

A New Scheme for the Tensor Representation

W. Hackbusch · S. Kühn

Received: 18 December 2008 / Revised: 29 June 2009 / Published online: 22 October 2009
© Birkhäuser Boston 2009

Abstract The paper presents a new scheme for the representation of tensors which is well-suited for high-order tensors. The construction is based on a hierarchy of tensor product subspaces spanned by orthonormal bases. The underlying binary tree structure makes it possible to apply standard Linear Algebra tools for performing arithmetical operations and for the computation of data-sparse approximations. In particular, a truncation algorithm can be implemented which is based on the standard matrix singular value decomposition (SVD) method.

Keywords Multilinear algebra · Tensor representation · Singular value decomposition

Mathematics Subject Classification (2000) 15A69

1 Introduction

1.1 General Remarks Concerning Representations

The representation of mathematical objects in numerical implementations is by no ways unique. In most of the cases one needs representations which are fitted to the special structure of the mathematical object to be characterised. The arising questions can be explained for the example of matrices. The representation of a matrix as full matrix, i.e. as array of all n^2 entries, is only practical for moderate n . For large-scale

Dedicated to the 60th birthday of Wolfgang Dahmen.

Communicated by Reinhard Hochmuth.

W. Hackbusch (✉) · S. Kühn

MPI Mathematik in den Naturwissenschaften, Inselstr. 22, 04103 Leipzig, Germany
e-mail: wh@mis.mpg.de

matrices this approach fails because of the too large storage requirements. However, exploiting structure properties, e.g., of Toeplitz matrices or sparse matrices, one is led to alternative representations which are optimal concerning storage. The problem remains that not all matrices appearing in practice possess such utilisable properties.

When representing a mathematical object one has to distinguish between an exact and an approximative representation. Already for the representation of real numbers one needs approximative representations. For matrices appearing, e.g., in wavelet applications one uses the fact that although the matrices are fully populated, most of the entries are very small and can be replaced by zero leading to a sparse matrix, i.e. the representation as sparse matrix is an approximative representation of the exact matrix. Similarly, the format of hierarchical matrices serves as an approximative representation. In this context several questions arise. Since the accuracy of the approximation behaves reciprocal to the storage of the representation, one likes to know:

- (1) What accuracy is reachable for a certain amount of storage?
- (2) What storage is needed for a given accuracy?

Both are rather theoretical questions. The essential question in practice is the constructive realisation:

- (3) How to determine an approximation of the best approximation?

In general, methods solving the last item have to work with incomplete data. Therefore, error estimates require hypotheses concerning to given data. In the case of the wavelet compression of matrices mentioned above, one obviously must not use the full matrix and test all its coefficients whether they are small enough, but one needs a priori information about their size. The criteria are, e.g., based on smoothness assumptions (cf. [3]).

Having answered the previous questions, one has only solved the storage problem. The next problem concerns the operations associated with the respective mathematical object. In the matrix case, the matrix-vector multiplication and further matrix operations (addition, multiplication, inversion, LU decomposition etc.) are of interest. Here the problem arises that, e.g., the sum of a sparse matrix and a Toeplitz matrix in general neither is sparse nor Toeplitz and even cannot be approximated in one of the mentioned representations.

In the case of matrices, these difficulties have led to the situation that one tries to reduce to sparse matrices and the matrix-vector multiplication. The technique of hierarchical matrices is now a new possibility which applies to a large class of matrices. The reason is a representation format of matrices which requires not only almost linear storage size, but also supports all operations, since they can be performed approximately with almost linear computation work (cf. [7]).

1.2 Representation of Tensors

The difficulties of the matrix representation multiply, when we want to deal with tensors of higher order $d \geq 3$. The representation as a full tensor is practically impossible. Two representation formats have been used, which in many cases are helpful for the approximative representation. The representation as rank- r -tensor described

in Sect. 3.1 can also be used for large d , whereas the subspace tensor representation (“Tucker” representation, cf. Sect. 3.2) becomes unfavourable for large d . In both cases, the third question concerning an (almost) optimal approximation is incomparably more difficult than in the matrix case. As a third representation format we mention the sparse-grid representation (see Sect. 3.3). In typical applications the latter format has a complexity containing a factor $(\log n)^d$ which limits applications for large d .

To be concrete we mention some of the tensor operations:

- addition of tensors,
- matrix-vector multiplication employing the Kronecker product for matrices,
- Hadamard product of tensors (component-wise products).

All these operations lead to a result with an increased tensor rank. Therefore, an essential tool is the

- tensor truncation, i.e. the tensor \mathbf{v} is replaced by an approximation $\tilde{\mathbf{v}}$ with an improved representation size (e.g., smaller tensor rank).

In the special case $d = 2$, the singular value decomposition (SVD) is the method of choice to solve the truncation problem. Unfortunately, the situation becomes very unpleasant for $d \geq 3$. There are methods available (cf. Espig [4]), but the solution of the underlying complicated optimisation problems is by far more complicated¹ than SVD.

Having in mind the above mentioned problems in the case $d \geq 3$ we derive some desired properties of a tensor representation scheme:

- (1) it should be applicable also for large d (avoid curse of dimensionality),
- (2) it should possess at least as good approximation properties as the previously mentioned representations,
- (3) it should support the tensor operations including the very important truncation,
- (4) it should rely on easy performable Linear Algebra tools.

The new format which is supposed to have these properties is described in Sect. 2. In Sect. 3 we show that the proposed format contains the three ones mentioned above, i.e. a tensor given in one of these formats can be represented exactly by the new format. Section 5 is devoted to the realisation of the tensor operations. The truncation procedure is discussed in Sect. 6.

The application of the multiplication by Kronecker products and the realisation of Hadamard product will be described in forthcoming papers. The problem to get information about the tensor from selected partial data is solved differently in [5] and [10]. Concerning the construction of tensor representations for certain applications we refer to [8] and [9]. Applications in the field of quantum chemistry can be found in [2].

¹For instance, the performance of the iteration depends on the choice of initial values. Different starting values may lead to different local minima.

1.3 Notation

We consider the tensor product space

$$\mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$$

where V^j ($1 \leq j \leq d$) are assumed to be Hilbert spaces, i.e. scalar products $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{V^j}$ are available. Then \mathbf{V} is again a Hilbert space with the induced scalar product defined by the property

$$\langle \mathbf{v}, \mathbf{w} \rangle = \prod_{j=1}^d \langle v^j, w^j \rangle_{V^j} \quad \text{for } \mathbf{v} = \bigotimes_{j=1}^d v^j, \mathbf{w} = \bigotimes_{j=1}^d w^j.$$

The degree (or dimension) d is assumed to be at least 2.

