ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Mathematics Institute of Computational Science and Engineering

School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report

Nr. 12.2012

March 2012

The geometry of algorithms using hierarchical tensors

André Uschmajew, Bart Vandereycken

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

The geometry of algorithms using hierarchical tensors

André Uschmajew · Bart Vandereycken

16 January 2012

Abstract In this paper, the differential geometry of the novel hierarchical Tucker format for tensors is derived. The set $\mathcal{H}_{T,\mathbf{k}}$ of tensors with fixed tree T and hierarchical rank $\mathbf{k}$ is shown to be a smooth quotient manifold, namely the set of orbits of a Lie group action corresponding to the non-unique basis representation of these hierarchical tensors. Explicit characterizations of the quotient manifold, its tangent space and the tangent space of $\mathcal{H}_{T,\mathbf{k}}$ are derived, suitable for high-dimensional problems. The usefulness of a complete geometric description is demonstrated by two typical applications. First, new convergence results for the nonlinear Gauss–Seidel method on $\mathcal{H}_{T,\mathbf{k}}$ are given. Notably and in contrast to earlier works on this subject, the task of minimizing the Rayleigh quotient is also addressed. Second, evolution equations for dynamic tensor approximation are formulated in terms of an explicit projection operator onto the tangent space of $\mathcal{H}_{T,\mathbf{k}}$. In addition, a numerical comparison is made between this dynamical approach and the standard one based on truncated singular value decompositions.

Keywords High-dimensional tensors · low-rank approximation · hierarchical Tucker · differential geometry · Lie groups · nonlinear Gauss–Seidel · eigenvalue problems · time-varying tensors

Mathematics Subject Classification (2010) 15A69 · 18F15 · 57N35 · 41A46 · 65K10 · 15A18 · 65L05

1 Introduction

The development and analysis of efficient numerical methods for high-dimensional problems is a highly active area of research in numerical analysis. High dimensionality can occur in several scientific disciplines, the most prominent case being

André Uschmajew
Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany,
E-mail: uschmajew@math.tu-berlin.de

Bart Vandereycken
Chair of ANCHP, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland,
E-mail: bart.vandereycken@epfl.ch

when the computational domain of a mathematical model is embedded in a high-dimensional space, say $\mathbb{R}^d$ with $d > 3$. A prototypical example is the Schrödinger equation in quantum dynamics [6], but there is an abundance of other high-dimensional problems coming from radiative transport [55], option pricing in computational finance [32], and solving the master and Fokker–Planck equations [19]. High dimensionality may also be caused by parameters and uncertainties in the model [65].

In an abstract setting, all these examples involve a function u governed by an underlying continuous mathematical model on a domain $\Omega \subset \mathbb{R}^d$ with $d > 3$. For all but the simplest models, the approximation of u requires discretization. However, any straightforward approach to discretization inevitably leads to what is known as the *curse of dimensionality* [7]. For example, the approximation in a finite-dimensional basis with N degrees of freedom in each dimension leads to a total of N^d coefficients representing u. Already for moderate values of N and d, the storage of these coefficients becomes unmanageable, let alone their computation. Fortunately, and in contrast to low-dimensional models, high-dimensional models typically feature rather simple geometries, and sometimes admit highly regular solutions; see, e.g., [12, 73].

Existing approaches to the solution of high-dimensional problems exploit these structures to circumvent the curse of dimensionality by approximating the solution in a much lower dimensional space. Sparse grid techniques are well-suited and well-established for this purpose [11]. The setting in this paper stems from an alternative technique: the approximation of u by low-rank tensors [44]. In contrast to sparse grids, low-rank tensor techniques usually consider the full discretization from the start and face the challenges of high dimensionality at a later stage with linear algebra tools. In particular, we focus on tensors that can be expressed in the hierarchical Tucker format.

1.1 Rank-structured tensors

The *Hierarchical Tucker (HT) format* was introduced in [24] to overcome the disadvantages of earlier proposed low-rank tensor formats. Of these, the *Candecomp/Parafac (CP) decomposition* is the most classical one: assuming that u can be approximated by a short sum of separable functions,

$$u(x_1, x_2, \ldots, x_d) \approx \sum_{\mu=1}^{k} u_\mu^{(1)}(x_1) u_\mu^{(2)}(x_2) \cdots u_\mu^{(d)}(x_d),$$

its discretization $\mathbf{u}$ satisfies

$$\mathbf{u} \approx \sum_{\mu=1}^{k} \mathbf{u}_\mu^{(1)} \otimes \mathbf{u}_\mu^{(2)} \otimes \cdots \otimes \mathbf{u}_\mu^{(d)}, \tag{1.1}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{u}_\mu^{(l)}$ is the discretization of $u_\mu^{(l)}(x_l)$. Hence, instead of the N^d entries of $\mathbf{u}$, a CP decomposition of rank k only requires dkN entries.

Although a CP decomposition with small rank k has a favorable scaling with respect to the dimension and despite its initial success in high-dimensional applications (see, e.g., [8,9]), the approximation by CP tensors is notoriously ill-behaved. In particular, the set of tensors of bounded CP rank is not closed and the existence of a best approximation may fail to exist [15], which necessitates the use of regularization due to the ill-posedness of the approximation problem; see [44,3,17] for recent discussions. In fact, many basic concepts from linear algebra generalized to CP, like rank or best approximations, are NP hard, even to approximate [26].

On the other hand, the so-called *Tucker format* (see (2.2) for a definition) forms a mathematically more appealing set in the sense that it allows for the existence of best approximations. While the computation of a global solution for a best approximation is still an open problem, the fact that it always exists, makes the approximation problem well-posed and the Tucker format can be regarded as stable. Furthermore, it does admit quasi-optimal approximations by means of standard singular value decompositions (SVDs) [14], which for many applications is sufficient. However, its memory requirements still grow exponentially with the dimension d.

This disparity in properties has spurred the numerical analysis community to develop new tensor formats that aim at combining the advantages of CP (linear scaling in d) and Tucker (stable format for approximation): besides the HT format of [24] and its numerical treatment in [22], related decompositions termed the *(Q)TT format*, loosely abbreviating (quantized) tensor train or tree tensor, were proposed in [58,61,59,37,60].

Although developed independently at first, these formats are closely related to the notions of tensor networks (TNs) and matrix-product states (MPS) used in the density matrix renormalization group (DMRG) technique; see, e.g., [72,64,71,30]. While TNs and MPS have proven to be very effective in the computational physics community for many years, in this paper we focus only on networks that can be associated with a tree, that is, a graph without cycles. In particular, we restrict ourselves to general binary trees, but we conjecture that much of our derivations are readily extendible for arbitrary trees. The reason for restricting to trees is quite important: the loss of a tree structure in a tensor network not only significantly complicates their numerical treatment [71], it also makes them unattractive from a theoretical point of view. For example, the contraction of 2D tensor networks is known to be NP hard [5], while it was recently shown in [49] that a TN that is not a tree can fail to be Zariski closed, implying that the computation of a best approximation is ill-posed.

1.2 The differential geometry of rank-structured tensors

Besides their longstanding tradition in the computational physics community (see [71,31] for recent overviews), the (Q)TT and HT formats have more recently been used, amongst others, in [36,4,39,46,45,40,38,27] for solving a variety of high-dimensional problems ranging from the many-particle Schrödinger equation to parametric and stochastic PDEs. Notwithstanding the efficacy of all these methods, their theoretical understanding from a numerical analysis point of view is, however, rather limited.

For example, many of the existing algorithms rely on a so-called low-rank arithmetic that reformulates classic iterative algorithms to work with rank-structured tensors. Since the rank of the tensors will grow during the iterations, one needs to truncate the iterates back to low rank in order to keep the methods scalable and efficient. However, the impact of this low-rank truncation on the convergence and final accuracy has not been sufficiently analyzed for the majority of these algorithms. In theory, the success of such a strategy can only be guaranteed if the truncation error is kept negligible—in other words, on the level of roundoff error—but this is rarely feasible for practical problems due to memory and time constraints.

In the present paper, we want to make the case that treating the set of rank-structured tensors as a smooth manifold allows for theoretical and algorithmic improvements of these tensor-based algorithms.

First, there already exist geometry-inspired results in the literature for tensor-based methods. For example, [51] analyzes the multi-configuration time-dependent Hartree (MCTDH) method for many-particle quantum dynamics by relating this method to a flow problem on the manifold of fixed-rank Tucker tensors (in an infinite-dimensional setting). This allows to bound the error of this MCDTH method in terms of the best approximation error on the manifold. In [13] this analysis is refined to an investigation of the influence of certain singular values involved in the low-rank truncation and conditions for the convergence for MCTDH are formulated. A smooth manifold structure underlies this analysis; see also [52] for related methods. In a completely different context, [68] proves convergence of the popular alternating least-squares (ALS) algorithm applied to CP tensors by considering orbits of equivalent tensor representations in a geometric way. In the extension to the TT format [62] this approach is followed even more consequently.

Even if many of the existing tensor-based methods do not use the smooth structure explicitly, in some cases, they do so by the very nature of the algorithm. For example, [42,43] present a geometric approach to the problem of tracking a low-rank approximation to a time-varying matrix and tensor, respectively. After imposing a Galerkin condition for the update on the tangent space, a set of differential equations for the approximation on the manifold of fixed-rank matrices or Tucker tensors can be established. In addition to theoretical bounds on the approximation quality, the authors show in the numerical experiments that this continuous-time updating compares favorably to point-wise approximations like ALS; see also [56,33] for applications.

In addition, low-rank matrix nearness problems are a prototypical example for our geometric purpose since they can be solved by a direct optimization of a certain cost function on a suitable manifold of low-rank matrices. This allows the application of the framework of optimization on manifolds (see, e.g., [2] for an overview), which generalizes standard optimization techniques to manifolds. A number of large-scale applications [34,70,66,53,69] show that exploiting the smooth structure of the manifold indeed accelerates the algorithms compared to other, more ad-hoc ways of dealing with low-rank constraints. Moreover, there is no need for a heuristic truncation operator since the algorithm is constrained to a certain rank by construction. In all these cases it is important to establish a global smoothness on the manifold since one employs methods from smooth integration and optimization. A similar, but slightly different approach is taken in [27] where a modified ALS is adopted to exploit the tangent space of the TT manifold.

1.3 Contributions

The main contribution of this work is establishing a globally smooth differential structure for the set of tensors with fixed HT rank (which will be precisely defined in Section 3). The used tools from differential geometry are quite standard and not specifically related to hierarchical tensors. In fact, the whole theory could be perfectly explained with the special case of rank-one matrices in $\mathbb{R}^{2\times 2}$. But its application to HT tensors requires reasonably sophisticated matrix calculations, which in turn lead to important explicit formulas for the involved geometric objects. The theoretical and practical value of our geometric description is substantiated by two applications. First, local convergence results for the non-linear Gauss–Seidel method, which is popular in the context of optimization with tensors, are given. In particular we go further than the analogous results for TT tensors [62] by analyzing the computation of the minimal eigenvalue by a non-convex, self-adjoint Rayleigh quotient minimization, as proposed in [47]. Second, a dynamical tensor updating algorithm that integrates a flow on the manifold is derived.

In [27] it was conjectured that many of the there given derivations for the geometry of TT can be extended to HT. While this is in principle true, our derivation differs at crucial parts since the present analysis applies to the more challenging case of a generic HT decomposition (where the tree is no longer restricted to be linear or degenerate). In contrast to [27], we also establish a global smooth structure by employing the standard construction of a smooth Lie group action. Furthermore, we consider the resulting quotient manifold as a full-fledged manifold on its own and show how a horizontal space can be used to represent tangent vectors of this abstract manifold.

Since the applications mentioned in the introduction exploit low-rank structures to reduce the dimensionality, it is clear that the implementation of all the necessary geometric objects, like tangent vectors and projectors, needs to be scalable, that is, have a linear scaling in the order d and mode dimension N. In this paper, we shall make some effort to ensure that the implementation of our geometry can be done efficiently, although the actual treatment of numerical algorithms is not our main focus here and will be reported elsewhere.

2 Preliminaries

We briefly recall some necessary definitions and properties of tensors. A more comprehensive introduction can be found in the survey paper [44].

By a *tensor* $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ we mean a d-dimensional array with entries $\mathbf{X}_{i_1,i_2,\ldots,i_d} \in \mathbb{R}$. We call d the *order* of the tensor. Tensors of higher-order coincide with $d > 2$ and will always be denoted by boldface letter, e.g., $\mathbf{X}$. A *mode-k fiber* is defined as the vector by varying the k-th index of a tensor and fixing all the other ones. *Slices* are two-dimensional sections of a tensor, defined by fixing all but two indices.

A *mode-k unfolding* or *matricization*, denoted by $X^{(k)}$, is the result of arranging the mode-k fibers of $\mathbf{X}$ to be the columns of a matrix $X^{(k)}$ while the other modes are assigned to be rows. The ordering of these rows is taken to be *lexicographically*, that is, the index for n_i with the largest physical dimension i varies first[1]. Using multi-indices (that are also assumed to be enumerated lexicographically), the unfolding can be defined as

$$X^{(k)} \in \mathbb{R}^{n_k \times (n_1 \cdots n_{k-1} n_{k+1} \cdots n_d)} \quad \text{with} \quad X^{(k)}_{i_k,(i_1,\dots,i_{k-1},i_{k+1},\dots,i_d)} = \mathbf{X}_{i_1,i_2,\dots,i_d}.$$

The *vectorization* of a tensor $\mathbf{X}$, denoted by $\mathrm{vec}(\mathbf{X})$, is the rearrangement of all the fibers into the column vector

$$\mathrm{vec}(\mathbf{X}) \in \mathbb{R}^{n_1 n_2 \cdots n_d} \quad \text{with} \quad (\mathrm{vec}(\mathbf{X}))_{(i_1,i_2,\dots,i_d)} = \mathbf{X}_{i_1,i_2,\dots,i_d}.$$

Observe that when X is a matrix, $\mathrm{vec}(X)$ corresponds to collecting all the *rows* of X into one vector.

Let $X = A \otimes B$, defined with multi-indices as $X_{(i_1,i_2),(j_1,j_2)} = A_{i_1,j_1} B_{i_2,j_2}$, denote the Kronecker product of A and B. Then, given a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and matrices $A_k \in \mathbb{R}^{m_k \times n_k}$ for $k = 1, \dots, d$, the *multilinear multiplication*

$$(A_1, A_2, \dots, A_d) \circ \mathbf{X} = \mathbf{Y} \in \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_d}$$

is defined by

$$\mathbf{Y}_{i_1,i_2,\dots,i_d} = \sum_{j_1,j_2,\dots,j_d=1}^{n_1,n_2,\dots,n_d} (A_1)_{i_1,j_1} (A_2)_{i_2,j_2} \cdots (A_d)_{i_d,j_d} \mathbf{X}_{j_1,j_2,\dots,j_d},$$

which is equivalent to,

$$Y^{(k)} = A_k X^{(k)} (A_1 \otimes \cdots \otimes A_{k-1} \otimes A_{k+1} \otimes \cdots \otimes A_d)^{\mathsf{T}} \quad \text{for } k = 1, 2, \dots, d. \quad (2.1)$$

Our notation for the multilinear product adheres to the convention used in [15] and expresses that $\circ$ is a left action of $\mathbb{R}^{m_1 \times n_1} \times \mathbb{R}^{m_2 \times n_2} \times \cdots \times \mathbb{R}^{m_d \times n_d}$ onto $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. Denoting by I_n the $n \times n$ *identity matrix*, the following shorthand notation for the *mode-k product* will be convenient:

$$A_k \circ_k \mathbf{X} = (I_{n_1}, \dots, I_{n_{k-1}}, A_k, I_{n_{k+1}}, \dots, I_{n_d}) \circ \mathbf{X}.$$

We remark that the notation $\mathbf{X} \times_k A_k$ used in [44] coincides with $A_k \circ_k \mathbf{X}$ in our convention.

The *multilinear rank* of a tensor $\mathbf{X}$ is defined as the the tuple

$$\mathbf{k}(\mathbf{X}) = \mathbf{k} = (k_1, k_2, \dots, k_d) \quad \text{with} \quad k_i = \mathrm{rank}(X^{(i)}).$$

Let $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ have multilinear rank $\mathbf{k}$. Then $\mathbf{X}$ admits a *Tucker decomposition* of the form

$$\mathbf{X} = (U_1, U_2, \cdots, U_d) \circ \mathbf{C}, \quad (2.2)$$

with $\mathbf{C} \in \mathbb{R}^{k_1 \times k_2 \times \cdot \times k_d}$ and $U_i \in \mathbb{R}^{n_i \times k_i}$. It is well known (see, e.g. [15, (2.18)]) that the multilinear rank is invariant under a change of bases:

$$\mathbf{k}(\mathbf{X}) = \mathbf{k}((A_1, A_2, \dots, A_d) \circ \mathbf{X}) \quad \text{when} \quad A_k \in \mathrm{GL}_{n_k} \quad \text{for all } k = 1, 2, \dots, d, \quad (2.3)$$

where GL_n denotes the set of full-rank matrices in $\mathbb{R}^{n \times n}$, called *the general linear group*.

[1] In [44] a reverse lexicographical ordering is used.

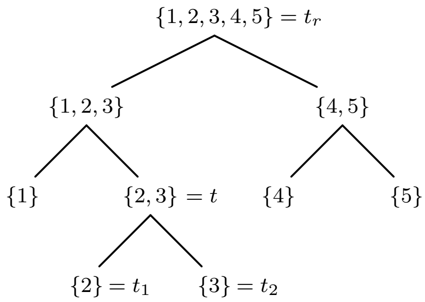

Fig. 3.1 A dimension tree of $\{1, 2, 3, 4, 5\}$.

3 The hierarchical Tucker decomposition

In this section, we define the set of HT tensors of fixed rank using the HT decomposition and the hierarchical rank. Much of this and related concepts were already introduced in [24, 22, 23] but we establish in addition some new relations on the parametrization of this set. In order to better facilitate the derivation of the smooth manifold structure in the next section, we have adopted a slightly different presentation compared to [24, 22].

3.1 The hierarchical Tucker format

Definition 3.1 Given the order d, a *dimension tree* T is a non-trivial, rooted binary tree whose nodes t can be labeled (and hence identified) by elements of the power set $\mathcal{P}(\{1, 2, \ldots, d\})$ such that

(i) the root has the label $t_r = \{1, 2, \ldots, d\}$; and,
(ii) every node $t \in T$, which is not a leaf, has two sons t_1 and t_2 that form an ordered partition of t, that is,

$$t_1 \cup t_2 = t \qquad \text{and} \qquad \mu < \nu \quad \text{for all } \mu \in t_1, \nu \in t_2. \tag{3.1}$$

The set of leafs is denoted by L. An example of a dimension tree for $t_r = \{1, 2, 3, 4, 5\}$ is depicted in Figure 3.1.

The idea of the HT format is to recursively factorize subspaces of $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ into tensor products of lower-dimensional spaces according to the index splittings in the tree T. If $\mathbf{X}$ is contained in such subspaces that allow for preferably low-dimensional factorizations, then $\mathbf{X}$ can be efficiently stored based on the next definition. Given dimensions $n_1, n_2, \ldots, n_d$, called *spatial* dimensions, and a node $t \subseteq \{1, 2, \ldots, d\}$, we define the dimension of t as $n_t = \prod_{\mu \in t} n_\mu$.

Definition 3.2 Let T be a dimension tree and $\mathbf{k} = (k_t)_{t \in T}$ a set of positive integers with $k_{t_r} = 1$. The *hierarchical Tucker (HT) format* for tensors $\mathbf{X} \in \mathbb{R}^{n_1, n_2, \ldots, n_d}$ is defined as follows.

(i) To each node $t \in T$, we associate a matrix $U_t \in \mathbb{R}^{n_t \times k_t}$.
(ii) For the root t_r, we define $U_{t_r} = \operatorname{vec}(\mathbf{X})$.

(iii) For each node t not a leaf with sons t_1 and t_2, there is a *transfer tensor* $\mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$ such that

$$U_t = (U_{t_1} \otimes U_{t_2})(B_t^{(1)})^\mathsf{T}, \tag{3.2}$$

where we recall that $B_t^{(1)}$ is the unfolding of $\mathbf{B}_t$ in the first mode.

When $\mathbf{X}$ admits such an HT decomposition in, we call $\mathbf{X}$ a $(T, \mathbf{k})$*-decomposable tensor.*

Remark 3.3 The restriction (3.1) to ordered splittings in the dimension tree guarantees that the recursive formula (3.2) produces matrices U_t whose rows are ordered lexicographically with respect to the indices of the spatial dimensions involved. The restriction to such splittings has been made for notational simplicity and is conceptually no loss in generality since relabeling nodes corresponds to permuting the modes (spatial dimensions) of $\mathbf{X}$.

