Rayleigh Quotient Based Optimization Methods For Eigenvalue Problems

Ren-Cang Li[*]

November 13, 2013

Abstract

Four classes of eigenvalue problems that admit similar min-max principles and the Cauchy interlacing inequalities as the symmetric eigenvalue problem famously does are investigated. These min-max principles pave ways for efficient numerical solutions for extreme eigenpairs by optimizing the so-called Rayleigh quotient functions. In fact, scientists and engineers have already been doing that for computing the eigenvalues and eigenvectors of Hermitian matrix pencils $A - \lambda B$ with B positive definite, the first class of our eigenvalue problems. But little attention has gone to the other three classes: positive semidefinite pencils, linear response eigenvalue problems, and hyperbolic eigenvalue problems, in part because most min-max principles for the latter were discovered only very recently and some more are being discovered. It is expected that they will drive the effort to design better optimization based numerical methods for years to come.

1 Introduction

Eigenvalue problems are ubiquitous. Eigenvalues explain many physical phenomena well such as vibrations and frequencies, (in)stabilities of dynamical systems, and energy levels in molecules or atoms. This article focuses on classes of eigenvalue problems that admit various min-max principles and the Cauchy interlacing inequalities as the symmetric eigenvalue problem famously does [4, 38, 47]. These results make it possible to efficiently calculate extreme eigenpairs of the eigenvalue problems by optimizing associated Rayleigh quotients.

Consider the generalized eigenvalue problem

$$Ax = \lambda Bx, \tag{1}$$

where both A and B are Hermitian. The first class of eigenvalue problems are those for which B is also positive definite. Such an eigenvalue problem is equivalent to a symmetric eigenvalue problem $B^{-1/2}AB^{-1/2}y = \lambda x$ and thus, not surprisingly, all min-max principles (Courant-Fischer, Ky Fan trace min/max, Wielandt-Lidskii) and the Cauchy interlacing inequalities have their counterparts in this eigenvalue problem. The associated *Rayleigh quotient* is

$$\rho(x) = \frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}Bx}. \tag{2}$$

[*]Department of Mathematics, University of Texas at Arlington, P.O. Box 19408, Arlington, TX 76019. E-mail: rcli@uta.edu. Supported in part by NSF grants DMS-1115834 and DMS-1317330, and a research gift grant from Intel Corporation.

When B is indefinite and even singular, (1) is no longer equivalent to a symmetric eigenvalue problem in general and it may even have complex eigenvalues which clearly admit no min-max representations. But if there is a real scalar λ_0 such that $A - \lambda_0 B$ is positive semidefinite, then the eigenvalue problem (1) has only real eigenvalues and they admit similar min-max principles and the Cauchy interlacing inequalities [25, 27, 29]. This is the second class of eigenvalue problems and it share the same Rayleigh quotient (2) as the first class. We call a matrix pencil in this class a *positive semidefinite pencil.* Opposite to the concept of a *positive semidefinite matrix pencil*, naturally, is that of a *negative semidefinite matrix pencil* $A - \lambda B$ by which we mean that A and B are Hermitian and there is a real λ_0 such that $A - \lambda_0 B$ is negative semidefinite. Evidently, if $A - \lambda B$ is a negative semidefinite matrix pencil, then $-(A - \lambda B) = (-A) - \lambda(-B)$ is a positive semidefinite matrix pencil because $(-A) - \lambda_0(-B) = -(A - \lambda_0 B)$. Therefore it suffices to only study either positive or negative semidefinite pencils.

The third class of eigenvalue problems is the so-called *linear response eigenvalue problem* or *random phase approximation eigenvalue problem*

$$\begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} y \\ x \end{bmatrix},$$

where K and M are Hermitian and positive semidefinite matrices and one of them is definite. Any eigenvalue problem in this class can be turned into one in the second class by permuting the first and second block rows to get

$$\begin{bmatrix} M & 0 \\ 0 & K \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix},$$

where I is the identity matrix of apt size. In this sense, the third class is a subclass of the second class, but with block substructures. The associated Rayleigh quotient is

$$\rho(x, y) = \frac{x^{\mathrm{H}} K x + y^{\mathrm{H}} M y}{2|x^{\mathrm{H}} y|}.$$

The first minimization principle for such eigenvalue problems was essentially published by Thouless [50, 1961], but more were obtained only very recently [2, 3].

The fourth class of eigenvalue problems is the *hyperbolic quadratic eigenvalue problem*

$$(\lambda^2 A + \lambda B + C)x = 0$$

arising from dynamical systems with friction, where A, B, and C are Hermitian and A is positive definite and

$$(x^{\mathrm{H}} B x)^2 - 4(x^{\mathrm{H}} A x)(x^{\mathrm{H}} C x) > 0 \quad \text{for any nonzero vector } x.$$

The associated Rayleigh quotients are

$$\rho_{\pm}(x) = \frac{-x^{\mathrm{H}} B x \pm \left[(x^{\mathrm{H}} B x)^2 - 4(x^{\mathrm{H}} A x)(x^{\mathrm{H}} C x)\right]^{1/2}}{2(x^{\mathrm{H}} A x)}.$$

Courant-Fischer type min-max principles were known to Duffin [10, 1955] and the Cauchy type interlacing inequalities to Veselić [52, 2010]. Other min-max principles (Wielandt-Lidskii type, Ky Fan trace min/max type) are being discovered [28].

In the rest of this article, we will explain the steepest descent/ascent methods and nonlinear conjugate gradient methods for the first class of eigenvalue problems, including the incorporation of preconditioning techniques, extending search spaces, and block implementations, in detail but only state min-max principles – old and new – for the other three classes. The interested reader can consult relevant references for the corresponding steepest descent/ascent methods and nonlinear conjugate gradient methods or design his own based on the min-max principles stated.

Notation. Throughout this paper, $\mathbb{C}^{n\times m}$ is the set of all $n\times m$ complex matrices, $\mathbb{C}^n=\mathbb{C}^{n\times 1}$, and $\mathbb{C}=\mathbb{C}^1$, and similarly $\mathbb{R}^{n\times m}$, $\mathbb{R}^n$, and $\mathbb{R}$ are for their real counterparts. I_n (or simply I if its dimension is clear from the context) is the $n\times n$ identity matrix, and e_j is its jth column. The superscript "$\cdot^{\mathrm{T}}$" and "$\cdot^{\mathrm{H}}$" take transpose and complex conjugate transpose of a matrix/vector, respectively. For a matrix X, $\mathcal{R}(X)$ and $\mathcal{N}(X)$ are the column space and null space of X, respectively.

We shall also adopt MATLAB-like convention to access the entries of vectors and matrices. Let $i:j$ be the set of integers from i to j inclusive. For a vector u and an matrix X, $u_{(j)}$ is u's jth entry, $X_{(i,j)}$ is X's (i,j)th entry; X's submatrices $X_{(k:\ell,i:j)}$, $X_{(k:\ell,:)}$, and $X_{(:,i:j)}$ consist of intersections of row k to row ℓ and column i to column j, row k to row ℓ, and column i to column j, respectively.

For $A\in\mathbb{C}^{n\times n}$, $A\succ 0$ ($A\succeq 0$) means that A is Hermitian and positive (semi-)definite, and $A\prec 0$ ($A\preceq 0$) means $-A\succ 0$ ($-A\succeq 0$).

2 Hermitian Pencil $A-\lambda B$ with Definite B

In this section, we consider the generalized eigenvalue problem

$$Ax=\lambda Bx, \tag{3}$$

where $A,B\in\mathbb{C}^{n\times n}$ are Hermitian with $B\succ 0$. When the equation (3) for a scalar $\lambda\in\mathbb{C}$ and $0\neq x\in\mathbb{C}^n$ holds, λ is called an *eigenvalue* and x a corresponding *eigenvector*. Theoretically, it is equivalent to the standard Hermitian eigenvalue problem

$$B^{-1/2}AB^{-1/2}y=\lambda y. \tag{4}$$

Both have the same eigenvalues with eigenvectors related by $y=B^{1/2}x$, where $B^{-1/2}=(B^{-1})^{1/2}$ is the positive definite square root of B^{-1} (also $B^{-1/2}=(B^{1/2})^{-1}$) [5, 19].

Numerically, if it has to be done (usually for modest n, up to a few thousands), the conversion of (3) to a standard Hermitian eigenvalue problem is usually accomplished through B's Cholesky decomposition: $B=R^{\mathrm{H}}R$, where R is upper triangular, rather than B's square root which is much more expensive to compute but often advantageous for theoretical investigations. The converted eigenvalue problem then is

$$R^{-\mathrm{H}}AR^{-1}y=\lambda y \tag{5}$$

with eigenvectors related by $y=Rx$, and can be solved as a dense eigenvalue problem by LAPACK [1] for modest n.

But calculating the Cholesky decomposition can be very expensive, too, for large n, not to mention possible fill-ins for unstructured sparse B. In this section, we are concerned with Rayleigh Quotient based optimization methods to calculate a few extreme eigenvalues of (3).

By the theoretical equivalence of (3) to the standard Hermitian eigenvalue problem (4) or (5), we know that (3) has n real eigenvalues and B-orthonormal eigenvectors.

Throughout the rest of this section, $A-\lambda B$ will be assumed a Hermitian matrix pencil of order n with $B \succ 0$, and its eigenvalues, eigenvectors, and eigen-decomposition are given by (6).

$$\boxed{\begin{array}{rl} \text{eigenvalues:} & \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n, \text{ and} \\ & \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \\ B\text{-orthonormal eigenvectors:} & u_1, u_2, \dots, u_n, \text{ and} \\ & U = [u_1, u_2, \dots, u_n], \\ \text{eigen-decomposition:} & U^{\mathrm{H}} A U = \Lambda \text{ and } U^{\mathrm{H}} B U = I_n. \end{array}} \tag{6}$$

In what follows, our focus is on computing the first few smallest eigenvalues and their associated eigenvectors. The case for the largest few eigenvalues can be dealt with in the same way by replacing A by $-A$, i.e., considering $(-A) - \lambda B$ instead.

2.1 Basic Theory

Given $x \in \mathbb{C}^n$, the ***Rayleigh Quotient*** for the generalized eigenvalue problem $Ax = \lambda Bx$ is defined by

$$\rho(x) = \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}. \tag{7}$$

Similarly for $X \in \mathbb{C}^{n\times k}$ with $\operatorname{rank}(X) = k$, the ***Rayleigh Quotient Pencil*** is

$$X^{\mathrm{H}} A X - \lambda X^{\mathrm{H}} B X. \tag{8}$$

Theorem 2.1 collects important min-max results and the Cauchy interlacing inequalities for the eigenvalue problem (3). They can be derived via the corresponding results for (4) or (5), the theoretical equivalence of (3) [4, 38, 47].

Theorem 2.1. *Let $A - \lambda B$ be a Hermitian matrix pencil of order n with $B \succ 0$.*

1 (Courant-Fischer min-max principles) *For $j = 1, 2, \dots, n$,*

$$\lambda_j = \min_{\dim \mathcal{X} = j} \max_{x \in \mathcal{X}} \rho(x), \tag{9a}$$

$$\lambda_j = \max_{\operatorname{codim} \mathcal{X} = j-1} \min_{x \in \mathcal{X}} \rho(x). \tag{9b}$$

In particular,

$$\lambda_1 = \min_x \rho(x), \quad \lambda_n = \max_x \rho(x). \tag{10}$$

2 (Ky Fan trace min/max principles) *For $1 \le k \le n$,*

$$\sum_{i=1}^{k} \lambda_i = \min_{X^{\mathrm{H}} B X = I_k} \operatorname{trace}(X^{\mathrm{H}} A X), \tag{11a}$$

$$\sum_{i=n-k+1}^{n} \lambda_i = \max_{X^{\mathrm{H}} B X = I_k} \operatorname{trace}(X^{\mathrm{H}} A X). \tag{11b}$$

Furthermore if $\lambda_k < \lambda_{k+1}$, then $\mathcal{R}(X) = \mathcal{R}(U_{(:,1:k)})$ for any minimizing $X \in \mathbb{C}^{n\times k}$ for (11a)*; if $\lambda_{n-k} < \lambda_{n-k+1}$, then $\mathcal{R}(X) = \mathcal{R}(U_{(:,n-k+1:n)})$ for any maximizing $X \in \mathbb{C}^{n\times k}$ for* (11b).

