

MAJORIZATION-TYPE CLUSTER ROBUST BOUNDS FOR BLOCK FILTERS AND EIGENSOLVERS

MING ZHOU, MERICO E. ARGENTATI, ANDREW V. KNYAZEV, AND KLAUS NEYMEYR

ABSTRACT. Convergence analysis of block iterative solvers for Hermitian eigenvalue problems and the closely related research on properties of matrix-based signal filters are challenging, and attract increasing attention due to their recent applications in spectral data clustering and graph-based signal processing. We combine majorization-based techniques pioneered for investigating the Rayleigh-Ritz method in [SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1521–1537] with tools of classical analysis of the block power method by Rutishauser [Numer. Math., 13 (1969), pp. 4–13] to derive convergence rate bounds of an abstract block iteration, wherein tuples of tangents of principal angles or relative errors of Ritz values are bounded using majorization in terms of arranged partial sums and tuples of convergence factors. Our novel bounds are robust in presence of clusters of eigenvalues, improve some previous results, and are applicable to most known block iterative solvers and matrix-based filters, e.g., to block power, Chebyshev, and Lanczos methods combined with shift-and-invert approaches and polynomial filtering.

1. Introduction

Matrix eigenvalue problems have historically appeared in computational mechanics and then in quantum physics as a tool to approximate solutions of wave and other time-dependent differential equations. Recent applications in machine learning, e.g., for spectral clustering and image semantic segmentation, often require numerical solution of eigenvalue problems for matrices of large sizes and in some cases in real time, e.g., for autonomous driving. Only a fraction of the eigenvectors is of practical interest, where the eigenvectors need to be computed just with certain accuracy. Such computations of specific invariant subspaces can be efficiently performed by iterative eigenvalue problem solvers (eigensolvers) that simultaneously iterate several vectors in a block. Recently graph-based signal processing has gained attention. Therein one utilizes signal filters that are functions, often a polynomial like Chebyshev, of a matrix with a goal of amplifying the components of a signal corresponding to selected eigenvalues of the matrix. Such filters can be interpreted as rudimentary eigensolvers with a fixed accuracy. Block filters and solvers match well with modern hardware that can optimize computing matrix-matrix products, which commonly is the main operation in block algorithms.

The convergence speed of block iterative solvers, or the reduction/amplification quality of signal filters, determines the efficiency of calculations and needs to be estimated in advance a priori to bound the computational time for streaming data processing. Deriving a priori convergence rate bounds is a classical area of research. The convergence speed of block iterative eigensolvers is determined by the quality of an initial approximation to target eigenvectors and the distribution of eigenvalues of the given matrix. Block iterative eigensolvers are known to be robust with respect to possible clustering of target eigenvalues – this property is called “cluster

2010 *Mathematics Subject Classification*. Primary 65F15, 65N12, 65N25.

Key words and phrases. Majorization, principal angles, Ritz values, subspace eigensolvers, cluster robustness, block Lanczos method.

December 2021.

robustness". Sharp bounds of convergence rates of these eigensolvers thus also need to be cluster robust, which is the main topic of our investigation.

Majorization techniques appear naturally for cluster robust bounds involving eigenvalue approximations [12, 17]. We explore relations between Ritz value errors and principal angles in the context of subspace iterates of block iterative eigensolvers. We refine the majorization-type analysis of tangents of principal angles from [12] by using auxiliary vectors from the classical analysis of the block power method by Rutishauser [22]. Our new results applied to the block Lanczos method improve existing ones from [15, 23]. The improvement is evident due to the cluster robustness of our majorization based approach.

1.1. Known cluster robust error bounds

The cluster robustness is a typical feature of block iterations for solving matrix eigenvalue problems. A simple example is the block power method that aims at the largest magnitude eigenvalues of a matrix $A \in \mathbb{C}^{n \times n}$ and significantly improves the simple power method in the case where the target eigenvalues are clustered. The convergence of the subspace iterates $\mathcal{Y}^{(0)}, \mathcal{Y}^{(1)}, \dots$ given by $\mathcal{Y}^{(\ell+1)} = A\mathcal{Y}^{(\ell)}$ toward an invariant subspace can be analyzed in terms of various angle measures. Classical bounds by Rutishauser [22] and Parlett [21, Chapter 14] can be extended to a normal A with the eigenvalue arrangement $|\lambda_1| \geq \dots \geq |\lambda_n|$, namely,

$$(1.1) \quad \tan \angle(x_i, \mathcal{Y}^{(\ell)}) \leq \left| \frac{\lambda_{p+1}}{\lambda_i} \right|^\ell \tan \angle(\mathcal{X}, \mathcal{Y}^{(0)}), \quad i = 1, \dots, p.$$

Therein $\angle(\cdot, \cdot)$ denotes Euclidean angles, and the separation of targeted eigenvalues from the rest is described by the assumption $|\lambda_p| > |\lambda_{p+1}|$ for the block size $p = \dim \mathcal{Y}^{(0)}$. Moreover, \mathcal{X} denotes the invariant subspace spanned by orthonormal eigenvectors x_1, \dots, x_p associated with the first p eigenvalues. In the nontrivial case $\angle(\mathcal{X}, \mathcal{Y}^{(0)}) < \pi/2$, all further subspace iterates have the same dimension p since the first p eigenvalues are nonzero due to $|\lambda_p| > |\lambda_{p+1}| \geq 0$. Evidently, the convergence factor $|\lambda_{p+1}/\lambda_i|$ is robust with respect to possible clustering of eigenvalues indexed by $i \leq p$. In contrast, the convergence factor for computing x_i with the simple power method is $|\lambda_{i+1}/\lambda_i|$ which is unimprovable and might be close to 1. If $i > 1$, a deflation with respect to the first $i-1$ eigenvalues is required.

An essential argument by Rutishauser [22] adapted to (1.1) is that $\mathcal{Y}^{(\ell)}$ contains a nonzero vector which is orthogonal to an invariant subspace associated with $\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_p$. Thus λ_i and λ_{p+1} become the only two relevant eigenvalues in the derivation. This argument is directly applicable to the abstract block iteration

$$(1.2) \quad \mathcal{Y}' = f(A)\mathcal{Y}$$

that describes various practical methods. Since A and $f(A)$ share the same eigenvectors, one attempts to select a proper function $f(\cdot)$ such that the desired eigenvalues of A , if denoted by $\lambda_1, \dots, \lambda_p$, get mapped into the top p magnitude eigenvalues $f(\lambda_1), \dots, f(\lambda_p)$ of $f(A)$. Then (1.1) turns into

$$(1.3) \quad \tan \angle(x_i, \mathcal{Y}') \leq \left| \frac{f(\lambda_{p+1})}{f(\lambda_i)} \right| \tan \angle(\mathcal{X}, \mathcal{Y}), \quad i = 1, \dots, p$$

analogously to [7, Lemma 2.3.1]. One can use a shifted Chebyshev polynomial $f(\cdot)$ in order to investigate Chebyshev signal filters and the block Lanczos method. Let eigenvalues of a Hermitian matrix A be arranged as $\lambda_1 \geq \dots \geq \lambda_n$ with the assumption $\lambda_p > \lambda_{p+1}$, and $f(\cdot)$ be defined by

$$(1.4) \quad f(\alpha) = T_{k-1} \left(1 + 2 \frac{\alpha - \lambda_{p+1}}{\lambda_{p+1} - \lambda_n} \right)$$

with the Chebyshev polynomial T_{k-1} of the first kind with degree $k-1$. Then $|f(\lambda_i)| > 1 \geq |f(\lambda_j)|$ holds for arbitrary indices $i \leq p$ and $j > p$ so that the values $f(\lambda_1), \dots, f(\lambda_p)$ are indeed the top p magnitude eigenvalues of $f(A)$. Thus, substituting (1.4) in (1.3) immediately leads to

$$(1.5) \quad \tan \angle(x_i, \mathcal{Y}') \leq \left[T_{k-1} \left(1 + 2 \frac{\lambda_i - \lambda_{p+1}}{\lambda_{p+1} - \lambda_n} \right) \right]^{-1} \tan \angle(\mathcal{X}, \mathcal{Y}), \quad i = 1, \dots, p,$$

where the reciprocal Chebyshev term is the ratio $|f(\lambda_{p+1})/f(\lambda_i)|$. Subsequently, since f is a polynomial of degree $k-1$ by (1.4), the subspace \mathcal{Y}' is a subset of the block Krylov subspace $\mathcal{K} = \mathcal{Y} + A\mathcal{Y} + \dots + A^{k-1}\mathcal{Y}$ (spanned by k blocks) arising from the block Lanczos method. Thus $\tan \angle(x_i, \mathcal{K}) \leq \tan \angle(x_i, \mathcal{Y}')$ holds and extends (1.5) as a bound for $\tan \angle(x_i, \mathcal{K})$. The reciprocal Chebyshev term decreases rapidly with the degree k of \mathcal{K} provided that λ_i is not close to λ_{p+1} . This can be achieved by selecting an initial subspace of sufficiently high dimension.

Bound (1.5) demonstrates the quality of the shifted Chebyshev polynomial (1.4) as a filter in (1.2) amplifying the components of a signal $y \in \mathcal{Y}$ corresponding to the p largest eigenvalues of A . For computing these eigenvalues, one is also interested in bounds of the errors of the Ritz values of A in \mathcal{Y}' , arranged as $\eta'_1 \geq \dots \geq \eta'_p$. Some direct bounds have been presented by Knyazev in [8, Section 2]. In particular, applying [8, Theorem 2.5] to the specified function (1.4) yields the bound

$$(1.6) \quad \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \leq \left[T_{k-1} \left(1 + 2 \frac{\lambda_i - \lambda_{p+1}}{\lambda_{p+1} - \lambda_n} \right) \right]^{-2} \tan^2 \angle(\mathcal{X}, \mathcal{Y}), \quad i = 1, \dots, p$$

in the nontrivial case $\angle(\mathcal{X}, \mathcal{Y}) < \pi/2$. Although the convergence factor in (1.6) is the squared value of that in (1.5), the derivation of (1.6) is not simply based on (1.5). In addition, the relation $\mathcal{Y}' \subseteq \mathcal{K}$ leads to inequalities $\eta'_i \leq \psi_i$, $i = 1, \dots, p$ for the p largest Ritz values $\psi_1 \geq \dots \geq \psi_p$ of A in \mathcal{K} according to the Courant-Fischer principles. This extends (1.6) to the block Lanczos method. Bounds (1.5) and (1.6) are cluster robust, just as well as (1.1).

In comparison to [23, Theorems 5 and 6] by Saad, bounds (1.5) and (1.6) are much more accurate in the case where λ_i belongs to an eigenvalue cluster; cf. numerical examples in [25, Subsection 5.1] concerning a slightly modified form of (1.5). Indeed, the corresponding analysis in [23] can also be reformulated with respect to the abstract block iteration (1.2). The specified function f contains linear factors of the form $(\alpha - \lambda_j)$ or $(\alpha - \psi_j)$, $j = 1, \dots, i-1$ depending on eigenvalues or Ritz values in order to construct auxiliary vectors which are orthogonal to the associated eigenvectors or Ritz vectors. This leads to the terms $\prod_{j=1}^{i-1} (\lambda_j - \lambda_n)/(\lambda_j - \lambda_i)$ and $\prod_{j=1}^{i-1} (\psi_j - \lambda_n)/(\psi_j - \lambda_i)$ called “bulky” factors in [15], which could be very large for clustered eigenvalues and cannot be influenced by the block size.

1.2. Tupewise analysis using majorization

Majorization theory is a powerful tool that allows extend traditional bounds like (1.3), (1.5) and (1.6) of individual convergence measures to bounds of the corresponding tuples such as

$$[\tan \angle(x_c, \mathcal{Y}'), \dots, \tan \angle(x_d, \mathcal{Y}')] \quad \text{and} \quad [\lambda_c - \eta'_c, \dots, \lambda_d - \eta'_d]$$

for $1 \leq c \leq d \leq p$. Majorization theory has been first applied to analyzing some subspace iterations and the block Lanczos method in [12], where convergence bounds involve tuples of convergence measures and even the convergence factors can also be tuples. Similar bounds of tuples with only scalar convergence factors have been later independently derived in [15].

Certain bounds from [12] are extendable to the convergence analysis of (1.2). For instance, applying [12, Theorem 2.3] yields a “stationary” bound in comparison to (1.6). Let us consider the p Ritz value errors $(\lambda_i - \eta'_i)/(\eta'_i - \lambda_n)$, $i = 1, \dots, p$, and write them as a sorted p -tuple ε

whose components fulfill $\varepsilon_1 \geq \dots \geq \varepsilon_p$. For bounding ε , we use the arranged principal angles $\theta'_1 \geq \dots \geq \theta'_p$ between the subspaces \mathcal{X} and \mathcal{Y}' . Skipping the majorization notation, we get

$$\sum_{j=1}^d \varepsilon_j \leq \sum_{j=1}^d \tan^2 \theta'_j, \quad d = 1, \dots, p.$$

An extension in terms of the arranged principal angles $\theta_1 \geq \dots \geq \theta_p$ between \mathcal{X} and \mathcal{Y} reads

$$(1.7) \quad \sum_{j=1}^d \varepsilon_j \leq \sum_{j=1}^d \beta_j \tan^2 \theta_j, \quad d = 1, \dots, p \quad \text{with} \quad \beta_j = \left[T_{k-1} \left(1 + 2 \frac{\lambda_{p+1-j} - \lambda_{p+1}}{\lambda_{p+1} - \lambda_n} \right) \right]^{-2}.$$

Because $\theta_1 = \angle(\mathcal{X}, \mathcal{Y})$, the special form $\varepsilon_1 \leq \beta_1 \tan^2 \theta_1$ of (1.7) for $d = 1$ can imply (1.6) for $i = p$. Moreover, summing up (1.6) for all indices leads to the bound

$$\sum_{j=1}^p \varepsilon_j \leq \left(\sum_{j=1}^p \beta_j \right) \tan^2 \theta_1.$$

The alternative bound

$$\sum_{j=1}^p \varepsilon_j \leq \beta_1 \sum_{j=1}^p \tan^2 \theta_j$$

is analogous to [15, (8.11)] where the convergence factor is also a scalar. These two bounds can be improved by (1.7) for $d = p$, especially if the maximal component $\tan^2 \theta_1$ or β_1 is dominant.