It is instructive to regard (3.2) as a multilinear product operating on third-order tensors. Given $U_t \in \mathbb{R}^{n_t \times k_t}$ in a node that is not a leaf, define the third-order tensor

$$\widehat{\mathbf{U}}_t \in \mathbb{R}^{k_t \times n_{t_1} \times n_{t_2}} \quad \text{such that} \quad \widehat{U}_t^{(1)} = U_t^\mathsf{T}.$$

Then, from (3.2) and property (2.1) for the multilinear product, we get

$$\widehat{U}_t^{(1)} = U_t^\mathsf{T} = B_t^{(1)} (U_{t_1} \otimes U_{t_2})^\mathsf{T},$$

that is,

$$\widehat{\mathbf{U}}_t = (I_{k_t}, U_{t_1}, U_{t_2}) \circ \mathbf{B}_t.$$

We now explain the meaning of the matrices U_t. For $t \in T$, we denote by

$$t^c = \{1, 2, \ldots, d\} \setminus t$$

the set complimentary to t. A *mode-t unfolding* of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the result of reshaping $\mathbf{X}$ into a matrix by merging the indices belonging to $t = \{\mu_1, \mu_2, \ldots, \mu_p\}$ into row indices, and those belonging to $t^c = \{\nu_1, \nu_2, \ldots, \nu_{d-p}\}$ into column indices:

$$X^{(t)} \in \mathbb{R}^{n_t \times n_{t^c}} \quad \text{such that} \quad (X^{(t)})_{(i_{\mu_1}, i_{\mu_2}, \ldots, i_{\mu_p}), (i_{\nu_1}, i_{\nu_2}, \ldots, i_{\nu_{d-p}})} = \mathbf{X}_{i_1, \ldots, i_d}.$$

For the root t_r, the unfolding $X^{(t_r)}$ is set to be $\operatorname{vec}(\mathbf{X})$. The ordering of the multi-indices, both for the rows and columns of $X^{(t)}$, is again taken to be lexicographically.

By virtue of property (3.2) in Definition 3.2, the subspaces spanned by the columns of the U_t are nested along the tree. Since in the root U_{t_r} is fixed to be $X^{(t_r)} = \operatorname{vec}(\mathbf{X})$, this implies a relation between U_t and $X^{(t)}$ for the other nodes too.

Proposition 3.4 *For all $t \in T$ it holds* $\operatorname{span}(X^{(t)}) \subseteq \operatorname{span}(U_t)$.

Proof Assume that this holds at least for some node $t \in T \setminus L$ with sons t_1 and t_2. Then there exists a matrix $P_t \in \mathbb{R}^{k_t \times n_{t^c}}$ such that

$$X^{(t)} = U_t P_t = (U_{t_1} \otimes U_{t_2})(B_t^{(1)})^\mathsf{T} P_t. \tag{3.3}$$

First, define the third-order tensor $\mathbf{Y}_t$ as the result of reshaping $\mathbf{X}$:

$$\mathbf{Y}_t \in \mathbb{R}^{n_{t^c} \times n_{t_1} \times n_{t_2}} \quad \text{such that} \quad Y_t^{(1)} = (X^{(t)})^\mathsf{T} = (X^{(t_1 \cup t_2)})^\mathsf{T}.$$

Observe that by definition of a mode-k unfolding, the indices for the columns of $Y_t^{(1)}$ are ordered lexicographically, which means that the multi-indices of t_2 belong to the third mode of $\mathbf{Y}_t$. Hence, $Y_t^{(3)} = X^{(t_2)}$ and similarly, $Y_t^{(2)} = X^{(t_1)}$. Now we obtain from (3.3) that

$$Y_t^{(1)} = P_t^\mathsf{T} B_t^{(1)} (U_{t_1} \otimes U_{t_2})^\mathsf{T},$$

or, equivalently,

$$\mathbf{Y}_t = (P_t^\mathsf{T}, U_{t_1}, U_{t_2}) \circ \mathbf{B}_t.$$

Unfolding $\mathbf{Y}_t$ in the second or third mode, we get, respectively,

$$Y_t^{(2)} = X^{(t_1)} = U_{t_1} B_t^{(2)} (P_t^\mathsf{T} \otimes U_{t_2})^\mathsf{T}, \quad Y_t^{(3)} = X^{(t_2)} = U_{t_2} B_t^{(3)} (P_t^\mathsf{T} \otimes U_{t_1})^\mathsf{T}. \tag{3.4}$$

Hence, we have shown $\operatorname{span}(X^{(t_1)}) \subseteq \operatorname{span}(U_{t_1})$ and $\operatorname{span}(X^{(t_2)}) \subseteq \operatorname{span}(U_{t_2})$. Since the root vector U_{t_r} equals $X^{(t_r)} = \operatorname{vec}(\mathbf{X})$, the assertion follows by induction. □

Remark 3.5 In contrast to our definition (and the one in [24]), the hierarchical decomposition of [22] is *defined* to satisfy $\operatorname{span}(X^{(t)}) = \operatorname{span}(U_t)$. From a practical point of view, this condition is not restrictive since one can always choose a $(T, \mathbf{k})$-decomposition such that $\operatorname{span}(X^{(t)}) = \operatorname{span}(U_t)$ is also satisfied; see Proposition 3.6 below. Hence, the set of tensors allowing such decompositions is the same in both cases, but Definition 3.2 is more suitable for our purposes.

By virtue of relation (3.2), it is not necessary to know the U_t in all the nodes to reconstruct the full tensor $\mathbf{X}$. Instead, it is sufficient to store only the transfer tensors $\mathbf{B}_t$ in the nodes $t \in T \setminus L$ and the matrices U_t in the leafs $t \in L$. This is immediately obvious from the recursive definition but it is still instructive to inspect how the actual reconstruction is carried out. Let us examine this for the dimension tree of Figure 3.1.

The transfer tensors and matrices that need to be stored are visible in Figure 3.2. Let B_{123} be a shorthand notation for $B_{\{1,2,3\}}$ and I_{45} for $I_{k_{\{4,5\}}}$, and likewise for other indices. Then $\mathbf{X}$ can be reconstructed as follows:

$$\begin{aligned}
\operatorname{vec}(\mathbf{X}) &= (U_{123} \otimes U_{45})(B_{12345}^{(1)})^\mathsf{T} \\
&= [(U_1 \otimes U_{23})(B_{123}^{(1)})^\mathsf{T} \otimes (U_4 \otimes U_5)(B_{45}^{(1)})^\mathsf{T}](B_{12345}^{(1)})^\mathsf{T} \\
&= [U_1 \otimes (U_2 \otimes U_3)(B_{23}^{(1)})^\mathsf{T} \otimes U_4 \otimes U_5](B_{123}^{(1)} \otimes B_{45}^{(1)})^\mathsf{T} (B_{12345}^{(1)})^\mathsf{T} \\
&= (U_1 \otimes U_2 \otimes \cdots \otimes U_5)(I_1 \otimes B_{23}^{(1)} \otimes I_{45})^\mathsf{T} (B_{123}^{(1)} \otimes B_{45}^{(1)})^\mathsf{T} (B_{12345}^{(1)})^\mathsf{T}.
\end{aligned} \tag{3.5}$$

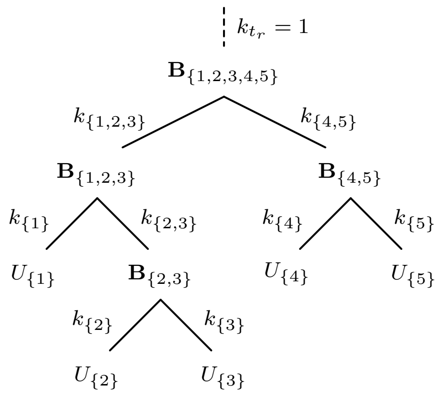

Fig. 3.2 The parameters of the HT format for the dimension tree of Figure 3.1.

Generalizing the example, we can parametrize all $(T, \mathbf{k})$-decomposable tensors by elements

$$x = (U_t, \mathbf{B}_t) = ((U_t)_{t \in L}, (\mathbf{B}_t)_{t \in T \setminus L}) \in M_{T,\mathbf{k}}$$

where

$$M_{T,\mathbf{k}} = \bigtimes_{t \in L} \mathbb{R}^{n_t \times k_t} \times \bigtimes_{t \in T \setminus L} \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}.$$

The hierarchical reconstruction of $\mathbf{X}$, given such an $x \in M_{T,\mathbf{k}}$, constitutes a mapping

$$f \colon M_{T,\mathbf{k}} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}.$$

For the tree of Figure 3.1, f is given by (3.5). Due to notational inconvenience, we refrain from giving an explicit definition of this reconstruction but the general pattern should be clear. Since the reconstruction involves only matrix multiplications and reshapings, it is smooth (C^∞). By definition, the image $f(M_{T,\mathbf{k}})$ consists of all $(T, \mathbf{k})$-decomposable tensors.

3.2 The hierarchical Tucker rank

A natural question arises in which cases the parametrization of $\mathbf{X}$ by $M_{T,\mathbf{k}}$ is *minimal*: given a dimension tree T, what are the nested subspaces with minimal dimension—in other words, what is the $x \in M_{T,\mathbf{k}}$ of smallest dimension such that $\mathbf{X} = f(x)$? The key concept turns out to be the *hierarchical Tucker rank* or *T-rank* of $\mathbf{X}$, denoted by $\operatorname{rank}_T(\mathbf{X})$, which is the tuple $\mathbf{k} = (k_t)_{t \in T}$ with

$$k_t = \operatorname{rank}(X^{(t)}).$$

Observe that by the dimensions of $X^{(t)}$ this choice of k_t implies

$$k_t \leq \min\{n_t, n_{t^c}\}. \tag{3.6}$$

In order to be able to write a tensor $\mathbf{X}$ in the $(T, \mathbf{k})$-format, it is by Proposition 3.4 necessary that $\operatorname{rank}_T(\mathbf{X}) \leq \mathbf{k}$, where this inequality is understood component-wise. As the next proposition shows, this condition is also sufficient. This is well known and can, for example, be found as Algorithm 1 in [22].

Proposition 3.6 *Every tensor* $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ *of* T*-rank bounded by* $\mathbf{k} = (k_t)_{t \in T}$ *can be written in the* $(T, \mathbf{k})$*-format. In particular, one can choose any set of matrices* $U_t \in \mathbb{R}^{n_t \times k_t}$ *satisfying*

$$\operatorname{span}(U_t) = \operatorname{span}(X^{(t)})$$

for all $t \in T$*. These can, for instance, be obtained as the left singular vectors of* $X^{(t)}$ *appended by* $k_t - \operatorname{rank}(X^{(t)})$ *zero columns.*

Proof If one chooses the U_t accordingly, the existence of transfer tensors $\mathbf{B}_t$ for a $(T, \mathbf{k})$-decomposition follows from the property

$$\operatorname{span}(X^{(t)}) \subseteq \operatorname{span}(X^{(t_1)} \otimes X^{(t_2)}),$$

which is shown in [21, Lemma 17] or [48, Lemma 2.1]. □

From now on we will mostly focus on the set of *hierarchical Tucker tensors of fixed* T*-rank* $\mathbf{k}$, denoted by

$$\mathcal{H}_{T,\mathbf{k}} = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} : \operatorname{rank}_T(\mathbf{X}) = \mathbf{k}\}.$$

This set is not empty only when (3.6) is satisfied. We emphasize that our definition again slightly differs from the definition of HT tensors of *bounded* T-rank $\mathbf{k}$ as used in [22], which is the union of all sets $\mathcal{H}_{T,\mathbf{r}}$ with $\mathbf{r} \leq \mathbf{k}$.

Before proving the next theorem, we state two basic properties involving rank. Let $A \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(A) = n$, then for all $B \in \mathbb{R}^{n \times p}$:

$$\operatorname{rank}(AB) = \operatorname{rank}(B). \tag{3.7}$$

In addition, for arbitrary matrices A, B it holds ([29, Theorem 4.2.15])

$$\operatorname{rank}(A \otimes B) = \operatorname{rank}(A) \operatorname{rank}(B). \tag{3.8}$$

Recall that $f : M_{T,\mathbf{k}} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ denotes the hierarchical construction of a tensor.

Theorem 3.7 *A tensor* $\mathbf{X}$ *is in* $\mathcal{H}_{T,\mathbf{k}}$ *if and only if for every* $x = (U_t, \mathbf{B}_t) \in M_{T,\mathbf{k}}$ *with* $f(x) = \mathbf{X}$ *the following holds:*

(i) *the matrices* U_t *have full column rank* k_t *(implying* $k_t \leq n_t$*) for* $t \in L$*, and*
(ii) *the tensors* $\mathbf{B}_t$ *have full multilinear rank* (k_t, k_{t_1}, k_{t_2}) *for all* $t \in T \setminus L$*.*

In fact, in that case the matrices U_t *have full column rank* k_t *for all* $t \in T$*.*

Interestingly, (i) and (ii) hence already guarantee $k_t \leq n_{t_c}$ for all $t \in T$.

Proof Assume that $\mathbf{X}$ has full T-rank $\mathbf{k}$. Then Proposition 3.4 gives that the matrices $U_t \in \mathbb{R}^{n_t \times k_t}$ have rank $k_t \leq n_t$ for all $t \in T$. So, by (3.8), all $U_{t_1} \otimes U_{t_2}$ have full column rank and, using (3.7), we obtain from (3.2) that $\operatorname{rank}(B_t^{(1)}) = k_t$ for all $t \in T \setminus L$. Additionally, the matrix P_t in (3.3) has to be of rank k_t (implying $k_t \leq n_{t^c}$) for all $t \in T$. From (3.4), we get the identities

$$P_{t_1} = B_t^{(2)}(P_t \otimes U_{t_2}^{\mathsf{T}}) \quad \text{and} \quad P_{t_2} = B_t^{(3)}(P_t \otimes U_{t_1}^{\mathsf{T}}). \tag{3.9}$$

Hence, using (3.8) again, $\operatorname{rank}(B_t^{(2)}) = k_{t_1}$ and $\operatorname{rank}(B_t^{(3)}) = k_{t_2}$ for all $t \in T \setminus L$.

Conversely, if the full rank conditions are satisfied for the leafs and the transfer tensors, it again follows, but this time from (3.2), that the U_t have full column rank (implying $k_t \leq n_t$) for all $t \in T$. Trivially, for the root node, relation (3.3) is satisfied for the full rank matrix $P_{t_r} = 1$ as scalar. Hence, by induction on (3.9), all P_t are of full column rank. Now, (3.3) implies $\operatorname{rank}(X^{(t)}) = k_t$ for all $t \in T$. □

Combining Theorem 3.7 with Proposition 3.4, one gets immediately the following result.

Corollary 3.8 *For every* $(T, \mathbf{k})$*-decomposition of* $\mathbf{X} \in \mathcal{H}_{T,\mathbf{k}}$*, it holds*

$$\operatorname{span}(X^{(t)}) = \operatorname{span}(U_t) \quad \textit{for all } t \in T.$$

Let $k_t \leq n_t$ for all $t \in T$. We denote by $\mathbb{R}_*^{n_t \times k_t}$ the matrices in $\mathbb{R}^{n_t \times k_t}$ of full column rank k_t and by $\mathbb{R}_*^{k_t \times k_{t_1} \times k_{t_2}}$ the tensors of full multilinear rank (k_t, k_{t_1}, k_{t_2}). We define

$$\mathcal{M}_{T,\mathbf{k}} = \bigtimes_{t \in L} \mathbb{R}_*^{n_t \times k_t} \times \bigtimes_{t \in T \setminus L} \mathbb{R}_*^{k_t \times k_{t_1} \times k_{t_2}}.$$

By the preceding theorem, $f(\mathcal{M}_{T,\mathbf{k}}) = \mathcal{H}_{T,\mathbf{k}}$ and $f^{-1}(\mathcal{H}_{T,\mathbf{k}}) = \mathcal{M}_{T,\mathbf{k}}$. Therefore, we call $\mathcal{M}_{T,\mathbf{k}}$ the *parameter space* of $\mathcal{H}_{T,\mathbf{k}}$. One can regard $\mathcal{M}_{T,\mathbf{k}}$ as an open and dense subset of $\mathbb{R}^D$ with

$$D = \dim(\mathcal{M}_{T,\mathbf{k}}) = \sum_{t \in L} n_t k_t + \sum_{t \in T \setminus L} k_t k_{t_1} k_{t_2}. \tag{3.10}$$

The restriction $f|_{\mathcal{M}_{T,\mathbf{k}}}$ will be denoted by

$$\phi\colon \mathcal{M}_{T,\mathbf{k}} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}, \ x \mapsto f(x). \tag{3.11}$$

Since $\mathcal{M}_{T,\mathbf{k}}$ is open (and dense) in M and f is smooth on M, ϕ is smooth on $\mathcal{M}_{T,\mathbf{k}}$.

Let $\mathbf{X}$ be a $(T, \mathbf{k})$-decomposable tensor with $K = \max\{k_t \colon t \in T\}$ and $N = \max\{n_1, n_2, \ldots, n_d\}$. Then, because of the binary tree structure, the dimension of the parameter space is bounded by

$$\dim(M_{T,\mathbf{k}}) = \dim(\mathcal{M}_{T,\mathbf{k}}) \leq dNK + (d-2)K^3 + K^2. \tag{3.12}$$

Compared to the full tensor $\mathbf{X}$ with N^d entries, this is indeed a significant reduction in the number of parameters to store when d is large and $K \ll N$.

In order not to overload notation in the rest of the paper, we will drop the explicit dependence on $(T, \mathbf{k})$ in the notation of $M_{T,\mathbf{k}}$, $\mathcal{M}_{T,\mathbf{k}}$, and $\mathcal{H}_{T,\mathbf{k}}$ where appropriate, and simply use M, $\mathcal{M}$, and $\mathcal{H}$, respectively. It will be clear from notation that the corresponding $(T, \mathbf{k})$ is silently assumed to be compatible.

3.3 The non-uniqueness of the decomposition

We now show that the HT decomposition is unique up to a change of bases. Although this seems to be well known (see, e.g., [21, Lemma 34]), we could not find a rigorous proof that this is the only kind of non-uniqueness in a full-rank decomposition.

Proposition 3.9 *Let $x = (U_t, \mathbf{B}_t) \in \mathcal{M}$ and $y = (V_t, \mathbf{C}_t) \in \mathcal{M}$. Then $\phi(x) = \phi(y)$ if and only if there exist (unique) invertible matrices $A_t \in \mathbb{R}_*^{k_t \times k_t}$ for every $t \in T \setminus t_r$ and $A_{t_r} = 1$ such that*

$$\begin{aligned} V_t &= U_t A_t \quad \text{for all } t \in L, \\ \mathbf{C}_t &= (A_t^{\mathsf{T}}, A_{t_1}^{-1}, A_{t_2}^{-1}) \circ \mathbf{B}_t \quad \text{for all } t \in T \setminus L. \end{aligned} \tag{3.13}$$

Proof Obviously, by (3.2), we have $\phi(x) = \phi(y)$ if y satisfies (3.13). Conversely, assume $\phi(x) = \phi(y) = \mathbf{X}$, then by Theorem 3.7, $\mathbf{X} \in \mathcal{H}$ and $\operatorname{rank}(U_t) = \operatorname{rank}(V_t) = k_t$ for all $t \in T$. Additionally, Corollary 3.8 gives $\operatorname{span}(U_t) = \operatorname{span}(X^{(t)}) = \operatorname{span}(V_t)$ for all $t \in T$. This implies

$$V_t = U_t A_t \quad \text{for all } t \in T \tag{3.14}$$

with unique invertible matrices A_t of appropriate sizes. Clearly, $A_{t_r} = 1$ since $V_{t_r} = U_{t_r} = \operatorname{vec}(\mathbf{X})$. By definition of the $(T, \mathbf{k})$-format, it holds

$$V_t = (V_{t_1} \otimes V_{t_2})(C_t^{(1)})^{\mathsf{T}} \quad \text{for all } t \in T \setminus L.$$

Inserting (3.14) into the above relation shows

$$U_t = (U_{t_1} \otimes U_{t_2})(A_{t_1} \otimes A_{t_2})(C_t^{(1)})^{\mathsf{T}} A_t^{-1}.$$

Applying (3.8) with the full column rank matrices U_t, we have that $U_{t_1} \otimes U_{t_2}$ is also of full column rank. Hence, due to (3.2),

$$B_t^{(1)} = A_t^{-\mathsf{T}} C_t^{(1)} (A_{t_1} \otimes A_{t_2})^{\mathsf{T}} \tag{3.15}$$

which together with (3.14) is (3.13). □

4 The smooth manifold of fixed rank

Our main goal is to show that the set $\mathcal{H} = \phi(\mathcal{M})$ of tensors of fixed T-rank $\mathbf{k}$ is an embedded submanifold of $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and describe its geometric structure.