We mention three concrete realisations of the tensor product:

1. For $V^j = \mathbb{R}^{n_j}$, vectors $v^j \in \mathbb{R}^{n_j}$ yield the tensor

$$\mathbf{v} = \bigotimes_{j=1}^d v^j = (\mathbf{v}_{i_1 i_2 \dots i_d})_{1 \leq i_1 \leq n_1, \dots, 1 \leq i_d \leq n_d}$$

with $\prod_{j=1}^d n_j$ entries of the form $(\mathbf{v}_{i_1 i_2 \dots i_d}) = v_{i_1}^1 \dots v_{i_d}^d$.

2. For the vector space $V^j = \mathbb{R}^{n_j \times n_j}$ of $n_j \times n_j$ matrices, the *Kronecker product* $\mathbf{A} = \bigotimes_{j=1}^d A^j$ is defined by the property

$$\mathbf{A}\mathbf{v} = \bigotimes_{j=1}^d A^j v^j \quad (\mathbf{v} \text{ from item 1}).$$

3. For the infinitely dimensional example $V^j = L^2(\Omega_j)$, the tensor product $f = \bigotimes_{j=1}^d f_j$ of functions $f_j \in V^j$ is the pointwise product $f(x_1, \dots, x_d) = \prod_{j=1}^d f_j(x_j)$.

In the finitely dimensional case, $\mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$ is defined by

$$\mathbf{V} = \text{span} \left\{ \bigotimes_{j=1}^d v^j : v^j \in V^j \right\}, \tag{1.1}$$

while in the infinitely dimensional case one must take the closure of the right-hand side in respect of the topology of the Hilbert space \mathbf{V} .

2 The Hierarchical Tensor Format

For the ease of presentation we assume in the following that the dimension d is a power of 2,

$$d = 2^L \geq 2. \tag{2.1}$$

In Sect. 7 we mention how to treat the case of general d . We want to approximate tensors \mathbf{v} from the tensor space $\mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$ introduced in (1.1).

2.1 Subspaces $U^{\ell,i}$

2.1.1 Level 0

We rewrite the vector spaces V^j appearing in $\mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$ by $V^{0,j}$ since these spaces will be associated with the level number 0. As in the subspace tensor representation we provide d subspaces

$$U^{0,j} \subset V^{0,j} = V^j \quad (1 \leq j \leq d)$$

of dimension $r^{0,j} < \infty$. These subspaces lead to the tensor subspace

$$\mathbf{V}^0 := U^{0,1} \otimes U^{0,2} \otimes \dots \otimes U^{0,d} \subset \mathbf{V}.$$

2.1.2 Levels 1, ..., L

Because of (2.1), d is even and we can rewrite \mathbf{V}^0 in the form

$$\mathbf{V}^0 = (U^{0,1} \otimes U^{0,2}) \otimes (U^{0,3} \otimes U^{0,4}) \otimes \dots \otimes (U^{0,d-1} \otimes U^{0,d}).$$

For each pair (denoted by $V^{1,j}$) we again introduce some finitely dimensional subspace

$$U^{1,j} \subset V^{1,j} := U^{0,2j-1} \otimes U^{0,2j} \quad (1 \leq j \leq 2^{L-1} = d/2).$$

This procedure can be repeated recursively. The transition from level $\ell - 1$ to ℓ is described by the choice of subspaces

$$U^{\ell,j} \subset V^{\ell,j} := U^{\ell-1,2j-1} \otimes U^{\ell-1,2j} \quad (1 \leq j \leq 2^{L-\ell}). \tag{2.2}$$

2.1.3 Level L

The recursion stops with $\ell = L$ (L from (2.1)) since $U^{L,1}$ is the only space of level L . In standard applications the purpose of $U^{L,1}$ is to contain an approximation $\tilde{\mathbf{v}}$ of some tensor $\mathbf{v} \in \mathbf{V}$. Therefore, a one-dimensional subspace

$$U^{L,1} = \text{span}\{\tilde{\mathbf{v}}\} \tag{2.3}$$

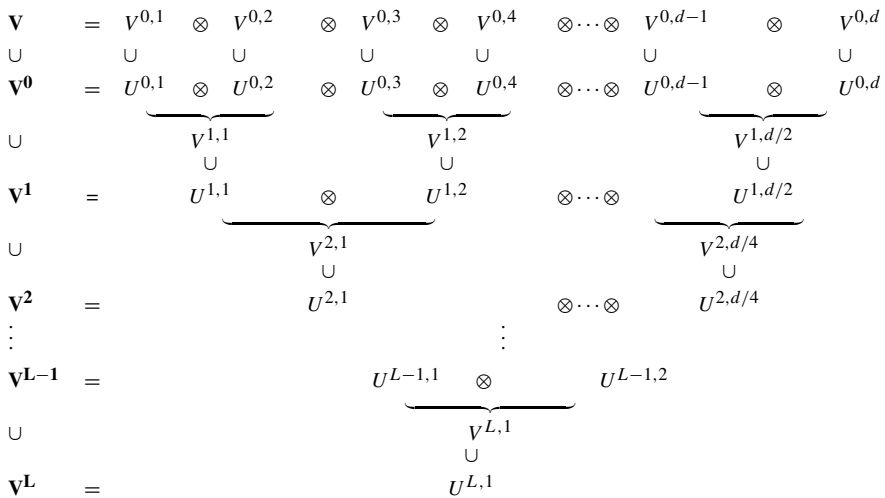
is the usual choice. Only when we want to approximate several tensors simultaneously, a larger dimension of $U^{L,1}$ is of interest.

The set of subspaces $U^{\ell,j}$ takes the form of a binary tree depicted below (see Table 1).

The diagram also shows the spaces $\mathbf{V}^\ell = \bigotimes_{j=1}^{2^{L-\ell}} U^{\ell,j}$ with the property

$$U^{L,1} = \mathbf{V}^L \subset \mathbf{V}^{L-1} \subset \dots \subset \mathbf{V}^0 \subset \mathbf{V}.$$

Table 1 Hierarchical tree structure for $d = 2^L$



2.2 Bases and Coefficients

2.2.1 Vector Space $V^{0,j}$ and Subspace $U^{0,j}$

The representation of elements of the fundamental vector spaces $V^{0,j} = V^j$ is assumed to be given together with V^j . It is not part of the representation scheme which we are going to describe here. If $V^j = \mathbb{R}^{n_j}$ with finite n_j , one may use a full vector representation. In that case, $storagesize(V^j) = n_j$ denotes the required storage of a vector $v^j \in V^j$. If $V^j \subset \mathbb{R}^{n_j \times n_j}$ contains matrices in the \mathcal{H}^2 -format, the required storage is $storagesize(V^j) = \mathcal{O}(n_j)$. If $V^j = L^2(\Omega_j)$ has infinite dimension, one needs other descriptions anyway. We also assume that for given vectors $v^j, w^j \in V^j$ we are able to compute the scalar product $\langle v^j, w^j \rangle$.

