Electronic Transactions on Numerical Analysis.
Volume 46, pp. 505–523, 2017.

ISSN 1068–9613.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

A BLOCK LANCZOS METHOD FOR THE LINEAR RESPONSE EIGENVALUE PROBLEM*

ZHONGMING TENG† AND LEI-HONG ZHANG‡

Abstract. In the linear response eigenvalue problem arising from computational quantum chemistry and physics one needs to compute a small portion of eigenvalues around zero together with the associated eigenvectors. Lanczos-type methods are particularly suitable for such a task. However, single-vector Lanczos methods can only find one copy of any multiple eigenvalue and can be very slow when the desired eigenvalues form a cluster. In this paper, we propose a block Lanczos-type implementation for the linear response eigenvalue problem, which is able to compute a cluster of eigenvalues much faster and more efficiently than the single-vector version. Convergence results are established and reveal the accuracy of the approximations of eigenvalues in a cluster and of the eigenspace. A practical thick-restart procedure is introduced to alleviate the increasing numerical difficulties of the block Lanczos method in computational costs, memory demands, and numerical stability. Numerical examples are presented to support the effectiveness of the thick-restart block Lanczos method.

Key words. linear response eigenvalue problems, block Lanczos methods, convergence analysis, thick-restart

AMS subject classifications. 65F15, 15A18

1. Introduction.

In this paper, we consider the *Linear Response Eigenvalue Problem* (LREP)

$$Hz := \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} y \\ x \end{bmatrix} =: \lambda z, \tag{1.1}$$

where K and M are $N \times N$ real symmetric matrices and one of them is positive definite. LREP (1.1) is also known as the *Random Phase Approximation* (RPA) eigenvalue problem. Such a problem is one of the most widely used in computational quantum chemistry and physics for studying the excitation energy of many-particle systems [13, 14, 18] which have applications for silicon and other nanoscale materials. There has been a great deal of recent work and interest in developing efficient numerical algorithms and simulation techniques for excitation response calculations of molecules for materials design in energy science [3, 4, 8, 11, 15, 27, 28].

In LREP, usually both K and M are positive semidefinite or definite [1, 2, 14, 20, 24]. But there are also cases where one of them may be indefinite [12]. To put it in a relatively general setting, in this paper, we assume

$$K \text{ and } M \text{ are } n \times n \text{ real symmetric and } M \text{ is positive definite,} \tag{1.2}$$

unless explicitly stated otherwise.

From (1.1), we have $Kx = \lambda y$ and $My = \lambda x$, and they together lead to

$$KMy = \lambda^2 y, \quad MKx = \lambda^2 x.$$

*Received May 4, 2017. Accepted July 3, 2017. Published online on January 10, 2018. Recommended by Wen-Wei Lin. The work of the first author is supported in part by National Natural Science Foundation of China NSFC-11601081, NSFC-11601347, NSFC-11461046, Natural Science Foundation of Fujian province No. 2015J01580, and the research fund for distinguished young scholars of Fujian Agriculture and Forestry University. The work of the second author is supported by the National Natural Science Foundation of China NSFC-11371102, NSFC-11671246, NSFC-91730303, and the Basic Academic Discipline Program, the 11th five year plan of 211 Project for Shanghai University of Finance and Economics.

†College of Computer and Information Science, Fujian Agriculture and Forestry University, Fuzhou, 350002, P. R. China (peter979@163.com).

‡School of Mathematics and Research School for Interdisciplinary Sciences, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, China (longzlh@163.com).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Recall (1.2). Since $KM = KM^{1/2}M^{1/2}$ has the same eigenvalues as the symmetric matrix $M^{1/2}KM^{1/2}$, all eigenvalues of KM are real. Denote these eigenvalues by ω_i $(1 \le i \le N)$ in ascending order, i.e.,

$$\omega_1 \le \omega_2 \le \cdots \le \omega_N. \tag{1.3}$$

The eigenvalues of MK are ω_i $(1 \le i \le N)$, too. Let $\iota = \sqrt{-1}$ denote the imaginary unit and set

$$\lambda_i = \begin{cases} \sqrt{\omega_i}, & \text{if } \omega_i \ge 0, \\ \iota\sqrt{-\omega_i}, & \text{if } \omega_i < 0. \end{cases} \tag{1.4}$$

The eigenvalues of H are $\pm\lambda_i$ for $i = 1, 2, \ldots, N$. This practice of enumerating the eigenvalues of H is also used later for the much smaller projection of H.

In this paper, we attempt to develop an efficient block Lanczos method for LREP. A Lanczos-type method for LREP was first introduced by Tsiper in [22, 23]. Tsiper's Lanczos method is a recursive process to reduce both K and M to tridiagonal form, given initial vectors v_0 and u_0 with $v_0^{\mathrm{T}}u_0 \neq 0$. A corresponding convergence theory of Tsiper's Lanczos method for LREP has been established in [19]. Besides the convergence analysis of Tsiper's method, [19] introduces a better implementation of the Lanczos-type method, which is called the first Lanczos method for LREP. This version of the Lanczos method reduces K to a tridiagonal matrix and M to a diagonal matrix. It can be regarded as a natural extension of the classical Lanczos method for the symmetric eigenvalue problem. The associated convergence analysis shows that Tsiper's Lanczos method may need up to twice as many Lanczos steps as the first Lanczos method in order to reach the same accuracy.

It is well known that a single-vector Lanczos method can only find one copy of any multiple eigenvalue, and it can be very slow when the desired eigenvalues lie in a cluster; see, e.g., [10]. To compute all or some of the copies of a multiple eigenvalue, one prefers a block Lanczos method that is able to compute a cluster of eigenvalues much faster and more efficiently on modern computer architecture than a single-vector Lanczos method. This is particularly important for LREP because only the first small portion of the eigenvalues, i.e., λ_i in (1.4) for $i = 1, \ldots, k$ with $k \ll N$, are of interest. Thus, algorithms that are capable of computing efficiently eigenvalues in a cluster of the interesting part and even all or some copies of a multiple eigenvalue are particularly desirable. This was the motivation to develop a block implementation of the first Lanczos method of [19].

In order to reflect the above mentioned advantages of the block Lanczos method, we establish a convergence theory to bound the approximation error of an eigenvalue cluster as well as of the entire approximate eigenspace associated with the cluster.

With increasing dimension of the Krylov subspace, the simple version of a block Lanczos method usually suffers from numerical difficulties affecting computational costs, memory demands, and numerical stability. To alleviate these and to make it more practical, we incorporate a restarting procedure to our block Lanczos method for LREP. There are several types of restarting schemes for the classic Lanczos method for the symmetric eigenvalue problem, including the implicitly restart method [9, 16], the Krylov-Schur method [17], and the thick-restart method [25, 26]; by considering the special structure of LREP, the thick-restart method of [25, 26] turns out to be efficient and is used in this paper.

The rest of this paper is organized as follows. In Section 2 we collect some basic results for LREP and M-canonical angles for two subspaces that are used frequently in our later developments. In Section 3, a block Lanczos method for LREP and the associated convergence theorems are established. Section 4 is devoted to the thick-restart block Lanczos method for

ETNA
Kent State University and
Johann Radon Institute (RICAM)

LREP. We present some numerical examples in Section 5 to show the numerical behaviors of the thick-restart block Lanczos method for LREP. Finally, conclusions are drawn in Section 6.

Throughout this paper, $\mathbb{R}^{n\times m}$ is the set of all $n \times m$ real matrices, $\mathbb{R}^n = \mathbb{R}^{n\times 1}$, and $\mathbb{R} = \mathbb{R}^1$. I_n (or simply I if its dimension is clear from the context) is the $n \times n$ identity matrix, and $0_{n\times m}$ is an $n \times m$ matrix of zeros. The superscript "H" denotes conjugate transpose, while "T" denotes transpose only. $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix. For $X \in \mathbb{R}^{m\times n}$, $\mathrm{rank}(X)$ is the rank of X and $\mathcal{R}(X) = \mathrm{span}(X)$ represents the column space of X; the submatrices $X_{(k:\ell,:)}$ and $X_{(:,i:j)}$ of X consist of row k to row ℓ and column i to column j, respectively. For matrices or scalars X_i, $\mathrm{diag}(X_1, \ldots, X_k)$ denotes the block diagonal matrix

$$\begin{bmatrix} X_1 & & \\ & \ddots & \\ & & X_k \end{bmatrix}.$$

Given $A \in \mathbb{R}^{N\times N}$ and $B \in \mathbb{R}^{N\times n_b}$, the nth Krylov subspace of A on B is defined by

$$\mathcal{K}_n(A, B) \stackrel{\text{def}}{=} \mathrm{span}\{B, AB, \ldots, A^{n-1}B\}.$$

2. Preliminaries. Given a symmetric positive definite $M \in \mathbb{R}^{N\times N}$, the *$M$-inner product* and its induced M-norm are defined by

$$\langle x, y\rangle_{M} \stackrel{\text{def}}{=} y^{\mathrm{T}}Mx, \quad \|x\|_{M} = \sqrt{\langle x, x\rangle}_{M}.$$

If $\langle x, y\rangle_{M} = 0$, then we say $x \perp_{M} y$ or $y \perp_{M} x$. The projector Π_{M} is termed the M-orthogonal projector onto $\mathcal{X}$ if for any vector $x \in \mathbb{R}^N$,

$$\Pi_{M}\, x \in \mathcal{X} \quad \text{and} \quad (I - \Pi_{M})x \perp_{M} \mathcal{X}.$$

Consider two subspaces $\mathcal{X}$ and $\mathcal{Y}$ of $\mathbb{R}^N$, and suppose $k = \dim(\mathcal{X}) \le \dim(\mathcal{Y}) = \ell$. Let $X \in \mathbb{R}^{N\times k}$ and $Y \in \mathbb{R}^{N\times \ell}$ be M-orthogonal basis matrices of $\mathcal{X}$ and $\mathcal{Y}$, respectively, i.e.,

$$X^{\mathrm{T}}MX = I_k, \quad \mathcal{X} = \mathcal{R}(X) \qquad \text{and} \qquad Y^{\mathrm{T}}MY = I_\ell, \quad \mathcal{Y} = \mathcal{R}(Y).$$

Denote the singular values of $Y^{\mathrm{T}}MX$ by σ_j for $1 \le j \le k$ in ascending order, i.e., $\sigma_1 \le \cdots \le \sigma_k$. The k M-canonical angles $\theta^{(j)}_{M}(\mathcal{X}, \mathcal{Y})$ from[1] $\mathcal{X}$ to $\mathcal{Y}$ are defined by

$$0 \le \theta^{(j)}_{M}(\mathcal{X}, \mathcal{Y}) = \arccos\sigma_j \le \frac{\pi}{2} \quad \text{for } 1 \le j \le k.$$

Set

$$\Theta_{M}(\mathcal{X}, \mathcal{Y}) = \mathrm{diag}(\theta^{(1)}_{M}(\mathcal{X}, \mathcal{Y}), \ldots, \theta^{(k)}_{M}(\mathcal{X}, \mathcal{Y})),$$

where $\theta^{(1)}_{M}(\mathcal{X}, \mathcal{Y}) \ge \cdots \ge \theta^{(k)}_{M}(\mathcal{X}, \mathcal{Y})$. In particular, when $k = 1$, X reduces to a vector, and there is only one M-canonical angle from $\mathcal{X}$ to $\mathcal{Y}$. In what follows, we sometimes place a vector or matrix at one or both arguments of $\Theta_{M}(.,.)$ with the understanding that this refers to the subspace spanned by the vector or the columns of the matrix argument.

