Acta Numerica (2023), pp. 1–121 Printed in the United Kingdom
doi:10.1017/S0962492922000125

Low-rank tensor methods for partial differential equations

Markus Bachmayr
Institut für Geometrie und Praktische Mathematik, RWTH Aachen University,
Templergraben 55, 52056 Aachen, Germany
E-mail: bachmayr@igpm.rwth-aachen.de

Low-rank tensor representations can provide highly compressed approximations of functions. These concepts, which essentially amount to generalizations of classical techniques of separation of variables, have proved to be particularly fruitful for functions of many variables. We focus here on problems where the target function is given only implicitly as the solution of a partial differential equation. A first natural question is under which conditions we should expect such solutions to be efficiently approximated in low-rank form. Due to the highly nonlinear nature of the resulting low-rank approximations, a crucial second question is at what expense such approximations can be computed in practice. This article surveys basic construction principles of numerical methods based on low-rank representations as well as the analysis of their convergence and computational complexity.

2020 Mathematics Subject Classification:
Primary 41A46, 41A63, 65D40
Secondary 65F55, 65J10, 65M12, 65N12, 65N25, 65Y20

CONTENTS

1. Introduction

A matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$ of rank r can be represented in low-rank format $\mathbf{M} = \mathbf{A}\mathbf{B}^\top$ with matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times r}$, which when $r \ll n$ substantially reduces the required storage from n^2 to $2nr$. Moreover, the low-rank format also leads to a reduction of the computational costs of many linear algebra operations, such as multiplication by matrices from the left or right. This basic idea can be generalized to tensors of order $d > 2$, represented – in the real-valued case that we focus on here – by elements of $\mathbb{R}^{n_1\times\ldots\times n_d}$ with $n_1, \ldots, n_d \in \mathbb{N}$. This leads to low-rank formats for higher-order tensors, which can be used to significantly reduce the dependence on d of storage requirements and computational complexity of numerical methods. One can thus hope to avoid the *curse of dimensionality*, that is, the exponential scaling with respect to d obtained with standard methods designed for lower-dimensional problems.

A main motivation for the development of numerical methods based on low-rank tensor representations is the application to partial differential equations posed on high-dimensional domains, which frequently arise in the context of quantum physics or probabilistic models. Closely related are methods for model order reduction of multi-parametric problems. In these settings, coefficient arrays of basis expansions of solutions can be regarded as high-order tensors. Low-rank approximations of these high-dimensional coefficient arrays lead to nonlinear approximation methods, that is, approximations are not chosen from linear spaces and are parametrized in a nonlinear manner. As above in the parametrization of $\mathbf{M}$ by the factors $\mathbf{A}$ and $\mathbf{B}$, this nonlinearity is of multiplicative type.

Certain types of low-rank tensor representations have proved to be particularly suited to the construction of numerical methods for problems in very high dimensions, such as hierarchical tensors, which are also known as tree tensor networks in physics. Since we consider approximations in such formats for functions that are given only implicitly as solutions of partial differential equations, not only the solution itself but also the problem data (such as domains, right-hand sides, boundary or initial conditions) need to have efficient low-rank approximations. A central issue in the construction of solvers is the adaptation of rank parameters in tensor representations: these ranks needs to be chosen sufficiently large to ensure convergence, but should also be kept as small as possible due to their strong influence on the computational costs of operating on low-rank representations.

The focus of this article is in particular on how low-rank tensor formats interact with discretizations of partial differential equations in the construction of numerical solvers. The remainder of this section gives a conceptual overview; further details and references are given in the respective later sections.

1.1. Notation

Vectors, matrices and tensors are denoted by bold-face letters. On the one hand, we consider tensors on finite index sets, that is, elements of $\mathbb{R}^{n_1\times\cdots\times n_d}$. These

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

are tensors of *order* d, where each dimension in the index set is called a *mode* of the tensor, with mode i having *mode size* n_i for $i = 1, \dots, d$. On the other hand, an important role is played by tensors with square-summable entries on countable d-dimensional index sets, which are sequences in $\ell_2(\mathbb{N}^d)$. In both cases, in order to avoid too many subscripts, for the entry of a tensor $\mathbf{t}$ of order d at index $(\nu_1, \dots, \nu_d) \in \mathbb{N}^d$, we write $\mathbf{t}[\nu_1, \dots, \nu_d]$.

For the unit cube in d dimensions, which frequently appears in high-dimensional model problems, we use the abbreviation $\square_d = (0,1)^d$. We write $\operatorname{supp} f$ for the support of the function f and use the same notation for the supports of sequences: for $\mathbf{v} \in \mathbb{R}^I$ with I a countable index set,

$$\operatorname{supp} \mathbf{v} = \{\nu \in I \colon \mathbf{v}[\nu] \neq 0\}.$$

We write $\operatorname{cond}(\mathbf{A})$ for the spectral norm condition number of the mapping defined by $\mathbf{A}$.

For scalar expressions A and B, we write $A \lesssim B$ to denote $A \leq CB$ with a constant $C > 0$ that does not depend on quantities that explicitly appear in A or B; $A \gtrsim B$ for $B \lesssim A$; and $A \eqsim B$ for $A \lesssim B$ and $A \gtrsim B$. Inequalities between tuples (such as tuples of rank parameters) are always to be understood componentwise.

1.2. Classes of model problems

Let us first consider several typical model cases of high-dimensional partial differential equations (PDEs). A first problem of frequent interest is the high-dimensional diffusion problem of the basic form

$$\begin{aligned} \partial_t u - \operatorname{div}_x(A \operatorname{grad}_x u + bu) &= f \quad \text{in } (0,T] \times \Omega, \\ u = 0 \quad \text{on } (0,T] \times \partial\Omega, \quad u(0,x) &= u_0(x) \quad \text{for } x \in \Omega. \end{aligned} \tag{1.1}$$

Here we consider a bounded domain Ω with homogeneous Dirichlet boundary conditions for simplicity, but problems of this type can also be considered on unbounded domains such as $\Omega = \mathbb{R}^d$. The sought solution u as well as the source term f and initial condition u_0 are scalar functions, whereas A and b take values in $\mathbb{R}^{d\times d}$ and $\mathbb{R}^d$, respectively. Equations of this type arise, for instance, as Kolmogorov equations describing the time evolutions of probability densities of stochastic processes.

A second source of high-dimensional problems is given by quantum-physical models, where the dimensionality of the problem scales with the number of particles. In classical quantum mechanics, the time evolution of quantum systems is modelled by the time-dependent Schrödinger equation for the complex-valued wavefunction u. For a given potential V and initial state u_0, it takes the form

$$\mathrm{i}\partial_t u = -\Delta u + Vu \text{ on } \mathbb{R}^d, \quad u|_{t=0} = u_0. \tag{1.2}$$

The eigenpairs of the Hamiltonian operator $-\Delta+V$ on the right-hand side correspond to the quantum states for the total energy of the system; the associated Schrödinger

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

eigenvalue problem reads

$$-\Delta u + Vu = \lambda u \quad \text{on } \mathbb{R}^d. \tag{1.3}$$

One example that is of central importance in quantum chemistry is the electronic Schrödinger eigenvalue problem, which provides the (non-relativistic) standard model of electronic structure of molecules for given positions of atomic nuclei. In this case we have $d = 3N$ for a system of N electrons, and V is a sum of Coulomb potentials.

We are also frequently interested in stationary states of (1.1), which are described by second-order elliptic equations. A frequently considered model problem, with the simplification $b = 0$, reads

$$-\operatorname{div}(A \operatorname{grad} u) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{1.4}$$

Such problems of elliptic type also arise as parts of algorithms for the parabolic problem (1.1) and the Schrödinger problems (1.2), (1.3) (with $A = \mathrm{I}$ and $\Omega = \mathbb{R}^d$) and are thus a natural starting point for the development of numerical methods.

The above model problems have in common that they involve differential operators on high-dimensional domains. A different type of high-dimensionality originates from problem data that depend on parameters, for instance some parameter y from a set Y. In the example of (1.4), with parameter-dependent diffusion tensor A_y and associated solutions u_y, this leads to the mapping $y \mapsto u_y$ defined for all $y \in Y$ by

$$-\operatorname{div}(A_y \operatorname{grad} u_y) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{1.5}$$

Instead of evaluating u_y independently for each required y, in the context of model order reduction or uncertainty quantification, we are frequently interested in economical approximations of the higher-dimensional function $u(x, y) = u_y(x)$ for $(x, y) \in \Omega \times Y$.

After treating the basics of low-rank tensor representations in Section 2, in the following sections we first focus on elliptic problems. Their solution is of interest in its own right, but also provides important building blocks for more involved problems. In particular, the norm in which solutions are to be approximated is important: in (1.4) and (1.5), we typically require approximations in the norm of the Sobolev space $H^1(\Omega)$, which is equivalent to the *energy norm* associated to the problem. How this needs to be taken into account in the construction of low-rank solvers is treated in Section 4. After treating low-rank solvers for elliptic problems in Section 5, we return to eigenvalue problems and time-dependent problems in Sections 6 and 7.

Concerning the application of low-rank approximations to solutions of PDE problems as above, there are several fundamental questions: Can we have algorithms with guaranteed convergence to the exact solution, that is, with the ability to produce approximations of the solution of the original PDE problem with any desired accuracy with respect to the appropriate norm? Can these algorithms be

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

ensured to be numerically stable? Can the costs of numerically computing low-rank approximations be quantified, and can we ensure that these costs are optimal in an appropriate sense? For problems in high dimensions d, we aim to avoid the curse of dimensionality, that is, an exponential dependence of the costs on d. This imposes requirements on both the approximability of solutions and the costs of numerical algorithms.

1.3. Separation of variables

A basic instance of low-rank methods is provided by the classical techniques of separation of variables in partial differential equations, where higher-dimensional problems are decomposed into simpler ones. As an example, consider the eigenvalue problem for the negative Laplacian

$$-\Delta = -\sum_{i=1}^{d} \partial_{x_i}^2$$

on $\square_d$ with homogeneous Dirichlet boundary values on $\partial\square_d$. From the solutions of the one-dimensional eigenvalue problem for the negative second derivative on $(0, 1)$ with homogeneous Dirichlet boundary conditions,

$$-\partial_x^2 \sin(\ell\pi x) = (\ell\pi)^2 \sin(\ell\pi x), \quad \ell \in \mathbb{N},$$

it is easy to see that for each $k = (k_1, \dots, k_d) \in \mathbb{N}^d$, the functions

$$\Phi_k(x_1, \dots, x_d) = \prod_{i=1}^{d} \sin(k_i \pi x_i)$$

satisfy $\Phi_k|_{\partial\square_d} = 0$ and

$$-\Delta\Phi_k = \pi^2 \left(k_1^2 + \cdots + k_d^2\right) \Phi_k.$$

Thus Φ_k is an eigenfunction of $-\Delta$ on $\square_d$, and since $\operatorname{span}\{\Phi_k\}_{k\in\mathbb{N}^d}$ is dense in $L_2(\square_d)$, all eigenfunctions in $L_2(\square_d)$ are of this form.

By a similar use of the structure of the Laplacian, certain multidimensional initial value problems can also be reduced to spatially one-dimensional equations. Let

$$g(x_1, \dots, x_d) = \prod_{i=1}^{d} g_i(x_i)$$

with

$$g_i(x) = \sum_{k=1}^{\infty} c_k^{(i)} \sin(k\pi x).$$

Then the solution of the heat equation $\partial_t u - \Delta u = 0$ on $\mathbb{R}^+ \times \square_d$ with homogeneous Dirichlet boundary conditions and initial condition $u(0, \cdot) = g$ on $\square_d$ has the product

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

form

$$u(t,x_1,\dots,x_d)=\prod_{i=1}^{d}u_i(t,x_i),\quad u_i(t,x)=\sum_{k=1}^{\infty}c_k^{(i)}e^{-(k\pi)^2t}\sin(k\pi x),$$

where u_i solve the spatially univariate problems $\partial_t u_i-\partial_x^2 u_i=0$, $u_i(0,\cdot)=g_i$.

This is an instance of the following more general structure: let $H_1,\dots,H_d$ be separable Hilbert spaces and consider the tensor product Hilbert space $H=\bigotimes_{i=1}^{d}H_i$ and an operator $A\colon H\to H$ of the Laplacian-type form

$$A=A_1\otimes \mathrm{I}\otimes\cdots\otimes\mathrm{I}+\mathrm{I}\otimes A_2\otimes\mathrm{I}\otimes\cdots\otimes\mathrm{I}+\cdots+\mathrm{I}\otimes\cdots\otimes\mathrm{I}\otimes A_d, \tag{1.6}$$

where for $i=1,\dots,d$, $A_i\colon H_i\to H_i$ are self-adjoint and positive definite, and where I denotes the identity mapping. Assume that the normalized eigenfunctions $\{\phi_k^{(i)}\}_{k\in\mathbb{N}}$ of A_i are an orthonormal basis of H_i for $i=1,\dots,d$. Then, for each $k=(k_1,\dots,k_d)$,

$$\Phi_k=\bigotimes_{i=1}^{d}\phi_{k_i}^{(i)}$$

is an eigenfunction of A. Since $\{\Phi_k\}_{k\in\mathbb{N}^d}$ is an orthonormal basis of H, every eigenfunction of A is of this form. As a further consequence, for $g\in H$ of the form $g=\bigotimes_{i=1}^{d}g_i$ with $g_i\in H_i$, we also have that the solution of the initial value problem $u'+Au=0$, $u(0)=g$ takes the form

$$u(t)=\bigotimes_{i=1}^{d}e^{-tA_i}g_i, \tag{1.7}$$

where $e^{-tA_i}g_i$ solves an analogous evolution problem for the operator A_i, with initial value g_i on H_i for each i. Thus, for A and f of this particular structure, both the eigenvalue problem and the associated homogeneous evolution problem can be reduced to problems on the component spaces $H_1,\dots,H_d$.

However, such simple product structures no longer apply when considering slightly different problems, such as non-homogeneous evolution equations $u'+Au=f$ or associated stationary problems $Au=f$. In these cases there is generally no longer a straightforward reduction to lower-dimensional problems. In some cases we can still find semi-analytical approximations that yield a reduction: given A of the Laplace-like form (1.6), for u solving $Au=f=\bigotimes_{i=1}^{d}f_i$ we have efficient *exponential sum approximations* of the form

$$u=A^{-1}f\approx\sum_{k=1}^{r}\omega_k\bigotimes_{i=1}^{d}e^{-\alpha_k A_i}f_i. \tag{1.8}$$

It can be shown, as detailed in Section 4.2, that there exist coefficients $\omega_k,\alpha_k>0$ for $k=1,\dots,r$ such that we can achieve approximation error ε for u in suitable norms with $r\approx|\log\varepsilon|^p$, with some $p\geq 1$ depending in particular on the chosen

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

norm. Note that the same construction can still be applied to f given by a *sum* of elementary tensor products.

Approximations as in (1.8) indicate that sums of separable terms can provide a more general framework for reducing high- to lower-dimensional problems. The above example, however, still depends quite strongly on the particular structure of the problem. This motivates more generally applicable computational methods based on low-rank tensor representations. In such approaches, solutions u of quite generic problems are approximated by more adaptable structures for separating coordinates which, however, still offer the advantage of reducing the computational effort.

1.4. Low-rank tensor representations

The notion of tensors can be approached in different ways. For a detailed treatment of different definitions, we refer to Lim (2021). Although tensors can be defined in a completely abstract manner as elements of tensor products of vector spaces, they can also be regarded as describing multilinear mappings on Cartesian products of vector spaces; with respect to bases of these vector spaces, tensors can be described by their coefficient arrays, which need to satisfy certain invariance properties with respect to changes of bases. Note that in physics, *tensor fields* (that is, tensor-valued mappings on manifolds) are also often referred to simply as tensors, but these geometric objects do not play a role in what follows.

The discrete analogue of separable functions, as considered in the previous subsection, is given by tensors of rank one, that is, *elementary tensors* of the form $\mathbf{t} = \mathbf{v}^{(1)} \otimes \mathbf{v}^{(2)} \otimes \cdots \otimes \mathbf{v}^{(d)}$ with $\mathbf{v}^{(i)} \in \mathbb{R}^{n_i}$ for $i = 1, \ldots, d$. The entries of $\mathbf{t}$ are then given by

$$\mathbf{t}[j_1, \ldots, j_d] = \prod_{i=1}^{d} \mathbf{v}^{(i)}[j_i].$$

A natural generalization is the *canonical tensor format*: clearly, each $\mathbf{t} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ can be represented as a sum of elementary tensors in the form

$$\mathbf{t} = \sum_{k=1}^{r} \bigotimes_{i=1}^{d} \mathbf{v}_k^{(i)}, \tag{1.9}$$

with $\mathbf{v}_k^{(i)} \in \mathbb{R}^{n_i}$ for $k = 1, \ldots, r$, $i = 1, \ldots, d$, where the smallest possible r in such a representation is called the *tensor rank* of $\mathbf{t}$. This tensor format, also called the canonical polyadic (CP) format, appeared already in Hitchcock (1927, 1928).

The use of representations (1.9) in iterative methods for solving high-dimensional partial differential equations was proposed in the seminal papers by Beylkin and Mohlenkamp (2002, 2005). An important requirement for such algorithms is a computational procedure for approximating a given tensor $\mathbf{t}^*$ by a tensor of lower rank, for instance the lowest rank permissible for a certain accuracy. In other words,

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

for some prescribed rank r, we aim to minimize $\|\mathbf{t}^* - \mathbf{a}\|_2$ over all tensors $\mathbf{a}$ of the same mode sizes as $\mathbf{t}^*$ with rank at most r.

As demonstrated by de Silva and Lim (2008), this problem of best approximation of tensor rank r is generally ill-posed for representations of the form (1.9). This is illustrated by the following example (de Silva and Lim 2008, Prop. 4.6): let $\mathbf{x}_i, \mathbf{y}_i$ be linearly independent vectors in $\mathbb{R}^{n_i}$ for $i = 1, 2, 3$, and define the tensors

$$\mathbf{t}_n = n\left(\mathbf{x}_1 + \frac{1}{n}\mathbf{y}_1\right) \otimes \left(\mathbf{x}_2 + \frac{1}{n}\mathbf{y}_2\right) \otimes \left(\mathbf{x}_3 + \frac{1}{n}\mathbf{y}_3\right) - n\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3, \tag{1.10}$$

which are of rank two. Then the limit of the tensors $\mathbf{t}_n$ as $n \to \infty$ is

$$\mathbf{t}^* = \mathbf{y}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 + \mathbf{x}_1 \otimes \mathbf{y}_2 \otimes \mathbf{x}_3 + \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_3,$$

which can be verified to have rank three.

This shows that, unlike the matrix rank for $d = 2$, for orders $d > 2$, the tensor rank is not lower semicontinuous, and in particular, the infimum

$$\inf\left\{\|\mathbf{t}^* - \mathbf{a}\|_2 \colon \mathbf{a} \in \mathbb{R}^{n_1 \times n_2 \times n_3} \text{ has tensor rank at most } 2\right\} = 0$$

is not attained. The properties of the tensor rank in the higher-order case are different from those of the matrix rank in other respects too. For instance, as shown by Hillar and Lim (2013), determining the tensor rank of a given tensor with $d > 2$ is generally an NP-hard problem.

The ill-posedness of the best tensor rank-r approximation problem poses a fundamental difficulty in the construction of reliable numerical approximation procedures, even when we are content with near-best approximations. These issues can be circumvented by different types of low-rank representations that can be characterized in terms of matrix ranks. The most classical instance of such a representation is the *Tucker format*, considered in the case $d = 3$ by Tucker (1964, 1966). In this particular case, for $\mathbf{t} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, a representation in Tucker format takes the form

$$\mathbf{t} = \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_2} \sum_{k_3=1}^{r_3} \mathbf{a}[k_1, k_2, k_3]\, \mathbf{U}_{k_1}^{(1)} \otimes \mathbf{U}_{k_2}^{(2)} \otimes \mathbf{U}_{k_3}^{(3)} \tag{1.11}$$

with the core tensor $\mathbf{a} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$ and vectors $\mathbf{U}_k^{(i)} \in \mathbb{R}^{n_i}$ for $k = 1, \dots, r_i$ and $i = 1, 2, 3$.

The entrywise smallest possible tuple (r_1, r_2, r_3) for representing $\mathbf{t}$ in the form (1.11) is called the *multilinear rank* of $\mathbf{t}$. It can be characterized in terms of matrix ranks: for the multilinear rank (r_1, r_2, r_3) of $\mathbf{t}$, with the matricizations

$$\begin{aligned}
\mathbf{M}_1 &= (\mathbf{t}[\nu_1, \nu_2, \nu_3])_{\nu_1 \in \{1,\dots,n_1\},(\nu_2,\nu_3) \in \{1,\dots,n_2\} \times \{1,\dots,n_3\}},\\
\mathbf{M}_2 &= (\mathbf{t}[\nu_1, \nu_2, \nu_3])_{\nu_2 \in \{1,\dots,n_2\},(\nu_1,\nu_3) \in \{1,\dots,n_1\} \times \{1,\dots,n_3\}},\\
\mathbf{M}_3 &= (\mathbf{t}[\nu_1, \nu_2, \nu_3])_{\nu_3 \in \{1,\dots,n_3\},(\nu_1,\nu_2) \in \{1,\dots,n_1\} \times \{1,\dots,n_2\}}
\end{aligned}$$

of $\mathbf{t}$, we have

$$r_1 = \operatorname{rank} \mathbf{M}_1, \quad r_2 = \operatorname{rank} \mathbf{M}_2, \quad r_3 = \operatorname{rank} \mathbf{M}_3.$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

In other words, we obtain the most efficient representations of the form (1.11) by choosing $\{\mathbf{U}_k^{(i)}\}_{k=1,\dots,r_i}$ as a basis of the subspace range$(\mathbf{M}_i) \subset \mathbb{R}^{n_i}$ for $i = 1, 2, 3$.

For the Tucker format, the existence of best approximations of prescribed rank is guaranteed (see Section 2.7). Moreover, standard linear algebra procedures for matrices can be used for the computation of near-best approximations by the *higher-order singular value decomposition*, which can serve as the basis of reliable computational methods. However, this format by itself is not suitable for avoiding the curse of dimensionality, since in general the number of entries $\prod_{i=1}^{d} r_i$ of the core tensor $\mathbf{a}$ still scales exponentially with respect to d.

This problem is addressed by more general *subspace tensor formats* that use further decompositions of $\mathbf{a}$ into lower-order tensors, which amounts to choosing subspaces not only for single tensor modes as in the Tucker format, but for groups of modes. Instances of such formats that are suitable for problems in very high dimensions are *hierarchical tensors*, also known as *tree tensor networks* in physics, and the special case of *tensor trains*, which are also known as *matrix product states*. For tensors that satisfy further constraints on additional rank parameters, representations in these tensor formats can achieve favourable scaling with respect to d. Note that there are two cases in which all formats considered below coincide: they reduce to standard representations of low-rank matrices for order $d = 2$, and at arbitrary order they contain elementary tensors as special cases.

In general, a tensor format defines a multilinear *representation mapping* τ that for certain fixed parameters (such as tensor order) maps a corresponding number N of representation component tensors $\mathbf{X}_1, \dots, \mathbf{X}_N$ to the represented tensor,

$$\mathbf{t} = \tau(\mathsf{X}), \quad \mathsf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N).$$

In the simplest case of low-rank matrices in $\mathbb{R}^{n_1 \times n_2}$ with factors $\mathbf{X}_1 \in \mathbb{R}^{n_1 \times r}$ and $\mathbf{X}_2 \in \mathbb{R}^{n_2 \times r}$, the usual representation mapping is

$$\tau(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{X}_1 \mathbf{X}_2^\top \tag{1.12}$$

for any rank parameter r. In the case of the Tucker format in d dimensions, the arguments of the representation mapping are $\mathbf{a} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$ and the matrices $\mathbf{U}^{(i)} \in \mathbb{R}^{n_i \times r_i}$ with columns $\mathbf{U}_k^{(i)}$, $k = 1, \dots, r_i$ as above for $i = 1, \dots, d$.

These representation mappings are not injective: for instance, for any given $\mathbf{X}_1$, $\mathbf{X}_2$ as in (1.12) and $\mathbf{G} \in \mathrm{GL}(r, \mathbb{R})$, we have $\tau(\mathbf{X}_1, \mathbf{X}_2) = \tau(\mathbf{X}_1 \mathbf{G}^{-1}, \mathbf{X}_2 \mathbf{G}^\top)$. A central feature of tensor formats is that they are closed under vector space operations, that is, under scaling and addition of tensors. For example, the addition of two low-rank matrices can again be written in the form (1.12) with summed rank parameters by a simple concatenation of components, as detailed in Section 2.2.

1.5. Low-rank approximations of functions

Low-rank tensors provide a natural tool for approximating functions on high-dimensional product domains. We exemplify this by the approximation of elements of $L_2(\square_d)$ by tensor product basis expansions.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Let $\{\varphi_\nu\}_{\nu\in\mathbb{N}}$ be an orthonormal basis of $L_2(0,1)$. Then, for any $f \in L_2(\square_d)$, we have the basis expansion

$$f = \sum_{(\nu_1,\dots,\nu_d)\in\mathbb{N}^d} \mathbf{f}[\nu_1,\dots,\nu_d] \bigotimes_{i=1}^{d} \varphi_{\nu_i},$$

where

$$\mathbf{f}[\nu_1,\dots,\nu_d] = \int_{\square_d} f \bigotimes_{i=1}^{d} \varphi_{\nu_i}\,\mathrm{d}x.$$

Note that the mapping $f \mapsto \mathbf{f}$ defined in this manner is an isometry between $L_2(\square_d)$ and $\ell_2(\mathbb{N}^d)$. Due to the tensor product structure of this mapping, if f has low-rank structure, it is inherited by the basis coefficients $\mathbf{f}$: in the simplest case, if $f = \bigotimes_{i=1}^{d} f_i$ with $f_i \in L_2(0,1)$ for $i = 1,\dots,d$, then

$$\mathbf{f} = \bigotimes_{i=1}^{d} \mathbf{f}_i, \quad \text{where } \mathbf{f}_i[\nu] = \int_0^1 f_i\varphi_\nu\,\mathrm{d}x,\ \nu\in\mathbb{N}.$$

This hints at the usefulness of considering tensor representations not only for tensors of finite mode sizes but also on sequence spaces defined on high-dimensional index sets such as $\ell_2(\mathbb{N}^d)$. In this setting, infinite basis expansions are covered in a natural way, and low-rank approximations can be studied independently of any particular choice of finitely many summands in the basis expansion.

The reduction to a problem on $\ell_2(\mathbb{N}^d)$ via a basis representation is also convenient for computational purposes, since standard algorithms for rank reduction in subspace tensor formats provide error estimates in this norm. However, an additional difficulty that is important for applications to partial differential equations is that for Sobolev spaces $H^s(\square_d)$, the isomorphism to $\ell_2(\mathbb{N}^d)$ provided by suitable bases is not of tensor product form, and thus does not preserve low-rank structures. As considered in detail in Section 4, this issue can be overcome by what amounts to a discretization-dependent low-rank approximation of this isomorphism.

Tensor approximations are of interest not only for problems of high dimensionality: although it is natural to assign tensor modes to independent variables in an approximation problem, entirely different parametrizations of functions by tensors can be useful for particular purposes. For instance, in the context of occupation number representations of symmetric or antisymmetric wavefunctions in quantum physics, each tensor mode corresponds to a lower-dimensional basis function. Another example is that of methods based on reinterpreting basis expansion coefficients of lower-dimensional functions as high-order tensors. In this case, by a suitable encoding of indices, tensor modes correspond to scales in the approximation problem. In particular, this leads to efficient generic approximations for problems with singularities. These examples are considered in detail in Section 3. In such a context, classical notions of dimensionality need no longer be the determining factor for the computational complexity of low-rank approximations.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

1.6. Low-rank methods as an instance of nonlinear approximation schemes

An important feature of low-rank tensor formats is that they are closed under addition, but addition of two tensor representations generally leads to an increase of the associated rank parameters. In the simplest case of matrices, the result of adding two matrices of ranks r_1 and r_2 may have rank up to $r_1 + r_2$. In low-rank methods, approximations of functions are thus not sought from linear spaces (spanned, for instance, by a fixed set of basis functions) but from sets with nonlinear structure, where the representation complexity of elements increases under vector space operations.

This difficulty arises in an analogous manner in *sparse* nonlinear approximations with dictionaries; see for instance the survey by DeVore (1998). Here, we choose a fixed dictionary of functions $\{\psi_\nu\}_{\nu\in\mathbb{N}}$, such as a Riesz basis or frame of the relevant function space H. For a given approximand $f \in H$ and error tolerance $\varepsilon > 0$, we choose a suitable finite subset $\Lambda_\varepsilon(f) \subset \mathbb{N}$ such that f has an approximation $f_\varepsilon \in \operatorname{span}\{\psi_\nu : \nu \in \Lambda_\varepsilon(f)\}$ with $\|f - f_\varepsilon\|_H \leq \varepsilon$. This notion of sparse approximation in some sense involves a lesser degree of nonlinearity than the low-rank approximations considered above, since f_ε is parametrized linearly by its expansion coefficients. However, we have a nonlinear dependence on f via the choice of non-zero entries of the coefficient sequence, and when adding sparse approximations of two different functions with n_1 and n_2 non-zero coefficients, respectively, the result may also have up to $n_1 + n_2$ non-zero coefficients. Hence, in contrast to approximations from fixed linear subspaces, the complexity of representations generally increases under vector space operations.

A natural benchmark for sparse approximations is provided by best n-term approximation, where in the above setting, for given $f \in H$ and for each budget of $n \in \mathbb{N}$ terms, we consider the error achievable by a best choice of $\Lambda_n(f) \subset \mathbb{N}$ with $\#\Lambda_n(f) \leq n$, that is, the best n-term errors

$$\min_{\substack{\Lambda_n(f)\subset\mathbb{N}\\ \#\Lambda_n(f)\leq n}} \min\{\|f - f_n\|_H : f_n \in \operatorname{span}\{\psi_\nu : \nu \in \Lambda_n(f)\}\}.$$

Ideally, numerical methods for computing sparse approximations should yield errors proportional to the best n-term errors when using n coefficients or (as a much stronger requirement) when using a multiple of n elementary operations. Such questions also arise in the context of low-rank methods when we need to identify suitable lower-dimensional basis expansions in each tensor mode. This combination of sparse and low-rank approximations is discussed in further detail in Sections 3.2 and 5.3.

Another important instance of nonlinear approximations is neural networks, where inputs are transformed by several layers of alternating linear mappings with entrywise nonlinear transformations such as $\mathrm{ReLU}(x) = \max\{x, 0\}$. Although neural networks can in principle approximate many classes of functions very efficiently, either matching or surpassing the performance of classical approximation

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

methods, finding such approximations as solutions of regression problems by machine learning methods can generally be quite a problematic task. In this regard, a distinctive feature of subspace tensor formats is that the generation of near-best low-rank approximations can be reduced to the computation of singular value decompositions, for which reliable numerical algorithms are available.

1.7. Basic algorithmic concepts

Standard methods for solving partial differential equations eventually amount to solving systems of equations for grid values or coefficients with respect to some basis. In the case of low-rank approximations, we are facing the additional difficulty that these basis coefficients are represented indirectly in terms of components of a tensor representation. Moreover, the rank parameters that are required for a certain approximation quality are typically not known *a priori*. Since the process of generating approximate solutions in low-rank format is not straightforward, the computational complexity of low-rank solvers is just as important as the basic approximability of solutions when considering the viability of such approximations.

With the exception of semi-analytical explicit low-rank approximations as in (1.8) that exist for problems with certain special structures (see Section 4.2), methods for computing low-rank approximations are generally iterative: starting from a certain initial approximation $\tau(\mathsf{X}^0)$ of a solution in low-rank format with representation X^0, each step from $\tau(\mathsf{X}^n)$ to $\tau(\mathsf{X}^{n+1})$ is performed by operating only on the representation components X^n.

Concerning the construction of such iterations, one can broadly distinguish two basic approaches. One common strategy is to realize standard iterative solvers, such as Krylov subspace methods, in tensor format. Perhaps the simplest example is Richardson iteration for solving systems of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ for positive definite $\mathbf{A}$. With starting vector $\mathbf{x}^0$ and a damping parameter $\omega > 0$, the iteration reads

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \omega(\mathbf{A}\mathbf{x}^n - \mathbf{b}). \tag{1.13}$$

To realize this iteration in low-rank form, the problem data $\mathbf{A}$ and $\mathbf{b}$ also need to be represented in a compatible format. For instance, let $d = 2$ with

$$\mathbf{x}^n = \sum_{k=1}^{r} \mathbf{v}_k^{(1)} \otimes \mathbf{v}_k^{(2)} \quad \text{and} \quad \mathbf{A} = \sum_{\ell=1}^{R} \mathbf{A}_\ell^{(1)} \otimes \mathbf{A}_\ell^{(2)}.$$

Then

$$\mathbf{A}\mathbf{x}^n = \sum_{\ell=1}^{R} \sum_{k=1}^{r} \mathbf{A}_\ell^{(1)} \mathbf{v}_k^{(1)} \otimes \mathbf{A}_\ell^{(2)} \mathbf{v}_k^{(2)}, \tag{1.14}$$

that is, we still have a representation in the same low-rank format, but with ranks multiplied by (at most) R. Except in special cases such as $R = 1$, without further

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

precautions, the rank parameters of the representations X^n of the iterates $\mathbf{x}^n$ in (1.13) will thus increase exponentially with respect to n.

The iteration (1.13) thus needs to be combined with a *rank reduction*: let $\mathsf{X}^{n+1} = \mathcal{F}(\mathsf{X}^n)$ be the iteration on the low-rank representations of the $\mathbf{x}^n$ defined by (1.13); we then replace this with $\mathsf{X}^{n+1} = (\mathcal{R} \circ \mathcal{F})(\mathsf{X}^n)$, where $\mathcal{R}$ is a suitable rank reduction procedure, such as a truncated singular value decomposition or one of its higher-order generalizations for subspace tensor formats. In devising criteria for performing this rank reduction, we need to balance computational costs against preserving the convergence of the iteration. When always truncating to some fixed rank parameter, the costs in each step are controlled, but the convergence of the iteration is difficult to guarantee; conversely, when truncating up to a prescribed error, it becomes a non-trivial problem to estimate the resulting ranks and hence the computational costs.

The second strategy is to exploit the particular structure of the tensor format and to directly optimize the components of the representation. Again assuming $\mathbf{A}$ to be positive definite, we can rewrite the original system of equations as a optimization problem: find a low-rank representation X such that

$$\frac{1}{2}\langle \mathbf{A}\tau(\mathsf{X}), \tau(\mathsf{X})\rangle - \langle \mathbf{b}, \tau(\mathsf{X})\rangle$$

is minimized. Variational formulations of this kind can be combined with alternating minimization procedures that operate on one or two components in X at a time. Solvers can also be constructed to use the Riemannian manifold structure of fixed-rank tensors.

There exist many different realizations of the two approaches described above, as well as methods that combine elements of both. An overview of some prominent examples is given in Section 5.1. In the remainder of Section 5 we then focus on methods that achieve guarantees on convergence and computational complexity by striking a balance between error reduction and controlled ranks.

1.8. Notes on the literature

There is now a large body of literature on low-rank tensor approximations. For a survey on classical tensor representations including canonical and Tucker formats, we refer to Kolda and Bader (2009). Tensors in the context of computational applications have also been considered in surveys by Grasedyck, Kressner and Tobler (2013), Hackbusch (2014), Khoromskij (2015), Bachmayr, Schneider and Uschmajew (2016) and Bachmayr and Dahmen (2020), as well as in the monographs by Hackbusch (2019) and Khoromskij (2018). The canonical tensor rank and its geometric and algebraic aspects are considered in a monograph by Landsberg (2012) and in a survey by Lim (2021).

In the present article, at the expense of developments that are aimed at other applications, such as machine learning, we consider low-rank methods specifically from the perspective of the numerical solution of partial differential equations.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Although we aim at a general overview of such methods, we focus on the use of subspace tensor formats in schemes with rigorous deterministic performance guarantees.

Concerning applications of tensor representations in signal processing and machine learning not covered by this article, we refer to the surveys by Signoretto, Tran Dinh, De Lathauwer and Suykens (2014), Cichocki *et al.* (2015), Sidiropoulos *et al.* (2017) and Janzamin *et al.* (2019), as well as to Stoudenmire and Schwab (2016), Chen, Batselier, Suykens and Wong (2017), Novikov, Trofimov and Oseledets (2018) and Michel and Nouy (2022). Concerning methods for solving PDEs using tensor representations with random sampling, which we do not consider in further detail, see for example Richter, Sallandt and Nüsken (2021), Oster, Sallandt and Schneider (2022) and Fackeldey, Oster, Sallandt and Schneider (2022).

2. Subspace tensor formats

In this section we consider in detail the particular class of low-rank representations of higher-order tensors that provide the foundation for all numerical methods that follow. We first introduce the basic notions of tensor products of abstract vector spaces. To speak about the asymptotic low-rank approximability of functions and construct numerical methods with discretization-independent performance, we need to consider tensor products of infinite-dimensional spaces, where we restrict ourselves to Hilbert spaces.

2.1. Tensor product Hilbert spaces

For continuous functions f and g, in analogy to the outer product of vectors, we have $(f \otimes g)(x, y) = f(x)g(y)$ for x and y from the respective domains. However, for Hilbert spaces such as L_2-spaces for which point evaluation is not defined, a more generally applicable construction of tensor products is required. Here we first define the *algebraic tensor product* of vector spaces, which contains all finite linear combinations of elementary tensors, and subsequently consider closures in suitable norms to obtain *topological tensor products*. Treatments of the basic notions considered in this section can also be found in Light and Cheney (1985), Hackbusch (2019) and Khoromskij (2018). For the more general case of Banach spaces, we refer to Defant and Floret (1993), Ryan (2002) and Hackbusch (2019).

For arbitrary vector spaces V, W, the algebraic tensor product $V \otimes_a W$ provides an abstract notion of finite sums of elementary tensor products of the form

$$\sum_{k=1}^{r} \lambda_k \, v_k \otimes w_k, \tag{2.1}$$

where $r \in \mathbb{N}$ and $\lambda_k \in \mathbb{R}$, $v_k \in V$, $w_k \in W$ for $k = 1, \dots, r$. To this end, we define

$$T_a(V, W) = \bigcup_{n=1}^{\infty} (\mathbb{R} \times V \times W)^n$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

with the natural addition operation by concatenation, and we identify expressions

$$t = ((\lambda_k, v_k, w_k))_{k=1,\dots,r} \in T_a(V, W) \tag{2.2}$$

with functionals on all bilinear forms b on $V \times W$ by

$$t(b) = \sum_{k=1}^{r} \lambda_k b(v_k, w_k).$$

We then define an equivalence relation $\sim_a$ on $T_a(V, W)$ by treating two expressions $t, \tilde{t} \in T_a(V, W)$ as equivalent if and only if they give rise to the same such functional, that is, if $t(b) = \tilde{t}(b)$ for all bilinear forms b on $V \times W$.

Definition 2.1. The *algebraic tensor product* $V \otimes_a W$ of two vector spaces is defined as the quotient vector space $T_a(V, W)/\sim_a$.

For the equivalence class in $V \otimes_a W$ represented by t as in (2.2), we use the notation (2.1). We have the following basic properties; see Light and Cheney (1985, Sec. 1) and Hackbusch (2019, Sec. 3.2).

Proposition 2.2.

(i) For $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, $v_1, v_2 \in V$ and $w_1, w_2 \in W$,

$$(\alpha v_1 + \beta v_2) \otimes (\gamma w_1 + \delta w_2) = \alpha\gamma v_1 \otimes w_1 + \alpha\delta v_1 \otimes w_2 + \beta\gamma v_2 \otimes w_1 + \beta\delta v_2 \otimes w_2.$$

(ii) If $\{v_1, \dots, v_{r_1}\} \subset V$ and $\{w_1, \dots, w_{r_2}\} \subset W$ with $r_1, r_2 \in \mathbb{N}$ are linearly independent, then $\{v_i \otimes w_j : i = 1, \dots, r_1, j = 1, \dots, r_2\} \subset V \otimes_a W$ is linearly independent.

The construction of algebraic tensor products can be iterated and is associative up to isomorphism, that is, $U \otimes_a (V \otimes_a W) \simeq (U \otimes_a V) \otimes_a W$ for vector spaces U, V, W; we thus simply write $U \otimes_a V \otimes_a W$.

Topological tensor products are obtained by taking closures of algebraic tensor products in specific norms. For the purposes of this article, it will be sufficient to consider topological tensor products of Hilbert spaces. In contrast to the case of more general Banach spaces, we then have a single canonical choice of tensor product norm. Let $H_1, \dots, H_d$ be Hilbert spaces. For elementary tensors $v = \bigotimes_{i=1}^{d} v_i$, $w = \bigotimes_{i=1}^{d} w_i \in H_1 \otimes_a \cdots \otimes_a H_d$, we define the inner product

$$\langle v, w \rangle_H = \prod_{i=1}^{d} \langle v_i, w_i \rangle_{H_i}, \tag{2.3}$$

extend this to sums of elementary tensors by bilinearity, and define the tensor product Hilbert space $H = H_1 \otimes \cdots \otimes H_d$ as the closure of $H_1 \otimes_a \cdots \otimes_a H_d$ with respect to the induced norm. By definition, the norm induced by (2.3) is a

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

cross-norm, that is,

$$\|v_1 \otimes \cdots \otimes v_d\|_H = \prod_{i=1}^{d} \|v_i\|_{H_i} \quad \text{for } v_i \in H_i,\ i = 1, \dots, d. \tag{2.4}$$

Unless otherwise stated, for tensor product Hilbert spaces, we assume this canonical choice of inner product (2.3) and induced norm in what follows.

For tensor product Hilbert spaces $H = \bigotimes_{i=1}^{d} H_i$ and $G = \bigotimes_{i=1}^{d} G_i$ and bounded linear operators $A_i \colon H_i \to G_i$ for $i = 1, \dots, d$, we define the *tensor product operator* $A_1 \otimes \cdots \otimes A_d$ by its action on elementary tensor products as

$$(A_1 \otimes \cdots \otimes A_d)(v_1 \otimes \cdots \otimes v_d) = (A_1 v_1) \otimes \cdots \otimes (A_d v_d). \tag{2.5}$$

This definition is uniquely extended to finite sums of elementary tensors by linearity and to infinite sums by the unique extension of bounded operators defined on dense subsets. The following identity can be obtained immediately from Light and Cheney (1985, Lem. 1.30).

Theorem 2.3. For $A_1, \dots, A_d$ as in (2.5),

$$\|A_1 \otimes \cdots \otimes A_d\|_{H \to G} = \prod_{i=1}^{d} \|A_i\|_{H_i \to G_i}.$$

When each of the mappings $A_1, \dots, A_d$ in (2.5) is an isomorphism, the corresponding tensor product operator $A_1 \otimes \cdots \otimes A_d$ is called a *tensor space isomorphism*, and we have $(A_1 \otimes \cdots \otimes A_d)^{-1} = A_1^{-1} \otimes \cdots \otimes A_d^{-1}$.

Remark 2.4. For countable sets $\mathcal{I}_1, \mathcal{I}_2$, we have $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2) = \ell_2(\mathcal{I}_1) \otimes \ell_2(\mathcal{I}_2)$. In other words, the norm $\|\cdot\|_{\ell_2(\mathcal{I}_1) \otimes \ell_2(\mathcal{I}_2)}$ defined by (2.4) coincides with the norm of $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$ and $\operatorname{span}\{\mathbf{v}_1 \otimes \mathbf{v}_2 \colon \mathbf{v}_1 \in \ell_2(\mathcal{I}_1), \mathbf{v}_2 \in \ell_2(\mathcal{I}_2)\}$ is dense in $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$. Similarly, for measurable sets Ω_1, Ω_2, we have $L_2(\Omega_1 \times \Omega_2) = L_2(\Omega_1) \otimes L_2(\Omega_2)$; for a proof, see Light and Cheney (1985, Thm 1.39).

It will suffice to restrict ourselves to *separable* Hilbert spaces in what follows. Any such space, by any choice of orthonormal basis, is isometrically isomorphic to $\ell_2(\mathbb{N})$. In general, such an isomorphism does not respect tensor product structures. These are preserved, however, in the case of tensor product orthonormal bases, as formulated in the following theorem; for a proof, see Weidmann (1980, Thm 3.12).

Theorem 2.5. Let $H_1, \dots, H_d$ be separable Hilbert spaces, and for $i = 1, \dots, d$, let $\{\varphi_\nu^{(i)}\}_{\nu \in \mathbb{N}}$ be an orthonormal basis of H_i. Then

$$\left\{ \bigotimes_{i=1}^{d} \varphi_{\nu_i}^{(i)} \right\}_{(\nu_1, \dots, \nu_d) \in \mathbb{N}^d}$$

is an orthonormal basis of $\bigotimes_{i=1}^{d} H_i$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Remark 2.6. With the basis isomorphisms

$$F_i \colon \ell_2(\mathbb{N}) \to H_i, \quad \mathbf{v} \mapsto \sum_{\nu \in \mathbb{N}} \mathbf{v}[\nu]\, \varphi_\nu^{(i)} ,$$

as a consequence of the cross-norm property (2.4), the isomorphism associated to the tensor product basis

$$F \colon \ell_2(\mathbb{N}^d) \to \bigotimes_{i=1}^{d} H_i, \quad \mathbf{v} \mapsto \sum_{\nu \in \mathbb{N}^d} \mathbf{v}[\nu] \bigotimes_{i=1}^{d} \varphi_{\nu_i}^{(i)}$$

can be written in tensor product form as $F = \bigotimes_{i=1}^{d} F_i$ and is thus a tensor space isomorphism.

For countable sets $\mathcal{I}_1, \mathcal{I}_2$, elements of $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2) = \ell_2(\mathcal{I}_1) \otimes \ell_2(\mathcal{I}_2)$ can be identified with bounded linear operators from $\ell_2(\mathcal{I}_2)$ to $\ell_2(\mathcal{I}_1)$ by interpreting their entries as those of an infinite matrix: the operator associated to $\mathbf{u} \in \ell_2(\mathcal{I}_1) \otimes \ell_2(\mathcal{I}_2)$ is given by

$$\ell_2(\mathcal{I}_2) \ni \mathbf{v} \mapsto \left(\sum_{\nu \in \mathcal{I}_2} \mathbf{u}[\nu', \nu] \mathbf{v}[\nu] \right)_{\nu' \in \mathcal{I}_1}, \tag{2.6}$$

where the right-hand side is in $\ell_2(\mathcal{I}_1)$ as a consequence of the Cauchy–Schwarz inequality.

Definition 2.7. For countable sets $\mathcal{I}_1, \mathcal{I}_2$ and $\mathbf{u} \in \ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$, we define

$$\operatorname{rank}(\mathbf{u}) = \dim \operatorname{range}(\mathbf{u}) \in \mathbb{N}_0 \cup \{\infty\},$$

where we identify $\mathbf{u}$ with the operator given by (2.6).

For $\mathbf{u} \in \ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$, the operator defined by (2.6) belongs to the Hilbert–Schmidt class, where the Hilbert–Schmidt inner product is simply the inner product in $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$. In particular, Hilbert–Schmidt operators are compact. As a consequence, we obtain the following adaptation of a classical result that goes back to Schmidt (1907).

Theorem 2.8 (Hilbert–Schmidt decomposition). For each $\mathbf{u} \in \ell_2(\mathbb{N}^2)$, there exist orthonormal systems $\{\mathbf{U}_k^{(1)}\}_{k \in \mathbb{N}}$ and $\{\mathbf{U}_k^{(2)}\}_{k \in \mathbb{N}}$ in $\ell_2(\mathbb{N})$ as well as a non-increasing, non-negative sequence $(\sigma_k)_{k \in \mathbb{N}} \in \ell_2(\mathbb{N})$ such that

$$\mathbf{u} = \sum_{k=1}^{\infty} \sigma_k \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)} \tag{2.7}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

with convergence in $\ell_2(\mathbb{N}^2)$. Moreover, for any $r \in \mathbb{N}$, we have the best approximation property

$$\left\| \mathbf{u} - \sum_{k=1}^{r} \sigma_k \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)} \right\|_{\ell_2(\mathbb{N}^2)} = \left(\sum_{k=r+1}^{\infty} \sigma_k^2 \right)^{1/2} = \min\{ \|\mathbf{u} - \mathbf{v}\|_{\ell_2(\mathbb{N}^2)} : \operatorname{rank}(\mathbf{v}) \le r \}, \tag{2.8}$$

and the best approximation is unique precisely when $\sigma_r > \sigma_{r+1}$.

The expansion (2.7) is also called the *singular value decomposition* (SVD) of the operator induced by $\mathbf{u}$. In this context, the expansion is identified with the mapping on $\ell_2(\mathbb{N})$ given by

$$\sum_{k=1}^{\infty} \sigma_k \langle \mathbf{U}_k^{(2)}, \cdot \rangle \mathbf{U}_k^{(1)}.$$

The orthonormal systems $\{\mathbf{U}_k^{(1)}\}_{k\in\mathbb{N}}$ and $\{\mathbf{U}_k^{(2)}\}_{k\in\mathbb{N}}$ are called the *left* and *right singular vectors*, respectively, and $(\sigma_k)_{k\in\mathbb{N}}$ is the sequence of *singular values*.

Remark 2.9. Let us note the following immediate consequences of Theorem 2.8.

(i) The same statement applies to arbitrary countable index sets $\mathcal{I}_1, \mathcal{I}_2$, in which case it yields orthonormal systems in $\ell_2(\mathcal{I}_1)$ and $\ell_2(\mathcal{I}_2)$. In particular, in this section we also use it with $\mathcal{I}_1 = \mathbb{N}^{d_1}$ and $\mathcal{I}_2 = \mathbb{N}^{d_2}$ with $d_1, d_2 \in \mathbb{N}$.

(ii) When $\operatorname{supp} \mathbf{u} \subseteq \mathcal{I}_1 \times \mathcal{I}_2$ with *finite* $\mathcal{I}_1, \mathcal{I}_2$, which corresponds to the case of the singular value decomposition of a matrix in $\mathbb{R}^{\mathcal{I}_1 \times \mathcal{I}_2}$ representing a linear mapping from $\mathbb{R}^{\mathcal{I}_2}$ to $\mathbb{R}^{\mathcal{I}_1}$, the sum in (2.7) has at most $\min\{\#\mathcal{I}_1, \#\mathcal{I}_2\}$ summands. In this case Theorem 2.8 is also known as the Eckart–Young theorem (Eckart and Young 1936).