By definition the subspace $U^{0,j}$ has a finite dimension $\dim(U^{0,j}) = r^{0,j}$. Together with the subspace we have to fix a basis $\{\phi^{0,j,\gamma} : 1 \leq \gamma \leq r^{0,j}\}$. Since the scalar product is available, we can apply, e.g., the Gram-Schmidt orthogonalisation. Therefore, w.l.o.g. we may assume that the basis is *orthonormal*.

Having fixed the orthonormal basis, any vector $v \in U^{0,j}$ is represented by means of its coefficients $(a_\gamma)_{\gamma=1, \dots, r^{0,j}}$ in $v = \sum a_\gamma \phi^{0,j,\gamma}$, whereas the basis vectors $\phi^{0,j,\gamma}$ are provided by the internal representation system of V^j .

The choice of an orthonormal basis has two reason. The first one is stability as stated later in Remark 3.2. The second one will become obvious in Sect. 6, when singular value decompositions yield orthonormal bases anyway.

2.2.2 Tensor Space $V^{\ell,j} = U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$

By induction we assume that each $U^{\ell-1,j}$ is spanned by the *orthonormal basis*

$$\left\{ \phi^{\ell-1,j,\gamma} : 1 \leq \gamma \leq r^{\ell-1,j} \right\}.$$

Then the tensor product space $V^{\ell,j} = U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$, which is of dimension $\dim V^{\ell,j} = \dim(U^{\ell-1,2j-1}) \cdot \dim(U^{\ell-1,2j}) = r^{\ell-1,2j-1} r^{\ell-1,2j}$, has the canonical orthonormal basis

$$\left\{ \phi^{\ell-1,2j-1,\alpha} \otimes \phi^{\ell-1,2j,\beta} : 1 \leq \alpha \leq r^{\ell-1,2j-1}, 1 \leq \beta \leq r^{\ell-1,2j} \right\}. \tag{2.4}$$

A vector (tensor) $v \in V^{\ell,j}$ is represented by means of its coefficients $a_{\alpha,\beta}$ in

$$v = \sum_{\alpha=1}^{r^{\ell-1,2j-1}} \sum_{\beta=1}^{r^{\ell-1,2j}} a_{\alpha,\beta} \phi^{\ell-1,2j-1,\alpha} \otimes \phi^{\ell-1,2j,\beta}. \tag{2.5}$$

Since these coefficients have two indices α, β , they form a matrix

$$A = (a_{\alpha,\beta}) \in \mathbb{R}^{r^{\ell-1,2j-1} \times r^{\ell-1,2j}}. \tag{2.6}$$

As a consequence of the orthonormality of the basis the following result holds.

Lemma 2.1 *Let $v, w \in V^{\ell,j}$ have the respective coefficients $A = (a_{\alpha,\beta})$ and $B = (b_{\alpha,\beta})$. Then*

$$\langle v, w \rangle = \sum_{\alpha,\beta} a_{\alpha,\beta} b_{\alpha,\beta} = \langle A, B \rangle_F,$$

where $\langle \cdot, \cdot \rangle_F$ is the Frobenius scalar product of matrices.

2.2.3 Subspace $U^{\ell,j}$

For the subspace $U^{\ell,j}$ we choose a basis $\{\phi^{\ell,j,\gamma}\}$ which again w.l.o.g. is organised as an orthonormal basis:

$$U^{\ell,j} = \text{span} \left\{ \phi^{\ell,j,\gamma} : 1 \leq \gamma \leq r^{\ell,j} \right\}.$$

Each basis vector $\phi^{\ell,j,\gamma}$ ($1 \leq \ell \leq L, 1 \leq j \leq 2^{L-\ell}, 1 \leq \gamma \leq r^{\ell,j}$) is represented by means of its coefficient matrix

$$A^{\ell,j,\gamma} = \left(a_{\alpha,\beta}^{\ell,j,\gamma} \right)_{\alpha,\beta} \in \mathbb{R}^{r^{\ell-1,2j-1} \times r^{\ell-1,2j}}, \tag{2.7}$$

where $a_{\alpha,\beta}^{\ell,j,\gamma}$ are the coefficients in (2.5) for $v = \phi^{\ell,j,\gamma}$:

$$\phi^{\ell,j,\gamma} = \sum_{\alpha=1}^{r^{\ell-1,2j-1}} \sum_{\beta=1}^{r^{\ell-1,2j}} a_{\alpha,\beta}^{\ell,j,\gamma} \phi^{\ell-1,2j-1,\alpha} \otimes \phi^{\ell-1,2j,\beta}. \tag{2.8}$$

Once the basis vectors are described, a vector $u \in U^{\ell,j}$ is associated with the usual coefficient vector $c = (c_\gamma)_{\gamma=1,\dots,r^{\ell,j}} \in \mathbb{R}^{r^{\ell,j}}$ with c_γ from

$$u = \sum_{\gamma=1}^{r^{\ell,j}} c_\gamma \phi^{\ell,j,\gamma}. \tag{2.9}$$

2.2.4 Final Representation

The subspaces $U^{\ell,j}$ for $0 \leq \ell \leq L - 1$ are of auxiliary purpose, whereas $U^{L,1}$ is used to represent the desired approximation. As mentioned above (cf. (2.3)), $\mathbf{V}^L = U^{L,1} \subset \mathbf{V}$ may have dimension $r^{L,1} = 1$, leading to a trivial representation in (2.9).

The connection with the vectors from the original vector spaces V^j is given indirectly by the recursive definitions of the basis vectors:

$$\begin{aligned} \phi^{L,1,\gamma} &= \sum_{\alpha=1}^{r^{L-1,1}} \sum_{\beta=1}^{r^{L-1,2}} a_{\alpha,\beta}^{L,1,\gamma} \phi^{L-1,1,\alpha} \otimes \phi^{L-1,2,\beta} \quad (1 \leq \gamma \leq r^{L,1}), \\ &\vdots \\ \phi^{\ell,j,\gamma} &= \sum_{\alpha=1}^{r^{\ell-1,2j-1}} \sum_{\beta=1}^{r^{\ell-1,2j}} a_{\alpha,\beta}^{\ell,j,\gamma} \phi^{\ell-1,2j-1,\alpha} \otimes \phi^{\ell-1,2j,\beta} \quad (1 \leq j \leq 2^{L-\ell}, 1 \leq \gamma \leq r^{\ell,j}), \\ &\vdots \\ \phi^{1,j,\gamma} &= \sum_{\alpha=1}^{r^{0,2j-1}} \sum_{\beta=1}^{r^{0,2j}} a_{\alpha,\beta}^{1,j,\gamma} \phi^{0,2j-1,\alpha} \otimes \phi^{0,2j,\beta} \quad (1 \leq j \leq 2^{L-1}, 1 \leq \gamma \leq r^{1,j}), \end{aligned}$$

and $\phi^{0,j,\alpha} \in V^j$. The latter scheme reflects the tree structure from Table 1.

