SIAM J. MATRIX ANAL. APPL.
Vol. 32, No. 2, pp. 642–663

PERTURBATION OF PARTITIONED HERMITIAN DEFINITE GENERALIZED EIGENVALUE PROBLEMS*

REN-CANG LI†, YUJI NAKATSUKASA‡, NINOSLAV TRUHAR§, AND SHUFANG XU¶

Abstract. This paper is concerned with the Hermitian definite generalized eigenvalue problem $A - \lambda B$ for block diagonal matrices $A = \mathrm{diag}(A_{11}, A_{22})$ and $B = \mathrm{diag}(B_{11}, B_{22})$. Both A and B are Hermitian, and B is positive definite. Bounds on how its eigenvalues vary when A and B are perturbed by Hermitian matrices are established. These bounds are generally of linear order with respect to the perturbations in the diagonal blocks and of quadratic order with respect to the perturbations in the off-diagonal blocks. The results for the case of no perturbations in the diagonal blocks can be used to bound the changes of eigenvalues of a Hermitian definite generalized eigenvalue problem after its off-diagonal blocks are dropped, a situation that occurs frequently in eigenvalue computations. The presented results extend those of Li and Li [*Linear Algebra Appl.*, 395 (2005), pp. 183–190]. It was noted by Stewart and Sun [*Matrix Perturbation Theory*, Academic Press, Boston, 1990] that different copies of a multiple eigenvalue may exhibit quite different sensitivities towards perturbations. We establish bounds to reflect that feature, too. We also derive quadratic eigenvalue bounds for diagonalizable non-Hermitian pencils subject to off-diagonal perturbations.

Key words. quadratic eigenvalue perturbation bound, generalized eigenvalue problem, multiple eigenvalue

AMS subject classifications. 15A22, 15A42, 65F15

DOI. 10.1137/100808356

1. Introduction.

The generalized eigenvalue problem (GEP) for the matrix pencil $A - \lambda B$ of two square matrices A and B is to determine nonzero vectors x and scalars μ such that $Ax = \mu Bx$. When it holds, we call μ a *(generalized) eigenvalue* and x its associated *eigenvector*.

In this paper, we are concerned with perturbations of a *Hermitian definite GEP* $A - \lambda B$ by which we mean both A and B are Hermitian and B is positive definite. Further we assume that A and B are already block diagonal:

$$A = \begin{array}{c} m \\ n \end{array} \overset{\begin{array}{cc} m & n \end{array}}{\begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix}}, \qquad B = \begin{array}{c} m \\ n \end{array} \overset{\begin{array}{cc} m & n \end{array}}{\begin{pmatrix} B_{11} & \\ & B_{22} \end{pmatrix}}. \tag{1.1}$$

When A and B are perturbed to

$$\tilde{A} \overset{\mathrm{def}}{=} A + E = \begin{pmatrix} A_{11} + E_{11} & E_{12} \\ E_{21} & A_{22} + E_{22} \end{pmatrix}, \qquad \tilde{B} \overset{\mathrm{def}}{=} B + F = \begin{pmatrix} B_{11} + F_{11} & F_{12} \\ F_{21} & B_{22} + F_{22} \end{pmatrix} \tag{1.2}$$

*Received by the editors September 13, 2010; accepted for publication (in revised form) April 20, 2011; published electronically June 30, 2011.

http://www.siam.org/journals/simax/32-2/80835.html

†Department of Mathematics, University of Texas at Arlington, P.O. Box 19408, Arlington, TX 76019-0408 (rcli@uta.edu). This author's research was supported in part by NSF grant DMS-0810506.

‡Department of Mathematics, University of California, Davis, CA 95616 (ynakatsukasa@ucdavis.edu). This author's research was supported in part by NSF grant OCI-0749217 and DOE grant DE-FC02-06ER25794.

§Department of Mathematics, University J. J. Strossmayer, Trg Ljudevita Gaja 6, 31000 Osijek, Croatia (ntruhar@mathos.hr). This author's research was supported in part by the "Passive control of mechanical models," grant 235-2352818-1042 of the Croatian MZOS.

¶School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China (xsf@math.pku.edu.cn). This author's research was supported in part by the NSF of China grant 10731060.

by two Hermitian matrices E and F, we are interested in bounding how much the eigenvalues of $A - \lambda B$ change. Two kinds of bounds will be established:

- bounds on the difference between the eigenvalues of $A - \lambda B$ and those of $\tilde{A} - \lambda\tilde{B}$;
- bounds on the difference between the eigenvalues of $A_{11} - \lambda B_{11}$ and some m eigenvalues of $\tilde{A} - \lambda\tilde{B}$.

There are two immediate applicable situations of such bounds. The first situation is when we have a GEP having A and B almost block diagonal, namely, in (1.2) $E_{ii} = F_{ii} = 0$ and all E_{ij} and F_{ij} for $i \neq j$ are tiny in magnitude relative to A_{ii} and B_{ii}. Such a situation can arise in practice, for example, when one uses a Jacobi-type algorithm [13, p. 353] that iteratively reduces both A and B to the diagonal form. Then it is natural to drop E_{ij} and F_{ij} for $i \neq j$ to decouple the GEP when they become relatively tiny. What is the effect of doing this? The second situation arises from (large scale) eigenvalue computations of a GEP, where one may have an approximate eigenspace. Projecting the GEP onto the approximate eigenspace would lead to (1.2) with again $E_{ii} = F_{ii} = 0$ for $i = 1, 2$ and some norm estimates on E_{ij} and F_{ij} for $i \neq j$ but usually unknown A_{22} and B_{22}. In such a case, we would like to estimate how well the eigenvalues of $A_{11} - \lambda B_{11}$ approximate some of those of $\tilde{A} - \lambda\tilde{B}$.

In the special case when $B = \tilde{B} = I_N$, the identity matrix can be well dealt with by some existing theories. For example, if all blocks in E have similar magnitudes, we may simply bound the eigenvalue changes using the norms of E by the Weyl–Lidskii theorem [1], [13], [14] (see also Lemma 2.1(a)); if E_{ij} for $i \neq j$ have much smaller magnitudes relative to E_{ii} for $i = 1, 2$, we may write

$$\hat{A} = A + \begin{pmatrix} & E_{12} \\ E_{21} & \end{pmatrix}, \qquad \tilde{A} = \hat{A} + \begin{pmatrix} E_{11} & \\ & E_{22} \end{pmatrix};$$

then the eigenvalue differences for A and $\tilde{A}$ can be bounded in two steps: bounding the differences for A and $\hat{A}$ and the differences for $\hat{A}$ and $\tilde{A}$. The eigenvalue differences for A and $\hat{A}$ are potentially of the second order in E_{ij} ($i \neq j$) and are no worse than of the first order in E_{ij} ($i \neq j$) [2], [6], [10], [15], while the eigenvalue differences for $\hat{A}$ and $\tilde{A}$ can be again bounded using the norms of E by the Weyl–Lidskii theorem.

The rest of the paper is organized as follows. In section 2 we present our main results: error bounds on the differences between the eigenvalues of GEPs (1.1) and (1.2). These error bounds are usually quadratic in the norms of E_{21} and F_{21}. The case when $B = I$ and $F_{21} = 0$ has been investigated in [6], and, in fact, our bounds here degenerate to the ones there in this case. A distinguished feature of GEPs is that different copies of a multiple eigenvalue may exhibit quite different sensitivities towards perturbations as noted in [14, p. 300]. This is studied in section 3. In section 4 we briefly consider perturbation of partitioned matrices in the non-Hermitian case.

Notation. Throughout this paper, $\|\cdot\|_2$, $\|\cdot\|_F$, $\sigma_{\min}(\cdot)$, and $\sigma_{\max}(\cdot)$ stand for the spectral norm, the Frobenius norm, the smallest singular value, and the largest singular value of a matrix, respectively. The superscript "$\cdot^*$" takes the complex conjugate transpose of a matrix or a vector. I_n (or simply I if its dimension is clear from the context) is the $n \times n$ identity matrix. We will use $\operatorname{eig}(A, B)$ to denote the set of eigenvalues of the matrix pencil $A - \lambda B$ and $\operatorname{eig}(A) \equiv \operatorname{eig}(A, I)$.

2. Perturbation bounds. Recall that B is assumed to be Hermitian and positive definite. The GEP $A - \lambda B$ is closely related to the *standard* Hermitian eigenvalue problem for $(B^{-1/2} A B^{-1/2}) - \lambda I_n$ in that

$$Ax = \mu Bx \text{ is equivalent to } (B^{-1/2}AB^{-1/2})(B^{1/2}x) = \mu(B^{1/2}x).$$

This implies, in particular, that all the eigenvalues of the GEP are real. If $\|F\|_2$ in (1.2) is sufficiently small, then $\tilde{A} - \lambda\tilde{B}$ will be Hermitian definite, too, and thus has only real eigenvalues.

Set $N = m + n$, and denote the eigenvalues of $A - \lambda B$ and $\tilde{A} - \lambda\tilde{B}$ by

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \quad \text{and} \quad \tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_N, \tag{2.1}$$

respectively. Define

$$\eta_i \overset{\text{def}}{=} \begin{cases} \min_{\mu_2 \in \text{eig}(A_{22},B_{22})} |\lambda_i - \mu_2| & \text{if } \lambda_i \in \text{eig}(A_{11}, B_{11}), \\ \min_{\mu_1 \in \text{eig}(A_{11},B_{11})} |\lambda_i - \mu_1| & \text{if } \lambda_i \in \text{eig}(A_{22}, B_{22}), \end{cases} \tag{2.2a}$$

$$\eta \overset{\text{def}}{=} \min_{1\leq i\leq N} \eta_i = \min_{\mu_1 \in \text{eig}(A_{11},B_{11}),\mu_2 \in \text{eig}(A_{22},B_{22})} |\mu_1 - \mu_2|. \tag{2.2b}$$

For the sake of this definition, we treat a multiple eigenvalue as different copies of the same value. If the multiple eigenvalue comes from both $\text{eig}(A_{11}, B_{11})$ and $\text{eig}(A_{22}, B_{22})$, each copy is considered as an eigenvalue of only one of $A_{ii} - \lambda B_{ii}$ for $i = 1, 2$ but not both.

In our analysis below, the following known results from the standard Hermitian eigenvalue problem will be repeatedly used.

