KRYLOV SUBSPACE METHODS FOR LINEAR SYSTEMS WITH TENSOR PRODUCT STRUCTURE[*]

DANIEL KRESSNER[†] AND CHRISTINE TOBLER[†]

Abstract. The numerical solution of linear systems with certain tensor product structures is considered. Such structures arise, for example, from the finite element discretization of a linear PDE on a d-dimensional hypercube. Linear systems with tensor product structure can be regarded as linear matrix equations for $d = 2$ and appear to be their most natural extension for $d > 2$. A standard Krylov subspace method applied to such a linear system suffers from the curse of dimensionality and has a computational cost that grows exponentially with d. The key to breaking the curse is to note that the solution can often be very well approximated by a vector of low tensor rank. We propose and analyse a new class of methods, so called *tensor Krylov subspace methods*, which exploit this fact and attain a computational cost that grows linearly with d.

1. Introduction. This paper is concerned with certain linear systems that can be written as the sum of d Kronecker products of matrices. More specifically, we consider for $d = 2$,

$$\big(A_1 \otimes I_{n_2} + I_{n_1} \otimes A_2\big)x = b_1 \otimes b_2, \tag{1.1}$$

and for $d = 3$,

$$\big(A_1 \otimes I_{n_2} \otimes I_{n_3} + I_{n_1} \otimes A_2 \otimes I_{n_3} + I_{n_1} \otimes I_{n_2} \otimes A_3\big)x = b_1 \otimes b_2 \otimes b_3, \tag{1.2}$$

where $A_s \in \mathbb{R}^{n_s \times n_s}$, $b_s \in \mathbb{R}^{n_s}$, and I_{n_s} denotes the $n_s \times n_s$ identity matrix. For general $d \in \mathbb{N}$, the linear system takes the form

$$\mathcal{A}x = b, \tag{1.3}$$

with

$$\mathcal{A} = \sum_{s=1}^{d} I_{n_1} \otimes \cdots \otimes I_{n_{s-1}} \otimes A_s \otimes I_{n_{s+1}} \otimes \cdots \otimes I_{n_d}, \tag{1.4}$$

$$b = b_1 \otimes \cdots \otimes b_d. \tag{1.5}$$

Classical Krylov subspace methods for solving linear systems, such as conjugate gradient or GMRES, are not well suited for solving (1.3). To illustrate this, let us consider the case of constant dimensions, $n_s \equiv n$. Then every vector in the Krylov subspace basis has length n^d and a single scalar product requires $2n^d$ operations. The purpose of this paper is to develop Krylov subspace methods having computational costs and memory requirements that scale *linearly*, rather than *exponentially*, in d.

The following model problem from [9] shall illustrate the type of applications leading to (1.3). Consider the partial differential equation

$$-\triangle u = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0, \tag{1.6}$$

where $\Omega = [0,1]^d$ is the d-dimensional hypercube. Choosing a tensorized finite element basis for discretzing the variational formulation of (1.6) yields the mass and stiffness matrices

$$\mathcal{M} = M_1 \otimes \cdots \otimes M_d, \quad \mathcal{B} = \sum_{s=1}^{d} M_1 \otimes \cdots \otimes M_{s-1} \otimes B_s \otimes M_{s+1} \otimes \cdots \otimes M_d.$$

Hence, $\mathcal{A} = \mathcal{M}^{-1/2}\mathcal{B}\mathcal{M}^{-1/2}$ is of the form (1.4) with $A_s = (M_s)^{-1/2}B_s(M_s)^{-1/2}$. If f is separable, $f = f_1(y_1)f_2(y_2)\cdots f_d(y_d)$, then the discretized right hand side takes the form (1.5). Otherwise,

the function f might be well approximated by a short sum of separable functions, in which case the solution of the discretized equation can be obtained from linear systems of the type (1.3) by superposition. This is possible, for example, if f is sufficiently smooth and d is moderate, see, e.g., [4, 5, 9].

For $d = 2$, the equation (1.1) can be reformulated as follows:

$$A_1 X + X A_2^\top = b_1 (b_2)^\top, \tag{1.7}$$

where $x = \text{rowvec}(X)$, with the rowvec operator stacking the rows of a matrix $X \in \mathbb{R}^{n_1 \times n_2}$ into a single column vector $x \in \mathbb{R}^{n_1 \cdot n_2}$. The linear matrix equation (1.7) is usually called *Sylvester equation*, which has been studied quite intensively, often motivated by applications in systems and control theory. In fact, most results and algorithms presented in this paper are already known for $d = 2$. In particular, several variants of Krylov subspace algorithms for solving (1.7) have been developed and analysed, see [16, 17, 20, 22, 23]. The novelty of our work is in the extension to $d > 2$; we will point out relevant connections to the case $d = 2$ whenever suitable. A notable exception is the convergence bound for extended Krylov subspace methods we give in Section 6; this result addresses an open question for the case $d = 2$.

Grasedyck [9] has combined an integral representation of the solution x to (1.3) with quadrature based on sinc interpolation [25] to show that x can be well approximated by vectors of low tensor rank and to develop a numerical algorithm that scales linearly with d. To the best of our knowledge, this was the first and so far the only algorithm for efficiently approximating x for high dimensions. Somewhat a drawback, the algorithm relies on computing matrix exponentials of scalar multiples of A_s, which might become expensive for larger matrices. In contrast, the approach proposed in this paper solely relies on matrix-vector multiplications with A_s. If available, matrix-vector products with $(A_s)^{-1}$ can be used to speed up convergence. A variant of Grasedyck's algorithm is still invoked for solving smaller subsystems.

The rest of this paper is organized as follows. Section 2 contains some preliminary results, mainly concerning tensor notation and approximations of low tensor rank to the solution of (1.3). In Section 3, we will describe the newly proposed tensor Krylov subspace method and discuss some implementation details, such as the efficient computation of the residual. The convergence of this method is analysed in Section 4 for the (symmetric and non-symmetric) positive definite case. Section 5 provides a discussion on solving the compressed systems needed in the course of the tensor Krylov subspace method. In Section 6, we propose an extension of the tensor Krylov subspace method, which is suitable if matrix-vector products not only with A_s but also with $(A_s)^{-1}$ can be performed. Section 7 contains some numerical experiments with academic examples to illustrate the theoretical results obtained in this paper. Finally, some conclusions are outlined in Section 8.

2. Preliminaries. The following lemma is a consequence of well-known properties of the Kronecker product [15].

LEMMA 2.1. *Consider the matrix $\mathcal{A}$ defined in (1.4). Then $\Lambda(\mathcal{A})$, the set of eigenvalues of $\mathcal{A}$, is given by all possible sums of eigenvalues of $A_1, A_2, \ldots, A_d$:*

$$\Lambda(\mathcal{A}) = \left\{ \lambda_1 + \lambda_2 + \cdots \lambda_d : \ \lambda_s \in \Lambda(A_s) \right\}. \tag{2.1}$$

The linear system (1.3) has a unique solution if and only if $\Lambda(\mathcal{A})$ contains no zero eigenvalues, which – by Lemma 2.1 – is equivalent to

$$\lambda_1 + \lambda_2 + \cdots + \lambda_d \neq 0, \quad \forall \lambda_s \in \Lambda(A_s). \tag{2.2}$$

In the case of the Sylvester equation (1.7), this corresponds to the well-known condition $\Lambda(A_1) \cap \Lambda(-A_2) = \emptyset$.

We recall that a non-symmetric matrix A is called *positive definite* if its symmetric part $(A + A^\top)/2$ is positive definite. By Lemma 2.1, the matrix $\mathcal{A}$ is positive definite if and only if

$$\xi_1 + \xi_2 + \cdots + \xi_d > 0, \quad \forall \xi_s \in \Lambda(A_s + A_s^\top)/2. \tag{2.3}$$

The following lemma recalls an integral representation of the solution x from [9].

LEMMA 2.2. *If $\mathcal{A}$ is positive definite then the solution of the linear system (1.3) admits the representation*

$$x = -\int_0^\infty \big(\exp(-tA_1)b_1 \otimes \cdots \otimes \exp(-tA_d)b_d \big) \, \mathrm{d}t.$$

Note that Lemma 2.2 still holds if we impose the less restrictive condition that the eigenvalues of $\mathcal{A}$ have positive real part, but this is not needed for the developments of our paper.

2.1. Tensor arithmetic and decompositions. This section provides a brief overview of tensor arithmetic concepts needed in the rest of the paper. We refer to the recent survey [19] for more details.

A d-way tensor $v \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is an element of the tensor product of the vector spaces $\mathbb{R}^{n_1}, \mathbb{R}^{n_2}, \ldots, \mathbb{R}^{n_d}$ for fixed integers $n_1, \ldots, n_d$. The coordinates of v (with respect to a choice of bases) form a multi-dimensional array. The element at the multi-index $\mathfrak{I} = (i_1, i_2, \ldots, i_n)$ in such an array is denoted by $v_{\mathfrak{I}}$. A tensor can be represented as a vector in $\mathbb{R}^{n_1 n_2 \cdots n_d}$ by simply stacking the elements $v_{\mathfrak{I}}$ in lexicographical order. *In the following, we will identify tensors with their vector representations.* For $d = 2$, this means that a matrix $A \in \mathbb{R}^{n_1 \times n_2}$ is identified with the vector $\mathrm{rowvec}(A) \in \mathbb{R}^{n_1 n_2}$, where rowvec stacks the transposed rows of A on top of each other. This identification of tensors with vectors is unambiguous as soon as the order d and the dimensions $n_1, \ldots, n_d$ are fixed.

A tensor $v \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is of *tensor rank one* if its vector representation can be written as a Kronecker product of d vectors:

$$v = v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(d)}, \quad v^{(s)} \in \mathbb{R}^{n_s}. \tag{2.4}$$

Note that this implies that the element at the multi-index $\mathfrak{I} = (i_1, i_2, \ldots i_n)$ takes the form $v_{\mathfrak{I}} = v_{i_1}^{(1)} v_{i_2}^{(2)} \cdots v_{i_d}^{(d)}$.

A tensor is called *supersymmetric* if it is invariant under any permutations of the indices. For matrices, supersymmetry coincides with the usual notion of symmetry. The following lemma generalizes a well-known result on the symmetry of solutions to Lyapunov matrix equations.

LEMMA 2.3. *Consider the linear system $\mathcal{A}x = b$, where $\mathcal{A}$ takes the form (1.4) with constant coefficients $A = A_1 = \cdots = A_d$. If $\mathcal{A}$ is invertible and b represents a supersymmetric tensor then the solution x is also the representation of a supersymmetric tensor.*

Proof. Assume that x is the solution of $\mathcal{A}x = b$ and that x is not supersymmetric. Then there is a permutation $\pi : \{1, \ldots, d\} \to \{1, \ldots, d\}$ such that the vector x_π, the representation of the tensor with elements $x_{\pi(\mathfrak{I})}$ for every multi-index $\mathfrak{I}$, is different from x. The structure of $\mathcal{A}$ implies $\mathcal{A}x_\pi = b_\pi$. Since b is supersymmetric, $\mathcal{A}x_\pi = b_\pi = b$ contradicting the unique solvability of $\mathcal{A}x = b$. $\square$

The *CANDECOMP/PARAFAC (CP) decomposition* represents a tensor as a sum of rank one tensors. In vector language, this means

$$v = \sum_{i=1}^k v_i^{(1)} \otimes v_i^{(2)} \otimes \cdots \otimes v_i^{(d)}, \quad v_i^{(s)} \in \mathbb{R}^{n_s}. \tag{2.5}$$

If v admits a representation (2.5) then we say that v has *tensor rank at most k*. In our context we do not need the concept of exact tensor rank, which is much more subtle than the usual rank concept for matrices. Defining the $n_s \times k$ matrices $V_s = \big[v_1^{(s)}, v_2^{(s)}, \ldots, v_k^{(s)}\big]$, a more compact way of writing (2.5) is

$$v = [\![V_1, V_2, \ldots, V_d]\!]$$

The *Tucker decomposition* is another popular tensor decomposition. For an integer tuple $\mathfrak{K} = (k_1, \ldots, k_d)$ it takes the form

$$v = \sum_{\mathfrak{I} \leq \mathfrak{K}} c_{\mathfrak{I}} \, v_{i_1}^{(1)} \otimes v_{i_2}^{(2)} \otimes \cdots \otimes v_{i_d}^{(d)} =: [\![c; V_1, V_2, \ldots, V_d]\!], \tag{2.6}$$

3

where the sum is taken over all multi-indices $\mathfrak{I} = (i_1, \ldots, i_d)$ that are elementwise not larger than $\mathfrak{K}$. It is worth emphasizing that $V_s = \left[v_1^{(s)}, v_2^{(s)}, \ldots, v_{k_s}^{(s)}\right]$ is now an $n_s \times k_s$ matrix, i.e., the number of columns of V_s may vary with s. The $k_1 \times \cdots \times k_d$ tensor formed from the elements $c_{\mathfrak{I}}$ is called the *core tensor*. Note that the CP decomposition (2.5) is a special case of (2.6) for constant $k_s \equiv k$ and an "identity" core tensor that is zero except for $c_{i,i,\ldots,i} = 1$ for $i = 1, \ldots, k$. However, the Tucker format usually assume the matrices V_s to be column-orthogonal, which is not the case for the CP format.

Alternatively, the Tucker decomposition can be written as

$$v = \left(V_1 \otimes \cdots \otimes V_d\right) c, \tag{2.7}$$

where c is to be understood as the vector representation of the core tensor in (2.6). This also reveals that v is in the subspace spanned by $V_1 \otimes V_2 \otimes \cdots \otimes V_d$.