2.2. *Operations on low-rank matrices*

Before turning to higher-order tensors, it may be instructive to first take a closer look at low-rank matrices from the computational point of view, that is, at the procedures for performing typical operations such as addition or singular value decomposition in low-rank form.

There are different possibilities for representing matrices in low-rank form that leave different types of redundancy in the representation. As mentioned in Section 1.4, matrices of rank at most r in $\mathbb{R}^{m\times n}$ can be represented by $\mathbf{A} \in \mathbb{R}^{m\times r}$ and $\mathbf{B} \in \mathbb{R}^{n\times r}$ via the representation mapping

$$\tau(\mathbf{A}, \mathbf{B}) = \mathbf{A}\mathbf{B}^\top,$$

where $\tau(\mathbf{A}, \mathbf{B}) = \tau(\mathbf{A}\mathbf{G}^{-1}, \mathbf{B}\mathbf{G}^\top)$ for any $\mathbf{G} \in \mathrm{GL}(r, \mathbb{R})$. In addition, we can demand that either $\mathbf{A}$ or $\mathbf{B}$ have orthonormal columns, in which case $\tau(\mathbf{A}, \mathbf{B}) = \tau(\mathbf{A}\mathbf{G}, \mathbf{B}\mathbf{G})$ for any $\mathbf{G} \in \mathrm{O}(r)$. An alternative representation is via

$$\tau(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \mathbf{U}\mathbf{S}\mathbf{V}^\top, \tag{2.9}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

where $\mathbf{U}, \mathbf{V}$ are required to have orthonormal columns and $\mathbf{S} \in \mathbb{R}^{r\times r}$. In this case $\tau(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \tau(\mathbf{U}\mathbf{G}_1, \mathbf{G}_1^\top \mathbf{S}\mathbf{G}_2, \mathbf{V}\mathbf{G}_2)$ for all $\mathbf{G}_1, \mathbf{G}_2 \in \mathrm{O}(r)$. Here $\mathbf{S}$ is diagonal precisely when the columns of $\mathbf{U}$ and $\mathbf{V}$ are left and right singular vectors of $\tau(\mathbf{U}, \mathbf{S}, \mathbf{V})$, respectively, in which case the diagonal entries of $\mathbf{S}$ are (up to sign) the corresponding singular values.

In the following description of operations, we focus on representations of the form (2.9) with $\mathbf{S}$ not necessarily diagonal, since this case is conceptually closest to higher-order tensor representations.

2.2.1. *Orthogonalization*

Let $\tilde{\mathbf{U}} \in \mathbb{R}^{m\times r}$, $\tilde{\mathbf{V}} \in \mathbb{R}^{n\times r}$ and $\tilde{\mathbf{S}} \in \mathbb{R}^{r\times r}$ be given. Then a representation satisfying the orthogonality requirement can be found as follows. Perform QR decompositions $\tilde{\mathbf{U}} = \mathbf{U}\mathbf{R}_1$, $\tilde{\mathbf{V}} = \mathbf{V}\mathbf{R}_2$ to obtain $\mathbf{U} \in \mathbb{R}^{m\times r}$ and $\mathbf{V} \in \mathbb{R}^{n\times r}$ with orthonormal columns, as well as (triangular) $\mathbf{R}_1, \mathbf{R}_2 \in \mathbb{R}^{r\times r}$; then set $\mathbf{S} = \mathbf{R}_1 \tilde{\mathbf{S}} \mathbf{R}_2^\top$ to obtain $\tau(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \tau(\tilde{\mathbf{U}}, \tilde{\mathbf{S}}, \tilde{\mathbf{V}})$, where $\mathbf{U}, \mathbf{V}$ have orthonormal columns. Note that in this form the Hilbert–Schmidt and spectral norms of $\tau(\mathbf{U}, \mathbf{S}, \mathbf{V})$ can be obtained as the respective norms of the matrix $\mathbf{S}$, which in this case is generally not diagonal.

Remark 2.10. To perform orthogonalization as outlined above, performing the QR decompositions requires $O((m + n)r^2)$ operations, and computing $\mathbf{S}$ in addition requires $O(r^3)$ operations. These QR decompositions can be computed by a numerically stable direct method, for instance by applying at most r orthogonal Householder reflectors.

2.2.2. *Addition of low-rank matrices*

For $\lambda_1, \lambda_2 \in \mathbb{R}$ and two given representations $\tau(\mathbf{U}_1, \mathbf{S}_1, \mathbf{V}_1)$ and $\tau(\mathbf{U}_2, \mathbf{S}_2, \mathbf{V}_2)$ of compatible sizes with ranks r_1 and r_2, respectively, to obtain a low-rank representation of $\lambda_1\tau(\mathbf{U}_1, \mathbf{S}_1, \mathbf{V}_1) + \lambda_2\tau(\mathbf{U}_2, \mathbf{S}_2, \mathbf{V}_2)$ we assemble the block matrices

$$\tilde{\mathbf{U}} = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{bmatrix}, \quad \tilde{\mathbf{S}} = \begin{bmatrix}\lambda_1\mathbf{S}_1 & \\ & \lambda_2\mathbf{S}_2\end{bmatrix}, \quad \tilde{\mathbf{V}} = \begin{bmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{bmatrix}$$

corresponding to a representation with rank at most $r_1 + r_2$, and subsequently orthogonalize with results $\mathbf{U}, \mathbf{S}, \mathbf{V}$ as above to arrive at $\tau(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \lambda_1\tau(\mathbf{U}_1, \mathbf{S}_1, \mathbf{V}_1) + \lambda_2\tau(\mathbf{U}_2, \mathbf{S}_2, \mathbf{V}_2)$.

2.2.3. *Singular value decomposition*

To obtain the SVD form of a given representation $\tau(\mathbf{U}, \mathbf{S}, \mathbf{V})$ with $\mathbf{U}, \mathbf{V}$ having orthonormal columns, we compute an SVD of the matrix $\mathbf{S}$,

$$\mathbf{S} = \mathbf{Q}\mathbf{\Sigma}\mathbf{P}^\top,$$

where $\mathbf{Q}, \mathbf{P} \in \mathrm{O}(r)$ and $\mathbf{\Sigma}$ is diagonal. Then, as summarized for later use in the following simple observation, the sought SVD form is given by $\tau(\mathbf{U}\mathbf{Q}, \mathbf{\Sigma}, \mathbf{V}\mathbf{P})$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Proposition 2.11. Let $\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$, where $\mathbf{U} \in \mathbb{R}^{m\times r}$ and $\mathbf{V} \in \mathbb{R}^{n\times r}$ have orthonormal columns and $\mathbf{S} \in \mathbb{R}^{r\times r}$. Let $\mathbf{S} = \mathbf{Q}\mathbf{\Sigma}\mathbf{P}^\top$ be a singular value decomposition of $\mathbf{S}$; then $\mathbf{M} = \hat{\mathbf{U}}\mathbf{\Sigma}\hat{\mathbf{V}}^\top$ with $\hat{\mathbf{U}} = \mathbf{U}\mathbf{Q}$, $\hat{\mathbf{V}} = \mathbf{V}\mathbf{P}$ is a singular value decomposition of $\mathbf{M}$.

Remark 2.12 (computation of the SVD). Truncation of the SVD yields solutions of best low-rank approximation problems by Theorem 2.8, and the computation of SVDs as outlined above is thus of central importance. Note that the SVD of a rectangular matrix $\mathbf{A}$ can be obtained from the eigendecompositions of the symmetric matrices $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$, which necessarily require an iterative scheme. These latter eigenvalue problems can be solved by transformation to tridiagonal form by a direct method, followed by a QR iteration for the resulting symmetric tridiagonal eigenvalue problems that has been shown by Wilkinson (1968) to be globally at least linearly convergent; see also Parlett (1998, Sec. 8.10). This implies that we obtain an algorithm for approximating the SVD with guaranteed convergence and complexity bounds.

When performed with machine precision, the approach via symmetric eigenvalue problems has the disadvantage that a loss of relative precision can occur for the smallest singular values. This is avoided by the specialized numerical SVD methods of Golub and Kahan (1965) and Golub and Reinsch (1970). There are also variants of the SVD algorithm with potentially improved quantitative performance based on divide-and-conquer strategies (Gu and Eisenstat 1995), but these can occasionally be observed to fail to converge in practical computations.

For each fixed relative precision, the work required for the SVD of an $r\times r$-matrix scales as $O(r^3)$, and thus the above procedure for the SVD of low-rank matrices requires

$$O\left(r^3 + (m+n)r^2\right)$$

operations in total, which is of the same order as the costs of orthogonalization, despite the different type of numerical algorithms.

2.3. Matricizations of tensors and product notation

We now introduce basic notions that play a role in the definitions of the different tensor formats that we consider below. We consider these for tensors of order d with countable index sets, that is, for tensors in $\ell_2(\mathbb{N}^d)$. Considering tensors in infinite-dimensional spaces is essential for discretization-independent notions of low-rank approximability of functions and for formulating algorithms that are robust under discretization refinement. Note that the corresponding definitions contain tensors of finite size as special cases with support on finite subsets of $\mathbb{N}^d$, and practical computations are carried out on such finite subsets as described above for the matrix case.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

In what follows, for the set of all modes of a tensor of order d, we use the abbreviation

$$\alpha^* = \{1, \dots, d\}.$$

For a multi-index $\nu \in \mathbb{N}^d$ and $\alpha \subseteq \alpha^*$, we write $\nu_\alpha = (\nu_i)_{i\in\alpha}$ for its sub-index for the modes in α and set

$$\alpha^{\mathsf{c}} = \alpha^* \setminus \alpha.$$

Definition 2.13. For non-empty $\alpha \subseteq \alpha^*$, the *$\alpha$-matricization* of $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ (also referred to as *unfolding* or *flattening* of $\mathbf{v}$) is given by

$$\operatorname{mat}_\alpha(\mathbf{v}) = (\mathbf{v}[\nu])_{\nu_\alpha \in \mathbb{N}^{\#\alpha},\, \nu_{\alpha^{\mathsf{c}}} \in \mathbb{N}^{d-\#\alpha}} \in \ell_2(\mathbb{N}^{\#\alpha}) \otimes \ell_2(\mathbb{N}^{d-\#\alpha}). \tag{2.10}$$

For matricizations with respect to single coordinates $i = 1, \dots, d$, we also write $\operatorname{mat}_i(\mathbf{v})$ in place of $\operatorname{mat}_{\{i\}}(\mathbf{v})$. The associated *$\alpha$-rank* of $\mathbf{v}$ is defined by

$$\operatorname{rank}_\alpha(\mathbf{v}) = \operatorname{rank} \operatorname{mat}_\alpha(\mathbf{v}) \in \mathbb{N}_0 \cup \{\infty\}$$

with the abbreviation $\operatorname{rank}_i(\mathbf{v}) = \operatorname{rank}_{\{i\}}(\mathbf{v})$ for $i = 1, \dots, d$.

For $\alpha \subseteq \alpha^*$, the bijective linear mapping $\operatorname{mat}_\alpha$ provides an identification of tensors in $\ell_2(\mathbb{N}^d)$ with infinite matrices, where indices of modes in α are used as row indices and the remaining ones as column indices. As in (2.6), these infinite matrices $\operatorname{mat}_\alpha(\mathbf{v})$ are identified with linear operators $\operatorname{mat}_\alpha(\mathbf{v})\colon \ell_2(\mathbb{N}^{d-\#\alpha}) \to \ell_2(\mathbb{N}^{\#\alpha})$. Note that $\operatorname{mat}_\alpha(\mathbf{v})^\top = \operatorname{mat}_{\alpha^{\mathsf{c}}}(\mathbf{v})$ and thus

$$\operatorname{rank}_{\alpha^{\mathsf{c}}}(\mathbf{v}) = \operatorname{rank} \operatorname{mat}_{\alpha^{\mathsf{c}}}(\mathbf{v}) = \operatorname{rank} \operatorname{mat}_\alpha(\mathbf{v})^\top = \operatorname{rank} \operatorname{mat}_\alpha(\mathbf{v}) = \operatorname{rank}_\alpha(\mathbf{v}).$$

Remark 2.14. For general separable Hilbert spaces H_i, $i = 1, \dots, d$, an analogous notion of matricizations of elements of $H = \bigotimes_{i=1}^d H_i$ can be introduced; the α-matricization is then an element of $\bigotimes_{i\in\alpha} H_i \otimes \bigotimes_{i\in\alpha^{\mathsf{c}}} H_i$, identified with an operator from $\bigotimes_{i\in\alpha} H_i$ to $\bigotimes_{i\in\alpha^{\mathsf{c}}} H_i$. The corresponding α-ranks are invariant under tensor space isomorphisms. In view of Remark 2.6, by choosing a suitable orthonormal basis for each space H_i, it thus suffices to consider tensors in $\ell_2(\mathbb{N}^d) = \bigotimes_{i=1}^d \ell_2(\mathbb{N})$.

Definition 2.15. Let $\mathbf{t} \in \ell_2(\mathbb{N}^d)$, let $\mathbf{U}\colon \ell_2(\mathbb{N}) \to \ell_2(\mathbb{N})$ be linear and bounded, and let $i \in \{1, \dots, d\}$. The (ith) *mode product* of $\mathbf{t}$ and $\mathbf{U}$ is defined by

$$\mathbf{t} \times_i \mathbf{U} = \operatorname{mat}_i^{-1}(\mathbf{U} \operatorname{mat}_i(\mathbf{t})),$$

or entrywise, with the identification (2.6),

$$(\mathbf{t} \times_i \mathbf{U})[\nu_1, \dots, \nu_d] = \sum_{\nu_i' \in \mathbb{N}} \mathbf{U}[\nu_i, \nu_i']\mathbf{t}[\nu_1, \dots, \nu_{i-1}, \nu_i', \nu_{i+1}, \dots, \nu_d]$$

for all $\nu_1, \dots, \nu_d \in \mathbb{N}$. In addition, for non-empty $\alpha \subset \alpha^*$ and $\mathbf{W}\colon \ell_2(\mathbb{N}^{\#\alpha}) \to \ell_2(\mathbb{N}^{\#\alpha})$, we introduce the α-mode product

$$\mathbf{t} \times_\alpha \mathbf{W} = \operatorname{mat}_\alpha^{-1}(\mathbf{W} \operatorname{mat}_\alpha(\mathbf{t}))$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

acting on the modes in α.

2.4. *Tucker format*

We have already briefly considered the Tucker tensor format in Section 1.4. In the general definition of the multilinear rank of a tensor in $\ell_2(\mathbb{N}^d)$, for $i = 1, \dots, d$, we rely on mode-wise matricizations

$$\mathrm{mat}_i(\mathbf{v}) = (\mathbf{v}[\nu])_{\nu_i \in \mathbb{N}, (\nu_1, \dots, \nu_{i-1}, \nu_{i+1}, \dots, \nu_d) \in \mathbb{N}^{d-1}} \tag{2.11}$$

as special cases of Definition 2.13.

Definition 2.16. The *multilinear rank* of a tensor $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ is the tuple

$$\mathrm{rank}_{\mathrm{ML}}(\mathbf{v}) = (\mathrm{rank}_i(\mathbf{v}))_{i=1,\dots,d} \in (\mathbb{N}_0 \cup \{\infty\})^d.$$

For each $i \in \{1, \dots, d\}$, since $\mathrm{rank}_i(\mathbf{v}) = \dim \mathrm{range}\, \mathrm{mat}_i(\mathbf{v})$, we can choose an orthonormal system $\{\mathbf{U}_k^{(i)}\}_{k=1,\dots,\mathrm{rank}_i(\mathbf{v})}$ in $\ell_2(\mathbb{N})$ that is an orthonormal basis of $\overline{\mathrm{range}\, \mathrm{mat}_i(\mathbf{v})}$, referred to as *orthonormal mode frame*. Let $\mathbf{U}^{(i)}$ be the matrix with columns $\mathbf{U}_k^{(i)}$. In terms of the mode product, the representation in Tucker format of $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ generalizing (1.11) can then be written in the form

$$\mathbf{v} = \mathbf{a} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \cdots \times_d \mathbf{U}^{(d)}, \tag{2.12}$$

where the core tensor $\mathbf{a}$ of size $\mathrm{rank}_1(\mathbf{v}) \times \cdots \times \mathrm{rank}_d(\mathbf{v})$ satisfies

$$\mathbf{a}[k_1, \dots, k_d] = \left\langle \mathbf{v}, \mathbf{U}_{k_1}^{(1)} \otimes \cdots \otimes \mathbf{U}_{k_d}^{(d)} \right\rangle.$$

Note that since an isometric isomorphism on ℓ_2 is applied in each mode, we have

$$\|\mathbf{v}\|_{\ell_2} = \|\mathbf{a}\|_{\ell_2}. \tag{2.13}$$

Remark 2.17 (storage costs). Let $\mathrm{rank}_{\mathrm{ML}}(\mathbf{v}) = (r_i)_{i=1,\dots,d}$ with $r_i \in \mathbb{N}$, $i = 1, \dots, d$, and $\mathrm{supp}\, \mathbf{v} \subseteq \{1, \dots, n_1\} \times \cdots \times \{1, \dots, n_d\}$. Then the number of storage locations used for the corresponding representation of $\mathbf{v}$ in Tucker format is

$$\prod_{i=1}^{d} r_i + \sum_{i=1}^{d} r_i n_i.$$

The multilinear rank refers to the minimum number of basis vectors required in each mode frame. Tensors can also be given in the form (2.12) in a potentially redundant representation with non-orthogonal mode frames. In this case we refer to the corresponding rank parameters (that is, the mode sizes of the core tensor) as *representation ranks*.

Remark 2.18 (basic operations). For the operations considered in Section 2.2 for low-rank matrices, we have the following analogous results for the Tucker format.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

(i) *Addition of tensors* in Tucker format can be performed analogously to the matrix case described in Section 2.2.2 by concatenation of mode frames and diagonal stacking of core tensors, so that the multilinear rank of the sum is at most the componentwise sum of multilinear ranks.

(ii) *Orthogonalization.* Given $\tilde{\mathbf{a}} \times_1 \tilde{\mathbf{U}}^{(1)} \cdots \times_d \tilde{\mathbf{U}}^{(d)}$, where $\tilde{\mathbf{U}}^{(1)}, \dots, \tilde{\mathbf{U}}^{(d)}$ have non-orthogonal columns, one can orthonormalize in a similar way to the matrix case in Section 2.2.1: perform a QR decomposition $\tilde{\mathbf{U}}^{(1)} = \mathbf{U}^{(1)}\mathbf{R}_1$ and replace the core tensor by $\tilde{\mathbf{a}} \times_1 \mathbf{R}_1$; carry out the same steps for $i = 2, \dots, d$. The total computational costs scale as

$$O\left(\sum_{i=1}^{d} r_i^2 \prod_{s \neq i} r_s + \sum_{i=1}^{d} r_i^2 n_i\right).$$

(iii) *Higher-order singular value decomposition.* In addition, the mode frames can be chosen to be left singular vectors of the respective matricizations. The resulting representation has been termed higher-order singular value decomposition (HOSVD) in De Lathauwer, De Moor and Vandewalle (2000); its properties, which to some extent parallel those of the SVD of matrices, are considered in Section 2.7. To compute the HOSVD numerically, assume that a representation $\mathbf{v} = \mathbf{a} \times_1 \mathbf{U}^{(1)} \cdots \times_d \mathbf{U}^{(d)}$ with orthonormal mode frames is given. Perform a QR factorization $\mathrm{mat}_1(\mathbf{a})^\top = \mathbf{V}_1\mathbf{R}_1$ with $\mathbf{V}_1$ having orthonormal columns and $\mathbf{R}_1 \in \mathbb{R}^{r_1 \times r_1}$, and compute the SVD $\mathbf{R}_1^\top = \mathbf{Q}\mathbf{\Sigma}\mathbf{P}^\top$. By Proposition 2.11, the columns of $\hat{\mathbf{U}}^{(1)} = \mathbf{U}^{(1)}\mathbf{Q}$ are then left singular vectors of $\mathrm{mat}_i(\mathbf{v})$. After replacing the core tensor by $\mathbf{a} \times_1 \mathbf{Q}^\top$, the same can be carried out for the further tensor modes, which eventually yields a decomposition $\hat{\mathbf{a}} \times_1 \hat{\mathbf{U}}^{(1)} \cdots \times_d \hat{\mathbf{U}}^{(d)}$, where the columns of each $\hat{\mathbf{U}}^{(i)}$ are left singular vectors of the respective matricization. In view of Remark 2.12, the total number of operations for this procedure scales as

$$O\left(\sum_{i=1}^{d} r_i^2 \prod_{s \neq i} r_s + \sum_{i=1}^{d} r_i^2 (n_i + r_i)\right).$$

Note that for $d > 2$, in contrast to the classical SVD, $\hat{\mathbf{a}}$ is in general a fully populated tensor without diagonal structure.

In summary, the Tucker format can be used to express a given tensor, separately in each mode, with respect to adapted basis vectors given by the mode frames. However, the number of entries of the core tensor still scales exponentially with respect to the number of modes.

2.5. *Tensor trains or matrix product states*

The *tensor train* (TT) format, introduced in Oseledets (2009*b*), Oseledets and Tyrtyshnikov (2009*a*,*b*) and Oseledets (2011*b*), is based on a different set of matricizations such that an application to very high tensor orders becomes feasible. In

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

quantum physics these tensor representations are known as *matrix product states*. In this context they were introduced in Vidal (2003) for the purpose of approximating wavefunctions; for an overview of corresponding applications, see Schollwöck (2011).

Definition 2.19. The *tensor train rank* of a tensor $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ is given by the vector

$$\operatorname{rank}_{\mathrm{TT}}(\mathbf{v}) = (\operatorname{rank}_{\{1,\dots,i\}}(\mathbf{v}))_{i=1,\dots,d-1}.$$

We now again deduce from this notion of rank a corresponding tensor representation. To this end, we use that for $\mathbf{v} \in \ell_2(\mathbb{N}^d)$,

$$\overline{\operatorname{range} \operatorname{mat}_{\{1,\dots,i-1,i\}}(\mathbf{v})} \subseteq \overline{\operatorname{range} \operatorname{mat}_{\{1,\dots,i-1\}}(\mathbf{v})} \otimes \ell_2(\mathbb{N}). \tag{2.14}$$

This can be deduced from Hackbusch (2019, Thm 6.31).

To simplify notation, we abbreviate $r_i(\mathbf{v}) = \operatorname{rank}_{\{1,\dots,i\}}(\mathbf{v})$. By the definition of $r_i(\mathbf{v})$, for $i = 1, \dots, d-1$ there exist orthonormal bases $\{\mathbf{U}_k^{\{1,\dots,i\}}\}_{k=1,\dots,r_i(\mathbf{v})}$ of $\overline{\operatorname{range} \operatorname{mat}_{\{1,\dots,i\}}(\mathbf{v})} \subseteq \ell_2(\mathbb{N}^i)$. In particular,

$$\mathbf{v}[j_1, \dots, j_d] = \sum_{k_{d-1}=1}^{r_{d-1}(\mathbf{v})} \mathbf{U}_{k_{d-1}}^{\{1,\dots,d-1\}}[j_1, \dots, j_{d-1}] \mathbf{V}_d[k_{d-1}, j_d]$$

with the coefficients

$$\mathbf{V}_d[k_{d-1}, j_d] = \sum_{(j_1,\dots,j_{d-1}) \in \mathbb{N}^{d-1}} \mathbf{U}_{k_{d-1}}^{\{1,\dots,d-1\}}[j_1, \dots, j_{d-1}]\, \mathbf{v}[j_1, \dots, j_{d-1}, j_d].$$

We now proceed recursively: by (2.14), for $1 < i < d$ we obtain a third-order coefficient tensor $\mathbf{V}_i$ such that

$$\mathbf{U}_k^{\{1,\dots,i\}}[j_1, \dots, j_i] = \sum_{k_{i-1}=1}^{r_{i-1}(\mathbf{v})} \mathbf{U}_{k_{i-1}}^{\{1,\dots,i-1\}}[j_1, \dots, j_{i-1}]\, \mathbf{V}_i[k_{i-1}, j_i, k].$$

Setting $\mathbf{V}_1[j_1, k] = \mathbf{U}_k^{\{1\}}[j_1]$, we arrive at the *tensor train representation*

$$\mathbf{v}[j_1, \dots, j_d] = \sum_{k_1=1}^{r_1(\mathbf{v})} \cdots \sum_{k_{d-1}=1}^{r_{d-1}(\mathbf{v})} \mathbf{V}_1[j_1, k_1] \mathbf{V}_2[k_1, j_2, k_2] \cdots \mathbf{V}_d[k_{d-1}, j_d]. \tag{2.15}$$

Remark 2.20 (storage costs). Let $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ such that $\operatorname{rank}_{\mathrm{TT}}(\mathbf{v}) = (r_i)_{i=1,\dots,d-1}$ with $r_i \in \mathbb{N}$, $i = 1, \dots, d-1$, and $\operatorname{supp} \mathbf{v} \subseteq \{1, \dots, n_1\} \times \cdots \times \{1, \dots, n_d\}$, with $n_1, \dots, n_d \in \mathbb{N}$. Then the number of entries to be stored for the representation (2.15) is

$$\sum_{i=1}^{d} r_{i-1} n_i r_i,$$

where we set $r_0 = r_d = 1$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Although the above bound has only a linear explicit dependence on d, note that the scaling with respect to d also depends on the ranks, which can in principle still scale exponentially with respect to d. Suitable rank bounds are thus crucial for actually avoiding the curse of dimensionality.

As a shorthand notation for explicitly specifying tensor train representations, individual cores $\mathbf{V} \in \mathbb{R}^{r_1 \times n \times r_2}$ with $\mathbf{V}[k_1, j, k_2] = \mathbf{a}_{k_1,k_2}[j]$ for certain vectors $\mathbf{a}_{k_1,k_2}$ can be written in the block form

$$\mathbf{V} = \begin{bmatrix} \mathbf{a}_{1,1} & \cdots & \mathbf{a}_{1,r_2} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{r_1,1} & \cdots & \mathbf{a}_{r_1,r_2} \end{bmatrix}. \tag{2.16}$$

This can be combined with the following special notion of product, introduced by De Launey and Seberry (1994), to obtain a compact notation.

Definition 2.21 (strong Kronecker product). For $\mathbf{V}_1 \in \mathbb{R}^{r_1 \times n_1 \times r_2}$ and $\mathbf{V}_2 \in \mathbb{R}^{r_2 \times n_2 \times r_3}$, we define

$$(\mathbf{V}_1 \bowtie \mathbf{V}_2)[k_1, j_1, j_2, k_3] = \sum_{k_2=1}^{r_2} \mathbf{V}_1[k_1, j_1, k_2]\mathbf{V}_2[k_2, j_2, k_3].$$

This product thus acts like standard matrix multiplication on the block form (2.16), where individual entries are combined by the tensor product. As a specific example,

$$\begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{a}_{22} \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{b}_{11} & \mathbf{b}_{12} \\ \mathbf{b}_{21} & \mathbf{b}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{11} \otimes \mathbf{b}_{11} + \mathbf{a}_{12} \otimes \mathbf{b}_{21} & \mathbf{a}_{11} \otimes \mathbf{b}_{12} + \mathbf{a}_{12} \otimes \mathbf{b}_{22} \\ \mathbf{a}_{21} \otimes \mathbf{b}_{11} + \mathbf{a}_{22} \otimes \mathbf{b}_{21} & \mathbf{a}_{21} \otimes \mathbf{b}_{12} + \mathbf{a}_{22} \otimes \mathbf{b}_{22} \end{bmatrix},$$

where the product of two rank-two blocks of tensor order one leads to a new rank-two block of order two. With this notation, we write (2.15) concisely as

$$\mathbf{v} = \mathbf{V}_1 \bowtie \mathbf{V}_2 \bowtie \cdots \bowtie \mathbf{V}_K. \tag{2.17}$$

2.6. Hierarchical tensors and tree tensor networks

We now consider hierarchical tensors as introduced in the context of numerical analysis by Hackbusch and Kühn (2009); see also Hackbusch (2019) and Falcó, Hackbusch and Nouy (2021). In the physics literature, these are a special case of *tree tensor networks*; see Shi, Duan and Vidal (2006) and Murg, Verstraete, Legeza and Noack (2010).

Definition 2.22. A set $\mathbb{T} \subset 2^{\{1,\dots,d\}}$ is called a binary *dimension tree* (for dimension d) if the following conditions hold.

(i) $\{1, \dots, d\} \in \mathbb{T}$ and $\emptyset \notin \mathbb{T}$.

(ii) For each $\alpha, \beta \in \mathbb{T}$ with $\alpha \neq \beta$, either $\alpha \subset \beta$ or $\beta \subset \alpha$ or $\alpha \cap \beta = \emptyset$.

(iii) For each $\alpha \in \mathbb{T}$ with $\#\alpha > 1$, there exists precisely one subset $\{\beta_1, \beta_2\} \subseteq \mathbb{T}$ with $\beta_1 \cap \beta_2 = \emptyset$ such that $\alpha = \beta_1 \cup \beta_2$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

As a consequence of this definition, if $\mathbb{T}$ is a binary dimension tree, then for all $i \in \{1, \dots, d\}$ we have $\{i\} \in \mathbb{T}$; these singleton elements are referred to as *leaves* of the tree.

As in Bachmayr and Schneider (2017), we introduce a set of *effective edges* $\mathbb{E}$ of $\mathbb{T}$ as the pairs

$$\mathbb{E} = \{\{\alpha, \alpha^c\} : \alpha \in \mathbb{T} \setminus \{\alpha^*\}\},$$

for which we have $\#\mathbb{E} = 2d - 3$. For each $e \in \mathbb{E}$, we define the representer $[e]$ as the $\alpha \in e$ such that $\alpha \in \mathbb{T}$; if this element is not unique, we make an arbitrary choice of $[e]$ among the elements of e.

Definition 2.23. Let $\mathbb{T}$ be a dimension tree with associated effective edges $\mathbb{E}$. The *hierarchical rank* of a tensor $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ is given by

$$\operatorname{rank}_{\mathbb{E}}(\mathbf{v}) = (\operatorname{rank}_{[e]}(\mathbf{v}))_{e \in \mathbb{E}}.$$

Remark 2.24. While for many purposes the elements of the dimension tree $\mathbb{T}$ are the decisive quantities, the ranks are indexed by the effective edges $\mathbb{E}$. These correspond to the different matricizations associated to $\mathbb{T}$, where those that can be obtained from each other by transposition are treated as equal (since these necessarily have the same rank). In general, there are several different $\mathbb{T}$ that lead to the same set of matricizations $\mathbb{E}$. For example, for each of the five different $\mathbb{T}$ for $d = 4$ shown in Figure 2.1, we have the same effective edges

$$\begin{aligned}\mathbb{E} = \big\{&\{\{1\}, \{2,3,4\}\}, \{\{2\}, \{1,3,4\}\},\\ &\{\{3\}, \{1,2,4\}\}, \{\{4\}, \{1,2,3\}\}, \{\{1,2\}, \{3,4\}\}\big\}.\end{aligned}$$

These $\mathbb{T}$ are thus in principle equivalent in that they lead to tensor representations with the same ranks; the difference lies in the placement of the root element $\{1,2,3,4\}$ in the tree. Intuitively, when looking only at the connectivity of the tree between modes (rather than its labels), the different $\mathbb{T}$ arise by placing the root element on different effective edges.

Similarly to the tensor train format, the notion of hierarchical rank leads to a representation of high-order tensors in terms of lower-order tensors, which can be obtained as follows. Let $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ be given. For each $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$, there exists an orthonormal system $\{\mathbf{U}^\alpha_k\}_{k=1,\dots,\operatorname{rank}_\alpha(\mathbf{v})}$ in $\ell_2(\mathbb{N}^{\#\alpha})$ that is an orthonormal basis of $\overline{\operatorname{range} \operatorname{mat}_\alpha(\mathbf{v})}$. Such an orthonormal system is again called an *orthonormal (α-)mode frame*. In what follows, for each α, we write $\mathbf{U}^\alpha$ for the (potentially infinite) matrix that has columns $\mathbf{U}^\alpha_k$ for $k = 1, \dots, \operatorname{rank}_\alpha(\mathbf{v})$.

In a first step, for $\mathbf{v}$ as above, we have the decomposition

$$\mathbf{v} = \sum_{\ell_1=1}^{\operatorname{rank}_{\alpha_1^*}(\mathbf{v})} \sum_{\ell_2=1}^{\operatorname{rank}_{\alpha_2^*}(\mathbf{v})} \mathbf{b}^{\alpha^*}[\ell_1, \ell_2]\, \mathbf{U}^{\alpha_1^*}_{\ell_1} \otimes \mathbf{U}^{\alpha_2^*}_{\ell_2}, \tag{2.18}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

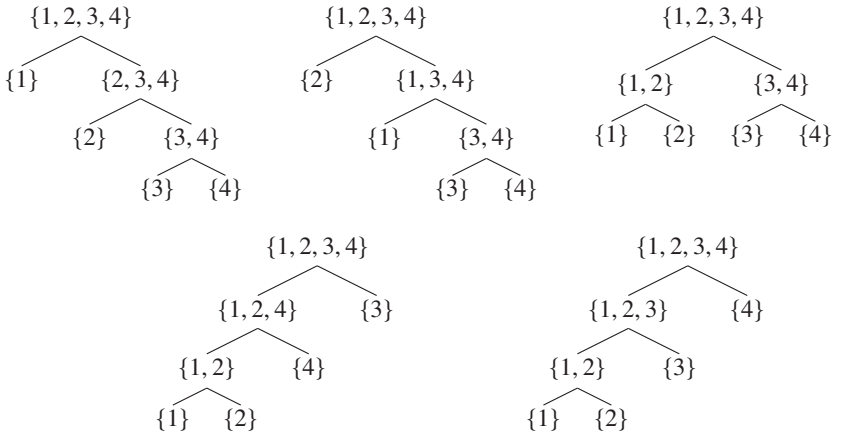

Figure 2.1. Examples of dimension trees $\mathbb{T}$ with $d = 4$ that lead to the same effective edges $\mathbb{E}$.

where α_1^*, α_2^* are the children of the root element α^* and where

$$\mathbf{b}^{\alpha^*}[\ell_1, \ell_2] = \left\langle \mathbf{v}, \mathbf{U}_{\ell_1}^{\alpha_1^*} \otimes \mathbf{U}_{\ell_2}^{\alpha_2^*} \right\rangle;$$

note that in this particular case $\operatorname{rank}_{\alpha_1^*}(\mathbf{v}) = \operatorname{rank}_{\alpha_2^*}(\mathbf{v})$, since the children of the root share (up to transposition) the same matricization. For $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$ with $\#\alpha > 1$, for the children $\alpha_1, \alpha_2 \in \mathbb{T}$ of α, we have

$$\overline{\operatorname{range} \operatorname{mat}_\alpha(\mathbf{v})} \subseteq \overline{\operatorname{range} \operatorname{mat}_{\alpha_1}(\mathbf{v})} \otimes \overline{\operatorname{range} \operatorname{mat}_{\alpha_2}(\mathbf{v})}, \tag{2.19}$$

which can be shown analogously to (2.14). As a consequence

$$\mathbf{U}_k^\alpha = \sum_{\ell_1=1}^{\operatorname{rank}_{\alpha_1}(\mathbf{v})} \sum_{\ell_2=1}^{\operatorname{rank}_{\alpha_2}(\mathbf{v})} \mathbf{b}^\alpha[k, \ell_1, \ell_2]\, \mathbf{U}_{\ell_1}^{\alpha_1} \otimes \mathbf{U}_{\ell_2}^{\alpha_2}, \quad k = 1, \ldots, \operatorname{rank}_\alpha(\mathbf{v}), \tag{2.20}$$

with the *transfer tensors* $\mathbf{b}^\alpha$ of size $\operatorname{rank}_\alpha(\mathbf{v}) \times \operatorname{rank}_{\alpha_1}(\mathbf{v}) \times \operatorname{rank}_{\alpha_2}(\mathbf{v})$ given by

$$\mathbf{b}^\alpha[k, \ell_1, \ell_2] = \left\langle \mathbf{U}_k^\alpha, \mathbf{U}_{\ell_1}^{\alpha_1} \otimes \mathbf{U}_{\ell_2}^{\alpha_2} \right\rangle.$$

The nestedness property (2.19) implies the restriction

$$\operatorname{rank}_\alpha(\mathbf{v}) \le \operatorname{rank}_{\alpha_1}(\mathbf{v}) \operatorname{rank}_{\alpha_2}(\mathbf{v}).$$

We let $\operatorname{Ranks}(\mathbb{E})$ denote the set of feasible rank vectors in $(\mathbb{N}_0 \cup \{\infty\})^{\mathbb{E}}$ of hierarchical tensors.

In summary, applying (2.20) recursively, from (2.18) we obtain a representation of $\mathbf{v}$ by the transfer tensors $\mathbf{b}^\alpha$ for all $\alpha \in \mathbb{T}$ with $\#\alpha > 1$ and the mode frames $\{\mathbf{U}_k^{\{i\}}\}_{k=1,\ldots,\operatorname{rank}_i(\mathbf{v})}$ for $i = 1, \ldots, d$. All operations on hierarchical tensors are then performed exclusively on these representation components.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Remark 2.25. One can also specify tensors in this form with non-orthogonal mode frames in a potentially redundant representation. In this case we refer to the mode sizes of the transfer tensors (which do not necessarily reflect the true hierarchical ranks) as *representation ranks*.

For $d = 4$, in the case of the balanced binary tree

$$\mathbb{T} = \big\{\{1,2,3,4\},\{1,2\},\{3,4\},\{1\},\{2\},\{3\},\{4\}\big\},$$

the decomposition can be written as

$$\mathbf{v} = \sum_{k_1=1}^{\mathrm{rank}_1(\mathbf{v})} \cdots \sum_{k_4=1}^{\mathrm{rank}_4(\mathbf{v})} \mathbf{a}[k_1,\dots,k_4]\, \mathbf{U}_{k_1}^{\{1\}} \otimes \cdots \otimes \mathbf{U}_{k_4}^{\{4\}}, \tag{2.21}$$

where, with $r_{12} = \mathrm{rank}_{\{1,2\}}(\mathbf{v})$, $r_{34} = \mathrm{rank}_{\{3,4\}}(\mathbf{v})$,

$$\mathbf{a}[k_1,\dots,k_4] = \sum_{\ell_1=1}^{r_{12}} \sum_{\ell_2=1}^{r_{34}} \mathbf{b}^{\{1,2,3,4\}}[\ell_1,\ell_2]\, \mathbf{b}^{\{1,2\}}[\ell_1,k_1,k_2]\, \mathbf{b}^{\{3,4\}}[\ell_2,k_3,k_4].$$

Since it can be interpreted as a direct generalization of the Tucker format – with additional constraints on ranks of matricizations and a resulting further decomposition of the core tensor – this tensor format is also referred to in the literature as the *hierarchical Tucker format*.

Remark 2.26 (storage costs). The number of coefficients that need to be stored for a tensor in hierarchical format of size $n_1 \times \cdots \times n_d$, with the abbreviations $r_\alpha = \mathrm{rank}_\alpha(\mathbf{v})$ and $r_i = \mathrm{rank}_i(\mathbf{v})$, is

$$\sum_{\substack{\alpha\in\mathbb{T}\\ \#\alpha>1}} r_\alpha\, r_{\alpha_1}\, r_{\alpha_2} + \sum_{i=1}^{d} r_i\, n_i,$$

where α_1, α_2 denote the children of each α.

Remark 2.27 (relation to the tensor train format). The tensor train format has the effective edges

$$\mathbb{E} = \big\{\{\{1,\dots,i\},\{i+1,\dots,d\}\} : i = 1,\dots,d-1\big\}, \tag{2.22}$$

which do not directly correspond to a dimension tree according to Definition 2.22, due to missing leaf elements. From a practical point of view, this means that pairs of transfer tensors and mode frames that are kept separate in the hierarchical format are contracted to a single component in the tensor train format. When keeping this separation, the linear dimension tree

$$\big\{\{1,\dots,d\},\{1\},\{2,\dots,d\},\{2\},\{3,\dots,d\},\dots,\{d-1,d\},\{d-1\},\{d\}\big\}$$

yields a hierarchical tensor representation that is analogous to the tensor train format. By constraining the mode frames to be identity mappings, the tensor train

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

format can alternatively be seen as a special case of this representation (Uschmajew and Vandereycken 2013). For a comparison of the respective notions of ranks, see Grasedyck and Hackbusch (2011).

Remark 2.28 (orthogonalization). A representation of a finitely supported $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ in hierarchical format by non-orthonormal mode frames $\tilde{\mathbf{U}}^{\{i\}}$ for $i = 1, \dots, d$ and transfer tensors $\tilde{\mathbf{b}}^\alpha$ for $\alpha \subseteq \alpha^*$ with $\#\alpha > 1$ can be transformed such that all α-mode frames are orthonormal; see Grasedyck (2009/10) and Hackbusch and Kühn (2009). First, for each leaf $\{i\}$ with parent $\beta_i \in \mathbb{T}$, perform a QR decomposition $\tilde{\mathbf{U}}^{\{i\}} = \mathbf{U}^{\{i\}}\mathbf{R}_i$. With $\tilde{\mathbf{U}}^{\{i\}}$ attached to mode m of the third-order tensor $\tilde{\mathbf{b}}^{\beta_i}$, replace $\tilde{\mathbf{b}}^{\beta_i}$ by $\tilde{\mathbf{b}}^{\beta_i} \times_m \mathbf{R}_i$. For $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$ with parent β and children α_1, α_2 such that $\mathbf{U}^{\alpha_1}$ and $\mathbf{U}^{\alpha_2}$ are already orthonormalized, perform a QR factorization $\mathrm{mat}_{\{2,3\}}(\tilde{\mathbf{b}}^\alpha) = \mathbf{C}_\alpha \mathbf{R}_\alpha$; with $\tilde{\mathbf{b}}^\alpha$ attached to mode m of $\tilde{\mathbf{b}}^\beta$, replace $\tilde{\mathbf{b}}^\beta$ by $\tilde{\mathbf{b}}^\beta \times_m \mathbf{R}_\alpha$, and set $\mathbf{b}^\alpha = \mathrm{mat}^{-1}_{\{2,3\}}(\mathbf{C}_\alpha)$.

After performing this recursively for all elements of $\mathbb{T} \setminus \{\alpha^*\}$, the computed $\mathbf{b}^\alpha$ and $\mathbf{U}^{\{i\}}$ yield a hierarchical tensor representation of $\mathbf{v}$ such that all associated mode frames $\mathbf{U}^\alpha$ for $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$ are orthonormal, that is, $(\mathbf{U}^\alpha)^\top \mathbf{U}^\alpha = \mathbf{I}$. To assess the total computational costs of this procedure, we again assume finite mode sizes $n_1, \dots, n_d$. We assume the input representation to be given with matricization ranks $(r_\alpha)_{\alpha \in \mathbb{T}}$. This means that for $\alpha \in \mathbb{T}$ with children α_1, α_2, the transfer tensor $\tilde{\mathbf{b}}^\alpha$ has size $r_\alpha \times r_{\alpha_1} \times r_{\alpha_2}$. In addition, we set $r_{\alpha^*} = 1$. The number of operations then scales as

$$O\left(\sum_{\substack{\alpha \in \mathbb{T} \\ \#\alpha > 1}} \left(r_\alpha^2 \, r_{\alpha_1} \, r_{\alpha_2} + r_\alpha \, r_{\alpha_1}^2 \, r_{\alpha_2} + r_\alpha \, r_{\alpha_1} \, r_{\alpha_2}^2 \right) + \sum_{i=1}^{d} r_{\{i\}}^2 \, n_i \right). \tag{2.23}$$

Here, orthogonalization is performed with respect to the root element α^*. Hence, as a consequence of (2.18), we have

$$\|\mathbf{v}\|_{\ell_2} = \|\mathbf{b}^{\alpha^*}\|_{\ell_2}.$$

We can also orthogonalize towards different edges in the dimension tree, which amounts to switching between alternative dimension trees that have the same effective edges $\mathbb{E}$ (as illustrated in Figure 2.1); see Bachmayr and Schneider (2017) for details. Tensor train representations can be orthogonalized in a similar fashion, where we distinguish between left and right orthogonality; see Oseledets (2011*b*) and Holtz, Rohwedder and Schneider (2012*b*).

Addition of tensors in hierarchical or tensor train format can be done by concatenation of representations similarly to the case of low-rank matrices as in Section 2.2.2 and Tucker tensors in Remark 2.18(i). The hierarchical or tensor train ranks of such sums are then again bounded componentwise by the sums of the ranks of summands.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Proposition 2.29. With $\mathbf{v}_k^{(i)} \in \ell_2(\mathbb{N})$ for $k = 1, \dots, r$ and $i = 1, \dots, d$, let

$$\mathbf{v} = \sum_{k=1}^{r} \mathbf{v}_k^{(1)} \otimes \cdots \otimes \mathbf{v}_k^{(d)}.$$

Then we have the componentwise estimate $\operatorname{rank}_{\mathbb{E}}(\mathbf{v}) \le r$.

Proof. Each elementary tensor product $\mathbf{v}_k = \mathbf{v}_k^{(1)} \otimes \cdots \otimes \mathbf{v}_k^{(d)}$ can be represented in hierarchical format with $\operatorname{rank}_{\mathbb{E}}(\mathbf{v}_k) \equiv 1$ for any $\mathbb{T}$, where all transfer tensors have size $1 \times 1 \times 1$. The statement thus follows by addition of these representations. □

Remark 2.30. Tree tensor networks can be considered based on trees of more general arities than the binary ones used Definition 2.22. These arities, however, enter exponentially into the storage costs (with the Tucker format as an extreme case), and thus binary trees are typically the most favourable choice.

In certain application scenarios, it can also be of interest to consider tensor networks without tree structure, where components are connected according to a more general graph with cycles. A simple example is given by *tensor chains* (which are also known as *cyclic MPS* in physics) with three components in the form

$$\mathbf{t}[j_1, j_2, j_3] = \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_2} \sum_{k_3=1}^{r_3} \mathbf{U}[k_1, j_1, k_2]\mathbf{V}[k_2, j_2, k_3]\mathbf{W}[k_3, j_3, k_1].$$

However, tensor networks with cycles lead to computational issues similar to those described for the canonical format in Section 1.4 (Landsberg, Qi and Ye 2012), and the algorithmic framework for subspace formats based on matricizations is not applicable to such networks. For a detailed discussion, see Hackbusch (2019, Sec. 12.4) and Hackbusch (2014, Sec. 9.4).

2.7. Hierarchical singular value decomposition and low-rank approximations

For tensor order $d = 2$, by Theorem 2.8, low-rank best approximations in $\ell_2(\mathbb{N}^2)$ (and thus in any separable tensor product Hilbert space) can be obtained by truncation of the SVD. As noted in Remark 2.12, for finite matrices, these best approximations can be computed by a polynomial-time algorithm that is actually not much more costly than other standard matrix factorizations.

As we saw in Section 1.4, in the case of the canonical tensor format (1.9), low-rank best approximations generally do not exist. This is different in subspace tensor formats, where we have the following result.

Theorem 2.31. For any $\mathbf{v} \in \ell_2(\mathbb{N}^d)$, any dimension tree $\mathbb{T}$ and any $\mathsf{r} = (r_e)_{e \in \mathbb{E}} \in (\mathbb{N} \cup \{\infty\})^{\mathbb{E}}$, there exists a best approximation with respect to the ℓ_2-norm in $\{\mathbf{w} \in \ell_2(\mathbb{N}^d) \colon \operatorname{rank}_{\mathbb{E}}(\mathbf{w}) \le \mathsf{r}\}$.

Proof. Let $S_{\mathsf{r}} = \{\mathbf{w} \in \ell_2(\mathbb{N}^d) \colon \operatorname{rank}_{\mathbb{E}}(\mathbf{w}) \le \mathsf{r}\}$. Then

$$S_{\mathsf{r}} = \bigcap_{e \in \mathbb{E}} \{\mathbf{w} \in \ell_2(\mathbb{N}^d) \colon \operatorname{rank}_{[e]}(\mathbf{w}) \le r_e\},$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

and for each e, the set $\{\mathbf{w} \in \ell_2(\mathbb{N}^d)\colon \operatorname{rank}_{[e]}(\mathbf{w}) \leq r_e\}$ with constraint on $\operatorname{rank}_{[e]}(\mathbf{w}) = \operatorname{rank} \operatorname{mat}_{[e]}(\mathbf{w})$ is weakly sequentially closed. This can be deduced from the lower semicontinuity of the rank of finite-dimensional matrices; see for example Uschmajew (2013, Lem. 6.5) or Hackbusch (2014, Sec. 6.4). Thus S_{r} is weakly sequentially closed. Since $\ell_2(\mathbb{N}^d)$ is reflexive, this allows us to obtain the weak limit of a minimizing sequence in S_{r}, which is a minimizer by weak lower semicontinuity of the norm. □

Related results, including the case of the Tucker format, have been obtained in Uschmajew (2010). In the case of general Banach spaces, the existence of best low-rank approximations has been treated by Falcó and Hackbusch (2012) using the notion of minimal subspaces.

In the case of subspace tensor formats for $d > 2$ considered in this section, there exist normal forms of tensor representations that share certain features of the SVD of matrices, such as the higher-order SVD discussed in Remark 2.18 in the case of the Tucker format. For general hierarchical tensors, we have an analogous notion developed in Grasedyck (2009/10) and Hackbusch and Kühn (2009); see also Hackbusch (2019, Sec. 11.3.3).

Definition 2.32. A hierarchical tensor representation of $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ with dimension tree $\mathbb{T}$ is a *hierarchical singular value decomposition* (HSVD) if, for all $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$, the mode frame $\{\mathbf{U}_k^\alpha\}_{k=1,\dots,\operatorname{rank}_\alpha(\mathbf{v})}$ is made up of left singular vectors of $\operatorname{mat}_\alpha(\mathbf{v})$, ordered such that the corresponding sequence of singular values, denoted by $\sigma^\alpha(\mathbf{v}) = \left(\sigma_k^\alpha(\mathbf{v})\right)_{k=1,\dots,\operatorname{rank}_\alpha(\mathbf{v})}$, is non-increasing.

In the case of the tensor train representation, the analogous form in terms of left singular vectors of matricizations is referred to as *tensor train SVD* (TT-SVD), introduced in Oseledets (2011*b*). In physics, an equivalent representation is known as the *Vidal representation* of matrix product states (Vidal 2003).

