SCIENCE CHINA
Mathematics

• ARTICLES •

August 2016 Vol. 59 No. 8: 1443–1460
doi: 10.1007/s11425-016-0297-1

Linear response eigenvalue problem solved by extended locally optimal preconditioned conjugate gradient methods

BAI ZhaoJun[1,*], LI RenCang[2] & LIN WenWei[3]

[1]*Department of Computer Science and Department of Mathematics, University of California at Davis, Davis, CA* 95616*, USA;*
[2]*Department of Mathematics, University of Texas at Arlington, Arlington, TX* 76019*, USA;*
[3]*Department of Applied Mathematics, Taiwan Chiao Tung University, Hsinchu* 300*, China*

Email: bai@cs.ucdavis.edu, rcli@uta.edu, wwlin@am.nctu.edu.tw

Received November 6, 2015; accepted May 3, 2016; published online June 2, 2016

Abstract The locally optimal block preconditioned 4-d conjugate gradient method (LOBP4dCG) for the linear response eigenvalue problem was proposed by Bai and Li (2013) and later was extended to the generalized linear response eigenvalue problem by Bai and Li (2014). We put forward two improvements to the method: A shifting deflation technique and an idea of extending the search subspace. The deflation technique is able to deflate away converged eigenpairs from future computation, and the idea of extending the search subspace increases convergence rate per iterative step. The resulting algorithm is called the extended LOBP4dCG (ELOBP4dCG). Numerical results of the ELOBP4dCG strongly demonstrate the capability of deflation technique and effectiveness the search space extension for solving linear response eigenvalue problems arising from linear response analysis of two molecule systems.

Keywords eigenvalue problem, linear response, deflation, conjugate-gradient, deflation

MSC(2010) 65L15, 15A18, 81Q15

Citation: Bai Z J, Li R C, Lin W W. Linear response eigenvalue problem solved by extended locally optimal preconditioned conjugate gradient methods. Sci China Math, 2016, 59: 1443–1460, doi: 10.1007/s11425-016-0297-1

1 Introduction

Consider the linear response eigenvalue problem (LREP) of the form

$$Hz \equiv \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} E_+ & 0 \\ 0 & E_- \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} \equiv \lambda E z, \tag{1.1}$$

where K and M are $n \times n$ Hermitian and positive semidefinite and one of them is definite, $E_\pm$ are $n \times n$ and nonsingular. In [4], minimization principles and Cauchy-type interlacing inequalities were obtained for (1.1). As a result, a locally optimal block preconditioned 4-d conjugate gradient method (LOBP4dCG) was designed to compute the first few positive eigenvalues and their associated eigenvectors of interest.

*Corresponding author

Rigorously speaking, the LREP (1.1) is not quite the linear response eigenvalue problem in its original form [8, 23, 24]:

$$\begin{bmatrix} A & B \\ -B & -A \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} \Sigma & \Delta \\ \Delta & \Sigma \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}, \tag{1.2}$$

where A and B are $n \times n$ Hermitian matrices and Σ and Δ are also $n \times n$ with Σ being Hermitian while Δ skew-Hermitian (i.e., $\Delta^{\mathrm{H}} = -\Delta$). Additionally, $\left[\begin{smallmatrix} A & B \\ B & A \end{smallmatrix}\right]$ is symmetric positive definite and $\left[\begin{smallmatrix} \Sigma & \Delta \\ \Delta & \Sigma \end{smallmatrix}\right]$ is nonsingular [26, 32]. Although (1.1) is not the same as (1.2), the two eigenvalue problems are equivalent through an orthogonal similarity transformation [4]. In the case of (1.2), its equivalent (1.1) has the property that both $K = A - B$ and $M = A + B$ are definite and $E_{\pm} = \Sigma \pm \Delta$ are nonsingular with $E_+^{\mathrm{H}} = E_-$. The linear response eigenvalue problem (1.1) includes the one in [2, 3] as a special case: $E = I_{2n}$. Therefore any developments in this paper apply to the LREP there.

There was little discussion in [4] on deflation as to how to effectively deflate away known or already computed eigenpairs for numerical efficiency in the process of computing more eigenpairs while avoiding the known ones. Although demonstrated feasible in [3], LOBP4dCG behaves delicately and can converge slowly when one of K and M is singular because of the appearance of eigenvalue 0. Conceivably, in this "singular" case, this known eigenvalue 0 should be deflated before applying LOBP4dCG.

The major goals of this paper are two-fold: to propose a deflation technique and an idea of extending search subspace for fast convergence rate. For deflation, we will develop a shifting deflation technique by a low-rank update to either K or M so that the resulting K or M performs at the about comparable cost as the original K or M when it comes to do matrix-vector multiplication operations. This deflation technique also allows us to deflate away known eigenvalue 0 if any for better numerical efficiency. For search subspace extension, we combine the inverse-free idea of Golub and Ye [10] (also known as augmented projection subspace in [13, 34]) with our 4-d searching idea in [3, 4]. Both deflation and search subspace extension techniques bring significant improvements to LOBP4DCG in numerical performance as our later examples will show. The resulting algorithm is called the extended LOBPCG (ELOBP4dCG). Along the way, we also propose a new implementation of the method to solve the projected eigenvalue problem having the same form as (1.1) but of much smaller scale has to be solved. In the new implementation, the projected eigenvalue problem is converted to a singular value problem of half the size and the latter can be solved more efficiently and robustly, namely computed approximate eigenvalues are always real.

The rest of this paper is organized as follows. Section 2 reviews theoretical foundations of the LREP. Section 3 presents the theory of deflation. The algorithmic framework and implementation remarks of the new ELOBP4dCG method are in Section 4. Numerical examples are presented in Section 5. Concluding remarks are given in Section 6.

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 j-th column. The superscripts "T" and "H" take transpose and complex conjugate transpose of a matrix/vector, respectively. $\mathcal{R}(X)$ is the column space of a matrix X. 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 a matrix X, $u_{(j)}$ is u's j-th 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 Theoretical foundations

We assume that K and M are $n \times n$ Hermitian and positive semidefinite and one of them is definite. Without loss of generality, for the sake of presentation, in what follows we assume $K \succeq 0$ and $M \succ 0$, unless explicitly otherwise. The interchangeable roles played by K and M make it rather straightforward

to create a version for the case when K is definite by simply swapping K and M in each of their appearances and E_+ and E_- in each of their appearances.

Decompose $E_\pm$ as

$$E_-^{\mathrm{H}} = E_+ = CD^{\mathrm{H}}, \tag{2.1}$$

where C, $D \in \mathbb{C}^{n\times n}$ are nonsingular[1)]. With (2.1), the LREP (1.1) is equivalent to

$$\boldsymbol{H}\boldsymbol{z} \equiv \begin{bmatrix} 0 & \boldsymbol{K} \\ \boldsymbol{M} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{bmatrix} = \lambda \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{bmatrix}, \tag{2.2a}$$

where

$$\boldsymbol{K} = C^{-1}KC^{-\mathrm{H}}, \quad \boldsymbol{M} = D^{-1}MD^{-\mathrm{H}}, \quad \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{bmatrix} = \Gamma^{\mathrm{H}} \begin{bmatrix} y \\ x \end{bmatrix} \quad \text{and} \quad \Gamma = \begin{bmatrix} D & 0 \\ 0 & C \end{bmatrix}. \tag{2.2b}$$

We now have two equivalent eigenvalue problems (1.1) and (2.2a) in the sense that both have the same eigenvalues and their eigenvectors are related by the relation shown in (2.2b).

For the eigenvalue problem (2.2a), we know that $\boldsymbol{K}$, $\boldsymbol{M} \succeq 0$ because K, $M \succeq 0$. As argued in [2] for (1.1), the eigenvalues for (2.2a) are real and come in $\pm\lambda$ pairs. More precisely, denote the eigenvalues of $\boldsymbol{K}\boldsymbol{M}$ by λ_i^2 $(1 \leqslant i \leqslant n)$ in the ascending order,

$$0 \leqslant \lambda_1^2 \leqslant \lambda_2^2 \leqslant \cdots \leqslant \lambda_n^2, \tag{2.3}$$

where $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. The eigenvalues of $\boldsymbol{M}\boldsymbol{K}$ are λ_i^2 $(1 \leqslant i \leqslant n)$, too. The eigenvalues of $H - \lambda E$ are then $\pm\lambda_i$ for $i = 1, 2, \ldots, n$ with the ordering

$$-\lambda_n \leqslant \cdots \leqslant -\lambda_1 \leqslant +\lambda_1 \leqslant \cdots \leqslant +\lambda_n. \tag{2.4}$$

For convenience, we shall associate half of eigenvalues 0, if any, with the positive sign and the other half with the negative sign, as argued in [2]. Doing so legitimizes the use of the phrase "the first k smallest eigenvalues with the positive sign of $H - \lambda E$" to refer to λ_i for $1 \leqslant i \leqslant k$ without ambiguity even when $\lambda_1 = +0$. Throughout this paper, we will stick to using $\pm\lambda_i$ for $1 \leqslant i \leqslant n$ in the order of (2.4) to denote the eigenvalues of $H - \lambda E$.

In the case of $K \succ 0$, $M \succ 0$, and $E_+ = E_- = I_n$, Thouless [31] essentially proved the following minimizing characterization[2)] for λ_1,

$$\lambda_1 = \min_{x,y} \rho(x, y) \quad \text{with} \quad \rho(x, y) := \frac{x^{\mathrm{H}}Kx + y^{\mathrm{H}}My}{2|x^{\mathrm{H}}y|} \tag{2.5}$$

which is now known as *Thouless' minimization principle*, and $\rho(x, y)$ is known as *Thouless' functional*. But in general, $K \succeq 0$ and without assuming $E_+ = E_- = I_n$, it is shown [4] that

$$\lambda_1 = \inf_{x,y} \rho(x, y) \quad \text{with} \quad \rho(x, y) := \frac{x^{\mathrm{H}}Kx + y^{\mathrm{H}}My}{2|x^{\mathrm{H}}E_+y|}, \tag{2.6}$$

where "inf" can be replaced by "min" if and only if $K \succ 0$. In view of (2.6), naturally $\rho(x, y)$ in (2.6) should be regarded as a Thouless' functional for the case. Without confusion, we still call it *Thouless' functional*.