Since the parametrization by $\mathcal{M}$ is not unique, the map ϕ is not injective. Fortunately, Proposition 3.9 allows us to identify the equivalence class of all possible $(T, \mathbf{k})$-decompositions of a tensor in $\mathcal{H}$ as the orbit of a Lie group action on $\mathcal{M}$. Using standard tools from differential geometry, we will see that the corresponding quotient space (the set of orbits) possesses itself a smooth manifold structure. It then remains to show that it is diffeomorphic to $\mathcal{H}$.

4.1 Orbits of equivalent representations

Let $\mathcal{G}$ be the Lie group

$$\mathcal{G} = \{\mathsf{A} = (A_t)_{t\in T} \colon A_t \in \mathrm{GL}_{k_t},\ A_{t_r} = 1\} \tag{4.1}$$

with the component-wise action of GL_{k_t} as group action. Let

$$\theta \colon \mathcal{M}\times\mathcal{G} \to \mathcal{M},\ (x,\mathsf{A}) := ((U_t,\mathbf{B}_t),(A_t)) \mapsto \theta_x(\mathsf{A}) := (U_t A_t, (A_t^\mathsf{T}, A_{t_1}^{-1}, A_{t_2}^{-1})\circ\mathbf{B}_t) \tag{4.2}$$

be a smooth, right action on $\mathcal{M}$. Observe that by property (2.3), $(A_t^\mathsf{T}, A_{t_1}^{-1}, A_{t_2}^{-1})\circ\mathbf{B}_t$ is of full multilinear rank. Hence, θ indeed maps to $\mathcal{M}$. In addition, the group $\mathcal{G}$ acts freely on $\mathcal{M}$, which means that the identity on $\mathcal{G}$ is the only element that leaves the action unchanged.

By virtue of Proposition 3.9, it is clear that the orbit of x,

$$\mathcal{G}x = \{\theta_x(\mathsf{A}) \colon \mathsf{A} \in \mathcal{G}\} \subseteq \mathcal{M},$$

contains all elements in $\mathcal{M}$ that map to the same tensor $\phi(x)$. This defines an equivalence relation on the parameterization as

$$x \sim y \text{ if and only if } y \in \mathcal{G}x.$$

The equivalence class of x is denoted by $\hat{x}$. Taking the quotient of $\sim$, we obtain the quotient space

$$\mathcal{M}/\mathcal{G} = \{\hat{x} \colon x \in \mathcal{M}\}$$

and the quotient map

$$\pi \colon \mathcal{M} \to \mathcal{M}/\mathcal{G},\ x \mapsto \hat{x}.$$

Finally, pushing ϕ down via π we obtain the injection

$$\widehat{\phi} \colon \mathcal{M}/\mathcal{G} \to \mathbb{R}^{n_1\times n_2\times\cdots\times n_d},\ \hat{x} \mapsto \phi(x), \tag{4.3}$$

whose image is $\mathcal{H}$.

Remark 4.1 As far as theory is concerned, the specific choice of parameterization $y \in \mathcal{G}x$ will be irrelevant as long as it is in the orbit. However, for numerical reasons, one usually chooses parametrizations such that

$$U_t^\mathsf{T} U_t = I_{k_t} \quad \text{for all } t \in T \setminus t_r. \tag{4.4}$$

This is called the *orthogonalized HT decomposition* and is advantageous for numerical stability [22,48] when truncating tensors or forming inner products (and more general contractions) between tensors, for example. On the other hand, the HT decomposition is inherently defined with subspaces which is clearly emphasized by our Lie group action.

4.2 Smooth quotient manifold

We now establish that the quotient space $\mathcal{M}/\mathcal{G}$ has a smooth manifold structure. By well-known results from differential geometry, we only need to establish that θ is a proper action; see, e.g., [16, Theorem 16.10.3] or [50, Theorem 9.16].

To show the properness of θ, we first observe that for $x = (U_t, \mathbf{B}_t)$ the inverse

$$\theta_x^{-1} \colon \mathcal{G}x \to \mathcal{G},\ y = (V_t, \mathbf{B}_t) \mapsto \theta_x^{-1}(y) = (A_t)_{t \in T} \tag{4.5}$$

is given by

$$\begin{aligned} A_t &= U_t^\dagger V_t \quad \text{for all } t \in L, \\ A_t &= [(B_t^{(1)})^\mathsf{T}]^\dagger (A_{t_1} \otimes A_{t_2})(C_t^{(1)})^\mathsf{T} \quad \text{for all } t \in T \setminus L, \end{aligned} \tag{4.6}$$

where $X^\dagger$ denotes the Moore–Penrose pseudo-inverse [20, Section 5.5.4] of X. In deriving (4.6), we have used the identities (3.14) and (3.15) and the fact that the unfolding $B_t^{(1)}$ has full row rank, since $\mathbf{B}_t$ has full multilinear rank.

Lemma 4.2 *Let (x_n) be a convergent sequence in $\mathcal{M}$ and (A_n) a sequence in $\mathcal{G}$ such that $(\theta_{x_n}(\mathsf{A}_n))$ converges in $\mathcal{M}$. Then (A_n) converges in $\mathcal{G}$.*

Proof Since U_t and $B_t^{(1)}$ have full rank, the pseudo-inverses in (4.6) are continuous. Hence, it is easy to see from (4.6) that $\theta_x^{-1}(y)$ is continuous with respect to $x = (U_t, \mathbf{B}_t)$ and $y = (V_t, \mathbf{C}_t)$. Hence $(\mathsf{A}_n) = (\theta_{x_n}^{-1}(\theta_{x_n}(\mathsf{A}_n)))$ converges in $\mathcal{G}$. □

Theorem 4.3 *The space $\mathcal{M}/\mathcal{G}$ possesses a unique smooth manifold structure such that the quotient map $\pi \colon \mathcal{M} \to \mathcal{M}/\mathcal{G}$ is a smooth submersion. Its dimension is*

$$\dim \mathcal{M}/\mathcal{G} = \dim \mathcal{M} - \dim \mathcal{G} = \sum_{t \in L} n_t k_t + \sum_{t \in T \setminus L} k_{t_1} k_{t_2} k_t - \sum_{t \in T \setminus \{t_r\}} k_t^2.$$

In addition, every orbit $\mathcal{G}x$ is an embedded submanifold in $\mathcal{M}$.

Proof From Lemma 4.2 we get that θ is a proper action [50, Proposition 9.13]. Since $\mathcal{G}$ acts properly, freely and smoothly on $\mathcal{M}$, it is well known that $\mathcal{M}/\mathcal{G}$ has a unique smooth structure with π a submersion; see, e.g., [50, Theorem 9.16]. The dimension of $\mathcal{M}/\mathcal{G}$ follows directly from counting the dimensions of $\mathcal{M}$ and $\mathcal{G}$. The assertion that $\pi^{-1}(x) = \mathcal{G}x$ is an embedded submanifold is direct from the property that π is a submersion [16, 16.8.8]. □

By our construction of $\mathcal{M}/\mathcal{G}$ as the quotient of a free right action, we have obtained a so-called *principal fiber bundle over $\mathcal{M}/\mathcal{G}$ with group $\mathcal{G}$ and total space* $\mathcal{M}$, see [54, Lemma 18.3]. Although fiber bundles are an important topic in differential geometry, we shall only use certain properties of them, like the existence of a principal connection in the next section and the gauge map in Section 4.5.

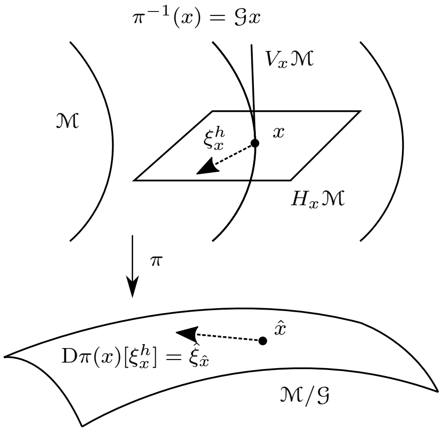

Fig. 4.1 A tangent vector $\hat{\xi}_{\hat{x}}$ and its unique horizontal lift ξ_x^h.

4.3 The horizontal space

Later on, we need the tangent space of $\mathcal{M}/\mathcal{G}$. Since $\mathcal{M}/\mathcal{G}$ is an abstract quotient of the concrete matrix manifold $\mathcal{M}$, we want use the tangent space of $\mathcal{M}$ to represent tangent vectors in $\mathcal{M}/\mathcal{G}$. Obviously, such a representation is not one-to-one since $\dim(\mathcal{M}) > \dim(\mathcal{M}/\mathcal{G})$. Fortunately, using the concept of horizontal lifts, this can be done rigorously as follows.

Since $\mathcal{M}$ is a dense and open subset of M, its tangent space is isomorphic to M,

$$T_x\mathcal{M} \simeq M = \bigtimes_{t\in L} \mathbb{R}^{n_t\times k_t} \times \bigtimes_{t\in T\setminus L} \mathbb{R}^{k_t\times k_{t_1}\times k_{t_2}},$$

and $\dim(T_x\mathcal{M}) = D$, with D defined in (3.10). Tangent vectors in $T_x\mathcal{M}$ will be denoted by ξ_x. The *vertical space*, denoted by $V_x\mathcal{M}$, is the subspace of $T_x\mathcal{M}$ consisting of the vectors tangent to the orbit $\mathcal{G}x$ through x. Since $\mathcal{G}x$ is an embedded submanifold of $\mathcal{M}$, this becomes

$$V_x\mathcal{M} = \left\{ \frac{\mathrm{d}}{\mathrm{d}s}\gamma(s)|_{s=0} \colon\ \gamma(s) \text{ smooth curve in } \mathcal{G}x \text{ with } \gamma(0) = x \right\}.$$

See also Figure 4.1, which we adapted from Figure 3.8 in [2].

Let $x = (U_t, \mathbf{B}_t) \in \mathcal{M}$. Then, taking the derivative in (4.2) for parameter-dependent matrices $A_t(s) = I_{k_t} + sD_t$ with arbitrary $D_t \in \mathbb{R}^{k_t\times k_t}$ and using the identity

$$\frac{\mathrm{d}}{\mathrm{d}s}(X(s)^{-1}) = -X^{-1}\frac{\mathrm{d}}{\mathrm{d}s}(X(s))X^{-1},$$

we get that the vertical vectors,

$$\xi_x^v = ((U_t^v \in \mathbb{R}^{n_t\times k_t})_{t\in L}, (\mathbf{B}_t^v \in \mathbb{R}^{k_t\times k_{t_1}\times k_{t_2}})_{t\in T\setminus L}) = (U_t^v, \mathbf{B}_t^v) \in V_x\mathcal{M},$$

have to be of the following form:

$$\begin{aligned}
&\text{for } t \in L: && U_t^v = U_t D_t, \quad D_t \in \mathbb{R}^{k_t\times k_t},\\
&\text{for } t \notin L \cup \{t_r\}: && \mathbf{B}_t^v = D_t^{\mathsf{T}} \circ_1 \mathbf{B}_t - D_{t_1} \circ_2 \mathbf{B}_t - D_{t_2} \circ_3 \mathbf{B}_t, \quad D_t \in \mathbb{R}^{k_t\times k_t},\\
&\text{for } t = t_r: && \mathbf{B}_t^v = -D_{t_1} \circ_2 \mathbf{B}_t - D_{t_2} \circ_3 \mathbf{B}_t.
\end{aligned} \tag{4.7}$$

Counting the degrees of freedom in the above expression, we obtain

$$\dim(V_x\mathcal{M}) = \sum_{t\in T\setminus\{t_r\}} k_t^2 = \dim(\mathcal{G}),$$

which shows that (4.7) indeed parametrizes the whole vertical space.

Next, the *horizontal space*, denoted by $H_x\mathcal{M}$, is any subspace of $T_x\mathcal{M}$ complementary to $V_x\mathcal{M}$. Horizontal vectors will be denoted by

$$\xi_x^h = ((U_t^h)_{t\in L}, (\mathbf{B}_t^h)_{t\in T\setminus L}) = (U_t^h, \mathbf{B}_t^h) \in H_x\mathcal{M}.$$

Since $H_x\mathcal{M} \subset T_x\mathcal{M} \simeq M$, the operation $x + \xi_x^h$ with $\xi_x^h \in H_x\mathcal{M}$ is well-defined as a partitioned addition of matrices. Then, the geometrical meaning of the affine space $x + H_x\mathcal{M}$ is that, for fixed x, it intersects every orbit in a neighborhood of x exactly once, see again Figure 4.1. Thus it can be interpreted as a local realization of the orbit manifold $\mathcal{M}/\mathcal{G}$.

Proposition 4.4 *Let $x \in \mathcal{M}$. Then $\pi|_{x+H_x\mathcal{M}}$ is a local diffeomorphism in a neighborhood of $\hat{x} = \pi(x)$ in $\mathcal{M}/\mathcal{G}$.*

Proof Observe that $x + H_x\mathcal{M}$ and $\mathcal{M}/\mathcal{G}$ have the same dimension. Since π is a submersion (Theorem 4.3), its rank is $\dim(\mathcal{M}/\mathcal{G})$. Being constant on $\mathcal{G}x$ implies $\mathrm{D}\pi(x)|_{V_x\mathcal{M}} = 0$. Hence, the map $\pi|_{x+H_x\mathcal{M}}$ has rank $\dim(\mathcal{M}/\mathcal{G}) = \dim(H_x\mathcal{M})$ and is therefore a submersion (also an immersion). The claim is then clear from the inverse function theorem, see [50, Corollary 7.11]. □

In light of the forthcoming derivations, we choose the following particular horizontal space:

$$H_x\mathcal{M} = \left\{ (U_t^h, \mathbf{B}_t^h) \colon \begin{array}{rl} (U_t^h)^{\mathsf{T}} U_t = 0 & \text{for } t \in L, \\ (B_t^h)^{(1)} (U_{t_1}^{\mathsf{T}} U_{t_1} \otimes U_{t_2}^{\mathsf{T}} U_{t_2}) (B_t^{(1)})^{\mathsf{T}} = 0 & \text{for } t \notin L \cup \{t_r\} \end{array} \right\}. \tag{4.8}$$

Observe that there is no condition on $\mathbf{B}_{t_r}^h$ (which is actually a matrix). Since all $U_{t_1}^{\mathsf{T}} U_{t_1} \otimes U_{t_2}^{\mathsf{T}} U_{t_2}$ are symmetric and positive definite, it is obvious that the parametrization above defines a linear subspace of $T_x\mathcal{M}$. To determine its dimension, observe that the orthogonality constraints for U_t^h and $\mathbf{B}_t^h$ take away k_t^2 degrees of freedom for all the nodes t except the root (due to the full ranks of the U_t and $B_t^{(1)}$). Hence, we have

$$\begin{aligned} \dim(H_x\mathcal{M}) &= \sum_{t\in L} (n_t k_t - k_t^2) \;+ \sum_{t\in T\setminus(L\cup\{t_r\})} (k_{t_1} k_{t_2} k_t - k_t^2) \;+\; k_{(t_r)_1} k_{(t_r)_2} \\ &= \sum_{t\in L} n_t k_t \;+ \sum_{t\in T\setminus L} k_{t_1} k_{t_2} k_t - \sum_{t\in T\setminus\{t_r\}} k_t^2 \\ &= \dim(T_x\mathcal{M}) - \dim(V_x\mathcal{M}), \end{aligned}$$

from which we can conclude that $V_x\mathcal{M} \oplus H_x\mathcal{M} = T_x\mathcal{M}$. (We omit the proof that the sum is direct, for it will follow from Lemma 4.9 and $\mathrm{D}\phi(x)|_{V_x\mathcal{M}} = 0$.)

Our choice of horizontal space has the following interesting property, which we will need for the main result of this section.

Proposition 4.5 *The horizontal space $H_x\mathcal{M}$ defined in* (4.8) *is invariant under the right action θ in* (4.2)*, that is,*

$$\mathrm{D}\theta(x,\mathsf{A})[H_x\mathcal{M},\mathsf{0}] = H_{\theta_x(\mathsf{A})}\mathcal{M} \quad \text{for any } x\in\mathcal{M} \text{ and } \mathsf{A}\in\mathcal{G}.$$

Proof Let $x=(U_t,\mathbf{B}_t)\in\mathcal{M}$, $\xi_x^h\in H_x\mathcal{M}$, $\mathsf{A}\in\mathcal{G}$ and $y=(V_t,\mathbf{C}_t)=\theta_x(\mathsf{A})$. Since θ depends linearly on x it holds

$$\eta=(V_t^\eta,\mathbf{C}_t^\eta)=\mathrm{D}\theta(x,\mathsf{A})[\xi_x^h,0]=\theta_{\xi_x^h}(\mathsf{A})=(U_t^hA_t,(A_t^\mathsf{T},A_{t_1}^{-1},A_{t_2}^{-1})\circ\mathbf{B}_t^h).$$

Verifying

$$\begin{aligned} V_t^\mathsf{T}V_t^\eta &= A_t^\mathsf{T}U_t^\mathsf{T}U_t^hA_t=0\\ (C_t^\eta)^{(1)}(V_{t_1}^\mathsf{T}V_{t_1}\otimes V_{t_2}^\mathsf{T}V_{t_2})(C_t^{(1)})^\mathsf{T} &= A_t^\mathsf{T}(B_t^h)^{(1)}(U_{t_1}^\mathsf{T}U_{t_1}\otimes U_{t_2}^\mathsf{T}U_{t_2})(B_t^{(1)})^\mathsf{T}A_t=0,\end{aligned}$$

we see from (4.8) applied to y that $\eta\in H_y\mathcal{M}$. Thus $\mathrm{D}\theta(x,\mathsf{A})[H_x\mathcal{M},\mathsf{0}]\subseteq H_y\mathcal{M}$. Since $H_x\mathcal{M}$ and $H_y\mathcal{M}$ have the same dimension, the assertion follows from the injectivity of the map $\xi\mapsto\mathrm{D}\theta(x,\mathsf{A})[\xi,0]=\theta_\xi(\mathsf{A})$, which is readily established from the full rank properties of A. □

Remark 4.6 For t not the root, the tensors $\mathbf{B}_t^h$ in the horizontal vectors in (4.8) satisfy a certain orthogonality condition with respect to a Euclidean inner product weighted by $U_{t_1}^\mathsf{T}U_{t_1}\otimes U_{t_2}^\mathsf{T}U_{t_2}$. In case the representatives are normalized in the sense of (4.4), this inner product becomes the standard Euclidean one.

The horizontal space we just introduced is in fact a principal connection on the principal $\mathcal{G}$-bundle. The set of all horizontal subspaces constitutes a distribution in $\mathcal{M}$. Its usefulness in the current setting lies in the following theorem, which is again visualized in Figure 4.1.

Theorem 4.7 *Let $\hat\xi$ be a smooth vector field on $\mathcal{M}/\mathcal{G}$. Then there exists a unique smooth vector field ξ^h on $\mathcal{M}$, called the* horizontal lift *of $\hat\xi$, such that*

$$\mathrm{D}\pi(x)[\xi_x^h]=\hat\xi_{\hat x} \quad \text{and} \quad \xi_x^h\in H_x\mathcal{M}, \tag{4.9}$$

for every $x\in\mathcal{M}$. In particular, for any smooth function $\hat h\colon\mathcal{M}/\mathcal{G}\to\mathbb{R}$, it holds

$$\mathrm{D}\hat h(\hat x)[\hat\xi_{\hat x}]=\mathrm{D}h(x)[\xi_x^h], \tag{4.10}$$

with $h=\hat h\circ\pi\colon\mathcal{M}\to\mathbb{R}$ a smooth function that is constant on the orbits $\mathcal{G}x$.

Proof Observe that $H_x\mathcal{M}$ varies smoothly in x since the orthogonality conditions in (4.8) involve full-rank matrices. In that case the existence of unique smooth horizontal lifts is a standard result for fiber bundles where the connection (that is, our choice of horizontal space $H_x\mathcal{M}$) is right-invariant (as shown in Proposition 4.5); see, e.g, [41, Proposition II.1.2]. Relation (4.10) is trivial after applying the chain rule to $h=\hat h\circ\pi$ and using (4.9). □

4.4 The embedding

In this section, we finally prove that $\mathcal{H} = \widehat{\phi}(\mathcal{G}/\mathcal{M})$ is an embedded submanifold in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ by showing that $\widehat{\phi}$ defined in (4.3) is an embedding, that is, an injective homeomorphism onto its image with injective differential. Since $\widehat{\phi} = \phi \circ \pi$, we will perform our derivation via ϕ.