3. **(Cauchy interlacing inequalities)** *Let $X \in \mathbb{C}^{n\times k}$ with* $\operatorname{rank}(X) = k$*, and denote by $\mu_1 \le \mu_2 \le \cdots \le \mu_k$ the eigenvalues of the Rayleigh quotient pencil* (8)*. Then*

$$\lambda_j \le \mu_j \le \lambda_{n-k+j} \quad \textit{for } 1 \le j \le k. \tag{12}$$

Furthermore if $\lambda_j = \mu_j$ for $1 \le j \le k$ and $\lambda_k < \lambda_{k+1}$, then $\mathcal{R}(X) = \mathcal{R}(U_{(:,1:k)})$; if $\mu_j = \lambda_{n-k+j}$ for $1 \le j \le k$ and $\lambda_{n-k} < \lambda_{n-k+1}$, then $\mathcal{R}(X) = \mathcal{R}(U_{(:,n-k+1:n)})$.

The computational implications of these results are as follows. The equations in (10) or (11) naturally leads to applications of optimization approaches to computing the first/last or first/last few eigenvalues and their associated eigenvectors, while the inequalities in (12) suggest that judicious choices of X can push μ_j either down to λ_j or up to λ_{n-k+j} for the purpose of computing them.

In pertinent to deflation, i.e., avoiding computing known or already computed eigenpairs, we have the following results.

Theorem 2.2. *Let integer $1 \le k < n$ and $\xi \in \mathbb{R}$.*

1. *We have*

$$\lambda_{k+1} = \min_{x \perp_B u_i,\, 1\le i\le k} \rho(x), \quad \lambda_{k+1} = \max_{x \perp_B u_i,\, k+2\le i\le n} \rho(x),$$

where $\perp_B$ stands for B-orthogonality, i.e., $x \perp_B y$ means $\langle x, y\rangle_B \equiv x^{\mathrm{H}}By = 0$.

2. *let $V = [u_1, u_2, \dots, u_k]$. The eigenvalues of matrix pencil $[A + \xi(BV)(BV)^{\mathrm{H}}] - \lambda B$ are*

$$\lambda_j + \xi \textit{ for } 1 \le j \le k \textit{ and } \lambda_j \textit{ for } k+1 \le j \le n$$

with the corresponding eigenvectors u_j for $1 \le j \le n$. In particular,

$$U^{\mathrm{H}}[A + \xi(BV)(BV)^{\mathrm{H}}]U = \begin{bmatrix} \Lambda_1 + \xi I_k & 0 \\ 0 & \Lambda_2 \end{bmatrix}, \quad U^{\mathrm{H}}BU = I_n,$$

where $\Lambda_1 = \Lambda_{(1:k,1:k)}$ and $\Lambda_2 = \Lambda_{(k+1:n,k+1:n)}$.

The concept of invariant subspace is very important in the standard eigenvalue problem and, more generally, operator theory. In a loose sense, computing a few eigenvalues of a large scale matrix $H \in \mathbb{C}^{n\times n}$ is equivalent to calculating a relevant invariant subspace $\mathcal{X}$ of H, i.e., a subspace $\mathcal{X} \subseteq \mathbb{C}^n$ such that $H\mathcal{X} \subseteq \mathcal{X}$. This concept naturally extends to the generalized eigenvalue problem for $A - \lambda B$ that we are interested in, i.e., A and B are Hermitian and $B \succ 0$.

Definition 2.1. A $\mathcal{X} \subseteq \mathbb{C}^n$ is called a *generalized invariant subspaces* of $A - \lambda B$ if

$$A\mathcal{X} \subseteq B\mathcal{X}.$$

Sometimes, it is simply called an *invariant subspace.*

Some important properties of an invariant subspaces are summarized into the following theorem whose proof is left as an exercise.

Theorem 2.3. *Let $\mathcal{X} \subseteq \mathbb{C}^n$ and* $\dim \mathcal{X} = k$*, and let $X \in \mathbb{C}^{n\times k}$ be a basis matrix of $\mathcal{X}$.*

1. *$\mathcal{X}$ is an invariant subspace of $A - \lambda B$ if and only if there is $A_1 \in \mathbb{C}^{k\times k}$ such that*

$$AX = BXA_1. \tag{13}$$

2. *Suppose $\mathcal{X}$ is an invariant subspace of $A - \lambda B$ and* (13) *holds. Then the following statements are true.*

   (a) *$A_1 = (X^{\rm H}BX)^{-1}(X^{\rm H}AX)$ and thus it has the same eigenvalues as $X^{\rm H}AX - \lambda X^{\rm H}BX$. If X has B-orthonormal columns, i.e., $X^{\rm H}BX = I_k$ (one can always pick a basis matrix like this), then $A_1 = X^{\rm H}AX$ which is also Hermitian.*

   (b) *For any eigenpair $(\hat\lambda, \hat x)$ of A_1: $A_1\hat x = \hat\lambda \hat x$, $(\hat\lambda, X\hat x)$ is an eigenpair of $A - \lambda B$.*

   (c) *Let $X_\perp \in \mathbb{C}^{n\times(n-k)}$ such that $Z := [X, X_\perp]$ is nonsingular and $X^{\rm H}BX_\perp = 0$. We have*

$$Z^{\rm H}AZ = \begin{bmatrix} X^{\rm H}AX & \\ & X_\perp^{\rm H}AX_\perp \end{bmatrix}, \quad Z^{\rm H}BZ = \begin{bmatrix} X^{\rm H}BX & \\ & X_\perp^{\rm H}BX_\perp \end{bmatrix}.$$

2.2 Rayleigh-Ritz Procedure

Theorem 2.3 says that partial spectral information can be extracted from an invariant subspace if known. But an exact invariant subspace is hard to come by in practice. Through computations we often end up with subspaces $\mathcal{X}$ that

1. are accurate approximate invariant subspaces themselves, or

2. have a nearby lower dimensional invariance subspace.

For the former, it means that $\|AX - BXA_1\|$ is tiny for some matrix A_1, where X is a basis matrix $\mathcal{X}$ and $\|\cdot\|$ is some matrix norm. For the latter, it means there is an invariant subspace $\mathcal{U}$ of a lower dimension than $\mathcal{X}$ such that the canonical angles from $\mathcal{U}$ to $\mathcal{X}$ are all tiny.

The Rayleigh-Ritz procedure is a way to extract approximate spectral information on $A - \lambda B$ for a given subspace that satisfies either one of the two requirements.

Algorithm 2.1 Rayleigh-Ritz procedure

Given a computed subspace $\mathcal{X}$ of dimension ℓ in the form of a basis matrix $X \in \mathbb{C}^{n\times\ell}$, this algorithm computes approximate eigenpairs of $A - \lambda B$.

```
1: compute the projection matrix pencil $X^{\rm H}AX - \lambda X^{\rm H}BX$ which is $\ell \times \ell$;
2: solve the eigenvalue problem for $X^{\rm H}AX - \lambda X^{\rm H}BX$ to obtain its eigenpairs $(\hat\lambda_i, \hat x_i)$
   which yield approximate eigenpairs $(\hat\lambda_i, X\hat x_i)$, called Rayleigh-Ritz pairs, for the original
   pencil $A - \lambda B$. These $\hat\lambda_i$ are called Ritz values and $X\hat x_i$ Ritz vectors.
```

If $\mathcal{X}$ is a true invariant subspace, the *Ritz values* and *Ritz vectors* as rendered by this Rayleigh-Ritz procedure are exact eigenvalues and eigenvectors of $A - \lambda B$ in the absence of roundoff errors, as guaranteed by Theorem 2.3. So in this sense, this Rayleigh-Ritz procedure is a natural thing to do. On the other hand, as for the standard symmetric eigenvalue problem, the procedure retains several optimality properties as we shall now explain.

By Theorem 2.1,

$$\lambda_i = \min_{\substack{\mathcal{Y}\subseteq\mathbb{C}^n \\ \dim\mathcal{Y}=i}} \max_{y\in\mathcal{Y}} \rho(y), \tag{14}$$

where the minimization is taken over all $\mathcal{Y} \subset \mathbb{C}^n$ with $\dim \mathcal{Y} = i$. So given $\mathcal{X} \subset \mathbb{C}^n$, the natural definition of the best approximation α_i to λ_i is to replace $\mathbb{C}^n$ by $\mathcal{X}$ to get

$$\alpha_i = \min_{\substack{\mathcal{Y}\subseteq\mathcal{X}\\ \dim\mathcal{Y}=i}} \max_{y\in\mathcal{Y}} \rho(y). \tag{15}$$

Any $\mathcal{Y} \subseteq \mathcal{X}$ with $\dim \mathcal{Y} = i$ can be represented by its basis matrix $Y \in \mathbb{C}^{n\times i}$ which in turn can be uniquely represented by $\widehat{Y} \in \mathbb{C}^{\ell\times i}$ with $\operatorname{rank}(\widehat{Y}) = i$ such that $Y = X\widehat{Y}$. So $y \in \mathcal{Y}$ is equivalent to $y = Y\hat{y} = X\widehat{Y}\hat{y} =: Xz$ for some unique $z \in \widehat{\mathcal{Y}} := \mathcal{R}(\widehat{Y}) \subseteq \mathbb{C}^\ell$. We have by (15)

$$\begin{aligned}\alpha_i &= \min_{\substack{\mathcal{Y}\subseteq\mathcal{X}\\ \dim\mathcal{Y}=i}} \max_{y\in\mathcal{Y}} \frac{y^{\mathrm{H}}Ay}{y^{\mathrm{H}}By}\\ &= \min_{\substack{\widehat{\mathcal{Y}}\subseteq\mathbb{C}^\ell\\ \dim\widehat{\mathcal{Y}}=i}} \max_{z\in\widehat{\mathcal{Y}}} \frac{z^{\mathrm{H}}X^{\mathrm{H}}AXz}{z^{\mathrm{H}}X^{\mathrm{H}}BXz} = \hat{\lambda}_i,\end{aligned}$$

the ith eigenvalues of $X^{\mathrm{H}}AX - \lambda X^{\mathrm{H}}BX$. This is the first optimality of the Rayleigh-Ritz procedure.

Suppose we are seeking λ_i for $1 \le i \le k$. By Theorem 11, we have

$$\sum_{i=1}^k \lambda_i = \min_{\substack{\mathcal{R}(Y)\subseteq\mathbb{C}^n\\ Y^{\mathrm{H}}BY=I_k}} \operatorname{trace}(Y^{\mathrm{H}}AY) \tag{16}$$

where the minimization is taken over all $Y \in \mathbb{C}^{n\times k}$ satisfying $Y^{\mathrm{H}}BY = I_k$. So given $\mathcal{X} \subset \mathbb{C}^n$, the natural definition for the best approximation is to replace $\mathbb{C}^n$ by $\mathcal{X}$ to achieve

$$\min_{\substack{\mathcal{R}(Y)\subseteq\mathcal{X}\\ Y^{\mathrm{H}}BY=I_k}} \operatorname{trace}(Y^{\mathrm{H}}AY). \tag{17}$$

Any $\mathcal{R}(Y) \subseteq \mathcal{X}$ with $Y^{\mathrm{H}}BY = I_k$ can be represented uniquely by $Y = X\widehat{Y}$ for some $\widehat{Y} \in \mathbb{C}^{\ell\times k}$ such that $\widehat{Y}^{\mathrm{H}}(X^{\mathrm{H}}BX)\widehat{Y} = I_k$. So (16) becomes

$$\min_{\substack{\mathcal{R}(Y)\subseteq\mathcal{X}\\ Y^{\mathrm{H}}BY=I_k}} \operatorname{trace}(Y^{\mathrm{H}}AY) = \min_{\widehat{Y}^{\mathrm{H}}(X^{\mathrm{H}}BX)\widehat{Y}=I_k} \operatorname{trace}(\widehat{Y}^{\mathrm{H}}(X^{\mathrm{H}}BX)\widehat{Y}) = \sum_{i=1}^k \hat{\lambda}_i.$$

This gives the second optimality of the Rayleigh-Ritz procedure.

The third optimality is concerned with the residual matrix

$$\mathscr{R}(A_1) := AX - BXA_1.$$

If $\mathscr{R}(A_1) = 0$, then $\mathcal{X}$ is an exact invariant subspace. So it would make sense to make $\|\mathscr{R}(A_1)\|$ as small as possible for certain matrix norm $\|\cdot\|$. The next theorem says the optimal A_1 is $X^{\mathrm{H}}AX$ when X is a B-orthonormal basis matrix of $\mathcal{X}$.

