Numer. Math. (2012) 120:701–731
DOI 10.1007/s00211-011-0419-7

Numerische Mathematik

On manifolds of tensors of fixed TT-rank

Sebastian Holtz · Thorsten Rohwedder · Reinhold Schneider

Received: 2 November 2010 / Revised: 12 April 2011 / Published online: 29 September 2011

Abstract Recently, the format of TT tensors (Hackbusch and Kühn in J Fourier Anal Appl 15:706–722, 2009; Oseledets in SIAM J Sci Comput 2009, submitted; Oseledets and Tyrtyshnikov in SIAM J Sci Comput 31:5, 2009; Oseledets and Tyrtyshnikov in Linear Algebra Appl 2009, submitted) has turned out to be a promising new format for the approximation of solutions of high dimensional problems. In this paper, we prove some new results for the TT representation of a tensor $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and for the manifold of tensors of TT-rank $\underline{r}$. As a first result, we prove that the TT (or compression) ranks r_i of a tensor U are unique and equal to the respective separation ranks of U if the components of the TT decomposition are required to fulfil a certain maximal rank condition. We then show that the set $\mathbb{T}$ of TT tensors of fixed rank $\underline{r}$ locally forms an embedded manifold in $\mathbb{R}^{n_1 \times \cdots \times n_d}$, therefore preserving the essential theoretical properties of the Tucker format, but often showing an improved scaling behaviour. Extending a similar approach for matrices (Conte and Lubich in M2AN 44:759, 2010), we introduce certain gauge conditions to obtain a unique representation of the tangent space $\mathcal{T}_U \mathbb{T}$ of $\mathbb{T}$ and deduce a local parametrization of the TT manifold. The parametrisation of $\mathcal{T}_U \mathbb{T}$ is often crucial for an algorithmic treatment of high-dimensional time-dependent PDEs and minimisation problems (Lubich in From quantum to classical molecular dynamics: reduced methods and numerical analysis, 2008). We conclude with remarks on those applications and present some numerical examples.

S. Holtz · T. Rohwedder · R. Schneider (✉)
TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany
e-mail: schneidr@math.tu-berlin.de

S. Holtz
e-mail: sholtz@math.tu-berlin.de

T. Rohwedder
e-mail: rohwedde@math.tu-berlin.de

Mathematics Subject Classification (2010) 15A69 · 90C06 · 41A46 · 41A63 · 57N35

1 Introduction

The treatment of high-dimensional problems, typically of problems involving quantities from $\mathbb{R}^d$ for larger dimensions d, is still a challenging task for numerical approximation. This is owed to the principal problem that classical approaches for their treatment normally scale exponentially in the dimension d in both needed storage and computational time and thus quickly become computationally infeasible for sensible discretizations of problems of interest. To circumvent this "curse of dimensionality" [2], alternative paradigms in their treatment are needed. Recent developments, motivated by problems in data compression and data analysis, indicate that concepts of tensor product approximation, i.e. the approximation of multivariate functions depending on d variables $x_1, \ldots, x_d$ by sums and products of lower-dimensional quantities, often offer a flexible tool for the data sparse approximation of quantities of interest.

In particular, this approach sheds new perspectives on the numerical treatment of PDEs in high dimensions, turning up in various applications from natural sciences as for example in the simulation of chemical reactions and in quantum dynamics, in the treatment of the Fokker–Planck equation or of boundary value problems with stochastic data. In particular, the treatment of the stationary electronic Schrödinger equation has recently received a lot of attention. In this case, the incorporation of antisymmetry constraints stemming from the Pauli principle causes additional technical difficulties; in the treatment of the nuclear Schrödinger equation in quantum dynamics, the symmetry constraints are therefore often disregarded, i.e. atoms are treated as distinguishable particles [1].

Unfortunately, besides from the elementary (matrix) case $d = 2$, the two classical concepts from tensor product approximation [26], i.e. the *canonical decomposition* also known as *CANDECOMP*, *PARAFAC*, *canonical polyadic* or *Kronecker decomposition* on the one hand, and so-called *Tucker decomposition* on the other, suffer from different shortcomings. The canonical format, although surely scaling linearly with respect to the order d, the dimension n of the vector space and the canonical rank r, thus being ideal with regard to complexity, carries a lot of theoretical and practical drawbacks: The set of tensors of fixed canonical rank is not closed, and the existence of a best approximation is not guaranteed [8]. Although in some cases the approximation works quite well [10], optimization methods [11] often fail to converge as a consequence of uncontrollable redundancies in the parametrisation, and an actual computation of a low-rank approximation can thus be a numerically hazardous task. In contrast to this, the Tucker format, in essence corresponding to orthogonal projections into optimal subspaces of $\mathbb{R}^n$, still scales exponentially with the order d, only reducing the basis from n to the Tucker rank r. It thus still suffers from the curse of dimensionality—but provides a stable format from the perspective of practical computations: A quasi-optimal approximation of a given tensor can be computed by higher order SVD (HOSVD) [6]; a local best approximation of a given tensor can be computed by higher order orthogonal iteration (HOOI [7]), or by the Newton–Grassmann approach introduced in [9,39]. Alternatively, usage of alternating least square (ALS) approaches is also recommendable for computation of a local best approximation [21,22,27]. Computation of a global best approximation is an open problem though.

Most importantly, the Tucker format is also well applicable to the discretization of differential equations, e.g. in the context of the MRSCF approach [18] to quantum chemical problems or of multireference Hartree and Hartree-Fock methods (MR-HF) in quantum dynamics. From a theoretical point of view, the set of Tucker tensors of fixed rank forms an embedded manifold [25], and the numerical treatment therefore follows the general concepts of the numerical treatment of differential equations on manifolds [17].

On the whole, one is faced with an unsatisfactory situation, confirmed by experiences made through the past decade: On one hand, the canonical format gives an unstable representation of ideal complexity, which cannot be recommended without serious warnings. On the other hand, the stable Tucker format provides the basis for systematic discretization of e.g. initial value problems, but still carries the curse of dimensionality. Recent developments in the field of tensor approximation now seem to offer a way out of this dilemma: Based on the framework of subspace approximation, Hackbusch and Kühn [16] have recently introduced a hierarchical Tucker (HT) format in which only tensors of at most order 3 are used for representation of an order-d-tensor. An according decomposition algorithm using hierarchical singular value decompositions and providing a quasi-optimal approximation in the ℓ_2-sense has been introduced by Grasedyck [15]; Hackbusch has also recently shown the existence of a best approximation [12]. Independently, the *TT format* (abbreviating "tree tensor" or "tensor train"), a special case of the above HT structure, was recently introduced by Oseledets and Tyrtyshnikov [33,37,38]. This format offers one of the simplest kinds of representation of a tensor in the HT format, and although we conjecture that the results given in this paper can be generalized to the hierarchical tensor format, we will confine ourselves to the particular case of the TT format throughout this work for sake of simplicity.

Without reference to these recent developments, the basic ideas of these new formats have interestingly enough already been used for the treatment of various problems of many particle quantum physics since almost 20 years, e.g. in quantum dynamics [1], in the computation of finitely correlated states (FCS [13]) and valence bond solid states (VBS [23]), in the context of the DMRG algorithm [46], and under the name of matrix product states (MPS [45]) utilized in the description of spin systems in quantum information theory [41]. A generalization of these ideas are the so-called tensor networks [20,32], and although the viewpoints on the problems generally posed and solved in physics may be quite different, we think that it is not only of historical interest to mention and utilize the intimate relationship between those developments in different communities.

With this paper, we wish to make a contribution to the approximation of solutions of high-dimensional problems by TT tensors of fixed rank $\underline{r}$ (see Sect. 2 for the definition). After a review of the TT decomposition and the introduction of some notation, we will start in Sect. 3 by showing that for a tensor $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, its TT rank (sometimes also termed compression rank) can be uniquely defined in terms of in some sense optimal TT decompositions of U (i.e. decompositions of minimal rank), and equals the so-called separation rank of U, see Sect. 2.5 and Theorem 1. We then continue by analysing the set $\mathbb{T}$ of tensors of fixed TT rank $\underline{r}$, forming a nonlinear manifold. For a formulation of algorithms on $\mathbb{T}$, it is helpful to understand the analytical structure of $\mathbb{T}$, in particular that of its tangent space: If we take the approach persued e.g. in [24] for the low rank approximation of solutions of high dimensional differential equations, or as for other manifolds in [25,31], the original problem is in each iteration step solved on the tangent space $\mathcal{T}_U\mathbb{T}$ of $\mathbb{T}$, taken at the current iterate U—an approach which can be viewed as a Galerkin approximation with the approximation space $\mathcal{T}_U\mathbb{T}$ depending on the current iterate U. Therefore, we show in Sect. 4 that the tangent space $\mathcal{T}_U\mathbb{T}$ of the TT manifold $\mathbb{T}$ of fixed maximal rank $\underline{r}$, taken at some $U \in \mathbb{T}$, can be uniquely represented by introducing gauge conditions similar to those used in [17,31], see Theorem 2. From our result, we also deduce a unique local parametrization of the TT manifold $\mathbb{T}$ in Sect. 5, Theorem 3. Roughly speaking, our main result states that the manifold of tensors of fixed TT-rank locally provides an embedded manifold, therefore preserving the essential properties of the Tucker format with a complexity scaling linearly with respect to n and d, but only quadratically with respect to the ranks. Section 6 uses the results of Sects. 4 and 5 to exemplify the scope of algorithmic applications of the tangent space $\mathcal{T}_U\mathbb{T}$ in the context of approximating solutions of optimization problems and differential equations in high-dimensional spaces; in Sect. 7 we finally demonstrate that best approximations on $\mathbb{T}$ can be computed stably in practice.

2 Review of the TT-tensor representation: notations and definitions

In this section, we will first of all review the basic idea and the various formats used in the different contexts utilizing the TT approximation. As the treatment of tensors naturally involves multi-index quantities, notational matters often tend to make the basic ideas hard to grasp. We will therefore—alongside with the "classical" notation—use a graphical notation inspired by [44] and similar approaches in quantum chemistry, see e.g. [5,42]. This graphical denotation of tensors and tensor manipulations, as for instance of partial summations over a subset of indices, contraction of mutual indices of different tensors etc., has in many cases in this paper led us to the ideas of proof in the first place and helped us towards a much clearer understanding of the subject. Therefore, we will introduce this notation to the reader here, stress its usefulness and hope that the reader feels the same.

2.1 General notes: graphical representation

In the present paper, we are dealing with the representation or approximation of tensors $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ by tensors given in the TT format. In this, the order (or dimension) $d \in \mathbb{N}$ as well as $n_1, \ldots, n_d$, determining finite index sets $\mathcal{I}_i := \{1, \ldots, n_i\}$ for all $i \in \{1, \ldots, d\}$, will be fixed in the following. For convenience of exposition, we will only treat real-valued tensors here. We will often regard a tensor U as a multivariate function depending on d variables $x_1, \ldots, x_d$, $x_i \in \mathcal{I}_i$, and write in the form

$$U : \mathcal{I}_1 \times \cdots \times \mathcal{I}_d \to \mathbb{R}, \qquad \underline{x} = (x_1, \ldots, x_d) \;\mapsto\; U(x_1, \ldots, x_d). \tag{1}$$

If not prone to arouse confusion, we will sometimes denote $U : \underline{x} \mapsto U(\underline{x})$ by $U(\underline{x})$ for sake of brevity.

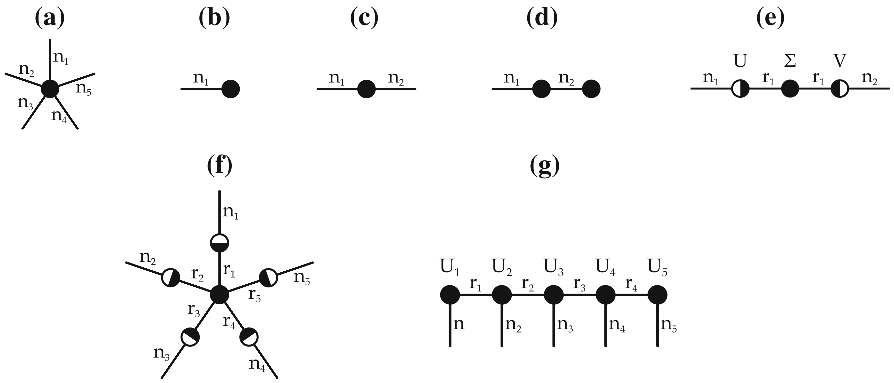

Fig. 1 Examples of graphical representations of tensors, see Sect. 2.1

Using the graph notation introduced in [44], a tensor U of the above form (1) is represented by a dot with d "arms", depicting the d free variables $x_1, \ldots, x_d$. For example, a tensor (1) for which $d = 5$ can be symbolized as in Fig. 1a. Figure 1b, c illustrates the special cases $d = 1$ and $d = 2$, i.e. that of a vector $\mathbf{x} \in \mathbb{R}^{n_1}$ and a matrix $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$, respectively. Many operations of multilinear algebra involve summations over one or more of the indices (or variables) $x_i = 1, \ldots, n_i$, and such summations are conveniently depicted in the graph representation by joining the respective "arms" of the involved tensors. Thus, a matrix–vector multiplication looks as in (d), yielding again an one-index quantity corresponding to one "free arm", while an SVD may be depicted as in (e), with the white two-armed dots representing orthogonal matrices. A globalisation of the SVD is the so-called Tucker decomposition [43] as depicted in (f).

