product representation* for the entries of A:

$$A(i_1, \ldots, i_d) = \prod_{k=1}^{d} M^{i_k},$$

where M^{i_k} is a matrix of size $r_{k-1} \times r_k$ with the entries

$$M^{i_k}(\alpha_{k-1}, \alpha_k) = g_k(\alpha_{k-1}, i_k, \alpha_k).$$

Given a d-level matrix A with sizes subject to (3), we may view it as a block matrix with $(m_1 \ldots m_s) \times (n_1 \ldots n_s)$ block entries denoted by $A_{i_{1:s}j_{1:s}}$ with multi-indices

$$i_{1:s} = i_1 \ldots i_s, \quad j_{1:s} = j_1 \ldots j_s.$$

The blocks $A_{i_{1:s}j_{1:s}}$ are referred to as blocks of the level s of A [11].

Theorem 2.1 *For any tensor-train representation (6) of a d-level matrix A with sizes subject to (3) the TT-ranks r_s satisfy inequalities*

$$r_s \geq \dim \mathcal{L}_s, \quad \textit{where} \quad \mathcal{L}_s = \mathrm{span}\{A_{i_{1:s}j_{1:s}}\},$$

and a tensor train exists with the equalities to be held for each s.

Proof. We can single out the blocks of level s by rearranging the terms of (6) as follows:

$$A = \sum_{\alpha_s=1}^{r_s} \left(\sum_{\alpha_1, \ldots, \alpha_{s-1}} \bigotimes_{k=1}^{s} G^k_{\alpha_{k-1}\alpha_k} \right) \left(\sum_{\alpha_{s+1}, \ldots, \alpha_{d-1}} \bigotimes_{k=s+1}^{d} G^k_{\alpha_{k-1}\alpha_k} \right).$$

The inequalities for r_s follow then directly from Lemma 1.1. They become equalities as soon as we take a common basis for the blocks of every level s at each step of the recursive construction of (6). $\square$

In the case $m = n = 2^d$, Theorem 2.1 was presented in [23]. Tensor-train representations in this special d-level case were proposed in [17] and called TTM-representations. In [17,18] it was first disclosed that TTM can be used for amazingly efficient implementations of basic operations with those matrices and vectors that enjoy reasonably small tensor-train ranks. Eventually, using TTM may lead to the complexity logarithmic in n.

The case of vectors in our exposition correponds to the values $m = 2^d$ and $n = 1$. Similar representations for vectors were advocated in [8] and called therein QTT-representations, the name emphasizing the quantization of the original size of a vector by introducing as many nontrivial mode-indices as possible. Nontrivial means attaining more than one value. Note that the very idea of artificial increasing the number of levels was earlier proposed and even somewhat studied in [21]. For various applications of QTT we refer to [9,10].

3 TT-ranks for matrices and their inverses

From now on we suppose that A is a d-level matrix with $m = n = 2^d$ and $m_s = n_s = 2$ for $1 \le s \le d$. By TT-ranks of A we mean the minimal possible values over all TT-representations available for A. The maximum of all TT-ranks for A in this case will be called a QTT-rank of A and denoted by $\mathrm{qttr}(A)$.

Theorem 3.1 *Assume that a matrix A is nonsingular and* $\mathrm{qttr}(A) = 1$. *Then* $\mathrm{qttr}(A^{-1}) = 1$.

Proof. It suffices to note that the equation $A = U \otimes V$ and nonsingularity of A imply that U and V are both nonsingular and $A^{-1} = U^{-1} \otimes V^{-1}$. □

Theorem 3.2 *Assume that a matrix A is nonsingular and* $\mathrm{qttr}(A) = 2$. *Then* $\mathrm{qttr}(A^{-1}) \le \sqrt{n}$.

Proof. The claim stems from a result proved in [20]: if $A = U_1 \otimes V_1 + U_2 \otimes V_2$, where U_1, U_2 are of size $p \times p$ and V_1, V_2 are of size $q \times q$, then A^{-1} admits a similar representation with at most $\min(p, q) \le \sqrt{n}$ terms. □

Unfortunately, the estimate of Theorem 3.2 is sharp (it follows from considerations of [20]). In order to have a better estimate we need to require of A something else, e.g. some agreeable structure. Below we consider Toeplitz banded matrices.

Recall that A is called a *Toeplitz matrix* if any its entry $A(i, j)$ depends only on $i - j$, i.e. we may set

$$A(i, j) = a_{i-j}.$$

Bandedness means that there are nonnegative integers l and u such that $a_k = 0$ whenever $k > l$ or $k < -u$. The numbers l and u are called the *lower* and *upper bandwidths* of A, respectively.