As a motivation of the present paper, it is remarkable that the terms θ_j and β_j in (1.7) are defined within p -tuples concerning principal angles between two p -dimensional subspaces and the first p eigenvalues. If (1.7) is applied to some $d \ll p$, a significant overestimation can occur since the utilized θ_j and β_j are the d largest components of the respective p -tuples. Therefore, we expect to improve (1.7) by using certain terms which simply depend on d -dimensional subspaces and the first d eigenvalues.

1.3. Aim and outline

Deriving majorization bounds for the block Lanczos method via the abstract block iteration (1.2) is a project originated from [1]. Some previous results such as (1.7) are included in Section 4 of the technical report [11] but due to space limits removed in the final version [12]. The published results in [12] concern relations between Ritz value errors and principal angles, but do not depend on (1.2). These results can also be extended to Hermitian operators on infinite dimensional spaces; cf. [12, Subsection 2.8] and a recent application in a filtered subspace iteration [4]. Moreover, some improvements or alternatives of the sine-based bound [12, (2.5)] have been achieved in [17, 18] and devoted to a partial confirmation of the conjecture [12, (2.4)]. The remaining results in [11] actually enable a majorization-type generalization of bound (1.6). Furthermore, the intermediate bound [11, (4.1)] is comparable with [3, Theorem 2.1] concerning block Krylov subspace methods for solving singular value problems. Another potential application is the convergence analysis of a subspace iteration for polynomial eigenvalue problems [5]. The present paper extends [11] and makes further progress concerning certain biorthogonal vectors which have been utilized implicitly for investigating the block power method by Rutishauser; cf. Lemma 3.1.

In the remaining part of this paper, we extend the existing convergence theory of the abstract block iteration (1.2) by majorization arguments. In Section 2, we introduce basic settings and necessary definitions. Section 3 provides main results after discussing some previous results that provide some motivation for our investigation. Applications to concrete methods are formulated in Section 4. We mainly consider the block Lanczos method and compare our new bounds with those from [15]. Further applications are related to shift-and-invert eigensolvers for certain discretized partial differential operators. In Section 5, several numerical examples illustrate the accuracy of the new bounds, followed by concluding remarks. Section 6 (Appendix) presents some detailed proofs.

2. Preliminaries

This section serves as preparation for our majorization-type convergence analysis of block iterative eigensolvers. Subsection 2.1 contains some basic settings which are consistently required in the analysis. In Subsection 2.2, we introduce principal angles and their tangent definition as a useful tool for deriving majorization-type bounds. Subsection 2.3 provides tuple notations and majorization statements for some singular value tuples. An overview of frequently used symbols is as follows.

- Subsection 2.1: matrices A , I_d , eigenpairs (λ_i, x_i) of A , subspaces \mathcal{X} , \mathcal{Y} , \mathcal{Y}' .
- Subsection 2.2: principal angles, notation $\angle(\cdot, \cdot)$ for some special forms.
- Subsection 2.3: tuple a^\downarrow with arranged components, majorization relations \prec_w , \prec , singular value tuple $S(\cdot)$, eigenvalue tuple $\Lambda(\cdot)$, principal angle tuple $\Theta(\cdot, \cdot)$.

2.1. Basic settings

We consider a normal matrix $A \in \mathbb{C}^{n \times n}$ with the eigenvalues $\lambda_1, \dots, \lambda_n$ and the associated orthonormal eigenvectors x_1, \dots, x_n . Identity matrices are denoted by I_d with the corresponding dimension $d \in \{1, \dots, n\}$. Some column vectors or “tall” block matrices are represented with their components or submatrices separated by “;” within a row as in MATLAB/Octave notation. The abstract block iteration $\mathcal{Y}' = f(A)\mathcal{Y}$ introduced in (1.2) serves to compute a moderate number of the first eigenvalues of A .

It is possible to frame our analysis with a Hermitian operator \mathcal{A} in a finite dimensional inner product space \mathcal{H} together with orthogonal projectors of subspaces avoiding introducing any bases and thus shortening the notation; cf. Remark 2.1. An extension to the infinite dimensional case can be made as in [12, Subsection 2.8]. We choose the matrix-based formulation since the new results in this paper are devoted to the convergence theory of matrix eigensolvers where the standard notation usually begins with matrices (even if the implementation is actually matrix-free). In addition, the resulting new bounds can easily be compared with some existing ones for the block Lanczos method, e.g., those from [15, 23].

In practical applications, the dimension p of the initial subspace $\mathcal{Y} \subset \mathbb{C}^n$ is larger than the number of the target eigenvalues, and much smaller than n . We define the invariant subspace $\mathcal{X} = \text{span}\{x_1, \dots, x_p\}$, and assume that the eigenvalue sets $\{\lambda_1, \dots, \lambda_p\}$ and $\{\lambda_{p+1}, \dots, \lambda_n\}$ are disjoint so that \mathcal{X} is unique. Moreover, we assume $\angle(\mathcal{X}, \mathcal{Y}) < \pi/2$ in order to exclude trivial terms such as $\tan \angle(\mathcal{X}, \mathcal{Y}) = \infty$. We reuse the following assumption on the function f made in [7, 8],

$$(2.1) \quad \max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)| < \min_{j \in \{1, \dots, p\}} |f(\lambda_j)|,$$

i.e., f serves as a filter enlarging the components with $i \leq p$ relative to those with $i > p$ and thus ensures that the corresponding convergence factors are smaller than 1.

Our Ritz value analysis deals with a Hermitian A and the approximation of its largest eigenvalues. Therein we arrange the eigenvalues as $\lambda_1 \geq \dots \geq \lambda_n$ and simply assume $\lambda_p > \lambda_{p+1}$ to ensure the uniqueness of \mathcal{X} . The analysis can easily be transformed by the substitution $A \rightarrow -A$ to the consideration of the smallest eigenvalues.

2.2. Principal angles

We use principal angles to measure the distance from-to or between subspaces of the invariant subspace \mathcal{X} and the initial subspace \mathcal{Y} . In general, for two subspaces $\mathcal{U}, \mathcal{V} \subset \mathbb{C}^n$ with $\dim \mathcal{U} \leq \dim \mathcal{V}$ and their arbitrary orthonormal basis matrices U and V , the cosine values of the principal

angles from \mathcal{U} to \mathcal{V} are defined by the singular values of $V^H U$. Therein the largest principal angle is the Euclidean angle $\angle(\mathcal{U}, \mathcal{V})$ which can also be defined by $\max_{u \in \mathcal{U} \setminus \{0\}} \min_{v \in \mathcal{V} \setminus \{0\}} \angle(u, v)$ with angles between nonzero vectors.

In the case $\dim \mathcal{U} = \dim \mathcal{V}$, one also says the principal angles “between \mathcal{U} and \mathcal{V} ” and uses $\angle(\mathcal{U}, \mathcal{V}) = \angle(\mathcal{V}, \mathcal{U})$ for denoting the largest principal angle. In the case $\dim \mathcal{U} = 1 \leq \dim \mathcal{V}$, there is only one principal angle. This coincides with the angle $\angle(u, \mathcal{V}) = \min_{v \in \mathcal{V} \setminus \{0\}} \angle(u, v)$ for arbitrary $u \in \mathcal{U} \setminus \{0\}$. If also $\dim \mathcal{V} = 1$, the associated cosine and tangent values are actually $|\cos \angle(u, v)|$ and $|\tan \angle(u, v)|$ for arbitrary $v \in \mathcal{V} \setminus \{0\}$.

For constructing majorization bounds in Section 3, we utilize tangent values of principal angles where an angle equal to $\pi/2$ would cause a trivial infinity bound. Therefore we restrict the consideration to the case where $V^H U$ has full rank. Then all singular values of $V^H U$ are nonzero so that all principal angles from \mathcal{U} to \mathcal{V} are smaller than $\pi/2$.

Remark 2.1. *A general consideration of tangent values of principal angles within an operator-based formulation has been presented in [27, Theorem 3.1]. Alternatively, one can apply the elegant form $P_{\mathcal{V}\perp}(P_{\mathcal{V}}P_{\mathcal{U}})^\dagger$ with the corresponding orthogonal projectors; cf. [27, Theorem 4.1]. Nevertheless, the following description with a matrix product $V_\perp^H \tilde{U}(V^H \tilde{U})^\dagger$ is more appropriate for estimating error reductions with respect to tangent values in a concise way without additional zero components.*

Lemma 2.1. *Consider the subspace $\mathcal{V} \subset \mathbb{C}^n$ and its orthogonal complement \mathcal{V}_\perp with their arbitrary orthonormal basis matrices $V \in \mathbb{C}^{n \times t}$ and $V_\perp \in \mathbb{C}^{n \times (n-t)}$. Let $\tilde{U} \in \mathbb{C}^{n \times s}$ with $s \leq \min\{t, n-t\}$ be an arbitrary (but not necessarily orthonormal) basis matrix of the subspace $\mathcal{U} \subset \mathbb{C}^n$ for which $V^H \tilde{U}$ has full rank. Then the s largest singular values of the $(n-t) \times t$ matrix $V_\perp^H \tilde{U}(V^H \tilde{U})^\dagger$ coincide with the tangent values of the principal angles from \mathcal{U} to \mathcal{V} where the symbol \dagger denotes the Moore-Penrose pseudoinverse. In particular, $\|V_\perp^H \tilde{U}(V^H \tilde{U})^\dagger\| = \tan \angle(\mathcal{U}, \mathcal{V})$ holds with the 2-norm $\|\cdot\|$.*

Proof. \rightarrow Subsection 6.1.

The condition $s \leq \min\{t, n-t\}$ in Lemma 2.1 is consistent with the analysis in further sections. Therein the subspace iterates or their subsets are considered as \mathcal{U} , whereas an invariant subspace of the same or higher dimension corresponds to \mathcal{V} , i.e., $s \leq t$. This leads to reasonable characteristics of bounds such as the gap ratio $(\lambda_s - \lambda_{t+1})/(\lambda_{t+1} - \lambda_n) > 0$ in a Chebyshev term. Moreover, $s \leq n-t$ is naturally fulfilled in the context of computing several eigenvalues of a large matrix or a matrix pair.

2.3. Majorization

We use majorization arguments for deriving bounds in terms of certain tuples of real numbers concerning tangent values of principal angles and Ritz value errors. For notational convenience we treat these tuples as row vectors, e.g., $a = [a_1, \dots, a_d]$ with the corresponding dimension d . In addition, we consider a rearrangement of the components of a in (not strictly) decreasing order and denote the resulting vector by a^\downarrow , i.e., $a_1^\downarrow \geq \dots \geq a_d^\downarrow$. Let b be another tuple with dimension d . If the components of a^\downarrow and b^\downarrow fulfill

$$\sum_{i=1}^k a_i^\downarrow \leq \sum_{i=1}^k b_i^\downarrow \quad \forall k \in \{1, \dots, d\},$$

one says that b weakly majorizes (submajorizes) a . The weak (additive) majorization is denoted by $a \prec_w b$. The corresponding strong majorization with the notation $a \prec b$ additionally means that the sum inequality for $k = d$ is actually an equality. For nonnegative tuples a and b with

dimension d , the weak multiplicative majorization is defined by

$$\prod_{i=1}^k a_i^\downarrow \leq \prod_{i=1}^k b_i^\downarrow \quad \forall k \in \{1, \dots, d\}$$

and denoted by $\log a \prec_w \log b$. The strong version $\log a \prec \log b$ requires the additional condition $\prod_{i=1}^d a_i^\downarrow = \prod_{i=1}^d b_i^\downarrow$. Moreover, nonnegative tuples with different dimensions can be compared by adding zeros to the shorter tuple.

Specific tuples in our analysis mostly consist of singular values or eigenvalues of matrices. We generally consider a matrix $B \in \mathbb{C}^{d \times s}$, and denote by $S(B)$ the tuple of arranged singular values of B in decreasing order. Multiple singular values are counted repeatedly so that $S(B)$ has dimension $\min\{d, s\}$. Similarly, we denote by $\Lambda(C)$ the tuple of arranged eigenvalues of a Hermitian matrix $C \in \mathbb{C}^{t \times t}$ in decreasing order. In particular, it holds that $S(B) = \sqrt{\Lambda(B^H B)}$ for $d \geq s$ and $S(B) = \sqrt{\Lambda(BB^H)}$ for $d \leq s$ with componentwise square roots.

The arranged principal angles $\theta_1 \geq \dots \geq \theta_s$ from \mathcal{U} to \mathcal{V} give the angle tuple $\Theta(\mathcal{U}, \mathcal{V}) = [\theta_1, \dots, \theta_s]$. The corresponding tangent tuple $\tan \Theta(\mathcal{U}, \mathcal{V}) = [\tan(\theta_1), \dots, \tan(\theta_s)]$ is the leading s -subtuple of $S(V_\perp^H \tilde{U} (V^H \tilde{U})^\dagger)$ according to Lemma 2.1. In addition, by using an arbitrary orthonormal basis matrix U of \mathcal{U} , the cosine-type definition of principal angles shows $(\cos \Theta(\mathcal{U}, \mathcal{V}))^\downarrow = S(V^H U)$.

We complete this subsection by a basic argument for the majorization-type analysis.

Lemma 2.2. *Consider the matrices $B_1 \in \mathbb{C}^{d_1 \times d_2}$, $B_2 \in \mathbb{C}^{d_2 \times d_3}$ and $B_3 \in \mathbb{C}^{d_3 \times d_4}$. Let $S_t(B)$ be the leading t -subtuple of the singular value tuple $S(B)$ for $B \in \{B_1, B_2, B_3, B_1 B_2 B_3\}$ and $t \leq \min\{d_1, d_2, d_3, d_4\}$. Then it holds with componentwise multiplication, division and power for $c \in \mathbb{N}$ of tuples, that $S_t^c(B_1 B_2 B_3) \prec_w S_t^c(B_1) S_t^c(B_2) S_t^c(B_3)$ and*

$$S_t^c(B_1 B_2 B_3) / S_t^c(B_2) \prec_w S_t^c(B_1) S_t^c(B_3) \quad \text{for } S_t(B_2) > 0.$$

Proof. Applying [12, Theorem 4.4] shows the expression

$$\log S(B_1 B_2 B_3) - \log S(B_2) \prec \log (S(B_1) S(B_3))$$

which means that the singular value inequality

$$\prod_{j=1}^k \sigma_{i_j}(B_1 B_2 B_3) \leq \prod_{j=1}^k (\sigma_j(B_1) \sigma_j(B_3)) \sigma_{i_j}(B_2)$$

holds for each $k \in \{1, \dots, d\}$ with $d = \max\{d_1, d_2, d_3, d_4\}$ and for each index set $\{i_1, \dots, i_k\} \subseteq \{1, \dots, d\}$ with $i_1 < \dots < i_k$. Therein zeros are occasionally added to shorter singular value tuples to match the sizes, and the equality is attained for $k = d$. By considering this singular value inequality to the c th power for $k \in \{1, \dots, t\}$ and $\{i_1, \dots, i_k\} = \{1, \dots, k\}$, we get the weak multiplicative majorizations $\log S_t^c(B_1 B_2 B_3) \prec_w \log (S_t^c(B_1) S_t^c(B_2) S_t^c(B_3))$ and

$$\log (S_t^c(B_1 B_2 B_3) / S_t^c(B_2)) \prec_w \log (S_t^c(B_1) S_t^c(B_3)) \quad \text{for } S_t(B_2) > 0.$$

These imply the weak (additive) majorizations in the claim of Lemma 2.2 by applying the exponential function “exp” which is convex and increasing; see [16, Proposition 4.B.2]. \square

In the above proof, the composition “exp \circ log” is only applied to positive components whereas zero components remain unchanged as they lead to a trivial case. Alternatively, one can derive the underlying additive inequality directly by the corresponding multiplicative inequality; cf. [6, Corollary 3.3.10] and [2, Example II.3.5].