Theorem 2.1 (See [4]). *Suppose that $M \succ 0$, and define C and D by* (2.1). *Then the following statements are true*:

(1) *There exist nonsingular Φ, $\Psi \in \mathbb{C}^{n\times n}$ such that*

$$K = C\Psi\Lambda^2\Psi^{\mathrm{H}}C^{\mathrm{H}}, \quad M = D\Phi\Phi^{\mathrm{H}}D^{\mathrm{H}}, \tag{2.7}$$

[1)] How this factorization is done is not mathematically essential. It is included for the case when such a factorization may be naturally available and can be exploited for numerical gains. In general, we can simply let one of C and D be I_n.

[2)] Original Thouless' minimization principle is for (1.2) with $\Sigma = I_n$ and $\Delta = 0$. The form of (2.5) is obtained by the equivalence between (1.1) and (1.2) (see [33]).

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *and* $\Phi = \Psi^{-\mathrm{H}}$.

(2) *If* K *is also definite, then all* $\lambda_i > 0$ *and* $H - \lambda E$ *is diagonalizable*:

$$H\Upsilon = E\Upsilon \begin{bmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{bmatrix}, \tag{2.8}$$

where

$$\Upsilon = \Gamma^{-\mathrm{H}} \begin{bmatrix} \Psi\Lambda & \Psi\Lambda \\ \Phi & -\Phi \end{bmatrix}.$$

(3) $H - \lambda E$ *is not diagonalizable if and only if* $\lambda_1 = 0$ *which happens when and only when* K *is singular.*

(4) *The* i*-th column of* Υ *is the eigenvector of* $H - \lambda E$ *corresponding to* λ_i*, where* $1 \leqslant i \leqslant n$*, and it is unique if*

(i) λ_i *is a simple eigenvalue of* (1.1)*, or*

(ii) $i = 1$, $\lambda_1 = +0 < \lambda_2$*. In this case,* 0 *is a double eigenvalue of* $H - \lambda E$ *but there is only one eigenvector associated with it.*

(5) *If* $0 = \lambda_1 = \cdots = \lambda_{n_0} < \lambda_{n_0+1}$*, then the Kronecker canonical form of* $H - \lambda E$ *is*

$$\underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{n_0} \oplus \mathrm{diag}(\lambda_{n_0+1}, -\lambda_{n_0+1}, \ldots, \lambda_n, -\lambda_n) - \lambda I_{2n}. \tag{2.9}$$

Thus $H - \lambda E$ *has* 0 *as an eigenvalue of algebraic multiplicity* $2n_0$ *with only* n_0 *linearly independent eigenvectors which are the columns of*

$$\Gamma^{-\mathrm{H}} \begin{bmatrix} 0 \\ \Phi_{(:,1:\ell)} \end{bmatrix}.$$

Set

$$\mathscr{I} = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}, \quad \mathscr{I}_E = \begin{bmatrix} 0 & E_- \\ E_+ & 0 \end{bmatrix} = \Gamma \mathscr{I} \Gamma^{\mathrm{H}}, \tag{2.10}$$

where Γ is given by (2.2b). Both are Hermitian but indefinite. The matrices $\mathscr{I}_E$ and $\mathscr{I}$ induce *indefinite inner products* on $\mathbb{C}^{2n}$,

$$\langle z_1, z_2 \rangle_{\mathscr{I}_E} := z_1^{\mathrm{H}} \mathscr{I}_E z_2 = \langle \boldsymbol{z}_1, \boldsymbol{z}_2 \rangle_{\mathscr{I}} =: \boldsymbol{z}_1^{\mathrm{H}} \mathscr{I} \boldsymbol{z}_2,$$

where $\boldsymbol{z}_i = \Gamma^{\mathrm{H}} z_i$. The following two theorems tell us orthogonality properties among eigenvectors and invariant subspaces of $H - \lambda E$.

Theorem 2.2 (See [4]). (1) *Let* (α, z) *be an eigenpair of* $H - \lambda E$*, where*

$$z = \begin{bmatrix} y \\ x \end{bmatrix} \neq 0$$

and x, $y \in \mathbb{C}^n$*. Suppose* $\alpha \neq 0$*. Then* $\alpha \langle z, z \rangle_{\mathscr{I}_E} = 2\alpha\, x^{\mathrm{H}} E_+ y > 0$ *which implies* $\langle z, z \rangle_{\mathscr{I}_E} = 2x^{\mathrm{H}} E_+ y \neq 0$.

(2) *Let* (α_i, z_i) $(i = 1, 2)$ *be two eigenpairs of* $H - \lambda E$*. Partition*

$$z_i = \begin{bmatrix} y_i \\ x_i \end{bmatrix} \neq 0,$$

where x_i, $y_i \in \mathbb{C}^n$*. The following statements are true*:

(i) *If* $\alpha_1 \neq \alpha_2$*, then* $\langle z_1, z_2 \rangle_{\mathscr{I}_E} = y_1^{\mathrm{H}} E_- x_2 + x_1^{\mathrm{H}} E_+ y_2 = 0$.

(ii) *If* $\alpha_1 \neq \pm\alpha_2$ *and* α_1 *and* α_2 *are nonzero, then* $y_1^{\mathrm{H}} E_- x_2 = x_1^{\mathrm{H}} E_+ y_2 = 0$.

Theorem 2.3. *Let*

$$Z_i = \begin{bmatrix} Y_i \\ X_i \end{bmatrix}$$

(*conformally partitioned*) *be the basis matrices of two invariant subspaces of* $H - \lambda E$, *i.e.*,

$$HZ_i = EZ_iA_i \tag{2.11}$$

for some square matrices A_i *for* $i = 1, 2$.

(1) *If* $\operatorname{eig}(A_1) \cap \operatorname{eig}(A_2) = \emptyset$, *then*

$$Z_1^{\mathrm{H}} \mathscr{I}_E Z_2 \equiv X_1^{\mathrm{H}} E_+ Y_2 + Y_1^{\mathrm{H}} E_- X_2 = 0. \tag{2.12}$$

(2) *If* $\operatorname{eig}(A_1) \cap [\operatorname{eig}(A_2) \cup \operatorname{eig}(-A_2)] = \emptyset$, *then*

$$X_1^{\mathrm{H}} E_+ Y_2 = Y_1^{\mathrm{H}} E_- X_2 = 0. \tag{2.13}$$

Proof. Let $\boldsymbol{Z}_i = \Gamma^{\mathrm{H}} Z_i$. Then $\boldsymbol{H}\boldsymbol{Z}_i = \boldsymbol{Z}_i A_i$ by (2.11). We have $\boldsymbol{Z}_1^{\mathrm{H}}(\mathscr{I}\boldsymbol{H})^{\mathrm{H}}\boldsymbol{Z}_2 = \boldsymbol{Z}_1^{\mathrm{H}}(\mathscr{I}\boldsymbol{H})\boldsymbol{Z}_2$ since $\mathscr{I}\boldsymbol{H}$ is Hermitian, and also

$$\begin{aligned} \boldsymbol{Z}_1^{\mathrm{H}}(\mathscr{I}\boldsymbol{H})^{\mathrm{H}}\boldsymbol{Z}_2 &= (\boldsymbol{H}\boldsymbol{Z}_1)^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2 = A_1^{\mathrm{H}}\boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2, \\ \boldsymbol{Z}_1^{\mathrm{H}}(\mathscr{I}\boldsymbol{H})\boldsymbol{Z}_2 &= \boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}(\boldsymbol{H}\boldsymbol{Z}_2) = \boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2 A_2. \end{aligned}$$

Therefore $A_1^{\mathrm{H}}\boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2 - \boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2 A_2 = 0$ which, as a Sylvester equation in $\boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2$, has a unique solution $\boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2 = 0$ since $\operatorname{eig}(A_1) \cap \operatorname{eig}(A_2) = \emptyset$. Now notice

$$\boldsymbol{Z}_1^{\mathrm{H}}\mathscr{I}\boldsymbol{Z}_2 = Z_1^{\mathrm{H}}\Gamma\mathscr{I}\Gamma^{\mathrm{H}}Z_2 = Z_1^{\mathrm{H}}\mathscr{I}_E Z_2$$

to get (2.12). It can be verified that $H\hat{Z}_i = E\hat{Z}_i(-A_i)$, where

$$\hat{Z}_i = \begin{bmatrix} Y_i \\ -X_i \end{bmatrix}.$$

Thus for the same reason $\operatorname{eig}(A_1) \cap \operatorname{eig}(-A_2) = \emptyset$ gives $X_1^{\mathrm{H}} E_+ Y_2 - Y_1^{\mathrm{H}} E_- X_2 = 0$, which together with (2.12), yields $X_1^{\mathrm{H}} E_+ Y_2 = Y_1^{\mathrm{H}} E_- X_2 = 0$. □

The following theorem justifies the way we normalize the basis matrix of an approximate eigenspace associated with positive eigenvalues of $H - \lambda E$ in our numerical implementation later.

Theorem 2.4. *Suppose* $K \succ 0$ *and* $M \succ 0$, *and let* $\mathcal{Z}$ *be an eigenspace of* $H - \lambda E$ *associated with its* k *positive eigenvalues. Then* $U^{\mathrm{H}}KU = V^{\mathrm{H}}MV$ *for any basis matrix*

$$Z = \begin{bmatrix} V \\ U \end{bmatrix}$$

(*conformally partitioned*) *of* $\mathcal{Z}$.