Lemma 2.1. *Let A and $\tilde{A}$ be two $N \times N$ Hermitian matrices, and denote their eigenvalues in descending order by λ_i $(1 \leq i \leq N)$ and by $\tilde{\lambda}_i$ $(1 \leq i \leq N)$, respectively.*

(a) *See* [1], [13], [14]. *$|\tilde{\lambda}_j - \lambda_j| \leq \|\tilde{A} - A\|_2$ for $1 \leq j \leq N$.*

(b) *See* [12] *(see also* [4, pp. 224–225]*). If $\tilde{A} = W^*AW$, then there exist t_j $(1 \leq j \leq N)$ satisfying $[\sigma_{\min}(W)]^2 \leq t_j \leq [\sigma_{\max}(W)]^2$ such that $\tilde{\lambda}_j = t_j\lambda_j$ for $1 \leq j \leq N$.*

(c) *See* [6]. *Suppose that*

$$A = \begin{matrix} m \\ n \end{matrix} \overset{\begin{matrix} m & & n \end{matrix}}{\begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix}}, \qquad \tilde{A} = \begin{pmatrix} A_{11} & E_{12} \\ E_{21} & A_{22} \end{pmatrix},$$

i.e., in (1.1) *and* (1.2) *$B = \tilde{B} = I_N$ and $E_{ii} = 0$ for $i = 1, 2$, and define the gaps η_i and η as in* (2.2a) *and* (2.2b)*, respectively. Then for $1 \leq j \leq N$,*

$$|\tilde{\lambda}_j - \lambda_j| \leq \frac{2\|E_{21}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|E_{21}\|_2^2}} \leq \frac{2\|E_{21}\|_2^2}{\eta + \sqrt{\eta^2 + 4\|E_{21}\|_2^2}}.$$

(d) *See* [5]. *Let Q be $N \times m$ with orthonormal columns, i.e., $Q^*Q = I_m$, and denote the eigenvalues of $H = Q^*AQ$ by θ_j $(1 \leq i \leq m)$ in descending order. Then there are m eigenvalues $\mu_1 \geq \cdots \geq \mu_m$ of A such that*

$$|\mu_j - \theta_j| \leq \|AQ - QH\|_2 \quad \text{for } 1 \leq j \leq m,$$

$$\sqrt{\sum_{j=1}^{m} |\mu_j - \theta_j|^2} \leq \|AQ - QH\|_{\text{F}}.$$

The inequality in Lemma 2.1(a) has a long history and is well known; see, e.g., [13, p. 208], [14, p. 203], and [1, p. 71], where its extension to all unitarily invariant norms can also be found. Better bounds than those in Lemma 2.1(d) are possible if further information becomes available. The interested reader is referred to the short survey at the end of the article [2] (see also [15]).

2.1. Special case. We will start by considering the special case

$$E_{ii} = F_{ii} = 0, \quad B_{11} = I_m, \quad B_{22} = I_n. \tag{2.3}$$

For this case,

$$\tilde{A} = \begin{pmatrix} A_{11} & E_{21}^* \\ E_{21} & A_{22} \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} I_m & F_{21}^* \\ F_{21} & I_n \end{pmatrix}. \tag{2.4}$$

We shall bound the differences $\tilde{\lambda}_j - \lambda_j$ via three different approaches. Throughout this subsection we assume that

$$\|F_{21}\|_2 < 1$$

so that $\tilde{B}$ is Hermitian positive definite.

Method I. Noting that $I - F_{21}F_{21}^*$ is Hermitian definite, we let

$$X = \begin{pmatrix} I_m & -F_{21}^* \\ 0 & I_n \end{pmatrix}, \quad W = \begin{pmatrix} I_m & 0 \\ 0 & [I - F_{21}F_{21}^*]^{1/2} \end{pmatrix}. \tag{2.5}$$

Then

$$\hat{B} \stackrel{\text{def}}{=} X^* \tilde{B} X = \begin{pmatrix} I_m & 0 \\ 0 & 1 - F_{21}F_{21}^* \end{pmatrix} = W^2, \tag{2.6a}$$

$$\hat{A} \stackrel{\text{def}}{=} X^* \tilde{A} X = \begin{pmatrix} A_{11} & -A_{11}F_{21}^* + E_{21}^* \\ -F_{21}A_{11} + E_{21} & A_{22} - E_{21}F_{21}^* - F_{21}E_{21}^* + F_{21}A_{11}F_{21}^* \end{pmatrix}. \tag{2.6b}$$

We now consider the following four eigenvalue problems:
EIG (a): $\tilde{A} - \lambda\tilde{B}$ (which has the same eigenvalues as $W^{-1}\hat{A}W^{-1} - \lambda I_N$),
EIG (b): $\hat{A} - \lambda I_N$,
EIG (c): $\begin{pmatrix} A_{11} & -A_{11}F_{21}^* + E_{21}^* \\ -F_{21}A_{11} + E_{21} & A_{22} \end{pmatrix} - \lambda I_N$,
EIG (d): $A - \lambda I_N$.
Denote the eigenvalues for EIG(x) by $\lambda_j^{(\mathrm{x})}$ in descending order, i.e.,

$$\lambda_1^{(\mathrm{x})} \geq \lambda_2^{(\mathrm{x})} \geq \cdots \geq \lambda_N^{(\mathrm{x})}. \tag{2.7}$$

Then we have $\lambda_j^{(\mathrm{a})} = \tilde{\lambda}_j$ and $\lambda_j^{(\mathrm{d})} = \lambda_j$ for all j, recalling (2.1) and (2.3). There are existing perturbation bounds as given in Lemma 2.1 for any two adjacent eigenvalue problems in the above list.

(a)–(b) By Lemma 2.1(b), there exist t_j $(1 \leq j \leq N)$ satisfying

$$1/[\sigma_{\max}(W)]^2 = [\sigma_{\min}(W^{-1})]^2 \leq t_j \leq [\sigma_{\max}(W^{-1})]^2 = 1/[\sigma_{\min}(W)]^2$$

such that

$$\lambda_j^{(a)} = t_j \lambda_j^{(b)}, \text{ or equivalently, } \lambda_j^{(b)} = t_j^{-1}\lambda_j^{(a)} \quad \text{for } 1 \le j \le N.$$

It can be seen that $\sigma_{\max}(W) = 1$ and $\sigma_{\min}(W) = \sqrt{1 - \|F_{21}\|_2^2}$. Thus

$$0 \le 1 - t_j^{-1} \le \|F_{21}\|_2^2.$$

(b)–(c) By Lemma 2.1(a), $|\lambda_j^{(b)} - \lambda_j^{(c)}| \le \|E_{21}F_{21}^* + F_{21}E_{21}^* - F_{21}A_{11}F_{21}^*\|_2$ for $1 \le j \le N$.

(c)–(d) By Lemma 2.1(c), for $1 \le j \le N$,

$$|\lambda_j^{(c)} - \lambda_j^{(d)}| \le \frac{2\|A_{11}F_{21}^* - E_{21}^*\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|A_{11}F_{21}^* - E_{21}^*\|_2^2}} \le \frac{2\|E_{21} - F_{21}A_{11}\|_2^2}{\eta + \sqrt{\eta^2 + 4\|E_{21} - F_{21}A_{11}\|_2^2}}.$$

Combining these three bounds we get for $1 \le j \le N$,

$$\begin{aligned}|\tilde{\lambda}_j - \lambda_j| &= |\lambda_j^{(a)} - \lambda_j^{(d)}| \\ &= |\lambda_j^{(a)} - \lambda_j^{(b)} + \lambda_j^{(b)} - \lambda_j^{(c)} + \lambda_j^{(c)} - \lambda_j^{(d)}| \\ &\le |1 - t_j^{-1}||\lambda_j^{(a)}| + |\lambda_j^{(b)} - \lambda_j^{(c)}| + |\lambda_j^{(c)} - \lambda_j^{(d)}| \\ &\le \|F_{21}\|_2^2|\tilde{\lambda}_j| + \|E_{21}F_{21}^* + F_{21}E_{21}^* - F_{21}A_{11}F_{21}^*\|_2 \\ &\quad + \frac{2\|E_{21} - F_{21}A_{11}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|E_{21} - F_{21}A_{11}\|_2^2}}. \end{aligned} \tag{2.8}$$

Remark 2.1. If $\eta_j > 0$, the right-hand side of (2.8) is of $O(\max\{\|E_{21}\|_2^2, \|F_{21}\|_2^2\})$ for that j. If $\eta > 0$, it is of $O(\max\{\|E_{21}\|_2^2, \|F_{21}\|_2^2\})$ for all $1 \le j \le N$. We can establish more bounds between $\tilde{\lambda}_j$ and λ_j by using various other existing bounds available to bound the differences among $\lambda_j^{(x)}$'s, other than those listed in Lemma 2.1. Interested readers may find them in [1], [3], [7], [8], [9], [10], [13], [14].

Symmetrically permute $\tilde{A}$ and $\tilde{B}$ in (1.2) to

$$\begin{pmatrix} A_{22} & E_{21} \\ E_{21}^* & A_{11} \end{pmatrix}, \qquad \begin{pmatrix} I_n & F_{21} \\ F_{21}^* & I_m \end{pmatrix},$$

and then apply (2.8) to get

$$\begin{aligned}|\tilde{\lambda}_j - \lambda_j| &\le \|F_{21}\|_2^2|\tilde{\lambda}_j| + \|E_{21}^*F_{21} + F_{21}^*E_{21} - F_{21}^*A_{22}F_{21}\|_2 \\ &\quad + \frac{2\|E_{21} - A_{22}F_{21}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|E_{21} - A_{22}F_{21}\|_2^2}}. \end{aligned} \tag{2.9}$$

The following theorem summarizes what we have obtained so far.

THEOREM 2.2. *Assume* (2.3) *and* $\|F_{21}\|_2 < 1$ *for the Hermitian definite GEPs* $A - \lambda B$ *and* $\tilde{A} - \lambda\tilde{B}$ *with* A, B, $\tilde{A}$, *and* $\tilde{B}$ *as in* (1.1) *and* (1.2)*. Denote their eigenvalues as in* (2.1)*, and define gaps* η_i *and* η *as in* (2.2a) *and* (2.2b)*. Then both* (2.8) *and* (2.9) *hold for all* $1 \le j \le N$.

We now investigate how accurate the eigenvalues of A_{11} are as approximations to a subset of the eigenvalues of $\tilde{A} - \lambda\tilde{B}$. Recall (2.3). We have

$$R \stackrel{\text{def}}{=} \tilde{A}\begin{pmatrix} I_m \\ 0 \end{pmatrix} - \tilde{B}\begin{pmatrix} I_m \\ 0 \end{pmatrix} A_{11} = \begin{pmatrix} 0 \\ E_{21} - F_{21}A_{11} \end{pmatrix}. \tag{2.10}$$

Note that $\tilde{B} = X^{-*}W^2X^{-1}$ and $WX^{-1}\binom{I_m}{0} = XW^{-1}\binom{I_m}{0} = \binom{I_m}{0}$ by (2.5), (2.6a), and (2.6b). Thus $\operatorname{eig}(\tilde{A}, \tilde{B}) = \operatorname{eig}(W^{-1}X^*\tilde{A}XW^{-1})$, and

$$\begin{aligned} W^{-1}X^*R &\equiv [W^{-1}X^*\tilde{A}XW^{-1}]\begin{pmatrix} I_m \\ 0 \end{pmatrix} - \begin{pmatrix} I_m \\ 0 \end{pmatrix} A_{11} \\ &= \begin{pmatrix} 0 \\ [I - F_{21}F_{21}^*]^{-1/2}(E_{21} - F_{21}A_{11}) \end{pmatrix}. \end{aligned} \tag{2.11}$$

Lemma 2.1(d) is applicable to (2.11) (to give Theorem 2.3 below). So are the results in [15] since A_{11} is the Rayleigh quotient matrix of $W^{-1}X^*\tilde{A}XW^{-1}$ generated by $(I_m, 0)^*$, but details are omitted because of the straightforwardness.