3. Main results

We motivate our new analysis of the abstract block iteration (1.2) by introducing some auxiliary terms and arguments. The basic ingredients are certain biorthogonal vectors from \mathcal{X} and \mathcal{Y} inspired by the classical analysis of the block power method by Rutishauser [22, Theorem 2] (do not confuse these biorthogonal vectors with those mentioned before [22, Section 1] which form two sets of iterates). We note that the corresponding biorthogonality is represented by a submatrix in [22, (10)] after some substitutions and results in an asymptotic bound. Indeed, this biorthogonality also enables a simpler proof of the nonasymptotic bound [21, (14.11)] by Parlett. A similar argument in [23, Lemma 4] by Saad is the starting point of an analysis of the block Lanczos method.

Lemma 3.1. *With the settings from Subsection 2.1, consider the orthonormal basis matrix $X = [x_1, \dots, x_p]$ of the invariant subspace \mathcal{X} together with an arbitrary orthonormal basis matrix $\tilde{Y} \in \mathbb{C}^{n \times p}$ of \mathcal{Y} . Then the $p \times p$ matrix $X^H \tilde{Y}$ is invertible, and the vectors*

$$y_i = \tilde{Y} c_i, \quad i = 1, \dots, p \quad \text{with the columns } c_1, \dots, c_p \text{ of } (X^H \tilde{Y})^{-1}$$

are evidently linearly independent and form a basis matrix $Y = [y_1, \dots, y_p]$. Moreover, x_1, \dots, x_p and y_1, \dots, y_p are biorthogonal, i.e.,

$$(3.1) \quad x_i^H y_j = \delta_{ij} \quad \text{for } i, j \in \{1, \dots, p\}.$$

Proof. The smallest singular value of $X^H \tilde{Y}$ coincides with $\cos \angle(\mathcal{X}, \mathcal{Y})$ which is nonzero because $\angle(\mathcal{X}, \mathcal{Y}) < \pi/2$. Thus $X^H \tilde{Y}$ is invertible. The biorthogonality is verified by

$$x_i^H y_j = (X e_i)^H \tilde{Y} c_j = e_i^H X^H \tilde{Y} ((X^H \tilde{Y})^{-1} e_j) = e_i^H e_j = \delta_{ij}$$

with the columns e_1, \dots, e_p of I_p . □

We introduce some single-angle bounds in Subsection 3.1 and present the corresponding multi-angle majorization-type bounds in Subsection 3.2. Some relatively long proofs are given in Section 6.

3.1. Single-angle bounds

In [21, 22], the auxiliary vectors defined in Lemma 3.1 are utilized separately. The simple power method applied to them can be observed within proper invariant subspaces according to the biorthogonality (3.1). An essentially analogous approach with the tangent description introduced in Lemma 2.1 leads to the following single-angle bound which corresponds to an abstract form of (1.5).

Lemma 3.2. (cf. [7, Lemma 2.3.1]) *With the settings from Subsection 2.1 and Lemma 3.1, it holds that*

$$(3.2) \quad \tan \angle(x_i, \mathcal{Y}') \leq \sigma_i \tan \angle(x_i, y_i) \leq \sigma_i \tan \angle(\mathcal{X}, \mathcal{Y})$$

with $\sigma_i = \frac{\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)|}{|f(\lambda_i)|}, \quad i = 1, \dots, p.$

Proof. → Subsection 6.2.

For the angle-dependent bound (1.6) on the Ritz values $\eta'_1 \geq \dots \geq \eta'_p$ of A in \mathcal{Y}' , an abstract form reads

$$(3.3) \quad \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \leq \frac{\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)|^2}{\min_{j \in \{1, \dots, i\}} |f(\lambda_j)|^2} \tan^2 \angle(\mathcal{X}, \mathcal{Y}), \quad i = 1, \dots, p.$$

Therein $\dim \mathcal{Y}' = p$ is ensured by the common assumption (2.1); see Lemma 3.3 below. For deriving (3.3), we can put several auxiliary vectors from Lemma 3.1 together in order to construct proper auxiliary subspaces. This improves the approach by Saad [23, Theorem 6] which constructs auxiliary vectors with Ritz values and causes less convenient terms in the bound.

The first step of the derivation of (3.3) produces an angle bound for auxiliary subspaces.

Lemma 3.3. (cf. [7, Lemma 2.3.1]) *With the settings from Subsection 2.1 and Lemma 3.1, consider for an index $i \in \{1, \dots, p\}$ the subspace \mathcal{X}_i with the orthonormal basis matrix $X_i = [x_1, \dots, x_i]$ and the subspace \mathcal{Y}_i with the basis matrix $Y_i = [y_1, \dots, y_i]$. Then the matrix $Y_i' = f(A)Y_i$ has rank i so that the subspace $\mathcal{Y}_i' = \text{span}\{Y_i'\}$ has dimension i . In particular, the subspace iterate \mathcal{Y}' of (1.2) coincides with \mathcal{Y}'_p and thus has dimension p . Moreover, it holds that*

$$(3.4) \quad \tan \angle(\mathcal{X}_i, \mathcal{Y}'_i) \leq \frac{\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)|}{\min_{j \in \{1, \dots, i\}} |f(\lambda_j)|} \tan \angle(\mathcal{X}_i, \mathcal{Y}_i).$$

Proof. → Subsection 6.3.

The next step of the derivation of (3.3) is to verify a direct relation between the convergence measures $\tan \angle(\mathcal{X}_i, \mathcal{Y}'_i)$ and $(\lambda_i - \eta'_i)/(\eta'_i - \lambda_n)$.

Lemma 3.4. (cf. [7, Lemmas 2.2.5 and 2.2.6]) *With the settings from Subsection 2.1, Lemma 3.1 and Lemma 3.3, consider the Ritz values $\eta'_1 \geq \dots \geq \eta'_p$ of A in \mathcal{Y}' . Then*

$$(3.5) \quad \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \leq \tan^2 \angle(\mathcal{X}_i, \mathcal{Y}'_i)$$

holds for each $i \in \{1, \dots, p\}$.

Proof. → Subsection 6.4.

Based on (3.4) and (3.5), we get the bound

$$(3.6) \quad \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \leq \frac{\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)|^2}{\min_{j \in \{1, \dots, i\}} |f(\lambda_j)|^2} \tan^2 \angle(\mathcal{X}_i, \mathcal{Y}_i), \quad i = 1, \dots, p.$$

If the auxiliary subspaces \mathcal{X}_i and \mathcal{Y}_i need to be eliminated, one can extend (3.6) as (3.3); see Subsection 6.5.

3.2. Multi-angle majorization-type bounds

Our new bounds for the abstract block iteration (1.2) are generalizations of those introduced in Subsection 3.1. We use partial sums of certain tuples containing (squared) tangent values of principal angles instead of the corresponding maxima.

We first generalize the single-angle bound (3.2) to principal angles concerning a possible eigenvalue cluster in the eigenvalue set $\{\lambda_1, \dots, \lambda_p\}$.

Theorem 3.5. *With the settings from Subsection 2.1 and Lemma 3.1, consider an index set $\tau = \{i_1, \dots, i_t\} \subseteq \{1, \dots, p\}$ with $i_1 < \dots < i_t$ as well as the subspaces*

$$\mathcal{X}_\tau = \text{span}\{x_{i_1}, \dots, x_{i_t}\} \quad \text{and} \quad \mathcal{Y}_\tau = \text{span}\{y_{i_1}, \dots, y_{i_t}\}.$$

By using the notations from Subsection 2.3 for tuples and majorization, it holds that

$$(3.7) \quad \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}') \prec_w \Phi_\tau \widehat{\Phi}_t \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau)$$

with $\Phi_\tau = [|f(\lambda_{i_1})|^{-1}, \dots, |f(\lambda_{i_t})|^{-1}]^\downarrow$, $\widehat{\Phi} = [|f(\lambda_{p+1})|, \dots, |f(\lambda_n)|]^\downarrow$ and $\widehat{\Phi}_t = \widehat{\Phi}(1:t)$.

Proof. → Subsection 6.6.

Theorem 3.5 is especially suitable for the case where the eigenvalues $\lambda_{i_1}, \dots, \lambda_{i_t}$ are consecutive and clustered. The components of the tuplewise convergence factor $\Phi_\tau \widehat{\Phi}_t$ can be bounded away from 1 by selecting a proper f such that $\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)| \ll \min_{j \in \{1, \dots, p\}} |f(\lambda_j)|$. An application to nonconsecutive eigenvalues is also feasible, but not of practical interest.

Remark 3.1. *A direct application of Lemmas 2.1 and 2.2 implies the more abstract bound*

$$(3.8) \quad \tan \Theta(F\mathcal{U}, \mathcal{V}) \prec_w S_s((V^H F V)^{-1}) S_s(V_\perp^H F V_\perp) \tan \Theta(\mathcal{U}, \mathcal{V})$$

for a normal matrix $F \in \mathbb{C}^{n \times n}$ provided that \mathcal{V} and \mathcal{V}_\perp are invariant subspaces of F^H and that $V^H F V$ is invertible; see Subsection 6.7 for the derivation.

We note that bound (3.8) can directly lead to bound (3.7) only in the special case $\tau = \{1, \dots, p\}$ because the biorthogonality (3.1) concerning the index set τ cannot be added to (3.8) afterwards. Alternatively, we can formulate a restricted form of (3.8) with respect to an invariant subspace of A associated with the eigenvalues $\lambda_{i_1}, \dots, \lambda_{i_t}, \lambda_{p+1}, \dots, \lambda_n$. Nevertheless, we prefer the easily understandable formulation in Theorem 3.5.

In order to eliminate \mathcal{Y}_τ , bound (3.7) can be modified as

$$(3.9) \quad \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}') \prec_w \Phi_\tau \widehat{\Phi}_t \tan \Theta_t(\mathcal{X}, \mathcal{Y})$$

with the leading t -subtuple $\Theta_t(\mathcal{X}, \mathcal{Y})$ of $\Theta(\mathcal{X}, \mathcal{Y})$. Therein Lemma 2.1 shows

$$\tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau) = S(\widetilde{G}) \quad \text{for} \quad \widetilde{G} = [x_{p+1}^H Y_\tau; \dots; x_n^H Y_\tau] \in \mathbb{C}^{(n-p) \times t}$$

$$\text{and} \quad \tan \Theta(\mathcal{X}, \mathcal{Y}) = S(\widehat{G}) \quad \text{for} \quad \widehat{G} = [x_{p+1}^H Y; \dots; x_n^H Y] \in \mathbb{C}^{(n-p) \times p}$$

as in Subsection 6.5. Moreover, the Courant-Fischer principles ensure the tuple inequality

$$S(\widetilde{G}) = \sqrt{\Lambda(\widetilde{G}^H \widetilde{G})} = \sqrt{\Lambda(E_\tau^H \widehat{G}^H \widehat{G} E_\tau)} \leq \sqrt{\Lambda_t(\widehat{G}^H \widehat{G})} = S_t(\widehat{G}),$$

where E_τ is a $p \times t$ orthonormal matrix consisting of the columns of I_p with indices in τ , and Λ_t or S_t denotes the leading t -subtuple of the corresponding eigenvalue tuple or singular value tuple. This leads to $\Phi_\tau \widehat{\Phi}_t \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau) \leq \Phi_\tau \widehat{\Phi}_t \tan \Theta_t(\mathcal{X}, \mathcal{Y})$ and verifies (3.9).

For deriving a majorization version of the angle-dependent bound (3.3), we proceed in two steps as in Subsection 3.1. The first step gives an intermediate multi-angle bound for auxiliary subspaces, and the second step enables a combination between two convergence measures.

We formulate the first step as a generalization of Lemma 3.3 in square form.

Lemma 3.6. *(majorization update of [7, Lemma 2.3.1]) With the settings from Subsection 2.1, Lemma 3.1 and Lemma 3.3, it holds that*

$$(3.10) \quad \tan^2 \Theta(\mathcal{X}_i, \mathcal{Y}'_i) \prec_w \Phi_i^2 \widehat{\Phi}_i^2 \tan^2 \Theta(\mathcal{X}_i, \mathcal{Y}_i)$$

with $\Phi_i = [|f(\lambda_1)|^{-1}, \dots, |f(\lambda_i)|^{-1}]^\downarrow$, $\widehat{\Phi} = [|f(\lambda_{p+1})|, \dots, |f(\lambda_n)|]^\downarrow$ and $\widehat{\Phi}_i = \widehat{\Phi}(1:i)$.

Proof. Bound (3.10) is derived by adapting the proof of Theorem 3.5 (see Subsection 6.5) to $\tau = \{1, \dots, i\}$ and $t = i$. Most arguments before the intermediate bound (6.4) are used in the same way. Only the application of Lemma 2.2 is slightly different, namely, by using $c = 2$ instead of $c = 1$. \square

The second step corresponds to a generalization of Lemma 3.4.

Lemma 3.7. (majorization update of [7, Lemmas 2.2.5 and 2.2.6]) With the settings from Subsection 2.1, Lemma 3.1, Lemma 3.3 and Lemma 3.4,

$$(3.11) \quad \left[\frac{\lambda_1 - \eta'_1}{\eta'_1 - \lambda_n}, \dots, \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \right] \prec_w \tan^2 \Theta(\mathcal{X}_i, \mathcal{Y}'_i)$$

holds for each $i \in \{1, \dots, p\}$.

Proof. → Subsection 6.8.

Now the majorization version of (3.3) can be shown.