2.2 The TT format

A representation of a tensor $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ in the TT format rewrites an order-d tensor of the form (1) as a suitable product of two matrices

$$\begin{aligned} U_1 : (x_1, k_1) &\mapsto U_1(x_1, k_1) \in \mathbb{R}^{n_1 \times r_1}, \\ U_d : (k_{d-1}, x_d) &\mapsto U_d(k_{d-1}, x_d) \in \mathbb{R}^{r_{d-1} \times n_d} \end{aligned} \tag{2}$$

and $d - 2$ tensors $U_i \in \mathbb{R}^{r_{i-1} \times n_i \times r_i}$, $\ 2 \le i \le d-1$ of order 3,

$$U_i : (k_{i-1}, x_i, k_i) \mapsto U_i(k_{i-1}, x_i, k_i) \in \mathbb{R}^{r_{i-1} \times n_i \times r_i}. \tag{3}$$

In terms of these quantities, an order-d-tensor U of the form (1) can be decomposed into a TT tensor, that is, $U(\underline{x}) = U(x_1, \ldots, x_d)$ can be written as

$$U(\underline{x}) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} U_1(x_1, k_1) \Big(\prod_{\mu=2}^{d-1} U_\mu(k_{\mu-1}, x_\mu, k_\mu) \Big) U_d(k_{d-1}, x_d). \tag{4}$$

The numbers r_i are the so-called *compression ranks*, determining the sparsity of the representation (4). The (probably more accessible) graphical representation of the TT tensor (4) is given in Fig. 1g.

Note that auxiliary variables $k_i, i = 1, \ldots, d-1$ have been introduced, "connecting" the single components $U_i(k_{i-1}, x_i, k_i)$ depending on solely one of the old variables x_i. By the above TT decomposition, the storage requirements can usually be reduced dramatically, e.g. from n^d to no more than $r_{max}^2 nd$, where r_{max} is the maximum over $r_i, i = 1, \ldots, d-1$. A decomposition of U of the form (4) can for instance be computed by successive singular value decompositions, see [33] or Sect. 3.2. Note also that a decomposition of U of the form (4) is highly non-unique, see the remarks in Sect. 2.3.

2.3 Component functions and matrix product representation

The matrices U_1, U_d from (2) can be interpreted as vector-valued functions, defined by

$$\begin{aligned} \mathbf{U}_1 : \mathcal{I}_1 \to \mathbb{R}^{r_1}, \quad & \mathbf{U}_1(x_1) : \quad k_1 \mapsto U_1(x_1, k_1), \\ \mathbf{U}_d : \mathcal{I}_d \to \mathbb{R}^{r_{d-1}}, \quad & \mathbf{U}_d(x_d) : \quad k_d \mapsto U_d(k_{d-1}, x_d) \end{aligned}$$

respectively (Note that for fixed x_1, $\mathbf{U}_1(x_1)$ is a vector, here written as univariate functions themselves; the same holds for $\mathbf{U}_d(x_d)$.) Analogously, for $2 \le i \le d-1$, the 3-d tensors U_i from (3) can be seen as matrix-valued functions

$$\mathbf{U}_i : \mathcal{I}_i \to \mathbb{R}^{r_{i-1} \times r_i}, \qquad \mathbf{U}_i(x_i) : (k_{i-1}, k_i) \mapsto U_i(k_{i-1}, x_i, k_i). \tag{5}$$

We will call these functions the *component functions* of the TT representation (4) of the tensor U. The cases $i = 1$ and $i = d$ are formally included in the notation (5) by letting $r_0 = 1$ and $r_d = 1$, respectively. We will also from time to time use the sets

$$C_i := \{\mathbf{U}_i : \mathcal{I}_i \to \mathbb{R}^{r_{i-1} \times r_i}\} \tag{6}$$

of all i-th component functions. The value $U(\underline{x})$ of U at $\underline{x} = (x_1, \ldots, x_d) \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_d$ can now conveniently be written in the *matrix product representation*,

$$U(\underline{x}) = \mathbf{U}_1(x_1)\mathbf{U}_2(x_2) \cdot \ldots \cdot \mathbf{U}_{d-1}(x_{d-1})\mathbf{U}_d(x_d), \tag{7}$$

and we will rather use the notation (7) than that in (4) in the following. Note that for fixed $\underline{x}$, $\mathbf{U}_1(x_1) \in \mathbb{R}^{1 \times r_1}$ is a row vector and $\mathbf{U}_d(x_d) \in \mathbb{R}^{r_{d-1} \times 1}$ is a column vector, so that $U(\underline{x})$ can be evaluated by a repeated computation of matrix–vector products, explaining the terminology. As introduced here, we will use bold-faced letters for all matrices, vectors and matrix- or vector-valued functions throughout this work.

In the representation (7) of a tensor U, multiplication of the component function $\mathbf{U}_i$ from the right with any invertible matrix $\mathbf{A} \in \mathbb{R}^{r_i \times r_i}$ and simultaneously of $\mathbf{U}_{i+1}$ with $\mathbf{A}^{-1}$ from the left yields a different TT decomposition for U, showing that a

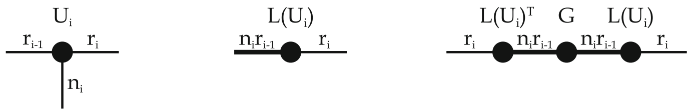

Fig. 2 Illustration of the quantities $\mathbf{U}_i$, $\mathbf{L}(\mathbf{U}_i)$, $[\mathbf{U}_i, \mathbf{U}_i]_{\mathbf{G}}$

decomposition of U of the form (4) is highly non-unique. In Sect. 3, we will formulate conditions under which we obtain a certain uniqueness of the TT decomposition of a tensor U, see Theorem 1.

2.4 Left and right unfoldings

For the treatment of TT tensors due in the following chapters, we will use some terminology concerned with unfoldings of the above component functions, given in the following.

The *left unfolding* of a component function (5) is obtained by unfolding the tensor $(k_{i-1}, x_i, k_i) \mapsto U_i(k_{i-1}, x_i, k_i)$, taking the indices k_{i-1}, x_i of as row indices and the indices k_i as column indices (using a fixed ordering of indices). This yields a matrix we denote by

$$\mathbf{L}(\mathbf{U}_i) \in \mathbb{R}^{(r_{i-1}n_i)\times r_i}.$$

See also Fig. 2 for an illustration of the quantities $\mathbf{U}_i$ and $\mathbf{L}(\mathbf{U}_i)$. The unfolding mapping

$$\mathbf{L}: \ \mathbf{U}_i \ \mapsto \ \mathbf{L}(\mathbf{U}_i) \tag{8}$$

defines a linear bijection between C_i and $\mathbb{R}^{(r_{i-1}n_i)\times r_i}$. We define the *left rank* of a component function $\mathbf{U}_i(\cdot)$ as the rank of the matrix $\mathbf{L}(\mathbf{U}_i) \in \mathbb{R}^{r_i \times r_i}$.

Analogously to the above, we define the *right unfolding* $\mathbf{R}(\mathbf{U}_i) \in \mathbb{R}^{r_{i-1}\times(n_i r_i)}$ of $\mathbf{U}_i$ as the matrix obtained by taking the indices k_{i-1} of $\mathbf{U}_i$ as row indices and the indices x_i, k_i as column indices, and the *right rank* of $\mathbf{U}_i$ as the rank of $\mathbf{R}(\mathbf{U}_i)$.

2.5 TT decompositions of minimal rank, full-rank-condition, TT rank and separation rank

Let $U \in \mathbb{R}^{n_1\times\cdots\times n_d}$ be an arbitrary tensor. A TT decomposition

$$U(\underline{x}) = \mathbf{U}_1(x_1)\mathbf{U}_2(x_2)\cdot\ldots\cdot\mathbf{U}_{d-1}(x_{d-1})\mathbf{U}_d(x_d) \tag{9}$$

of the tensor U will be called *minimal* or *fulfilling the full-rank-condition* if all component functions $\mathbf{U}_i$ have full left and right rank, i.e. the rank of the left unfolding of $\mathbf{U}_i$ is r_i and the rank of the right unfolding of $\mathbf{U}_i$ is r_{i-1}. For sake of brevity we will sometimes denote the tensor U given pointwise by (9) as $U = \mathbf{U}_1 \cdot \ldots \cdot \mathbf{U}_d$. If

(9) is a given minimal TT decomposition of U, consisting of component functions $\mathbf{U}_i : \mathcal{I}_i \to \mathbb{R}^{r_{i-1}\times r_i}$ having full left rank r_i and full right rank r_{i-1}, we will call (9) a *decomposition of TT rank*

$$\underline{r} := (r_1, \ldots, r_{d-1}) \tag{10}$$

of U.

Note that *a priori*, there may be different TT decompositions (9) of U, having different TT ranks $\underline{r}$, so that the TT rank is a property of a particular decomposition of U, not of U itself. In the next section, we will show though that for any minimal TT decomposition of U, the TT rank of this minimal decomposition $\underline{r}$ is equal to the separation rank $\underline{s}$ of U, defined next—thus, a minimal $\underline{r}$ is uniquely determined for each tensor $U \in \mathbb{R}^{n_1\times\cdots\times n_d}$, and the TT rank of a tensor U will be defined as this minimal $\underline{r}$ in Theorem 1.

By $\mathbf{A}_i \in \mathbb{R}^{(n_1\cdot\ldots\cdot n_i)\times(n_{i+1}\cdot\ldots\cdot n_d)}$, we denote the i-th canonical unfolding (or matrification) of U, i.e. the matrix obtained by taking the indices $x_1, \ldots, x_i$ of U as row indices and the indices $x_{i+1}, \ldots, x_d$ as column indices, see e.g. [26] for details. The *i-th separation rank s_i* of U is then defined as the rank of the i-th canonical unfolding $\mathbf{A}_i$. The vector

$$\underline{s} := (s_1, \ldots, s_{d-1}) \tag{11}$$

will be called the *separation rank* of U.

2.6 Orthogonality constraints

We will say that $\mathbf{U}_i$ and $\mathbf{W}_i$ are *mutually left-orthogonal* with respect to an inner product induced by some symmetric positive definite matrix $\mathbf{G} \in \mathbb{R}^{(r_{i-1}n_i)\times(r_{i-1}n_i)}$ iff

$$[\mathbf{U}_i, \mathbf{W}_i]_{\mathbf{G}} := \big(\mathbf{L}(\mathbf{U}_i)\big)^T \mathbf{G}\mathbf{L}(\mathbf{W}_i) = \mathbf{0} \in \mathbb{R}^{r_i\times r_i}, \tag{12}$$

i.e. the columns of the left unfoldings of $\mathbf{U}_i$ and $\mathbf{W}_i$ are mutually orthogonal with respect to the inner product induced by $\mathbf{G}$. See also Fig. 2 for a graphical representation of $[\mathbf{U}_i, \mathbf{W}_i]_{\mathbf{G}}$.

For the case $\mathbf{G} = \mathbf{I} \in \mathbb{R}^{r_{i-1}n_i\times r_{i-1}n_i}$, a component function $\mathbf{U}_i$ will be called *left-orthogonal* iff

$$[\mathbf{U}_i, \mathbf{U}_i] := [\mathbf{U}_i, \mathbf{U}_i]_{\mathbf{I}} = \mathbf{I} \in \mathbb{R}^{r_i\times r_i} \tag{13}$$

holds, i.e. if the columns of $\mathbf{L}(\mathbf{U}_i)$ are an orthonormal system. In the same way we define that $\mathbf{U}_i$ is *right-orthogonal* iff

$$\mathbf{R}(\mathbf{U}_i)\big(\mathbf{R}(\mathbf{U}_i)\big)^T = \mathbf{I} \in \mathbb{R}^{r_{i-1}\times r_{i-1}}. \tag{14}$$

holds.

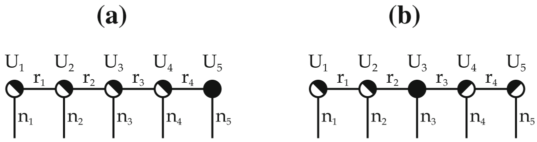

Fig. 3 Different formats for TT decompositions

Without going into much detail, we note that a tensor may be represented in various TT formats being special cases or equivalent representations of (7). For example, the canonical proceeding of SVD from left to right gives a tensor with the components $\mathbf{U}_i$ being left-orthogonal for $i = 1, \ldots, d$, as depicted in Fig. 3a, where the dots being white at the arms belonging to the indices k_{i-1}, x_i indicate left-orthogonality. More globally, one can pick $i \in \{1, \ldots, d\}$ and decompose U by SVDs from the left and the right into a TT tensor the components of which are left-orthogonal for $j < i$ and right-orthogonal for $j > i$, cf. (b). Subsequent QR-decomposition of $\mathbf{U}_i$ yields another equivalent representation with invertible D, as given in (c). The index i chosen above may even vary during some algorithmic applications, as is for instance the case in the intermediate stages of the DMRG algorithm used in quantum physics (cf. [46]).

3 Uniqueness statements for the TT rank $\underline{r}$ and for TT decompositions

In this section, we prove the following theorem. Part (a) is an extension of results from the two publications [34,35], where existence of TT decompositions with TT ranks $r_i \leq s_i$ and $r_i = s_i$, respectively, is proven. It shows that for each minimal TT decomposition, its TT rank coincides with the separation rank and thus is a uniquely defined quantity for each tensor U. In (b), we then give a certain uniqueness statement for TT decompositions of minimal rank of a tensor U. Part (c) shows that in practice, a minimal rank TT decomposition can be computed by successive SVDs, i.e. by the algorithm proposed by Oseledets in [33].

Theorem 1 (Uniqueness of TT decompositions; the TT rank of a tensor) *Let* $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ *an arbitrary tensor.*

(a) *There is exactly one rank vector* $\underline{r}$ *such that* U *admits for a TT decomposition*

$$U : \underline{x} \mapsto U(\underline{x}) = \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_d(x_d),$$

of minimal rank $\underline{r}$. *If* $\underline{s} = \underline{s}(U)$ *denotes the (unique) separation rank of* U, *there holds*

$$\underline{r} = \underline{s}. \tag{15}$$

Therefore, the separation rank $\underline{r} = \underline{s}$ *will also be called the* TT rank *of the tensor* U.