Proof. Since H is diagonalizable, we start by letting

$$Z = \begin{bmatrix} V \\ U \end{bmatrix}$$

be the eigenvector matrix for the k positive eigenvalues. Then $KU = E_+V\Omega$ and $MV = E_-U\Omega$, where $\Omega = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_k)$. Therefore,

$$U^{\mathrm{H}}KU = U^{\mathrm{H}}E_+V\Omega, \quad V^{\mathrm{H}}MV = V^{\mathrm{H}}E_-U\Omega.$$

First consider the case when the k positive eigenvalues μ_i are distinct. Use Theorem 2.2 to conclude that $U^{\rm H}E_+V$ is diagonal and thus real; so is $V^{\rm H}E_-U$. Thus they commutes with Ω. Hence,

$$U^{\rm H}KU = (U^{\rm H}KU)^{\rm H} = (U^{\rm H}E_+V\Omega)^{\rm H} = \Omega\,(V^{\rm H}E_-U) = V^{\rm H}E_-U\,\Omega = V^{\rm H}MV, \tag{2.14}$$

as expected. Consider now that the k positive eigenvalues μ_i are not all distinct. Without loss of generality, we may assume μ_i are labelled in such a way that any μ_i of equal value are labelled consecutively. By Theorem 2.3, $U^{\rm H}E_+V = (V^{\rm H}E_-U)^{\rm H}$ is block diagonal with each diagonal block corresponding to a μ_i of equal value. Thus it still commutes with Ω. Hence the equations in (2.14) remain valid.

Finally any basis matrix of $\mathcal{Z}$ takes the form ZQ for some nonsingular Q. By what we just proved, we have

$$(UQ)^{\rm H}K(UQ) = Q^{\rm H}U^{\rm H}KUQ = Q^{\rm H}V^{\rm H}MVQ = (VQ)^{\rm H}M(VQ),$$

as was to be shown. □

3 Deflation

We first present the following theorem that lends itself to designing a deflation technique.

Theorem 3.1. *Let Ψ be the one in* (2.7), $\mathbb{J} = \{i_j : 1 \leqslant j \leqslant k\} \subset \{1, 2, \ldots, n\}$, *and let $V \in \mathbb{C}^{n\times k}$ with* $\mathrm{rank}(V) = k$ *satisfy*

$$\mathcal{R}(D^{\rm H}V) = \mathcal{R}(\Psi_{(:,\mathbb{J})}), \tag{3.1}$$

or equivalently $D^{\rm H}V = \Psi_{(:,\mathbb{J})}Q$ for some nonsingular $Q \in \mathbb{C}^{k\times k}$, where D is defined as (2.1). *Let $\xi > 0$, and define*

$$\underline{H} = \begin{bmatrix} 0 & K + \xi(E_+V)(E_+V)^{\rm H} \\ M & 0 \end{bmatrix}.$$

Then $H - \lambda E$ and $\underline{H} - \lambda E$ share the same eigenvalues $\pm\lambda_i$ for $i \notin \mathbb{J}$ and the corresponding eigenvectors, and the rest of eigenvalues of $\underline{H} - \lambda E$ are the square roots of the eigenvalues of $\Lambda_1^2 + \xi QQ^{\rm H}$, where $\Lambda_1 = \mathrm{diag}(\lambda_{i_1}, \ldots, \lambda_{i_k})$[3].

Proof. Let $\mathbb{J}_c$ be the complement of $\mathbb{J}$ in $\{1, 2, \ldots, n\}$, and write

$$\Phi_1 = \Phi_{(:,\mathbb{J})}, \quad \Phi_2 = \Phi_{(:,\mathbb{J}_c)}, \quad \Psi_1 = \Psi_{(:,\mathbb{J})}, \quad \Psi_2 = \Psi_{(:,\mathbb{J}_c)}, \tag{3.2}$$

and $\Lambda_2 = \mathrm{diag}(\lambda_i)_{i\in\mathbb{J}_c}$. It follows from (2.7) and $E_+ = CD^{\rm H}$ that

$$\begin{aligned} K &= C[\Psi_1, \Psi_2]\begin{bmatrix}\Lambda_1^2 & \\ & \Lambda_2^2\end{bmatrix}[\Psi_1, \Psi_2]^{\rm H}C^{\rm H}, \\ M &= D[\Phi_1, \Phi_2][\Phi_1, \Phi_2]^{\rm H}D^{\rm H}, \\ K + \xi(E_+V)(E_+V)^{\rm H} &= C[\Psi_1, \Psi_2]\begin{bmatrix}\Lambda_1^2 + \xi QQ^{\rm H} & 0 \\ 0 & \Lambda_2^2\end{bmatrix}[\Psi_1, \Psi_2]^{\rm H}C^{\rm H}. \end{aligned}$$

Therefore $\pm\lambda_i$ for $i \in \mathbb{J}_c$ are part of eigenvalues of $\underline{H} - \lambda E$ with the same eigenvectors as the ones for $H - \lambda E$ associated with eigenvalues $\pm\lambda_i$ for $i \in \mathbb{J}_c$. □

In an optimization approach like LOBP4dCG and other variations in [4], usually the eigenvalues of $H - \lambda E$ emerge in sequential order: From the smallest λ_1 upwards λ_2, etc. Once one or several of them converge to a preset accuracy, they should be deflated so that they will not be re-computed in the following iterations. Theorem 3.1 points to one deflation idea of doing this. Suppose we are interested in computing λ_i for $1 \leqslant i \leqslant \ell$ and their associated eigenvectors, and suppose k of them have been computed,

[3] The conclusion of this theorem is actually valid for any scalar ξ (real or complex). We specify $\xi > 0$ because this is what we need later.

where $k < \ell$. If at the same time we have $V \in \mathbb{C}^{n\times k}$ such that (3.1) holds with $\mathbb{J}$ being the collection of the indices of the k computed eigenpairs and have chosen a sufficiently large $\xi > 0$ such that

$$\lambda_1^2 + \xi[\sigma_{\min}(Q)]^2 \geqslant \lambda_{\ell+1}^2, \tag{3.3}$$

where $\sigma_{\min}(Q)$ denotes the smallest singular value of Q, then we can apply the optimization approach to $\underline{H} - \lambda E$, instead of the original $H - \lambda E$, to compute the rest of the desired eigenvalues and eigenvectors. There are two issues here for this deflation idea to work: (a) We know such a ξ exists, but precisely how large ξ should be may not be known *a priori*. On the other hand, it is not necessary to know the precise value so long that (3.3) is satisfied. In practice, some rough estimate of $\lambda_{\ell+1}$ may be available or guessed. (b) We need a $V \in \mathbb{C}^{n\times k}$ such that (3.1) holds with $\mathbb{J}$.

We now prove two theoretical results about what V should be in order to satisfy (3.1). There are two cases to consider, one for the eigenvalue 0, if any, and the other for nonzero eigenvalues.

There are eigenvalues 0 only if K is singular. In fact, if

$$n > r = \operatorname{rank}(K) > 0, \tag{3.4}$$

then $\lambda_i = 0$ for $1 \leqslant i \leqslant n - r$. How can we perturb K to move these eigenvalues 0 away without affecting all nonzero eigenvalues?

By Theorem 2.1(5), 0 as an eigenvalue of $H - \lambda E$ has algebraic multiplicity $2(n - r)$ and geometric multiplicity $n - r$. Now let us look for a basis of the invariant subspace associated with the eigenvalue 0. To this end, we first seek the invariant subspace associated with the eigenvalue 0 for the eigenvalue problem (2.2). By Theorem 2.1(5), the invariant subspace is the same as the null space of $\boldsymbol{H}^2$. So we need to solve

$$\boldsymbol{H}^2\boldsymbol{z} \equiv \begin{bmatrix} \boldsymbol{K}\boldsymbol{M} & 0 \\ 0 & \boldsymbol{M}\boldsymbol{K} \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{bmatrix} = 0 \Rightarrow \boldsymbol{K}\boldsymbol{M}\boldsymbol{y} = 0 \quad \text{and} \quad \boldsymbol{M}\boldsymbol{K}\boldsymbol{x} = 0.$$

Since $\boldsymbol{M} \succ 0$, the second equation $\boldsymbol{M}\boldsymbol{K}\boldsymbol{x} = 0$ is equivalent to $\boldsymbol{K}\boldsymbol{x} = 0$ which has $n - r$ linearly independent solutions. Let $\boldsymbol{U}_0 \in \mathbb{C}^{n\times(n-r)}$ be a basis solution matrix, i.e.,

$$\operatorname{rank}(\boldsymbol{U}_0) = n - r, \quad \boldsymbol{K}\boldsymbol{U}_0 = 0. \tag{3.5}$$

Such a $\boldsymbol{U}_0$ is not unique, but for our purpose any one of them is good enough. Now the equation $\boldsymbol{K}\boldsymbol{M}\boldsymbol{y} = 0$ is equivalent to $\boldsymbol{M}\boldsymbol{y} \in \mathcal{R}(\boldsymbol{U}_0)$ which has $n - r$ linearly independent solutions, too. Again for our purpose any one set of $n - r$ linearly independent solutions to $\boldsymbol{M}\boldsymbol{y} \in \mathcal{R}(\boldsymbol{U}_0)$ is good enough; so we may take the $n - r$ columns of

$$\boldsymbol{V}_0 = \boldsymbol{M}^{-1}\boldsymbol{U}_0. \tag{3.6}$$

It is not difficult to see that a basis of $\boldsymbol{H}$'s invariant subspace corresponding the eigenvalue 0 consists of the columns of

$$\begin{bmatrix} 0 \\ \boldsymbol{U}_0 \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{V}_0 \\ 0 \end{bmatrix}. \tag{3.7}$$

Now we shall translate them back to for $H - \lambda E$. Let $U_0 = C^{-\mathrm{H}}\boldsymbol{U}_0$ and $V_0 = D^{-\mathrm{H}}\boldsymbol{V}_0$. Equations (3.5) and (3.6) become

$$U_0 \in \mathbb{C}^{n\times(n-r)}, \quad \operatorname{rank}(U_0) = n - r, \quad KU_0 = 0, \quad MV_0 = E_-U_0, \tag{3.8}$$

and a basis of the generalized invariant subspace corresponding the eigenvalue 0 for $H - \lambda E$ consists of the columns of

$$\begin{bmatrix} 0 \\ U_0 \end{bmatrix}, \quad \begin{bmatrix} V_0 \\ 0 \end{bmatrix}. \tag{3.9}$$

Note it is implied that $V_0 \in \mathbb{C}^{n\times(n-r)}$ by (3.8).