Recall from Section 3.2 that the smooth mapping $\phi \colon \mathcal{M} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ represents the hierarchical construction of a tensor in $\mathcal{H}$. Let $x = (U_t, \mathbf{B}_t) \in \mathcal{M}$, then $U_{t_r} = \mathrm{vec}(\phi(x))$ can be computed recursively from

$$U_t = (U_{t_1} \otimes U_{t_2})(B_t^{(1)})^\mathsf{T} \quad \text{for all } t \in T \setminus L. \tag{4.11}$$

First, we show that the derivative of $\phi(x)$,

$$\mathrm{D}\phi(x) \colon T_x\mathcal{M} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d},\ \xi_x \mapsto \mathrm{D}\phi(x)[\xi_x],$$

can be computed using a similar recursion. Namely, differentiating (4.11) with respect to $(U_{t_1}, U_{t_2}, \mathbf{B}_t)$ gives

$$\delta U_t = (\delta U_{t_1} \otimes U_{t_2})(B_t^{(1)})^\mathsf{T} + (U_{t_1} \otimes \delta U_{t_2})(B_t^{(1)})^\mathsf{T} + (U_{t_1} \otimes U_{t_2})(\delta B_t^{(1)})^\mathsf{T} \ \text{ for } t \in T \setminus L, \tag{4.12}$$

where the δU_t denotes a standard differential (an infinitesimal variation) of U_t. Since this relation holds at all inner nodes, the derivative $\mathrm{vec}(\mathrm{D}\phi(x)[\xi_x]) = \delta U_{t_r}$ can be recursively calculated from the variations of the leafs and of the transfer tensors, which will be collected in the tangent vector

$$\xi_x = (\delta U_t, \delta \mathbf{B}_t) \in T_x\mathcal{M}.$$

Next, we state three lemmas in preparation for the main theorem. The first two deal with $\mathrm{D}\phi$ when its argument is restricted to the horizontal space $H_x\mathcal{M}$ in (4.8), while the third states that $\widehat{\phi}$ is an immersion.

Lemma 4.8 *Let $x = (U_t, \mathbf{B}_t) \in \mathcal{M}$ and $\xi_x^h = (U_t^h, \mathbf{B}_t^h) \in H_x\mathcal{M}$. Apply recursion (4.12) for evaluating $\mathrm{D}\phi(x)[\xi_x^h]$ and denote for $t \in T \setminus L$ the intermediate matrices δU_t by U_t^h. Then it holds*

$$U_t^\mathsf{T} U_t^h = 0_{k_t \times k_t} \quad \textit{for all } t \in T \setminus \{t_r\}. \tag{4.13}$$

Proof By definition of $H_x\mathcal{M}$, this holds for the leafs $t \in L$. Now take any $t \notin L \cup \{t_r\}$ for which (4.13) is satisfied for the sons t_1 and t_2. Then by (4.11) and (4.12), we have

$$U_t = (U_{t_1} \otimes U_{t_2})(B_t^{(1)})^\mathsf{T},$$
$$U_t^h = (U_{t_1}^h \otimes U_{t_2})(B_t^{(1)})^\mathsf{T} + (U_{t_1} \otimes U_{t_2}^h)(B_t^{(1)})^\mathsf{T} + (U_{t_1} \otimes U_{t_2})((B_t^h)^{(1)})^\mathsf{T}.$$

Together with the definition of $\mathbf{B}_t^h$ in (4.8), it is immediately clear that $U_t^\mathsf{T} U_t^h = 0$. The assertion follows by induction. □

Lemma 4.9 *Let $x = (U_t, \mathbf{B}_t) \in \mathcal{M}$. Then $\mathrm{D}\phi(x)|_{H_x\mathcal{M}}$ is injective.*

Proof Let $\xi_x^h = (U_t^h, \mathbf{B}_t^h) \in H_x\mathcal{M}$ with $\mathrm{D}\phi(x)[\xi_x^h] = 0$. Applying (4.12) to the root node, we have

$$\mathrm{vec}(\mathrm{D}\phi(x)[\xi_x^h]) = (U^h_{(t_r)_1} \otimes U_{(t_r)_2})(B^{(1)}_{t_r})^\mathsf{T} + (U_{(t_r)_1} \otimes U^h_{(t_r)_2})(B^{(1)}_{t_r})^\mathsf{T} + (U_{(t_r)_1} \otimes U_{(t_r)_2})((B^h_{t_r})^{(1)})^\mathsf{T}, \quad (4.14)$$

where we again used the notation U_t^h instead of δU_t for the inner nodes. According to Lemma 4.8, matrix $U^h_{(t_r)_1}$ is perpendicular to $U_{(t_r)_1}$, and similar for $(t_r)_2$. Hence, the Kronecker product matrices in the above relation span mutually linearly independent subspaces, so that the condition $\mathrm{D}\phi(x)[\xi_x^h] = 0$ is equivalent to

$$(U^h_{(t_r)_1} \otimes U_{(t_r)_2})(B^{(1)}_{t_r})^\mathsf{T} = 0 \quad (4.15)$$

$$(U_{(t_r)_1} \otimes U^h_{(t_r)_2})(B^{(1)}_{t_r})^\mathsf{T} = 0 \quad (4.16)$$

$$(U_{(t_r)_1} \otimes U_{(t_r)_2})((B^h_{t_r})^{(1)})^\mathsf{T} = 0. \quad (4.17)$$

Since $U_{(t_r)_1}$ and $U_{(t_r)_2}$ both have full column rank (see Theorem 3.7), we get immediately from (4.17) that $(B^h_{t_r})^{(1)}$ and hence $\mathbf{B}^h_{t_r} = 0$ need to vanish. Next, we can rewrite (4.15) as

$$(I_{k_{t_r}}, U^h_{(t_r)_1}, U_{(t_r)_2}) \circ \mathbf{B}_{t_r} = 0,$$

or,

$$(U^h_{(t_r)_1}) B^{(2)}_{t_r} (I_{k_{t_r}} \otimes U_{(t_r)_2})^\mathsf{T} = 0.$$

Since $B^{(2)}_{t_r}$ has full rank, one gets $U^h_{(t_r)_1} = 0$. In the same way one shows that (4.16) reduces to $U^h_{(t_r)_2} = 0$. Applying induction, one obtains $U_t^h = 0$ for all $t \in T$ and $\mathbf{B}_t = 0$ for all $t \in T \setminus L$. In particular, $\xi_x^h = 0$. □

Lemma 4.10 *The map* $\widehat{\phi} : \mathcal{M}/\mathcal{G} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ *as defined in* (4.3) *is an injective immersion.*

Proof Smoothness of $\widehat{\phi}$ follows by pushing down the smooth map ϕ through the quotient, see [50, Prop. 7.17]. Injectivity is trivial by construction. Fix $x \in \mathcal{M}$. To show the injectivity of $\mathrm{D}\widehat{\phi}(\hat{x})$ let $\hat{\xi}_{\hat{x}} \in T_{\hat{x}}\mathcal{G}/\mathcal{M}$ such that $\mathrm{D}\widehat{\phi}(\hat{x})[\hat{\xi}_{\hat{x}}] = 0$. We can assume [50, Lemma 4.5] that $\hat{\xi}_{\hat{x}}$ is part of a smooth vector field on $\mathcal{M}/\mathcal{G}$. By Theorem 4.7, there exists a unique horizontal lift $\xi_x^h \in H_x\mathcal{M}$ which satisfies

$$\mathrm{D}\phi(x)[\xi_x^h] = \mathrm{D}\widehat{\phi}(\hat{x})[\hat{\xi}_{\hat{x}}] = 0.$$

According to Lemma 4.9, this implies $\xi_x^h = 0$, which by (4.9) means $\hat{\xi}_{\hat{x}} = 0$. □

We are now ready to prove the embedding of $\mathcal{H}$ as a submanifold.

Theorem 4.11 *The set* $\mathcal{H}$ *is a smooth, embedded submanifold of* $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. *Its dimension is*

$$\dim(\mathcal{H}_{T,\mathbf{k}}) = \dim(\mathcal{M}/\mathcal{G}) = \sum_{t \in L} n_t k_t + \sum_{t \in T \setminus L} k_{t_1} k_{t_2} k_t - \sum_{t \in T \setminus \{t_r\}} k_t^2.$$

In particular, the map $\widehat{\phi} \colon \mathcal{M}/\mathcal{G} \to \mathcal{H}$ *is a diffeomorphism.*

Proof Due to Lemma 4.10 we only need to establish that $\widehat{\phi}$ is a homeomorphism onto its image $\mathcal{H}_{T,\mathbf{k}}$ in the subspace topology. The theorem then follows from standard results; see, e.g., [50, Theorem 8.3]. The dimension has been determined in Theorem 4.3.

The continuity of $\widehat{\phi}$ is immediate since it is smooth. We have to show that $\widehat{\phi}^{-1} : \mathcal{H} \to \mathcal{M}/\mathcal{G}$ is continuous. Let $(\mathbf{X}_n) \subseteq \mathcal{H}$ be a sequence that converges to $\mathbf{X}^* \in \mathcal{H}$. This means that every unfolding also converges,

$$X_n^{(t)} \to \ X_*^{(t)} \quad \text{for all } t \in T.$$

By definition of the T-rank, every sequence $(X_n^{(t)})$ and its limit are of rank k_t. Hence it holds

$$\operatorname{span}(X_n^{(t)}) \to \ \operatorname{span}(X_*^{(t)}) \quad \text{for all } t \in T \tag{4.18}$$

in the sense of subspaces [20, Section 2.6.3]. In particular, we can interpret the previous sequence as a converging sequence in $\mathrm{Gr}(n_t, k_t)$, the Grassmann manifold of k_t-dimensional linear subspaces in $\mathbb{R}^{n_t}$. It is well known [16] that $\mathrm{Gr}(n_t, k_t)$ can be seen as the quotient of $\mathbb{R}_*^{n_t \times k_t}$ by GL_{k_t}. Therefore, we can alternatively take matrices

$$U_t^* \in \mathbb{R}_*^{n_t \times k_t} \quad \text{with} \quad \operatorname{span}(U_t^*) = \operatorname{span}(X_*^{(t)}) \quad \text{for } t \in T \setminus \{t_r\}$$

as representatives of the limits, while for t_r we choose

$$U_{t_r}^* = X_*^{(t_r)} = \operatorname{vec}(\mathbf{X}^*). \tag{4.19}$$

Now, it can be shown [1, Eq. (7)] that the map

$$S_t \colon \mathrm{Gr}(n_t, k_t) \to \mathbb{R}_*^{n_t \times k_t}, \ \operatorname{span}(V) \mapsto V[(U_t^*)^\mathsf{T} V]^{-1} (U_t^*)^\mathsf{T} U_t^*,$$

where $V \in \mathbb{R}_*^{n_t \times k_t}$ is any matrix representation of $\operatorname{span}(V)$, is a local diffeomorphism onto its range. Thus, by (4.18), it holds for all $t \in T \setminus \{t_r\}$ that

$$U_t^n = S_t(\operatorname{span}(X_n^{(t)})) \to S_t(\operatorname{span}(X_*^{(t)})) = S_t(\operatorname{span}(U_t^*)) = U_t^*. \tag{4.20}$$

Also observe from the definition of S_t that we have

$$\operatorname{span}(U_t^n) = \operatorname{span}(X_n^{(t)}) \quad \text{for all } n \in \mathbb{N} \text{ and } t \in T \setminus \{t_r\}.$$

We again treat the root separately by setting

$$U_{t_r}^n = X_n^{(t_r)} = \operatorname{vec}(\mathbf{X}_n). \tag{4.21}$$

If we now choose transfer tensors $\mathbf{B}_t^n$ and $\mathbf{B}_t^*$ as

$$((B_t^n)^{(1)})^\mathsf{T} = (U_{t_1}^n \otimes U_{t_2}^n)^\dagger U_t^n \quad \text{and} \quad ((B_t^*)^{(1)})^\mathsf{T} = (U_{t_1}^* \otimes U_{t_2}^*)^\dagger U_t^*$$

for all $t \in T \setminus L$, then the nestedness property (3.2) will be satisfied. Moreover, (4.20) and (4.21) imply

$$\mathbf{B}_t^n \to \mathbf{B}_t^* \quad \text{for all } t \in T \setminus L. \tag{4.22}$$

Taking (4.19) and (4.21) into account, we see that $x_n = (U_t^n, \mathbf{B}_t^n)$ and $x^* = (U_t^*, \mathbf{B}_t^*)$ are $(T, \mathbf{k})$-decompositions of $\mathbf{X}_n$ and $\mathbf{X}^*$, respectively. According to Theorem 3.7, all x_n and x^* belong to $\mathcal{M}$. Hence, using (4.19)–(4.22) and the continuity of π (Theorem 4.3), we have proven

$$\widehat{\phi}^{-1}(\mathbf{X}_n) = \pi(x_n) \to \pi(x_*) = \widehat{\phi}^{-1}(\mathbf{X}_*),$$

that is, $\widehat{\phi}^{-1}$ is continuous. □

The embedding of $\mathcal{H}$ via $\widehat{\phi}$, although global, is somewhat too abstract. From Theorem 4.3 we immediately see that $\phi = \widehat{\phi} \circ \pi$, when regarded as a map onto $\mathcal{H}$, is a submersion. From Proposition 4.4 we obtain the following result.

Proposition 4.12 *Let $x \in \mathcal{M}$. Then*

$$\psi \colon H_x\mathcal{M} \to \mathcal{H}, \ \xi_x^h \mapsto \phi(x + \xi_x^h)$$

is a local diffeomorphism in a neighborhood of $\mathbf{X} = \phi(x)$ in $\mathcal{H}$.

4.5 The tangent space and its unique representation by gauging.

The following is immediate from Proposition 4.12, Lemma 4.8 and the definition of the HT format. We state it here explicitly to emphasize how tangent vectors can be constructed.

Corollary 4.13 *Let $x = (U_t, \mathbf{B}_t) \in \mathcal{M}$ and $\mathbf{X} = \phi(x) \in \mathcal{H}$, then the map $\mathrm{D}\phi(x)|_{H_x\mathcal{M}}$ is an isomorphism onto $T_{\mathbf{X}}\mathcal{H}$ for any horizontal space $H_x\mathcal{M}$. In particular, for $H_x\mathcal{M}$ from (4.8), every $\delta\mathbf{X} \in T_{\mathbf{X}}\mathcal{H} \subseteq \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ admits a unique minimal representation*

$$\mathrm{vec}(\delta\mathbf{X}) = (\delta U_{(t_r)_1} \otimes U_{(t_r)_2} + U_{(t_r)_1} \otimes \delta U_{(t_r)_2})(B_{t_r}^{(1)})^{\mathsf{T}} + (U_{(t_r)_1} \otimes U_{(t_r)_2})(\delta B_{t_r}^{(1)})^{\mathsf{T}}$$

where the matrices $U_t \in \mathbb{R}^{n_t \times k_t}$ and $\delta U_t \in \mathbb{R}_*^{n_t \times k_t}$ satisfy for $t \notin L \cup \{t_r\}$ the recursions

$$U_t = (U_{t_1} \otimes U_{t_2})(B_t^{(1)})^{\mathsf{T}},$$
$$\delta U_t = (\delta U_{t_1} \otimes U_{t_2} + U_{t_1} \otimes \delta U_{t_2})(B_t^{(1)})^{\mathsf{T}} + (U_{t_1} \otimes U_{t_2})(\delta B_t^{(1)})^{\mathsf{T}}, \quad U_t^{\mathsf{T}} \delta U_t = 0,$$

such that $\xi_x^h = (U_t^h, \mathbf{B}_t^h) = (\delta U_t, \delta \mathbf{B}_t)$ is a horizontal vector.

This corollary shows that while $\delta\mathbf{X} \in T_{\mathbf{X}}\mathcal{H}$ is a tensor in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ of possibly very large dimension, it it structured in a specific way. In particular, when a tensor $\mathbf{X}$ has low T-rank, $\delta\mathbf{X}$ can be represented parsimoniously with very few degrees of freedom, namely by a horizontal vector in $H_x\mathcal{M}$. In other words, if $\mathbf{X}$ can be stored efficiently in $(T, \mathbf{k})$-format, so can its tangent vectors.

Additionally, despite the non-uniqueness when parameterizing $\mathbf{X}$ by x, the representation by horizontal lifts obtained via $(\mathrm{D}\phi(x)|_{H_x\mathcal{M}})^{-1}$ is unique. In contrary to the abstract tangent space of $\mathcal{M}/\mathcal{G}$, a great benefit is that these horizontal lifts are standard matrix-valued quantities, so we can perform standard (Euclidean) arithmetics with them. We will show some applications of this in Section 6.

Our choice of the horizontal space as (4.8) is arguably arbitrary—yet it turned out very useful when proving Theorem 4.11. This freedom is known as *gauging* of a principal fibre bundle.

More precisely, one introduces a *gauge map* γ_x defined for every $x \in \mathcal{M}$ such that $\gamma_x : T_x\mathcal{M} \to T_I\mathcal{G}$ is linear, and the *extended tangent map*,

$$\Gamma_x : T_x\mathcal{M} \to T_{\phi(x)}\mathcal{H} \times T_I\mathcal{G}, \ \xi_x \mapsto (\mathrm{D}\phi(x)[\xi_x], \gamma_x(\xi_x))$$

is an isomorphism. So, after imposing a certain *gauge condition*, like $\gamma_x(\xi_x) = 0$, the tangent vector $\xi_x \in T_x\mathcal{M}$ can be uniquely determined from x and $\delta\mathbf{X} \in T_{\mathbf{X}}\mathcal{H}$, where $\mathbf{X} = \phi(x)$.

Compared to the horizontal lifts of Section 4.3, gauging is another possibility to construct a unique representation of tangent vectors. We now apply it to our setting. Recall from (4.1) that

$$T_I\mathcal{G} = \bigtimes_{t \in T \setminus \{t_r\}} \mathbb{R}^{k_t \times k_t},$$

since there is no Lie group action associated to the root t_r. The gauge map can be taken as

$$\gamma_x : T_x\mathcal{M} \to T_I\mathcal{G},$$
$$(\delta U_t, \delta \mathbf{B}_t) \mapsto \Big((U_t^{\mathsf{T}} \delta U_t)_{t \in L}, (\delta B_t^{(1)} (U_{t_1}^{\mathsf{T}} U_{t_1} \otimes U_{t_2}^{\mathsf{T}} U_{t_2}) (B_t^{(1)})^{\mathsf{T}})_{t \in T \setminus \{t_r\}} \Big),$$

which is obviously linear. Furthermore, by the construction of $V_x\mathcal{M}$ and $H_x\mathcal{M}$, we have

$$\gamma_x(\xi_x^h) = 0 \text{ for all } \xi_x^h \in H_x\mathcal{M} \quad \text{and} \quad \gamma_x(\xi_x^v) \neq 0 \text{ for all } \xi_x^v \in V_x\mathcal{M} \setminus \{0\},$$

which block-diagonalizes Γ_x in the following sense:

$$\Gamma_x(H_x\mathcal{M}) = (T_{\phi(x)}\mathcal{H}, 0) \quad \text{and} \quad \Gamma_x(V_x\mathcal{M}) = (0, \gamma_x(V_x\mathcal{M})).$$

Using our previous derivations, one shows that Γ_x is injective and hence, since it is a mapping between spaces of equal dimension, bijective. Consequently, γ_x is a valid gauging and tangent vectors can be uniquely represented by the gauge condition $\gamma_x(H_x\mathcal{M}) = 0$. See also [51,28] for gauging in the context of the Tucker and TT formats, respectively.

4.6 The closure of $\mathcal{H}_{T,\mathbf{k}}$

Since $\mathcal{M}_{T,\mathbf{k}} = f^{-1}(\mathcal{H}_{T,\mathbf{k}})$ is not closed and f is continuous, the manifold $\mathcal{H}_{T,\mathbf{k}}$ cannot be closed in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. This can be a problem when approximating tensors in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ by elements of $\mathcal{H}_{T,\mathbf{k}}$. As one might expect, the closure of $\mathcal{H}_{T,\mathbf{k}}$ consists of all tensors with T-rank bounded by $\mathbf{k}$, that is, of all $(T, \mathbf{k})$-decomposable tensors. This is covered by a very general result in [18] on the closedness of minimal subspace representations in Banach tensor spaces. We shall give a simple proof for the finite dimensional case.