Input: $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$; **Output:** $\mathbf{U}_i \in C_i$, $i = 1, \ldots, d$.
Set $\mathbf{B}_{(1)} = \mathbf{A}_1$ ($\mathbf{A}_1$ is the first canonical unfolding)
Set $r_1 := \operatorname{rank} \mathbf{B}_{(1)}$
Compute the SVD $\mathbf{B}_{(1)} = \mathbf{U}_1 \mathbf{V}_{(1)}$, $\quad \mathbf{U}_1 \in \mathbb{R}^{n_1 \times r_1}, \mathbf{V}_{(1)} \in \mathbb{R}^{r_1 \times (n_2 \cdot \ldots \cdot n_d)}$
with left-orthogonal $\mathbf{U}_1$
for $i = 2, \ldots, d-1$ **do**
Rearrange $\mathbf{V}_{(i-1)} \in \mathbb{R}^{r_{i-1} \times (n_i \cdot \ldots \cdot n_d)}$ to $\mathbf{B}_{(i)} \in \mathbb{R}^{(r_{i-1} \cdot n_i) \times (n_{i+1} \cdot \ldots \cdot n_d)}$
Compute the SVD
$\mathbf{B}_{(i)} = \mathbf{L}(\mathbf{U}_i) \mathbf{V}_{(i)}$, $\quad \mathbf{L}(\mathbf{U}_i) \in \mathbb{R}^{(r_{i-1} \cdot n_i) \times r_i}, \mathbf{V}_{(i)} \in \mathbb{R}^{r_i \times (n_{i+1} \cdot \ldots \cdot n_d)}$.
giving left-orthogonal $\mathbf{U}_i$, TT-rank $r_i := \operatorname{rank} \mathbf{B}_{(i)}$
end for
Set $\mathbf{U}_d(k_{d-1}, x_d) := \mathbf{V}_{(d)}(k_{d-1}, x_d)$

Fig. 4 The successive SVD-algorithm for computing a TT decomposition, taken from Oseledets [33]

(b) *The TT decomposition* (15) *of* U *of minimal rank can be chosen such that the component functions are left-orthogonal,*

$$[\mathbf{U}_i, \mathbf{U}_i] = \mathbf{I} \in \mathbb{R}^{r_i \times r_i} \tag{16}$$

for all $i = 1, \ldots, d-1$ (*see* [33] *for the constructive proof*). *Under this condition, the decomposition* (15) *is unique up to insertion of orthogonal matrices: For any two left-orthogonal minimal decompositions of* U *for which*

$$U(\underline{x}) = \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_d(x_d) = \mathbf{V}_1(x_1) \cdot \ldots \cdot \mathbf{V}_d(x_d) \tag{17}$$

holds for all $\underline{x} = (x_1, \ldots, x_d) \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_d$, *there exist orthogonal* $\mathbf{Q}_1, \ldots, \mathbf{Q}_{d-1}$, $\mathbf{Q}_i \in \mathbb{R}^{r_i \times r_i}$ *such that*

$$\begin{aligned} &\mathbf{U}_1(x_1)\mathbf{Q}_1 = \mathbf{V}_1(x_1), \qquad \mathbf{Q}_{d-1}^T \mathbf{U}_d(x_d) = \mathbf{V}_d(x_d), \\ &\mathbf{Q}_{i-1}^T \mathbf{U}_i(x_i) \mathbf{Q}_i = \mathbf{V}_i(x_i) \;\; \textit{for} \;\; i = 2, \ldots, d-1, \;\; \underline{x} \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_d. \end{aligned} \tag{18}$$

(c) *Let* $U \neq 0$. *In Fig.* 4, *we reproduced the SVD-based TT decomposition algorithm introduced in Oseledets* [33]. *This algorithm, when applied to* U *without truncation steps and in exact arithmetic, returns a minimal TT decomposition with left-orthogonal component functions* $\mathbf{U}_i$ (*see* (16)).

Remark 1 Theorem 1 uniquely defines a minimal TT rank $\underline{r}$ for any tensor U. This rank $\underline{r}$ is attained by any TT decomposition of U fulfilling the full-rank-condition (see Sect. 2.5). Note that $\underline{r}$ is not invariant under permutations of indices (because the separation rank is not). In particular, the arrangement of indices might influence the complexity of the storage needed to represent U in the TT format.

Also, Theorem 1 implies a new (without further information sharp) bound on the TT rank of a tensor U: Because

$$m_i := \min\Big\{\prod_{j=1}^{i} n_j, \prod_{j=i+1}^{d} n_j\Big\} \tag{19}$$

defines the maximal rank possible for an SVD of the i-th canonical unfolding, Theorem 1 shows

$$r_i \le m_i \tag{20}$$

for all $i = 1, \ldots, d-1$.

Before we approach the proof of Theorem 1, we finally note that the uniqueness statements of Theorem 1 hold analogously if for fixed $i \in \{1, \ldots, d\}$, the left-orthogonality conditions (16) are replaced by left-orthogonality in the first $j < i$ components and by the according right-orthogonality condition $\mathbf{R}(\mathbf{U}_j)\big(\mathbf{R}(\mathbf{U}_j)\big)^T = \mathbf{I}$ for $j > i$, so that tensors are required to be of the form depicted in Fig. 3b. Analogous globalisations hold for Theorem 2, Theorem 3 (with the gauge condition (35) modified appropriately).

3.1 Notations for left and right parts of a TT tensor

For the proof of Theorem 1 and also in the later chapters, we will need some more formal notation concerned with certain unfoldings of parts of U, and they are introduced in the following.

Let $i \in \{1, \ldots, d\}$ and the tensor $U^{\le i} \in \mathbb{R}^{n_1 \times \cdots \times n_i \times r_i}$ be defined via U, given pointwise by

$$U^{\le i}(x_1, \ldots, x_i, k_i) := \sum_{k_1=1}^{r_1} \ldots \sum_{k_{i-1}=1}^{r_{i-1}} \prod_{\mu=1}^{i} U_\mu(k_{\mu-1}, x_\mu, k_\mu), \tag{21}$$

i.e. by summing up the first i component functions. Aside from $U^{\le i}$, we will also often use a specific unfolding of $U^{\le i}$ obtained by taking $x_1, \ldots, x_i$ as row indices and k_i as column index. We denote this matrix by

$$\mathbf{U}^{\le i} \in \mathbb{R}^{(n_1 \cdot \ldots \cdot n_i) \times r_i}$$

and refer to as the *i-th left part* of U.

Analogously, we define $U^{\ge i} \in \mathbb{R}^{r_{i-1} \times n_i \times \cdots \times n_d}$ pointwise by

$$U^{\ge i}(k_{i-1}, x_i, \ldots, x_d) := \sum_{k_i=1}^{r_i} \ldots \sum_{k_{d-1}=1}^{r_{d-1}} \prod_{\mu=i}^{d} U_\mu(k_{\mu-1}, x_\mu, k_\mu), \tag{22}$$

i.e. as summing up of the i-th to d-th component functions. We then define the *i-th right part* $\mathbf{U}^{\ge i} \in \mathbb{R}^{r_{i-1} \times (n_i \cdot \ldots \cdot n_d)}$ to be the unfolding of the tensor $U^{\ge i}$ that takes

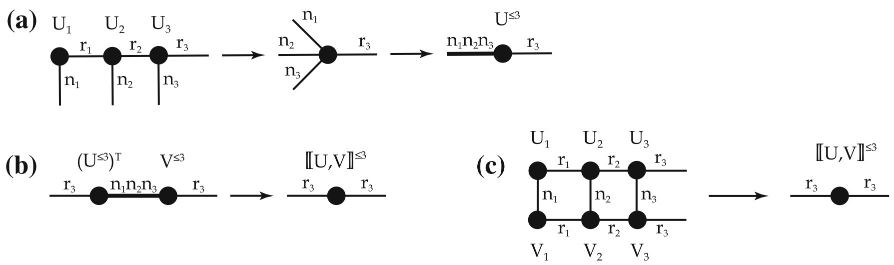

Fig. 5 **a** Illustration of $U^{\leq i}$. **b**, **c** Two ways to obtain $[\![U, V]\!]^{\leq i}$

$x_i, \ldots, x_d$ as column indices and k_i as row index. For formal reasons, we additionally define $\mathbf{U}^{\leq 0} := (1) =: \mathbf{U}^{\geq d+1}$. Note that $\mathbf{U}^{\leq d}$ and $\mathbf{U}^{\geq 1}$ yield a vectorization of the tensor U as a column vector and as a row vector, respectively.

For two given TT representations of tensors U, V and $i \in \{0, \ldots, d\}$, we define the *i-th left half product matrix* as

$$[\![U, V]\!]^{\leq i} := (\mathbf{U}^{\leq i})^T \mathbf{V}^{\leq i} \in \mathbb{R}^{r_i \times r_i},$$

i.e. pointwise by

$$[\![U, V]\!]^{\leq i}(k_i, k_i') = \sum_{x_1=1}^{n_1} \cdots \sum_{x_i=1}^{n_i} U^{\leq i}\big(x_1, \ldots, x_i, k_i\big) V^{\leq i}\big(x_1, \ldots, x_i, k_i'\big) \tag{23}$$

for $k_i, k_i' \in \{1, \ldots, r_i\}$.

Analogously, we introduce for $i \in \{1, \ldots, d+1\}$ the *i-th right half product matrix*,

$$[\![U, V]\!]^{\geq i} \; := \; \mathbf{U}^{\geq i}(\mathbf{V}^{\geq i})^T \in \mathbb{R}^{r_{i-1} \times r_{i-1}}.$$

Figure 5a, b illustrates the quantities $\mathbf{U}^{\leq i}$ and $[\![U, V]\!]^{\leq i}$.

Using this notation, we derive following recursive formula will play a central role later: For $i \in \{1, \ldots, d\}$, for $k_i, k_i' \in \{1, \ldots, r_i\}$ and using $U^{\leq i}, U^{\geq i}$ from (21),(22), we observe that

$$\begin{aligned} [\![U, U]\!]^{\leq i}(k_i, k_i') &= \sum_{x_1=1}^{n_1} \cdots \sum_{x_i=1}^{n_i} U^{\leq i}(x_1, \ldots, x_i, k_i) U^{\leq i}(x_1, \ldots, x_i, k_i') \\ &= \sum_{k_{i-1}, k_{i-1}'=1}^{r_{i-1}} \sum_{x_i=1}^{n_i} [\![U, U]\!]^{\leq i-1}(k_{i-1}, k_{i-1}') U_i(k_{i-1}, x_i, k_i) U_i(k_{i-1}', x_i, k_i') \end{aligned}$$

and that

$$\sum_{x_i=1}^{n_i} U_i(k_{i-1}, x_i, k_i) U_i(k_{i-1}', x_i, k_i') = \sum_{x_i, x_i'=1}^{n_i} U_i(k_{i-1}, x_i, k_i) \mathbf{I}(x_i, x_i') U_i(k_{i-1}', x_i', k_i'),$$

so that abbreviating $\mathbf{G}_i := [\![U, U]\!]^{\leq i-1} \otimes \mathbf{I}_{n_i \times n_i}$, we obtain the important formula

$$[\![U, U]\!]^{\leq i} = \mathbf{L}(\mathbf{U}_i)^T \left([\![U, U]\!]^{\leq i-1} \otimes \mathbf{I}_{n_i \times n_i}\right)\mathbf{L}(\mathbf{U}_i) = \mathbf{L}(\mathbf{U}_i)^T \mathbf{G}_i \mathbf{L}(\mathbf{U}_i). \tag{24}$$

3.2 Proof of Theorem 1(a)

We start the proof of Theorem 1(a) by showing the existence of a TT decomposition of minimal rank. First of all, we note that the zero tensor $0 \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ can be written as an elementary tensor built up from the zero vectors $0^{(i)} \in \mathbb{R}^{n_i}$, giving a minimal TT decomposition for 0. We now show that for a non-trivial tensor $U \neq 0$, a TT representation of minimal rank can be computed by the algorithm given in Fig. 4.

Lemma 1 *For a given non-trivial tensor $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, the algorithm in Fig. 4 returns (in exact arithmetic) a TT decomposition $(\mathbf{U}_1, \ldots, \mathbf{U}_d) \in C_1 \times \cdots \times C_d$ with full left and right rank for each component function $\mathbf{U}_i$, $i = 1, \ldots, d$.*
Also, the component functions $\mathbf{U}_i$, $i = 1, \ldots, d-1$ are left-orthogonal.

Proof From the properties of the SVD, it is clear that the algorithm gives back a representation with full left rank and left-orthogonal component functions $\mathbf{U}_1, \ldots, \mathbf{U}_{d-1}$. Assume that for some $i \in \{1, \ldots, d\}$, a component function $\mathbf{U}_i$ is without full right rank. If $i = 1$, U is the zero tensor in contradiction to the assumption that U is non-trivial. If we have $\hat{r} := \operatorname{rank} \mathbf{R}(\mathbf{U}_i) < r_{i-1}$ for $i > 1$, an SVD of $\mathbf{R}(\mathbf{U}_i)$ yields a decomposition

$$\mathbf{R}(\mathbf{U}_i) = \mathbf{V}\mathbf{W}, \qquad \mathbf{V} \in \mathbb{R}^{r_{i-1} \times \hat{r}}, \quad \mathbf{W} \in \mathbb{R}^{\hat{r} \times n_i r_i}$$

with $\mathbf{V}$ orthogonal and $\mathbf{W}$ of rank $\hat{r}$. Setting $\widehat{\mathbf{U}}_{i-1} := \mathbf{L}^{-1}\big(\mathbf{L}(\mathbf{U}_{i-1})\mathbf{V}\big)$ and $\widehat{\mathbf{U}}_i := \mathbf{R}^{-1}(\mathbf{W})$, there holds

$$\widehat{\mathbf{U}}_{i-1}(x_{i-1})\widehat{\mathbf{U}}_i(x_i) = \mathbf{U}_{i-1}(x_{i-1})\mathbf{U}_i(x_i).$$

This implies that the unfolded matrix $\mathbf{B}_{(i-1)}$ from the algorithm in Fig. 4 may be decomposed to

$$\sum_{\hat{k}=1}^{\hat{r}} \hat{U}_{i-1}(k_{i-2}, x_{i-1}, \hat{k}) \left(\sum_{k_i=1}^{r_i} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} \hat{U}_i(\hat{k}, x_i, k_i) \prod_{\mu=i+1}^{d} U_\mu(k_{\mu-1}, x_\mu, k_\mu) \right)$$

i.e. as

$$\mathbf{B}_{(i-1)} = \mathbf{M}\mathbf{N} \quad \text{with} \quad \mathbf{M} \in \mathbb{R}^{(r_{i-2} \cdot n_{i-1}) \times \hat{r}}, \quad \mathbf{N} \in \mathbb{R}^{\hat{r} \times (n_i \cdot \ldots \cdot n_d)}$$

which can be rearranged to give a rank-$\hat{r}$-SVD of $\mathbf{B}_{(i-1)}$. Because of the uniqueness of SVD ranks, this would mean that the left rank r_{i-1} of $\mathbf{U}_{i-1}$ equals $\hat{r}$, a contradiction. □

The preceding lemma already proves part (c) of Theorem 1. For the proof of (a), it remains to show uniqueness of the TT rank $\underline{r} = \underline{r}(U)$ for a given tensor U. The proof bases essentially on the observation (a) in the next lemma. The statement made in (b) will also be used later to prove the central Theorem 2 of Sect. 4.