Lemma 3.1 *Assume that A is a Toeplitz banded matrix with lower and upper bandwidths l and u. Then* $\mathrm{qttr}(A) \le 1 + \lceil \sqrt{2l} \rceil + \lceil \sqrt{2u} \rceil$.

Proof. If q is a divisor of n, then A can be considered as a block matrix with blocks of size $q \times q$. From the structure of A it emanates that there could be at most $\mu(q) = 1 + \lceil l/q \rceil + \lceil u/q \rceil$ different blocks. Note also that these blocks are Toeplitz matrices defined by $2q - 1$ parameters. Hence, the dimension of linear span of these blocks cannot exceed $2q - 1$. We obtain the result by maximization of $\min(\mu(q), q)$ over q. □

Lemma 3.2 *Assume that A is a Toeplitz banded matrix of order $n = pq$ with lower and upper bandwidths l and u. Then A can be written as*

$$A = U_0 \otimes V_0 + \sum_{k=1}^{l+u} U_k \otimes V_k, \tag{7}$$

where U_k and V_k are of size $p \times p$ and $q \times q$ respectively, and additionally

$$\mathrm{rank} V_k = 1, \quad 1 \le k \le l + u. \tag{8}$$

Proof. Introduce a down-shift matrix Z_p of order p, it reads

$$Z_p = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix},$$

and let I_p denote the identity matrix of order p. Evidently,

$$A = a_0 \, I + \sum_{k=1}^{l} a_k \, Z_n^k + \sum_{k=1}^{u} a_{-k} \, (Z_n^\top)^k,$$

and similarly, when regarding A as a block Toeplitz matrix with $p \times p$ block entries $A_{i_1 - j_1}$, $1 \le i_1, j_1 \le p$, we obtain

$$A = I_p \otimes A_0 + \sum_{k=1}^{\lceil l/q \rceil} Z_p^k \otimes A_k + \sum_{k=1}^{\lceil u/q \rceil} (Z_p^\top)^k \otimes A_{-k}.$$

For illustration, let $n = 8$, $l = 3$ and $u = 1$. Then

$$A = \left[\begin{array}{cc|cc|cc|cc}
a & f & & & & & & \\
b & a & f & & & & & \\
\hline
c & b & a & f & & & & \\
d & c & b & a & f & & & \\
\hline
 & d & c & b & a & f & & \\
 & & d & c & b & a & f & \\
\hline
 & & & d & c & b & a & f \\
 & & & & d & c & b & a
\end{array}\right]$$

and with $p = 4$, $q = 2$ we obtain

$$A = I_4 \otimes \begin{bmatrix} a & f \\ b & a \end{bmatrix} + Z_4 \otimes \begin{bmatrix} c & b \\ d & c \end{bmatrix} + Z_4^2 \otimes \begin{bmatrix} 0 & d \\ 0 & 0 \end{bmatrix} + Z_4^\top \otimes \begin{bmatrix} 0 & 0 \\ f & 0 \end{bmatrix}.$$

To complete the proof, observe that the sum of ranks of the involved blocks A_k, $1 \le k \le l$, does not exceed l, and the sum of ranks of A_{-k}, $1 \le k \le u$, does not exceed u. $\square$

Lemma 3.3 [13] *If a matrix of the form*

$$K = I + \sum_{i=1}^{r} A_i \otimes u_i v_i^\top \tag{9}$$

with $p \times p$ matrices A_i and column vectors u_i, v_i with q entries is nonsingular, then

$$K^{-1} = I + \sum_{i=1}^{r} \sum_{j=1}^{r} A_{ij} \otimes u_i v_j^\top \tag{10}$$

with some $p \times p$ matrices A_{ij}.

Theorem 3.3 *Assume that A is a nonsingular banded Toeplitz matrix with the lower and upper bandwidths l and u. Then* $\mathrm{qttr}(A^{-1}) \le 1 + (l + u)^2$.

Proof. Let q be any divisor of n, and set $n = pq$. By Lemma 3.2 we represent A by a sum of two-factor Kronecker products (7). For all sufficiently small $\varepsilon > 0$ it is possible to find a nonsingular $p \times p$ matrix U_ε and a nonsingular $q \times q$ matrix V_ε with the following properties:

(1) $||U_\varepsilon - U_0||_2 \le \varepsilon$, $\quad ||V_\varepsilon - V_0|| \le \varepsilon$,

(2) $A_\varepsilon \equiv U_\varepsilon \otimes V_\varepsilon + \sum_{k=1}^{l+u} U_k \otimes V_k$ is a nonsingular matrix.