Let $\Phi = (A^{\ell,j,\gamma})$ be the tuple of all basis coefficients characterising the described representation. Then we say that a tensor from (2.9) (in particular for $\ell = L$) is described by the Φ -system.

2.3 Storage

Remark 2.2

(a) The storage for all matrices $A^{\ell,j,\gamma}$ characterising the basis vectors $\phi^{\ell,j,\gamma}$ amounts to

$$N_{\text{basis}} := \sum_{\ell=1}^L \sum_{j=1}^{2^{L-\ell}} r^{\ell,j} r^{\ell-1,2j-1} r^{\ell-1,2j}.$$

- (b) The storage for $\phi^{0,j,\gamma}$ at level 0 requires additionally $\sum_{j=1}^d r^{0,j} \text{storage size}(V^j)$.
- (c) Let $r := \max\{r^{\ell,j} : 1 \leq j \leq 2^{L-\ell}, 1 \leq \ell \leq L\}$. Then $N_{\text{basis}} \leq \sum_{\ell=1}^L 2^{L-\ell} r^3 = (d - 1)r^3$ holds. Note that the dependence on the dimension d is linear.
- (d) Vectors in $U^{\ell,j}$ require a storage of size $r^{\ell,j}$ (cf. (2.9)).

3 Comparison with Other Formats

3.1 r -term Representation

Because of (1.1), every $\mathbf{v} \in \mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$ has a representation

$$\mathbf{v} = \sum_{v=1}^r v^{1,v} \otimes v^{2,v} \otimes \dots \otimes v^{d,v} \quad (v^{j,v} \in V^j = V^{0,j}), \tag{3.1}$$

which is called *r-term representation*. (3.1) is of practical interest if r is sufficiently small. The minimal r of all representations (3.1) defines the *tensor rank* of \mathbf{v} .

Proposition 3.1 *Let $\mathbf{v} \in \mathbf{V}$ satisfy (3.1). Then there is a Φ -system such that $\mathbf{v} \in U^{L,1}$ holds exactly. The subspaces $U^{\ell,j}$ are given by*

$$U^{\ell,j} := \text{span} \left\{ \bigotimes_{i=2^{\ell}(j-1)+1}^{2^{\ell}j} v^{i,v} : 1 \leq v \leq r \right\} \quad \text{for } 0 \leq \ell \leq L, 1 \leq j \leq 2^{L-\ell}.$$

Their dimensions are bounded by $r^{\ell,j} \leq r$. The last space $U^{L,1}$ may be replaced by $U^{L,1} = \text{span}\{\mathbf{v}\}$. For $\ell = 0$, the subspaces are $U^{0,j} = \text{span}\{v^{j,v} : 1 \leq v \leq r\}$. Their storage (see Remark 2.2b) is as large as for the r -term representation.

Proof The property $U^{0,j} \subset V^j$ is trivial. Next, we have to check that $U^{\ell,j} \subset U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$. By the definition of $U^{\ell,j}$, $\bigotimes_{i=2^{\ell}(j-1)+1}^{2^{\ell}j} v^{i,v} \in U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$ is to be shown. Note that $\bigotimes_{i=2^{\ell}(j-1)+1}^{2^{\ell}j} v^{i,v} = (\bigotimes_{i=2^{\ell}(j-1)+1}^{2^{\ell}(j-1/2)} v^{i,v}) \otimes (\bigotimes_{i=2^{\ell}(j-1/2)+1}^{2^{\ell}j} v^{i,v})$. The first factor equals $\bigotimes_{i=2^{\ell-2}(2j-1)}^{2^{\ell-2}(2j-1)} v^{i,v} \in U^{\ell-1,2j-1}$, while the second factor is one of the vectors spanning $U^{\ell-1,2j}$. This proves that the defined spaces define a Φ -system. Since $U^{L,1}$ contains all terms $v^{1,v} \otimes v^{2,v} \otimes \dots \otimes v^{d,v}$, $\mathbf{v} \in U^{L,1}$ holds.

All subspaces are spanned by only r vectors (which may be linear dependent). Therefore $r^{\ell,j} \leq r$ follows. \square

In order to define the bases of the Φ -system, one has to orthonormalise the vectors spanning $U^{\ell,j}$. According to Remark 2.2c modified for the case of $\dim U^{L,1} = 1$, the storage of the Φ -system is

$$N_{\text{basis}} \leq (d-2)r^3 + r^2$$

plus the storage for the components $v^{j,v}$ from (3.1).

The comparison shows that the Φ -system requires a storage overhead of size N_{basis} , which is a disadvantage. However, for a hopefully small tensor rank r this additional term may be negligible compared with storage for the components $v^{j,v}$. On the other hand, besides the advantages which we describe later, both representation system have quite different *stability properties*.

Remark 3.2 For $d \geq 3$, the r -term representation may behave unstable, since the size $\sum_{v=1}^r \|v^{1,v} \otimes v^{2,v} \otimes \dots \otimes v^{d,v}\|$ may be much larger than $\|\mathbf{v}\|$ (“cancellation effect”). On the other hand, the Φ -system uses orthonormal basis representations which guarantees stability. Only the orthonormalisation of the original data, represented by the input on level 0, can be influenced by the cancellation effect. Once the orthonormal bases are constructed, cancellation cannot occur.

3.2 Subspace Representation

The subspace representation (or Tucker representation, see [11]) of $\mathbf{v} \in \mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$ is characterised by subspaces $U^j \subset V^j$ ($1 \leq j \leq d$) with the property

$$\mathbf{v} \in \mathbf{U} := U^1 \otimes U^2 \otimes \dots \otimes U^d. \tag{3.2}$$

Each U^j is spanned by an orthonormal basis $\{\phi^{0,j,\gamma} : 1 \leq \gamma \leq \dim U^j\}$. Then \mathbf{v} is described by

$$\mathbf{v} = \sum_{v_1=1}^{\dim U^1} \sum_{v_2=1}^{\dim U^2} \dots \sum_{v_d=1}^{\dim U^d} \mathbf{a}_{v_1, v_2, \dots, v_d} \phi^{1, v_1} \otimes \phi^{2, v_2} \otimes \dots \otimes \phi^{d, v_d}$$

requiring the storage for the so-called core tensor $\mathbf{A} = (\mathbf{a}_{v_1, v_2, \dots, v_d})_{v_1, v_2, \dots, v_d} \in \otimes_{j=1}^d \mathbb{R}^{\dim U^j}$ of size $\prod_{j=1}^d \dim U^j$. The basis vectors $\phi^{j,v}$ need the storage $\sum_{j=1}^d \text{storage size}(V^j)$. For the particular case $\dim U^j = r$ for all j , the core tensor needs the storage r^d and the storage of basis vectors is the same as for the r -term representation.

