

# PRECONDITIONED LOW-RANK RIEMANNIAN OPTIMIZATION FOR LINEAR SYSTEMS WITH TENSOR PRODUCT STRUCTURE\*

DANIEL KRESSNER<sup>†</sup>, MICHAEL STEINLECHNER<sup>†</sup>, AND BART VANDEREYCKEN<sup>‡</sup>

**Abstract.** The numerical solution of partial differential equations on high-dimensional domains gives rise to computationally challenging linear systems. When using standard discretization techniques, the size of the linear system grows exponentially with the number of dimensions, making the use of classic iterative solvers infeasible. During the last few years, low-rank tensor approaches have been developed that allow one to mitigate this *curse of dimensionality* by exploiting the underlying structure of the linear operator. In this work, we focus on tensors represented in the Tucker and tensor train formats. We propose two preconditioned gradient methods on the corresponding low-rank tensor manifolds: a Riemannian version of the preconditioned Richardson method as well as an approximate Newton scheme based on the Riemannian Hessian. For the latter, considerable attention is given to the efficient solution of the resulting Newton equation. In numerical experiments, we compare the efficiency of our Riemannian algorithms with other established tensor-based approaches such as a truncated preconditioned Richardson method and the alternating linear scheme. The results show that our approximate Riemannian Newton scheme is significantly faster in cases when the application of the linear operator is expensive.

**Key words.** tensors, tensor train, matrix product states, Riemannian optimization, low rank, high dimensionality

**AMS subject classifications.** 65F10, 15A69, 65K05, 58C05

**DOI.** 10.1137/15M1032909

**1. Introduction.** This work is concerned with the approximate solution of large-scale linear systems  $Ax = f$  with  $A \in \mathbb{R}^{n \times n}$ . In certain applications, such as the structured discretization of  $d$ -dimensional partial differential equations (PDEs), the size of the linear system naturally decomposes as  $n = n_1 n_2 \cdots n_d$  with  $n_\mu \in \mathbb{N}$  for  $\mu = 1, \dots, d$ . This allows us to view  $Ax = f$  as a tensor equation

$$(1.1) \quad \mathcal{A}\mathbf{X} = \mathbf{F},$$

where  $\mathbf{F}, \mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$  are tensors of order  $d$  and  $\mathcal{A}$  is a linear operator on  $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ . In the following, we restrict ourselves to the symmetric positive definite case, although some of the developments can, in principle, be generalized to indefinite and nonsymmetric matrices.

Assuming  $\mathcal{A}$  to be symmetric positive definite allows us to reformulate (1.1) as an optimization problem

$$(1.2) \quad \min_{\mathbf{X} \in \mathbb{R}^{n_1 \times \cdots \times n_d}} \frac{1}{2} \langle \mathbf{X}, \mathcal{A}\mathbf{X} \rangle - \langle \mathbf{X}, \mathbf{F} \rangle.$$

\*Submitted to the journal's Methods and Algorithms for Scientific Computing section July 28, 2015; accepted for publication (in revised form) May 10, 2016; published electronically July 6, 2016.  
<http://www.siam.org/journals/sisc/38-4/M103290.html>

<sup>†</sup>MATHICSE-ANCHP, Section de Mathématiques, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (daniel.kressner@epfl.ch, michael.steinlechner@epfl.ch). The work of the second author was supported by the SNSF research module *Riemannian optimization for solving high-dimensional problems with low-rank tensor techniques* within the SNSF ProDoc *Efficient Numerical Methods for Partial Differential Equations*. Parts of this work were conducted during a research stay at PACM, Princeton University, with a mobility grant from SNSF.

<sup>‡</sup>Section de Mathématiques, Université de Genève, 2–4 rue du Lièvre, 1211 Genève, Switzerland (bart.vandereycken@unige.ch).

It is well-known that the above problem is equivalent to minimizing the  $\mathcal{A}$ -induced norm of the error  $\|\mathbf{X} - \mathcal{A}^{-1}\mathbf{F}\|_{\mathcal{A}}$ . Neither (1.1) nor (1.2) is computationally tractable for larger values of  $d$ . During the last decade, low-rank tensor techniques have been developed that aim at dealing with this curse of dimensionality by approximating  $\mathbf{F}$  and  $\mathbf{X}$  in a compressed format; see [18, 20] for overviews. One approach consists of restricting (1.2) to a subset  $\mathcal{M} \subset \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$  of compressed tensors:

$$(1.3) \quad \min_{\mathbf{X} \in \mathcal{M}} f(\mathbf{X}) := \frac{1}{2} \langle \mathbf{X}, \mathcal{A}\mathbf{X} \rangle - \langle \mathbf{X}, \mathbf{F} \rangle.$$

Examples for  $\mathcal{M}$  include the Tucker format [57, 31], the tensor train (TT) format [49], the matrix product states (MPS) format [4], or the hierarchical Tucker format [17, 22]. Assuming that the corresponding ranks are fixed,  $\mathcal{M}$  is a smooth embedded submanifold of  $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$  for each of these formats [25, 58, 59, 23]. This property does not hold for the CANDECOMP/PARAFAC (CP) format, which we will therefore not consider.

**1.1. Riemannian optimization.** When  $\mathcal{M}$  is a manifold, Riemannian optimization techniques [1] can be used to address (1.3). In a related context, first-order methods, such as Riemannian steepest descent and nonlinear CG, have been successfully applied to matrix completion [10, 42, 46, 61] and tensor completion [11, 34, 50, 54].

Similar to Euclidean optimization, the condition number of the Riemannian Hessian of the objective function is instrumental in predicting the performance of first-order optimization algorithms on manifolds; see, e.g., [41, Thm. 2] and [1, Thm. 4.5.6]. As will be evident from (4.2) in section 4.1, an ill-conditioned operator  $\mathcal{A}$  can be expected to yield an ill-conditioned Riemannian Hessian. As this is the case for the applications we consider, any naive first-order method will be prohibitively slow and noncompetitive with existing methods.

For Euclidean optimization, it is well known that preconditioning or, equivalently, adapting the underlying metric can be used to address the slow convergence of such first-order methods. Combining steepest descent with the Hessian as a (variable) preconditioner yields the Newton method with (local) second-order convergence [45, sect. 1.3.1]. To overcome the high computational cost associated with Newton's method, several approximate Newton methods exist that use cheaper second-order models. For example, Gauss–Newton is a particularly popular approximation when solving nonlinear least-squares problems.

For Riemannian optimization, the connection between preconditioning and adapting the metric is less immediate and we explore both directions to speed up first-order methods. On the one hand, we will employ a rather ad hoc way to precondition the Riemannian gradient direction. On the other hand, we will consider an approximate Newton method that can be interpreted as a constrained Gauss–Newton method. This requires setting up and solving linear systems with the Riemannian Hessian or an approximation thereof. In [62], it was shown that neglecting curvature terms in the Riemannian Hessian leads to an efficient low-rank solver for Lyapunov matrix equations. We will extend these developments to more general equations with tensors approximated in the Tucker and the TT/MPS formats.

In passing, we mention that there is another notion of preconditioning for Riemannian optimization on low-rank matrix manifold; see, e.g., [43, 44, 46]. These techniques address the ill-conditioning of the manifold parametrization, an aspect that is not related and relevant to our developments, as we do not directly work with the parametrization.

**1.2. Other methods.** Riemannian optimization is by no means the only sensible approach to finding low-rank tensor approximations to the solution of the linear system (1.1). Exponential sum approximations [16, 21] and tensorized Krylov subspace methods [35] have been proven to be effective for Laplace-like operators  $\mathcal{L}$ , having a matrix representation

$$(1.4) \quad L = I_{n_d} \otimes \cdots \otimes I_{n_2} \otimes L_1 + I_{n_d} \otimes \cdots \otimes I_{n_3} \otimes L_2 \otimes I_{n_1} + \cdots + L_d \otimes I_{n_{d-1}} \otimes \cdots \otimes I_1,$$

with symmetric positive definite  $L_\mu \in \mathbb{R}^{n_\mu \times n_\mu}$  and identity matrices  $I_{n_\mu}$ . For more general equations, a straightforward approach is to apply standard iterative methods, such as the Richardson iteration or the CG method, to (1.1) and represent all iterates in the low-rank tensor format; see [6, 13, 27, 29, 36] for examples. For instance, such a truncated iteration can be applied to

$$(1.5) \quad \mathcal{A} = \mathcal{L} + \mathcal{V},$$

using the Laplace-like operator  $\mathcal{L}$  as a preconditioner to ensure reasonable convergence. Equations of this form arise, for example, from the discretization of the Schrödinger Hamiltonian [39], for which  $\mathcal{L}$  and  $\mathcal{V}$  correspond to the discretization of the kinetic and the potential energy terms, respectively. Apart from preconditioning, another critical issue in truncated iterations is to strike a balance between maintaining convergence and avoiding excessive intermediate rank growth of the iterates. Only recently, this has been analyzed in more detail [5].

A very different approach consists of applying alternating optimization techniques to the constrained optimization problem (1.3). Such methods have originated in quantum physics, most notably the so-called density matrix renormalization group (DMRG) method to address eigenvalue problems in the context of strongly correlated quantum lattice systems; see [52] for an overview. The ideas of DMRG have been extended to linear systems in the numerical analysis community in [14, 15, 24, 48]. The fixed-rank versions are generally referred to as alternating linear schemes (ALS) while the alternating minimal energy scheme (AMEn) from [15] allows for rank-adaptivity. While such methods often exhibit fast convergence, especially for operators of the form (1.5), their global convergence properties are poorly understood. Even the existing local convergence results for ALS [51, 59] offer little intuition on the convergence *rate*. The efficient implementation of ALS for low-rank tensor formats can be a challenge. In the presence of larger ranks, the (dense) subproblems that need to be solved in every step of ALS are large and tend to be ill-conditioned. In [33, 37], this issue has been addressed by combining an iterative solver with a preconditioner tailored to the subproblem. The design of such a preconditioner is by no means simple, even the knowledge of an effective preconditioner for the full-space problem (1.1) is generally not sufficient. So far, the only known effective preconditioners are based on exponential sum approximations for operators with Laplace-like structure (1.4), which is inherited by the subproblems.

**1.3. Contributions and outline.** Compared to existing approaches, the preconditioned low-rank Riemannian optimization methods proposed in this paper have a number of advantages. Due to imposing the manifold constraint, the issue of rank growth is completely avoided. Our methods have a global nature, all components of the low-rank tensor format are improved at once, and hence the stagnation typically observed during ALS sweeps is avoided. Moreover, we completely avoid the need for solving subproblems very accurately. One of our methods can make use of preconditioners for the full-space problem (1.1), while for the other methods preconditioners

are implicitly obtained from approximating the Riemannian Hessian. A disadvantage shared with existing approaches is that our methods strongly rely on structural properties of the operator, such as the one shown in (1.5), in order to construct effective preconditioners.

The rest of this paper is structured as follows. In section 2, we briefly review the Tucker and TT tensor formats and the structure of the corresponding manifolds. Then, in section 3, a Riemannian variant of the preconditioned Richardson method is introduced. In section 4, we incorporate second-order information using an approximation of the Riemannian Hessian of the cost function and solving the corresponding Newton equation. Finally, numerical experiments comparing the proposed algorithms with existing approaches are presented in section 5.

**2. Manifolds of low-rank tensors.** In this section, we discuss two different representations for tensors  $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ , namely, *Tucker* and *tensor train/matrix product states* (which we will simply call the TT format), along with their associated notions of low-rank structure and their geometry. We will only mention the main results here and refer to the articles by Kolda and Bader [31] and Oseledets [49] for more details. More elaborate discussions on the manifold structure and computational efficiency considerations can be found in [30, 34] for the Tucker format and in [40, 54, 59] for the TT format, respectively.

### 2.1. Tucker format.

*Format.* The *multilinear rank* of a tensor  $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$  is defined as the  $d$ -tuple

$$\text{rank}_{\text{ML}}(\mathbf{X}) = (r_1, r_2, \dots, r_d) = (\text{rank}(\mathbf{X}_{(1)}), \text{rank}(\mathbf{X}_{(2)}), \dots, \text{rank}(\mathbf{X}_{(d)}))$$

with

$$\mathbf{X}_{(\mu)} \in \mathbb{R}^{n_\mu \times (n_1 \cdots n_{\mu-1} n_{\mu+1} \cdots n_d)}, \quad \mu = 1, \dots, d,$$

the  $\mu$ th matricization of  $\mathbf{X}$ ; see [31] for more details.

Any tensor  $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$  of multilinear rank  $\mathbf{r} = (r_1, r_2, \dots, r_d)$  can be represented as

$$(2.1) \quad \mathbf{X}(i_1, \dots, i_d) = \sum_{j_1=1}^{r_1} \cdots \sum_{j_d=1}^{r_d} \mathbf{S}(j_1, j_2, \dots, j_d) U_1(i_1, j_1) U_2(i_2, j_2) \cdots U_d(i_d, j_d)$$

for some *core tensor*  $\mathbf{S} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$  and *factor matrices*  $U_\mu \in \mathbb{R}^{n_\mu \times r_\mu}$ ,  $\mu = 1, \dots, d$ . In the following, we always choose the factor matrices to have orthonormal columns,  $U_\mu^\top U_\mu = I_{r_\mu}$ .

Using the  $\mu$ th mode product  $\times_\mu$  (see [31]), one can write (2.1) more compactly as

$$(2.2) \quad \mathbf{X} = \mathbf{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d.$$

*Manifold structure.* It is known [30, 20, 58] that the set of tensors having multilinear rank  $\mathbf{r}$  forms a smooth submanifold embedded in  $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ . This manifold  $\mathcal{M}_{\mathbf{r}}$  is of dimension

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

For  $\mathbf{X} \in \mathcal{M}_{\mathbf{r}}$  represented as in (2.2), any tangent vector  $\xi \in T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  can be written as

$$(2.3) \quad \xi = \mathbf{S} \times_1 \delta U_1 \times_2 U_2 \cdots \times_d U_d + \mathbf{S} \times_1 U_1 \times_2 \delta U_2 \cdots \times_d U_d \\ + \cdots + \mathbf{S} \times_1 U_1 \times_2 U_2 \cdots \times_d \delta U_d + \delta \mathbf{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$$

for some first-order variations  $\delta \mathbf{S} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$  and  $\delta U_\mu \in \mathbb{R}^{n_\mu \times r_\mu}$ . This representation of tangent vectors allows us to decompose the tangent space  $T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  orthogonally as

$$(2.4) \quad T_{\mathbf{X}}\mathcal{M} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_d \oplus \mathcal{V}_{d+1} \quad \text{with } \mathcal{V}_\mu \perp \mathcal{V}_\nu \quad \forall \mu \neq \nu,$$

where the subspaces  $\mathcal{V}_\mu$  are given by

$$(2.5) \quad \mathcal{V}_\mu = \left\{ \mathbf{S} \times_\mu \delta U_\mu \times_{\substack{\nu=1 \\ \nu \neq \mu}}^d U_\nu : \delta U_\mu \in \mathbb{R}^{n_\mu \times r_\mu}, \delta U_\mu^\top U_\mu = 0 \right\}, \quad \mu = 1, \dots, d,$$

and

$$\mathcal{V}_{d+1} = \left\{ \delta \mathbf{S} \times_{\nu=1}^d U_\nu : \delta \mathbf{S} \in \mathbb{R}^{r_1 \times \cdots \times r_d} \right\}.$$

In particular, this decomposition shows that given the core tensor  $\mathbf{S}$  and factor matrices  $U_\mu$  of  $\mathbf{X}$ , the tangent vector  $\xi$  is uniquely represented in terms of  $\delta \mathbf{S}$  and *gauged*  $\delta U_\mu$ .