Theorem 4.14 *The closure of $\mathcal{H}_{T,\mathbf{k}}$ in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is given by*

$$\begin{aligned}\overline{\mathcal{H}_{T,\mathbf{k}}} &= f(M_{T,\mathbf{k}}) \\ &= \bigcup_{\mathbf{r} \le \mathbf{k}} \mathcal{H}_{T,\mathbf{r}} = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \colon \operatorname{rank}_T(\mathbf{X}) \le \mathbf{k}\}.\end{aligned}$$

Proof Since $\mathcal{M}_{T,\mathbf{k}}$ is dense in $M_{T,\mathbf{k}}$, the continuity of f implies that $f(M_{T,\mathbf{k}})$ is contained in the closure of $f(\mathcal{M}_{T,\mathbf{k}}) = \phi(\mathcal{M}_{T,\mathbf{k}}) = \mathcal{H}_{T,\mathbf{k}}$. It thus suffices to show that $f(M_{T,\mathbf{k}})$ is closed in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. Now, since this set consists of tensors for which each mode-t unfolding is at most of rank k_t, and since each such unfolding is an isomorphism, the first part of the claim follows immediately from the lower semicontinuity of the matrix rank function (level sets are closed). The second part is immediate by definition of $\mathcal{M}_{T,\mathbf{k}}$ and enumerating all possible ranks. □

A consequence of the preceding theorem is that every tensor of $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ possesses a best approximation in $\overline{\mathcal{H}_{T,\mathbf{k}}}$, that is, it has a best $(T,\mathbf{k})$-decomposable approximant.

5 Tensors of fixed TT-rank

In this short section we show how the TT format of [61, 60] is obtained as a special case of the HT format. Slightly extending the results in [28], an analysis similar to that of the previous sections gives that the manifold of tensors of fixed TT-rank (see below) is a *globally* embedded submanifold.

Let $\mathbf{r} = (r_1, r_2, \ldots, r_{d-1}) \in \mathbb{N}^{d-1}$ be given and $r_0 = r_d = 1$. The TT$_{\mathbf{r}}$-decomposition of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is an HT decomposition with $(T,\mathbf{k})$ having the following properties:

(i) The tree T is degenerate (linear) in the sense that, at each level, the first of the remaining spatial indices is split from the others to the left son and the rest to the right son:

$$T = \{\{1,\ldots,d\},\{1\},\{2,\ldots d\},\{2\},\ldots,\{d-1,d\},\{d-1\},\{d\}\}. \tag{5.1}$$

(ii) The rank vector $\mathbf{k}$ is given by

$$k_{\{\mu\}} = n_\mu \quad \text{for } \mu = 1,2,\ldots,d-1, \qquad k_{\{\mu,\ldots,d\}} = r_{\mu-1} \quad \text{for } \mu = 1,2,\ldots,d.$$

(iii) The matrices in the first $d-1$ leafs are the identity:

$$U_{\{\mu\}} = I_{n_\mu} \quad \text{for } \mu = 1,2,\ldots,d-1.$$

To simplify the notation, we abbreviate the inner transfer tensors $\mathbf{B}_{\{\mu,\ldots,d\}} \in \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu}$ by $\mathbf{B}_\mu$. For notational convenience, we regard the last leaf matrix $U_{\{d\}}$ as the result of reshaping an additional transfer tensor $\mathbf{B}_d \in \mathbb{R}^{r_{d-1} \times n_d \times 1}$ such that

$$U_{\{d\}}^{\mathsf{T}} = \mathbf{B}_d^{(1)}.$$

An illustration is given in Figure 5.1.

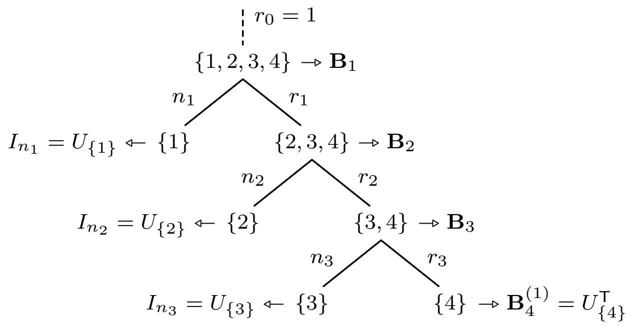

Fig. 5.1 A $\mathrm{TT}_{\mathbf{r}}$-decomposition as a constrained HT-tree.

By the nestedness relation (3.2) of a hierarchical $(T, \mathbf{k})$-decomposition, the subspace belonging to an inner node $t = \{\mu, \dots, d\}$ is spanned by the columns of

$$U_{\{\mu,\dots,d\}} = (I_{n_\mu} \otimes U_{\{\mu+1,\dots,d\}})(B_\mu^{(1)})^{\mathsf{T}}. \tag{5.2}$$

Denote by $B_\mu[\nu] \in \mathbb{R}^{r_{\mu-1} \times r_\mu}$ the ν-th lateral slice of $\mathbf{B}_\mu$ for any $\nu = 1, 2, \dots, n_\mu$. We then have that $B_\mu^{(1)}$ satisfies the partitioning

$$B_\mu^{(1)} = \big(B_\mu[1] \; B_\mu[2] \; \cdots \; B_\mu[n_\mu]\big) \quad \text{for } \mu = 1, 2, \dots, d,$$

so that (5.2) an be written as

$$U_{\{d\}} = \begin{bmatrix} B_d[1]^{\mathsf{T}} \\ B_d[2]^{\mathsf{T}} \\ \vdots \\ B_d[n_d]^{\mathsf{T}} \end{bmatrix} \quad \text{and} \quad U_{\{\mu,\dots,d\}} = \begin{bmatrix} U_{\{\mu+1,\dots,d\}} B_\mu[1]^{\mathsf{T}} \\ U_{\{\mu+1,\dots,d\}} B_\mu[2]^{\mathsf{T}} \\ \vdots \\ U_{\{\mu+1,\dots,d\}} B_\mu[n_\mu]^{\mathsf{T}} \end{bmatrix} \tag{5.3}$$

for $\mu = 1, 2, \dots, d-1$. Recursively applying (5.3) reveals that the $(i_\mu, \dots, i_d)$-th row of $U_{\{\mu,\dots,d\}}$ is given by

$$(B_d[i_d])^{\mathsf{T}} (B_{d-1}[i_{d-1}])^{\mathsf{T}} \cdots (B_\mu[i_\mu])^{\mathsf{T}}.$$

In particular, for $U_{\{1,2,\dots,d\}} = \mathrm{vec}(\mathbf{X})$ we obtain, after taking a transpose, that

$$\mathbf{X}_{i_1,\dots,i_d} = B_1[i_1]\, B_2[i_2] \cdots B_d[i_d].$$

This is the classical matrix product representation of the TT format.

We emphasize again that the TT format is not only specified by the linear tree (5.1), but also by the requirement that the first $d-1$ leafs contain identity matrices. From a practical point of view, one does not need to store these leaf matrices since they are known and always the same. Thus, all $\mathrm{TT}_{\mathbf{r}}$–decomposable tensors can be parametrized by a mapping f acting on tuples (recall $r_0 = r_d = 1$)

$$x = (\mathbf{B}_\mu) = (\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_d) \in M_{\mathbf{r}} = \bigtimes_{\mu=1}^{d} \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu}.$$

On the other hand, while only tensor $\mathbf{B}_d$ has to be stored for the leafs, the inner transfer tensors might be larger compared to those of an HT format with the same linear tree, but minimal ranks in all leafs. In the TT format the rank parameter $\mathbf{r}$ can be chosen only for the inner nodes and the last leaf. Therefore, the TT-rank of a tensor $\mathbf{X}$ is defined as that $\mathbf{r}$ which satisfies

$$r_\mu = \operatorname{rank}(X^{\{1,\dots,\mu\}}) = \operatorname{rank}(X^{\{\mu+1,\dots,d\}}).$$

Letting $R = \max\{r_1, r_2, \cdots, r_{d-1}\}$ and $N = \max\{n_1, n_2, \dots, n_d\}$, we see that

$$\dim(M_{\mathbf{r}}) \le (d-2)NR^2 + 2NR, \tag{5.4}$$

which should be compared to (3.12). Depending on the application (and primarily the sizes of K and R), one might prefer storing a tensor in HT or in TT format. Bounds on the TT-rank in terms of the hierarchical rank for a canonical binary dimension tree, and vice versa, can be found in [23].

Similar to Proposition 3.6 it holds that a tensor can be represented as a $\mathrm{TT}_{\mathbf{r}}$-decomposition if and only if its TT-rank is bounded by $\mathbf{r}$. Denoting

$$\mathcal{T}_{\mathbf{r}} = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \colon \text{TT-rank}(\mathbf{X}) = \mathbf{r}\},$$

the analog of Theorem 3.7 reads as follows.

Theorem 5.1 [28, Theorem 1(a)] *A tensor $\mathbf{X}$ is in $\mathcal{T}_{\mathbf{r}}$ if and only if for every $x = (\mathbf{B}_\mu) \in M_{\mathbf{r}}$ with $f(x) = \mathbf{X}$ the tensors $\mathbf{B}_\mu$ satisfy*

$$\operatorname{rank}(B_\mu^{(1)}) = \operatorname{rank}(B_{\mu-1}^{(3)}) = r_{\mu-1}$$

for $\mu = 2, \dots, d$.

Based on this theorem, one can describe the set $\mathcal{T}_{\mathbf{r}}$ as a quotient manifold along similar lines as for the HT format. The parameter space is now given by

$$\mathcal{M}_{\mathbf{r}} = \{(\mathbf{B}_\mu) \in M_{\mathbf{r}} \colon \operatorname{rank}(B_\mu^{(1)}) = \operatorname{rank}(B_{\mu-1}^{(3)}) = r_{\mu-1} \text{ for } \mu = 2, \dots, d\}.$$

Let again ϕ denote the restriction of f to $\mathcal{M}_{\mathbf{r}}$. The non-uniqueness of the $\mathrm{TT}_{\mathbf{r}}$-decomposition is described in the following proposition which we state without proof. See also [28, Theorem 1(b)] for the orthogonalized case.

Proposition 5.2 *Let $x = (\mathbf{B}_\mu) \in \mathcal{M}_{\mathbf{r}}$ and $y = (\mathbf{C}_\mu) \in \mathcal{M}_{\mathbf{r}}$. Then $\phi(x) = \phi(y)$ if and only if there exist invertible matrices $A_1, A_2, \dots, A_{d-1}$ of appropriate size such that*

$$C_1[i_1] = B_1[i_1]A_2^{-\mathsf{T}}, \quad C_d[i_d] = A_d^{\mathsf{T}} B_d[i_d], \quad C_\mu[i_\mu] = A_\mu^{\mathsf{T}} B_\mu[i_\mu] A_{\mu+1}^{-\mathsf{T}} \tag{5.5}$$

holds for all multi-indices $(i_1, i_2, \dots, i_d)$ and $\mu = 2, \dots, d-1$.

In terms of the Lie group action (4.2), relation (5.5) can be written as

$$y = \theta_x(\mathsf{A}), \quad \mathsf{A} \in \mathcal{G}_{\mathbf{r}}$$

(slightly abusing notation by extending θ to $\mathcal{M}_{\mathbf{r}}$), where $\mathcal{G}_{\mathbf{r}}$ is the subgroup of $\mathcal{G}$ that leaves the first $d-1$ identity leafs unchanged, or formally,

$$\mathcal{G}_{\mathbf{r}} = \{\mathsf{A} \in \mathcal{G} \colon A_{\{\mu\}} = I_{n_\mu} \text{ for } \mu = 1, \dots, d-1\}.$$

One now could start the same machinery as in Section 4. After showing that $\mathcal{G}_{\mathbf{r}}$ acts properly on $\mathcal{M}_{\mathbf{r}}$, one would obtain that $\mathcal{M}_{\mathbf{r}}/\mathcal{G}_{\mathbf{r}}$ is a smooth orbit manifold. Since the approach should be clear, we skip further details of the proof that $\widehat{\phi} \colon \mathcal{M}_{\mathbf{r}}/\mathcal{G}_{\mathbf{r}} \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is an embedding and only formulate the main result.

Theorem 5.3 *The set $\mathcal{T}_{\mathbf{r}}$ of tensors of TT-rank $\mathbf{r}$ is a globally embedded submanifold of $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ which is diffeomorphic to $\mathcal{M}_{\mathbf{r}}/\mathcal{G}_{\mathbf{r}}$. Its dimension is*

$$\dim(\mathcal{T}_{\mathbf{r}}) = \dim(\mathcal{M}_{\mathbf{r}}) - \dim(\mathcal{G}_{\mathbf{r}}) = \sum_{\mu=1}^{d} r_{\mu-1} n_\mu r_\mu - \sum_{\mu=1}^{d-1} r_\mu^2.$$

This extends Theorem 3 in [28], where it only has been shown that $\mathcal{T}_{\mathbf{r}}$ is locally embedded. We also refer the reader to this source for a characterization of the tangent space of $\mathcal{T}_{\mathbf{r}}$ via a horizontal space obtained by orthogonality conditions similar to (4.8).

In analogy to Theorem 4.14, it holds that the closure of $\mathcal{T}_{\mathbf{r}}$ is the set of tensors whose TT-rank is bounded by $\mathbf{r}$, which actually are all $\mathrm{TT}_{\mathbf{r}}$-decomposable tensors.

6 Applications

After our theoretical investigations of the HT format in the previous sections, we should not forget that it has been initially proposed as a promising tool for dealing with problems of high dimensionality. As outlined in the introduction, the HT format is accompanied by a list of concrete problems with accompanying algorithms to solve them. Besides the aesthetic satisfaction, our theory of the format's geometry also has practical value in understanding and improving these algorithms. We hope to support this case with the following two examples: convergence theories for local optimization methods and a dynamical updating algorithm for time-varying tensors.

6.1 Alternating optimization in the hierarchical Tucker format

Consider a C^2-function $J \colon \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \supseteq D \to \mathbb{R}$ on an open domain D. The task is to find a minimizer of J. Two important examples are the approximate solution of linear equations with

$$J(\mathbf{X}) = \|\mathcal{A}(\mathbf{X}) - \mathbf{Y}\|_{\mathrm{F}}^2 = \min, \tag{6.1}$$

and the approximate calculation of the smallest eigenvalue by

$$J(\mathbf{X}) = \frac{\langle \mathbf{X}, \mathcal{A}(\mathbf{X}) \rangle_{\mathrm{F}}}{\langle \mathbf{X}, \mathbf{X} \rangle_{\mathrm{F}}} = \min, \tag{6.2}$$

where in both cases $\mathcal{A}$ is a (tensor-structured) self-adjoint, positive linear operator and $\langle \cdot, \cdot \rangle_{\mathrm{F}}$ is the Frobenius (Euclidian) inner product.

Let T be a dimension tree and $\mathbf{k}$ be a T-rank, such that the manifold $\mathcal{H} = \mathcal{H}_{T,\mathbf{k}}$ is not empty and contained in D. It is assumed that, on the one hand, $\mathcal{H}$ is a good model to minimize J while, on the other hand, $\mathbf{k}$ is still small enough to make the HT format efficient. We can then circumvent the high-dimensionality of the domain D by restricting to the problem

$$J(\mathbf{X}) = \min, \quad \mathbf{X} \in \mathcal{H},$$

or to the handy, but redundant formulation

$$j(x) = J(\phi(x)) = \min, \quad x \in \mathcal{M} = \mathcal{M}_{T,\mathbf{k}}.^{2}$$

More generally, we want to solve for $\mathbf{X} \in \mathcal{H}$ such that

$$\mathrm{D}J(\mathbf{X})[\delta\mathbf{X}] = 0 \text{ for all } \delta\mathbf{X} \in T_{\mathbf{X}}\mathcal{H}. \tag{6.3}$$

Alternatively, we consider

$$\mathrm{D}j(x) = 0, \quad x \in \mathcal{M}. \tag{6.4}$$

Since $\mathrm{D}\phi(x)$ is a surjection onto $T_{\mathbf{X}}\mathcal{M}$ by Corollary 4.13, problems (6.3) and (6.4) are equivalent in the sense that $\mathbf{X}^* = \phi(x^*)$ solves (6.3) if and only if x^* is a solution of (6.4). In particular, in that case every $\hat{x}^* \in \mathcal{G}x^*$ solves (6.4). We hence call $\mathcal{G}x^*$ a *solution orbit* of (6.4). In the following, x^* always denotes a solution of (6.4) and $\mathbf{X}^* = \phi(x^*)$.

The idea of nonlinear relaxation [63] or the Gauss–Seidel method [57] is to solve (6.4) only with respect to one node in the tree at a time while keeping the others fixed. For the TT format this idea has been realized in [27], for the HT format in [47].

To describe the method further, let $t^1, t^2, \dots, t^{|T|}$ be an enumeration of the nodes of T. For notational simplicity, we now partition $x \in \mathcal{M}$ into block variables, $x = (x_1, x_2, \dots, x_{|T|})$, where

$$x_i = \begin{cases} U_{t^i} \in V_i = \mathbb{R}^{n_{t^i} \times k_{t^i}}, & \text{if } t^i \text{ is a leaf,} \\ \mathbf{B}_{t^i} \in V_i = \mathbb{R}^{k_{t^i} \times k_{t^i_1} \times k_{t^i_2}}, & \text{if } t^i \text{ is an inner node.} \end{cases}$$

For $x \in M = M_{T,\mathbf{k}}$ and $i = 1, 2, \dots, |T|$ we define embeddings

$$p_{x,i} \colon V_i \to M, \ \eta \mapsto (x_1, \dots, x_{i-1}, \eta, x_{i+1}, \dots, x_{|T|})$$

and denote by

$$E_{x,i} = p_{x,i}(V_i)$$

their ranges. The elements in $E_{0,i}$ play a particular role and will be called *block coordinate vectors*. Obviously, $E_{x,i} = x + E_{0,i}$.

Let $\mathrm{D}_i\, j(x) = \mathrm{D}j(x) \circ p_{0,i} \colon V_i \to \mathbb{R}$ denote the partial derivative of j at x with regard to the block variable x_i. For x the current iterate, we define the result of one micro-step of Gauss–Seidel for a node t^i as

$$s_i(x) \in E_{x,i} \cap \mathcal{M},$$

with the property that

$$\mathrm{D}_i\, j(s_i(x)) = \mathrm{D}j(s_i(x)) \circ p_{0,i} = 0. \tag{6.5}$$

[2] In practice, one would prefer minimizing j over the closure $M = M_{T,\mathbf{k}}$, that is, over all $(T, \mathbf{k})$-decomposable tensors. For the theory, we would have to assume then that the solution is in $\mathcal{M}$, anyway, since otherwise there is little we can say about it. To give conditions under which this is automatically true, that is, all disposable ranks exploited, seems far from trivial.

In other words, $\eta = p_{x,i}^{-1}(s_i(x))$ is a critical point of $j \circ p_{x,i}$; the update equals $\xi_{0,i} = s_i(x) - x = p_{0,i}(\eta - x_i)$. The nonlinear Gauss-Seidel iteration now informally reads

$$x^{(n+1)} = s(x^{(n)}) = (s_{|T|} \circ s_{|T|-1} \circ \cdots \circ s_1)(x^{(n)}). \tag{6.6}$$

Our aim is to give conditions under which this sequence can be uniquely defined in a neighborhood of a solution of (6.3) and (6.4), respectively, and is convergent. The convergence analysis does not differ much from recent results for the CP and TT format [68,62], so our exposition will be brief yet dense. The results of Section 6.1.3 on the Rayleigh-quotient minimization, however, are new. For numerical results we refer to [27], where the nonlinear Gauss-Seidel algorithm has been studied for the TT format, and to [47], where the Rayleigh quotient minimization has been applied in the HT format.

6.1.1 General existence and convergence criteria

In view of the non-uniqueness of the HT representation, it is reasonable to formulate the convergence results in terms of the sequence $\mathbf{X}_n = \phi(x^{(n)})$. This will be possible in certain neighborhoods of x^* thanks to the next lemma. For fixed $\mathsf{A} \in \mathcal{G}$, let θ_{A} denote the linear map $x \mapsto \theta_x(\mathsf{A})$. This is an isomorphism on M (extending the domain of θ), the inverse being $\theta_{\mathsf{A}^{-1}}$.