Proposition 3.3 *Let $\mathbf{v} \in \mathbf{V}$ satisfy (3.2). Then there is a Φ -system such that $\mathbf{v} \in U^{L,1}$ holds exactly. The subspaces $U^{\ell,j}$ are given by*

$$U^{\ell,j} := \bigotimes_{i=2^\ell(j-1)+1}^{2^\ell j} U^i \quad \text{for } 0 \leq \ell \leq L, 1 \leq j \leq 2^{L-\ell}$$

and have the dimension $\prod_{i=2^\ell(j-1)+1}^{2^\ell j} \dim(U^i)$. The last space $U^{L,1}$ may be replaced by $U^{L,1} = \text{span}\{\mathbf{v}\}$. For $0 < \ell < L$, the subspaces $U^{\ell,j}$ are maximal since the equality $U^{\ell,j} = U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$ holds.

The fact that the subspaces $U^{\ell,j}$ are chosen maximally, yields a less advantageous Φ -system reflecting the higher storage requirement of the subspace representation.

For simplicity, we consider the case $\dim(U^i) = r$ for all i . Then

$$N_{\text{basis}} = \sum_{\ell=1}^{L-1} 2^{L-\ell} r^{2^{\ell+1}} + r^{2^L} = 3r^d \left(1 + \mathcal{O}(r^{-d/2})\right)$$

(cf. Remark 2.2(a)), when we make use of $U^{L,1} = \text{span}\{\mathbf{v}\}$. This is the same order as the storage of the core tensor of size r^d .

3.3 Sparse Grids

The sparse grid structure (for an overview see [1]) requires that each vector space V^j from $\mathbf{V} = V^1 \otimes V^2 \otimes \dots \otimes V^d$ has a subspace hierarchy

$$V_0^j \subset V_1^j \subset \dots \subset V_M^j = V^j$$

(for instance, V_ℓ^j is associated with the grid size $h_\ell = 2^{-\ell}h_0$). Then the sparse-grid space is defined by

$$U = \sum_{m_1+m_2+\dots+m_d=m} V_{m_1} \otimes V_{m_2} \otimes \dots \otimes V_{m_d}$$

for some $m \geq M$.

Proposition 3.4 *Let $\mathbf{v} \in \mathbf{V}$ satisfy $\mathbf{v} \in U$. Then there is a Φ -system such that $\mathbf{v} \in U^{L,1}$ holds exactly. The subspaces $U^{\ell,j}$ are given by*

$$U^{\ell,j} := \sum_{\sum_{i=2^\ell(j-1)+1}^{2^\ell j} m_i \leq m} \bigotimes_{i=2^\ell(j-1)+1}^{2^\ell j} V_{m_i}^j \quad \text{for } 0 \leq \ell \leq L, 1 \leq j \leq 2^{L-\ell}$$

(the sum is taken over all $0 \leq m_i \leq M$ for $2^\ell(j-1) + 1 \leq i \leq 2^\ell j$ with the side condition $\sum_{i=2^\ell(j-1)+1}^{2^\ell j} m_i \leq m$).

Proof For $\ell = 0$, $U^{\ell,j} = V_M^j = V^j$ holds. The desired inclusion $U^{\ell,j} \subset U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$ follows from $\bigotimes_{i=2^\ell(j-1)+1}^{2^\ell j} V_{m_i}^j \subset U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$ for m_i satisfying $\sum_{i=2^\ell(j-1)+1}^{2^\ell j} m_i \leq m$.

We rewrite $\bigotimes_{i=2^\ell(j-1)+1}^{2^\ell j} V_{m_i}^j$ as $(\bigotimes_{i=2^\ell(j-1)+1}^{2^\ell(j-1/2)} V_{m_i}^j) \otimes (\bigotimes_{i=2^\ell(j-1/2)+1}^{2^\ell j} V_{m_i}^j)$. Obviously, also the partial sums $\sum_{i=2^\ell(j-1)+1}^{2^\ell(j-1/2)} m_i$ and $\sum_{i=2^\ell(j-1/2)+1}^{2^\ell j} m_i$ are bounded by m , i.e. $\bigotimes_{i=2^\ell(j-1)+1}^{2^\ell(j-1/2)} V_{m_i}^j \subset U^{\ell-1,2j-1}$ and $\bigotimes_{i=2^\ell(j-1/2)+1}^{2^\ell j} V_{m_i}^j \subset U^{\ell-1,2j}$ hold. □

3.4 An Advantageous Example for the New Representation

The new representation can take advantage of separation or factorisation properties, that cannot be exploited by other representations, leading to even lower storage requirements and thus better performance compared to other representation schemes. To exemplify this, we consider the approximation of a function f by a separable expression. Let $d = 2^2$ and assume that f is of the form

$$f(x_1, x_2, x_3, x_4) = \alpha(x_1, x_2)\beta(x_3, x_4),$$

i.e., the pairs x_1, x_2 and x_3, x_4 are already perfectly separated. In quantum chemistry such a situation occurs if both groups of variables belong to particles which are sufficiently far apart. For the given example we have to determine separable approximations $f \approx f_r := \alpha_r \otimes \beta_r$ with

$$\alpha_r(x_1, x_2) = \sum_{\nu=1}^r \gamma_{\nu,1}(x_1)\gamma_{\nu,2}(x_2), \quad \beta_r(x_3, x_4) = \sum_{\nu=1}^r \gamma_{\nu,3}(x_3)\gamma_{\nu,4}(x_4),$$

i.e., $\alpha_r = \sum_{v=1}^r \gamma_{v,1} \otimes \gamma_{v,2}$ and $\beta_r = \sum_{v=1}^r \gamma_{v,3} \otimes \gamma_{v,4}$. The described new representation system is characterised by r -dimensional subspaces $U^{0,j} = \text{span}\{\gamma_{v,j}\}$ for $1 \leq j \leq 4$ on level 0, while on level 1 the subspaces are one-dimensional: $U^{1,1} = \text{span}\{\alpha_r\}$ and $U^{1,2} = \text{span}\{\beta_r\}$ as well as $U^{2,1} = \text{span}\{\alpha_r \otimes \beta_r\}$ on level 2. Hence, the subspace dimensions which are characteristic for the storage and computational costs are r at level 0 and only 1 otherwise.