Lemma 2 *Let $U \in \mathbb{T}$ be a minimal TT-tensor of rank $\underline{r}$.*

(a) *For all $i \in \{1, \dots, d-1\}$, the matrices $\mathbf{U}^{\leq i}, \mathbf{U}^{\geq i+1}$ have rank r_i.*
(b) *The matrices*

$$\mathbf{G}_i := [\![U, U]\!]^{\leq i-1} \otimes \mathbf{I}_{n_i \times n_i} \in \mathbb{R}^{(r_{i-1}n_i)\times(r_{i-1}n_i)} \tag{25}$$

and

$$\mathbf{P}_i := [\![U, U]\!]^{\geq i} \in \mathbb{R}^{r_{i-1}\times r_{i-1}} \tag{26}$$

are symmetric and positive definite for all $i \in \{1, \dots, d\}$.

Proof If $U \in \mathbb{T}$ is a minimal TT-tensor, the component functions $\mathbf{U}_i$ have full left and right ranks r_i resp. r_{i-1} for all $1 \leq i \leq d$. We only show that if $\mathbf{U}_j$ has full left rank for all $1 \leq j < i$ (i.e. $\mathbf{L}(\mathbf{U}_j)^T\mathbf{L}(\mathbf{U}_j)$ has full rank r_j), then $[\![U, U]\!]^{\leq i-1} := (\mathbf{U}^{\leq i-1})^T\mathbf{U}^{\leq i-1}$ and $\mathbf{G}_i$ are positive definite; in particular, we then obtain that $\mathbf{U}^{\leq i}$ has full rank r_i for all $1 \leq i \leq d-1$. An analogous argument applies to $\mathbf{R}(\mathbf{U}_i)\big(\mathbf{R}(\mathbf{U}_i)\big)^T$, finishing the proof. We proceed by induction. For $i = 1$, $[\![U, U]\!]^{\leq 0} = (1)$ and $\mathbf{G}_1$ are positive definite. For the induction step, let the hypothesis hold for $(\mathbf{U}^{\leq i-1})^T\mathbf{U}^{\leq i-1}$ and $\mathbf{G}_i$, and let $\mathbf{L}(\mathbf{U}_i)^T\mathbf{L}(\mathbf{U}_i)$ have full rank. Then, the columns of $\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i)$ are linearly independent. This means that the Gramian matrix $(\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i))^T(\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i))$ is positive definite, i.e. there holds

$$\big(\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i)\big)^T(\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i)) > 0.$$

Using the recursion formula (24), $[\![U, U]\!]^{\leq i}$ can be rewritten inductively as

$$\begin{aligned} [\![U, U]\!]^{\leq i} &= \mathbf{L}(\mathbf{U}_i)^T\big([\![U, U]\!]^{\leq i-1} \otimes \mathbf{I}_{n_i\times n_i}\big)\mathbf{L}(\mathbf{U}_i) \\ &= (\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i))^T(\mathbf{G}_i^{1/2}\mathbf{L}(\mathbf{U}_i)) > 0, \end{aligned}$$

which also implies that $\mathbf{G}_{i+1} := [\![U, U]\!]^{\leq i} \otimes \mathbf{I}_{n_{i+1}\times n_{i+1}} > 0$ holds, completing the proof. □

We are now in the position to show that for any minimal TT decomposition, its TT rank is equal to the separation rank of U: The i-th canonical unfolding $\mathbf{A}_i$ of U can be written as the matrix product

$$\mathbf{A}_i = \mathbf{U}^{\leq i}\mathbf{U}^{\geq i+1}$$

of the i-th left part and the $(i+1)$-th right part of U, both having full rank r_i according to Lemma 2(a). We use the QR-decompositions

$$\mathbf{U}^{\leq i} = \mathbf{Q}_i \mathbf{S}_i, \qquad (\mathbf{U}^{\geq i+1})^T = \mathbf{Q}'_{i+1}\mathbf{S}'_{i+1},$$

where $\mathbf{S}_i, \mathbf{S}'_{i+1} \in \mathbb{R}^{r_i \times r_i}$ have full rank r_i, to obtain

$$\mathbf{A}_i = \mathbf{Q}_i \mathbf{S}_i (\mathbf{S}'_{i+1})^T (\mathbf{Q}'_{i+1})^T \quad =: \quad \mathbf{Q}_i \mathbf{S} (\mathbf{Q}'_{i+1})^T \tag{27}$$

with $\mathbf{Q}_i$, $\mathbf{Q}'_{i+1}$ having orthonormal columns, and $\mathbf{S}$ having full rank r_i. Diagonalization of the right hand side of (27) yields an SVD of A_i; thus

$$s_i = \text{rank } \mathbf{A}_i = r_i.$$

□

3.3 Proof of Theorem 1 (b)

The proof of part (b) uses the following simple lemma.

Lemma 3 *Let* $\mathbf{M}_1, \mathbf{N}_1 \in \mathbb{R}^{p \times r}$, $\mathbf{M}_2, \mathbf{N}_2 \in \mathbb{R}^{r \times q}$ *matrices of rank* r. *If*

$$\mathbf{M}_1\mathbf{M}_2 = \mathbf{N}_1\mathbf{N}_2 \quad \textit{and} \quad \mathbf{M}_1^T\mathbf{M}_1 = \mathbf{N}_1^T\mathbf{N}_1 = \mathbf{I} \in \mathbb{R}^{r\times r}, \tag{28}$$

there is an orthogonal matrix $\mathbf{Q}$ *such that*

$$\mathbf{M}_1 = \mathbf{N}_1\mathbf{Q}, \qquad \mathbf{M}_2 = \mathbf{Q}^T\mathbf{N}_2. \tag{29}$$

Proof There holds $\mathbf{M}_2 = \mathbf{M}_1^T\mathbf{N}_1\mathbf{N}_2$ and therefore

$$\begin{aligned}\mathbf{N}_2 &= \mathbf{N}_1^T\mathbf{N}_1\mathbf{N}_2 = \mathbf{N}_1^T\mathbf{M}_1\mathbf{M}_2 = \mathbf{N}_1^T\mathbf{M}_1\mathbf{M}_1^T\mathbf{N}_1\mathbf{N}_2 \\ &= \mathbf{N}_1^T\mathbf{M}_1(\mathbf{N}_1^T\mathbf{M}_1)^T\mathbf{N}_2.\end{aligned}$$

This implies that $\mathbf{Q} := \mathbf{N}_1^T\mathbf{M}_1 \in \mathbb{R}^{r\times r}$ is an orthogonal matrix because the columns of $\mathbf{N}_2$ span $\mathbb{R}^r$. Using $\mathbf{M}_2 = \mathbf{M}_1^T\mathbf{N}_1\mathbf{N}_2$ again yields $\mathbf{M}_2 = \mathbf{Q}^T\mathbf{N}_2$ and also

$$\mathbf{N}_1\mathbf{N}_2 = \mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_1\mathbf{Q}^T\mathbf{N}_2$$

which implies $\mathbf{M}_1\mathbf{Q}^T = \mathbf{N}_1$ due to the full rank of $\mathbf{N}_2$. □

The assertion of Theorem part (b) now follows by inductively applying the result of the Lemma 3 to the sequence of matrices $(\mathbf{Q}_{i-1}^T \otimes \mathbf{I})\mathbf{L}(\mathbf{U}_i) \in \mathbb{R}^{r_{i-1}\cdot n_i \times r_i}$, $\mathbf{U}^{\geq i+1} \in \mathbb{R}^{r_i \times (n_{i+1}\cdot\ldots\cdot n_d)}$ formed from the component functions $\mathbf{U}_i$ belonging to the left representation of U in (17) (where $\mathbf{Q}_0 = (1)$), and to $\mathbf{V}_i$, $\mathbf{V}^{\geq i+1}$ analogously formed from the right hand representation: We at first note that Lemma 2(a) ensures that these matrices have rank r_i. Further, if $\mathbf{Q}_{i-1}$ is orthogonal, $\mathbf{Q}_{i-1} \otimes \mathbf{I}$ is orthogonal. Thus, the conditions (28) of Lemma 3 are satisfied. If (17) holds, we thus get inductively that

$$\begin{aligned}&\mathbf{L}(\mathbf{U}_1) = \mathbf{L}(\mathbf{V}_1)\mathbf{Q}_1, \quad \mathbf{L}(\mathbf{U}_i) = (\mathbf{Q}_{i-1}^T \otimes \mathbf{I})\mathbf{L}(\mathbf{V}_i)\mathbf{Q}_i\\ &U^{\geq 2} = \mathbf{U}_2(\cdot) \cdot \ldots \cdot \mathbf{U}_d(\cdot) = \big(\mathbf{Q}_1^T \mathbf{V}_2(\cdot)\big) \cdot \mathbf{V}_3(\cdot) \cdot \ldots \cdot \mathbf{V}(\cdot)_d;\\ &U^{\geq i+1} = \mathbf{U}_{i+1}(\cdot) \cdot \ldots \cdot \mathbf{U}_d(\cdot) = \big(\mathbf{Q}_i^T \mathbf{V}_{i+1}(\cdot)\big)\mathbf{V}_{i+2}(\cdot) \cdot \ldots \cdot \mathbf{V}_d(\cdot),\end{aligned}$$

which proves the assertion by the observation that

$$\mathbf{L}(\mathbf{U}_i) = (\mathbf{Q}_{i-1}^T \otimes \mathbf{I})\mathbf{L}(\mathbf{V}_i)\mathbf{Q}_i \iff \begin{array}{l}\mathbf{U}_i(x_i) = \mathbf{Q}_{i-1}^T \mathbf{V}_i(x_i)\mathbf{Q}_i\\ \text{for all } \; x_i \in \{1, \ldots, n_i\}.\end{array}$$

In the final step $i = d-1$, we obtain $\mathbf{U}^{\geq d} = \mathbf{R}(\mathbf{U}_d) = \mathbf{Q}_{d-1}^T \mathbf{R}(\mathbf{V}_d)$, finishing the proof. □

4 The manifold of tensors of fixed TT rank and its tangent space

4.1 The manifold of tensors of fixed TT rank $\underline{r}$

From this point on, we will be concerned with the manifold $\mathbb{T}_{\underline{r}}$ formed by the d-dimensional TT tensors $U \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ of fixed rank $\underline{r} = (r_1, \ldots, r_{d-1})$,

$$\mathbb{T} := \mathbb{T}_{\underline{r}} := \{U \in \mathbb{R}^{n_1 \times \cdots \times n_d} \;\text{ is tensor of TT rank }\; \underline{r}\}. \tag{30}$$

Theorem 1 implies that $\mathbb{R}^{n_1 \times \cdots \times n_d}$ is the disjoint union of the manifolds $\mathbb{T}_{\underline{r}}$, with the possible values of $\underline{r}$ restricted by (19).