Theorem 3.2. *Suppose* (3.4) *holds and let* U_0, $V_0 \in \mathbb{C}^{n\times(n-r)}$ *be determined by* (3.8), *and*[4)]

$$\mathbb{J} = \{1, 2, \ldots, n-r\}.$$

Then (3.1) *holds with* $k = n - r$.

Proof. The equations in (3.8) yield

$$KU_0 = C\Psi\Lambda^2\Psi^{\mathrm{H}}C^{\mathrm{H}}U_0 = 0, \quad MV_0 = D\Phi\Phi^{\mathrm{H}}D^{\mathrm{H}}V_0 = E_-U_0. \tag{3.10}$$

Partition Φ and Ψ as in (3.2) with $\mathbb{J} = \{1, 2, \ldots, n-r\}$, and accordingly

$$\Lambda = \mathrm{diag}(\underbrace{0, \ldots, 0}_{n-r}, \Lambda_2) \quad \text{with} \quad \Lambda_2 = \mathrm{diag}(\lambda_{n-r+1}, \ldots, \lambda_n).$$

We get

$$\Psi_2^{\mathrm{H}}C^{\mathrm{H}}U_0 = 0, \quad D^{\mathrm{H}}V_0 = \Phi^{-\mathrm{H}}\Phi^{-1}D^{-1}E_-U_0 = \Psi\Psi^{\mathrm{H}}C^{\mathrm{H}}U_0 = \Psi_1\Psi_1^{\mathrm{H}}C^{\mathrm{H}}U_0.$$

Since $n - r = \mathrm{rank}(\Psi^{\mathrm{H}}C^{\mathrm{H}}U_0) = \mathrm{rank}(\Psi_1^{\mathrm{H}}C^{\mathrm{H}}U_0)$, we conclude

$$\mathcal{R}(D^{\mathrm{H}}V_0) = \mathcal{R}(\Psi_1) = \mathcal{R}(\Psi_{(:,1:n-r)}),$$

as expected. □

Next, we consider positive eigenvalues.

Theorem 3.3. *Let* $Z = \left[\begin{smallmatrix} V \\ U \end{smallmatrix}\right]$ *be a basis matrix of an eigenspace subspace* $\mathcal{Z}$ *of* $H - \lambda E$ *corresponding to its* k *positive eigenvalues* λ_{i_j} *for* $1 \leqslant j \leqslant k$, *and let* $\mathbb{J}$, $\mathbb{J}_c$, Φ_i, *and* Ψ_i *be the ones in* Theorem 3.1 *and its proof. Then* (3.1) *holds.*

Proof. First consider the case when $\lambda_i \neq \lambda_j$ for all $i \in \mathbb{J}$ and $j \in \mathbb{J}_c$. Then the invariant subspace associated with all λ_{i_j} is unique. By Theorem 2.1, $\Gamma^{-\mathrm{H}}\left[\begin{smallmatrix} \Psi_1\Lambda_1 \\ \Phi_1 \end{smallmatrix}\right]$ is also a basis matrix of the invariant subspace, where $\Lambda_1 = \mathrm{diag}(\lambda_{i_1}, \ldots, \lambda_{i_k})$. Therefore there exists a nonsingular $Q \in \mathbb{R}^{k\times k}$ such that

$$Z = \begin{bmatrix} V \\ U \end{bmatrix} = \Gamma^{-\mathrm{H}} \begin{bmatrix} \Psi_1\Lambda_1 \\ \Phi_1 \end{bmatrix} Q, \tag{3.11}$$

which implies $V = D^{-\mathrm{H}}\Psi_1\Lambda_1 Q$, i.e., $D^{\mathrm{H}}V = \Psi_{(:,\mathbb{J})}\Lambda_1 Q$. Hence (3.1) holds because $\Lambda_1 Q$ is nonsingular.

Now suppose some λ_i for $i \in \mathbb{J}$ are equal to certain λ_j for $j \in \mathbb{J}_c$. If such a λ_i repeats s times among $\{\lambda_{i_1}, \ldots, \lambda_{i_k}\}$, then there are s linearly independent eigenvectors associated with this λ_i that live in $\mathcal{Z}$ and we can always adjust the decompositions in (2.7) and thus (2.8) such that $\Gamma^{-\mathrm{H}}\left[\begin{smallmatrix} \Psi_1\Lambda_1 \\ \Phi_1 \end{smallmatrix}\right]$ has s columns that span the same subspace as the ℓ linearly independent eigenvectors associated with this λ_i. We perform such an adjustment for every λ_i as such. At the end, (3.11) still holds. So does (3.1). □

Combining Theorems 3.2 and 3.3, we propose the following deflation strategy. Assume, as usual, that $K \succeq 0$, $M \succ 0$. Again, there are two cases.

Case 1. *Deflate eigenvalue* 0. Let $r = \mathrm{rank}(K)$ which either is known *a priori* in which case a U_0 satisfying the first three requirements in (3.8) may or may not be known, or has to be numerically determined. The latter case and the first case without the knowledge of a U_0 satisfying the first three requirements in (3.8) entail an efficient numerical way to determine r and/or a U_0. An (adaptive) symmetric block Lanczos method or a variation on K can be used to see if its smallest eigenvalues are zeros. The use of a block version is to make sure the multiplicity $n - r$ of 0 as an eigenvalue of K is correctly found if K is singular. For this to happen, the block size has to be no smaller than $n - r$, and therefore some adaptive strategy in updating block size may have to be used [35]. The symmetric Lanczos methods are usually very fast in computing extreme eigenpairs [18, 29, 36]. Once a U_0 satisfying the first

[4)] We know $0 = \lambda_1 = \lambda_2 = \cdots = \lambda_{n-r} < \lambda_{n-r+1} \leqslant \cdots \leqslant \lambda_n$.

three requirements in (3.8) is computed, the last equation there can be solved for V_0 by the linear CG method.

In any case, we have $r = \mathrm{rank}(K)$, and in the case $0 < r < n$ (the usual case is $n - r$ is small, like under 10) we have U_0 and V_0 that solve (3.8). Let

$$V_{\mathrm{cvgd}} \text{ be an empty array if } r = n \text{ or } V_0 \text{ if } 0 < r < n, \tag{3.12}$$

and symbolically, let

$$\underline{K} = K + \xi(E_+ V_{\mathrm{cvgd}})(E_+ V_{\mathrm{cvgd}})^{\mathrm{H}}, \quad \underline{H} = \begin{bmatrix} 0 & \underline{K} \\ M & 0 \end{bmatrix}, \tag{3.13}$$

where $\xi > 0$ is selected sufficiently large. By "*symbolically*", we mean $\underline{K}$ should never be explicitly computed, but it exists in this form through K, E_+, and V_{cvgd}.

Case 2. *Deflate positive eigenvalue.* Now we work with $\underline{H} - \lambda E$ to compute its few smallest positive eigenvalues and associated eigenvectors. Anytime, a new computed eigenpair $(\mu, \boldsymbol{z})$ is determined to converge within a preset tolerance, V_{cvgd} is expanded as

$$V_{\mathrm{cvgd}} \leftarrow [V_{\mathrm{cvgd}}, \boldsymbol{v}]$$

and update $\underline{K}$ and $\underline{H}$ accordingly as in (3.13), where $\boldsymbol{v}$ is the first half of $\boldsymbol{z}$.

4 Extended LOBP4dCG

4.1 Projection subspace extension

In [3,4], we proposed the *locally optimal block preconditioned* 4*-d CG method* (LOBP4dCG) for simultaneous computation of the first few smallest eigenpairs of $H - \lambda E$ with positive sign. In what follows, we will introduce a variation of it through incorporation of the idea of augmented subspaces introduced in [10] (see also [17,22,25,34]). This new variation will be called the *extended locally optimal block preconditioned* 4*-d CG* (*ELOB4dCG*) *algorithm.*

Recall $K \succeq 0$ and $M \succ 0$. We shall first deflate away eigenvalue 0, if any, of $H - \lambda E$, and then set V_{cvgd} as in (3.12). We update K to $\underline{K}$ and H to $\underline{H}$ as in (3.13), accordingly. Now and from this point forward, both $\underline{K} \succ 0$ and $M \succ 0$ hold.

To explain the ELOBP4dCG method, we switch the LREP (1.1) to the equivalent eigenvalue problem for

$$\boldsymbol{A} - \lambda \boldsymbol{B} \equiv \begin{bmatrix} M & 0 \\ 0 & K \end{bmatrix} - \lambda \begin{bmatrix} 0 & E_- \\ E_+ & 0 \end{bmatrix}. \tag{4.1}$$

This is a positive semidefinite pencil in the sense that $\boldsymbol{A} - \lambda_0 \boldsymbol{B} \succeq 0$ for $\lambda_0 = 0$ [19, 20] since we assume $K \succeq 0$ and $M \succ 0$. When convenient, we will turn to $\boldsymbol{A} - \lambda \boldsymbol{B}$ in our later development.