*Projection onto tangent space.* Given  $\mathbf{Z} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ , the components  $\delta U_\mu$  and  $\delta \mathbf{S}$  of the orthogonal projection  $\xi = P_{T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}}(\mathbf{Z})$  are given by (see [30, eqn. (2.7)])

$$(2.6) \quad \delta \mathbf{S} = \mathbf{Z} \times_{\mu=1}^d U_\mu^\top, \\ \delta U_\mu = (I_{n_\mu} - U_\mu U_\mu^\top) \left[ \mathbf{Z} \times_{\substack{\nu=1 \\ \nu \neq \mu}}^d U_\nu^\top \right]_{(1)} \mathbf{S}_{(\mu)}^\dagger,$$

where  $\mathbf{S}_{(\mu)}^\dagger = \mathbf{S}_{(\mu)}^\top (\mathbf{S}_{(\mu)} \mathbf{S}_{(\mu)}^\top)^{-1}$  is the Moore–Penrose pseudoinverse of  $\mathbf{S}_{(\mu)}$ . The projection of a Tucker tensor of multilinear rank  $\tilde{\mathbf{r}}$  into  $T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  can be performed in  $O(dn\tilde{r}r^{d-1} + \tilde{r}^d r)$  operations, where we set  $\tilde{r} := \max_\mu \tilde{r}_\mu$ ,  $r := \max_\mu r_\mu$ , and  $\tilde{r} \geq r$ .

## 2.2. Representation in the TT format.

*Format.* The TT format is (implicitly) based on matricizations that merge the first  $\mu$  modes into row indices and the remaining indices into column indices:

$$\mathbf{X}^{<\mu>} \in \mathbb{R}^{(n_1 \cdots n_\mu) \times (n_{\mu+1} \cdots n_d)}, \quad \mu = 1, \dots, d-1.$$

The *TT rank* of  $\mathbf{X}$  is the tuple  $\text{rank}_{\text{TT}}(\mathbf{X}) := (r_0, r_1, \dots, r_d)$  with  $r_\mu = \text{rank}(\mathbf{X}^{<\mu>})$ . By definition,  $r_0 = r_d = 1$ .

A tensor  $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$  of TT rank  $\mathbf{r} = (r_0, r_1, \dots, r_d)$  admits the representation

$$(2.7) \quad \mathbf{X}(i_1, \dots, i_d) = U_1(i_1)U_2(i_2) \cdots U_d(i_d),$$

where each  $U_\mu(i_\mu)$  is a matrix of size  $r_{\mu-1} \times r_\mu$  for  $i_\mu = 1, 2, \dots, n_\mu$ . By stacking the matrices  $U_\mu(i_\mu)$ ,  $i_\mu = 1, 2, \dots, n_\mu$ , into third-order tensors  $\mathbf{U}_\mu$  of size  $r_{\mu-1} \times n_\mu \times r_\mu$ , the so-called TT cores, we can also write (2.7) as

$$\mathbf{X}(i_1, \dots, i_d) = \sum_{j_1=1}^{r_1} \cdots \sum_{j_{d-1}=1}^{r_{d-1}} \mathbf{U}_1(1, i_1, j_1) \mathbf{U}_2(j_1, i_2, j_2) \cdots \mathbf{U}_d(j_{d-1}, i_d, 1).$$

To access and manipulate individual cores, it is useful to introduce the *interface matrices*

$$\begin{aligned}\mathbf{X}_{\leq\mu} &= [U_1(i_1)U_2(i_2)\cdots U_\mu(i_\mu)] \in \mathbb{R}^{n_1 n_2 \cdots n_\mu \times r_\mu}, \\ \mathbf{X}_{\geq\mu} &= [U_\mu(i_\mu)U_{\mu+1}(i_{\mu+1})\cdots U_d(i_d)]^\top \in \mathbb{R}^{n_\mu n_{\mu+1} \cdots n_d \times r_{\mu-1}},\end{aligned}$$

and

$$(2.8) \quad \mathbf{X}_{\neq\mu} = \mathbf{X}_{\geq\mu+1} \otimes I_{n_\mu} \otimes \mathbf{X}_{\leq\mu-1} \in \mathbb{R}^{n_1 n_2 \cdots n_d \times r_{\mu-1} n_\mu r_\mu}.$$

In particular, this allows us to pull out the  $\mu$ th core as  $\text{vec}(\mathbf{X}) = \mathbf{X}_{\neq\mu} \text{vec}(\mathbf{U}_\mu)$ , where  $\text{vec}(\cdot)$  denotes the vectorization of a tensor.

There is some freedom in choosing the cores in the representation (2.7). In particular, we can orthogonalize parts of  $\mathbf{X}$ . We say that  $\mathbf{X}$  is  $\mu$ -orthogonal if  $\mathbf{X}_{\leq\nu}^\top \mathbf{X}_{\leq\nu} = I_{r_\nu}$  for all  $\nu = 1, \dots, \mu-1$  and  $\mathbf{X}_{\geq\nu} \mathbf{X}_{\geq\nu}^\top = I_{r_{\nu-1}}$  for all  $\nu = \mu+1, \dots, d$ ; see, e.g., [54] for more details.

*Manifold structure.* The set of tensors having fixed TT rank,

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

forms a smooth embedded submanifold of  $\mathbb{R}^{n_1 \times \cdots \times n_d}$  (see [25, 20, 59]) of dimension

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

Similar to the Tucker format, the tangent space  $T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  at  $\mathbf{X} \in \mathcal{M}_{\mathbf{r}}$  admits an orthogonal decomposition:

$$(2.9) \quad T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_d \quad \text{with } \mathcal{V}_\mu \perp \mathcal{V}_\nu \quad \forall \mu \neq \nu.$$

Assuming that  $\mathbf{X}$  is  $d$ -orthogonal, the subspaces  $\mathcal{V}_\mu$  can be represented as

$$(2.10) \quad \begin{aligned}\mathcal{V}_\mu &= \left\{ \mathbf{X}_{\neq\mu} \text{vec}(\delta \mathbf{U}_\mu) : \delta \mathbf{U}_\mu \in \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu}, (\mathbf{U}_\mu^\mathbf{L})^\top \delta \mathbf{U}_\mu^\mathbf{L} = 0 \right\}, \quad \mu = 1, \dots, d-1, \\ \mathcal{V}_d &= \left\{ \mathbf{X}_{\neq d} \text{vec}(\delta \mathbf{U}_d) : \delta \mathbf{U}_d \in \mathbb{R}^{r_{d-1} \times n_d \times r_d} \right\}.\end{aligned}$$

Here,  $\mathbf{U}_\mu^\mathbf{L} \equiv \mathbf{U}_\mu^{<2>} \in \mathbb{R}^{r_{\mu-1} n_\mu \times r_\mu}$  is called the *left unfolding* of  $\mathbf{U}_\mu$  and it has orthonormal columns for  $\mu = 1, \dots, d-1$ , due to the  $d$ -orthogonality of  $\mathbf{X}$ . The *gauge conditions*  $(\mathbf{U}_\mu^\mathbf{L})^\top \delta \mathbf{U}_\mu^\mathbf{L} = 0$  for  $\mu \neq d$  ensure the mutual orthogonality of the subspaces  $\mathcal{V}_\mu$  and thus yield a unique representation of a tangent vector  $\xi$  in terms of gauged  $\delta \mathbf{U}_\mu$ . Hence, we can write any tangent vector  $\xi \in T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  in the convenient form

$$(2.11) \quad \xi = \sum_{\mu=1}^d \mathbf{X}_{\neq\mu} \text{vec}(\delta \mathbf{U}_\mu) \in \mathbb{R}^{n_1 n_2 \cdots n_d} \quad \text{s.t.} \quad (\mathbf{U}_\mu^\mathbf{L})^\top \delta \mathbf{U}_\mu^\mathbf{L} = 0 \quad \forall \mu \neq d.$$

*Projection onto tangent space.* The orthogonal projection  $P_{T_{\mathbf{X}}\mathcal{M}}$  onto the tangent space  $T_{\mathbf{X}}\mathcal{M}$  can be decomposed in accordance with (2.9):

$$P_{T_{\mathbf{X}}\mathcal{M}} = P^1 + P^2 + \cdots + P^d,$$

where  $P^\mu$  are orthogonal projections onto  $\mathcal{V}_\mu$ . Let  $\mathbf{X} \in \mathcal{M}_\mathbf{r}$  be  $d$ -orthogonal and  $\mathbf{Z} \in \mathbb{R}^{n_1 \times \dots \times n_d}$ . Then the projection can be written as

$$(2.12) \quad P_{T_{\mathbf{X}}\mathcal{M}_\mathbf{r}}(\mathbf{Z}) = \sum_{\mu=1}^d P^\mu(\mathbf{Z}), \quad \text{where} \quad P^\mu(\mathbf{Z}) = \mathbf{X}_{\neq \mu} \text{vec}(\delta \mathbf{U}_\mu).$$

For  $\mu = 1, \dots, d-1$ , the components  $\delta \mathbf{U}_\mu$  in this expression are given by [40, p. 924]

$$(2.13) \quad \delta \mathbf{U}_\mu^\mathbf{L} = (I_{n_\mu r_{\mu-1}} - P_\mu^\mathbf{L})(I_{n_\mu} \otimes \mathbf{X}_{\leq \mu-1}^\mathbf{T}) \mathbf{Z}^{<\mu>} \mathbf{X}_{\geq \mu+1} (\mathbf{X}_{\geq \mu+1}^\mathbf{T} \mathbf{X}_{\geq \mu+1})^{-1}$$

with  $P_\mu^\mathbf{L} = \mathbf{U}_\mu^\mathbf{L} (\mathbf{U}_\mu^\mathbf{L})^\mathbf{T}$  the orthogonal projector onto the range of  $\mathbf{U}_\mu^\mathbf{L}$ . For  $\mu = d$ , we have

$$(2.14) \quad \delta \mathbf{U}_d^\mathbf{L} = (I_{n_d} \otimes \mathbf{X}_{\leq d-1}^\mathbf{T}) \mathbf{Z}^{<d>}.$$

The projection of a tensor of TT rank  $\tilde{\mathbf{r}}$  into  $T_{\mathbf{X}}\mathcal{M}_\mathbf{r}$  can be performed in  $O(dnr\tilde{r}^2)$  operations, where we set again  $\tilde{r} := \max_\mu \tilde{r}_\mu$ ,  $r := \max_\mu r_\mu$  and  $\tilde{r} \geq r$ .

*Remark 1.* Equation (2.13) is not well-suited for numerical calculations due to the presence of the inverse of the Gram matrix  $\mathbf{X}_{\geq \mu+1}^\mathbf{T} \mathbf{X}_{\geq \mu+1}$ , which is typically severely ill-conditioned. In [28, 54], it was shown that by  $\mu$ -orthogonalizing the  $\mu$ th summand of the tangent vector representation, these inverses can be avoided at no extra costs. To keep the notation short, we do not include this individual orthogonalization in the equations above but make use of it in the implementation of the algorithm and the numerical experiments discussed in section 5.

**2.3. Retractions.** Riemannian optimization algorithms produce search directions that are contained in the tangent space  $T_{\mathbf{X}}\mathcal{M}_\mathbf{r}$  of the current iterate. To obtain the next iterate on the manifold, tangent vectors are mapped back to the manifold by application of a *retraction* map  $R$  that satisfies certain properties; see [3, Def. 1] for a formal definition.

It has been shown in [34] that the higher-order SVD (HOSVD) [12], which aims at approximating a given tensor of rank  $\tilde{\mathbf{r}}$  by a tensor of lower rank  $\mathbf{r}$ , constitutes a retraction on the Tucker manifold  $\mathcal{M}_\mathbf{r}$  that can be computed efficiently in  $O(dn\tilde{r}^2 + \tilde{r}^{d+1})$  operations. For the TT manifold, we will use the analogous TT-SVD [49, sect. 3] for a retraction with a computational cost of  $O(dn^3)$ ; see [54]. For both manifolds, we will denote by  $R(\mathbf{X} + \xi)$  the retraction<sup>1</sup> of  $\xi \in T_{\mathbf{X}}\mathcal{M}_\mathbf{r}$  that is computed by the HOSVD/TT-SVD of  $\mathbf{X} + \xi$ .

**3. First-order Riemannian optimization and preconditioning.** In this section, we discuss ways to incorporate preconditioners into simple first-order Riemannian optimization methods.

**3.1. Riemannian gradient descent.** To derive a first-order optimization method on a manifold  $\mathcal{M}_\mathbf{r}$ , we first need to construct the Riemannian gradient. For the cost function (1.3) associated with linear systems, the Euclidean gradient is given by

$$\nabla f(\mathbf{X}) = \mathcal{A}\mathbf{X} - \mathbf{F}.$$

<sup>1</sup>Note that the domain of definition of  $R$  is the affine tangent space  $\mathbf{X} + T_{\mathbf{X}}\mathcal{M}_\mathbf{r}$ , which departs from the usual convention to define  $R$  on  $T_{\mathbf{X}}\mathcal{M}_\mathbf{r}$  but makes sense for this particular type of retraction.