Theorem 3.8. With the settings from Subsection 2.1, Lemma 3.1, Lemma 3.3 and Lemma 3.4, it holds that

$$\left[\frac{\lambda_1 - \eta'_1}{\eta'_1 - \lambda_n}, \dots, \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \right] \prec_w \Phi_i^2 \widehat{\Phi}_i^2 \tan^2 \Theta(\mathcal{X}_i, \mathcal{Y}_i) \leq \Phi_i^2 \widehat{\Phi}_i^2 \tan^2 \Theta_i(\mathcal{X}, \mathcal{Y})$$

with $\Phi_i = [|f(\lambda_1)|^{-1}, \dots, |f(\lambda_i)|^{-1}]^\downarrow$, $\widehat{\Phi} = [|f(\lambda_{p+1})|, \dots, |f(\lambda_n)|]^\downarrow$, $\widehat{\Phi}_i = \widehat{\Phi}(1:i)$ and the leading i -subtuple $\Theta_i(\mathcal{X}, \mathcal{Y})$ of $\Theta(\mathcal{X}, \mathcal{Y})$.

Proof. The weak majorization is verified by combining (3.10) with (3.11). The tuple inequality is shown analogously to the derivation of (3.9). \square

Theorem 3.8 additionally provides a majorization bound with $\Phi_i^2 \widehat{\Phi}_i^2 \tan^2 \Theta_i(\mathcal{X}, \mathcal{Y})$ which does not depend on auxiliary subspaces.

Remark 3.2. A more abstract bound concerning Lemma 2.1 in the case $\dim \mathcal{U} = s = t = \dim \mathcal{V}$ can be derived by combining (3.8) with an analogue of (3.11). Therein \mathcal{V} is assumed to be an invariant subspace of A associated with the t largest eigenvalues $\lambda_1 \geq \dots \geq \lambda_t$. Then it holds that

$$(3.12) \quad \left[\frac{\lambda_1 - \psi_1}{\psi_1 - \psi}, \dots, \frac{\lambda_t - \psi_t}{\psi_t - \psi} \right] \prec_w S^2((V^H FV)^{-1}) S_t^2(V_\perp^H FV_\perp) \tan^2 \Theta(\mathcal{U}, \mathcal{V})$$

for the Ritz values $\psi_1 \geq \dots \geq \psi_t$ of A in $F\mathcal{U}$ and the smallest Ritz value ψ of A in $F\mathcal{U} + \mathcal{V}$. Proving Theorem 3.8 based on (3.12) is only feasible for $i = p$ since the terms $|f(\lambda_{i+1})|, \dots, |f(\lambda_p)|$ cannot easily be dropped afterwards.

Furthermore, we can use [12, Theorem 2.1] for formulating an alternative bound under the weaker assumption that \mathcal{V} is an invariant subspace of A but not necessarily associated with the t largest eigenvalues. More precisely, we reformulate [12, (2.2)] as

$$\zeta^{-1} |\Lambda(V^H AV) - \Lambda(W^H AW)| \prec_w \sin^2 \Theta(F\mathcal{U}, \mathcal{V}).$$

with an orthonormal basis matrix W of $F\mathcal{U}$ and the spread ζ of the Ritz value set of A in $F\mathcal{U} + \mathcal{V}$. Subsequently, we apply the convex and increasing function $\alpha/(1-\alpha)$ defined for $\alpha \in [0, 1)$ to the tuples in the above majorization-type bound. Then

$$\frac{|\Lambda(V^H AV) - \Lambda(W^H AW)|}{[\zeta, \dots, \zeta] - |\Lambda(V^H AV) - \Lambda(W^H AW)|} \prec_w \tan^2 \Theta(F\mathcal{U}, \mathcal{V}).$$

Combining this with (3.8) results in

$$(3.13) \quad \frac{|\Lambda(V^H AV) - \Lambda(W^H AW)|}{[\zeta, \dots, \zeta] - |\Lambda(V^H AV) - \Lambda(W^H AW)|} \prec_w S^2((V^H FV)^{-1}) S_t^2(V_\perp^H FV_\perp) \tan^2 \Theta(\mathcal{U}, \mathcal{V}).$$

The weaker assumption on \mathcal{V} enables the application of (3.13) to interior eigenvalues. However, it is somewhat challenging to select an optimal filter F without knowing good approximations to undesired eigenvalues for polynomial filters and/or to desired eigenvalues for rational filters.

The well-known Chebyshev terms for this type of bounds actually require that \mathcal{V} corresponds to the contiguous set of the t largest eigenvalues.

4. Applications to specific filters

We apply our new bounds presented in Subsection 3.2 to the convergence analysis of the block Lanczos method and its inverted version. The construction of the specified bounds essentially consists of specifying the filter f in the abstract block iteration (1.2) and adapting intermediate bounds to explicit convergence measures based on subspace inclusions. The practical choices of f are normally limited to polynomials and rational functions. The polynomial version of (1.2) is easier to implement, since its essential part is the multiplication of vectors or basis matrices by A . For implementing the rational version of (1.2), one requires linear system solvers which make a single step more expensive with respect to computational time and storage requirements. Nevertheless, the total computational time can significantly be reduced.

In Subsection 4.1, we specify the filter f as various shifted Chebyshev polynomials so that (1.2) simulates certain underlying iterations with almost optimal convergence rates within the block Lanczos method. This is comparable with the standard approach from [7, 21, 23] and the recent approach from [15]. A remarkable feature of our majorization-type bounds is that the convergence factors are tuples instead of scalars as in the existing bounds. In Subsection 4.2, we introduce an adaption of the main results to some shift-and-invert eigensolvers.

4.1. Block Lanczos method

With the settings from Subsection 2.1, the block Lanczos method constructs block Krylov subspaces of the form $\mathcal{K} = \mathcal{Y} + A\mathcal{Y} + \dots + A^{k-1}\mathcal{Y}$. Thus the abstract block iteration (1.2), with an arbitrary real polynomial of degree $k-1$ as f , produces a subset \mathcal{Y}' of \mathcal{K} . The inclusion $\mathcal{Y}' \subseteq \mathcal{K}$ leads to simple inequalities for principal angles and Ritz values.

In order to specify f as a reasonable polynomial, we consider the multi-angle bound (3.7) as an example. Ideally, we want to construct f in an optimal way, namely, minimizing the tuplewise convergence factor $\Phi_\tau \widehat{\Phi}_t$ defined by

$$\Phi_\tau = [|f(\lambda_{i_1})|^{-1}, \dots, |f(\lambda_{i_t})|^{-1}]^\downarrow, \quad \widehat{\Phi} = [|f(\lambda_{p+1})|, \dots, |f(\lambda_n)|]^\downarrow \quad \text{and} \quad \widehat{\Phi}_t = \widehat{\Phi}(1:t)$$

with respect to the included eigenvalues. Following the standard approach from [7, 21, 23], we simplify this minimization problem with respect to the eigenvalue interval $[\lambda_n, \lambda_{p+1}]$. Therein we begin with the tuple inequality

$$(4.1) \quad \Phi_\tau \widehat{\Phi}_t \leq \Phi_\tau \varphi = [|f(\lambda_{i_1})|^{-1}\varphi, \dots, |f(\lambda_{i_t})|^{-1}\varphi]^\downarrow \quad \text{with} \quad \varphi = \max_{\lambda \in [\lambda_n, \lambda_{p+1}]} |f(\lambda)|$$

and then determine an f minimizing the upper bound $\Phi_\tau \varphi$. Fortunately, a shifted Chebyshev polynomial simultaneously minimizes all the components of $\Phi_\tau \varphi$ as shown in the following lemma inspired by [7, Lemma 2.4.1].

Lemma 4.1. *Every component of the tuple $\Phi_\tau \varphi$ in (4.1) is minimized in the class of real polynomials of degree $k-1$ at*

$$(4.2) \quad f(\alpha) = T_{k-1} \left(\frac{2\alpha - \lambda_{p+1} - \lambda_n}{\lambda_{p+1} - \lambda_n} \right) = T_{k-1} \left(1 + 2 \frac{\alpha - \lambda_{p+1}}{\lambda_{p+1} - \lambda_n} \right)$$

where T_l with $l \in \mathbb{N}$ denotes the Chebyshev polynomials (of the first kind). With this choice of f , we have $\Phi_\tau \varphi = [\sigma_{i_t}, \dots, \sigma_{i_1}]$ where

$$(4.3) \quad \sigma_j = \left[T_{k-1} \left(\frac{1 + \xi_j}{1 - \xi_j} \right) \right]^{-1} = [T_{k-1}(1 + 2\gamma_j)]^{-1}$$

with $\xi_j = \frac{\lambda_j - \lambda_{p+1}}{\lambda_j - \lambda_n}$ and $\gamma_j = \frac{\lambda_j - \lambda_{p+1}}{\lambda_{p+1} - \lambda_n}$.

Proof. The absolute value of the Chebyshev polynomial T_l is bounded above by 1 on the interval $[-1, 1]$ and exceeds 1 outside of this interval where the growth is faster in comparison to any other polynomial of degree l whose absolute value is also bounded above by 1 on $[-1, 1]$. In our context, after mapping the interval $[-1, 1]$ linearly to the interval $[\lambda_n, \lambda_{p+1}]$, the above well-known property indicates that function (4.2) solves the minimization problem described in the lemma. The verification of $\Phi_\tau \varphi = [\sigma_{i_t}, \dots, \sigma_{i_1}]$ is straightforward where the arrangement of the components is based on the monotonicity of T_l on $(1, \infty)$. \square

Several implementations of subspace iterations directly using Chebyshev polynomials are available; see, e.g., [21] for the three-term recurrence and [8] for the two-term recurrence iterative formulas. Therein reasonable quality bounds of λ_{p+1} and λ_n are normally required.

The convergence theory of subspace iterations using Chebyshev polynomials is a standard way to derive the convergence rate bounds for the block Lanczos method. We note that the ξ_j -notation and the γ_j -notation of the Chebyshev term (4.3) are suggested in [8] and [23], respectively. In the γ_j -notation, one can easily see that the bound decreases if the so-called gap ratio γ_j increases. The sharpness of scalar Chebyshev bounds for the single-vector version of the Lanczos method has been discussed in [14, 24]. Some more accurate bounds are constructed by interpolating polynomials in [24] and adapted to the block Lanczos method in [25].

We continue with the reformulation of majorization bounds. Based on Lemma 4.1, we specify the bounds from Subsection 3.2 for the block Lanczos method as follows.

Theorem 4.2. *With the settings from Subsection 2.1 and Lemma 3.1, consider an index set $\tau = \{i_1, \dots, i_t\} \subseteq \{1, \dots, p\}$ with $i_1 < \dots < i_t$ as well as the subspaces*

$$\mathcal{X}_\tau = \text{span}\{x_{i_1}, \dots, x_{i_t}\}, \quad \mathcal{Y}_\tau = \text{span}\{y_{i_1}, \dots, y_{i_t}\} \quad \text{and} \quad \mathcal{K} = \mathcal{Y} + A\mathcal{Y} + \dots + A^{k-1}\mathcal{Y}.$$

By using the notations from Subsection 2.3 for tuples and majorization, and the parameter definition (4.3), it holds that

$$(4.4) \quad \tan \Theta(\mathcal{X}_\tau, \mathcal{K}) \prec_w [\sigma_{i_t}, \dots, \sigma_{i_1}] \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau) \leq [\sigma_{i_t}, \dots, \sigma_{i_1}] \tan \Theta_t(\mathcal{X}, \mathcal{Y})$$

with the leading t -subtuple $\Theta_t(\mathcal{X}, \mathcal{Y})$ of $\Theta(\mathcal{X}, \mathcal{Y})$.

Next, consider the p largest Ritz values $\psi_1 \geq \dots \geq \psi_p$ of A in \mathcal{K} together with the subspaces $\mathcal{X}_i = \text{span}\{x_1, \dots, x_i\}$ and $\mathcal{Y}_i = \text{span}\{y_1, \dots, y_i\}$ for an index $i \in \{1, \dots, p\}$. Then

$$(4.5) \quad \left[\frac{\lambda_1 - \psi_1}{\psi_1 - \lambda_n}, \dots, \frac{\lambda_i - \psi_i}{\psi_i - \lambda_n} \right] \prec_w [\sigma_i^2, \dots, \sigma_1^2] \tan^2 \Theta(\mathcal{X}_i, \mathcal{Y}_i) \leq [\sigma_i^2, \dots, \sigma_1^2] \tan^2 \Theta_i(\mathcal{X}, \mathcal{Y})$$

holds with the leading i -subtuple $\Theta_i(\mathcal{X}, \mathcal{Y})$ of $\Theta(\mathcal{X}, \mathcal{Y})$.

Proof. For any real polynomial f of degree $k-1$ we have $\mathcal{Y}' = f(A)\mathcal{Y} \subseteq \mathcal{K}$ and thus

$$\tan \Theta(\mathcal{X}_\tau, \mathcal{K}) \leq \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}'), \quad \lambda_j \geq \psi_j \geq \eta'_j \quad \forall j \in \{1, \dots, p\},$$

where the tangent tuple inequality can be proved analogously to the end of the proof of Theorem 3.5, and the Ritz value inequalities are ensured by the Courant-Fischer principles. Then the left-hand sides of the bounds from Theorem 3.5, its supplement (3.9) and Theorem 3.8 can be

extended as the left-hand sides of the corresponding specified bounds. The extension of the right-hand sides is justified by Lemma 4.1. \square

Remark 4.1. *The multi-angle bound (4.4) is a majorization-type generalization of the single-angle bound [25, (15)] which corresponds to (1.5) and improves [23, (3.4)]. A similar multi-angle bound reads*

$$(4.6) \quad \sum_{j=1}^l \tan \theta_j(\mathcal{X}_\tau, \mathcal{K}) \leq \sum_{j=1}^l \sigma_{i_t} \tan \theta_j(\mathcal{X}_\tau, \mathcal{Y}_\tau), \quad l = 1, \dots, t$$

by applying [15, Theorem 8.1] where the underlying unitarily invariant norm is specified as the Ky Fan l -norm for $l = 1, \dots, t$, i.e., the sum of the l largest singular values; cf. [2, (IV.33)]. In comparison to (4.6), the new bound (4.4) is more accurate by considering that σ_{i_t} is the maximal component of the tuplewise convergence factor $[\sigma_{i_t}, \dots, \sigma_{i_1}]$.