Remark 2.33 (existence and uniqueness). We can always choose the mode frames for each $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$ as left singular vectors of the α-matricization, whose existence is ensured by Theorem 2.8. Thus a hierarchical singular value decomposition exists for every $\mathbf{v} \in \ell_2(\mathbb{N}^d)$. It is unique precisely when the sequences $\sigma^\alpha(\mathbf{v})$ are strictly decreasing for all α.

A procedure for numerically transforming a hierarchical tensor representation to HSVD form was introduced by Grasedyck (2009/10). Starting from a representation with orthonormalized mode frames, the matrices transforming these mode frames can be extracted from eigenvalue decompositions of certain recursively assembled Gramians associated to the nodes in $\mathbb{T}$. The computational costs are of the same order (2.23) as the orthogonalization, which for $\mathbf{v}$ of mode sizes $n_1 \times \cdots \times n_d$ is bounded by

$$O\left(d \max_{\alpha \in \mathbb{T}} \operatorname{rank}_\alpha^4(\mathbf{v}) + \sum_{i=1}^d \operatorname{rank}_i(\mathbf{v})^2 n_i\right). \tag{2.24}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

For the details of this procedure, we refer to Grasedyck (2009/10) and Hackbusch (2019, Sec. 11.4.2). Oseledets (2011*b*) has introduced a procedure for computing TT-SVD representations of tensor trains following similar lines to the HOSVD in Remark 2.18(iii).

2.7.1. Truncated hierarchical SVD

We now consider rank truncation of HSVD representations in analogy to the truncated SVD of matrices in Theorem 2.8.

Let $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ be given in HSVD form with orthonormal mode frames $\mathbf{U}^\alpha$ for $\alpha \in \mathbb{T} \setminus \{\alpha^*\}$. For each α and r, we write $\mathbf{U}^{\alpha,r}$ for $\mathbf{U}^\alpha$ truncated to its first r columns. For each $e \in \mathbb{E}$, we define the orthogonal projector

$$\mathrm{P}^e_{\mathbf{v},r}\mathbf{v} = \mathbf{v} \times_{[e]} \mathbf{U}^{[e],r} \left(\mathbf{U}^{[e],r}\right)^\top.$$

We now introduce a notion of *level* of nodes in $\mathbb{T}$ as the distance to the root element α^*, so that α^* has level zero and the children of α^* have level one. We next choose an ordering $e(1), e(2), \dots, e(\#\mathbb{E})$ of $\mathbb{E}$ such that $[e(1)], [e(2)], \dots$ have non-decreasing level. With this ordering, for $\mathsf{r} = (r_e)_{e\in\mathbb{E}} \in \mathrm{Ranks}(\mathbb{E})$ we define the sequence of projections

$$\mathrm{P}_{\mathbf{v},\mathsf{r}} = \mathrm{P}^{e(\#\mathbb{E})}_{\mathbf{v},r_{e(\#\mathbb{E})}} \cdots \mathrm{P}^{e(2)}_{\mathbf{v},r_{e(2)}} \mathrm{P}^{e(1)}_{\mathbf{v},r_{e(1)}}, \tag{2.25}$$

which projects onto truncated mode frames starting from the root of $\mathbb{T}$. This sequence of projections realizes a truncation to hierarchical rank at most r.

Lemma 2.34. Let $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ and $\mathsf{r} \in \mathrm{Ranks}(\mathbb{E})$ with $\mathsf{r} \le \mathrm{rank}_{\mathbb{T}}(\mathbf{v})$. Then $\mathrm{rank}_{\mathbb{E}}(\mathrm{P}_{\mathbf{v},\mathsf{r}}\mathbf{v}) \le \mathsf{r}$.

Proof. As a consequence of the ordering in (2.25), for each $e = \{\alpha, \alpha^\mathsf{c}\} \in \mathbb{E}$, once the projection $\mathrm{P}^e_{\mathbf{v},r_e}$ has been applied, all projections that follow in the ordering act only on modes that are subsets of either α or α^c and thus act from either the left or right on $\mathrm{mat}_\alpha(\mathbf{v})$. The subsequent projections can thus only decrease $\mathrm{rank}_\alpha(\mathbf{v})$. □

Note that the ordering of the projections in (2.25) is essential: without the ordering from root to leaves, applying the projections may indeed increase the ranks beyond the prescribed r; see Grasedyck (2009/10).

Remark 2.35 (applying the truncation). It is easy to see that applying $\mathrm{P}_{\mathbf{v},\mathsf{r}}$ to an HSVD representation amounts to selecting a subset of basis vectors in each mode frame. Since the mode frames are ordered by non-increasing singular values, this is accomplished by simply truncating all transfer tensors and mode frames in the HSVD to match the hierarchical rank r. A completely analogous procedure can be performed for the tensor train format (Oseledets 2011*b*).

Truncating the HSVD does not provide the best approximation with the given ranks from Theorem 2.31. However, we instead obtain a *quasi-best* approximation: the truncation error can be estimated by a fixed multiple of the best approximation

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

error. The following result to this effect was obtained for hierarchical tensors in Grasedyck (2009/10).

Theorem 2.36. Let $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ and $\mathsf{r} = (r_e)_{e\in\mathbb{E}} \in \mathrm{Ranks}(\mathbb{E})$ with $\mathsf{r} \le \mathrm{rank}_{\mathbb{E}}(\mathbf{v})$. Then

$$\begin{aligned}\|\mathbf{v} - \mathrm{P}_{\mathbf{v},\mathsf{r}}\mathbf{v}\|_{\ell_2} &\le \left(\sum_{e\in\mathbb{E}}\sum_{k>r_e}\big(\sigma_k^{[e]}(\mathbf{v})\big)^2\right)^{1/2}\\ &\le \sqrt{\#\mathbb{E}}\,\min\big\{\|\mathbf{v}-\mathbf{w}\|_{\ell_2} : \mathbf{w}\in\ell_2(\mathbb{N}^d),\ \mathrm{rank}_{\mathbb{E}}(\mathbf{w})\le\mathsf{r}\big\}.\end{aligned}$$

Proof. Let P be an orthogonal projection and Q a bounded operator on $\ell_2(\mathbb{N}^d)$, respectively. Then, since $\mathrm{range}(I-P)\perp\mathrm{range}(P)$ and $\|P\|_{\ell_2\to\ell_2} = 1$,

$$\begin{aligned}\|\mathbf{v} - PQ\mathbf{v}\|_{\ell_2}^2 &= \|(I-P)\mathbf{v} + P(I-Q)\mathbf{v}\|_{\ell_2}^2\\ &= \|(I-P)\mathbf{v}\|_{\ell_2}^2 + \|P(I-Q)\mathbf{v}\|_{\ell_2}^2\\ &\le \|(I-P)\mathbf{v}\|_{\ell_2}^2 + \|(I-Q)\mathbf{v}\|_{\ell_2}^2.\end{aligned}$$

Applying this inductively, starting from

$$P = \mathrm{P}_{\mathbf{v},r_{e(\#\mathbb{E})}}^{e(\#\mathbb{E})} \quad\text{and}\quad Q = \mathrm{P}_{\mathbf{v},r_{e(\#\mathbb{E}-1)}}^{e(\#\mathbb{E}-1)}\cdots\mathrm{P}_{\mathbf{v},r_{e(1)}}^{e(1)},$$

we arrive at

$$\|\mathbf{v} - \mathrm{P}_{\mathbf{v},\mathsf{r}}\mathbf{v}\|_{\ell_2}^2 \le \sum_{e\in\mathbb{E}}\|\mathbf{v} - \mathrm{P}_{\mathbf{v},r_e}^e\mathbf{v}\|_{\ell_2}^2.$$

By Theorem 2.8, for each $e\in\mathbb{E}$,

$$\|\mathbf{v} - \mathrm{P}_{\mathbf{v},r_e}^e\mathbf{v}\|_{\ell_2}^2 = \sum_{k>r_e}\big(\sigma_k^{[e]}(\mathbf{v})\big)^2 = \min\big\{\|\mathbf{v}-\mathbf{w}\|_{\ell_2}^2 :\ \mathrm{rank}_{[e]}(\mathbf{w})\le r_e\big\}.$$

Since $\min\{\|\mathbf{v}-\mathbf{w}\|_{\ell_2} :\ \mathrm{rank}_{[e]}(\mathbf{w})\le r_e\} \le \min\{\|\mathbf{v}-\mathbf{w}\|_{\ell_2} :\ \mathrm{rank}_{\mathbb{E}}(\mathbf{w})\le\mathsf{r}\}$, the statement follows. □

Remark 2.37. Theorem 2.36 applies to all subspace tensor formats discussed so far.

(i) In the case of the Tucker format, which corresponds to the *non-binary* dimension tree $\mathbb{T} = \{\{1,\dots,d\},\{1\},\dots,\{d\}\}$ with $\#\mathbb{E} = d$, Theorem 2.36 reduces to the error bounds for truncated higher-order singular value decompositions obtained by De Lathauwer *et al.* (2000).

(ii) For matricizations (2.22) of the tensor train format, we have $\#\mathbb{E} = d-1$. In this case Theorem 2.36 yields the result for the tensor train format obtained by Oseledets (2011*b*).

Remark 2.38. The rank truncation described here is based on transforming the entire representation to HSVD form and then performing the truncation. Variants that sequentially transform single matricizations into SVD form and directly

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

perform the corresponding truncation have been investigated in Kühn (2012) (see also Hackbusch 2019, Sec. 11.4.2.2); the procedure for truncating the TT-SVD proposed by Oseledets (2011*b*) is also of this type. An analogous method for general tree tensor networks has been analysed by Ceruti, Lubich and Sulz (2023). Due to the decreased ranks of intermediate results, in practice this sequential truncation can reduce the total computational costs. A potential disadvantage is that error tolerances need to be allocated to matricizations before all singular values are known.

2.7.2. Soft thresholding of hierarchical tensor representations

We now turn to an alternative strategy for rank reduction of hierarchical tensors with the additional feature of *non-expansiveness*, which can be exploited in the construction of iterative methods.

Let $\mathbf{v} \in \ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$, where $\mathcal{I}_1, \mathcal{I}_2$ are countable index sets, have the singular value decomposition

$$\mathbf{v} = \sum_{k=1}^{\infty} \sigma_k \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)} \tag{2.26}$$

according to Theorem 2.8. For $\delta \geq 0$ and $x \in \mathbb{R}$, let $s_\delta(x) = \operatorname{sgn}(x) \max\{|x| - \delta, 0\}$, which satisfies

$$s_\delta(x) = \operatorname*{arg\,min}_{y \in \mathbb{R}} \left\{ \frac{1}{2}|x - y|^2 + \delta|y| \right\}. \tag{2.27}$$

We define the nonlinear *soft thresholding operator* S_δ on $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$ in terms of the singular value decomposition (2.26) by

$$S_\delta(\mathbf{v}) = \sum_{k=1}^{\infty} s_\delta(\sigma_k)\, \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)}. \tag{2.28}$$

Lemma 2.39. For each $\delta \geq 0$, the operator S_δ is non-expansive, that is,

$$\|S_\delta(\mathbf{v}) - S_\delta(\mathbf{w})\|_{\ell_2} \leq \|\mathbf{v} - \mathbf{w}\|_{\ell_2} \quad \text{for all } \mathbf{v}, \mathbf{w} \in \ell_2(\mathcal{I}_1 \times \mathcal{I}_2).$$

Proof. We have the characterization

$$S_\delta(\mathbf{v}) = \operatorname*{arg\,min}_{\mathbf{w} \in \ell_2(\mathcal{I}_1 \times \mathcal{I}_2)} \left\{ \frac{1}{2}\|\mathbf{v} - \mathbf{w}\|_{\ell_2}^2 + \delta \|\sigma(\mathbf{w})\|_{\ell_1} \right\}, \tag{2.29}$$

where $\sigma(\mathbf{w})$ denotes the sequence of singular values of $\mathbf{w}$. This is shown in Bachmayr and Schneider (2017, Lem. 2.3) by a reduction to the scalar case (2.27); an alternative proof for $\#\mathcal{I}_1, \#\mathcal{I}_2 < \infty$ based on subgradient characterizations is given in Cai, Candès and Shen (2010). The right-hand side (2.29) is the definition of the *proximity operator* on $\ell_2(\mathcal{I}_1 \times \mathcal{I}_2)$ of the functional $\|\sigma(\cdot)\|_{\ell_1}$. As shown by Moreau (1965), every proximity operator is non-expansive. □

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Based on (2.28), we now define a soft thresholding operation for hierarchical tensors in $\ell_2(\mathbb{N}^d)$ that preserves the non-expansiveness. For a given dimension tree $\mathbb{T}$ and associated effective edges $\mathbb{E}$, for each $e \in \mathbb{E}$, we set

$$S_\delta^e = \mathrm{mat}_{[e]}^{-1} \circ S_\delta \circ \mathrm{mat}_{[e]}.$$

We now choose an enumeration $e(1), \dots, e(\#\mathbb{E})$ of the elements of $\mathbb{E}$ and define the *hierarchical tensor soft thresholding operator*

$$\mathrm{ST}_\delta = S_\delta^{e(\#\mathbb{E})} \circ \cdots \circ S_\delta^{e(1)} \tag{2.30}$$

of Bachmayr and Schneider (2017). By repeated application of Lemma 2.39, we immediately obtain non-expansiveness of ST_δ on $\ell_2(\mathbb{N}^d)$.

Proposition 2.40. For any $\delta \geq 0$,

$$\|\mathrm{ST}_\delta(\mathbf{v}) - \mathrm{ST}_\delta(\mathbf{w})\|_{\ell_2} \leq \|\mathbf{v} - \mathbf{w}\|_{\ell_2} \quad \text{for all } \mathbf{v}, \mathbf{w} \in \ell_2(\mathbb{N}^d).$$

Remark 2.41. For any dimension tree, by a suitable enumeration of $\mathbb{E}$, the action of ST_δ can be implemented with the same asymptotic complexity (2.24) as the computation of the HSVD; see Bachmayr and Schneider (2017, Alg. 1).

The following characterization of the δ-dependence of the error incurred by application of ST_δ in terms of singular values of matricizations is proved in Bachmayr and Schneider (2017, Lem. 3.4).

Theorem 2.42. For $\delta > 0$, $e \in \mathbb{E}$ and $\mathbf{v} \in \ell_2(\mathbb{N}^d)$, with

$$r_\delta^e(\mathbf{v}) = \#\big\{k \in \mathbb{N} \colon \sigma_k^{[e]}(\mathbf{v}) > \delta\big\}, \quad D_\delta^e(\mathbf{v}) = \sqrt{\delta^2 r_\delta^e(\mathbf{v}) + \sum_{k > r_\delta^e(\mathbf{v})} \big|\sigma_k^{[e]}(\mathbf{v})\big|^2},$$

we have the bounds

$$\max_{e \in \mathbb{E}} D_\delta^e(\mathbf{v}) \leq \|\mathrm{ST}_\delta(\mathbf{v}) - \mathbf{v}\|_{\ell_2} \leq \sum_{e \in \mathbb{E}} D_\delta^e(\mathbf{v}).$$

Moreover, $D_\delta^e(\mathbf{v}) \to 0$ as $\delta \to 0$ for each $e \in \mathbb{E}$.

2.8. *Matrices and operators in tensor formats*

When applying operators on $\ell_2(\mathbb{N}^d)$ to subspace representations, we again need to represent the results in the same format. For this to be efficiently realizable we require a representation of operators that is compatible with the considered low-rank format. In the case $d = 2$ corresponding to low-rank matrices, as outlined in (1.14), sums of tensor product operators can be applied in this format and are in this sense compatible: we obtain a representation of the same form, generally with increased ranks.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

In the case of the Tucker format, corresponding representations of operators on $\ell_2(\mathbb{N}^d)$ are of the analogous form

$$\sum_{\ell_1=1}^{R_1} \cdots \sum_{\ell_d=1}^{R_d} \mathbf{c}[\ell_1, \dots, \ell_d]\, \mathbf{A}^{(i)}_{\ell_1} \otimes \cdots \otimes \mathbf{A}^{(i)}_{\ell_d} \tag{2.31}$$

with component operators $\mathbf{A}^{(i)}_{\ell}$, $\ell = 1, \dots, R_i$, for $i = 1, \dots, d$. Applying this representation to a Tucker representation of multilinear ranks $(r_1, \dots, r_d)$ then leads to a (potentially redundant) representation of the result with representation ranks $(R_1 r_1, \dots, R_d r_d)$.

The case of hierarchical tensors is similar: we have the same basic form as in (2.31), but the core tensor $\mathbf{c}$ needs to be decomposed in terms of transfer tensors as in (2.21) according to the given dimension tree $\mathbb{T}$. In the particular case of the tensor train representation, the shorthand notation in terms of the block form (2.16) can also be applied to operators, for instance

$$\begin{bmatrix} \mathbf{I} & \mathbf{A}_1 \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{I} & \mathbf{A}_2 \\ 0 & \mathbf{I} \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{A_3} \\ \mathbf{I} \end{bmatrix} = \mathbf{A}_1 \otimes \mathbf{I} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}_2 \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{I} \otimes \mathbf{A}_3. \tag{2.32}$$

Applying this representation to a tensor train of representation rank $(2, 2)$,

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{v}_{1,1} & \mathbf{v}_{1,2} \\ \mathbf{v}_{2,1} & \mathbf{v}_{2,2} \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix},$$

yields the result with representation rank $(4, 4)$,

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{A}_1\mathbf{u}_1 & \mathbf{A}_1\mathbf{u}_2 \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{v}_{1,1} & \mathbf{v}_{1,2} & \mathbf{A}_2\mathbf{v}_{1,1} & \mathbf{A}_2\mathbf{v}_{1,2} \\ \mathbf{v}_{2,1} & \mathbf{v}_{2,2} & \mathbf{A}_2\mathbf{v}_{2,1} & \mathbf{A}_2\mathbf{v}_{2,2} \\ 0 & 0 & \mathbf{v}_{1,1} & \mathbf{v}_{1,2} \\ 0 & 0 & \mathbf{v}_{2,1} & \mathbf{v}_{2,2} \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{A}_3\mathbf{w}_1 \\ \mathbf{A}_3\mathbf{w}_2 \\ \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}.$$

2.9. Condition numbers of tensor representations

The subspace tensor formats considered in this section avoid some of the difficulties associated with the canonical format that are outlined in Section 1.4. However, under certain conditions these representations can still be sensitive to round-off errors. This can happen in particular when the components are in non-orthogonalized form with redundancies that lead to large cancellations, which can occur in intermediate results in numerical methods.

Example 2.43. In order to illustrate how this issue can arise in tensor decompositions, let us consider the tensor $\mathbf{v} \in \bigotimes_{\ell=1}^{d} \mathbb{R}^2$ with all entries equal to one, which can be represented in TT format with rank one in the form

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

However, one can also write $\mathbf{v}$ in a redundant rank-two representation

$$\mathbf{v} = \mathbf{X}_1 \bowtie \cdots \bowtie \mathbf{X}_d$$

with, for any $R > 0$,

$$\begin{aligned} \mathbf{X}_1 &= \begin{bmatrix} (1 + R^{-d}) \begin{pmatrix} R \\ R \end{pmatrix} & - \begin{pmatrix} R \\ R \end{pmatrix} \end{bmatrix}, \\ \mathbf{X}_2 = \cdots = \mathbf{X}_{d-1} &= \begin{bmatrix} \begin{pmatrix} R \\ R \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} R \\ R \end{pmatrix} \end{bmatrix}, \quad \mathbf{X}_d = \begin{bmatrix} \begin{pmatrix} R \\ R \end{pmatrix} \\ \begin{pmatrix} R \\ R \end{pmatrix} \end{bmatrix}. \end{aligned} \tag{2.33}$$

We now define $\mathbf{v}_\varepsilon$ by replacing precisely one of the cores $\mathbf{X}_i$ with $1 < i < d$ by

$$\mathbf{X}_{i,\varepsilon} = \begin{bmatrix} (1 + \varepsilon) \begin{pmatrix} R \\ R \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} R \\ R \end{pmatrix} \end{bmatrix},$$

so that $\mathbf{v}_\varepsilon$ has all entries equal to $1 + (R^d + 1)\varepsilon$. For the relative error in the represented tensor, we thus obtain

$$\frac{\|\mathbf{v} - \mathbf{v}_\varepsilon\|_2}{\|\mathbf{v}\|_2} = (R^d + 1)\varepsilon.$$

Even when R is of moderate size, a significant loss of precision can thus occur for large d.

Redundancies similar to those in Example 2.43 can arise in particular when applying operators in low-rank representation. Although one can subsequently apply orthogonalization to remove the redundancies, in general these orthogonalizations will already lead to large numerical errors. This effect can be quantified by the following notion introduced in Bachmayr and Kazeev (2020).

Definition 2.44. Let $\tau\colon \mathcal{C} \to \ell_2(\mathbb{N}^d)$ be the representation mapping of a tensor format, where $\mathcal{C}$ is the set of admissible tuples of components, and let $\mathsf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N) \in \mathcal{C}$. For $n = 1, \dots, N$, we define the representation condition numbers of X by

$$\begin{aligned} \operatorname{rcond}_n(\mathsf{X}) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \sup \Bigg\{ & \frac{\|\tau(\mathsf{X}) - \tau(\mathsf{Y})\|_{\ell_2(\mathbb{N}^d)}}{\|\tau(\mathsf{X})\|_{\ell_2(\mathbb{N}^d)}} : \\ & \mathsf{Y} = (\mathbf{X}_1, \dots, \mathbf{X}_{n-1}, \mathbf{Y}_n, \mathbf{X}_{n+1}, \dots, \mathbf{X}_N), \\ & \|\mathbf{X}_n - \mathbf{Y}_n\|_{\ell_2} \le \varepsilon \|\mathbf{X}_n\|_{\ell_2} \Bigg\}. \end{aligned} \tag{2.34}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

For tensor trains and more general hierarchical tensors, the quantities defined in (2.34) can be computed directly from norms of components; see Bachmayr and Kazeev (2020, Prop. 1). By multilinearity of τ for these tensor formats, if $\mathbf{x} = \tau(\mathsf{X})$ and $\mathbf{y} = \tau(\mathsf{Y})$ with $\mathsf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N)$ and $\mathsf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_N)$ such that

$$\|\mathbf{X}_n - \mathbf{Y}_n\|_{\ell_2} \le \varepsilon \|\mathbf{X}_n\|_{\ell_2} \quad \text{for } n = 1, \dots, N,$$

we obtain

$$\frac{\|\mathbf{x} - \mathbf{y}\|_{\ell_2}}{\|\mathbf{x}\|_{\ell_2}} \le \sum_{n=1}^{N} \operatorname{rcond}_n(\mathsf{X})\, \varepsilon + O(\varepsilon^2).$$

Definition 2.45. With the notation of Definition 2.44, let A be an operator representation acting on elements of $\mathcal{C}$ and let $\mathsf{A} \bullet \mathsf{X}$ be the result of applying A to $\mathsf{X} \in \mathcal{C}$. For $n = 1, \dots, N$, we then define the operator representation condition number of A as

$$\operatorname{oprcond}_n(\mathsf{A}) = \sup_{\mathsf{X} \in \mathcal{C}} \frac{\operatorname{rcond}_n(\mathsf{A} \bullet \mathsf{X})}{\operatorname{rcond}_n(\mathsf{X})}. \tag{2.35}$$

Hence operator representations that are ill-conditioned in the sense of large condition numbers (2.35) can lead to a correspondingly large deterioration of the representation condition numbers of the tensors that they are acting on, for instance by introducing cancellations in the representation as in Example 2.43. This can occur in certain representations of differential operators considered in Sections 3.4 and 4.5.

2.10. Manifold structure

The set of matrices of a given size $m \times n$ of fixed rank $r \in \mathbb{N}$ is a smooth embedded submanifold of $\mathbb{R}^{m \times n}$ (Helmke and Shayman 1995). This means that techniques of Riemannian optimization can be applied for solving minimization problems on fixed-rank matrices (Absil, Mahony and Sepulchre 2008), considered in further detail in Section 5.1.3. The structure of such fixed-rank matrix manifolds and their tangent spaces is also utilized in *dynamical low-rank approximation* (Koch and Lubich 2007*a*) of initial value problems; see Section 7.1.

Analogous manifold properties hold for Tucker tensors, tensor trains and hierarchical tensors with fixed ranks. This has been established for tensors with finite mode sizes for Tucker tensors in Koch and Lubich (2010), for tensor trains in Holtz *et al.* (2012*b*) and, using a different approach, for hierarchical tensors and tensor trains in Uschmajew and Vandereycken (2013); see also Arnold and Jahnke (2014). These constructions can be transferred to the tensor product Hilbert space setting of $\ell_2(\mathbb{N}^d)$; see Uschmajew (2013). The Tucker format in Banach spaces is considered in Falcó, Hackbusch and Nouy (2019). For a general discussion of manifold properties in the context of tensor networks, we refer to Bachmayr *et al.* (2016, Sec. 3.8).

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

3. Approximations of functions and operators in low-rank tensor formats

When using tensor formats in approximation problems, a natural approach is to directly assign tensor modes to coordinates in the given problem. In this case entries of a tensor representing a function typically correspond to values on a product grid or to coefficients in a basis expansion. However, tensors can also be used for encoding functions in completely different ways, where the modes are associated to basis functions or to scales in the problem. We now consider some common ways of representing functions by tensors and corresponding representations of operators. As a particular example, we consider representations of the Laplacian as the most basic example of an elliptic operator and the simplest corresponding boundary value problem.

Example 3.1. On the bounded domain Ω, the classical formulation of the *Poisson problem* reads: for a given $f \in C(\Omega)$, find $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{3.1}$$

The corresponding weak formulation is: for $f \in H^{-1}(\Omega)$, find $u \in H^1_0(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = f(v) \quad \text{for all } v \in H^1_0(\Omega). \tag{3.2}$$

We first consider grid-based discretizations by finite differences or finite elements in Section 3.1. In Section 3.2 we then turn to representations in terms of Riesz bases, which play a central role in the following sections in obtaining well-conditioned problems and in adaptively refining discretizations. More specialized uses of low-rank tensors are considered in Section 3.3 on occupation number representations and Section 3.4 on tensorized representations, where tensor modes correspond to basis functions and to scales in the approximation problem, respectively.

3.1. Uniform grids

The most common way of representing function approximations in low-rank tensor methods is by point values on a regular tensor product grid. As an example, let us consider a uniform product grid in $\Omega = \square_d$ with N points in each coordinate direction. With the grid spacing $h = 1/(N+1)$, the set of grid points in Ω is given by

$$\Omega_h = \left\{x_{j_1,\dots,j_d} : j \in \{1,\dots,N\}^d\right\}, \quad \text{where } x_{j_1,\dots,j_d} = (j_1 h, \dots, j_d h).$$

3.1.1. Finite differences

In finite difference discretizations, the original solution u of the PDE on Ω is replaced by a *grid function* $u_h : \Omega_h \to \mathbb{R}$ given by the coefficient tensor $\mathbf{u}_h \in \mathbb{R}^{N\times\cdots\times N}$ with entries

$$\mathbf{u}_h[j_1,\dots,j_d] = u_h(x_{j_1,\dots,j_d}) \approx u(x_{j_1,\dots,j_d}).$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Using low-rank tensor approximations of $\mathbf{u}_h$ in finite difference methods requires suitable representations of the corresponding discretization matrices. In Example 3.1 with $\Omega = \square_d$ and $f \in C(\Omega)$, we obtain a linear system of equations $\mathbf{A}_{\mathrm{FD}}\mathbf{u}_h = \mathbf{f}_h$, where $\mathbf{f}_h[j] = f(x_j)$ for $j \in \{1,\dots,N\}^d$ and

$$\begin{aligned}\mathbf{A}_{\mathrm{FD}} = h^{-2}\boldsymbol{\Delta}_2 \otimes \mathbf{I} \otimes \mathbf{I} \otimes \cdots \otimes \mathbf{I} + \mathbf{I} \otimes h^{-2}\boldsymbol{\Delta}_2 \otimes \mathbf{I} \otimes \cdots \otimes \mathbf{I} \\ + \cdots + \mathbf{I} \otimes \mathbf{I} \otimes \cdots \otimes \mathbf{I} \otimes h^{-2}\boldsymbol{\Delta}_2,\end{aligned} \tag{3.3}$$

where the $\mathbf{I}$ are $N \times N$ identity matrices and where we set

$$\boldsymbol{\Delta}_2 = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}. \tag{3.4}$$

Although the right-hand side of (3.3) has d summands, the matrix $\mathbf{A}_{\mathrm{FD}}$ can be written in the operator versions of the Tucker, tensor train or hierarchical tensor formats with all ranks equal to two. In particular, as observed by Kazeev and Khoromskij (2012), in the tensor train format we have, analogously to (2.32),

$$\mathbf{A}_{\mathrm{FD}} = \begin{bmatrix} \mathbf{I} & h^{-2}\boldsymbol{\Delta}_2 \end{bmatrix} \bowtie \begin{bmatrix} \mathbf{I} & h^{-2}\boldsymbol{\Delta}_2 \\ 0 & \mathbf{I} \end{bmatrix}^{\bowtie(d-2)} \bowtie \begin{bmatrix} h^{-2}\boldsymbol{\Delta}_2 \\ \mathbf{I} \end{bmatrix},$$

where for a core $\mathbf{X}$ we write $\mathbf{X}^{\bowtie n}$ for the n-fold strong Kronecker product of $\mathbf{X}$ with itself. It is easy to see that for the condition number with respect to the spectral norm, we have $\mathrm{cond}(\mathbf{A}_{\mathrm{FD}}) = \mathrm{cond}(\boldsymbol{\Delta}_2)$ independently of d, where $\mathrm{cond}(\boldsymbol{\Delta}_2) = O(h^{-2})$.

We thus obtain a discretization with simple structure that serves as a benchmark problem in many works, where it is typically considered with rather coarse grids of sizes $N = 2^5$ to $N = 2^7$ that can be treated without preconditioning. For large grid sizes, the use of a preconditioner in low-rank format becomes necessary due to the large condition number of $\mathbf{A}_{\mathrm{FD}}$. However, an important limitation of such finite difference discretizations is that they offer no mechanism for obtaining a d-robust computable error bound for the discretization error. This is especially important in the high-dimensional setting, since the scaling with respect to d of the quantities arising in *a priori* error estimates is typically unknown. Moreover, finite differences are less suitable for problems with non-smooth data (such as $f \notin C(\Omega)$).

3.1.2. Finite elements

An alternative discretization that also uses the values on the grid Ω_h, but that is suitable for problems with low regularity and offers more advanced means for error control, is provided by lowest-order multilinear finite elements. In this case, with the univariate piecewise linear hat functions $\varphi_j(x) = \max\{1 - |x - jh|/h, 0\}$, $j \in \{1,\dots,N\}$, with $h = 1/(N+1)$ as above, we consider a Galerkin discretization

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

of (3.2): find $u_h \in V_h$ such that

$$\int_\Omega \nabla u_h \cdot \nabla v_h \, \mathrm{d}x = f(v_h) \quad \text{for all } v_h \in V_h, \tag{3.5}$$

with the subspace

$$V_h = \operatorname{span}\left\{ \bigotimes_{i=1}^{d} \varphi_{j_i} : j \in \{1, \dots, N\}^d \right\} \subset H_0^1(\Omega).$$

The coefficients $\mathbf{u}_h \in \mathbb{R}^{N \times \cdots \times N}$ such that

$$u_h = \sum_{j \in \{1,\dots,N\}^d} \mathbf{u}_h[j] \bigotimes_{i=1}^{d} \varphi_{j_i}$$

then again correspond to the values of u_h at the grid points. The system of equations resulting from (3.5) is $\mathbf{A}_{\mathrm{FE}}\mathbf{u}_h = \mathbf{f}_h$ with

$$\mathbf{f}_h[j] = f\left(\bigotimes_{i=1}^{d} \varphi_{j_i} \right)$$

and

$$\begin{aligned} \mathbf{A}_{\mathrm{FE}} = {} & \mathbf{A}_1 \otimes \mathbf{M}_1 \otimes \mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_1 + \mathbf{M}_1 \otimes \mathbf{A}_1 \otimes \mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_1 \\ & + \cdots + \mathbf{M}_1 \otimes \mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_1 \otimes \mathbf{A}_1, \end{aligned} \tag{3.6}$$

where

$$\mathbf{A}_1 = \left(\int_0^1 \varphi_i' \varphi_j' \, \mathrm{d}x \right)_{i,j=1,\dots,N} = h^{-1} \mathbf{\Delta}_2$$

and

$$\mathbf{M}_1 = \left(\int_0^1 \varphi_i \varphi_j \, \mathrm{d}x \right)_{i,j=1,\dots,N} = \frac{h}{6} \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 4 \end{pmatrix}.$$

Concerning representation ranks, $\mathbf{A}_{\mathrm{FE}}$ can be represented analogously to $\mathbf{A}_{\mathrm{FD}}$ with all ranks equal to two in the formats considered.

Remark 3.2. In contrast to the finite difference case, since $\operatorname{cond}(\mathbf{M}_1) > 1$, it is easy to see that the condition number of $\mathbf{A}_{\mathrm{FE}}$ deteriorates exponentially with increasing d. This means in particular that iterative methods applied to $\mathbf{A}_{\mathrm{FE}}$ become extremely inefficient for large d. Note that once the discrete problem is represented in the basis of multilinear hat functions, this problem is not easy to circumvent, since any application of $\mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_1$ or its inverse may incur a relative error increasing exponentially with respect to d.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

3.2. *Basis representations*

Rather than working with grid values, in many approximation problems it is more advantageous to consider expansions in terms of bases of the relevant function spaces that can serve as dictionaries for nonlinear sparse approximation as outlined in Section 1.6.

Example 3.3. In parametric and stochastic problems, we are often interested in approximation in spaces $L_2(\mathbb{R}^{\mathcal{N}}, \mu)$ with $\mu = \bigotimes_{i\in\mathcal{N}} \mu_i$, where $\mathcal{N}$ can be finite or countable and μ_i is a probability measure on $\mathbb{R}$ for each $i \in \mathcal{N}$. To approximate smooth functions in such spaces, a natural choice of basis is given by tensor products of orthonormal polynomials $(P_\nu^{(i)})_{\nu\in\mathbb{N}_0}$ in $L_2(\mathbb{R}, \mu_i)$ for $i \in \mathcal{N}$, so that each $u \in L_2(\mathbb{R}^{\mathcal{N}}, \mu)$ has an expansion of the form

$$u = \sum_{\nu\in\mathcal{F}} \mathbf{u}[\nu] \bigotimes_{i\in\mathcal{N}} P_{\nu_i}^{(i)}, \quad \text{where } \mathbf{u}[\nu] = \int_{\mathbb{R}^{\mathcal{N}}} u \bigotimes_{i\in\mathcal{N}} P_{\nu_i}^{(i)} \,\mathrm{d}\mu,\ \nu \in \mathcal{F},$$

with the index set

$$\mathcal{F} = \big\{\nu \in \mathbb{N}_0^{\mathcal{N}} \colon \nu_i \neq 0 \text{ for finitely many } i \in \mathcal{N}\big\}.$$

The coefficient sequences $\mathbf{u}$, restricted to suitable finite Cartesian product subsets of $\mathcal{F}$, can be approximated in low-rank formats. For a detailed treatment of such product polynomial expansions, we refer to Schwab and Gittelson (2011) and Cohen and DeVore (2015).

In many scenarios, the notion of orthonormal bases is too restrictive, and we instead work with bases satisfying the following weaker requirements.

Definition 3.4. Let H be a separable Hilbert space, let $\mathcal{I}$ be a countable set and let $\gamma_\nu \in H$ for $\nu \in \mathcal{I}$. Then $\{\gamma_\nu\}_{\nu\in\mathcal{I}}$ is a *Riesz sequence* if there exist $c, C > 0$ such that for all $\mathbf{v} \in \ell_2(\mathcal{I})$,

$$c\|\mathbf{v}\|_{\ell_2(\mathcal{I})} \leq \Big\|\sum_{\nu\in\mathcal{I}} \mathbf{v}[\nu]\,\gamma_\nu\Big\|_H \leq C\|\mathbf{v}\|_{\ell_2(\mathcal{I})}, \tag{3.7}$$

and a *Riesz basis* of H if in addition $\operatorname{span}\{\gamma_\nu \colon \nu \in \mathcal{I}\}$ is dense in H.

In other words, in addition to the density requirement, for a Riesz basis the mapping taking coefficient sequences to represented elements in the Hilbert space is an isomorphism; when $c = C = 1$, the Riesz basis is in fact an orthonormal basis, where this mapping is an *isometric* isomorphism.

When approximating an arbitrary $u = \sum_{\nu\in\mathcal{I}} \mathbf{u}[\nu]\,\gamma_\nu$ with coefficient sequence $\mathbf{u} \in \ell_2(\mathcal{I})$ by an expansion $\tilde{u} = \sum_{\nu\in\Lambda} \mathbf{u}[\nu]\,\gamma_\nu$ with $\Lambda \subset \mathcal{I}$, by (3.7) we have in particular

$$c\|\mathbf{u} - \tilde{\mathbf{u}}\|_{\ell_2} \leq \|u - \tilde{u}\|_H \leq C\|\mathbf{u} - \tilde{\mathbf{u}}\|_{\ell_2}, \tag{3.8}$$

which means that up to the fixed constants c, C, approximation in H-norm is equivalent to approximation of coefficient sequences in $\ell_2(\mathcal{I})$. This is particularly useful

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

in the context of low-rank methods, since the methods discussed in Section 2.7 for rank truncation of tensors based on the HSVD yield error bounds in ℓ_2-norm.

For Riesz bases in tensor product Hilbert spaces, we have the following analogue of Theorem 2.5 for orthonormal bases.

Proposition 3.5. For $i = 1, \dots, d$, let $\{\gamma_\nu^{(i)}\}_{\nu\in\mathbb{N}}$ be a Riesz basis of the separable Hilbert space H_i with

$$c_i\|\mathbf{v}\|_{\ell_2(\mathbb{N})} \le \left\|\sum_{\nu\in\mathbb{N}} \mathbf{v}[\nu]\,\gamma_\nu^{(i)}\right\|_{H_i} \le C_i\|\mathbf{v}\|_{\ell_2(\mathbb{N})} \quad \text{for all } \mathbf{v}\in\ell_2(\mathbb{N}).$$

Then $\{\gamma_{\nu_1}^{(1)}\otimes\cdots\otimes\gamma_{\nu_d}^{(d)}\}_{\nu\in\mathbb{N}^d}$ is a Riesz basis of $H = \bigotimes_{i=1}^d H_i$, and for all $\mathbf{v}\in\ell_2(\mathbb{N}^d)$,

$$\left(\prod_{i=1}^d c_i\right)\|\mathbf{v}\|_{\ell_2(\mathbb{N}^d)} \le \left\|\sum_{\nu\in\mathbb{N}^d} \mathbf{v}[\nu]\bigotimes_{i=1}^d \gamma_{\nu_i}^{(i)}\right\|_H \le \left(\prod_{i=1}^d C_i\right)\|\mathbf{v}\|_{\ell_2(\mathbb{N}^d)}.$$

The natural spaces for treating variational formulations of linear elliptic PDEs are the Hilbertian Sobolev spaces $H^k(\Omega)$ with $k\in\mathbb{N}$ on the given domain Ω, in particular $H^1(\Omega)$ in the case of second-order problems as in Example 3.1. In particular, the weak formulation of a problem of this type with homogeneous Dirichlet boundary conditions as in (3.2) can be written as an operator equation $Au = f$ for $u\in H_0^1(\Omega)$, where $f\in H^{-1}(\Omega)$, and $A\colon H_0^1(\Omega)\to H^{-1}(\Omega)$ is defined by

$$\langle Av, w\rangle = \int_\Omega \nabla v\cdot\nabla w\,\mathrm{d}x, \quad v, w\in H_0^1(\Omega). \tag{3.9}$$

Note that by the Lax–Milgram theorem, A is an isomorphism.

With the aid of a Riesz basis of $H_0^1(\Omega)$, such operator equations can be reduced to problems on ℓ_2-sequence spaces; see Cohen, Dahmen and DeVore (2001) and Dahmen (1997).

Proposition 3.6. Let V be a separable Hilbert space, let $A\colon V\to V'$ be an isomorphism satisfying $\|A\|_{V\to V'}\le C_A$ and $\|A^{-1}\|_{V'\to V}\le c_A^{-1}$ with $c_A, C_A > 0$, and let $\{\gamma_\nu\}_{\nu\in\mathcal{I}}$ be a Riesz basis of V such that with $c_\Gamma, C_\Gamma > 0$,

$$c_\Gamma\|\mathbf{v}\|_{\ell_2(\mathcal{I})} \le \left\|\sum_{\nu\in\mathcal{I}} \mathbf{v}[\nu]\,\gamma_\nu\right\|_H \le C_\Gamma\|\mathbf{v}\|_{\ell_2(\mathcal{I})} \quad \text{for all } \mathbf{v}\in\ell_2(\mathcal{I}).$$

In addition, let the bi-infinite matrix $\mathbf{A}\in\mathbb{R}^{\mathcal{I}\times\mathcal{I}}$ and the sequence $\mathbf{f}\in\mathbb{R}^{\mathcal{I}}$ be defined by

$$\mathbf{A}[\nu,\nu'] = \langle A\gamma_{\nu'},\gamma_\nu\rangle, \quad \mathbf{f}[\nu] = f(\gamma_\nu), \quad \text{for } \nu,\nu'\in\mathcal{I}. \tag{3.10}$$

Then $\mathbf{A}$ is an isomorphism on $\ell_2(\mathcal{I})$ with

$$c_A c_\Gamma^2\|\mathbf{v}\|_{\ell_2(\mathcal{I})} \le \|\mathbf{A}\mathbf{v}\|_{\ell_2(\mathcal{I})} \le C_A C_\Gamma^2\|\mathbf{v}\|_{\ell_2(\mathcal{I})} \quad \text{for all } \mathbf{v}\in\ell_2(\mathcal{I}). \tag{3.11}$$

Moreover, $\mathbf{f}\in\ell_2(\mathcal{I})$ and the unique solution $u\in V$ of $Au = f$ is given by $u = \sum_{\nu\in\mathcal{I}}\mathbf{u}[\nu]\gamma_\nu$, where $\mathbf{u}\in\ell_2(\mathcal{I})$ is the unique solution of $\mathbf{A}\mathbf{u} = \mathbf{f}$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Proof. With the isomorphism $S_\Gamma\colon \ell_2(\mathcal{I}) \to V$, $\mathbf{v} \mapsto \sum_{\nu\in\mathcal{I}} \mathbf{v}[\nu]\gamma_\nu$, we have $\mathbf{A} = S_\Gamma' A S_\Gamma$ and $\mathbf{f} = S_\Gamma' f$. Since $c_\Gamma\|\mathbf{v}\|_{\ell_2} \le \|S_\Gamma\mathbf{v}\|_V \le C_\Gamma\|\mathbf{v}\|_{\ell_2}$ for all $\mathbf{v} \in \ell_2(\mathcal{I})$ and $S_\Gamma'\colon V' \to \ell_2(\mathcal{I})$ is also an isomorphism with $c_\Gamma\|g\|_{V'} \le \|S_\Gamma' g\|_{\ell_2} \le C_\Gamma\|g\|_{V'}$ for all $g \in V'$, the statements follow. □

Remark 3.7. The spaces $H^1(\Omega)$ and $H^1_0(\Omega)$ for $\Omega = \square_d$ cannot be characterized as tensor product Hilbert spaces, which means that no simple tensor product constructions of Riesz bases as in Proposition 3.5 are applicable. Note that

$$\bigotimes_{i=1}^{d} H^1(0,1) = H^1_{\mathrm{mix}}(\square_d),$$

where the cross-norm defined by (2.4) reads

$$\|v\|^2_{H^1_{\mathrm{mix}}} = \sum_{\substack{\alpha\in\mathbb{N}_0^d \\ \max_i \alpha_i \le 1}} \int_{\square_d} \left|\partial^\alpha v\right|^2 \mathrm{d}x.$$

In contrast, $H^1(\square_d)$ can be obtained by taking the closure of $H^1_{\mathrm{mix}}(\square_d)$ with respect to the standard H^1-norm

$$\|v\|^2_{H^1} = \sum_{\substack{\alpha\in\mathbb{N}_0^d \\ \alpha_1+\cdots+\alpha_d \le 1}} \int_{\square_d} \left|\partial^\alpha v\right|^2 \mathrm{d}x, \tag{3.12}$$

and $H^1_{\mathrm{mix}} \subsetneq H^1$. The norm (3.12) can also be characterized as the canonical norm on an *intersection* of tensor product Hilbert spaces, that is,

$$H^1(\square_d) = \bigcap_{i=1}^{d} \bigotimes_{j=1}^{d} X_{ij}, \quad X_{ij} = \begin{cases} H^1(0,1) & i = j, \\ L_2(0,1) & i \ne j, \end{cases}$$

and the analogous characterization holds for the subspace $H^1_0(\square_d)$ with $H^1(0,1)$ replaced by $H^1_0(0,1)$. A detailed treatment of related questions is given in Ali and Nouy (2020*a*).

In order to still obtain Riesz bases of $H^1(\square_d)$ with some degree of product structure, one can consider tensor product bases of $L_2(\square_d)$ that become Riesz bases of $H^1(\square_d)$ by appropriate normalization.

Example 3.8. Let $\phi_j(x) = \sqrt{2}\,\sin(j\pi x)$ for $j \in \mathbb{N}$, $x \in (0,1)$. Then $\{\phi_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of $L_2(0,1)$ and by Theorem 2.5, setting $\Phi_j = \bigotimes_{i=1}^d \phi_{j_i}$ for $j \in \mathbb{N}^d$, $\{\Phi_j\}_{j\in\mathbb{N}^d}$ is an orthonormal basis of $L_2(\square_d)$. We easily verify that

$$\left\{\|\Phi_j\|^{-1}_{H^1}\Phi_j\right\}_{j\in\mathbb{N}^d}, \quad \text{where } \|\Phi_j\|_{H^1} = \sqrt{1+\pi^2\left(j_1^2+\cdots+j_d^2\right)}, \tag{3.13}$$

is a Riesz basis (in fact, even an orthonormal basis) of $H^1_0(\square_d)$ with the norm (3.12). Here the tensor product functions Φ_j are modified by the factor $\|\Phi_j\|^{-1}_{H^1}$, which

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

is *not separable* with respect to $j_1, \dots, j_d$. This lack of separability subsequently needs to be dealt with in any low-rank approximation method using the basis (3.13).

One construction of Riesz bases that is more suitable for locally refined approximations of potentially less regular functions is that of *wavelets*. A classical example of such a basis is the *Haar wavelet*: we define the *scaling function* φ and the *mother wavelet* ψ by

$$\varphi = \chi_{[0,1]}, \quad \psi = \chi_{[0,1/2]} - \chi_{(1/2,1)}$$

and, for $j \in \mathbb{N}_0$ and $k = 0, \dots, 2^j - 1$, we introduce the notation

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k). \tag{3.14}$$

Then, for each $J \in \mathbb{N}$,

$$\begin{aligned} &\operatorname{span}\big(\{\varphi_{0,0}\} \cup \{\psi_{j,k} : j \in \{0, \dots, J-1\}, k \in \{0, \dots, 2^j - 1\}\}\big) \\ &\qquad = \operatorname{span}\{\varphi_{J,k} : k \in \{0, \dots, 2^J - 1\}\}. \end{aligned} \tag{3.15}$$

Since the union over J of the spaces of piecewise constant functions on the right is dense in $L_2(0,1)$, and by the orthonormality properties of the functions defined in (3.14), we obtain that

$$\{\varphi_{0,0}\} \cup \{\psi_{j,k} : j \in \mathbb{N}_0, k \in \{0, \dots, 2^j - 1\}\}$$

is an orthonormal basis of $L_2(0,1)$. Due to their lack of regularity, Haar wavelets are not suitable for the construction of Riesz bases of H^1-spaces.

Based on more general hierarchies of nested spaces, referred to as *multiresolution analyses*, wavelet bases of any desired regularity can be constructed. Here we give a brief overview, and refer to the survey by Dahmen (1997) and the monographs by Daubechies (1992) and Cohen (2003) for further details.

The wavelets constructed by Daubechies (1988) provide orthonormal bases of $L_2(\mathbb{R})$ with compactly supported scaling function and mother wavelet, and they can be constructed to have arbitrarily high Hölder and Sobolev regularity (at the price of increasing support size). The Haar wavelet arises as the case of lowest order of this construction. The wavelets in this family of higher regularity, however, are fractals that are given only in terms of refinement relations, and are thus difficult to use in numerical methods.

The biorthogonal wavelets constructed by Cohen, Daubechies and Feauveau (1992) are also compactly supported and can be constructed to have arbitrarily high regularity, but are obtained as linear combinations of spline functions and are thus easy to handle numerically. However, the price to pay is that these wavelets yield non-orthonormal Riesz bases of $L_2(\mathbb{R})$. The use of such wavelets on bounded intervals requires modifications to basis elements that touch the boundary. Such adaptations to wavelet Riesz bases of $L_2(0,1)$ that are suitable for numerical methods have been constructed by Dahmen, Kunoth and Urban (1999) and Primbs (2010).

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

For such wavelet bases, we use the following customary generic notation: we write wavelet bases as $\{\psi_\nu\}_{\nu\in\vee}$, where $\vee$ denotes a suitable countable index set such that the functions $\{\psi_\nu\}_{\nu\in\vee}$ comprise both the scaling functions on the lowest level and the wavelets that form the basis. In addition, each ν includes a level j, denoted by $|\nu|$, and a translation parameter k in (3.14). Without loss of generality, for the lowest level in the basis, we assume $\min_{\nu\in\vee}|\nu| = 0$. The boundary adaptations mentioned above then lead to wavelet bases that in particular have the following properties: with some $t > 0$, for all $\nu \in \vee$,

$$\psi_\nu \in H^{1+t}(0,1), \quad \operatorname{diam}\operatorname{supp}\psi_\nu \lesssim 2^{-|\nu|}, \tag{3.16}$$

and there exist $c_0, c_1 > 0$ and $C_0, C_1 > 0$ such that for all $\mathbf{v} \in \ell_2(\vee)$,

$$c_0\|\mathbf{v}\|_{\ell_2(\vee)} \le \Big\|\sum_{\nu\in\vee}\mathbf{v}[\nu]\,\psi_\nu\Big\|_{L^2(0,1)} \le C_0\|\mathbf{v}\|_{\ell_2(\vee)} \tag{3.17}$$

as well as

$$c_1\|\mathbf{v}\|_{\ell_2(\vee)} \le \Big\|\sum_{\nu\in\vee}\mathbf{v}[\nu]\,\|\psi_\nu\|^{-1}_{H^1(0,1)}\,\psi_\nu\Big\|_{H^1(0,1)} \le C_1\|\mathbf{v}\|_{\ell_2(\vee)}. \tag{3.18}$$

Remark 3.9. In constructing Riesz bases of $H^1(\square_d)$ using properties (3.17), (3.18), we face a similar issue to that observed for finite elements in Remark 3.2: as shown by Dijkema, Schwab and Stevenson (2009), for the H^1-normalized tensor product wavelets

$$\Psi_\nu = \Big\|\bigotimes_{i=1}^d \psi_{\nu_i}\Big\|^{-1}_{H^1(\square_d)} \bigotimes_{i=1}^d \psi_{\nu_i}, \quad \nu \in \vee^d, \tag{3.19}$$

we obtain the bounds

$$c_1 c_0^{d-1}\|\mathbf{v}\|_{\ell_2(\vee^d)} \le \Big\|\sum_{\nu\in\vee^d}\mathbf{v}[\nu]\,\Psi_\nu\Big\|_{H^1(\square_d)} \le C_1 C_0^{d-1}\|\mathbf{v}\|_{\ell_2(\vee^d)} \tag{3.20}$$

for all $\mathbf{v} \in \ell_2(\vee^d)$. Similarly to Remark 3.2, one can show that whenever $c_0 < C_0$, the bounds in (3.20) and in the equivalence of errors (3.8) deteriorate exponentially with respect to d. In general, this equally affects the condition numbers of operator representations (3.11).