Later in this paper, we need the Chebyshev polynomials of the 1st kind defined as follows,

$$\mathscr{T}_n(\tau) = \begin{cases} \cos(n\arccos\tau) & \text{for } |\tau| \le 1, \\ \frac{1}{2}\left(\tau + \sqrt{\tau^2 - 1}\right)^n + \frac{1}{2}\left(\tau - \sqrt{\tau^2 - 1}\right)^n & \text{for } |\tau| \ge 1. \end{cases}$$

[1] If $\ell = k$, we may say that these angles are between $\mathcal{X}$ and $\mathcal{Y}$ [10].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

They frequently show up in numerical analysis and computations because of their numerous nice properties, for example, $|\mathscr{T}_n(\tau)| \leq 1$ for $|\tau| \leq 1$, and $|\mathscr{T}_n(\tau)|$ grows extremely fast for $|\tau| > 1$.

The following results are critical for our later developments. The reader is referred to [1, 10, 19] for proofs and details.

LEMMA 2.1 ([10, Proposition 2.1]). *Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^N$ with $k = \dim(\mathcal{X}) \leq \dim(\mathcal{Y}) = \ell$. For any $\widehat{\mathcal{Y}} \subseteq \mathcal{Y}$ with $\dim(\widehat{\mathcal{Y}}) = \dim(\mathcal{X}) = k$, we have $\theta_M^{(j)}(\mathcal{X}, \mathcal{Y}) \leq \theta_M^{(j)}(\mathcal{X}, \widehat{\mathcal{Y}})$ for $1 \leq j \leq k$.*

LEMMA 2.2 ([1, Theorem 2.3]). *The following statements hold for any symmetric matrices M, $K \in \mathbb{R}^{N\times N}$ with M being positive definite.*

(a) There exists a nonsingular $Y = [y_1, y_2, \ldots, y_N] \in \mathbb{R}^{N\times N}$ such that

$$K = Y\Lambda^2 Y^T, \quad M = XX^T, \quad \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N),$$

where $\lambda_1^2 \leq \lambda_2^2 \leq \cdots \leq \lambda_N^2$ and $X = Y^{-T} = [x_1, x_2, \ldots, x_N]$.

(b) The ith column of $Z = \left[\begin{smallmatrix} Y\Lambda \\ X \end{smallmatrix}\right]$ is the eigenvector corresponding to the eigenvalue λ_i of (1.1).

(c) Let (λ_i, z_i) $(i = 1, 2)$ be two eigenpairs of H, and partition $z_i = \left[s_i^{\mathrm{H}}, t_i^{\mathrm{H}}\right]^{\mathrm{H}}$. Then,

(i) if $\lambda_1 \neq \bar{\lambda}_2$, then $s_1^{\mathrm{H}} t_2 + s_2^{\mathrm{H}} t_1 = 0$.

(ii) if $\lambda_1 \neq \pm\lambda_2$, then $s_1^{\mathrm{H}} t_2 = s_2^{\mathrm{H}} t_1 = 0$.

LEMMA 2.3 ([19, Theorem 2.2]). *Given $0 \neq v_0 \in \mathbb{R}^N$ and $0 \neq u_0 \in \mathbb{R}^N$ such that $Mv_0 = u_0$. There exist nonsingular U, $V \in \mathbb{R}^{N\times N}$ such that $Ve_1 = \alpha v_0$ and $Ue_1 = \beta u_0$ for some $\alpha, \beta \in \mathbb{R}$, and*

$$U^T V = I_N, \quad U^T K U = T, \quad V^T M V = D,$$

where T is tridiagonal and D is diagonal.

3. Block Lanczos method for LREP.

3.1. Block Lanczos process for LREP. The first Lanczos process for LREP presented in [19] is a partial realization of the decomposition in Lemma 2.3. The block Lanczos process for LREP mentioned in this section is actually a block implementation of the first Lanczos process. We summarize its simple form in Algorithm 3.1, where no further treatment is given when $\operatorname{rank}(\widetilde{V}_{i+1}) < n_b$ happens in Line 8. Given $V_0 \in \mathbb{R}^{N\times n_b}$ with $\operatorname{rank}(V_0) = n_b$ and $U_0 = MV_0$ where $n_b \geq 1$ is the block size. According to [19], if $\operatorname{rank}(V_i) = n_b$ for $i = 1, 2, \ldots, n$, then we know that the recursively computed $P_n, Q_n \in \mathbb{R}^{N\times nn_b}$, the symmetric block-tridiagonal matrix $T_n \in \mathbb{R}^{nn_b\times nn_b}$, and the block-diagonal matrix $D_n \in \mathbb{R}^{nn_b\times nn_b}$ in Algorithm 3.1 satisfy

$$P_n^{\mathrm{T}} Q_n = I_{nn_b}, \quad KP_n = Q_n T_n + V_{n+1} B_n E_n^{\mathrm{T}}, \quad MQ_n = P_n D_n, \tag{3.1}$$

where

$$P_n = [U_1, U_2, \ldots, U_n], \qquad Q_n = [V_1, V_2, \ldots, V_n],$$

$$T_n = \begin{bmatrix} A_1 & B_1^{\mathrm{T}} & & \\ B_1 & A_2 & \ddots & \\ & \ddots & \ddots & B_{n-1}^{\mathrm{T}} \\ & & B_{n-1} & A_n \end{bmatrix}, \quad D_n = \operatorname{diag}(\Gamma_1, \Gamma_2, \ldots, \Gamma_n), \tag{3.2}$$

and $E_n^{\mathrm{T}} = \left[0_{n_b\times(n-1)\,n_b}, I_{n_b}\right]$. Here, D_n is symmetric positive definite because every diagonal block Γ_j for $1 \leq j \leq n$ is symmetric positive definite. It is noted that $\|V_{i(:,j)}\|_2 = 1$ for

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$j = 1, 2, \ldots, n_b$ in Algorithm 3.1 is enforced for all i. This is for convenience rather than necessity. It is possible to use $\|U_{i(:,j)}\|_2 = 1$ for $j = 1, 2, \ldots, n_b$ instead or enforce neither. When $n_b = 1$, Algorithm 3.1 reduces to the single-vector Lanczos process for LREP in [19]. Let

$$\mathcal{P}_n = \operatorname{span}(P_n), \quad \mathcal{Q}_n = \operatorname{span}(Q_n), \quad H_n = \begin{bmatrix} 0 & T_n \\ D_n & 0 \end{bmatrix}.$$

Then, the following lemma generalizes [19, Lemma 3.1].

LEMMA 3.1. *In* Algorithm 3.1*, if Γ_i is nonsingular for $i = 1, 2, \ldots, n$, then we have*

$$\mathcal{K}_n(KM, V_0) = \mathcal{Q}_n, \quad \mathcal{K}_n(MK, U_0) = \mathcal{P}_n.$$

Basically the block Lanczos method for LREP is this block Lanczos process followed by solving the small scale LREP for H_n to obtain approximate eigenpairs for H in (1.1). Let the eigenvalues $\pm\mu_j$ $(1 \le j \le nn_b)$ of H_n be enumerated in the same way as for H in (1.4) and let the corresponding eigenvectors be $\hat{z}_j$, i.e.,

$$H_n\hat{z}_j = \mu_j\hat{z}_j, \quad \hat{z}_j = \begin{bmatrix} \hat{y}_j \\ \hat{x}_j \end{bmatrix}, \quad \hat{x}_j,\ \hat{y}_j \in \mathbb{R}^{nn_b}. \tag{3.3}$$

Approximate eigenpairs of H, i.e., Ritz pairs, are then taken to be

$$(\mu_j, \tilde{z}_j), \quad \text{where} \quad \tilde{z}_j = \begin{bmatrix} Q_n\hat{y}_j \\ P_n\hat{x}_j \end{bmatrix}. \tag{3.4}$$

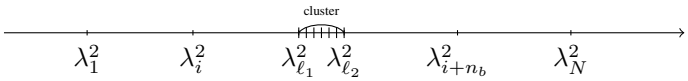
Algorithm 3.1 A block Lanczos process for LREP.

Input: Choose $U_0, V_0 \in \mathbb{R}^{N\times n_b}$ such that $\operatorname{rank}(V_0) = n_b$, $MV_0 = U_0$ and an integer $n \ge 1$.
Output: P_n, Q_n, T_n, D_n and V_{n+1} in (3.1) and (3.2).

1: Let $\beta_j = \|V_{0(:,j)}\|_2$ $(j = 1 : n_b)$, $B = \operatorname{diag}(\beta_1, \ldots, \beta_{n_b})$, $V_1 = V_0B^{-1}$, $W = V_0^{\mathrm{T}}U_0$, $U_1 = (U_0)W^{-1}B$, and $\Gamma_1 = B^{-1}WB^{-1}$.
2: $A_1 = U_1^{\mathrm{T}}KU_1$, $\widetilde{V}_2 = KU_1 - A_1V_1$.
3: If $\operatorname{rank}(\widetilde{V}_2) < n_b$, stop;
4: Let $\beta_j = \|\widetilde{V}_{2(:,j)}\|_2$ $(j = 1 : n_b)$, $B_1 = \operatorname{diag}(\beta_1, \ldots, \beta_{n_b})$, and $V_2 = \widetilde{V}_2B_1^{-1}$.
5: **for** $i = 2, 3, \ldots, n$ **do**
6: $\quad\Gamma_i = V_i^{\mathrm{T}}MV_i$, $U_i = MV_i\Gamma_i^{-1}$.
7: $\quad A_i = U_i^{\mathrm{T}}KU_i$, $\widetilde{V}_{i+1} = KU_i - V_iA_i - V_{i-1}B_{i-1}^{\mathrm{T}}$.
8: $\quad$If $\operatorname{rank}(\widetilde{V}_{i+1}) < n_b$, break;
9: $\quad$Let $\beta_j = \|\widetilde{V}_{i+1(:,j)}\|_2$ $(j = 1 : n_b)$, $B_i = \operatorname{diag}(\beta_1, \ldots, \beta_{n_b})$, and $V_{i+1} = \widetilde{V}_{i+1}B_i^{-1}$.
10: **end for**

3.2. Convergence analysis. Naturally we would use the first few μ_j as approximations to the first few λ_j. In this section we investigate how accurate such approximations could be. As we know, compared to a single-vector Lanczos method, a block Lanczos method with a block size that is not smaller than the multiplicity of an eigenvalue can be used to compute all copies of that eigenvalue. Therefore, motivated by [10], we directly analyze the convergence to a cluster of eigenvalues including multiple eigenvalues, and consider to bound the errors of the approximate eigenpairs belonging to the eigenvalue cluster together rather than separately for each individual eigenpair. Consider λ_i to $\lambda_{(i+n_b-1)}$ and their corresponding eigenvectors of LREP, in which λ_{ℓ_1} to λ_{ℓ_2} form a cluster as illustrated in the following figure.

ETNA
Kent State University and
Johann Radon Institute (RICAM)



Here, the cluster is described in terms of the squares of the eigenvalues since the eigenvalues of LREP come in pairs $\{-\lambda, \lambda\}$ and they may be purely imaginary numbers.