Remark 4.2. *We can modify (4.4) by restricting its derivation to an invariant subspace which is orthogonal to the eigenvectors associated with a number of the largest eigenvalues and occasionally also the smallest eigenvalues; cf. [23] and [21, Section 12.5]. Then additional terms like $\prod_{j=1}^{i-1} (\lambda_j - \lambda_n) / (\lambda_j - \lambda_i)$ occur in the bound so that this modification is only meaningful for well-separated eigenvalues. If the target eigenvalues are clustered and the initial subspace in the block Lanczos method is relatively small, we can interpret (4.4) in another way, namely, let \mathcal{Y} be a block Krylov subspace $\tilde{\mathcal{K}} = \tilde{\mathcal{Y}} + A\tilde{\mathcal{Y}} + \dots + A^{l-1}\tilde{\mathcal{Y}}$ (therein $\dim \tilde{\mathcal{K}} \leq l \dim \tilde{\mathcal{Y}}$ holds, not necessarily with equality), then $\mathcal{K} = \mathcal{Y} + A\mathcal{Y} + \dots + A^{k-1}\mathcal{Y}$ is related to a block Krylov subspace with the initial subspace $\tilde{\mathcal{Y}}$ and degree $k+l-1$. This improves the applicability of (4.4) concerning small initial subspaces; cf. [15, Section 8].*

In addition, the Ritz vector bounds [15, (4.10) and (8.8)] are generalizations of [23, (2.15)] related to subspaces spanned by Ritz vectors with a unitarily invariant norm. We note that a generalization with a standard operator norm is presented in [9, Theorem 4.3]. This bound type requires Ritz values. A drawback is that the bound can decrease if some estimated Ritz values are utilized instead of the corresponding exact Ritz values so that a combination with Ritz value bounds is not meaningful. For this reason, we recommend bound [8, (2.7)]. Moreover, direct and concise Ritz vector bounds are perhaps only known for the single-vector version of the Lanczos method; see [19, (3.3)]. Improving Ritz vector bounds is a potential topic in our future research.

We now turn our attention to the angle-dependent Ritz value bound (4.5). Similarly to Remark 6.1, a slightly sharper form of (4.5) with the smallest Ritz value of A in $\mathcal{X} + \mathcal{K}$ instead of λ_n can be constructed by a restricted analysis in $\mathcal{X} + \mathcal{K}$. Indeed, also the parameter definition (4.3) can be modified for (4.5) by using certain Ritz values of A in $\mathcal{X} + \mathcal{K}$ instead of λ_{p+1} and λ_n . This is enabled by the following fact: With an arbitrary orthonormal basis matrix V of $\mathcal{X} + \mathcal{K}$, the relation

$$\begin{aligned} V^H A^j \mathcal{Y} &= V^H A(A^{j-1} \mathcal{Y}) = V^H A(VV^H)(A^{j-1} \mathcal{Y}) \\ &= (V^H AV)V^H A^{j-1} \mathcal{Y} = \dots = (V^H AV)^j (V^H \mathcal{Y}) \end{aligned}$$

holds for each $j \in \{1, \dots, k-1\}$ by using the orthogonal projector VV^H . Thus

$$\begin{aligned} V^H \mathcal{K} &= V^H \mathcal{Y} + V^H A\mathcal{Y} + \dots + V^H A^{k-1} \mathcal{Y} \\ &= V^H \mathcal{Y} + (V^H AV)(V^H \mathcal{Y}) + \dots + (V^H AV)^{k-1} (V^H \mathcal{Y}) \end{aligned}$$

is a block Krylov subspace with respect to $V^H AV$, and analogous bounds can be achieved with eigenvalues of $V^H AV$, i.e., Ritz values of A in $\mathcal{X} + \mathcal{K}$. This fact is a direct generalization of that for Krylov subspaces; cf. [7, pp. 36] and [19, Lemma 3.3].

Remark 4.3. Bound (4.5) generalizes [8, (2.20)] which can be reformulated as (1.6) or [25, (19)] and improves [23, (3.10)]. A similar bound based on [15, Theorem 8.2] reads

$$(4.7) \quad \sum_{j=1}^l \frac{\lambda_j - \psi_j}{\lambda_1 - \lambda_n} \leq \sum_{j=1}^l \sigma_i^2 \tan^2 \theta_j(\mathcal{X}_i, \mathcal{Y}_i), \quad l = 1, \dots, i.$$

Comparing the new bound (4.5) with (4.7) indicates a twofold improvement because $(\lambda_1 - \lambda_n)^{-1} \leq (\psi_j - \lambda_n)^{-1}$ and $\sigma_i^2 \geq \sigma_j^2 \quad \forall j \in \{1, \dots, i\}$.

In addition, setting auxiliary vectors orthogonal to the Ritz vectors associated with a number of the largest Ritz values leads to an alternative bound with terms like $\prod_{j=1}^{i-1} (\psi_j - \lambda_n) / (\psi_j - \lambda_i)$ which are suboptimal for clustered eigenvalues. Furthermore, the interpretation of (4.4) in Remark 4.2 concerning small initial subspaces also fits (4.5).

4.2. Shift-and-invert eigensolvers

We consider a generalized eigenvalue problem $Lv = \alpha Sv$ arising from the finite element discretization of an operator eigenvalue problem. Therein L and S are $n \times n$ Hermitian matrices, and S is positive definite. This formally includes the case of the finite difference discretization with $S = I_n$. Usually one only needs to compute a moderate number of eigenvalues.

By using a proper shift β , the shifted matrix $L_\beta = L - \beta S$ is invertible, and computing eigenvalues of (L, S) close to β corresponds to computing extremal eigenvalues of (S, L_β) . The latter problem can be reformulated (implicitly) as computing the largest eigenvalues of the Hermitian matrix pair

$$(\tilde{L}, M) \quad \text{with} \quad \tilde{L} = \pm L_\beta \quad \text{and} \quad M = L_\beta S^{-1} L_\beta$$

where M is positive definite. This is equivalent to computing the largest eigenvalues of

$$A = M^{-1/2} \tilde{L} M^{-1/2}.$$

A reverse transformation toward the original problem can be used to construct shift-and-invert versions of various iterative eigensolvers.

For instance, a block Krylov subspace $\mathcal{K} = \mathcal{Y} + A\mathcal{Y} + \dots + A^{k-1}\mathcal{Y}$ can be transformed as

$$\hat{\mathcal{K}} = \mathcal{Z} + M^{-1} \tilde{L} \mathcal{Z} + \dots + (M^{-1} \tilde{L})^{k-1} \mathcal{Z} \quad \text{with} \quad \hat{\mathcal{K}} = M^{-1/2} \mathcal{K} \quad \text{and} \quad \mathcal{Z} = M^{-1/2} \mathcal{Y}.$$

Such block Krylov subspaces with respect to $M^{-1} \tilde{L}$ correspond to a shift-and-invert version of the block Lanczos method. For an implementation with practical construction of $\hat{\mathcal{K}}$, we can solve linear systems of the form $Mw = r$ for certain Ritz vector residuals r similarly to the generalized Davidson method. The linear system $Mw = r$ is actually $L_\beta(S^{-1}L_\beta w) = r$ and can thus be solved as two successive linear systems for the shifted matrix L_β .

For the convergence analysis, we can reformulate Lemma 4.1 and Theorem 4.2 based on the above substitutions. Therein the parameter definition (4.3) and the Ritz value measures are reformulated by $\lambda = \pm(\alpha - \beta)^{-1}$ for eigenvalues. The transformation of angle terms is related to the inner product induced by M .

5. Numerical examples

We discuss the accuracy of our new results with several numerical examples. Examples 1 and 2 deal with the block Lanczos method applied to real diagonal matrices following a classical example from [23]. The bounds from Theorem 4.2 and their counterparts from [15] can directly be applied and demonstrated with a comparative illustration. The associated MATLAB/Octave codes (including validity checks for Theorems 3.5 and 3.8) are available on

https://github.com/lobpcg/MAJORIZATION_TYPE_CLUSTER_ROBUST_BOUNDS_FOR_BLOCK_FILTERS_AND_EIGENSOLVERS

5.1. Example 1

Similarly to [23, Subsection 4.2] and [15, Example 7.3], we consider the diagonal matrix $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ with $n = 900$ and the eigenvalues

$$\lambda_1 = 2, \quad \lambda_2 = 1.6, \quad \lambda_3 = 1.4, \quad \lambda_j = 1 - (j - 3)/n \quad \text{for } j = 4, \dots, n.$$

Combining this with the settings from Subsection 2.1 and Lemma 3.1, the invariant subspace \mathcal{X} associated with the p largest eigenvalues is spanned by the first p columns of the identity matrix I_n . We set $p = 3$ and construct the initial subspace $\mathcal{Y} \in \mathbb{R}^{n \times p}$ of the block Lanczos method using a random matrix $Y = [\text{orth}(\text{randn}(p, p)); \text{randn}(n-p, p)]$ such that the principal angles between \mathcal{X} and \mathcal{Y} are evenly distributed. We note that the initial subspace in the related examples in [15, 23] is spanned by a fixed matrix $Y = [V; \dots; V]$ with $V = [1, 1, 1; 1, 0, -2; 1, -1, 1]$ for which the principal angles between \mathcal{X} and \mathcal{Y} are actually equal. We use 1000 randomly constructed initial subspaces instead. For each of them, the corresponding block Krylov subspace $\mathcal{K} = \mathcal{Y} + A\mathcal{Y} + \dots + A^{k-1}\mathcal{Y}$ is determined up to $k = 15$ where full orthogonalization is used in order to reduce instability. Indeed, the target eigenvalues $\lambda_1, \lambda_2, \lambda_3$ in this classical example are well separated, and the gap $\lambda_3 - \lambda_4 \approx 0.4$ is sufficiently large to ensure meaningful Chebyshev terms in the bounds. Figure 1 shows the numerical comparison between our new bounds (4.4), (4.5) and the reformulated bounds (4.6), (4.7) based on [15]. In the left subfigure, the multi-angle bounds (4.4) and (4.6) are compared in the special form

$$\sum_{j=1}^t \tan \theta_j(\mathcal{X}_\tau, \mathcal{K}) \leq \sum_{j=1}^t \sigma \tan \theta_j(\mathcal{X}_\tau, \mathcal{Y}_\tau) \quad \text{with } \sigma \in [\sigma_{i_t}, \dots, \sigma_{i_1}] \quad \text{or } \sigma = \sigma_{i_t}.$$

The ‘‘Lanczos’’ curve illustrates the mean value (cf. [15]) of $\sum_{j=1}^t \tan \theta_j(\mathcal{X}_\tau, \mathcal{K})$ among 1000 samples for $\tau = \{1, 2, 3\}$ and $t = 3$ for each $k \in \{1, \dots, 15\}$. The ‘‘Chebyshev’’ curve presents the corresponding data determined for $f(A)\mathcal{Y}$ instead of \mathcal{K} by using the shifted Chebyshev polynomial defined in (4.2). The other curves contain the mean values of bounds. The accuracy of (4.6) is at the same level as that observed in [15, Table 4] concerning the measure $(\sum_{j=1}^t \sin^2 \theta_j(\mathcal{X}_\tau, \mathcal{K}))^{1/2}$. The improvement achieved by (4.4) is evident, and is also observed in a comparison with respect to maxima instead of mean values. In the right subfigure, we evaluate

$$\sum_{j=1}^i \frac{\lambda_j - \psi_j}{\lambda_1 - \lambda_n} \leq \sum_{j=1}^i \sigma^2 \tan^2 \theta_j(\mathcal{X}_i, \mathcal{Y}_i) \quad \text{with } \sigma \in [\sigma_i, \dots, \sigma_1] \quad \text{or } \sigma = \sigma_i$$

to compare the Ritz value bounds (4.5) and (4.7) for $i = 3$. Therein the denominator $\psi_j - \lambda_n$ in (4.5) is simplified as $\lambda_1 - \lambda_n$. This reduces the accuracy of (4.5), but only slightly after several iteration steps. A possible overestimation in the first iteration steps is avoided by additionally using the trivial bound $\sum_{j=1}^i \frac{\lambda_j - \psi_j}{\lambda_1 - \lambda_n} \leq i$. We observe that (4.5) is more accurate than (4.7). However, their relative accuracy is slightly worse than that of (4.4) and (4.6). This reflects the fact that the Ritz value bounds are essentially derived by combining the multi-angle bounds with further inequalities which are not necessarily sharp at the same time. Furthermore, the reader may refer to [15, Section 7] and [25, Section 5] that discuss the drawbacks of the classical bounds from [23].

5.2. Example 2

We use the diagonal matrix $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ with $n = 3600$ and the eigenvalues

$$\begin{aligned} \lambda_1 = 2.05, \quad \lambda_2 = 2, \quad \lambda_3 = 1.95, \quad \lambda_4 = 1.65, \quad \lambda_5 = 1.6, \quad \lambda_6 = 1.55, \\ \lambda_7 = 1.45, \quad \lambda_8 = 1.4, \quad \lambda_9 = 1.35, \quad \lambda_j = 1 - (j - 9)/n \quad \text{for } j = 10, \dots, n. \end{aligned}$$

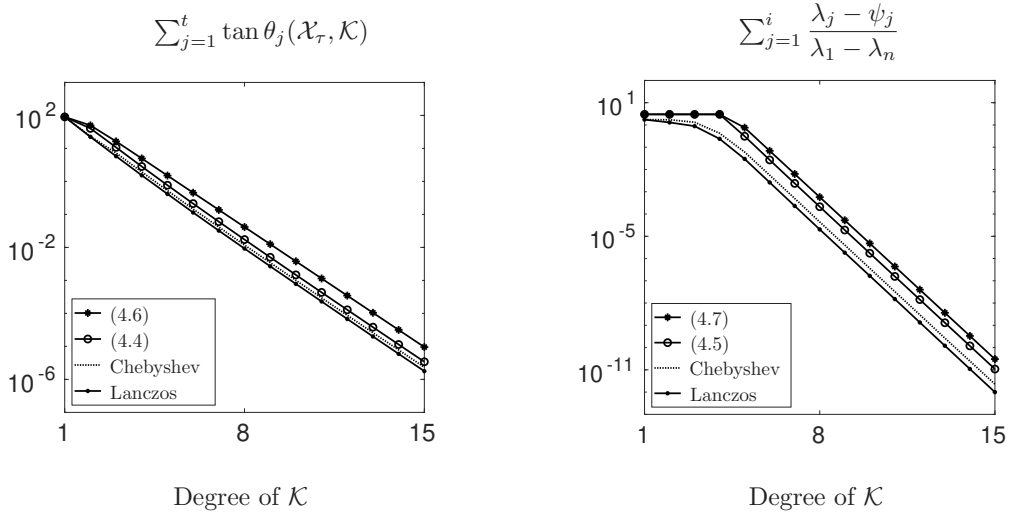


FIGURE 1. Numerical comparison of the new bounds (4.4) and (4.5) with the relevant bounds (4.6) and (4.7) based on [15] accompanying the block Chebyshev and Lanczos methods; see Example 1 with $p = 3$, $\tau = \{1, 2, 3\}$, $t = 3$ and $i = 3$.