Before we mainly turn our attention to the tangent space of $\mathbb{T}$ in the subsequent sections, we make the following observation:

Lemma 4 *Let a rank vector $\underline{r} = (r_1, \ldots, r_{d-1})$ be given, $r_0 = r_d = 1$ as before. For fixed $d, n_1, \ldots, n_d \in \mathbb{N}$, $\mathbb{T}_{\underline{r}}$ is a manifold of dimension*

$$dim\ \mathbb{T} = \sum_{i=1}^{d} r_{i-1} n_i r_i - \sum_{i=1}^{d-1} r_i^2. \tag{31}$$

Proof The assertion follows from Theorem 1: Let us define

$$f : \times_{i=1}^{d} C_i \to \mathbb{R}^{n_1 \times \cdots \times n_d},$$

given by

$$\big(x_1 \mapsto \mathbf{U}_1(x_1), \ldots, x_d \mapsto \mathbf{U}_d(x_d)\big) \;\; \mapsto \Big((x_1, \ldots, x_d) \mapsto \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_d(x_d)\Big),$$

and an equivalence relation $\sim$ on $\times_{i=1}^{d} C_i$ by

$$\begin{aligned} & (\mathbf{U}_1(\cdot), \ldots, \mathbf{U}_d(\cdot)) \;\sim\; (\mathbf{V}_1(\cdot), \ldots, \mathbf{V}_d(\cdot)) \\ \Longleftrightarrow \; & (18) \quad \text{holds for some orthogonal } \; \mathbf{Q}_i \in \mathbb{R}^{r_i \times r_i}, i = 1, \ldots, d. \end{aligned}$$

Note that f is also well defined on the factorized space $\mathcal{V} := \times_{i=1}^{d} C_i / \sim$ (by application to arbitrary representants). Let us define an auxiliary manifold $\mathcal{M} \subseteq \mathcal{V}$ by the constraint conditions

$$g_i(\mathbf{U}_1, \ldots, \mathbf{U}_d) \; := \; [\mathbf{U}_i, \mathbf{U}_i] \; = \; \mathbf{I} \in \mathbb{R}^{r_i \times r_i}, \;\; i = 1, \ldots, d-1. \tag{32}$$

The set $\mathrm{GL}(r_i \times r_i)$ of invertible matrices $A \in \mathbb{R}^{r_i \times r_i}$ is open and the mappings $\mathbf{U}_i \mapsto [\mathbf{U}_i, \mathbf{U}_i]$, $\mathbf{U}_i \mapsto \mathbf{R}(\mathbf{U}_i)\mathbf{R}(\mathbf{U}_i)^T$ are continuous; thus, for any $(\mathbf{U}_1, \ldots, \mathbf{U}_d)$ for which the $\mathbf{U}_i$ have full left and right rank, there is a neighbourhood N_δ such that for $(\mathbf{W}_1, \ldots, \mathbf{W}_d) \in N_\delta$, each $\mathbf{W}_i$ still has full left and right rank. For $D := \mathcal{M} \cap N_\delta$, the restriction $f|_D : D \to f(D)$ therefore maps to $\mathbb{T}$. Obviously, $f|_D$ is continuous, and bijective by Theorem 1. Thus, a further restriction of $f|_D$ to a compact subset K of D with nonempty interior possesses a continuous inverse. Yet another restriction $f|_L$ to an open subset $L \subseteq K$ containing U gives a local chart for $\mathbb{T}$ containing U, and a suitable collection of such restrictions for all $U \in \mathbb{T}$ thus constitutes an $\mathcal{M}$-atlas [29] of $\mathbb{T}$. Thus, the dimensions of $\mathbb{T}$ and $\mathcal{M}$ coincide. To determine the latter, we first note that orthogonal $\mathbb{R}^{r_i \times r_i}$-matrices are determined by $r_i(r_i - 1)/2$ degrees of freedom, so that the factorization with respect to $\sim$ gives that for $\mathcal{V}$,

$$\dim \mathcal{V} = \sum_{i=1}^{d} r_{i-1} n_i r_i - \sum_{i=1}^{d-1} \frac{r_i(r_i - 1)}{2}.$$

In analogy to the constraint conditions imposed on the Stiefel manifold (see e.g. [17]), it is not hard to see that the constraint (32) yields $r_i(r_i + 1)/2$ independent constraint conditions in each component, so that the Jacobian of accordingly constructed constraint function g has full rank $\sum_{i=1}^{d-1} r_i(r_i + 1)/2$. This implies (see also [17], IV.5.1) that

$$\dim \mathbb{T} = \dim \mathcal{M} = \dim \mathcal{V} \; - \sum_{i=1}^{d-1} \frac{r_i(r_i + 1)}{2} = \sum_{i=1}^{d} r_{i-1} n_i r_i - \sum_{i=1}^{d-1} r_i^2,$$

completing the proof. □

4.2 Representations for the tangent space of $\mathbb{T}$: main result

In approximative algorithmic treatment of high dimensional problems by means of tensor approximation, it is often the general ansatz to fix a certain tensor rank and thus a corresponding manifold $\mathcal{M}$, i.e. $\mathbb{T}_{\underline{r}}$ in our case, and then to compute a—in some sense—best approximation of the solution of the problem on the given manifold $\mathcal{M}$. In many cases, knowledge of a (non-redundant) representation of the tangent space $\mathcal{T}_U\mathcal{M}$ of this manifold is needed for the design of according algorithms, see e.g. [17,31,40] and also Sect. 6 for examples.In the remainder of this section, we shall prove the below Theorem 2, which gives a unique representation of the tangent space $\mathcal{T}_U\mathbb{T}$ taken at U.

We will proceed as in [17,31], introducing gauge conditions to obtain uniqueness of the representation. Although these conditions are similar to the ones used for the matrix case treated in [24] and for the Tucker format [25], the proof of existence and uniqueness in the present situation will be a little more subtle due to the more complicated structure of the TT tensors.

Theorem 2 *Let $U \in \mathbb{T}$, $\delta U \in \mathcal{T}_U\mathbb{T}$, and let $(\mathbf{G}_i)_{i=1}^{d-1}$ a gauge sequence, i.e. a sequence of symmetric positive definite matrices $\mathbf{G}_i \in \mathbb{R}^{(r_{i-1}n_i)\times(r_{i-1}n_i)}$, $i = 1,\ldots,d-1$.*

There are unique component functions $\mathbf{W}_i(\cdot) \in C_i$, $i \in \{1,\ldots,d\}$, such that the tensors δU_i, given pointwise by

$$\delta U_i(\underline{x}) := \mathbf{U}_1(x_1)\ldots\mathbf{U}_{i-1}(x_{i-1})\mathbf{W}_i(x_i)\mathbf{U}_{i+1}(x_{i+1})\ldots\mathbf{U}_d(x_d), \tag{33}$$

fulfil both

$$\delta U = \delta U_1 + \cdots + \delta U_d \tag{34}$$

and the gauge conditions

$$[\mathbf{U}_i, \mathbf{W}_i]_{\mathbf{G}_i} = \mathbf{0} \in \mathbb{R}^{r_i\times r_i}, \tag{35}$$

for all $i = 1,\ldots,d-1$, i.e. the column vectors of the left unfoldings of $\mathbf{U}_i$ and $\mathbf{W}_i$ are orthogonal in the inner products induced by the gauging matrices $\mathbf{G}_i$. These unique component functions $(\mathbf{W}_1,\ldots,\mathbf{W}_d)$ will be called the $(\mathbf{G}_i)$-gauged representation of $\delta U \in \mathcal{T}_U\mathbb{T}$.

Remark 2 Note that in particular, by choosing $\mathbf{G}_i = \mathbf{I} \in \mathbb{R}^{(r_{i-1}n_i)\times(r_{i-1}n_i)}$ for all $i = 1,\ldots,d-1$, elements of the tangent space have a unique $(\mathbf{I})$-gauged representation given by component functions $\mathbf{W}_i : x_i \mapsto \mathbf{W}_i(x_i)$ for which

$$[\mathbf{U}_i, \mathbf{W}_i] := \mathbf{L}(\mathbf{U}_i)^T\mathbf{L}(\mathbf{W}_i) = \mathbf{0} \in \mathbb{R}^{r_i\times r_i}, \tag{36}$$

i.e. the column vectors of $\mathbf{L}(\mathbf{U}_i)$ and $\mathbf{L}(\mathbf{W}_i)$ are orthogonal with respect to the standard inner product on $\mathbb{R}^{r_{i-1}n_i}$. □

To begin the proof of Theorem 2, let $U \in \mathbb{T}$ and a gauge sequence $(\mathbf{G}_i)_{i=1}^{d-1}$ be given. We remind the reader that any $\delta U \in \mathcal{T}_U\mathbb{T}$ can be represented as the derivative of a continuously differentiable curve γ on $\mathbb{T}$, i.e.

$$\begin{aligned}\mathcal{T}_U\mathbb{T} = \Big\{&\gamma'(t)|_{t=0} \mid \gamma \in C^1(]-\delta,\delta[,\mathbb{T}),\\ &\gamma(t) = \mathbf{U}_1(\cdot,t)\cdot\ldots\cdot\mathbf{U}_d(\cdot,t),\quad \gamma(0) = U(\underline{x})\Big\}\end{aligned}$$

up to isomorphisms. For technical reasons, we will at first work with the set

$$\hat{\mathcal{T}}_U\mathbb{T} := \Big\{ \gamma'(t)|_{t=0} \in \mathcal{T}_U\mathbb{T} \;\; t \mapsto \mathbf{U}_i(\cdot, t) \;\text{ is }\; C^1(]-\delta, \delta[, C_i) \\ \text{for all } \; i = 1, \ldots, d \Big\} \subseteq \mathcal{T}_U\mathbb{T}$$

(with C_i the spaces of component functions from (6)) and prove Theorem 2 for all $\delta U \in \hat{\mathcal{T}}_U\mathbb{T}$. At the end of this section, a dimensional argument will prove $\hat{\mathcal{T}}_U\mathbb{T} = \mathcal{T}_U\mathbb{T}$ (in contrast to the fact that there are C^1-curves γ the components of which are not all C^1)—Theorem 2 thus holds for all $\delta U \in \mathcal{T}_U\mathbb{T}$ as asserted.

In Sect. 4.3, we prove the existence of a $(\mathbf{G}_i)$-gauged representation of $\delta U \in \hat{\mathcal{T}}_U\mathbb{T}$; uniqueness will be proven in Sect. 4.4. Finally, the equality of $\hat{\mathcal{T}}_U\mathbb{T}$ and $\mathcal{T}_U\mathbb{T}$ is subject to Sect. 4.5.

4.3 Proof of existence of a $(\mathbf{G}_i)$-gauged representation

Let $\delta U \in \hat{\mathcal{T}}_U\mathbb{T}$ be given. There holds for $\underline{x} = (x_1, \ldots, x_d)$,

$$\begin{aligned} \delta U(\underline{x}) \;\simeq\; (\gamma'(t)|_{t=0})(\underline{x}) \;&= \mathbf{U}_1'(x_1, 0)\mathbf{U}_2(x_2) \cdot \ldots \cdot \mathbf{U}_d(x_d) \\ &+ \mathbf{U}_1(x_1)\mathbf{U}_2'(x_2, 0) \cdot \ldots \cdot \mathbf{U}_d(x_d) \\ &+ \cdots + \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_{d-1}(x_{d-1})\mathbf{U}_d'(x_d, 0). \end{aligned} \tag{37}$$

This yields a representation of the form (34) for δU; alas, the gauge condition (35) does not need to be satisfied. Therefore, we now utilize the following basic lemma to transform the component functions $\mathbf{U}_1'(\cdot, 0), \ldots, \mathbf{U}_d'(\cdot, 0)$ to a $(\mathbf{G}_i)$-gauged representation of δU.

Lemma 5 *For $i \in \{1, \ldots, d\}$, let $\mathbf{M}, \mathbf{U} : \mathcal{I}_i \to \mathbb{R}^{r_{i-1} \times r_i}$ component functions, and let $\mathbf{G} \in \mathbb{R}^{(r_{i-1}n_i) \times (r_{i-1}n_i)}$ a symmetric positive definite matrix. Then there exists a unique component function $\mathbf{W} \in C_i$, $\mathbf{W} : \mathcal{I}_i \to \mathbb{R}^{r_{i-1} \times r_i}$ and a matrix $\Lambda \in \mathbb{R}^{r_i \times r_i}$ such that*

$$\mathbf{M}(x_i) = \mathbf{U}(x_i)\Lambda + \mathbf{W}(x_i) \tag{38}$$

and

$$[\mathbf{U}, \mathbf{W}]_{\mathbf{G}} = \mathbf{0} \;\in\; \mathbb{R}^{r_i \times r_i}, \tag{39}$$

i.e. the column vectors of the left unfoldings of $\mathbf{U} : x_i \mapsto \mathbf{U}(x_i)$ and $\mathbf{W} : x_i \mapsto \mathbf{W}(x_i)$ are mutually orthogonal in the inner product induced by $\mathbf{G}$. If $\mathbf{U}$ has full left rank, Λ is also unique.

Proof Let $\mathbf{m}_1, \ldots, \mathbf{m}_r \in \mathbb{R}^m$ denote the columns of the left unfolding $\mathbf{L}(\mathbf{M})$ and $\mathbf{l}_1, \ldots, \mathbf{l}_r \in \mathbb{R}^m$ denote the columns of $\mathbf{L}(\mathbf{U})$. For each $i \in \{1, \ldots, r\}$, we express $\mathbf{m}_i$ as

$$\mathbf{m}_i = \sum_{j=1}^{r} \lambda_{i,j} \mathbf{l}_j + \mathbf{w}_i,$$

with suitably chosen coefficients $\lambda_{i,j} \in \mathbb{R}$ and unique $\mathbf{w}_i \in \mathbb{R}^m$ from the $\mathbf{G}$-orthogonal complement of $\text{span}\{\mathbf{l}_1, \ldots, \mathbf{l}_r\}$. Letting $\Lambda = (\lambda_{j,i})_{i,j=1}^r$ and $\mathbf{L}(\mathbf{W}) = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$ yields

$$\mathbf{L}(\mathbf{M}) = \mathbf{L}(\mathbf{U})\Lambda + \mathbf{L}(\mathbf{W})$$

and thus, by applying the inverse of the left unfolding mapping $\mathbf{L}(\cdot)$, the representation (38). Finally, we note that if $\mathbf{U}$ has full left rank, the coefficients $\lambda_{i,j}$ also are unique. $\square$

We can now continue the proof of Theorem 2. We apply Lemma 5 to $\mathbf{G} = \mathbf{G}_1$, $x_i \mapsto \mathbf{M}(x_i) := U_1'(x_1, 0)$, $x_i \mapsto \mathbf{U}(x_1) := \mathbf{U}_1(x_1)$, and obtain that for suitable $\Lambda_1 \in \mathbb{R}^{r_1 \times r_1}$ and a component function $\mathbf{W}_1 \in \mathbb{R}^{n_1 \times r_1}$ for which $\mathbf{W}_1$ is left-orthogonal to $\mathbf{U}_1$ in the $\mathbf{G}_1$-inner product (see (39)), the relation

$$\mathbf{M}_1(x_1) = \mathbf{U}_1(x_1)\Lambda_1 + \mathbf{W}_1(x_1)$$

holds; thus, for $\underline{x} = (x_1, \ldots x_d)$,

$$\begin{aligned} \delta U(\underline{x}) = & \underbrace{\mathbf{W}_1(x_1)\mathbf{U}_2(x_2) \cdot \ldots \cdot \mathbf{U}_d(x_d)}_{=: \, \delta U_1(\underline{x})} \\ & + \mathbf{U}_1(x_1)\big(\Lambda_1 \mathbf{U}_2(x_2) + \mathbf{U}_2'(x_2, 0)\big) \cdot \ldots \cdot \mathbf{U}_d(x_d) \\ & + \cdots + \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_{d-1}(x_{d-1})\mathbf{U}_d'(x_d, 0). \end{aligned}$$