The key idea for the i-th iterative step in the LOBP4dCG method is to seek the best possible approximations in the subspace

$$\mathrm{span}\left\{ \begin{bmatrix} y_j^{(i)} \\ 0 \end{bmatrix}, \begin{bmatrix} y_j^{(i-1)} \\ 0 \end{bmatrix}, \begin{bmatrix} q_j \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ p_j \end{bmatrix}, \begin{bmatrix} 0 \\ x_j^{(i)} \end{bmatrix}, \begin{bmatrix} 0 \\ x_j^{(i-1)} \end{bmatrix} \text{ for } 1 \leqslant j \leqslant n_b \right\}, \tag{4.2}$$

where n_b is the block size, the superscripts $(i-1)$ and (i) indicate that they are for the $(i-1)$-th and i-th iterative steps, respectively, and

$$\begin{bmatrix} q_j \\ p_j \end{bmatrix} = \Pi_j \left. \begin{bmatrix} \nabla_y \rho \\ \nabla_x \rho \end{bmatrix} \right|_{(x,y)=(x_j^{(i)}, y_j^{(i)})}, \tag{4.3}$$

and Π_j is a preconditioner that aims at speeding up the convergence for j-th eigenpair of $\underline{H}-\lambda E$, $\nabla_x\rho$ and $\nabla_y\rho$ are partial gradients of $\rho(x,y)$:

$$\nabla_x\rho=\frac{1}{x^{\mathrm{H}}E_+y}[Kx-\rho(x,y)\,E_+y],\quad \nabla_y\rho=\frac{1}{x^{\mathrm{H}}E_+y}[My-\rho(x,y)\,E_-x]. \tag{4.4}$$

But for the first iterative step (i.e., $i=0$), those vectors with the superscript $(i-1)=-1$ must be removed from the list in (4.2) as they are not available yet.

There is a close relation between these two partial gradients and the residual,

$$\boldsymbol{A}z-\rho(x,y)\boldsymbol{B}z\equiv\begin{bmatrix}M&0\\0&K\end{bmatrix}\begin{bmatrix}y\\x\end{bmatrix}-\rho(x,y)\begin{bmatrix}0&E_-\\E_+&0\end{bmatrix}\begin{bmatrix}y\\x\end{bmatrix}=x^{\mathrm{H}}E_+y\begin{bmatrix}\nabla_y\rho\\\nabla_x\rho\end{bmatrix}. \tag{4.5}$$

Namely the block vector obtained by stacking $\nabla_x\rho$ over $\nabla_y\rho$ is parallel to the residual. Write

$$z_j^{(i)}=\begin{bmatrix}y_j^{(i)}\\x_j^{(i)}\end{bmatrix}.$$

The spanning vectors in (4.2) are basically obtained by "breaking" up each vector of the following vectors

$$z_j^{(i)},\ \Pi_j[\boldsymbol{A}-\rho(x_j^{(i)},y_j^{(i)})\boldsymbol{B}]z_j^{(i)},\ z_j^{(i-1)}\quad\text{for}\quad 1\leqslant j\leqslant n_b \tag{4.6}$$

into two "half" vectors. Observe that the first two vectors are the spanning vectors of Krylov subspace $\mathcal{K}_2(\Pi_j[\boldsymbol{A}-\rho(x_j^{(i)},y_j^{(i)})\boldsymbol{B}],z_j^{(i)})$ of order 2. This is an important observation because it points to a natural way to expand the search subspaces: using a higher order Krylov subspace [10, 17]. The ELOBP4dCG follows the framework of LOBP4dCG [3, 4] but replace the searching subspace (4.2) by the one spanned by the vectors obtained by "breaking" up each vector in

$$\begin{aligned}&\text{the spanning vectors of Krylov subspace } \mathcal{K}_m(\Pi_j[\boldsymbol{A}-\rho(x_j^{(i)},y_j^{(i)})\boldsymbol{B}],z_j^{(i)})\\&\quad\text{of order } m,\text{and } z_j^{(i-1)} \text{ for } 1\leqslant j\leqslant n_b\end{aligned} \tag{4.7}$$

into two "half" vectors, where $m\geqslant 2$ is a parameter to be selected.

To seek the best possible approximations in the extended subspace, we first compute a basis matrix $\left[\begin{smallmatrix}V_1\\U_1\end{smallmatrix}\right]$ of

$$\bigcup_{j=1}^{n_b}\mathcal{K}_m(\Pi_j[\boldsymbol{A}-\rho(x_j^{(i)},y_j^{(i)})\boldsymbol{B}],z_j^{(i)}), \tag{4.8}$$

and then compute two basis matrices V and U for the subspaces

$$\mathcal{V}=\mathcal{R}(V_1)\cup\operatorname{span}\{y_j^{(i-1)},\ \text{for } 1\leqslant j\leqslant n_b\}, \tag{4.9a}$$

$$\mathcal{U}=\mathcal{R}(U_1)\cup\operatorname{span}\{x_j^{(i-1)},\ \text{for } 1\leqslant j\leqslant n_b\}, \tag{4.9b}$$

respectively, and finally solve the projected eigenvalue problem for

$$\begin{aligned}H_{\mathrm{SR}}-\lambda E_{\mathrm{SR}}:&=\begin{bmatrix}U&0\\0&V\end{bmatrix}^{\mathrm{H}}(H-\lambda E)\begin{bmatrix}V&0\\0&U\end{bmatrix}\\&=\begin{bmatrix}0&U^{\mathrm{H}}KU\\V^{\mathrm{H}}MV&0\end{bmatrix}-\lambda\begin{bmatrix}U^{\mathrm{H}}E_+V&0\\0&V^{\mathrm{H}}E_-U\end{bmatrix}\end{aligned} \tag{4.10}$$

to construct new approximations [4, Theorem 8.2]. For small m and modest n_b, the major cost in computing new approximation lies in computing the basis matrix V and U and the projected matrices H_{SR} and E_{SR}.

4.2 The projected eigenvalue problem

In [3, 4], the eigenvalue problem (4.10) is solved as follows. Suppose that $W = U^{\mathrm{H}}E_+V$ is square and nonsingular, and factorize $W = W_1^{\mathrm{H}}W_2$, where W_i are nonsingular for $i = 1, 2$. Compute the eigenvalues of [5)]

$$\hat{H}_{\mathrm{SR}} = \begin{bmatrix} 0 & W_1^{-\,\mathrm{H}}U^{\mathrm{H}}KUW_1^{-1} \\ W_2^{-\,\mathrm{H}}V^{\mathrm{H}}MVW_2^{-1} & 0 \end{bmatrix} \tag{4.11}$$

as the new approximations for some eigenvalues of $H - \lambda E$, and construct approximate eigenvectors accordingly. This is done by calling, e.g., LAPACK's subroutine `xSYGVD` [1]. As assured by [4, Theorem 8.2], computed eigenvalue and eigenvector approximations this way are the same as those from (4.10) in the case when $W = U^{\mathrm{H}}E_+V$ is square and nonsingular. There is a complicated procedure to deal with the case when $W = U^{\mathrm{H}}E_+V$ is either not square or singular in [3, 4], too. It turns out the resulting approximations are also the same as would be obtained from working with (4.10).

But here we propose to solve the eigenvalue problem (4.10) differently. We make the basis matrix U with $\underline{K}$-orthonormal columns and the basis matrix V with M-orthonormal columns. This can be done by the modified Gram-Schmidt orthogonalization in the $\underline{K}$-inner product and M-inner product, respectively.

Generically, $\dim(\mathcal{U}) = \dim(\mathcal{V}) = mn_b$ for $i = 0$ and $(m + 1)n_b$ for $i > 0$, and thus U and V make $U^{\mathrm{H}}\underline{K}U = V^{\mathrm{H}}MV = I$. There is actually a good reason to enforce $U^{\mathrm{H}}\underline{K}U = V^{\mathrm{H}}MV$ as suggested by Theorem 2.4. But in general, possibly

$$k := \dim(\mathcal{U}) \neq \dim(\mathcal{V}) =: \ell.$$

In what follows, we make no assumption on whether $k = \ell$ or not.

The next approximations are then computed by solving the eigenvalue problem for

$$\begin{aligned} \underline{H}_{\mathrm{SR}} - \lambda E_{\mathrm{SR}} :&= \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix}^{\mathrm{H}} (\underline{H} - \lambda E) \begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} \\ &= \begin{bmatrix} 0 & U^{\mathrm{H}}\underline{K}U \\ V^{\mathrm{H}}MV & 0 \end{bmatrix} - \lambda \begin{bmatrix} U^{\mathrm{H}}E_+V & \\ & V^{\mathrm{H}}E_-U \end{bmatrix} \\ &= \begin{bmatrix} 0 & I_k \\ I_\ell & 0 \end{bmatrix} - \lambda \begin{bmatrix} U^{\mathrm{H}}E_+V & 0 \\ 0 & V^{\mathrm{H}}E_-U \end{bmatrix}. \end{aligned} \tag{4.12}$$

The eigenvalue problem for (4.12) is equivalent to the eigenvalue problem for

$$I_{\ell+k} - \lambda \begin{bmatrix} 0 & V^{\mathrm{H}}E_-U \\ U^{\mathrm{H}}E_+V & 0 \end{bmatrix},$$

which can be solved through the singular value decomposition (SVD) of $V^{\mathrm{H}}E_-U \in \mathbb{C}^{\ell\times k}$. Let the SVD of $V^{\mathrm{H}}E_-U$ be

$$V^{\mathrm{H}}E_-U = \hat{V}\Omega\hat{U}^{\mathrm{H}}, \quad \Omega = \mathrm{diag}(\omega_1, \omega_2, \ldots),$$

where $\hat{V} \in \mathbb{C}^{\ell\times\ell}$ and $\hat{U} \in \mathbb{C}^{k\times k}$ is unitary, $\Omega \in \mathbb{C}^{\ell\times k}$ is a leading diagonal matrix, where $\omega_1 \geqslant \omega_2 \geqslant \cdots$ are the singular values. It is reasonable to assume

$$\min\{\ell, k\} \geqslant n_b \tag{4.13}$$

under which $\omega_{n_b} > 0$. Write $\hat{V} = [\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_\ell]$ and $\hat{U} = [\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_k]$. Then the first n_b smallest positive eigenvalues of $\underline{H}_{\mathrm{SR}} - \lambda E_{\mathrm{SR}}$ are

$$0 < \omega_1^{-1} \leqslant \omega_2^{-1} \leqslant \cdots \leqslant \omega_{n_b}^{-1}$$

[5)] This is inconsistent with what we did in [2–4], where this is H_{SR} which we have used in (4.10) for something (subtly) different.

with corresponding eigenvectors

$$\begin{bmatrix}\hat{v}_j\\ \hat{u}_j\end{bmatrix} \quad \text{for} \quad 1 \leqslant j \leqslant n_b.$$

Consequently, the new approximate eigenvalues for $H - \lambda E$ are ω_j^{-1} with corresponding approximate eigenvectors

$$\hat{z}_j = \begin{bmatrix}V & 0\\ 0 & U\end{bmatrix}\begin{bmatrix}\hat{v}_j\\ \hat{u}_j\end{bmatrix} = \begin{bmatrix}V\hat{v}_j\\ U\hat{u}_j\end{bmatrix} \quad \text{for} \quad 1 \leqslant j \leqslant n_b. \tag{4.14}$$

The deflation strategy discussed in Section 3 can be implemented to deflate away any converged eigenpairs $(\omega_j^{-1}, \hat{z}_j)$ by appending to V_{cvgd} the first "half" of all converged $\hat{z}_j$ and updating $\underline{K}$ and $\underline{H}$ accordingly as in (3.13) symbolically. Now there is another important decision to be made anytime a deflation is completed. That is to either reduce the block size n_b by the number of eigenpairs deflated or append vectors to $\{y_j^{(i)}\}$ and to $\{x_j^{(i)}\}$ to maintain the block size n_b the same as before deflation. But what vectors should we use for appending? There are natural choices: if s is the number of converged eigenpairs at the current iterative step, we should use the top half and bottom half of $\hat{z}_j$ in (4.14) for $n_b + 1 \leqslant j \leqslant n_b + s$ to append $\{y_j^{(i)}\}$ and $\{x_j^{(i)}\}$, respectively.