Note that by (3.3) and (3.4) we get

$$\begin{aligned} T_n\hat{x}_j &= \mu_j\hat{y}_j, \qquad & D_n\hat{y}_j &= \mu_j\hat{x}_j, \\ T_nD_n\hat{y}_j &= \mu_j^2\hat{y}_j, \qquad & D_nT_n\hat{x}_j &= \mu_j^2\hat{x}_j. \end{aligned}$$

Since D_n is symmetric positive definite, the eigenvalues of $D_n^{1/2}T_nD_n^{1/2}$ are μ_j^2 with the corresponding eigenvectors $D_n^{-1/2}\hat{x}_j$ for $1 \le j \le nn_b$, i.e.,

$$D_n^{1/2}T_nD_n^{1/2}\left(D_n^{-1/2}\hat{x}_j\right) = \mu_j^2\left(D_n^{-1/2}\hat{x}_j\right). \tag{3.5}$$

We first present some technical lemmas for our later developments. These lemmas are critical in our main theorem.

LEMMA 3.2 ([10, Proposition 2.4]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two subspaces in $\mathbb{R}^N$ with equal dimensions $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Suppose $\theta_M^{(I)}(\mathcal{X}, \mathcal{Y}) < \pi/2$. Then, for any set $\{y_1, y_2, \ldots, y_{k_1}\}$ of the basis vectors in $\mathcal{Y}$ where $1 \le k_1 \le k$, there is a set $\{x_1, x_2, \ldots, x_{k_1}\}$ of linearly independent vectors in $\mathcal{X}$ such that $\Pi_M x_j = y_j$ for $1 \le j \le k_1$, where Π_M is the M-orthogonal projector onto $\mathcal{Y}$.*

LEMMA 3.3. *Let $f \in \mathbb{P}_n$ where $\mathbb{P}_n$ is the collection of all polynomials of degree no higher than n and $\mathcal{V}_0 = \mathcal{R}(V_0)$. Then, for any $\hat{v} \in \mathcal{V}_0$, if $f(\mu_j^2) = 0$, then $f(KM)\hat{v} \perp_M Q_nD_n^{-1}\hat{x}_j$, where μ_j and $\hat{x}_j$ are defined by (3.3) and $1 \le j \le nn_b$.*

Proof. First, by assumptions, for any $\hat{v} \in \mathcal{V}_0$, we have $M\hat{v} = U_1c$ for some $c \in \mathbb{R}^{n_b}$. Now, for any integer $0 \le m \le n$, note from (3.1)

$$\begin{aligned} \hat{v}^{\mathrm{T}}(MK)^mMQ_nD_n^{-1}\hat{x}_j &= \hat{v}^{\mathrm{T}}M(KM)^{m-1}KMQ_nD_n^{-1}\hat{x}_j \\ &= c^{\mathrm{T}}U_1^{\mathrm{T}}(KM)^{m-1}(Q_nT_nD_n + V_{n+1}B_nE_n^{\mathrm{T}}D_n)D_n^{-1}\hat{x}_j \\ &= c^{\mathrm{T}}U_1^{\mathrm{T}}(KM)^{m-1}Q_n(T_nD_n)D_n^{-1}\hat{x}_j + c^{\mathrm{T}}U_1^{\mathrm{T}}(KM)^{m-1}V_{n+1}B_nE_n^{\mathrm{T}}\hat{x}_j \\ &= c^{\mathrm{T}}U_1^{\mathrm{T}}(KM)^{m-1}Q_n(T_nD_n)D_n^{-1}\hat{x}_j \qquad (\text{by } (MK)^{m-1}U_1 \in \mathcal{P}_n \perp V_{n+1}) \\ &= \cdots = c^{\mathrm{T}}U_1^{\mathrm{T}}Q_n(T_nD_n)^mD_n^{-1}\hat{x}_j = c^{\mathrm{T}}[I_{n_b}, 0, \ldots, 0](T_nD_n)^mD_n^{-1}\hat{x}_j; \end{aligned}$$

thereby, for any $f \in \mathbb{P}_n$,

$$\begin{aligned} \hat{v}^{\mathrm{T}}[f(KM)]^{\mathrm{T}}MQ_nD_n^{-1}\hat{x}_j &= c^{\mathrm{T}}[I_{n_b}, 0, \ldots, 0]f(T_nD_n)D_n^{-1}\hat{x}_j \\ &= f(\mu_j^2)c^{\mathrm{T}}[I_{n_b}, 0, \ldots, 0]D_n^{-1}\hat{x}_j = f(\mu_j^2)c^{\mathrm{T}}\Gamma_1^{-1}\hat{x}_{j,1}, \end{aligned}$$

where $\hat{x}_{j,1}$ is the sub-vector consisting of the first n_b components of $\hat{x}_j$. Thus, if $f(\mu_j^2) = 0$, $f(KM)\hat{v} \perp_M Q_nD_n^{-1}\hat{x}_j$. ∎

LEMMA 3.4 ([19, Lemma 3.5]). *We have, for $1 \le j \le nn_b$, $\lambda_j^2 \le \mu_j^2 \le \lambda_{N-nn_b+j}^2$.*

LEMMA 3.5 ([5, Problem III.6.15]). *For $N \times N$ Hermitian matrices A and B, we have, for any unitarily invariant norm,*

$$|||\Lambda(A) - \Lambda(B)||| \le |||A - B|||$$

where $\Lambda(A)$ is the diagonal matrix whose diagonal elements are the eigenvalues of A in descending order, i.e., $\Lambda(A) = \operatorname{diag}(\lambda_1(A), \ldots, \lambda_N(A))$ with $\lambda_1(A) \ge \cdots \ge \lambda_N(A)$, and $\Lambda(B)$ is defined similarly.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Partition Y and Λ^2 in Lemma 2.2 as

$$Y = \begin{bmatrix} \overset{i-1}{Y_1} & \overset{n_b}{Y_2} & \overset{N-n_b-i+1}{Y_3} \end{bmatrix}, \quad \Lambda^2 = \begin{matrix} i-1 \\ n_b \\ N-n_b-i+1 \end{matrix} \begin{bmatrix} \overset{i-1}{\Lambda_1^2} & \overset{n_b}{} & \overset{N-n_b-i+1}{} \\ & \Lambda_2^2 & \\ & & \Lambda_3^2 \end{bmatrix}.$$

Let $\widehat{\Lambda}_2^2 = \operatorname{diag}(\lambda_{\ell_1}^2, \ldots, \lambda_{\ell_2}^2)$, $\widehat{Y}_2 = Y_{(:,\ell_1:\ell_2)}$, $\widehat{Y}_2^{\perp}$ be the M-orthogonal complement of $\widehat{Y}_2$, and Π_M be the M-orthogonal projector onto $\mathcal{R}(Y_2)$. In the rest of this section, we always assume

$$\theta_M^{(1)}(V_0, Y_2) < \pi/2, \tag{3.6}$$

i.e., the matrix $V_0^{\mathrm{T}} M Y_2$ is nonsingular, which is generically true for a randomly chosen starting V_0. By Lemma 3.2, there exists $\Psi \in \mathbb{R}^{N \times (\ell_2 - \ell_1 + 1)}$ with $\mathcal{R}(\Psi) \subset \mathcal{R}(V_0)$ such that

$$\Pi_M \Psi = \widehat{Y}_2, \tag{3.7}$$

or equivalently, $Y_2 Y_2^{\mathrm{T}} M \Psi = \widehat{Y}_2$.

THEOREM 3.6. *Let Ψ satisfy (3.7). We have*

$$\begin{aligned} &\| \operatorname{diag}(\mu_{\ell_1}^2 - \lambda_{\ell_1}^2, \ldots, \mu_{\ell_2}^2 - \lambda_{\ell_2}^2) \|_F \\ &\qquad \le (\lambda_N^2 - \lambda_{\ell_1}^2) \times \left(\frac{\Delta_{\ell_1}}{\Xi_{\ell_1,\ell_2} \times \Upsilon_{n,\ell_1,\ell_2}} \right)^2 \times \| \tan^2 \Theta_M(\widehat{Y}_2, \Psi) \|_F, \end{aligned} \tag{3.8}$$

where

$$\delta_{\ell_2} = \frac{\lambda_{i+n_b}^2 - \lambda_{\ell_2}^2}{\lambda_N^2 - \lambda_{\ell_2}^2}, \qquad \Upsilon_{n,\ell_1,\ell_2} = \left| \mathscr{T}_{n-\ell_1} \left(\frac{\delta_{\ell_2} + 1}{\delta_{\ell_2} - 1} \right) \right|,$$

$$\Delta_{\ell_1} = \max_{i+n_b \le j \le N} \prod_{t=1}^{\ell_1 - 1} |\lambda_j^2 - \mu_t^2|, \quad \Xi_{\ell_1,\ell_2} = \min_{\ell_1 \le j \le \ell_2} \prod_{t=1}^{\ell_1 - 1} |\lambda_j^2 - \mu_t^2|.$$

In particular, if also $\lambda_{\ell_1} > \mu_{\ell_1 - 1}$, then

$$\frac{\Delta_{\ell_1}}{\Xi_{\ell_1,\ell_2}} = \left| \frac{(\lambda_N^2 - \mu_1^2) \cdots (\lambda_N^2 - \mu_{\ell_1-1}^2)}{(\lambda_{\ell_1}^2 - \mu_1^2) \cdots (\lambda_{\ell_1}^2 - \mu_{\ell_1-1}^2)} \right|.$$

Proof. Let $n_c = \ell_2 - \ell_1 + 1$ and $n_r = N - n_b - i + 1$ for convenience. It follows from (3.7) that

$$\begin{aligned} \Psi = Y X^{\mathrm{T}} \Psi = Y Y^{\mathrm{T}} M \Psi &= [Y_1, Y_2, Y_3] \begin{bmatrix} Y_1^{\mathrm{T}} \\ Y_2^{\mathrm{T}} \\ Y_3^{\mathrm{T}} \end{bmatrix} M \Psi \\ &= Y_1 Y_1^{\mathrm{T}} M \Psi + Y_2 Y_2^{\mathrm{T}} M \Psi + Y_3 Y_3^{\mathrm{T}} M \Psi = Y_1 Y_1^{\mathrm{T}} M \Psi + \widehat{Y}_2 \widehat{Y}_2^{\mathrm{T}} M \Psi + Y_3 Y_3^{\mathrm{T}} M \Psi. \end{aligned}$$

The last equality holds because of (3.7) and $\widehat{Y}_2^{\mathrm{T}} M \Psi = I_{n_c}$. Let $\Psi_0 = \Psi (\Psi^{\mathrm{T}} M \Psi)^{-1/2}$, $f \in \mathbb{P}_{n-1}$ with $f(\mu_j^2) = 0$ for $1 \le j \le \ell_1 - 1$, and

$$Z = f(KM)\Psi_0 = Y_1 f(\Lambda_1^2) Y_1^{\mathrm{T}} M \Psi_0 + \widehat{Y}_2 f(\widehat{\Lambda}_2^2) \widehat{Y}_2^{\mathrm{T}} M \Psi_0 + Y_3 f(\Lambda_3^2) Y_3^{\mathrm{T}} M \Psi_0.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Note that $\widehat{Y}_2^{\mathrm{T}} M Z = f(\widehat{\Lambda}_2^2)\widehat{Y}_2^{\mathrm{T}} M \Psi_0$ and $\widehat{Y}_2^{\mathrm{T}} M \Psi_0$ is nonsingular; if $f(\widehat{\Lambda}_2^2)$ is also nonsingular (which is true for the one we choose later), we have

$$Z_0 = Z\left(\widehat{Y}_2^{\mathrm{T}} M Z\right)^{-1} = Z\left(\widehat{Y}_2^{\mathrm{T}} M \Psi_0\right)^{-1}\left[f(\widehat{\Lambda}_2^2)\right]^{-1} = Y_1 R_1 + \widehat{Y}_2 + Y_3 R_3$$

and $\operatorname{rank}(Z_0) = n_c$, where

$$R_j = f(\Lambda_j^2) Y_j^{\mathrm{T}} M \Psi_0 \left(\widehat{Y}_2^{\mathrm{T}} M \Psi_0\right)^{-1}\left[f(\widehat{\Lambda}_2^2)\right]^{-1}, \text{ for } j = 1, 3.$$