THEOREM 2.3. *Assume the conditions of Theorem* 2.2. *There are m eigenvalues $\mu_1 \geq \cdots \geq \mu_m$ of $\tilde{A} - \lambda\tilde{B}$ such that*

$$|\mu_j - \theta_j| \leq \frac{\|E_{21} - F_{21}A_{11}\|_2}{\sqrt{1 - \|F_{21}\|_2^2}} \quad \text{for } 1 \leq j \leq m,$$

$$\sqrt{\sum_{j=1}^{m} |\mu_j - \theta_j|^2} \leq \frac{\|E_{21} - F_{21}A_{11}\|_F}{\sqrt{1 - \|F_{21}\|_2^2}},$$

where $\theta_1 \geq \cdots \geq \theta_m$ are the m eigenvalues of A_{11}.

Method II. This approach is very much adapted from [6] for the standard eigenvalue problem. We shall begin by seeking motivation from a 2-by-2 example.

Example 2.1. Consider the 2×2 Hermitian matrices

$$\tilde{A} = \begin{pmatrix} \alpha & \epsilon^* \\ \epsilon & \beta \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} 1 & \delta^* \\ \delta & 1 \end{pmatrix}, \tag{2.12}$$

where $\alpha > \beta$ and $1 - |\delta|^2 > 0$ (i.e., B is positive definite). The eigenvalues of $\tilde{A} - \lambda\tilde{B}$, denoted by $\lambda_\pm$, satisfy

$$(1 - |\delta|^2)\lambda^2 - (\alpha + \beta - \epsilon^*\delta - \epsilon\delta^*)\lambda + \alpha\beta - |\epsilon|^2 = 0.$$

Let

$$\begin{aligned} \Delta &= (\alpha + \beta - \epsilon^*\delta - \epsilon\delta^*)^2 - 4(1 - |\delta|^2)(\alpha\beta - |\epsilon|^2) \\ &= (\alpha - \beta)^2 + 2[(\alpha\delta - \epsilon)^*(\beta\delta - \epsilon) + (\alpha\delta - \epsilon)(\beta\delta - \epsilon)^*] + (\epsilon^*\delta - \epsilon\delta^*)^2. \end{aligned}$$

The eigenvalues of $\tilde{A} - \lambda\tilde{B}$ are

$$\lambda_{\pm} = \frac{(\alpha + \beta - \epsilon^* \delta - \epsilon \delta^*) \pm \sqrt{\Delta}}{2(1 - |\delta|^2)}.$$

It is not obvious to see $\lambda_+ \geq \alpha$ and $\lambda_- \leq \beta$ from this expression. A better way to see $\lambda_+ \geq \alpha$ and $\lambda_- \leq \beta$ is through the min-max principle (see Appendix A). Namely,

$$\lambda_+ = \max_x \frac{x^* \tilde{A} x}{x^* \tilde{B} x} \geq \frac{e_1^* \tilde{A} e_1}{e_1^* \tilde{B} e_1} = \alpha, \qquad \lambda_- = \min_x \frac{x^* \tilde{A} x}{x^* \tilde{B} x} \leq \beta,$$

where e_1 is the first column of the identity matrix. Consider

$$\tilde{A} - \lambda_+ \tilde{B} = \begin{pmatrix} \alpha - \lambda_+ & (\epsilon - \lambda_+ \delta)^* \\ \epsilon - \lambda_+ \delta & \beta - \lambda_+ \end{pmatrix}.$$

Since $\lambda_+ \geq \alpha > \beta$, we can define

$$X = \begin{pmatrix} 1 & 0 \\ -(\beta - \lambda_+)^{-1}(\epsilon - \lambda_+ \delta) & 1 \end{pmatrix}.$$

We have

$$X^*(\tilde{A} - \lambda_+ \tilde{B}) X = \begin{pmatrix} \alpha - \lambda_+ - (\epsilon - \lambda_+ \delta)^* (\beta - \lambda_+)^{-1} (\epsilon - \lambda_+ \delta) & 0 \\ 0 & \beta - \lambda_+ \end{pmatrix}$$

which must be singular. Since $\lambda_+ \geq \alpha > \beta$, we have

$$\begin{aligned} &\alpha - \lambda_+ - (\epsilon - \lambda_+ \delta)^* (\beta - \lambda_+)^{-1} (\epsilon - \lambda_+ \delta) = 0, \\ &\alpha - \lambda_+ = \frac{|\epsilon - \lambda_+ \delta|^2}{(\beta - \alpha) + (\alpha - \lambda_+)}, \\ &(\lambda_+ - \alpha)^2 + (\alpha - \beta)(\lambda_+ - \alpha) - |\epsilon - \lambda_+ \delta|^2 = 0. \end{aligned}$$

The last equation gives, upon noticing that $\lambda_+ - \alpha \geq 0$, that

$$\lambda_+ - \alpha = \frac{2|\epsilon - \lambda_+ \delta|^2}{(\alpha - \beta) + \sqrt{(\alpha - \beta)^2 + 4|\epsilon - \lambda_+ \delta|^2}}. \tag{2.13}$$

Apply (2.13) to $(-\tilde{A}) - \lambda \tilde{B}$ to get

$$\beta - \lambda_- = \frac{2|\epsilon - \lambda_- \delta|^2}{(\alpha - \beta) + \sqrt{(\alpha - \beta)^2 + 4|\epsilon - \lambda_- \delta|^2}}. \tag{2.14}$$

Surprisingly an inequality similar to (2.13) and (2.14) holds for the general case as stated in Theorem 2.4 below. ∎

THEOREM 2.4. *Under the conditions of Theorem 2.2, we have, for all $1 \leq i \leq N$,*

$$|\tilde{\lambda}_i - \lambda_i| \leq \frac{2\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2}{\eta_i + \sqrt{\eta_i^2 + 4\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2}} \leq \frac{2\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2}{\eta + \sqrt{\eta^2 + 4\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2}}. \tag{2.15}$$

Proof. We shall prove (2.15) by induction. We may assume that A_{11} and A_{22} are diagonal with their diagonal entries arranged in descending order, respectively; otherwise replace $\tilde{A}$ and $\tilde{B}$ by

$$(U \bigoplus V)^* \tilde{A} (U \bigoplus V) \quad \text{and} \quad (U \bigoplus V)^* \tilde{B} (U \bigoplus V),$$

respectively, where U and V are unitary such that $U^* A_{11} U$ and $V^* A_{22} V$ are in such diagonal forms.

We may perturb the diagonal of A so that all entries are distinct and apply the continuity argument for the general case.

If $N = 2$, the result is true by Example 2.1. Assume that $N > 2$, and assume that the result is true for Hermitian matrices of size $N - 1$.

First, we show that (2.15) holds for $i = 1$. Assume that the $(1,1)$th entry of A_{11} equals λ_1. By the min-max principle, we have

$$\tilde{\lambda}_1 \geq \frac{e_1^* \tilde{A} e_1}{e_1^* \tilde{B} e_1} = \lambda_1.$$

No proof is necessary if $\tilde{\lambda}_1 = \lambda_1$. Assume $\tilde{\lambda}_1 > \lambda_1$, and let

$$X = \begin{pmatrix} I_m & 0 \\ -(A_{22} - \tilde{\lambda}_1 I_n)^{-1}(E_{21} - \tilde{\lambda}_1 F_{21}) & I_n \end{pmatrix}.$$

Then

$$X^* (\tilde{A} - \tilde{\lambda}_1 \tilde{B}) X = \begin{pmatrix} M(\tilde{\lambda}_1) & 0 \\ 0 & A_{22} - \tilde{\lambda}_1 I_n \end{pmatrix},$$

where

$$M(\lambda) = A_{11} - \lambda I_m - (E_{21} - \lambda F_{21})^* (A_{22} - \lambda I_n)^{-1} (E_{21} - \lambda F_{21}).$$

Since $\tilde{A} - \tilde{\lambda}_1 \tilde{B}$ and $X^*(\tilde{A} - \tilde{\lambda}_1 \tilde{B})X$ have the same inertia, we see that $M(\tilde{\lambda}_1)$ has zero as the largest eigenvalue. Notice that the largest eigenvalue of $A_{11} - \tilde{\lambda}_1 I$ is $\lambda_1 - \tilde{\lambda}_1 \leq 0$. Thus, for $\delta_1 = |\tilde{\lambda}_1 - \lambda_1| = \tilde{\lambda}_1 - \lambda_1$, we have by Lemma 2.1(a)

$$\begin{aligned} \delta_1 = |0 - (\lambda_1 - \tilde{\lambda}_1)| &\leq \|(E_{21} - \tilde{\lambda}_1 F_{21})^* (A_{22} - \tilde{\lambda}_1 I_n)^{-1} (E_{21} - \tilde{\lambda}_1 F_{21})\|_2 \\ &\leq \|E_{21} - \tilde{\lambda}_1 F_{21}\|_2^2 \|(A_{22} - \tilde{\lambda}_1 I_n)^{-1}\|_2 \\ &\leq \frac{\|E_{21} - \tilde{\lambda}_1 F_{21}\|_2^2}{\delta_1 + \eta_1}, \end{aligned}$$

where we have used $[\|(A_{22} - \tilde{\lambda}_1 I_n)^{-1}\|_2]^{-1} = \tilde{\lambda}_1 - \max_{\mu \in \mathrm{eig}(A_{22})} \mu = \delta_1 + \eta_1$. Consequently,

$$\delta_1 \leq \frac{2\|E_{21} - \tilde{\lambda}_1 F_{21}\|_2^2}{\eta_1 + \sqrt{\eta_1^2 + 4\|E_{21} - \tilde{\lambda}_1 F_{21}\|_2^2}},$$

as asserted. Similarly, we can prove the result if the $(1,1)$th entry of A_{22} equals λ_1. In this case, we will apply the inertia arguments to $\tilde{A} - \tilde{\lambda}_1 \tilde{B}$ and $Y(\tilde{A} - \tilde{\lambda}_1 \tilde{B})Y^*$ with

$$Y = \begin{pmatrix} I_m & 0 \\ -(E_{21} - \tilde{\lambda}_1 F_{21})(A_{11} - \tilde{\lambda}_1 I_m)^{-1} & I_n \end{pmatrix}.$$

Applying the result of the last paragraph to $(-\tilde{A}) - \lambda\tilde{B}$, we see that (2.15) holds for $i = N$.