4.3 Algorithm

Algorithm 1 presents the framework of ELOBP4dCG, with the shifting deflation technique included. A few comments on implementation details are in order.

1. Often K is known to be positive definite as those from applications [27]. Otherwise, at line 1, as we mentioned in Section 3, an (adaptive) block symmetric Lanczos method on K is usually rather efficient for the purpose. It will compute $r = \text{rank}(K)$ and, in the case $r < n$, U_0 as well. For numerical stability, we can make U_0 to have orthonormal columns. The linear equation $MV_0 = E_-U_0$ can be solved by the linear conjugrate gradient method.

2. Line 2 is about processing the initial input

$$Z_0 = \begin{bmatrix}Y_0\\ X_0\end{bmatrix}$$

with $X_0 \in \mathbb{C}^{n\times n_b}$;

(a) $\underline{K}$-orthogonalize the columns of X_0 and M-orthogonalize the columns of Y_0 to get X and Y, respectively;

(b) compute $W = (X^{\text{H}}Y + Y^{\text{H}}X)/2$,

$$R = \begin{bmatrix}M & \\ & \underline{K}\end{bmatrix}\begin{bmatrix}Y\\ X\end{bmatrix}W - \begin{bmatrix} & E_-\\ E_+ & \end{bmatrix};$$

the rationale to justify using this W at line 4 is the fact that

$$\underset{W}{\text{argmin}}\left\|\begin{bmatrix}M & \\ & \underline{K}\end{bmatrix}\begin{bmatrix}Y\\ X\end{bmatrix}W - \begin{bmatrix} & E_-\\ E_+ & \end{bmatrix}\begin{bmatrix}Y\\ X\end{bmatrix}\right\|_{\text{F}} = \frac{1}{2}(X^{\text{H}}Y + Y^{\text{H}}X),$$

where $\|\cdot\|_{\text{F}}$ is the matrix Frobenius norm, assuming $X^H\underline{K}X = Y^HMY = I$;

(c) precondition R; Since R is the residual corresponding to a subspace, not individual eigenpairs, using a single preconditioner is recommended such as one of those mention in the next comment: $R \leftarrow \Pi R$;

(d) calculate a $\underline{K}$-orthonormal basis matrix $U \in \mathbb{C}^{n\times k}$ of $\mathcal{R}([X, R_{(n+1:2n,:)}])$ and M-orthonormal basis matrix $V \in \mathbb{C}^{n\times \ell}$ of $\mathcal{R}([Y, R_{(1:n,:)}])$ such that $\mathcal{R}(U_{(:,1:n_b)}) = \mathcal{R}(X)$ and $\mathcal{R}(V_{(:,1:n_b)}) = \mathcal{R}(Y)$;

(e) complete the rest as in lines 6–15.

3. At line 4, these preconditioners Π_j in the most general way are also dependent upon the iterative index i, but they are suppressed for the sake to presentation. But having a different preconditioner for each

Algorithm 1: The ELOBP4dCG method

Input: $Z_0 \in \mathbb{C}^{2n\times n_b}$ with $\mathrm{rank}(Z_0) = n_b$ such that $\mathcal{R}(Z_0)$ approximates the eigenspace of $H - \lambda E$ associated with its first n_b smallest positive eigenvalues with positive sign, sufficiently large $\xi > 0$ for use in deflation;

Output: the first n_{need} smallest nonnegative eigenpairs (λ_j, z_j) for $1 \leqslant j \leqslant n_{\mathrm{need}}$ (usually $n_{\mathrm{need}} \geqslant n_b$).

1: if K is not singular, set $n_{\mathrm{cvgd}} = 0$, $\Lambda_{\mathrm{cvgd}} = [\,]$, and $Z_{\mathrm{cvgd}} = [\,]$ (empty arrays);
or if K is singular, then solve (3.8) for U_0 and V_0, and set

$$n_{\mathrm{cvgd}} = n - r, \quad \Lambda_{\mathrm{cvgd}} = [\underbrace{0, \ldots, 0}_{n-r}], \quad V_{\mathrm{cvgd}} = V_0, \quad Z_{\mathrm{cvgd}} = \begin{bmatrix} V_0 \\ U_0 \end{bmatrix}$$

and update K to $\underline{K}$ in form and accordingly H to $\underline{H}$ in form as in (3.13);

2: preprocess initial approximation Z_0 by one step steepest descent method to give new approximations Z, $\Omega = \mathrm{diag}(\omega_1, \omega_2, \ldots)$, "difference" $\hat{Z}$ between Z and Z_0 (see detail in Comment 2);

3: **for** $i = 0, 1, \ldots$ while $n_{\mathrm{cvgd}} < n_{\mathrm{need}}$ **do**

4: compute a basis matrix $\left[\begin{smallmatrix} V_1 \\ U_1 \end{smallmatrix}\right]$ of $\bigcup_{j=1}^{n_b} \mathcal{K}_m(\Pi_j[\boldsymbol{A} - \omega_j^{-1}\boldsymbol{B}], Z_{(:,j)})$ such that the first n_b columns of the basis matrix span the same subspace as the one by the columns of Z;

5: calculate a $\underline{K}$-orthonormal basis matrix $U \in \mathbb{C}^{n\times k}$ of $\mathcal{R}([U_1, \hat{Z}_{(1:n,:)}])$ and M-orthonormal basis matrix $V \in \mathbb{C}^{n\times m}$ of $\mathcal{R}([V_1, \hat{Z}_{(n+1:2n,:)}])$ such that $\mathcal{R}(U_{(:,1:n_b)}) = \mathcal{R}([U_1]_{(:,1:n_b)})$ and $\mathcal{R}(V_{(:,1:n_b)}) = \mathcal{R}([V_1]_{(:,1:n_b)})$;

6: compute SVD $V^{\mathrm{H}} E_- U = \hat{V}\Omega\hat{U}^{\mathrm{H}}$, where $\Omega = \mathrm{diag}(\omega_1, \omega_2, \ldots)$ and $\omega_1 \geqslant \omega_2 \geqslant \cdots$;

7:
$$Z \leftarrow \begin{bmatrix} V\hat{V}_{(:,1:n_b)} \\ U\hat{U}_{(:,1:n_b)} \end{bmatrix}, \quad \hat{Z} \leftarrow \begin{bmatrix} V_{(:,n_b+1:\ell)}\hat{V}_{(n_b+1:\ell,1:n_b)} \\ U_{(:,n_b+1:k)}\hat{U}_{(n_b+1:k,1:n_b)} \end{bmatrix};$$

8: $R = \underline{\boldsymbol{A}}Z - \boldsymbol{B}Z\,\mathrm{diag}(\omega_1^{-1}, \ldots, \omega_{n_b}^{-1})$;

9: test convergence and let s be the number of converged eigenpairs $(\omega_j^{-1}, Z_{(:,j)})$;

10: **if** $s \geqslant 1$ **then**

11: lock up the converged, and update $\Lambda_{\mathrm{cvgd}} \leftarrow [\Lambda_{\mathrm{cvgd}}, [\cdots \omega_j^{-1} \cdots]]$, $Z_{\mathrm{cvgd}} \leftarrow [Z_{\mathrm{cvgd}}, [\cdots Z_{(:,j)} \cdots]]$, and $V_{\mathrm{cvgd}} \leftarrow [V_{\mathrm{cvgd}}, [\cdots Z_{(1:n,j)} \cdots]]$, and accordingly $\underline{K}$ and $\underline{H}$ as a result;

12: drop the s converged columns from Z, and replace the dropped columns in Z by

$$\begin{bmatrix} V\hat{V}_{(:,n_b+1:n_b+s)} \\ U\hat{U}_{(:,n_b+1:n_b+s)} \end{bmatrix};$$

(These replacing columns are approximate eigenvectors associated with approximate eigenvalues ω_j^{-1} for $n_b + 1 \leqslant j \leqslant n_b + s$.)