Let $E = Z_0 (Z_0^{\mathrm{T}} M Z_0)^{-1/2}$ and $F = ME$. Then, E has full column rank and $\mathcal{R}(E) \subset \mathcal{Q}_n$ by Lemma 3.1. Write $E = Q_n \widehat{E}$ with $\widehat{E} \in \mathbb{R}^{nn_b \times n_c}$. It is true that $\widehat{E}^{\mathrm{T}} D_n \widehat{E} = I_{n_c}$ by $E^{\mathrm{T}} M E = I_{n_c}$. Denote the eigenvalues of $F^{\mathrm{T}} K F$ by γ_j^2 where $1 \le j \le n_c$. Then, by (3.1),

$$\begin{aligned} F^{\mathrm{T}} K F &= E^{\mathrm{T}} M K M E = \widehat{E}^{\mathrm{T}} Q_n^{\mathrm{T}} M K M Q_n \widehat{E} \\ &= \widehat{E}^{\mathrm{T}} D_n P_n^{\mathrm{T}} K P_n D_n \widehat{E} = \widehat{E}^{\mathrm{T}} D_n^{1/2} (D_n^{1/2} T_n D_n^{1/2}) D_n^{1/2} \widehat{E}. \end{aligned}$$

Because $f(\mu_j^2) = 0$ for $1 \le j \le \ell_1 - 1$, according to Lemma 3.3, we know

$$E \perp_M Q_n D_n^{-1} \hat{x}_j \quad \text{for} \quad 1 \le j \le \ell_1 - 1.$$

That means

$$\hat{x}_j^{\mathrm{T}} D_n^{-1} Q_n^{\mathrm{T}} M Q_n \widehat{E} = \hat{x}_j^{\mathrm{T}} D_n^{-1} D_n \widehat{E} = \hat{x}_j^{\mathrm{T}} \widehat{E} = 0,$$

i.e., $(D_n^{1/2}\widehat{E})^{\mathrm{T}} D_n^{-1/2} \hat{x}_j = 0$ for $0 \le j \le \ell_1 - 1$. It is noted from (3.5) that $D_n^{-1/2}\hat{x}_j$ is the eigenvector of $D_n^{1/2} T_n D_n^{1/2}$ associated to the eigenvalue μ_j^2. Hence, by Cauchy's interlacing inequality,

$$\mu_{\ell_1+j-1}^2 \le \gamma_j^2 \quad \text{for} \quad 1 \le j \le n_c. \tag{3.9}$$

For any vector $g \in \mathbb{R}^{n_c}$, let $\hat{g} = (Z_0^{\mathrm{T}} M Z_0)^{-1/2} g$, and consider the Rayleigh quotient of $F^{\mathrm{T}} K F - \lambda_{\ell_1}^2 I$,

$$\begin{aligned} \frac{g^{\mathrm{T}} F^{\mathrm{T}} K F g - \lambda_{\ell_1} g^{\mathrm{T}} g}{g^{\mathrm{T}} g} &= \frac{\hat{g}^{\mathrm{T}} Z_0^{\mathrm{T}} (MKM - \lambda_{\ell_1} M) Z_0 \hat{g}}{\hat{g}^{\mathrm{T}} Z_0^{\mathrm{T}} M Z_0 \hat{g}} \\ &= \frac{\hat{g}^{\mathrm{T}} \left[R_1^{\mathrm{T}} (\Lambda_1^2 - \lambda_{\ell_1}^2 I_{i-1}) R_1 + (\widehat{\Lambda}_2^2 - \lambda_{\ell_1}^2 I_{n_c}) + R_3^{\mathrm{T}} (\Lambda_3^2 - \lambda_{\ell_1}^2 I_{n_r}) R_3 \right] \hat{g}}{\hat{g}^{\mathrm{T}} \left[R_1^{\mathrm{T}} R_1 + I_{n_c} + R_3^{\mathrm{T}} R_3 \right] \hat{g}} \\ &\le \frac{\hat{g}^{\mathrm{T}} \left[(\widehat{\Lambda}_2^2 - \lambda_{\ell_1}^2 I_{n_c}) + R_3^{\mathrm{T}} (\Lambda_3^2 - \lambda_{\ell_1}^2 I_{n_r}) R_3 \right] \hat{g}}{\hat{g}^{\mathrm{T}} \hat{g}}. \end{aligned} \tag{3.10}$$

The last inequality in (3.10) holds because

$$\hat{g}^{\mathrm{T}} R_1^{\mathrm{T}} (\Lambda_1^2 - \lambda_{\ell_1}^2 I_{i-1}) R_1 \hat{g} \le 0 \quad \text{and} \quad \hat{g}^{\mathrm{T}} (R_1^{\mathrm{T}} R_1 + R_3^{\mathrm{T}} R_3) \hat{g} \ge 0.$$

Denote by $\hat{\gamma}_j^2$ for $1 \le j \le n_c$ the eigenvalues of $(\widehat{\Lambda}_2^2 - \lambda_{\ell_1}^2 I_{n_c}) + R_3^{\mathrm{T}} (\Lambda_3^2 - \lambda_{\ell_1}^2 I_{n_r}) R_3$. We have $\gamma_j^2 - \lambda_{\ell_1} \le \hat{\gamma}_j^2$ for $1 \le j \le n_c$ by (3.10). Then, it follows from (3.9), Lemma 3.4, and

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Lemma 3.5 that

$$\begin{aligned}\|\operatorname{diag}(\mu_{\ell_1}^2-\lambda_{\ell_1}^2,\dots,\mu_{\ell_2}^2-\lambda_{\ell_2}^2)\|_{\mathrm{F}} &\le \|\operatorname{diag}(\gamma_1^2-\lambda_{\ell_1}^2,\dots,\gamma_{n_c}^2-\lambda_{\ell_2}^2)\|_{\mathrm{F}}\\ &= \|\operatorname{diag}\left((\gamma_1^2-\lambda_{\ell_1}^2)-(\lambda_{\ell_1}^2-\lambda_{\ell_1}^2),\dots,(\gamma_{n_c}^2-\lambda_{\ell_1}^2)-(\lambda_{\ell_2}^2-\lambda_{\ell_1}^2)\right)\|_{\mathrm{F}}\\ &\le \|\operatorname{diag}\left(\hat\gamma_1^2-(\lambda_{\ell_1}^2-\lambda_{\ell_1}^2),\dots,\hat\gamma_{n_c}^2-(\lambda_{\ell_2}^2-\lambda_{\ell_1}^2)\right)\|_{\mathrm{F}}\\ &\le \|R_3^{\mathrm{T}}(\Lambda_3^2-\lambda_{\ell_1}^2 I_{n_r})R_3\|_{\mathrm{F}} \quad \text{(by Lemma 3.5)}\\ &\le (\lambda_N^2-\lambda_{\ell_1}^2)\|R_3^{\mathrm{T}}R_3\|_{\mathrm{F}}.\end{aligned}$$

Since

$$\|Y_3^{\mathrm{T}}M\Psi_0(\widehat{Y}_2^{\mathrm{T}}M\Psi_0)^{-1}\|_{\mathrm{F}} \le \left\|\begin{bmatrix}Y_1^{\mathrm{T}}M\Psi_0(\widehat{Y}_2^{\mathrm{T}}M\Psi_0)^{-1}\\ Y_3^{\mathrm{T}}M\Psi_0(\widehat{Y}_2^{\mathrm{T}}M\Psi_0)^{-1}\end{bmatrix}\right\|_{\mathrm{F}} = \|\tan\Theta_M(\widehat{Y}_2,\Psi_0)\|_{\mathrm{F}},$$

we have

$$\begin{aligned}(3.11)\quad (\lambda_N^2-\lambda_{\ell_1}^2)\|R_3^{\mathrm{T}}R_3\|_{\mathrm{F}} &= (\lambda_N^2-\lambda_{\ell_1}^2)\\ \times\left\|\left[f(\widehat{\Lambda}_2^2)\right]^{-1}\left[Y_3^{\mathrm{T}}M\Psi_0(\widehat{Y}_2^{\mathrm{T}}M\Psi_0)^{-1}\right]^{\mathrm{T}}\left[f(\Lambda_3^2)\right]^2\times Y_3^{\mathrm{T}}M\Psi_0(\widehat{Y}_2^{\mathrm{T}}M\Psi_0)^{-1}\left[f(\widehat{\Lambda}_2^2)\right]^{-1}\right\|_F&\\ \le (\lambda_N^2-\lambda_{\ell_1}^2)\max_{i+n_b\le j\le N}\left[f(\lambda_j^2)\right]^2\times\max_{\ell_1\le j\le\ell_2}\frac{1}{\left[f(\lambda_j^2)\right]^2}\times\|\tan^2\Theta_M(\widehat{Y}_2,\Psi_0)\|_{\mathrm{F}}.&\end{aligned}$$

Take

$$f(t)=(t-\mu_1^2)\cdots(t-\mu_{\ell_1-1}^2)\times\mathscr{T}_{n-\ell_1}(\tau)\in\mathbb{P}_{n-1},\quad \text{where } \tau=\frac{2t-(\lambda_{i+n_b}^2+\lambda_N^2)}{\lambda_N^2-\lambda_{i+n_b}^2}.$$

Note that $f(\widehat{\Lambda}_2^2)$ is nonsingular, otherwise some of the exact eigenvalues have been found. Then $-1\le\tau\le 1$ for $\lambda_{i+n_b}^2\le t\le\lambda_N^2$, and

$$\tau|_{t=\lambda_{\ell_2}^2}=\frac{2\lambda_{\ell_2}^2-(\lambda_{i+n_b}^2+\lambda_N^2)}{\lambda_N^2-\lambda_{i+n_b}^2}=\frac{\delta_{\ell_2}+1}{\delta_{\ell_2}-1},$$

where

$$\delta_{\ell_2}=\frac{\lambda_{i+n_b}^2-\lambda_{\ell_2}^2}{\lambda_N^2-\lambda_{\ell_2}^2}.$$

Therefore we have

$$(3.12)\qquad \min_{\ell_1\le j\le\ell_2}|f(\lambda_j^2)|\ge\Upsilon_{n,\ell_1,\ell_2}\times\min_{\ell_1\le j\le\ell_2}\prod_{t=1}^{\ell_1-1}|\lambda_j^2-\mu_t^2|,$$

$$(3.13)\qquad |f(\lambda_j^2)|\le\max_{i+n_b\le j\le N}\prod_{t=1}^{\ell_1-1}|\lambda_j^2-\mu_t^2|,\quad \text{for } i+n_b\le j\le N.$$

Inequality (3.8) is now a consequence of (3.11), (3.12), and (3.13). □

Theorem 3.6 is particularly useful to bound the approximate eigenvalue errors when a cluster of eigenvalues, including the case of multiple eigenvalues, occurs. Nevertheless, it is also applicable to the simple eigenvalue case. Specifically, set $\ell_2=\ell_1=i$ to obtain the following corollary.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

COROLLARY 3.7. *For* $1 \le i \le nn_b$, *there exists a vector* $\psi_i \in \mathcal{R}(V_0)$ *such that* $\Pi_M \psi_i = y_i$, *and*

$$0 \le \mu_i^2 - \lambda_i^2 \le (\lambda_N^2 - \lambda_i^2) \times \widetilde{\Delta}_i^2 \times \Upsilon_{n,i}^{-2} \times (\tan \Theta_M(y_i, \psi_i))^2,$$

where

$$\delta_i = \frac{\lambda_{i+n_b}^2 - \lambda_i^2}{\lambda_N^2 - \lambda_i^2}, \quad \Upsilon_{n,i} = \left| \mathscr{T}_{n-i}\left(\frac{\delta_i + 1}{\delta_i - 1}\right)\right|, \quad \textit{and} \quad \widetilde{\Delta}_i = \max_{i+n_b \le j \le N} \prod_{t=1}^{i-1} \left| \frac{\lambda_j^2 - \mu_t^2}{\lambda_i^2 - \mu_t^2}\right|.$$

Next we treat the eigenspace approximations. Lemma 2.2(b) says that $z_j = \left[\lambda_j y_j^{\mathrm{T}}, x_j^{\mathrm{T}}\right]^{\mathrm{T}}$ is the eigenvector of H associated with its eigenvalue λ_j. That means the eigenspaces consist of two components which are spanned by the columns of Y and X, respectively. Similarly, the approximate eigenspaces are composed by $\mathcal{Q}_n$ and $\mathcal{P}_n$. Thus, we are interested in bounding the angles from $\mathcal{R}\left(Y_{(:,\ell_1:\ell_2)}\right)$ to $\mathcal{Q}_n = \mathcal{K}_n(KM, V_0)$ and the angles from $\mathcal{R}\left(X_{(:,\ell_1:\ell_2)}\right)$ to $\mathcal{P}_n = \mathcal{K}_n(MK, U_0)$. This is established in the following theorem.