The above observation shows that to obtain a wavelet Riesz basis of $H^1(\square_d)$ with d-robust constants we essentially need to start from an orthonormal wavelet basis of $L_2(0,1)$ satisfying (3.17) with $c_0 = C_0 = 1$. By the wavelet constructions mentioned above, we do not obtain the regularity and support properties (3.16) at the same time as orthonormality in $L_2(0,1)$ when using wavelets that are at the same time piecewise polynomials. However, this is achieved by certain spline *multiwavelets*, which are generated by dilations and translations of more than one mother wavelet, constructed by Donovan, Geronimo and Hardin (1996, 1999). For such multiwavelets we keep the above notation, where each index in $\vee$ now also

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

encodes the respective mother wavelet. The Riesz basis $\{\Psi_\nu\}_{\nu\in\vee^d}$ of $H^1(\square_d)$ obtained from these spline multiwavelets according to (3.19) then satisfies (3.20) with $c_0 = C_0 = 1$, and thus with constants independent of d.

Remark 3.10. If, in addition to (3.16), (3.17) with $c_0 = C_0 = 1$, and (3.18), we have $\psi_\nu \in H^1_0(0,1)$ for all $\nu \in \vee$, then with

$$\Psi_\nu = \left(\sum_{i=1}^d \|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)^{-1/2} \bigotimes_{i=1}^d \psi_{\nu_i},$$

$\{\Psi_\nu\}_{\nu\in\vee^d}$ is a Riesz basis of $H^1_0(\square_d)$.

For the isomorphism $A\colon H^1_0(\square_d) \to H^{-1}(\square_d)$ defined by the Laplacian in (3.9), let us now consider basis representations as in (3.10) with respect to such a Riesz basis. By (3.11) and (3.20), for $\mathbf{A}\colon \ell_2(\vee^d) \to \ell_2(\vee^d)$ defined by

$$\mathbf{A}[\nu,\nu'] = \langle A\Psi_{\nu'}, \Psi_\nu\rangle = \int_{\square_d} \nabla\Psi_{\nu'}\cdot\nabla\Psi_\nu \,\mathrm{d}x, \quad \nu,\nu' \in \vee^d, \tag{3.21}$$

we then have $c_1\|\mathbf{v}\|_{\ell_2} \le \|\mathbf{A}\mathbf{v}\|_{\ell_2} \le C_1\|\mathbf{v}\|_{\ell_2}$ with c_1, C_2 from (3.18).

Moreover, we have the decomposition $\mathbf{A} = \mathbf{D}\mathbf{T}\mathbf{D}$ with the infinite diagonal matrix

$$\mathbf{D}[\nu,\nu'] = \left(\sum_{i=1}^d \|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)^{-1/2} \delta_{\nu,\nu'}, \quad \nu,\nu' \in \vee^d,$$

and

$$\mathbf{T}[\nu,\nu'] = \sum_{i=1}^d \int_0^1 \psi'_{\nu_i}\psi'_{\nu'_i}\,\mathrm{d}x \prod_{j\ne i}\int_0^1 \psi_{\nu_j}\psi_{\nu'_j}\,\mathrm{d}x = \sum_{i=1}^d \int_0^1 \psi'_{\nu_i}\psi'_{\nu'_i}\,\mathrm{d}x \prod_{j\ne i}\delta_{\nu_j,\nu'_j},$$

so that with $\mathbf{T}_1[\nu,\nu'] = \int_0^1 \psi'_\nu\psi'_{\nu'}\,\mathrm{d}x$ for $\nu,\nu' \in \vee$,

$$\mathbf{T} = \mathbf{T}_1\otimes\mathbf{I}\otimes\cdots\otimes\mathbf{I} + \cdots + \mathbf{I}\otimes\cdots\otimes\mathbf{I}\otimes\mathbf{T}_1. \tag{3.22}$$

Here we have used the L_2-orthonormality of $\{\psi_\nu\}_{\nu\in\vee}$. Note that $\mathbf{D}$ can be regarded as a two-sided diagonal preconditioner for the operator $\mathbf{T}$, which is unbounded on $\ell_2(\vee^d)$. For the operator equation $Au = f$ with $f \in H^{-1}(\square_d)$, the basis representation $\mathbf{A}\mathbf{u} = \mathbf{f}$ can be written in the form

$$\mathbf{D}\mathbf{T}\mathbf{D}\mathbf{u} = \mathbf{D}\mathbf{g}, \quad \text{where } \mathbf{g} = \left(f\left(\bigotimes_{i=1}^d \psi_{\nu_i}\right)\right)_{\nu\in\vee^d}. \tag{3.23}$$

As already observed in Section 3.1.1, operators of the form (3.22) can be represented with all ranks equal to two in Tucker, tensor train or hierarchical tensor formats. However, similarly to Example 3.8, $\mathbf{D}$ does not have an explicit low-rank

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

representation, which reflects the lack of tensor structure of $H^1_0(\square_d)$ noted in Remark 3.7. Note that since $\|\psi_\nu\|_{H^1(0,1)} \eqsim 2^{|\nu|}$ for the wavelets under consideration,

$$\mathbf{D}[\nu,\nu] = \left(\sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)^{-1/2} \eqsim \left(\sum_{i=1}^{d} 2^{2|\nu_i|}\right)^{-1/2} \eqsim 2^{-\max_i |\nu_i|} \tag{3.24}$$

uniformly for all $\nu \in \vee^d$. Approaches for obtaining efficient low-rank approximations of $\mathbf{D}$ are considered in Section 4.

Remark 3.11. For $\{\Psi_\nu\}_{\nu\in\vee^d}$ constructed from L_2-orthonormal univariate wavelets as above, by (3.8), for any $u, \tilde{u} \in H^1_0(\square_d)$ with

$$u = \sum_{\nu\in\vee^d} \mathbf{u}[\nu]\Psi_\nu, \quad \tilde{u} = \sum_{\nu\in\vee^d} \tilde{\mathbf{u}}[\nu]\Psi_\nu,$$

we have in particular

$$c_1\|\mathbf{u}-\tilde{\mathbf{u}}\|_{\ell_2} \le \|u-\tilde{u}\|_{H^1} \le C_1\|\mathbf{u}-\tilde{\mathbf{u}}\|_{\ell_2}.$$

This means that rank truncations of coefficient tensors can be performed by standard HSVD-based procedures with controlled H^1-error. Without the use of a Riesz basis, error control in H^1 could in principle also be achieved by different means, for instance direct optimization without the aid of the SVD (which is in general difficult to do reliably), or by alternative approaches considered in Ali and Nouy (2020*a*,*b*) which, however, involve additional restrictions.

The preconditioning of operators provided by Riesz basis representations also yields a proportionality between solution error in H^1 and residual: with A and $\mathbf{A}$ as in (3.21), if u is the solution of $Au = f \in H^{-1}(\square_d)$ with basis representation $\mathbf{Au} = \mathbf{f}$, then

$$\|u-\tilde{u}\|_{H^1} \eqsim \|\mathbf{A}(\mathbf{u}-\tilde{\mathbf{u}})\|_{\ell_2} = \|\mathbf{f}-\mathbf{A}\tilde{\mathbf{u}}\|_{\ell_2}$$

with constants independent of d. This means that the total approximation error in H^1 with respect to exact solutions of the considered PDE can be estimated from finitely supported approximations of the residuals $\mathbf{f} - \mathbf{A}\tilde{\mathbf{u}}$, which generally have infinite support in $\vee^d$.

Remark 3.12. Adaptively refined Riesz basis expansions are used in combination with hierarchical tensor representations in infinite-dimensional Hilbert spaces in Bachmayr (2012*a*), Bachmayr and Dahmen (2015) and subsequently in Bachmayr and Dahmen (2016*b*) and Ali and Urban (2020). A similar formalism focused on continuous variables rather than basis expansions has been developed independently in Bigoni, Engsig-Karup and Marzouk (2016) based on the tensor train format, termed *functional tensor train* representation. This is combined with heuristic adaptivity based on spectral basis functions in Gorodetsky, Karaman and Marzouk (2019); see also Dolgov, Kressner and Strössner (2021).

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

3.3. Occupation number representations

In the Schrödinger equations (1.2) and (1.3), solutions describing systems with N particles are functions depending on N particle coordinates. We additionally need to ensure *symmetry* of solutions under exchange of particle coordinates when considering bosonic particles (due to their indistinguishability), or *antisymmetry* in case of fermionic particles (corresponding to the Pauli exclusion principle).

One case of particular importance in chemistry is the *electronic Schrödinger equation*, as a special case of the eigenvalue problem (1.3) for a Hamiltonian operator H of the basic form

$$H = -\frac{1}{2}\Delta + \sum_i V_1(x_i) + \sum_{i<j} V_2(x_i, x_j), \tag{3.25}$$

where x_i, $i = 1, \dots, N$ are the coordinates of each particle and V_2 is a symmetric function. Hamiltonians of this structure arise frequently in quantum physics. In the case of the electronic Schrödinger case, V_2 takes the form of a Coulomb interaction between electrons, and V_1 denotes external Coulomb potentials of atomic nuclei. The associated eigenfunctions, referred to as *wavefunctions*, describe the quantum states of electrons for given nuclear positions. Unless the model under consideration has certain special features (such as the presence of magnetic fields), the wavefunctions are real-valued. Since electrons are fermions, they are required to be antisymmetric under exchange of each pair of electron coordinates. Here each electron has coordinates in $\Omega = \mathbb{R}^3 \times \mathbb{Z}_2$, where the binary degree of freedom in $\mathbb{Z}_2$ accounts for electron spin.

The hierarchical tensor format is not well suited to the direct representation of antisymmetric functions (Hackbusch 2018). However, this issue can be circumvented by an indirect parametrization provided by the formalism of *second quantization* that is widely used in quantum physics and chemistry. The starting point for this formalism is an orthonormal basis $\{\phi_\nu\}_{\nu\in\mathbb{N}}$ of $L_2(\Omega)$. These basis functions are also referred to as *orbitals*. From these functions, we form antisymmetrized tensor products, called *Slater determinants*, for all $\nu_1, \dots, \nu_N \in \mathbb{N}$ with $\nu_1 < \nu_2 < \cdots < \nu_N$,

$$\begin{aligned}\Phi_{\nu_1,\dots,\nu_N}(x_1, \dots, x_N) &= \frac{1}{\sqrt{N!}} \sum_{\pi\in S_N} \operatorname{sgn}(\pi) \prod_{i=1}^{N} \phi_{\nu_i}(x_{\pi(i)}) \\ &= \frac{1}{\sqrt{N!}} \det\big(\phi_{\nu_i}(x_j)\big)_{i,j=1,\dots,N}\end{aligned}$$

for $x_1, \dots, x_N \in \Omega$, which provide an orthonormal basis of the subspace $\mathcal{V}_N$ of antisymmetric functions in $L_2(\Omega^N)$.

We now restrict to a subset of $K \geq N$ orbitals $\phi_1, \dots, \phi_K$. These finitely many orbitals may also be optimized for the problem under consideration; see for instance Krumnow, Veis, Legeza and Eisert (2016). The corresponding Slater determinants

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

are an orthonormal basis of the subspace

$$\mathcal{V}_N^K = \operatorname{span}\{\Phi_\nu : \nu \in \{1,\dots,K\}^N, \nu_1 < \cdots < \nu_N\} \subset \mathcal{V}_N.$$

We next introduce an injective mapping from $\mathcal{V}_N^K$ to the space

$$\mathcal{F}^K = \bigotimes_{i=1}^K \mathbb{R}^2.$$

To this end, let $\mathbf{e}_0 = (1,0)^\top, \mathbf{e}_1 = (0,1)^\top \in \mathbb{R}^2$. For each single Slater determinant $\Phi_\nu \in \mathcal{V}_N^K$ and scalar coefficient c_ν, we define

$$c_\nu \Phi_\nu \;\mapsto\; c_\nu \bigotimes_{i=1}^K \mathbf{e}_{b_i(\nu)}, \qquad \text{where } b_i(\nu) = \begin{cases} 1 & i \in \{\nu_1,\dots,\nu_d\}, \\ 0 & \text{otherwise,} \end{cases}$$

and extend this by linearity to a mapping from $\mathcal{V}_N^K$ to $\mathcal{F}^K$. We obtain an identification of linear combinations of Slater determinants with tensors in $\mathcal{F}^K$ that encode the occupations of orbitals in each Slater determinant. This representation is thus also called *occupation number representation*. It has the advantage that the elements of $\mathcal{F}^K$, referred to as occupation numbers, do not need to satisfy any antisymmetry requirement.

Operators on $\mathcal{F}^K$ can be expressed in terms of the *annihilation operators* $\mathbf{a}_i$ and *creation operators* $\mathbf{a}_i^*$ for $i = 1,\dots,K$, which are given by

$$\mathbf{a}_i = \mathbf{S}^{\otimes(i-1)} \otimes \mathbf{J} \otimes \mathbf{I}^{\otimes(K-i)}, \quad \mathbf{a}_i^* = \mathbf{S}^{\otimes(i-1)} \otimes \mathbf{J}^\top \otimes \mathbf{I}^{\otimes(K-i)}$$

in terms of the components

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.26}$$

In the occupation number representation, antisymmetry of Slater determinants is reflected by the anticommutation relations satisfied by annihilation and creation operators: for $i, j \in \{1,\dots,K\}$,

$$\mathbf{a}_i \mathbf{a}_j^* + \mathbf{a}_j^* \mathbf{a}_i = \delta_{ij}, \quad \mathbf{a}_i^* \mathbf{a}_j^* + \mathbf{a}_j^* \mathbf{a}_i^* = \mathbf{a}_i \mathbf{a}_j + \mathbf{a}_j \mathbf{a}_i = 0. \tag{3.27}$$

Making use of the antisymmetry properties of Slater determinants, which lead to the so-called *Slater–Condon rules* (Helgaker, Jørgensen and Olsen 2000, Chap. 1), for H as in (3.25) we obtain the representation

$$\mathbf{H} = \sum_{i,j=1}^K t_{ij} \mathbf{a}_i^* \mathbf{a}_j + \sum_{i,j,k,l=1}^K v_{ijkl} \mathbf{a}_i^* \mathbf{a}_j^* \mathbf{a}_k \mathbf{a}_l \tag{3.28}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

with the coefficients

$$t_{ij} = \int \varphi_i(x)\big\{-\tfrac{1}{2}\Delta + V_1(x)\big\}\varphi_j(x)\,\mathrm{d}x,$$
$$v_{ijkl} = \int \varphi_i(x)\,\varphi_j(y)\,V_2(x,y)\,\varphi_k(x)\,\varphi_l(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Note that the number of particles is not fixed *a priori* for elements of $\mathcal{F}^K$ and thus needs to be added as a constraint in this formulation. The occupation numbers $\mathbf{x} \in \mathcal{F}^K$ that represent wavefunctions with N particles are precisely the eigenfunctions with eigenvalue N of the *particle number operator*

$$\mathbf{P} = \sum_{i=1}^{K} \mathbf{a}_i^* \mathbf{a}_i, \tag{3.29}$$

that is, they lie in the subspace

$$\mathcal{F}_N^K = \{\mathbf{x} \in \mathcal{F}^K : \mathbf{P}\mathbf{x} = N\mathbf{x}\}.$$

The Ritz–Galerkin approximation in $\mathcal{V}_N^K$ of the original eigenvalue problem for H is then equivalent to

$$\mathbf{H}\mathbf{x} = \lambda\mathbf{x}, \quad \text{where } \mathbf{x} \in \mathcal{F}_N^K, \mathbf{x} \neq 0, \tag{3.30}$$

and in particular, the approximation in $\mathcal{V}_N^K$ of the lowest eigenvalue of H is characterized by

$$\min_{0\neq\mathbf{x}\in\mathcal{F}_N^K} \frac{\langle \mathbf{H}\mathbf{x}, \mathbf{x}\rangle}{\langle \mathbf{x}, \mathbf{x}\rangle}.$$

The problem (3.30) in occupation number representation is directly amenable to approximations in hierarchical tensor format, in this context typically referred to as tree tensor networks. A frequent choice is tensor trains, in this context known as matrix product states, corresponding to representing or approximating the sought $\mathbf{x} \in \mathcal{F}^K$ in the form

$$\mathbf{x} = \mathbf{X}_1 \bowtie \cdots \bowtie \mathbf{X}_K \tag{3.31}$$

as in (2.17). For a more detailed survey of this approach, which has been successfully applied in quantum chemistry and many other applications of quantum physics, we refer to Szalay *et al.* (2015). Note that in this formulation the dimensionality N of the original problem enters only indirectly, and the main parameters determining the computational complexity are the ranks and the number of orbitals K.

Remark 3.13. It is clear that $\mathbf{H}$ can be represented with ranks $O(K^4)$, since each operator $\mathbf{a}_i^*\mathbf{a}_j^*\mathbf{a}_k\mathbf{a}_l$ is of rank one. However, this is not sharp: as observed in Crosswhite and Bacon (2008) and Keller, Dolfi, Troyer and Reiher (2015), $\mathbf{H}$ can generically be represented with ranks $O(K^2)$. For a detailed mathematical treatment, see Bachmayr, Götte and Pfeffer (2022); further rank reductions are possible, for instance with appropriately localized orbitals.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Remark 3.14. Since (3.31) yields representations of general elements of $\mathcal{F}^K$, the constraint to the N-particle eigenspace $\mathcal{F}^K_N$ of $\mathbf{P}$ needs to be handled appropriately. There are certain cases where this constraint does not need to be enforced, since it is automatically satisfied by the sought eigenvectors of $\mathbf{H}$. In general, however, $\mathbf{x} \in \mathcal{F}^K$ is in $\mathcal{F}^K_N$ if and only if it has a representation (3.31) where each of the cores $\mathbf{X}_i$, $i = 1, \dots, K$, has a certain *block-sparse structure*. In a physics context, this observation goes back to Singh, Pfeifer and Vidal (2011). A mathematical proof, extending this to more general operators than $\mathbf{P}$ with similar structure, is given in Bachmayr *et al.* (2022). As a consequence, membership in $\mathcal{F}^K_N$ can be enforced by working directly on the non-zero blocks of such block-sparse representations. The operator $\mathbf{H}$ can be applied to such block-sparse representations in a matrix-free manner that preserves the block-sparse structure (Bachmayr *et al.* 2022, Sec. 5.2).

3.4. Multilevel tensorized representations

We now turn to a different way of approximating solutions of PDEs with low-rank tensors that is aimed at highly compressed approximations of simple discretizations. Here grid vectors are efficiently encoded in low-rank format such that the use of extremely fine uniform grids becomes feasible. For PDEs on domains of moderate dimensionality, this leads to generic approximations for problems with singularities or multiscale structure that otherwise require specialized approaches. In this case tensor modes correspond to *scales* in the approximation problem.

The basic idea is that hierarchical tensor representations can be applied to arbitrary vectors when these are reinterpreted as high-order tensors. This is achieved by a suitable encoding of indices. For instance, for $L \in \mathbb{N}$, we introduce the isomorphism

$$\mathcal{J} \colon \{0,1\}^L \to \{1, \dots, 2^L\}, \quad (b_1, \dots, b_L) \mapsto \sum_{\ell=1}^{L} b_\ell 2^{L-\ell} + 1. \tag{3.32}$$

For a given vector $\mathbf{v}$ of length 2^L, we define the tensor $\mathbf{u} \in \bigotimes_{\ell=1}^{L} \mathbb{R}^2$ indexed by $\{0,1\}^L$ as

$$\mathbf{u}[b_1, \dots, b_L] = \mathbf{v}[\mathcal{J}(b_1, \dots, b_L)], \quad b \in \{0,1\}^L. \tag{3.33}$$

In an analogous manner, a vector in $\mathbb{R}^N$ with $N = \prod_{\ell=1}^{L} n_\ell$ can be mapped to a tensor of mode sizes $n_1 \times \cdots \times n_L$.

We can then approximate $\mathbf{u}$ in a suitable tensor format to obtain a compressed version of $\mathbf{v}$. Specifically, we can obtain a tensor train representation $\mathsf{V} = (\mathbf{V}_1, \dots, \mathbf{V}_L)$ of $\mathbf{u}$ of the form

$$\mathbf{u} = \tau(\mathsf{V}) = \mathbf{V}_1 \bowtie \mathbf{V}_2 \bowtie \cdots \bowtie \mathbf{V}_L. \tag{3.34}$$

Since individual modes are associated to scales, such a representation can be interpreted as a type of *multilevel TT representation* of the underlying function. In the

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

literature this is referred to as *quantized tensor train* (QTT) or as *tensorized* representation, introduced by Khoromskij (2011) and Grasedyck (2010). Precursors can be found in similar ideas based on canonical tensor representations by Tyrtyshnikov (2003) and in numerical observations on tensor train representations of matrices by Oseledets (2009*a*). Detailed accounts of these techniques are given in Khoromskij (2018) and Hackbusch (2019, Chap. 14).

Example 3.15. Consider the vector $\mathbf{v} \in \mathbb{R}^N$ with $N = 2^L$, $L \in \mathbb{N}$, defined by $\mathbf{v}[j] = \exp(\lambda x_j)$ with $\lambda \in \mathbb{C}$ and $x_j = (j-1)/N$ for $j = 1, \ldots, N$, corresponding to grid values of an exponential function on $[0, 1]$. Then, as observed by Khoromskij (2011), for the tensor $\mathbf{u}$ defined in (3.33), we have the rank-one representation

$$\mathbf{u} = \bigotimes_{\ell=1}^{L} \begin{pmatrix} 1 \\ \exp(\lambda 2^{L-\ell}/N) \end{pmatrix},$$

which requires only $2L$ coefficients as opposed to the 2^L entries of the vector $\mathbf{v}$. As an immediate consequence, grid values of trigonometric functions have representations (3.34) of ranks at most two.

Remark 3.16. A variety of further results on representations and approximations of more general functions in the multilevel form (3.34) have been obtained, including the following.

(i) Values of polynomials of degree p on uniform grids can be represented in analogous form with ranks bounded by $p + 1$; see Grasedyck (2010) and Khoromskij (2011).

(ii) In Grasedyck (2010), this representation of polynomials is applied on subintervals in combination with local refinement to obtain approximations of the form (3.34) of functions on intervals with isolated singularities. These approximations converge exponentially, uniformly on an equispaced grid of points, with respect to the number of parameters in the representation, similarly to hp-type piecewise polynomial approximations.

(iii) Explicit representations (3.34) of continuous multivariate functions, with functions as components, have been constructed in Oseledets (2013).

(iv) A general formulation of tensorized approximations on a function space level has been developed in Ali and Nouy (2023), where convergence rates for functions with Besov regularity are also obtained.

Various transformations with particular structure have been shown to be efficiently realizable for multilevel representations as in (3.34), including convolutions (Hackbusch 2011), Toeplitz matrices (Kazeev, Khoromskij and Tyrtyshnikov 2013*a*) and wavelet transforms (Khoromskij and Miao 2014).

Finding representation (3.34) as approximate solutions of differential equations requires suitable representations of discretized differential operators in a compatible

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

format. In the case of the second derivative on an interval, we have the following basic result.

Example 3.17. As shown by Kazeev and Khoromskij (2012), the matrix $\mathbf{\Delta}_2$ defined in (3.4) that appears in the finite difference or finite element discretization of the second derivative can be represented as $\mathbf{L}_1 \bowtie \cdots \bowtie \mathbf{L}_L$, where

$$\mathbf{L}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{J}^\top & \mathbf{J} \end{bmatrix}, \quad \mathbf{L}_2 = \cdots = \mathbf{L}_{L-1} = \begin{bmatrix} \mathbf{I} & \mathbf{J}^\top & \mathbf{J} \\ 0 & \mathbf{J} & 0 \\ 0 & 0 & \mathbf{J}^\top \end{bmatrix}, \quad \mathbf{L}_L = \begin{bmatrix} 2\mathbf{I} - \mathbf{J} - \mathbf{J}^\top \\ -\mathbf{J} \\ -\mathbf{J}^\top \end{bmatrix}$$

with component matrices as in (3.26). This representation can be applied to grid vectors represented in the format (3.34).

Concerning the operator representation condition number, as verified in Bachmayr and Kazeev (2020, Prop. 4.10), the above representation $\mathsf{L} = (\mathbf{L}_1, \dots, \mathbf{L}_L)$ satisfies

$$\max_{\ell=1,\dots,L} \operatorname{oprcond}_\ell(\mathsf{L}) \eqsim 2^{2L}.$$

This means that applying the representation L to some V as in (3.34) may generate redundancies leading to cancellations (as in Example 2.43) with a loss of relative precision by a factor proportional to 2^{2L}. This is indeed confirmed by numerical tests (Bachmayr and Kazeev 2020, Ex. 4.3), which means that the direct application of L is limited to moderately large L (where with double precision, all digits of accuracy are lost when $L \approx 25$). This issue, which affects such representations of differential operators in general, can be dealt with by suitable preconditioners that also improve the representation condition number of the representation, rather than just the usual condition number of the represented matrix. A construction of such preconditioners is considered in Section 4.5.

Remark 3.18. The concepts outlined above for univariate functions can also be applied in the multivariate case. For problems in d dimensions, one can devise an isomorphism analogously to (3.32) but depending on parameters $(b_{i,1}, \dots, b_{i,L})$ for $i = 1, \dots, d$. There are two canonical ways of ordering this larger set of parameters in an isomorphism as in (3.32): either the modes in the TT representation are ordered in the form

$$(b_{1,1}, \dots, b_{1,L},\ b_{2,1}, \dots, b_{2,L},\ \dots,\ b_{d,1}, \dots, b_{d,L}),$$

which has been used, for instance, in Khoromskij and Oseledets (2011), closely related to a standard tensor representation in d dimensions where multilevel TT representations are applied in each mode (Dolgov and Khoromskij 2013); or the modes are arranged in the so-called *transposed ordering*

$$(b_{1,1}, b_{2,1}, \dots, b_{d,1},\ b_{1,2}, b_{2,2}, \dots, b_{d,2},\ \dots,\ b_{1,L}, \dots, b_{d,L}),$$

used, for instance, by Kazeev and Schwab (2018), which corresponds to a parametrization in terms of a hierarchy of d-dimensional patches. This latter form has

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

proved to be particularly suitable for efficient approximations in dimensions $d = 2$ or $d = 3$; see also Ali and Nouy (2021).

Corresponding representations of operators in the general d-dimensional case are studied in Khoromskij and Oseledets (2011), Dolgov and Khoromskij (2013), Kazeev, Reichmann and Schwab (2013*b*) and Bachmayr and Kazeev (2020). Applications to elliptic boundary value problems are considered in further detail in Section 5.5.3.

4. Approximate inverses and preconditioning of elliptic operators

For bounded elliptic operators $A : V \to V'$ on Hilbert spaces V, where we focus in particular on the case $V = H^1_0(\square_d)$ as in (3.9), in this section we consider two closely related tasks. The first is finding approximate inverses of A, that is, approximating A^{-1} by low-rank expressions in suitable operator norms. Such approximations can be used for the direct construction of approximate solutions in low-rank format; however, they are applicable only to operators with particular structure such as the Laplacian-type form (1.6).

The second is the construction of preconditioners, where for a discretization or basis representation $\mathbf{A}$ of A, we aim to find $\mathbf{C}$ such that for all $\mathbf{v}$,

$$c\langle \mathbf{C}\mathbf{v}, \mathbf{v}\rangle \le \langle \mathbf{A}^{-1}\mathbf{v}, \mathbf{v}\rangle \le C\langle \mathbf{C}\mathbf{v}, \mathbf{v}\rangle$$

with uniform constants $c, C > 0$, where $\mathbf{C}$ can be applied efficiently. Although this second aim of establishing a spectral equivalence can be accomplished by sufficiently accurate approximate inverses, more direct constructions of preconditioners can generally provide more efficiency, in particular concerning more favourable representation ranks. Low-rank preconditioners are again obtained for operators with Laplacian-type structure, but are equally applicable to more general elliptic operators.

4.1. Exponential sums

Let us first consider the case of a tensor product Hilbert space $H = \bigotimes_{i=1}^d H_i$, with $A : \operatorname{dom}(A) \subseteq H \to H$ a potentially unbounded densely defined, self-adjoint, positive definite operator with Laplacian-type structure as in (1.6), that is,

$$A = \sum_{i=1}^d \bigotimes_{j=1}^d B_{i,j}, \quad B_{i,j} = \begin{cases} A_i, & i = j, \\ \mathrm{I}, & i \ne j, \end{cases} \tag{4.1}$$

with densely defined, self-adjoint and positive definite $A_i : \operatorname{dom}(A_i) \subseteq H_i \to H_i$. An instance of such an operator would be the Laplacian regarded as an unbounded operator on $L_2(\square_d)$, but we will also apply the following to operators on $\ell_2(\mathbb{N}^d)$ arising from basis representations. The observation that is fundamental henceforth

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

is that for $\alpha \geq 0$,

$$e^{-\alpha A} = \bigotimes_{i=1}^{d} e^{-\alpha A_i}.$$

The basic strategy is now to find $\omega_k, \alpha_k > 0$ for $k = 1, \dots, r$ such that

$$t^{-1} \approx \sum_{k=1}^{r} \omega_k e^{-\alpha_k t} \quad \text{for } t \in \sigma(A) \tag{4.2}$$

in an appropriate sense, and by applying this exponential sum approximation to A via its spectral decomposition to conclude

$$A^{-1} \approx \sum_{k=1}^{r} \omega_k e^{-\alpha_k A} = \sum_{k=1}^{r} \omega_k \bigotimes_{i=1}^{d} e^{-\alpha_k A_i},$$

which provides an approximation in terms of a sum of r tensor product operators. By the continuous spectral calculus for self-adjoint operators (or, in the finite-dimensional case, the eigenvalue decomposition of A), we have the following estimates.

Proposition 4.1. Let A be as in (4.1) and let $f, \tilde{f} \colon \mathbb{R}^+ \to \mathbb{R}^+$ be continuous functions on an interval containing the spectrum $\sigma(A)$.

(i) Under these conditions, we have the estimate

$$\|f(A) - \tilde{f}(A)\|_{H \to H} \leq \sup_{t \in \sigma(A)} |f(t) - \tilde{f}(t)|.$$

(ii) Assume that for some $\delta \in (0, 1)$ we have $|f(t) - \tilde{f}(t)| \leq \delta f(t)$ for all $t \in \sigma(A)$. Then

$$(1 - \delta)\langle f(A)v, v\rangle \leq \langle \tilde{f}(A)v, v\rangle \leq (1 + \delta)\langle f(A)v, v\rangle \quad \text{for all } v \in H.$$

For related results under slightly different assumptions, see also Hackbusch (2014, Sec. 5.9).

4.1.1. Inverse Laplace transform and quadrature approximation
A first construction principle for exponential sum approximations as in (4.2) for a given function f, such as $f(t) = t^{-1}$, is applying a quadrature rule to its representation as a Laplace transform of a function g,

$$f(t) = \int_0^\infty e^{-st} g(s)\, \mathrm{d}s.$$

This can be done by *sinc quadrature* methods: after applying a suitable substitution $\sigma \colon \mathbb{R} \to \mathbb{R}^+$, we use the trapezoidal rule on $\mathbb{R}$ with step size $\tau > 0$ to obtain the approximations

$$f(t) = \int_{-\infty}^{\infty} e^{-\sigma(x)t} g\big(\sigma(x)\big)\sigma'(x)\, \mathrm{d}x \approx \tau \sum_{k \in \mathbb{Z}} e^{-\sigma(k\tau)t} g\big(\sigma(k\tau)\big)\sigma'(k\tau). \tag{4.3}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

We then exploit decay properties of the integrand to truncate the sum on the right.

For $f(t) = t^{-\lambda}$ with $\lambda > 0$, the following result of this kind, which provides the conditions of Proposition 4.1(ii), is obtained in Beylkin and Monzón (2010); see also Beylkin and Monzón (2005) and Scholz and Yserentant (2017).

Theorem 4.2. Let $\tau, \lambda > 0$ and

$$\delta(\tau, \lambda) = \frac{1}{\Gamma(\lambda)} \sum_{\substack{n \in \mathbb{Z} \\ n \neq 0}} |\Gamma(\lambda + 2\pi i n/\tau)|.$$

Then, for all $t > 0$,

$$\left| \frac{1}{t^\lambda} - \frac{\tau}{\Gamma(\lambda)} \sum_{k \in \mathbb{Z}} e^{\lambda k \tau} \exp\bigl(-e^{k\tau} t\bigr) \right| \leq \frac{\delta(\tau, \lambda)}{t^\lambda}, \tag{4.4}$$

and for each $\hat{\delta} \in (0, 1]$, if

$$\tau \leq \frac{2\pi}{\ln 3 + \lambda |\ln(\cos 1)| + |\ln \hat{\delta}|},$$

then $\delta(\tau, \lambda) \leq \hat{\delta}$.

We thus obtain approximations on all of $(0, \infty)$ with infinitely many exponential terms. In the restriction to finitely many summands, we arrive at approximations with prescribed relative error on finite intervals as required in Proposition 4.1(ii). The following result is adapted from Bachmayr and Dahmen (2016*b*).

Corollary 4.3. Let $\lambda \in (0, 1]$, $\delta = \delta_0 + \eta < 1$ with $\delta_0, \eta \in (0, 1)$, and $T > 1$. With

$$\begin{aligned} \tau &= \frac{2\pi}{\ln 3 + \lambda |\ln(\cos 1)| + |\ln(\delta_0/2)|}, \\ m &= \bigl\lceil \tau^{-1} \ln |\ln(\delta_0/2)| \bigr\rceil, \quad n = \bigl\lceil \tau^{-1} \bigl(\lambda^{-1} |\ln \eta| + \ln T \bigr) \bigr\rceil, \end{aligned} \tag{4.5}$$

define the truncated exponential sums

$$S_{\lambda, \delta_0}(t) = \frac{\tau}{\Gamma(\lambda)} \sum_{k=-\infty}^{m} e^{\lambda k \tau} \exp\bigl(-e^{k\tau} t\bigr),$$

$$S_{\lambda, \delta_0, \eta, T}(t) = \frac{\tau}{\Gamma(\lambda)} \sum_{k=-n}^{m} e^{\lambda k \tau} \exp\bigl(-e^{k\tau} t\bigr).$$

Then

$$|t^{-\lambda} - S_{\lambda, \delta_0}(t)| \leq \delta_0 t^{-\lambda}, \quad \text{for all } t \in [1, \infty), \tag{4.6a}$$

$$|S_{\lambda, \delta_0}(t) - S_{\lambda, \delta_0, \eta, T}(t)| \leq \eta t^{-\lambda}, \quad \text{for all } t \in [1, T], \tag{4.6b}$$

$$|t^{-\lambda} - S_{\lambda, \delta_0, \eta, T}(t)| \leq \delta t^{-\lambda}, \quad \text{for all } t \in [1, T]. \tag{4.6c}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Proof. For m as in (4.5),

$$\tau \sum_{k>m} e^{\lambda k\tau} \exp\bigl(-e^{k\tau}t\bigr) \le \int_{m\tau}^{\infty} \exp(-te^x + \lambda x)\,\mathrm{d}x = \frac{1}{t^\lambda} \int_{t\exp(m\tau)}^{\infty} y^{\lambda-1}e^{-y}\,\mathrm{d}y,$$

where we have used monotonicity and the substitution $x = \ln(y/t)$. Since $\lambda \le 1$, for $t \ge 1$ we have

$$\int_{t\exp(m\tau)}^{\infty} y^{\lambda-1}e^{-y}\,\mathrm{d}y \le \exp\bigl(-e^{m\tau}\bigr),$$

and thus (4.6a) follows from Theorem 4.2 and the choice of τ and m. With the bound

$$\tau \sum_{k<-n} e^{\lambda k\tau} \exp\bigl(-e^{k\tau}t\bigr) \le \int_{n\tau}^{\infty} e^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}e^{-n\lambda\tau}$$

and using $\lambda\Gamma(\lambda) = \Gamma(\lambda+1) \in (0,1]$, we obtain the two further estimates. □

Note that with τ, m, n as in (4.5), for the cases $\lambda = 1$ and $\lambda = \frac{1}{2}$ of particular interest we obtain the exponential sums

$$S_{1,\delta_0,\eta,T} = \tau \sum_{k=-n}^{m} e^{k\tau} \exp\bigl(-e^{k\tau}t\bigr), \quad S_{1/2,\delta_0,\eta,T} = \frac{\tau}{\sqrt{\pi}} \sum_{k=-n}^{m} e^{(1/2)k\tau} \exp\bigl(-e^{k\tau}t\bigr)$$

with $m + n + 1$ terms. The parameters τ, m and n can be optimized further for particular cases; see Scholz and Yserentant (2017, Sec. 5).

Results analogous to Corollary 4.3 are obtained in Bachmayr and Dahmen (2016*a*,*b*). The construction used there, however, is slightly different from the one in (4.4) and based on general results on sinc quadrature by Stenger (1993). These results can be applied whenever the integrands in (4.3) are holomorphic on $\{z \in \mathbb{C}\colon |\mathrm{Im}\, z| < \eta\}$ for some $\eta > 0$ and integrable on the boundary of this strip. This technique can also be used for obtaining approximations with bounds on $\sup_{t\in[a,b]}|t^{-\lambda} - S_r(t)|$ for a given interval $[a,b]$; see Hackbusch and Khoromskij (2006*a*) and Hackbusch (2015*a*, Sec. D.4). However, stronger results for uniform approximation can be obtained by a different approach, which we consider next.

4.1.2. Uniform best approximation

The application of Proposition 4.1(i) requires uniform approximations on intervals. To this end, rather than aiming for explicit constructions, we can consider best approximations by exponential sums. These are in general not analytically computable, but lead to stronger results. The existence of such best approximations for a class of functions including $f(t) = t^{-\lambda}$ for any $\lambda > 0$ is shown in Braess (1986). The following result on exponential convergence is proved in Braess and Hackbusch (2005, 2009).

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Theorem 4.4. For each $R > 1$, $\lambda > 0$ and $r \in \mathbb{N}$, there exist $\omega_k, \alpha_k > 0$, $k = 1, \dots, r$, such that

$$\sup_{t\in[1,R]} \left| \frac{1}{t^\lambda} - \sum_{k=1}^{r} \omega_k e^{-\alpha_k t} \right| \le 2^\lambda 8 \exp\left(-\frac{\pi^2 r}{\ln(8R)} \right).$$

In contrast to Theorem 4.2, we thus obtain a bound on the absolute rather than the relative approximation error. By choosing suitable R and rescaling, from the above we obtain the following result on intervals $[a, \infty)$ with $a > 0$; see Braess and Hackbusch (2009) and Dahmen, DeVore, Grasedyck and Süli (2016).

Corollary 4.5. Let $a > 0$. Then, for each $r \in \mathbb{N}$, there exist $\omega_k, \alpha_k > 0$, $k = 1, \dots, r$ such that with $\hat{S}_r(t) = \sum_{k=1}^{r} \omega_k e^{-\alpha_k t}$,

$$\sup_{t\in[a,\infty)} \left| \frac{1}{t} - \hat{S}_r(t) \right| \le \frac{16}{a} \exp(-\pi\sqrt{r}).$$

Such best approximations are computed in Hackbusch (2005) by a Remez-type algorithm based on alternation properties of uniform best approximation errors.

4.2. Low-rank approximate inverses of Laplacian-type operators

For the finite difference discretization matrix $\mathbf{A}_{\mathrm{FD}}$ as in (3.3), combining Corollary 4.5 and Proposition 4.1(i), we obtain a bound in the spectral norm and thus

$$\left\| \mathbf{A}_{\mathrm{FD}}^{-1}\mathbf{f} - \hat{S}_r(\mathbf{A}_{\mathrm{FD}})\mathbf{f} \right\|_2 \le \frac{16}{\min \sigma(\mathbf{A}_{\mathrm{FD}})} \exp\left(-\pi\sqrt{r}\right) \|\mathbf{f}\|_2 \tag{4.7}$$

for all right-hand sides $\mathbf{f}$, where

$$\hat{S}_r(\mathbf{A}_{\mathrm{FD}}) = \sum_{k=1}^{r} \omega_k \bigotimes_{i=1}^{d} \exp\left(-\alpha_k h^{-2} \mathbf{\Delta}_2\right).$$

Results of this type, including further matrix functions, are shown in Grasedyck (2004), Gavrilyuk, Hackbusch and Khoromskij (2005), Hackbusch and Khoromskij (2006*b*) and Khoromskij (2009). As proposed in these works, the required matrix exponentials $\exp(-\alpha_k h^{-2}\mathbf{\Delta}_2)$ for each k can be evaluated, for instance, by diagonalization or by hierarchical matrix approximations (Hackbusch 2015*a*).

However, estimates as in (4.7) do not reflect the natural norms of the problem. This can also be seen on the level of function spaces: for $A = -\Delta\colon H_0^1(\square_d) \to H^{-1}(\square_d)$, considering A as an unbounded operator on $\square_d$ of the form (4.1), from Proposition 4.1(i) we also obtain

$$\|A^{-1}f - \hat{S}_r(A)f\|_{L_2(\square_d)} \le \frac{16}{d\pi^2} \exp(-\pi\sqrt{r}) \|f\|_{L_2(\square_d)} \quad \text{for all } f \in L_2(\square_d).$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Here the norm for f is too strong (L_2 rather than H^{-1}), whereas the approximation error is measured in a norm that is too weak (L_2 rather than H^1); in particular, A^{-1} is compact as an operator on $L_2(\square_d)$.

The approximation error achieved by $\hat{S}_r$ in appropriate norms is investigated in Dahmen *et al.* (2016). Specializing a result from this work that holds for more general scales of norms (Dahmen *et al.* 2016, Prop. 1), for the present choice of A and approximation in H^1-norm, we obtain the following.

Theorem 4.6. For $s \in (0,1)$, for all $f \in H^{-1+s}(\square_d)$,

$$\|A^{-1}f - \hat{S}_r(A)f\|_{H^{-1+s}(\square_d)\to H^1_0(\square_d)} \le 8\exp\left(-\frac{s}{2}\pi\sqrt{r}\right)\|f\|_{H^{-1+s}(\square_d)}.$$

The convergence rate is thus strongly dependent on additional regularity of corresponding right-hand sides f higher than H^{-1}, and thus on compactness of A^{-1}. In Dahmen *et al.* (2016), the approximations $\hat{S}_r(A)$ are combined with quadrature approximations of contour integral representations of required operator exponentials to obtain bounds on the total computational complexity of approximating $A^{-1}f$ in $H^1_0(\square_d)$; provided that f has suitable low-rank structure, these bounds are in particular polynomial with respect to d.

4.3. Low-rank preconditioners for finite difference and finite element discretizations

Let us again consider $\mathbf{A}_{\mathrm{FD}} \in \mathbb{R}^{N\times N}$ as in (3.3), rescaled without loss of generality such that $\sigma(\mathbf{A}_{\mathrm{FD}}) \subset [1, \mathrm{cond}(\mathbf{A}_{\mathrm{FD}})]$; recall that $\mathrm{cond}(\mathbf{A}_{\mathrm{FD}}) = O(h^{-2})$ with $h = 1/(N+1)$. We then obtain the following from Proposition 4.1(ii) and Corollary 4.3.

Proposition 4.7. With $S_{\lambda,\delta_0,\eta,T}$ as in Corollary 4.3, with $\delta \in (0,1)$ and $T = \mathrm{cond}(\mathbf{A}_{\mathrm{FD}})$, for $\lambda \in (0,1]$ let

$$\mathbf{C}_\lambda = S_{\lambda,\delta/2,\delta/2,T}(\mathbf{A}_{\mathrm{FD}}).$$

Then, for $\lambda = 1$ and $\lambda = 1/2$, we have

$$\mathrm{cond}(\mathbf{C}_1\mathbf{A}_{\mathrm{FD}}) \le \frac{1+\delta}{1-\delta}, \quad \mathrm{cond}(\mathbf{C}_{1/2}\mathbf{A}_{\mathrm{FD}}\mathbf{C}_{1/2}) \le \left(\frac{1+\delta}{1-\delta}\right)^2,$$

where $\mathbf{C}_1$ and $\mathbf{C}_{1/2}$ are each sums of $O(|\log\delta| + |\log h|)$ elementary tensor products.

Exponential sum preconditioners have also been constructed in Khoromskij (2009) based on uniform approximation bounds as in Section 4.1.2; however, this leads to ranks scaling with respect to h as $O(\log^2 h)$ instead of $O(|\log h|)$.

Remark 4.8. Compared to the commonly used left preconditioning $\mathbf{C}_1\mathbf{A}_{\mathrm{FD}}$, the two-sided preconditioning $\mathbf{C}_{1/2}\mathbf{A}_{\mathrm{FD}}\mathbf{C}_{1/2}$ based on the exponential sum for $t \mapsto t^{-1/2}$ can be of particular interest in the context of low-rank methods, for the following reason: starting from the original system $\mathbf{A}_{\mathrm{FD}}\mathbf{u}_h = \mathbf{f}_h$, when using two-sided

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

preconditioning we solve the equivalent system $\mathbf{C}_{1/2}\mathbf{A}_{\mathrm{FD}}\mathbf{C}_{1/2}\tilde{\mathbf{u}}_h = \mathbf{C}_{1/2}\mathbf{f}_h$ with $\tilde{\mathbf{u}}_h = \mathbf{C}_{1/2}^{-1}\mathbf{u}_h$. This means that $\|\tilde{\mathbf{u}}_h\|_2^2 \eqsim \langle \mathbf{A}_{\mathrm{FD}}\mathbf{u}_h, \mathbf{u}_h\rangle$; similarly to Riesz basis representations as in Remark 3.11, when working with two-sided preconditioning, the Euclidean norm of solution coefficients (which is controlled in HSVD rank truncations) is thus equivalent to the discrete energy norm induced by $\mathbf{A}_{\mathrm{FD}}$.

Remark 4.9. As an alternative preconditioner for finite difference and finite element discretizations, standard multigrid methods performed in tensor representation are proposed in Ballani and Grasedyck (2013) and Hackbusch (2015*b*). When such methods are used on a grid hierarchy on product domains such as $\square_d$, prolongation and restriction operations are elementary tensor products of corresponding one-dimensional grid hierarchies. Addition of fine- and coarse-grid corrections as well as iterative smoothers performed in tensor representations generally increase the representation ranks of iterates, and these operations thus need to be combined with rank truncations. The convergence analysis of this approach with truncation errors is an open problem.

4.4. Low-rank diagonal preconditioners

Recall the wavelet Riesz basis representation $\mathbf{A}$ of the Laplacian defined in (3.21) with the decomposition $\mathbf{A} = \mathbf{DTD}$. Here $\mathbf{T}$ is the representation of the Laplacian with respect to the orthonormal basis

$$\left\{ \bigotimes_{i=1}^{d} \psi_{\nu_i} \right\}_{\nu \in \vee^d}$$

of $L_2(\square_d)$, and

$$\mathbf{D}[\nu, \nu'] = \left(\sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)} \right)^{-1/2}, \quad \nu, \nu' \in \vee^d. \tag{4.8}$$

In this setting, $\mathbf{D}$ thus acts as a diagonal preconditioner on the sequence space $\ell_2(\vee^d)$, ensuring that $\mathbf{A}\colon \ell_2(\vee^d) \to \ell_2(\vee^d)$ is an isomorphism. The issue we are facing here is that (4.8) does not have low-rank form. We thus consider how $\mathbf{D}$ can be replaced by low-rank substitutes.

Remark 4.10. For any sequence $(d_\nu)_{\nu \in \vee^d}$ of positive numbers with

$$d_\nu \eqsim \left(\sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)} \right)^{-1/2}, \quad \nu \in \vee^d,$$

$\{d_\nu \bigotimes_{i=1}^{d} \psi_{\nu_i}\}_{\nu \in \vee^d}$ is a Riesz basis of $H^1_0(\square_d)$.

We now choose such d_ν to obtain a substitute with efficient low-rank approximations for $\mathbf{D}$. For any fixed $\delta_0 \in (0,1)$, with $S_{1/2,\delta_0}$ as in Corollary 4.3 and

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

defining

$$M_0 = \min_{\nu\in\vee^d} \sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)},$$

let

$$d_\nu = M_0^{-1/2}\, S_{1/2,\delta_0}\left(M_0^{-1} \sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right).$$

By Corollary 4.3,

$$\left|\left(\sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)^{-1/2} - d_\nu\right| \le \delta_0 \left(\sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)^{-1/2}, \quad \nu \in \vee^d.$$

Thus $\{\tilde\Psi_\nu\}_{\nu\in\vee^d}$ with $\tilde\Psi_\nu = d_\nu \bigotimes_{i=1}^{d} \psi_{\nu_i}$ is indeed a Riesz basis of $H^1_0(\square_d)$.

Representing A in terms of this basis, we obtain the new infinite matrix representation

$$\tilde{\mathbf{A}} = \left(\langle A\tilde\Psi_{\nu'}, \tilde\Psi_\nu\rangle\right)_{\nu,\nu'\in\vee^d} = \tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}},$$

where

$$\tilde{\mathbf{D}}[\nu,\nu'] = d_\nu \delta_{\nu,\nu'} \quad \text{for all } \nu,\nu' \in \vee^d.$$

Note that with τ and m as in (4.5), with

$$\omega_k = \tau\sqrt{\frac{e^{k\tau}}{M_0\pi}}, \quad \alpha_k = \frac{e^{k\tau}}{M_0},$$

we have

$$\tilde{\mathbf{D}}[\nu,\nu] = d_\nu = \sum_{k=-\infty}^{m} \omega_k \prod_{i=1}^{d} e^{-\alpha_k \|\psi'_{\nu_i}\|^2_{L_2(0,1)}}, \tag{4.9}$$

in other words $\tilde{\mathbf{D}}$ has an explicit expansion in terms of tensor product diagonal operators with infinitely many terms.