13: drop the converged ω_j from Ω and relabel the remaining ω_j to match well with the columns of the updated Z;

14: $n_{\mathrm{cvgd}} \leftarrow n_{\mathrm{cvgd}} + s$;

15: **end if**

16: **end for**

17: **return** n_{cvgd} approximate eigenpairs as extracted from Λ_{cvgd} and Z_{cvgd}.

targeted eigenpairs may not be the most numerically efficient way, not to mention letting preconditioners vary from one iterative step to the next, because effective preconditioners can be expensive to construct. A *generic* preconditioner to compute the eigenvalues of $\underline{H} - \lambda E$, or equivalently $\underline{\boldsymbol{A}} - \lambda\boldsymbol{B}$, near a prescribed point μ is $\Pi = (\underline{\boldsymbol{A}} - \mu\boldsymbol{B})^{-1}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are given by (4.1) and $\underline{\boldsymbol{A}}$ is $\boldsymbol{A}$ with K replaced by $\underline{K}$. Since we are particularly interested in the smallest eigenvalues with the positive sign, $\mu = 0$ is often an obvious choice. Then

$$\Pi \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} M^{-1} & 0 \\ 0 & \underline{K}^{-1} \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} M^{-1}q \\ \underline{K}^{-1}p \end{bmatrix} \tag{4.15}$$

which can be computed by using the linear conjugate gradient method [7, 11] on M and $\underline{K}$ separately. For $0 < \mu <$ (the smallest positive eigenvalue of $\underline{H} - \lambda E$), $\Pi \succ 0$ (see [15,16,20]), $\Pi[q^{\mathrm{H}}, p^{\mathrm{H}}]^{\mathrm{H}}$ can still be computed by the conjugate gradient method. Since we are updating $\underline{K}$ at line 11, the preconditioner Π actually changes automatically every time an eigenpair converges.

4. With all preconditioners Π_j being the same, there is a compact way to express the subspace at line 4. Define the linear operator

$$\mathscr{R} : Z \in \mathbb{C}^{n\times n_b} \to \mathscr{R}(Z) = \underline{\boldsymbol{A}}Z - \boldsymbol{B}Z\hat{\Lambda} \in \mathbb{C}^{n\times n_b},$$

where

$$\hat{\Lambda} = \mathrm{diag}(\omega_1^{-1}, \ldots, \omega_{n_b}^{-1}).$$

Then the subspace can be compactly written as

$$\mathcal{K}_m(\Pi\mathscr{R}, Z) = \mathrm{span}\{Z, \Pi\mathscr{R}(Z), \ldots, (\Pi\mathscr{R})^{m-1}(Z)\}, \tag{4.16}$$

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

$$(\Pi\mathscr{R})^2(Z) = \Pi\mathscr{R}(\Pi\mathscr{R}(Z)).$$

In our preliminary testing, we use only modest m, say $m \leqslant 5$, and compute V_1 and U_1 without doing any orthogonalization for speed consideration, i.e., simply as the top and bottom half submatrices of $[Z, \Pi\mathscr{R}(Z), \ldots, (\Pi\mathscr{R})^{m-1}(Z)]$, respectively. This is the similar notion of the recent work [5, 6] on communication-avoiding Krylov subspace methods.

5. At line 5, U and V can be computed by the modified Gram-Schmidt orthogonalization process in the $\underline{K}$- and M-inner product, respectively. This action is justified by Theorem 2.4.

6. $\hat{Z}$ at line 7 is a kind of difference of the newly computed approximation Z from the very previous approximation. This idea is borrowed from [12, 14].

7. Testing convergence at line 11 can be done by computing the normalized residual norms

$$\epsilon_j := \frac{\|\underline{H}z_j - \omega_j^{-1}Ez_j\|_1}{(\|\underline{H}\|_1 + \omega_j^{-1}\|E\|_1)\|z_j\|_1} \equiv \frac{\|R_{(:,j)}\|_1}{(\|\underline{\boldsymbol{A}}\|_1 + \omega_j^{-1}\|\boldsymbol{B}\|_1)\|z_j\|_1} \quad \text{with} \quad z_j = Z_{(:,j)},$$

where $\|\cdot\|_1$ is the ℓ_1-norm of a vector or the ℓ_1-operator norm of a matrix, and R is the one at line 8. If $\epsilon_j \leqslant$ rtol (a preset relative tolerance), then we claim (ω_j^{-1}, z_j) is converged.

5 Numerical examples

In this section, we present numerical results obtained within the MATLAB environment to illustrate the essential convergence behaviors of ELOBP4dCG. We show two testing examples: one from linear response analysis for Na_2 and the other for silane (SiH_4) compound, generated by the turboTDDFT code in QUANTUM ESPRESSO (QE), an electronic structure calculation code that implements density functional theory (DFT) using plane-waves as the basis set and pseudopotentials [9]. Both examples have $E_{\pm} = I$. Such small molecules are often used as benchmark tests to assess various simulation models, functionals and methods (see [21]). For Na_2, the order of the symmetric positive definite matrices K and M is 1,862. Consequently, the dimension of H is 3,724. For SiH_4, the order of the symmetric positive definite matrices K and M is 5,660. Consequently, the dimension of H is 11,320.

Our goal is to compute ten smallest positive eigenvalues $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_{10}$ and corresponding eigenvectors $z_1, z_2, \ldots, z_{10}$ of H, with block size $n_b = 4$, with or without a preconditioner, and $m = 2$ (simply LOBP4dCG) or 3. When with a preconditioner, we simply use the generic one

$$\Pi = \underline{\boldsymbol{A}}^{-1} = \begin{bmatrix} M^{-1} & 0 \\ 0 & \underline{K}^{-1} \end{bmatrix}. \tag{5.1}$$

The preconditioned search vectors q_i and p_i are computed by using the linear CG method [7,11] to solve the associated linear systems of equations (see (4.15)). Often very crude approximations of q_i and p_i are good enough. In this example, the linear CG iterations are set with the stopping tolerance 10^{-2} or maximal 20 iterations.

We should note that the generic preconditioner Π is not the natural preconditioner for these examples. For the plane wave-based calculations, it is more natural to use a proper scaled diagonal-like preconditioner proposed in [30]. Excellent performance of such preconditioner in the the turboTDDFT code has been reported [28]. But K and M in our test were outputted unnaturally from turboTDDFT runs because K and M were not needed explicitly and they only existed in certain structural form to allow matrix-vector products with them be performed at very fast speed. The scaled diagonal preconditioner in [30], however, is only possible when K and M live in the structural form which is lost once K and M are outputted explicitly as matrices. It will be a future work of ours to implement our solver here with a scaled diagonal-like preconditioner similar to the one in [30].

Figures 1 and 2 show iterative history plots of LOBP4dCG and ELOBP4dCG on Na_2 and SiH_4 problems, respectively. We see dramatic reductions in the numbers of iterations required in going from "without preconditioning" to "with preconditioning" and in going from $m = 2$ to $m = 3$. The powers of using a preconditioner and extending the searching subspace are in display prominently.

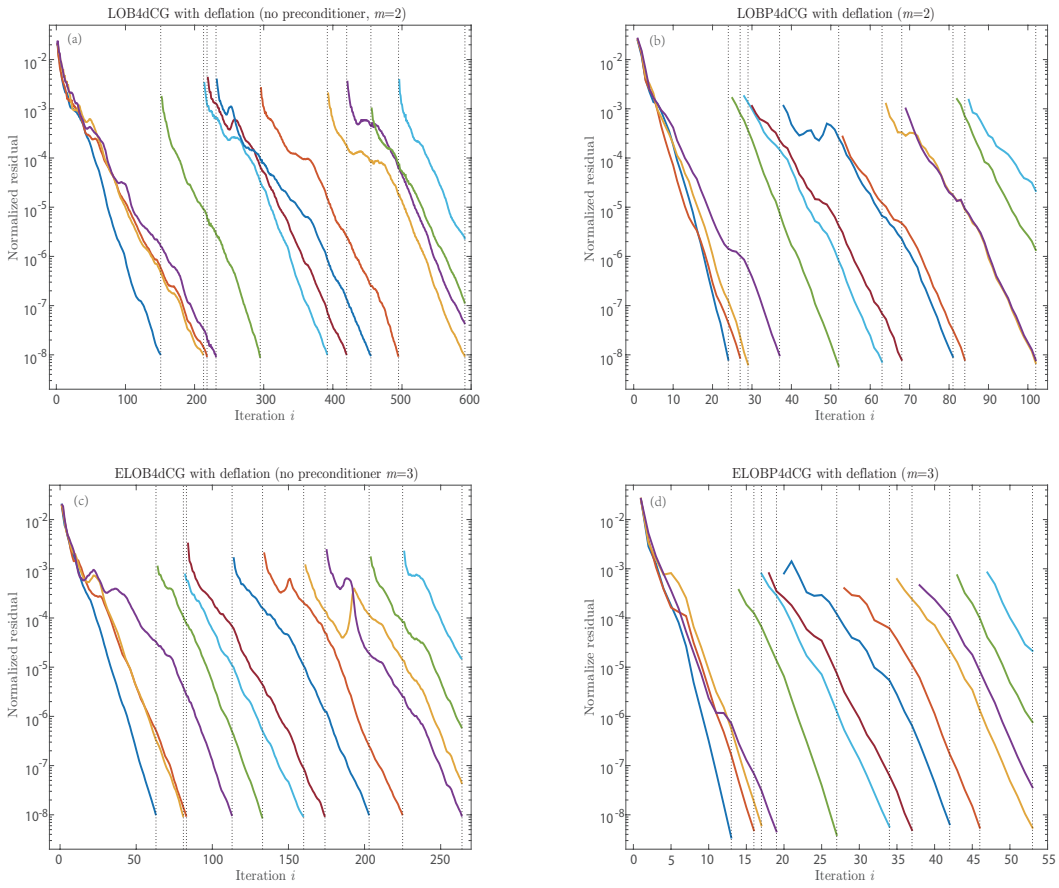

Figure 1 Na_2 linear response eigenvalue problem. (a) LOB4dCG ($m = 2$, without preconditioning); (b) LOBP4dCG ($m = 2$, with preconditioning); (c) ELOB4dCG ($m = 3$, without preconditioning); (d) ELOBP4dCG ($m = 3$, with preconditioning). For each iteration i, there are 4 normalized residuals which move down as i goes. As soon as one reaches 10^{-8}, the corresponding eigenpair is deflated and locked away, and a new residual shows up at the top