Notation 2.4. *We write the multi-dimensional Kronecker product as $\bigotimes_{s=1}^{d} v_s := v_1 \otimes \cdots \otimes v_d$. Note that the order in which the index s is evaluated is important, as the Kronecker product does not commute.*

2.2. Low tensor rank approximations. Solving (1.3) for larger d requires to work with a data sparse representation of x. For this purpose, x will be approximated by a low rank tensor. The following theorem provides a fundamental connection between approximations of x by low rank tensors and separable approximations to the reciprocal of a sum of d variables.

THEOREM 2.5. *Consider the linear system (1.3) with coefficient matrices $A_s \in \mathbb{R}^{n_s \times n_s}$, assume that the system is solvable and that the matrices A_s are diagonalizable: $P_s^{-1} A_s P_s = \Lambda_s$ with invertible P_s and diagonal Λ_s. Let $f_j^{(s)} : \Omega_s \to \mathbb{C}$, $j = 1, \ldots, k$, be analytic functions such that Ω_s contains the eigenvalues of A_s and*

$$\left\| \frac{1}{\mu_1 + \mu_2 + \cdots + \mu_d} - \sum_{i=1}^{k} f_i^{(1)}(\mu_1) f_i^{(2)}(\mu_2) \cdots f_i^{(d)}(\mu_d) \right\|_{\Omega} \leq \epsilon(k), \tag{2.8}$$

where $\| \cdot \|_{\Omega}$ denotes the supremum norm on $\Omega = \Omega_1 \times \cdots \times \Omega_d$ Then there is a rank-k tensor x_k such that

$$\|x - x_k\|_2 \leq \kappa_{\mathcal{P}} \, \epsilon(k) \, \|b\|_2,$$

where $\kappa_{\mathcal{P}} = \kappa_2(P_1) \kappa_2(P_2) \cdots \kappa_2(P_d)$ and $\kappa_2(\cdot)$ denotes the 2-norm condition number of a matrix.

Proof. By a similarity transformation with the matrix $\mathcal{P} = P_1 \otimes \cdots \otimes P_d$, the linear system (1.3) is transformed into

$$\left(\sum_{s=1}^{d} I_{n_1} \otimes \cdots \otimes I_{n_{s-1}} \otimes \Lambda_s \otimes I_{n_{s+1}} \otimes \cdots \otimes I_{n_d} \right) \tilde{x} = \tilde{b}.$$

with $\tilde{x} = \mathcal{P}^{-1} x$ and $\tilde{b} = \mathcal{P}^{-1} b$. This is a diagonal linear system and the entry of the solution $\tilde{x}$ at the multi-index $\mathfrak{I} = (i_1, \ldots, i_d)$ is given by

$$\tilde{x}_{\mathfrak{I}} = \frac{\tilde{b}_{\mathfrak{I}}}{\lambda_{i_1}^{(1)} + \lambda_{i_2}^{(2)} + \cdots + \lambda_{i_d}^{(d)}},$$

where $\lambda_{i_1}^{(1)} + \cdots + \lambda_{i_d}^{(d)} \neq 0$ from Lemma 2.1. Similarly, if we define the rank-k tensor

$$x_k = \sum_{j=1}^{k} f_j^{(1)}(A_1) b_1 \otimes f_j^{(2)}(A_2) b_2 \otimes \cdots \otimes f_j^{(d)}(A_d) b_d, \tag{2.9}$$

the entry of the correspondingly transformed tensor $\tilde{x}_k = \mathcal{P}^{-1} x_k$ at $\mathfrak{I}$ is given by

$$\tilde{b}_{\mathfrak{I}} \, f_j^{(1)}\left(\lambda_{i_1}^{(1)}\right) f_j^{(2)}\left(\lambda_{i_2}^{(2)}\right) \cdots f_j^{(d)}\left(\lambda_{i_d}^{(d)}\right).$$

4

Hence, with $\mathfrak{K} = (k, \ldots, k)$,

$$\|\tilde{x} - \tilde{x}_k\|_2^2 = \sum_{\mathfrak{I} \leq \mathfrak{K}} |\tilde{b}_{\mathfrak{I}}|^2 \left| \frac{1}{\lambda_{i_1}^{(1)} + \lambda_{i_2}^{(2)} + \cdots + \lambda_{i_d}^{(d)}} - \sum_{j=1}^{k} f_j^{(1)}\big(\lambda_{i_1}^{(1)}\big) \cdots f_j^{(d)}\big(\lambda_{i_d}^{(d)}\big) \right|^2$$
$$\leq \epsilon(k)^2 \sum_{\mathfrak{I} \leq \mathfrak{K}} |\tilde{b}_{\mathfrak{I}}|^2 = \epsilon(k)^2 \|\tilde{b}\|_2^2.$$

Combining this bound with $\|x - x_k\|_2 = \|\mathcal{P}(\tilde{x} - \tilde{x}_k)\|_2 \leq \|\mathcal{P}\|_2 \|\tilde{x} - \tilde{x}_k\|_2$, $\|\tilde{b}\|_2 \leq \|\mathcal{P}^{-1}\|_2 \|b\|_2$ and $\kappa_{\mathcal{P}} = \|\mathcal{P}\|_2 \|\mathcal{P}^{-1}\|_2$ yields the statement of the theorem. $\square$

Theorem 2.5 provides an upper bound on the error for the best approximation of x by a rank-k tensor. To be practically useful, we still need to address the approximation problem (2.8). The following technical lemma by Braess and Hackbusch [6, Sec. 2] will turn out to be very helpful for this purpose.

LEMMA 2.6 ([6]). *Let $s_k(\mu) = \sum_{i=1}^{k} \omega_i \exp(-\alpha_i \mu)$ with $\alpha_i, \omega_i \in \mathbb{R}$. Then there is a choice of $\alpha_i > 0, \omega_i > 0$ (depending on k and $R > 1$) such that*

$$\sup_{\mu \in [1,R]} \left| \frac{1}{\mu} - s_k(\mu) \right| \leq 16 \exp\left(\frac{-k\pi^2}{\log(8R)} \right).$$

COROLLARY 2.7. *Consider the linear system $\mathcal{A}x = b$ with $\mathcal{A}$ and b of the form (1.4)–(1.5). If $\mathcal{A}$ is symmetric positive definite then there exists an approximation x_k of tensor rank at most k, such that*

$$\|x - x_k\|_2 \leq \frac{16}{\lambda_{\min}(\mathcal{A})} \exp\left(\frac{-k\pi^2}{\log(8\kappa(\mathcal{A}))} \right) \|b\|_2. \tag{2.10}$$

Proof. As $\mathcal{A}$ is symmetric positive definite, all coefficient matrices A_s are symmetric and $\mathcal{A}x = b$ has a unique solution. Therefore, Theorem 2.5 can be applied with P_s orthogonal (hence, $\kappa_{\mathcal{P}} = 1$). The result follows directly from this theorem combined with the following observation. Applying Lemma 2.6, we use the substitution $\mu = \frac{\mu_1 + \ldots + \mu_d}{\lambda_{\min}(\mathcal{A})}, y \in [1, \frac{\lambda_{\max}(\mathcal{A})}{\lambda_{\min}(\mathcal{A})}] =: [1, R]$ and obtain

$$\lambda_{\min}(\mathcal{A}) \left| \frac{1}{\mu_1 + \ldots + \mu_d} - \sum_{i=1}^{k} \frac{\omega_i}{\lambda_{\min}(\mathcal{A})} e^{-\alpha_i \mu_1 / \lambda_{\min}(\mathcal{A})} \cdots e^{-\alpha_i \mu_d / \lambda_{\min}(\mathcal{A})} \right| \leq 16 \exp\left(\frac{-k\pi^2}{\log(8\kappa(\mathcal{A}))} \right),$$

yielding a bound on the quantity $\epsilon(k)$ defined in (2.8). $\square$

Corollary 2.7 shows that the solution x can be well approximated by a low-rank tensor, provided that $\mathcal{A}$ is symmetric positive definite. In comparison, the bound in ([12], Appendix D3.4.1) yields

$$\|x - x_{2k+1}\|_2 \leq \frac{C_{\mathrm{St}}}{\lambda_{\min}(\mathcal{A})} \exp(-\pi\sqrt{k}) \|b\|_2, \tag{2.11}$$

where C_{St} is independent of $\mathcal{A}$ and k. Experimentally, we found $C_{\mathrm{St}} \approx 2.75$. It is important to emphasize that the convergence rate predicted by (2.11) does not depend on the condition number of $\mathcal{A}$. However, this comes at the expense of having $\sqrt{k}$ instead of k in the exponent. It is therefore of interest to compare (2.11) with the bound of Corollary 2.7 for different $\kappa(\mathcal{A})$, see Figure 2.1. It turns out that the bound of Corollary 2.7 is often significantly better, except for very large values of $\kappa(\mathcal{A})$ and smaller k.

Remark 2.8. *Note that Lemma 2.2 also suggests an algorithm for calculating x_k:*

$$x_k = \sum_{j=1}^{k} \tilde{\omega}_j \bigotimes_{s=1}^{d} \exp(-\tilde{\alpha}_j A_s) b_s, \tag{2.12}$$

with $\tilde{\alpha}_j = \alpha_j / \lambda_{\min}(\mathcal{A}), \tilde{\omega}_j = \omega_j / \lambda_{\min}(\mathcal{A})$, and α_j, ω_j as in Lemma 2.6. The coefficients α_j, ω_j only depend on k and $R = \kappa(\mathcal{A}) > 1$. The resulting method will be discussed in somewhat more detail in Section 5.
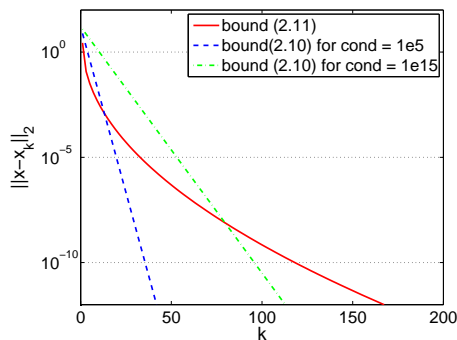
5

FIG. 2.1. *Comparison of the convergence bounds (2.10) and (2.11), assuming $\lambda_{\min}(\mathcal{A}) = 1$*

3. The tensor Krylov subspace method. In the following, we will develop numerical algorithms for approximating the solution x to the linear system (1.3). Note that Section 2.2, in particular Remark 2.8, already provides a rather effective method for computing low tensor rank approximations as the computational effort grows linearly with d and the convergence rate depends very mildly, at most logarithmically on the conditioning of $\mathcal{A}$. Indeed, this method is used for solving compressed systems arising in our algorithms, see Section 5 below. However, for the purpose of this method, the expressions $\exp(-\tilde{\alpha}_j A_s)b_s$ must be computed rather exactly to guarantee a good accuracy of x, which may be regarded expensive for larger matrices when compared to, say, a simple matrix-vector multiplication.

3.1. Tensorized Krylov subspaces. We let

$$\mathcal{K}_{k_s}(A_s, b_s) = \mathrm{span}\big\{b_s,\, A_s b_s,\, \ldots,\, A_s^{k_s-1} b_s\big\}, \quad s = 1, \ldots, d,$$

denote the *Krylov subspace* obtained from $k_s - 1$ successive matrix-vector products of A_s with b_s. In view of the PDE (1.6), each $\mathcal{K}_{k_s}(A_s, b_s)$ could be seen as a subspace corresponding to one coordinate of the domain. To obtain a subspace for all d coordinates, we tensorize and take the linear hull.

DEFINITION 3.1. *Let $\mathcal{A}$, b be as in equations (1.3)–(1.4) and consider a multi-index $\mathfrak{K} = (k_1, \ldots, k_d)$ with $k_s \in \mathbb{N}$. Then*

$$\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b) := \mathrm{span}\big(\mathcal{K}_{k_1}(A_1, b_1) \otimes \cdots \otimes \mathcal{K}_{k_d}(A_d, b_d)\big)$$

is called the **tensorized Krylov subspace** *associated with $\mathcal{A}$ and b.*

Equivalently, the tensorized Krylov subspace can be defined as

$$\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b) = \mathrm{span}\big\{A_1^{i_1-1} b_1 \otimes \cdots \otimes A_d^{i_d-1} b_d : \mathfrak{I} \leq \mathfrak{K}\big\}. \tag{3.1}$$

A more computationally oriented definition is obtained as follows. Define d matrices U_s, $s = 1, \ldots, d$, such that the columns of each U_s span $\mathcal{K}_{k_s}(A_s, b_s)$. Then $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ is spanned by the columns of $\mathcal{U} = U_1 \otimes \cdots \otimes U_d$. Combined with equations (2.6)–(2.7), this shows that $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ is spanned by the Tucker decompositions $[\![c;\, U_1, U_2, \ldots, U_d]\!]$ for all possible core tensors c.