The experiment introduced in Example 1 is run for this matrix together with $p = 9$, $\tau = \{3, \dots, 8\}$, $t = 6$ and $i = 8$; see Figure 2. The smaller distances between target eigenvalues do not deteriorate the convergence rates due to the large gap $\lambda_9 - \lambda_{10} \approx 0.35$. This also results in suitable Chebyshev terms so that the slopes of the bound curves accurately reflect the convergence rates and the cluster robustness, at least in the final phase. Our new bounds (4.4) and (4.5) provide visible improvements in comparison to (4.6) and (4.7) based on [15]. The corresponding bounds from [15] already dramatically improve the classical bounds from [23] involving the “bulky” factors; see [15, Example 7.2].

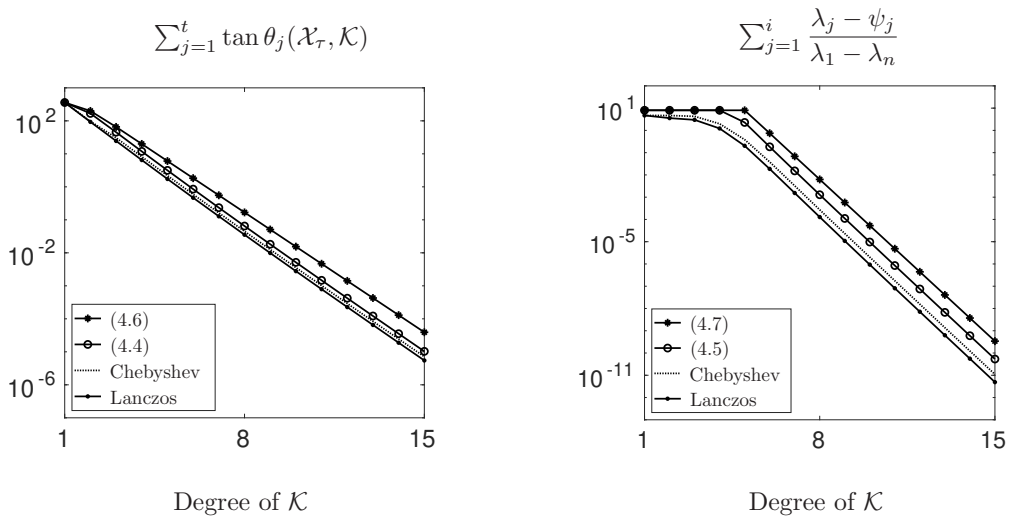


FIGURE 2. Numerical comparison of the new bounds (4.4) and (4.5) with the relevant bounds (4.6) and (4.7) based on [15] accompanying the block Chebyshev and Lanczos methods; see Example 2 with $p = 9$, $\tau = \{3, \dots, 8\}$, $t = 6$ and $i = 8$.

Conclusions

Majorization techniques exploring relations between Ritz value errors and principal angles are extended to block signal filters and subspace iterates of block eigensolvers. The majorization-type analysis of tangents of principal angles from [12] is improved by using auxiliary vectors from the classical analysis of the block power method by Rutishauser [22]. This leads to novel majorization bounds for the Rayleigh-Ritz method applied to the final iterative subspace in terms of principal angles between an initial subspace and a target invariant subspace. Our results improve the existing approaches from [15, 23] that are used for the block Lanczos method. Our majorization technique is especially advantageous in situations where the concerned principal angles are evenly distributed, which is common for random initial subspaces. Our forthcoming work will focus on angle-free bounds for Ritz value errors which are applicable to restarted iterations, such as the stochastic block descent. Simultaneous consideration of several Ritz values approximating clustered eigenvalues enables meaningful bounds and has potential for investigating block preconditioned eigensolvers. A future aim is to cover the popular locally optimal block preconditioned conjugate gradient (LOBPCG) method [10, 13], taking advantage of techniques developed in [20, 26].

6. Appendix: Detailed proofs

6.1. Proof of Lemma 2.1

By using the orthonormal basis matrix $U = \tilde{U}G$ with $G = (\tilde{U}^H\tilde{U})^{-1/2} \in \mathbb{C}^{s \times s}$, the cosine values of the principal angles $\theta_1 \geq \dots \geq \theta_s$ from \mathcal{U} to \mathcal{V} are given by the singular values of $V^H U \in \mathbb{C}^{t \times s}$. More precisely, a standard singular value decomposition $V^H U = W \Sigma Z^H$ provides unitary matrices $W \in \mathbb{C}^{t \times t}$ and $Z \in \mathbb{C}^{s \times s}$ together with the diagonal (rectangle) matrix

$$\Sigma = \begin{bmatrix} D \\ O \end{bmatrix} \in \mathbb{R}^{t \times s} \quad \text{consisting of} \quad D = \text{diag}(\cos(\theta_s), \dots, \cos(\theta_1)) \in \mathbb{R}^{s \times s}$$

and the zero matrix $O \in \mathbb{R}^{(t-s) \times s}$. Then $\hat{V} = VW$ and $\hat{U} = UZ = \tilde{U}GZ$ are orthonormal basis matrices of \mathcal{U} and \mathcal{V} with the property $\Sigma = \hat{V}^H \hat{U} = W^H V^H \tilde{U} G Z$. Since $V^H \tilde{U}$ has full rank, so does Σ . In addition, it holds that

$$\begin{aligned} V_{\perp}^H \hat{U} \Sigma^{\dagger} &= V_{\perp}^H \hat{U} (\hat{V}^H \hat{U})^{\dagger} = V_{\perp}^H \tilde{U} G Z (W^H V^H \tilde{U} G Z)^{\dagger} = V_{\perp}^H \tilde{U} G Z (G Z)^{\dagger} (V^H \tilde{U})^{\dagger} (W^H)^{\dagger} \\ &= V_{\perp}^H \tilde{U} G Z (G Z)^{-1} (V^H \tilde{U})^{\dagger} (W^H)^{-1} = (V_{\perp}^H \tilde{U} (V^H \tilde{U})^{\dagger}) W \end{aligned}$$

so that the singular values of $V_{\perp}^H \hat{U} \Sigma^{\dagger}$ coincide with those of $V_{\perp}^H \tilde{U} (V^H \tilde{U})^{\dagger}$. Their squared values are the $\min\{t, n-t\}$ largest eigenvalues of the $t \times t$ matrix

$$\begin{aligned} (V_{\perp}^H \hat{U} \Sigma^{\dagger})^H V_{\perp}^H \hat{U} \Sigma^{\dagger} &= (\Sigma^{\dagger})^H \hat{U}^H V_{\perp} V_{\perp}^H \hat{U} \Sigma^{\dagger} = (\Sigma^{\dagger})^H \hat{U}^H (I_n - \hat{V} \hat{V}^H) \hat{U} \Sigma^{\dagger} \\ (6.1) \quad &= (\Sigma^{\dagger})^H \hat{U}^H \hat{U} \Sigma^{\dagger} - (\Sigma^{\dagger})^H \hat{U}^H \hat{V} \hat{V}^H \hat{U} \Sigma^{\dagger} \\ &= (\Sigma^{\dagger})^H \Sigma^{\dagger} - (\Sigma^{\dagger})^H \Sigma^H \Sigma \Sigma^{\dagger} = (\Sigma^{\dagger})^H \Sigma^{\dagger} - (\Sigma \Sigma^{\dagger})^H (\Sigma \Sigma^{\dagger}). \end{aligned}$$

Therein $V_{\perp} V_{\perp}^H$ is the orthogonal projector on \mathcal{V}_{\perp} and thus coincides with $I_n - \hat{V} \hat{V}^H$ for the orthogonal projector $\hat{V} \hat{V}^H$ on \mathcal{V} . Moreover, $\Sigma^{\dagger} = [D^{-1} \ O^H]$ holds as Σ has full rank. Then (6.1) implies

$$(V_{\perp}^H \hat{U} \Sigma^{\dagger})^H V_{\perp}^H \hat{U} \Sigma^{\dagger} = \begin{bmatrix} D^{-2} & \\ & \tilde{O} \end{bmatrix} - \begin{bmatrix} I_s & \\ & \tilde{O} \end{bmatrix}^H \begin{bmatrix} I_s & \\ & \tilde{O} \end{bmatrix} = \begin{bmatrix} D^{-2} - I_s & \\ & \tilde{O} \end{bmatrix}$$

with the zero matrix $\tilde{O} = OO^H$. Thus the s largest singular values of $V_{\perp}^H \tilde{U} (V^H \tilde{U})^{\dagger}$ or $V_{\perp}^H \hat{U} \Sigma^{\dagger}$ coincide with the square roots of the diagonal entries of the diagonal matrix

$$D^{-2} - I_s = \text{diag}\left(\left(\cos(\theta_s)\right)^{-2} - 1, \dots, \left(\cos(\theta_1)\right)^{-2} - 1\right) = \text{diag}\left(\tan^2(\theta_s), \dots, \tan^2(\theta_1)\right),$$

i.e., the tangent values of the principal angles from \mathcal{U} to \mathcal{V} . The largest singular value also coincides with the 2-norm so that $\|V_{\perp}^H \tilde{U} (V^H \tilde{U})^{\dagger}\| = \tan(\theta_1) = \tan \angle(\mathcal{U}, \mathcal{V})$. \square

6.2. Proof of Lemma 3.2

Applying Lemma 2.1 to

$$V = x_i, \quad V_{\perp} = [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n], \quad \tilde{U} = y_i$$

shows that the only singular value of the $(n-1) \times 1$ matrix

$$w = [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]^H y_i (x_i^H y_i)^{\dagger} \stackrel{(3.1)}{=} [0; \dots; 0; x_{p+1}^H y_i; \dots; x_n^H y_i].$$

coincides with the tangent value of the only principal angle from $\text{span}\{y_i\}$ to $\text{span}\{x_i\}$. Then

$$\|w\| = |\tan \angle(y_i, x_i)| = |\tan \angle(x_i, y_i)| = \tan \angle(x_i, y_i)$$

holds since $x_i^H y_i = 1 > 0$ leads to $\tan \angle(x_i, y_i) > 0$. In addition, the common assumption (2.1) ensures $f(\lambda_i) \neq 0$ and $x_i^H f(A) y_i = ((f(A))^H x_i)^H y_i = (\overline{f(\lambda_i)} x_i)^H y_i = f(\lambda_i) x_i^H y_i = f(\lambda_i) \neq 0$. Thus $f(A) y_i$ is a nonzero vector. Applying Lemma 2.1 to

$$V = x_i, \quad V_{\perp} = [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n], \quad \tilde{U} = f(A) y_i$$

analogously implies $|\tan \angle(x_i, f(A) y_i)| = \|w'\|$ for

$$\begin{aligned} w' &= [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]^H f(A) y_i (x_i^H f(A) y_i)^{\dagger} \\ &\stackrel{(3.1)}{=} [0; \dots; 0; f(\lambda_{p+1}) x_{p+1}^H y_i; \dots; f(\lambda_n) x_n^H y_i] (f(\lambda_i))^{-1}. \end{aligned}$$

Then a simple comparison between $\|w'\|$ and $\|w\|$ shows

$$|\tan \angle(x_i, f(A) y_i)| \leq \frac{\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)|}{|f(\lambda_i)|} \tan \angle(x_i, y_i).$$

This implies the first inequality in (3.2) according to $f(A) y_i \in f(A) \mathcal{Y} = \mathcal{Y}'$.

Moreover, the term $\tan \angle(x_i, y_i)$ can be enlarged by $\tan \angle(\mathcal{X}, \mathcal{Y})$ for eliminating the auxiliary vector y_i ; cf. (1.5). This modification is enabled by projection arguments or by applying Lemma 2.1 to

$$V = X = [x_1, \dots, x_p], \quad V_{\perp} = [x_{p+1}, \dots, x_n], \quad \tilde{U} = y_i.$$

Therein $\tan \angle(y_i, \mathcal{X}) = \|W\|$ holds for

$$W = [x_{p+1}, \dots, x_n]^H y_i ([x_1, \dots, x_p]^H y_i)^{\dagger} = \tilde{w} e_i^{\dagger} = \tilde{w} e_i^H$$

with $\tilde{w} = [x_{p+1}^H y_i; \dots; x_n^H y_i]$ and the i th column e_i of the identity matrix I_p . Then

$$\|W\| = \|\tilde{w} e_i^H\| = \|\tilde{w}\| \|e_i\| = \|\tilde{w}\| = \|w\|$$

holds by comparing the components of w and \tilde{w} . Thus

$$\tan \angle(x_i, y_i) = \|w\| = \|W\| = \tan \angle(y_i, \mathcal{X}) \leq \tan \angle(\mathcal{Y}, \mathcal{X}) = \tan \angle(\mathcal{X}, \mathcal{Y})$$

leads to the second inequality in (3.2). \square

Bound (3.2) can also be shown in an elementary way; cf. [25, Theorem 1] concerning the specified function (1.4). The proof of Lemma 3.2 aims at motivating that of Theorem 3.5 where majorization bounds are to be derived.

6.3. Proof of Lemma 3.3

The common assumption (2.1) ensures $f(\lambda_j) \neq 0$ for each $j \in \{1, \dots, i\}$. The $i \times i$ matrix

$$X_i^H Y_i' = X_i^H f(A) Y_i = \text{diag}(f(\lambda_1), \dots, f(\lambda_i)) X_i^H Y_i \stackrel{(3.1)}{=} \text{diag}(f(\lambda_1), \dots, f(\lambda_i))$$

is then invertible. Thus Y_i' has rank i , and $\dim \mathcal{Y}_i' = i$. Moreover, the relation $\mathcal{Y}' = f(A) \text{span}\{Y\} = \text{span}\{f(A)Y\} = \text{span}\{f(A)Y_p\} = \text{span}\{Y_p'\} = \mathcal{Y}'_p$ holds so that $\dim \mathcal{Y}' = p$.