Lemma 6.1 *Let $x^* \in \mathcal{M}$ be a solution of* (6.4). *Partition the Hessian (matrix) $\mathrm{D}^2 j(x^*)$ according to the block variables x_i into $\mathrm{D}^2 j(x^*) = L + \Delta + U$, with L, Δ and $U = L^{\mathsf{T}}$ being the lower block triangular, block diagonal and upper block triangular part, respectively.*

(i) *Assume that x^* possesses neighborhoods $\mathcal{U}_i^* \subseteq \mathcal{V}_i^* \subseteq \mathcal{M}$ such that, for $i = 1, 2, \ldots, |T|$, continuously differentiable operators $s_i \colon \mathcal{U}_i^* \to \mathcal{V}_i^*$ satisfying* (6.5) *can be* uniquely *defined. Then there exist possibly smaller neighborhoods $\mathcal{U}_i \subseteq \mathcal{V}_i$ with this property and an open neighborhood $\mathcal{G}_0 \subseteq \mathcal{G}$ of the identity with the following property: For every $\mathsf{A} \in \mathcal{G}_0$ and $x \in \mathcal{U}_i$ it holds for $y = \theta_{\mathsf{A}}(x)$ that $\theta_{\mathsf{A}^{-1}}(E_{y,i} \cap \theta(\mathcal{V}_i, \mathcal{G}_0)) \subseteq E_{x,i} \cap \mathcal{V}_i^*$.*
For all $\mathcal{U}_i, \mathcal{V}_i$ and $\mathcal{G}_0$ with this property the s_i can be uniquely extended to maps $\theta(\mathcal{U}_i, \mathcal{G}_0) \to \theta(\mathcal{V}_i, \mathcal{G}_0)$ via the relation

$$s_i(\theta_{\mathsf{A}}(x)) = \theta_{\mathsf{A}}(s_i(x)), \quad \mathsf{A} \in \mathcal{G}_0, \tag{6.7}$$

which then holds for all $x \in \theta(\mathcal{U}_i, \mathcal{G}_0)$ and appropriate A. *For some possibly smaller neighborhood $\mathcal{U}^* \subseteq \mathcal{M}$ of x^* the operator s given by* (6.6) *can be defined on $\mathcal{U} = \theta(\mathcal{U}^*, \mathcal{G}_0)$ and satisfies the same relation.*

(ii) *Notably, if all $\mathcal{V}_i^*$ can be chosen as $\mathcal{M}$, then one can choose $\mathcal{G}_0 = \mathcal{G}$ and $\mathcal{U} = \theta(\mathcal{U}^*, \mathcal{G})$, which is a neighborhood of the whole solution orbit $\mathcal{G}x^*$.*

(iii) *Furthermore we have*

$$(L + \Delta)\mathrm{D}s(x^*) = -U = (L + \Delta) - \mathrm{D}^2 j(x^*). \tag{6.8}$$

(iv) *Assume that Δ is invertible, then the condition of* (i) *(existence of unique s_i) is fulfilled and we have*

$$\mathrm{D}s(x^*) = -(L + \Delta)^{-1} U = I - (L + \Delta)^{-1} \mathrm{D}^2 j(x^*). \tag{6.9}$$

Remark 6.2 We did not succeed to dispose with the complicated condition in (i) compared to the much simpler on in (ii). Before approaching the proof we highlight an important special case in which one is in the favorable situation (ii) [62]. Namely, when J is strictly convex and possesses a critical point in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, then J is coercive and thus has a unique critical point on the range of the *linear* operator

$$P_{x,i}\colon V_i \to \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}, \ \eta \mapsto \phi(p_{x,i}(\eta)). \tag{6.10}$$

It can be easily verified that $P_{x,i}$ is injective for $x \in \mathcal{M}$ (use (3.2), cf. [27]). Hence $\eta \mapsto j(p_{x,i}(\eta)) = J(P_{x,i}(\eta))$ is strictly convex and possesses a unique critical point η^*, so that $s_i(x) = p_{x,i}(\eta^*)$ can be uniquely defined for every $x \in \mathcal{M}$. It is of vital importance that one can guarantee $s_i(x) \in \mathcal{M}$ only for x in a neighborhood $\mathcal{U}$ of the orbit $\mathcal{G}x^* \subseteq \mathcal{M}$, which under the given conditions contains the fixed points of s_i.

Remark 6.3 One can show that the non-singularity of Δ in (iv) is in fact a condition on $\mathbf{X}^* = \phi(x^*)$, that is, it has to hold for all $x \in \mathcal{G}x^*$.

Proof of Lemma 6.1 To show (i), let $y = \theta_{\mathsf{A}}(x)$ as described in the lemma and $z \in E_{y,i} \cap \theta(\mathcal{V}_i, \mathcal{G})$. Then $\theta_{\mathsf{A}^{-1}}(z)$ is in $E_{x,i} \cap \mathcal{V}_i$. It holds

$$\mathrm{D}_i\, j(z) = \mathrm{D}j(z) \circ p_{0,i} = \mathrm{D}j(\theta_{\mathsf{A}^{-1}}(z)) \circ \theta_{\mathsf{A}^{-1}} \circ p_{0,i},$$

where we used $j = j \circ \theta_{\mathsf{A}^{-1}}$ on $\mathcal{M}$ for the second equality. Since $\theta_{\mathsf{A}^{-1}}$ is an isomorphism on $E_{0,i}$, we conclude from the assumption that $\mathrm{D}_i\, j(z) = 0$ if and only if $z = \theta_{\mathsf{A}}(s_i(x))$. This proves the main assertion of (i). The existence of the neighborhoods $\mathcal{U}_i$, $\mathcal{V}_i$ and $\mathcal{G}_0$ and the existence of a definition domain for s (using that x^* is a fixed point of each s_i) follows from continuity arguments. Conclusion (ii) is direct. Formula (6.8) can be found in [10, Eq. (16)]. In the case (iv), the existence of a unique continuously differentiable $s_i\colon \mathcal{U}_i^* \to \mathcal{V}_i^*$ of x^* follows from the implicit function theorem; see also [57, 10.3.5]. □

In the situation of the preceding lemma, the nonlinear Gauss-Seidel method, although formally an algorithm on $\mathcal{M}$, can be regarded as an algorithm on $\mathcal{H}$ too, since equivalent representations are mapped onto equivalent ones. The counterpart of the iteration s in (6.6) is given in the following proposition.

Proposition 6.4 *Under the conditions of Lemma 6.1,* $\mathcal{O} = \phi(\mathcal{U})$ *is an open neighborhood of* $\mathbf{X}^* = \phi(x^*)$ *in* $\mathcal{H}$*. Furthermore, the operator*

$$S\colon \mathcal{O} \to \mathcal{H}, \ \mathbf{X} = \phi(x) \mapsto \phi(s(x))$$

is a well-defined C^1 *map;* $\mathbf{X}^*$ *being one of its fixed points.*

Proof We first have to note that ϕ, when regarded as a map from $\mathcal{M}$ onto $\mathcal{H}$, is a submersion between manifolds and, as such, an open map. Hence, $\mathcal{O}$ is open [16, 16.7.5] and we can pass smoothly to the quotient [50, Proposition 5.20]. □

In the end, one is interested in the sequence

$$\mathbf{X}_{n+1} = S(\mathbf{X}_n), \tag{6.11}$$

but the computations are performed in $\mathcal{M}$ via s_i. Fortunately, Lemma 6.1 ensures that moderately (or in (ii) even arbitrarily) changing representation along an orbit during the iteration process (say, to a norm-balanced or orthogonalized HT) is allowed and will not affect the sequence (6.11). This is important for an efficient implementation of the alternating optimization schemes as done in [27, 47].

We now calculate the derivative of S. From $S \circ \phi = \phi \circ s$ (on $\mathcal{U}$) and $s(x^*) = x^*$ we obtain

$$\mathrm{D}S(\mathbf{X}^*) \circ \mathrm{D}\phi(x^*) = \mathrm{D}\phi(x^*) \circ \mathrm{D}s(x^*). \tag{6.12}$$

The vertical space $V_{x^*}\mathcal{M}$, which is the tangent space to $\mathcal{G}x^*$ at x^* (see Section 4.3), is the null space of $\mathrm{D}\phi(x^*)$ by Corollary 4.13, and hence, by the above relation, has to be an invariant subspace of $\mathrm{D}s(x^*)$.[3] Thus, if $H_{x^*}\mathcal{M}$ is *any* horizontal space and q the projection onto $H_{x^*}\mathcal{M}$ with regard to the splitting $M = V_{x^*}\mathcal{M} \oplus H_{x^*}\mathcal{M}$, we may regard $\mathrm{D}\phi$ as isomorphism between $H_{x^*}\mathcal{M}$ and $T_{\mathbf{X}^*}\mathcal{H}$ (similar to Corollary 4.13 but with another choice for $H_{x^*}\mathcal{M}$), and replace (6.12) by the equivalent equation

$$\mathrm{D}S(\mathbf{X}^*) = \mathrm{D}\phi(x^*) \circ q \circ \mathrm{D}s(x^*) \circ (\mathrm{D}\phi(x^*))^{-1}. \tag{6.13}$$

Let ρ denote the spectral radius. By the contraction principle, sequence (6.11) will be locally linearly convergent to $\mathbf{X}^*$ when $\rho(\mathrm{D}S(\mathbf{X}^*)) < 1$. Since, by (6.13), $\rho(\mathrm{D}S(\mathbf{X}^*)) = \rho(q \circ \mathrm{D}s(x^*))$ and information on $\mathrm{D}s(x^*)$ is available through (6.8), we have the following result for the iteration S.

Proposition 6.5 *Under the conditions of Lemma 6.1 assume that* $\rho(q \circ \mathrm{D}s(x^*)) < 1$ *for some horizontal space* $H_{x^*}\mathcal{M}$*. Then, the sequence* (6.11) *is locally linearly convergent to* $\mathbf{X}^*$*.*

Again, the condition does not depend on the specific choice of $H_{x^*}\mathcal{M}$ as long as it is complementary to $V_{x^*}\mathcal{M}$. Also note that the convergence region might be much smaller than $\mathcal{O}$.

6.1.2 The case of an invertible block diagonal

Suppose we are in the situation of (iv) in Lemma 6.1. Then, by (6.9), $\mathrm{D}s(x^*)$ is the error iteration matrix of the linear block Gauss-Seidel iteration applied to the Hessian $\mathrm{D}^2 j(x^*)$. Since $V_{x^*}\mathcal{M}$ is an invariant subspace of $\mathrm{D}s(x^*)$, we have

$$(q \circ \mathrm{D}s(x^*))^n = q \circ (\mathrm{D}s(x^*))^n. \tag{6.14}$$

If $\mathrm{D}^2 j(x^*)$ is positive semidefinite (which is the case if x^* is a local minimum of (6.4)) and the block diagonal Δ is positive definite, then, for every $\xi \in M$ the sequence $(\mathrm{D}s(x^*))^n[\xi]$ will converge to an element in the null space of $\mathrm{D}^2 j(x^*)$; see [35, Theorem 2].

Theorem 6.6 *Assume the Hessian* $\mathrm{D}^2 j(x^*)$ *is positive semidefinite and its null space equals* $V_{x^*}\mathcal{M}$ *(cf. footnote 3), or, equivalently,* $\operatorname{rank}(\mathrm{D}^2 j(x^*)) = \dim(\mathcal{H})$*. Then condition* (iv) *of Lemma 6.1 that* Δ *should be positive definite is fulfilled and the sequence* (6.11) *is locally linearly convergent to* $\mathbf{X}^*$*.*

---

[3] If Δ is invertible, it even follows from (6.9) that $\mathrm{D}s(x^*)$ is the identity on $V_{x^*}\mathcal{M}$. Namely, since j is constant on $\mathcal{G}x^*$ and $\mathrm{D}j(x^*) = 0$, $V_{x^*}\mathcal{M}$ is in the null space of $\mathrm{D}^2 j(x^*)$.

Proof Combining the convergence of $\mathrm{D}s(x^*)$ with (6.14) implies $(q\circ \mathrm{D}s(x^*))^n \to 0$, which means $\rho(q \circ \mathrm{D}s(x^*)) < 1$. To prove that Lemma 6.1,(ii) indeed holds, one has to verify that block coordinate vectors $\xi_{0,i} = p_{0,i}(\eta)$ are not in $V_{x^*}\mathcal{M}$ unless they are zero. But since $\phi(x^* + \xi_{0,i}) = \mathbf{X}^* + P_{x^*,i}(\eta)$ (see (6.10)), $\mathrm{D}\phi(x^*)[\xi_{0,i}] = 0$ implies $P_{x^*,i}(\eta) = 0$ which is false unless $\eta = 0$ (see also Remark 6.2). □

Written out explicitly, the quadratic form $\mathrm{D}^2 j(x^*)$ satisfies

$$\mathrm{D}^2 j(x^*)[\xi,\xi] = \mathrm{D}^2 J(\mathbf{X}^*)[\mathrm{D}\phi(x^*)[\xi], \mathrm{D}\phi(x^*)[\xi]] + \mathrm{D}J(\mathbf{X}^*)(\mathrm{D}^2\phi(x^*)[\xi,\xi]). \quad (6.15)$$

Hence, to prevent misunderstandings, the condition formulated in Theorem 6.6 is *not* that $\mathrm{D}^2 J(\mathbf{X}^*)$ has to be positive definite on $T_{\mathbf{X}^*}\mathcal{H}$. One also has to take the second term in (6.15) into account, which is related to the curvature of $\mathcal{H}$. However, let $\mathbf{Z}$ be a critical point of J in $\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ and $\mathrm{D}^2 J(\mathbf{Z})$ be positive definite, then $\mathrm{D}^2 j(x^*)$ will be positive definite too if $\mathbf{X}^* = \mathbf{Z}$. In that case, the manifold $\mathcal{H}$ has been perfectly chosen for solving (6.3).

One now could ask whether the same conclusion holds when $\mathcal{H}$ is "good enough" by which we mean that a critical point $\mathbf{X}^*$ on $\mathcal{H}$ is close to a critical point $\mathbf{Z}$ on $\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$, or equivalently, $\mathrm{D}J(\mathbf{X}^*)$ almost zero on $\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$. Deriving estimates using (6.15) without further knowledge of J, however, is not very helpful since one obtains that the distance one has to achieve between $\mathbf{X}^*$ and $\mathbf{Z}$ depends on $\mathbf{X}^*$ itself, cf. [62, Theorem 4.1]. Fortunately, for the important problem (6.1) of approximately solving a linear equation, it is possible to formulate a condition on $\|\mathbf{X}^* - \mathbf{Z}\|_{\mathrm{F}}$ in terms of $\mathbf{Z}$ (and $\mathcal{A}$) only, since for local minima of (6.1) it necessarily holds

$$\|\mathcal{A}(\mathbf{X}^*)\|_{\mathrm{F}}^2 = \|\mathbf{Y}\|_{\mathrm{F}}^2 - \|\mathcal{A}(\mathbf{X}^*) - \mathbf{Y}\|_{\mathrm{F}}^2 = \|\mathcal{A}(\mathbf{Z})\|_{\mathrm{F}}^2 - \|\mathcal{A}(\mathbf{X}^*) - \mathcal{A}(\mathbf{Z})\|_{\mathrm{F}}^2.$$

Quantitative conditions in that spirit (for $\mathcal{A}$ being the identity) have been obtained for the TT format in [62, Theorem 4.2]. For the Tucker format something similar has been previously done in [43].

6.1.3 Rayleigh quotient minimization[4]

Theorem 6.6 assumes that the function j satisfies a certain rank condition involving $V_{x^*}\mathcal{M}$, which cannot be relaxed under the given circumstances. While this condition is likely reasonable for certain functions, it does not hold for the minimization of the Rayleigh quotient as in (6.2). The reason is that J in (6.2) is constant for any $\mathbf{X} \in \mathcal{H}$ on the subspace $\mathrm{span}(\mathbf{X}) \setminus \{0\}$. We therefore adapt the previous convergence proof for such type of functions.

Let $x \in \mathcal{M}$. We define for $i = 1, 2, \dots, |T|$, the subspaces

$$W_{x,i} = \mathrm{span}((0,\dots,0,x_i,0,\dots,0)) \subseteq E_{0,i}$$

and

$$W_x = W_{x,1} \oplus W_{x,2} \oplus \cdots \oplus W_{x,|T|}.$$

Then j is constant on $\mathcal{W}_x = W_{x,1}\setminus\{0\} \oplus W_{x,2}\setminus\{0\} \oplus \cdots \oplus W_{x,|T|}\setminus\{0\}$, since, by multilinearity,

$$\phi(\mathcal{W}_x) = \mathrm{span}(\phi(x)) \setminus \{0\}. \quad (6.16)$$

[4] Some ideas in this section are inspired by similar ones in [74].

Therefore, $\mathrm{D}^2 j(x)$ vanishes on W_x. Fix i and assume t^j is either a son or the father of t^i. Then it is easy to deduce from (4.7) that $W_{x,j} \subseteq V_x\mathcal{M} \oplus W_{x,i}$. By traversing through the tree, we see that this in fact holds for all $j \neq i$. Therefore,

$$V_x\mathcal{M} + W_x = V_x\mathcal{M} \oplus W_{x,i}, \tag{6.17}$$

which shows $\mathrm{rank}(\mathrm{D}^2 j(x)) \leq \dim(\mathcal{H}) - 1$, so Theorem 6.6 indeed is not applicable. We want to establish a convergence result for the case $\mathrm{rank}(\mathrm{D}^2 j(x)) = \dim(\mathcal{H}) - 1$.

At a lower level, we have the technical difficulty of not being able to define the operators s_i in a unique way. More precisely, let $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{|T|})$ be an invertible diagonal matrix, such that $\Lambda x = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_{|T|} x_{|T|})$. Then $j = j \circ \Lambda$ on $\mathcal{M}$, which gives

$$\mathrm{D}_i j(x) = \mathrm{D} j(x) \circ p_{0,i} = \mathrm{D} j(\Lambda x) \circ \Lambda \circ p_{0,i}$$

for all $x \in \mathcal{M}$. Since Λ is an isomorphism on $E_{0,i}$, we conclude from this equation that (for $s_i(x) \in \mathcal{M}$)

$$\mathrm{D}_i j(s_i(x)) = 0 \quad \text{if and only if} \quad \mathrm{D}_i j(\Lambda s_i(x)) = 0. \tag{6.18}$$

In particular, consider $\Lambda = \mathrm{diag}(1, \ldots, 1, \lambda, 1, \ldots, 1)$ with $\lambda \neq 0$ at the i-th position. Then Λ is an isomorphism on $E_{x,i}$, which shows that we have infinitely many possibilities to choose $s_i(x)$ according to (6.5).

On the other hand, every reasonable application of nonlinear Gauss-Seidel to problem (6.2) would incorporate a normalization step of the iterates to remove the above ambiguity (at least locally). Without loss of generalization, we may restrict these normalized iterates to belong to the submanifold

$$\mathcal{N} = \{x \in \mathcal{M} \colon \|x_i\|_{\mathrm{F}} = 1 \text{ for } i = 1, 2, \ldots, |T|\}.$$

Define

$$g \colon \mathcal{M} \to \mathcal{N},\ x \mapsto (x_1/\|x_1\|_{\mathrm{F}}, x_2/\|x_2\|_{\mathrm{F}}, \ldots, x_{|T|}/\|x_{|T|}\|_{\mathrm{F}})$$

and

$$\vartheta \colon \mathcal{N} \times \mathcal{G} \to \mathcal{N},\ (x, \mathsf{A}) \mapsto \vartheta_{\mathsf{A}}(x) = g(\theta_{\mathsf{A}}(x)).$$

Then we can state the following analog to Lemma 6.1.

Lemma 6.7 *Let $x^* \in \mathcal{N}$ be a solution of* (6.4). *Assume x^* possesses neighborhoods $\mathcal{U}_i^* \subseteq \mathcal{V}_i^* \subseteq \mathcal{N}$ such that, for $i = 1, 2, \ldots, |T|$, continuously differentiable operators $\bar{s}_i \colon \mathcal{U}_i^* \to \mathcal{V}_i^*$ satisfying* (6.5) *can be* uniquely *defined. Then there hold similar statements as in Lemma 6.1,*(i) *and* (ii). *In particular, there exists a neighborhood $\mathcal{U} \subseteq \mathcal{N}$ of x^* and a neighborhood $\mathcal{G}_0 \subseteq \mathcal{G}$ of the identity, such that $\bar{s} = \bar{s}_{|T|} \circ \bar{s}_{|T|-1} \circ \cdots \circ \bar{s}_1$ can be defined on $\vartheta(\mathcal{U}, \mathcal{G}_0)$ and satisfies*

$$\bar{s}(\vartheta_{\mathsf{A}}(x)) = \vartheta_{\mathsf{A}}(\bar{s}(x))$$

for appropriate A.

Proof The proof is almost the same as of Lemma 6.1. One first has to replace $\theta_{\mathsf{A}^{-1}}$ by the map $\Lambda^{-1} \circ \theta_{\mathsf{A}^{-1}}$, where Λ is the diagonal matrix which satisfies $y = \vartheta_{\mathsf{A}}(x) = \Lambda\theta_{\mathsf{A}}(x)$. One then will find $\mathrm{D}_i\, j(\Lambda^{-1}\theta_{\mathsf{A}^{-1}}(z)) = 0$, if $\mathrm{D}_i\, j(z) = 0$. By (6.18), this then shows $\mathrm{D}_i\, j(\vartheta_{\mathsf{A}^{-1}}(z)) = 0$, so one can deduce $z = \vartheta_{\mathsf{A}}(\bar{s}_i(x))$. □

Remark 6.8 As described in [47], the restriction of minimization task (6.2) to $E_{x,i}$ (that is, to V_i) is a generalized eigenvalue problem of certain auxiliary matrices and can therefore be solved by algebraic methods. These auxiliary matrices are obtained by contracting over the fixed directions in the scalar product of (6.2). Such a contraction can be efficiently computed if the nodes of the tree are visited in an adequate order (depth-first search) and if the HT representation is kept orthogonalized.