For both the Tucker and the TT formats,  $\mathcal{M}_{\mathbf{r}}$  is an embedded submanifold of  $\mathbb{R}^{n_1 \times \cdots \times n_d}$  and hence the Riemannian gradient can be obtained by projecting  $\nabla f$  onto the tangent space:

$$\text{grad } f(\mathbf{X}) = P_{T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}}(\mathcal{A}\mathbf{X} - \mathbf{F}).$$

Together with the retraction  $R$  of section 2.3, this yields the basic Riemannian gradient descent algorithm:

$$(3.1) \quad \mathbf{X}_{k+1} = R(\mathbf{X}_k + \alpha_k \xi_k) \quad \text{with} \quad \xi_k = -P_{T_{\mathbf{X}_k}\mathcal{M}} \nabla f(\mathbf{X}_k).$$

As usual, a suitable step size  $\alpha_k$  is obtained by standard Armijo backtracking line-search. Following [61], a good initial guess for the backtracking procedure is constructed by an exact linearized linesearch on the tangent space alone (that is, by neglecting the retraction):

$$(3.2) \quad \underset{\alpha}{\text{argmin}} f(\mathbf{X}_k + \alpha \xi_k) = -\frac{\langle \xi_k, \nabla f(\mathbf{X}_k) \rangle}{\langle \xi, \mathcal{A}\xi \rangle}.$$

### 3.2. Truncated preconditioned Richardson iteration.

*Truncated Richardson iteration.* The Riemannian gradient descent defined by (3.1) closely resembles a truncated Richardson iteration for solving linear systems:

$$(3.3) \quad \mathbf{X}_{k+1} = R(\mathbf{X}_k + \alpha_k \xi_k) \quad \text{with} \quad \xi_k = -\nabla f(\mathbf{X}_k) = \mathbf{F} - \mathcal{A}\mathbf{X}_k,$$

which was proposed for the CP tensor format in [29]. For the hierarchical Tucker format, a variant of the TT format, the iteration (3.3) has been analyzed in [5]. In contrast to manifold optimization, the rank does not need to be fixed but can be adjusted to strike a balance between low rank and convergence speed. It has been observed, for example in [32], that such an iterate-and-truncate strategy greatly benefits from preconditioners, not only to attain an acceptable convergence speed but also to avoid excessive rank growth of the intermediate iterates.

*Preconditioned Richardson iteration.* For the standard Richardson iteration defined by  $\mathbf{X}_{k+1} = \mathbf{X}_k - \alpha_k \xi_k$ , a symmetric positive definite preconditioner  $\mathcal{B}$  for  $\mathcal{A}$  can be incorporated as follows:

$$(3.4) \quad \mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathcal{B}^{-1} \xi_k \quad \text{with} \quad \xi_k = \mathbf{F} - \mathcal{A}\mathbf{X}_k.$$

Using the Cholesky factorization  $\mathcal{B} = \mathcal{C}\mathcal{C}^T$ , this iteration turns out to be equivalent to applying the Richardson iteration to the transformed symmetric positive definite linear system

$$\mathcal{C}^{-1}\mathcal{A}\mathcal{C}^{-T}\mathbf{Y} = \mathcal{C}^{-1}\mathbf{F}$$

after changing coordinates by  $\mathcal{C}^T\mathbf{X}_k$ . At the same time, (3.4) can be viewed as applying gradient descent in the inner product induced by  $\mathcal{B}$ .

*Truncated preconditioned Richardson iteration.* The most natural way of combining truncation with preconditioning leads to the *truncated preconditioned Richardson iteration*

$$(3.5) \quad \mathbf{X}_{k+1} = R(\mathbf{X}_k + \alpha_k \mathcal{B}^{-1} \xi_k) \quad \text{with} \quad \xi_k = \mathbf{F} - \mathcal{A}\mathbf{X}_k;$$

see also [29]. In view of Riemannian gradient descent (3.1), it appears natural to project the search direction to the tangent space, leading to the “geometric” variant

$$(3.6) \quad \mathbf{X}_{k+1} = R(\mathbf{X}_k + \alpha_k P_{T_{\mathbf{X}_k}\mathcal{M}_{\mathbf{r}}} \mathcal{B}^{-1} \xi_k) \quad \text{with} \quad \xi_k = \mathbf{F} - \mathcal{A}\mathbf{X}_k,$$

which we call the *truncated Riemannian preconditioned Richardson iteration*.



In terms of convergence, we have observed that the scheme (3.6) behaves similarly to (3.5); see section 5.4. However, it can be considerably cheaper per iteration: Since only tangent vectors need to be retracted in (3.6), the computation of the HOSVD/TT-SVD in  $R$  involves only tensors of bounded rank, regardless of the rank of  $\mathcal{B}^{-1}\xi_k$ . In particular, with  $\mathbf{r}$  the Tucker/TT rank of  $\mathbf{X}_k$ , the corresponding rank of  $\mathbf{X}_k + \alpha_k P_{T_{\mathbf{X}_k}\mathcal{M}_r} \mathcal{B}^{-1}\xi_k$  is at most  $2\mathbf{r}$ ; see [34, sect. 3.3] and [54, Prop. 3.1]. On the other hand, in (3.5) the rank of  $\mathbf{X}_k + \alpha_k \mathcal{B}^{-1}\xi_k$  is determined primarily by the quality of the preconditioner  $\mathcal{B}$  and can possibly be very large.

Another advantage of (3.6) occurs for the special but important case when  $\mathcal{B}^{-1} = \sum_{\alpha=1}^s \mathcal{B}_\alpha$ , where each term  $\mathcal{B}_\alpha$  is relatively cheap to apply. This is, for example, the case for the exponential sum preconditioner considered in section 5.4, with  $s$  terms and each  $\mathcal{B}_\alpha$  a Kronecker product of small matrices. By the linearity of  $P_{T_{\mathbf{X}_k}\mathcal{M}_r}$ , we have

$$(3.7) \quad P_{T_{\mathbf{X}_k}\mathcal{M}_r} \mathcal{B}^{-1}\xi_k = \sum_{\alpha=1}^s P_{T_{\mathbf{X}_k}\mathcal{M}_r} \mathcal{B}_\alpha \xi_k,$$

for which the expression on the right is usually cheaper to evaluate. To see this, suppose that for the TT format  $\mathcal{B}_\alpha \xi$  has TT ranks  $r_p$ . Then the preconditioned direction  $\mathcal{B}^{-1}\xi_k$  can be expected to have TT ranks  $sr_p$ . Hence, the straightforward application of  $P_{T_{\mathbf{X}_k}\mathcal{M}_r}$  to  $\mathcal{B}^{-1}\xi_k$  requires  $O(dn(sr_p)^2r)$  operations. Using the expression on the right-hand side of (3.7) instead reduces the cost to  $O(dnsr_p^2r)$  operations, since the summation of tangent vectors amounts to simply adding their parametrizations. In contrast, since retraction is a nonlinear operation, trying to achieve similar cost savings in (3.5) by simply truncating the accumulated sum subsequently may lead to severe cancellation [38, sect. 6.3].

**4. Riemannian optimization using a quadratic model.** As we will see in the numerical experiments in section 5, the convergence of the first-order methods presented above crucially depends on the availability of a good preconditioner  $\mathcal{B}$  for the full problem  $\mathcal{A}$ .

In this section, we present Riemannian optimization methods based on minimizing a quadratic model of  $f(\mathcal{X})$  on  $\mathcal{M}_r$ . This model is built by approximating the Riemannian Hessian of  $f$ . In this approximation, we also allow for replacing  $\mathcal{A}$  by a positive definite preconditioner  $\mathcal{B}$ . In contrast to the previous section, where the action of the inverse of  $\mathcal{B}$  was required in each iteration, we now require knowledge of  $\mathcal{B}$  itself. Moreover, the structure of  $\mathcal{B}$  needs to admit for an efficient solution of the quadratic model.

**4.1. Approximate Newton method.** The Riemannian Newton method [1] applied to (1.3) determines the search direction  $\xi_k$  from the equation

$$(4.1) \quad \mathcal{H}_{\mathbf{X}_k} \xi_k = -P_{T_{\mathbf{X}_k}\mathcal{M}_r} \nabla f(\mathbf{X}_k),$$

where the symmetric linear operator  $\mathcal{H}_{\mathbf{X}_k} : T_{\mathbf{X}_k}\mathcal{M}_r \rightarrow T_{\mathbf{X}_k}\mathcal{M}_r$  is the Riemannian Hessian of (1.3). Using [2], we have

$$(4.2) \quad \begin{aligned} \mathcal{H}_{\mathbf{X}_k} &= P_{T_{\mathbf{X}_k}\mathcal{M}_r} [\nabla^2 f(\mathbf{X}_k) + J_{\mathbf{X}_k} \nabla f(\mathbf{X}_k)] P_{T_{\mathbf{X}_k}\mathcal{M}_r} \\ &= P_{T_{\mathbf{X}_k}\mathcal{M}_r} [\mathcal{A} + J_{\mathbf{X}_k}(\mathcal{A}\mathbf{X}_k - \mathbf{F})] P_{T_{\mathbf{X}_k}\mathcal{M}_r} \end{aligned}$$

with the Fréchet derivative<sup>2</sup>  $J_{\mathbf{X}_k}$  of  $P_{T_{\mathbf{X}_k}\mathcal{M}_r}$ .

<sup>2</sup> $J_{\mathbf{X}_k}$  is an operator from  $\mathbb{R}^{n \times n \times \dots \times n}$  to the space of self-adjoint linear operators  $T_{\mathbf{X}_k}\mathcal{M}_r \rightarrow T_{\mathbf{X}_k}\mathcal{M}_r$ .

As usual, the Newton equation is only well-defined near a strict local minimizer and solving it exactly is prohibitively expensive in a large-scale setting. We therefore approximate the linear system (4.1) in two steps: First, we drop the term containing  $J_{\mathbf{X}_k}$ , and second, we replace  $\mathcal{A}$  by a positive definite preconditioner  $\mathcal{B}$ . The first approximation can be interpreted as neglecting the curvature of  $\mathcal{M}_{\mathbf{r}}$  or, equivalently, as linearizing the manifold at  $\mathbf{X}_k$ . Indeed, this term is void if  $\mathcal{M}_{\mathbf{r}}$  would be a (flat) linear subspace. This approximation is also known as constrained Gauss–Newton (see, e.g., [9]) since it replaces the constraint  $\mathbf{X} \in \mathcal{M}_{\mathbf{r}}$  with its linearization  $\mathbf{X} \in T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  and neglects the constraints in the Lagrangian.

The result is an *approximate Newton method* where the search direction  $\xi_k$  is determined from

$$(4.3) \quad P_{T_{\mathbf{X}_k}\mathcal{M}_{\mathbf{r}}} \mathcal{B} P_{T_{\mathbf{X}_k}\mathcal{M}_{\mathbf{r}}} \xi_k = P_{T_{\mathbf{X}_k}\mathcal{M}_{\mathbf{r}}} (\mathbf{F} - \mathcal{A}\mathbf{X}_k).$$

Since  $\mathcal{B}$  is positive definite, this equation is always well-defined for any  $\mathbf{X}_k$ . In addition,  $\xi_k$  is also gradient-related and hence the iteration

$$\mathbf{X}_{k+1} = R(\mathbf{X}_k + \alpha_k \xi_k)$$

is guaranteed to converge globally to a stationary point of the cost function if  $\alpha_k$  is determined from Armijo backtracking [1].

Despite all the simplifications, the numerical solution of (4.3) turns out to be a nontrivial task and needs to be tailored to the situation at hand. In the following sections, we explain how this task can be addressed for different situations. In section 4.2, we consider the Tucker manifold and discuss Laplacian-like as well as rank-one structures, nicely extending techniques from [62]. In section 4.3, we address the TT manifold and first present a general approach based on CG with an overlapping block-Jacobi preconditioner and then specialize it to Laplacian-like as well as rank-one structures.

**4.2. The approximate Riemannian Hessian in the Tucker case.** The solution of the linear system (4.3) was addressed for the matrix case ( $d = 2$ ) in [62, sect. 7.2]. In the following, we extend this approach to tensors in the Tucker format. To keep the presentation concise, we restrict ourselves to  $d = 3$ ; the extension to  $d > 3$  is straightforward.

For tensors of order 3 in the Tucker format, we write (4.3) as follows:

$$(4.4) \quad P_{T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}} \mathcal{B} P_{T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}} \xi = \eta,$$

where

- $\mathbf{X} \in \mathcal{M}_{\mathbf{r}}$  is parametrized by factor matrices  $U_1, U_2, U_3$  having *orthonormal columns* and the core tensor  $\mathbf{S}$ ;
- the right-hand side  $\eta \in T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  is given in terms of its *gauged* parametrization  $\delta U_1^\eta, \delta U_2^\eta, \delta U_3^\eta$ , and  $\delta \mathbf{S}^\eta$ , as in (2.3) and (2.5);
- the unknown  $\xi \in T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  is to be determined in terms of its *gauged* parametrization  $\delta U_1, \delta U_2, \delta U_3$ , and  $\delta \mathbf{S}$ , again as in (2.3) and (2.5).

To derive equations for  $\delta U_\mu$  with  $\mu = 1, 2, 3$  and  $\delta \mathbf{S}$  we exploit that  $T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  decomposes orthogonally into  $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_4$ ; see (2.4). This allows us to split (4.4) into a system of four coupled equations by projecting onto  $\mathcal{V}_\mu$  for  $\mu = 1, \dots, 4$ .