Theorem 2.4. *Suppose X has B-orthonormal columns, i.e., $X^{\mathrm{H}}BX = I_k$, and let $H = X^{\mathrm{H}}AX$. Then for any unitarily invariant norm*[1] $\|\cdot\|_{\mathrm{ui}}$

$$\|B^{-1/2}\mathscr{R}(H)\|_{\mathrm{ui}} \le \|B^{-1/2}\mathscr{R}(A_1)\|_{\mathrm{ui}} \quad \textit{for all } k\textit{-by-}k\ A_1. \tag{18}$$

[1] Two common used unitarily invariant norms are the spectral norm $\|\cdot\|_2$ and the Frobenius norm $\|\cdot\|_{\mathrm{F}}$. It is natural to think $\|B^{-1/2}(\,\cdot\,)\|_{\mathrm{ui}}$ as a B^{-1}-unitarily invariant norm induced by a given unitarily invariant norm $\|\cdot\|_{\mathrm{ui}}$. For example, the usual Frobenius norm can be defined by $\|C\|_{\mathrm{F}} := \sqrt{\operatorname{trace}(C^{\mathrm{H}}C)}$. Correspondingly, we may define the B^{-1}-Frobenius norm by $\|C\|_{B^{-1};\mathrm{F}} = \sqrt{\operatorname{trace}(C^{\mathrm{H}}B^{-1}C)}$.

2.3 Steepest Descent Methods

The basic idea of the steepest descent (SD) method to minimize a function value is to perform a line-search along the (opposite) direction of the gradient of the function at each iteration step. Our function is $\rho(x)$ defined by (7) whose gradient is given by

$$\nabla\rho(x) = \frac{2}{x^{\rm H}Bx}\,r(x), \tag{19}$$

where $r(x) := Ax - \rho(x)\,Bx$ is the residual of $(\rho(x), x)$ as an approximate eigenpair of $A - \lambda B$. Notice that $\nabla\rho(x)$ points to the same direction as $r(x)$. Therefore, given an approximation $\boldsymbol{x}$ to u_1 and $\|\boldsymbol{x}\|_B = 1$, one step of the steepest descent method for computing (λ_1, u_1) is simply to perform a line-search:

$$\inf_{t\in\mathbb{C}} \rho(\boldsymbol{x} + t\boldsymbol{r}), \tag{20}$$

where $\boldsymbol{r} = r(x)$. Since $\boldsymbol{x}^{\rm H}\boldsymbol{r} = 0$, $\boldsymbol{x}$ and $\boldsymbol{r}$ are linearly independent unless $\boldsymbol{r} = 0$ which implies $(\rho(\boldsymbol{x}), \boldsymbol{x})$ is already an exact eigenpair. An easy to use stopping criteria is to check if

$$\frac{\|r(\boldsymbol{x})\|_2}{\|A\boldsymbol{x}\|_2 + |\rho(\boldsymbol{x})|\,\|B\boldsymbol{x}\|_2} \le \texttt{rtol}, \tag{21}$$

where `rtol` is a given relative tolerance. When it is satisfied, $(\rho(\boldsymbol{x}), \boldsymbol{x})$ will be accepted as a computed eigenpair.

We have to solve the line-search (20). Since such a problem arises often in the conjugate gradient methods for $A - \lambda B$, we consider the following more general line-search:

$$\inf_{t\in\mathbb{C}} \rho(x + tp), \tag{22}$$

where the search direction p is the residual $r(x)$ in the SD method but will be different in the conjugate gradient method, for example. Suppose that x and p are linearly independent; otherwise $\rho(x + tp) \equiv \rho(x) = \rho(p)$. It is not difficult to show that

$$\inf_{t\in\mathbb{C}} \rho(x + tp) = \min_{|\xi|^2+|\zeta|^2>0} \rho(\xi x + \zeta p). \tag{23}$$

Therefore the infimum in (22) is the smaller eigenvalue μ of the 2×2 matrix pencil $X^{\rm H}AX - \lambda X^{\rm H}BX$, where $X = [x, p]$. Let $v = [\nu_1, \nu_2]^{\rm T}$ be the corresponding eigenvector. Then $\rho(Xv) = \mu$. Note $Xv = \nu_1 x + \nu_2 p$. We conclude

$$\operatorname*{arginf}_{t\in\mathbb{C}} \rho(x + tp) =: t_{\rm opt} = \begin{cases} \nu_2/\nu_1, & \text{if } \nu_1 \ne 0, \\ \infty, & \text{if } \nu_1 = 0. \end{cases} \tag{24}$$

Here $t_{\rm opt} = \infty$ should be interpreted in the sense of taking $t \to \infty$:

$$\lim_{t\to\infty} \rho(x + tp) = \rho(p).$$

Accordingly, we have

$$\rho(y) = \inf_{t\in\mathbb{C}} \rho(x + tp), \quad y = \begin{cases} x + t_{\rm opt}p & \text{if } t_{\rm opt} \text{ is finite}, \\ p & \text{otherwise}. \end{cases} \tag{25}$$

Now the simple SD method can be readily stated. We leave it to the reader.

This simple SD method can be slowly convergent in practice. This happens when the contours of $\rho(x)$ on the sphere $\{x : x^{\mathrm{H}}x = 1\}$ near the eigenvector u_1 is very flat: very long stretched in one or a few directions but very short compressed in other directions. So rarely, this plain version is used in practice, but rather as a starting point for designing faster variations of the method. In what follows, we will present three ideas some or all of which can be combined to improve the method in practice. The three ideas are

- extending the search space,
- preconditioning the search direction,
- introducing block implementation.

We now explain the three ideas in detail.

Extending the search space. The key step of the SD method is the line-search (20) which can be interpreted as seeking the best possible approximation $\boldsymbol{\rho}_{\text{new}}$:

$$\boldsymbol{\rho}_{\text{new}} = \min_{z \in \text{span}\{\boldsymbol{x}, \boldsymbol{r}\}} \rho(z) \tag{26}$$

to λ_1 through projecting $A - \lambda B$ to the 2-dimensional subspace spanned by

$$\boldsymbol{x}, \quad \boldsymbol{r} = A\boldsymbol{x} - \boldsymbol{\rho}B\boldsymbol{x} = (A - \boldsymbol{\rho}B)\boldsymbol{x},$$

where $\boldsymbol{\rho} = \rho(\boldsymbol{x})$. This subspace is in fact the 2nd order Krylov subspace $\mathcal{K}_2(A - \boldsymbol{\rho}B, \boldsymbol{x})$ of $A - \boldsymbol{\rho}B$ on $\boldsymbol{x}$. Naturally, a way to accelerate the simple SD method is to use a larger Krylov subspace, i.e., the mth order Krylov subspace $\mathcal{K}_m(A - \boldsymbol{\rho}B, \boldsymbol{x})$ which is spanned by

$$\boldsymbol{x}, (A - \boldsymbol{\rho}B)\boldsymbol{x}, \dots, (A - \boldsymbol{\rho}B)^{m-1}\boldsymbol{x}.$$

A better approximation to λ_1 is then obtained for $m \geq 3$ since now they are achieved by minimizing $\rho(x)$ over a larger subspace that contains $\text{span}\{\boldsymbol{x}, \boldsymbol{r}\} = \mathcal{K}_2(A - \boldsymbol{\rho}B, \boldsymbol{x})$:

$$\boldsymbol{\rho}_{\text{new}} = \min_{z \in \mathcal{K}_m(A - \boldsymbol{\rho}B, \boldsymbol{x})} \rho(z) \tag{27}$$

This leads to the *inverse free Krylov subspace method* of Golub and Ye [12] but we will call it the *extended steepest descent method* (ESD).

Theorem 2.5 ([12]). *Suppose λ_1 is simple, i.e., $\lambda_1 < \lambda_2$, and $\lambda_1 < \boldsymbol{\rho} < \lambda_2$. Let $\omega_1 < \omega_2 \leq \cdots \leq \omega_n$ be the eigenvalues of $A - \boldsymbol{\rho}B$ and v_1 be an eigenvector corresponding to ω_1, and let $\boldsymbol{\rho}_{\text{new}}$ be defined by* (27)*. Then*

$$\boldsymbol{\rho}_{\text{new}} - \lambda_1 \leq (\boldsymbol{\rho} - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho} - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + O(|\boldsymbol{\rho} - \lambda_1|^2), \tag{28}$$

where

$$\epsilon_m := \min_{f \in \mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)| \leq 2\left[\varDelta_\eta^{m-1} + \varDelta_\eta^{-(m-1)}\right]^{-1}, \tag{29}$$

$\eta = \frac{\omega_2 - \omega_1}{\omega_n - \omega_1}$ *and* $\varDelta_\eta = \frac{1+\sqrt{\eta}}{1-\sqrt{\eta}}$.

There are a few other existing results for $m = 2$ and $B = I$. Kantorovich and Akilov [21, p.617,1964] established

$$(\boldsymbol{\rho}_{\text{new}} - \lambda_1)/(\boldsymbol{\rho} - \lambda_1) \lessapprox \epsilon_m^2$$

for completely continuous operators. Knyazev and Skorokhodov [22, 1991] obtained something that is stronger in the sense that it is a strict inequality (i.e., without the need of ignoring high order terms). Samokish [45] presented an estimate on convergence rate for the preconditioned steepest descent method. Although his technique was for the case $B = I$, but can be made to work for the case $B \neq I$ after minor changes (see also [24, 37]). We omit stating them to limit the length of this paper.

Preconditioning the search direction. The idea of preconditioning a linear system $Ax = b$ to $KAx = Kb$ such that KA is "almost" the identity matrix before it is iteratively solved is quite natural. After all if $KA = I$, we would have $x = Kb$ immediately. Here that KA is "almost" the identity matrix is understood either $\|KA - I\|$ is relatively small or $KA - I$ is near a low rank matrix.

But there is no such an obvious and straightforward way to precondition the eigenvalue problem $Ax = \lambda Bx$. How could any direction be more favorable than the steepest descent one when it comes to minimize $\rho(x)$? After all, we are attempting to minimize the objective function $\rho(x)$.

In what follows, we shall offer two view points as to understand preconditioning an eigenvalue problem and how an effective preconditioner should be approximately constructed.

The first view point is more intuitive. The rationale lies as follows. It is well-known that when the contours of the objective function near its optimum are extremely elongated, at each step of the conventional steepest descent method, following the search direction which is the opposite of the gradient gets closer to the optimum on the line for a very short while and then starts to get away because the direction doesn't point "towards the optimum", resulting in a long zigzag path of a large number of steps. The ideal search direction p is therefore the one such that with its starting point at $\boldsymbol{x}$, p points to the optimum, i.e., the optimum is on the line $\{\boldsymbol{x} + tp \, : \, t \in \mathbb{C}\}$. Specifically, expand $\boldsymbol{x}$ as a linear combination of eigenvectors u_j

$$\boldsymbol{x} = \sum_{j=1}^{n} \alpha_j u_j =: \alpha_1 u_1 + \boldsymbol{v}, \quad \boldsymbol{v} = \sum_{j=2}^{n} \alpha_j u_j. \tag{30}$$

Then the ideal search direction is

$$p = \alpha u_1 + \beta \boldsymbol{v}$$

for some scalar α and $\beta \neq 0$ such that $\alpha_1 \beta - \alpha \neq 0$ (otherwise $p = \beta \boldsymbol{x}$). Of course, this is impractical because we don't know u_1 and $\boldsymbol{v}$. But we can construct one that is close to it. One such p is

$$p = (A - \sigma B)^{-1} \boldsymbol{r} = (A - \sigma B)^{-1}(A - \boldsymbol{\rho} B)\boldsymbol{x},$$

where[2] σ is some shift near λ_1 but not equal to $\boldsymbol{\rho}$. Let us analyze this p. By (6), we find

$$p = \sum_{j=1}^{n} \mu_j \alpha_j u_j, \quad \mu_j := \frac{\lambda_j - \boldsymbol{\rho}}{\lambda_j - \sigma}. \tag{31}$$

[2] We reasonably assume also $\sigma \neq \lambda_j$ for all other j, too.