THEOREM 3.8. *Let* Ψ *be defined in* (3.7) *and* $\Phi = M\Psi$. *We have*

$$\|\tan \Theta_M(Y_{(:,\ell_1:\ell_2)}, Q_n)\|_F \le \prod_{j=1}^{i-1} \frac{\lambda_N^2 - \lambda_j^2}{\lambda_{\ell_1}^2 - \lambda_j^2} \times \Upsilon_{n,i,\ell_2}^{-1} \times \|\tan \Theta_M(Y_{(:,\ell_1:\ell_2)}, \Psi)\|_F, \tag{3.14}$$

$$\begin{aligned}\|\tan \Theta_{M^{-1}}(X_{(:,\ell_1:\ell_2)}, P_n)\|_F \\ \le \prod_{j=1}^{i-1} \frac{\lambda_N^2 - \lambda_j^2}{\lambda_{\ell_1}^2 - \lambda_j^2} \times \Upsilon_{n,i,\ell_2}^{-1} \times \|\tan \Theta_{M^{-1}}(X_{(:,\ell_1:\ell_2)}, \Phi)\|_F,\end{aligned} \tag{3.15}$$

where

$$\delta_{\ell_2} = \frac{\lambda_{i+n_b}^2 - \lambda_{\ell_2}^2}{\lambda_N^2 - \lambda_{\ell_2}^2} \quad \textit{and} \quad \Upsilon_{n,i,\ell_2} = \left| \mathscr{T}_{n-i}\left(\frac{\delta_{\ell_2} + 1}{\delta_{\ell_2} - 1}\right)\right|.$$

Proof. Take

$$f(t) = (t - \lambda_1^2) \cdots (t - \lambda_{i-1}^2) \times \mathscr{T}_{n-i}(\tau), \quad \text{where } \tau = \frac{2t - (\lambda_{i+n_b}^2 + \lambda_N^2)}{\lambda_N^2 - \lambda_{i+n_b}^2}.$$

It follows that

$$|f(\lambda_j^2)| = 0 \quad \text{for} \quad 1 \le j \le i-1, \tag{3.16}$$

$$\min_{\ell_1 \le j \le \ell_2} |f(\lambda_j^2)| \ge \Upsilon_{n,i,\ell_2} \times \prod_{t=1}^{i-1} |\lambda_{\ell_1}^2 - \lambda_t^2|, \tag{3.17}$$

$$\max_{i+n_b \le j \le N} |f(\lambda_j^2)| \le \prod_{t=1}^{i-1} |\lambda_N^2 - \lambda_t^2|, \tag{3.18}$$

and $f(\widehat{\Lambda}_2^2)$ is nonsingular. Recall the proof of Theorem 3.6. Let $Z = f(KM)\Psi_0$ where $\Psi_0 = \Psi(\Psi^{\mathrm{T}} M \Psi)^{-1/2}$. Then,

$$Z = f(KM)\Psi_0 = Y_1 f(\Lambda_1^2) Y_1^{\mathrm{T}} M \Psi_0 + \widehat{Y}_2 f(\widehat{\Lambda}_2^2) \widehat{Y}_2^{\mathrm{T}} M \Psi_0 + Y_3 f(\Lambda_3^2) Y_3^{\mathrm{T}} M \Psi_0.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We have, by (3.16), (3.17), (3.18), and Lemma 2.1,

$$\begin{aligned}
\|\tan\Theta_M(Y_{(:,\ell_1:\ell_2)},Q_n)\|_{\mathrm{F}} &\le \|\tan\Theta_M(\widehat{Y}_2,Z)\|_{\mathrm{F}}\\
&= \left\|(\widehat{Y}_2^{\perp})^{\mathrm{T}}MZ(Z^{\mathrm{T}}MZ)^{-1/2}\left[\widehat{Y}_2^{\mathrm{T}}MZ(Z^{\mathrm{T}}MZ)^{-1/2}\right]^{-1}\right\|_{\mathrm{F}}\\
&= \left\|(\widehat{Y}_2^{\perp})^{\mathrm{T}}MZ\left(\widehat{Y}_2^{\mathrm{T}}MZ\right)^{-1}\right\|_{\mathrm{F}} = \left\|\begin{bmatrix}Y_1^{\mathrm{T}}MZ\\ Y_3^{\mathrm{T}}MZ\end{bmatrix}\left(\widehat{Y}_2^{\mathrm{T}}MZ\right)^{-1}\right\|_{\mathrm{F}}\\
&= \left\|\begin{bmatrix}f(\Lambda_1^2)Y_1^{\mathrm{T}}M\Psi_0\\ f(\Lambda_3^2)Y_3^{\mathrm{T}}M\Psi_0\end{bmatrix}\left(\widehat{Y}_2^{\mathrm{T}}M\Psi_0\right)^{-1}\left[f(\widehat{\Lambda}_2^2)\right]^{-1}\right\|_{\mathrm{F}}\\
&= \left\|\begin{bmatrix}f(\Lambda_1^2) & \\ & f(\Lambda_3^2)\end{bmatrix}\begin{bmatrix}Y_1^{\mathrm{T}}M\Psi_0\left(\widehat{Y}_2^{\mathrm{T}}M\Psi_0\right)^{-1}\\ Y_3^{\mathrm{T}}M\Psi_0\left(\widehat{Y}_2^{\mathrm{T}}M\Psi_0\right)^{-1}\end{bmatrix}\left[f(\widehat{\Lambda}_2^2)\right]^{-1}\right\|_{\mathrm{F}}\\
&\le \max_{\substack{1\le j\le i-1\\ i+n_b\le j\le N}} f(\lambda_j^2)\times\max_{\ell_1\le j\le\ell_2}\frac{1}{f(\lambda_j^2)}\times\|\tan\Theta_M(\widehat{Y}_2,\Psi_0)\|_{\mathrm{F}}.\\
&\le \prod_{j=1}^{i-1}\frac{\lambda_N^2-\lambda_j^2}{\lambda_{\ell_1}^2-\lambda_j^2}\times\Upsilon_{n,i,\ell_2}^{-1}\times\|\tan\Theta_M(\widehat{Y}_2,\Psi)\|_{\mathrm{F}},
\end{aligned}$$

which gives (3.14). Similarly we can prove (3.15). ∎

Similarly to Corollary 3.7, the following corollary bounds the eigenvector approximations in the case of simple eigenvalues.

COROLLARY 3.9. *Use the notation of Corollary 3.7 and let* $\phi_i = M\psi_i$. *We have, for* $1 \le i \le nn_b$,

$$\tan\Theta_M(y_i,\mathcal{Q}_n) \le \prod_{j=1}^{i-1}\frac{\lambda_N^2-\lambda_j^2}{\lambda_i^2-\lambda_j^2}\times\Upsilon_{n,i}^{-1}\times\tan\Theta_M(y_i,\psi_i),$$

$$\tan\Theta_{M^{-1}}(x_i,\mathcal{P}_n) \le \prod_{j=1}^{i-1}\frac{\lambda_N^2-\lambda_j^2}{\lambda_i^2-\lambda_j^2}\times\Upsilon_{n,i}^{-1}\times\tan\Theta_{M^{-1}}(x_i,\phi_i).$$

REMARK 3.10. Listed below are some comments for Theorems 3.6 and 3.8.

(a) Similarly to Theorem 3.6 and 3.8, by slight modifications of the above proofs, we can obtain a bound for

$$\|\operatorname{diag}(\lambda_{N-nn_b+\ell_1}^2-\mu_{\ell_1}^2,\dots,\lambda_{N-nn_b+\ell_2}^2-\mu_{\ell_2}^2)\|_{\mathrm{F}}$$

and bounds for the associated eigenspace approximations, i.e., bounds for the M-canonical angles from $\mathcal{R}(Y_{(:,N-nn_b+\ell_1:N-nn_b+\ell_2)})$ to $\mathcal{Q}_n$ and the M^{-1}-canonical angles from $\mathcal{R}(X_{(:,N-nn_b+\ell_1:N-nn_b+\ell_2)})$ to $\mathcal{P}_n$, respectively.

(b) Although we use the Frobenius norm in Theorems 3.6 and 3.8 to measure the accuracy of eigenvalues and eigenspace approximations, the arguments in the proofs work for any unitarily invariant norm (see [10] for some properties of the unitarily invariant norm).

(c) Compared to the single-vector version of the first Lanczos method in [19], our convergence results have already reflected, to some extent, the advantages of this block Lanczos version. For example, it is shown in Corollary 3.7 that the bound for the convergence speed of the block version for the approximate eigenvalue μ_i is proportional to $\left|\mathscr{T}_{n-i}\left(\frac{\delta_i+1}{\delta_i-1}\right)\right|^{-2}$ where $\delta_i = \frac{\lambda_{i+n_b}^2-\lambda_i^2}{\lambda_N^2-\lambda_i^2}$, which is better than $\left|\mathscr{T}_{n-i}\left(\frac{\tilde{\delta}_i+1}{\tilde{\delta}_i-1}\right)\right|^{-2}$ with $\tilde{\delta}_i = \frac{\lambda_{i+1}^2-\lambda_i^2}{\lambda_N^2-\lambda_i^2}$ established in the single-vector Lanczos method for LREP [19]; although each Lanczos

ETNA
Kent State University and
Johann Radon Institute (RICAM)

step in the block version requires more computational work than in the single-vector one, we argue that the improvements induced in the factor δ_i can pay for the additional computational work, especially when the desired eigenpairs are from a well-separated cluster. The same comment is still valid for the bound on the convergence speed of the eigenvector approximations in Corollary 3.9.