The strategy pursued in Bachmayr and Dahmen (2016*b*) is now to formulate the problem in terms of the basis $\{\tilde\Psi_\nu\}_{\nu\in\vee^d}$ with operator representation $\tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}$, with the representation (3.23) of the operator equation $Au = f \in H^{-1}(\square_d)$ modified to

$$\tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}\mathbf{u} = \tilde{\mathbf{D}}\mathbf{g}, \tag{4.10}$$

and introduce additional low-rank approximations of $\tilde{\mathbf{D}}$. This can be done for any finite subset $\Lambda \subset \vee^d$, for which we define

$$T(\Lambda) = M_0^{-1} \max_{\nu\in\Lambda} \sum_{i=1}^{d} \|\psi'_{\nu_i}\|^2_{L_2(0,1)}.$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

For $\eta > 0$ such that $\delta_0 + \eta < 1$, we define the diagonal operator

$$\tilde{\mathbf{D}}_{\eta,\Lambda}[\nu,\nu] = M_0^{-1/2}\, S_{1/2,\delta_0,\eta,T(\Lambda)}\left(M_0^{-1}\sum_{i=1}^{d}\|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)$$
$$= \sum_{k=-n}^{m}\omega_k\prod_{i=1}^{d}e^{-\alpha_k\|\psi'_{\nu_i}\|^2_{L_2(0,1)}}, \tag{4.11}$$

with $S_{1/2,\delta_0,\eta,T(\Lambda)}$ and n as in Corollary 4.3, where n from (4.5) satisfies

$$n = \left\lceil \tau^{-1}(2|\ln\eta| + \ln T(\Lambda))\right\rceil$$

and the ω_k, α_k are as in (4.9). The modified rescaling $\tilde{\mathbf{D}}$ and its low-rank approximations $\tilde{\mathbf{D}}_{\eta,\Lambda}$ have the following properties by Corollary 4.3 and Proposition 4.1.

Proposition 4.11. Let $\delta_0 \in (0,1)$; then, for $\mathbf{D}$ as in (4.8) and $\tilde{\mathbf{D}}$ defined in (4.9), we have $\|\mathbf{D}^{-1}\tilde{\mathbf{D}}\|_{\ell_2\to\ell_2} \le 1+\delta_0$, $\|\tilde{\mathbf{D}}^{-1}\mathbf{D}\|_{\ell_2\to\ell_2} \le (1-\delta_0)^{-1}$, and

$$\langle\tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}\mathbf{v},\mathbf{v}\rangle \eqsim \|\mathbf{v}\|^2_{\ell_2} \quad \text{for all } \mathbf{v}\in\ell_2(\vee^d).$$

Moreover, for $\eta > 0$ such that $\delta = \delta_0+\eta < 1$ and finite $\Lambda \subset \vee^d$, let $\tilde{\mathbf{D}}_{\eta,\Lambda}$ be defined as in (4.11). Then

$$\langle\tilde{\mathbf{D}}_{\eta,\Lambda}\mathbf{T}\tilde{\mathbf{D}}_{\eta,\Lambda}\mathbf{v},\mathbf{v}\rangle \eqsim \|\mathbf{v}\|^2_{\ell_2} \quad \text{for all } \mathbf{v}\in\ell_2(\vee^d) \text{ with } \operatorname{supp}\mathbf{v}\subseteq\Lambda$$

with constants independent of Λ. In addition,

$$\sup_{\substack{\|\mathbf{v}\|_{\ell_2}=1\\ \operatorname{supp}\mathbf{v}\subseteq\Lambda}}\|\mathbf{D}^{-1}(\tilde{\mathbf{D}}-\tilde{\mathbf{D}}_{\eta,\Lambda})\mathbf{v}\|_{\ell_2} \le \eta, \qquad \sup_{\substack{\|\mathbf{v}\|_{\ell_2}=1\\ \operatorname{supp}\mathbf{v}\subseteq\Lambda}}\|\mathbf{D}^{-1}(\mathbf{D}-\tilde{\mathbf{D}}_{\eta,\Lambda})\mathbf{v}\|_{\ell_2} \le \delta.$$

For the number of summands in $\tilde{\mathbf{D}}_{\eta,\Lambda}$ in terms of η and Λ, we have

$$m+n+1 \lesssim 1+|\log\eta| + \log\left(\max_{\nu\in\Lambda}\max_{i=1,\dots,d}|\nu_i|\right). \tag{4.12}$$

In summary, we obtain low-rank approximations of the wavelet Riesz basis representation of A. In this particular case, the evaluation of exponential sums requires only trivial matrix exponentials of diagonal matrices that can be evaluated entrywise as in (4.11).

Remark 4.12. Note that by (4.12) we have a dependence of the representation rank of $\tilde{\mathbf{D}}_{\eta,\Lambda}$ on the maximum wavelet level $\max_i|\nu_i|$ in the index set $\Lambda\subset\vee^d$. This dependence becomes more transparent when using, based on the observation (3.24), the alternative rescaling by $d_\nu = 2^{-\max_i|\nu_i|}$, $\nu\in\vee^d$, which satisfies

$$2^{-\max_i|\nu_i|} \le \left(\sum_{i=1}^{d}\|\psi'_{\nu_i}\|^2_{L_2(0,1)}\right)^{-1/2} \le \sqrt{d}\,2^{-\max_i|\nu_i|}.$$

Note that the spectral bounds for the resulting operator representations may then

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

deteriorate by a factor $\sqrt{d}$, but for any finite $\Lambda \subset \vee^d$, we obtain a simple explicit low-rank representation as follows. For $S \subset \vee$, let $\mathbf{1}_S \in \ell_2(\vee)$ denote the sequence with entries equal to one on S, and zero otherwise. Moreover, let $J(\Lambda) = \max_{\nu\in\Lambda} \max_i |\nu_i|$. Then

$$\left(2^{-\max_i |\nu_i|}\right)_{\nu\in\vee^d} = \frac{1}{2} \sum_{l=1}^{J(\Lambda)-1} 2^{-l} \bigotimes_{i=1}^{d} \mathbf{1}_{\{\nu\in\vee:\ |\nu|\le l\}} + 2^{-J(\Lambda)} \bigotimes_{i=1}^{d} \mathbf{1}_{\{\nu\in\vee:\ |\nu|\le J(\Lambda)\}},$$

with $\max_{\nu\in\Lambda} \max_i |\nu_i|$ separable terms.

A further approach based on this alternative rescaling sequence is to construct a partition of $\vee^d = \bigcup_{k=1}^{\infty} \Lambda_k$ such that for each Λ_k, the vector $(2^{-\max_i |\nu_i|})_{\nu\in\Lambda_k}$ is an elementary tensor product. When using a *separate* tensor representation for each Λ_k, applying these rescalings then leaves the respective ranks unchanged. Partitions with this property of finite subsets $\{\nu \in \vee^d :\ \max_i |\nu_i| \le J\} \subset \vee^d$ for any $J \in \mathbb{N}$ into $O(J^{d-2}2^J)$ elements are constructed in Bachmayr (2012*a*).

4.5. Preconditioning tensorized discretizations

Let us now return to the tensorized multilevel representations considered in Section 3.4. As noted in Example 3.17, for high discretization levels L corresponding (in d-dimensional problems) to grids of size $O(2^{dL})$, naive tensorized representations of differential operators have large representation condition numbers according to Definition 2.45.

In certain special cases, this issue can be circumvented by explicit inverses, constructed for instance for the matrix $\mathbf{\Delta}_2$ considered in Example 3.17 in Kazeev and Khoromskij (2012), but for more general problems these are not available and suitable preconditioning is required.

The unusual aspect here is that such a preconditioner needs to remedy not only the large standard condition number of the represented matrix but also, as noted in Example 3.17, the equally large representation condition number of the underlying low-rank representation. Such a preconditioner is proposed and analysed in Bachmayr and Kazeev (2020), with finite element discretizations of diffusion problems on $\square_d$ (where in this context we aim at $d = 2$ or $d = 3$) with diffusion coefficient $\hat{M}$ as model problems. With a dyadically refined hierarchy of tensor product finite element discretizations with nested spaces V_ℓ on levels $\ell = 0, \dots, L$, where $\{\varphi_i\}_{i\in\mathcal{N}_L}$ with $\#\mathcal{N}_L = O(2^{dL})$ is the basis of hat functions on level L, the considered discretization matrices read

$$\mathbf{A} = \left(\int_{\square_d} \hat{M} \nabla\varphi_i \cdot \nabla\varphi_j \, \mathrm{d}x \right)_{i,j\in\mathcal{N}_L}$$

with a uniformly positive definite $\hat{M} \in \mathbb{R}^{d\times d}$. These discretization matrices are represented in multilevel tensorized format in transposed ordering, according to Remark 3.18, in the form

$$\mathbf{A} = \tau(\mathbf{A}_0, \dots, \mathbf{A}_L) = \mathbf{A}_0 \bowtie \cdots \bowtie \mathbf{A}_L,$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

where each $\mathbf{A}_\ell$ for $\ell = 0, \dots, L$ acts on modes of size 2^d with indices in $\{0,1\}^d$.

We have a decomposition $\mathbf{A} = \mathbf{G}^\top \mathbf{M} \mathbf{G}$, where $\mathbf{M} = \tau(\mathsf{M})$ is a matrix depending on the diffusion coefficient $\hat{M}$ with well-conditioned representation M, and $\mathbf{G} = \tau(\mathsf{G})$ has an ill-conditioned representation similarly to Example 3.17, where $\max_\ell \operatorname{oprcond}_\ell(\mathsf{G}) \eqsim 2^L$. Here $\mathbf{G}$ can be interpreted as a discrete gradient that maps $\operatorname{span}\{\varphi_i\}_{i \in \mathcal{N}_L}$ to a space of discontinuous piecewise polynomials.

As a first step towards a suitable preconditioner, we have the following result (Bachmayr and Kazeev 2020, Thm 2, Thm 3).

Theorem 4.13. Let $\mathbf{C} = \sum_{\ell=0}^{L} 2^{-\ell} \mathbf{P}_{\ell,L} \mathbf{P}_{\ell,L}^\top$, where $\mathbf{P}_{\ell,L}$ is the matrix representation of the canonical prolongation operator from V_ℓ to V_L. Then $\langle \mathbf{CACv}, \mathbf{v} \rangle \eqsim \|\mathbf{v}\|_2^2$ for all $\mathbf{v}$, with constants that depend only on M and d. Moreover, there exists C with $\mathbf{C} = \tau(\mathsf{C})$ with ranks bounded independently of L.

Note that the matrix $\mathbf{C}$ provides a two-sided version of the classical BPX preconditioner (Bramble, Pasciak and Xu 1990). Although $\mathbf{CAC}$ is well-conditioned as a matrix uniformly in L, performing the multiplication by $\mathbf{CAC}$ by separately applying $\mathbf{C}$ and $\mathbf{A}$ does not remove the representation ill-conditioning. To arrive at a well-conditioned representation, however, we can combine G and C into a new rank-reduced representation B (again with ranks bounded independently of L) such that $\tau(\mathsf{B}) = \mathbf{GC}$. We arrive at

$$\mathbf{CAC} = \tau(\mathsf{B})^\top \tau(\mathsf{M}) \tau(\mathsf{B}),$$

which provides a well-conditioned representation of $\mathbf{CAC}$ for all L. For the technical details of the construction, we refer to Bachmayr and Kazeev (2020). Based on this preconditioner, numerically stable computations with $L \approx 50$, corresponding to represented grid sizes near machine precision, are demonstrated to be feasible.

5. Solving operator equations in low-rank format

In this section we turn to the central task of computing low-rank approximations of solutions of elliptic operator equations $Au = f$, where $A : V \to V'$, $f \in V'$ with a separable Hilbert space V. Here our aim is, for each given error tolerance $\varepsilon > 0$, to find an approximation u_ε to the exact solution $u \in V$ such that $\|u - u_\varepsilon\|_V \le \varepsilon$. A guiding model case will again be the Poisson problem as in Example 3.1, where $V = H_0^1(\square_d)$.

In the context of low-rank approximations, guaranteeing convergence to the exact solution with respect to the V-norm involves two central issues: on the one hand, we need to find highly nonlinear representations of approximate solutions, where the required rank parameters are not known *a priori*. This is addressed by iterative solvers operating on suitable basis representations of the PDE problem. On the other hand, we need to ensure sufficiently small discretization errors, which requires *a posteriori* error estimates computable in high dimensions as well as adaptive discretization refinement that is compatible with the low-rank format.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Convergence guarantees alone, however, are not very informative: to give an extreme example, one could use a standard solver to obtain a sufficiently accurate approximate solution in full format, and only then compute a compressed low-rank approximation. With such an approach, one can easily ensure convergence, but without any benefits concerning computational complexity of using low-rank approximations. An essential question is therefore how the computational costs of a low-rank solver compare to the minimal required costs, which depend on the intrinsic low-rank approximability of solutions and problem data. In particular, can the curse of dimensionality be avoided in the total computational complexity?

In Section 5.1 we first survey different approaches for computing low-rank approximations of solutions of linear systems of equations. In Section 5.2 we consider iterative methods for finding low-rank approximate solutions with two main features: first, they can always be guaranteed to converge, and second, for certain typical model classes of low-rank approximability, the approximations produced by the iterations have near-optimal ranks in relation to the currently achieved accuracy. Two approaches for achieving this are discussed, based on HSVD truncation (see Section 2.7.1) and on soft thresholding of hierarchical tensors (see Section 2.7.2), respectively. At this point, these methods are formulated as operating on fixed discretizations or on function spaces.

In Section 5.3 we turn to the issue of combining rank adaptation with adaptive refinement of discretizations. Again we obtain methods that can be guaranteed to converge, and that, for solutions in certain model classes of approximability, exhibit near-optimal performance in terms of ranks and discretization sizes. Note that the resulting algorithms are *universal*: the respective degree of approximability of the exact solution does not need to be known but is automatically detected by the adaptive scheme. Altogether, this leads to bounds on the total computational complexity of approximately solving elliptic PDEs in high dimensions by such methods, which we discuss in Section 5.4. In Section 5.5 we survey what is known on the low-rank approximability of particular classes of elliptic problems.

5.1. Iterative solvers in low-rank format

Iterative solvers operating on low-rank representations are often introduced in the literature for discretized problems on finite-dimensional spaces, such as the finite difference discretization $\mathbf{A}_{\mathrm{FD}}\mathbf{u}_h = \mathbf{f}_h$ discussed in Section 3.1.1. However, the methods considered below can also be formulated to operate on tensors in sequence spaces $\ell_2(\mathbb{N}^d)$, which with an appropriate choice of bases is equivalent to performing iterations on *function spaces*. This is the case with the Riesz basis representations $\mathbf{Au} = \mathbf{f}$, where $\mathbf{A}$ is an isomorphism on $\ell_2(\mathbb{N}^d)$, as obtained in Section 3.2. Based on such formulations, we obtain iterative methods that are robust with respect to the choice of discretization. Note that by ellipticity of A we also have ellipticity of $\mathbf{A}$ on $\ell_2(\mathbb{N}^d)$, that is,

$$\langle \mathbf{Av}, \mathbf{v} \rangle \gtrsim \|\mathbf{v}\|_{\ell_2}^2 \quad \text{for all } \mathbf{v} \in \ell_2(\mathbb{N}^d).$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

5.1.1. Krylov subspace methods with rank truncation

In Krylov subspace methods, approximate solutions are generated by repeated application of the linear mapping $\mathbf{A}_{\mathrm{FD}}$ or $\mathbf{A}$ under consideration. One of the simplest instances of such methods, outlined in Section 1.7, is Richardson iteration. Applied to a fixed discretization $\mathbf{A}_{\mathrm{FD}}$, with a suitable left preconditioner $\mathbf{C}$, in its basic form the method reads

$$\mathbf{u}_h^{n+1} = \mathbf{u}_h^n - \omega \mathbf{C}\big(\mathbf{A}_{\mathrm{FD}}\mathbf{u}_h^n - \mathbf{f}_h\big). \tag{5.1}$$

As described in Section 1.7, when applying the method in this form directly in low-rank format, the representation ranks of $\mathbf{u}_h^n$ will in general increase exponentially with respect to n. It is thus natural to combine the basic iteration (5.1) with a suitable rank reduction operation $\mathcal{R}_n$,

$$\mathbf{u}_h^{n+1} = \mathcal{R}_n\big(\mathbf{u}_h^n - \omega \mathbf{C}\big(\mathbf{A}_{\mathrm{FD}}\mathbf{u}_h^n - \mathbf{f}_h\big)\big). \tag{5.2}$$

One option is to choose $\mathcal{R}_n(\mathbf{v})$ as the truncation of the HSVD of the given $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ to *fixed hierarchical rank*. With different tensor formats, such strategies are used in Khoromskij and Schwab (2011) and Billaud-Friess, Nouy and Zahm (2014). However, the resulting truncated iterative scheme can generally be guaranteed to converge only under extremely restrictive conditions on the error reduction achieved in each step of (5.1); see Hackbusch, Khoromskij and Tyrtyshnikov (2008), Billaud-Friess *et al.* (2014) and Bachmayr *et al.* (2016, Prop. 6.1).

A generally more practical alternative is to choose $\mathcal{R}_n$ as the truncation of the HSVD to lower ranks up to a *prescribed error tolerance*, potentially dependent on n, based on the computable error bounds in terms of matricization singular values of Theorem 2.36. In combination with Richardson iteration, such a strategy is used in Matthies and Zander (2012). Error-controlled truncations can also be combined with more advanced Krylov space methods, such as the preconditioned conjugate gradient (PCG) method with rank truncation proposed in Kressner and Tobler (2011*a*) and Tobler (2012) for positive definite problems, or the generalized minimal residual (GMRES) method with truncation proposed in Ballani and Grasedyck (2013). The variants are typically observed to require fewer iterations than the truncated Richardson iteration, although the orthogonality properties that are crucial for the standard convergence analysis of PCG and GMRES are not preserved by the truncations. However, the analysis by Ali and Urban (2020) of rank-truncated PCG with appropriate truncation tolerances as a perturbed descent method shows that worst-case performance comparable to (5.2) can still be ensured.

In iterations with truncations up to specified tolerances, however, the crucial question is how the ranks of the iterates evolve, that is, whether the truncation tolerances are sufficiently large to ensure the appropriate rank reduction. Ideally, the ranks of iterates should behave similarly to the ranks of best approximations of comparable accuracy. How this can be achieved in a general setting is treated in Section 5.2.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

To enable a useful comparison to low-rank best approximations, rank truncations need to be performed with errors with respect to the appropriate norm, in our present setting in V-norm. As discussed in Remark 4.8, this can be achieved by two-sided preconditioning and is inherent in Richardson iteration performed on a Riesz basis representation $\mathbf{Au} = \mathbf{f}$ with $\mathbf{A} = \mathbf{DTD}$, $\mathbf{f} = \mathbf{Dg}$. For such a Riesz basis representation, $\|\mathbf{u} - \mathbf{u}^n\|_{\ell_2} \eqsim \|\mathbf{f} - \mathbf{Au}^n\|_{\ell_2}$, and both quantities are proportional to the error of approximating the exact solution u in V-norm. In this case (5.2) takes the form

$$\mathbf{u}^{n+1} = \mathcal{R}_n(\mathbf{u}^n - \omega \mathbf{r}^n),$$

where $\mathbf{r}^n$ is a sufficiently accurate low-rank approximation of the residual $\mathbf{Au}^n - \mathbf{f} = \mathbf{D}(\mathbf{TDu}^n - \mathbf{g})$, with the infinite diagonal scaling matrix $\mathbf{D}$ acting as a two-sided preconditioner. In this form, Richardson iteration serves as the basis of low-rank methods with adaptively refined discretizations, considered in detail in Section 5.3.2.

5.1.2. Alternating optimization

Solution strategies that sequentially optimize individual components of tensor representations are based on reformulations of the original systems of equations as minimization problems. In particular, when $\mathbf{A}$ is symmetric, the solution $\mathbf{u}$ of $\mathbf{Au} = \mathbf{f}$ can be characterized as the unique minimizer of the functional J defined by

$$J(\mathbf{v}) = \frac{1}{2}\langle \mathbf{Av}, \mathbf{v}\rangle - \langle \mathbf{f}, \mathbf{v}\rangle. \tag{5.3}$$

Methods of this kind have been considered for the canonical tensor format (1.9), where the issues discussed in Section 1.4 apply; see for instance Uschmajew (2012). In the context of subspace tensor formats, although alternating optimization approaches are applicable to general hierarchical tensors, they have been investigated primarily for the tensor train format, for which they are easiest to formulate.

A first classical method of this type is known, depending on context, as *alternating least squares* (ALS), the *alternating linear scheme* (Holtz, Rohwedder and Schneider 2012*a*) or as *one-site density matrix renormalization group* (White 2005). Starting from a tensor train representation

$$\mathbf{u}^n = \tau\big(\mathbf{X}_1^n, \dots, \mathbf{X}_d^n\big) = \mathbf{X}_1^n \bowtie \cdots \bowtie \mathbf{X}_d^n, \tag{5.4}$$

the next iterate $\mathbf{u}^{n+1} = \mathbf{X}_1^{n+1} \bowtie \cdots \bowtie \mathbf{X}_d^{n+1}$ is obtained by sequentially solving the substeps

$$\begin{aligned}
\mathbf{X}_1^{n+1} &= \arg\min_{\mathbf{Y}} J\big(\tau\big(\mathbf{Y}, \mathbf{X}_2^n, \mathbf{X}_3^n, \dots, \mathbf{X}_d^n\big)\big),\\
\mathbf{X}_2^{n+1} &= \arg\min_{\mathbf{Y}} J\big(\tau\big(\mathbf{X}_1^{n+1}, \mathbf{Y}, \mathbf{X}_3^n, \dots, \mathbf{X}_d^n\big)\big),\\
&\vdots\\
\mathbf{X}_d^{n+1} &= \arg\min_{\mathbf{Y}} J\big(\tau\big(\mathbf{X}_1^{n+1}, \mathbf{X}_2^{n+1}, \mathbf{X}_3^{n+1}, \dots, \mathbf{Y}\big)\big).
\end{aligned} \tag{5.5}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Note that since τ is multilinear and J is quadratic, the solution of the minimization problem in each substep amounts to a linear system of equations for $\mathbf{Y}$. In this method, only one component is modified at a time and all rank parameters are determined by the remaining components. As a consequence, in this basic form, there is no mechanism for the adaptation of approximation ranks, and the ranks of $\mathbf{u}^n$ remain bounded by those of the initial iterate $\mathbf{u}^0$ for all n. Effectively, we are thus minimizing J under a rank constraint. The iteration may in general fail to converge to the constrained solution. Convergence can be shown under additional conditions that, however, cannot generally be verified; see Rohwedder and Uschmajew (2013), Wang and Chu (2014) and Oseledets, Rakhuba and Uschmajew (2018).

In the (two-site) *density matrix renormalization group* (DMRG) algorithm introduced by White (1992) in the context of eigenvalue problems, and formulated by Vidal (2003) in terms of matrix product states, the iteration (5.5) is modified by optimizing two components at a time as follows: in the first substep, starting again from $\mathbf{u}^n$ as in (5.4), we combine $\mathbf{X}_1^n \bowtie \mathbf{X}_2^n$ into a single component carrying two modes of the represented tensor, and minimize over this fourth-order component to obtain

$$\tilde{\mathbf{X}}_{1,2}^{n+1} = \arg\min_{\tilde{\mathbf{Y}}} J\big(\tilde{\mathbf{Y}} \bowtie \mathbf{X}_3^n \bowtie \mathbf{X}_4^n \bowtie \cdots \bowtie \mathbf{X}_d^n\big).$$

We then perform a low-rank approximation, for instance by truncated SVD, to split the intermediate result again into two components:

$$\tilde{\mathbf{X}}_{1,2}^{n+1} \approx \mathbf{X}_1^{n+1} \bowtie \tilde{\mathbf{X}}_2^{n+1}.$$

In the next substep, we combine $\tilde{\mathbf{X}}_2^{n+1} \bowtie \mathbf{X}_3^n$, and so forth. As a solver for linear systems formulated as a minimization problem (5.3), this approach is considered in Holtz *et al.* (2012*a*) and Oseledets and Dolgov (2012); see also Oseledets (2011*a*). Although this approach offers a mechanism for the adaptation of approximation ranks, it may fail to converge in certain cases. An explicit example is given in Pfeffer (2018, Ex. 5.12).

A further variant that can achieve guaranteed convergence is the *alternating minimal energy* (AMEn) method of Dolgov and Savostyanov (2014). The basic iteration is performed as in (5.5), but in outer iteration n, before performing the ith substep for $i = 2, \dots, d$, the previously updated component $\mathbf{X}_{i-1}^{n+1}$ is augmented by a certain number of vectors from a projected residual approximation. The available analysis of this methods guarantees convergence, but with an error reduction that may in the worst case deteriorate exponentially with respect to the tensor order d. With additional practical modifications proposed in Dolgov and Savostyanov (2014), including rank truncation of iterates, the method is demonstrated to perform far better in practice than indicated by the analysis, in particular on test problems with coarse discretizations and low ranks. Although with fine discretizations its performance can degrade for ill-conditioned problems requiring large ranks

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

(Bachmayr and Dahmen 2016*a*, Sec. 4.3), it shows very good practical efficiency when combined with optimal preconditioning (Bachmayr and Kazeev 2020, Sec. 7).

5.1.3. Riemannian optimization

Solving linear systems by minimization of functionals J as in (5.3) can also be approached differently, using the manifold structure of tensors of fixed hierarchical ranks discussed in Section 2.10. Specifically, for a given dimension tree $\mathbb{T}$, for each $\mathsf{r} \in \mathrm{Ranks}(\mathbb{E})$, the set $\{\mathbf{v} \in \ell_2(\mathcal{I}_1 \times \cdots \times \mathcal{I}_d) \colon \operatorname{rank}_{\mathbb{E}}(\mathbf{v}) = \mathsf{r}\}$ can be regarded as an embedded submanifold $\mathcal{M}_\mathsf{r}$ with Riemannian metric inherited from the ambient space $\ell_2(\mathcal{I}_1 \times \cdots \times \mathcal{I}_d)$. Note that while results using this manifold structure are in most cases formulated for finite $\mathcal{I}_1, \dots, \mathcal{I}_d$, the approach can be generalized to the case of general separable Hilbert spaces corresponding to countable index sets.

For given $\mathbf{v} \in \mathcal{M}_\mathsf{r}$ we let $\mathcal{T}_\mathbf{v}\mathcal{M}_\mathsf{r}$ denote the tangent space of $\mathcal{M}_\mathsf{r}$ at $\mathbf{v}$, and let $\Pi_\mathbf{v}$ denote the orthogonal projection onto this tangent space. Basic versions of Riemannian optimization methods can be obtained using, for a given iterate $\mathbf{u}^n$, the negative *Riemannian gradient*

$$\mathbf{p}^n = -\Pi_{\mathbf{u}^n} \nabla J(\mathbf{u}^n)$$

as a search direction. To obtain a new iterate on the manifold, we then use a *retraction*, that is, a mapping $R_{\mathcal{M}_\mathsf{r}} \colon \mathcal{M}_\mathsf{r} \times \mathcal{T}_\mathbf{v}\mathcal{M}_\mathsf{r} \to \mathcal{M}_\mathsf{r}$ with the property

$$\|\mathbf{v} + \mathbf{p} - R_{\mathcal{M}_\mathsf{r}}(\mathbf{v}, \mathbf{p})\| = o(\|\mathbf{p}\|) \quad \text{for } \mathbf{v} \in \mathcal{M}_\mathsf{r},\ \mathbf{p} \in \mathcal{T}_\mathbf{v}\mathcal{M}_\mathsf{r}.$$

With such a mapping, we set $\mathbf{u}^{n+1} = R_{\mathcal{M}_\mathsf{r}}(\mathbf{u}^n, \omega_n \mathbf{p}^n)$ with a suitable damping parameter $\omega_n > 0$. As shown in Steinlechner (2016), the HSVD truncation to rank r is a retraction. For a comparative overview of several different types of retractions in the case of low-rank matrices, see Absil and Oseledets (2015). In addition to gradient information, the *Riemannian Hessian* can be used for the construction of Newton-type methods.

For an overview of such methods in the case of low-rank matrices, see Absil *et al.* (2008). In their basic form operating on fixed ranks, they are applied to tensor completion problems in Kressner, Steinlechner and Vandereycken (2014*b*) and Da Silva and Herrmann (2015), where tensors are to be recovered from a limited number of entries under rank constraints. Riemannian optimization methods for solving linear systems, in particular finite difference discretizations, are studied in Kressner, Steinlechner and Vandereycken (2016). In this case rank adaptation is performed by applying the solver repeatedly with successively increased ranks.

While the representation ranks that can arise during the iteration are fixed, this comes at the price that convergence to the minimizer of J with ranks constrained to r is not generally guaranteed. For a detailed discussion and numerical comparisons demonstrating accelerated convergence of Riemannian methods in certain cases, see Uschmajew and Vandereycken (2020). In the case of finite-dimensional matrices, systematic approaches to rank adaptation are considered in Uschmajew

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

and Vandereycken (2015), Schneider and Uschmajew (2015), Zhou *et al.* (2016) and Gao and Absil (2022).

5.1.4. Greedy methods

A further construction principle of optimization-based methods is given by greedy strategies of the basic form

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \mathbf{w}^n, \quad \mathbf{w}^n = \min_{\mathbf{w}\in\mathcal{S}} J(\mathbf{u}^n + \mathbf{w}),$$

where J can be quadratic as in (5.3) or a more general functional, and $\mathcal{S}$ is a restricted set of tensors, for instance the set of elementary tensor products. Methods of this type are also known as *proper generalized decomposition*; see Ammar, Chinesta and Falcó (2010) and Falcó and Nouy (2011, 2012). Closely related approaches are analysed in Cancès, Ehrlacher and Lelièvre (2011, 2013, 2014). Convergence of such methods can be obtained from variational arguments for fairly general classes of functionals on Banach spaces. Unless combined with an additional rank reduction strategy, the ranks of approximations produced by this approach may in general be substantially higher than best approximation ranks for similar accuracies.

5.2. Maintaining quasi-optimal ranks

A crucial point when analysing the computational complexity of low-rank solvers is to derive bounds on the ranks of iterates. These have a major impact on the computational costs, since for hierarchical tensor representations, the costs of orthogonalization and HSVD – as in (2.23) and (2.24) – scale like the *fourth* power of the maximum hierarchical ranks. The natural point of reference for the ranks arising in an iterative method is given by the ranks required for low-rank best approximations of comparable accuracy. The best behaviour of the ranks of iterates one can hope for is that they remain of similar size to these best approximation ranks.

We next consider two constructions of methods that achieve this goal. We formulate these on $\ell_2(\mathbb{N}^d)$, corresponding to methods operating on the infinite-dimensional function spaces under consideration. The methods then automatically remain robust under a suitable selection of discretizations. Starting from an initial iterate $\mathbf{u}^0$ represented in a subspace tensor format, both constructions take the basic form

$$\mathbf{u}^{n+1} = (\mathcal{R}_n \circ \mathcal{F}_n)(\mathbf{u}^n).$$

Here, for each $n \in \mathbb{N}_0$, $\mathcal{F}_n$ provides an *error reduction* with respect to the exact solution $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ of $\mathbf{A}\mathbf{u} = \mathbf{f}$ in the sense that

$$\|\mathbf{u} - \mathcal{F}_n(\mathbf{u}^n)\|_{\ell_2} \le \rho \|\mathbf{u} - \mathbf{u}^n\|_{\ell_2} \tag{5.6}$$

with $\rho \in (0, 1)$ independent of n. The mapping $\mathcal{R}_n$ provides a *rank reduction* that may increase the error with respect to $\mathbf{u}$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

In principle, the error reduction operation $\mathcal{F}_n$ can be realized by any iterative solver that guarantees the property (5.6). However, to eventually allow for complexity estimates, this reduction should be sufficiently robust with respect to the dimensionality and other features of the problem. A simple suitable choice for $\mathcal{F}_n$ – as used in Bachmayr (2012*a*), Bachmayr and Dahmen (2015) and Bachmayr and Schneider (2017) – is one or more steps of Richardson iteration, which may also be performed inexactly such that (5.6) is guaranteed.

5.2.1. HSVD truncation

We first consider an approach for obtaining approximations with near-optimal ranks by HSVD truncation with sufficiently large error tolerance. In the case of low-rank tensor approximation, this was analysed in Bachmayr and Dahmen (2015) based on the iterative scheme developed in Bachmayr (2012*a*). The basic idea for establishing a connection to ranks of best approximations is applicable in a more general context, and we therefore first state this result in abstract terms.

Let H be a Hilbert space and let $(\mathcal{B}_r)_{r\in\mathbb{N}}$ with $\mathcal{B}_r \subset H$ be a sequence of subsets, with $\bigcup_{r\in\mathbb{N}} \mathcal{B}_r$ dense in H, from which approximations are selected, where increasing r corresponds to increasing complexity of approximations. For $v \in H$ and $\varepsilon > 0$, we define

$$b(v,\varepsilon) = \min\{r \in \mathbb{N} \colon \exists w \in \mathcal{B}_r \colon \|v - w\|_H \le \varepsilon\}.$$

In other words, $b(v,\varepsilon)$ is the minimal complexity parameter such that $\mathcal{B}_{b(v,\varepsilon)}$ contains a best approximation of accuracy ε. In addition, for each $v \in H$ and $\varepsilon > 0$, we assume that we have a linear orthogonal projection $B_{v,\varepsilon}$ onto $\mathcal{B}_{b(v,\varepsilon)}$ such that $\|v - B_{v,\varepsilon} v\|_H \le \varepsilon$.

Depending on the sets $\mathcal{B}_r$, it may be infeasible to compute best approximations, or even to exactly evaluate corresponding approximation errors. We thus assume that for all $r \in \mathbb{N}$ we have mappings $A_r \colon H \to \mathcal{B}_r$ and $E_r \colon H \to [0,\infty)$ such that for some $\kappa \ge 1$,

$$\|v - A_r(v)\|_H \le E_r(v) \le \kappa \inf_{w\in\mathcal{B}_r} \|v - w\|_H. \tag{5.7}$$

Here E_r plays the role of a computable quasi-optimal error bound for the approximations provided by A_r, which play the role of computable substitutes of best approximations. In our present setting of hierarchical tensor approximations, the HSVD truncation provides such a quasi-optimal substitute for low-rank best approximations, where a computable error bound is provided by the hierarchical singular values as in Theorem 2.36.

Lemma 5.1. For each $\varepsilon > 0$, let $a(v,\varepsilon) \in \mathbb{N}$ be chosen as the minimal integers such that $E_{a(v,\varepsilon)}(v) \le \varepsilon$. Then, for any $\alpha, \eta > 0$ and any $u, v \in H$ with $\|u - v\|_H \le \eta$,

$$\begin{aligned} \|u - A_{a(v,\kappa(1+\alpha)\eta)}(v)\|_H &\le (1 + \kappa(1+\alpha))\eta, \\ a(v,\kappa(1+\alpha)\eta) &\le b(u,\alpha\eta). \end{aligned} \tag{5.8}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Proof. Concerning the first estimate in (5.8), by the triangle inequality

$$\begin{aligned}\|u - A_{a(v,\kappa(1+\alpha)\eta)}(v)\|_H &\le \|u - v\|_H + \|v - A_{a(v,\kappa(1+\alpha)\eta)}(v)\|_H \\ &\le \eta + E_{a(v,\kappa(1+\alpha)\eta}(v),\end{aligned}$$

and $E_{a(v,\kappa(1+\alpha)\eta)}(v) \le \kappa(1+\alpha)\eta$ by definition.

For the proof of the second estimate, note that by linearity and orthogonality of $B_{u,\alpha\eta}$,

$$\|v - B_{u,\alpha\eta}v\|_H \le \|(v-u) - B_{u,\alpha\eta}(v-u)\|_H + \|u - B_{u,\alpha\eta}u\|_H \le (1+\alpha)\eta.$$

With (5.7), we obtain

$$E_{b(u,\alpha\eta)}(v) \le \kappa \inf_{w\in\mathcal{B}_{b(u,\alpha\eta)}} \|v - w\|_H \le \kappa\|v - B_{u,\alpha\eta}v\|_H \le \kappa(1+\alpha)\eta.$$

As a consequence,

$$a(v,\kappa(1+\alpha)\eta) = \min\{r\in\mathbb{N}\colon E_r(v) \le \kappa(1+\alpha)\eta\} \le b(u,\alpha\eta),$$

which is the second estimate in (5.8). □

In other words, Lemma 5.1 states that when an approximation v to u is re-approximated with a sufficiently large error tolerance, the complexity parameter of this reapproximation can be estimated by that of a best approximation of u with proportional error.

In its basic form, Lemma 5.1 goes back to Cohen, Dahmen and DeVore (2001, 2002), where it was used in the context of *best n-term approximation*; see also Cohen (2003, Thm 4.9.1). In this case we take $H = \ell_2(\mathbb{N})$ and for each $n \in \mathbb{N}$, $\mathcal{B}_n = \{\mathbf{v} \in \ell_2(\mathbb{N})\colon \#\operatorname{supp}\mathbf{v} \le n\}$. The mapping $A_n(\mathbf{v})$ is realized by retaining n entries of $\mathbf{v}$ of largest absolute value, so that here

$$\|\mathbf{v} - A_n(\mathbf{v})\|_{\ell_2} = \inf_{\mathbf{v}\in\mathcal{B}_n} \|\mathbf{v} - \mathbf{w}\|_{\ell_2}, \tag{5.9}$$

and thus (5.7) holds with $\kappa = 1$. The orthogonal projections realizing best approximations $B_{v,\varepsilon}$ in Lemma 5.1 are then the mappings that restrict to the respective supports and extend by zero.

For later use, we state the definition of approximation spaces for quantifying algebraic rates of best n-term approximation; see for example Dahmen (1997), DeVore (1998) and Cohen *et al.* (2001).

Definition 5.2. For $s > 0$ and $\mathbf{v} \in \ell_2(\mathbb{N})$, we set

$$|\mathbf{v}|_{\mathcal{A}^s(\mathbb{N})} = \sup_{n\in\mathbb{N}_0} n^s \min\{\|\mathbf{v} - \mathbf{w}\|_{\ell_2}\colon \#\operatorname{supp}\mathbf{w} \le n\}$$

and $\|\mathbf{v}\|_{\mathcal{A}^s(\mathbb{N})} = \|\mathbf{v}\|_{\ell_2(\mathbb{N})} + |\mathbf{v}|_{\mathcal{A}^s(\mathbb{N})}$, and define

$$\mathcal{A}^s(\mathbb{N}) = \{\mathbf{v} \in \ell_2(\mathbb{N})\colon |\mathbf{v}|_{\mathcal{A}^s(\mathbb{N})} < \infty\}.$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

The same definition applies to other countable index sets, and we simply write $\mathcal{A}^s$ when no confusion can arise. In other words, the linear spaces $\mathcal{A}^s$ comprise those elements $\mathbf{v} \in \ell_2(\mathbb{N})$ that satisfy

$$\min_{\#\operatorname{supp}\mathbf{w}\le n} \|\mathbf{v}-\mathbf{w}\|_{\ell_2} \le Cn^{-s}$$

for some $C > 0$, where the smallest possible such C is precisely $|\mathbf{v}|_{\mathcal{A}^s(\mathbb{N})}$. Conversely, these are the sequences that can be approximated to ℓ_2-error $\varepsilon > 0$ using $|\mathbf{v}|^{1/s}_{\mathcal{A}^s(\mathbb{N})}\varepsilon^{-1/s}$ non-zero coefficients. As a particular conclusion from Lemma 5.1, for elements of $\mathcal{A}^s$ we obtain approximations using at most a fixed multiple of the optimal number of coefficients.

In our setting of low-rank approximations by hierarchical tensors, we now apply Lemma 5.1 on the Hilbert space $\ell_2(\mathbb{N}^d)$ with

$$\mathcal{B}_r = \{\mathbf{v} \in \ell_2(\mathbb{N}^d)\colon \|\operatorname{rank}_{\mathbb{E}}(\mathbf{v})\|_\infty \le r\}.$$

A quasi-optimal error bound of the form (5.7) is then provided for HSVD truncation by Theorem 2.36. For $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ and $\mathsf{r} = (r_e)_{e\in\mathbb{E}}$, it yields

$$\begin{aligned}\|\mathbf{v}-\mathrm{P}_{\mathbf{v},\mathsf{r}}\mathbf{v}\|_{\ell_2} &\le \left(\sum_{e\in\mathbb{E}}\sum_{k>r_e}\big(\sigma_k^{[e]}(\mathbf{v})\big)^2\right)^{1/2}\\ &\le \kappa_{\mathbb{E}}\min\{\|\mathbf{v}-\mathbf{w}\|_{\ell_2}\colon \mathbf{w}\in\ell_2(\mathbb{N}^d),\ \operatorname{rank}_{\mathbb{E}}(\mathbf{w})\le\mathsf{r}\} \qquad (5.10)\end{aligned}$$

with $\kappa_{\mathbb{E}} = \sqrt{\#\mathbb{E}}$, and where $\mathrm{P}_{\mathbf{v},\mathsf{r}}\mathbf{v}$ is the HSVD truncation of $\mathbf{v}$ to hierarchical rank r. We adapt this to ensure prescribed truncation errors as follows.

Definition 5.3. For each $\eta > 0$ and $\mathbf{v} \in \ell_2(\mathbb{N}^d)$, we choose $\mathsf{r} = (r_e)_{e\in\mathbb{E}} \in \operatorname{Ranks}(\mathbb{E})$ with minimal $\|\mathsf{r}\|_\infty$ such that

$$\left(\sum_{e\in\mathbb{E}}\sum_{k>r_e}\big(\sigma_k^{[e]}(\mathbf{v})\big)^2\right)^{1/2} \le \eta. \qquad (5.11)$$

With such r, we define $\operatorname{Trunc}_\eta(\mathbf{v})$ as the truncation of the HSVD representation of $\mathbf{v}$ to rank r.

Our point of reference is provided by the maximum hierarchical ranks of best approximations with error at most $\eta > 0$,

$$r_{\mathrm{best}}(\mathbf{v},\eta) = \min\{r\in\mathbb{N}_0\colon (\exists \mathbf{w}\in\ell_2(\mathbb{N}^d)\colon \|\operatorname{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r \wedge \|\mathbf{v}-\mathbf{w}\|_{\ell_2}\le\eta)\},$$

where best approximations can trivially be written as orthogonal projections. Then, for the computable HSVD truncation operation defined in Definition 5.3, combining (5.10) and Lemma 5.1, we obtain the following rank bound in terms of best approximation ranks. Note that here the role of the quasi-optimal computable error bound is played by the expression on the left in (5.11), and the parameter that controls the approximation complexity is the maximum entry of the hierarchical rank as in Definition 5.3.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Lemma 5.4. Let $\kappa_{\mathbb{E}} = \sqrt{\#\mathbb{E}}$ and $\alpha > 0$. Then, for any $\mathbf{u}, \mathbf{v} \in \ell_2(\mathbb{N}^d)$ with $\|\mathbf{u} - \mathbf{v}\|_{\ell_2} \le \eta$ and

$$\hat{\mathbf{v}} = \mathrm{Trunc}_{\kappa_{\mathbb{E}}(1+\alpha)\eta}(\mathbf{v}),$$

we have

$$\|\mathbf{u} - \hat{\mathbf{v}}\|_{\ell_2} \le \big(1 + \kappa_{\mathbb{E}}(1+\alpha)\big)\eta, \quad \|\mathrm{rank}_{\mathbb{E}}(\hat{\mathbf{v}})\|_\infty \le r_{\mathrm{best}}(\mathbf{u}, \alpha\eta).$$

For low-rank approximations of operators corresponding to $d = 2$, a similar result is found independently in Dölz, Egger and Schlottbom (2021, Lem. 2.7).

We now turn to the construction of an iterative scheme using Lemma 5.4 based on a mapping $\mathcal{F}_n$ with the error reduction property (5.6). Here we assume that $\mathcal{F}_n$ satisfies (5.6) with a reduction factor

$$\rho < \frac{1}{2\big(1 + (1+\alpha)\kappa_{\mathbb{E}}\big)}, \tag{5.12}$$

where $\kappa_{\mathbb{E}} = \sqrt{\#\mathbb{E}} = \sqrt{2d-3}$ in the case of the hierarchical tensor format. This can always be achieved when $\mathcal{F}_n$ is realized by sufficiently many steps of (inexact) Richardson iteration.

Then, starting from $\mathbf{u}^0 = 0$ with $\|\mathbf{u} - \mathbf{u}^0\|_{\ell_2} = \|\mathbf{u}\|_{\ell_2} \le \|\mathbf{A}^{-1}\|_{\ell_2\to\ell_2}\|\mathbf{f}\|_{\ell_2} = \eta_0$, for $n = 0, 1, 2, \dots$, iterate

$$\mathbf{u}^{n+1} = \mathrm{Trunc}_{\theta\eta_n}\big(\mathcal{F}_n(\mathbf{u}^n)\big) \quad \text{with } \theta = \frac{(1+\alpha)\kappa_{\mathbb{E}}}{2(1+(1+\alpha)\kappa_{\mathbb{E}})}, \tag{5.13}$$

and set $\eta_{n+1} = \frac{1}{2}\eta_n$. We then have the following convergence result with rank estimates, which is used implicitly in Bachmayr and Dahmen (2015).

Theorem 5.5. For the sequence $(\mathbf{u}^n)_{n\in\mathbb{N}_0}$ defined by (5.13), for each $n \in \mathbb{N}_0$ we have

$$\|\mathbf{u} - \mathbf{u}^n\|_{\ell_2} \le \eta_n, \quad \|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \le r_{\mathrm{best}}\big(\mathbf{u}, (1+(1+\alpha)\kappa_{\mathbb{E}})^{-1}\alpha\eta_n\big). \tag{5.14}$$

Proof. For $n = 0$, the statement is clear. Assume now that it holds for an $n \in \mathbb{N}$. By (5.12), we have

$$\|\mathbf{u} - \mathcal{F}_n(\mathbf{u}^n)\|_{\ell_2} \le \frac{\eta_n}{2(1+(1+\alpha)\kappa_{\mathbb{E}})}.$$

Then, by combining (5.13) and Lemma 5.4, we obtain $\|\mathbf{u} - \mathbf{u}^{n+1}\|_{\ell_2} \le \frac{1}{2}\eta_n = \eta_{n+1}$ as well as

$$\begin{aligned}\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^{n+1})\|_\infty &\le r_{\mathrm{best}}\big(\mathbf{u}, \tfrac{1}{2}(1+(1+\alpha)\kappa_{\mathbb{E}})^{-1}\alpha\eta_n\big)\\ &\le r_{\mathrm{best}}\big(\mathbf{u}, (1+(1+\alpha)\kappa_{\mathbb{E}})^{-1}\alpha\eta_{n+1}\big),\end{aligned}$$

and the statement thus follows by induction. □

Note that Theorem 5.5 yields rank bounds only for each $\mathbf{u}^n$, $n \in \mathbb{N}_0$, but not for the intermediate results produced by $\mathcal{F}_n$. The possible growth of such ranks of

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

intermediate results depends strongly on the problem and on the particular method used for $\mathcal{F}_n$. We return to this point in Section 5.2.3.

Remark 5.6. The precise meaning of the rank bound in (5.14) depends on the low-rank approximability of $\mathbf{u}$.

(i) Under the assumption of *algebraic decay* of best low-rank approximation errors with respect to ranks, that is,

$$\inf_{\|\mathrm{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r} \|\mathbf{u}-\mathbf{w}\|_{\ell_2} \le C r^{-s}$$

with $C, s > 0$, from (5.14) we obtain

$$\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \lesssim d^{1/(2s)} \eta_n^{-1/s}.$$

Ranks are thus ensured to increase at the optimal rate with respect to the achieved accuracy, where a moderate algebraic dependence on d is possible.

(ii) The most interesting case for low-rank methods is that of best approximation errors with *exponential-type decay* with respect to ranks, as obtained for certain elliptic problems in Section 4.2. Assuming that

$$\inf_{\|\mathrm{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r} \|\mathbf{u}-\mathbf{w}\|_{\ell_2} \le C e^{-c r^\beta} \tag{5.15}$$

with some $c, C > 0$ and $\beta > 0$, recalling that $\kappa_{\mathbb{E}} = \sqrt{2d-3}$, we obtain

$$\begin{aligned}\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty &\le \left[\frac{1}{c}\ln\big(C\big(1+(1+\alpha)\sqrt{2d-3}\big)(\alpha\eta_n)^{-1}\big)\right]^{1/\beta} \\ &\lesssim (\ln d + |\ln \eta_n|)^{1/\beta}.\end{aligned}$$

Remark 5.7. To quantify low-rank approximability, we can also introduce approximation classes generalizing Definition 5.2 based on strictly increasing sequences $\gamma = (\gamma_n)_{n\in\mathbb{N}_0}$ with $\gamma_0 = 1$, $\gamma_n \to \infty$, with membership of $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ in the corresponding class $\mathcal{A}_{\mathbb{E}}(\gamma)$ determined by the condition

$$\|\mathbf{v}\|_{\mathcal{A}_{\mathbb{E}}(\gamma)} = \sup_{r\in\mathbb{N}_0} \gamma_r \inf\{\|\mathbf{v}-\mathbf{w}\|_{\ell_2} : \|\mathrm{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r\} < \infty.$$

This allows for a more fine-grained analysis of low-rank approximability: provided that γ grows at most exponentially, which corresponds to a restriction to $\beta \le 1$ in (5.15), we have in particular the additional estimate

$$\|\mathbf{u}^n\|_{\mathcal{A}_{\mathbb{E}}(\gamma)} \lesssim \|\mathbf{u}\|_{\mathcal{A}_{\mathbb{E}}(\gamma)}, \quad n \in \mathbb{N}_0,$$

which shows that the generated low-rank approximations inherit the low-rank approximability of $\mathbf{u}$. For details, see Bachmayr and Dahmen (2015, Sec. 3.1).