Now, suppose $1 < i < N$. The result trivially holds if $\tilde{\lambda}_i = \lambda_i$. Suppose $\tilde{\lambda}_i \neq \lambda_i$. We may assume that $\lambda_i > \tilde{\lambda}_i$. Otherwise, replace $\{(A, B), (\tilde{A}, \tilde{B}), i\}$ by $\{(-A, B), (-\tilde{A}, \tilde{B}), N - i + 1\}$. Delete the row and column of $\tilde{A}$ that contain the diagonal entry λ_N, and delete the corresponding row and column of $\tilde{B}$ as well. Let $\hat{A}$ and $\hat{B}$ be the resulting matrices. Write the eigenvalues of $\hat{A} - \lambda\hat{B}$ as $\nu_1 \geq \cdots \geq \nu_{N-1}$. By the interlacing inequalities (A.3), we have

$$\tilde{\lambda}_i \geq \nu_i \quad \text{and hence } \lambda_i - \tilde{\lambda}_i \leq \lambda_i - \nu_i. \tag{2.16}$$

Note that λ_i is the ith largest diagonal entry of $\hat{A}$. Let $\hat{\eta}_i$ be the minimum distance between λ_i and the diagonal entries in the diagonal block $\hat{A}_{jj}$ in $\hat{A}$ not containing λ_i, where $j \in \{1, 2\}$. Then

$$\hat{\eta}_i \geq \eta_i$$

because $\hat{A}_{jj}$ may have one fewer diagonal entries than A_{jj}. Let $\hat{E}_{21}$ and $\hat{F}_{21}$ be the off-diagonal block of $\hat{A}$ and $\hat{B}$, respectively. Then

$$\|\hat{E}_{21} - \tilde{\lambda}_i \hat{F}_{21}\|_2 \leq \|E_{21} - \tilde{\lambda}_i F_{21}\|_2. \tag{2.17}$$

Finally, we have

$$\begin{aligned}
|\tilde{\lambda}_i - \lambda_i| &= \lambda_i - \tilde{\lambda}_i && \text{because } \lambda_i > \tilde{\lambda}_i \\
& && \leq \lambda_i - \nu_i \qquad \text{by (2.16)} \\
&\leq \frac{2\|\hat{E}_{21} - \tilde{\lambda}_i \hat{F}_{21}\|_2^2}{\hat{\eta}_i + \sqrt{\hat{\eta}_i^2 + 4\|\hat{E}_{21} - \tilde{\lambda}_i \hat{F}_{21}\|_2^2}} && \text{by induction assumption} \\
&\leq \frac{2\|\hat{E}_{21} - \tilde{\lambda}_i \hat{F}_{21}\|_2^2}{\eta_i + \sqrt{\eta_i^2 + 4\|\hat{E}_{21} - \tilde{\lambda}_i \hat{F}_{21}\|_2^2}} && \text{because } \hat{\eta}_i > \eta_i \\
&= \frac{1}{2}\left(\sqrt{\eta_i^2 + 4\|\hat{E}_{21} - \tilde{\lambda}_i \hat{F}_{21}\|_2^2} - \eta_i\right) \\
&\leq \frac{1}{2}\left(\sqrt{\eta_i^2 + 4\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2} - \eta_i\right) && \text{by (2.17)} \\
&= \frac{2\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2}{\eta_i + \sqrt{\eta_i^2 + 4\|E_{21} - \tilde{\lambda}_i F_{21}\|_2^2}},
\end{aligned}$$

as expected. ∎

Method III. We now consider the following three eigenvalue problems:
EIG (a): $\tilde{A} - \lambda\tilde{B}$ (which has the same eigenvalues as $\tilde{B}^{-1/2}\tilde{A}\tilde{B}^{-1/2} - \lambda I_N$),
EIG (b): $\tilde{A} - \lambda I_N$,
EIG (c): $A - \lambda I_N$.
Denote the eigenvalues for EIG(x) by $\lambda_j^{(\mathrm{x})}$ in descending order as in (2.7). Then we have $\lambda_j^{(\mathrm{a})} = \tilde{\lambda}_j$ and $\lambda_j^{(\mathrm{c})} = \lambda_j$ for all j, recalling (2.1). Existing perturbation bounds as given in Lemma 2.1 can be used for any two adjacent eigenvalue problems in the above list.

(a)–(b) By Lemma 2.1(b), there exist t_j $(1 \le j \le N)$ satisfying

$$1/\sigma_{\max}(\tilde{B}) = \sigma_{\min}(\tilde{B}^{-1}) \le t_j \le \sigma_{\max}(\tilde{B}^{-1}) = 1/\sigma_{\min}(\tilde{B})$$

such that

$$\lambda_j^{(\mathrm{a})} = t_j \lambda_j^{(\mathrm{b})} \quad (\text{or equivalently } \lambda_j^{(\mathrm{b})} = t_j^{-1}\lambda_j^{(\mathrm{a})}) \quad \text{for } 1 \le j \le N.$$

It can be seen that $1 - \sigma_{\max}(F_{21}) = \sigma_{\min}(\tilde{B}) \le \sigma_{\max}(\tilde{B}) = 1 + \sigma_{\max}(F_{21})$. Thus $|1 - t_j^{-1}| \le \|F_{21}\|_2$, which will be used later.

(b)–(c) By Lemma 2.1(c), for $1 \le j \le N$,

$$|\lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}| \le \frac{2\|E_{21}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|E_{21}\|_2^2}} \le \frac{2\|E_{21}\|_2^2}{\eta + \sqrt{\eta^2 + 4\|E_{21}\|_2^2}}.$$

Finally, for $1 \le j \le N$,

$$\begin{aligned} |\tilde{\lambda}_j - \lambda_j| = |\lambda_j^{(\mathrm{a})} - \lambda_j^{(\mathrm{c})}| &= |\lambda_j^{(\mathrm{a})} - \lambda_j^{(\mathrm{b})} + \lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}| \\ &\le |1 - t_j^{-1}||\lambda_j^{(\mathrm{a})}| + |\lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}| \\ &\le \|F_{21}\|_2|\tilde{\lambda}_j| + \frac{2\|E_{21}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|E_{21}\|_2^2}}. \end{aligned} \tag{2.18}$$

Remark 2.2. The derivation here is the shortest (and simplest) among the three methods that lead to (2.8) and (2.9), (2.15), and (2.18), but not without a sacrifice; namely, it is likely the weakest when $\|F_{21}\|_2$ has a much bigger magnitude than $\|E_{21}\|_2^2$ because $\|F_{21}\|_2$ appears linearly in (2.18) vs. quadratically in (2.8), (2.9), and (2.15). Note the similarity among the third term in the right-hand side of (2.8), the second and last terms in (2.18), and the second term in the right-hand side of (2.18).

THEOREM 2.5. *Under the conditions of Theorem* 2.2, *we have* (2.18) *for all* $1 \le j \le N$.

We point out in passing that Theorems 2.2, 2.4, and 2.5 all reduce to the main result in [6], i.e., Lemma 2.1(c), for the standard eigenvalue problem if $F_{21} = 0$.

2.2. General case. We are now looking into the general case, i.e., without assuming (2.3).

LEMMA 2.6. *Let* Δ *be a Hermitian matrix. If* $\delta \stackrel{\mathrm{def}}{=} \|\Delta\|_2 < 1$, *then* $I + \Delta$ *is positive definite, and*

$$\|(I + \Delta)^{-1/2} - I\|_2 \le \frac{1}{\sqrt{1-\delta}} - 1 = \frac{\delta}{\sqrt{1-\delta}(1 + \sqrt{1-\delta})}.$$

Proof. Any eigenvalue of $I+\Delta$ is no smaller than $1-\delta>0$, so $I+\Delta$ is positive definite. We have

$$\begin{aligned}\|(I+\Delta)^{-1/2}-I\|_2 &= \max_{\mu\in\mathrm{eig}(\Delta)}|(1+\mu)^{-1/2}-1|\\ &\le \max\{(1-\delta)^{-1/2}-1, 1-(1+\delta)^{-1/2}\}\\ &= (1-\delta)^{-1/2}-1,\end{aligned}$$

as was to be shown. ∎

Recall that A, $\tilde{A}$, B, $\tilde{B}$, E, and F are all Hermitian. Set

$$\Delta_{ij}=B_{ii}^{-1/2}F_{ij}B_{jj}^{-1/2}, \tag{2.19a}$$

$$Y=\mathrm{diag}([I+\Delta_{11}]^{-1/2}B_{11}^{-1/2},[I+\Delta_{22}]^{-1/2}B_{22}^{-1/2}), \tag{2.19b}$$

$$\hat{F}_{ij}=[I+\Delta_{ii}]^{-1/2}\Delta_{ij}[I+\Delta_{jj}]^{-1/2}\quad\text{for } i\neq j, \tag{2.19c}$$

$$\hat{A}_{ii}=B_{ii}^{-1/2}A_{ii}B_{ii}^{-1/2}, \tag{2.19d}$$

$$\hat{E}_{ij}=[I+\Delta_{ii}]^{-1/2}B_{ii}^{-1/2}E_{ij}B_{jj}^{-1/2}[I+\Delta_{jj}]^{-1/2}\quad\text{for } i\neq j, \tag{2.19e}$$

$$\begin{aligned}\hat{E}_{ii}&=[I+\Delta_{ii}]^{-1/2}B_{ii}^{-1/2}(A_{ii}+E_{ii})B_{ii}^{-1/2}[I+\Delta_{ii}]^{-1/2}-\hat{A}_{ii}\\ &=([I+\Delta_{ii}]^{-1/2}-I)\hat{A}_{ii}([I+\Delta_{ii}]^{-1/2}-I)\\ &\quad+\hat{A}_{ii}([I+\Delta_{ii}]^{-1/2}-I)+([I+\Delta_{ii}]^{-1/2}-I)\hat{A}_{ii} &(2.19\text{f})\\ &\quad+[I+\Delta_{ii}]^{-1/2}B_{ii}^{-1/2}E_{ii}B_{ii}^{-1/2}[I+\Delta_{ii}]^{-1/2}, &(2.19\text{g})\end{aligned}$$

and set

$$\delta_{ij}=\|\Delta_{ij}\|_2\le\sqrt{\|B_{ii}^{-1}\|_2\|B_{jj}^{-1}\|_2}\|F_{ij}\|_2,\quad \gamma_{ij}=(1-\delta_{ij})^{-1/2}-1. \tag{2.20}$$

In obtaining (2.20), we have used $\|B_{ii}^{-1/2}\|_2=\sqrt{\|B_{ii}^{-1}\|_2}$. To ensure that $\tilde{B}$ is positive definite, throughout this subsection we assume that

$$\max\{\delta_{11},\delta_{22}\}<1,\qquad \delta_{12}^2<(1-\delta_{11})(1-\delta_{22}). \tag{2.21}$$

We can bound $\hat{E}_{ij}$ and $\hat{F}_{ij}$ as follows:

$$\|\hat{E}_{ij}\|_2\le\frac{\|B_{ii}^{-1/2}E_{ij}B_{jj}^{-1/2}\|_2}{\sqrt{(1-\delta_{ii})(1-\delta_{jj})}} \tag{2.22a}$$

$$\le\sqrt{\frac{\|B_{ii}^{-1}\|_2\|B_{jj}^{-1}\|_2}{(1-\delta_{ii})(1-\delta_{jj})}}\|E_{ij}\|_2\quad\text{for } i\neq j, \tag{2.22b}$$

$$\|\hat{E}_{ii}\|_2 \le \gamma_{ii}(2+\gamma_{ii})\|\hat{A}_{ii}\|_2 + \frac{\|B_{ii}^{-1/2}E_{ii}B_{ii}^{-1/2}\|_2}{1-\delta_{ii}} \tag{2.22c}$$

$$\le \gamma_{ii}(2+\gamma_{ii})\|\hat{A}_{ii}\|_2 + \frac{\|B_{ii}^{-1}\|_2}{1-\delta_{ii}}\|E_{ii}\|_2, \tag{2.22d}$$

$$\|\hat{F}_{ij}\|_2 \le \frac{\delta_{ij}}{\sqrt{(1-\delta_{ii})(1-\delta_{jj})}} \quad \text{for } i \ne j. \tag{2.22e}$$

We have used Lemma 2.6 to get (2.22c) from (2.19g). We then have

$$\hat{A} \overset{\text{def}}{=} Y^*\tilde{A}Y = \begin{pmatrix} \hat{A}_{11}+\hat{E}_{11} & \hat{E}_{12} \\ \hat{E}_{21} & \hat{A}_{22}+\hat{E}_{22} \end{pmatrix}, \qquad \hat{B} \overset{\text{def}}{=} Y^*\tilde{B}Y = \begin{pmatrix} I_m & \hat{F}_{12} \\ \hat{F}_{21} & I_n \end{pmatrix}.$$

We now consider the following three eigenvalue problems:
EIG (a): $\tilde{A}-\lambda\tilde{B}$ (which is equivalent to $\hat{A}-\lambda\hat{B}$),
EIG (b): $\begin{pmatrix} \hat{A}_{11} & \hat{E}_{12} \\ \hat{E}_{21} & \hat{A}_{22} \end{pmatrix} - \lambda\begin{pmatrix} I_m & \hat{F}_{12} \\ \hat{F}_{21} & I_n \end{pmatrix}$,
EIG (c): $\begin{pmatrix} \hat{A}_{11} & \\ & \hat{A}_{22} \end{pmatrix} - \lambda I_N$.