In the following, we discuss an extension of the well-known relation between Krylov subspaces and matrix polynomials. Given a multi-index $\mathfrak{K}$ we call $p : \mathbb{R}^d \to \mathbb{R}$ a *multivariate polynomial* of degree less than $\mathfrak{K}$ if p is a polynomial of degree at most $k_s - 1$ in the sth variable. The space of all such multivariate polynomials is denoted by $\Pi_{\mathfrak{K}}^{\otimes}$. Each $p \in \Pi_{\mathfrak{K}}^{\otimes}$ can be written as

$$p(\mu_1, \ldots, \mu_d) = \sum_{\mathfrak{L} \leq \mathfrak{K}} c_{\mathfrak{L}}\, \mu_1^{l_1-1} \mu_2^{l_2-1} \cdots \mu_d^{l_d-1}, \quad c_{\mathfrak{L}} \in \mathbb{R},$$

where the sum is taken over multi-indices $\mathfrak{L}$ that satisfy $1 \leq l_s \leq k_s$ The *evaluation* of p at the

6

matrix $\mathcal{A}$, defined in (1.4) and represented by the matrix tuple $(A_1, \ldots, A_d)$, is defined as

$$
\begin{aligned}
p(A_1, \ldots, A_d) &:= \sum_{\mathfrak{L} \le \mathfrak{K}} c_{\mathfrak{L}} \, \hat{A}_1^{l_1 - 1} \hat{A}_2^{l_2 - 1} \cdots \hat{A}_d^{l_d - 1} \\
&= \sum_{\mathfrak{L} \le \mathfrak{K}} c_{\mathfrak{L}} \, A_1^{l_1 - 1} \otimes A_2^{l_2 - 1} \otimes \cdots \otimes A_d^{l_d - 1},
\end{aligned} \tag{3.2}
$$

where $\hat{A}_s = I_{n_1} \otimes \cdots \otimes I_{n_{s-1}} \otimes A_s \otimes I_{n_{s+1}} \otimes \cdots \otimes I_{n_d}$ and $\mathcal{A} = \hat{A}_1 + \cdots + \hat{A}_d$.

LEMMA 3.2. *With the notation introduced above,* $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b) = \mathrm{span}\{p(A_1, \ldots, A_d)b : p \in \Pi_{\mathfrak{K}}^{\otimes}\}$.

Proof. Let $g = p(A_1, \ldots, A_d)b$ for some $p \in \Pi_{\mathfrak{K}}^{\otimes}$. By (3.2), this is equivalent to

$$
g = \sum_{\mathfrak{L} \le \mathfrak{K}} c_{\mathfrak{L}} \, A_1^{l_1 - 1} b_1 \otimes \cdots \otimes A_d^{l_d - 1} b_d, \tag{3.3}
$$

i.e., g is a linear combination of elements from $\mathcal{K}_{k_1}(A_1, b_1) \otimes \cdots \otimes \mathcal{K}_{k_d}(A_d, b_d)$ and therefore – by definition – $g \in \mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$. For the other direction, we note that (3.1) implies that any $g \in \mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ can be written in the form (3.3), which concludes the proof. $\square$

Remark 3.3. Lemma 3.2 reveals an important difference between standard and tensorized Krylov subspace. For $k_0 \in \mathbb{N}$, the standard Krylov subspace satisfies

$$
\mathcal{K}_{k_0}(\mathcal{A}, b) = \{p(\mathcal{A})b : p \in \Pi_{k_0}\},
$$

where Π_{k_0} denotes the space of all *univariate* polynomials of degree at most k_0. For a given univariate polynomial $p \in \Pi_{k_0}$, we define the multivariate polynomial

$$
\tilde{p}(\mu_1, \mu_2, \ldots, \mu_d) := p(\mu_1 + \mu_2 + \cdots + \mu_d). \tag{3.4}
$$

By direct computation $\tilde{p}(A_1, \ldots, A_d)b = p(\mathcal{A})b$, which – together with Lemma 3.2 – shows

$$
\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b) \supset \mathcal{K}_{k_0}(\mathcal{A}, b)
$$

for $\mathfrak{K} = (k_0, \ldots, k_0)$. On the other hand, it is obvious that not every multivariate polynomial takes the particular form (3.4) and hence $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b) \ne \mathcal{K}_{k_0}(\mathcal{A}, b)$ for $d > 1$ and nontrivial choices of $\mathcal{A}, b$. To summarize: *Tensorized Krylov subspaces are richer than standard Krylov subspaces.* This is no surprise when taking into account that the dimension of $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ grows exponentially with d while the dimension of $\mathcal{K}_{k_0}(\mathcal{A}, b)$ grows only linearly with d.

3.2. Basic algorithm. In this section, we present the basics of the newly proposed tensor Krylov subspace algorithm. This algorithm approximates the solution x of the linear system (1.3) by an element from $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$.

To start with, we require a basis of $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$. For this purpose, the standard Arnoldi method is used to compute matrices $U_s \in \mathbb{R}^{n_s \times k_s}$ such that the columns of each U_s form an orthonormal basis of the Krylov subspace $\mathcal{K}_{k_s}(A_s, b_s)$. A brief description of the Arnoldi method is provided in Algorithm 1; more algorithmic details can be found, e.g., in [21, 26]. We assume that a suitable reorthogonalization strategy is performed such that the columns of U_s are also numerically orthonormal.

Upon successful completion of Algorithm 1, one obtains the so called *Arnoldi decomposition*

$$
A_s U_s = U_s H_s + h_{k_s+1, k_s}^{(s)} u_{k_s+1}^{(s)} e_{k_s}^{\top}, \tag{3.5}
$$

where the upper Hessenberg matrix H_s collects the coefficients $h_{ij}^{(s)}$:

$$
H_s = \begin{bmatrix}
h_{11}^{(s)} & h_{12}^{(s)} & \cdots & & h_{1,k_s}^{(s)} \\
h_{21}^{(s)} & \ddots & & \ddots & \vdots \\
& \ddots & & \ddots & h_{k_s-1,k_s}^{(s)} \\
& & h_{k_s,k_s-1}^{(s)} & & h_{k_s,k_s}^{(s)}
\end{bmatrix} = U_s^{\top} A_s U_s. \tag{3.6}
$$

Algorithm 1 Arnoldi method

Input: Matrix $A_s \in \mathbb{R}^{n_s \times n_s}$, Vector $b_s \in \mathbb{R}^{n_s}$, $k_s \in \mathbb{N}$.
Output: Matrix $U_s \in \mathbb{R}^{n \times k_s}$ containing an orthonormal basis of $\mathcal{K}_{k_s}(A_s, b_s)$.

$\quad u_1^{(s)} \leftarrow b_s / \|b_s\|_2$
$\quad$ **for** $j = 1, \ldots, k_s$ **do**
$\quad\quad w \leftarrow A_s u_j^{(s)}$
$\quad\quad h_{1:j,j}^{(s)} \leftarrow \left[u_1^{(s)}, \ldots, u_j^{(s)} \right]^\top w$
$\quad\quad \tilde{u}_{j+1}^{(s)} \leftarrow w - \left[u_1^{(s)}, \ldots, u_j^{(s)} \right] h_{1:j,j}^{(s)}$
$\quad\quad h_{j+1,j}^{(s)} = \| \tilde{u}_{j+1}^{(s)} \|_2$
$\quad\quad u_{j+1}^{(s)} = \tilde{u}_{j+1}^{(s)} / \| \tilde{u}_{j+1}^{(s)} \|_2$
$\quad$ **end for**
$\quad$ Set $U_s = \left[u_1^{(s)}, \ldots, u_{k_s}^{(s)} \right]$.

Note that if A_s is symmetric then H_s inherits this symmetry and becomes a tridiagonal matrix.

As discussed above, the tensor Krylov subspace $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ is spanned by the columns of $\mathcal{U} = U_1 \otimes \ldots \otimes U_d$. To extract an approximation to the solution of the linear system (1.3) from the tensor Krylov subspace, we define $x_{\mathfrak{K}} = \mathcal{U}y$ where y solves the *compressed linear system*

$$\mathcal{H}y = \tilde{b}, \quad \text{with } \mathcal{H} = \mathcal{U}^\top \mathcal{A}\mathcal{U} \text{ and } \tilde{b} = \mathcal{U}^\top b. \tag{3.7}$$

Then, $x_{\mathfrak{K}}$ satisfies the following *Galerkin condition*:

$$\mathcal{A}x_{\mathfrak{K}} - b \perp \mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b).$$

The solvability of (3.7) will be discussed in more detail in Section 4. Note that $\mathcal{H}$ and $\tilde{b}$ take the form

$$\mathcal{H} = \mathcal{U}^\top \mathcal{A}\mathcal{U} = \bigotimes_{s=1}^{d} U_s^\top \left(\sum_{s=1}^{d} I_{n_1} \otimes \cdots \otimes I_{n_{s-1}} \otimes A_s \otimes I_{n_{s+1}} \otimes \cdots \otimes I_{n_d} \right) \bigotimes_{s=1}^{d} U_s$$

$$= \sum_{s=1}^{d} I_{k_1} \otimes \cdots \otimes I_{k_{s-1}} \otimes H_s \otimes I_{k_{s+1}} \otimes \cdots \otimes I_{k_d},$$

$$\tilde{b} = \mathcal{U}^\top \bigotimes_{s=1}^{d} b_s = \bigotimes_{s=1}^{d} U_s^\top b_s = \prod_{s=1}^{d} \|b_s\|_2 \bigotimes_{s=1}^{d} e_1.$$

It is important to note that the compressed system $\mathcal{H}y = \tilde{b}$ inherits the Kronecker product structure from the original linear system. Solution methods applicable to the original system can therefore also be applied to the compressed system, see also Section 5. Note that the computational effort for building up and storing the bases of the tensorized Krylov subspace grows only linearly with the number of dimensions. Assuming that the cost for solving the compressed system admits the same growth, we therefore obtain a numerical method with an overall cost that scales linearly with d.

For small dimensions d, it might be feasible to store an explicit representation of the solution y to (3.7). In this case, the approximation $x_{\mathfrak{K}}$ is represented by the Tucker decomposition

$$x_{\mathfrak{K}} = \mathcal{U}y = [\![y; U_1, U_2, \ldots, U_d]\!]$$

with core tensor y. If y itself is represented by a Tucker decomposition, then $x_{\mathfrak{K}}$ admits again a Tucker decomposition with the same core tensor as y. For high dimensions d, such a representation is not admissible and we will discuss in Section 5 how to represent (or rather approximate) y by a CP decomposition,

$$y = \sum_{i=1}^{t} y_i^{(1)} \otimes \cdots \otimes y_i^{(d)}. \tag{3.8}$$

Then $x_{\mathfrak{K}}$ can also be represented by a CP decomposition:

$$x_{\mathfrak{K}} \approx \mathcal{U} \sum_{i=1}^{t} y_i^{(1)} \otimes \cdots \otimes y_i^{(d)} = \sum_{i=1}^{t} U_1 y_i^{(1)} \otimes \cdots \otimes U_d y_i^{(d)}.$$

Algorithm 2 summarizes the proposed tensor Krylov subspace method for solving (1.3).

Algorithm 2 Tensor Krylov subspace method

Input: Coefficients $A_s \in \mathbb{R}^{n_s \times n_s}$ and $b_s \in \mathbb{R}^{n_s}$ of the linear system (1.3).
Output: Approximation $x_{\mathfrak{K}} = \mathcal{U}y$ to the solution x of (1.3).
   for s = 1 to d **do**
      Apply Algorithm 1 to A_s and b_s to compute U_s, H_s.
   end for
   Compute/approximate solution y to the compressed equation $\mathcal{H}y = \tilde{b}$ in (3.7).

3.3. Computation of the residual. To monitor the convergence of Algorithm 2, one can compute the norm of the residual $r_{\mathfrak{K}} = b - \mathcal{A}x_{\mathfrak{K}}$. The following lemma extends known results for Lyapunov and Sylvester equations [17] to compute this norm in a cheaper way when $x_{\mathfrak{K}}$ is obtained by a Krylov subspace method.