If we try to write f_r in the form of Sect. 3.1, it becomes an r^2 -term approximation: $f_r = \sum_{\sigma=1}^r \sum_{\sigma=1}^r \gamma_{\sigma,4}$. The subspace representation is anyway more costly.

Increasing d , one easily finds examples where for the new scheme the dimensions are r on level 0 and 1 for level $\ell \geq 1$, whereas the s -term representation requires an exponentially increasing tensor rank $s := r^{d/2}$.

4 Basis Organisation

4.1 Basis Transformation

Later we will see that there is a reason to change a basis. Since we always require orthonormality, a new basis $\{\hat{\phi}^{\ell,j,\gamma}\}$ in $U^{\ell,j}$ satisfies

$$\hat{\phi}^{\ell,j,\gamma} = \sum_{\alpha=1}^{r^{\ell,j}} q_{\gamma\alpha} \phi^{\ell,j,\alpha}, \quad \phi^{\ell,j,\alpha} = \sum_{\gamma=1}^{r^{\ell,j}} q_{\gamma\alpha} \hat{\phi}^{\ell,j,\gamma},$$

with an unitary matrix $Q = (q_{\gamma\alpha}) \in \mathbb{R}^{r^{\ell,j} \times r^{\ell,j}}$. As a consequence, the coefficients c of a vector $u \in U^{\ell,j}$ (see (2.9)) change into $\hat{c} = Qc$.

Now we assume that the bases of the subspaces $U^{\ell-1,2j-1}$ and $U^{\ell-1,2j}$ from level $\ell - 1$ have changed and that the respective associated unitary matrices are $Q^{\ell-1,2j-1}$ and $Q^{\ell-1,2j}$ (if only one basis has changed, the other matrix Q equals I). The basis vector $\phi^{\ell,j,\gamma}$ from level ℓ has the coefficient matrix $A^{\ell,j,\gamma}$ with respect to the old bases of $U^{\ell-1,2j-1}$ and $U^{\ell-1,2j}$ (see (2.7)). One easily finds that the coefficient matrix $\hat{A}^{\ell,j,\gamma}$ with respect to the new bases of $U^{\ell-1,2j-1}$ and $U^{\ell-1,2j}$ is given by

$$\hat{A}^{\ell,j,\gamma} = Q^{\ell-1,2j-1} A^{\ell,j,\gamma} (Q^{\ell-1,2j})^\top.$$

Therefore, the computation of the new coefficient matrices $\{\hat{A}^{\ell,j,\gamma}\}$ costs $2r^{\ell,j}$ matrix multiplications. Changing all bases of level $\ell - 1$, the update at level ℓ requires

$$N_{\text{transf},\ell} \leq 2 \sum_{j=1}^{2^{L-\ell}} r^{\ell,j} r^{\ell-1,2j-1} r^{\ell-1,2j} (r^{\ell-1,2j-1} + r^{\ell-1,2j})$$

arithmetical operations.

Let $r := \max_j \{r^{\ell,j}, r^{\ell-1,j}\}$. Then $N_{\text{transf},\ell} \leq 4 \cdot 2^{L-\ell} r^4$ and

$$N_{\text{transf}} := \sum_{\ell=1}^L N_{\text{transf},\ell} \leq 4(d-1)r^4.$$

4.2 Orthonormalisation

As we have seen from Proposition 3.1, the space $U^{\ell,j}$ may be spanned by a general basis $\{\psi^{\ell,j,\gamma} : 1 \leq \gamma \leq r^{\ell,j}\}$ which needs to be orthonormalised. If $\ell = 0$, the orthonormalisation requires the scalar products of the ansatz spaces V^j . The corresponding cost depends on the nature of V^j .

Next we consider the case $\ell \geq 1$, where $U^{\ell,j} \subset U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$. We have to start from the representation of $\psi^{\ell,j,\gamma}$ by the coefficient matrix $B^{\ell,j,\gamma} \in \mathbb{R}^{r^{\ell-1,2j-1} \times r^{\ell-1,2j}}$ (see (2.5)):

$$\psi^{\ell,j,\gamma} = \sum_{\alpha=1}^{r^{\ell-1,2j-1}} \sum_{\beta=1}^{r^{\ell-1,2j}} b_{\alpha,\beta}^{\ell,j,\gamma} \phi^{\ell-1,2j-1,\alpha} \otimes \phi^{\ell-1,2j,\beta}, \quad 1 \leq \gamma \leq r^{\ell,j}.$$

The scalar products $\langle \psi^{\ell,j,\delta}, \psi^{\ell,j,\gamma} \rangle$ are realised by the sums $\sum_{\alpha,\beta} b_{\alpha,\beta}^{\ell,j,\delta} b_{\alpha,\beta}^{\ell,j,\gamma} = \langle B^{\ell,j,\delta}, B^{\ell,j,\gamma} \rangle_F$ (cf. Lemma 2.1), which cost $2r^{\ell-1,2j-1}r^{\ell-1,2j}$ operations. The orthonormalisation (QR -decomposition) requires in total

$$N_{\text{ortho}} \leq (r^{\ell,j} + 1)r^{\ell,j}r^{\ell-1,2j-1}r^{\ell-1,2j}$$

operations (the Gram-Schmidt method is of the same order).

Let $r := \max_{\ell,j} r^{\ell,j}$. Then the orthonormalisation for all $\ell \geq 1$ and j leads to $d(r + 1)r^3 = \mathcal{O}(dr^4)$ operations.

5 Basic Operations

5.1 Union of Two Φ -systems

Consider a φ -system and a ψ -system characterised by the respective bases

$$\begin{aligned} \{\varphi^{\ell,j,\gamma} : 1 \leq \ell \leq L, 1 \leq j \leq 2^{L-\ell}, 1 \leq \gamma \leq r^{\ell,j}\}, \\ \{\psi^{\ell,j,\gamma} : 1 \leq \ell \leq L, 1 \leq j \leq 2^{L-\ell}, 1 \leq \gamma \leq s^{\ell,j}\}, \end{aligned}$$

which span the subspaces $\varphi U^{\ell,j}$ and $\psi U^{\ell,j}$. We define the union of both system by means of the subspaces generated by the direct sums

$$U^{\ell,j} := \varphi U^{\ell,j} + \psi U^{\ell,j},$$

which obviously satisfy the requirement $U^{\ell,j} \subset U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$. The new dimension $\dim U^{\ell,j}$ is bounded by $r^{\ell,j} + s^{\ell,j}$. Concerning the basis of $U^{\ell,j}$, one can enrich $\{\varphi^{\ell,j,\gamma}\}_\gamma$ from $\varphi U^{\ell,j}$ by linear independent contributions from $\{\psi^{\ell,j,\gamma}\}_\gamma$ which need to be orthonormalised (for $s^{\ell,j} > r^{\ell,j}$ one should start from $\{\psi^{\ell,j,\gamma}\}_\gamma$). In principle, similar transformations are needed as for an orthonormalisation (see Sect. 4.2), but since $\{\varphi^{\ell,j,\gamma}\}_\gamma$ is already orthonormal, only a partial orthonormalisation is necessary, which simplifies the task.