4. Restart. Recalling (3.1) and (3.2), we know that the quantities computed by the block Lanczos method (Algorithm 3.1) satisfy the following relationship for LREP,

$$\begin{cases} KP_n = Q_n T_n + V_{n+1} B_n E_n^{\mathrm{T}}, \\ MQ_n = P_n D_n, \end{cases} \tag{4.1}$$

where $P_n^{\mathrm{T}} Q_n = I_{nn_b}$. A problem with Algorithm 3.1 is that, as the iteration proceeds, computational and memory costs increase rapidly, and numerical stability deteriorates gradually. This is also true for the classical Lanczos method for the standard symmetric eigenvalue problem. To resolve these issues, a restarting strategy usually turns out to be an efficient remedy. Several restarting schemes (e.g., [9, 17, 25, 26]) have been proposed in the literature. For our case, the LREP, the thick-restart technique [25, 26] appears to be an effective one, and we describe the detailed procedure in this section.

Note that T_n and D_n are symmetric and D_n is positive definite. By Lemma 2.2(a), there exist nonsingular matrices $S, R \in \mathbb{R}^{nn_b \times nn_b}$ with $S = R^{-T}$ such that

$$T_n = S\Omega_n^2 S^{\mathrm{T}} \quad \text{and} \quad D_n = RR^{\mathrm{T}}, \tag{4.2}$$

where $\Omega_n^2 = \operatorname{diag}(\mu_1^2, \ldots, \mu_{nn_b}^2)$ and $\mu_1^2 \le \cdots \le \mu_{nn_b}^2$. Let $S = [s_1, s_2, \ldots, s_{nn_b}]$ and $R = [r_1, r_2, \ldots, r_{nn_b}]$. To save the costs of forming larger subspaces in the block Lanczos process and to reduce the costs in the Ritz procedure for solving the resulting LREP for larger nn_b, the iteration is restarted after the basis vector V_{n+1} has been computed. Since the eigenvalues of interest lie in the left part of the spectrum (1.3), the eigen-information of the wanted Ritz values (appearing in the top-left of Ω_n^2) and Ritz vectors (the corresponding ones appearing to the left of S and R) should be maintained as much as possible. Suppose $k \times n_b$ is the number of Ritz values to be kept in the top-left of Ω_n^2. Let S_k and R_k be the submatrices consisting of the first kn_b columns of S and R, respectively, i.e.,

$$S_k = [s_1, s_2, \ldots, s_{kn_b}] \quad \text{and} \quad R_k = [r_1, r_2, \ldots, r_{kn_b}].$$

Then, by (4.2), it follows that

$$T_n R_k = S_k \Omega_k^2 \quad \text{and} \quad D_n S_k = R_k I_{kn_b}, \tag{4.3}$$

where $\Omega_k^2 = \operatorname{diag}(\mu_1^2, \ldots, \mu_{kn_b}^2)$.

For the thick-restart technique [25, 26], post-multiply by R_k and S_k in both equations (4.1), respectively, gives

$$\begin{cases} KP_n R_k = Q_n T_n R_k + V_{n+1} B_n E_n^{\mathrm{T}} R_k, \\ MQ_n S_k = P_n D_n S_k. \end{cases} \tag{4.4}$$

By (4.3), we can rewrite (4.4) as

$$\begin{cases} KP_n R_k = Q_n S_k \Omega_k^2 + V_{n+1} B_n E_n^{\mathrm{T}} R_k, \\ MQ_n S_k = P_n R_k. \end{cases} \tag{4.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Let

$$\widehat{P}_k = P_n R_k,\ \widehat{Q}_k = Q_n S_k,\ \widehat{D}_k = I_{kn_b},\ W = R_k^{\mathrm{T}} E_n,\ \widehat{V}_{k+1} = V_{n+1},\ \widehat{B}_k = B_n,\ \widehat{T}_k = \Omega_k^2.$$

Then, (4.5) can be expressed as

$$\begin{cases} K\widehat{P}_k = \widehat{Q}_k \widehat{T}_k + \widehat{V}_{k+1}\widehat{B}_k W^{\mathrm{T}}, \\ M\widehat{Q}_k = \widehat{P}_k \widehat{D}_k, \end{cases} \tag{4.6}$$

and $\widehat{P}_k^{\mathrm{T}}\widehat{Q}_k = R_k^{\mathrm{T}} P_n^{\mathrm{T}} Q_n S_k = I_{kn_b}$.

The restarting begins with $\widehat{P}_k$ and $\widehat{Q}_k$ as the first kn_b basis vectors, and $\widehat{V}_{k+1}$ as the $(k+1)$st block. To compute $\widehat{U}_{k+1}$, according to the block Lanczos process in Algorithm 3.1, we compute

$$\widehat{\Gamma}_{k+1} = \widehat{V}_{k+1}^{\mathrm{T}} M \widehat{V}_{k+1} \quad \text{and} \quad \widehat{U}_{k+1} = M\widehat{V}_{k+1}\widehat{\Gamma}_{k+1}^{-1};$$

thus, $\widehat{P}_k$ and $\widehat{Q}_k$ are expanded to

$$\widehat{P}_{k+1} = [\widehat{P}_k, \widehat{U}_{k+1}] \quad \text{and} \quad \widehat{Q}_{k+1} = [\widehat{Q}_k, \widehat{V}_{k+1}],$$

respectively, which satisfy $\widehat{P}_{k+1}^{\mathrm{T}}\widehat{Q}_{k+1} = I_{(k+1)n_b}$. For the next $\widehat{V}_{k+2}$, we first compute

$$\begin{aligned} \widetilde{V}_{k+2} &= K\widehat{U}_{k+1} - \widehat{V}_{k+1}\widehat{U}_{k+1}^{\mathrm{T}} K \widehat{U}_{k+1} - \widehat{Q}_k \widehat{P}_k^{\mathrm{T}} K \widehat{U}_{k+1} \\ &= K\widehat{U}_{k+1} - \widehat{V}_{k+1}\widehat{A}_{k+1} - \widehat{Q}_k W \widehat{B}_k^{\mathrm{T}}, \end{aligned}$$

where $\widehat{A}_{k+1} = \widehat{U}_{k+1}^{\mathrm{T}} K \widehat{U}_{k+1}$. By (4.6) and $\widehat{P}_{k+1}^{\mathrm{T}}\widehat{Q}_{k+1} = I_{(k+1)n_b}$, $\widetilde{V}_{k+2}^{\mathrm{T}}\widehat{P}_{k+1} = 0$. Set $\beta_j = \|\widetilde{V}_{k+2(:,j)}\|_2$ for $j = 1, 2, \ldots, n_b$, $\widehat{B}_{k+1} = \operatorname{diag}(\beta_1, \ldots, \beta_{n_b})$, and $\widehat{V}_{k+2} = \widetilde{V}_{k+2}\widehat{B}_{k+1}^{-1}$. Compute

$$\widehat{U}_{k+2} = M\widehat{V}_{k+2}\widehat{\Gamma}_{k+2}^{-1} \quad \text{with} \quad \widehat{\Gamma}_{k+2} = \widehat{V}_{k+2}^{\mathrm{T}} M \widehat{V}_{k+2},$$

and expand $\widehat{P}_{k+1}$ and $\widehat{Q}_{k+1}$ to $\widehat{P}_{k+2} = [\widehat{P}_{k+1}, \widehat{U}_{k+2}]$ and $\widehat{Q}_{k+2} = [\widehat{Q}_{k+1}, \widehat{V}_{k+2}]$, respectively. Consequently, we have

$$\begin{cases} K\widehat{P}_{k+1} = \widehat{Q}_{k+1} \begin{bmatrix} \widehat{T}_k & W\widehat{B}_k^{\mathrm{T}} \\ \widehat{B}_k W^{\mathrm{T}} & \widehat{A}_{k+1} \end{bmatrix} + \widehat{V}_{k+2}\widehat{B}_{k+1}E_{k+1}^{\mathrm{T}}, \\ M\widehat{Q}_{k+1} = \widehat{P}_{k+1} \begin{bmatrix} \widehat{D}_k & \\ & \widehat{\Gamma}_{k+1} \end{bmatrix}, \end{cases}$$

where $E_{k+1}^{\mathrm{T}} = [0_{n_b \times k\, n_b}, I_{n_b}]$.

Continue the procedure for $i \geq 2$ to obtain

$$\begin{aligned} \widetilde{V}_{k+i+1} &= K\widehat{U}_{k+i} - \widehat{V}_{k+i}\widehat{U}_{k+i}^{\mathrm{T}} K\widehat{U}_{k+i} - \widehat{V}_{k+i-1}\widehat{U}_{k+i-1}^{\mathrm{T}} K\widehat{U}_{k+i} - \widehat{Q}_{k+i-2}\widehat{P}_{k+i-2}^{\mathrm{T}} K\widehat{U}_{k+i} \\ &= K\widehat{U}_{k+i} - \widehat{V}_{k+i}\widehat{A}_{k+i} - \widehat{V}_{k+i-1}\widehat{B}_{k+i-1}^{\mathrm{T}}, \end{aligned}$$

where $\widehat{A}_{k+i} = \widehat{U}_{k+i}^{\mathrm{T}} K \widehat{U}_{k+i}$. The last equality holds because

$$\widehat{P}_{k+i-2}^{\mathrm{T}} K \widehat{U}_{k+i} = 0_{(k+i-2)n_b \times n_b} \quad \text{and} \quad \widehat{U}_{k+i-1}^{\mathrm{T}} K = \widehat{B}_{k+i-1}^{\mathrm{T}} \widehat{V}_{k+i}^{\mathrm{T}} + \widehat{A}^{\mathrm{T}} \widehat{Q}_{k+i-1}^{\mathrm{T}},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where

$$\widehat{A}^{\mathrm{T}} = \begin{cases} [\widehat{B}_k W^{\mathrm{T}},\ \widehat{A}_{k+i-1}^{\mathrm{T}}], & i = 2, \\ [0_{n_b\times(k+i-3)n_b},\ \widehat{B}_{k+i-2}^{\mathrm{T}},\ \widehat{A}_{k+i-1}^{\mathrm{T}}], & i > 2. \end{cases}$$

We summarize what we have obtained in this section in Algorithm 4.1, the thick-restart block Lanczos Algorithm for LREP. We denote it by `BlanLR(`n`,`k`)` where the indices n and k are the parameters for the thick-restart. We point out that after the first restart of `BlanLR(`n`,`k`)`, i.e., after the execution of line 19 of Algorithm 4.1, the new computed basis matrices $\widehat{P}_n, \widehat{Q}_n \in \mathbb{R}^{N\times nn_b}$, the symmetric matrix $\widehat{T}_n \in \mathbb{R}^{nn_b\times nn_b}$ and the block-diagonal matrix $\widehat{D}_n \in \mathbb{R}^{nn_b\times nn_b}$ in Algorithm 4.1 satisfy the relationship $\widehat{P}_n^{\mathrm{T}}\widehat{Q}_n = I_{nn_b}$ and

$$\begin{cases} K\widehat{P}_n = \widehat{Q}_n\widehat{T}_n + \widehat{V}_{n+1}\widehat{B}_n E_n^{\mathrm{T}}, \\ M\widehat{Q}_n = \widehat{P}_n\widehat{D}_n, \end{cases}$$

where

$$\widehat{T}_n = \begin{bmatrix} \widehat{T}_k & W\widehat{B}_k^{\mathrm{T}} & & & \\ \widehat{B}_k W^{\mathrm{T}} & \widehat{A}_{k+1} & \widehat{B}_{k+1}^{\mathrm{T}} & & \\ & \widehat{B}_{k+1} & \widehat{A}_{k+2} & \ddots & \\ & & \ddots & \ddots & \widehat{B}_{n-1}^{\mathrm{T}} \\ & & & \widehat{B}_{n-1} & \widehat{A}_n \end{bmatrix}, \tag{4.7}$$

$$\widehat{D}_n = \mathrm{diag}(I_{kn_b},\ \widehat{\Gamma}_{k+1}, \ldots, \widehat{\Gamma}_n).$$

Note that $\widehat{T}_n$ is no longer a block-tridiagonal matrix.