In order to show (3.4), the property $\dim \mathcal{Y}'_i = i = \dim \mathcal{X}_i$ enables the representation

$$\tan \angle(\mathcal{X}_i, \mathcal{Y}'_i) = \tan \angle(\mathcal{Y}'_i, \mathcal{X}_i) = \max_{y' \in \mathcal{Y}'_i \setminus \{0\}} \tan \angle(y', \mathcal{X}_i) = \tan \angle(\hat{y}', \mathcal{X}_i)$$

with a maximizer $\hat{y}' \in \mathcal{Y}'_i \setminus \{0\}$. Because $\mathcal{Y}'_i = \text{span}\{f(A)Y_i\} = f(A)\mathcal{Y}_i$, we can represent \hat{y}' by $\hat{y}' = f(A)\hat{y}$ with a certain $\hat{y} \in \mathcal{Y}_i \setminus \{0\}$. Applying Lemma 2.1 to

$$V = X_i, \quad V_\perp = [x_{i+1}, \dots, x_n], \quad \tilde{U} = \hat{y}' = f(A)\hat{y}$$

(there is only one principal angle in this case) yields $\tan \angle(\hat{y}', \mathcal{X}_i) = \|W'\|$ for

$$W' = [x_{i+1}, \dots, x_n]^H f(A)\hat{y} (X_i^H f(A)\hat{y})^\dagger = w' e'^\dagger$$

$$\text{with } w' = [0; \dots; 0; f(\lambda_{p+1})x_{p+1}^H \hat{y}; \dots; f(\lambda_n)x_n^H \hat{y}]$$

$$\text{and } e' = [f(\lambda_1)x_1^H \hat{y}; \dots; f(\lambda_i)x_i^H \hat{y}].$$

The vector w' initially consists of $x_j^H f(A)\hat{y} = f(\lambda_j)x_j^H \hat{y}$ for $j = i+1, \dots, n$, but the first components up to the index p are simply zero due to (3.1) and that \hat{y} belongs to $\mathcal{Y}_i = \text{span}\{y_1, \dots, y_i\}$. Moreover, \hat{y} can be represented by $\hat{y} = Y_i g$ with a certain $g \in \mathbb{C}^i \setminus \{0\}$ so that

$$e' = \text{diag}(f(\lambda_1), \dots, f(\lambda_i)) X_i^H Y_i g \stackrel{(3.1)}{=} \text{diag}(f(\lambda_1), \dots, f(\lambda_i)) g \neq 0.$$

Then $e'^\dagger = e'^H / \|e'\|^2$, and

$$\tan \angle(\hat{y}', \mathcal{X}_i) = \|W'\| = \|w' e'^H / \|e'\|^2\| = \frac{\|w' e'^H\|}{\|e'\|^2} = \frac{\|w'\| \|e'\|}{\|e'\|^2} = \frac{\|w'\|}{\|e'\|}.$$

Analogously, $\tan \angle(\hat{y}, \mathcal{X}_i)$ has the representation

$$\tan \angle(\hat{y}, \mathcal{X}_i) = \frac{\|w\|}{\|e\|} \quad \text{with } w = [0; \dots; 0; x_{p+1}^H \hat{y}; \dots; x_n^H \hat{y}]$$

$$\text{and } e = [x_1^H \hat{y}; \dots; x_i^H \hat{y}].$$

Summarizing the above together with

$$\|w'\| \leq (\max_{j \in \{p+1, \dots, n\}} |f(\lambda_j)|) \|w\|, \quad \|e'\| \geq (\min_{j \in \{1, \dots, i\}} |f(\lambda_j)|) \|e\|$$

and $\tan \angle(\hat{y}, \mathcal{X}_i) \leq \tan \angle(\mathcal{Y}_i, \mathcal{X}_i) = \tan \angle(\mathcal{X}_i, \mathcal{Y}_i)$ yields (3.4). \square

The above proof focuses on the rank-1 matrices $w' e'^\dagger$ and $w e^\dagger$. An alternative proof using rank- i matrices is involved in the corresponding majorization-type analysis; see the proofs of Theorem 3.5 and Lemma 3.6.

6.4. Proof of Lemma 3.4

We use the smallest Ritz value of A in \mathcal{Y}'_i and an associated normalized Ritz vector which are denoted by η'_i and y'_i . Then

$$\mathcal{Y}'_i \subseteq \mathcal{Y}' \subset \mathbb{C}^n \quad \Rightarrow \quad \lambda_i \geq \eta'_i \geq \eta' \geq \lambda_n$$

according to the Courant-Fischer principles and $\dim \mathcal{Y}'_i = i$. Moreover, (3.5) can be implied by

$$(6.2) \quad \frac{\lambda_i - \eta'_i}{\eta'_i - \lambda_n} \leq \tan^2 \angle(y'_i, \mathcal{X}_i)$$

since y' is a nonzero vector in \mathcal{Y}'_i and thus $\tan^2 \angle(y', \mathcal{X}_i) \leq \tan^2 \angle(\mathcal{Y}'_i, \mathcal{X}_i) = \tan^2 \angle(\mathcal{X}_i, \mathcal{Y}'_i)$. An elementary trigonometric transformation shows that (6.2) is equivalent to

$$(6.3) \quad \cos^2 \angle(y', \mathcal{X}_i) \leq \frac{\eta'_i - \lambda_n}{\lambda_i - \lambda_n}.$$

Thus we only need to show (6.3). Therein we use the standard definition of principal angles which indicates that $\cos \angle(y', \mathcal{X}_i)$ coincides with the only singular value of $X_i^H y'$, i.e., $\|X_i^H y'\|$. In addition, we consider the Hermitian matrices

$$\tilde{A} = A - \lambda_n I_n \quad \text{and} \quad \tilde{D} = \text{diag}(\lambda_1 - \lambda_n, \dots, \lambda_i - \lambda_n).$$

Evidently, \tilde{A} is positive semidefinite, and \tilde{D} is positive definite because $\lambda_1 \geq \dots \geq \lambda_i \geq \lambda_p > \lambda_{p+1} \geq \lambda_n$. By using their square roots $\tilde{A}^{1/2}$ and $\tilde{D}^{1/2}$, we get $\tilde{A}^{1/2} X_i = X_i \tilde{D}^{1/2}$ so that

$$X_i^H y' = \tilde{D}^{-1/2} \tilde{D}^{1/2} X_i^H y' = \tilde{D}^{-1/2} (X_i \tilde{D}^{1/2})^H y' = \tilde{D}^{-1/2} (\tilde{A}^{1/2} X_i)^H y' = \tilde{D}^{-1/2} X_i^H \tilde{A}^{1/2} y'$$

$$\Rightarrow \cos \angle(y', \mathcal{X}_i) = \|X_i^H y'\| = \|\tilde{D}^{-1/2} X_i^H \tilde{A}^{1/2} y'\| \leq \|\tilde{D}^{-1/2}\| \|X_i^H \tilde{A}^{1/2} y'\|.$$

Combining this with

$$\begin{aligned} \|\tilde{D}^{-1/2}\| &= \max_{j=1, \dots, i} |(\lambda_j - \lambda_n)^{-1/2}| = (\lambda_i - \lambda_n)^{-1/2} \quad \text{and} \\ \|X_i^H \tilde{A}^{1/2} y'\| &\leq \|X_i^H\| \|\tilde{A}^{1/2} y'\| = \|\tilde{A}^{1/2} y'\| = (y'^H \tilde{A} y')^{1/2} \\ &= (y'^H A y' - \lambda_n)^{1/2} = (\eta'_i - \lambda_n)^{1/2} \leq (\eta'_i - \lambda_n)^{1/2} \end{aligned}$$

yields (6.3). \square

The proof of Lemma 3.4 uses inequalities with respect to the 2-norm, i.e., the largest singular value. In the corresponding majorization-type analysis, we consider tuples of singular values; see the proof of Lemma 3.7.

6.5. Extending (3.6) as (3.3)

For this extension, we apply Lemma 2.1 to

$$V = X_i = [x_1, \dots, x_i], \quad V_\perp = [x_{i+1}, \dots, x_n], \quad \tilde{U} = Y_i = [y_1, \dots, y_i].$$

This leads to $\tan \angle(\mathcal{X}_i, \mathcal{Y}_i) = \tan \angle(\mathcal{Y}_i, \mathcal{X}_i) = \|W\|$ for

$$W = [x_{i+1}, \dots, x_n]^H Y_i (X_i^H Y_i)^\dagger \stackrel{(3.1)}{=} [O; x_{p+1}^H Y_i; \dots; x_n^H Y_i] I_i^\dagger = \begin{bmatrix} O \\ G \end{bmatrix}$$

with $G = [x_{p+1}^H Y_i; \dots; x_n^H Y_i] \in \mathbb{C}^{(n-p) \times i}$ and the zero matrix $O \in \mathbb{R}^{(p-i) \times i}$. Moreover, applying Lemma 2.1 to

$$V = X = [x_1, \dots, x_p], \quad V_\perp = [x_{p+1}, \dots, x_n], \quad \tilde{U} = Y = [y_1, \dots, y_p]$$

yields $\tan \angle(\mathcal{X}, \mathcal{Y}) = \tan \angle(\mathcal{Y}, \mathcal{X}) = \|\hat{G}\|$ for

$$\hat{G} = [x_{p+1}, \dots, x_n]^H Y (X^H Y)^\dagger \stackrel{(3.1)}{=} [x_{p+1}^H Y; \dots; x_n^H Y] I_p^\dagger = [x_{p+1}^H Y; \dots; x_n^H Y] \in \mathbb{C}^{(n-p) \times p}.$$

In summary, it holds that

$$\tan \angle(\mathcal{X}_i, \mathcal{Y}_i) = \|W\| = \|G\| = \|\hat{G} E_i\| \leq \|\hat{G}\| = \tan \angle(\mathcal{X}, \mathcal{Y})$$

where $E_i \in \mathbb{R}^{p \times i}$ consists of the first i columns of I_p .

6.6. Proof of Theorem 3.5

Applying Lemma 2.1 to

$$V = X_\tau = [x_{i_1}, \dots, x_{i_t}], \quad V_\perp = [x_j \text{ for } j \in \{1, \dots, n\} \setminus \tau], \quad \tilde{U} = Y_\tau = [y_{i_1}, \dots, y_{i_t}]$$

and $s = t$ generates the $(n-t) \times t$ matrix

$$W = [x_j \text{ for } j \in \{1, \dots, n\} \setminus \tau]^H Y_\tau (X_\tau^H Y_\tau)^\dagger \stackrel{(3.1)}{=} [O; x_{p+1}^H Y_\tau; \dots; x_n^H Y_\tau]$$

with the zero matrix $O \in \mathbb{R}^{(p-t) \times t}$ where $(X_\tau^H Y_\tau)^\dagger = I_t^\dagger = I_t$ is eliminated. Moreover, $t \leq p \ll n$ holds according to Subsection 2.1 and ensures $t \leq n-t$ so that W has t singular values. By using the notations from Subsection 2.3, the statement in Lemma 2.1 yields

$$S(W) = \tan \Theta(\mathcal{Y}_\tau, \mathcal{X}_\tau) = \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau).$$

In addition, the common assumption (2.1) ensures $f(\lambda_j) \neq 0 \quad \forall j \in \tau$ so that

$$D_\tau = X_\tau^H f(A) Y_\tau = \text{diag}(f(\lambda_{i_1}), \dots, f(\lambda_{i_t})) X_\tau^H Y_\tau \stackrel{(3.1)}{=} \text{diag}(f(\lambda_{i_1}), \dots, f(\lambda_{i_t}))$$

is an invertible diagonal matrix. Consequently, the $n \times t$ matrix $f(A) Y_\tau$ has rank t , and the subspace $\mathcal{Y}'_\tau = \text{span}\{f(A) Y_\tau\}$ has dimension t . Applying Lemma 2.1 to

$$V = X_\tau, \quad V_\perp = [x_j \text{ for } j \in \{1, \dots, n\} \setminus \tau], \quad \tilde{U} = f(A) Y_\tau$$

implies $S(W') = \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}'_\tau)$ for the $(n-t) \times t$ matrix

$$\begin{aligned} W' &= [x_j \text{ for } j \in \{1, \dots, n\} \setminus \tau]^H f(A) Y_\tau (X_\tau^H f(A) Y_\tau)^\dagger \\ &\stackrel{(3.1)}{=} [O; f(\lambda_{p+1}) x_{p+1}^H Y_\tau; \dots; f(\lambda_n) x_n^H Y_\tau] D_\tau^\dagger \\ &= \hat{D} W D_\tau^{-1} \quad \text{with} \quad \hat{D} = \text{diag}(0, \dots, 0, f(\lambda_{p+1}), \dots, f(\lambda_n)). \end{aligned}$$

Furthermore, applying Lemma 2.2 to

$$B_1 = \hat{D}, \quad B_2 = W, \quad B_3 = D_\tau^{-1}$$

and $c = 1$ leads to

$$\tan \Theta(\mathcal{X}_\tau, \mathcal{Y}'_\tau) = S(W') = S_t(W') \prec_w S_t(\hat{D}) S_t(W) S_t(D_\tau^{-1}) = S(D_\tau^{-1}) S_t(\hat{D}) S(W).$$

Combining this with

$$\begin{aligned} S(D_\tau^{-1}) &= \left[|(f(\lambda_{i_1}))^{-1}|, \dots, |(f(\lambda_{i_t}))^{-1}| \right]^\downarrow = \Phi_\tau, \\ S(\hat{D}) &= [0, \dots, 0, |f(\lambda_{p+1})|, \dots, |f(\lambda_n)|]^\downarrow \Rightarrow S_t(\hat{D}) = \hat{\Phi}_t \end{aligned}$$

and $S(W) = \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau)$ yields

$$(6.4) \quad \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}'_\tau) \prec_w \Phi_\tau \hat{\Phi}_t \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}_\tau).$$

Then (3.7) is obtained by using the tuple inequality $\tan \Theta(\mathcal{X}_\tau, \mathcal{Y}') \leq \tan \Theta(\mathcal{X}_\tau, \mathcal{Y}'_\tau)$ based on $\mathcal{Y}'_\tau \subseteq \text{span}\{f(A) Y_\tau\} = \mathcal{Y}'$. This inequality can be shown in the equivalent form $(\cos^2 \Theta(\mathcal{X}_\tau, \mathcal{Y}'))^\downarrow \geq (\cos^2 \Theta(\mathcal{X}_\tau, \mathcal{Y}'_\tau))^\downarrow$ as follows. An arbitrary orthonormal basis matrix \tilde{Y} of \mathcal{Y}'_τ can be extended as an orthonormal basis matrix $[\tilde{Y} \quad \hat{Y}]$ of \mathcal{Y}' so that

$$\begin{aligned} (\cos^2 \Theta(\mathcal{X}_\tau, \mathcal{Y}'))^\downarrow &= S^2([\tilde{Y} \quad \hat{Y}]^H X_\tau) = \Lambda(([\tilde{Y} \quad \hat{Y}]^H X_\tau)^H [\tilde{Y} \quad \hat{Y}]^H X_\tau) \\ &= \Lambda(X_\tau^H \tilde{Y} \tilde{Y}^H X_\tau + X_\tau^H \hat{Y} \hat{Y}^H X_\tau) \geq \Lambda(X_\tau^H \tilde{Y} \tilde{Y}^H X_\tau) = (\cos^2 \Theta(\mathcal{X}_\tau, \mathcal{Y}'_\tau))^\downarrow. \end{aligned}$$