**4.2.1. Laplacian-like structure.** We first discuss the case of a Laplacian-like preconditioner  $\mathcal{B} = \mathcal{L}$  having the form (1.4). Since  $\xi \in T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}$  by assumption, we can insert  $\mathbf{Z} := \mathcal{L} P_{T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}}} \xi = \mathcal{L} \xi$  into (2.6). By exploiting the structure of  $\mathcal{L}$  and

the orthogonality of the gauged representation of tangent vectors (see (2.5)), we can simplify the expressions considerably and arrive at the equations

$$\begin{aligned}
 (4.5) \quad \delta U_1^\eta &= P_{U_1}^\perp \left( L_1 U_1 \delta S_{(1)} + L_1 \delta U_1 S_{(1)} + \delta U_1 S_{(1)} [I_{r_3} \otimes U_2^\top L_2 U_2 + U_3^\top L_3 U_3 \otimes I_{r_2}] \right) S_{(1)}^\dagger, \\
 \delta U_2^\eta &= P_{U_2}^\perp \left( L_2 U_2 \delta S_{(2)} + L_2 \delta U_2 S_{(2)} + \delta U_2 S_{(2)} [I_{r_3} \otimes U_1^\top L_1 U_1 + U_3^\top L_3 U_3 \otimes I_{r_1}] \right) S_{(2)}^\dagger, \\
 \delta U_3^\eta &= P_{U_3}^\perp \left( L_3 U_3 \delta S_{(3)} + L_3 \delta U_3 S_{(3)} + \delta U_3 S_{(3)} [I_{r_2} \otimes U_1^\top L_1 U_1 + U_2^\top L_2 U_2 \otimes I_{r_1}] \right) S_{(3)}^\dagger, \\
 \delta \mathbf{S}^\eta &= [U_1^\top L_1 U_1 \delta S_{(1)} + U_1^\top L_1 \delta U_1 S_{(1)}]^{(1)} + [U_2^\top L_2 U_2 \delta S_{(2)} + U_2^\top L_2 \delta U_2 S_{(2)}]^{(2)} \\
 &\quad + [U_3^\top L_3 U_3 \delta S_{(3)} + U_3^\top L_3 \delta U_3 S_{(3)}]^{(3)}.
 \end{aligned}$$

Additionally, the gauge conditions need to be satisfied:

$$(4.6) \quad U_1^\top \delta U_1 = U_2^\top \delta U_2 = U_3^\top \delta U_3 = 0.$$

In order to solve these equations, we will use the first three equations of (4.5), together with (4.6), to substitute  $\delta U_\mu$  in the last equation of (4.5) and determine a decoupled equation for  $\delta \mathbf{S}$ . Rearranging the first equation of (4.5), we obtain

$$\begin{aligned}
 P_{U_1}^\perp \left( L_1 \delta U_1 + \delta U_1 S_{(1)} [I_{r_3} \otimes U_2^\top L_2 U_2 + U_3^\top L_3 U_3 \otimes I_{r_2}] S_{(1)}^\dagger \right) \\
 = \delta U_1^\eta - P_{U_1}^\perp L_1 U_1 \delta S_{(1)} S_{(1)}^\dagger.
 \end{aligned}$$

Vectorization and adhering to (4.6) yield the saddle point system

$$(4.7) \quad \begin{bmatrix} G & I_{r_1} \otimes U_1 \\ I_{r_1} \otimes U_1^\top & 0 \end{bmatrix} \begin{bmatrix} \text{vec}(\delta U_1) \\ y_1 \end{bmatrix} = \begin{bmatrix} b_1 \\ 0 \end{bmatrix},$$

where

$$\begin{aligned}
 G &= I_{r_1} \otimes L_1 + (S_{(1)}^\dagger)^\top (I_{r_3} \otimes U_2^\top L_2 U_2 + U_3^\top L_3 U_3 \otimes I_{r_2}) S_{(1)}^\top \otimes I_{n_1}, \\
 b_1 &= \text{vec}(\delta U_1^\eta) - ((S_{(1)}^\dagger)^\top \otimes P_{U_1}^\perp L_1 U_1) \text{vec}(\delta S_{(1)}),
 \end{aligned}$$

and  $y_1 \in \mathbb{R}^{r_1^2}$  is the dual variable. The positive definiteness of  $L_1$  and the full rank conditions on  $U_1$  and  $\mathbf{S}$  imply that the above system is nonsingular; see, e.g., [8]. Using the Schur complement  $G_S = -(I_{r_1} \otimes U_1)^\top G^{-1} (I_{r_1} \otimes U_1)$ , we obtain the explicit expression

$$(4.8) \quad \text{vec}(\delta U_1) = \left( I_{n_1 r_1} + G^{-1} (I_{r_1} \otimes U_1) G_S^{-1} (I_{r_1} \otimes U_1^\top) \right) G^{-1} b_1 = w_1 - F_1 \text{vec}(\delta S_{(1)})$$

with

$$\begin{aligned}
 w_1 &:= \left( I_{n_1 r_1} + G^{-1} (I_{r_1} \otimes U_1) G_S^{-1} (I_{r_1} \otimes U_1^\top) \right) G^{-1} \text{vec}(\delta U_1^\eta), \\
 F_1 &:= \left( I_{n_1 r_1} + G^{-1} (I_{r_1} \otimes U_1) G_S^{-1} (I_{r_1} \otimes U_1^\top) \right) G^{-1} \left( (S_{(1)}^\dagger)^\top \otimes P_{U_1}^\perp L_1 U_1 \right).
 \end{aligned}$$

Expressions analogous to (4.8) can be derived for the other two factor matrices:

$$\begin{aligned}
 \text{vec}(\delta U_2) &= w_2 - F_2 \text{vec}(\delta S_{(2)}), \\
 \text{vec}(\delta U_3) &= w_3 - F_3 \text{vec}(\delta S_{(3)})
 \end{aligned}$$

with suitable analogues for  $w_2, w_3, F_2$ , and  $F_3$ . These expressions are now inserted into the last equation of (4.5) for  $\delta \mathbf{S}^\eta$ . To this end, define permutation matrices  $\Pi_{i \rightarrow j}$  that map the vectorization of the  $i$ th matricization to the vectorization of the  $j$ th matricization:

$$\Pi_{i \rightarrow j} \text{vec}(\delta \mathbf{S}_{(i)}) = \text{vec}(\delta \mathbf{S}_{(j)}).$$

By definition,  $\text{vec}(\delta S_{(1)}) = \text{vec}(\delta \mathbf{S})$ , and we finally obtain the following linear system for  $\text{vec}(\delta \mathbf{S})$ :

$$(4.9) \quad \begin{aligned} F \text{vec}(\delta \mathbf{S}) = & \text{vec}(\delta \mathbf{S}^\eta) - (S_{(1)}^\top \otimes U_1^\top L_1)w_1 - \Pi_{2 \rightarrow 1}(S_{(2)}^\top \otimes U_2^\top L_2)w_2 \\ & - \Pi_{3 \rightarrow 1}(S_{(3)}^\top \otimes U_3^\top L_3)w_3 \end{aligned}$$

with the  $r_1 r_2 r_3 \times r_1 r_2 r_3$  matrix

$$\begin{aligned} F := & I_{r_2 r_3} \otimes U_1^\top L_1 U_1 - (S_{(1)}^\top \otimes U_1^\top L_1)F_1 \\ & + \Pi_{2 \rightarrow 1} \left[ I_{r_1 r_3} \otimes U_2^\top L_2 U_2 - (S_{(2)}^\top \otimes U_2^\top L_2)F_2 \right] \Pi_{1 \rightarrow 2} \\ & + \Pi_{3 \rightarrow 1} \left[ I_{r_1 r_2} \otimes U_3^\top L_3 U_3 - (S_{(3)}^\top \otimes U_3^\top L_3)F_3 \right] \Pi_{1 \rightarrow 3}. \end{aligned}$$

For small ranks, the linear system (4.9) is solved by forming the matrix  $F$  explicitly and using a direct solver. Since this requires  $O(r_1^3 r_2^3 r_3^3)$  operations, it is advisable to use an iterative solver for larger ranks, in which the Kronecker product structure can be exploited when applying  $F$ ; see also [62]. Once we have obtained  $\delta \mathbf{S}$ , we can easily obtain  $\delta U_1, \delta U_2, \delta U_3$  using (4.8).

*Remark 2.* The application of  $G^{-1}$  needed in (4.8) as well as in the construction of  $G_S$  can be implemented efficiently by noting that  $G$  is the matrix representation of the Sylvester operator  $V \mapsto L_1 V + V \Gamma_1^\top$ , with the matrix

$$\Gamma_1 := (S_{(1)}^\dagger)^\top (I_{r_3} \otimes U_2^\top L_2 U_2 + U_3^\top L_3 U_3 \otimes I_{r_2}) S_{(1)}^\top.$$

The  $r_1 \times r_1$  matrix  $\Gamma_1$  is nonsymmetric but it can be diagonalized by first computing a QR decomposition  $S_{(1)}^\top = Q_S R_S$  such that  $Q_S^\top Q_S = I_{r_1}$  and then computing the spectral decomposition of the symmetric matrix

$$Q_S^\top (I_{r_3} \otimes U_2^\top L_2 U_2 + U_3^\top L_3 U_3 \otimes I_{r_2}) Q_S.$$

After diagonalization of  $\Gamma_1$ , the application of  $G^{-1}$  requires the solution of  $r_1$  linear systems with the matrices  $L_1 + \lambda I$ , where  $\lambda$  is an eigenvalue of  $\Gamma_1$ ; see also [53]. The Schur complement  $G_S \in \mathbb{R}^{r_1^2 \times r_1^2}$  is constructed explicitly by applying  $G^{-1}$  to the  $r_1^2$  columns of  $I_{r_1} \otimes U_1$ . Analogous techniques apply to the computation of  $w_2, F_2$ , and  $w_3, F_3$ .

Assuming, for example, that each  $L_\mu$  is a tridiagonal matrix, the solution of a linear system with the shifted matrix  $L_\mu + \lambda I$  can be performed in  $O(n)$  operations. Therefore, using Remark 2, the construction of the Schur complement  $G_S$  requires  $O(nr^3)$  operations. Hence, the approximate Newton equation (4.4) can be solved in  $O(nr^3 + r^9)$  operations. This cost dominates the complexity of the Riemannian gradient calculation and the retraction step.

**4.2.2. Rank-one structure.** Preconditioners of the form  $I_{n_1} \otimes I_{n_2} \otimes L_1$  with symmetric positive definite  $L_1 \in \mathbb{R}^{n_1 \times n_1}$  play an important role in parameter-dependent problems; see section 5.3. Since this corresponds to setting  $L_2 = 0$  and  $L_3 = 0$ , the derivations above simplify and lead to significantly lower cost.

In particular, it is straightforward to check that the linear system (4.9) simplifies to

$$F \operatorname{vec}(\delta \mathbf{S}) = \operatorname{vec}(\delta \mathbf{S}^\eta) + (S_{(1)}^\top \otimes (U_1^\top L_1^{-1} U_1)^{-1} U_1^\top L_1^{-1}) \operatorname{vec}(\delta U_1^\eta)$$

with

$$F = I_{r_2 r_3} \otimes U_1^\top L_1 U_1 - S_{(1)}^\top (S_{(1)}^\dagger)^\top \otimes (U_1^\top L_1 U_1 - (U_1^\top L_1^{-1} U_1)^{-1}).$$

Apart from  $r_1$  linear system solves and  $r_1$  matrix-vector multiplications with  $L_1$ , the solution of this equation requires  $O(nr^2 + r^4)$  operations, when first using a QR decomposition to reduce  $S_{(1)}^\top$  and then applying standard techniques for solving Sylvester equations (see also Remark 2). Once  $\delta \mathbf{S}$  is known, the computation of  $\delta U_1, \delta U_2, \delta U_3$  can be performed within  $O(nr^2 + r^4)$  operations.

**4.3. The approximate Riemannian Hessian in the TT case.** When using the TT format, it seems to be much harder to solve the approximate Newton equation (4.3) directly and we therefore resort to the preconditioned conjugate gradient (PCG) method for solving the linear system iteratively. As preconditioner, we use a specific parallel subspace correction (PSC) method [63] that generalizes standard block-Jacobi preconditioners. It turns out that by our choice of subspaces we obtain a preconditioner that coincides with a nonmultiplicative version of ALS for the approximate Newton equation.

Before we derive our preconditioner, denoted by  $\hat{\mathcal{B}}$ , let us first examine the approximate Newton equation (4.3) more closely. For  $d$ -dimensional tensors in the TT format, it takes the form

$$(4.10) \quad P_{T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}} \mathcal{B} P_{T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}} \xi = \eta,$$

where

- $\mathbf{X} \in \mathcal{M}_{\mathbf{r}}$  is parametrized by its cores  $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_d$  and is  $d$ -orthogonal;
- the right-hand side  $\eta \in T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}$  is represented in terms of its *gauged* parametrization  $\delta \mathbf{U}_1^\eta, \delta \mathbf{U}_2^\eta, \dots, \delta \mathbf{U}_d^\eta$ , as in (2.11);
- the unknown  $\xi \in T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}$  needs to be determined in terms of its *gauged* parametrization  $\delta \mathbf{U}_1, \delta \mathbf{U}_2, \dots, \delta \mathbf{U}_d$ , again as in (2.11).

When PCG is applied to (4.10) with our preconditioner  $\hat{\mathcal{B}}: T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}} \rightarrow T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}$ , we need to evaluate an expression of the form  $\xi = \hat{\mathcal{B}} \eta$  for a given, arbitrary vector  $\eta \in T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}$ . Again,  $\xi$  and  $\eta$  are represented using the gauged parametrization above.

It is well known that (4.10) does not need to be solved exactly when used to minimize  $f(\mathbf{X})$ . In particular, it suffices [47, Chap. 7.1] to use the stopping criterion

$$\|P_{T_{\mathbf{X}_k} \mathcal{M}_{\mathbf{r}}} [\mathcal{B} \tilde{\xi} - \nabla f(\mathbf{X}_k)]\| \leq \min \left( 0.5, \sqrt{\|P_{T_{\mathbf{X}_k} \mathcal{M}_{\mathbf{r}}} \nabla f(\mathbf{X}_k)\|} \right) \cdot \|P_{T_{\mathbf{X}_k} \mathcal{M}_{\mathbf{r}}} \nabla f(\mathbf{X}_k)\|$$

for accepting the approximation  $\tilde{\xi}$  produced by PCG.

**4.3.1. An overlapping block-Jacobi preconditioner.** PSC methods for (4.10) differ in their choice of splitting  $T_{\mathbf{X}} \mathcal{M}_{\mathbf{r}}$  into subspaces. The most immediate choice is to simply take the direct sum (2.9). The PSC method is then defined in terms of the local operators

$$\mathcal{B}_\mu: \mathcal{V}_\mu \rightarrow \mathcal{V}_\mu, \quad \mathcal{B}_\mu = P^\mu \mathcal{B} P^\mu|_{\mathcal{V}_\mu}, \quad \mu = 1, \dots, d,$$

where  $\mathbf{P}^\mu$  is the orthogonal projector onto  $\mathcal{V}_\mu$ ; see section 2.2. This decomposition leads to a standard block-Jacobi preconditioner; see [55] for a derivation. However, due to the gauging conditions in the definition of  $\mathcal{V}_\mu$ , computing such a preconditioner is considerably more expensive than a PSC method based on the subspaces

$$\widehat{\mathcal{V}}_\mu := \{ \mathbf{X}_{\neq\mu} \text{vec}(\delta\mathbf{U}_\mu) : \delta\mathbf{U}_\mu \in \mathbb{R}^{r_{\mu-1} \times n_\mu \times r_\mu} \} = \text{span } \mathbf{X}_{\neq\mu}, \quad \mu = 1, \dots, d.$$

The corresponding local operators now satisfy

$$\widehat{\mathcal{B}}_\mu : \widehat{\mathcal{V}}_\mu \rightarrow \widehat{\mathcal{V}}_\mu, \quad \widehat{\mathcal{B}}_\mu = \widehat{\mathbf{P}}^\mu \mathcal{B} \widehat{\mathbf{P}}^\mu \Big|_{\widehat{\mathcal{V}}_\mu}, \quad \mu = 1, \dots, d,$$

Since  $\mathcal{B}$  is symmetric and positive definite on  $\mathbb{R}^{n \times n \times \dots \times n}$ , the  $\widehat{\mathcal{B}}_\mu$  are symmetric and positive definite (hence, invertible) on  $\widehat{\mathcal{V}}_\mu$ . This allows us to express the resulting preconditioner as [64, sect. 3.2]

$$(4.11) \quad \widehat{\mathcal{B}} = \sum_{\mu=1}^d \widehat{\mathcal{B}}_\mu^{-1} \widehat{\mathbf{P}}^\mu = \sum_{\mu=1}^d \left( \widehat{\mathbf{P}}^\mu \mathcal{B} \widehat{\mathbf{P}}^\mu \Big|_{\widehat{\mathcal{V}}_\mu} \right)^{-1} \widehat{\mathbf{P}}^\mu,$$

where  $\widehat{\mathbf{P}}^\mu$  is the orthogonal projector onto  $\widehat{\mathcal{V}}_\mu$ . The action of the preconditioner  $\xi = \widehat{\mathcal{B}}\eta$  can be computed as  $\xi = \sum_{\mu=1}^d \xi_\mu$  with

$$(4.12) \quad \widehat{\mathbf{P}}^\mu \mathcal{B} \widehat{\mathbf{P}}^\mu \xi_\mu = \widehat{\mathbf{P}}^\mu \eta, \quad \xi_\mu \in \widehat{\mathcal{V}}_\mu, \quad \mu = 1, \dots, d.$$

*Remark 3.* Observe that  $\mathcal{V}_\mu \subsetneq \widehat{\mathcal{V}}_\mu$  for  $\mu \neq d$ . Hence, the decomposition  $T_{\mathbf{X}}\mathcal{M}_{\mathbf{r}} = \cup_{\mu=1}^d \widehat{\mathcal{V}}_\mu$  is no longer a direct sum as in (2.9) and one can indeed regard  $\widehat{\mathcal{B}}$  as an *overlapping* block-Jacobi preconditioner.