Denote the eigenvalues for EIG(x) by $\lambda_j^{(\mathrm{x})}$ in descending order as in (2.7). Then we have $\lambda_j^{(\mathrm{a})} = \tilde{\lambda}_j$ and $\lambda_j^{(\mathrm{c})} = \lambda_j$ for all j, recalling (2.1). Note $\|\hat{F}_{21}\|_2 < 1$ because of (2.21). The eigenvalue differences between any two adjacent eigenvalue problems in the above list can be bounded as follows.

(a)–(b) By Lemma 2.1(a), $|\lambda_j^{(\mathrm{a})} - \lambda_j^{(\mathrm{b})}| \le \frac{\max_i\|\hat{E}_{ii}\|_2}{1-\|\hat{F}_{21}\|_2}$ for $1 \le j \le N$. This is because

$$\hat{B}^{-1/2}\hat{A}\hat{B}^{-1/2} = \hat{B}^{-1/2}\begin{pmatrix} \hat{A}_{11} & \hat{E}_{12} \\ \hat{E}_{21} & \hat{A}_{22} \end{pmatrix}\hat{B}^{-1/2} + \hat{B}^{-1/2}\begin{pmatrix} \hat{E}_{11} & \\ & \hat{E}_{22} \end{pmatrix}\hat{B}^{-1/2},$$

and $\|\hat{B}^{-1}\|_2 = [1-\|\hat{F}_{21}\|_2]^{-1}$.

(b)–(c) Use Theorems 2.2, 2.4, or 2.5 to bound $\lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}$ to yield three bounds:

$$\begin{aligned} |\lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}| \le{}& \|\hat{F}_{21}\|_2^2|\tilde{\lambda}_j| + \|\hat{E}_{21}\hat{F}_{21}^* + \hat{F}_{21}\hat{E}_{21}^* - \hat{F}_{21}\hat{A}_{11}\hat{F}_{21}^*\|_2 \\ &+ \frac{2\|\hat{E}_{21}-\hat{F}_{21}\hat{A}_{11}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|\hat{E}_{21}-\hat{F}_{21}\hat{A}_{11}\|_2^2}}, \end{aligned} \tag{2.23}$$

$$|\lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}| \le \frac{2\|\hat{E}_{21}-\tilde{\lambda}_j\hat{F}_{21}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|\hat{E}_{21}-\tilde{\lambda}_j\hat{F}_{21}\|_2^2}}, \tag{2.24}$$

$$|\lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}| \le \|\hat{F}_{21}\|_2|\tilde{\lambda}_j| + \frac{2\|\hat{E}_{21}\|_2^2}{\eta_j + \sqrt{\eta_j^2 + 4\|\hat{E}_{21}\|_2^2}}. \tag{2.25}$$

Further bounds in terms of the norms of E, F, and B_{ii} can be obtained by using inequalities in (2.22a)–(2.22e). Finally we use $\tilde\lambda_j - \lambda_j = \lambda_j^{(\mathrm{a})} - \lambda_j^{(\mathrm{c})} = \lambda_j^{(\mathrm{a})} - \lambda_j^{(\mathrm{b})} + \lambda_j^{(\mathrm{b})} - \lambda_j^{(\mathrm{c})}$ to conclude what follows.

THEOREM 2.7. *For the Hermitian definite GEPs $A - \lambda B$ and $\tilde A - \lambda \tilde B$ with A, B, $\tilde A$, and $\tilde B$ as in* (1.1) *and* (1.2)*, assume* (2.21)*, where δ_{ij} are as defined in* (2.19a)–(2.19g) *and* (2.20)*. Denote their eigenvalues as in* (2.1)*, and define gaps η_i and η as in* (2.2a) *and* (2.2b)*. Then for all $1 \le j \le N$,*

$$|\tilde\lambda_j - \lambda_j| \le \frac{\max_i \|\hat E_{ii}\|_2}{1 - \|\hat F_{21}\|_2} + \epsilon_j,$$

where ϵ_j can be taken to be any one of the right-hand sides of (2.23)*,* (2.24)*, and* (2.25)*.*

Next we estimate the differences between the eigenvalues of $A_{11} - \lambda B_{11}$ (which is the same as those of $\hat A_{11}$) and some m eigenvalues of $\tilde A - \lambda \tilde B$. This will be done in two steps:

(a) bound the differences between the eigenvalues of $\hat A_{11} + \hat E_{11}$ and m of those of $\tilde A - \lambda \tilde B$;

(b) bound the differences between the eigenvalues of $\hat A_{11} + \hat E_{11}$ and those of $\hat A_{11}$.

The first step can be accomplished by using Theorem 2.3, while the second step can be done by using Lemma 2.1(a). We thereby get the following theorem.

THEOREM 2.8. *Assume the conditions of Theorem* 2.7*. There are m eigenvalues $\mu_1 \ge \cdots \ge \mu_m$ of $\tilde A - \lambda \tilde B$ such that*

$$|\mu_j - \theta_j| \le \|\hat E_{11}\|_2 + \frac{\|\hat E_{21} - \hat F_{21}(\hat A_{11} + \hat E_{11})\|_2}{\sqrt{1 - \|\hat F_{21}\|_2^2}} \quad for\ 1 \le j \le m,$$

$$\sqrt{\sum_{j=1}^{m} |\mu_j - \theta_j|^2} \le \|\hat E_{11}\|_{\mathrm{F}} + \frac{\|\hat E_{21} - \hat F_{21}(\hat A_{11} + \hat E_{11})\|\mathrm{F}}{\sqrt{1 - \|\hat F_{21}\|_2^2}},$$

where $\theta_1 \ge \cdots \ge \theta_m$ are the m eigenvalues of $A_{11} - \lambda B_{11}$.

Remark 2.3. Some comments for comparing Theorems 2.7 and 2.8 are in order.

(a) Theorem 2.7 bounds the changes in all the eigenvalues of $A - \lambda B$, while Theorem 2.8 bounds only a portion of them, namely, those of $A_{11} - \lambda B_{11}$.

(b) $\hat E_{22}$ does not appear in the bounds in Theorem 2.8, while both $\hat E_{ii}$ show up in those in Theorem 2.7. This could potentially make the bounds in Theorem 2.8 more favorable if $\hat E_{22}$ has much larger magnitude than $\hat E_{11}$. This point will be well manifested in our analysis in section 3. Also, note that the quantities $\|\hat E_{ij}\|$ are relative quantities as opposed to absolute quantities $\|E_{ij}\|$, because $\hat E_{ij}$ has been multiplied by $B_{ii}^{-1/2}$ and $B_{jj}^{-1/2}$.

(c) Except when ϵ_j is taken to be the right-hand side of (2.25), bounds in Theorem 2.7 are of quadratic order in $\hat E_{21}$ and $\hat F_{21}$, while those in Theorem 2.8 are of linear order in $\hat E_{21}$ and $\hat F_{21}$. All bounds are of linear order in $\hat E_{ii}$ and $\hat F_{ii}$. Thus when $\hat E_{ii} = F_{ii} = 0$ for $i = 1$, 2, Theorem 2.7 may lead to sharper bounds.

(d) Theorem 2.7 requires some gap information among the eigenvalues of $A_{ii} - \lambda B_{ii}$ for $i = 1$, 2, while Theorem 2.8 does not.

Example 2.2. To illustrate these comments in Remark 2.3, we consider the parameterized GEP

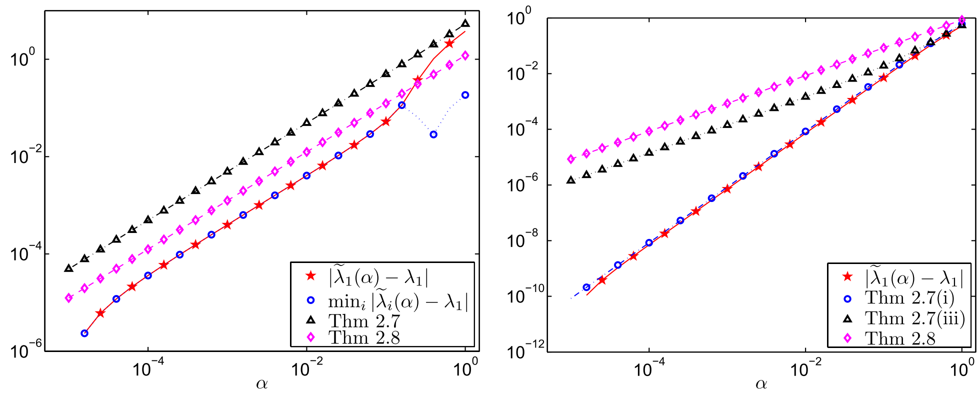

FIG. 2.1. $|\tilde\lambda_1(\alpha)-\lambda_1|$ *and its bounds by Theorems* 2.7 *and* 2.8. *Left: Under perturbation* (2.26), *the curves for the three bounds by Theorem* 2.7 *are indistinguishable, and the bound by Theorem* 2.8 *is smaller. It is interesting to notice that the curve for* $|\tilde\lambda_1(\alpha)-\lambda_1|$ *crosses above the bound by Theorem* 2.8 *for* α *about* 0.25 *or larger. This is because the bound by Theorem* 2.8 *is actually for* $\min_i|\tilde\lambda_i(\alpha)-\lambda_1|$. *Right: Under perturbation* (2.27), *the curve for Thm* 2.7(ii) *is the same as for* $|\tilde\lambda_1(\alpha)-\lambda_1|$, *and the bound by Theorem* 2.8 *is larger.*

$$\tilde A(\alpha)-\lambda\tilde B(\alpha)\equiv(A+\alpha E)-\lambda(B+\alpha F),$$

where α is a parameter ranging from 0 to 1, $A=\operatorname{diag}(4,1)$, and $B=\operatorname{diag}(2,1)$. Two types of perturbations E and F will be considered: the general dense perturbations

$$E=\begin{pmatrix}1&1\\1&5\end{pmatrix},\qquad F=\frac{1}{10}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}, \tag{2.26}$$

and the off-diagonal perturbations

$$E=\begin{pmatrix}0&1\\1&0\end{pmatrix},\qquad F=\frac{1}{10}\begin{pmatrix}0&-1\\-1&0\end{pmatrix}. \tag{2.27}$$

Denote by $\tilde\lambda_j(\alpha)$ the jth largest eigenvalue of $\tilde A(\alpha)-\lambda\tilde B(\alpha)$. Here we take $j=1$, so $\lambda_1(0)=\lambda_1=2$. Figure 2.1 shows log-log scale plots for the actual $|\tilde\lambda_1(\alpha)-\lambda_1|$, its bound by Theorem 2.8, and the three bounds by Theorem 2.7 corresponding to ϵ_j being the right-hand sides of (2.23), (2.24), and (2.25), respectively. These three bounds are shown as "Thm 2.7(i)," "Thm 2.7(ii)," and "Thm 2.7(iii)" in the plots. We assumed the gap $\eta_1=1$ is known.