LEMMA 3.4. *Let $x_{\mathfrak{K}}$ be computed by Algorithm 2. Then the residual $r_{\mathfrak{K}} = \mathcal{A}x_{\mathfrak{K}} - b$ satisfies*

$$\|r_{\mathfrak{K}}\|_2^2 = \sum_{s=1}^{d} \left| h_{k_s+1,k_s}^{(s)} \right|^2 \sum_{\substack{\mathfrak{L} \leq \mathfrak{K} \\ l_s = k_s}} |y_{\mathfrak{L}}|^2 + \|\mathcal{H}y - \tilde{b}\|_2^2.$$

Proof. For $\beta \in \{0,1\}^d$, we define $\mathcal{U}_\beta = U_{\beta,1} \otimes U_{\beta,2} \otimes \cdots \otimes U_{\beta,d}$, where $U_{\beta,s} = U_s$ if $\beta(s) = 0$ and the columns of $U_{\beta,s}$ form a basis of $\text{span}(U_s)^\perp$ otherwise. Note that $\text{span}(\mathcal{U}_\beta) \perp \text{span}(\mathcal{U}_\gamma)$ unless $\beta = \gamma$. Hence $\mathbb{R}^{n_1 n_2 \cdots n_d}$ can be written as the orthogonal sum of all $\text{span}(\mathcal{U}_\beta)$ and therefore

$$\|r_{\mathfrak{K}}\|_2^2 = \sum_{\beta \in \{0,1\}^d} \|\mathcal{U}_\beta^\top r\|_2^2 \tag{3.9}$$

For $\beta \equiv 0$, $\mathcal{U}_\beta = \mathcal{U}$ and $\mathcal{U}^\top r_{\mathfrak{K}} = \mathcal{U}^\top \mathcal{A}\mathcal{U}y - \mathcal{U}^\top b = \mathcal{H}y - \tilde{b}$, the error in solving the compressed system. For general β,

$$\mathcal{U}_\beta^\top \mathcal{A}\mathcal{U} = \sum_{s=1}^{d} U_{\beta,1}^\top U_1 \otimes \cdots \otimes U_{\beta,s-1}^\top U_{s-1} \otimes U_{\beta,s}^\top A_s U_s \otimes U_{\beta,s+1}^\top U_{s+1} \otimes \cdots \otimes U_{\beta,d}^\top U_d.$$

Note that $U_{\beta,j}^\top U_j = 0$ if $\beta(j) = 1$. Hence, $\mathcal{U}_\beta^\top \mathcal{A}\mathcal{U} = 0$ if β contains more than one entry 1. Combined with the fact that $U_\beta^\top b = 0$ unless $\beta \equiv 0$, this implies that only terms corresponding to β with exactly one entry 1 contribute to the sum (3.9). Let us consider such a $\beta_s = (0, \ldots, 0, 1, 0, \ldots, 0)$ with a single entry 1 at the sth position. Then

$$\begin{aligned}
\|\mathcal{U}_{\beta_s}^\top r_{\mathfrak{K}}\|_2^2 &= \left\| \left(I \otimes \cdots \otimes I \otimes U_{\beta,s}^\top A_s U_s \otimes I \otimes \cdots \otimes I \right) y \right\|_2^2 \\
&= \left\| \left(I \otimes \cdots \otimes I \otimes h_{k_s+1,k_s}^{(s)} U_{\beta,s}^\top u_{k_s+1}^{(s)} e_{k_s}^\top \otimes I \otimes \cdots \otimes I \right) y \right\|_2^2 \\
&= \left| h_{k_s+1,k_s}^{(s)} \right|^2 \left\| \left(I \otimes \cdots \otimes I \otimes e_{k_s}^\top \otimes I \otimes \cdots \otimes I \right) y \right\|_2^2 = \left| h_{k_s+1,k_s}^{(s)} \right|^2 \sum_{\substack{\mathfrak{L} \leq \mathfrak{K} \\ l_s = k_s}} |y_{\mathfrak{L}}|^2, \tag{3.10}
\end{aligned}$$

where we used the Arnoldi decomposition (3.5) and $\left\| U_{\beta,s}^\top u_{k_s+1}^{(s)} \right\|_2 = 1$. This completes the proof as $\|r\|_2^2$ is obtained by summing up the terms (3.10) for $s = 1, \ldots, d$. $\square$

Although the expression provided by Lemma 3.4 reduces the cost for computing the residual norm significantly, it still scales exponentially with d simply because almost all elements of y need to be accessed. This exponential growth can be avoided if y is represented in a data-sparse format. Consider, for example, a CP decomposition (3.8) of y. Then

$$\sum_{\substack{\mathfrak{L} \leq \mathfrak{K} \\ l_s = k_s}} |y_{\mathfrak{L}}|^2 = \Big\| \sum_{i=1}^{t} e_{k_s}^{\top} y_i^{(s)} \bigotimes_{j \neq s} y_i^{(j)} \Big\|_2^2,$$

where the cost for computing the latter expression scales linearly with d, assuming that t remains constant as d grows. Hence, the overall cost for evaluating $\|r_{\mathfrak{K}}\|_2^2$ scales quadratically with d.

4. Convergence analysis. In the following, we will develop a convergence analysis for the tensor Krylov subspace method in special cases. It is clear that this can only be performed in a meaningful way if the unique solvability of the compressed system (3.7) is guaranteed. The following lemma is an extension of the usual positive definiteness condition in the convergence analysis of FOM methods for standard linear systems, see [24] and the references therein.

LEMMA 4.1. *Given the equation (1.3), suppose that the eigenvalues of the symmetric parts $(A_s + A_s^{\top})/2$ are contained in intervals $[\alpha_s, \beta_s]$. Let $\mathcal{U} = U_1 \otimes \cdots \otimes U_d$ where the columns of each $U_s \in \mathbb{R}^{n_s \times k_s}$ form an orthonormal basis. Then the compressed matrix $\mathcal{U}^{\top} \mathcal{A} \mathcal{U}$ is invertible if*

$$\Big[\sum_{s=1}^{d} \alpha_s, \sum_{s=1}^{d} \beta_s \Big] \cap \{0\} = \emptyset. \tag{4.1}$$

Proof. The Cauchy interlacing theorem implies that the eigenvalues of the compressed symmetric parts $U_s^{\top} A_s U_s + U_s^{\top} A_s^{\top} U_s = U_s^{\top}(A_s + A_s^{\top})U_s$ are also contained in $[\alpha_s, \beta_s]$. Combined with Lemma 2.1, this shows that any eigenvalue of $\mathcal{U}^{\top} \mathcal{A} \mathcal{U} + \mathcal{U}^{\top} \mathcal{A}^{\top} \mathcal{U}$ can be written as $\mu = \sum_{s=1}^{d} \mu_s$, $\mu_s \in [\alpha_s, \beta_s]$ and therefore $\mu \in \big[\sum_{s=1}^{d} \alpha_s, \sum_{s=1}^{d} \beta_s \big]$. Then (4.1) implies that the set of all such μ is either negative or positive. Hence, the symmetric part of $\mathcal{U}^{\top} \mathcal{A} \mathcal{U}$ is positive or negative definite, which concludes the proof. $\square$

It is common to call a non-symmetric matrix to be positive/negative definite if its symmetric part is positive/negative definite. By Lemma 2.1, the condition (4.1) is equivalent to the definiteness of $\mathcal{A}$. For this condition to be satisfied it is sufficient but not necessary* that all coefficients A_s are either positive definite or negative definite. In particular, the compressed system is solvable in the special case when all A_s are symmetric positive definite.

For the development of our convergence analysis it is central to note that the *residual* $r_{\mathfrak{K}} = \mathcal{A} x_{\mathfrak{K}} - b$, with $x_{\mathfrak{K}}$ produced by Algorithm 2, satisfies

$$\mathcal{U}^{\top} r_{\mathfrak{K}} = \mathcal{U}^{\top} \mathcal{A} x_{\mathfrak{K}} - \mathcal{U}^{\top} b = \mathcal{U}^{\top} \mathcal{A} \mathcal{U} y - \tilde{b} = \mathcal{H} y - \tilde{b} = 0.$$

In other words, the following *Galerkin condition* holds:

$$\mathcal{A} x_{\mathfrak{K}} - b \perp \mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b). \tag{4.2}$$

For practical purposes, the compressed system cannot be solved exactly, but only a low-rank approximation to the solution y is found. However, as this approximation error can be reduced at will by increasing the rank of y, we will assume in this section that the compressed system is solved exactly.

4.1. The symmetric positive definite case. We first consider the case that $\mathcal{A}$ is *symmetric positive definite*. This allows us to introduce the weighted Euclidean norm

$$\|c\|_{\mathcal{A}} := \|\mathcal{A}^{1/2} c\|_2, \quad c \in \mathbb{R}^{n_1 n_2 \cdots n_s},$$

---

*On the other hand, it is easy to see that one can always shift the matrices A_s in a way such that each A_s inherits the definiteness of $\mathcal{A}$ but $\mathcal{A}$ itself is not altered. For positive definite $\mathcal{A}$, $\tilde{A}_s = A_s + (\bar{\alpha} - \alpha_s)I$ with $\bar{\alpha} = \frac{\alpha_1 + \cdots + \alpha_d}{d}$ is such a shift.

and relate the Galerkin condition (4.2) to a linear least-squares problem.

PROPOSITION 4.2. *Let x denote the solution of (1.3), where $\mathcal{A}$ is symmetric positive definite. Then the Galerkin condition (4.2) for an approximation $x_{\mathfrak{K}}$ implies*

$$x_{\mathfrak{K}} = \underset{\tilde{x}\in\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A},b)}{\arg\min} \|\tilde{x} - x\|_{\mathcal{A}}.$$

Proof. This result is well known and can be proven by standard techniques (see, e.g. [21]). □
An immediate consequence of Proposition 4.2, the error in the $\mathcal{A}$-norm of $x_{\mathfrak{K}}$ decreases monotonically as any of the individual dimensions k_s in the tensorized Krylov subspace $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ grows.

The following theorem turns Proposition 4.2 into a multivariate polynomial approximation problem. Let us recall from Section 3.1 that $\Pi_{\mathfrak{K}}^{\otimes}$ denotes the space of all multivariate polynomials of degree at most $k_s - 1$ in the sth variable.

THEOREM 4.3. *Under the assumptions of Proposition 4.2,*

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} \leq \sqrt{\|\mathcal{A}\|_2}\|b\|_2 \min_{p\in\Pi_{\mathfrak{K}}^{\otimes}} \max_{\substack{\lambda_s\in\Lambda(A_s)\\ s=1,\ldots,d}} \left| p(\lambda_1,\ldots,\lambda_d) - \frac{1}{\lambda_1 + \cdots + \lambda_d}\right|.$$

Proof. From Proposition 4.2, we have

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} = \min_{\tilde{x}\in\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A},b)} \|\tilde{x} - x\|_{\mathcal{A}} \leq \sqrt{\|\mathcal{A}\|_2} \min_{\tilde{x}\in\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A},b)} \|\tilde{x} - x\|_2.$$

By Lemma 3.2, any tensor $\tilde{x}$ in the tensorized Krylov subspace $\mathcal{K}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b)$ can be expressed as $\tilde{x} = p(A_1,\ldots,A_d)b$, where p is a multivariate polynomial in $p \in \Pi_{\mathfrak{K}}^{\otimes}$. Hence,

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} \leq \sqrt{\|\mathcal{A}\|_2}\,\|b\|_2 \min_{p\in\Pi_{\mathfrak{K}}^{\otimes}} \left\|p(A_1,\ldots,A_d) - \mathcal{A}^{-1}\right\|_2.$$

By a similarity transformation (parallel to the proof of Theorem 2.5), we find

$$\left\|p(A_1,\ldots,A_d) - \mathcal{A}^{-1}\right\|_2 = \max_{\substack{\lambda_s\in\Lambda(A_s)\\ s=1,\ldots,d}} \left|p(\lambda_1,\ldots,\lambda_d) - \frac{1}{\lambda_1 + \cdots + \lambda_d}\right|,$$

which concludes the proof. □
We follow the standard technique of relaxing the min-max problem of Theorem 4.3 such that the maximum is taken over the intervals $[\alpha_s, \beta_s] := [\lambda_{\min}(A_s), \lambda_{\max}(A_s)]$ instead of the discrete sets of eigenvalues. This can only increase the bound and we therefore obtain

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} \leq \sqrt{\|\mathcal{A}\|_2}\|b\|_2 \, E_{\Omega}(\mathfrak{K}),$$

where

$$E_{\Omega}(\mathfrak{K}) := \min_{p\in\Pi_{\mathfrak{K}}^{\otimes}} \left\|p(\mu_1,\ldots,\mu_d) - \frac{1}{\mu_1 + \cdots + \mu_d}\right\|_{\Omega}, \tag{4.3}$$

with $\|\cdot\|_{\Omega}$ defined as the supremum norm on $\Omega := [\alpha_1, \beta_1] \times \cdots \times [\alpha_d, \beta_d]$.

There are several ways to approach the multivariate polynomial approximation problem (4.3). Inspired by work in [5], we have first followed an approach based on interpolation by tensor Chebyshev polynomials in [27]. Unfortunately, the Lebesgue constants needed to be taken care of in this approach grow exponentially with d, leading to rather loose bounds for high dimensions. For example, if $k_1 = \cdots = k_d =: k$ a factor proportional to $\left(\frac{2}{\pi}\ln(k)\right)^{d-1}$ is introduced by the Lebesgue constants. This factor can be avoided when following a completely different approach, essentially mimicking the proof of [23] on a scalar level for arbitrary dimensions. Lemma A.1 in the Appendix contains the approximation result obtained in this manner. Combining Theorem 4.3 with Lemma A.1 yields the following convergence bound.

11

COROLLARY 4.4. *Under the assumptions of Proposition 4.2,*

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} \le \frac{\sqrt{\|\mathcal{A}\|_2}\,\|b\|_2}{\lambda_{\min}(\mathcal{A})} \sum_{s=1}^{d} \frac{\sqrt{\kappa_s}+1}{\sqrt{\kappa_s}} \cdot \left(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\right)^{k_s},$$

where $\kappa_s = 1 + \frac{\beta_s - \alpha_s}{\lambda_{\min}(\mathcal{A})}$.

For $d = 2$, a bound similar to the bound of Corollary 4.4 has been obtained in [23, Proposition 3.1]. In fact, the bound of [23] has a somewhat smaller constant by avoiding the detour via multivariate polynomial approximations. In principle, the proof techniques [23] could also be extended to $d > 2$ but our approach has the advantage of also admitting convergence bounds for the extended Krylov subspace method, see Section 6.

Remark 4.5. It is instructive to compare the error bound of Corollary 4.4 with the well-known error bound of the classical CG method applied to the linear system (1.3):

$$\|x_k - x\|_{\mathcal{A}} \le 2\|b\|_{\mathcal{A}}\left(\frac{\sqrt{\kappa(\mathcal{A})}-1}{\sqrt{\kappa(\mathcal{A})}+1}\right)^k \le 2\sqrt{\|\mathcal{A}\|_2}\|b\|_2\left(\frac{\sqrt{\kappa(\mathcal{A})}-1}{\sqrt{\kappa(\mathcal{A})}+1}\right)^k, \tag{4.4}$$

where $\kappa(\mathcal{A}) = \|\mathcal{A}\|_2\|\mathcal{A}^{-1}\|_2 = \frac{\beta_1+\cdots+\beta_d}{\alpha_1+\cdots+\alpha_d}$.