Let us now define

$$\mathcal{S} = \{\mathbf{X} \in \mathcal{H} \colon \|\mathbf{X}\|_{\mathrm{F}} = 1\}.$$

This is a $(\dim(\mathcal{H}) - 1)$-dimensional submanifold of $\mathcal{H}$. To see this, observe that $\operatorname{span}(\mathbf{X}) \subseteq T_{\mathbf{X}}\mathcal{H}$ for every $\mathbf{X} \in \mathcal{H}$. Thus, $G \colon \mathcal{H} \to \mathbb{R},\ \mathbf{X} \mapsto \|\mathbf{X}\|_{\mathrm{F}}^2 - 1$ is of rank one, and $\mathcal{S} = G^{-1}(0)$.

Lemma 6.9 *The map* $\psi \colon \mathcal{N} \to \mathcal{S},\ x \mapsto \phi(x)/\|\phi(x)\|_{\mathrm{F}}$ *is a surjective submersion.*

Proof The surjectivity follows from (6.16). Let $x \in \mathcal{N}$. Since $\mathrm{D}\phi(x)$ vanishes on $V_x\mathcal{M}$, but has rank one on every $W_{x,i}$, (6.17) shows that $\mathrm{D}\phi(x)$ has rank one on W_x. Thus, $\mathrm{D}\phi(x)$ is at least of of rank $\dim(\mathcal{H}) - 1 = \dim(\mathcal{S})$ on $T_x\mathcal{N} = W_x^{\perp}$ (simply because $\operatorname{rank}(\mathrm{D}\phi(x)) = \dim(\mathcal{H})$ on M). In particular, $\mathrm{D}\phi(x)[T_x\mathcal{N}]$ contains a $(\dim(\mathcal{H}) - 1)$-dimensional subspace complementary to $\operatorname{span}(\phi(x)) = \mathrm{D}\phi(x)[W_{x,i}]$, and obviously no larger one with this property. The derivative of $\mathbf{X} \mapsto \mathbf{X}/\|\mathbf{X}\|_{\mathrm{F}}$ at $\mathbf{X} = \phi(x)$ equals the orthogonal projection onto $(\operatorname{span}(\mathbf{X}))^{\perp} = T_{\mathbf{X}}\mathcal{S}$, which proves $\operatorname{rank}(\mathrm{D}\psi(x)) = \dim(\mathcal{S})$. □

Under the conditions of Lemma 6.7,

$$\bar{S} \colon \mathcal{O} \to \mathcal{S},\ \mathbf{X} = \psi(x) \mapsto \psi(s(x)),$$

is a well-defined C^1-map on the neighborhood $\mathcal{O} = \psi(\mathcal{U}) \subseteq \mathcal{S}$ of its fixed-point $\mathbf{X}^* = \psi(x^*)$. We now reduce the local convergence proof of the sequence

$$\mathbf{X}^{(n+1)} = \bar{S}(\mathbf{X}^{(n)}) \tag{6.19}$$

to the case of an invertible block diagonal as in Section 6.1.2 by replacing the role of M with $T_{x^*}\mathcal{N} = W_{x^*}^{\perp}$ and $\mathcal{H}$ with $\mathcal{S}$. To do so, first note that $g_* = g|_{T_{x^*}\mathcal{N}}$ is a local diffeomorphism, so that, by Lemma 6.9, $\psi_* = \psi \circ g_*$ is a local submersion. Therefore, the relation

$$\mathrm{D}\bar{S}(\mathbf{X}^*) \circ \mathrm{D}\psi(x^*) = \mathrm{D}\psi(x^*) \circ \mathrm{D}\bar{s}(x^*),$$

which we obtained from $\bar{S} \circ \psi = \psi \circ \bar{s}$ and $\bar{s}(x^*) = x^*$, is equivalent to

$$\mathrm{D}\bar{S}(\mathbf{X}^*) \circ \mathrm{D}\psi_*(x^*) = \mathrm{D}\psi_*(x^*) \circ \mathrm{D}(g_*^{-1} \circ \bar{s} \circ g_*)(x^*).$$

The key observation is now that, in virtue of (6.18), $g_*^{-1} \circ \bar{s} \circ g_*$ is a nonlinear Gauss-Seidel operator on the *linear* space $T_{x^*}\mathcal{N} = W_{x^*}^{\perp}$. We recall that the analogous equation equation (6.12) was the starting point of the general analysis. Therefore, given that $\mathrm{D}^2 j(x^*)|_{W_{x^*}^{\perp}}$ is semidefinite, the local convergence of (6.19) will follow in the same way as Theorem 6.6, if (i) the null space of $\mathrm{D}^2 j(x^*)|_{W_{x^*}^{\perp}}$ coincides with the null space of $\mathrm{D}\bar{\psi}(x^*)$, and (ii) the block diagonal of $\mathrm{D}^2 j(x^*)|_{W_{x^*}^{\perp}}$ is positive definite. It remains to show that both is the case, if $\operatorname{rank}(\mathrm{D}^2 j(x^*)) = \dim(\mathcal{H}) - 1$.

Theorem 6.10 *For every* $\mathbf{X} \in D$*, let* J *be constant on* $\mathrm{span}(\mathbf{X}) \setminus \{0\}$*. Under the conditions of Lemma 6.7, assume the Hessian* $\mathrm{D}^2 j(x^*)$ *is positive semidefinite and its null space equals* $W_{x^*} + V_{x^*}\mathcal{M}$*, or, in other words,* $\mathrm{rank}(\mathrm{D}^2 j(x^*)) = \dim(\mathcal{H}) - 1$ *(cf.* (6.17)*). Then, the sequence* (6.19) *is locally linearly convergent to* $\mathbf{X}^* = \psi(x^*)$*.*

Proof Restrict j onto $T_{x^*}\mathcal{N} = W_{x^*}^{\perp}$. Since ψ_* is a local submersion, $\psi_*^{-1}(\mathbf{X}^*)$ is a submanifold of dimension $\dim(\mathcal{S}) = \dim(\mathcal{H}) - 1$, and its tangent space at x^* is the null space of $\mathrm{D}\psi_*(x^*)$, see [16, 16.8.8]. Clearly, j is constant on $\psi_*^{-1}(\mathbf{X}^*)$ and thus, since $\mathrm{D}j(x^*) = 0$, $\mathrm{D}^2 j(x^*)$ vanishes on this null space. This shows (i) of the comment preceding the theorem, while (ii) holds by assumption, since, by (6.17), $\mathrm{D}^2 j(x^*)$ is positive definite on the block coordinate spaces $W_{x^*,i}^{\perp}$ in $W_{x^*}^{\perp}$ (which are not in $V_{x^*}\mathcal{M}$). □

Remark 6.11 The condition in the theorem is reasonable for the Rayleigh quotient minimization, if the eigenvalue one is looking for has multiplicity one. Namely, J is constant on the whole eigenspace, so if its dimension is larger than one, the Hessian of j likely will vanish on a space larger than W_{x^*}.

6.2 Dynamical hierarchical Tucker rank approximation

Consider a time- or parameter-varying tensor $\mathbf{Y}(s) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ where $0 \leq s \leq S$. Let $(T, \mathbf{k})$ represent an HT format which is believed to be useful for approximating $\mathbf{Y}(s)$ at each s. In other words, our aim would be solving the best approximation problem

$$\mathbf{X}(s) \in \mathcal{H} = \mathcal{H}_{T,\mathbf{k}} \quad \text{such that} \quad \|\mathbf{Y}(s) - \mathbf{X}(s)\|_{\mathrm{F}} = \min . \tag{6.20}$$

Unfortunately, problem (6.20) is hard to compute (even for fixed values of s) and, contrary to the matrix case, there is no explicit solution known. For most approximation problems, however, one is usually content with a quasi-optimal solution. For instance, the procedure in [22] based on several SVD approximations of unfoldings of $\mathbf{Y}(s)$ delivers a tensor $\hat{\mathbf{X}}(s)$ such that, for every s,

$$\|\mathbf{Y}(s) - \hat{\mathbf{X}}(s)\|_{\mathrm{F}} \leq \sqrt{2d-3}\, \|\mathbf{Y}(s) - \mathbf{X}(s)\|_{\mathrm{F}}. \tag{6.21}$$

On the other hand, computing all these SVDs for every value of s can be very costly: $O(dN^{d+1})$ operations with $N = \max\{n_1, n_2, \ldots, n_d\}$; see [22, Lemma 3.21]. Applied to TT, the bound (6.21) can be tightened using a factor $\sqrt{d-1}$; see, e.g., [60].

An alternative to solving (6.20), is the so-called *dynamical low-rank approximation* proposed in [42, 43] for time-dependent matrices and Tucker tensors, respectively. In this section, we will generalize this procedure to the HT tensors using the geometry of the previous section.

6.2.1 A non-linear initial value problem on $\mathcal{H}_{T,\mathbf{k}}$

Let $\dot{\mathbf{X}}$ denote $\mathrm{d}\mathbf{X}(s)/\mathrm{d}s$. The idea of dynamical low-rank approximation consists of determining an approximation $\mathbf{X}(s) \in \mathcal{H}$ such that, for every s, the derivative of $\mathbf{X}(s)$ is chosen as

$$\dot{\mathbf{X}}(s) \in T_{\mathbf{X}(s)}\mathcal{H} \quad \text{such that} \quad \|\dot{\mathbf{Y}}(s) - \dot{\mathbf{X}}(s)\|_{\mathrm{F}} = \min . \tag{6.22}$$

Together with an initial condition like $\mathbf{X}(0) = \mathbf{Y}(0) \in \mathcal{H}$, (6.22) is a flow problem on the manifold $\mathcal{H}$ since $\dot{\mathbf{X}}(s) \in T_{\mathbf{X}(s)}\mathcal{H}$ for every s. Similarly as in [42,43], we show how this flow can be formulated as a set of non-linear ordinary differential equations suitable for numerical integration.

It is readily seen that, at a particular value of s, the minimization in (6.22) is equivalent to the following Galerkin condition: find $\dot{\mathbf{X}}(s) \in T_{\mathbf{X}(s)}\mathcal{H}$ such that

$$\langle \dot{\mathbf{X}}(s) - \dot{\mathbf{Y}}(s), \delta\mathbf{X} \rangle_{\mathrm{F}} \quad \text{for all } \delta\mathbf{X} \in T_{\mathbf{X}(s)}\mathcal{H}.$$

Define the orthogonal projection (with respect to the Frobenius inner product)

$$P_{\mathbf{X}} \colon \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \to T_{\mathbf{X}}\mathcal{H} \quad \text{for any } \mathbf{X} \in \mathcal{H},$$

then (6.22) becomes

$$\dot{\mathbf{X}}(s) = P_{\mathbf{X}(s)} \dot{\mathbf{Y}}(s). \tag{6.23}$$

This makes that, together with an initial condition $\mathbf{X}(0) \in \mathcal{H}$, the flow (6.23) could be integrated by methods of geometric integration [25]. However, due to the size of the tensors at hand, it is preferable avoiding unstructured tensors in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and instead exploiting that $\mathcal{H}$ is diffeomorphic to the quotient manifold $\mathcal{M}/\mathcal{G}$.

Since $\widehat{\phi} \colon \mathcal{M}/\mathcal{G} \to \mathcal{H}$ in Theorem 4.11 is a diffeomorphism, (6.23) induces a flow on the quotient manifold $\mathcal{M}/\mathcal{G}$ too, namely,

$$\frac{\mathrm{d}\hat{x}(s)}{\mathrm{d}s} = \hat{\xi}_{\hat{x}}(s) = \mathrm{D}\widehat{\phi}^{-1}(\hat{x}(s))[P_{\widehat{\phi}(\hat{x}(s))}\dot{\mathbf{Y}}(s)], \quad \hat{\phi}(\hat{x}(0)) = \mathbf{X}(0),$$

with the property that $\widehat{\phi}(\hat{x}(s)) = \mathbf{X}(s)$ for all s. To integrate this flow on $\mathcal{M}/\mathcal{G}$, we lift the vector field $\hat{\xi}_{\hat{x}}(s)$ to a unique horizontal vector field on $\mathcal{M}$ once an element $x(0)$ in the orbit $\pi^{-1}(\hat{x}(0))$ is chosen. To do this, introduce the shorthand notation F_x for the restriction of $\mathrm{D}\phi$ onto the horizontal spaces $H_x\mathcal{M}$ in (4.8) (which form a bundle). Then we get another flow on $\mathcal{M}$,

$$\frac{\mathrm{d}x(s)}{\mathrm{d}s} = \xi_x^h(s) = F_{x(s)}^{-1}(P_{\phi(x(s))}\dot{\mathbf{Y}}(s)), \quad \phi(x(0)) = \mathbf{X}(0), \tag{6.24}$$

with the property that $\phi(x(s)) = \mathbf{X}(s)$ for all s. Put differently, the isomorphisms F_x give us a unique and well-defined horizontal vector field and smoothness is guaranteed by Theorem 4.7.

From Corollary 4.13, we know how tangent vectors are structured. All that remains now, is deriving how the result of the projection $P_{\mathbf{X}(s)}\dot{\mathbf{Y}}(s)$ can be computed directly as a horizontal vector, that is, the result of $F_{x(s)}^{-1}(P_{\phi(x(s))}\dot{\mathbf{Y}}(s))$. After that, a standard numerical integrator can be used to integrate (6.24) since $\mathcal{M}$ is a matrix manifold that is dense in the Euclidean space M. Contrary to the geometries in [42,43,28] that are based on U_t being orthonormalized, we only require U_t to be of full rank. This is much simpler to ensure during the integration.

Remark that in order to enhance the integrator's efficiency, a special step size control could be used like in [56, (9)]. We did not use such an estimator in our numerical experiments, however; `ode45` of MATLAB performed adequately enough.

So far, we did not mention why (6.22) is a good alternative for (6.20). We refer to [42,43] for theoretical results on specific approximation properties that are based on bounds on the curvature of the manifold and information on the singular values of unfoldings of $\mathbf{Y}(s)$ and $\mathbf{X}(s)$. This is generalized in [13] to an infinite dimensional setting suitable for MCDTH. We conjecture that most of these theoretical results are also extensible to the present geometry but we leave this for future work. Furthermore, in the numerical results later on, it will be clear that (6.22) is indeed sensible.

6.2.2 The orthogonal projection onto $T_{\mathbf{X}}\mathcal{H}$

Given an $(T, \mathbf{k})$-decomposable tensor $\mathbf{X} = \phi(x)$ with $x = (U_t, \mathbf{B}_t)$ and an arbitrary $\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, let $P_{\mathbf{X}}\mathbf{Z} = \delta\mathbf{X} \in T_{\mathbf{X}}\mathcal{H}$ be the desired tangent vector. From Corollary 4.13, we know that

$$\delta U_t = (\delta U_{t_1} \otimes U_{t_2} + U_{t_1} \otimes \delta U_{t_2})(B_t^{(1)})^\mathsf{T} + (U_{t_1} \otimes U_{t_2})(\delta B_t^{(1)})^\mathsf{T},$$

holds for $t = t_r$ and $\delta U_{t_r} = \operatorname{vec}(\delta\mathbf{X})$. We will outline the principle behind computing the corresponding horizontal vector of $\delta\mathbf{X}$, that is, $(\delta U_t, \delta\mathbf{B}_t) \in H_x\mathcal{M}$. Since the tangent space is defined recursively, we immediately formulate it in full generality for an arbitrary node. This has the benefit that, while our derivation starts at the root, it holds true for the children without change.

In order not to overload notation, we drop the explicit notation for the node t and denote $B^{(1)}$ by B. Then, in case of the root, $\delta U = \operatorname{vec}(\delta\mathbf{X})$ has to be of the form

$$\delta U = (\delta U_1 \otimes U_2 + U_1 \otimes \delta U_2)\, B^\mathsf{T} + (U_1 \otimes U_2)\, \delta B^\mathsf{T}, \tag{6.25}$$

such that $\delta U_1 \perp U_1$ and $\delta U_2 \perp U_2$. In addition, δB is required to be a horizontal vector B^h. Recalling once again definition (4.8) for the horizontal space, an equivalent requirement is that

$$\delta B^\mathsf{T} = M\delta C$$

such that

$$\begin{aligned} &M = (U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)^{-1}(B^\mathsf{T})^\perp, && \delta C \in \mathbb{R}^{(k_1k_2-k)\times k} && \text{for } t \notin L \cup t_r, \\ &M = I_n, && \delta C \in \mathbb{R}^{k_1k_2\times 1} && \text{for } t = t_r. \end{aligned} \tag{6.26}$$

In the above, we used $X^\perp \in \mathbb{R}_*^{n\times(n-k)}$ to denote a basis for the orthogonal complement of the matrix $X \in \mathbb{R}_*^{n\times k}$.

Let $Z \in \mathbb{R}^{n\times k}$ be an arbitrary matrix. Then we have the unique decomposition

$$\begin{aligned} Z = (U_1^\perp \otimes U_2)\, \delta C^1 + (U_1 \otimes U_2^\perp)\, \delta C^2 + (U_1 \otimes U_2)\, M\delta C^3 \\ +(U_1 \otimes U_2)\, M^\perp \delta C^4 + (U_1^\perp \otimes U_2^\perp)\, \delta C^5 \end{aligned}$$

where all five[5] subspaces are mutually orthogonal. Observe that the first three terms in this decomposition constitute the desired δU from (6.25). Hence, straightforward orthogonal projection onto these subspaces delivers δU as the sum of

$$(P_1^\perp \otimes P_2)\, Z = (\delta U_1 \otimes U_2)B^\mathsf{T}, \quad (P_1 \otimes P_2^\perp)\, Z = (U_1 \otimes \delta U_2)B^\mathsf{T}, \tag{6.27}$$

and

$$(U_1 \otimes U_2)M[M^\mathsf{T}(U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)M]^{-1} M^\mathsf{T}(U_1 \otimes U_2)^\mathsf{T}\, Z = (U_1 \otimes U_2)\delta B^\mathsf{T}, \tag{6.28}$$

[5] In the root t_r, there are only four subspaces since $M^\perp$ for $M = I_n$ is void.

where we have used the orthogonal projectors

$$P_t Z = U_t (U_t^\mathsf{T} U_t)^{-1} U_t^\mathsf{T} Z, \quad P_t^\perp Z = Z - P_t Z.$$

We are now ready to obtain δB from (6.28). If $t = t_r$, we get immediately that

$$\delta B^\mathsf{T} = (U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)^{-1} (U_1 \otimes U_2)^\mathsf{T} Z = [(U_1^\mathsf{T} U_1)^{-1} U_1^\mathsf{T} \otimes (U_2^\mathsf{T} U_2)^{-1} U_2^\mathsf{T}] Z.$$

For nodes that are not the root, we first denote $V = (B^\mathsf{T})^\perp$ and introduce the oblique projector onto $\operatorname{span}(U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)^{-1} V$ along the span of $V^\perp = B^\mathsf{T}$ (see, e.g., [67, Theorem 2.2]) as

$$P_M = (U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)^{-1} V [V^\mathsf{T} (U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)^{-1} V]^{-1} V^\mathsf{T}.$$

After some straightforward manipulation on (6.28), we obtain

$$\delta B^\mathsf{T} = P_M [(U_1^\mathsf{T} U_1)^{-1} U_1^\mathsf{T} \otimes (U_2^\mathsf{T} U_2)^{-1} U_2^\mathsf{T}] Z.$$

One can avoid $V = (B^\mathsf{T})^\perp$ by using $P_M^\perp$, the oblique projection onto the span of B^T along the span of $(U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2)^{-1} V$ as follows

$$P_M = I - P_M^\perp = I - B^\mathsf{T} [B (U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2) B^\mathsf{T}]^{-1} B (U_1^\mathsf{T} U_1 \otimes U_2^\mathsf{T} U_2).$$

In order to obtain δU_1, we first reshape $Z \in \mathbb{R}^{n_t \times k_t}$ into the third-order tensor

$$\widehat{\mathbf{Z}} \in \mathbb{R}^{k_t \times n_{t_1} \times n_{t_2}} \quad \text{such that} \quad \widehat{Z}^{(1)} = Z^\mathsf{T}.$$

Then using (2.1), we can write (6.27) as a multilinear product (recall $B = B_t^{(1)}$),

$$(I_{k_t}, P_1^\perp, P_2) \circ \widehat{\mathbf{Z}} = (I_{k_t}, \delta U_1, U_2) \circ \mathbf{B}_t, \quad (I_{k_t}, P_1, P_2^\perp) \circ \widehat{\mathbf{Z}} = (I_{k_t}, U_1, \delta U_2) \circ \mathbf{B}_t,$$

and after unfolding in the second mode, we get

$$P_1^\perp \widehat{Z}^{(2)} (I_{k_t} \otimes P_2) = \delta U_1 B_t^{(2)} (I_{k_t} \otimes U_2)^\mathsf{T}.$$

So, isolating for δU_1 we obtain

$$\delta U_1 = P_1^\perp \widehat{Z}^{(2)} (I \otimes U_2 (U_2^\mathsf{T} U_2)^{-1}) (B_t^{(2)})^\dagger$$

since $B_t^{(2)}$ has full rank. Similarly we have

$$\delta U_2 = P_2^\perp \widehat{Z}^{(3)} (I \otimes U_1 (U_1^\mathsf{T} U_1)^{-1}) (B_t^{(3)})^\dagger.$$

In the beginning of this derivation, we started with the root and $Z = \operatorname{vec}(\mathbf{Z})$. It should be clear that the same procedure can now be done for the children by setting $Z = \delta U_{(t_r)_1}$ and $Z = \delta U_{(t_r)_2}$ to determine $\delta B_{(t_r)_1}$ and $\delta B_{(t_r)_2}$, respectively. Continuing this recursion, we finally obtain all δB_t and, in the leafs, the δU_t representing $\delta \mathbf{X}$.