In the left plot of Figure 2.1 we plot only one curve for the three bounds by Theorem 2.7 because they are visually indistinguishable. The figure illustrates the first two comments. First, the bound by Theorem 2.7 becomes smaller than $|\tilde\lambda_1(\alpha)-\lambda_1|$ for $\alpha\gtrapprox 0.25$. This is not an error but is because the bound by Theorem 2.8 is for the distance between $\lambda_1=2$ and *an* eigenvalue of $\tilde A(\alpha)-\lambda\tilde B(\alpha)$, which may not necessarily be $\tilde\lambda_1(\alpha)$. In fact, for $\alpha\gtrapprox 0.25$ the eigenvalue of $\tilde A(\alpha)-\lambda\tilde B(\alpha)$ closer to 2 is $\tilde\lambda_2(\alpha)$. Second, Theorem 2.8 gives a smaller bound than Theorem 2.7, reflecting the fact that E_{22} is much larger than E_{11}.

The right plot of Figure 2.1 illustrates the third comment. Specifically, the first two bounds by Theorem 2.7 are much smaller than the other bounds. They reflect the quadratic scaling of $|\tilde\lambda_1(\alpha)-\lambda_1|$ under off-diagonal perturbations, as can be seen by the slope of the plots. ∎

We now specialize the results so far in this subsection to the case

$$E_{ii} = F_{ii} = 0 \quad \text{for } i = 1, 2. \tag{2.28}$$

This corresponds to a practical situation in eigenvalue computations: What is the effect of dropping off off-diagonal blocks with relatively small magnitudes? Assume (2.28); then

$$\begin{aligned} &\Delta_{ii} = 0, \qquad \hat{E}_{ii} = 0, \qquad \gamma_{ii} = \delta_{ii} = 0 \quad \text{for } i = 1, 2, \\ &\hat{E}_{21} = B_{22}^{-1/2} E_{21} B_{11}^{-1/2}, \qquad \hat{F}_{21} = B_{22}^{-1/2} F_{21} B_{11}^{-1/2}. \end{aligned} \tag{2.29}$$

Theorem 2.7 yields the following corollary.

COROLLARY 2.9. *To the conditions of Theorem* 2.7 *add* $E_{ii} = F_{ii} = 0$ *for* $i = 1, 2$. *Let* $\hat{E}_{21}$ *and* $\hat{F}_{21}$ *be given as in* (2.29), *and assume* $\|\hat{F}_{21}\|_2 < 1$. *Then for all* $1 \le j \le N$,

$$|\tilde{\lambda}_j - \lambda_j| \le \epsilon_j, \tag{2.30}$$

where ϵ_j *can be taken to be any one of the right-hand sides of* (2.23), (2.24), *and* (2.25).

At the same time Theorem 2.8 gives the following corollary.

COROLLARY 2.10. *Assume the conditions of Corollary* 2.9. *There are* m *eigenvalues* $\mu_1 \ge \cdots \ge \mu_m$ *of* $\tilde{A} - \lambda \tilde{B}$ *such that*

$$|\mu_j - \theta_j| \le \frac{\|\hat{E}_{21} - \hat{F}_{21}\hat{A}_{11}\|_2}{\sqrt{1 - \|\hat{F}_{21}\|_2^2}} \quad \textit{for } 1 \le j \le m,$$

$$\sqrt{\sum_{j=1}^{m} |\mu_j - \theta_j|^2} \le \frac{\|\hat{E}_{21} - \hat{F}_{21}\hat{A}_{11}\|_{\mathrm{F}}}{\sqrt{1 - \|\hat{F}_{21}\|_2^2}},$$

where $\theta_1 \ge \cdots \ge \theta_m$ *are the* m *eigenvalues of* $A_{11} - \lambda B_{11}$.

3. Application to perturbations of a multiple eigenvalue.

A distinguished feature of GEP eigenvalue perturbations is that different copies of a multiple eigenvalue may exhibit quite different sensitivities towards perturbations. Stewart and Sun [14, p. 300] used the example

$$A = \begin{pmatrix} 2 & \\ & 2000 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & \\ & 1000 \end{pmatrix} \tag{3.1}$$

to illustrate that the copy 2000/1000 is much less sensitive than the other copy 2/1. Stewart and Sun then wrote "how to make this observation precise is an open research question." This question was recently solved by Nakatsukasa [11], who established bounds to reflect the different sensitivities of different copies of the multiple eigenvalue. In this section, through applying Theorem 2.8, we give different bounds for the same purpose.

Suppose a Hermitian definite GEP $A - \lambda B$ has a multiple eigenvalue λ_0 of multiplicity m. Then there is an $(m+n)$-by-$(m+n)$ matrix $X = (X_1, X_2)$ with $X_1^* X_1 = I_m$ and $X_2^* X_2 = I_n$ such that

$$X^*AX = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{pmatrix} \lambda_0 B_{11} & \\ & A_{22} \end{pmatrix} \end{array}, \qquad X^*BX = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{pmatrix} B_{11} & \\ & B_{22} \end{pmatrix} \end{array}. \tag{3.2}$$

This can be seen by letting X_1 and X_2 be the orthogonal factors in the QR decompositions of $\hat{X}_1$ and $\hat{X}_2$, respectively, where $(\hat{X}_1 \hat{X}_2)$ is the nonsingular matrix that diagonalizes $A - \lambda B$ [13, p. 344].

We may assume that B_{11} is diagonal:

$$B_{11} \equiv \mathbf{\Omega} = \mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_m), \qquad \omega_1 \geq \omega_2 \geq \cdots \geq \omega_m > 0. \tag{3.3}$$

Otherwise, let the eigendecomposition of B_{11} (which is Hermitian and positive definite) be $B_{11} = U\mathbf{\Omega} U^*$, where $\mathbf{\Omega}$ is diagonal, and then perform substitutions $B_{11} \leftarrow \mathbf{\Omega}$ and $X_1 \leftarrow X_1 U$.

Suppose $A - \lambda B$ is perturbed to $\tilde{A} - \lambda \tilde{B} \equiv (A + E) - \lambda(B + F)$, where E and F are Hermitian. Write

$$X^* \tilde{A} X = \begin{pmatrix} \lambda_0 \mathbf{\Omega} + E_{11} & E_{12} \\ E_{21} & A_{22} + E_{22} \end{pmatrix}, \qquad X^* \tilde{B} X = \begin{pmatrix} \mathbf{\Omega} + F_{11} & F_{12} \\ F_{21} & B_{22} + F_{22} \end{pmatrix}. \tag{3.4}$$

For any given k ($1 \leq k \leq m$), we repartition X^*AX, X^*BX, X^*EX, and X^*FX with k-by-k $(1,1)$ blocks as follows:

$$\begin{pmatrix} A_{11}^{(k)} & 0 \\ 0 & A_{22}^{(k)} \end{pmatrix}, \quad \begin{pmatrix} B_{11}^{(k)} & 0 \\ 0 & B_{22}^{(k)} \end{pmatrix}, \quad \begin{pmatrix} E_{11}^{(k)} & E_{12}^{(k)} \\ E_{21}^{(k)} & E_{22}^{(k)} \end{pmatrix}, \quad \begin{pmatrix} F_{11}^{(k)} & F_{12}^{(k)} \\ F_{21}^{(k)} & F_{22}^{(k)} \end{pmatrix}. \tag{3.5}$$

It can be seen that

$$\begin{aligned} A_{11}^{(k)} &= \lambda_0 \mathbf{\Omega}_{(1:k,1:k)}, \quad & A_{22}^{(k)} &= \mathrm{diag}(\lambda_0 \mathbf{\Omega}_{(k+1:m,k+1:m)}, A_{22}), \\ B_{11}^{(k)} &= \mathbf{\Omega}_{(1:k,1:k)}, \quad & B_{22}^{(k)} &= \mathrm{diag}(\mathbf{\Omega}_{(k+1:m,k+1:m)}, B_{22}), \end{aligned}$$

where $\mathbf{\Omega}_{(i:j,i:j)}$ is the MATLAB-like notation for $\mathbf{\Omega}$'s submatrix consisting of the intersections of rows i through j and columns i through j. Similarly to those in (2.19a)–(2.19g) define

$$\Delta_{ij}^{(k)} = [B_{ii}^{(k)}]^{-1/2} F_{ij}^{(k)} [B_{jj}^{(k)}]^{-1/2}, \tag{3.6a}$$

$$Y^{(k)} = \mathrm{diag}([I + \Delta_{11}^{(k)}]^{-1/2} [B_{11}^{(k)}]^{-1/2}, [I + \Delta_{22}^{(k)}]^{-1/2} [B_{22}^{(k)}]^{-1/2}), \tag{3.6b}$$

$$\hat{F}_{ij}^{(k)} = [I + \Delta_{ii}^{(k)}]^{-1/2} \Delta_{ij}^{(k)} [I + \Delta_{jj}^{(k)}]^{-1/2} \quad \text{for } i \neq j, \tag{3.6c}$$

$$\hat{A}_{ii}^{(k)} = [B_{ii}^{(k)}]^{-1/2} A_{ii}^{(k)} [B_{ii}^{(k)}]^{-1/2} \quad (= \lambda_0 I \text{ when } i = 1), \tag{3.6d}$$

$$\hat{E}_{ij}^{(k)} = [I + \Delta_{ii}^{(k)}]^{-1/2} [B_{ii}^{(k)}]^{-1/2} E_{ij}^{(k)} [B_{jj}^{(k)}]^{-1/2} [I + \Delta_{jj}^{(k)}]^{-1/2} \quad \text{for } i \neq j, \tag{3.6e}$$

$$\hat{E}_{ii}^{(k)} = [I + \Delta_{ii}^{(k)}]^{-1/2} [B_{ii}^{(k)}]^{-1/2} (A_{ii}^{(k)} + E_{ii}^{(k)}) [B_{ii}^{(k)}]^{-1/2} [I + \Delta_{ii}^{(k)}]^{-1/2} - \hat{A}_{ii}^{(k)}, \tag{3.6f}$$

and

$$\delta_{ij}^{(k)} = \|\Delta_{ij}^{(k)}\|_2 \le \sqrt{\|[B_{ii}^{(k)}]^{-1}\|_2 \|[B_{jj}^{(k)}]^{-1}\|_2} \|F_{ij}^{(k)}\|_2, \qquad \gamma_{ij}^{(k)} = (1-\delta_{ij}^{(k)})^{-1/2} - 1. \tag{3.7}$$