Now, we successively apply Lemma 5 to

$$i = 2, \ldots, d, \quad \mathbf{G} = \mathbf{G}_i, \quad \mathbf{M}_i(x_i) = \Lambda_{i-1}\mathbf{U}_i(x_i) + \mathbf{U}_i'(x_i, 0), \quad \mathbf{U}(x_i) = \mathbf{U}_i(x_i).$$

We obtain

$$\begin{aligned} \delta U(\underline{x}) = & \underbrace{\mathbf{W}_1(x_1)\mathbf{U}_2(x_2) \cdot \ldots \cdot \mathbf{U}_d(x_d)}_{=: \, \delta U_1(\underline{x})} + \underbrace{\mathbf{U}_1(x_1)\mathbf{W}_2(x_2)\mathbf{U}_3(x_3) \cdot \ldots \cdot \mathbf{U}_d(x_d)}_{=: \, \delta U_2(\underline{x})} \\ & + \cdots + \underbrace{\mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_{d-1}(x_{d-1})(\Lambda_{d-1}\mathbf{U}_d(x_d) + \mathbf{U}_d'(x_d, 0))}_{=: \, \delta U_d(\underline{x})} \end{aligned}$$

for suitable $\mathbf{W}_2, \ldots, \mathbf{W}_{d-1}$ and Λ_{d-1}, where (35) is fulfilled for all $i = 1, \ldots, d-1$. Letting $\mathbf{W}_d(x_d) = \Lambda_{d-1}\mathbf{U}_d(x_d) + \mathbf{U}_d'(x_d, 0)$ completes the proof for the existence of a $(\mathbf{G}_i)$-gauged representation for any $\delta U \in \hat{\mathcal{T}}_U \mathbb{T}$ and any gauge sequence $(\mathbf{G}_i)_{i=1}^{d-1}$. $\square$

4.4 Proof of uniqueness of a $(\mathbf{G}_i)$-gauged representation

To prove uniqueness of a $(\mathbf{G}_i)$-gauged representation for $\delta U \in \hat{\mathcal{T}}_U\mathbb{T}$, we will use the following lemma, which shows that it suffices to show uniqueness of representation with respect to one gauging sequence.

Lemma 6 *Let $\delta U \in \hat{\mathcal{T}}_U\mathbb{T}$, $(\mathbf{G}_i)_{i=1}^{d-1}$, $(\mathbf{H}_i)_{i=1}^{d-1}$ be gauge sequences, and $(\mathbf{W}_1, \ldots, \mathbf{W}_d)$ and $(\mathbf{V}_1, \ldots, \mathbf{V}_d)$ be $(\mathbf{G}_i)$- and $(\mathbf{H}_i)$-gauged representations of δU, respectively. Then, if $(\mathbf{W}_1, \ldots, \mathbf{W}_d)$ is a unique $(\mathbf{G}_i)$-gauged representation, $(\mathbf{V}_1, \ldots, \mathbf{V}_d)$ also is a unique $(\mathbf{H}_i)$-gauged representation.*

Proof Suppose that there exist two distinct $(\mathbf{H}_i)$-gauged representations

$$(\mathbf{V}_1, \ldots, \mathbf{V}_d), \qquad (\tilde{\mathbf{V}}_1, \ldots, \tilde{\mathbf{V}}_d)$$

for δU. Analogously to the proceeding in the above existence proof, we use Lemma 5 to obtain two different $(\mathbf{G}_i)$-gauged representation for δU: We again write δU recursively as

$$\delta U(\underline{x}) = \sum_{j=1}^{i-1} \delta U_j(\underline{x}) + \mathbf{U}_1(x_1) \cdot \ldots \cdot (\Lambda_{i-1}\mathbf{U}_i(x_i) + \mathbf{V}_i(x_i)) \cdot \ldots \cdot \mathbf{U}_d(x_d)$$
$$+ \sum_{j=i+1}^{d} \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{V}_j(x_i) \cdot \ldots \cdot \mathbf{U}_d(x_d)$$

for $i = 1, \ldots, d$, where the summands δU_j are gauged with respect to the matrices $\mathbf{G}_i$, and as a corresponding expression for the second gauge sequence, with suitable $(\mathbf{G}_i)$-gauged $\delta\tilde{U}_j$, $j < i$ and $\tilde{\Lambda}_{i-1}$, and with $\tilde{\mathbf{V}}_j$ in place of $\mathbf{V}_j$ for $j \geq i$. For $i \in \{1, \ldots, d\}$ chosen minimal such that $\mathbf{V}_i \neq \tilde{\mathbf{V}}_i$, we have $\delta U_j = \delta\tilde{U}_j$ for $j < i$ and also $\Lambda_{i-1} = \tilde{\Lambda}_{i-1}$ due to the uniqueness of the expressions yielded by application of Lemma 5 (Note that $\mathbf{U}_i$ has full left rank by definition of $\mathbb{T}_{\underline{r}}$). Thus there holds

$$\mathbf{M}_i(x_i) := \Lambda_{i-1}\mathbf{U}_i(x_i) + \mathbf{V}_i(x_i) \neq \Lambda_{i-1}\mathbf{U}_i(x_i) + \tilde{\mathbf{V}}_i(x_i) =: \tilde{\mathbf{M}}_i(x_i).$$

Applying Lemma 5 to $\mathbf{M}_i$, $\tilde{\mathbf{M}}_i$ and the gauge matrix $\mathbf{G}_i$ gives left-orthogonal decompositions in the $\mathbf{G}_i$-product

$$\mathbf{M}_i = \mathbf{U}_i\Lambda_i + \mathbf{W}_i \neq \mathbf{U}_i\tilde{\Lambda}_i + \tilde{\mathbf{W}}_i = \tilde{\mathbf{M}}_i,$$

for which $\Lambda_i \neq \tilde{\Lambda}_i$ or $\mathbf{W}_i \neq \tilde{\mathbf{W}}_i$ due to left-orthogonality of the summands with respect to the $\mathbf{G}_i$-product. An easy inductive argument now shows that if we proceed as in the above existence proof to obtain $(\mathbf{G}_i)$-gauged representations, this implies $\delta U_j \neq \delta\tilde{U}_j$ for some $j \geq i$, finishing the proof. □

To complete the proof of Theorem 2, we note that the matrices

$$\mathbf{G}_i := [\![U, U]\!]^{\leq i-1} \otimes \mathbf{I}_{n_i \times n_i} \in \mathbb{R}^{(r_{i-1}n_i)\times(r_{i-1}n_i)}, \tag{40}$$

defined in (25) form a gauge sequence by Lemma 2; we will now show that for this particular gauge sequence, the $(\mathbf{G}_i)$-gauged representation of δU is unique. As a final preparation for the proof thereof, we give a decomposition for $\langle V, W\rangle$ in the following remark.

Remark 3 Let $V, W \in \mathbb{T}$ be tensors, and let $\langle \cdot, \cdot\cdot \rangle$ denote the usual Euclidean inner product on $\mathbb{R}^{n_1\times\cdots\times n_d}$. By splitting up $[\![V, W]\!]^{\leq d}$ into an i-th left and $i+1$-th right part, it is not hard to see (e.g. by writing out sums or by usage of the diagrams for $[\![V, W]\!]^{\leq i}$, $[\![V, W]\!]^{\geq i+1}$) that

$$\begin{aligned}\langle V, W\rangle = [\![V, W]\!]^{\leq d} &= \langle \underbrace{[\![V, W]\!]^{\leq i}}_{\in\mathbb{R}^{r_i\times r_i}}, \underbrace{[\![V, W]\!]^{\geq i+1}}_{\in\mathbb{R}^{r_i\times r_i}} \rangle \\ &= \Big\langle (\mathbf{L}(\mathbf{W}_i)^T ([\![V, W]\!]^{(\leq i-1)} \otimes \mathbf{I}_{n_i\times n_i})\mathbf{L}(\mathbf{V}_i),\ [\![V, W]\!]^{\geq i+1} \Big\rangle,\end{aligned}$$

in which the last line follows from (24). In particular, if for some $j \in \{1, \ldots, d\}$, V and W coincide in the first $j-1$ components, i.e.

$$\begin{aligned}V(\underline{x}) &= \mathbf{U}_1(x_1)\ldots\mathbf{U}_{j-1}(x_{j-1})\mathbf{V}_j(x_j)\ldots\mathbf{V}_d(x_d),\\ W(\underline{x}) &= \mathbf{U}_1(x_1)\ldots\mathbf{U}_{j-1}(x_{j-1})\mathbf{W}_j(x_j)\ldots\mathbf{W}_d(x_d)\end{aligned}$$

for some component functions $\mathbf{U}_i(x_i)$, then for all $i \leq j$

$$\begin{aligned}\langle V, W\rangle &= \Big\langle (\mathbf{L}(\mathbf{W}_i)^T\mathbf{G}_i\mathbf{L}(\mathbf{V}_i),\ [\![V, W]\!]^{\geq i+1} \Big\rangle\\ &= \Big\langle [\mathbf{V}_i\mathbf{W}_i]_{\mathbf{G}_i}, [\![V, W]\!]^{\geq i+1} \Big\rangle\end{aligned} \tag{41}$$

with $\mathbf{G}_i$ defined as by (40).

We are now in the position to prove uniqueness of the $(\mathbf{G}_i)$-gauged representation

$$\begin{aligned}\delta U &= \sum_{i=1}^{d} \delta U_i,\\ \delta U_i(\underline{x}) &= \mathbf{U}_1(x_1)\cdot\ldots\cdot\mathbf{U}_{i-1}(x_{i-1})\mathbf{W}_i(x_i)\mathbf{U}_{i+1}(x_{i+1})\cdot\ldots\cdot\mathbf{U}_d(x_d)\end{aligned} \tag{42}$$

for $\delta U \in \hat{\mathcal{T}}_U\mathbb{T}$, where the component functions $\mathbf{W}_i$ fulfil the gauge condition (35) with the gauge sequence $(\mathbf{G}_i)$ defined via (40) by the tensor U. To do so, we have to show that for $i = 1\ldots, d$, the component functions $\mathbf{W}_i$ are uniquely determined by δU. We start by noting that $\mathbf{W}_i$ is uniquely determined iff $\mathbf{L}(\mathbf{W}_i)$ is unique, which in turn by Lemma 2 is the case iff the matrix

$$\mathbf{G}_i\mathbf{L}(\mathbf{W}_i)\mathbf{P}_{i+1} \in \mathbb{R}^{(r_{i-1}\cdot n_i)\times r_i}$$

is uniquely determined (where we let $\mathbf{G}_1 = \mathbf{I}_{n_1 \times n_1}$ and $\mathbf{P}_{d+1} = (1) \in \mathbb{R}^{1\times 1}$ for convenience). Cast into a weak formulation, this is the case if and only if for any component function $\mathbf{V}_i : \mathbb{R}^{n_i} \to \mathbb{R}^{(r_{i-1}\times r_i)}$, the Euclidean inner product

$$\langle \mathbf{G}_i \mathbf{L}(\mathbf{W}_i)\mathbf{P}_{i+1}, \mathbf{L}(\mathbf{V}_i)\rangle = \langle \mathbf{L}(\mathbf{V}_i)^T \mathbf{G}_i \mathbf{L}(\mathbf{W}_i), \mathbf{P}_{i+1}\rangle, \tag{43}$$

taken on $\mathbb{R}^{r_{i-1}n_i r_{i-1}}$) is uniquely determined by δU. If we define for any component function $\mathbf{V}_i$ a corresponding tensor

$$V_i(\underline{x}) := \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_{i-1}(x_{i-1})\mathbf{V}_i(x_i)\mathbf{U}_{i+1}(x_{i+1}) \ldots \cdot \mathbf{U}_d(x_d), \tag{44}$$

we can use (41), (42) and (44) to rewrite (43) as

$$\langle (\mathbf{L}(\mathbf{V}_i))^T \mathbf{G}_i \mathbf{L}(\mathbf{W}_i), \mathbf{P}_{i+1}\rangle = \langle \delta U_i, V_i\rangle, \qquad \text{for all } \mathbf{V}_i \in C_i \tag{45}$$

and we now use the gauge condition to show that for fixed test component function $\mathbf{V}_i$, $\langle \delta U_i, V_i\rangle$ is indeed uniquely determined by δU. To this end, we observe that for all $1 \le j < i \le d$,

$$\langle \delta U_j, V_i\rangle = \langle\ (\mathbf{L}(\mathbf{W}_j))^T \mathbf{G}_j \mathbf{L}(\mathbf{U}_j), [\![\delta U_j, V_i]\!]^{\le j+1}\ \rangle = 0 \qquad \text{for all } \mathbf{V}_i \in C_i$$

by (41) and the gauge condition (35), and that therefore

$$\langle \delta U, V_i\rangle = \sum_{j=i}^{d} \langle \delta U_j, V_i\rangle, \qquad \text{for all } i = 1, \ldots, d. \tag{46}$$

Starting with $i = d$, we obtain the uniquely solvable linear equation

$$\langle (\mathbf{L}(\mathbf{V}_d))^T \mathbf{G}_d \mathbf{L}(\mathbf{W}_d), \mathbf{P}_{d+1}\rangle = \langle \delta U, V_d\rangle$$

for $\mathbf{L}(\mathbf{W}_d)$, with the right hand side fixed by δU. We now proceed recursively: Once $\mathbf{L}(\mathbf{W}_i)$ and thus $\mathbf{W}_i$, $i = d, d-1, \ldots,$ is computed, the values $\langle \delta U_i, V_i\rangle$ are computable. $\mathbf{W}_{i-1}$ is then fixed by (45) with

$$\langle \delta U_i, V_i\rangle = \langle \delta U, V_i\rangle - \sum_{j=i+1}^{d} \langle \delta U_j, V_i\rangle, \quad \text{for all } \ i = d, \ldots, 1, \tag{47}$$

uniquely determining the right hand side, so that $\mathbf{L}(\mathbf{W}_{i-1})$ and thus $\mathbf{W}_{i-1}$ are unique as solution of (45).