5.2 Addition of Tensors

If two tensors \mathbf{t} and \mathbf{t}' are represented by the same Φ -system, the addition is trivial, since only the corresponding coefficient vectors are to be added (see (2.9), $\ell = L$).

Next, we assume that \mathbf{t} is represented by a φ -system, while \mathbf{t}' uses a different ψ -system. First, we have to construct the union of both systems and to represent \mathbf{t} and \mathbf{t}' in the new Φ -system. Then, the addition can be performed as above.

6 Truncation

Since the computational cost depends critically on the dimensions $r^{\ell,j}$, these numbers should be as small as possible. Operations like the addition increase the dimensions. Therefore, we need a reduction of $r^{\ell,j}$ by means of a truncation. Such a procedure will map a tensor \mathbf{t} onto an approximation $\tilde{\mathbf{t}}$. Here, we have to control the error $\|\mathbf{t} - \tilde{\mathbf{t}}\|$.

In the following, we describe the truncation process for a Φ -system, which will start at level L and proceed downwards to level 0. Here, we assume $\dim U^{L,1} = 1$ (cf. (2.3)). Otherwise, the procedure applied to the level $\ell < L$ is to be used also here.

6.1 Level L

The tensor \mathbf{t} is given in the form $t = \sum_{\gamma=1}^{r^{L,1}} c_{\gamma} \phi^{L,1,\gamma}$, where

$$\phi^{L,1,\gamma} = \sum_{\alpha=1}^{r^{L-1,1}} \sum_{\beta=1}^{r^{L-1,2}} a_{\alpha,\beta}^{L,1,\gamma} \phi^{L-1,1,\alpha} \otimes \phi^{L-1,2,\beta}.$$

This leads to

$$\mathbf{t} = \sum_{\alpha=1}^{r^{L-1,1}} \sum_{\beta=1}^{r^{L-1,2}} \left(\sum_{\gamma=1}^{r^{L,1}} c_{\gamma} a_{\alpha,\beta}^{L,1,\gamma} \right) \phi^{L-1,1,\alpha} \otimes \phi^{L-1,2,\beta},$$

i.e., t has the coefficient matrix $T := (\sum_{\gamma} c_{\gamma} a_{\alpha,\beta}^{L,1,\gamma})_{\alpha,\beta}$ with respect to the basis $\{\phi^{L-1,1,\alpha} \otimes \phi^{L-1,2,\beta}\}$. Now we apply singular value decomposition (SVD) to T ,

$$T = U \Sigma V^{\top}, \quad U, V \text{ orthogonal}, \Sigma = \text{diag}\{\sigma_1, \sigma_2, \dots\},$$

implying

$$\begin{aligned} \mathbf{t} &= \sum_{\alpha=1}^{r^{L-1,1}} \sum_{\beta=1}^{r^{L-1,2}} \left(\sum_{\gamma=1}^{\min\{r^{L-1,1}, r^{L-1,2}\}} U_{\alpha\gamma} \sigma_{\gamma} V_{\beta\gamma} \right) \phi^{L-1,1,\alpha} \otimes \phi^{L-1,2,\beta} \\ &= \sum_{\gamma=1}^{\min\{r^{L-1,1}, r^{L-1,2}\}} \sigma_{\gamma} \hat{\phi}^{L-1,1,\gamma} \otimes \hat{\phi}^{L-1,2,\gamma} \end{aligned} \tag{6.1}$$

$$\text{with } \hat{\phi}^{L-1,1,\gamma} := \sum_{\alpha=1}^{r^{L-1,1}} U_{\alpha\gamma} \phi^{L-1,1,\alpha}, \quad \hat{\phi}^{L-1,2,\gamma} := \sum_{\beta=1}^{r^{L-1,2}} V_{\beta\gamma} \phi^{L-1,2,\beta}. \quad (6.2)$$

Note that $\{\hat{\phi}^{L-1,1,\gamma}\}$ and $\{\hat{\phi}^{L-1,2,\gamma}\}$ is a new orthonormal basis in $U^{L,1}$ and $U^{L,2}$. In the case $r^{L-1,1} \neq r^{L-1,2}$, one of the bases is shorter than before, since γ varies in $\{1, \dots, \min\{r^{L-1,1}, r^{L-1,2}\}\}$. In the latter case, there are superfluous components in $U^{L,1}$ or $U^{L,2}$ which do not help in the representation and can be omitted without any loss of accuracy. The next step is to check the singular values σ_γ . If σ_γ is below a threshold $\sigma_\gamma \leq \varepsilon$, we may omit the respective terms in (6.1) together with the basis vectors $\hat{\phi}^{L-1,1,\gamma}$ and $\hat{\phi}^{L-1,2,\gamma}$ for these γ . This second truncation changes the tensor \mathbf{t} into $\tilde{\mathbf{t}}$, but the Euclidean norm is only changed by $\sqrt{\sum \sigma_\gamma^2}$, where the sum refers to the omitted γ .

The final results are:

- (1) a truncated tensor $\tilde{\mathbf{t}}$,
- (2) reduced orthonormal bases $\{\hat{\phi}^{L-1,1,\gamma} : 1 \leq \gamma \leq \hat{r}\}$ and $\{\hat{\phi}^{L-1,2,\gamma} : 1 \leq \gamma \leq \hat{r}\}$, where \hat{r} is maximal with the property $\sigma_\gamma > \varepsilon$ for all $1 \leq \gamma \leq \hat{r}$,
- (3) a weight factor σ_γ of the basis vectors $\hat{\phi}^{L-1,1,\gamma}$ and $\hat{\phi}^{L-1,2,\gamma}$. This weight factor indicates how important the basis vector is.

For the following steps, we rename $\hat{\phi}^{\ell,j,\gamma}$ again by $\phi^{\ell,j,\gamma}$ and the reduced dimension $\hat{r}^{\ell,j}$ by $r^{\ell,j}$.