We can bound $\hat{E}_{ij}^{(k)}$ and $\hat{F}_{ij}^{(k)}$ as follows:

$$\|\hat{E}_{ij}^{(k)}\|_2 \le \frac{\|[B_{ii}^{(k)}]^{-1/2} E_{ij}^{(k)} [B_{jj}^{(k)}]^{-1/2}\|_2}{\sqrt{(1-\delta_{ii}^{(k)})(1-\delta_{jj}^{(k)})}} \tag{3.8a}$$

$$\le \sqrt{\frac{\|[B_{ii}^{(k)}]^{-1}\|_2 \|[B_{jj}^{(k)}]^{-1}\|_2}{(1-\delta_{ii}^{(k)})(1-\delta_{jj}^{(k)})}} \|E_{ij}^{(k)}\|_2 \quad \text{for } i \ne j, \tag{3.8b}$$

$$\|\hat{E}_{ii}^{(k)}\|_2 \le \gamma_{ii}^{(k)}(2+\gamma_{ii}^{(k)})\|\hat{A}_{ii}^{(k)}\|_2 \le \frac{\|[B_{ii}^{(k)}]^{-1/2} E_{ii}^{(k)} [B_{ii}^{(k)}]^{-1/2}\|_2}{1-\delta_{ii}^{(k)}} \tag{3.8c}$$

$$\le \gamma_{ii}^{(k)}(2+\gamma_{ii}^{(k)})\|\hat{A}_{ii}^{(k)}\|_2 + \frac{\|[B_{ii}^{(k)}]^{-1}\|_2}{1-\delta_{ii}^{(k)}} \|E_{ii}^{(k)}\|_2, \tag{3.8d}$$

$$\|\hat{F}_{ij}^{(k)}\|_2 \le \frac{\delta_{ij}^{(k)}}{\sqrt{(1-\delta_{ii}^{(k)})(1-\delta_{jj}^{(k)})}} \quad \text{for } i \ne j. \tag{3.8e}$$

The gaps η_j as previously defined, when applied to the current situation with the partitioning as in (3.5), are all zeros except when $k = m$ and λ_0 is not an eigenvalue of $A_{22} - \lambda B_{22}$. This makes Theorem 2.7 less favorable to apply than Theorem 2.8 because of the appearance of $\max_i \|\hat{E}_{ii}^{(k)}\|_2$. Also we are interested here only in how different copies of λ_0 change due to the perturbation. The following theorem is a consequence of Theorem 2.8.

THEOREM 3.1. *Suppose that the Hermitian definite GEP* (3.2) *is perturbed to* (3.4) *and assume* (3.3). *Let all assignments* (3.5)–(3.7) *hold. Then for any given* k $(1 \le k \le m)$, *there are* k *eigenvalues* $\mu_1 \ge \cdots \ge \mu_k$ *of* $\tilde{A} - \lambda \tilde{B}$ *such that*

$$|\mu_j - \lambda_0| \le \|\hat{E}_{11}^{(k)}\|_2 + \frac{\|\hat{E}_{21}^{(k)} - \hat{F}_{21}^{(k)}(\lambda_0 I + \hat{E}_{11}^{(k)})\|_2}{\sqrt{1 - \|\hat{F}_{21}^{(k)}\|_2^2}} \quad \textit{for } 1 \le j \le k, \tag{3.9}$$

$$\sqrt{\sum_{j=1}^{k} |\mu_j - \lambda_0|^2} \le \|\hat{E}_{11}^{(k)}\|_{\mathrm{F}} + \frac{\|\hat{E}_{21}^{(k)} - \hat{F}_{21}^{(k)}(\lambda_0 I + \hat{E}_{11}^{(k)})\|_{\mathrm{F}}}{\sqrt{1 - \|\hat{F}_{21}^{(k)}\|_2^2}}. \tag{3.10}$$

What makes this theorem interesting is that the right-hand sides of the inequalities may increase with k, illustrating different sensitivities of different copies of the multiple eigenvalue λ_0.

Example 3.1. Consider matrices A and B, slightly different from the ones in (3.1), by multiplying both 1000 and 2000 by 10 to avoid writing $\sqrt{10^{\pm 3}}$ later in the analysis for the sake of presentational convenience. It can be seen that

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad X^*AX = \begin{pmatrix} 2\times 10^4 & 0 \\ 0 & 2 \end{pmatrix}, \quad X^*BX = \begin{pmatrix} 10^4 & 0 \\ 0 & 1 \end{pmatrix}.$$

Perturb A and B by Hermitian matrices E and F with $\max_{i,j}\{|E_{(i,j)}|, |F_{(i,j)}|\} \le \varepsilon$. Because X here is a permutation matrix, $\max_{i,j}\{|(X^*EX)_{(i,j)}|, |(X^*FX)_{(i,j)}|\} \le \varepsilon$ as well. We shall now use Theorem 3.1 to bound how much the two copies of the multiple eigenvalue 2 may be perturbed. The application is done for $k = 1$ and 2. Recall that the right-hand sides of (3.9) and (3.10) depend on k; let ρ_k denote the right-hand side of (3.9).

$k = 1$: $\delta_{11}^{(k)} \le 10^{-4}\varepsilon$, $\gamma_{11}^{(k)} \le (\sqrt{1 - 10^{-4}\varepsilon})^{-1} - 1 \approx \frac{1}{2} \times 10^{-4}\varepsilon$, $\delta_{22}^{(k)} \le \varepsilon$, and

$$|\hat{E}_{11}^{(k)}| \le 2\gamma_{11}^{(k)}(2 + \gamma_{11}^{(k)}) + \frac{10^{-4}\varepsilon}{1 - 10^{-4}\varepsilon} \approx 3 \times 10^{-4}\varepsilon,$$

$$|\hat{E}_{21}^{(k)}|, |\hat{F}_{21}^{(k)}| \le \frac{10^{-2}\varepsilon}{(1 - 10^{-4}\varepsilon)(1 - \varepsilon)} \approx 10^{-2}\varepsilon.$$

Therefore, $\rho_1 \lessapprox 3 \times 10^{-4}\varepsilon + 3 \times 10^{-2}\varepsilon \approx 3 \times 10^{-2}\varepsilon$ after dropping higher order terms in ε. Here and in what follows, this "approximately less than" notation means the inequality holds up to the first order in ε.

$k = 2$: Now the blocks in the second row and column are empty. We have

$$\delta_{11}^{(k)} = \|\Delta_{11}^{(k)}\|_2 \le \|\Delta_{11}^{(k)}\|\mathrm{F} \le \sqrt{1 + 2 \cdot 10^{-4} + 10^{-8}}\varepsilon \approx (1 + 10^{-4})\varepsilon,$$

$$\gamma_{11}^{(k)} \le [1 - (1 + 10^{-4})\varepsilon]^{-1/2} - 1 \approx \frac{1}{2}(1 + 10^{-4})\varepsilon,$$

$$\|\hat{E}_{11}^{(k)}\|_2 \le 2\gamma_{11}^{(k)}(2 + \gamma_{11}^{(k)}) + \frac{(1 + 10^{-4})\varepsilon}{1 - \delta_{11}^{(k)}} \approx 3(1 + 10^{-4})\varepsilon.$$

Therefore, $\rho_2 \lessapprox 3(1 + 10^{-4})\varepsilon$, again after dropping higher order terms in ε. Putting these two facts together, we conclude that *the perturbed pencil has one eigenvalue that is away from* 2 *by approximately no more than* $3 \times 10^{-2}\varepsilon$*, while its other eigenvalue is away from* 2 *by approximately no more than* $3(1 + 10^{-4})\varepsilon$. Further detailed examination reveals that the copy 20000/10000 is much less sensitive to perturbations than the copy 2/1. The bounds are rather sharp. For example, in (3.1) if the $(1, 1)$th blocks of A and B are perturbed to $2 + \varepsilon$ and $1 - \varepsilon$, respectively, then the more sensitive copy 2/1 is changed to $(2 + \varepsilon)/(1 - \varepsilon) \approx 2 + 3\varepsilon$ whose first order term is 3ε, barely less than the bound on ρ_2. If A and B are perturbed to

$$A \to \begin{pmatrix} 2 & \varepsilon \\ \varepsilon & 20000 \end{pmatrix}, \qquad B \to \begin{pmatrix} 1 & -\varepsilon \\ -\varepsilon & 10000 \end{pmatrix},$$

where $\varepsilon \ge 0$, then the perturbed pencil has eigenvalues, to the first order of ε,

$$2 - 3 \times 10^{-2}\varepsilon, \qquad 2 + 3 \times 10^{-2}\varepsilon,$$

which suggests that our estimate on ρ_1 is also sharp. □

4. An extension to non-Hermitian pencil. In this section we will make an attempt to derive a quadratic eigenvalue bound for diagonalizable non-Hermitian pencils subject to off-diagonal perturbations. Specifically, let

$$A = \begin{matrix} & m & n \\ m & \multirow{2}{*}{} & \\ n & & \end{matrix} \begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix}, \qquad B = \begin{matrix} m \\ n \end{matrix} \overset{\begin{matrix} m & n \end{matrix}}{\begin{pmatrix} B_{11} & \\ & B_{22} \end{pmatrix}}, \tag{4.1a}$$

$$\tilde{A} = \begin{pmatrix} A_{11} & E_{12} \\ E_{21} & A_{22} \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} B_{11} & F_{12} \\ F_{21} & B_{22} \end{pmatrix} \tag{4.1b}$$

be non-Hermitian matrices. We assume that B is nonsingular and $A - \lambda B$ is diagonalizable. So $A - \lambda B$ has only finite eigenvalues, and there exist nonsingular matrices $X = \operatorname{diag}(X_1, X_2)$ and $Y = \operatorname{diag}(Y_1, Y_2)$ such that $YAX = \Lambda = \operatorname{diag}(\Lambda_1, \Lambda_2)$ and $YBX = I$, where X_1, Y_1, and Λ_1 are m-by-m and Λ is the diagonal matrix of eigenvalues. The last assumption loses little generality, since if $A - \lambda B$ is regular and diagonalizable but B is singular (hence infinite eigenvalues exist), we can apply the results below to $A - \lambda(B - \alpha A)$ for a suitable scalar α such that $B - \alpha A$ is nonsingular (the regularity assumption of $A - \lambda B$ ensures the existence of such α). The eigenvalues ν of $A - \lambda(B - \alpha A)$ and τ of $A - \lambda B$ are related by $\nu = \tau/(1 - \alpha\tau)$.

We will establish a bound on $|\mu - \tilde{\mu}|$, where μ is an eigenvalue of $A - \lambda B$ and $\tilde{\mu}$ is an eigenvalue of $\tilde{A} - \lambda\tilde{B}$.