It remains to explain how to solve the local equations (4.12). By definition of  $\widehat{\mathcal{V}}_\mu$ , we have  $\xi_\mu = \mathbf{X}_{\neq\mu} \text{vec}(\widehat{\delta\mathbf{U}}_\mu)$  for some  $\widehat{\delta\mathbf{U}}_\mu$ , which will generally differ from the gauged  $\delta\mathbf{U}_\mu$  parametrization of  $\xi$ . Likewise  $\widehat{\mathbf{P}}^\mu \eta = \mathbf{X}_{\neq\mu} \text{vec}(\widehat{\delta\mathbf{U}}_\mu^\eta)$  for some  $\widehat{\delta\mathbf{U}}_\mu^\eta$ . The local problem (4.12) can therefore be written as

$$\widehat{\mathbf{P}}^\mu \mathcal{B} \mathbf{X}_{\neq\mu} \text{vec}(\widehat{\delta\mathbf{U}}_\mu) = \mathbf{X}_{\neq\mu} \text{vec}(\widehat{\delta\mathbf{U}}_\mu^\eta), \quad \mu = 1, \dots, d.$$

Since  $\mathbf{X}$  is  $d$ -orthogonal, we can substitute the projector in this equation by

$$(4.13) \quad \begin{aligned} \widehat{\mathbf{P}}^\mu &= \mathbf{X}_{\neq\mu} (\mathbf{X}_{\neq\mu}^\top \mathbf{X}_{\neq\mu})^{-1} \mathbf{X}_{\neq\mu}^\top \\ &= \mathbf{X}_{\neq\mu} [(\mathbf{X}_{\geq\mu+1}^\top \mathbf{X}_{\geq\mu+1})^{-1} \otimes I_{n_\mu} \otimes I_{r_{\mu-1}}] \mathbf{X}_{\neq\mu}^\top, \end{aligned}$$

where we used (2.8). From this, we obtain a standard system of equations:

$$(4.14) \quad \mathbf{X}_{\neq\mu}^\top \mathcal{B} \mathbf{X}_{\neq\mu} \text{vec}(\widehat{\delta\mathbf{U}}_\mu) = [(\mathbf{X}_{\geq\mu+1}^\top \mathbf{X}_{\geq\mu+1}) \otimes I_{n_\mu} \otimes I_{r_{\mu-1}}] \text{vec}(\widehat{\delta\mathbf{U}}_\mu^\eta).$$

When  $n_\mu$  and  $r_\mu$  are small, one can simply solve for  $\widehat{\delta\mathbf{U}}_\mu$  by a direct method. For larger problems, as we will consider in the numerical experiments, explicitly computing  $\mathbf{X}_{\neq\mu}^\top \mathcal{B} \mathbf{X}_{\neq\mu}$  can be prohibitively expensive. Fortunately, when  $\mathcal{B}$  has the structure of the Laplacian (1.4), we show in the next section how to efficiently solve the local problems in the large-scale setting.

In both cases, we have obtained the solution as an ungauged parametrization:

$$(4.15) \quad \xi = \widehat{\mathcal{B}}\eta = \sum_{\mu=1}^d \mathbf{X}_{\neq\mu} \text{vec}(\delta\widehat{\mathbf{U}}_{\mu}).$$

To obtain the gauged parametrization  $\delta\mathbf{U}_{\mu}$  of  $\xi$  satisfying (2.11), we can simply apply (2.13) to compute  $P_{T_{\mathbf{X}}\mathcal{M}_r}(\xi)$  and exploit that  $\xi$  is a TT tensor (with doubled TT ranks compared to  $\mathbf{X}$ ).

**4.3.2. Laplacian-like structure.** We now elaborate on how to solve the local problems (4.14) for the case of  $\mathcal{B}$  having a Laplacian-like structure. Recall from (1.4),

$$\mathcal{B} = I_{n_d} \otimes \cdots \otimes I_{n_2} \otimes L_1 + I_{n_d} \otimes \cdots \otimes I_{n_3} \otimes L_2 \otimes I_{n_1} \\ + \cdots + L_d \otimes I_{n_{d-1}} \otimes \cdots \otimes I_1$$

with symmetric and positive definite matrices  $L_{\mu} \in \mathbb{R}^{n_{\mu} \times n_{\mu}}$  and identity matrices  $I_{n_{\mu}}$ . The main idea is to avoid forming  $\mathbf{X}_{\neq\mu}^{\top} \mathcal{B} \mathbf{X}_{\neq\mu}$  explicitly.

*Rank-structured form of local problems.* Using (2.8), the application of the Laplacian-like operator  $\mathcal{B}$  to  $\mathbf{X}_{\neq\mu}$  can be decomposed into three parts,

$$\mathbf{X}_{\neq\mu}^{\top} \mathcal{B} \mathbf{X}_{\neq\mu} = \mathcal{B}_{\geq\mu+1} \otimes I_{n_{\mu}} \otimes I_{r_{\mu-1}} + \mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1} \otimes L_{\mu} \otimes I_{r_{\mu-1}} \\ + \mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1} \otimes I_{n_{\mu}} \otimes \mathcal{B}_{\leq\mu-1}$$

with the reduced leading and trailing terms, respectively,

$$\mathcal{B}_{\leq\mu-1} = \mathbf{X}_{\leq\mu-1}^{\top} \left( \sum_{\nu=1}^{\mu-1} I_{n_{\mu-1}} \otimes \cdots \otimes L_{\nu} \otimes \cdots \otimes I_{n_1} \right) \mathbf{X}_{\leq\mu-1}$$

and

$$\mathcal{B}_{\geq\mu+1} = \mathbf{X}_{\geq\mu+1}^{\top} \left( \sum_{\nu=\mu+1}^d I_{n_d} \otimes \cdots \otimes L_{\nu} \otimes \cdots \otimes I_{n_{\mu+1}} \right) \mathbf{X}_{\geq\mu+1}.$$

Some manipulation establishes the identity

$$(\mathbf{X}_{\neq\mu}^{\top} \mathcal{B} \mathbf{X}_{\neq\mu} \text{vec}(\delta\widehat{\mathbf{U}}_{\mu}))^{\mathbf{L}} = \delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} \mathcal{B}_{\geq\mu+1} \\ + (L_{\mu} \otimes I_{r_{\mu-1}} + I_{n_{\mu}} \otimes \mathcal{B}_{\leq\mu-1}) \delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} \mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1}.$$

Hence, (4.14) can be written as

$$(4.16) \quad \delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} \mathcal{B}_{\geq\mu+1} (\mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1})^{-1} + (L_{\mu} \otimes I_{r_{\mu-1}} + I_{n_{\mu}} \otimes \mathcal{B}_{\leq\mu-1}) \delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} = (\delta\widehat{\mathbf{U}}_{\mu}^{\eta})^{\mathbf{L}}.$$

*Efficient solution of local problems.* Since  $\mathcal{B}_{\geq\mu+1}$  and  $\mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1}$  are symmetric positive definite, they admit a generalized eigenvalue decomposition: There is an invertible matrix  $Q$  such that  $\mathcal{B}_{\geq\mu+1} Q = (\mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1}) Q \Lambda$  with  $\Lambda$  diagonal and  $Q^{\top} (\mathbf{X}_{\geq\mu+1}^{\top} \mathbf{X}_{\geq\mu+1}) Q = I_{r_{\mu}}$ . This transforms (4.16) into

$$\delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} Q^{-\top} \Lambda + (L_{\mu} \otimes I_{r_{\mu}} + I_{n_{\mu}} \otimes \mathcal{B}_{\leq\mu-1}) \delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} Q^{-\top} = (\delta\widehat{\mathbf{U}}_{\mu}^{\eta})^{\mathbf{L}} Q^{-\top}.$$

Setting  $\widetilde{\delta\mathbf{U}}_{\mu}^{\mathbf{L}} = \delta\widehat{\mathbf{U}}_{\mu}^{\mathbf{L}} Q^{-\top}$  and  $(\widetilde{\delta\mathbf{U}}_{\mu}^{\eta})^{\mathbf{L}} = (\delta\widehat{\mathbf{U}}_{\mu}^{\eta})^{\mathbf{L}} Q^{-\top}$ , we can write these equations columnwise as

$$G_{\mu,i} \widetilde{\delta\mathbf{U}}_{\mu}^{\mathbf{L}}(:, i) = (\widetilde{\delta\mathbf{U}}_{\mu}^{\eta})^{\mathbf{L}}(:, i)$$

with the system matrix

$$(4.17) \quad G_{\mu,i} = \lambda_i I_{n_\mu} \otimes I_{r_\mu} + L_\mu \otimes I_{r_\mu} + I_{n_\mu} \otimes \mathcal{B}_{\leq \mu-1}, \quad \lambda_i = \Lambda(i, i).$$

*Remark 4.* Similar to Remark 2 and by vectorizing  $V$ , the matrix  $G_{\mu,i}$  represents the Sylvester operator

$$V \mapsto (L_\mu + \lambda_i I_{n_\mu})V + V \mathcal{B}_{\leq \mu-1}.$$

Hence, after diagonalization of  $\mathcal{B}_{\leq \mu-1}$ , the application of  $G_{\mu,i}^{-1}$  requires the solution of  $r_\mu$  linear systems with the matrices  $L_\mu + (\lambda_i + \beta)I_{n_\mu}$ , where  $\beta$  is an eigenvalue of  $\mathcal{B}_{\leq \mu-1}$ .

After forming  $\widehat{\delta \mathbf{U}}_\mu^L = \widetilde{\delta \mathbf{U}}_\mu^L Q^T$  for all  $\mu$ , we have then obtained the solution as (4.15). Assuming again that solving with  $L_\mu + (\lambda_i + \beta)I_{n_\mu}$  can be performed in  $O(n_\mu)$  operations, we end up with a total computational complexity of  $O(dnr^3)$  for applying  $\widehat{\mathcal{B}}$ .

*Remark 5.* By  $\mu$ -orthogonalizing  $\mathbf{X}$  and transforming  $\widehat{\delta \mathbf{U}}_\mu$ , as described in [54], the Gram matrix  $\mathbf{X}_{\geq \mu+1}^T \mathbf{X}_{\geq \mu+1}$  in (4.16) becomes the identity matrix. This leads to a more stable calculation of the corresponding unknown  $\widehat{\delta \mathbf{U}}_\mu$ ; see also Remark 1. We make use of this transformation in our implementations.

**4.3.3. Rank-one structure.** Once again, the equations simplify significantly for the special case  $\mathcal{B} = I_{n_d} \otimes \cdots \otimes I_{n_2} \otimes L_1$ . For  $\mu = 1$ , we obtain from (4.16) that  $L_1 \widehat{\delta \mathbf{U}}_1^L = (\widehat{\delta \mathbf{U}}_1^\eta)^L$ . For  $\mu > 1$ , we obtain

$$(I_{n_\mu} \otimes \mathbf{X}_{\leq \mu-1}^T (I_{n_{\mu-1}} \otimes \cdots \otimes I_{n_2} \otimes L_1) \mathbf{X}_{\leq \mu-1}) \widehat{\delta \mathbf{U}}_\mu^L = (\widehat{\delta \mathbf{U}}_\mu^\eta)^L.$$

Note that this assumes  $\mu$ -orthogonality of  $\mathbf{X}$ ; see Remark 5.

**4.3.4. Connection to ALS.** The overlapping block-Jacobi preconditioner  $\widehat{\mathcal{B}}$  defined in (4.11) is closely related to ALS applied to (1.2). There are, however, crucial differences explaining why  $\widehat{\mathcal{B}}$  is significantly cheaper per iteration than ALS.

Using  $\text{vec}(\mathbf{X}) = \mathbf{X}_{\neq \mu} \text{vec}(\mathbf{U}_\mu)$ , one microstep of ALS fixes  $\mathbf{X}_{\neq \mu}$  and replaces  $\mathbf{U}_\mu$  by the minimizer of (see, e.g., [24, Alg. 1])

$$\min_{\mathbf{U}_\mu} \frac{1}{2} \langle \mathbf{X}_{\neq \mu} \text{vec}(\mathbf{U}_\mu), \mathcal{A} \mathbf{X}_{\neq \mu} \text{vec}(\mathbf{U}_\mu) \rangle - \langle \mathbf{X}_{\neq \mu} \text{vec}(\mathbf{U}_\mu), \text{vec}(\mathbf{F}) \rangle.$$

After  $\mathbf{U}_\mu$  has been updated, ALS proceeds to the next core until all cores have eventually been updated in a particular order, for example,  $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_d$ . The solution to the above minimization problem is obtained from solving the ALS subproblem

$$(4.18) \quad \mathbf{X}_{\neq \mu}^T \mathcal{A} \mathbf{X}_{\neq \mu} \text{vec}(\mathbf{U}_\mu) = \mathbf{X}_{\neq \mu}^T \text{vec}(\mathbf{F}).$$

When  $\mathbf{X}$  is  $\mu$ -orthogonal,  $\mathbf{X}_{\geq \mu+1}^T \mathbf{X}_{\geq \mu+1} = I_{r_\mu}$  and the ALS subproblem has the same form as the subproblem (4.14) for  $\widehat{\mathcal{B}}$ . However, there are crucial differences:

- ALS directly optimizes for the cores and as such uses  $\mathcal{A}$  in the optimization problem. The approximate Newton method, on the other hand, updates (all) the cores using a search direction obtained from minimizing the quadratic model (4.3). It can therefore use any positive definite approximation  $\mathcal{B}$  to construct this model.