6.2 Levels $\ell < L$

While before only one tensor \mathbf{t} had to be described, we now have several $\phi^{\ell,j,\gamma}$ to be described. Each $\phi^{\ell,j,\gamma}$ is represented by means of its coefficient matrix $A^{\ell,j,\gamma}$ and has a weight factor $\sigma_\gamma^{\ell,j}$. Since $\phi^{\ell,j,\gamma} \in U^{\ell-1,2j-1} \otimes U^{\ell-1,2j}$, we have to determine the most important components of $U^{\ell-1,2j-1}$ and $U^{\ell-1,2j}$ which are needed to approximate $\phi^{\ell,j,1}, \phi^{\ell,j,2}, \dots$ simultaneously.

The components of $U^{\ell-1,2j-1}$ are responsible for the approximation of the images of the matrices $A^{\ell,j,\gamma}$. Therefore, we apply SVD to the agglomerated matrix

$$\mathcal{A} := \begin{bmatrix} \sigma_1^{\ell,j} A^{\ell,j,1} & \sigma_2^{\ell,j} A^{\ell,j,2} & \dots & \sigma_{r^{\ell,j}}^{\ell,j} A^{\ell,j,r^{\ell,j}} \end{bmatrix} = U \Sigma V^\top.$$

As we are only interested in the matrices U and Σ , this requires the diagonalisation of

$$\mathcal{A} \mathcal{A}^\top = \sum_{\gamma=1}^{r^{\ell,j}} (\sigma_\gamma^{\ell,j})^2 A^{\ell,j,\gamma} (A^{\ell,j,\gamma})^\top = U D U^\top \quad \text{with } D = \Sigma \Sigma^\top.$$

The orthogonal matrix U defines a new basis $\{\hat{\phi}^{\ell-1,2j-1,\gamma}\}$ in $U^{\ell-1,2j-1}$ (analogously to the left formula in (6.2)). Moreover, the new basis vector get the weight factors from Σ allowing the truncation of those basis vector with sufficiently small weight.

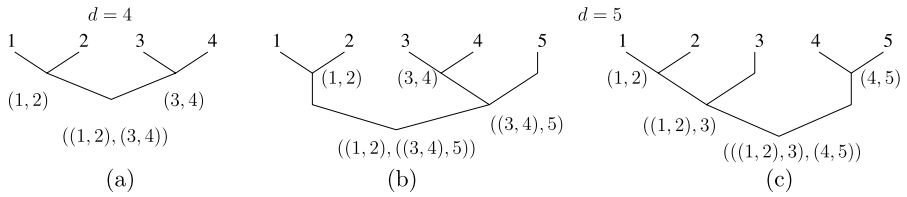


Fig. 1 Tree structures for $d = 4$ and $d = 5$

Concerning $U^{\ell-1,2j}$, one has to apply SVD to the matrix

$$B := \left[\sigma_1^{\ell,j} (A^{\ell,j,1})^\top \quad \sigma_2^{\ell,j} (A^{\ell,j,2})^\top \quad \dots \quad \sigma_{r^{\ell,j}}^{\ell,j} (A^{\ell,j,r^{\ell,j}})^\top \right]^\top = U \Sigma V^\top$$

yielding U, Σ, V different from those above. Again only V is of interest and is obtained from the diagonalisation

$$B^\top B = V D V^\top \quad \text{with } D = \Sigma^\top \Sigma.$$

The computation of the new basis $\{\hat{\phi}^{\ell-1,2j,\gamma}\}$ in $U^{\ell-1,2j}$ is similar to the right formula in (6.2). Again the weight factors from Σ can be used to omit some of the basis vectors.

The change of the bases is associated with the transformation of the coefficient matrices (cf. Sect. 4.1). If some of the basis vectors $\phi^{\ell-1,2j-1,\alpha}$ or $\phi^{\ell-1,2j,\beta}$ are omitted, they are missing in the representation (2.8) implying that the basis vectors $\phi^{\ell,j,\gamma}$ are no more orthonormal. Hence, an orthonormalisation may be needed.

If we assume that $r^{\ell,j} \leq r$ holds before the truncation and we truncate such that for the reduced dimensions $s^{\ell,j} \leq s$ hold afterwards, then the complexity of the truncation is given by $\mathcal{N}_{\text{trunc}} = \mathcal{O}(dr^3s)$. This can be improved, if the input tensor can be represented with less storage overhead.

An error analysis is given in [6], furthermore, numerical examples show, that one can treat problems up to dimensions $d = 1,000,000$.

7 Generalisation to Arbitrary d

For general integers d , the tree is not completely binary. Figure 1(b), (c) shows examples for $d = 5$. The suitable construction depends on the underlying problem. Even for $d = 4$ and the example from Sect. 3.4 the ordering $((1, 2), (3, 4))$ is essential.

References

1. Bungartz, H.-J., Griebel, M.: Sparse grids. Acta Numer. **13**, 147–269 (2004)
2. Chinnamsetty, S.R., Espig, M., Khoromskij, B.N., Hackbusch, W., Flad, H.-J.: Tensor product approximation with optimal rank in quantum chemistry. J. Chem. Phys. **127**, 084110 (2007), 14 pages
3. Dahmen, W., Prössdorf, S., Schneider, R.: Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution. Adv. Comput. Math. **1**, 259–335 (1993)
4. Espig, M.: Approximation mit Elementartensoren. Dissertation, Universität Leipzig (2008)

5. Espig, M., Grasedyck, L., Hackbusch, W.: Black box low tensor rank approximation using fibre-crosses. Preprint 60, Max-Planck-Institut für Mathematik, Leipzig (2008)
6. Grasedyck, L.: Hierarchical singular value decomposition of tensors. Preprint 27, Max-Planck-Institut für Mathematik, Leipzig (2009)
7. Grasedyck, L., Hackbusch, W.: Construction and arithmetics of \mathcal{H} -matrices. *Computing* **70**, 295–334 (2003)
8. Hackbusch, W., Khoromskij, B.N.: Tensor-product approximation to operators and functions in high dimensions. *J. Complex.* **23**, 697–714 (2007)
9. Khoromskij, B.N.: Structured rank- (r_1, \dots, r_d) decomposition of function-related tensors in \mathbb{R}^d . *Comp. Methods Appl. Math.* **6**, 194–220 (2006)
10. Khoromskij, B.N., Khoromskaia, V.: Low rank Tucker-type tensor approximation to classical potentials. *Cent. Eur. J. Math.* **5**(3), 1–18 (2007)
11. Tucker, L.R.: Implications of factor analysis of three-way matrices for measurement of change. In: Harris, C.W. (ed.) *Problems in Measuring Change*, pp. 122–137. University of Wisconsin Press, Madison (1963)