Finally, we have a few more remarks for `BlanLR(`n`,`k`)`:

1. In our numerical implementation of `BlanLR(`n`,`k`)`, we monitor the convergence of a Ritz pair $(\mu_j, \tilde{z}_j)$ by its relative residual norm

$$r(\mu_j) = \frac{\|H\tilde{z}_j - \mu_j\tilde{z}_j\|_1}{(\|H\|_1 + |\mu_j|)\|\tilde{z}_j\|_1}. \tag{4.8}$$

2. In order not to miss the wanted eigenvalues, we keep the converged Ritz values in ascending order in the restarting swap procedure. When K is indefinite, purely imaginary Ritz values might emerge. In this case we sort them by ascending order of their squares.
3. A deflation procedure is part of our numerical implementation. According to Lemma 2.2(c), deflations can be done by orthogonalizing the newly generated block $\widetilde{V}$ against the associated part of the converged eigenvectors.
4. While our block Lanczos method `BlanLR(`n`,`k`)` has been developed for real symmetric K and M, the algorithm can be rewritten to work for (complex) Hermitian K and M. This is done by simply replacing all $\mathbb{R}$ by $\mathbb{C}$ (the set of complex numbers) and each matrix/vector transpose by complex conjugate transpose.

5. Numerical examples. In this section, we present some numerical examples to illustrate the sharpness of our upper bounds for the convergence of the block Lanczos method and evaluate the effectiveness of the thick-restart block Lanczos method for LREP.

EXAMPLE 5.1. We first examine our upper bounds of Theorems 3.6 and 3.8. For simplicity, we consider diagonal matrices for K and M in this example. Set $N = 100$, take

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 4.1 The thick-restart block Lanczos Algorithm for LREP (`BlanLR(`n`,`k`)`).

Input: $U_0, V_0 \in \mathbb{R}^{N\times n_b}$ such that $MV_0 = U_0$ and integers $k, n \geq 1$.
Output: Converged Ritz pairs $(\mu_j, \tilde{z}_j)$ in (3.4).

Generate P_n, Q_n, T_n, D_n and V_{n+1} by Algorithm 3.1.
  Compute the approximate eigenpairs of LERP.
  if the stopping criterion is satisfied **then**
    return;
  else
`% the restart begins`
    Compute the decomposition $T_n = S\Omega_n^2 S^{\mathrm{T}}$ and $D_n = RR^{\mathrm{T}}$ as in (4.2).
    Let $S_k = S_{(:,1:kn_b)}$, $R_k = R_{(:,1:kn_b)}$, and $\Omega_k = \Omega_{n\,(1:kn_b,1:kn_b)}$.
    Compute $\widehat{P}_k = P_n R_k$, $\widehat{Q}_k = Q_n S_k$, and $W = R_k^{\mathrm{T}} E_n$.
    Set $\widehat{D}_k = I_{kn_b}$, $\widehat{V}_{k+1} = V_{n+1}$, $\widehat{B}_k = B_n$, and $\widehat{T}_k = \Omega_k^2$.
    Compute $\widehat{\Gamma}_{k+1} = \widehat{V}_{k+1}^{\mathrm{T}} M \widehat{V}_{k+1}$ and $\widehat{U}_{k+1} = M\widehat{V}_{k+1}\widehat{\Gamma}_{k+1}^{-1}$.
    Compute $\widehat{A}_{k+1} = \widehat{U}_{k+1}^{\mathrm{T}} K \widehat{U}_{k+1}$ and $\widetilde{V}_{k+2} = K\widehat{U}_{k+1} - \widehat{V}_{k+1}\widehat{A}_{k+1} - \widehat{Q}_k W \widehat{B}_k^{\mathrm{T}}$.
    $\beta_j = \|\widetilde{V}_{k+2\,(:,j)}\|_2$ $(j = 1 : n_b)$, $\widehat{B}_{k+1} = \mathrm{diag}(\beta_1, \ldots, \beta_{n_b})$, $\widehat{V}_{k+2} = \widetilde{V}_{k+2}\widehat{B}_{k+1}^{-1}$.
    Set $\widehat{P}_{k+1} = [\widehat{P}_k, \widehat{U}_{k+1}]$ and $\widehat{Q}_{k+1} = [\widehat{Q}_k, \widehat{V}_{k+1}]$.
  end if
`% the restart loop`
  for $i = k+2, \ldots, n$
    $\widehat{\Gamma}_i = \widehat{V}_i^{\mathrm{T}} M \widehat{V}_i$, $\widehat{U}_i = M\widehat{V}_i\widehat{\Gamma}_i^{-1}$, $\widehat{A}_i = \widehat{U}_i^{\mathrm{T}} K \widehat{U}_i$, $\widetilde{V}_{i+1} = K\widehat{U}_i - \widehat{V}_i\widehat{A}_i - \widehat{V}_{i-1}\widehat{B}_{i-1}^{\mathrm{T}}$.
    $\beta_j = \|\widetilde{V}_{i+1\,(:,j)}\|_2$ $(j = 1 : n_b)$, $\widehat{B}_i = \mathrm{diag}(\beta_1, \ldots, \beta_{n_b})$, $\widehat{V}_{i+1} = \widetilde{V}_{i+1}\widehat{B}_i^{-1}$.
    Set $\widehat{P}_i = [\widehat{P}_{i-1}, \widehat{U}_i]$ and $\widehat{Q}_i = [\widehat{Q}_{i-1}, \widehat{V}_i]$.
  end for
`% the restart ends`
Goto step 2 with $T_n = \widehat{T}_n$ and $D_n = \widehat{D}_n$ given in (4.7) and $P_n = \widehat{P}_n$, $Q_n = \widehat{Q}_n$, $V_{n+1} = \widehat{V}_{n+1}$.

$M = K = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$, where

$$\lambda_1 = 1 - \eta,\ \lambda_2 = 1,\ \lambda_3 = 1 + \eta, \quad \lambda_j = 4 + \frac{5j}{N},\ \text{for } j = 4, \ldots, N,$$

and set $i = \ell_1 = 1$, $\ell_2 = 3$ and $n_b = 3$. In such a case, there are two eigenvalue clusters: $\{\pm\lambda_1, \pm\lambda_2, \pm\lambda_3\}$ and $\{\pm\lambda_4, \ldots, \pm\lambda_N\}$, and $Y = K^{-1/2}$.

We seek the approximations associated with the first cluster $\{\pm\lambda_1, \pm\lambda_2, \pm\lambda_3\}$. In addition, we vary the parameter $\eta > 0$ to control the tightness among eigenvalues in the first cluster and check how it affects the upper bounds of the approximate eigenpair errors in the block Lanczos method for LREP. To make the numerical example repeatable, the initial block V_0 is chosen to be

$$V_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{N} & \sin 1 & \cos 1 \\ \vdots & \vdots & \vdots \\ \frac{N-n_b}{N} & \sin(N-n_b) & \cos(N-n_b) \end{bmatrix}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In such a way, V_0 satisfies the condition (3.6), i.e., $V_0^{\mathrm{T}} M Y_{(:,1:3)}$ is nonsingular. We implement the simple version of the block Lanczos method, i.e., Algorithm 3.1 for LREP in MATLAB with full reorthogonalization, and stop at $n = 20$, and then check the bounds for

$$\operatorname{diag}(\mu_1^2 - \lambda_1^2, \mu_2^2 - \lambda_2^2, \mu_3^2 - \lambda_3^2) \quad \text{and} \quad \tan\Theta_M(Y_{(:,1:3)}, Q_n)$$

given by (3.8) and (3.14), respectively. Since $i = \ell_1 = 1$, we know $\Delta_{\ell_1} = 1$ and $\Xi_{\ell_1,\ell_2} = 1$ in (3.8). Similarly, the first term in the right hand side of (3.14) also equals 1. For this reason, we compute the following factors

$$\begin{aligned}
\varepsilon_1 &= \|\operatorname{diag}(\mu_1^2 - \lambda_1^2, \mu_2^2 - \lambda_2^2, \mu_3^2 - \lambda_3^2)\|_{\mathrm{F}}, \\
\varepsilon_2 &= (\lambda_N^2 - \lambda_1^2) \times \Upsilon_{20,1,3}^{-2} \times \|\tan^2\Theta_M(Y_{(:,1:3)}, \Psi)\|_{\mathrm{F}}, \\
\varepsilon_3 &= \|\tan\Theta_M(Y_{(:,1:3)}, Q_n)\|_{\mathrm{F}}, \\
\varepsilon_4 &= \Upsilon_{20,1,3}^{-1} \times \|\tan\Theta_M(Y_{(:,1:3)}, \Psi)\|_{\mathrm{F}},
\end{aligned}$$

where Ψ can be computed by (3.7) and indeed $\Psi = V_0(Y_{(:,1:3)}^{\mathrm{T}} M V_0)^{-1}$ in this case. In fact, by (3.8) and (3.14), ε_2 and ε_4 are upper bounds for ε_1 and ε_3, respectively. As η goes to 0, Table 5.1 reports the numerical results of ε_i for $i = 1, 2, 3, 4$, from which we can see that our bounds for the eigenvalues for cluster and the associated eigenspace are rather sharp. In particular, the upper bounds ε_2 and ε_4 are comparable to the observed errors ε_1 and ε_3; furthermore, they appear to be insensitive to η when η goes to 0.

TABLE 5.1
ε_1, ε_3 together with their corresponding upper bounds ε_2 and ε_4 of Example 5.1.

η	ε_1	ε_2 (bound for ε_1)	ε_3	ε_4 (bound for ε_3)
10^{-1}	2.1366×10^{-12}	1.1430×10^{-11}	3.6500×10^{-8}	4.9611×10^{-7}
10^{-2}	2.4337×10^{-12}	9.4095×10^{-12}	2.1073×10^{-8}	4.4960×10^{-7}
10^{-3}	1.5237×10^{-12}	9.2447×10^{-12}	3.9425×10^{-8}	4.4555×10^{-7}
10^{-4}	4.5743×10^{-13}	9.2286×10^{-12}	3.3320×10^{-8}	4.4515×10^{-7}
10^{-5}	8.3923×10^{-13}	9.2269×10^{-12}	2.9802×10^{-8}	4.4511×10^{-7}

EXAMPLE 5.2. To test the effectiveness of the block Lanczos method with thick-restart technique for LREP, we choose 4 test problems TEST 1 to TEST 4 used previously in [21]. In particular, TEST 1 and TEST 2 come from the linear response analysis for Na_2 and silane (SiH4) compound, respectively, which are generated by the turboTDDFT code in QUANTUM ESPRESSO [7]. The matrices K and M of TEST 1 and TEST 2 are symmetric positive definite of order $N = 1862$ and 5660, respectively. TEST 3 and TEST 4 are then chosen to evaluate `BlanLR(`n`,`k`)` for the case when K is indefinite. TEST 3 and TEST 4 consist of matrices from the University of Florida Sparse Matrix Collection [6] to give K and M where M is definite but K indefinite. The features of these matrices are presented in Table 5.2. In the case when the two matrices from the collection have different dimensions, we extract the leading principal submatrix of the larger one to have K or M of equal size.