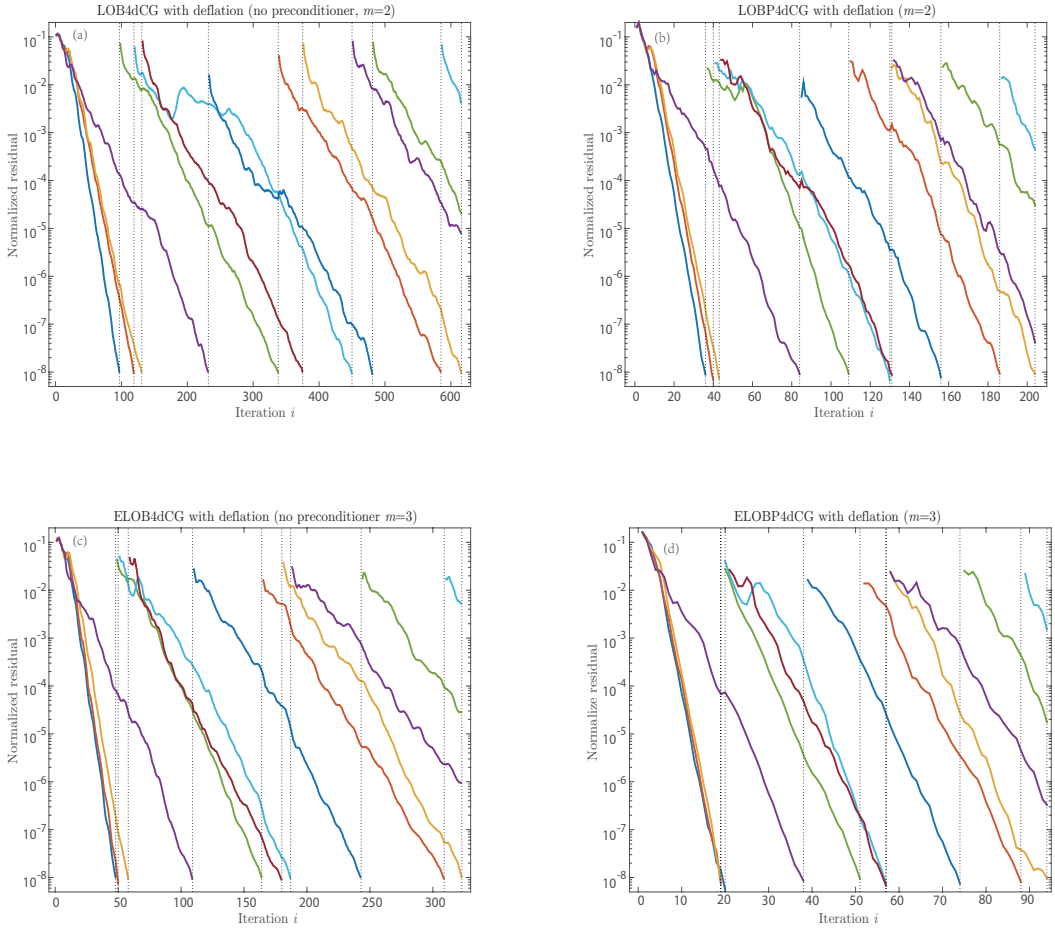

Figure 2 SiH_4 linear response eigenvalue problem. (a) LOB4dCG ($m = 2$, without preconditioning); (b) LOBP4dCG ($m = 2$, with preconditioning); (c) ELOB4dCG ($m = 3$, without preconditioning); (d) ELOBP4dCG ($m = 3$, with preconditioning). For each iteration i, there are 4 normalized residuals which move down as i goes. As soon as one reaches 10^{-8}, the corresponding eigenpair is deflated and locked away, and a new residual shows up at the top

6 Conclusions

Deflation techniques are ways to remove already computed eigenpairs from being possibly recomputed. It is critical for an efficient eigensolver. In order to preserve fast matrix-vector multiplications in large scale eigenvalue computations, a common framework is some kind of low-rank modification to the original matrix so as to displace the computed eigenvalue while keeping unknown eigenvalues unchanged or shifted in an explicitly known way. There are two major contributions in this paper that improve our earlier LOBP4dCG for the linear response eigenvalue problem [3, 4]. One is a shifting deflation technique and the other is an idea of extending the search subspace. Both have been incorporated into LOBP4dCG to give a new method called the ELOBP4dCG. Numerical examples prominently demonstrate the efficiency of the deflation technique and show dramatic increase of the rate of convergence.

Acknowledgements This work was supported by National Science Foundation of USA (Grant Nos. DMS-1522697, CCF-1527091, DMS-1317330 and CCF-1527091) and National Natural Science Foundation of China (Grant No. 11428104). The authors express their gratitude to the referees for their valuable comments and suggestions to improve the presentation of the paper.

References

1 Anderson E, Bai Z, Bischof C, et al. LAPACK Users' Guide, 3rd ed. Philadelphia: SIAM, 1999
2 Bai Z, Li R-C. Minimization principles for the linear response eigenvalue problem, I: Theory. SIAM J Matrix Anal Appl, 2012, 33: 1075–1100
3 Bai Z, Li R-C. Minimization principle for the linear response eigenvalue problem, II: Computation. SIAM J Matrix Anal Appl, 2013, 34: 392–416
4 Bai Z, Li R-C. Minimization principles and computation for the generalized linear response eigenvalue problem. BIT Numer Math, 2014, 54: 31–54
5 Carson E. Communication-avoiding Krylov Subspace Methods in Theory and Practice. PhD thesis. Berkeley: University of California, 2015
6 Carson E, Demmel J. Accuracy of the s-step Lanczos method for the symmetric eigenproblem in finite precision. SIAM J Matrix Anal Appl, 2015, 36: 793–819
7 Demmel J. Applied Numerical Linear Algebra. Philadelphia: SIAM, 1997
8 Flaschka U, Lin W-W, Wu J-L. A KQZ algorithm for solving linear-response eigenvalue equations. Linear Algebra Appl, 1992, 165: 93–123
9 Giannozzi P. QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. J Phys Condensed Matter, 2009, 21: 395502
10 Golub G, Ye Q. An inverse free preconditioned Krylov subspace methods for symmetric eigenvalue problems. SIAM J Sci Comput, 2002, 24: 312–334
11 Golub G H, Van Loan C F. Matrix Computations, 3rd ed. Baltimore-Maryland: Johns Hopkins University Press, 1996
12 Hetmaniuk U, Lehoucq R. Basis selection in LOBPCG. J Comput Phys, 2006, 218: 324–332
13 Imakura A, Du L, Sakurai T. Error bounds of Rayleigh-Ritz type contour integral-based eigensolver for solving generalized eigenvalue problems. Numer Algor, 2016, 71: 103–120
14 Knyazev A V. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. SIAM J Sci Comput, 2001, 23: 517–541
15 Kovač-Striko J, Veselić K. Trace minimization and definiteness of symmetric pencils. Linear Algebra Appl, 1995, 216: 139–158
16 Lancaster P, Ye Q. Variational properties and Rayleigh quotient algorithms for symmetric matrix pencils. Oper Theory Adv Appl, 1989, 40: 247–278
17 Li R-C. Rayleigh quotient based optimization methods for eigenvalue problems. In: Bai Z, Gao W, Su Y F, eds. Matrix Functions and Matrix Equations, vol. 19. Series in Contemporary Applied Mathematics. Singapore: World Scientific, 2015, 76–108
18 Li R-C, Zhang L-H. Convergence of block Lanczos method for eigenvalue clusters. Numer Math, 2015, 131: 83–113
19 Liang X, Li R-C. Extensions of Wielandt's min-max principles for positive semi-definite pencils. Linear Multilinear Algebra, 2014, 62: 1032–1048
20 Liang X, Li R-C, Bai Z. Trace minimization principles for positive semi-definite pencils. Linear Algebra Appl, 2013, 438: 3085–3106
21 Marques M A, Castro A, Rubio A. Assessment of exchange-correlation functionals for the calculation of dynamical properties of small clusters in time-dependent density functional theory. J Chem Phys, 2001, 115: 3006–3014
22 Money J, Ye Q. EIGIFP: A MATLAB program for solving large symmetric generalized eigenvalue problems. ACM Trans Math Software, 2005, 31: 270–279
23 Olsen J, Jensen Aa H J, Jørgensen P. Solution of the large matrix equations which occur in response theory. J Comput Phys, 1988, 74: 265–282
24 Olsen J, Jørgensen P. Linear and nonlinear response functions for an exact state and for an MCSCF state. J Chem Phys, 1985, 82: 3235–3264
25 Quillen P, Ye Q. A block inverse-free preconditioned Krylov subspace method for symmetric generalized eigenvalue problems. J Comput Appl Math, 2010, 233: 1298–1313
26 Ring P, Schuck P. The Nuclear Many-Body Problem. New York: Springer-Verlag, 1980
27 Rocca D. Time-Dependent Density Functional Perturbation Theory: New algorithms with Applications to Molecular Spectra. PhD thesis. Trieste: The International School for Advanced Studies, 2007
28 Rocca D. Iterative diagonalization of non-hermitian eigenproblems in time-dependent density functional and many-body perturbation theory. Boston: Presentation at Session B39, the APS Marching Meeting, 2012
29 Saad Y. On the rates of convergence of the Lanczos and the block-Lanczos methods. SIAM J Numer Anal, 1980, 15: 687–706
30 Teter M, Payne M, Allan D. Solution of Schrödinger equation for large systems. Phys Rev B, 1989, 40: 12255–12263
31 Thouless D J. Vibrational states of nuclei in the random phase approximation. Nuclear Phys, 1961, 22: 78–95
32 Thouless D J. The Quantum Mechanics of Many-Body Systems. New York: Academic, 1972
33 Tsiper E V. Variational procedure and generalized Lanczos recursion for small-amplitude classical oscillations. JETP

Letters, 1999, 70: 751–755
34 Wen Z, Zhang Y. Block algorithms with augmented Rayleigh-Ritz projections for large-scale eigenpair computation. Technical report. ArXiv:1507.06078, 2015
35 Ye Q. An adaptive block Lanczos algorithm. Numer Algor, 1996, 12: 97–110
36 Zhou Y, Li R-C. Bounding the spectrum of large Hermitian matrices. Linear Algebra Appl, 2011, 435: 480–493