5.2.2. *Soft thresholding*

An alternative approach to controlling hierarchical ranks of iterates, proposed and analysed in Bachmayr and Schneider (2017), is based on the soft thresholding

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

procedure ST_δ, with $\delta > 0$, discussed in Section 2.7.2. In this case the analysis is more strongly tied to the underlying fixed point iteration, and for the basic analysis we thus assume that $\mathcal{F}_n = \mathcal{F}$, where $\mathbf{u} = \mathcal{F}(\mathbf{u})$ and

$$\|\mathcal{F}(\mathbf{v}) - \mathcal{F}(\mathbf{w})\|_{\ell_2} \le \rho \|\mathbf{v} - \mathbf{w}\|_{\ell_2} \quad \text{for all } \mathbf{v}, \mathbf{w} \in \ell_2(\mathbb{N}^d) \tag{5.16}$$

with $\rho \in (0, 1)$. This immediately implies

$$\|\mathbf{u} - \mathcal{F}(\mathbf{v})\|_{\ell_2} \le \rho \|\mathbf{u} - \mathbf{v}\|_{\ell_2} \quad \text{for all } \mathbf{v} \in \ell_2(\mathbb{N}^d). \tag{5.17}$$

Note that the following applies with *any* $\rho \in (0, 1)$, and we do not need a stronger error reduction assumption as in (5.12). Thus, in the present case, $\mathcal{F}$ satisfying (5.17) can always be realized by a *single* step of Richardson iteration.

Let us first observe that by the non-expansiveness of ST_δ according to Proposition 2.40, the mapping $\mathrm{ST}_\delta \circ \mathcal{F}$ retains the same contractivity property as $\mathcal{F}$ for any $\delta > 0$. This leads to the following result; for a proof, see Bachmayr and Schneider (2017, Lem. 4.1).

Lemma 5.8. Let $\mathcal{F}$ satisfy (5.16) with $\rho \in (0, 1)$, and let $\delta > 0$. Then $\mathrm{ST}_\delta \circ \mathcal{F}$ is also a contraction with Lipschitz constant ρ on $\ell_2(\mathbb{N}^d)$, and the unique fixed points $\mathbf{u}$ of $\mathcal{F}$ and $\mathbf{u}^\delta$ of $\mathrm{ST}_\delta \circ \mathcal{F}$ satisfy

$$(1 + \rho)^{-1} \|\mathbf{u} - \mathrm{ST}_\delta(\mathbf{u})\|_{\ell_2} \le \|\mathbf{u} - \mathbf{u}^\delta\|_{\ell_2} \le (1 - \rho)^{-1} \|\mathbf{u} - \mathrm{ST}_\delta(\mathbf{u})\|_{\ell_2}.$$

As a consequence, on the one hand, by contractivity of $\mathrm{ST}_\delta \circ \mathcal{F}$ for each fixed $\delta > 0$ we obtain

$$\|\mathbf{u}^\delta - (\mathrm{ST}_\delta \circ \mathcal{F})(\mathbf{v})\|_{\ell_2} \le \rho \|\mathbf{u}^\delta - \mathbf{v}\|_{\ell_2} \quad \text{for all } \mathbf{v} \in \ell_2(\mathbb{N}^d).$$

In other words, the thresholded iteration always converges at the same rate as the original one, but to a modified fixed point. On the other hand, combining Lemma 5.8 with Theorem 2.42 yields

$$\|\mathbf{u} - \mathbf{u}^\delta\|_{\ell_2} \lesssim \|\mathbf{u} - \mathrm{ST}_\delta(\mathbf{u})\|_{\ell_2} \to 0 \quad \text{as } \delta \to 0.$$

Rank bounds can be obtained with the aid of the following lemma, which is obtained in Bachmayr and Schneider (2017, Lem. 4.3) by a generalization of an observation in Cohen *et al.* (2001, Lem. 5.1), adapted to soft thresholding of sequences as in Dahlke, Fornasier and Raasch (2012).

Lemma 5.9. Let $\mathbf{u}, \mathbf{v} \in \ell_2(\mathbb{N}^d)$, let $\delta > 0$ and let $\eta > 0$ be such that $\|\mathbf{u} - \mathbf{v}\|_{\ell_2} \le \eta$. Then, for all $e \in \mathbb{E}$,

$$\mathrm{rank}_{[e]}\big(\mathrm{ST}_\delta(\mathbf{v})\big) \le \frac{4\eta^2}{\delta^2} + \#\big\{k \in \mathbb{N} \colon \sigma_k^{[e]}(\mathbf{u}) > \delta/2\big\}. \tag{5.18}$$

This means that rank estimates can be obtained by balancing the two terms on the right in (5.18) depending on the currently achieved accuracy η. The second term is a function of δ that depends on decay of matricization singular values of the exact solution $\mathbf{u}$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

A scheme for appropriate adjustment of thresholding parameters that does not require explicit knowledge of the singular value decay is given in Bachmayr and Schneider (2017). It requires a functional $\mathcal{E}\colon \ell_2(\mathbb{N}^d) \to [0,\infty)$ such that with some $\gamma, \Gamma > 0$, we have

$$\gamma\|\mathbf{u}-\mathbf{v}\|_{\ell_2} \le \mathcal{E}(\mathbf{v}) \le \Gamma\|\mathbf{u}-\mathbf{v}\|_{\ell_2} \quad \text{for all } \mathbf{v} \in \ell_2(\mathbb{N}^d).$$

In particular, when solving $\mathbf{A}\mathbf{u} = \mathbf{f}$, where $\mathbf{A}$ is an isomorphism on $\ell_2(\mathbb{N}^d)$, we may take $\mathcal{E}(\mathbf{v}) = \|\mathbf{A}\mathbf{v}-\mathbf{f}\|_{\ell_2}$ with $\gamma = \|\mathbf{A}^{-1}\|^{-1}_{\ell_2\to\ell_2}$, $\Gamma = \|\mathbf{A}\|_{\ell_2\to\ell_2}$.

The scheme proceeds as follows. With $\mathbf{u}^0 = 0$ and $\delta_0 = (\#\mathbb{E})^{-1}\|\mathcal{F}(\mathbf{u}^0)\|_{\ell_2}$, for each $n = 0, 1, 2, \ldots$:

$$\text{Set } \mathbf{u}^{n+1} = (\mathrm{ST}_{\delta_n} \circ \mathcal{F})(\mathbf{u}^n). \tag{5.19a}$$

$$\text{If } \|\mathbf{u}^{n+1}-\mathbf{u}^n\|_{\ell_2} \le \frac{(1-\rho)}{2\Gamma\rho}\mathcal{E}(\mathbf{u}^{n+1}), \text{ then set } \delta_{n+1} = \frac{\delta_n}{2}; \text{ otherwise, set } \delta_{n+1} = \delta_n. \tag{5.19b}$$

The basic idea is that by contractivity of $\mathcal{F}$,

$$\frac{\rho}{1-\rho}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_{\ell_2}$$

is an upper bound for $\|\mathbf{u}^{\delta_n}-\mathbf{u}^{n+1}\|_{\ell_2}$, whereas $\mathcal{E}(\mathbf{u}^{n+1})$ is proportional to $\|\mathbf{u}-\mathbf{u}^{n+1}\|_{\ell_2}$. The condition in (5.19b) thus ensures that the thresholding parameter is adjusted when the iterates are closer to $\mathbf{u}^{\delta_n}$ than to $\mathbf{u}$. The following convergence result with rank estimates for this iteration is shown in Bachmayr and Schneider (2017, Thm 5.1, Remark 5.6). On $\mathcal{F}$, we assume here only (5.16) with any $\rho \in (0,1)$.

Theorem 5.10. The iterates defined by (5.19) satisfy $\|\mathbf{u}-\mathbf{u}^n\|_{\ell_2} \to 0$. In addition we have the following.

(i) If $\sigma^{[e]}(\mathbf{u}) \in \mathcal{A}^s$ for an $s > 0$ for all $e \in \mathbb{E}$, then

$$\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \le C_1 d^{2+1/s} \max_{e\in\mathbb{E}} \|\sigma^{[e]}(\mathbf{u})\|^{1/s}_{\mathcal{A}^s} \|\mathbf{u}-\mathbf{u}^n\|^{-1/s}_{\ell_2},$$

where C_1 depends on ρ, γ, Γ and s.

(ii) If there exist $c, C, \beta > 0$ such that $\sigma^{[e]}_k(\mathbf{u}) \le Ce^{-ck^\beta}$ for all k and $e \in \mathbb{E}$, then

$$\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \le C_2 d^2 (\ln d + |\ln\|\mathbf{u}-\mathbf{u}^n\|_{\ell_2}|)^{1/\beta},$$

where C_2 depends on ρ, γ, Γ, c, C and β.

In both cases (i) and (ii), there exists C_3 depending on the same parameters as C_1 and C_2, respectively, such that for each $\varepsilon > 0$ we have $\|\mathbf{u}-\mathbf{u}^n\|_{\ell_2} \le \varepsilon$ when $n \ge C_3(|\ln\varepsilon| + \ln d)\ln d$.

Note that by Theorem 2.36, the assumptions on singular value decay made in Theorem 5.10 are closely connected to the assumptions on the decay of total approximation errors made in Remark 5.6. Theorem 5.10 thus yields very similar

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

near-optimal rank estimates to Theorem 5.5, but, in view of Remark 5.6, involving less favourable additional constants. Moreover, (5.19) is linearly convergent, where the number of required iterations may deteriorate logarithmically with respect to d.

Remark 5.11. Theorem 5.10 can be extended to the case where the evaluation of $\mathcal{F}$ and $\mathcal{E}$ is done not exactly but with controlled relative errors, for instance when the action of $\mathbf{A}$ is performed only approximately (as required when working with Riesz basis representations on sequence spaces with approximate rescaling as in Section 4.4). For details we refer to Bachmayr and Schneider (2017, Sec. 5.1).

5.2.3. Comparison of rank control strategies

The two strategies for obtaining rank bounds discussed in Sections 5.2.1 and 5.2.2 can both be combined with quite general error reduction steps. As a basic example, let us now consider solving a linear system $\mathbf{Au} = \mathbf{f}$ on $\ell_2(\mathbb{N}^d)$, where $\mathbf{A}$ is a symmetric elliptic isomorphism on $\ell_2(\mathbb{N}^d)$, with error reduction done by Richardson iteration. Note that non-symmetric elliptic operators can be treated completely analogously to what follows, but lead to different convergence estimates for Richardson iteration.

In the case of the iteration (5.19) with soft thresholding, we may simply take

$$\mathcal{F}(\mathbf{v}) = \mathbf{v} - \omega(\mathbf{Av} - \mathbf{f}), \tag{5.20}$$

which, with appropriately chosen step size parameter $\omega > 0$, yields (5.16) and (5.17) with a fixed error reduction factor $\rho \in (0, 1)$. Thus, while applying $\mathcal{F}$ as in (5.20) generally increases the ranks of iterates, the subsequent soft thresholding step immediately returns this intermediate result to a new iterate that satisfies the rank bounds of Theorem 5.10.

The situation is more delicate in the case of the iteration (5.13) with HSVD truncation, since here ρ needs to satisfy condition (5.12), which becomes more stringent with increasing d. To satisfy this requirement, for each n we can define the mapping $\mathcal{F}_n$ in (5.13) by

$$\mathcal{F}_n(\mathbf{v}) = \mathbf{v}^J, \quad \text{where } \mathbf{v}^{j+1} = \mathbf{v}^j - \omega(\mathbf{Av}^j - \mathbf{f}) \text{ for } j \in \mathbb{N}_0 \text{ with } \mathbf{v}^0 = \mathbf{v}, \tag{5.21}$$

where the stopping index J is chosen to achieve the required total error reduction. With appropriately chosen ω, by the standard convergence analysis of Richardson iteration, (5.12) is satisfied when

$$J = \zeta \operatorname{cond}(\mathbf{A}) \ln d$$

with some $\zeta > 0$. In this form, (5.13) consists of an outer iteration with rank bounds and the inner iteration (5.21). Unlike those for the iterates of the outer iteration, the rank bounds that can be achieved for the iterates $\mathbf{v}^1, \dots, \mathbf{v}^J$ in (5.21) depend strongly on the available knowledge on $\mathbf{A}$ and $\mathbf{f}$.

Remark 5.12. When $\mathbf{A}$ is a basis representation of an operator, in general each column of $\mathbf{A}$ has infinitely many non-zero entries. For this reason, in this case

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

the iterations (5.20) and (5.21), which are then equivalent to iterating on the entire represented function spaces, cannot be carried out in practice in this form. However, as one possible strategy for the adaptive refinement of discretizations, these iterations can be carried out approximately, with the result of applying $\mathbf{A}$ (as well as $\mathbf{f}$) approximated in each step by a sequence with finitely many non-zero entries. Such adaptive methods are discussed in Section 5.3.2.

Example 5.13. To give a simple example, let $\mathbf{A}$ and $\mathbf{f}$ be represented with fixed rank parameters bounded componentwise by $R_{\mathbf{A}} \in \mathbb{N}$ and $r_{\mathbf{f}} \in \mathbb{N}$, respectively. Then the largest representation ranks of an intermediate result arising with soft thresholding via (5.20) are bounded for each n by $\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty (R_{\mathbf{A}} + 1) + r_{\mathbf{f}}$.

In the case of (5.13), for the largest representation rank of an iterate arising in (5.21), in general we only have the bound

$$\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \frac{R_{\mathbf{A}}^{J+1} - 1}{R_{\mathbf{A}} - 1} + r_{\mathbf{f}} \frac{R_{\mathbf{A}}^{J} - 1}{R_{\mathbf{A}} - 1} \lesssim (\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty + r_{\mathbf{f}}) R_{\mathbf{A}}^{\zeta\, \mathrm{cond}(\mathbf{A}) \ln d}.$$

Note that $R_{\mathbf{A}}^{\zeta\, \mathrm{cond}(\mathbf{A}) \ln d} = d^{\zeta\, \mathrm{cond}(\mathbf{A}) \ln R_{\mathbf{A}}}$, which means that in this case rank bounds that depend polynomially on d require d-independent $\mathrm{cond}(\mathbf{A})$ and $R_{\mathbf{A}}$.

Remark 5.14. The above discussion with uniform representation rank bounds for operators does not apply directly to preconditioned operators on Sobolev spaces. Recall that in terms of Riesz bases of $H_0^1(\square_d)$ as in Remark 3.10, for $A\colon H_0^1(\square_d) \to H^{-1}(\square_d)$ we obtain operator representations of the form $\mathbf{A} = \mathbf{DTD}$, where the diagonal operator $\mathbf{D}$ does not have an explicit low-rank form. By Proposition 4.11, $\mathbf{D}$ can be replaced by an equivalent diagonal scaling operator $\tilde{\mathbf{D}}$ that is still of infinite rank but has efficient low-rank approximations $\tilde{\mathbf{D}}_{\eta,\Lambda}$ of relative error $\eta > 0$ on the selected subset of basis indices Λ. We thus work with the low-rank representations $\tilde{\mathbf{D}}_{\eta,\Lambda}\mathbf{T}\tilde{\mathbf{D}}_{\eta,\Lambda}$ approximating $\tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}$. This means, on the one hand, that the error reduction procedures discussed above necessarily need to be applied inexactly. On the other hand, since the representation rank of $\tilde{\mathbf{D}}_{\eta,\Lambda}$ depends on Λ, it is desirable to refine the discretization determined by Λ in parallel to the low-rank approximation of the solution to avoid unnecessarily large ranks.

5.3. Adaptive refinement of discretizations

Refining discretizations based on computable *a posteriori* error estimates is of particular interest in high-dimensional problems: in addition to the necessity of adaptively resolving non-smooth solutions, known *a priori* estimates are often difficult to apply in the high-dimensional case due to unknown dimension dependences in the arising constants.

An important requirement for the practicality of *a posteriori* estimates for large d, however, is that both the costs of their evaluation and the constants arising in the error bounds depend only moderately on d. For error estimates that are d-robust in

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

this sense, the next question is whether one can use them to drive a similarly robust adaptive refinement of discretizations.

Remark 5.15. Whether standard mesh-based approaches for error estimation in finite element discretizations, such as the classical residual error estimator, can be suitably adapted to high-dimensional finite elements in low-rank formats remains largely an open problem. A key difficulty, closely related to Remark 3.2, is establishing d-robustness of estimates. A finite element error estimate for tensor train representations using a *global* auxiliary flux reconstruction problem was proposed in Dolgov and Vejchodský (2021); it has the disadvantage, however, that for large d the auxiliary problem can be substantially more expensive to solve than the original problem.

In what follows, we focus on error estimation and discretization refinement for low-rank methods using basis expansions, which are comparably well studied. Here we start from a Riesz basis representation of the problem on $\ell_2(\mathbb{N}^d)$. We then select finite index sets $\Lambda \subset \mathbb{N}^d$ such that the span of the corresponding basis elements provides sufficiently accurate approximations.

These concepts have been used in the context of low-rank approximations, for instance in Bachmayr (2012*a*), Bachmayr and Dahmen (2015, 2016*a*,*b*), and Ali and Urban (2020) with wavelet Riesz bases. They appear in a similar form in Eigel, Pfeffer and Schneider (2017), Bachmayr, Cohen and Dahmen (2018) and Eigel, Marschall, Pfeffer and Schneider (2020) with orthonormal polynomial expansions as in Remark 3.3. For an overview, see also Bachmayr and Dahmen (2020).

Remark 5.16. As an additional constraint in our present setting of tensor representations, we restrict ourselves to discretizations defined by subsets of basis indices that have *Cartesian product structure*. For instance, discretizations in terms of basis representations on $\ell_2(\mathbb{N}^d)$ can be defined by $\Lambda \subset \mathbb{N}^d$ of the form

$$\Lambda = \bigtimes_{i=1}^{d} \Lambda^{(i)} \tag{5.22}$$

with finite $\Lambda^{(i)} \subset \mathbb{N}$ for $i = 1, \dots, d$. In principle it is also conceivable to adapt discretizations individually to each basis vector in the mode frames of a tensor representation. However, since algorithms using HSVD representations require repeated orthogonalizations of mode frames, such a further adaptation is impractical: since the discretization subspaces of all vectors are merged by orthogonalizations, there is no benefit in terms of computational costs.

In what follows, we consider basic techniques for error estimation and discretization refinement in the context of low-rank approximation, as well as their use within iterative solvers based on HSVD truncation. The approaches for generating sparse approximations with tensor structure discussed in Section 5.3.1 and the residual approximations considered in Section 5.3.3 are in fact independent of any

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

particular adaptive scheme. The computational methods treated in Sections 5.3.2 and 5.3.4 both rely on an extension of the strategy based on HSVD truncation in Section 5.2.1. The combination with adaptive discretization refinement of the alternative soft thresholding-based approach for rank control of Section 5.2.2 is still open.

5.3.1. Mode-wise sparse approximations

We first consider a method for selecting index sets of the form (5.22) that capture the entries of largest absolute values of a given tensor, but avoid looking at all individual entries (which would necessarily scale exponentially with respect to the tensor order d). Here we assume tensors on $\mathbb{N}^d$ for ease of presentation, but the following is applicable to general countable Cartesian product index sets.

The following notion was introduced in Bachmayr (2012*a*) and analysed further in Bachmayr and Dahmen (2015).

Definition 5.17. For $i = 1, \dots, d$ and $\mathbf{v} \in \ell_2(\mathbb{N}^d)$, we define the sequences $\pi^{(i)}(\mathbf{v}) = (\pi^{(i)}_\nu(\mathbf{v}))_{\nu\in\mathbb{N}}$, called *contractions* of $\mathbf{v}$, by

$$\pi^{(i)}_{\nu_i}(\mathbf{v}) = \left(\sum_{\nu_1=1}^{\infty} \cdots \sum_{\nu_{i-1}=1}^{\infty} \sum_{\nu_{i+1}=1}^{\infty} \cdots \sum_{\nu_d=1}^{\infty} \left|\mathbf{v}[\nu_1, \dots, \nu_{i-1}, \nu_i, \nu_{i+1}, \dots, \nu_d]\right|^2 \right)^{1/2}, \tag{5.23}$$

and in addition, we define

$$\operatorname{supp}_i \mathbf{v} = \operatorname{supp} \pi^{(i)}(\mathbf{v}) \subseteq \mathbb{N}.$$

The contractions $\pi^{(i)}(\mathbf{v})$, $i = 1, \dots, d$, can also be interpreted as the sequences of ℓ_2-norms of the rows of the matricizations $\operatorname{mat}_i(\mathbf{v})$ of $\mathbf{v}$. They can serve as criteria for assessing where the ℓ_2-norm of $\mathbf{v}$ is concentrated on the high-dimensional index set $\mathbb{N}^d$ by considering only sequences on $\mathbb{N}$. An important point in this regard is the following observation, which states that the high-dimensional summation on the right in Definition 5.17 can be avoided by exploiting orthogonality properties of the HSVD form.

Proposition 5.18. For $\mathbf{v} \in \ell_2(\mathbb{N}^d)$, let $\mathbf{U}^\alpha$, $\alpha \in \mathbb{T}$, be the mode frames of an HSVD of $\mathbf{v}$, and let $\sigma^\alpha_k(\mathbf{v})$ be the associated sequences of singular values. Then

$$\pi^{(i)}_\nu(\mathbf{v}) = \left(\sum_{k=1}^{\operatorname{rank}_i(\mathbf{v})} \left|\mathbf{U}^{\{i\}}_k[\nu]\, \sigma^{\{i\}}_k\right|^2 \right)^{1/2}, \quad \nu \in \mathbb{N}.$$

The following result from Bachmayr and Dahmen (2015) shows how the jointly sorted sequences $\pi^{(i)}(\mathbf{v})$ can be used to obtain finite product index sets of the form (5.22) for approximating $\mathbf{v}$ in ℓ_2. These index sets are quasi-optimal among all those of product structure with respect to the sum of the numbers of indices in each tensor mode. Here, for $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ and $\Lambda \subset \mathbb{N}^d$, we define the restriction $\mathrm{R}_\Lambda \mathbf{v} \in \ell_2(\mathbb{N}^d)$ as the sequence that is equal to $\mathbf{v}$ on Λ and vanishes on $\mathbb{N}^d \setminus \Lambda$.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Proposition 5.19. For any $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ and $N \in \mathbb{N}$, select N largest values in $\{\pi_\nu^{(i)}(\mathbf{v})\colon i = 1, \dots, d, \nu \in \mathbb{N}\}$, collect the corresponding indices $\nu \in \mathbb{N}$ of selected $\pi_\nu^{(i)}(\mathbf{v})$ in the respective sets $\Lambda_N^{(i)} \subset \mathbb{N}$ for $i = 1, \dots, d$, and define

$$\Lambda_N(\mathbf{v}) = \bigtimes_{i=1}^{d} \Lambda_N^{(i)}(\mathbf{v}).$$

Then, for each N,

$$\begin{aligned}\|\mathbf{v} - \mathrm{R}_{\Lambda_N(\mathbf{v})}\mathbf{v}\|_{\ell_2} &\le \left(\sum_{i=1}^{d} \sum_{\nu \in \mathbb{N}\setminus\Lambda_N^{(i)}(\mathbf{v})} \left|\pi_\nu^{(i)}(\mathbf{v})\right|^2\right)^{1/2} \\ &\le \sqrt{d}\, \min\left\{\|\mathbf{v} - \mathrm{R}_{\hat\Lambda}\mathbf{v}\|_{\ell_2}\colon \hat\Lambda = \bigtimes_{i=1}^{d} \hat\Lambda^{(i)},\ \sum_{i=1}^{d} \#\hat\Lambda^{(i)} \le N\right\}. \quad (5.24)\end{aligned}$$

Proof. The first estimate follows from the observation that for any $\Lambda = \bigtimes_{i=1}^{d} \Lambda^{(i)}$,

$$\begin{aligned}\|\mathbf{v} - \mathrm{R}_\Lambda\mathbf{v}\|_{\ell_2}^2 &\le \|\mathbf{v} - \mathrm{R}_{\Lambda^{(1)}\times\mathbb{N}^{d-1}}\mathbf{v}\|_{\ell_2}^2 + \cdots + \|\mathbf{v} - \mathrm{R}_{\mathbb{N}^{d-1}\times\Lambda^{(d)}}\mathbf{v}\|_{\ell_2}^2 \\ &= \sum_{i=1}^{d} \sum_{\nu \in \mathbb{N}\setminus\Lambda^{(i)}} \left|\pi_\nu^{(i)}(\mathbf{v})\right|^2.\end{aligned}$$

By construction of $\Lambda_N(\mathbf{v})$,

$$\sum_{i=1}^{d} \sum_{\nu \in \mathbb{N}\setminus\Lambda_N^{(i)}(\mathbf{v})} \left|\pi_\nu^{(i)}(\mathbf{v})\right|^2 = \min_{\sum_{i=1}^{d} \#\hat\Lambda^{(i)} \le N} \sum_{i=1}^{d} \sum_{\nu \in \mathbb{N}\setminus\hat\Lambda^{(i)}} \left|\pi_\nu^{(i)}(\mathbf{v})\right|^2.$$

Finally, we note that for $\hat\Lambda = \bigtimes_{i=1}^{d} \hat\Lambda^{(i)}$, we have

$$\sum_{\nu \in \mathbb{N}\setminus\hat\Lambda^{(i)}} \left|\pi_\nu^{(i)}(\mathbf{v})\right|^2 \le \|\mathbf{v} - \mathrm{R}_{\hat\Lambda}\mathbf{v}\|_{\ell_2}^2$$

for each $i \in \{1, \dots, d\}$. □

Based on Proposition 5.19, we define a procedure for producing approximations with finitely many non-zero entries of any desired accuracy $\eta > 0$ for arbitrary input sequences.

Definition 5.20. For $\mathbf{v} \in \ell_2(\mathbf{v})$ and $\eta > 0$, with $\Lambda_N(\mathbf{v})$ defined as in Proposition 5.19, and with N chosen as

$$N = \min\left\{N \in \mathbb{N}\colon \left(\sum_{i=1}^{d} \sum_{\nu \in \mathbb{N}\setminus\Lambda_N^{(i)}(\mathbf{v})} \left|\pi_\nu^{(i)}(\mathbf{v})\right|^2\right)^{1/2} \le \eta\right\},$$

we define

$$\mathrm{Coarse}_\eta(\mathbf{v}) = \mathrm{R}_{\Lambda_N(\mathbf{v})}\mathbf{v}.$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Note that $\mathrm{Coarse}_\eta(\mathbf{v})$ requires the HSVD form of $\mathbf{v}$ for evaluating the contractions $\pi^{(i)}(\mathbf{v})$ via Proposition 5.18. The costs of computing the HSVD form dominate those for subsequently selecting $\Lambda_N(\mathbf{v})$ with the appropriate N. With Lemma 5.1, the quasi-optimality property (5.24) immediately yields a result analogous to the one for Trunc_η, as in Lemma 5.4 for the support sizes produced by Coarse_η. More importantly for the present purposes, however, the two procedures can also be combined, which leads to a construction of adaptive methods, considered next.

5.3.2. Adaptive basis refinement by approximate Richardson iteration

We now come to our first method to combine rank adaptation with adaptive discretization refinement, proposed in Bachmayr (2012*a*) and analysed in Bachmayr and Dahmen (2015). Here we consider general Riesz basis representations $\mathbf{Au} = \mathbf{f}$ on $\ell_2(\mathbb{N}^d)$, where $\mathbf{A}\colon \ell_2(\mathbb{N}^d) \to \ell_2(\mathbb{N}^d)$ is an elliptic isomorphism.

As noted in Remark 5.12, in this setting, the rank-controlled iteration (5.13) combined with error reduction by an inner Richardson iteration (5.21) becomes computationally feasible only when the required residuals are approximated by finitely supported sequences. We thus assume for the moment that we have some routine Res_η at our disposal that for each $\eta > 0$ and $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ with $\#\operatorname{supp}\mathbf{v} < \infty$ produces an approximation of the residual $\mathbf{Av} - \mathbf{f}$ (which generally has infinite support) such that

$$\|\mathrm{Res}_\eta(\mathbf{v}) - (\mathbf{Av} - \mathbf{f})\|_{\ell_2} \le \eta, \tag{5.25}$$

and where $\mathrm{Res}_\eta(\mathbf{v})$ is in low-rank representation with finite support.

The basic idea is now to perform Richardson iteration on the Riesz basis representation, but with each required residual evaluation replaced by Res_η with appropriate $\eta > 0$. Similarly to iterations in low-rank format, we face the problem that when the method is carried out in this form, the number of non-zero basis indices that are active in each tensor mode may increase too strongly. This can be remedied by complementing the rank truncation in (5.13) by an additional basis coarsening by the routine Coarse_η from Definition 5.20.

For this combination we have the following generalization of Lemma 5.4; for the proof, we refer to Bachmayr and Dahmen (2015, Thm 7). Note that for the exact solution $\mathbf{u}$, it is natural to assume that $\pi^{(i)}(\mathbf{u}) \in \mathcal{A}^s(\mathbb{N})$ for $i = 1, \dots, d$ with some $s > 0$, since wavelet basis expansions as discussed in Section 3.2 typically lead to such algebraic convergence rates. These assumptions can be modified for more rapidly converging spectral approximations.

Lemma 5.21. Let $\mathbf{u}, \mathbf{v} \in \ell_2(\mathbb{N}^d)$ with $\pi^{(i)}(\mathbf{u}) \in \mathcal{A}^s(\mathbb{N})$ for some $s > 0$ for $i = 1, \dots, d$, such that $\|\mathbf{u} - \mathbf{v}\|_{\ell_2} \le \eta$. Let $\kappa_{\mathbb{E}} = \sqrt{\#\mathbb{E}}$, $\kappa_d = \sqrt{d}$. Let $\alpha > 0$; then for

$$\mathbf{w} = \mathrm{Coarse}_{\kappa_d(\kappa_{\mathbb{E}}+1)(1+\alpha)\eta}(\mathrm{Trunc}_{\kappa_{\mathbb{E}}(1+\alpha)\eta}(\mathbf{v}))$$

we have

$$\|\mathbf{u} - \mathbf{w}\|_{\ell_2} \le C_1(\alpha, \kappa_{\mathbb{E}}, \kappa_d)\eta,$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

where $C_1(\alpha, \kappa_{\mathbb{E}}, \kappa_d) = (1 + \kappa_{\mathbb{E}}(1+\alpha) + \kappa_d(\kappa_{\mathbb{E}}+1)(1+\alpha))$, and moreover

$$\|\mathrm{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r_{\mathrm{best}}(\mathbf{u}, \alpha\eta)$$

as well as

$$\sum_{i=1}^{d} \#\,\mathrm{supp}_i\,\mathbf{w} \le 2d\alpha^{-1/s}\left(\sum_{i=1}^{d}\|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s}\right)^{1/s}\eta^{-1/s},$$
$$\sum_{i=1}^{d}\|\pi^{(i)}(\mathbf{w})\|_{\mathcal{A}^s} \le C_2(\alpha, \kappa_{\mathbb{E}}, \kappa_d)\sum_{i=1}^{d}\|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s},$$

where $C_2(\alpha, \kappa_{\mathbb{E}}, \kappa_d, d, s) = 2^s(1+3^s) + 2^{4s}\alpha^{-1}C_1(\alpha, \kappa_{\mathbb{E}}, \kappa_d)d^{\max\{1,s\}}$.

The resulting rank- and basis-adaptive iterative method can be regarded as a generalization of the adaptive method or sparse approximation introduced in Cohen *et al.* (2002), which follows the same basic strategy. Let $\omega > 0$ be chosen such that $\|\mathbf{I} - \omega\mathbf{A}\|_{\ell_2\to\ell_2} \le \rho < 1$. With $\kappa_{\mathbb{E}} = \sqrt{\#\mathbb{E}} = \sqrt{2d-3}$, $\kappa_d = \sqrt{d}$ and $\alpha > 0$, let

$$\hat{\rho} = \frac{1}{1 + (1+\alpha)(\kappa_{\mathbb{E}} + \kappa_d + \kappa_{\mathbb{E}}\kappa_d)},$$
$$\theta_1 = \kappa_d(\kappa_{\mathbb{E}}+1)(1+\alpha)\hat{\rho}, \quad \theta_2 = (1+\alpha)\kappa_{\mathbb{E}}\hat{\rho},$$

as well as

$$J = \min\{j \in \mathbb{N}\colon \rho^j(1+\omega j) \le \tfrac{1}{2}\hat{\rho}\}.$$

Starting with $\mathbf{u}^0 = 0$, setting $\eta_0 = \|\mathbf{A}^{-1}\|_{\ell_2\to\ell_2}\|\mathbf{f}\|_{\ell_2}$ and $\eta_n = 2^{-n}\eta_0$ for $n = 0, 1, 2, \ldots,$

$$\mathbf{v}^{n,0} = \mathbf{u}^n, \quad \mathbf{v}^{n,j+1} = \mathbf{v}^{n,j} - \omega\,\mathrm{Res}_{\rho^j\eta_n}(\mathbf{v}^{n,j}),\ j = 0, 1, \ldots, J-1, \tag{5.26a}$$
$$\mathbf{u}^{n+1} = \mathrm{Coarse}_{\theta_1\eta_{n+1}}(\mathrm{Trunc}_{\theta_2\eta_{n+1}}(\mathbf{v}^{n,J})). \tag{5.26b}$$

Remark 5.22. The inner iterations (5.26a) can be performed with additional rank truncation and basis coarsening steps with sufficiently small tolerances; see Bachmayr and Dahmen (2015, Alg. 5.1). These do not influence the analysis based on Lemma 5.21 but in general improve the quantitative performance.

The proof of the following convergence result with rank and discretization size bounds for (5.26) is given in Bachmayr and Dahmen (2015).

Theorem 5.23. The sequence $(\mathbf{u}^n)_{n\in\mathbb{N}_0}$ defined by (5.26) for each $n \in \mathbb{N}_0$ satisfies $\|\mathbf{u} - \mathbf{u}^n\|_{\ell_2} \le \eta_n = 2^{-n}\eta_0$ and

$$\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \le r_{\mathrm{best}}(\mathbf{u}, \alpha\hat{\rho}\eta_n) \tag{5.27}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

as well as

$$\sum_{i=1}^{d} \#\operatorname{supp}_i \mathbf{u}^n \lesssim \left(\sum_{i=1}^{d} \|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s}\right)^{1/s} \eta_n^{-1/s},$$
$$\sum_{i=1}^{d} \|\pi^{(i)}(\mathbf{u}^n)\|_{\mathcal{A}^s} \lesssim \sum_{i=1}^{d} \|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s}, \tag{5.28}$$

with constants that depend only on α, s and algebraically on d.

Concerning the implications of the rank bound (5.27) when assuming algebraically or exponentially decaying best low-rank approximation errors, Remark 5.6 applies also in this case, with minor modifications to the d-dependence of the bounds.

Altogether, (5.26) provides a numerical scheme that operates on finitely many indices of the underlying Riesz basis (that is, on finitely supported coefficient sequences) in each step, and yet is guaranteed to converge to the *exact* solution of the PDE problem, represented by the coefficient sequence $\mathbf{u}$. After every rank truncation and basis coarsening step (5.26b), the resulting iterates $\mathbf{u}^n$ are guaranteed to have near-optimal ranks and support sizes compared to the respective best approximations of $\mathbf{u}$. A crucial point for the total computational costs of the method, however, is how precisely the residual approximation Res_η is realized. Similarly to Example 5.13, in particular, the support sizes should not increase too strongly during the inner iteration (5.26a).

5.3.3. Residual approximations

The adaptive algorithm (5.26) relies on a routine $\mathrm{Res}_\eta(\mathbf{v})$ that for each $\eta > 0$ produces a *finitely supported* approximation of the residual $\mathbf{Av}-\mathbf{f}$ (which is generally a sequence on $\mathbb{N}^d$ with infinitely many non-zero entries) with error at most η. In addition to the direct use of $\mathrm{Res}_\eta(\mathbf{v})$ in an iterative scheme, this procedure also yields computable *a posteriori* error estimates. For example, for given $\mathbf{v}$, starting from a sufficiently large value of η, by successively halving η until $\eta \le \frac{1}{2}\|\mathrm{Res}_\eta(\mathbf{v})\|_{\ell_2}$ holds, we obtain

$$\|\mathrm{Res}_\eta(\mathbf{v})\|_{\ell_2} \eqsim \|\mathbf{Av}-\mathbf{f}\|_{\ell_2} \eqsim \|\mathbf{v}-\mathbf{u}\|_{\ell_2}. \tag{5.29}$$

Therefore, due to the Riesz basis property, such a residual approximation also yields a computable quantity that remains proportional to the total error in approximating the exact solution of the underlying PDE.

Whereas suitable approximations of $\mathbf{f}$ are problem-dependent but can typically be derived from properties of the represented functional f, the approximate application of operators can be performed by universal strategies, which we focus on here. For the efficient realization of such approximations, we consider a standard construction introduced in Cohen *et al.* (2001) that is adapted to the adaptivity in the low-rank context in Bachmayr and Dahmen (2015, 2016*b*). It is based on the following notion of sparse approximability of operators represented by bi-infinite matrices.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Definition 5.24. Let $\mathcal{I}$ be a countable index set, and let $s^* > 0$. An operator $\mathbf{B}\colon \ell_2(\mathcal{I}) \to \ell_2(\mathcal{I})$ is called *s^*-compressible* if for any $s \in (0, s^*)$ there exist summable positive sequences $(\alpha_q)_{q\in\mathbb{N}_0}, (\beta_q)_{q\in\mathbb{N}_0}$ such that for each $q \in \mathbb{N}_0$ there exists $\mathbf{B}_q$ with at most $\alpha_q 2^q$ non-zero entries per row and column satisfying $\|\mathbf{B} - \mathbf{B}_q\| \le \beta_q 2^{-sq}$.

In particular, such approximations can be obtained for representations of differential (or integral) operators with respect to wavelet Riesz bases; see Cohen *et al.* (2001). For the adaptive compressibility of operators, the following generic strategy for sparse approximations is developed there based on Definition 5.24.

Remark 5.25. To obtain approximations $\mathbf{w}_\eta$ of $\mathbf{Bv}$ such that $\|\mathbf{w}_\eta - \mathbf{Bv}\|_{\ell_2} \le \eta$, for $\mathbf{B}$ as in Definition 5.24 and $\mathbf{v} \in \ell_2(\mathcal{I})$ with $N = \#\operatorname{supp}\mathbf{v} < \infty$, for each fixed $s \in (0, s^*)$ we can proceed as follows. First, let $\mathcal{I}_{[q]} \subset \mathcal{I}$ be the supports of the best 2^q-term approximations of $\mathbf{v}$ for $q = 0, 1, \dots, \lceil \log_2 N \rceil$, so that $\|\mathbf{v} - \mathrm{R}_{\mathcal{I}_{[q]}}\mathbf{v}\|_{\ell_2} \le |\mathbf{v}|_{\mathcal{A}^s} 2^{-sq}$. Then, for each $Q \le \lceil \log_2 N \rceil$, with $\mathcal{I}_{[-1]} = \emptyset$, consider the approximations

$$\hat{\mathbf{B}}_Q(\mathbf{v}) = \sum_{q=0}^{Q} \mathbf{B}_{Q-q}\, \mathrm{R}_{\mathcal{I}_{[q]}\setminus\mathcal{I}_{[q-1]}}\mathbf{v},$$

where $\mathbf{B}_{Q-q}$ are the approximations of $\mathbf{B}$ from Definition 5.24. Note that in this approximation the most accurate (and expensive) approximations of the operator $\mathbf{B}$ are applied only to the largest entries of the input vector $\mathbf{v}$. By the triangle inequality, we have the error bound

$$\begin{aligned}&\|\mathbf{Bv} - \hat{\mathbf{B}}_Q(\mathbf{v})\|_{\ell_2}\\ &\quad\le \|\mathbf{B}\|_{\ell_2\to\ell_2}\|\mathbf{v} - \mathrm{R}_{\mathcal{I}_{[Q]}}\mathbf{v}\|_{\ell_2} + \sum_{q=0}^{Q}\|\mathbf{B} - \mathbf{B}_{Q-q}\|_{\ell_2\to\ell_2}\|\mathrm{R}_{\mathcal{I}_{[q]}\setminus\mathcal{I}_{[q-1]}}\mathbf{v}\|_{\ell_2}.\end{aligned}$$

We now insert the estimates $\|\mathrm{R}_{\mathcal{I}_{[q]}\setminus\mathcal{I}_{[q-1]}}\mathbf{v}\|_{\ell_2} \le \|\mathbf{v} - \mathrm{R}_{\mathcal{I}_{[q-1]}}\mathbf{v}\|_{\ell_2} \le |\mathbf{v}|_{\mathcal{A}^s} 2^{-s(q-1)}$ as well as $\|\mathbf{B} - \mathbf{B}_{Q-q}\|_{\ell_2\to\ell_2} \le \beta_{Q-q} 2^{-s(Q-q)}$ to obtain $\|\mathbf{Bv} - \hat{\mathbf{B}}_Q(\mathbf{v})\|_{\ell_2} \lesssim |\mathbf{v}|_{\mathcal{A}^s} 2^{-sQ}$. Taking Q as the smallest value such that $\|\mathbf{Bv} - \hat{\mathbf{B}}_Q(\mathbf{v})\|_{\ell_2} \le \eta$, we set $\mathbf{w}_\eta = \hat{\mathbf{B}}_Q(\mathbf{v})$. Then

$$\begin{aligned}\#\operatorname{supp}\mathbf{w}_\eta &\le \sum_{q=0}^{Q} 2^{Q-q}\#\big(\mathcal{I}_{[q]} \setminus \mathcal{I}_{[q-1]}\big)\\ &\le \sum_{q=1}^{Q} 2^{Q-q}2^{q-1} + 2^Q \lesssim 2^Q \lesssim |\mathbf{v}|_{\mathcal{A}^s}^{1/s}\eta^{-1/s}.\end{aligned}$$

In addition, we verify as in Cohen *et al.* (2001) that

$$\|\mathbf{w}_\eta\|_{\mathcal{A}^s} \le C_{\mathcal{A}^s}\|\mathbf{v}\|_{\mathcal{A}^s}$$

with a constant $C_{\mathcal{A}^s} > 0$ independent of $\mathbf{v}$ and η.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Next we turn to the generalization of this strategy to the setting of low-rank approximations, where we consider $\mathbf{A}\colon \ell_2(\mathbb{N}^d) \to \ell_2(\mathbb{N}^d)$ of the form

$$\mathbf{A} = \sum_{n_1=1}^{R_1} \cdots \sum_{n_d=1}^{R_d} \mathbf{c}[n_1, \dots, n_d] \bigotimes_{i=1}^{d} \mathbf{A}_{n_i}^{(i)}, \tag{5.30}$$

with the core tensor $\mathbf{c}$ represented in hierarchical format. Our aim is to approximate $\mathbf{Av}$ for finitely supported $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ that is given in the same hierarchical tensor format. We assume that all component operators $\mathbf{A}_n^{(i)}\colon \ell_2(\mathbb{N}) \to \ell_2(\mathbb{N})$ for $i = 1, \dots, d$, $n = 1, \dots, R_i$, are s^*-compressible, with the corresponding approximations according to Definition 5.24 denoted by $\mathbf{A}_{n,q}^{(i)}$, $q \in \mathbb{N}_0$. We then consider approximations of the form

$$\hat{\mathbf{A}} = \sum_{n_1=1}^{R_1} \cdots \sum_{n_d=1}^{R_d} \mathbf{c}[n_1, \dots, n_d] \bigotimes_{i=1}^{d} \hat{\mathbf{A}}_{n_i}^{(i)}. \tag{5.31}$$

For the definition of $\hat{\mathbf{A}}_n^{(i)}$ for $i = 1, \dots, d$, $n = 1, \dots, R_i$, we choose suitable $\Lambda_q^{(i)} \subset \mathbb{N}$ for $q = 0, \dots, Q_i$ such that $\bigcup_{q=0}^{Q_i} \Lambda_q^{(i)} = \operatorname{supp}_i \mathbf{v}$ and $\Lambda_q^{(i)} \cap \Lambda_p^{(i)} = \emptyset$ whenever $q \neq p$. We now take

$$\hat{\mathbf{A}}^{(i)} = \sum_{q=0}^{Q_i} \mathbf{A}_{n,q}^{(i)} \mathrm{R}_{\Lambda_q^{(i)}}. \tag{5.32}$$

For approximations of the form (5.31), (5.32), we have the following simple observation.

Proposition 5.26. Let $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ with $\#\operatorname{supp} \mathbf{v} < \infty$, let $\mathbf{A}$ be as in (5.30) and $\hat{\mathbf{A}}$ as in (5.31), (5.32). Then

$$\|\mathbf{Av} - \hat{\mathbf{A}}\mathbf{v}\|_{\ell_2(\mathbb{N}^d)} \leq \sum_{i=1}^{d} C_{\mathbf{A}}^{(i)} \sum_{n=1}^{R_i} \sum_{q=0}^{Q_i} \|\mathbf{A}_n^{(i)} - \mathbf{A}_{n,q}^{(i)}\|_{\ell_2(\mathbb{N})\to\ell_2(\mathbb{N})} \big\|\mathrm{R}_{\Lambda_q^{(i)}} \pi^{(i)}(\mathbf{v})\big\|_{\ell_2(\mathbb{N})}$$

with $C_{\mathbf{A}}^{(i)} > 0$, $i = 1, \dots, d$, which can be determined from $\mathbf{A}$.

For a proof and a discussion of the constants $C_{\mathbf{A}}^{(i)}$, we refer to Bachmayr and Dahmen (2015, Lem. 4). We now follow a similar strategy to Remark 5.25, combining the bounds from Definition 5.24 for $\|\mathbf{A}_n^{(i)} - \mathbf{A}_{n,q}^{(i)}\|_{\ell_2(\mathbb{N})\to\ell_2(\mathbb{N})}$ with choosing the sets $\Lambda_q^{(i)}$ for $i = 1, \dots, d$ based on 2^q-term approximations of the respective (lower-dimensional) contractions $\pi^{(i)}(\mathbf{v})$. Proceeding in this manner, assuming that $\pi^{(i)}(\mathbf{v}) \in \mathcal{A}^s$ for $i = 1, \dots, d$ with an $s < s^*$, for each $\eta > 0$ we obtain a

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

finitely supported $\mathbf{w}_\eta = \hat{\mathbf{A}}\mathbf{v}$ that maintains the structure and ranks of the low-rank representation of $\mathbf{Av}$, such that $\|\mathbf{w}_\eta - \mathbf{Av}\|_{\ell_2(\mathbb{N}^d)} \le \eta$ and

$$\begin{aligned} \#\operatorname{supp}_i \mathbf{w}_\eta &\le C_1 \left(\sum_{i=1}^{d} \|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s} \right)^{1/s} \eta^{-1/s}, \\ \|\pi^{(i)}(\mathbf{w}_\eta)\|_{\mathcal{A}^s} &\le C_2 \|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s} \end{aligned} \tag{5.33}$$

for $i = 1, \dots, d$, where $C_1, C_2 > 0$ depend only on $\mathbf{A}$ and s (Bachmayr and Dahmen 2015, Thm 8).

Altogether, the computation of approximate residuals is thus reduced to operations on single tensor components. Note that in this case the entries of $\pi^{(i)}(\mathbf{v})$ determine to what accuracy the respective individual columns of the component operators of $\mathbf{A}$ in mode i are approximated.

Remark 5.27. Some further modifications are required for wavelet Riesz basis representations on $\ell_2(\vee^d)$ of problems on Sobolev spaces, which take the form (3.23). As noted in Remark 5.14, this can be replaced by (4.10) with the operator representation $\tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}$ with diagonal operators $\tilde{\mathbf{D}}$ from Proposition 4.11. Here $\mathbf{T}$ depends on the coefficients in the differential operator. In the case of the Laplacian, $\mathbf{T}$ has the explicit low-rank representation (3.22) with rank two, but its lower-dimensional component operators $\mathbf{T}_1$ generally have infinitely many non-zero entries in each column. For given $\mathbf{v} \in \ell_2(\vee^d)$, we thus consider approximations $\mathbf{w}_\eta$ such that $\|\tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}\mathbf{v} - \mathbf{w}_\eta\| \le \eta$ of the form

$$\mathbf{w}_\eta = \tilde{\mathbf{D}}_{\eta_1,\Lambda_1} \tilde{\mathbf{T}} \tilde{\mathbf{D}}_{\eta_2,\Lambda_2} \mathbf{v}. \tag{5.34}$$

Here $\tilde{\mathbf{D}}_{\eta_j,\Lambda_j}$, $j = 1, 2$, is defined as in Proposition 4.11 with suitable $\eta_j \eqsim \eta$, where $\Lambda_2 \supseteq \operatorname{supp} \mathbf{v}$ and $\Lambda_1 \supseteq \operatorname{supp}(\tilde{\mathbf{T}}\tilde{\mathbf{D}}_{\eta_2,\Lambda_2}\mathbf{v})$. The operator $\tilde{\mathbf{T}}$ is obtained by componentwise sparse approximations analogously to (5.31), (5.32), additionally taking into account the effect of the basis rescaling by $\tilde{\mathbf{D}}$. We thus combine sparse approximations of $\mathbf{T}$ with low-rank approximations of $\tilde{\mathbf{D}}$. For details and the corresponding estimates replacing (5.33), we refer to Bachmayr and Dahmen (2016*b*).

Remark 5.28. The estimates (5.33) can now be combined with the estimates from Theorem 5.23 for the iterates of the adaptive scheme (5.26). As noted above, for the evolution of the ranks of the iterates $\mathbf{v}^{n,j}$ of the inner iterations (5.26a), the conclusions of Remark 5.6 apply. As shown in Bachmayr and Dahmen (2016*b*, Sec. 6.5), if also $\|\pi^{(i)}(\mathbf{f})\|_{\mathcal{A}^s} < \infty$ for $i = 1, \dots, d$ and certain technical assumptions on the low-rank representation of $\mathbf{A}$ hold true (which are satisfied in the case of

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Remark 5.27), combining (5.28) with (5.33) we obtain estimates

$$\sum_{i=1}^{d} \#\operatorname{supp}_i \mathbf{v}^{n,j} \le C_{1,j} \left(\sum_{i=1}^{d} \max\{ \|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s}, \|\pi^{(i)}(\mathbf{f})\|_{\mathcal{A}^s} \} \right)^{1/s} \eta_n^{-1/s},$$

$$\sum_{i=1}^{d} \|\pi^{(i)}(\mathbf{v}^{n,j})\|_{\mathcal{A}^s} \le C_{2,j} \sum_{i=1}^{d} \max\{ \|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s}, \|\pi^{(i)}(\mathbf{f})\|_{\mathcal{A}^s} \}.$$

Here $C_{1,j}, C_{2,j} > 0$ depend polynomially on d, but may in general increase exponentially with respect to the inner iteration number j. After a fixed number of inner iteration steps, however, by the complexity reduction step (5.26b) we return to the estimates (5.28). In summary, we thus retain quasi-optimal sums of mode-wise support sizes throughout the entire iteration (5.26).