To simplify the discussion, assume $\alpha_1 = \cdots = \alpha_d =: \alpha$ and $\beta_1 = \cdots = \beta_d =: \beta$, in which case it is reasonable to choose $\mathfrak{K} = (k, \ldots, k)$. Then the bound of Corollary 4.4 simplifies to

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} \le d\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}} \cdot \frac{\sqrt{\|\mathcal{A}\|_2}\,\|b\|_2}{\lambda_{\min}(\mathcal{A})}\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k, \tag{4.5}$$

where $\kappa = 1 + \frac{\beta-\alpha}{d\alpha} = \frac{d-1}{d} + \frac{\kappa(\mathcal{A})}{d}$. For larger $\kappa(\mathcal{A})$ this means that the effective condition number that determines the convergence rate in (4.5) is divided by d in comparison to (4.4). This indicates that the tensor Krylov subspace method can be expected to require $1/\sqrt{d}$ times the iterations needed by classical CG, see also Remark 3.3. More surprisingly, the convergence rate of the tensor Krylov subspace method appears to *improve* with increasing number of dimensions d, assuming that the condition number of $\mathcal{A}$ remains constant. This benefit from higher dimensions was already noted for $d = 2$ in [23] and is confirmed by the numerical experiments in Section 7.

Remark 4.6. To avoid unnecessary work, it is of interest to choose $\mathfrak{K} = (k_1, \ldots, k_d)$ such that the summands in the convergence bound of Corollary 4.4 are balanced, i.e., the terms $\left(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\right)^{k_s}$ are nearly constant for all s. Taking the logarithm yields

$$k_s \log\left(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\right) \approx -2\frac{k_s}{\sqrt{\kappa_s}}$$

for large κ_s. Hence it is reasonable to choose $\mathfrak{K} = (k_1, \ldots, k_d)$ such that $\frac{k_s}{\sqrt{\kappa_s}}$ is nearly constant across all dimensions s.

4.2. An approach to the non-symmetric positive definite case. To obtain convergence bounds in the case that $\mathcal{A}$ is *non-symmetric* positive definite, we follow the proof technique by Simoncini and Druskin [23] for Lyapunov equations. To simplify the presentation it is assumed that $\|b_s\|_2 = 1$ throughout the rest of this section.

LEMMA 4.7. *Let x denote the solution of (1.3), where $\mathcal{A}$ is positive definite. For the approximation $x_{\mathfrak{K}}$ obtained by Algorithm 2 it holds that*

$$\|x_{\mathfrak{K}} - x\|_2 \le \sum_{s=1}^{d} \int_0^{\infty} e^{-\hat{\alpha}_s t}\big\|x_{k_s}^{(s)} - x^{(s)}\big\|_2 \mathrm{d}t,$$

where $x^{(s)} = e^{-tA_s}b_s$, $x_{k_s}^{(s)} = U_s e^{-tH_s}e_1$, and $\hat{\alpha}_s = \sum_{j\ne s} \alpha_j$ with $\alpha_j = \lambda_{min}(A_j + A_j^\top)/2$.

12

Proof. By Lemma 2.2, both x and $x_{\mathfrak{K}}$ admit integral representations:

$$x = \int_0^\infty x^{(1)} \otimes \cdots \otimes x^{(d)} \; \mathrm{d}t, \quad x_{\mathfrak{K}} = \int_0^\infty x_{k_1}^{(1)} \otimes \cdots \otimes x_{k_d}^{(d)} \; \mathrm{d}t.$$

Note that $\|x^{(s)}\|_2 \leq e^{-\alpha_s t}$ as well as $\|x_{k_s}^{(s)}\|_2 \leq e^{-\alpha_s t}$. We have

$$\begin{aligned}
\|x_{\mathfrak{K}} - x\|_2 &\leq \int_0^\infty \Big\| \bigotimes_{j=1}^d x_{k_j}^{(j)} - \bigotimes_{j=1}^d x^{(j)} \Big\|_2 \; \mathrm{d}t \\
&= \int_0^\infty \Big\| \sum_{s=1}^d \Big(\bigotimes_{j=1}^{s-1} x_{k_j}^{(j)} \Big) \otimes \big(x_{k_s}^{(s)} - x^{(s)} \big) \otimes \Big(\bigotimes_{j=s+1}^d x^{(j)} \Big) \Big\|_2 \; \mathrm{d}t \\
&\leq \sum_{s=1}^d \int_0^\infty \Big(\prod_{j=1}^{s-1} \|x_{k_j}^{(j)}\|_2 \Big) \|x_{k_s}^{(s)} - x^{(s)}\|_2 \Big(\prod_{j=s+1}^d \|x^{(j)}\|_2 \Big) \; \mathrm{d}t \\
&\leq \sum_{s=1}^d \int_0^\infty e^{-\hat{\alpha}_s t} \|x_{k_s}^{(s)} - x^{(s)}\|_2 \; \mathrm{d}t,
\end{aligned}$$

which concludes the proof. ☐

Note that the term $\big\|x_{k_s}^{(s)} - x^{(s)}\big\|_2$ appearing in the bound of Lemma 4.7 corresponds to the approximation error of the usual Krylov subspace approximation to the matrix exponential $e^{-tA_s} b_s$. Any reasonably good bound on this approximation error could be inserted to yield a bound on $\|x_{\mathfrak{K}} - x\|_2$. In the following, we demonstrate this procedure for the case that the fields of values $F(A_s) = \{w^\top A_s w : w \in \mathbb{C}^{n_s}, \|w\|_2 = 1\}$ for $s = 1, \ldots, d$ are contained in ellipses.

THEOREM 4.8. *Additionally to the assumptions of Lemma 4.7, suppose that the field of values of each $A_s \in \mathbb{R}^{n_s \times n_s}$ is contained in an ellipse of center $(c_s, 0)$, foci $(c_s \pm f_s, 0)$ and semi-axes $a_s^{(1)}$ and $a_s^{(2)}$. Then*

$$\|x_{\mathfrak{K}} - x\|_2 \leq \sum_{s=1}^d \frac{4}{\sqrt{(\hat{\alpha}_s + c_s)^2 - f_s^2}} \frac{\rho_s}{\rho_s - 1} \rho_s^{-k_s},$$

where $r_s = \frac{a_s^{(1)} + a_s^{(2)}}{2}$, $\rho_s = \frac{c_s + \hat{\alpha}_s}{2 r_s} + \frac{1}{2 r_s} \sqrt{(c_s + \hat{\alpha}_s)^2 - f_s^2}$ for $s = 1, \ldots, d$.

Proof. By the proof of Proposition 4.1 in [23],

$$\|x_{k_s}^{(s)} - x^{(s)}\|_2 \leq 4 \sum_{j=k_s}^\infty e^{-c_s t} I_j(f_s t) \Big(\frac{2 r_s}{f_s} \Big)^j,$$

where I_j denotes the jth modified Bessel function, see also the proof of Lemma A.1. Combined with Lemma 4.7, this yields

$$\begin{aligned}
\|x_{\mathfrak{K}} - x\|_2 &\leq 4 \sum_{s=1}^d \sum_{j=k_s}^\infty \int_0^\infty e^{-(\hat{\alpha}_s + c_s) t} I_j(f_s t) \Big(\frac{2 r_s}{f_s} \Big)^j \; \mathrm{d}t \\
&= \sum_{s=1}^d \frac{4}{\sqrt{(\hat{\alpha}_s + c_s)^2 - f_s^2}} \sum_{j=k_s}^\infty \rho_s^{-j} = \sum_{s=1}^d \frac{4}{\sqrt{(\hat{\alpha}_s + c_s)^2 - f_s^2}} \frac{\rho_s}{\rho_s - 1} \rho_s^{-k_s}.
\end{aligned}$$

☐

In a manner analogous to the proof of Theorem 4.8 the other results from [23] for non-symmetric positive definite matrices, dealing for example with a field of values contained in a wedge-shaped set, might be extended to arbitrary dimensions.

13

5. Solving the compressed system. The tensor Krylov subspace method, see Algorithm 2, requires the solution of the compressed system

$$\mathcal{H}y = \tilde{b},$$

having the same Kronecker product structure as the original system (1.3), with the coefficients H_s in upper Hessenberg form. As mentioned in Section 3.2, this system might be solved explicitly by a direct method for small dimensions but for higher dimensions this will quickly become prohibitively expensive. The method already suggested in Remark 2.8 provides a viable alternative. An approximation y_t to the exact solution y is calculated as

$$y_t = \sum_{j=1}^{t} \frac{\omega_j}{\lambda_{\min}(\mathcal{H})} \bigotimes_{s=1}^{d} \exp\Big(- \frac{\alpha_j}{\lambda_{\min}(\mathcal{H})} H_s \Big) \tilde{b}_s,$$

with

$$1/\lambda \approx \sum_{j=1}^{t} \omega_j e^{-\alpha_j \lambda}, \quad \lambda \in \Lambda(\mathcal{H}).$$

The success of this approach depends of course crucially on the choice of the coefficients $\alpha_j > 0, \omega_j > 0$. Ideally, we would like to solve the min-max problem

$$\min_{\alpha_j, \omega_j \in \mathbb{R}^+} \max_{\mu \in \Omega} \Big| 1/\mu - \sum_{j=1}^{t} \omega_j e^{-\alpha_j \mu} \Big|,$$

where $\Omega \subseteq \mathbb{C}$ contains all eigenvalues of $\mathcal{H}$ scaled by some factor $1/\rho$. Section 5.1 discusses the case of symmetric positive definite $\mathcal{H}$, in which case $\rho = \lambda_{\min}(\mathcal{H})$ and $\Omega = [1, \kappa(\mathcal{H})]$. For this purpose, a finite upper bound on the condition number $\kappa(\mathcal{H})$ must be available, which can be determined from the eigenvalues of H_s and applying Lemma 2.1. For non-symmetric $\mathcal{H}$ (or in case no upper bound on $\kappa(\mathcal{H})$ can be computed) we have $\Omega = \{z \in \mathbb{C} : \Re(z) \geq 1\}$, assuming that $\mathcal{H}$ has only eigenvalues in the right-half complex plane and $\rho = \min\{\Re(\lambda) : \lambda \in \Lambda(\mathcal{H})\}$. This case is discussed in Section 5.2. A comparison of the coefficients α_j resulting in each case can be found in Figure 5.1.
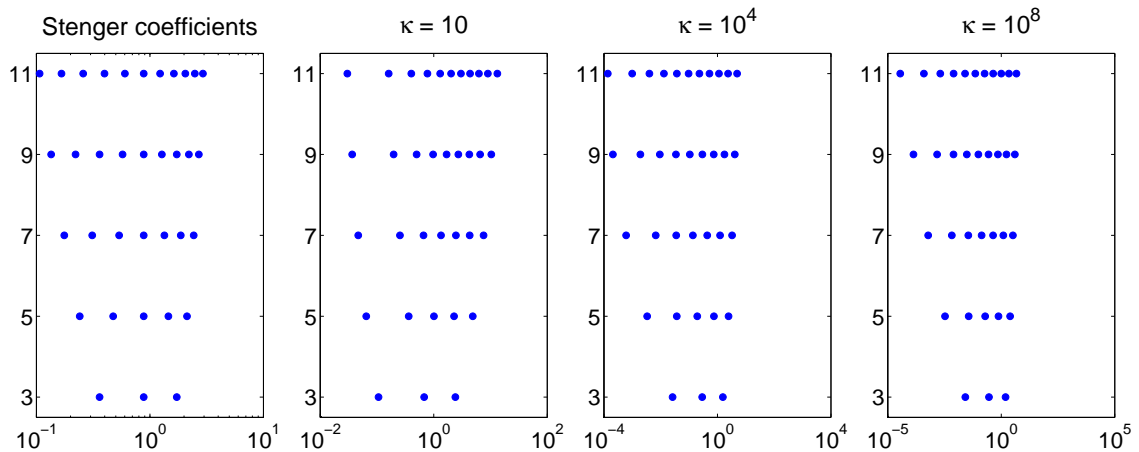

FIG. 5.1. *Coefficients α_j proposed in Section 5.2 (leftmost plot) and Section 5.1 (3 rightmost plots, for $\kappa(\mathcal{H}) = 10, 10^4, 10^8$), for $t = 3, 5, \ldots, 11$ coefficients.*

For both choices of coefficients, the choice of t needs to be determined in advance. We choose t such that the convergence bounds given below are not larger than a tolerance provided by the user. As this tolerance determines the overall quality of the approximation to the large linear system (1.3), it will usually be chosen rather small, say 10^{-9}.

14

5.1. Symmetric $\mathcal{H}$ and condition number known. As already mentioned in Section 2, the existence of coefficients $\alpha_j > 0, \omega_j > 0$ satisfying

$$\left| \frac{1}{\mu} - \sum_{j=1}^{t} \omega_j e^{-\alpha_j \mu} \right| \leq 16 \exp\left(\frac{-t\pi^2}{\log(8R)} \right) \quad \forall \mu \in [1, R]$$

is proved in [11]. By Corollary 2.7,

$$\|y - y_t\|_2 \leq \frac{16}{\lambda_{\min}(\mathcal{H})} \exp\left(\frac{-t\pi^2}{\log(8\kappa(\mathcal{H}))} \right) \|\tilde{b}\|_2.$$

Unfortunately, there is no explicit formula for determining such coefficients $\alpha_j > 0, \omega_j > 0$. The optimal choice of coefficients can be calculated by a Remez-like algorithm for specific values of t and R. This has been performed for a large set of pairs (t, R) by Hackbusch [10] and the resulting coefficients were used in this paper.