Now if $\lambda_1 \le \boldsymbol{\rho} < \lambda_2$ and σ is also near λ_1 but not equal to $\boldsymbol{\rho}$ and if the gap $\lambda_2 - \lambda_1$ is reasonably modest, then

$$\mu_j \approx 1 \quad \text{for } j > 1$$

to give a $p \approx \alpha u_1 + \boldsymbol{v}$, resulting in fast convergence. This rough but intuitive analysis suggests that $(A - \sigma B)^{-1}$ with a suitably chosen shift σ can be used to serve as a good preconditioner. Qualitatively, we have

Theorem 2.6. *Let $\boldsymbol{x}$ be given by* (30)*, and suppose $\alpha_1 \neq 0$. If $\sigma \neq \boldsymbol{\rho}$ such that*

$$\textit{either } \mu_1 < \mu_j \textit{ for } 2 \le j \le n \textit{ or } \mu_1 > \mu_j \textit{ for } 2 \le j \le n, \tag{32}$$

where μ_j are defined in (31)*, then*

$$\tan\theta_B(u_1, \mathcal{K}_m) \le 2\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-1} \tan\theta_B(u_1, \boldsymbol{x}), \tag{33}$$

$$0 \le \boldsymbol{\rho}_{\text{new}} - \lambda_1 \le 4\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-2} \tan\theta_B(u_1, \boldsymbol{x}), \tag{34}$$

where $\mathcal{K}_m := \mathcal{K}_m([A - \sigma B]^{-1}(A - \boldsymbol{\rho} B), \boldsymbol{x})$, and

$$\eta = \begin{cases} \frac{\lambda_n - \sigma}{\lambda_n - \lambda_1} \cdot \frac{\lambda_2 - \lambda_1}{\lambda_2 - \sigma}, & \textit{if } \mu_1 < \mu_j \textit{ for } 2 \le j \le n, \\ \frac{\lambda_2 - \sigma}{\lambda_2 - \lambda_1} \cdot \frac{\lambda_n - \lambda_1}{\lambda_n - \sigma}, & \textit{if } \mu_1 > \mu_j \textit{ for } 2 \le j \le n, \end{cases} \qquad \Delta_\eta = \frac{1 + \sqrt{\eta}}{1 - \sqrt{\eta}}.$$

Proof. The proof is similar to the one in Saad [43] for the symmetric Lanczos method. □

The assumption (32) is one of the two criteria for selecting a shift σ, and the other is to make η close to 1. Three interesting cases are

- $\sigma < \lambda_1 \le \rho < \lambda_2$ under which μ_1 is smallest
- $\lambda_1 < \sigma < \rho < \lambda_2$ under which μ_1 is biggest
- $\lambda_1 < \rho < \sigma < \lambda_2$ under which μ_1 is smallest.

Often σ is selected as a lower bound of λ_1 as in the first case above, but it does not have to be. As for η, it is 1 for $\sigma = \lambda_1$, but since λ_1 is unknown, the best one can hope is to make $\sigma \approx \lambda_1$ through some kind of estimation.

In practice, because of high cost associated with $(A - \sigma B)^{-1}$, some forms of approximations to $(A - \sigma B)^{-1}$, such as those by incomplete decompositions LDL^{H} of $A - \sigma B$ or by iterative methods [9, 13, 14] CG, MINRES, or GMRES, are widely used.

The second view point is proposed by Golub and Ye [12], based on Theorem 2.5 which reveals that the rate of convergence depend on the distribution of the eigenvalues ω_j of $A - \boldsymbol{\rho} B$, not those of the pencil $A - \lambda B$ as in the Lanczos algorithm. In particular, if all $\omega_2 = \cdots = \omega_n$, then $\epsilon_m = 0$ for $m \ge 2$ and thus

$$\boldsymbol{\rho}_{\text{new}} - \lambda_1 = O(|\boldsymbol{\rho} - \lambda_1|^2),$$

suggesting quadratic convergence. Such an extreme case, though highly welcome, is unlikely to happen in practice, but it gives us an idea that if somehow we could transform an eigenvalue problem towards such an extreme case, the transformed problem would be easier to solve. Specifically we should seek equivalent transformations that change the eigenvalues of $A - \boldsymbol{\rho} B$ as much as possible to,

$$\boxed{\begin{array}{l}\text{one smallest isolated eigenvalue } \omega_1\text{, and the rest}\\ \omega_j \ (2 \le j \le n) \text{ tightly clustered,}\end{array}} \tag{35}$$

but leave those of $A - \lambda B$ unchanged. This goal is much as the one for preconditioning a linear system $Ax = b$ to $KAx = Kb$ for which a similar eigenvalue distribution for KA like (35) will result in swift convergence by most iterative methods.

We would like to equivalently transform the eigenvalue problem for $A - \lambda B$ to $L^{-\mathrm{H}}(A - \lambda B)L^{-1}$ by some nonsingular L (whose inverse or any linear system with L is easy to solve) so that the eigenvalues of $L^{-1}(A - \boldsymbol{\rho} B)L^{-\mathrm{H}}$ distribute more or less like (35). Then apply one step of ESD to the pencil $L^{-1}(A - \lambda B)L^{-\mathrm{H}}$ to find the next approximation $\boldsymbol{\rho}_{\text{new}}$. The process repeats.

Borrowed from the incomplete decomposition idea for preconditioning a linear system, such an L can be constructed using the LDL^{H} decomposition of $A - \boldsymbol{\rho} B$ [13, p.139] if the decomposition exists: $A - \boldsymbol{\rho} B = LDL^{\mathrm{H}}$, where L is lower triangular and $D = \operatorname{diag}(\pm 1)$. Then $L^{-1}(A - \boldsymbol{\rho} B)L^{-\mathrm{H}} = D$ has the ideal eigenvalue distribution that gives $\epsilon_m = 0$ for any $m \ge 2$. Unfortunately, this simple solution is impractical in practice for the following reasons:

1. The decomposition may not exist at all. In theory, the decomposition exists if all of the leading principle submatrices of $A - \boldsymbol{\rho} B$ are nonsingular.

2. If the decomposition does exist, it may not be numerically stable to compute, especially when $\boldsymbol{\rho}$ comes closer and closer to λ_1.

3. The sparsity in A and B is most likely destroyed, leaving L significantly denser than A and B combined. This makes all ensuing computations much more expensive.

A more practical solution is, however, through an incomplete LU factorization (see [44, Chapter 10]), to get

$$A - \boldsymbol{\rho} B \approx LDL^{\mathrm{H}},$$

where "$\approx$" includes not only the usual "approximately equal", but also the case when $(A - \boldsymbol{\rho} B) - LDL^{\mathrm{H}}$ is approximately a low rank matrix, and $D = \operatorname{diag}(\pm 1)$. Such an L changes from one step of the algorithm to another. In practice, often we may use one fixed preconditioner for all or several iterative steps. Using a constant preconditioner is certainly not optimal: it likely won't give the best rate of convergence per step and thus increases the number of total iterative steps but it can reduce overall cost because it saves work in preconditioner constructions and thus reduces cost per step. The basic idea of using a step-independent preconditioner is to find a σ that is close to λ_1, and perform an incomplete LDL^{H} decomposition of

$$A - \sigma B \approx LDL^{\mathrm{H}}$$

and transform $A - \lambda B$ accordingly before applying SD or ESD. Now the rate of convergence is determined by the eigenvalues of

$$\widehat{C} = L^{-1}(A - \sigma B)L^{-\mathrm{H}} + (\sigma - \boldsymbol{\rho})L^{-1}BL^{-\mathrm{H}} \approx D$$

which would have a better spectral distribution so long as $(\sigma - \boldsymbol{\rho})L^{-1}BL^{-\mathrm{H}}$ is small relative to $\widehat{C}$. When $\sigma < \lambda_1$, $A - \sigma B \succ 0$ and the incomplete LDL^{H} factorization becomes incomplete Cholesky factorization.

We have insisted so far about applying SD or ESD straightforwardly to the transformed problem. There is another way, perhaps, better: only symbolically applying SD or ESD to the transformed problem as a derivation stage for a preconditioned method that always projects the original pencil $A - \lambda B$ directly every step. The only difference is now the projecting subspaces are preconditioned.

Suppose $A - \lambda B$ is transformed to $\widehat{A} - \lambda\widehat{B} := L^{-1}(A - \lambda B)L^{-\mathrm{H}}$. Consider a typical step of ESD applied to $\widehat{A} - \lambda\widehat{B}$. For the purpose of distinguishing notational symbols, we will put hats on all those for $\widehat{A} - \lambda\widehat{B}$. The typical step of ESD is

$$\boxed{\begin{array}{l}\text{compute the smallest eigenvalue } \mu \text{ and corresponding eigen-}\\ \text{vector } v \text{ of } \widehat{Z}^{\mathrm{H}}(\widehat{A} - \lambda\widehat{B})\widehat{Z}, \text{ where } \widehat{Z} \in \mathbb{C}^{n\times m} \text{ is a basis matrix}\\ \text{of Krylov subspace } \mathcal{K}_m(\widehat{A} - \hat{\boldsymbol{\rho}}\widehat{B}, \hat{\boldsymbol{x}}).\end{array}} \tag{36}$$

Notice $\left[\widehat{A} - \hat{\boldsymbol{\rho}}\widehat{B}\right]^j \hat{\boldsymbol{x}} = L^{\mathrm{H}}\left[(LL^{\mathrm{H}})^{-1}(A - \hat{\boldsymbol{\rho}}B)\right]^j (L^{-\mathrm{H}}\hat{\boldsymbol{x}})$ to see

$$L^{-\mathrm{H}} \cdot \mathcal{K}_m(\widehat{A} - \hat{\boldsymbol{\rho}}\widehat{B}, \hat{\boldsymbol{x}}) = \mathcal{K}_m(\,K(A - \hat{\boldsymbol{\rho}}B), \boldsymbol{x}),$$

where $\boldsymbol{x} = L^{-\mathrm{H}}\hat{\boldsymbol{x}}$ and $K = (LL^{\mathrm{H}})^{-1}$. So $Z = L^{-\mathrm{H}}\widehat{Z}$ is a basis matrix of Krylov subspace $\mathcal{K}_m(\,K(A - \hat{\boldsymbol{\rho}}B), \boldsymbol{x})$. Since also

$$\widehat{Z}^{\mathrm{H}}(\widehat{A} - \lambda\widehat{B})\widehat{Z} = (L^{-\mathrm{H}}\widehat{Z})^{\mathrm{H}}(A - \lambda B)(L^{-\mathrm{H}}\widehat{Z}),$$
$$\hat{\boldsymbol{\rho}} = \frac{\hat{\boldsymbol{x}}^{\mathrm{H}}\widehat{A}\hat{\boldsymbol{x}}}{\hat{\boldsymbol{x}}^{\mathrm{H}}\widehat{B}\hat{\boldsymbol{x}}} = \frac{\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}B\boldsymbol{x}} = \boldsymbol{\rho},$$

the typical step (36) can be reformulated equivalently to

$$\boxed{\begin{array}{l}\text{compute the smallest eigenvalue } \mu \text{ and corresponding eigen-}\\ \text{vector } v \text{ of } Z^{\mathrm{H}}(A - \lambda B)Z, \text{ where } Z \in \mathbb{C}^{n\times m} \text{ is a ba-}\\ \text{sis matrix of Krylov subspace } \mathcal{K}_m(\,K(A - \boldsymbol{\rho}B), \boldsymbol{x}), \text{ where}\\ K = (LL^{\mathrm{H}})^{-1}.\end{array}} \tag{37}$$

Introducing block implementation. The convergence rate of ESD with a preconditioner $K \succ 0$ is determined by the eigenvalues $\omega_1 < \omega_2 \le \cdots \le \omega_n$ of $K^{1/2}(A - \boldsymbol{\rho}B)K^{1/2}$ can still be very slow if λ_2 is very close to λ_1 relative to λ_n in which case $\omega_1 \approx \omega_2$.

Often in practice, there are needs to compute the first few eigenpairs, not just the first one. For that purpose, block variations of the methods become particularly attractive for at least the following reasons:

1. they can simultaneously compute the first k eigenpairs (λ_j, u_j);
2. they run more efficiently on modern computer architecture because more computations can be organized into matrix-matrix multiplication type;
3. they have better rates of convergence to the desired eigenpairs and save overall cost by using a block size that is slightly bigger than the number of asked eigenpairs.

In summary, the benefits of using a block variation are similar to those of using the simultaneous subspace iteration *vs.* the power method [46].

A block variation starts with $\boldsymbol{X} \in \mathbb{C}^{n\times n_b}$ with $\operatorname{rank}(\boldsymbol{X}) = n_b$, instead of just one vector $\boldsymbol{x} \in \mathbb{C}^n$ previously for the single-vector steepest descent methods. Here either the jth column of $\boldsymbol{X}$ is already an approximation to u_j or the subspace $\mathcal{R}(\boldsymbol{X})$ is a good approximation to the generalized invariant subspace spanned by u_j for $1 \le j \le n_b$ or the canonical angles from $\mathcal{R}([u_1, \dots, u_k])$ to $\mathcal{R}(\boldsymbol{X})$ are nontrivial, where $k \le n_b$ is the number of desired eigenpairs. In the latter two cases, a preprocessing is needed to turn the case into the first case:

1. solve the eigenvalue problem $\boldsymbol{X}^{\mathrm{H}}(A - \lambda B)\boldsymbol{X}$ to get $(\boldsymbol{X}^{\mathrm{H}}A\boldsymbol{X})\,W = (\boldsymbol{X}^{\mathrm{H}}B\boldsymbol{X})\,W\varOmega$, where $\varOmega = \operatorname{diag}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_{n_b})$ is the diagonal matrix of eigenvalues in ascending order, and W is the eigenvector matrix;

2. reset $\boldsymbol{X} := \boldsymbol{X}W$.