- ALS updates each core immediately and it is a block version of *nonlinear* Gauss–Seidel for the nonlinear problem (1.2). On the other hand,  $\hat{\mathcal{B}}$  updates all cores simultaneously and it constitutes a block version of *linear* Jacobi for the linear system (4.10). In addition, the linear problem allows us to use PCG as a cheap and locally optimal acceleration method.
- Even in the large-scale setting of  $n_\mu \gg 10^3$ , the subproblems (4.14) can be solved efficiently in closed form as long as  $L_\mu + \lambda I_{n_\mu}$  allows for efficient system solves, e.g., for tridiagonal  $L_\mu$ . This is not possible in ALS, where the subproblems have to be formulated with  $\mathcal{A}$  and typically need to be solved iteratively using PCG.

*Remark 6.* Instead of PSC, we experimented with a symmetrized version of a successive subspace correction (SSC) preconditioner, also known as a back and forth ALS sweep. However, the higher computational cost per iteration of SSC was not offset by a possibly improved convergence behavior.

**5. Numerical experiments.** In this section, we compare the performance of the different preconditioned optimization techniques discussed in this paper for three representative test cases.

We have implemented all algorithms in MATLAB. For the TT format, we used the TTeMPS toolbox; see <http://anchp.epfl.ch/TTeMPS>. All numerical experiments and timings are performed on a 12-core Intel Xeon X5675, 3.07 GHz, 192 GB RAM using MATLAB 2014a, running on Linux kernel 3.2.0-0.

To simplify the discussion, we assume throughout this section that the tensor size and ranks are equal along all modes and therefore state them as scalar values:  $n = \max_\mu n_\mu$  and  $r = \max_\mu r_\mu$ .

**5.1. Test case 1: Newton potential.** As a standard example leading to a linear system of the form (1.5), we consider the PDE

$$\begin{aligned} -\Delta u(x) + V(x) &= f(x), & x \in \Omega = (-10, 10)^d, \\ u(x) &= 0, & x \in \partial\Omega, \end{aligned}$$

with the Laplace operator  $\Delta$ , the *Newton potential*  $V(x) = \|x\|^{-1}$ , and the source function  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ . Equations of this type are used to describe the energy of a charged particle in an electrostatic potential.

We discretize the domain  $\Omega$  by a uniform tensor grid with  $n^d$  grid points and corresponding mesh width  $h$ . Then, by finite difference approximation on this tensor grid, we obtain a tensor equation of the type (1.1), where the linear operator  $\mathcal{A}$  is the sum of the  $d$ -dimensional Laplace operator as in (1.4) with  $L_\mu = \frac{1}{h^2} \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{n \times n}$ , and the discretized Newton potential  $\mathbf{V}$ . To create a low-rank representation of the Newton potential,  $V(x)$  is approximated by a rank 10 tensor  $\mathbf{V}$  using exponential sums [19]. The application of  $\mathcal{A}$  to a tensor  $\mathbf{X}$  is given by

$$\mathcal{A}\mathbf{X} = \mathcal{L}\mathbf{X} + \mathbf{V} \circ \mathbf{X},$$

where  $\circ$  denotes the Hadamard (elementwise) product. The application of this operator increases the ranks significantly: If  $\mathbf{X}$  has rank  $r$ , then  $\mathcal{A}\mathbf{X}$  has rank  $(2 + 10)r = 12r$ .

**5.2. Test case 2: Anisotropic diffusion equation.** As a second example, we consider the anisotropic diffusion equation

$$\begin{aligned} -\text{div}(D\nabla u(x)) &= f(x), & x \in \Omega = (-10, 10)^d, \\ u(x) &= 0, & x \in \partial\Omega, \end{aligned}$$

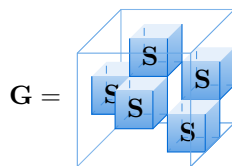


FIG. 1. Structure of the core tensor  $\mathbf{G}$  for the case  $d = 3$  resulting from an application of the anisotropic diffusion operator.

with a tridiagonal diffusion matrix  $D = \text{tridiag}(\alpha, 1, \alpha) \in \mathbb{R}^{d \times d}$ . The discretization on a uniform tensor grid with  $n^d$  grid points and mesh width  $h$  yields a linear equation with system matrix  $A = L + V$  consisting of the potential term

$$V = I_n \otimes \cdots \otimes I_n \otimes B_2 \otimes 2\alpha B_1 + I_n \otimes \cdots \otimes I_n \otimes B_3 \otimes 2\alpha B_2 \otimes I_n \\ + \cdots + B_d \otimes 2\alpha B_{d-1} \otimes I_n \otimes \cdots \otimes I_n$$

and the Laplace part  $L$  defined as in the previous example. The matrix  $B_\mu = \frac{1}{2h} \text{tridiag}(-1, 0, 1) \in \mathbb{R}^{n \times n}$  represents the one-dimensional central finite difference matrix for the first derivative.

The corresponding linear operator  $\mathcal{A}$  acting on  $\mathbf{X} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$  can be represented as a TT operator of rank three, with the cores given by

$$A_1(i_1, j_1) = [L_1(i_1, j_1) \quad 2\alpha B_1(i_1, j_1) \quad I_{n_1}(i_1, j_1)], \quad A_d(i_d, j_d) = \begin{bmatrix} I_{n_d}(i_d, j_d) \\ B_d(i_d, j_d) \\ L_d(i_d, j_d) \end{bmatrix},$$

and

$$A_\mu(i_\mu, j_\mu) = \begin{bmatrix} I_{n_\mu}(i_\mu, j_\mu) & 0 & 0 \\ B_\mu(i_\mu, j_\mu) & 0 & 0 \\ L_\mu(i_\mu, j_\mu) & 2\alpha B_\mu(i_\mu, j_\mu) & I_{n_\mu}(i_\mu, j_\mu) \end{bmatrix}, \quad \mu = 2, \dots, d-1.$$

In the Tucker format, this operator is also of rank three. Given a tensor  $\mathbf{X}$  in the representation (2.1), the result  $\mathbf{Y} = \mathcal{A}\mathbf{X}$  is explicitly given by  $\mathbf{Y} = \mathbf{G} \times_1 V_1 \times_2 \cdots \times_d V_d$  with

$$V_\mu = [L_\mu U_\mu \quad U_\mu \quad B_\mu U_\mu] \in \mathbb{R}^{n \times 3r_\mu}$$

and core tensor  $\mathbf{G} \in \mathbb{R}^{3r_1 \times \cdots \times 3r_d}$  which has a block structure shown in Figure 1 for the case  $d = 3$ .

The rank of  $\mathcal{A}$  increases linearly with the band width of the diffusion matrix  $D$ . For example, a pentadiagonal structure would yield an operator of rank 4. See also [26] for more general bounds in terms of certain properties of  $D$ .

**5.3. Test case 3: Parametric PDE.** We consider the so-called cookie problem [36, 7], an elliptic PDE on  $\Omega = (0, 1)^2$  of the form

$$(5.1) \quad \begin{aligned} -\operatorname{div}(a(x, p)\nabla u(x, p)) &= 1, \quad x \in \Omega, \\ u(x, p) &= 0, \quad x \in \partial\Omega. \end{aligned}$$

The piecewise constant coefficient  $a(x, p) \in \mathbb{R}$  depends on a parameter vector  $p \in \mathbb{R}^9$ , with  $p_1, \dots, p_9 \in [0, 10]$ , as follows. We place nine disks  $\Omega_{s,t}$  with radius  $\rho = \frac{1}{14}$  and

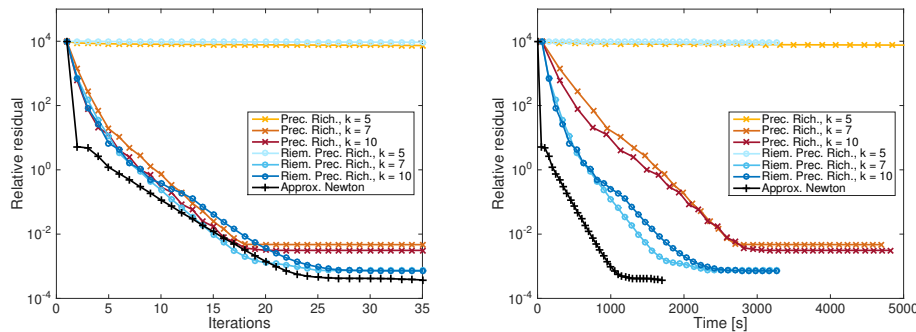


FIG. 2. *Newton potential with  $d = 3$ . Comparison of truncated preconditioned Richardson, truncated Riemannian preconditioned Richardson, and the approximate Newton scheme when applied to the Newton potential in the Tucker format. For the Richardson iterations, exponential sum approximations with  $k \in \{5, 7, 10\}$  terms are compared. Left: Relative residual versus number of iterations. Right: Relative residual versus execution time.*

midpoints  $(\rho(4s - 1), \rho(4t - 1))$ ,  $s, t = 1, \dots, 3$ , inside  $\Omega$ . We then define

$$a(x, p) := \begin{cases} 1 + p_\mu & \text{if } x \in \Omega_{s,t}, \mu = 3(t - 1) + s, \\ 1 & \text{otherwise.} \end{cases}$$

As described, e.g., in [36], discretizing (5.1) in space (using linear finite elements with  $m$  degrees of freedom) and sampling each parameter  $p_\mu$  in  $n_\mu$  equally spaced points in  $[0, 10]$  leads to a linear system, for which the solution vector can be reshaped into a tensor  $\mathbf{U} \in \mathbb{R}^{m \times n_1 \times \dots \times n_9}$ . The coefficient matrix takes the form

$$A = I_{n_9} \otimes \dots \otimes I_{n_1} \otimes A_0 + \sum_{\mu=1}^9 I_{n_9} \otimes \dots \otimes I_{n_{\mu+1}} \otimes D_\mu \otimes I_{n_{\mu-1}} \otimes \dots \otimes I_{n_1} \otimes A_\mu,$$

where  $A_0$  corresponds to the stiffness matrix for  $\Delta$  on the whole domain  $\Omega$ , each  $A_\mu$  corresponds to the stiffness matrix on one of the disks, and each  $D_\mu = \text{diag}(p_\mu^1, \dots, p_\mu^{n_\mu})$  contains the samples of  $p_\mu$ . The application of the corresponding linear operator  $\mathcal{A}$  to a tensor of rank  $\mathbf{r}$  results in a tensor of rank  $10\mathbf{r}$ . Setting the heat coefficient to 1 on the entire domain  $\Omega$  yields the simple rank-one preconditioner

$$B = I_{n_9} \otimes \dots \otimes I_{n_1} \otimes A_0.$$

**5.4. Results for the Tucker format.** For tensors represented in the Tucker format we want to investigate the convergence of the truncated preconditioned Richardson (3.5) and its Riemannian variant (3.6) and compare them to the approximate Newton scheme discussed in section 4.2. Figure 2 displays the obtained results for the first test case, the Newton potential, where we set  $d = 3$ ,  $n = 100$ , and used multilinear ranks  $r = 15$ . Figure 3 displays the results for the second test case, the anisotropic diffusion operator with  $\alpha = \frac{1}{4}$ , using the same settings. In both cases, the right-hand side is given by a random rank-one Tucker tensor. Here and in the following, we call a tensor *random* if the entries of  $U_\mu$  and  $\mathbf{S}$  are chosen independently from the uniform distribution on  $[0, 1]$ . To create a full space preconditioner for both Richardson approaches, we approximate the inverse Laplacian by an exponential sum

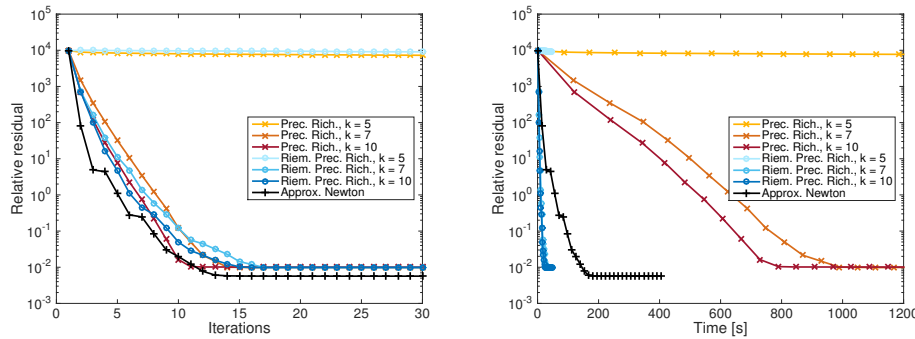


FIG. 3. Anisotropic diffusion with  $d = 3$ . Comparison of truncated preconditioned Richardson, truncated Riemannian preconditioned Richardson, and the approximate Newton scheme when applied to the anisotropic diffusion in the Tucker format. For the Richardson iterations, exponential sum approximations with  $k \in \{5, 7, 10\}$  terms are compared. Left: Relative residual versus number of iterations. Right: Relative residual versus execution time.

of  $k \in \{5, 7, 10\}$  terms. It can be clearly seen that the quality of the preconditioner has a strong influence on the convergence. For  $k = 5$ , convergence is extremely slow. Increasing  $k$  yields a drastic improvement on the convergence.

With an accurate preconditioner, the truncated Richardson scheme converges fast with regard to the number of iterations but suffers from very long computation times due to the exceedingly high intermediate ranks. In comparison, the Riemannian Richardson scheme yields similar convergence speed, but with significantly reduced computation time due to the additional projection into the tangent space. The biggest saving in computational effort comes from relation (3.7) which allows us to avoid having to form the preconditioned residual  $\mathcal{P}^{-1}(\mathbf{F} - \mathcal{A}\mathbf{X}_k)$  explicitly, a quantity with very high rank. Note that for both Richardson approaches, it is necessary to round the Euclidean gradient to lower rank using a tolerance of, say,  $10^{-5}$  before applying the preconditioner to avoid excessive intermediate ranks.

The approximate Newton scheme converges equally well as the best Richardson approaches with regard to the number of iterations and does not require setting up a preconditioner. For the first test case, it only needs about half of the time as the best Richardson approach. For the second test case, it is significantly slower than Riemannian preconditioned Richardson. Since this operator is of lower rank than the Newton potential, the additional complexity of constructing the approximate Hessian does not pay off in this case.