Obviously, $A_\varepsilon \to A$ as $\varepsilon \to 0$. At the same time, for $\varepsilon > 0$ we can come from A to a new matrix

$$K = A\,(U_\varepsilon^{-1} \otimes V_\varepsilon^{-1}),$$

which is exactly of the form (9). Then we directly apply Lemma 3.3 established in [13] and remark that the Kronecker rank in the case of two-factor products cannot increase when we get to a limit. $\square$

Theorem 3.3 adds an essential new feature to the already known properties of the inverses to Toeplitz banded matrices. It has been known to date that they possess a remarkable structure: the Toeplitz structure induces that the inverse is a sum of two products of Toeplitz triangular matrices, the bandedness implies that the defining parameters of those Toeplitz triangular matrices are determined through linear combinations of some vectors whose entries comprise a geometrical progression. We might have obtained some estimate of QTT-ranks by employing this structure, but we have opted for a more direct approach based on Theorem 2.1.

When performing operations with matrices, we can use tensor trains as basic data structures and deal only with them during all computations. Of course, TT-ranks may grow and we need to approximate some tensor trains by another ones with smaller TT-ranks. Such an approach is viable indeed with many other data structures. In theory we impose pretty mild requirements on these structures [5], in practice we surely need efficient approximation algorithms. However, it is crucial that the results of computations *admit* low-parametric approximations in the selected data formats. In this work we proved that this is so for the inversion of a Toeplitz banded matrix. Estimates for other structured matrices are the subject of on-going research.

References

[1] B. Bader, T. Kolda, Tensor decompositions and applications, SIAM Rev., 51 (2009), pp. 455-500.

[2] G. Beylkin, M. J. M. Mohlenkamp, Numerical operator calculus in higher dimensions, Proc. Nat. Acad. Sci. USA., 99 (2002), pp. 10246-10251.

[3] L. Grasedyck, Hierarchical singular value decomposition of tensors, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2029-2054.

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

[5] W. Hackbusch, B. N. Khoromskij, E. Tyrtyshnikov, Approximate iterations for structured matrices, *Numer. Math.*, vol.109, no. 3, pp. 365–383 (2008).

[6] W. Hackbusch, S. Kuhn, A new scheme for the tensor representation, J. Fourier Anal. Appl., 15 (2009), pp. 706-722.

[7] J. Kamm and J. G. Nagy, Optimal Kronecker Product Approximations of Block Toeplitz Matrices, *SIAM J. Matrix Anal. Appl.*, Vol. 22, No. 1, pp. 155-172 (2000).

[8] B. N. Khoromskij, $O(dlogN)$-quantics approximation of $N-d$ tensors in high-dimensional numerical modelling, Preprint 40, MPI MIS, 2009.

[9] B. N. Khoromskij, I. V. Oseledets, DMRG+QTT approach to quantum molecular dynamics, preprint, MIS MPI, 2010.

[10] B. N. Khoromskij, I. V. Oseledets, Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDE, Preprint 37, MIS MPI, 2010.

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

[12] V. Olshevsky, I. Oseledets, E. Tyrtyshnikov, Superfast inversion of two-level Toeplitz matrices using Newton iteration and tensor-displacement structure, *Operator Theory Advances and Applications*, vol. 179, pp. 229–240 (2007).

[13] I. Oseledets, E. Tyrtyshnikov, N. Zamarashkin, Matrix inversion cases with size-independent tensor rank estimates, Linear Algebra Appl., 431 (2009), 558–570.

[14] I. Oseledets, E. Tyrtyshnikov, Recursive decomposition of multidimensional tensors, Doklady Mathematics, vol. 80, no. 1 (2009), pp. 460–462. (In Russian: Doklady Akademii Nauk, vol. 427, no. 1 (2009), pp. 14–16)

[15] I. Oseledets, E. Tyrtyshnikov, Breaking the curse of dimensionality, or how to use SVD in many dimensions. SIAM J. Sci. Comput., vol 31, no. 5 (2009), pp. 3744–3759.

[16] I. V. Oseledets, A new tensor decomposition, Doklady Mathematics, vol. 80, no. 1 (2009), pp. 495-496.

[17] I. V. Oseledets, On the approximation of matrices with logarithmical number of parameters, Doklady Mathematics, vol. 80, no. 2 (2009), pp. 653-654.

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

[19] I. Oseledets, E. Tyrtyshnikov, TT-cross approximation for multidimensional arrays, Linear Algebra Appl., 432 (2010), pp. 70–88.

[20] E. Tyrtyshnikov, Tensor ranks for the inversion of tensor-product binomials, J. Comput. Appl. Math., vol. 234, no. 11 (2010) 3170-3174.

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

[22] C. F. Van Loan, N. P. Pitsianis, Approximation with Kronecker products, *NATO Adv. Sci. Ser E Appl. Sci.* 232, Kluwer: Dordrecht, pp. 293–314 (1993).

[23] 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. (In Russian: Doklady Akademii Nauk, vol. 428, no. 2 (2009), pp. 161–162)