So we will assume henceforth the jth column of the given $\boldsymbol{X}$ is an approximation to u_j. Now consider generalizing the steepest descent method to a block one. Its typical iterative step may well look like the following. Let

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_{n_b}] \in \mathbb{C}^{n\times n_b}$$

whose jth column $\boldsymbol{x}_j$ approximates u_j and

$$\varOmega = \operatorname{diag}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_{n_b})$$

whose jth diagonal entry $\boldsymbol{\rho}_j = \rho(\boldsymbol{x}_j)$ approximates λ_j. We may well assume $\boldsymbol{X}$ has B-orthonormal columns, i.e., $\boldsymbol{X}^{\mathrm{H}}B\boldsymbol{X} = I$. Define the residual matrix

$$\boldsymbol{R} = [r(\boldsymbol{x}_1), r(\boldsymbol{x}_2), \dots, r(\boldsymbol{x}_{n_b})] = A\boldsymbol{X} - B\boldsymbol{X}\varOmega.$$

The key iterative step of the block steepest descent (BSD) method for computing the next set of approximations is as follows:

1. compute a basis matrix Z of $\mathcal{R}([\boldsymbol{X}, \boldsymbol{R}])$ by, e.g., MGS in the B-inner product, keeping in mind that $\boldsymbol{X}$ has B-orthonormal columns already;

2. find the first n_b eigenpairs of $Z^{\mathrm{H}}AZ - \lambda Z^{\mathrm{H}}BZ$ by, e.g., one of LAPACK's subroutines [1, p.25] because of its small scale, to get $(Z^{\mathrm{H}}AZ)W = (Z^{\mathrm{H}}BZ)W\varOmega_{\mathrm{new}}$, where $\varOmega_{\mathrm{new}} = \operatorname{diag}(\boldsymbol{\rho}_{\mathrm{new};1}, \boldsymbol{\rho}_{\mathrm{new};2}, \dots, \boldsymbol{\rho}_{\mathrm{new};n_b})$;

3. set $\boldsymbol{X}_{\mathrm{new}} = ZW$.

This is in fact the stronger version of ***Simultaneous Rayleigh Quotient Minimization Method***, called SIRQIT-G2, in Longsine and McCormick [30]. To introduce the block extended steepest descent (BESD) method, we notice that $r(\boldsymbol{x}_j) = (A - \boldsymbol{\rho}_j B)\boldsymbol{x}_j$ and thus

$$\begin{aligned}\mathcal{R}([\boldsymbol{X}, \boldsymbol{R}]) &= \sum_{j=1}^{n_b} \mathcal{R}([\boldsymbol{x}_j, (A - \boldsymbol{\rho}_j B)\boldsymbol{x}_j]) \\ &= \sum_{j=1}^{n_b} \mathcal{K}_2(A - \boldsymbol{\rho}_j B, \boldsymbol{x}_j).\end{aligned}$$

BESD is simply to extend each Krylov subspace $\mathcal{K}_2(A-\boldsymbol{\rho}_jB,\boldsymbol{x}_j)$ to a high order one, and of course different Krylov subspaces can be expanded to different orders. For simplicity, we will expand each to the mth order. The new extended search subspace now is

$$\sum_{j=1}^{n_b}\mathcal{K}_m(A-\boldsymbol{\rho}_jB,\boldsymbol{x}_j). \tag{38}$$

Define the linear operator

$$\mathscr{R}\,:\,X\in\mathbb{C}^{n\times n_b}\to\mathscr{R}(X)=AX-BX\Omega\in\mathbb{C}^{n\times n_b}.$$

Then the subspace in (38) can be compactly written as

$$\mathcal{K}_m(\mathscr{R},X)=\operatorname{span}\{\boldsymbol{X},\mathscr{R}(\boldsymbol{X}),\ldots,\mathscr{R}^{m-1}(\boldsymbol{X})\}, \tag{39}$$

where $\mathscr{R}^i(\,\cdot\,)$ is understood as successively applying the operator $\mathscr{R}$ i times, e.g., $\mathscr{R}^2(X)=\mathscr{R}(\mathscr{R}(X))$.

As to incorporate suitable preconditioners, in light of our extensive discussions before, the search subspace should be modified to

$$\sum_{j=1}^{n_b}\mathcal{K}_m(K_j(A-\boldsymbol{\rho}_jB),\boldsymbol{x}_j), \tag{40}$$

where K_j are the preconditioners, one for each approximate eigenpair $(\boldsymbol{\rho}_j,\boldsymbol{x}_j)$ for $1\le j\le n_b$. As before, K_j can be constructed in one of the following two ways:

- K_j is an approximate inverse of $A-\tilde{\boldsymbol{\rho}}_jB$ for some $\tilde{\boldsymbol{\rho}}_j$ different from $\boldsymbol{\rho}_j$, ideally closer to λ_j than to any other eigenvalue of $A-\lambda B$. But this requirement on $\tilde{\boldsymbol{\rho}}_j$ is impractical because the eigenvalues of $A-\lambda B$ are unknown. A compromise would be to make $\tilde{\boldsymbol{\rho}}_j$ close but not equal to $\boldsymbol{\rho}_j$ than to any other $\boldsymbol{\rho}_j$.
- Perform an incomplete LDL^{H} factorization (see [44, Chapter 10]) $A-\rho_jB\approx L_jD_jL_j^{\mathrm{H}}$, where "$\approx$" includes not only the usual "approximately equal", but also the case when $(A-\boldsymbol{\rho}_jB)-L_jD_jL_j^{\mathrm{H}}$ is approximately a low rank matrix, and $D_j=\operatorname{diag}(\pm1)$. Finally set $K_j=L_jL_j^{\mathrm{H}}$.

Algorithm 2.2 is the general framework of a Block Preconditioned Extended Steepest Descent method (BPESD) which embeds many steepest descent methods into one. In particular,

1. With $n_b=1$, it gives various single-vector steepest descent methods:
    - Steepest Descent method (SD): $m=2$ and all preconditioners $K_{\ell;j}=I$;
    - Preconditioned Steepest Descent method (PSD): $m=2$;
    - Extended Steepest Descent method (ESD): all preconditioners $K_{\ell;j}=I$;
    - Preconditioned Extended Steepest Descent method (PESD).
2. With $n_b>1$, various block steepest descent methods are born:
    - Block Steepest Descent method (BSD): $m=2$ and all preconditioners $K_{\ell;j}=I$;
    - Block Preconditioned Steepest Descent method (BPSD): $m=2$;

Algorithm 2.2 Extended Block Preconditioned Steepest Descent method

Given an initial approximation $X_0 \in \mathbb{C}^{n\times n_b}$ with $\operatorname{rank}(X_0) = n_b$, and an integer $m \geq 2$, the algorithm attempts to compute approximate eigenpairs to (λ_j, u_j) for $1 \leq j \leq n_b$.

```
1: compute the eigen-decomposition: $(X_0^{\rm H} A X_0)W = (X_0^{\rm H} B X_0)W\Omega_0$,
   where $W^{\rm H}(X_0^{\rm H} B X_0)W = I$, $\Omega_0 = \operatorname{diag}(\rho_{0;1}, \rho_{0;2}, \ldots, \rho_{0;n_b})$;
2: $X_0 \equiv [x_{0;1}, x_{0;2}, \ldots, x_{0;n_b}] = X_0 W$;
3: for $\ell = 0, 1, \ldots$ do
4:    test convergence and lock up the converged (detail to come later);
5:    construct preconditioners $K_{\ell;j}$ for $1 \leq j \leq n_b$;
6:    compute a basis matrix $Z$ of the subspace (40) with $\boldsymbol{\rho}_j = \rho_{\ell;j}$ and $\boldsymbol{x}_j = x_{\ell+1;j}$;
7:    compute the $n_b$ smallest eigenvalues and corresponding eigenvectors of $Z^{\rm H}(A - \lambda B)Z$ to get $(Z^{\rm H} A Z)W = (Z^{\rm H} B Z)W\Omega_\ell$, where $W^{\rm H}(Z^{\rm H} B Z)W = I$, $\Omega_{\ell+1} = \operatorname{diag}(\rho_{\ell+1;1}, \rho_{\ell+1;2}, \ldots, \rho_{\ell+1;n_b})$;
8:    $X_{\ell+1} \equiv [x_{\ell+1;1}, x_{\ell+1;2}, \ldots, x_{\ell+1;n_b}] = ZW$;
9: end for
10: return approximate eigenpairs to $(\lambda_j, u_j)$ for $1 \leq j \leq n_b$.
```

- Block Extended Steepest Descent method (BESD): all preconditioners $K_{\ell;j} = I$;
- Block Preconditioned Extended Steepest Descent method (BPESD).

This framework is essentially the one implied in [40, section 4].

There are four important implementation issues to worry about in turning Algorithm 2.2 into a piece of working code.

1. In (40), a different preconditioner is used for each and every approximate eigenpair $(\rho_{\ell;j}, x_{\ell;j})$ for $1 \leq j \leq n_b$. While, conceivably, doing so will speed up convergence for each approximate eigenpair because each preconditioner can be constructed to make that approximate eigenpair converge faster, but the cost in constructing these preconditioners may likely be too heavy to bear. A more practical approach would be to use one preconditioner K_ℓ for all $K_{\ell;j}$ aiming at speeding up the convergence of $(\rho_{\ell;1}, x_{\ell;1})$ (or the first few approximate eigenpairs for tightly clustered eigenvalues). Once it (or the first few in the case of a tight cluster) is determined to be sufficiently accurate, the converged eigenpairs are locked up and deflated and a new preconditioner is computed to aim at the next non-converged eigenpairs, and the process continues. We will come back to discuss the deflation issue, i.e., Line 4 of Algorithm 2.2.

2. Consider implementing Line 6, i.e., generating a basis matrix for the subspace (40). In the most general case, Z can be gotten by packing the basis matrices of all

$$\mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j}) \quad \text{for } 1 \leq j \leq n_b$$

together. There could be two problems with this: 1) such Z could be ill-conditioned, i.e., the columns of Z may not be sufficiently numerically linearly independent, and 2) the arithmetic operations in building a basis for each $\mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j})$ are mostly matrix-vector multiplications, straying from one of the purposes: performing most arithmetic operations through matrix-matrix multiplications in order to achieve high performance on modern computers. To address these two problems, we do a tradeoff of using $K_{\ell;j} \equiv K_\ell$ for all j. This may likely degrade the effectiveness of the preconditioner per step in terms of rate of convergence for all approximate eigenpairs $(\rho_{\ell;j}, x_{\ell;j})$ but may achieve overall gain in using less time because then the code will run much faster in matrix-matrix operations, not to mention the saving in constructing just one preconditioner K_ℓ instead of n_b different ones. To simplify our discussion below, we will drop the subscript ℓ for readability. Since $K_{\ell;j} \equiv K$ for all j, (40) is the same as

$$\mathcal{K}_m(K\mathscr{R}, X) = \operatorname{span}\{X, K\mathscr{R}(X), \ldots, [K\mathscr{R}]^{m-1}(X)\}, \tag{41}$$

where $[K\mathscr{R}]^i(\cdot)$ is understood as successively applying the operator $K\mathscr{R}$ i times, e.g., $[K\mathscr{R}]^2(X) = K\mathscr{R}_\ell(K\mathscr{R}(X))$. A basis matrix

$$Z = [Z_1, Z_2, \ldots, Z_m]$$

can be computed by the following block Arnoldi-like process in the B-inner product [40, Algorithm 5].

```
1: $Z_1 = X$ (recall $X^{\mathrm{H}}BX = I_{n_b}$ already);
2: for $i = 2$ to $m$ do
3:    $Y = K(AZ_{i-1} - B\Omega Z_{i-1})$;
4:    for $j = 1$ to $i - 1$ do
5:       $T = Z_j^{\mathrm{H}}BY$, $Y = Y - Z_jT$;
6:    end for
7:    $Z_iT = Y$ (MGS in the $B$-inner product);
8: end for
```

(42)

There is a possibility that at Line 7 of (42), Y is numerically not of full column rank. If it happens, it poses no difficulty at all. In running MGS on Y's columns, anytime if a column is deemed linearly dependent on previous columns, that column should be deleted, along with the corresponding ρ_j from Ω to shrink its size by 1 as well. At the completion of MGS, Z_i will have fewer columns than Y and the size of Ω is shrunk accordingly. Finally, at the end, the columns of Z are B-orthonormal, i.e., $Z^{\mathrm{H}}BZ = I$ (of apt size) which may fail to an unacceptably level due to roundoff; so some form of re-orthogonalization should be incorporated.