We emphasize that the implementation of the procedure above does not form Kronecker products explicitly. Instead, one can formulate most computations as multilinear products. Our MATLAB implementation of the projection operator is available at `http://sma.epfl.ch/~vanderey/geom_ht`. It uses the `htucker` toolbox [48] to represent and operate with HT tensors and their tangent vectors.

We remark that the current implementation uses $\mathbf{Z}$ as a full tensor. Clearly, this is not suitable for high-dimensional applications. When $\mathbf{Z}$ is available in HT format, it is possible to exploit this structure. It is however beyond the scope of the current paper to give the details and we leave this for future work.

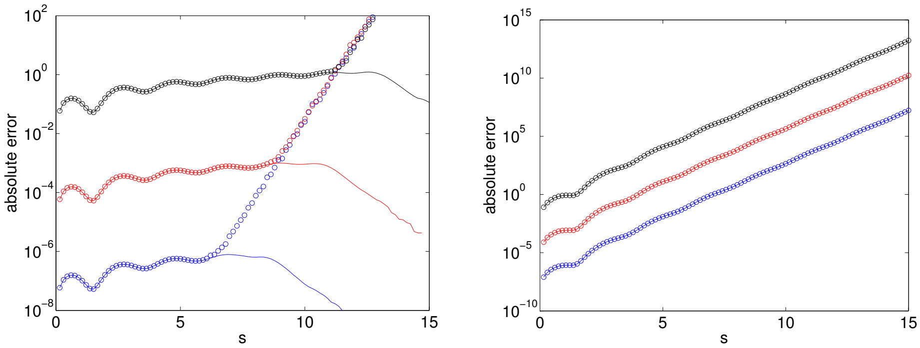

Fig. 6.1 Approximation error for dynamical low-rank (full line) and truncated SVD (circles) with $\epsilon \in \{10^{-1}, 10^{-4}, 10^{-7}\}$ from bottom to top. On the left $c = 0$ and on the right $c = \log_2(d) = 2$.

6.2.3 Numerical example

Let $(T, \mathbf{k})$ be an HT format for d-dimensional tensors in $\mathbb{R}^{N\times N\times\cdots\times N}$ with $k_t = K$ for $t \in T \setminus t_r$. Consider the set $x(s) = (U_t(s), \mathbf{B}_t(s))$ of parametrizations in $\mathcal{M}$ given by

$$U_t(s) = (2 + \sin(s))\, \widehat{U}_t, \quad \mathbf{B}_t(s) = \exp(s)\, \widehat{\mathbf{B}}_t,$$

where the $\widehat{U}_t$ and $\widehat{\mathbf{B}}_t$ are fixed. We will apply a dynamical low-rank approximation to the parameter-dependent tensor

$$\mathbf{Y}(s) = \phi(x(s)) + \omega\epsilon \exp(c\, s)\,(s + \sin(3s))\mathbf{1}, \quad 0 \le s \le 15, \tag{6.29}$$

with ω a random perturbation of unit norm, c and ϵ constants and $\mathbf{1}$ the tensor containing all ones. Observe that the norm of $\phi(x(s))$ grows like $\exp(s \log_2 d)$ when T is a balanced tree. A similar tensor was used in the context of Tucker tensors in [43]. As initial value we take $\mathbf{Y}(0)$.

In order to compare our approach with point-wise approximation by SVDs, we display the absolute approximation error in Figure 6.1 for $N = 45$, $d = 4$, $K = 3$, and different values of $\epsilon \in \{10^{-1}, 10^{-4}, 10^{-7}\}$. The dynamical low-rank approximations are the result of integrating (6.24) with `ode45` in MATLAB using a relative tolerance of 10^{-3}. The SVD approximations are computed in every sample point of s with explicit knowledge of the rank K.

First, observe in Figure 6.1 that the absolute errors for the SVD approach grow in both cases. Since the norm of $\mathbf{Y}(s)$ grows too, this is normal and the relative accuracy is in line with (6.21). In both cases, we also observe that, as expected, the approximation error decreases with smaller ϵ.

For dynamical low-rank, the approximation errors are virtually the same as for the SVD approach for most of the values of s. On the left panel, however, we see that the error for dynamical low-rank stays more or less constant and even goes to zero. This does not happen on the right. Hence, the dynamical low-rank algorithm achieves very high relative accuracy when $c = 0$. An explanation for this difference is that the factor $c = 0$ on the left ensures that $\phi(x(s))$ becomes bigger in norm compared to the noise factor ω. If s is large enough, $\mathbf{Y}(s)$ will have a numerically exact $(T, \mathbf{k})$-decomposition. The dynamical low-rank algorithm seems

			dynamical			SVD	
N	d	K	total time	scaled time	steps	total time	scaled time
25	4	3	9.62	0.050	193	7.22	0.72
35	4	3	19.1	0.10	187	28.4	0.28
45	4	3	39.8	0.21	187	79.2	0.80
25	4	6	10.4	0.056	187	7.13	0.71
35	4	6	23.8	0.13	187	30.9	0.31
45	4	6	49.4	0.26	193	82.8	0.83
25	4	9	11.3	0.062	181	7.73	0.78
35	4	9	24.3	0.12	205	29.3	0.30
45	4	9	50.5	0.25	205	80.2	0.81

Table 6.1 Computational results applied to (6.29) with $c = 0$ and $\epsilon = 10^{-4}$. Times are in seconds.

to detect this and approximates $\mathbf{Y}(s)$ exactly. This happens at some point during the integration when $c < 2$. On the right, however, c is chosen such that both tensors are equally large in norm. Now $\mathbf{Y}(s)$ fails to be an exact $(T, \mathbf{k})$ tensor and dynamical low-rank is as accurate as the SVD approach.

In the next experiment, we investigate the computational time applied to the problem from above with $c = 0$ and $\epsilon = 10^{-4}$. In Table 6.1, we see the results for several sizes N and ranks K with $d = 4$. In addition to the total time, the number of steps that the numerical integrator took is also visible. Observe that this is fairly constant. The scaled time is the total time divided by the number of function evaluations, that is, computing the horizontal vector $\xi_x^h(s)$.

Next, since the dynamical rank approximation is able to form $\mathbf{X}(s)$ for every s, we also computed 100 point-wise SVD-based approximations throughout the interval $0 \leq s \leq 15$ using `htensor.truncate_ltr` of [48]. The indicated scaled time is the mean. Since the solution is smooth, one can probably reduce this number somewhat, for example, by Chebyshev interpolation. Nevertheless, we can see from the table that there is a significant gap between the computational times of both approaches. It is also clear that dynamical approximation becomes faster with increasing N compared to the SVD approach. In addition, the rank K does not seem to have much influence.

Acknowledgements A.U. was supported by the DFG-Priority Program 1324. The authors would like to thank Daniel Kressner, Thorsten Rohwedder, and Christine Tobler for helpful discussions about parts of this paper.

References

1. P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, Acta Appl. Math. **80** (2004), no. 2, 199–220.
2. ———, *Optimization algorithms on matrix manifolds*, Princeton University Press, 2008.
3. E. Acar, D. M. Dunlavy, and T. G. Kolda, *A scalable optimization approach for fitting canonical tensor decompositions*, J. Chemometrics (2010).
4. J. Ballani and L. Grasedyck, *A projection method to solve linear systems in tensor format*, Tech. Report 22, MPI Leipzig, 2010.
5. F. Barahona, *On the computational complexity of Ising spin glass models*, J. Phys. A: Math. Gen. **15** (1982).
6. M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer, *The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets*, Phys. Reports **324** (2000), no. 1, 1–105.

7. R. E. Bellman, *Dynamic programming*, Princeton University Press, 1957.
8. G. Beylkin and M. J. Mohlenkamp, *Numerical operator calculus in higher dimensions*, Proc. Natl. Acad. Sci. USA **99** (2002), no. 16.
9. ———, *Algorithms for numerical analysis in high dimensions*, SIAM J. Scient. Comput. **26** (2005), no. 6, 2133–2159.
10. J. C. Bezdek and R. J. Hathaway, *Convergence of alternating optimization*, Neural Parallel Sci. Comput. **11** (2003), 351–368.
11. H.-J. Bungartz and M. Griebel, *Sparse grids*, Acta Numerica **13** (2004), no. 1–123.
12. A. Cohen, R. DeVore, and Ch. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic pdes*, Technical report 2010-03, Seminar for applied mathematics, 2010.
13. D. Conte and Ch. Lubich, *An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics*, ESAIM: M2AN **44** (2010), 759—780.
14. L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl. **21** (2000), no. 4, 1253–1278.
15. V. de Silva and L.-H. Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Mat. Anal. Appl. **20** (2008), no. 3, 1084—1127.
16. J. Dieudonn'e, *Treatise on analysis*, vol. III, Academic Press, 1973.
17. M. Espig and W. Hackbusch, *A regularized Newton method for the efficient approximation of tensors represented in the canonical tensor format*, 78/2010, MPI Leipzig, 2010.
18. A. Falcó and W. Hackbusch, *Minimal subspaces in tensor representations*, Found. Comput. Math. (2011).
19. C. Gardiner, *Handbook of stochastic methods: for physics, chemistry and the natural sciences*, 3rd ed., Springer Series in Synergetics, Springer, 2004.
20. G. H. Golub and C. F. Van Loan, *Matrix computations*, 3rd ed., Johns Hopkins Studies in Mathematical Sciences, 1996.
21. L. Grasedyck, *Hierarchical low rank approximation of tensors and multivariate functions*, Lecture notes of the Zürich summer school on Sparse Tensor Discretizations of High-Dimensional Problems, 2010.
22. ———, *Hierarchical singular value decomposition of tensors*, SIAM J. Matrix Anal. Appl. **31** (2010), no. 4, 2029–2054.
23. L. Grasedyck and W. Hackbusch, *An introduction to hierarchical ($\mathcal{H}$-) rank and TT-rank of tensors with examples*, Comp. Methods. Appl. Math. **11** (2011), no. 3, 291–304.
24. W. Hackbusch and S. Kühn, *A new scheme for the tensor representation*, J. Fourier Anal. Appl. (2009).
25. E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration*, second ed., Springer-Verlag, 2006.
26. C. J. Hillar and L.-H. Lim, *Most tensor problems are NP hard*, Tech. report, `http://arxiv.org/abs/0911.1393`, 2010.
27. S. Holtz, T. Rohwedder, and R. Schneider, *The alternating linear scheme for tensor optimisation in the TT format*, SIAM J. Numer. Anal. (submitted) (2011).
28. ———, *On manifolds of tensors of fixed TT-rank*, Num. Math. (2012).
29. Roger Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, 1991.
30. R. Hübener, V. Nebendahl, and W. Dür, *Concatenated tensor network states*, New J. Phys. **10** (2010), 025004.
31. T. Huckle, K. Waldherra, and T. Schulte-Herbrüggen, *Computations in quantum tensor networks*, Lin. Alg. Appl. (submitted), 2011.
32. J. Hull, *Options, futures and other derivatives*, 6th ed., Pearson/Prentice Hall, 2009.
33. T. Jahnke and W. Huisinga, *A dynamical low-rank approach to the chemical master equation*, Bull. Math. Biol. **70** (2008), no. 8, 2283—2302.
34. M. Journée, F. Bach, P.-A. Absil, and R. Sepulchre., *Low-rank optimization on the cone of positive semidefinite matrices*, SIAM J. Optim. **20** (2010), no. 5, 2327–2351.
35. H. B. Keller, *On the solution of singular and semidefinite linear systems by iteration*, SIAM J. Numer. Anal. **2** (1965), no. 2, 281–290.
36. B. N. Khoromskij, *Tensor-structured preconditioners and approximate inverse of elliptic operators in* $\mathbb{R}^d$, Constr. Approx. **30** (2009), no. 599–620.
37. ———, *$O(d \log N)$-quantics approximation of N-d tensors in high-dimensional numerical modeling*, Constr. Approx. **34** (2011), no. 257—280.
38. B. N. Khoromskij, V. Khoromskaia, and H.-J. Flad, *Numerical solution of the Hartree–Fock equation in the multilevel tensor-structured format*, SIAM J. Scient. Computing **33** (2011), no. 1, 45–65.

39. B. N. Khoromskij and C. Schwab, *Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs*, Tech. Report 2010-04, Seminar for applied mathematics, ETH Zurich, 2010.
40. B.N. Khoromskij and I. Oseledets, *DMRG + QTT approach to high-dimensional quantum molecular dynamics*, Tech. Report 69/2010, MPI Leipzig, 2010.
41. Shoshichi Kobayashi and Katsumi Nomizu, *Foundations of differential geometry*, Interscience Publishers, a division of John Wiley & Sons, New York-London, 1963, Volumes 1 and 2.
42. O. Koch and Ch. Lubich, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl. **29** (2007), no. 2, 434–454.
43. ———, *Dynamical tensor approximation*, SIAM J. Matrix Anal. Appl. **31** (2010), no. 5, 2360–2375.
44. T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Review **51** (2009), no. 3, 455–500.
45. D. Kressner and C. Tobler, *Krylov subspace methods for linear systems with tensor product structure*, SIAM J. Matrix Anal. Appl. **31** (2010), no. 4, 1688–1714.
46. ———, *Low-rank tensor Krylov subspace methods for parametrized linear systems*, Tech. Report 2010-16, Seminar for applied mathematics, ETH Zurich, 2010.
47. ———, *Preconditioned low-rank methods for high-dimensional elliptic pde eigenvalue problems*, Comp. Methods. Appl. Math. **11** (2011), no. 3, 363—381.
48. ———, `htucker` – *a* MATLAB *toolbox for tensors in hierarchical Tucker format (in preparation)*, Tech. report, Seminar for applied mathematics, ETH Zurich, 2012.
49. J. M. Landsberg, Y. Qi, and K. Ye, *On the geometry of tensor network states*, Tech. report, Department of Mathematics, Texas A&M, 2011.
50. John M. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics, vol. 218, Springer-Verlag, New York, 2003.
51. Ch. Lubich, *On variational approximations in quantum molecular dynamics*, Math. Comp. **74** (2005), no. 250.
52. ———, *From quantum to classical molecular dynamics: reduced models and numerical analysis*, European Math. Soc., 2008.
53. G. Meyer, S. Bonnabel, and R. Sepulchre, *Regression on fixed-rank positive semidefinite matrices: a Riemannian approach*, JMLR **12** (2011), 593–625.
54. P. W. Michor, *Topics in differential geometry*, Amer. Math. Soc., 2008.
55. M. F. Modest, *Radiative heat transfer*, Academic Press, Amsterdam, 2003.
56. A. Nonnenmacher and Ch. Lubich, *Dynamical low-rank approximation: applications and numerical experiments*, Math. Comput. Simulation **79** (2008), no. 4, 1346–1357.
57. J. M. Ortega and W. C. Reinholdt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970.
58. I. V. Oseledets, *Compact matrix form of the d-dimensional tensor decomposition*, Tech. Report 09-01, Institute of Numerical Mathematics, RAS, 2009.
59. ———, *Approximation of* $2^d \times 2^d$ *matrices using tensor decomposition*, SIAM J. Matrix Anal. Appl. **31** (2010), no. 4, 2130–2145.
60. ———, *Tensor-Train decomposition*, SIAM J. on Scient. Comp. **33** (2011), no. 5, 2295–2317.
61. I. V. Oseledets and E. E. Tyrtyshnikov, *Breaking the curse of dimensionality, or how to use SVD in many dimensions*, SIAM J. Scient. Comp. **31** (2009), no. 5, 3744–3759.
62. T. Rohwedder and A. Uschmajew, *Local convergence of alternating schemes for optimization of convex problems in the TT format*, Tech. report, TU Berlin, 2011.
63. S. Schechter, *Iteration methods for nonlinear problems*, Trans. Am. Math. Soc. **104** (1962), no. 179–189.
64. U. Schollwöck, *The density-matrix renormalization group*, Rev. Mod. Phys. **77** (2005), no. 1, 259–315.
65. Ch. Schwab and C. J. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numerica **20** (2011), 291–467.
66. U. Shalit, D. Weinshall, and G. Chechik, *Online learning in the manifold of low-rank matrices*, Neural Information Processing Systems, 2010.
67. G. W. Stewart, *On the numerical analysis of oblique projectors*, SIAM J. Matrix Anal. Appl. **32** (2011), no. 1, 309–348.
68. A. Uschmajew, *Local convergence of the alternating least squares algorithm for canonical tensor approximation.*, SIAM J. Matrix Anal. Appl. (submitted) (2011).

69. B. Vandereycken, *Low-rank matrix completion by riemannian optimization*, Tech. report, ANCHP-MATHICSE, Mathematics Section, École Polytechnique Fédérale de Lausanne, 2011.
70. B. Vandereycken and S. Vandewalle, *A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations*, SIAM J. Matrix Anal. Appl. **31** (2010), no. 5, 2553–2579.
71. F. Verstraete, V. Murg, and J. I. Cirac, *Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems*, Adv. Phys. **57** (2008), no. 2, 143–224.
72. S. R. White, *Density-matrix algorithms for quantum renormalization groups*, Phys. Rev. B **48** (1993), no. 14.
73. Harry Yserentant, *Regularity and approximability of electronic wave functions*, Lecture Notes In Mathematics, Springer, 2010.
74. T. Zhang and G. H. Golub, *Rank-one approximation to high order tensors*, SIAM J. Matrix Anal. Appl. **23** (2001), no. 2, 534–550.

Recent publications :

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

01.2012 A. ABDULLE, A. NONNENMACHER:
A posteriori error estimate in quantities of interest for the finite element heterogeneous multiscale method

02.2012 F. NOBILE, M. POZZOLI, C. VERGARA:
Time accurate partitioned algorithms for the solution of fluid-structure interaction problems in haemodynamics

03.2012 T. LASSILA, A. MANZONI, A. QUARTERONI, G. ROZZA:
Boundary control and shape optimization for the robust design of bypass anastomoses under uncertainty

04.2012 D. KRESSNER, C. TOBLER:
htucker – A Matlab toolbox for tensors in hierarchical Tucker format

05.2012 A. ABDULLE, G. VILLMART, KONSTANTINOS C. ZYGALAKIS:
Second weak order explicit stabilized methods for stiff stochastic differential equations.

06.2012 A. CABOUSSAT, S. BOYAVAL, A. MASSEREY:
Three-dimensional simulation of dam break flows.

07.2012 J BONNEMAIN, S. DEPARIS, A. QUARTERONI:
Connecting ventricular assist devices to the aorta: a numerical model.

08.2012 J BONNEMAIN, ELENA FAGGIANO, A. QUARTERONI, S. DEPARIS:
A framework for the analysis of the haemodynamics in patient with ventricular assist device.

09.2012 T. LASSILA, A. MANZONI, G. ROZZA:
Reduction strategies for shape dependent inverse problems in haemodynamics.

10.2012 C. MALOSSI, P. BLANCO, P. CROSETTO, S. DEPARIS, A. QUARTERONI:
Implicit coupling of one-dimensional and three-dimensional blood flow models with compliant vessels.

11.2012 S. FLOTRON J. RAPPAZ:
Conservation schemes for convection-diffusion equations with Robin's boundary conditions.

12.2012 A. USCHMAJEW, B. VANDEREYCKEN:
The geometry of algorithms using hierarchical tensors.