THEOREM 4.1. *Let A, B, $\tilde{A}$, $\tilde{B}$ be as in* (4.1a) *and* (4.1b). *Suppose that there exist nonsingular matrices $X = \operatorname{diag}(X_1, X_2)$ and $Y = \operatorname{diag}(Y_1, Y_2)$ such that $YAX = \Lambda$ is diagonal and $YBX = I$. If $\tilde{\mu}$ is an eigenvalue of $\tilde{A} - \lambda\tilde{B}$ such that*

$$\eta_k \overset{\text{def}}{=} \min_{\mu \in \operatorname{eig}(A_{kk}, B_{kk})} |\tilde{\mu} - \mu| > 0$$

for $k = 1$ or 2, then $A - \lambda B$ has an eigenvalue μ such that

$$|\tilde{\mu} - \mu| \le \|X\|_2 \|Y\|_2 \|E_{12} - \tilde{\mu} F_{12}\|_2 \|E_{21} - \tilde{\mu} F_{21}\|_2 \|(A_{kk} - \tilde{\mu} B_{kk})^{-1}\|_2 \tag{4.2a}$$

$$\le \frac{\kappa_2(X)\kappa_2(Y)\|E_{12} - \tilde{\mu} F_{12}\|_2 \|E_{21} - \tilde{\mu} F_{21}\|_2}{\eta_k}. \tag{4.2b}$$

Proof. We prove the result only for $k = 2$. The proof for $k = 1$ is entirely analogous.

Suppose that $\tilde{\mu} \notin \operatorname{eig}(A_{22}, B_{22})$ is an eigenvalue of $\tilde{A} - \lambda\tilde{B}$. Thus $\tilde{A} - \tilde{\mu}\,\tilde{B}$ is singular. Defining the nonsingular matrices

$$W_L = \begin{pmatrix} I & -(E_{12} - \tilde{\mu} F_{12})(A_{22} - \tilde{\mu} B_{22})^{-1} \\ 0 & I \end{pmatrix},$$
$$W_R = \begin{pmatrix} I & 0 \\ -(A_{22} - \tilde{\mu} B_{22})^{-1}(E_{21} - \tilde{\mu} F_{21}) & I \end{pmatrix},$$

we see that

$$
\begin{aligned}
&\operatorname{diag}(Y_1, I_n)\, W_L(\tilde{A} - \tilde{\mu}\,\tilde{B})\, W_R \operatorname{diag}(X_1, I_n)\\
&\quad = \operatorname{diag}(Y_1, I_n)\\
&\qquad \times \begin{pmatrix} A_{11} - \tilde{\mu} B_{11} - (E_{12} - \tilde{\mu} F_{12})(A_{22} - \tilde{\mu} B_{22})^{-1}(E_{21} - \tilde{\mu} F_{21}) & 0 \\ 0 & A_{22} - \tilde{\mu} B_{22} \end{pmatrix}\\
&\qquad \times \operatorname{diag}(X, I_n)\\
&\quad = \begin{pmatrix} \Lambda_1 - \tilde{\mu} I_m - Y_1(E_{12} - \tilde{\mu} F_{12})(A_{22} - \tilde{\mu} B_{22})^{-1}(E_{21} - \tilde{\mu} F_{21}) X_1 & 0 \\ 0 & A_{22} - \tilde{\mu} B_{22} \end{pmatrix}.
\end{aligned}
$$

Since this matrix is also singular, it follows that $A_{11} - \lambda B_{11}$ must have an eigenvalue μ that satisfies

$$
\begin{aligned}
|\tilde{\mu} - \mu| &\leq \|Y_1\|_2 \|(E_{12} - \tilde{\mu} F_{12})(A_{22} - \tilde{\mu} B_{22})^{-1}(E_{21} - \tilde{\mu} F_{21})\|_2 \|X_1\|_2\\
&\leq \frac{1}{\eta_k} \|X_1\|_2 \|X_2^{-1}\|_2 \|Y_1\|_2 \|Y_2^{-1}\|_2 \|E_{12} - \tilde{\mu} F_{12}\|_2 \|E_{21} - \tilde{\mu} F_{21}\|_2,
\end{aligned}
$$

where we have used

$$
\|(A_{22} - \tilde{\mu} B_{22})^{-1}\|_2 = \|X_2^{-1}(\Lambda_2 - \tilde{\mu} I)^{-1} Y_2^{-1}\|_2 \leq \|X_2^{-1}\|_2 \|Y_2^{-1}\|_2 / \eta_k.
$$

Now use $\|X\|_2 = \max\{\|X_1\|_2, \|X_2\|_2\}$ and $\|X^{-1}\|_2 = \max\{\|X_1^{-1}\|_2, \|X_2^{-1}\|_2\}$ to get (4.2b) for the case $k = 2$. ∎

When the pencils $A - \lambda B$ and $\tilde{A} - \lambda\tilde{B}$ are Hermitian definite and (2.3) holds, we have $\kappa_2(X) = \kappa_2(Y) = 1$; so (4.2b) reduces to $|\tilde{\mu} - \mu| \leq \|E_{12} - \tilde{\mu} F_{12}\|_2^2 / \eta_k$. This is similar to our earlier result (2.15), except for the slight difference in the denominator. If we further let $F_{12} = 0$, then the expression (4.2b) becomes exactly that of the quadratic residual bound in [10] for Hermitian matrices. However, (4.2b) does not give a one-to-one pairing between the eigenvalues of $A - \lambda B$ and $\tilde{A} - \lambda\tilde{B}$.

The assumption that $\eta_k > 0$ is a reasonable one when there is a gap between $\operatorname{eig}(A_{kk}, B_{kk})$ for $k = 1, 2$, because then it is reasonable to expect that $\tilde{\mu}$ is very near one of them but away from the other.

5. Concluding remarks.

In this paper we have shown three different approaches for constructing perturbation bounds between the eigenvalues of $A - \lambda B$ and those of $\tilde{A} - \lambda\tilde{B}$ as well as the eigenvalues of $A_{11} - \lambda B_{11}$ and some of those of $\tilde{A} - \lambda\tilde{B}$. Our bounds work well regardless of eigenvalue gaps, and they are sharper than existing ones. The distinguished feature that different copies of a multiple eigenvalue may exhibit quite different sensitivities can also be explained by our bounds. As an attempt to extend our earlier quadratic perturbation bounds for Hermitian definite pencils, we also investigated a diagonalizable matrix pencil and obtained a quadratic perturbation bound for it.

Appendix A. The min-max principle.

The results in this section are known and hold for any GEP $A - \lambda B$ for which A and B are Hermitian and B is positive definite, not necessarily having the form as in (1.1). Assume A and B are such $N \times N$ matrices and the eigenvalues of $A - \lambda B$ are

$$
\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N.
$$

The min-max principle [1], [13], [14] is often stated for the standard Hermitian eigenvalue problem, i.e., $B = I_N$. But it can be easily extended to the GEP. Namely,

$$\lambda_j = \min_{\mathcal{S}_{N-j+1}} \max_{x \in \mathcal{S}_{N-j+1}} \frac{x^* A x}{x^* B x}, \qquad \lambda_j = \max_{\mathcal{S}_j} \min_{x \in \mathcal{S}_j} \frac{x^* A x}{x^* B x}, \tag{A.1}$$

where $\mathcal{S}_j$ denotes a j-dimensional subspace of $\mathbb{C}^N$. In fact, since

$$B^{1/2}\{\mathcal{S}_j \subset \mathbb{C}^N\} = \{\mathcal{S}_j \subset \mathbb{C}^N\}, \qquad \operatorname{eig}(A, B) = \operatorname{eig}(B^{-1/2} A B^{-1/2}),$$

we have

$$\max_{\mathcal{S}_j} \min_{x \in \mathcal{S}_j} \frac{x^* A x}{x^* B x} = \max_{\mathcal{S}_j} \min_{y \in \mathcal{S}_j} \frac{y^* B^{-1/2} A B^{-1/2} y}{y^* y} = \lambda_j, \tag{A.2}$$

where the last equality is due to the max-min principle for the standard Hermitian eigenvalue problem. The second equation in (A.1) is thus obtained. In the same way, we can get the first equation in (A.1). In particular, (A.1) gives

$$\lambda_1 = \max_x \frac{x^* A x}{x^* B x}, \qquad \lambda_N = \min_x \frac{x^* A x}{x^* B x}.$$

The Cauchy interlacing property [13] can be extended to the GEP, too. Let A_1 and B_1 be obtained by deleting the kth rows and columns from both A and B, respectively. Then A_1 and B_1 are still Hermitian, and B_1 is still positive definite. Denote the eigenvalues of $A_1 - \lambda B_1$ by

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{N-1}.$$

Then by (A.1) and the same argument as that in the proof for the standard case [4, p. 186], one can prove

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \cdots \geq \lambda_j \geq \mu_j \geq \lambda_{j+1} \geq \cdots \geq \mu_{N-1} \geq \lambda_N. \tag{A.3}$$

Acknowledgment. The authors wish to thank the referees for their careful readings of the manuscript and constructive suggestions that improved the presentation considerably.

REFERENCES

[1] R. Bhatia, *Matrix Analysis*, Grad. Texts in Math. 169, Springer, New York, 1996.

[2] Z.-H. Cao, J.-J. Xie, and R.-C. Li, *A sharp version of Kahan's theorem on clustered eigenvalues*, Linear Algebra Appl., 245 (1996), pp. 147–155.

[3] S. C. Eisenstat and I. C. F. Ipsen, *Relative perturbation techniques for singular value problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1972–1988.

[4] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[5] W. Kahan, *Inclusion Theorems for Clusters of Eigenvalues of Hermitian Matrices*, Technical report, Computer Science Department, University of Toronto, 1967.

[6] C.-K. Li and R.-C. Li, *A note on eigenvalues of perturbed hermitian matrices*, Linear Algebra Appl., 395 (2005), pp. 183–190.

[7] C.-K. Li and R. Mathias, *The Lidskii–Mirsky–Wielandt theorem—additive and multiplicative versions*, Numer. Math., 81, pp. 377–413.

[8] R.-C. Li, *Relative perturbation theory* I. *Eigenvalue and singular value variations*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 956–982.

[9] R.-C. Li, *Matrix perturbation theory*, in Handbook of Linear Algebra, L. Hogben, R. Brualdi, A. Greenbaum, and R. Mathias, eds., CRC Press, Boca Raton, FL, 2006.
[10] R. Mathias, *Quadratic residual bounds for the Hermitian eigenvalue problem*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 541–550.
[11] Y. Nakatsukasa, *Perturbation behavior of a multiple eigenvalue in generalized Hermitian eigenvalue problems*, BIT, 50 (2010), pp. 109–121.
[12] A. M. Ostrowski, *A quantitative formulation of Sylvester's law of inertia*, Proc. Natl. Acad. Sci. U.S.A., 45 (1959), pp. 740–744.
[13] B. N. Parlett, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.
[14] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[15] J.-G. Sun, *Eigenvalues of Rayleigh quotient matrices*, Numer. Math., 59 (1991), pp. 603–614.