5.2. Condition number unknown and/or non-symmetric $\mathcal{H}$. If Ω is a subset of the right-half complex plane, the following bound from [12, Appendix D3.4.1] applies:

$$\left| \frac{1}{\mu} - \sum_{j=-t}^{t} \omega_j e^{-\alpha_j \mu} \right| \leq C_{\mathrm{St}} \exp\left(|\Im(\mu)|/\pi \right) \exp(-\pi\sqrt{t}), \quad \forall \mu \in \mathbb{C}, \Re(\mu) \geq 1,$$

where C_{St} does not depend on t or μ. This yields

$$\|y - y_{2t+1}\|_2 \leq C_{\mathrm{St}} \frac{\kappa_{\mathcal{P}}}{\beta} \exp(\gamma/\pi) \exp(-\pi\sqrt{t}) \|\tilde{b}\|_2,$$

where $\beta = \min\{\Re(\lambda) : \lambda \in \Lambda(\mathcal{H})\}$, $\gamma = \max\{|\Im(\lambda)| : \lambda \in \Lambda(\mathcal{H})\}$ and $\kappa_{\mathcal{P}}$ as defined in Theorem 2.5. In this case, there are explicit formulas for suitable coefficients:

$$h_{\mathrm{St}} = \pi/\sqrt{t},$$
$$\alpha_j = \log\left(\exp(jh_{\mathrm{St}}) + \sqrt{1 + \exp(2jh_{\mathrm{St}})} \right),$$
$$\omega_j = h_{\mathrm{St}}\left(1 + \exp(-2jh_{\mathrm{St}})\right)^{-1/2}, \text{ for } j = -t, \ldots, t.$$

6. An extended tensor Krylov subspace method. In many cases of practical interest, the convergence of the tensor Krylov subspace method can be enhanced by including matrix-vector products with A_s^{-1} for all or some s, for example by means of sparse direct factorizations. Although the computational cost for such an operation is significantly higher than a matrix-vector multiplication, this may be offset by accelerated convergence. Moreover, the tensor rank of the resulting approximate solution is reduced, potentially offering advantages to subsequent computations.

In the following we propose such an *extended tensor Krylov subspace method*, in the spirit of the closely related extended Krylov subspace methods for matrix functions [8] and linear matrix equations [22, 14]. The convergence of this method for approximating matrix functions has been recently discussed in [3, 18].

In contrast to the tensor Krylov subspace method described in Section 3, some (or all) of the matrices U_s now represent orthonormal bases for the *extended* Krylov subspace

$$\widetilde{\mathcal{K}}_{k_s}(A_s, b_s) := \mathcal{K}_{k_s}(A_s, b_s) + \mathcal{K}_{k_s+1}(A_s^{-1}, b_s),$$

assuming of course that A_s is invertible. Generically, the dimension of the extended Krylov subspace is $2k_s$ and hence $U_s \in \mathbb{R}^{n_s \times 2k_s}$. Arnoldi-type algorithms for computing U_s are discussed, for example, in [18, 22]. The rest of the extended tensor Krylov subspace method is identical with Algorithm 2.

For simplicity, we assume that all A_s are symmetric positive definite and extended Krylov subspaces are used for all $s = 1, \ldots, d$. Then

$$\widetilde{\mathcal{K}}_{k_s}(A_s, b_s) = \text{span}\{A_s^{-k_s}b_s, \ldots, A_s^{-1}b_s, b_s, A_s b_s, \ldots, A_s^{k_s-1}b_s\}$$
$$= \text{span}\{\ell(A_s)b_s : \ell \in \mathbb{L}_{k_s}\},$$

where $\mathbb{L}_{k_s}$ denotes the linear space of Laurent polynomials $\ell(\mu) = \sum_{j=-k_s}^{k_s-1} c_j \mu^j$. Similarly for the *tensorized extended Krylov subspace* it holds that

$$\widetilde{\mathcal{K}}_{\mathfrak{K}}^{\otimes}(\mathcal{A}, b) := \text{span}\big(\widetilde{\mathcal{K}}_{k_1}(A_1, b_1) \otimes \cdots \otimes \widetilde{\mathcal{K}}_{k_d}(A_d, b_d)\big)$$
$$= \text{span}\{\ell(A_1, \ldots, A_s)b : \ell \in \mathbb{L}_{k_s}^{\otimes}\}. \tag{6.1}$$

Here, $L_{k_s}^{\otimes}$ denotes the space of all multivariate Laurent polynomials

$$\ell(\mu_1, \ldots, \mu_d) = \sum_{-\mathfrak{K} \le \mathfrak{L} < \mathfrak{K}} c_{\mathfrak{L}} \mu_1^{l_1} \mu_2^{l_2} \cdots \mu_d^{l_d}, \quad c_{\mathfrak{L}} \in \mathbb{R},$$

where $-\mathfrak{K} \le \mathfrak{L} < \mathfrak{K}$ is to be understood as $-k_s \le l_s \le k_s - 1$ for $s = 1, \ldots, d$. The evaluation of ℓ at a matrix tuple $(A_1, \ldots, A_d)$ is then – analogously to (3.2) – defined as

$$\ell(A_1, \ldots, A_d) = \sum_{-\mathfrak{K} \le \mathfrak{L} < \mathfrak{K}} c_{\mathfrak{L}} A_1^{l_1} \otimes A_2^{l_2} \otimes \cdots \otimes A_d^{l_d}, \quad c_{\mathfrak{L}} \in \mathbb{R}.$$

The identity (6.1) can be shown in a similar way as Lemma 3.2.

Much of the convergence analysis of Section 4 can be extended in a straightforward way. In particular, the convergence bound of Theorem 4.3 becomes

$$\|x_{\mathfrak{K}} - x\|_{\mathcal{A}} \le \sqrt{\|\mathcal{A}\|_2} \|b\|_2 \min_{\ell \in \mathbb{L}_{\mathfrak{K}}^{\otimes}} \max_{\lambda_s \in \Lambda(A_s)} \left| \ell(\lambda_1, \ldots, \lambda_d) - \frac{1}{\lambda_1 + \cdots + \lambda_d} \right|$$
$$\le \sqrt{\|\mathcal{A}\|_2} \|b\|_2 \min_{\ell \in \mathbb{L}_{\mathfrak{K}}^{\otimes}} \max_{\mu_s \in [\alpha_s, \beta_s]} \left| \ell(\mu_1, \ldots, \mu_d) - \frac{1}{\mu_1 + \cdots + \mu_d} \right|, \tag{6.2}$$

where $[\alpha_s, \beta_s] = [\lambda_{\min}(A_s), \lambda_{\max}(A_s)]$. To proceed further, we need to address the multivariate Laurent polynomial approximation problem

$$\widetilde{E}_{\Omega} := \min_{\ell \in \mathbb{L}_{\mathfrak{K}}^{\otimes}} \left\| \ell(\mu_1, \ldots, \mu_d) - \frac{1}{\mu_1 + \cdots + \mu_d} \right\|_{\Omega}, \tag{6.3}$$

where $\| \cdot \|_{\Omega}$ denotes the supremum norm on $\Omega := [\alpha_1, \beta_1] \times \cdots \times [\alpha_d, \beta_d]$.

LEMMA 6.1. *Consider $\widetilde{E}_{\Omega}$ defined in (6.3) for constant $\alpha_s \equiv \alpha$, $\beta_s \equiv \beta$, and $\mathfrak{K} = (k, \ldots, k)$. Then*

$$\widetilde{E}_{\Omega} \le \left(1 + \frac{\beta}{\alpha}\right) \frac{1}{\tilde{\alpha}} \frac{\sqrt{\tilde{\kappa}} + 1}{\sqrt{\tilde{\kappa}}} \cdot \left(\frac{\sqrt{\tilde{\kappa}} - 1}{\sqrt{\tilde{\kappa}} + 1}\right)^k,$$

where
1. for $d = 2$: $\tilde{\alpha} = 2\sqrt{\alpha\beta}$, $\tilde{\beta} = \alpha + \beta$, $\tilde{\kappa} = (\tilde{\alpha} + \tilde{\beta})/(2\tilde{\alpha})$;
2. for $d > 2$: $\tilde{\alpha} = d^2(d-1)^{\frac{1-d}{d}} \alpha^{\frac{d-1}{d}} \beta^{\frac{1}{d}}$, $\tilde{\beta} = d(\alpha + \beta)$, $\tilde{\kappa} = \tilde{\beta}/\tilde{\alpha}$.
Proof. Part 1. Setting $\xi_s(\mu_s) = \mu_s + \gamma/\mu_s$ for $s = 1, 2$, we have

$$\frac{1}{\mu_1 + \mu_2} = \left(1 + \gamma\mu_1^{-1}\mu_2^{-1}\right) \frac{1}{\xi_1(\mu_1) + \xi_2(\mu_2)}. \tag{6.4}$$

The choice $\gamma = \alpha\beta$ balances the maximal value of ξ_s at the boundaries of the interval: $\xi_s(\alpha) = \xi_s(\beta) = \alpha + \beta = \tilde{\beta}$. It is straightforward to see that the minimum of ξ_s is attained at $\sqrt{\gamma}$ with

16

$\xi_s(\sqrt{\gamma}) = 2\sqrt{\alpha\beta} = \tilde{\alpha}$. Hence $\tilde{\alpha} \leq \xi_s(\mu_s) \leq \tilde{\beta}$ for $\mu_s \in [\alpha, \beta]$. By Lemma A.1 there is a bivariate polynomial $p(\xi_1, \xi_2) \in \Pi_{\mathfrak{K}}^{\otimes}$ such that

$$E_{\widetilde{\Omega}} := \left\| p(\xi_1, \xi_2) - \frac{1}{\xi_1 + \xi_2} \right\|_{\widetilde{\Omega}} \leq \frac{1}{\tilde{\alpha}} \frac{\sqrt{\tilde{\kappa}} + 1}{\sqrt{\tilde{\kappa}}} \cdot \left(\frac{\sqrt{\tilde{\kappa}} - 1}{\sqrt{\tilde{\kappa}} + 1} \right)^k,$$

where $\widetilde{\Omega} = [\tilde{\alpha}, \tilde{\beta}]^d$ and $\tilde{\kappa}$ is defined as in the statement of the lemma. By (6.4), the bivariate Laurent polynomial $\ell(\mu_1, \mu_2) := \left(1 + \gamma\mu_1^{-1}\mu_2^{-1}\right)p\left(\xi_1(\mu_1), \xi_2(\mu_2)\right) \in \mathbb{L}_{\mathfrak{K}}^{\otimes}$ satisfies

$$\widetilde{E}_{\Omega} \leq \left\| \ell(\mu_1, \mu_2) - \frac{1}{\mu_1 + \mu_2} \right\|_{\Omega} \leq \left\| 1 + \gamma\mu_1^{-1}\mu_2^{-1} \right\|_{\Omega} E_{\widetilde{\Omega}} = (1 + \beta/\alpha)E_{\widetilde{\Omega}},$$

which concludes the proof of Part 1.

Part 2. Trivially, $(\mu_1 + \cdots + \mu_d)^{-1} = \left(1 + \gamma\prod_{s=1}^d \mu_s^{-1}\right)\xi(\mu_1, \ldots, \mu_d)^{-1}$, with

$$\xi(\mu_1, \ldots, \mu_d) = \left(1 + \gamma\prod_{s=1}^d \mu_s^{-1}\right)\sum_{s=1}^d \mu_s = \sum_{s=1}^d \mu_s + \gamma\sum_{s=1}^d \prod_{j \neq s} \mu_j^{-1}$$

for some $\gamma > 0$. Note that ξ is a sum of convex functions and therefore itself convex. The maximum of ξ is attained at the extrema of the convex polytope $[\alpha, \beta]^d$. The choice $\gamma = \alpha^{d-1}\beta$ approximately balances the values at the extrema $\mu_s \equiv \alpha$ and $\mu_s \equiv \beta$, which leads to

$$\max_{\mu_s \in \{\alpha, \beta\}} \xi(\mu_1, \ldots, \mu_d) = \xi(\alpha, \ldots, \alpha) = d(\alpha + \beta) = \tilde{\beta}.$$

Since $\operatorname{grad}\xi = 0$ at $\mu_s \equiv \left((d-1)\gamma\right)^{1/d}$, the minimum of ξ is given by $d^2(d-1)^{\frac{1-d}{d}}\gamma^{1/d} = \tilde{\alpha}$. By Lemma A.1 there is a polynomial p of degree at most $k - 1$ such that

$$E_{\widetilde{\Omega}} := \left\| p(\xi) - \xi^{-1} \right\|_{[\tilde{\alpha}, \tilde{\beta}]} \leq \frac{1}{\tilde{\alpha}} \frac{\sqrt{\tilde{\kappa}} + 1}{\sqrt{\tilde{\kappa}}} \cdot \left(\frac{\sqrt{\tilde{\kappa}} - 1}{\sqrt{\tilde{\kappa}} + 1} \right)^k.$$

Similarly as in the proof of Part 1, the multivariate Laurent polynomial $\ell(\mu_1, \ldots, \mu_d) := \left(1 + \gamma\prod_{s=1}^d \mu_s^{-1}\right)p(\xi(\mu_1, \ldots, \mu_d)) \in \mathbb{L}_{\mathfrak{K}}^{\otimes}$ yields

$$\widetilde{E}_{\Omega} \leq \left\| \ell(\mu_1, \ldots, \mu_d) - \frac{1}{\mu_1 + \cdots + \mu_d} \right\|_{\Omega} \leq \left\| 1 + \gamma\prod_{s=1}^d \mu_s^{-1} \right\|_{\Omega} E_{\widetilde{\Omega}} = \left(1 + \beta/\alpha\right)E_{\widetilde{\Omega}},$$

which concludes the proof. $\square$

For $d = 2$, the factor $\tilde{\kappa}$ that determines the asymptotics of the convergence bound in Lemma 6.1 satisfies $\tilde{\kappa} \approx \sqrt{\kappa(\mathcal{A})}/4$ for $\beta \gg \alpha$. This compares favorably with the factor $\kappa = \kappa(\mathcal{A})/2$ that determines the asymptotics of the convergence bound (4.5) for the (standard) tensor Krylov subspace method. The linear convergence rate of the extended Krylov subspace method for solving Sylvester equations is therefore bounded by $\sqrt{\tilde{\kappa}} \approx \kappa(\mathcal{A})^{1/4}/2$, which was also observed experimentally in [22]. For $d = 3$, $\tilde{\kappa} \approx \left(2\kappa(\mathcal{A})\right)^{2/3}/3$ for $\beta \gg \alpha$. As d increases, the bound of Lemma 6.1 becomes more pessimistic: $\tilde{\kappa} \overset{d\to\infty}{\to} 1 + \kappa(\mathcal{A})$. It is not clear to us whether the bound of Lemma 6.1 could be improved to also reflect the significantly better convergence of the extended tensor Krylov subspace method we observed for higher dimensions.