We compare the thick-restart block Lanczos method in Algorithm 4.1 (`BlanLR(`n`,`k`)`) with the block Lanczos method without restart in Algorithm 3.1 (denote by `BlanLR`). Our goal is to compute the first 5 eigenvalues, i.e., λ_i in (1.4) for $i = 1, \ldots, 5$, and the corresponding eigenvectors. A computed approximate eigenpair $(\mu_j, \tilde{z}_j)$ is considered as converged when its relative residual norm is bounded by 10^{-8},

$$r(\mu_j) = \frac{\|H\tilde{z}_j - \mu_j\tilde{z}_j\|_1}{(\|H\|_1 + |\mu_j|)\|\tilde{z}_j\|_1} \le 10^{-8}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.2
The matrices K and M of TEST 3 *and* TEST 4.

Problem	N	K	M
TEST 3	5832	`Na5`	`fv1`
TEST 4	74752	SiO_2	`finan512`

In this example, the block size is chosen as $n_b = 3$, initially $V_0 = \texttt{eye}(N, 3)$ where `eye` is MATLAB's built-in function; the parameters $n = 30$ and $k = 20$ are used for the restart, which means that the restart will be triggered whenever the dimension of the projection subspace is larger than 90, and then 60 basis vectors are kept. We carried out our testing in MATLAB version 8.5 (R2015a) on a laptop with 8G memory and CPU Intel core i5-3210M@2.50GHz.

The approximate eigenvalues (Ritz values) and the associated eigenvectors (Ritz vectors) are computed when the dimension of the projection subspace fulfills the condition of the restart in `BlanLR(`n`,`k`)`. For Algorithm 3.1 (`BlanLR`), since there is no restarting and the Lanczos process continues, we then choose to calculate the approximate eigenpairs (Ritz pairs) whenever the same amount of Lanczos steps is carried out as in `BlanLR(`n`,`k`)`, i.e., we compute the approximate eigenpairs whenever the Lanczos steps equal to $30 + 10 \times (j - 1)$ for $j = 1, 2, \ldots$. We report the total number of Lanczos steps and the CPU time in seconds for `BlanLR(`n`,`k`)` and `BlanLR` in Table 5.3. One can see from Table 5.3 that the thick-restart block Lanczos method and the block Lanczos method without restart for LREP are competitive in the number of Lanczos steps. But the thick-restart block Lanczos method reduces remarkably the computation time (i.e., an indication of the reduction in the computational costs), which is mainly due to the saving in the orthogonalization procedure and in solving much smaller projected LREP's.

To illustrate the accuracy of computed approximations, finally, we also calculate the relative eigenvalue error

$$e(\mu_j) = \frac{|\mu_j - \lambda_j|}{|\lambda_j|},$$

as well as the relative residual norm $r(\mu_j)$ given in (4.8) for the jth approximate eigenpair $(\mu_j, \tilde{z}_j)$. The accuracy of the first two computed eigenpairs of TEST 1 are compared between `BlanLR(`n`,`k`)` and `BlanLR`. The corresponding numerical results are plotted in Figure 5.1. It is clearly shown by Figure 5.1 that, compared to the simple version of the block Lanczos method, the thick-restart block Lanczos method just needs one or two more restarts to obtain the first two eigenpairs of TEST 1 in the same accuracy. Since the dimension of the projected problem is bounded by nn_b in `BlanLR(`n`,`k`)`, the savings from the orthogonalization and from the Ritz procedure for computing the resulting much smaller projected LREP's outweigh the additional restart steps.

TABLE 5.3
The number of Lanczos steps and CPU time in seconds for computing the first 5 eigenpairs of TEST 1 *to* TEST 4 *by the* `BlanLR` *and* `BlanLR(30,20)`.

	`BlanLR(30,20)`		`BLanLR`	
	CPU time(s)	Lanczos steps	CPU time(s)	Lanczos steps
TEST 1	3.816	173	6.541	149
TEST 2	65.760	393	117.711	349
TEST 3	3.586	253	20.916	229
TEST 4	106.579	553	776.920	469

ETNA
Kent State University and
Johann Radon Institute (RICAM)

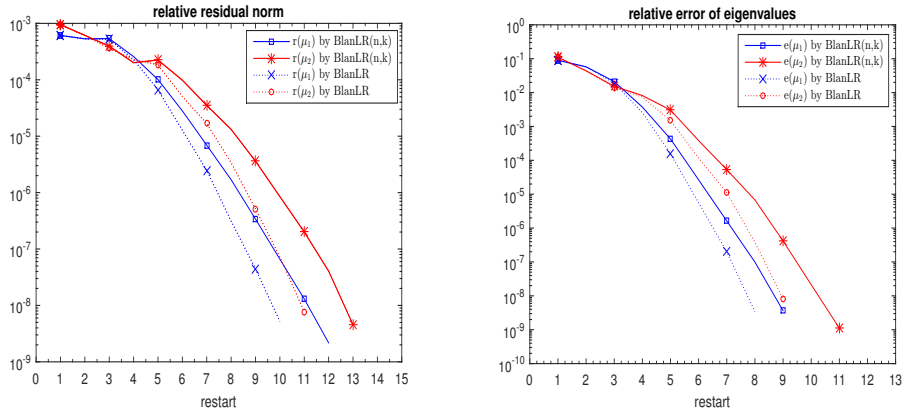


FIG. 5.1. *Convergence behavior of* `BlanLR(n,k)` *with* $(n,k) = (30,20)$ *and* `BlanLR` *for computing the first 2 eigenpairs of* TEST 1.

6. Concluding remarks. In this paper, motivated by the fact that in LREP only a small portion of eigenpairs near zero are required to be computed efficiently, we proposed a block Lanczos method for (1.1). Theoretical bounds for the eigenvalue and eigenvector approximations are established in Theorems 3.6 and 3.8, respectively. These theorems are tailored particularly to bound the errors in approximate eigenpairs belonging to a cluster of eigenvalues, including the case of multiple eigenvalues; they are also applicable in the case of simple eigenpairs. These theoretical convergence results reveal the accuracy of the approximations of both eigenvalues in a cluster and eigenspace and show, to some extent, the advantages of the block Lanczos method over the single-vector version. To make this block Lanczos method more practical, we discussed in detail a thick-restart procedure to reduce memory and orthogonalization costs. Numerical examples are presented to demonstrate that the final thick-restart block Lanczos method can compute the desired eigenvalues in a cluster around zero efficiently.

Acknowledgements. The authors are grateful to the anonymous referees for their careful reading, useful comments, and suggestions for improving the presentation of this paper.

REFERENCES

[1] Z. BAI AND R.-C. LI, *Minimization principles for the linear response eigenvalue problem I: Theory*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1075–1100.
[2] ———, *Minimization principles for the linear response eigenvalue problem II: Computation*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 392–416.
[3] ———, *Minimization principles and computation for the generalized linear response eigenvalue problem*, BIT, 54 (2014), pp. 31–54.
[4] Z. BAI, R.-C. LI, AND W.-W. LIN, *Linear response eigenvalue problem solved by extended locally optimal preconditioned conjugate gradient methods*, Sci. China Math., 59 (2016), pp. 1443–1460.
[5] R. BHATIA, *Matrix Analysis*, Springer, New York, 1997.
[6] T. A. DAVIS AND Y. HU, *The University of Florida sparse matrix collection*, ACM Trans. Math. Software, 38 (2011), Art. 1, 25 pages.
[7] P. GIANNOZZI ET AL., *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*, J. Phys.: Condens. Matter, 21 (2009), Art. 395502, 19 pages.
[8] M. GRÜNING, A. MARINI, AND X. GONZE, *Exciton-plasmon states in nanoscale materials: breakdown of the Tamm-Dancoff approximation*, Nano Lett., 9 (2009), pp. 2820–2824.
[9] R. B. LEHOUCQ AND D. C. SORENSEN, *Deflation techniques for an implicitly restarted Arnoldi iteration*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 789–821.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[10] R.-C. LI AND L.-H. ZHANG, *Convergence of the block Lanczos method for eigenvalue clusters*, Numer. Math., 131 (2015), pp. 83–113.
[11] M. T. LUSK AND A. E. MATTSSON, *High-performance computing for materials design to advance energy science*, MRS Bull. 36 (2011), pp. 169–174.
[12] P. PAPAKONSTANTINOU, *Reduction of the RPA eigenvalue problem and a generalized Cholesky decomposition for real-symmetric matrices*, Europhys. Lett. EPL, 78 (2007), Art. 12001, 5 pages.
[13] P. RING, Z.-Y. MA, N. V. GIAI, D. VRETENAR, A. WANDELT, AND L.-G. GAO, *The time-dependent relativistic mean-field theory and the random phase approximation*, Nuclear Phys. A, 694 (2001), pp. 249–268.
[14] D. ROCCA, *Time-Dependent Density Functional Perturbation Theory: New algorithms with applications to molecular spectra*, PhD. Thesis, International School for Advanced Studies, Trieste, Italy, 2007.
[15] D. ROCCA, D. LU, AND G. GALLI, *Ab initio calculations of optical absorpation spectra: solution of the Bethe-Salpeter equation within density matrix perturbation theory*, J. Chem. Phys., 133 (2010), Art. 164109, 10 pages.
[16] D. C. SORENSEN, *Implicit application of polynomial filters in a k-step Arnoldi method*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 357–385.
[17] G. W. STEWART, *A Krylov-Schur algorithm for large eigenproblems*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 601–614.
[18] R. E. STRATMANN, G. E. SCUSERIA, AND M. J. FRISCH, *An efficient implementation of time-dependent density-functional theory for the calculation of excitation of large molecules*, J. Chem. Phys., 109 (1998), pp. 8218–8824.
[19] Z. TENG AND R.-C. LI, *Convergence analysis of Lanczos-type methods for the linear response eigenvalue problem*, J. Comput. Appl. Math., 247 (2013), pp. 17–33.
[20] Z. TENG, L. LU, AND R.-C. LI, *Perturbation of partitioned linear response eigenvalue problems*, Electron. Trans. Numer. Anal., 44 (2015), pp. 624–638.
http://etna.ricam.oeaw.ac.at/vol.44.2015/pp624-638.dir/pp624-638.pdf
[21] Z. TENG, Y. ZHOU, AND R.-C. LI, *A block Chebyshev-Davidson method for linear response eigenvalue problems*, Adv. Comput. Math., 42 (2016), pp. 1103–1128.
[22] E. V. TSIPER, *Variational procedure and generalized Lanczos recursion for small-amplitude classical oscillations*, JETP Lett., 70 (1999), pp. 751–755.
[23] ———, *A classical mechanics technique for quantum linear response*, J. Phys. B, 34 (2001), pp. L401–L407.
[24] W.-G. WANG, L.-H. ZHANG, AND R.-C. LI, *Error bounds for approximate deflating subspaces for linear response eigenvalue problems*, Linear Algebra Appl., 528 (2017), pp. 273–289.
[25] K. WU AND H. SIMON, *Thick-restart Lanczos method for large symmetric eigenvalue problems*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 602–616.
[26] I. YAMAZAKI, Z. BAI, H. SIMON, L.-W. WANG, AND K. WU, *Adaptive projection subspace dimension for the thick-restart Lanczos method*, ACM Trans. Math. Software, 37 (2010), Art. 27, 18 pages.
[27] L.-H. ZHANG, W.-W. LIN, AND R.-C. LI, *Backward perturbation analysis and residual-based error bounds for the linear response eigenvalue problem*, BIT, 55 (2015), pp. 869–896.
[28] L.-H. ZHANG, J. XUE, AND R.-C. LI, *Rayleigh-Ritz approximation for the linear response eigenvalue problem*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 765–782.