4. At Line 4, a test for convergence are required. The same criteria (21) can be used: $(\rho_{\ell;j}, x_{\ell;j})$ is considered acceptable if

$$\frac{\|r_{\ell;j}\|_2}{\|Ax_{\ell;j}\|_2 + |\rho_{\ell;j}|\,\|Bx_{\ell;j}\|_2} \le \texttt{rtol}$$

where `rtol` is a pre-set relative tolerance. Usually the eigenvalues λ_j are converged to in order, i.e., the smallest eigenvalues emerge first. All acceptable approximate eigenpairs should be locked in, say, a $k_{\mathrm{cvgd}} \times k_{\mathrm{cvgd}}$ diagonal matrix[3] $\boldsymbol{D}$ for converged eigenvalues and an $n \times k_{\mathrm{cvgd}}$ tall matrix $\boldsymbol{U}$ for eigenvectors such that

$$A\boldsymbol{U} \approx B\boldsymbol{U}\boldsymbol{D}, \quad \boldsymbol{U}^{\mathrm{H}}B\boldsymbol{U} \approx I$$

to an acceptable level of accuracy. Every time a converged eigenpair is detected, delete the converged $\rho_{\ell;j}$ and $x_{\ell;j}$ from Ω_ℓ and X_ℓ, respectively, and expand $\boldsymbol{D}$ and $\boldsymbol{U}$ to lock up the pair, accordingly. At the same time, either reduce n_b by 1 or append a (possibly random) B-orthogonal column to X to maintain n_b unchanged. There are two different ways to avoid recomputing any of the converged eigenpairs – a process called **deflation**.

[3] In actual programming code, it is likely an 1-D array. But we use a diagonal matrix for the sake of presentation.

1. At Line 7 in the above block Arnoldi-like process (42), each column of Z_{j+1} is B-orthogonalized against $\boldsymbol{U}$.

2. Modify $A - \lambda B$ in form, but not explicitly, to $(A + \zeta B\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}}B) - \lambda B$, where ζ is a real number intended to move λ_j for $1 \le j \le k_{\mathrm{cvgd}}$ to $\lambda_j + \zeta$; so it should be selected such that $\zeta + \lambda_1 \ge \lambda_{k_{\mathrm{cvgd}}+n_b+1}$.

But if there is a good way to pick a ζ such that $\zeta + \lambda_1 \ge \lambda_{k_{\mathrm{cvgd}}+n_b+1}$, the second approach is easier to use in implementation than the first one for which, if not carefully implemented, rounding errors can make $\mathcal{R}(Z)$ crawl into $\mathcal{R}(\boldsymbol{U})$ unnoticed.

2.4 Locally Optimal CG Methods

As is well-known, the slow convergence of the plain steepest descent method is due to the extreme flat contours of the objective function near (sometimes local) optimal points. The nonlinear conjugate gradient method is another way, besides preconditioning technique, to move the searching direction away from the steepest descent direction. Originally, the conjugate gradient (CG) method was invented in 1950s by Hestenes and Stiefel [17, 1952] for solving linear system $Hx = b$ with Hermitian and positive definite H, and later was interpreted as an iterative method for large scale linear systems. This is so-called the *linear CG* method [9, 13, 35]. In the 1960s, it was extended by Fletcher and Reeves [11, 1964] as an iterative method for solving nonlinear optimization problems (see also [35, 48]). We shall call the resulting method the *nonlinear CG* method. Often we leave out the word "linear" and "nonlinear" and simply call either method the CG method when no confusion can arise from this.

Because of the optimality properties (10) of the Rayleigh quotient $\rho(x)$, it is natural to apply the nonlinear CG method to compute the first eigenpair and, with the aid of deflation, the first few eigenpairs of $A - \lambda B$. The article [7, 1966] by Bradbury and Fletcher seems to be the first one to do just that.

However, it is suggested [23] that the local optimal CG (LOCG) method [39, 49] is more suitable for the symmetric eigenvalue problem. In its simplest form, LOCG for our eigenvalue problem $A - \lambda B$ is obtained by simply modifying the line-search (26) for the SD method to

$$\boldsymbol{\rho}_{\mathrm{new}} = \min_{x \in \mathrm{span}\{\boldsymbol{x}, \boldsymbol{x}_{\mathrm{old}}, \boldsymbol{r}\}} \rho(x), \tag{43}$$

where $\boldsymbol{x}_{\mathrm{old}}$ is the approximate eigenvector to u_1 from the previous iterative step.

The three ideas we explained in the previous subsection to improve the plain SD method can be introduced to improve the approximation given by (43), too, upon noticing the search space in (43) is

$$\mathcal{K}_2(A - \boldsymbol{\rho}B, \boldsymbol{x}) + \mathcal{R}(\boldsymbol{x}_{\mathrm{old}}),$$

making it possible for us to 1) extend the search space, 2) precondition the search direction $\boldsymbol{r}$, and 3) introduce block implementation, in the same way as we did for the plain SD method.

All things considered, we now present an algorithmic framework: Algorithm 2.3, *Locally Optimal Block Preconditioned Extended Conjugate Gradient method* (LOBPECG) which has implementation choices:

- block size n_b;

- preconditioners varying with iterative steps, with approximate eigenpairs, or not;

- the dimension m of Krylov subspaces in extending the search subspace at each iterative step. It may also vary with iterative steps, too.

Algorithm 2.3 Locally Optimal Block Preconditioned Extended Conjugate Gradient method (LOBPECG)

Given an initial approximation $X_0 \in \mathbb{C}^{n\times n_b}$ with $\operatorname{rank}(X_0) = n_b$, and an integer $m \geq 2$, the algorithm attempts to compute approximate eigenpairs to (λ_j, u_j) for $1 \leq j \leq n_b$.

```
 1: compute the eigen-decomposition: $(X_0^{\mathrm{H}} A X_0)W = (X_0^{\mathrm{H}} B X_0)W\Omega_0$, where
    $W^{\mathrm{H}}(X_0^{\mathrm{H}} B X_0)W = I$, $\Omega_0 = \operatorname{diag}(\rho_{0;1}, \rho_{0;2}, \ldots, \rho_{0;n_b})$;
 2: $X_0 \equiv [x_{0;1}, x_{0;2}, \ldots, x_{0;n_b}] = X_0 W$, $X_{-1} = 0$;
 3: for $\ell = 0, 1, \ldots$ do
 4:    test convergence and lock up the converged;
 5:    construct preconditioners $K_{\ell;j}$ for $1 \leq j \leq n_b$;
 6:    compute a basis matrix $Z$ of the subspace
           $$\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j}) + \mathcal{R}(X_{\ell-1}); \qquad (44)$$
 7:    compute the $n_b$ smallest eigenvalues and corresponding eigenvectors of $Z^{\mathrm{H}}(A - \lambda B)Z$ to get $(Z^{\mathrm{H}} A Z)W = (Z^{\mathrm{H}} B Z)W\Omega_\ell$, where $W^{\mathrm{H}}(Z^{\mathrm{H}} B Z)W = I$, $\Omega_{\ell+1} = \operatorname{diag}(\rho_{\ell+1;1}, \rho_{\ell+1;2}, \ldots, \rho_{\ell+1;n_b})$;
 8:    $X_{\ell+1} \equiv [x_{\ell+1;1}, x_{\ell+1;2}, \ldots, x_{\ell+1;n_b}] = ZW$;
 9: end for
10: return approximate eigenpairs to $(\lambda_j, u_j)$ for $1 \leq j \leq n_b$.
```

The four important implementation issues we discussed for Algorithm 2.2 (BPESD) after its introduction essentially apply here, except some changes are needed in the computation of Z at Line 6 of Algorithm 2.3.

First $X_{\ell-1}$ can be replaced by something else while the subspace (44) remains the same. Specifically, we modify Lines 2, 6, and 8 of Algorithm 2.3 to

```
2: $X_0 \equiv [x_{0;1}, x_{0;2}, \ldots, x_{0;n_b}] = X_0 W$, and $Y_0 = 0$;
6:     compute a basis matrix $Z$ of the subspace
           $$\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j}) + \mathcal{R}(Y_\ell) \qquad (45)$$
       such that $\mathcal{R}(Z_{(:,1:n_b)}) = \mathcal{R}(X_\ell)$, and let $n_Z$ be the number of the columns of $Z$;
8:     $X_{\ell+1} \equiv [x_{\ell+1;1}, x_{\ell+1;2}, \ldots, x_{\ell+1;n_b}] = ZW$, $Y_{\ell+1} = Z_{(:,n_b+1:n_Z)} W_{(n_b+1:n_Z,:)}$;
```

This idea is basically the same as the one in [18, 23]. Next we will compute a basis matrix for the subspace (45) (or (44)). For better performance (by using more matrix-matrix multiplications), we will assume $K_{\ell;j} \equiv K_\ell$ for all j for simplification. Dropping the subscript ℓ for readability, we see (45) is the same as

$$\mathcal{K}_m(K\mathscr{R}, X) + \mathcal{R}(Y) = \operatorname{span}\{X, K\mathscr{R}(X), \ldots, [K\mathscr{R}]^{m-1}(X)\} + \mathcal{R}(Y). \tag{46}$$

We will first compute a basis matrix $[Z_1, Z_2, \ldots, Z_m]$ for $\mathcal{K}_m(K\mathscr{R}, X)$ by the Block Arnoldi-like process in the B-inner product (42). In particular, $Z_1 = X$. Then B-orthogonalize

Y against $[Z_1, Z_2, \dots, Z_m]$ to get Z_{m+1} satisfying $Z_{m+1}^{\mathrm{H}} B Z_{m+1} = I$. Finally take $Z = [Z_1, Z_2, \dots, Z_{m+1}]$.

3 Min-Max Principles for a Positive Semidefinite Pencil

Let $A - \lambda B$ be an $n \times n$ positive semidefinite pencil, i.e., A and B are Hermitian and there is a real scalar λ_0 such that $A - \lambda_0 B$ is positive semidefinite. Note that this does not demand anything on the regularity of $A - \lambda B$, i.e., a positive semidefinite matrix pencil can be either regular (meaning $\det(A - \lambda B) \not\equiv 0$) or singular (meaning $\det(A - \lambda B) \equiv 0$ for all $\lambda \in \mathbb{C}$).

Let the integer triplet (n_-, n_0, n_+) be the *inertia* of B, meaning B has n_- negative, n_0 0, and n_+ positive eigenvalues, respectively. Necessarily

$$r := \operatorname{rank}(B) = n_+ + n_-. \tag{47}$$

We say $\mu \neq \infty$ is a *finite eigenvalue* of $A - \lambda B$ if

$$\operatorname{rank}(A - \mu B) < \max_{\lambda \in \mathbb{C}} \operatorname{rank}(A - \lambda B), \tag{48}$$

and $x \in \mathbb{C}^n$ is a corresponding *eigenvector* if $0 \neq x \notin \mathcal{N}(A) \cap \mathcal{N}(B)$ satisfies

$$Ax = \mu Bx, \tag{49}$$

or equivalently, $0 \neq x \in \mathcal{N}(A - \mu B) \backslash (\mathcal{N}(A) \cap \mathcal{N}(B))$. Let k_+ and k_- be two nonnegative integers such that $k_+ \leq n_+$, $k_- \leq n_-$, and $k_+ + k_- \geq 1$, and set

$$J_k = \begin{bmatrix} I_{k_+} & \\ & -I_{k_-} \end{bmatrix} \in \mathbb{C}^{k \times k}, \quad k = k_+ + k_-. \tag{50}$$

Theorem 3.1 ([29])**.** *If $A - \lambda B$ is positive semidefinite, then $A - \lambda B$ has $r = \operatorname{rank}(B)$ finite eigenvalues all of which are real.*

In what follows, if $A - \lambda B$ is positive semidefinite, we will denote its finite eigenvalues by $\lambda_i^{\pm}$ arranged in the order:

$$\lambda_{n_-}^- \leq \cdots \leq \lambda_1^- \leq \lambda_1^+ \leq \cdots \leq \lambda_{n_+}^+. \tag{51}$$

For the case of a regular Hermitian pencil $A - \lambda B$ (i.e., $\det(A - \lambda B) \not\equiv 0$), Theorem 3.2 is a special case of the ones considered in [6, 34]. For a diagonalizable positive semidefinite Hermitian pencil $A - \lambda B$ with nonsingular B, Theorem 3.2 was implied in [26, 53]. Recall that a positive semidefinite Hermitian pencil $A - \lambda B$ can be possibly a singular pencil; so the condition of Theorem 3.2 does not exclude a singular pencil $A - \lambda B$ which was not considered before [29, 2013], not to mention that B may possibly be singular.