Remark 4 We note that the preceding proof mainly relies on the fact that the full-rank-condition imposed on the TT decomposition of U implies that the matrices $\mathbf{G}_i$, $\mathbf{P}_{i+1}$ are symmetric and positive definite by Lemma 2, thus allowing for a unique solution of the weak equations yielded by (46). Those equations can be restated as

$$\langle (\mathbf{G}_i \otimes \mathbf{P}_{i+1})\, \mathrm{vec}(\mathbf{W}_i), \mathrm{vec}(\mathbf{V}_i)\rangle = \langle \delta U_i, V_i\rangle, \tag{48}$$

in which $\mathrm{vec}(\mathbf{U}_i)$ denotes the vectorisation of a component function $\mathbf{U}_i$. The matrices $\mathbf{G}_i, \mathbf{P}_{i+1}$, often termed density matrices in the context of quantum physics and DMRG calculations, thus play a role similar to that of the analogous density matrices in [25].

In practical TT computations, the treatment of problems as computation of best approximations or solution of linear equations often boils down to solution of equations similar to (48); in particular, when computing the i-th component of a best approximation, one can choose the component functions $\mathbf{U}_j$ left-orthogonal for $j < i$ and right-orthogonal for $j > i$, so that the density matrices are given by the identity. See the forthcoming publication [19] for further details.

4.5 A parametrization of the tangent space $\mathcal{T}_U\mathbb{T}$

Finally, we now show $\hat{\mathcal{T}}_U\mathbb{T} = \mathcal{T}_U\mathbb{T}$, by which Theorem 2 is then proven. To this end, observe at first that $\dim\mathcal{T}_U\mathbb{T} = \dim\mathbb{T}$ is given by (31). We show that $\hat{\mathcal{T}}_U\mathbb{T} \subseteq \mathcal{T}_U\mathbb{T}$ possesses the same dimension, thus proving $\hat{\mathcal{T}}_U\mathbb{T} = \mathcal{T}_U\mathbb{T}$. The more general statement of the below Lemma 7 will also be useful in the next section. At first, we define within the spaces C_i of component functions the (linear) left-orthogonal spaces of $\mathbf{U}_i$,

$$U_i^\ell := \{\mathbf{W}_i \in C_i,\ \ [\mathbf{U}_i, \mathbf{W}_i]_{\mathbf{G}_i} = \mathbf{0}\}.$$

for $1 \leq i \leq d-1$, and

$$X := U_1^\ell \times \cdots \times U_{d-1}^\ell \times C_d. \tag{49}$$

Obviously, due to the r_i orthogonality constraints imposed on each of the column vectors of $\mathbf{L}(\mathbf{W}_i)$ for $i \in \{1, \ldots, d-1\}$

$$\dim U_i^\ell = r_{i-1}n_i r_i - r_i^2; \quad \text{and} \quad \dim C_d = r_{d-1}n_d,$$
$$\dim X = \sum_{i=1}^{d} r_{i-1}n_i r_i - \sum_{i=1}^{d-1} r_i^2.$$

The following lemma is now an immediate corollary of the existence and uniqueness results proven in Sects. 4.3 and 4.4.

Lemma 7 *The mapping*

$$\tau : X \to \hat{\mathcal{T}}_U\mathbb{T},\ (\mathbf{W}_1, \ldots, \mathbf{W}_d) \mapsto \delta U \tag{50}$$

where

$$\delta U(\underline{x}) = \sum_{i=1}^{d} \mathbf{U}_1(x_1) \cdot \ldots \cdot \mathbf{U}_{i-1}(x_{i-1})\mathbf{W}_i(x_i)\mathbf{U}_{i+1}(x_{i+1}) \cdot \ldots \cdot \mathbf{U}_d(x_d),$$

is a linear bijection between X and the tangent space $\mathcal{T}_U\mathbb{T}$, taken at U. In particular,

$$dim\ \hat{\mathcal{T}}_U\mathbb{T} = dim\ \mathcal{T}_U\mathbb{T}, \qquad \hat{\mathcal{T}}_U\mathbb{T} = \mathcal{T}_U\mathbb{T}. \tag{51}$$

5 A local parametrization for the manifold $\mathbb{T}$

In this section, we use the just proven representation

$$(\mathbf{W}_1, \ldots, \mathbf{W}_d) \in X, \quad \text{i.e.} \quad [\mathbf{U}_i, \mathbf{W}_i]_{\mathbf{G}_i} = \mathbf{0}$$

for $\mathcal{T}_U\mathbb{T}$ and the statement given in Lemma 7 to set up a local parametrization of $\mathbb{T}$ and to define local charts for $\mathbb{T}$. The results are collected in Theorem 3. As before, we fix $U \in \mathbb{T}$ and a gauging sequence $(\mathbf{G}_i)$.

Lemma 8 *Let X be defined as in (49), and let*

$$\Psi : X \to \mathbb{R}^{n_1\times\cdots\times n_d}, \ (\mathbf{W}_1, \ldots, \mathbf{W}_d) \ \mapsto \ (\mathbf{U}_1 + \mathbf{W}_1) \cdot \ldots \cdot (\mathbf{U}_d + \mathbf{W}_d), \tag{52}$$

where the matrix product representation is understood pointwise as above. There is an open neighbourhood $N_\delta = N_\delta(0)$ of $0 \in X$ such that $\Psi(N_\delta)$ is an open subset of $\mathbb{T}$, and such that the restriction

$$\Psi|_{N_\delta} : N_\delta \mapsto \Psi(N_\delta) \tag{53}$$

is a diffeomorphism.

Proof We utilize the inverse mapping theorem for manifolds, cf. e.g. [30], Theorem 2.25. We note at first that with the same arguments as in the proof of Lemma 4, restriction of Ψ to a suitable open neighbourhood $\tilde{N}$ of 0 gives a mapping $\Psi : \tilde{N} \to \Psi(\tilde{N}) \subseteq \mathbb{T}$. A straightforward computation shows that the tangent map $T_U\Psi$ belonging to Ψ (also see e.g. [30]) is given by the bijection (50). Thus the conditions of the inverse mapping theorem for manifolds are fulfilled, yielding the asserted statement. □

Because of Lemma 7, any set $\mathcal{C} = (\mathcal{C}_1, \ldots, \mathcal{C}_d)$ of coordinate mappings $\mathcal{C}_i : U_i^\ell \to \mathbb{R}^{r_{i-1}n_ir_i - r_i^2}$ for $i = 1, \ldots, d-1$, $\mathcal{C}_d : C_d \to \mathbb{R}^{r_{d-1}n_d}$, defines an isomorphism between X and $\mathbb{R}^D$,

$$D := \sum_{i=1}^{d} r_{i-1}n_ir_i - \sum_{i=1}^{d-1} r_i^2.$$

We can therefore combine (52) with the coordinate mappings to obtain a local parametrization of $\mathbb{T}$ by letting $\mathcal{C} = (\mathcal{C}_1, \ldots, \mathcal{C}_d)$, $\psi := \Psi \circ \mathcal{C}^{-1} : \mathbb{R}^D \to \mathbb{T}$. The properties of this parametrization and some more implications are collected in the next theorem, finishing this section.

Theorem 3 *For all $U \in \mathbb{T}$, there exists an open neighbourhood $N_U \subset \mathbb{R}^{n_1 \times \cdots \times n_d}$ of U, an open neighbourhood $N_\delta(0)$ of $0 \in \mathbb{R}^D$ and differentiable functions*

$$\psi = \psi_U : \quad N_\delta(0) \to \mathbb{R}^{n_1 \times \cdots , \times n_d}, \qquad g = g_U : N_U \to \mathbb{R}^c, \quad c := \sum_{i=1}^{d-1} r_i^2$$

such that

$$N_U \cap \mathbb{T} = \psi(N_\delta(0)) = \{U \in \mathbb{R}^{n_1 \times \cdots , \times n_d} \ : \ g(U) = 0\} \tag{54}$$

and the above parametrisation ψ is an embedding (i.e. an immersion that is a homeomorphism onto its image), that is, $N_U \cap \mathbb{T}$ is a regular submanifold of $\mathbb{R}^{n_1 \times \cdots \times n_d}$.

Proof The local parametrization ψ was already constructed above. The proof of Lemma 8 shows that the tangent mapping of ψ is an injection, making ψ an immersion. Lemma 8 also states that ψ is a homeomorphism onto its image. This also implies the existence of a local constraint function g characterizing $\mathbb{T}$ on a neighbourhood of U, see e.g. [17], IV.5.1. The last statement follows from the fact that ψ is an embedding, cf. e.g. [30], Theorem 3.5. □

6 Examples of problems posed on the manifold of TT tensors

To illustrate the range of applications, we now review a well-known ansatz [17,24,31] for how partial differential equations and optimisation problems posed in high-dimensional spaces $\mathbb{R}^{n_1 \times \cdots \times n_d}$ may be solved approximately on a given approximation manifold, in our case on the manifold $\mathbb{T}$ of TT tensors.

Let a differential equation

$$\frac{du}{dt} = f(u) \tag{55}$$

for a high dimensional function $u : t \mapsto u(t) \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ be given, e.g. stemming from the Galerkin discretization of a time-dependent PDE formulated on a high-dimensional function space. Our goal is the approximation of the solution u by a (usually much more sparsely representable) TT tensor-valued function $U(t) \in \mathbb{T} = \mathbb{T}(\underline{r})$ of fixed rank $\underline{r}$ chosen in advance. To this end, we replace the equation (55) by a related differential equation posed *on the approximation manifold* $\mathbb{T}$, i.e. for all starting values $U(0) = U_0 \in \mathbb{T}$, the solution fulfils $U(t) \in \mathbb{T}$. This is achieved as follows: If a manifold $\mathbb{T}$ possesses a local embedding $\psi = \psi_U$ for each $U \in \mathbb{T}$, as the TT manifold does by Theorem 3, a differential equation $\dot{U} = F(U)$ is a differential equation on $\mathbb{T}$ if and only if

$$F(U) \in \mathcal{T}_U \mathbb{T}$$

holds for all $U \in \mathbb{T}$, see e.g. [17], Theorem 5.2; in particular, we have $U(t) \in \mathbb{T}$ for all $t > 0$. Therefore, (55) can be solved approximately by projecting $f(U)$ on the tangent space $\mathcal{T}_U \mathbb{T}$ for each $U \in \mathbb{T}$, i.e. by defining

$$F(U) := P_{\mathcal{T}_U} f(U) \in \mathcal{T}_U \mathbb{T},$$

where $P_{\mathcal{T}_U}$ projects on the tangent space $\mathcal{T}_U\mathbb{T}$, and by then solving the projected differential equation

$$\frac{dU}{dt} = F(U), \tag{56}$$

that, by the above reasoning, is a differential equation posed on the manifold $\mathbb{T}$. Note that the solution $U(t)$ is a curve in $\mathbb{T}$, so by definition dU/dt is contained in the tangent space $\mathcal{T}_U\mathbb{T}$. Thus $P_{\mathcal{T}_U} dU/dt = dU/dt$, which implies that

$$\frac{dU}{dt} - F(U) = 0 \Longleftrightarrow \left\langle \frac{dU}{dt} - f(U), \delta U \right\rangle = 0 \quad \text{for all} \quad \delta U \in \mathcal{T}_U\mathbb{T}, \tag{57}$$

i.e. (56) can therefore be interpreted as the Galerkin projection of the original problem onto the state dependent test space $\mathcal{T}_U\mathbb{T}$. In the context of time-dependent quantum chemistry, the above proceeding is well-known as the *Dirac-Frenkel time dependent variational principle*, see [31]. The projected problem can now be solved by applying standard methods [28] to the equivalent differential equation

$$\frac{dz}{dt} = \tau^{-1}\big(F(\psi(z))\big), \tag{58}$$

with τ from (50) and ψ being a local parametrization. On the whole, problem (55) on $\mathbb{R}^{n_1\times\cdots\times n_d}$ is thus replaced by a differential equation in the (linear, much lower-dimensional) coordinate space $\mathbb{R}^D$, also cf. [17] for details.

The above problem of solving the partial differential equation (55) includes the noteworthy special case of the standard minimisation problem

$$\mathcal{J}(u) \to \min \qquad \text{for differentiable} \quad \mathcal{J} : \mathbb{R}^{n_1\times\cdots\times n_d} \to \mathbb{R}, \tag{59}$$

which by the choices

$$\mathcal{J}(u) = \frac{\langle Au, u\rangle}{\langle u, u\rangle}, \qquad \mathcal{J}(u) = \|Au - b\|^2, \qquad \mathcal{J}(u) = \|f(u)\|^2$$

for instance includes the problem of finding the lowest eigenvalue for a symmetric positive definite A, solution of linear equations for such A, nonlinear equations and also, by letting $A = \hat{A}^{\frac{1}{2}}$ for $\hat{A} > 0$, $b = \hat{A}^{\frac{1}{2}}\hat{b}$, the problem of finding a best approximation to given $\hat{b}$ with respect to the $\hat{A}$ inner product. Defining the gradient flow of $\mathcal{J}$ by the differential equation

$$\frac{du}{dt} = -\mathcal{J}'(u), \tag{60}$$

the problem (59) is equivalent to computation of the long-term behaviour $\lim_{t\to\infty} u(t)$ of a solution of (60). The variational formulation on $\mathbb{T}$, obtained in the above way, reads

$$\left\langle \frac{dU}{dt} + \mathcal{J}'(U), \delta U \right\rangle = 0 \qquad \text{for all} \quad \delta U \in \mathcal{T}_U \mathbb{T}, \tag{61}$$

and a solution $U(t) \in \mathbb{T}$ to (60) can be computed by the above mentioned methods.