*Quadratic convergence.* In Figure 4 we investigate the convergence of the approximate Newton scheme when applied to a pure Laplace operator,  $A = L$ , and to the anisotropic diffusion operator  $A = L + V$ . In order to have an exact solution of known rank  $r = 4$ , we construct the right-hand side by applying  $A$  to a random rank 4 tensor. For the dimension and tensor size we have chosen  $d = 3$  and  $n = 200$ , respectively. By construction, the exact solution lies on the manifold. Hence, if the approximate Newton method converges to this solution, we have zero residual and our Gauss–Newton approximation of (4.2) is an exact second-order model despite only containing the  $A$  term. In other words, we expect quadratic convergence when  $A = L$  but only linear when  $A = L + V$  since our approximate Newton method (4.3) only solves with  $L$ . This is indeed confirmed in Figure 4.

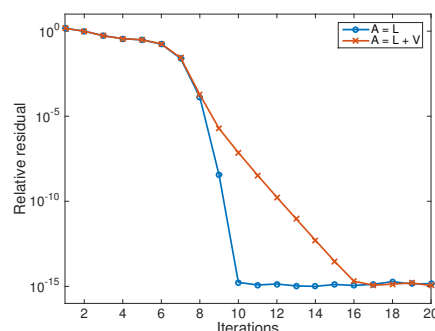


FIG. 4. Convergence of the approximate Newton method for the zero-residual case when applied to a pure Laplace operator  $L$  and to the anisotropic diffusion operator  $L + V$ .

**5.5. Results for the TT format.** In the TT format, we compare the convergence of our approximate Newton scheme (with the overlapping block-Jacobi preconditioner described in section 4.3.1) to standard approaches: ALS and AMEn from [15]. Both ALS and AMEn require the solution of subproblems (4.18) with the dense matrix  $\mathbf{X}_{\neq\mu}\mathcal{A}\mathbf{X}_{\neq\mu}$  of size  $r_{\mu-1}n_{\mu}r_{\mu} \times r_{\mu-1}n_{\mu}r_{\mu}$ . To solve it efficiently, we use PCG with the same preconditioner  $\mathbf{X}_{\mu}^{\top}\mathcal{B}\mathbf{X}_{\mu}$  that is used in the approximate Newton scheme.

As in [33], we noticed in our experiments that a *preconditioned residual* step generally improves the convergence of AMEn. In particular, the new iterate is obtained as

$$\text{vec}(\tilde{\mathbf{X}}) = \text{vec}(\mathbf{X}) - \alpha \mathbf{X}_{\mu,\mu+1} \mathcal{B}_{\mu,\mu+1}^{-1} \mathbf{X}_{\mu,\mu+1}^{\top} \mathcal{T}(\text{vec}(\mathbf{F}) - \mathcal{A} \text{vec}(\mathbf{X})).$$

The matrix  $\mathcal{B}_{\mu,\mu+1}^{-1} = (\mathbf{X}_{\neq\mu,\mu+1} \mathcal{B} \mathbf{X}_{\neq\mu,\mu+1})^{-1}$  represents a preconditioner for the DMRG two-core system  $\mathbf{X}_{\neq\mu,\mu+1} \mathcal{A} \mathbf{X}_{\neq\mu,\mu+1}$  and can be efficiently applied in a similar way as described in sections 4.3.2 and 4.3.3. In addition, as is customary (see [15, sect. 6.2]),  $\mathcal{T}(\text{vec}(\mathbf{F}) - \mathcal{A} \text{vec}(\mathbf{X}))$  indicates that the high-rank residual is truncated to a fixed rank by TT-SVD (we use ranks 4 and 8).

*Laplacian-like structure.* In this setting, we have chosen  $d = 60$ ,  $n = 100$ , and a random rank-one right-hand side of norm one. In the first test case, the Newton potential, we have chosen TT ranks  $r = 10$  for the approximate solution. The corresponding convergence curves are shown in Figure 5. We observe that the approximate Newton scheme needs significantly less time to converge than the ALS and AMEn schemes. As a reference, we have also included a steepest descent method using the overlapping block-Jacobi scheme directly as a preconditioner for every gradient step instead of using it to solve the approximate Newton equation (4.10). The additional effort of solving the Newton equation approximately clearly pays off.

In Figure 6, we show results for the anisotropic diffusion case. To obtain a good accuracy of the solution, we have to choose a relatively high rank of  $r = 25$  in this case. Here, the approximate Newton scheme is still faster, especially at the beginning of the iteration, but the final time needed to reach a residual of  $10^{-4}$  is similar to ALS.

Note that in Figures 5 and 6 the plots with regard to the number of iterations are to be read with care due to the different natures of the algorithms. One ALS or AMEn iteration corresponds to the optimization of one core. In the plots, the beginning of each half-sweep of ALS or AMEn is denoted by a circle. To assess the performance of both schemes as fairly as possible, we have taken considerable care to

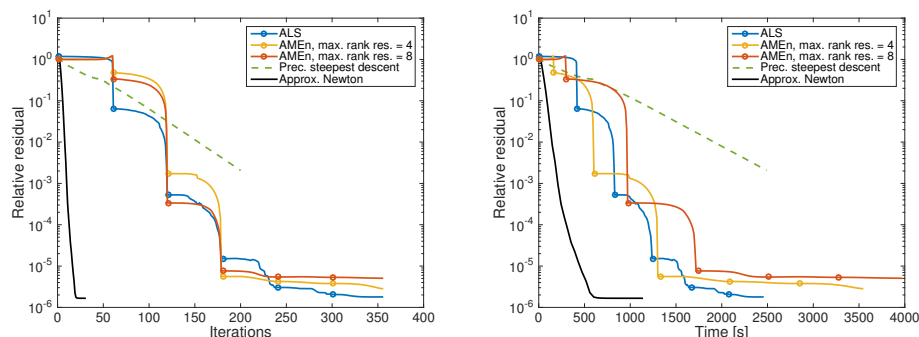


FIG. 5. *Newton potential with  $d = 60$ . Convergence of ALS and AMEn compared to preconditioned steepest descent with overlapping block-Jacobi as preconditioner and the approximate Newton scheme. Left: Relative residual versus number of iterations. Right: Relative residual versus execution time.*

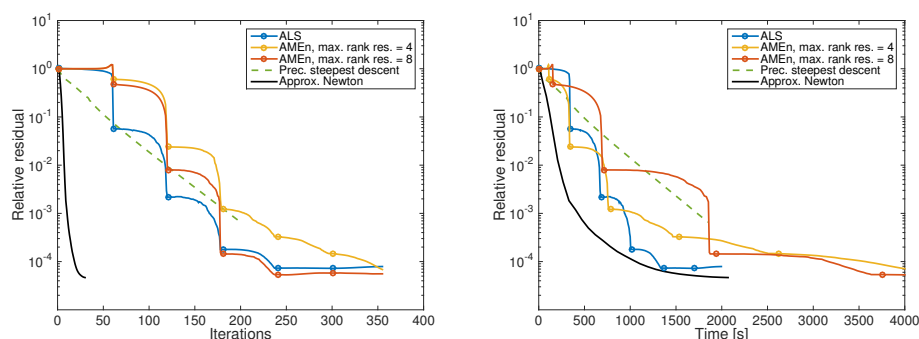


FIG. 6. *Anisotropic diffusion with  $d = 60$ . Convergence of ALS and AMEn compared to preconditioned steepest descent with overlapping block-Jacobi as preconditioner and the approximate Newton scheme. Left: Relative residual versus number of iterations. Right: Relative residual versus execution time.*

provide the same level of optimization to the implementations of ALS, AMEn, and the approximate Newton scheme.

*Rank-one structure.* For the third test case, we chose a Galerkin discretization of the unit square using quadrilateral elements and piecewise linear basis function, leading to Galerkin matrices  $A_\mu$ ,  $\mu = 0, \dots, 9$  of size  $2796 \times 2796$ . The parameters  $p_\mu$  are discretized using 50 equally spaced points on  $[0, 10]$ . As also observed in [36], the resulting solution tensor is of very high rank. We have used  $\mathbf{r} = (1, 100, 80, 60, 40, 30, 30, 30, 10, 1)$  to obtain an accuracy of at least  $10^{-4}$ . The convergence of both ALS and the approximate Newton scheme is shown in Figure 7. It turns out that approximate Newton needs fewer iterations to reach the final accuracy, but this advantage is offset by the cheaper and simpler ALS procedure which is around three times faster.

*Mesh-dependence of the preconditioner.* To investigate how the performance of the preconditioner depends on the mesh width of the discretization, we look again at the anisotropic diffusion operator and measure the convergence as the mesh width  $h$  and therefore the tensor size  $n \in \{60, 120, 180, 240, 360, 420, 480, 540, 600\}$  changes by one order of magnitude. As in the test for quadratic convergence, we construct the right-hand side by applying  $A$  to a random rank-three tensor.

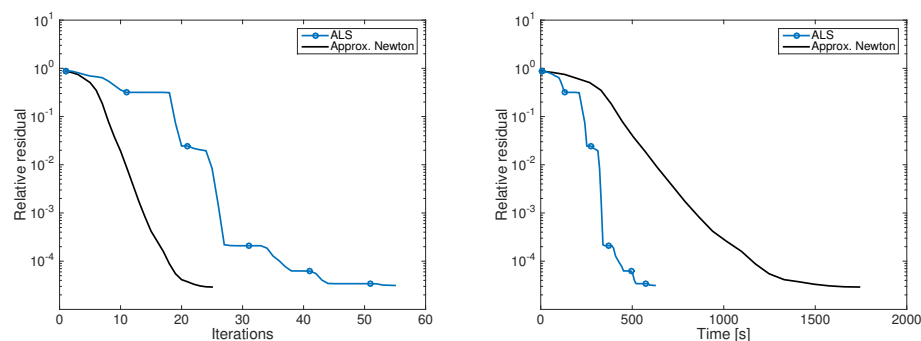


FIG. 7. *Parameter-dependent PDE. Convergence of ALS compared to preconditioned steepest descent with overlapping block-Jacobi as preconditioner and the approximate Newton scheme. Left: Relative residual versus number of iterations. Right: Relative residual versus execution time.*

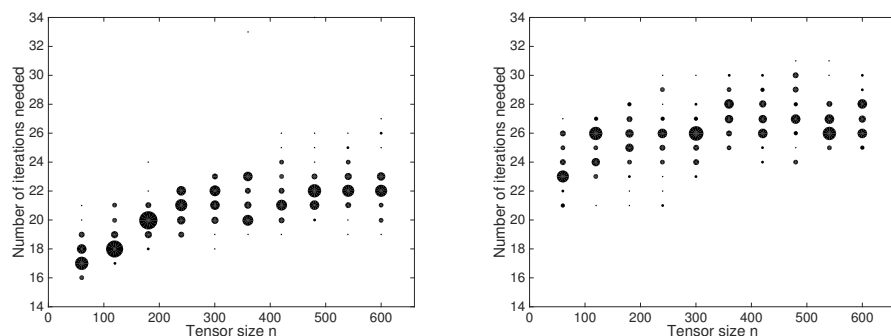


FIG. 8. *Number of iterations that the proposed approximate Newton scheme needs to reach a relative residual of  $10^{-6}$  for different mesh widths  $h = 1/n$ . The solution is of rank  $r = 3$ . We perform 30 runs for each size. The radii of the circles correspond to the number of runs achieving this number of iterations. Left: Dimension  $d = 10$ . Right: Dimension  $d = 30$ .*

To measure the convergence, we take the number of iterations needed to converge to a relative residual of  $10^{-6}$ . For each tensor size, we perform 30 runs with random starting guesses of rank  $r = 3$ . The result is shown in Figure 8, where circles are drawn for each combination of size  $n$  and number of iterations needed. The radius of each circle denotes how many runs have achieved a residual of  $10^{-6}$  for this number of iterations.

On the left plot of Figure 8 we see the results of dimension  $d = 10$ , whereas on the right plot we have  $d = 30$ . We see that the number of iterations needed to converge changes only mildly as the mesh width varies over one order of magnitude. In addition, the dependence on  $d$  is also not very large.

**5.6. Rank-adaptivity.** Note that in many applications, rank-adaptivity of the algorithm is a desired property. For the Richardson approach, this would result in replacing the fixed-rank truncation with a tolerance-based rounding procedure. In the alternating optimization, this would lead to the DMRG or AMEn algorithm. In the framework of Riemannian optimization, rank-adaptivity can be introduced by successive runs of increasing rank, using the previous solution as a warm start for the next rank. For a recent discussion of this approach, see [60]. A basic example

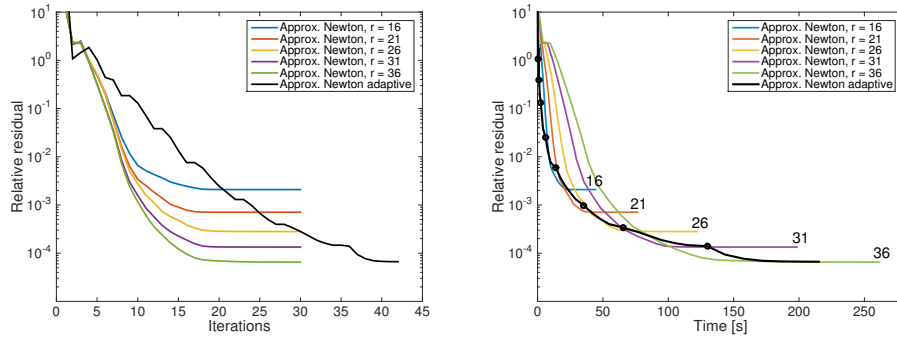


FIG. 9. Rank-adaptivity for approximate Newton applied to the anisotropic diffusion equation with  $n = 100$ ,  $d = 10$ . Starting from rank one, the rank is increased by 5 after stagnation is detected. Each rank increase is denoted by a black circle. The other curves show the convergence when running approximate Newton directly with the target rank. Left: Relative residual versus number of iterations. Right: Relative residual versus execution time.

of introducing rank-adaptivity to the approximate Newton scheme is shown in Figure 9. Starting from ranks  $r^{(0)} = 1$ , we run the approximate Newton scheme and use this result to warm start the algorithm with ranks  $r^{(i)} = r^{(i-1)} + 5$ . At each rank, we run the approximate Newton scheme until stagnation is detected: Let  $\xi^{(i-1)}$  be the Riemannian gradient after optimizing at rank  $r^{(i-1)}$ . Then, we run the approximate Newton scheme until  $\|\xi^{(i)}\| < \frac{1}{4}\|\xi^{(i-1)}\|$  holds for the current Riemannian gradient  $\xi^{(i)} := P_{T_{\mathbf{X}_k} \mathcal{M}_{r^{(i)}}} \nabla f(\mathbf{X}_k)$ . The result is compared to the convergence of approximate Newton when starting directly with the target rank  $r^{(i)}$ . Although the adaptive rank scheme is not faster for a desired target rank, it offers more flexibility when we want to instead prescribe a desired accuracy. For a relative residual of  $10^{-3}$ , the adaptive scheme needs about half the time than for using the (too large) rank  $r = 36$ .