Theorem 3.2. *Let $A - \lambda B$ be a positive semidefinite Hermitian pencil. Then for $1 \leq i \leq n_+$*

$$\lambda_i^+ = \sup_{\substack{\mathcal{X} \\ \operatorname{codim} \mathcal{X} = i-1}} \inf_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x = 1}} x^{\mathrm{H}} A x = \sup_{\substack{\mathcal{X} \\ \operatorname{codim} \mathcal{X} = i-1}} \inf_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x > 0}} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}, \tag{52a}$$

$$\lambda_i^+ = \inf_{\substack{\mathcal{X} \\ \dim \mathcal{X} = i}} \sup_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x = 1}} x^{\mathrm{H}} A x = \inf_{\substack{\mathcal{X} \\ \dim \mathcal{X} = i}} \sup_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x > 0}} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}, \tag{52b}$$

and for $1 \le i \le n_-$,

$$\lambda_i^- = -\sup_{\substack{\mathcal{X}\\ \operatorname{codim}\mathcal{X}=i-1}} \inf_{\substack{x\in\mathcal{X}\\ x^{\mathrm{H}}Bx=-1}} x^{\mathrm{H}}Ax = \inf_{\substack{\mathcal{X}\\ \operatorname{codim}\mathcal{X}=i-1}} \sup_{\substack{x\in\mathcal{X}\\ x^{\mathrm{H}}Bx<0}} \frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}Bx}, \tag{52c}$$

$$\lambda_i^- = -\inf_{\substack{\mathcal{X}\\ \dim\mathcal{X}=i}} \sup_{\substack{x\in\mathcal{X}\\ x^{\mathrm{H}}Bx=-1}} x^{\mathrm{H}}Ax = \sup_{\substack{\mathcal{X}\\ \dim\mathcal{X}=i}} \inf_{\substack{x\in\mathcal{X}\\ x^{\mathrm{H}}Bx<0}} \frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}Bx}. \tag{52d}$$

In particular, setting $i = 1$ *in* (52) *gives*

$$\lambda_1^+ = \inf_{x^{\mathrm{H}}Bx>0} \frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}Bx}, \quad \lambda_1^- = \sup_{x^{\mathrm{H}}Bx<0} \frac{x^{\mathrm{H}}Ax}{x^{\mathrm{H}}Bx}. \tag{53}$$

*All "*inf*" and "*sup*" can be replaced by "*min*" and "*max*" if* $A - \lambda B$ *is positive definite or positive semidefinite but diagonalizable*[4]

The following theorem for the case when B is also nonsingular is due to Kovač-Striko and Veselić [25, 1995]. But in this general form, it is due to [29].

Theorem 3.3 ([29])**.** *Let* $A - \lambda B$ *be a Hermitian pencil of order* n

1. *Suppose* $A-\lambda B$ *is positive semidefinite. Let* $X \in \mathbb{C}^{n\times k}$ *satisfying* $X^{\mathrm{H}}BX = J_k$*, and denote by* $\mu_i^\pm$ *the eigenvalues of* $X^{\mathrm{H}}AX - \lambda X^{\mathrm{H}}BX$ *arranged in the order:*

$$\mu_{k_-}^- \le \cdots \le \mu_1^- \le \mu_1^+ \le \cdots \le \mu_{k_+}^+. \tag{54}$$

Then

$$\lambda_i^+ \le \mu_i^+ \le \lambda_{i+n-k}^+, \quad \text{for } 1 \le i \le k_+, \tag{55}$$

$$\lambda_{j+n-k}^- \le \mu_i^- \le \lambda_i^-, \quad \text{for } 1 \le j \le k_-, \tag{56}$$

where we set $\lambda_i^+ = \infty$ *for* $i > n_+$ *and* $\lambda_j^- = -\infty$ *for* $j > n_-$.

2. *If* $A - \lambda B$ *is positive semidefinite, then*

$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) = \sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^-. \tag{57}$$

   (a) *The infimum is attainable, if there exists a matrix* $X_{\min}$ *that satisfies* $X_{\min}^{\mathrm{H}}BX_{\min} = J_k$ *and whose first* k_+ *columns consist of the eigenvectors associated with the eigenvalues* λ_j^+ *for* $1 \le j \le k_+$ *and whose last* k_- *columns consist of the eigenvectors associated with the eigenvalues* λ_i^- *for* $1 \le i \le k_-$.

   (b) *If* $A - \lambda B$ *is positive definite or positive semidefinite but diagonalizable, then the infimum is attainable.*

   (c) *When the infimum is attained by* $X_{\min}$*, there is a Hermitian* $A_0 \in \mathbb{C}^{k\times k}$ *whose eigenvalues are* $\lambda_i^\pm$*,* $i = 1, 2, \dots, k_\pm$ *such that*

$$X_{\min}^{\mathrm{H}}BX_{\min} = J_k, \quad AX_{\min} = BX_{\min}A_0.$$

[4] Hermitian pencil $A - \lambda B$ is *diagonalizable* if there exists a nonsingular matrix W such that both $W^{\mathrm{H}}AW$ and $W^{\mathrm{H}}BW$ are diagonal.

3. *$A-\lambda B$ is a positive semidefinite pencil if and only if*
$$\inf_{X^{\mathrm{H}}BX=J_k} \operatorname{trace}(X^{\mathrm{H}}AX) > -\infty. \tag{58}$$

4. *If* $\operatorname{trace}(X^{\mathrm{H}}AX)$ *as a function of X subject to $X^{\mathrm{H}}BX = J_k$ has a local minimum, then $A-\lambda B$ is a positive semidefinite pencil and the minimum is global.*

4 Linear Response Eigenvalue Problem

We are interested in solving the standard eigenvalue problem of the form:

$$\begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} y \\ x \end{bmatrix}, \tag{59}$$

where K and M are $n \times n$ real symmetric positive semidefinite matrices and one of them is definite. We referred to it as a *linear response (LR) eigenvalue problem* because it is equivalent to the original LR eigenvalue problem

$$\begin{bmatrix} A & B \\ -B & -A \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix} \tag{60}$$

via a simple orthogonal similarity transformation [2], where A and B are $n \times n$ real symmetric matrices such that the symmetric matrix $\begin{bmatrix} A & B \\ B & A \end{bmatrix}$ is symmetric positive definite[5] [41, 51]. In computational physics and chemistry literature, it is this eigenvalue problem that is referred to as the linear response eigenvalue problem (see, e.g., [36]), or *random phase approximation* (RPA) eigenvalue problem (see, e.g., [15]).

While (59) is not a symmetric eigenvalue problem, it has the symmetric structure in its submatrices and many optimization principles that are similar to those one usually find in the symmetric eigenvalue problem. For example, (59) has only real eigenvalues. But more can be said: its eigenvalues come in $\pm\lambda$ pairs. Denote its eigenvalues by

$$-\lambda_n \le \cdots \le -\lambda_1 \le +\lambda_1 \le \cdots \le +\lambda_n.$$

In practice, the first few positive eigenvalues and their corresponding eigenvectors are needed. In 1961, Thouless [50] obtained a minimization principle for λ_1, now known as *Thouless' minimization principle*, which equivalently stated for (59) is

$$\lambda_1 = \min_{x,y} \frac{x^{\mathrm{T}}Kx + y^{\mathrm{T}}My}{2|x^{\mathrm{T}}y|}, \tag{61}$$

provided both $K \succ 0$ and $M \succ 0$. This very minimization principle, reminiscent of the first equation in (10), has been seen in action recently in, e.g., [8, 31, 33], for calculating λ_1 and, with aid of deflation, other λ_j.

Recently, Bai and Li [2] obtained Ky Fan trace min type principle, as well as Cauchy interlacing inequalities.

Theorem 4.1 (Bai and Li [2]). *Suppose that one of K, $M \in \mathbb{R}^{n\times n}$ is definite.*

[5] This condition is equivalent to that both $A \pm B$ are positive definite. In [2, 3] and this article, we focus on very much this case, except that one of $A \pm B$ is allowed to be positive semidefinite.

1. *We have*
$$\sum_{i=1}^{k} \lambda_i = \frac{1}{2} \inf_{U^{\mathrm{T}}V=I_k} \operatorname{trace}(U^{\mathrm{T}}KU + V^{\mathrm{T}}MV). \tag{62}$$
Moreover, "$\inf$*" can be replaced by "*$\min$*" if and only if both K and M are definite. When they are definite and if also $\lambda_k < \lambda_{k+1}$, then for any U and V that attain the minimum can be used to recovered λ_j for $1 \le j \le k$ and their corresponding eigenvectors.*

2. *Let U, $V \in \mathbb{R}^{n\times k}$ such that $U^{\mathrm{T}}V$ is nonsingular. Write $W = U^{\mathrm{T}}V = W_1^{\mathrm{T}}W_2$, where $W_i \in \mathbb{R}^{k\times k}$ are nonsingular, and define*
$$H_{\mathrm{SR}} = \begin{bmatrix} 0 & W_1^{-\mathrm{T}}U^{\mathrm{T}}KUW_1^{-1} \\ W_2^{-\mathrm{T}}V^{\mathrm{T}}MVW_2^{-1} & 0 \end{bmatrix}. \tag{63}$$
Denote by $\pm\mu_i$ ($1 \le i \le k$) the eigenvalues of H_{SR}, where $0 \le \mu_1 \le \cdots \le \mu_k$. Then
$$\lambda_i \le \mu_i \le \frac{\sqrt{\min\{\kappa(K), \kappa(M)\}}}{\cos\angle(\mathcal{U}, \mathcal{V})}\, \lambda_{i+n-k} \quad \text{for } 1 \le i \le k, \tag{64}$$
where $\mathcal{U} = \mathcal{R}(U)$ and $\mathcal{V} = \mathcal{R}(V)$, and $\kappa(K) = \|K\|_2\|K^{-1}\|_2$ and $\kappa(M) = \|M\|_2\|M^{-1}\|_2$ are the spectral condition numbers.

Armed with these minimization principles, we can work out extensions of the previously discussed steepest descent methods in subsection 2.3 and conjugate gradient methods in subsection 2.4 for the linear response eigenvalue problem (59). In fact, some extensions have been given in [2, 3, 42].

5 Hyperbolic Quadratic Eigenvalue Problem

It was argued in [20] that the hyperbolic quadratic eigenvalue problem (HQEP) is the closest analogue of the standard Hermitian eigenvalue problem $Hx = \lambda x$ when it comes to the quadratic eigenvalue problem
$$(\lambda^2 A + \lambda B + C)x = 0. \tag{65}$$
In many ways, both problems share common properties: the eigenvalues are all real, and for HQEP there is a version of the min-max principles [10] that is very much like the Courant-Fischer min-max principles.

When (65) is satisfied for a scalar λ and nonzero vector x, we call λ a *quadratic eigenvalue*, x an associated *quadratic eigenvector*, and (λ, x) a *quadratic eigenpair*.

One source of HQEP (65) is dynamical systems with friction, where A, C are associated with the kinetic-energy and potential-energy quadratic form, respectively, and B is associated with the Rayleigh dissipation function. When A, B, and C are Hermitian, and A and B are positive definite and C positive semidefinite, we say the dynamical system is *overdamped* if
$$(x^{\mathrm{H}}Bx)^2 - 4(x^{\mathrm{H}}Ax)(x^{\mathrm{H}}Cx) > 0 \quad \text{for any nonzero vector } x. \tag{66}$$
An HQEP is slightly more general than an overdamped QEP in that B and C are no longer required positive definite or positive semidefinite, respectively. However, a suitable shift in λ can turn an HQEP into an overdamped HQEP [16].