7. Numerical experiments. We have implemented the (extended) tensor Krylov subspace methods in MATLAB, using the Tensor Toolbox [1, 2] for storing tensors in CP decomposition and for multiplying matrices with tensors. For solving the compressed systems, the coefficients described in Section 5 are used. A tolerance of $\varepsilon = 10^{-9}$ is chosen as an upper bound on the accuracy of the solution to the compressed system.

7.1. Symmetric example. As a first example, we consider the d-dimensional Poisson equation (1.6) with separable right-hand side $f(\mu_1, \mu_2, \ldots, \mu_d) = g(\mu_1)g(\mu_2)\cdots g(\mu_d)$. A standard finite-difference discretization on equidistant nodes leads to the linear system $\mathcal{A}x = b$, with

$$
A_s = \frac{1}{h^2}\begin{pmatrix}
2 & -1 & & & \\
-1 & 2 & -1 & & \\
& \ddots & \ddots & \ddots & \\
& & -1 & 2 & -1 \\
& & & -1 & 2
\end{pmatrix}, \quad
b_s = \begin{pmatrix}
g(z_1^{(s)}) \\
\vdots \\
g(z_n^{(s)})
\end{pmatrix}.
$$

It is well-known that the minimal and maximal eigenvalues of A_s are given by

$$
\lambda_{\min} = \frac{2}{h^2}\left(1 - \cos\left(\frac{\pi}{n_s + 1}\right)\right), \quad \lambda_{\max} = \frac{2}{h^2}\left(1 - \cos\left(\frac{\pi n_s}{n_s + 1}\right)\right),
$$

resulting in the convergence factor (see Corollary 4.4),

$$
\kappa = 1 + \frac{\lambda_{\max} - \lambda_{\min}}{d\lambda_{\min}} = 1 + \frac{2\cos(\frac{\pi}{n+1})}{d(1 - \cos(\frac{\pi}{n+1}))} \approx \frac{4(n+1)^2}{\pi^2 d}. \tag{7.1}
$$

For simplicity, we have used right-hand side vectors b_s composed of uniformly distributed pseudo-random numbers.

The tensor Krylov subspace method was used with multi-index $\mathfrak{K} = (k, \ldots, k)$, as the size n_s and condition number are identical for all A_s. All convergence plots display the convergence of the *relative residual* $\|\mathcal{A}x_k - b\|_2/\|b\|_2$. As the solution x_k cannot be stored explicitly, the norm of the residual must be calculated directly from its CP decomposition. For this purpose, an efficient method is proposed in Section 3.3. As Lemma 3.4 shows, the error is ultimately bounded by the residual induced by the approximation error from the compressed system.

Figures 7.1 and 7.2 show the convergence of the tensor Krylov subspace method for systems of size $n_s = 200$ and $n_s = 1000$, respectively, with $s = 1, \ldots, d$ and various dimensions d. The observed convergence rates correspond reasonably well to the theoretically predicted convergence rates; in particular, the convergence rates improve for higher dimensions. The plots in Figures 7.1 and 7.2 are remarkably similar apart from the different scaling of the x-axis, caused by the increase of the condition number as n_s grows. The convergence curves also nicely confirm the fact that $x_k = x$ holds for $k \equiv n_s$ up to the approximation error from the compressed system.

In Figure 7.3, we apply the extended tensor Krylov subspaces method to a system of size $n_s = 200$. At $k = 40$, the method converges up to the approximation error in the compressed system for all dimensions. The observed convergence is significantly better than the convergence rate $\left(\frac{\sqrt{\tilde{\kappa}}-1}{\sqrt{\tilde{\kappa}}+1}\right)^k$ predicted by Lemma 6.1. We suspect that this can be attributed to eigenvalues converging at both ends of the spectrum rather quickly, leading to a rapid decrease of the effective condition number in the course of the iteration.

7.2. Non-symmetric example. We next consider the convection-diffusion equation

$$
\begin{aligned}
-\Delta u + c^\top \nabla u &= f \quad \text{in } \Omega = [0,1]^d \\
u &= 0 \quad \text{on } \Gamma := \partial\Omega,
\end{aligned}
$$

where f is again a separable function. A standard finite-difference discretization on equidistant nodes, combined now with a second order convergent scheme for the convection term, leads to

$$
A_s = \frac{1}{h^2}\begin{pmatrix}
2 & -1 & & & \\
-1 & 2 & -1 & & \\
& \ddots & \ddots & \ddots & \\
& & -1 & 2 & -1 \\
& & & -1 & 2
\end{pmatrix} + \frac{c_s}{4h}\begin{pmatrix}
3 & -5 & 1 & & \\
1 & 3 & -5 & \ddots & \\
& \ddots & \ddots & \ddots & 1 \\
& & 1 & 3 & -5 \\
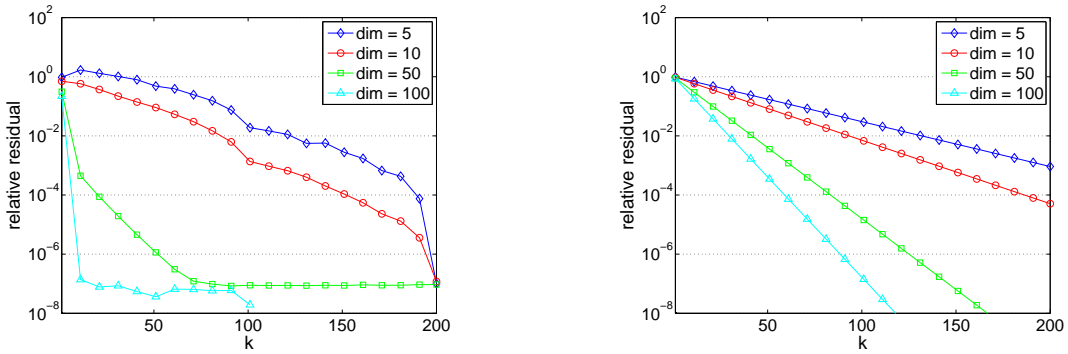& & & 1 & 3
\end{pmatrix}.
$$

FIG. 7.1. *Relative residual of the tensor Krylov subspace method for symmetric example and $n_s = 200$ (left). Predicted convergence rate $\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k$ (right).*
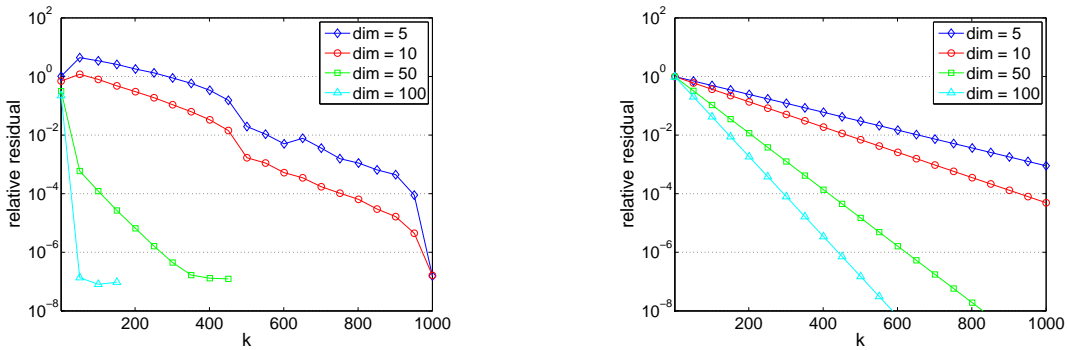

FIG. 7.2. *Relative residual of the tensor Krylov subspace method for symmetric example and $n_s = 1000$ (left). Predicted convergence rate $\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k$ (right).*

Note that $\mathcal{A}$ is identical with the system matrix in [9, Example 22]. Again, the right-hand side is composed of vectors b_s containing uniformly distributed pseudo-random numbers. Since we have chosen the system size $n_s = 200$ to be constant over all dimensions, it is again reasonable to choose $\mathfrak{K} = \{k, \ldots, k\}$.

Figure 7.4 displays the convergence of the tensor Krylov subspace method applied to this example for various dimensions. The convection coefficients are $c_s = 10$ for the left plot and significantly larger, $c_s = 100$, for the right plot. The convergence is significantly slower compared to the symmetric case but the qualitative behavior is similar: the convergence rate improves for higher dimensions. As expected, the convergence rate slightly deteriorates as the convection coefficient grows.

We have also performed experiments with the extended tensor Krylov subspace method applied to the non-symmetric example. The observed convergence rates improve similarly as for the symmetric example.

8. Conclusions. A tensor Krylov subspace method has been proposed that deals well with high-dimensional linear systems that exhibit a certain, rather particular Kronecker product structure. We envision the methods developed in this paper as building blocks in future algorithms for addressing more general high-dimensional problems. For example, they could be used in a preconditioned inverse iteration for high-dimensional PDE eigenvalue problems [13].

9. Acknowledgments. The authors thank Lars Grasedyck for helpful and inspiring discussions on the subject of this paper and both referees for carefully checking the paper and providing a number of helpful remarks.
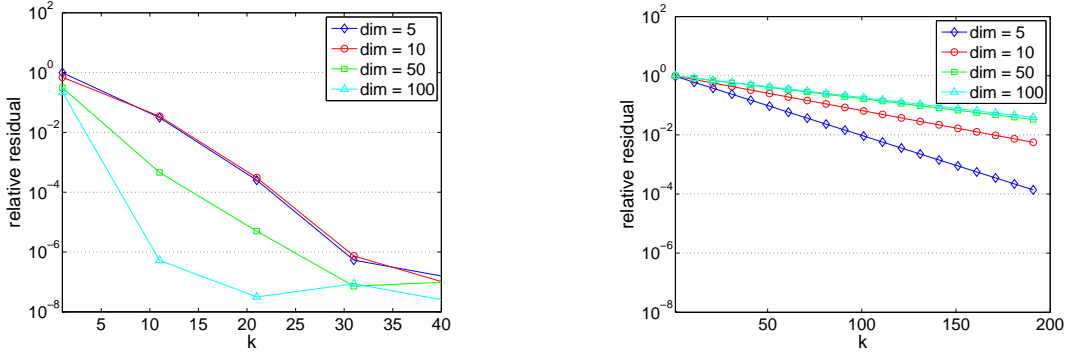
Appendix A. Appendix.

19

Fig. 7.3. *Relative residual of the extended tensor Krylov subspace method for symmetric example and $n_s = 200$ (left). Predicted convergence rate $\left(\frac{\sqrt{\tilde{\kappa}}-1}{\sqrt{\tilde{\kappa}}+1}\right)^k$ (right)*
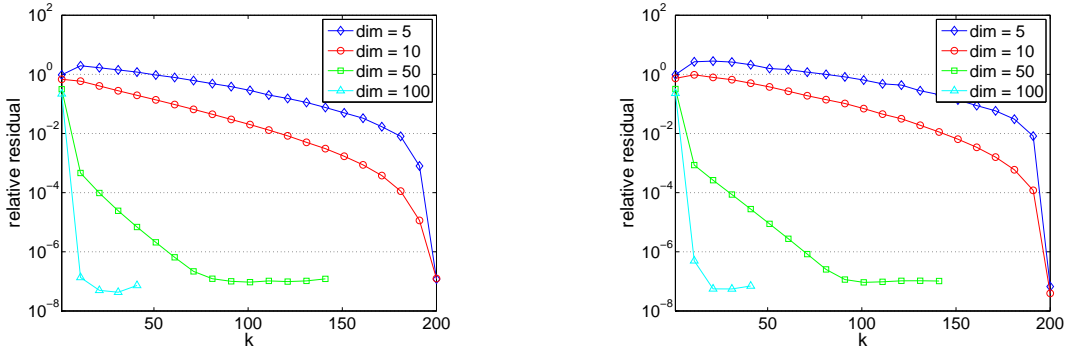

Fig. 7.4. *Relative residual of the tensor Krylov subspace method for non-symmetric example and $n_s = 200$, with factor $c_s = 10$ (left) and $c_s = 100$ (right).*

The following lemma contains the approximation result needed in Section 4 for the convergence analysis of the tensor Krylov subspace method.