Note that in the case of tensor completion, rank-adaptivity becomes a crucial ingredient to avoid overfitting and to steer the algorithm in the right direction; see, e.g., [61, 33, 56, 60, 54]. For difficult completion problems, careful core-by-core rank increases become necessary. Here, for linear systems, such a core-by-core strategy does not seem to be necessary, as the algorithms will converge even if we directly optimize using rank  $r = 36$ . This is likely due to the preconditioner which acts globally over all cores.

**6. Conclusions.** We have investigated different ways of introducing preconditioning into Riemannian gradient descent. As a simple but effective approach, we have seen the Riemannian truncated preconditioned Richardson scheme. Another approach used second-order information by means of approximating the Riemannian Hessian. In the Tucker case, the resulting approximate Newton equation could be solved efficiently in closed form, whereas in the TT case, we have shown that this equation can be solved iteratively in a very efficient way using PCG with an overlapping block-Jacobi preconditioner. The numerical experiments show favorable performance of the proposed algorithms when compared to standard non-Riemannian approaches, such as truncated preconditioned Richardson and ALS. The advantages of the approximate Newton scheme become especially pronounced in cases when the linear operator is expensive to apply, e.g., the Newton potential.



**Acknowledgments.** The authors would like to thank the referees for their valuable input and Jonas Ballani for providing the matrices in section 5.3.

## REFERENCES

- [1] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [2] P.-A. ABSIL, R. MAHONY, AND J. TRUMPF, *An extrinsic look at the Riemannian Hessian*, in Geometric Science of Information, F. Nielsen and F. Barbaresco, eds., of Lecture Notes in Comput. Sci. 8085, Springer, Berlin, 2013, pp. 361–368, [http://dx.doi.org/10.1007/978-3-642-40020-9\\_39](http://dx.doi.org/10.1007/978-3-642-40020-9_39).
- [3] P.-A. ABSIL AND J. MALICK, *Projection-like retractions on matrix manifolds*, SIAM J. Control Optim., 22 (2012), pp. 135–158.
- [4] I. AFFLECK, T. KENNEDY, E. H. LIEB, AND H. TASAKI, *Rigorous results on valence-bond ground states in antiferromagnets*, Phys. Rev. Lett., 59 (1987), pp. 799–802.
- [5] M. BACHMAYR AND W. DAHMEN, *Adaptive near-optimal rank tensor approximation for high-dimensional operator equations*, Found. Comput. Math., 15 (2015), pp. 839–898.
- [6] J. BALLANI AND L. GRASEDYCK, *A projection method to solve linear systems in tensor format*, Numer. Linear Algebra Appl., 20 (2013), pp. 27–43, <http://dx.doi.org/10.1002/nla.1818>.
- [7] J. BALLANI AND L. GRASEDYCK, *Hierarchical Tensor Approximation of Output Quantities of Parameter-Dependent PDEs*, Tech. report, ANCHP, EPF Lausanne, Switzerland, 2014.
- [8] M. BENZI, G. H. GOLUB, AND J. LIESEN, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.
- [9] H. G. BOCK, *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*, Bonner Math. Schriften, University of Bonn, 1987.
- [10] N. BOUMAL AND P.-A. ABSIL, *RTRMC: A Riemannian trust-region method for low-rank matrix completion*, in Proceedings of the Neural Information Processing Systems Conference (NIPS), 2011.
- [11] C. DA SILVA AND F. J. HERRMANN, *Hierarchical Tucker tensor optimization—applications to tensor completion*, Linear Algebra Appl., 481 (2015), pp. 131–173.
- [12] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.
- [13] S. V. DOLGOV, *TT-GMRES: Solution to a linear system in the structured tensor format*, Russian J. Numer. Anal. Math. Modelling, 28 (2013), pp. 149–172, <http://arxiv.org/abs/1206.5512>.
- [14] S. V. DOLGOV AND I. V. OSELEDETS, *Solution of linear systems and matrix inversion in the TT-format*, SIAM J. Sci. Comput., 34 (2012), pp. A2718–A2739, <http://dx.doi.org/10.1137/110833142>.
- [15] S. V. DOLGOV AND D. V. SAVOSTYANOV, *Alternating minimal energy methods for linear systems in higher dimensions*, SIAM J. Sci. Comput., 36 (2014), pp. A2248–A2271, <http://dx.doi.org/10.1137/140953289>.
- [16] L. GRASEDYCK, *Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure*, Computing, 72 (2004), pp. 247–265, <http://dx.doi.org/10.1007/s00607-003-0037-z>.
- [17] L. GRASEDYCK, *Hierarchical singular value decomposition of tensors*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2029–2054, <http://dx.doi.org/10.1137/090764189>.
- [18] L. GRASEDYCK, D. KRESSNER, AND C. TOBLER, *A literature survey of low-rank tensor approximation techniques*, GAMM-Mitt., 36 (2013), pp. 53–78.
- [19] W. HACKBUSCH, *Entwicklungen nach Exponentialsummen*, Technical Report 4/2005, MPI MIS Leipzig, 2010; also available online from <http://www.mis.mpg.de/preprints/tr/report-0405.pdf>.
- [20] W. HACKBUSCH, *Tensor Spaces and Numerical Tensor Calculus*, Springer, Heidelberg, 2012, <http://dx.doi.org/10.1007/978-3-642-28027-6>.
- [21] W. HACKBUSCH AND B. N. KHOROMSKIJ, *Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators. I. Separable approximation of multi-variate functions*, Computing, 76 (2006), pp. 177–202, <http://dx.doi.org/10.1007/s00607-005-0144-0>.
- [22] W. HACKBUSCH AND S. KÜHN, *A new scheme for the tensor representation*, J. Fourier Anal. Appl., 15 (2009), pp. 706–722, <http://dx.doi.org/10.1007/s00041-009-9094-9>.
- [23] J. HAEGEMAN, M. MARIËN, T. J. OSBORNE, AND F. VERSTRAETE, *Geometry of matrix product states: Metric, parallel transport and curvature*, J. Math. Phys., 55 (2014).

- [24] S. HOLTZ, T. ROHWEDDER, AND R. SCHNEIDER, *The alternating linear scheme for tensor optimization in the tensor train format*, SIAM J. Sci. Comput., 34 (2012), pp. A683–A713, <http://dx.doi.org/10.1137/100818893>.
- [25] S. HOLTZ, T. ROHWEDDER, AND R. SCHNEIDER, *On manifolds of tensors of fixed TT-rank*, Numer. Math., 120 (2012), pp. 701–731, <http://dx.doi.org/10.1007/s00211-011-0419-7>.
- [26] V. KAZEYEV, O. REICHMANN, AND C. SCHWAB, *Low-rank tensor structure of linear diffusion operators in the TT and QTT formats*, Linear Algebra Appl., 438 (2013), pp. 4204–4221.
- [27] B. N. KHOROMSKIJ AND I. V. OSELEDETS, *Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDEs*, Comput. Methods Appl. Math., 10 (2010), pp. 376–394; also available online from <http://www.degruyter.com/view/j/cmam.2010.10.issue-4/cmam-2010-0023/cmam-2010-0023.xml?format=INT>.
- [28] B. N. KHOROMSKIJ, I. V. OSELEDETS, AND R. SCHNEIDER, *Efficient Time-Stepping Scheme for Dynamics on TT-Manifolds*, Technical Report 24, MPI MIS Leipzig, 2012; also available online from <http://www.mis.mpg.de/preprints/2012/preprint2012.24.pdf>.
- [29] B. N. KHOROMSKIJ AND C. SCHWAB, *Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs*, SIAM J. Sci. Comput., 33 (2011), pp. 364–385, <http://dx.doi.org/10.1137/100785715>.
- [30] O. KOCH AND C. LUBICH, *Dynamical tensor approximation*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2360–2375.
- [31] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.
- [32] D. KRESSNER, M. PLESINGER, AND C. TOBLER, *A preconditioned low-rank CG method for parameter-dependent Lyapunov matrix equations*, Numer. Linear Algebra Appl., 21 (2014), pp. 666–684, <http://dx.doi.org/10.1002/nla.1919>.
- [33] D. KRESSNER, M. STEINLECHNER, AND A. USCHMAJEV, *Low-rank tensor methods with subspace correction for symmetric eigenvalue problems*, SIAM J. Sci. Comput., 36 (2014), pp. A2346–A2368, <http://dx.doi.org/10.1137/130949919>.
- [34] D. KRESSNER, M. STEINLECHNER, AND B. VANDEREYCKEN, *Low-rank tensor completion by Riemannian optimization*, BIT, 54 (2014), pp. 447–468.
- [35] D. KRESSNER AND C. TOBLER, *Krylov subspace methods for linear systems with tensor product structure*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1688–1714, <http://dx.doi.org/10.1137/090756843>.
- [36] D. KRESSNER AND C. TOBLER, *Low-rank tensor Krylov subspace methods for parametrized linear systems*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1288–1316, <http://dx.doi.org/10.1137/100799010>.
- [37] D. KRESSNER AND C. TOBLER, *Preconditioned low-rank methods for high-dimensional elliptic PDE eigenvalue problems*, Comput. Methods Appl. Math., 11 (2011), pp. 363–381.
- [38] D. KRESSNER AND C. TOBLER, *Algorithm 941: htucker—a MATLAB toolbox for tensors in hierarchical Tucker format*, ACM Trans. Math. Software, 40 (2014).
- [39] C. LUBICH, *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*, Zur. Lect. Adv. Math., European Mathematical Society, Zürich, 2008, <http://dx.doi.org/10.4171/067>.
- [40] C. LUBICH, I. V. OSELEDETS, AND B. VANDEREYCKEN, *Time integration of tensor trains*, SIAM J. Numer. Anal., 53 (2015), pp. 917–941, <http://dx.doi.org/10.1137/140976546>.
- [41] D. G. LUENBERGER, *The gradient projection method along geodesics*, Management Sci., 18 (1970), pp. 620–631.
- [42] B. MISHRA, G. MEYER, S. BONNABEL, AND R. SEPULCHRE, *Fixed-rank matrix factorizations and Riemannian low-rank optimization*, Comput. Statist., 29 (2014), pp. 591–621, <http://dx.doi.org/10.1007/s00180-013-0464-z>.
- [43] B. MISHRA AND R. SEPULCHRE, *R3MC: A Riemannian three-factor algorithm for low-rank matrix completion*, in Proceedings of the 53rd Annual Conference on Decision and Control, IEEE, 2014, pp. 1137–1142, <http://dx.doi.org/10.1109/CDC.2014.7039534>.
- [44] B. MISHRA AND R. SEPULCHRE, *Riemannian preconditioning*, SIAM J. Optim., 26 (2016), pp. 635–660.
- [45] Y. NESTEROV, *Introductory Lectures on Convex Optimization: A Basic Course*, Appl. Optim. 87, Kluwer Academic Publishers, Boston, MA, 2004, <http://dx.doi.org/10.1007/978-1-4419-8853-9>.
- [46] T. NGO AND Y. SAAD, *Scaled gradients on Grassmann manifolds for matrix completion*, in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds., Curran Associates, 2012, pp. 1412–1420.
- [47] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, Springer Ser. Oper. Res., 2nd ed., Springer, New York, 2006.

- [48] I. V. OSELEDETS, *DMRG approach to fast linear algebra in the TT-format*, Comput. Methods Appl. Math, 11 (2011), pp. 382–393.
- [49] I. V. OSELEDETS, *Tensor-train decomposition*, SIAM J. Sci. Comput., 33 (2011), pp. 2295–2317, <http://dx.doi.org/10.1137/090752286>.
- [50] H. RAUHUT, R. SCHNEIDER, AND Z. STOJANAC, *Tensor completion in hierarchical tensor representations*, in Compressed Sensing and Its Applications: MATHEON Workshop 2013, H. Boche, R. Calderbank, G. Kutyniok, and J. Vybřal, eds., Appl. Numer. Harmon. Anal., Birkhäuser, New York, 2015, pp. 419–450.
- [51] T. ROHWEDDER AND A. USCHMAJEV, *On local convergence of alternating schemes for optimization of convex problems in the tensor train format*, SIAM J. Numer. Anal., 51 (2013), pp. 1134–1162, <http://dx.doi.org/10.1137/110857520>.
- [52] U. SCHOLLWÖCK, *The density-matrix renormalization group in the age of matrix product states*, Ann. Physics, 326 (2011), pp. 96–192, <http://dx.doi.org/10.1016/j.aop.2010.09.012>.
- [53] V. SIMONCINI, *Computational methods for linear matrix equations*, SIAM Rev., to appear; available online from <http://www.dm.unibo.it/~simoncin/list.html>.
- [54] M. STEINLECHNER, *Riemannian Optimization for High-Dimensional Tensor Completion*, Technical Report MATHICSE 5.2015, EPF Lausanne, Switzerland, 2015.
- [55] M. STEINLECHNER, *Riemannian Optimization for Solving High-Dimensional Problems with Low-Rank Tensor Structure*, Ph.D. thesis, EPF Lausanne, 2015.
- [56] M. TAN, I. TSANG, L. WANG, B. VANDEREYCKEN, AND S. PAN, *Riemannian pursuit for big matrix recovery*, in Proceedings of ICML 2014, vol. 32, 2014, pp. 1539–1547; also available online from <http://jmlr.org/proceedings/papers/v32/tan14.html>.
- [57] L. TUCKER, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311, <http://dx.doi.org/10.1007/BF02289464>.
- [58] A. USCHMAJEV, *Zur Theorie der Niedrigrangapproximation in Tensorprodukten von Hilberträumen*, Ph.D. thesis, Technische Universität Berlin, 2013.
- [59] A. USCHMAJEV AND B. VANDEREYCKEN, *The geometry of algorithms using hierarchical tensors*, Linear Algebra Appl., 439 (2013), pp. 133–166.
- [60] A. USCHMAJEV AND B. VANDEREYCKEN, *Greedy rank updates combined with Riemannian descent methods for low-rank optimization*, in Proceedings of International Conference on Sampling Theory and Applications, IEEE, 2015, pp. 420–424.
- [61] B. VANDEREYCKEN, *Low-rank matrix completion by Riemannian optimization*, SIAM J. Optim., 23 (2013), pp. 1214–1236.
- [62] B. VANDEREYCKEN AND S. VANDEWALLE, *A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2553–2579, <http://dx.doi.org/10.1137/090764566>.
- [63] J. XU, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.
- [64] J. XU, *The method of subspace corrections*, J. Comput. Appl. Math., 128 (2001), pp. 335–362.