7 Numerical examples

For the practical realization of solutions of differential equations and optimization problems by approximation on the manifold $\mathbb{T}_{\underline{r}}$, an important algorithmic ingredient is the computation of a best rank-$\underline{r}$-approximation for a given tensor U. In this final section, we demonstrate that for this task we have two stable algorithms at hand, both providing reasonable alternatives while differing in their theoretical properties. The practical treatment of linear and nonlinear equations on $\mathbb{T}$ by various algorithms, often using the best approximation step investigated here, is an extensive topic of its own that will be treated in a forthcoming publication [19].

The first algorithm tested is the `full_to_tt` method from the TT-Toolbox, introduced in [33] and reproduced in this paper in Fig. 4. The method creates a TT tensor by computation of successive SVDs, afterwards truncating them at a chosen bound relative to the Frobenius norm of the full tensor. The `full_to_tt` method yields the best approximation of rank r up to a factor of $\sqrt{d-1}$ [15]. We compare `full_to_tt` to a modified alternating least square algorithm (MALS) inspired by the DMRG algorithm [46] used in quantum physics. It computes a stationary point of the functional measuring the distance between U and the approximation in the Frobenius norm, and is therefore in contrast to the SVD-approach of `full_to_tt` capable of finding the exact best approximation (if not a local minimum). The MALS algorithm is introduced at all length in [19], we therefore only note here that a MALS step consists of $2d-3$ inner iterations (one "left and right sweep"), in each of which two component functions are contracted, optimized and then decomposed again by truncated SVDs to update one of the component functions. This corresponds to a computation on the tangent space $\mathbb{T}_{\underline{r}'}$ of a reordering of the current iterate to a tensor $\tilde{U}^{(n)} \in \mathbb{R}^{n_1 \times \cdots \times (n_i n_{i+1}) \times \cdots \times n_d}$.

In the experiment, we approximated synthetic data obtained from the combination of continuous functions. We use a variation of the Friedman1 data set [14] for example used in [3]. For $d = 5$ and $n_i = n$ with $n \in \{3, \ldots, 20\}$, we use index sets

$$\mathcal{I}_1^{(n)} = \cdots = \mathcal{I}_5^{(n)} = \left\{ \frac{k}{n} \mid k = 0, \ldots, n \right\}.$$

We approximate the tensor given by

$$U^{(n)} : \mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_5^{(n)} \to \mathbb{R},$$
$$U(\underline{x}) = 10 \sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5.$$

The bound for the SVD truncations was set to 10^{-12} in `full_to_tt` and to 10^{-9} for `MALS`. All algorithms were implemented in Matlab using only the build-in functions and the TT-Toolbox [36], using an AMD processor with four 2,6 GHz cores and a total

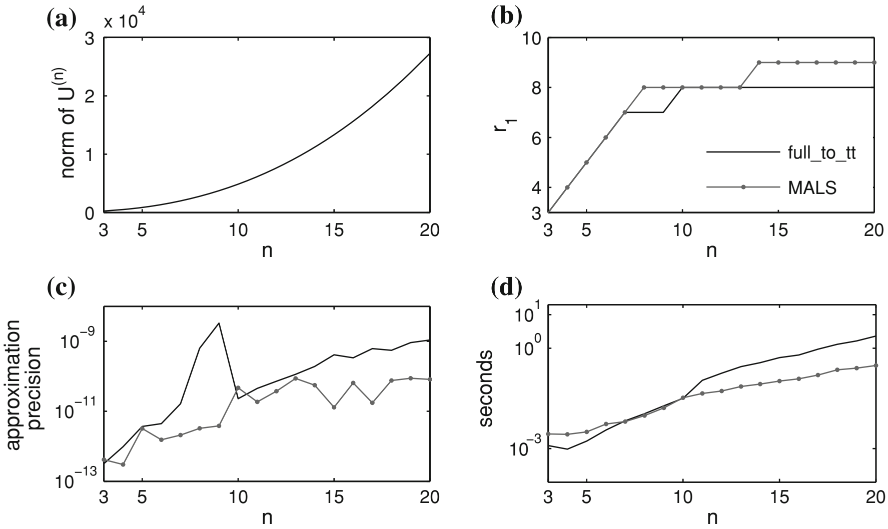

Fig. 6 TT approximations of Friedman data sets

16 GB RAM. For the MALS algorithm, we started with a random TT approximation of rank one.

Note that the Frobenius norm of U grows exponentially with n, see Fig. 6a. The algorithms `full_to_tt` and `MALS` both produce TT approximations with ranks $r_2 = r_3 = r_4 = 2$. The results for r_1 are depicted in Fig. 6b. In (c), the absolute approximation errors (in the Frobenius norm) are displayed. The plot (b) shows that for some values of n, `MALS` computes an approximation for which the rank r_1 is bigger than that computed by `full_to_tt`, corresponding a better approximation precision in (c). Figure 6d compares the computation times needed for both decompositions.

Our experiments demonstrate that as noted in the literature before, the `full_to_tt` method works stably due to the use of SVDs; also, in contrast to ALS for the canonical format, the `MALS` variation for TT works very well in many cases: It seems to be as stable as the direct SVD `full_to_tt` method, while often being much faster than `full_to_tt`.

8 Conclusion

We have shown that the set $\mathbb{T}_{\underline{r}}$ of tensors of fixed TT rank, being the simplest special case of the hierarchical HT format [16], locally forms an embedded submanifold of $\mathbb{R}^{n_1 \times \cdots \times n_d}$. This is analogous to according results for the Tucker format [25], and similar methods for the approximation of high dimensional problems as differential equations and optimisation problems may now be utilized. In particular, we have parametrized the tangent spaces of $\mathbb{T}$ uniquely by introducing appropriate gauge conditions similar to those in [31]. Thus, persuing the quasi-Galerkin approach introduced in Sect. 6, the according projectors on the respective tangent spaces—needed for a numerical treatment in this vein—may now be computed explicitly. Although we have only recently started the development and implementation of such algorithms, our preliminary numerical examples given in this paper show that in practice, the stability of TT format is competitive with the Tucker format. Its potential for the treatment of high-dimensional problems will be explored further in a forthcoming publication.

Acknowledgments This work has been supported by the DFG SPP 1324 "Mathematische Methoden zur Extraktion quantifizierbarer Information aus komplexen Systemen". The authors also thank the DFG research center MATHEON and the Berlin Mathematical School (BMS) for providing travel means used in connection with the development of this work.

References

1. Beck, M.H., Jäckle, A., Worth, G.A., Meyer, H.D.: The multiconfiguration time dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. Phys. Rep. **324**, 1 (2000)
2. Bellman, R.E.: Adaptive Control Processes. Princeton University Press, Princeton (1961)
3. Beylkin, G., Garcke, J., Mohlenkamp, M.J.: Multivariate regression and machine learning with sums of separable functions. SIAM J. Sci. Comput. **31**(3), 1840 (2009)
4. Conte, D., Lubich, C.: An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics. M2AN **44**, 759 (2010)
5. Crawford, T.D., Schaeffer, H.F. III.: An introduction to coupled cluster theory for computational chemists. Rev. Comput. Chem. **14**, 33 (2000)
6. De Lathauwer, L., De Moor, B., Vandewalle, J.: A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. **21**(4), 1253–1278 (2000)
7. De Lathauwer, L., De Moor, B., Vandewalle, J.: On the best rank-1 and rank-$(R1, R2, \ldots, RN)$ approximation and applications of higher-order tensors. SIAM J. Matrix Anal. Appl. **21**(4), 1324 (2000)
8. de Silva, V., Lim, L.-H.: Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM J. Matrix Anal. Appl.: Special Issue on Tensor Decompositions and Applications **30**(3), 1084–1127 (2008)
9. Eldén, L., Savas, B.: A Newton–Grassmann method for computing the best multi-linear rank-(r_1, r_2, r_3) approximation of a tensor. SIAM J. Matrix Anal. Appl. **31**, 248 (2009)
10. Espig, M.: Effziente Bestapproximation mittels Summen von Elementartensoren in hohen Dimensionen. Ph.D. thesis (2007)
11. Espig, M., Hackbusch, W., Rohwedder, T., Schneider, R.: Variational calculus with sums of elementary tensors of fixed rank. Numer. Math. (submitted)
12. Falcó, A., Hackbusch, W.: On minimal subspaces in tensor representations. Preprint, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig. http://www.mis.mpg.de/publications/preprints/2010/prepr2010-70.html (2010)
13. Fannes, M., Nachtergaele, B., Werner, R.F.: Finitely correlated states on quantum spin chains. Commun. Math. Phys. **144**, 443 (1992)
14. Friedman, J.H.: Multivariate adaptive regression splines (with discussion). Ann. Stat. **191**, 1 (1991)
15. Grasedyck, L.: Hierarchical singular value decomposition of tensors. SIAM. J. Matrix Anal. Appl. **31**, 2029 (2010)
16. Hackbusch, W., Kühn, S.: A new scheme for the tensor representation. J. Fourier Anal. Appl. **15**, 706–722 (2009)
17. Hairer, E., Lubich, C., Wanner, G.: Geometrical Numerical Integration—Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edn. Springer, Berlin (2006)
18. Helgaker, T., Jørgensen, P., Olsen, J.: Molecular Electronic-Structure Theory. Wiley, New York (2000)
19. Holtz, S., Rohwedder, T., Schneider, R.: The alternating linear scheme for tensor optimisation in the TT format. SISC (2010, submitted)
20. Huckle, T., Waldherr, K., Schulte-Herbrüggen, T.: Computations in quantum tensor networks. Linear Algebra Appl.: Special Issue on Tensors. http://www5.in.tum.de/pub/CompQuantTensorNetwork.pdf (2010, submitted)
21. Kapteyn, A., Neudecker, H., Wansbeek, T.: An approach to n-mode components analysis. Psychometrika **51**, 269 (1986)

22. Khoromskij, B.N., Khoromskaia, V.: Multigrid accelerated tensor approximation of function related multidimensional arrays. SIAM J. Sci. Comput. **31**(4), 3002 (2009)
23. Klümper, A., Schadschneider, A., Zittartz, J.: Groundstate properties of a generalized VBS-model. Z. Phys. B: Condensed Matter **87**, 281–287 (1992)
24. Koch, O., Lubich, C.: Dynamical low rank approximation. SIAM J. Matrix Anal. Appl. **29**(2), 434 (2008)
25. Koch, O., Lubich, C.: Dynamical low-rank approximation of tensors. SIAM J. Matrix Anal. Appl. **31**, 2360 (2010)
26. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM Rev. **51**(3), 455–500 (2008)
27. Kroonenberg, P.M., De Leeuw, J.: Principal component analysis of three-mode data by means of alternating least squares algorithms. Psychometrika **45**, 69 (1980)
28. Kunkel, P., Mehrmann, V.: Differential-Algebraic Equations Analysis and Numerical Solution. EMS Publishing House, Zurich (2006)
29. Lang, S.: Fundamentals of Differential Geometry. Springer (2001)
30. Lee, J.M.: Manifolds and differential geometry. In: Graduate Studies in Mathematics, vol. 107. AMS (2009)
31. Lubich, C.: From quantum to classical molecular dynamics: reduced methods and numerical analysis. In: Zürich Lectures in Advanced Mathematics. EMS (2008)
32. Marti, K.H., Bauer, B., Reiher, M., Troyer, M., Verstraete, F.: Complete-graph tensor network states: a new fermionic wave function ansatz for molecules. New J. Phys. **12**, 103008 (2010)
33. Oseledets, I.: Compact matrix form of the d-dimensional tensor decomposition. SIAM J. Sci. Comput. (2009, submitted)
34. Oseledets, I.: On a new tensor decomposition. Doklady Math. **80**(1), 495–496 (2009)
35. Oseledets, I.: Tensors inside matrices give logarithmic complexity. Preprint 2009-04, IMA RAS April 2009. SIAM J. Matrix Anal. Appl. (accepted)
36. Oseledets, I.: TT Toolbox 1.0: Fast multidimensional array operations in MATLAB. Preprint 2009-06, INM RAS, August 2009
37. Oseledets, I., Tyrtyshnikov, E.E.: Breaking the curse of dimensionality, or how to use SVD in many dimensions. SIAM J. Sci. Comput. **31**, 5 (2009)
38. Oseledets, I.V., Tyrtyshnikov, E.E.: Tensor tree decomposition does not need a tree. Linear Algebra Appl. (2009, submitted)
39. Savas, B., Lim, L.-H.: Quasi-Newton methods on Grassmannians and multilinear approximations of tensors. ARXIV, eprint arXiv:0907.2214, http://arxiv.org/abs/0907.2214, 2009. Also accepted in SIAM J. Sci. Comput. (2010)
40. Schneider, R., Rohwedder, T., Blauert, J., Neelov, A.: Direct minimization for calculating invariant subspaces in density functional computations of the electronic structure. J. Comput. Math. **27**, 360 (2009)
41. Schollwöck, U.: The density-matrix renormalization group. Rev. Mod. Phys. **77**(1), 259 (2005)
42. Szabo, A., Ostlund, N.S.: Modern Quantum Chemistry. Dover, New York (1992)
43. Tucker, L.R.: Some mathematical notes on three-mode factor analysis. Psychometrica **31**(3), 279–311 (1966)
44. Van Loan, C.F.: Tensor network computations in quantum chemistry. www.cs.cornell.edu/cv/OtherPdf/ZeuthenCVL.pdf (2008)
45. Vidal, G.: Efficient classical simulation of slightly entangled quantum computation. Phys. Rev. Lett. **91**(14), 147902 (2003)
46. White, S.: Density matrix formulation for quantum renormalization groups. Phys. Rev. Lett. **69**, 2863 (1992)