In what follows, A, B, $C \in \mathbb{C}^{n\times n}$ are Hermitian, $A \succ 0$, and (66) holds. Thus (65) is a HQEP for $\boldsymbol{Q}(\lambda) = \lambda^2 A + \lambda B + C \in \mathbb{C}^{n\times n}$. Denote its quadratic eigenvalues by $\lambda_i^{\pm}$ and arrange them in the order of

$$\lambda_1^- \le \cdots \le \lambda_n^- < \lambda_1^+ \le \cdots \le \lambda_n^+. \tag{67}$$

Consider the following equation in λ

$$f(\lambda, x) := x^{\mathrm{H}}\boldsymbol{Q}(\lambda)x = \lambda^2(x^{\mathrm{H}}Ax) + \lambda(x^{\mathrm{H}}Bx) + (x^{\mathrm{H}}Cx) = 0, \tag{68}$$

given $x \ne 0$. Since $\boldsymbol{Q}(\lambda)$ is hyperbolic, this equation always has two distinct real roots (as functions of x)

$$\rho_{\pm}(x) = \frac{-x^{\mathrm{H}}Bx \pm \left[(x^{\mathrm{H}}Bx)^2 - 4(x^{\mathrm{H}}Ax)(x^{\mathrm{H}}Cx)\right]^{1/2}}{2(x^{\mathrm{H}}Ax)}. \tag{69}$$

We shall call $\rho_+(x)$ the *pos-type Rayleigh quotient* of $\boldsymbol{Q}(\lambda)$ on x, and $\rho_-(x)$ the *neg-type Rayleigh quotient* of $\boldsymbol{Q}(\lambda)$ on x.

Theorem 5.1 below is a restatement of [32, Theorem 32.10, Theorem 32.11 and Remark 32.13]. However, it is essentially due to Duffin [10] whose proof, although for overdamped $\boldsymbol{Q}$, works for the general hyperbolic case. They can be considered as a generalization of the Courant-Fischer min-max principles (see [38, p.206], [47, p.201]).

Theorem 5.1 ([10])**.** *We have*

$$\lambda_i^+ = \max_{\substack{\mathcal{X}\subseteq\mathbb{C}^n \\ \operatorname{codim}\mathcal{X}=i-1}} \min_{\substack{x\in\mathcal{X} \\ x\ne 0}} \rho_+(x), \quad \lambda_i^+ = \min_{\substack{\mathcal{X}\subseteq\mathbb{C}^n \\ \dim\mathcal{X}=i}} \max_{\substack{x\in\mathcal{X} \\ x\ne 0}} \rho_+(x), \tag{70a}$$

$$\lambda_i^- = \max_{\substack{\mathcal{X}\subseteq\mathbb{C}^n \\ \operatorname{codim}\mathcal{X}=i-1}} \min_{\substack{x\in\mathcal{X} \\ x\ne 0}} \rho_-(x), \quad \lambda_i^- = \min_{\substack{\mathcal{X}\subseteq\mathbb{C}^n \\ \dim\mathcal{X}=i}} \max_{\substack{x\in\mathcal{X} \\ x\ne 0}} \rho_-(x). \tag{70b}$$

In particular,

$$\lambda_1^+ = \min_{x\ne 0} \rho_+(x), \quad \lambda_n^+ = \max_{x\ne 0} \rho_+(x), \tag{71a}$$

$$\lambda_1^- = \min_{x\ne 0} \rho_-(x), \quad \lambda_n^- = \max_{x\ne 0} \rho_-(x). \tag{71b}$$

To generalize Ky Fan trace min/max principle and Cauchy's interlacing inequalities, we introduce the following notations. For $X \in \mathbb{C}^{n\times k}$ with $\operatorname{rank}(X) = k$, $X^{\mathrm{H}}\boldsymbol{Q}(\lambda)X$ is a $k \times k$ hyperbolic quadratic matrix polynomial. Hence its quadratic eigenvalues are real. Denote them by $\lambda_{i,X}^{\pm}$ arranged as

$$\lambda_{1,X}^- \le \cdots \le \lambda_{k,X}^- \le \lambda_{1,X}^+ \le \cdots \le \lambda_{k,X}^+.$$

Theorem 5.2. *1.* [28] *We have*

$$\min_{\operatorname{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^{\pm} = \sum_{j=1}^{k} \lambda_j^{\pm}, \quad \max_{\operatorname{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^{\pm} = \sum_{j=1}^{k} \lambda_{n-k+j}^{\pm}. \tag{72}$$

2. [52] *For* $X \in \mathbb{C}^{n\times k}$ *with* $\operatorname{rank}(X) = k$,

$$\lambda_i^+ \le \lambda_{i,X}^+ \le \lambda_{i+n-k}^+, \quad i = 1, \cdots, k, \tag{73a}$$

$$\lambda_j^- \le \lambda_{j,X}^- \le \lambda_{j+n-k}^-, \quad j = 1, \cdots, k. \tag{73b}$$

Armed with these minimization principles, we can work out extensions of the previously discussed steepest descent methods in subsection 2.3 and conjugate gradient methods in subsection 2.4 for the HQEP $\boldsymbol{Q}(\lambda)x = 0$. Details, among others, can be found in [28].

References

[1] E. Anderson, Z. Bai, C. H. Bischof, S. Blackford, J. W. Demmel, J. J. Dongarra, J. J. Du Croz, A. Greenbaum, S. J. Hammarling, A. McKenney, and D. C. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, third edition, 1999.

[2] Zhaojun Bai and Ren-Cang Li. Minimization principle for linear response eigenvalue problem, I: Theory. *SIAM J. Matrix Anal. Appl.*, 33(4):1075–1100, 2012.

[3] Zhaojun Bai and Ren-Cang Li. Minimization principle for linear response eigenvalue problem, II: Computation. *SIAM J. Matrix Anal. Appl.*, 34(2):392–416, 2013.

[4] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics, vol. 169. Springer, New York, 1996.

[5] Rajendra Bhatia. *Positive Definite Matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, New Jersey, 2007.

[6] Paul Binding, Branko Najman, and Qiang Ye. A variational principle for eigenvalues of pencils of Hermitian matrices. *Integr. Eq. Oper. Theory*, 35:398–422, 1999.

[7] W. W. Bradbury and R. Fletcher. New iterative methods for solution of the eigenproblem. *Numer. Math.*, 9(3):259–267, 1966.

[8] M. Challacombe. Linear scaling solution of the time-dependent self-consisten-field equations. e-print arXiv:1001.2586v2, 2010.

[9] J. Demmel. *Applied Numerical Linear Algebra*. SIAM, Philadelphia, PA, 1997.

[10] R. Duffin. A minimax theory for overdamped networks. *Indiana Univ. Math. J.*, 4:221–233, 1955.

[11] R. Fletcher and C. M.Reeves. Function minimization by conjugate gradients. *Comput. J.*, 7:149–154, 1964.

[12] G. Golub and Q. Ye. An inverse free preconditioned Krylov subspace methods for symmetric eigenvalue problems. *SIAM J. Sci. Comput.*, 24:312–334, 2002.

[13] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[14] Anne Greenbaum. *Iterative Methods for Solving Linear Systems*. SIAM, Philadelphia, 1997.

[15] M. Grüning, A. Marini, and X. Gonze. Implementation and testing of Lanczos-based algorithms for random-phase approximation eigenproblems. Technical report, arXiv:1102.3909v1, February 2011.

[16] C.-H. Guo and P. Lancaster. Algorithms for hyperbolic quadratic eigenvalue problems. *Math. Comp.*, 74:1777–1791, 2005.

[17] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, 49:409436, 1952.

[18] U. Hetmaniuk and R. Lehoucq. Basis selection in LOBPCG. *J. Comput. Phys.*, 218(1):324–332, 2006.

[19] Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[20] Nicholas J. Higham, Françoise Tisseur, and Paul M. Van Dooren. Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems. *Linear Algebra Appl.*, 351-352:455–474, 2002.

[21] L. V. Kantorovich and G. P. Akilov. *Functional Analysis in Normed Spaces.* MacMillian, New York, 1964.

[22] A. V. Knyazev and A. L. Skorokhodov. On exact estimates of the convergence rate of the steepest ascent method in the symmetric eigenvalue problem. *Linear Algebra Appl.*, 154-156:245–257, 1991.

[23] Andrew V. Knyazev. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. *SIAM J. Sci. Comput.*, 23(2):517–541, 2001.

[24] Andrew V. Knyazev and Klaus Neymeyr. A geometric theory for preconditioned inverse iteration III: A short and sharp convergence estimate for generalized eigenvalue problems. *Linear Algebra Appl.*, 358(1-3):95–114, 2003.

[25] J. Kovač-Striko and K. Veselić. Trace minimization and definiteness of symmetric pencils. *Linear Algebra Appl.*, 216:139–158, 1995.

[26] P. Lancaster and Q. Ye. Variational properties and Rayleigh quotient algorithms for symmetric matrix pencils. *Oper. Theory: Adv. Appl.*, 40:247–278, 1989.

[27] Xin Liang and Ren-Cang Li. Extensions of Wielandt's min-max principles for positive semi-definite pencils. *Linear and Multilinear Algebra*, 2013. Published online: 07 Jun 2013.

[28] Xin Liang and Ren-Cang Li. The hyperbolic quadratic eigenvalue problem. work-in-progress, 2013.

[29] Xin Liang, Ren-Cang Li, and Zhaojun Bai. Trace minimization principles for positive semi-definite pencils. *Linear Algebra Appl.*, 438:3085–3106, 2013.

[30] D. E. Longsine and S. F. McCormick. Simultaneous Rayleigh-quotient minimization methods for $Ax = \lambda Bx$. *Linear Algebra Appl.*, 34:195–234, 1980.

[31] M. J. Lucero, A. M. N. Niklasson, S. Tretiak, and M. Challacombe. Molecular-orbital-free algorithm for excited states in time-dependent perturbation theory. *J. Chem. Phys.*, 129(6):064114, 2008.

[32] A.S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils.* Translations of mathematical monographs, vol. 71. AMS, Providence, RI, 1988.

[33] Atsushi Muta, Jun-Ichi Iwata, Yukio Hashimoto, and Kazuhiro Yabana. Solving the RPA eigenvalue equation in real-space. *Progress Theoretical Physics*, 108(6):1065–1076, 2002.

[34] B. Najman and Q. Ye. A minimax characterization of eigenvalues of Hermitian pencils II. *Linear Algebra Appl.*, 191:183–197, 1993.

[35] Jorge Nocedal and Stephen Wright. *Numerical Optimization.* Springer, 2nd edition, 2006.

[36] J. Olsen, H. J. Aa. Jensen, and P. Jørgensen. Solution of the large matrix equations which occur in response theory. *J. Comput. Phys.*, 74(2):265–282, 1988.

[37] E. E. Ovtchinnikov. Sharp convergence estimates for the preconditioned steepest descent method for Hermitian eigenvalue problems. *SIAM J. Numer. Anal.*, 43(6):2668–2689, 2006.

[38] B. N. Parlett. *The Symmetric Eigenvalue Problem.* SIAM, Philadelphia, 1998.

[39] Boris Teodorovich Polyak. *Introduction to Optimization.* Optimization Software, New York, 1987.

[40] Patrick Quillen and Qiang Ye. A block inverse-free preconditioned Krylov subspace method for symmetric generalized eigenvalue problems. *J. Comput. Appl. Math.*, 233(5):1298–1313, 2010.

[41] P. Ring and P. Schuck. *The Nuclear Many-Body Problem.* Springer-Verlag, New York, 1980.

[42] D. Rocca, Z. Bai, R.-C. Li, and G. Galli. A block variational procedure for the iterative diagonalization of non-Hermitian random-phase approximation matrices. *J. Chem. Phys.*, 136:034111, 2012.

[43] Y. Saad. On the rates of convergence of the Lanczos and the block-Lanczos methods. *SIAM J. Numer. Anal.*, 15(5):687–706, October 1980.

[44] Yousef Saad. *Iterative Methods for Sparse Linear Systems.* SIAM, Philadelphia, 2nd edition, 2003.

[45] B. Samokish. The steepest descent method for an eigenvalue problem with semi-bounded operators. *Izv. Vyssh. Uchebn. Zaved. Mat.*, 5:105–114, 1958. in Russian.

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

[47] G. W. Stewart and Ji-Guang Sun. *Matrix Perturbation Theory.* Academic Press, Boston, 1990.

[48] Wenyu Sun and Ya-Xiang Yuan. *Optimization Theory and Methods – Nonlinear Programming.* Springer, New York, 2006.

[49] I. Takahashi. A note on the conjugate gradient method. *Inform. Process. Japan*, 5:45–49, 1965.

[50] D. J. Thouless. Vibrational states of nuclei in the random phase approximation. *Nuclear Physics*, 22(1):78–95, 1961.

[51] D. J. Thouless. *The Quantum Mechanics of Many-Body Systems.* Academic, 1972.

[52] K. Veselić. Note on interlacing for hyperbolic quadratic pencils. In Jussi Behrndt, Karl-Heinz Förster, and Carsten Trunk, editors, *Recent Advances in Operator Theory in Hilbert and Krein Spaces*, volume 198 of *Oper. Theory: Adv. Appl.*, pages 305–307. 2010.

[53] Qiang Ye. *Variational Principles and Numerical Algorithms for Symmetric Matrix Pencils.* PhD thesis, University of Calgary, Calgary, Canada, 1989.