5.3.4. Iteratively refined Galerkin discretizations

An alternative construction of adaptive methods that is conceptually closer to adaptive finite element methods is based on solving successively refined Galerkin discretizations. Methods of this type with quasi-optimal computational costs for sparse approximation by wavelets were developed and analysed in Cohen *et al.* (2001).

We first sketch the basic approach. Here we need to assume, in addition to boundedness and ellipticity, that $\mathbf{A}$ is a symmetric operator, so $\|\mathbf{v}\|_{\mathbf{A}} = \langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle^{1/2}$ for $\mathbf{v} \in \ell_2(\mathbb{N}^d)$ defines a norm on $\ell_2(\mathbb{N}^d)$; note that symmetry is not required for the scheme (5.26). The method proceeds by stepwise refinement of finite index sets $\Lambda^n \subset \mathbb{N}^d$, $n = 0, 1, 2, \ldots$, starting from a given Λ^0. For each n, we first determine the Galerkin solution $\mathbf{u}_{\Lambda^n}$, which is uniquely defined by the conditions $\operatorname{supp} \mathbf{u}_{\Lambda^n} \subseteq \Lambda^n$ and

$$\mathrm{R}_{\Lambda^n} \mathbf{A} \mathbf{u}_{\Lambda^n} = \mathrm{R}_{\Lambda^n} \mathbf{f}. \tag{5.35a}$$

With $\mathbf{r}_{\Lambda^n} = \mathbf{A}\mathbf{u}_{\Lambda^n} - \mathbf{f}$ and a fixed $\alpha \in (0, 1]$, we select the smallest $\hat{\Lambda} \supseteq \Lambda^n$ such that

$$\|\mathrm{R}_{\hat{\Lambda}} \mathbf{r}_{\Lambda^n}\|_{\ell_2} \ge \alpha \|\mathbf{r}_{\Lambda^n}\|_{\ell_2} \tag{5.35b}$$

and set $\Lambda^{n+1} = \hat{\Lambda}$. Here, for computational purposes, $\mathbf{r}_{\Lambda^n}$ is again replaced by a finitely supported approximation. As shown in Cohen *et al.* (2001, Lem. 4.1), (5.35b) ensures

$$\|\mathbf{u} - \mathbf{u}_{\Lambda^{n+1}}\|_{\mathbf{A}} \le \rho \|\mathbf{u} - \mathbf{u}_{\Lambda^n}\|_{\mathbf{A}},$$

with $\rho \in (0, 1)$ that depends on α and $\operatorname{cond}(\mathbf{A})$. Cohen *et al.* (2001) combined this error reduction with a basis coarsening by application of Lemma 5.1 to best n-term approximation as in (5.9). Assuming $\mathbf{u} \in \mathcal{A}^s$ for some $s > 0$, this again leads to a method using a quasi-optimal number of non-zero coefficients for each iterate.

Gantumur, Harbrecht and Stevenson (2007) adapted the method to avoid the coarsening steps: provided that α in (5.35b) is chosen in $(0, \operatorname{cond}^{-1/2}(\mathbf{A}))$, the

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

index sets Λ^n obtained from (5.35b) have quasi-optimal cardinality without further coarsening. This variation of the method, termed the *adaptive wavelet-Galerkin method*, yields improved quantitative performance. For an overview, see also Stevenson (2009).

The above approach, based on successively refined Galerkin discretizations, has been adapted to the setting of low-rank tensor approximations in Ali and Urban (2020). The Galerkin solutions $\mathbf{u}_{\Lambda^n}$ are approximated by a rank-truncated PCG method, and finitely supported low-rank approximations of $\mathbf{r}_{\Lambda^n}$ are computed by the techniques discussed in Section 5.3.3. Index sets $\hat{\Lambda}$ with product structure in (5.35b) are obtained by applying Coarse_η to $\mathbf{r}_{\Lambda^n}$ with a suitable tolerance $\eta > 0$, as in Definition 5.3.

To obtain rank and discretization size estimates as in Theorem 5.23, this variant also relies on the combination of Coarse_η and Trunc_η as in (5.26b). In principle, the scheme is thus of a similar structure to the adaptive low-rank Richardson iteration (5.26), but with the inner iteration (5.26a) replaced by sufficiently many steps of solving a Galerkin discretization and selecting new degrees of freedom via (5.35b). The conclusions concerning rank and discretization size bounds for all iterates obtained in Ali and Urban (2020) are analogous to those for (5.26), but with quantitative advantages concerning the practical performance.

In contrast to the setting of best n-term approximations, the question remains open whether it is possible to obtain quasi-optimal discretization sizes without applying a basis coarsening step as with Coarse_η in (5.26b). The reason is that the effective value of α obtained in (5.35b) when selecting $\hat{\Lambda}$ using Coarse_η cannot in general be chosen sufficiently small for larger d, since the selected index sets are only quasi-optimal up to a factor $\sqrt{d}$ as in (5.24). For a detailed discussion, see Ali and Urban (2020, Sec. 4.7).

5.4. *Computational complexity*

In Section 5.2 we consider two strategies, based on HSVD truncation and on soft thresholding, for maintaining quasi-optimal ranks in iterative low-rank solvers. While these approaches can be formulated on infinite-dimensional sequence spaces, in computationally realizable methods they need to be combined with suitable discretizations. A basic approach for guaranteeing prescribed discretization errors in high dimensions is given in Section 5.3, based on combined low-rank and sparse approximations of residuals. On this basis, iterative solvers that complement HSVD truncation with basis coarsening can be constructed, with error reduction realized by approximate Richardson iteration as in (5.26) or by solving successively refined Galerkin discretizations as in (5.35).

Remark 5.29. An important feature of these adaptive solvers is their *universality*, in the sense that they can be carried out without any knowledge of the approximability of the exact solution $\mathbf{u}$: we only need to know suitable approximability of

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

the problem data $\mathbf{A}$ and $\mathbf{f}$. In particular, plain convergence of the methods to $\mathbf{u}$ is obtained without any assumptions on low-rank or sparse approximability of $\mathbf{u}$.

Further statements on the performance of the methods then require additional assumptions on the approximability of $\mathbf{u}$. As a consequence of Theorem 5.23 and Remark 5.28, if $\mathbf{u}$ has a certain low-rank and sparse approximability, this will be reflected in the ranks and support sizes of iterates. The two crucial approximability assumptions concern intrinsic properties of $\mathbf{u}$: on the one hand, convergence of best approximations of prescribed maximum hierarchical rank, which is determined by the decay of singular value sequences $\sigma^{[e]}(\mathbf{u})$ for $e \in \mathbb{E}$, and on the other hand, near-sparsity of the contractions $\pi^{(i)}(\mathbf{u})$, $i = 1, \dots, d$, quantified by finiteness of $\mathcal{A}^s$-norms for some $s > 0$.

Based on such rank and support size estimates for iterates, we arrive at estimates for the total computational costs of these adaptive solvers; in other words, we obtain estimates for the computational complexity of deterministically solving high-dimensional PDEs up to a guaranteed total approximation error.

Remark 5.30. The computational costs of the iteration (5.26) are generally dominated by those of the computation of HSVD forms of the iterates $\mathbf{v}^{n,j}$. As noted in (2.24), the number of arithmetic operations this requires for each n and j is bounded by

$$C\left(d\|\mathrm{rank}_{\mathbb{E}}(\mathbf{v}^{n,j})\|_\infty^4 + \|\mathrm{rank}_{\mathbb{E}}(\mathbf{v}^{n,j})\|_\infty^2 \sum_{i=1}^{d} \#\,\mathrm{supp}_i\, \mathbf{v}^{n,j}\right) \tag{5.36}$$

with a fixed d-independent constant $C > 0$. Assuming $\pi^{(i)}(\mathbf{u}) \in \mathcal{A}^s$, $i = 1, \dots, d$, for some $s > 0$, we compare the implications of this bound for the two different types of low-rank approximability considered in Remark 5.6. For the iterates $\mathbf{u}^n$ with $\|\mathbf{u} - \mathbf{u}^n\|_{\ell_2} \le \eta_n$ of the outer iteration (5.26b), as a consequence of Theorem 5.23, the complexity bound (5.36) takes the following forms.

(i) Assuming $\inf_{\|\mathrm{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r} \|\mathbf{u} - \mathbf{w}\|_{\ell_2} \le Cr^{-p}$ with $C, p > 0$, we obtain $\|\mathrm{rank}_{\mathbb{E}}(\mathbf{u}^n)\|_\infty \lesssim d^{1/(2p)}\eta_n^{-1/p}$, and thus the HSVD of $\mathbf{u}^n$ costs at most

$$C_1\left(\eta_n^{-4/p} + \eta_n^{-2/p-1/s}\right)$$

arithmetic operations, where C_1 depends on $\mathbf{u}$ and polynomially on d. Depending on p, the costs may thus scale much less favourably than the sizes $O\left(\eta_n^{-1/s}\right)$ of the discretizations in each tensor mode.

(ii) Under the stronger assumption

$$\inf_{\|\mathrm{rank}_{\mathbb{E}}(\mathbf{w})\|_\infty \le r} \|\mathbf{u} - \mathbf{w}\|_{\ell_2} \le Ce^{-cr^\beta} \tag{5.37}$$

with $c, C, \beta > 0$, the costs of the HSVD of $\mathbf{u}^n$ can be estimated by

$$C_2\left(|\log \eta_n|^4 + |\log \eta_n|^2 \eta_n^{-1/s}\right),$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

where C_2 depends again on $\mathbf{u}$ and polynomially on d. In this more favourable case, the second summand in the above bound dominates and the costs differ from those of the lower-dimensional discretizations in each tensor mode only by a logarithmic factor.

The ranks of the iterates $\mathbf{v}^{n,j}$ of the inner iteration additionally depend on the representation ranks of residual approximations and are thus more problem-dependent.

Example 5.31. To consider the total computational costs, we return to the example of a second-order elliptic operator $A\colon H^1_0(\square_d) \to H^{-1}(\square_d)$ with wavelet Riesz basis representation $\tilde{\mathbf{A}} = \tilde{\mathbf{D}}\mathbf{T}\tilde{\mathbf{D}}$ on $\ell_2(\vee^d)$ and residual approximation as outlined in Remark 5.27. We assume $\mathbf{T}$ to be of fixed rank such that $\operatorname{cond}(\tilde{\mathbf{A}})$ is bounded independently of d. Moreover, we assume $\mathbf{u}$, motivated by the results for the Poisson problem from Section 4.2, to have low-rank best approximations with exponential-type convergence (5.37). In addition, we assume $\mathbf{f}$ to have the same type of low-rank approximability, and that for some $s > 0$, $\|\pi^{(i)}(\mathbf{u})\|_{\mathcal{A}^s}, \|\pi^{(i)}(\mathbf{f})\|_{\mathcal{A}^s} < \infty$ for $i = 1, \dots, d$. With some further technical assumptions, a bound for the total computational costs in this setting is obtained in Bachmayr and Dahmen (2016*b*). Here, a major issue is taking into account the connection between discretization and preconditioning ranks in the residual approximations (5.34). Altogether, (5.26) under these conditions is guaranteed to produce for each given $\varepsilon > 0$ an approximation $\mathbf{u}_\varepsilon$ such that $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell_2} \le \varepsilon$ using a total number of floating-point operations bounded by

$$C_0 d^{c_1 \ln d}(1 + |\ln \varepsilon|)^{c_2 \ln d + 2/\beta} \varepsilon^{-1/s}, \tag{5.38}$$

where $C_0, c_1, c_2 > 0$ are independent of d and ε. Recall that for the functions $u, u_\varepsilon \in H^1(\square_d)$ represented by $\mathbf{u}$ and $\mathbf{u}_\varepsilon$, respectively, we have $\|u - u_\varepsilon\|_{H^1} \eqsim \|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell_2}$. Up to logarithmic factors, we thus obtain the costs $O(\varepsilon^{-1/s})$ that are optimal for the approximation of each one-dimensional component. Moreover, the curse of dimensionality is avoided: the leading d-dependence $d^{c_1 \ln d} = e^{c_1 \ln^2 d}$ grows faster than any polynomial, but substantially slower than exponentially in d.

This shows that for this typical class of problems, the curse of dimensionality can be avoided while achieving (deterministic) computable *a posteriori* error bounds for the approximate solutions. However, the above complexity bounds are likely not sharp: the numerical tests in Bachmayr and Dahmen (2016*b*) indicate that for the Poisson problem and certain variants, the total numerical costs exhibit low-order polynomial scaling with respect to d.

For the approach using Galerkin discretizations, analogous conclusions for Example 5.31 are obtained in Ali and Urban (2020). For the alternative rank control by soft thresholding discussed in Section 5.2.2, the analysis of a combination with adaptively refined discretizations is still open, but one may conjecture that this can

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

lead to an improved d-dependence in the total cost estimates compared to (5.38). For discussions of the total computational complexity of (5.26) under different assumptions, see also Bachmayr and Dahmen (2015, 2020).

5.5. Approximability and applications to high-dimensional elliptic problems

We now turn to applications to elliptic PDEs beyond the guiding model problems considered so far. As we have seen, the efficiency of low-rank solvers is determined in particular by the low-rank approximability of the corresponding exact solutions. On the one hand, the approximability of solutions can be studied experimentally by solvers with rank guarantees, as in Theorems 5.5 or 5.10. On the other hand, it is also of interest to obtain analytical results that yield efficient low-rank approximations. Such investigations can also shed some light on the mechanisms and structural features of solutions that lead to rapidly convergent best low-rank approximations.

General low-rank approximability results for functions of classical or mixed Sobolev regularity are obtained in Schneider and Uschmajew (2014), Griebel and Harbrecht (2014, 2019, 2023) and Bachmayr, Nouy and Schneider (2021*b*), where results for functions with compositional structure are also obtained, and in Griebel, Harbrecht and Schneider (2022). These results based on regularity properties only yield algebraic convergence of low-rank best approximations. In many applications, however, we observe exponential-type decay of best approximation errors, which in view of Remark 5.30 leads to substantially more favourable complexity of solvers. Such stronger approximability is often tied to particular structural features of solutions.

5.5.1. High-dimensional diffusion problems

The results discussed in Section 4.2, especially Theorem 4.6 from Dahmen *et al.* (2016), show low-rank approximability of solutions of Poisson problems as in Example 3.1 on $\square_d$ with potentially large d. In this case we obtain exponential-type decay of best approximation errors with respect to ranks, or conversely, best approximation ranks that grow like a power of a logarithm of the achieved error.

More generally, one may be interested in problems of finding $u \in H^1_0(\square_d)$ such that

$$\int_{\square_d} \hat{M}\nabla u \cdot \nabla v \,\mathrm{d}x = \int_{\square_d} f v \,\mathrm{d}x \quad \text{for all } v \in H^1_0(\square_d)$$

with a diffusion coefficient $\hat{M} \in \mathbb{R}^{d\times d}$. The above-mentioned results apply only to diagonal $\hat{M}$. The results for discretized problems in Kressner and Uschmajew (2016) are applicable to more general $\hat{M}$, but yield only algebraic decay of best approximation errors with bounds that deteriorate under discretization refinement. Discretization-independent bounds for the approximability of solutions with general $\hat{M}$ are still an open problem. Experimentally, for problems with certain tridiagonal $\hat{M}$, Bachmayr and Dahmen (2016*b*) and Kressner *et al.* (2016) have

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

observed low-rank approximability similar to that for diagonal $\hat{M}$. Slightly more efficient approximations are obtained when errors are controlled only in $L_2(\square_d)$ rather than in $H^1_0(\square_d)$; see Bachmayr and Dahmen (2016*a*).

5.5.2. Parameter-dependent elliptic problems

In parameter-dependent PDEs as outlined in (1.5), the problems considered are typically posed on domains $\Omega \subset \mathbb{R}^d$ with $d \in \{1, 2, 3\}$, but with solutions that are still high-dimensional functions due to an additional dependence on a parameter from some set Y. A frequently treated model case is that of second-order elliptic problems with diffusion coefficients $a(\cdot, y) \in L_\infty(\Omega)$ with a parameter $y \in Y$, where for each y, the solution $u(\cdot, y) \in H^1_0(\Omega)$ for fixed right-hand side $f \in L_2(\Omega)$ solves

$$\int_\Omega a(x, y)\nabla u(x, y) \cdot \nabla v(x)\,\mathrm{d}x = \int_\Omega f(x)\, v(x)\,\mathrm{d}x \quad \text{for all } v \in H^1_0(\Omega). \tag{5.39}$$

Different types of parameter dependences of a are considered in the literature.

Example 5.32. A frequently used class of model problems is that of coefficients with affine parametrizations of the form

$$a(x, y) = \bar{a}(x) + \sum_{j\in\mathcal{N}} y_j \theta_j(x) \tag{5.40}$$

with $y = (y_1, y_2, \ldots) \in Y = [-1, 1]^{\mathcal{N}}$, where $\mathcal{N}$ may be finite or countably infinite and $\bar{a}$ as well as θ_j, $j \in \mathcal{N}$, are fixed functions in $L_\infty(\Omega)$. Two examples in which ellipticity holds uniformly in $y \in Y$ are as follows:

(A) $\#\mathcal{N} < \infty$, $\bar{a} = 1$ and $\theta_j = c_j \chi_{\Omega_j}$, where $\Omega_j \subset \Omega$ are disjoint subdomains and c_j are constants with $|c_j| < 1$;

(B) $\mathcal{N} \simeq \mathbb{N}$ and $\|\theta_j\|_{L_\infty} \lesssim j^{-r}$ with some $r > 0$ such that $\bar{a} - \sum_{j\in\mathcal{N}} |\theta_j| \geq a_0 > 0$ in Ω for some $a_0 \in \mathbb{R}$.

Particularly in a stochastic context, where the entries of y correspond to scalar random variables, it is natural to consider weak formulations in the parametric domain as well. For simplicity, we consider the uniform distribution on Y: let μ_1 be the uniform measure on $[-1, 1]$, so that $\mu = \bigotimes_{i\in\mathcal{N}} \mu_1$ is the uniform measure on $Y = [-1, 1]^{\mathcal{N}}$. We then seek $u \in \mathcal{V} = H^1_0(\Omega) \otimes L_2(Y)$ such that for all $v \in \mathcal{V}$,

$$\int_Y \int_\Omega a(x, y)\nabla u(x, y) \cdot \nabla v(x, y)\,\mathrm{d}x\,\mathrm{d}\mu(y) = \int_Y \int_\Omega f(x)\, v(x, y)\,\mathrm{d}x\,\mathrm{d}\mu(y). \tag{5.41}$$

Here the space $\mathcal{V}$ is a tensor product; in the case $\#\mathcal{N} < \infty$, we have in particular

$$\mathcal{V} = H^1_0(\Omega) \otimes L_2(Y) = H^1_0(\Omega) \otimes \bigotimes_{i\in\mathcal{N}} L_2(-1, 1).$$

A natural choice of basis for $L_2(Y)$ is that of tensor product orthonormal polynomials as in Example 3.3, which can be combined with any discretization or Riesz

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

basis for $H_0^1(\Omega)$. Since the spatial coordinate is usually treated as a single tensor mode here, in solving (5.41) one can work with *tensor product Riesz bases* of $\mathcal{V}$. This means that the issues of low-rank preconditioning discussed in Section 4.4 do not arise here.

The low-rank approximability of solutions depends strongly on the particular problem, especially on the structure of the parametrization of a. Bachmayr and Cohen (2017) have obtained bounds on best low-rank approximation errors for case (A) of Example 5.32. They show favourable exponential decay of errors and, for problems with certain structural features, a substantial improvement over direct product polynomial approximations in the parametric variable of the solution. This is confirmed by numerical tests in Bachmayr *et al.* (2018).

The situation is different, however, in problems with infinitely many parameters, as in case (B) of Example 5.32. The anisotropy in the parametric variables, due to the algebraic decay of $\|\theta_j\|_{L_\infty}$, can be exploited by sparse product polynomial approximations that converge independently of any parametric dimensionality parameter; see Cohen, DeVore and Schwab (2010, 2011) and Bachmayr, Cohen, Dũng and Schwab (2017). Low-rank approximability of such problems is considered in Bachmayr *et al.* (2018). As these results show, the singular values of the particular matricization with spatial degrees of freedom in its rows and parametric ones in its columns can generally be expected to decay only algebraically. In such cases, concerning the asymptotic total computational complexity of solvers, we can generally not expect low-rank approximations to offer an advantage over direct sparse approximations.

Low-rank approximation methods based on stochastic Galerkin formulations (5.41) with fixed discretizations are considered, for instance, in Khoromskij and Schwab (2011), Kressner and Tobler (2011*a*), Tobler (2012), Matthies and Zander (2012) and Lee and Elman (2017). Methods with adaptive refinement of spatial and parametric discretizations are considered in Eigel *et al.* (2017) based on adaptive finite elements for the spatial degrees of freedom, and in Bachmayr *et al.* (2018) using wavelets in space. The more involved case of lognormal coefficient parametrizations is considered in Eigel *et al.* (2020).

Remark 5.33. As an alternative to variational formulations (5.41), approximations of parameter-dependent solutions can also be constructed from point evaluations with respect to y, that is, from solutions of (5.39) for certain well-chosen parameter values. In the low-rank context, this can be done by the *TT-cross* method for tensor trains (Oseledets and Tyrtyshnikov 2010) or its counterpart for hierarchical tensors (Ballani, Grasedyck and Kluge 2013). Such approaches are used for parametric problems, for instance in Ballani and Grasedyck (2015) and Dolgov and Scheichl (2019). A different approach based on interpolation of parametric point evaluations is taken in Khoromskij and Oseledets (2010). Compared to (5.41), such methods offer additional flexibility, especially in problems with a depending

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

nonlinearly on y, but we do not have the mechanisms for error control as offered by (5.41).

Remark 5.34. The application of low-rank methods to parameter-dependent problems is strongly connected to certain methods for model order reduction, in particular to *proper orthogonal decomposition* (Kahlbacher and Volkwein 2007) and to *reduced basis methods* (Prud'homme and Patera 2004, Rozza, Huynh and Patera 2008, Hesthaven, Rozza and Stamm 2016); for a recent overview, see Benner, Cohen, Ohlberger and Willcox (2017). In reduced basis methods, case (A) in Example 5.32 is a typical model problem. With N well-chosen parameter values $y^1, \dots, y^N \in Y$, we compute the so-called solution *snapshots* $v_k = u(\cdot, y^k)$, $k = 1, \dots, N$, and obtain efficient approximate solutions for arbitrary parameters $y \in Y$ by solving the Galerkin projections of (5.39) onto $\operatorname{span}\{v_k\}_{k=1,\dots,N}$. This leads to approximations of the form

$$u(x, y) \approx \sum_{k=1}^{N} v_k(x)\, \phi_k(y),$$

where the functions ϕ_k, $k = 1, \dots, N$, are defined implicitly by Galerkin projection.

5.5.3. *Tensorized approximations for boundary value problems*

We briefly comment on second-order elliptic boundary value problems on domains $\Omega \subset \mathbb{R}^d$ with $d \in \{1, 2, 3\}$. The efficient approximation of solutions of such problems can require specialized methods, such as hp-finite elements for problems with gradient singularities caused by non-smooth problem data or domains, or homogenization approaches for problems with high-frequency oscillations. The technique of multilevel tensorized approximations, as discussed in Sections 3.4 and 4.5, can be used to obtain generic approximations for such problems by providing efficient parametrizations of simple low-order methods (such as piecewise multilinear finite elements) on extremely fine grids. For diffusion problems parametrized in tensorized low-rank form as outlined in Section 4.5, this is considered by Kazeev (2015), Kazeev and Schwab (2018) and Marcati, Rakhuba and Schwab (2022*a*). Similarly to convergence results for hp-FEM, they establish exponential-type convergence of approximate solutions with respect to the total number of representation parameters. Comparably efficient approximations are obtained for multiscale problems in Kazeev, Oseledets, Rakhuba and Schwab (2017, 2022) and for convection–diffusion problems on intervals in Marcati, Rakhuba and Ulander (2022*b*). The technique is extended to isogeometric analysis in Markeeva, Tsybulin and Oseledets (2021).

Combined with preconditioning as described in Section 4.5, low-rank parametrizations of approximations on very fine grids (corresponding to grid spacings near machine precision) can be used, so that adaptive refinement of discretizations typically becomes unnecessary in this context. Approximate solutions can be obtained by the methods discussed in Section 5.1 with preconditioning or by the more problem-specific alternating direction implicit method of Rakhuba (2021).

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

6. Eigenvalue problems

Solvers based on low-rank approximations for eigenvalue problems are especially important in quantum physics applications, where they are applied both in the classical formulation in terms of particle coordinates, as in (3.25), and in the occupation number formulation discussed in Section 3.3. In the latter case, low-rank structures in wavefunction approximations are closely connected to measures of entanglement in quantum systems (Orús 2014). A physical motivation for the application of low-rank methods in this context is that in view of so-called *area laws* (Eisert, Cramer and Plenio 2010), the most relevant states can be expected to be those of relatively low ranks in tensor representations.

Many of the considerations of the preceding sections for elliptic partial differential equations apply equally to associated eigenvalue problems. The corresponding convergence theory, however, is not as well developed in the latter case, especially concerning error estimation and adaptive refinement of discretizations.

6.1. Applications in quantum physics

A high-dimensional eigenvalue problem that is of central importance in quantum chemistry is the electronic Schrödinger equation. The corresponding Hamiltonian operator is of the general form (3.25). Writing $x_1, \dots, x_N \in \mathbb{R}^3$ for the spatial coordinates of the N electrons in the system under consideration, for a molecule composed of K atomic nuclei with positions $R_k \in \mathbb{R}^3$ and charges $Z_k > 0$ for $k = 1, \dots, K$, this Hamiltonian reads

$$H = -\frac{1}{2}\sum_{i=1}^{N} \Delta_{x_i} - \sum_{i=1}^{N}\sum_{k=1}^{K} \frac{Z_k}{\|x_i - R_k\|_2} + \frac{1}{2}\sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{1}{\|x_i - x_j\|_2}.$$

For an in-depth treatment of the analysis of electronic Schrödinger eigenvalue problems, we refer to Yserentant (2010). Here, the eigenfunctions (called wavefunctions) of interest are typically those corresponding to the lowest eigenvalues, that is, to the lowest energies. By standard techniques, approximation errors in eigenvalues can be estimated in terms of H^1-errors in the corresponding eigenfunction approximations.

As noted in Section 3.3, fermionic wavefunctions need to be antisymmetric under exchange of electron coordinates. Although the direct application of the hierarchical tensor format is problematic due to this requirement, direct approximations of wavefunctions by antisymmetrized canonical tensor representations are considered in Beylkin, Mohlenkamp and Pérez (2008) and Mohlenkamp (2010). However, with this approach we face the difficulties outlined in Section 1.4.

There exist a variety of approximations of the electronic Schrödinger equation by simplified models. The *Hartree–Fock* (HF) equations are derived by forming the Rayleigh quotient of H over a single N-electron Slater determinant and minimizing over the corresponding N orbitals $\phi_1, \dots, \phi_N$, which are functions of a single

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

electron coordinate. Constraining the orbitals $\phi_1, \dots, \phi_N$ to be orthonormal, the necessary optimality conditions lead to the HF equations

$$F(\phi_1, \dots, \phi_N)\phi_i = \mu_i \phi_i, \quad i = 1, \dots, N, \tag{6.1}$$

with the *Fock operator* $F(\phi_1, \dots, \phi_N)$, which is an integrodifferential operator acting on functions on $\mathbb{R}^3$.

While standard methods for (6.1) are based on approximations by a small number of atom-centred Gaussian functions (Helgaker *et al.* 2000, Chap. 8), numerical methods that can reach higher accuracies are frequently of interest. In this case grid-based methods for the HF equations (and related problems in density functional theory) that approximate the orbitals as low-rank third-order tensors are proposed in Khoromskij, Khoromskaia, Chinnamsetty and Flad (2009), Khoromskij, Khoromskaia and Flad (2011) and Rakhuba and Oseledets (2016). In Bischoff and Valeev (2011), approximations in low-rank tensor format of an integral equation formulation of (6.1) are combined with locally refined discontinuous wavelets expansions for one-dimensional components; these wavelets, however, are not sufficiently regular for a rigorous error control of computed energies. In Bachmayr (2012*a*), adaptively refined higher-order Daubechies wavelet expansions of one- and two-electron wavefunctions are approximated in Tucker format.

These approaches rely on explicit low-rank approximations of potential terms, which can be obtained from exponential sum approximations as considered in Section 4.1. In particular, applying approximations of $t \mapsto t^{-1/2}$ for $t > 0$ provided by Theorems 4.2 or Corollary 4.5 leads to approximations of Coulomb potentials by Gaussians,

$$\frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \approx \sum_{k=1}^{r} \omega_k \prod_{i=1}^{3} e^{-\alpha_k x_i^2}.$$

Such approximations are considered, for instance, in Chinnamsetty *et al.* (2007), Zeiser (2010), Bachmayr (2012*a*,*b*) and Scholz and Yserentant (2017). Related issues include the efficient evaluation of two-electron potentials (Khoromskaia, Khoromskij and Schneider 2013) and tensor-based summation of potentials on lattices (Khoromskaia and Khoromskij 2016). For an overview, see also Khoromskaia and Khoromskij (2015, 2018).

A variety of refined models for the electronic Schrödinger equations take HF approximations as their starting points. Instances of such so-called *post-Hartree–Fock* methods are the *Møller–Plesset perturbation theory* and the *coupled cluster* method. The use of low-rank tensor approximations in the context of such methods is investigated in Benedikt, Auer, Espig and Hackbusch (2011), Benedikt, Böhm and Auer (2013*b*) and Benedikt *et al.* (2013*a*).

The concepts of matrix product states (tensor trains) and tree tensor networks (hierarchical tensors), as well as the DMRG algorithm outlined in Section 5.1, have their origins in the study of *spin systems*. These are quantum systems of particles

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

on lattices with locally interacting spins, where N-particle systems typically lead to wavefunctions on $\{0, 1\}^N$; concerning such applications, see Schollwöck (2011) and Orús (2014).

Starting with White and Martin (1999), such methods have also been successfully applied to the occupation number representation of the electronic Schrödinger eigenvalue problem as discussed in Section 3.3. For an overview, we refer to Wouters and Van Neck (2014). Such approaches based on tree tensor networks, which can be regarded as *multi-configurational self-consistent field methods*, are of particular interest in cases where the Hartree–Fock equations yield poor approximations. They can also serve as the basis of modified post-HF methods such as the *tailored coupled cluster* method (Faulstich *et al.* 2019).

Concerning the low-rank approximability of operators in this context, see Remark 3.13. The low-rank approximability of electronic wavefunctions in occupation number representation is investigated in Graswald and Friesecke (2021), Friesecke and Graswald (2022) and Friesecke, Graswald and Legeza (2022). Here, the optimization of the underlying orbitals plays an important role in improving the efficiency of low-rank approximations; see for instance Krumnow *et al.* (2016).

6.2. *Solving eigenvalue problems in low-rank formats*

One strategy for constructing eigensolvers operating on low-rank representations is the adaptation of standard methods by additional rank truncations. Power methods and preconditioned inverse iteration (PINVIT) for discretizations of differential operators are considered by Hackbusch, Khoromskij, Sauter and Tyrtyshnikov (2012). Kressner and Tobler (2011*b*) proposed an LOBPCG eigensolver modified by additional rank truncations and tested it numerically. Cancès *et al.* (2014) considered the application of the greedy approach outlined in Section 5.1.4. Rakhuba and Oseledets (2018) proposed a Jacobi–Davidson method adapted to low-rank matrix manifolds. Rakhuba, Novikov and Oseledets (2019) developed a Riemannian version of LOBPCG for higher-order tensors. Concerning the combination of low-rank approximations with adaptivity, Bachmayr (2012*a*) have shown that a low-rank eigensolver based on PINVIT with adaptive wavelet discretization has guaranteed local convergence in H^1-norm.

A second construction principle for eigensolvers is given by methods based on alternating optimization. A prominent example is the DMRG algorithm introduced in White (1992, 2005) and formulated in terms of matrix product states in Vidal (2003). These methods are investigated from a numerical perspective in Holtz *et al.* (2012*a*). The adaptation of such methods to the simultaneous approximation of several eigenspaces by attaching eigenfunction indices as additional tensor modes is considered in Kressner, Steinlechner and Uschmajew (2014*a*) and Dolgov, Khoromskij, Oseledets and Savostyanov (2014). Krumnow, Pfeffer and Uschmajew (2021) propose direct energy minimization instead of SVD for the rank truncation in DMRG substeps.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

7. Time-dependent problems

In Section 1.2 we discussed two examples of time-dependent problems of potentially high spatial dimensionality: the parabolic problem (1.1) and the instationary Schrödinger equation (1.2). To obtain approximate solutions in low-rank tensor formats, one strategy is to apply standard time-stepping procedures. For implicit methods, as required for parabolic problems, the problems on the spatial domain that need to be solved in each time step can be treated by low-rank solvers for elliptic problems, as in Section 5. Such approaches are pursued, for instance, in Dolgov, Khoromskij and Oseledets (2012) and Cho, Venturi and Karniadakis (2016).

However, it is a delicate matter to adjust low-rank approximation tolerances in each time step to achieve a desired total error over the given time interval, and no informative bounds on the arising ranks are available. Moreover, in certain problems we are also interested in preserving conservation properties of the continuous problem, which is generally difficult to reconcile with rank truncations in every time step. These limitations are addressed by alternative methods for low-rank approximation of time-dependent problems, on which we focus in this section.

7.1. Dynamical low-rank approximation

In dynamical low-rank approximations, evolution problems are approximated on manifolds of fixed-rank matrices or tensors, which leads to a reduction to lower-dimensional evolution problems for the individual components of low-rank representations.

For matrices of fixed rank, this concept is introduced in Koch and Lubich (2007*a*). Assume that we are given a differential equation for a time-dependent matrix $\mathbf{Y}(t) \in \mathbb{R}^{m\times n}$,

$$\dot{\mathbf{Y}}(t) = F(\mathbf{Y}(t)), \quad \mathbf{Y}(0) = \mathbf{Y}_0,$$

where $F\colon \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ is a Lipschitz continuous function. Discretizations of instationary PDEs on two-dimensional spatial domains lead naturally to matrix differential equations of this structure.

Let $\mathcal{M}$ be the submanifold of matrices of fixed rank $r \in \mathbb{N}$ (see also Sections 2.10 and 5.1.3). We now approximate $\mathbf{Y}(t)$ by $\mathbf{X}(t) \in \mathcal{M}$ defined by the *Dirac–Frenkel variational principle* (Dirac 1930, Frenkel 1934):

$$\langle \dot{\mathbf{X}}(t) - F(\mathbf{X}(t)), \Xi \rangle = 0 \quad \text{for all } \Xi \in \mathcal{T}_{\mathbf{X}(t)}\mathcal{M}.$$

Since $\dot{\mathbf{X}}(t) \in \mathcal{T}_{\mathbf{X}(t)}\mathcal{M}$, we thus solve the modified problem

$$\dot{\mathbf{X}}(t) = \Pi_{\mathbf{X}(t)} F(\mathbf{X}(t)), \quad \mathbf{X}(0) = \mathbf{X}_0, \tag{7.1}$$

where $\mathbf{X}_0$ is the best rank-r approximation of $\mathbf{Y}_0$ and where again $\Pi_{\mathbf{X}(t)}$ denotes the orthogonal projection onto $\mathcal{T}_{\mathbf{X}(t)}\mathcal{M}$.

With the representation (2.9) in the form $\mathbf{X}(t) = \mathbf{U}(t)\mathbf{S}(t)\mathbf{V}(t)^\top$, where $\mathbf{U}, \mathbf{V}$ are required to have orthonormal columns for all times and $\mathbf{S} \in \mathbb{R}^{r\times r}$ is invertible, as

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

shown in Koch and Lubich (2007*a*), (7.1) can be rewritten componentwise as

$$\dot{\mathbf{S}} = \mathbf{U}^\top F(\mathbf{X})\mathbf{V}, \tag{7.2a}$$

$$\dot{\mathbf{U}} = (\mathbf{I} - \mathbf{U}\mathbf{U}^\top)F(\mathbf{X})\mathbf{V}\mathbf{S}^{-1}, \tag{7.2b}$$

$$\dot{\mathbf{V}} = (\mathbf{I} - \mathbf{V}\mathbf{V}^\top)F(\mathbf{X})\mathbf{U}\mathbf{S}^{-\top}. \tag{7.2c}$$

Provided that F has suitable structure, the original problem is thus reduced to differential equations for lower-dimensional components. An analogous scheme for Tucker tensors is considered in Koch and Lubich (2010). Similarly to (7.2), we obtain separate but coupled evolution equations for the core tensor and for the orthonormal mode frames. The method can also be formulated on function spaces, with a natural generalization to Hilbert spaces. A formulation on Banach spaces is given in Falcó *et al.* (2019).

In the context of quantum dynamics described by instationary Schrödinger equations (1.2), the method for Tucker tensors was introduced as the *multi-configuration time-dependent Hartree* (MCTDH) method by Meyer, Manthe and Cederbaum (1990); for an overview of such models and related methods, we refer to Lubich (2008) and Meyer, Gatti and Worth (2009). Applied to (1.2), approximations are obtained by solving

$$\mathrm{i}\partial_t u = -\Delta u + \Pi_u V u$$

with appropriate low-rank approximations of initial data, where Π_u denotes the projection onto the tangent space at u of the manifold of fixed-rank tensors. Note that the Laplacian Δ maps elements of the manifold to the tangent space, and thus no projection is required on this term.[1]

MCTDH is applied to models for the evolution for distinguishable particles such as atomic nuclei, where we do not have the symmetry or antisymmetry requirements mentioned in Section 3.3. However, similar approaches can be considered for electron dynamics, where the antisymmetry requirement applies, which leads to the *multi-configuration time-dependent Hartree–Fock* method; see Bardos, Catto, Mauser and Trabelsi (2009, 2010) and Koch and Lubich (2011).

Applying the above approach to hierarchical tensor approximations leads to the *multilayer MCTDH* method (Wang and Thoss 2003). In a general context, dynamical low-rank approximations for tensor trains and hierarchical tensors are analysed in Lubich, Rohwedder, Schneider and Vandereycken (2013), Uschmajew and Vandereycken (2013) and Arnold and Jahnke (2014). The application to the time evolution of matrix product states in spin systems is investigated in Haegeman, Osborne and Verstraete (2013). Projection methods for time integration of such formulations for matrices and tensor trains are analysed in Kieri and Vandereycken (2019).

[1] This corresponds to the observation that for operators of the form (1.6), the solution (1.7) is exactly reproduced by dynamical low-rank approximation.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

For solutions u of time-dependent PDEs with an additional dependence on a parameter y (which may be a random variable), approximations of the low-rank form

$$u(t,x,y) \approx \bar{u}(t,x) + \sum_{k=1}^{r} v_k(t,x) w_k(t,y), \tag{7.3}$$

where $\{v_k(t,\cdot)\}_{k=1,\dots,r}$ are L_2-orthonormal on the spatial domain for each t, can be obtained by a variation of (7.2) termed *dynamically orthogonal approximation* that is introduced in Sapsis and Lermusiaux (2009). This approach is analysed in Musharbash, Nobile and Zhou (2015) with a particular focus on the treatment of parabolic PDEs with random parameters, and also concerning its close relation to dynamical low-rank approximation (7.2). A general analysis is given in Feppon and Lermusiaux (2018, 2019), including a characterization of dynamical low-rank approximation as a limit of alternating time steps and low-rank truncations. Further applications to Navier–Stokes problems and wave equations are considered in Musharbash and Nobile (2018) and Musharbash, Nobile and Vidličková (2020).

Remark 7.1. In the case of finite-dimensional matrices as in (7.2), local-in-time existence of dynamical low-rank approximations follows by the Picard–Lindelöf theorem. In the case of infinite-dimensional function spaces, the situation is more delicate. Existence, uniqueness and regularity of the MCTDH equations for sufficiently regular initial data are shown in Koch and Lubich (2007*b*). For approximations of parabolic problems depending on a random parameter y as in (7.3), well-posedness in the mild and strong sense is shown in Kazashi and Nobile (2021). In the case of parabolic problems in two dimensions, existence and uniqueness for a weak formulation of dynamical low-rank approximation are demonstrated in Bachmayr, Eisenmann, Kieri and Uschmajew (2021*a*), and robustness results for corresponding spatially discretized problems are obtained in Conte (2020).

Remark 7.2. Let $\mathcal{M}$ be a manifold of rank-one tensors in a Hilbert space H, and consider

$$\dot{u}(t) = f, \quad u(0) = u_0 \in \mathcal{M},$$

with fixed $f \in \mathcal{M}$. Then, whenever f is orthogonal to $\mathcal{T}_{u_0}\mathcal{M}$, the dynamical low-rank approximation equals u_0 for all t. The best rank-one approximation of the exact solution $u(t) = u_0 + tf$, however, will be tf when t is sufficiently large.

Local quasi-optimality properties of MCTDH approximations, compared to best approximations of the same rank, are established in Lubich (2005), and an error analysis of MCTDH is conducted in Conte and Lubich (2010). The assumptions required in the analysis in particular ensure that the above issue is avoided, but these conditions are typically not verifiable in practice.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

7.2. Splitting and Basis Update & Galerkin integrators

The system of differential equations (7.2) can be difficult to solve numerically, especially in the presence of small singular values of $\mathbf{S}$. An integrator for dynamical low-rank approximation with particularly favourable properties can be obtained by a splitting approach. Writing the projector onto the tangent space at $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ as

$$\Pi_{\mathbf{X}}\mathbf{Z} = \mathbf{Z}\mathbf{V}\mathbf{V}^\top - \mathbf{U}\mathbf{U}^\top\mathbf{Z}\mathbf{V}\mathbf{V}^\top + \mathbf{U}\mathbf{U}^\top\mathbf{Z}$$

and decomposing the right-hand side in (7.1) accordingly, one can apply standard Lie–Trotter or Strang splitting to arrive at the *projector-splitting* integrator that is proposed and analysed in Lubich and Oseledets (2014). This integrator is shown to enjoy particularly favourable exactness properties with improved robustness concerning small singular values of $\mathbf{S}$.

Integrators based on the same principles are introduced for MCTDH in Lubich (2015) and for general dynamical low-rank approximations of tensor trains in Lubich, Oseledets and Vandereycken (2015). Haegeman *et al.* (2016) established a connection to the DMRG scheme for matrix product states. The robustness properties of the splitting integrators with respect to small singular values are investigated further in Kieri, Lubich and Walach (2016) and Ostermann, Piazzola and Walach (2019). Einkemmer and Lubich (2018, 2019) constructed projector-splitting integrators for Vlasov–Poisson equations. For random parabolic equations, a scheme of this type is analysed in Kazashi, Nobile and Vidličková (2021). Variants including strategies for the adaptation of approximation ranks are proposed in Yang and White (2020), Dektor, Rodgers and Venturi (2021) and Dunnett and Chin (2021).

Integrators for fixed-rank matrices and Tucker tensors with similar structure and robustness properties, but not derived from a splitting of the projector, are introduced and analysed in Ceruti and Lubich (2020, 2022). In this approach, termed *Basis Update & Galerkin*, in the matrix case considered above a time step is applied to the bases given by $\mathbf{U}$ and $\mathbf{V}$, and $\mathbf{S}$ is evolved by a Galerkin method using these bases. An integrator that includes rank adaptivity is developed in Ceruti, Kusch and Lubich (2022) and extended to tree tensor networks in Ceruti *et al.* (2023); it joins the basis vectors from current and previous time steps in each step for $\mathbf{S}$ followed by SVD-based rank truncation.

7.3. Space–time methods

Solving time-dependent problems in formulations on the full space–time domain leads to approximation problems with an additional dimension, compared to the spatial approximations required in time-stepping schemes. Such methods are considered in the literature, in particular for parabolic problems. A classical model problem is the heat equation on the d-dimensional unit cube with homogeneous Dirichlet boundary conditions,

$$\partial_t u - \Delta u = f \quad \text{in } (0,T)\times\square_d, \quad u|_{t=0} = u_0,\ u|_{(0,T)\times\partial\square_d} = 0, \tag{7.4}$$

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

as a special case of (1.1).

An *algebraic* construction of space–time solvers is obtained in Dolgov *et al.* (2012), where the entire coupled system resulting from a standard time-stepping discretization is solved jointly for all time steps in a multilevel tensorized representation. Although approximability results of solutions are obtained for certain cases, robustness with respect to discretization refinement and spatial dimensionality remain unclear in this approach. A related method based on certain series expansions is considered in Gavrilyuk and Khoromskij (2019). Further heuristic space–time schemes based on Chebyshev interpolation are considered in Sun and Kumar (2014) and Chertkov and Oseledets (2021).

To control discretization errors, methods derived from well-posed space–time weak formulations are of particular interest. Based on corresponding residuals measured in appropriate norms, discretization errors can be estimated jointly in space and time for such formulations. For second-order parabolic problems such as (7.4), a classical weak formulation using the space $\mathcal{W} = L_2(0,T;V) \cap H^1(0,T;V')$ with $V = H^1_0(\square_d)$ reads as follows: find $u \in \mathcal{W}$ such that

$$\begin{aligned} &\int_0^\top \langle \partial_t u, v\rangle_{V'\times V} + \int_{\square_d} \nabla u \cdot \nabla v \,\mathrm{d}x\,\mathrm{d}t + \int_{\square_d} u|_{t=0}\, w \,\mathrm{d}x \\ &\quad = \int_0^\top \langle f, v\rangle_{V'\times V} \,\mathrm{d}t + \int_{\square_d} u_0 w \,\mathrm{d}x \end{aligned} \tag{7.5}$$

for all $v \in L_2(0,T;V)$ and $w \in L_2(\square_d)$. The left-hand side of (7.5) can be shown to define an isomorphism from $\mathcal{W}$ to $L_2(0,T;V')\times L_2(\square_d)$; see Schwab and Stevenson (2009).

Remark 7.3. Schwab and Stevenson (2009) have constructed an adaptive wavelet method based on a Riesz basis representation of (7.5). Applying such a Riesz basis representation in the context of low-rank approximations leads to additional difficulties. To see these, assume that ψ_ν, $\nu \in \vee$, are as in Remark 3.10, and that in addition, $\omega_\mu \in H^1(0,T)$ are such that $\{\omega_\mu\}_{\mu\in\hat\vee}$ and $\{\|\omega_\mu\|^{-1}_{H^1}\omega_\mu\}_{\mu\in\hat\vee}$ are Riesz bases of $L_2(0,T)$ and $H^1(0,1)$, respectively. Then a Riesz basis for the test space $L_2(0,T;V) \simeq L_2(0,T)\otimes V$ is given by $\{\omega_\mu \otimes \Psi_\nu\}_{(\mu,\nu)\in\hat\vee\times\vee^d}$ with Ψ_ν, $\nu \in \vee^d$, as in Remark 3.10. This can be handled as described in Section 4.4. For the space $\mathcal{W}$, however, Riesz bases are given by $\{d_{\mu,\nu}\,\omega_\mu \otimes \psi_{\nu_1} \otimes \cdots \otimes \psi_{\nu_d}\}_{(\mu,\nu)\in\hat\vee\times\vee^d}$, with the requirement

$$d_{\mu,\nu_1,\dots,\nu_d} \eqsim \left(\sum_{i=1}^d 2^{2|\nu_i|} + 2^{2|\mu|}\left(\sum_{i=1}^d 2^{2|\nu_i|}\right)^{-1}\right)^{-1/2}, \quad \mu \in \hat\vee, \nu \in \vee^d \tag{7.6}$$

(compare (3.24)). Low-rank approximations of the expressions on the right in (7.6) that are suitable for diagonal preconditioning in low-rank formats are substantially more difficult to construct than in the case of H^1-norms considered in Section 4.4.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Andreev and Tobler (2015) have solved a fixed discretization of (7.5) by multilinear finite elements in space and time in least-squares form with low-rank approximations in hierarchical format, where time and each spatial variable are treated as separate tensor modes. The formulation is based on the stable Petrov–Galerkin discretization by Andreev (2013), which relies on a careful adjustment of trial and test spaces. The appropriate norms are realized by a BPX preconditioner, which – as in the case of Riesz bases – reduces preconditioning to diagonal scaling. However, this also requires scaling by the values in (7.6), which is realized heuristically in Andreev and Tobler (2015).

A space–time method based on a slightly different least-squares formulation is considered in Boiveau, Ehrlacher, Ern and Nouy (2019). In this case the analysis is based on a semidiscretization in the spatial variables. The low-rank approximation uses tensors of order two, separating time from the spatial variables; in this form, this method is thus not aimed at high-dimensional problems. The approximations are constructed by a greedy method as outlined in Section 5.1.4.

A different approach combining sparse wavelet approximations in time with adaptive low-rank approximations in space is used in Bachmayr and Faldum (2023). To this end, (7.5) is represented in a Riesz basis representation as in Remark 7.6 in terms of spline wavelets, using $\{\tilde{\Psi}_{\mu,\nu}\}_{(\mu,\nu)\in\hat{\vee}\times\vee^d}$ with $\tilde{\Psi}_{\mu,\nu} = d_{\mu,\nu}\,\omega_\mu \otimes \psi_{\nu_1} \otimes \cdots \otimes \psi_{\nu_d}$, and $d_{\mu,\nu}$ satisfying (7.6), as a Riesz basis of $\mathcal{W}$. The solution u of (7.5) then has the expansion

$$u = \sum_{\mu\in\hat{\vee}} \sum_{\nu\in\vee^d} \mathbf{u}[\mu,\nu_1,\dots,\nu_d]\,\tilde{\Psi}_{\mu,\nu}.$$

With a fixed dimension tree $\mathbb{T}$ and a suitable index set $\hat{\Lambda} \subset \hat{\vee}$ defining a temporal discretization, we now aim to find *separately for each* $\mu \in \hat{\Lambda}$ hierarchical tensor approximations $\tilde{\mathbf{u}}_\mu$ of the dth-order tensors

$$\mathbf{u}_\mu = (\mathbf{u}[\mu,\nu_1,\dots,\nu_d])_{\nu_1,\dots,\nu_d\in\vee}.$$

These approximations are chosen with $\operatorname{supp}\tilde{\mathbf{u}}_\mu \subseteq \Lambda_\mu = \bigtimes_{i=1}^d \Lambda_\mu^{(i)} \subset \vee^d$, where the finite index sets Λ_μ also need to be determined.