Lemma A.1. *For the multivariate polynomial approximation error $E_\Omega(\mathfrak{K})$ defined in (4.3) it holds that*

$$E_\Omega(\mathfrak{K}) \leq \frac{1}{\lambda_{\min}(\mathcal{A})} \sum_{s=1}^{d} \frac{\sqrt{\kappa_s}+1}{\sqrt{\kappa_s}} \cdot \left(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\right)^{k_s},$$

where $\kappa_s = 1 + \frac{\beta_s - \alpha_s}{\lambda_{\min}(\mathcal{A})}$.

Proof. Since $\mathcal{A}$ is assumed to be symmetric positive definite, $\alpha_1 + \cdots + \alpha_d > 0$ and hence every $(\mu_1, \ldots, \mu_d) \in \Omega$ also satisfies $\mu_1 + \cdots + \mu_d > 0$. This allows us to write

$$\frac{1}{\mu_1 + \cdots + \mu_d} = \int_0^\infty e^{-(\mu_1 + \cdots + \mu_d)t} \, \mathrm{d}t = \int_0^\infty \prod_{s=1}^d e^{-\mu_s t} \, \mathrm{d}t.$$

Following [7, 23], we expand $e^{-\mu_s t}$ in terms of Chebyshev series:

$$e^{-\mu_s t} = e^{-\frac{\alpha_s + \beta_s}{2}t} \sum_{j=0}^\infty a_j^{(s)}(t) \, T_j(-g_s(\mu_s)), \tag{A.1}$$

where T_j denotes the jth Chebyshev polynomial of the first kind, and the parameter transformations $g_s : [\alpha_s, \beta_s] \to [-1, 1]$ are defined as

$$g_s(\mu_s) = \frac{2}{\beta_s - \alpha_s} \mu_s - \frac{\beta_s + \alpha_s}{\beta_s - \alpha_s}$$

20

The Chebyshev coefficients $a_j^{(s)}$ take the form

$$a_j^{(s)}(t) = \begin{cases} I_j\left(\frac{\beta_s - \alpha_s}{2}t\right), & j = 0, \\ 2I_j\left(\frac{\beta_s - \alpha_s}{2}t\right), & \text{otherwise,} \end{cases}$$

where I_j denotes the jth modified Bessel function. We define polynomials $p_{k_s}^{(s)}(\mu_s, t)$ in μ_s by truncating the series (A.1) after the $(k_s - 1)$th term. Obviously,

$$\left\|p_{k_s}^{(s)}(\mu_s, t) - e^{-\mu_s t}\right\|_\Omega \leq e^{-\frac{\alpha_s + \beta_s}{2}t} \sum_{j=k_s}^{\infty} a_j^{(s)}(t). \tag{A.2}$$

Moreover,

$$\left\|p_{k_s}^{(s)}(\mu_s, t)\right\|_\Omega \leq \left|p_{k_s}^{(s)}(\alpha_s, t)\right| = e^{-\frac{\alpha_s + \beta_s}{2}t} \sum_{j=0}^{k_s - 1} a_j^{(s)}(t) \leq e^{-\frac{\alpha_s + \beta_s}{2}t} \sum_{j=0}^{\infty} a_j^{(s)}(t) = e^{-\alpha_s t}.$$

Hence,

$$\left\|\prod_{s=1}^{d} e^{-\mu_s t} - \prod_{s=1}^{d} p_{k_s}^{(s)}(\mu_s, t)\right\|_\Omega = \left\|\sum_{s=1}^{d}\left[\prod_{j=1}^{s-1} e^{-\mu_j t}\right]\left(e^{-\mu_s t} - p_{k_s}^{(s)}(\mu(s), t)\right)\left[\prod_{j=s+1}^{d} p_{k^{(j)}}^{(j)}(\mu_j, t)\right]\right\|_\Omega$$

$$\leq \sum_{s=1}^{d}\left\|e^{-\mu_s t} - p_{k_s}^{(s)}(\mu_s, t)\right\|_\Omega \prod_{\substack{j=1 \\ j \neq s}}^{d} e^{-\alpha_j t}$$

$$= \sum_{s=1}^{d}\left\|e^{-\mu_s t} - p_{k_s}^{(s)}(\mu_s, t)\right\|_\Omega e^{-(\lambda_{\min}(\mathcal{A}) - \alpha_s)t} \tag{A.3}$$

To use these results for our multivariate interpolation problem, we set

$$p(\mu_1, \ldots, \mu_d) = \sum_{\mathfrak{I} \leq \mathfrak{K}} c_{\mathfrak{I}}\, p_{i_1}^{(1)}(\mu_1) p_{i_2}^{(2)}(\mu_2) \cdots p_{i_d}^{(d)}(\mu_d),$$

where $\mathfrak{I} = (i_1, \ldots, i_d)$ with $1 \leq i_s \leq k_s$, and

$$p_{i_s}^{(s)}(\mu_s) = T_{i_s - 1}(-g_s(\mu_s)), \quad c_{\mathfrak{I}} = \int_0^\infty \prod_{s=1}^{d} e^{-\frac{\alpha_s + \beta_s}{2}t} a_{i_s - 1}^{(s)}(t)\, \mathrm{d}t.$$

21

With this choice,

$$
\begin{aligned}
E_\Omega(\mathfrak{K}) \;&\leq\; \Big\| p(\mu_1,\ldots,\mu_s) - \frac{1}{\mu_1 + \cdots + \mu_s} \Big\|_\Omega \\
&=\; \Big\| \sum_{\mathfrak{I}\leq\mathfrak{K}} c_{\mathfrak{I}} \prod_{s=1}^{d} p_{i_s}^{(s)}(\mu_s) - \frac{1}{\mu_1 + \cdots + \mu_s} \Big\|_\Omega \\
&=\; \Big\| \int_0^\infty \sum_{\mathfrak{I}\leq\mathfrak{K}} \prod_{s=1}^{d} e^{-\frac{\alpha_s+\beta_s}{2}t} a_{i_s-1}^{(s)}(t) p_{i_s}^{(s)}(\mu_s) \,\mathrm{d}t - \int_0^\infty \prod_{s=1}^{d} e^{-\mu_s t} \,\mathrm{d}t \Big\|_\Omega \\
&=\; \Big\| \int_0^\infty \prod_{s=1}^{d} \underbrace{\sum_{i_s=1}^{k_s} e^{-\frac{\alpha_s+\beta_s}{2}t} a_{i_s-1}^{(s)}(t) p_{i_s}^{(s)}(\mu_s)}_{=p_{k_s}^{(s)}(\mu_s,t)} \,\mathrm{d}t - \int_0^\infty \prod_{s=1}^{d} e^{-\mu_s t} \,\mathrm{d}t \Big\|_\Omega . \\
&\leq\; \int_0^\infty \Big\| \prod_{s=1}^{d} p_{k_s}^{(s)}(\mu_s,t) - \prod_{s=1}^{d} e^{-\mu_s t} \Big\|_\Omega \,\mathrm{d}t \\
&\overset{(A.3)}{\leq}\; \sum_{s=1}^{d} \int_0^\infty e^{-(\lambda_{\min}(\mathcal{A})-\alpha_s)t} \big\| e^{-\mu_s t} - p_{k_s}^{(s)}(\mu_s,t) \big\|_\Omega \,\mathrm{d}t \\
&\overset{(A.2)}{\leq}\; \sum_{s=1}^{d} \underbrace{\sum_{j=k_s}^{\infty} \int_0^\infty e^{-(\lambda_{\min}(\mathcal{A})+\frac{\beta_s-\alpha_s}{2})t} a_j^{(s)}(t) \,\mathrm{d}t}_{=:E_s(k_s)} .
\end{aligned}
\tag{A.4}
$$

Using $\int_0^\infty e^{-\delta t} I_j(\gamma t)\,\mathrm{d}t = \frac{\gamma^j}{\sqrt{\delta^2-\gamma^2}(\delta+\sqrt{\delta^2-\gamma^2})^j}$, which holds for $\delta > \gamma$ [23] and $j \neq 0$, we obtain after some algebraic manipulation

$$
\int_0^\infty e^{-\left(\lambda_{\min}(\mathcal{A})+\frac{\beta_s-\alpha_s}{2}\right)t} 2 I_j\Big(\frac{\beta_s-\alpha_s}{2}t\Big) \,\mathrm{d}t = \frac{2}{\lambda_{\min}(\mathcal{A})\sqrt{\kappa_s}} \cdot \Big(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\Big)^j
$$

with $\kappa_s = 1 + \frac{\beta_s-\alpha_s}{\lambda_{\min}(\mathcal{A})}$. Hence,

$$
E_s(k_s) = \frac{2}{\lambda_{\min}(\mathcal{A})\sqrt{\kappa_s}} \sum_{j=k_s}^{\infty} \Big(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\Big)^j = \frac{\sqrt{\kappa_s}+1}{\lambda_{\min}(\mathcal{A})\sqrt{\kappa_s}} \cdot \Big(\frac{\sqrt{\kappa_s}-1}{\sqrt{\kappa_s}+1}\Big)^{k_s}
$$

which, together with (A.4), concludes the proof. $\square$

REFERENCES

[1] B. W. Bader and T. G. Kolda. Efficient MATLAB computations with sparse and factored tensors. *SIAM J. Sci. Comput.*, 30(1):205–231, 2007.

[2] B. W. Bader and T. G. Kolda. MATLAB tensor toolbox version 2.2, January 2007. Available from http://csmr.ca.sandia.gov/~tgkolda/TensorToolbox/.

[3] B. Beckermann and L. Reichel. Error estimation and evaluation of matrix functions via the Faber transform. Technical report, Université de Lille, 2008.

[4] G. Beylkin and M. J. Mohlenkamp. Algorithms for numerical analysis in high dimensions. *SIAM J. Sci. Comput.*, 26(6):2133–2159, 2005.

[5] S. Börm. $\mathcal{H}_2$-matrices – an efficient tool for the treatment of dense matrices. Habilitationsschrift, Christian-Albrechts-Universität zu Kiel, 2006.

[6] D. Braess and W. Hackbusch. Approximation of $1/x$ by exponential sums in $[1,\infty)$. *IMA J. Numer. Anal.*, 25(4):685–697, 2005.

[7] V. L. Druskin and L. A. Knizhnerman. Two polynomial methods for calculating functions of symmetric matrices. *Zh. Vychisl. Mat. i Mat. Fiz.*, 29(12):1763–1775, 1989.

[8] V. L. Druskin and L. A. Knizhnerman. Extended Krylov subspaces: approximation of the matrix square root and related functions. *SIAM J. Matrix Anal. Appl.*, 19(3):755–771, 1998.

[9] L. Grasedyck. Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure. *Computing*, 72(3-4):247–265, 2004.

[10] W. Hackbusch. Approximation of 1/x by exponential sums. Available from `http://www.mis.mpg.de/scicomp/EXP_SUM/1_x/tabelle`. Retrieved August 2008.

[11] W. Hackbusch. Entwicklungen nach Exponentialsummen. Technical Report, Max-Planck-Institut für Mathematik in den Naturwissenschaften, 2009. Revised version. See `http://www.mis.mpg.de/preprints/tr/report-0405.pdf`.

[12] W. Hackbusch. *Hierarchische Matrizen: Algorithmen und Analysis*. Springer, Berlin, 2009.

[13] W. Hackbusch, B. N. Khoromskij, S. A. Sauter, and E. E. Tyrtyshnikov. Use of tensor formats in elliptic eigenvalue problems. Preprint 78/2008, Max-Planck-Institut für Mathematik in den Naturwissenschaften, 2008.

[14] M. Heyouni. Extended Arnoldi methods for large Sylvester matrix equations. Technical report L.M.P.A., 2008.

[15] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[16] D. Y. Hu and L. Reichel. Krylov-subspace methods for the Sylvester equation. *Linear Algebra Appl.*, 172:283–313, 1992.

[17] I. M. Jaimoukha and E. M. Kasenally. Krylov subspace methods for solving large Lyapunov equations. *SIAM J. Numer. Anal.*, 31:227–251, 1994.

[18] L. A. Knizhnerman and V. Simoncini. A new investigation of the extended Krylov subspace method for matrix function evaluations. Technical report, 2008. To appear in Numer. Linear Algebra Appl.

[19] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[20] Y. Saad. Numerical solution of large Lyapunov equations. In *Signal processing, scattering and operator theory, and numerical methods (Amsterdam, 1989)*, volume 5 of *Progr. Systems Control Theory*, pages 503–511. Birkhäuser Boston, Boston, MA, 1990.

[21] Y. Saad. *Iterative Methods for Sparse Linear Systems, 2nd edition*. SIAM, Philadelpha, PA, 2003.

[22] V. Simoncini. A new iterative method for solving large-scale Lyapunov matrix equations. *SIAM J. Sci. Comput.*, 29(3):1268–1288, 2007.

[23] V. Simoncini and V. L. Druskin. Convergence analysis of projection methods for the numerical solution of large Lyapunov equations. *SIAM Journal on Numerical Analysis*, 47(2):828–843, 2009.

[24] G. Starke. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Numer. Math.*, 78(1):103–117, 1997.

[25] F. Stenger. *Numerical methods based on sinc and analytic functions*, volume 20 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1993.

[26] G. W. Stewart. *Matrix Algorithms. Vol. II*. SIAM, Philadelphia, PA, 2001. Eigensystems.

[27] C. Tobler. Krylov subspace methods for large linear systems with tensor product structure. Master's thesis, ETH Zürich, September 2008.