Therein the intermediate inequality uses the Weyl's inequality and the fact that $X_\tau^H \hat{Y} \hat{Y}^H X_\tau$ is Hermitian positive semidefinite. \square

6.7. Derivation of (3.8)

For deriving (3.8), we use the given basis matrix \tilde{U} of \mathcal{U} together with the orthogonal projectors VV^H on \mathcal{V} so that

$$V^H F \tilde{U} = (F^H V)^H \tilde{U} = ((VV^H)F^H V)^H \tilde{U} = (V^H F V) V^H \tilde{U}.$$

Analogously, $V_{\perp}^H F \tilde{U} = (V_{\perp}^H F V_{\perp}) V_{\perp}^H \tilde{U}$. Moreover, $V^H F V \in \mathbb{C}^{t \times t}$ is invertible and $V^H \tilde{U} \in \mathbb{C}^{t \times s}$ has full rank by setting, thus $V^H F \tilde{U}$ has full rank. Then $F \tilde{U}$ also has full rank and can be treated as a basis matrix of $F\mathcal{U}$ with Lemma 2.1. By using the matrices

$$G = V^H F V, \quad \tilde{W} = V^H \tilde{U}, \quad D = (\tilde{W}^H \tilde{W})^{1/2} \quad \text{and} \quad W = \tilde{W} D^{-1},$$

it holds that

$$(V^H F \tilde{U})^{\dagger} = (G \tilde{W})^{\dagger} = \tilde{W}^{\dagger} \tilde{W} (G W D)^{\dagger} = \tilde{W}^{\dagger} \tilde{W} D^{-1} (G W)^{\dagger} = (V^H \tilde{U})^{\dagger} W (G W)^{\dagger}.$$

In summary, we get

$$V_{\perp}^H F \tilde{U} (V^H F \tilde{U})^{\dagger} = (V_{\perp}^H F V_{\perp}) (V_{\perp}^H \tilde{U} (V^H \tilde{U})^{\dagger}) (W (G W)^{\dagger}).$$

Subsequently, we consider the singular values $\alpha_1 \leq \dots \leq \alpha_t$ of G and the singular values $\beta_1 \leq \dots \leq \beta_s$ of $G W$. These values are evidently all positive, and it holds that $\alpha_i \leq \beta_i$, $i = 1, \dots, s$ according to $W^H W = I_s$ and the Courant-Fischer principles. This leads to the tuple inequality

$$S_s(W (G W)^{\dagger}) = S((G W)^{\dagger}) = [\beta_1^{-1}, \dots, \beta_s^{-1}] \leq [\alpha_1^{-1}, \dots, \alpha_s^{-1}] = S_s(G^{-1}) = S_s((V^H F V)^{-1})$$

which can easily be combined with Lemma 2.2 for showing

$$S_s(V_{\perp}^H F \tilde{U} (V^H F \tilde{U})^{\dagger}) \prec_w S_s(V_{\perp}^H F V_{\perp}) S_s(V_{\perp}^H \tilde{U} (V^H \tilde{U})^{\dagger}) S_s((V^H F V)^{-1}),$$

i.e., an equivalent form of (3.8) according to Lemma 2.1.

6.8. Proof of Lemma 3.7

As in the proof of Lemma 3.4, we use the Hermitian positive semidefinite matrix $\tilde{A} = A - \lambda_n I_n$ and the Hermitian positive definite matrix $\tilde{D} = \text{diag}(\lambda_1 - \lambda_n, \dots, \lambda_i - \lambda_n)$ together with their square roots $\tilde{A}^{1/2}$ and $\tilde{D}^{1/2}$. In addition, we define the orthonormal basis matrix $\tilde{Y} = Y_i' (Y_i'^H Y_i')^{-1/2}$ of \mathcal{Y}_i' by using the basis matrix $Y_i' = f(A) Y_i$ introduced in Lemma 3.3. Then the Ritz values $\tilde{\eta}_1 \geq \dots \geq \tilde{\eta}_i$ of A in \mathcal{Y}_i' coincide with the eigenvalues of $\tilde{Y}^H A \tilde{Y}$ so that

$$\Lambda(\tilde{Y}^H A \tilde{Y}) = \Lambda(\tilde{Y}^H A \tilde{Y} - \lambda_n I_i) = [\tilde{\eta}_1 - \lambda_n, \dots, \tilde{\eta}_i - \lambda_n] \leq [\eta_1' - \lambda_n, \dots, \eta_i' - \lambda_n]$$

holds according to $\mathcal{Y}_i' \subseteq \mathcal{Y}'$ and the Courant-Fischer principles.

Moreover, $X_i^H \tilde{Y}$ is invertible since the invertibility of $X_i^H Y_i'$ is verified at the beginning of the proof of Lemma 3.3. Then we transform $X_i^H \tilde{Y}$ as

$$X_i^H \tilde{Y} = \tilde{D}^{-1/2} \tilde{D}^{1/2} X_i^H \tilde{Y} = \tilde{D}^{-1/2} (X_i \tilde{D}^{1/2})^H \tilde{Y} = \tilde{D}^{-1/2} (\tilde{A}^{1/2} X_i)^H \tilde{Y} = \tilde{D}^{-1/2} X_i^H \tilde{A}^{1/2} \tilde{Y}$$

so that

$$\tilde{D}^{1/2} = X_i^H (\tilde{A}^{1/2} \tilde{Y}) (X_i^H \tilde{Y})^{-1}.$$

Then applying Lemma 2.2 to

$$B_1 = X_i^H, \quad B_2 = \tilde{A}^{1/2} \tilde{Y}, \quad B_3 = (X_i^H \tilde{Y})^{-1},$$

$t = i$ and $c = 2$ yields a majorization relation which can be extended as

$$\begin{aligned} & \left[\frac{\lambda_1 - \lambda_n}{\eta'_1 - \lambda_n}, \dots, \frac{\lambda_i - \lambda_n}{\eta'_i - \lambda_n} \right] \leq \left[\frac{\lambda_1 - \lambda_n}{\tilde{\eta}_1 - \lambda_n}, \dots, \frac{\lambda_i - \lambda_n}{\tilde{\eta}_i - \lambda_n} \right] \\ & = \Lambda(\tilde{D})/\Lambda(\tilde{Y}^H \tilde{A} \tilde{Y}) = S_i^2(\tilde{D}^{1/2})/S_i^2(\tilde{A}^{1/2} \tilde{Y}) \prec_w S_i^2(X_i^H) S_i^2((X_i^H \tilde{Y})^{-1}) \\ & = [1, \dots, 1] \left[(\cos(\theta_1))^{-2}, \dots, (\cos(\theta_i))^{-2} \right] = [1 + \tan^2(\theta_1), \dots, 1 + \tan^2(\theta_i)] \end{aligned}$$

with the associated principal angles $\theta_1 \geq \dots \geq \theta_i$. Thus

$$\left[\frac{\lambda_1 - \lambda_n}{\eta'_1 - \lambda_n}, \dots, \frac{\lambda_i - \lambda_n}{\eta'_i - \lambda_n} \right] \prec_w [1 + \tan^2(\theta_1), \dots, 1 + \tan^2(\theta_i)]$$

holds and implies bound (3.11) by subtracting 1 from each component. \square

Remark 6.1. Bound (3.11) is comparable with [12, Theorems 2.3 and 2.4] where actually (despite different notations) the smallest Ritz value of A in $\mathcal{X}_i + \mathcal{Y}'_i$ or $\mathcal{X}_i + \mathcal{Y}'$ is used instead of the smallest eigenvalue λ_n . Such a modification can be made by restricting the proof of Lemma 3.7 to the associated subspace. In comparison to the corresponding derivation in [12], our new derivation is more direct (e.g., without shifting) and the factorization $\tilde{D}^{1/2} = X_i^H (\tilde{A}^{1/2} \tilde{Y}) (X_i^H \tilde{Y})^{-1}$ enables a simpler formulation.

REFERENCES

- [1] M. E. ARGENTATI AND A. V. KNYAZEV, *Majorization-based convergence rate bounds for subspace iterations and the block Lanczos method*, Talk at the 10th Copper Mountain Conference on Iterative Methods, 2008. doi:10.13140/RG.2.2.31726.74565
- [2] R. BHATIA, *Matrix Analysis*, Springer, New York, NY, 1997. ISBN: 978-0-387-94846-1
- [3] P. DRINEAS, I. C. F. IPSEN, E. M. KONTOPOULOU, AND M. MAGDON-ISMAIL, *Structural convergence results for approximation of dominant subspaces from block Krylov spaces*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 567–586. doi:10.1137/16M1091745
- [4] J. GOPALAKRISHNAN, L. GRUBIŠIĆ, AND J. O'VALL, *Spectral discretization errors in filtered subspace iteration*, Math. Comp., 89 (2020), pp. 203–228.
- [5] J. GOPALAKRISHNAN, B. Q. PARKER, AND P. VANDENBERGE, *Computing leaky modes of optical fibers using a FEAST algorithm for polynomial eigenproblems*, Wave Motion, 108 (2022), 102826.
- [6] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, New York, NY, 1994. ISBN: 978-0521467131
- [7] A. V. KNYAZEV, *Computation of Eigenvalues and Eigenvectors for Mesh Problems: Algorithms and Error Estimates*, Dept. Numerical Math. USSR Academy of Sciences, Moscow, 1986. (In Russian).
- [8] A. V. KNYAZEV, *Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem*, Soviet J. Num. Anal. Math. Modelling, 2 (1987), pp. 371–396.
- [9] A. V. KNYAZEV, *New estimates for Ritz vectors*, Math. Comp., 66 (1997), pp. 985–995.
- [10] A. V. KNYAZEV, *Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method*, SIAM J. Sci. Comput., 23 (2001), pp. 517–541. doi:10.1137/S1064827500366124
- [11] A. V. KNYAZEV AND M. E. ARGENTATI, *Rayleigh-Ritz majorization error bounds with applications to FEM and subspace iterations*, Technical Report <https://arxiv.org/abs/math/0701784v2>, 2008.
- [12] A. V. KNYAZEV AND M. E. ARGENTATI, *Rayleigh-Ritz majorization error bounds with applications to FEM*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1521–1537. doi:10.1137/08072574X

- [13] A. V. KNYAZEV, M. E. ARGENTATI, I. LASHUK, AND E. E. OVTCHINNIKOV, *Block Locally Optimal Preconditioned Eigenvalue Solvers (BLOPEX) in HyPre and PETSc*, SIAM J. Sci. Comput., 29 (2007), pp. 2224–2239. doi:10.1137/060661624
- [14] R. C. LI, *Sharpness in rates of convergence for the symmetric Lanczos method*, Math. Comp., 79 (2010), pp. 419–435.
- [15] R. C. LI AND L. H. ZHANG, *Convergence of the block Lanczos method for eigenvalue clusters*, Numer. Math., 131 (2015), pp. 83–113.
- [16] A. W. MARSHALL, I. OLKIN, AND B. C. ARNOLD, *Inequalities: Theory of Majorization and Its Applications*, Springer, New York, NY, 2011. ISBN: 978-0-387-40087-7
- [17] P. G. MASSEY, D. STOJANOFF, AND S. ZARATE, *Majorization bounds for Ritz values of self-adjoint matrices*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 554–572. doi:10.1137/19M1263996
- [18] P. G. MASSEY, D. STOJANOFF, AND S. ZARATE, *Absolute variation of Ritz values, principal angles and spectral spread*, Technical Report <https://arxiv.org/abs/2012.09018>, 2021.
- [19] K. NEYMEYR AND M. ZHOU, *Convergence analysis of restarted Krylov subspace eigensolvers*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 955–975. doi:10.1137/16M1056481
- [20] E. E. OVTCHINNIKOV, *Cluster robustness of preconditioned gradient subspace iteration eigensolvers*, Linear Algebra Appl., 415 (2006), pp. 140–166.
- [21] B. N. PARLETT, *The Symmetric Eigenvalue Problem*. SIAM, Philadelphia, PA, 1998. ISBN: 978-0898714029
- [22] H. RUTISHAUSER, *Computational aspects of F. L. Bauer’s simultaneous iteration method*, Numer. Math., 13 (1969), pp. 4–13.
- [23] Y. SAAD, *On the rates of convergence of the Lanczos and the block-Lanczos methods*, SIAM J. Numer. Anal., 17 (1980), pp. 687–706.
- [24] M. ZHOU AND K. NEYMEYR, *Sharp Ritz value estimates for restarted Krylov subspace iterations*, Electron. Trans. Numer. Anal., 46 (2017), pp. 424–446.
- [25] M. ZHOU, *Convergence estimates of nonrestarted and restarted block-Lanczos methods*, Numer. Linear Algebra Appl., 25 (2018), e2182.
- [26] M. ZHOU AND K. NEYMEYR, *Cluster robust estimates for block gradient-type eigensolvers*, Math. Comp., 88 (2019), pp. 2737–2765.
- [27] P. ZHU AND A. V. KNYAZEV, *Angles between subspaces and their tangents*, J. Numer. Math., 21 (2013), pp. 325–340.

UNIVERSITÄT ROSTOCK, INSTITUT FÜR MATHEMATIK, ULMENSTRASSE 69, 18055 ROSTOCK, GERMANY

E-mail address: ming.zhou at uni-rostock (dot) de

DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES, UNIVERSITY OF COLORADO DENVER, DENVER, CO 80217-3364, USA

E-mail address: merico.argentati at ucdenver (dot) edu

DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES, UNIVERSITY OF COLORADO DENVER, DENVER, CO 80217-3364, USA

E-mail address: andrew.knyazev at ucdenver (dot) edu

UNIVERSITÄT ROSTOCK, INSTITUT FÜR MATHEMATIK, ULMENSTRASSE 69, 18055 ROSTOCK, GERMANY

E-mail address: klaus.neymeyr at uni-rostock (dot) de