As shown in Bachmayr and Faldum (2023), the basic techniques described in Section 5.3 can be modified for computing such families $(\tilde{\mathbf{u}}_\mu)_{\mu\in\hat{\Lambda}}$ of hierarchical tensor approximations with adaptively selected discretizations. This approach requires efficient low-rank approximations of (7.6) only separately for each $\mu \in \hat{\vee}$, which can be achieved by exponential sum approximations similar to those considered in Section 4.1.1. The resulting method yields guaranteed convergence to u with computable error bounds in the $\mathcal{W}$-norm, as well as bounds on the arising ranks, discretization sizes and total computational costs analogous to those for the elliptic case discussed in Section 5.4. The method is demonstrated to be applicable to large spatial dimensions d in practice.

https://doi.org/10.1017/S0962492922000125 Published online by Cambridge University Press

Acknowledgements

The author thanks Wolfgang Dahmen, Henrik Eisenmann, Manfred Faldum, Vladimir Kazeev, Sebastian Krämer, Christian Lubich, Reinhold Schneider and André Uschmajew for reading earlier versions of this article and for important hints and suggestions for improvements.

The author acknowledges funding by Deutsche Forschungsgemeinschaft (German Research Foundation), project number 442047500, through the Collaborative Research Center 'Sparsity and Singular Structures' (SFB 1481).

References

P.-A. Absil and I. V. Oseledets (2015), Low-rank retractions: A survey and new results, *Comput. Optim. Appl.* **62**, 5–29.

P.-A. Absil, R. Mahony and R. Sepulchre (2008), *Optimization Algorithms on Matrix Manifolds*, Princeton University Press.

M. Ali and A. Nouy (2020*a*), Singular value decomposition in Sobolev spaces, Part I, *Z. Anal. Anwend.* **39**, 349–369.

M. Ali and A. Nouy (2020*b*), Singular value decomposition in Sobolev spaces, Part II, *Z. Anal. Anwend.* **39**, 371–394.

M. Ali and A. Nouy (2021), Approximation theory of tree tensor networks: Tensorized multivariate functions. Available at arXiv:2101.11932.

M. Ali and A. Nouy (2023), Approximation theory of tree tensor networks: Tensorized univariate functions, *Constr. Approx.* Available at doi:10.1007/s00365-023-09620-w.

M. Ali and K. Urban (2020), HT-AWGM: A hierarchical Tucker-adaptive wavelet Galerkin method for high-dimensional elliptic problems, *Adv. Comput. Math.* **46**, 59.

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

R. Andreev (2013), Stability of sparse space–time finite element discretizations of linear parabolic evolution equations, *IMA J. Numer. Anal.* **33**, 242–260.

R. Andreev and C. Tobler (2015), Multilevel preconditioning and low-rank tensor iteration for space–time simultaneous discretizations of parabolic PDEs, *Numer. Linear Algebra Appl.* **22**, 317–337.

A. Arnold and T. Jahnke (2014), On the approximation of high-dimensional differential equations in the hierarchical Tucker format, *BIT Numer. Math.* **54**, 305–341.

M. Bachmayr (2012*a*), Adaptive low-rank wavelet methods and applications to two-electron Schrödinger equations. PhD thesis, RWTH Aachen.

M. Bachmayr (2012*b*), Hyperbolic wavelet discretization of the two-electron Schrödinger equation in an explicitly correlated formulation, *ESAIM Math. Model. Numer. Anal.* **46**, 1337–1362.

M. Bachmayr and A. Cohen (2017), Kolmogorov widths and low-rank approximations of parametric elliptic PDEs, *Math. Comp.* **86**, 701–724.

M. Bachmayr and W. Dahmen (2015), Adaptive near-optimal rank tensor approximation for high-dimensional operator equations, *Found. Comput. Math.* **15**, 839–898.

M. Bachmayr and W. Dahmen (2016*a*), Adaptive low-rank methods for problems on Sobolev spaces with error control in L_2, *ESAIM Math. Model. Numer. Anal.* **50**, 1107–1136.

M. Bachmayr and W. Dahmen (2016*b*), Adaptive low-rank methods: Problems on Sobolev spaces, *SIAM J. Numer. Anal.* **54**, 744–796.
M. Bachmayr and W. Dahmen (2020), Adaptive low-rank approximations for operator equations: Accuracy control and computational complexity, in *75 Years of Mathematics of Computation* (S. C. Brenner *et al.*, eds), Vol. 754 of Contemporary Mathematics, American Mathematical Society, pp. 1–44.
M. Bachmayr and M. Faldum (2023), A space-time adaptive low-rank method for high-dimensional parabolic partial differential equations. Available at arXiv:2302.01658.
M. Bachmayr and V. Kazeev (2020), Stability of low-rank tensor representations and structured multilevel preconditioning for elliptic PDEs, *Found. Comput. Math.* **20**, 1175–1236.
M. Bachmayr and R. Schneider (2017), Iterative methods based on soft thresholding of hierarchical tensors, *Found. Comput. Math.* **17**, 1037–1083.
M. Bachmayr, A. Cohen and W. Dahmen (2018), Parametric PDEs: Sparse or low-rank approximations?, *IMA J. Numer. Anal.* **38**, 1661–1708.
M. Bachmayr, A. Cohen, D. Dũng and C. Schwab (2017), Fully discrete approximation of parametric and stochastic elliptic PDEs, *SIAM J. Numer. Anal.* **55**, 2151–2186.
M. Bachmayr, H. Eisenmann, E. Kieri and A. Uschmajew (2021*a*), Existence of dynamical low-rank approximations to parabolic problems, *Math. Comp.* **90**, 1799–1830.
M. Bachmayr, M. Götte and M. Pfeffer (2022), Particle number conservation and block structures in matrix product states, *Calcolo* **59**, 24.
M. Bachmayr, A. Nouy and R. Schneider (2021*b*), Approximation by tree tensor networks in high dimensions: Sobolev and compositional functions. Available at arXiv:2112.01474. To appear in *Pure Appl. Funct. Anal.*
M. Bachmayr, R. Schneider and A. Uschmajew (2016), Tensor networks and hierarchical tensors for the solution of high-dimensional partial differential equations, *Found. Comput. Math.* **16**, 1423–1472.
J. Ballani and L. Grasedyck (2013), A projection method to solve linear systems in tensor format, *Numer. Linear Algebra Appl.* **20**, 27–43.
J. Ballani and L. Grasedyck (2015), Hierarchical tensor approximation of output quantities of parameter-dependent PDEs, *SIAM/ASA J. Uncertain. Quantif.* **3**, 852–872.
J. Ballani, L. Grasedyck and M. Kluge (2013), Black box approximation of tensors in hierarchical Tucker format, *Linear Algebra Appl.* **438**, 639–657.
C. Bardos, I. Catto, N. Mauser and S. Trabelsi (2010), Setting and analysis of the multi-configuration time-dependent Hartree–Fock equations, *Arch. Ration. Mech. Anal.* **198**, 273–330.
C. Bardos, I. Catto, N. J. Mauser and S. Trabelsi (2009), Global-in-time existence of solutions to the multiconfiguration time-dependent Hartree–Fock equations: A sufficient condition, *Appl. Math. Lett.* **22**, 147–152.
U. Benedikt, A. A. Auer, M. Espig and W. Hackbusch (2011), Tensor decomposition in post-Hartree–Fock methods, I: Two-electron integrals and MP2, *J. Chem. Phys.* **134**, 054118.
U. Benedikt, H. Auer, M. Espig, W. Hackbusch and A. A. Auer (2013*a*), Tensor representation techniques in post-Hartree–Fock methods: Matrix product state tensor format, *Molecular Phys.* **111**, 2398–2413.
U. Benedikt, K.-H. Böhm and A. A. Auer (2013*b*), Tensor decomposition in post-Hartree–Fock methods, II: CCD implementation, *J. Chem. Phys.* **139**, 224101.

P. Benner, A. Cohen, M. Ohlberger and K. Willcox, eds (2017), *Model Reduction and Approximation: Theory and Algorithms*, Vol. 15 of Computational Science & Engineering, SIAM.

G. Beylkin and M. J. Mohlenkamp (2002), Numerical operator calculus in higher dimensions, *Proc. Natl. Acad. Sci. USA* **99**, 10246–10251.

G. Beylkin and M. J. Mohlenkamp (2005), Algorithms for numerical analysis in high dimensions, *SIAM J. Sci. Comput.* **26**, 2133–2159.

G. Beylkin and L. Monzón (2005), On approximation of functions by exponential sums, *Appl. Comput. Harmon. Anal.* **19**, 17–48.

G. Beylkin and L. Monzón (2010), Approximation by exponential sums revisited, *Appl. Comput. Harmon. Anal.* **28**, 131–149.

G. Beylkin, M. J. Mohlenkamp and F. Pérez (2008), Approximating a wavefunction as an unconstrained sum of Slater determinants, *J. Math. Phys.* **49**, 032107.

D. Bigoni, A. P. Engsig-Karup and Y. M. Marzouk (2016), Spectral tensor-train decomposition, *SIAM J. Sci. Comput.* **38**, A2405–A2439.

M. Billaud-Friess, A. Nouy and O. Zahm (2014), A tensor approximation method based on ideal minimal residual formulations for the solution of high-dimensional problems, *ESAIM Math. Model. Numer. Anal.* **48**, 1777–1806.

F. A. Bischoff and E. F. Valeev (2011), Low-order tensor approximations for electronic wave functions: Hartree–Fock method with guaranteed precision, *J. Chem. Phys.* **134**, 104104.

T. Boiveau, V. Ehrlacher, A. Ern and A. Nouy (2019), Low-rank approximation of linear parabolic equations by space–time tensor Galerkin methods, *ESAIM Math. Model. Numer. Anal.* **53**, 635–658.

D. Braess (1986), *Nonlinear Approximation Theory*, Vol. 7 of Springer Series in Computational Mathematics, Springer.

D. Braess and W. Hackbusch (2005), Approximation of $1/x$ by exponential sums in $[1, \infty)$, *IMA J. Numer. Anal.* **25**, 685–697.

D. Braess and W. Hackbusch (2009), On the efficient computation of high-dimensional integrals and the approximation by exponential sums, in *Multiscale, Nonlinear and Adaptive Approximation* (R. DeVore and A. Kunoth, eds), Springer, pp. 39–74.

J. H. Bramble, J. E. Pasciak and J. Xu (1990), Parallel multilevel preconditioners, *Math. Comp.* **55**, 1–22.

J.-F. Cai, E. J. Candès and Z. Shen (2010), A singular value thresholding algorithm for matrix completion, *SIAM J. Optim.* **20**, 1956–1982.

E. Cancès, V. Ehrlacher and T. Lelièvre (2011), Convergence of a greedy algorithm for high-dimensional convex nonlinear problems, *Math. Models Methods Appl. Sci.* **21**, 2433–2467.

E. Cancès, V. Ehrlacher and T. Lelièvre (2013), Greedy algorithms for high-dimensional non-symmetric linear problems, *ESAIM Proc.* **41**, 95–131.

E. Cancès, V. Ehrlacher and T. Lelièvre (2014), Greedy algorithms for high-dimensional eigenvalue problems, *Constr. Approx.* **40**, 387–423.

G. Ceruti and C. Lubich (2020), Time integration of symmetric and anti-symmetric low-rank matrices and Tucker tensors, *BIT Numer. Math.* **60**, 591–614.

G. Ceruti and C. Lubich (2022), An unconventional robust integrator for dynamical low-rank approximation, *BIT Numer. Math.* **62**, 23–44.

G. Ceruti, J. Kusch and C. Lubich (2022), A rank-adaptive robust integrator for dynamical low-rank approximation, *BIT Numer. Math.* **62**, 1149–1174.

G. Ceruti, C. Lubich and D. Sulz (2023), Rank-adaptive time integration of tree tensor networks, *SIAM J. Numer. Anal.* **61**, 194–222.

Z. Chen, K. Batselier, J. A. K. Suykens and N. Wong (2017), Parallelized tensor train learning of polynomial classifiers, *IEEE Trans. Neural Netw. Learn. Syst.* **29**, 4621–4632.

A. Chertkov and I. V. Oseledets (2021), Solution of the Fokker–Planck equation by cross approximation method in the tensor train format, *Front. Artif. Intell.* **4**, 668215.

S. R. Chinnamsetty, M. Espig, B. N. Khoromskij, W. Hackbusch and H.-J. Flad (2007), Tensor product approximation with optimal rank in quantum chemistry, *J. Chem. Phys.* **127**, 084110.

H. Cho, D. Venturi and G. E. Karniadakis (2016), Numerical methods for high-dimensional probability density function equations, *J. Comput. Phys.* **305**, 817–837.

A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa and H. A. Phan (2015), Tensor decompositions for signal processing applications: From two-way to multiway component analysis, *IEEE Signal Process. Mag.* **32**, 145–163.

A. Cohen (2003), *Numerical Analysis of Wavelet Methods*, Vol. 32 of Studies in Mathematics and its Applications, North-Holland.

A. Cohen and R. DeVore (2015), Approximation of high-dimensional parametric PDEs, *Acta Numer.* **24**, 1–159.

A. Cohen, W. Dahmen and R. DeVore (2001), Adaptive wavelet methods for elliptic operator equations: Convergence rates, *Math. Comp.* **70**, 27–75.

A. Cohen, W. Dahmen and R. DeVore (2002), Adaptive wavelet methods, II: Beyond the elliptic case, *Found. Comput. Math.* **2**, 203–245.

A. Cohen, I. Daubechies and J.-C. Feauveau (1992), Biorthogonal bases of compactly supported wavelets, *Commun. Pure Appl. Math.* **45**, 485–560.

A. Cohen, R. DeVore and C. Schwab (2010), Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs, *Found. Comput. Math.* **10**, 615–646.

A. Cohen, R. DeVore and C. Schwab (2011), Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's, *Anal. Appl. (Singap.)* **9**, 11–47.

D. Conte (2020), Dynamical low-rank approximation to the solution of parabolic differential equations, *Appl. Numer. Math.* **156**, 377–384.

D. Conte and C. Lubich (2010), An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics, *M2AN Math. Model. Numer. Anal.* **44**, 759–780.

G. M. Crosswhite and D. Bacon (2008), Finite automata for caching in matrix product algorithms, *Phys. Rev. A* **78**, 012356.

C. Da Silva and F. J. Herrmann (2015), Optimization on the hierarchical Tucker manifold: Applications to tensor completion, *Linear Algebra Appl.* **481**, 131–173.

S. Dahlke, M. Fornasier and T. Raasch (2012), Multilevel preconditioning and adaptive sparse solution of inverse problems, *Math. Comp.* **81**, 419–446.

W. Dahmen (1997), Wavelet and multiscale methods for operator equations, *Acta Numer.* **6**, 55–228.

W. Dahmen, R. DeVore, L. Grasedyck and E. Süli (2016), Tensor-sparsity of solutions to high-dimensional elliptic partial differential equations, *Found. Comput. Math.* **16**, 813–874.

W. Dahmen, A. Kunoth and K. Urban (1999), Biorthogonal spline wavelets on the interval: Stability and moment conditions, *Appl. Comput. Harmon. Anal.* **6**, 132–196.
I. Daubechies (1988), Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.* **41**, 909–996.
I. Daubechies (1992), *Ten Lectures on Wavelets*, Vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM.
L. De Lathauwer, B. De Moor and J. Vandewalle (2000), A multilinear singular value decomposition, *SIAM J. Matrix Anal. Appl.* **21**, 1253–1278.
W. De Launey and J. Seberry (1994), The strong Kronecker product, *J. Combin. Theory Ser. A* **66**, 192–213.
V. de Silva and L.-H. Lim (2008), Tensor rank and the ill-posedness of the best low-rank approximation problem, *SIAM J. Matrix Anal. Appl.* **30**, 1084–1127.
A. Defant and K. Floret (1993), *Tensor Norms and Operator Ideals*, Vol. 176 of North-Holland Mathematics Studies, North-Holland.
A. Dektor, A. Rodgers and D. Venturi (2021), Rank-adaptive tensor methods for highdimensional nonlinear pdes, *J. Sci. Comput.* **88**, 36.
R. A. DeVore (1998), Nonlinear approximation, *Acta Numer.* **7**, 51–150.
T. J. Dijkema, C. Schwab and R. Stevenson (2009), An adaptive wavelet method for solving high-dimensional elliptic PDEs, *Constr. Approx.* **30**, 423–455.
P. A. M. Dirac (1930), Note on exchange phenomena in the Thomas atom, *Math. Proc. Camb. Phil. Soc.* **26**, 376–385.
S. V. Dolgov and B. N. Khoromskij (2013), Two-level QTT-Tucker format for optimized tensor calculus, *SIAM J. Matrix Anal. Appl.* **34**, 593–623.
S. V. Dolgov and D. V. Savostyanov (2014), Alternating minimal energy methods for linear systems in higher dimensions, *SIAM J. Sci. Comput.* **36**, A2248–A2271.
S. V. Dolgov and R. Scheichl (2019), A hybrid alternating least squares-TT-cross algorithm for parametric PDEs, *SIAM/ASA J. Uncertain. Quantif.* **7**, 260–291.
S. V. Dolgov and T. Vejchodský (2021), Guaranteed *a posteriori* error bounds for low-rank tensor approximate solutions, *IMA J. Numer. Anal.* **41**, 1240–1266.
S. V. Dolgov, B. N. Khoromskij and I. V. Oseledets (2012), Fast solution of parabolic problems in the tensor train/quantized tensor train format with initial application to the Fokker–Planck equation, *SIAM J. Sci. Comput.* **34**, A3016–A3038.
S. V. Dolgov, B. N. Khoromskij, I. V. Oseledets and D. V. Savostyanov (2014), Computation of extreme eigenvalues in higher dimensions using block tensor train format, *Comput. Phys. Commun.* **185**, 1207–1216.
S. V. Dolgov, D. Kressner and C. Strössner (2021), Functional Tucker approximation using Chebyshev interpolation, *SIAM J. Sci. Comput.* **43**, A2190–A2210.
J. Dölz, H. Egger and M. Schlottbom (2021), A model reduction approach for inverse problems with operator valued data, *Numer. Math.* **148**, 889–917.
G. C. Donovan, J. S. Geronimo and D. P. Hardin (1996), Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets, *SIAM J. Math. Anal.* **27**, 1791–1815.
G. C. Donovan, J. S. Geronimo and D. P. Hardin (1999), Orthogonal polynomials and the construction of piecewise polynomial smooth wavelets, *SIAM J. Math. Anal.* **30**, 1029–1056.
A. J. Dunnett and A. W. Chin (2021), Efficient bond-adaptive approach for finite-temperature open quantum dynamics using the one-site time-dependent variational principle for matrix product states, *Phys. Rev. B* **104**, 214302.

C. Eckart and G. Young (1936), The approximation of one matrix by another of lower rank, *Psychometrika* **1**, 211–218.
M. Eigel, M. Marschall, M. Pfeffer and R. Schneider (2020), Adaptive stochastic Galerkin FEM for lognormal coefficients in hierarchical tensor representations, *Numer. Math.* **145**, 655–692.
M. Eigel, M. Pfeffer and R. Schneider (2017), Adaptive stochastic Galerkin FEM with hierarchical tensor representations, *Numer. Math.* **136**, 765–803.
L. Einkemmer and C. Lubich (2018), A low-rank projector-splitting integrator for the Vlasov–Poisson equation, *SIAM J. Sci. Comput.* **40**, B1330–B1360.
L. Einkemmer and C. Lubich (2019), A quasi-conservative dynamical low-rank algorithm for the Vlasov equation, *SIAM J. Sci. Comput.* **41**, B1061–B1081.
J. Eisert, M. Cramer and M. B. Plenio (2010), Colloquium: Area laws for the entanglement entropy, *Rev. Mod. Phys.* **82**, 277–306.
K. Fackeldey, M. Oster, L. Sallandt and R. Schneider (2022), Approximative policy iteration for exit time feedback control problems driven by stochastic differential equations using tensor train format, *Multiscale Model. Simul.* **20**, 379–403.
A. Falcó and W. Hackbusch (2012), On minimal subspaces in tensor representations, *Found. Comput. Math.* **12**, 765–803.
A. Falcó and A. Nouy (2011), A proper generalized decomposition for the solution of elliptic problems in abstract form by using a functional Eckart–Young approach, *J. Math. Anal. Appl.* **376**, 469–480.
A. Falcó and A. Nouy (2012), Proper generalized decomposition for nonlinear convex problems in tensor Banach spaces, *Numer. Math.* **121**, 503–530.
A. Falcó, W. Hackbusch and A. Nouy (2019), On the Dirac–Frenkel variational principle on tensor Banach spaces, *Found. Comput. Math.* **19**, 159–204.
A. Falcó, W. Hackbusch and A. Nouy (2021), Tree-based tensor formats, *SeMA J.* **78**, 159–173.
F. M. Faulstich, A. Laestadius, O. Legeza, R. Schneider and S. Kvaal (2019), Analysis of the tailored coupled-cluster method in quantum chemistry, *SIAM J. Numer. Anal.* **57**, 2579–2607.
F. Feppon and P. F. J. Lermusiaux (2018), A geometric approach to dynamical model order reduction, *SIAM J. Matrix Anal. Appl.* **39**, 510–538.
F. Feppon and P. F. J. Lermusiaux (2019), The extrinsic geometry of dynamical systems tracking nonlinear matrix projections, *SIAM J. Matrix Anal. Appl.* **40**, 814–844.
J. Frenkel (1934), *Wave Mechanics: Advanced General Theory*, Clarendon Press.
G. Friesecke and B. R. Graswald (2022), Two-electron wavefunctions are matrix product states with bond dimension three, *J. Math. Phys.* **63**, 091901.
G. Friesecke, B. R. Graswald and Ö. Legeza (2022), Exact matrix product state representation and convergence of a fully correlated electronic wavefunction in the infinite-basis limit, *Phys. Rev. B* **105**, 165144.
T. Gantumur, H. Harbrecht and R. Stevenson (2007), An optimal adaptive wavelet method without coarsening of the iterands, *Math. Comp.* **76**, 615–629.
B. Gao and P.-A. Absil (2022), A Riemannian rank-adaptive method for low-rank matrix completion, *Comput. Optim. Appl.* **81**, 67–90.
I. Gavrilyuk and B. N. Khoromskij (2019), Quasi-optimal rank-structured approximation to multidimensional parabolic problems by Cayley transform and Chebyshev interpolation, *Comput. Methods Appl. Math.* **19**, 55–71.

I. P. Gavrilyuk, W. Hackbusch and B. N. Khoromskij (2005), Hierarchical tensor-product approximation to the inverse and related operators for high-dimensional elliptic problems, *Computing* **74**, 131–157.
G. H. Golub and W. Kahan (1965), Calculating the singular values and pseudo-inverse of a matrix, *J. Soc. Indust. Appl. Math. Ser. B Numer. Anal.* **2**, 205–224.
G. H. Golub and C. Reinsch (1970), Handbook Series Linear Algebra: Singular value decomposition and least squares solutions, *Numer. Math.* **14**, 403–420.
A. Gorodetsky, S. Karaman and Y. Marzouk (2019), A continuous analogue of the tensor-train decomposition, *Comput. Methods Appl. Mech. Engrg* **347**, 59–84.
L. Grasedyck (2004), Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure, *Computing* **72**, 247–265.
L. Grasedyck (2009/10), Hierarchical singular value decomposition of tensors, *SIAM J. Matrix Anal. Appl.* **31**, 2029–2054.
L. Grasedyck (2010), Polynomial approximation in hierarchical Tucker format by vector-tensorization. DFG SPP 1324 Preprint 43.
L. Grasedyck and W. Hackbusch (2011), An introduction to hierarchical ($\mathcal{H}$-) rank and TT-rank of tensors with examples, *Comput. Methods Appl. Math.* **11**, 291–304.
L. Grasedyck, D. Kressner and C. Tobler (2013), A literature survey of low-rank tensor approximation techniques, *GAMM-Mitt.* **36**, 53–78.
B. R. Graswald and G. Friesecke (2021), Electronic wavefunction with maximally entangled mps representation, *Europ. Phys. J. D* **75**, 1–4.
M. Griebel and H. Harbrecht (2014), Approximation of bi-variate functions: Singular value decomposition versus sparse grids, *IMA J. Numer. Anal.* **34**, 28–54.
M. Griebel and H. Harbrecht (2019), Singular value decomposition versus sparse grids: Refined complexity estimates, *IMA J. Numer. Anal.* **39**, 1652–1671.
M. Griebel and H. Harbrecht (2023), Analysis of tensor approximation schemes for continuous functions, *Found. Comput. Math.* **23**, 219–240.
M. Griebel, H. Harbrecht and R. Schneider (2022), Low-rank approximation of continuous functions in Sobolev spaces with dominating mixed smoothness. Available at arXiv:2203.04100.
M. Gu and S. C. Eisenstat (1995), A divide-and-conquer algorithm for the bidiagonal SVD, *SIAM J. Matrix Anal. Appl.* **16**, 79–92.
W. Hackbusch (2005), Entwicklungen nach Exponentialsummen. Technical Report 4, MPI MIS Leipzig.
W. Hackbusch (2011), Tensorisation of vectors and their efficient convolution, *Numer. Math.* **119**, 465–488.
W. Hackbusch (2014), Numerical tensor calculus, *Acta Numer.* **23**, 651–742.
W. Hackbusch (2015*a*), *Hierarchical Matrices: Algorithms and Analysis*, Vol. 49 of Springer Series in Computational Mathematics, Springer.
W. Hackbusch (2015*b*), Solution of linear systems in high spatial dimensions, *Comput. Vis. Sci.* **17**, 111–118.
W. Hackbusch (2018), On the representation of symmetric and antisymmetric tensors, in *Contemporary Computational Mathematics: A Celebration of the 80th Birthday of Ian Sloan* (J. Dick *et al.*, eds), Springer, pp. 483–515.
W. Hackbusch (2019), *Tensor Spaces and Numerical Tensor Calculus*, Vol. 56 of Springer Series in Computational Mathematics, second edition, Springer.

W. Hackbusch and B. N. Khoromskij (2006*a*), Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators, I: Separable approximation of multi-variate functions, *Computing* **76**, 177–202.

W. Hackbusch and B. N. Khoromskij (2006*b*), Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators, II: HKT representation of certain operators, *Computing* **76**, 203–225.

W. Hackbusch and S. Kühn (2009), A new scheme for the tensor representation, *J. Fourier Anal. Appl.* **15**, 706–722.

W. Hackbusch, B. N. Khoromskij and E. E. Tyrtyshnikov (2008), Approximate iterations for structured matrices, *Numer. Math.* **109**, 365–383.

W. Hackbusch, B. N. Khoromskij, S. Sauter and E. E. Tyrtyshnikov (2012), Use of tensor formats in elliptic eigenvalue problems, *Numer. Linear Algebra Appl.* **19**, 133–151.

J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken and F. Verstraete (2016), Unifying time evolution and optimization with matrix product states, *Phys. Rev. B* **94**, 165116.

J. Haegeman, T. J. Osborne and F. Verstraete (2013), Post-matrix product state methods: To tangent space and beyond, *Phys. Rev. B* **88**, 075133.

T. Helgaker, P. Jørgensen and J. Olsen (2000), *Molecular Electronic-Structure Theory*, Wiley.

U. Helmke and M. A. Shayman (1995), Critical points of matrix least squares distance functions, *Linear Algebra Appl.* **215**, 1–19.

J. S. Hesthaven, G. Rozza and B. Stamm (2016), *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics, Springer; BCAM Basque Center for Applied Mathematics, Bilbao.

C. J. Hillar and L.-H. Lim (2013), Most tensor problems are NP-hard, *J. Assoc. Comput. Mach.* **60**, Art. 45, 39.

F. L. Hitchcock (1927), The expression of a tensor or a polyadic as a sum of products, *J. Math. Phys.* **6**, 164–189.

F. L. Hitchcock (1928), Multiple invariants and generalized rank of a p-way matrix or tensor, *J. Math. Phys.* **7**, 39–79.

S. Holtz, T. Rohwedder and R. Schneider (2012*a*), The alternating linear scheme for tensor optimization in the tensor train format, *SIAM J. Sci. Comput.* **34**, A683–A713.

S. Holtz, T. Rohwedder and R. Schneider (2012*b*), On manifolds of tensors of fixed TT-rank, *Numer. Math.* **120**, 701–731.

M. Janzamin, R. Ge, J. Kossaifi, A. Anandkumar *et al.* (2019), Spectral learning on matrices and tensors, *Found. Trends Mach. Learn.* **12**, 393–536.

M. Kahlbacher and S. Volkwein (2007), Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems, *Discuss. Math. Differ. Incl. Control Optim.* **27**, 95–117.

Y. Kazashi and F. Nobile (2021), Existence of dynamical low rank approximations for random semi-linear evolutionary equations on the maximal interval, *Stoch. Partial Differ. Equ. Anal. Comput.* **9**, 603–629.

Y. Kazashi, F. Nobile and E. Vidličková (2021), Stability properties of a projector-splitting scheme for dynamical low rank approximation of random parabolic equations, *Numer. Math.* **149**, 973–1024.

V. A. Kazeev (2015), Quantized tensor structured finite elements for second-order elliptic PDEs in two dimensions. PhD thesis, ETH Zürich.

V. A. Kazeev and B. N. Khoromskij (2012), Low-rank explicit QTT representation of the Laplace operator and its inverse, *SIAM J. Matrix Anal. Appl.* **33**, 742–758.
V. A. Kazeev and C. Schwab (2018), Quantized tensor-structured finite elements for second-order elliptic PDEs in two dimensions, *Numer. Math.* **138**, 133–190.
V. A. Kazeev, B. N. Khoromskij and E. E. Tyrtyshnikov (2013*a*), Multilevel Toeplitz matrices generated by tensor-structured vectors and convolution with logarithmic complexity, *SIAM J. Sci. Comput.* **35**, A1511–A1536.
V. A. Kazeev, I. V. Oseledets, M. Rakhuba and C. Schwab (2017), QTT-finite-element approximation for multiscale problems, I: Model problems in one dimension, *Adv. Comput. Math.* **43**, 411–442.
V. A. Kazeev, I. V. Oseledets, M. V. Rakhuba and C. Schwab (2022), Quantized tensor FEM for multiscale problems: Diffusion problems in two and three dimensions, *Multiscale Model. Simul.* **20**, 893–935.
V. A. Kazeev, O. Reichmann and C. Schwab (2013*b*), Low-rank tensor structure of linear diffusion operators in the TT and QTT formats, *Linear Algebra Appl.* **438**, 4204–4221.
S. Keller, M. Dolfi, M. Troyer and M. Reiher (2015), An efficient matrix product operator representation of the quantum chemical Hamiltonian, *J. Chem. Phys.* **143**, 244118.
V. Khoromskaia and B. N. Khoromskij (2015), Tensor numerical methods in quantum chemistry: From Hartree–Fock to excitation energies, *Phys. Chem. Chem. Phys.* **17**, 31491–31509.
V. Khoromskaia and B. N. Khoromskij (2016), Fast tensor method for summation of long-range potentials on 3D lattices with defects, *Numer. Linear Algebra Appl.* **23**, 249–271.
V. Khoromskaia and B. N. Khoromskij (2018), *Tensor Numerical Methods in Quantum Chemistry*, De Gruyter.
V. Khoromskaia, B. N. Khoromskij and R. Schneider (2013), Tensor-structured factorized calculation of two-electron integrals in a general basis, *SIAM J. Sci. Comput.* **35**, A987–A1010.
B. N. Khoromskij (2009), Tensor-structured preconditioners and approximate inverse of elliptic operators in $\mathbb{R}^d$, *Constr. Approx.* **30**, 599–620.
B. N. Khoromskij (2011), $O(d \log N)$-quantics approximation of N-d tensors in high-dimensional numerical modeling, *Constr. Approx.* **34**, 257–280.
B. N. Khoromskij (2015), Tensor numerical methods for multidimensional PDEs: Theoretical analysis and initial applications, *ESAIM Proc. Surveys* **48**, 1–28.
B. N. Khoromskij (2018), *Tensor Numerical Methods in Scientific Computing*, Vol. 19 of Radon Series on Computational and Applied Mathematics, De Gruyter.
B. N. Khoromskij and S. Miao (2014), Superfast wavelet transform using quantics-TT approximation, I: Application to Haar wavelets, *Comput. Methods Appl. Math.* **14**, 537–553.
B. N. Khoromskij and I. V. Oseledets (2010), Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDEs, *Comput. Methods Appl. Math.* **10**, 376–394.
B. N. Khoromskij and I. V. Oseledets (2011), QTT approximation of elliptic solution operators in higher dimensions, *Russian J. Numer. Anal. Math. Modelling* **26**, 303–322.
B. N. Khoromskij and C. Schwab (2011), Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs, *SIAM J. Sci. Comput.* **33**, 364–385.
B. N. Khoromskij, V. Khoromskaia and H.-J. Flad (2011), Numerical solution of the Hartree–Fock equation in multilevel tensor-structured format, *SIAM J. Sci. Comput.* **33**, 45–65.

B. N. Khoromskij, V. Khoromskaia, S. R. Chinnamsetty and H.-J. Flad (2009), Tensor decomposition in electronic structure calculations on 3D Cartesian grids, *J. Comput. Phys.* **228**, 5749–5762.

E. Kieri and B. Vandereycken (2019), Projection methods for dynamical low-rank approximation of high-dimensional problems, *Comput. Methods Appl. Math.* **19**, 73–92.

E. Kieri, C. Lubich and H. Walach (2016), Discretized dynamical low-rank approximation in the presence of small singular values, *SIAM J. Numer. Anal.* **54**, 1020–1038.

O. Koch and C. Lubich (2007*a*), Dynamical low-rank approximation, *SIAM J. Matrix Anal. Appl.* **29**, 434–454.

O. Koch and C. Lubich (2007*b*), Regularity of the multi-configuration time-dependent Hartree approximation in quantum molecular dynamics, *M2AN Math. Model. Numer. Anal.* **41**, 315–331.

O. Koch and C. Lubich (2010), Dynamical tensor approximation, *SIAM J. Matrix Anal. Appl.* **31**, 2360–2375.

O. Koch and C. Lubich (2011), Variational-splitting time integration of the multi-configuration time-dependent Hartree–Fock equations in electron dynamics, *IMA J. Numer. Anal.* **31**, 379–395.

T. G. Kolda and B. W. Bader (2009), Tensor decompositions and applications, *SIAM Rev.* **51**, 455–500.

D. Kressner and C. Tobler (2011*a*), Low-rank tensor Krylov subspace methods for parametrized linear systems, *SIAM J. Matrix Anal. Appl.* **32**, 1288–1316.

D. Kressner and C. Tobler (2011*b*), Preconditioned low-rank methods for high-dimensional elliptic PDE eigenvalue problems, *Comput. Methods Appl. Math.* **11**, 363–381.

D. Kressner and A. Uschmajew (2016), On low-rank approximability of solutions to high-dimensional operator equations and eigenvalue problems, *Linear Algebra Appl.* **493**, 556–572.

D. Kressner, M. Steinlechner and A. Uschmajew (2014*a*), Low-rank tensor methods with subspace correction for symmetric eigenvalue problems, *SIAM J. Sci. Comput.* **36**, A2346–A2368.

D. Kressner, M. Steinlechner and B. Vandereycken (2014*b*), Low-rank tensor completion by Riemannian optimization, *BIT Numer. Math.* **54**, 447–468.

D. Kressner, M. Steinlechner and B. Vandereycken (2016), Preconditioned low-rank Riemannian optimization for linear systems with tensor product structure, *SIAM J. Sci. Comput.* **38**, A2018–A2044.

C. Krumnow, M. Pfeffer and A. Uschmajew (2021), Computing eigenspaces with low rank constraints, *SIAM J. Sci. Comput.* **43**, A586–A608.

C. Krumnow, L. Veis, O. Legeza and J. Eisert (2016), Fermionic orbital optimization in tensor network states, *Phys. Rev. Lett.* **117**, 210402.

S. Kühn (2012), Hierarchische Tensordarstellung. PhD thesis, Universität Leipzig.

J. M. Landsberg (2012), *Tensors: Geometry and Applications*, American Mathematical Society.

J. M. Landsberg, Y. Qi and K. Ye (2012), On the geometry of tensor network states, *Quantum Inf. Comput.* **12**, 346–354.

K. Lee and H. C. Elman (2017), A preconditioned low-rank projection method with a rank-reduction scheme for stochastic partial differential equations, *SIAM J. Sci. Comput.* **39**, S828–S850.

W. A. Light and E. W. Cheney (1985), *Approximation Theory in Tensor Product Spaces*, Vol. 1169 of Lecture Notes in Mathematics, Springer.
L.-H. Lim (2021), Tensors in computations, *Acta Numer.* **30**, 555–764.
C. Lubich (2005), On variational approximations in quantum molecular dynamics, *Math. Comp.* **74**, 765–779.
C. Lubich (2008), *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS).
C. Lubich (2015), Time integration in the multiconfiguration time-dependent Hartree method of molecular quantum dynamics, *Appl. Math. Res. Express. AMRX* **2015**, 311–328.
C. Lubich and I. V. Oseledets (2014), A projector-splitting integrator for dynamical low-rank approximation, *BIT Numer. Math.* **54**, 171–188.
C. Lubich, I. V. Oseledets and B. Vandereycken (2015), Time integration of tensor trains, *SIAM J. Numer. Anal.* **53**, 917–941.
C. Lubich, T. Rohwedder, R. Schneider and B. Vandereycken (2013), Dynamical approximation by hierarchical Tucker and tensor-train tensors, *SIAM J. Matrix Anal. Appl.* **34**, 470–494.
C. Marcati, M. Rakhuba and C. Schwab (2022*a*), Tensor rank bounds for point singularities in $\mathbb{R}^3$, *Adv. Comput. Math.* **48**, 18.
C. Marcati, M. Rakhuba and J. E. M. Ulander (2022*b*), Low-rank tensor approximation of singularly perturbed boundary value problems in one dimension, *Calcolo* **59**, 2.
L. Markeeva, I. Tsybulin and I. V. Oseledets (2021), QTT-isogeometric solver in two dimensions, *J. Comput. Phys.* **424**, 109835.
H. G. Matthies and E. Zander (2012), Solving stochastic systems with low-rank tensor compression, *Linear Algebra Appl.* **436**, 3819–3838.
H.-D. Meyer, F. Gatti and G. A. Worth (2009), *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*, Wiley.
H.-D. Meyer, U. Manthe and L. S. Cederbaum (1990), The multi-configurational time-dependent Hartree approach, *Chem. Phys. Lett.* **165**, 73–78.
B. Michel and A. Nouy (2022), Learning with tree tensor networks: Complexity estimates and model selection, *Bernoulli* **28**, 910–936.
M. J. Mohlenkamp (2010), A center-of-mass principle for the multiparticle Schrödinger equation, *J. Math. Phys.* **51**, 022112.
J.-J. Moreau (1965), Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France* **93**, 273–299.
V. Murg, F. Verstraete, O. Legeza and R. M. Noack (2010), Simulating strongly correlated quantum systems with tree tensor networks, *Phys. Rev. B* **82**, 205105.
E. Musharbash and F. Nobile (2018), Dual dynamically orthogonal approximation of incompressible Navier Stokes equations with random boundary conditions, *J. Comput. Phys.* **354**, 135–162.
E. Musharbash, F. Nobile and E. Vidličková (2020), Symplectic dynamical low rank approximation of wave equations with random parameters, *BIT Numer. Math.* **60**, 1153–1201.
E. Musharbash, F. Nobile and T. Zhou (2015), Error analysis of the dynamically orthogonal approximation of time dependent random PDEs, *SIAM J. Sci. Comput.* **37**, A776–A810.

A. Novikov, M. Trofimov and I. Oseledets (2018), Exponential machines, *Bull. Pol. Acad. Sci. Math.* **66**, 789–797.
R. Orús (2014), A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys.* **349**, 117–158.
I. V. Oseledets (2009*a*), Approximation of matrices with logarithmic number of parameters, *Dokl. Math.* **80**, 653–654.
I. V. Oseledets (2009*b*), On a new tensor decomposition, *Dokl. Akad. Nauk* **427**, 168–169. In Russian; English translation in *Dokl. Math.* **80**, 495–496 (2009).
I. V. Oseledets (2011*a*), DMRG approach to fast linear algebra in the TT-format, *Comput. Methods Appl. Math.* **11**, 382–393.
I. V. Oseledets (2011*b*), Tensor-train decomposition, *SIAM J. Sci. Comput.* **33**, 2295–2317.
I. V. Oseledets (2013), Constructive representation of functions in low-rank tensor formats, *Constr. Approx.* **37**, 1–18.
I. V. Oseledets and S. V. Dolgov (2012), Solution of linear systems and matrix inversion in the TT-format, *SIAM J. Sci. Comput.* **34**, A2718–A2739.
I. V. Oseledets and E. E. Tyrtyshnikov (2009*a*), Breaking the curse of dimensionality, or how to use SVD in many dimensions, *SIAM J. Sci. Comput.* **31**, 3744–3759.
I. V. Oseledets and E. E. Tyrtyshnikov (2009*b*), Recursive decomposition of multidimensional tensors, *Dokl. Akad. Nauk* **427**, 14–16. In Russian; English translation in *Dokl. Math.* **80**, 460–462 (2009).
I. V. Oseledets and E. E. Tyrtyshnikov (2010), TT-cross approximation for multidimensional arrays, *Linear Algebra Appl.* **432**, 70–88.
I. V. Oseledets, M. V. Rakhuba and A. Uschmajew (2018), Alternating least squares as moving subspace correction, *SIAM J. Numer. Anal.* **56**, 3459–3479.
M. Oster, L. Sallandt and R. Schneider (2022), Approximating optimal feedback controllers of finite horizon control problems using hierarchical tensor formats, *SIAM J. Sci. Comput.* **44**, B746–B770.
A. Ostermann, C. Piazzola and H. Walach (2019), Convergence of a low-rank Lie–Trotter splitting for stiff matrix differential equations, *SIAM J. Numer. Anal.* **57**, 1947–1966.
B. N. Parlett (1998), *The Symmetric Eigenvalue Problem*, Vol. 20 of Classics in Applied Mathematics, SIAM. Corrected reprint of the 1980 original.
M. Pfeffer (2018), Tensor methods for the numerical solution of high-dimensional parametric partial differential equations. PhD thesis, Technische Universität Berlin.
M. Primbs (2010), New stable biorthogonal spline-wavelets on the interval, *Results Math.* **57**, 121–162.
C. Prud'homme and A. T. Patera (2004), Reduced-basis output bounds for approximately parameterized elliptic coercive partial differential equations, *Comput. Vis. Sci.* **6**, 147–162.
M. Rakhuba (2021), Robust alternating direction implicit solver in quantized tensor formats for a three-dimensional elliptic PDE, *SIAM J. Sci. Comput.* **43**, A800–A827.
M. Rakhuba, A. Novikov and I. V. Oseledets (2019), Low-rank Riemannian eigensolver for high-dimensional Hamiltonians, *J. Comput. Phys.* **396**, 718–737.
M. V. Rakhuba and I. V. Oseledets (2016), Grid-based electronic structure calculations: The tensor decomposition approach, *J. Comput. Phys.* **312**, 19–30.
M. V. Rakhuba and I. V. Oseledets (2018), Jacobi–Davidson method on low-rank matrix manifolds, *SIAM J. Sci. Comput.* **40**, A1149–A1170.

L. Richter, L. Sallandt and N. Nüsken (2021), Solving high-dimensional parabolic PDEs using the tensor train format, in *38th International Conference on Machine Learning*, Vol. 139 of Proceedings of Machine Learning Research, PMLR, pp. 8998–9009.

T. Rohwedder and A. Uschmajew (2013), On local convergence of alternating schemes for optimization of convex problems in the tensor train format, *SIAM J. Numer. Anal.* **51**, 1134–1162.

G. Rozza, D. B. P. Huynh and A. T. Patera (2008), Reduced basis approximation and *a posteriori* error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics, *Arch. Comput. Methods Engrg* **15**, 229–275.

R. A. Ryan (2002), *Introduction to Tensor Products of Banach Spaces*, Springer Monographs in Mathematics, Springer.

T. P. Sapsis and P. F. Lermusiaux (2009), Dynamically orthogonal field equations for continuous stochastic dynamical systems, *Phys. D* **238**, 2347–2360.

E. Schmidt (1907), Zur Theorie der linearen und nichtlinearen Integralgleichungen, *Math. Ann.* **63**, 433–476.

R. Schneider and A. Uschmajew (2014), Approximation rates for the hierarchical tensor format in periodic Sobolev spaces, *J. Complexity* **30**, 56–71.

R. Schneider and A. Uschmajew (2015), Convergence results for projected line-search methods on varieties of low-rank matrices via Łojasiewicz inequality, *SIAM J. Optim.* **25**, 622–646.

U. Schollwöck (2011), The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96–192.

S. Scholz and H. Yserentant (2017), On the approximation of electronic wavefunctions by anisotropic Gauss and Gauss–Hermite functions, *Numer. Math.* **136**, 841–874.

C. Schwab and C. J. Gittelson (2011), Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs, *Acta Numer.* **20**, 291–467.

C. Schwab and R. Stevenson (2009), Space–time adaptive wavelet methods for parabolic evolution problems, *Math. Comp.* **78**, 1293–1318.

Y.-Y. Shi, L.-M. Duan and G. Vidal (2006), Classical simulation of quantum many-body systems with a tree tensor network, *Phys. Rev. A* **74**, 022320.

N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis and C. Faloutsos (2017), Tensor decomposition for signal processing and machine learning, *IEEE Trans. Signal Process.* **65**, 3551–3582.

M. Signoretto, Q. Tran Dinh, L. De Lathauwer and J. A. K. Suykens (2014), Learning with tensors: A framework based on convex optimization and spectral regularization, *Mach. Learn.* **94**, 303–351.

S. Singh, R. N. C. Pfeifer and G. Vidal (2011), Tensor network states and algorithms in the presence of a global $U(1)$ symmetry, *Phys. Rev. B* **83**, 115125.

M. Steinlechner (2016), Riemannian optimization for high-dimensional tensor completion, *SIAM J. Sci. Comput.* **38**, S461–S484.

F. Stenger (1993), *Numerical Methods Based on Sinc and Analytic Functions*, Vol. 20 of Springer Series in Computational Mathematics, Springer.

R. Stevenson (2009), Adaptive wavelet methods for solving operator equations: An overview, in *Multiscale, Nonlinear and Adaptive Approximation* (R. DeVore and A. Kunoth, eds), Springer, pp. 543–597.

E. Stoudenmire and D. J. Schwab (2016), Supervised learning with tensor networks, in *Advances in Neural Information Processing Systems 29* (D. Lee *et al.*, eds), Curran Associates.
Y. Sun and M. Kumar (2014), Numerical solution of high dimensional stationary Fokker–Planck equations via tensor decomposition and Chebyshev spectral differentiation, *Comput. Math. Appl.* **67**, 1960–1977.
S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider and Ö. Legeza (2015), Tensor product methods and entanglement optimization for *ab initio* quantum chemistry, *Internat. J. Quantum Chem.* **115**, 1342–1391.
C. Tobler (2012), Low-rank tensor methods for linear systems and eigenvalue problems. PhD thesis, ETH Zürich.
L. R. Tucker (1964), The extension of factor analysis to three-dimensional matrices, in *Contributions to Mathematical Psychology* (H. Gulliksen and N. Frederiksen, eds), Holt, Rinehart & Winston, pp. 109–127.
L. R. Tucker (1966), Some mathematical notes on three-mode factor analysis, *Psychometrika* **31**, 279–311.
E. E. Tyrtyshnikov (2003), Tensor approximations of matrices generated by asymptotically smooth functions, *Mat. Sb.* **194**, 147–160.
A. Uschmajew (2010), Well-posedness of convex maximization problems on Stiefel manifolds and orthogonal tensor product approximations, *Numer. Math.* **115**, 309–331.
A. Uschmajew (2012), Local convergence of the alternating least squares algorithm for canonical tensor approximation, *SIAM J. Matrix Anal. Appl.* **33**, 639–652.
A. Uschmajew (2013), Zur Theorie der Niedrigrangapproximation in Tensorprodukten von Hilberträumen. PhD thesis, Technische Universität Berlin.
A. Uschmajew and B. Vandereycken (2013), The geometry of algorithms using hierarchical tensors, *Linear Algebra Appl.* **439**, 133–166.
A. Uschmajew and B. Vandereycken (2015), Greedy rank updates combined with Riemannian descent methods for low-rank optimization, in *2015 International Conference on Sampling Theory and Applications (SampTA)*, pp. 420–424.
A. Uschmajew and B. Vandereycken (2020), Geometric methods on low-rank matrix and tensor manifolds, in *Handbook of Variational Methods for Nonlinear Geometric Data* (P. Grohs *et al.*, eds), Springer, pp. 261–313.
G. Vidal (2003), Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902.
H. Wang and M. Thoss (2003), Multilayer formulation of the multiconfiguration time-dependent Hartree theory, *J. Chem Phys.* **119**, 1289–1299.
L. Wang and M. T. Chu (2014), On the global convergence of the alternating least squares method for rank-one approximation to generic tensors, *SIAM J. Matrix Anal. Appl.* **35**, 1058–1072.
J. Weidmann (1980), *Linear Operators in Hilbert Spaces*, Vol. 68 of Graduate Texts in Mathematics, Springer.
S. R. White (1992), Density matrix formulation for quantum renormalization groups, *Phys. Rev. Lett.* **69**, 2863–2866.
S. R. White (2005), Density matrix renormalization group algorithms with a single center site, *Phys. Rev. B* **72**, 180403.
S. R. White and R. L. Martin (1999), *Ab initio* quantum chemistry using the density matrix renormalization group, *J. Chem. Phys.* **110**, 4127–4130.

J. H. Wilkinson (1968), Global convergence of tridiagonal QR algorithm with origin shifts, *Linear Algebra Appl.* **1**, 409–420.
S. Wouters and D. Van Neck (2014), The density matrix renormalization group for *ab initio* quantum chemistry, *Europ. Phys. J. D* **68**, 1–20.
M. Yang and S. R. White (2020), Time-dependent variational principle with ancillary Krylov subspace, *Phys. Rev. B* **102**, 094315.
H. Yserentant (2010), *Regularity and Approximability of Electronic Wave Functions*, Springer.
A. Zeiser (2010), Direkte Diskretisierung der Schrödinger-Gleichung auf dünnen Gittern. PhD thesis, TU Berlin.
G. Zhou, W. Huang, K. A. Gallivan, P. Van Dooren and P.-A. Absil (2016), A Riemannian rank-adaptive method for low-rank optimization, *Neurocomput.* **192**, 72–80.