

## ON PERTURBATIONS OF MATRIX PENCILS WITH REAL SPECTRA, A REVISIT

REN-CANG LI

ABSTRACT. This paper continues earlier studies by Bhatia and Li on eigenvalue perturbation theory for diagonalizable matrix pencils having real spectra. A unifying framework for creating crucial perturbation equations is developed. With the help of a recent result on generalized commutators involving unitary matrices, new and much sharper bounds are obtained.

### 1. INTRODUCTION

Consider an  $n \times n$  matrix pencil  $A - \lambda B$ , where  $A, B$  are constant matrices. The pencil is *regular* if  $\det(A - \lambda B) \neq 0$  for all complex  $\lambda$ . The generalized eigenvalue problem for the pencil is to determine nonzero pairs  $(\alpha, \beta) \neq (0, 0)$  and nonzero  $n$ -dimensional vectors  $x, y$  such that

$$\beta Ax = \alpha Bx, \quad \beta y^* A = \alpha y^* B,$$

where the asterisk denotes matrix conjugate and transpose, and  $(\alpha, \beta)$  is called a (*generalized*) *eigenvalue* and  $x$  and  $y$  are called the corresponding (*right*) *eigenvector* and *left eigenvector*. In using a nonzero pair to represent an eigenvalue, one takes into account the case of singular  $B$  when the pencil has an eigenvalue  $\infty$ , which now is simply denoted as  $(0, 1)$ .

Throughout this paper, all pairs of interest are nonzero. Two pairs  $(\alpha, \beta)$  and  $(\tilde{\alpha}, \tilde{\beta})$  are *equivalent* if  $(\tilde{\alpha}, \tilde{\beta}) = (\xi\alpha, \xi\beta)$  for some  $\xi \neq 0$ . As far as eigenvalues are concerned, all members in an equivalence class represent the same one. We say that  $(\alpha, \beta)$  is *real* if both entries of a member in its equivalence class are real. An  $n \times n$  regular matrix pencil  $A - \lambda B$  is *diagonalizable*, if there exist  $n \times n$  invertible matrices  $X, Y$  such that

$$(1.1) \quad Y^*AX = \Lambda \equiv \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad Y^*BX = \Omega \equiv \text{diag}(\beta_1, \beta_2, \dots, \beta_n).$$

Definite matrix pencils, perhaps the most important class of matrix pencils, are always diagonalizable and have real spectra, see § 5 below.

The existing results (in the spectral norm) [5, 11] imply that if a diagonalizable pencil is perturbed to another one, also diagonalizable, and if both have real spectra, then there are a one-to-one pairing among the eigenvalues of the two pencils

---

Received by the editor January 10, 2001 and, in revised form, August 24, 2001.

2000 *Mathematics Subject Classification*. Primary 15A22, 15A42, 65F15.

*Key words and phrases*. Diagonalizable matrix pencil, definite pencil, real spectrum, unitarily invariant norm, perturbation bound.

This work was supported in part by the National Science Foundation under Grant No. ACI-9721388 and by the National Science Foundation CAREER award under Grant No. CCR-9875201.

and a uniform bound on the chordal differences of the paired eigenvalues, independent of the matrix dimensions. Roughly, the uniform bound is the square of the condition number of the eigenvector matrix times the perturbation errors in the perturbed pencil, see Remark 4.1 below. (There are versions of such results in the Frobenius norm, as well as all unitarily invariant norms.) Two crucial ingredients in the development of perturbation theory for diagonalizable pencils are the creation of right perturbation equations and the use of perturbation bounds for unitary matrices. This paper refines our previous developments and makes two contributions:

1. a unifying framework for creating perturbation equations that include those in [5, 11] and many more (thus provide room for optimizations as needed), and
2. significantly improved bounds in which roughly the square of the condition number of the eigenvector matrix in the old bounds is replaced by the condition number itself.

This second contribution is made possible by the use of some recent results on generalized commutators [4].

The rest of this paper is organized as follows. §2 presents the unifying framework for perturbation equations and shows how to pick (nearly) optimal ones. §3 establish the key result that makes new and sharper bounds possible. We derive perturbation results for general diagonalizable pencils with real spectra with comparisons to the existing ones in §4 and pay special attention to the most important class of matrix pencils – definite matrix pencils in §5. §6 summarizes this paper’s contributions, together with comments as well as open questions on diagonalizable pencils that have complex spectra.

*Notation.*  $\mathbb{C}$ ,  $\mathbb{C}^n$ , and  $\mathbb{C}^{m \times n}$  are the sets of complex numbers,  $n$ -dimensional complex vectors, and  $m \times n$  complex matrices, respectively.  $I_n \in \mathbb{C}^{n \times n}$  is the identity matrix, and when no ambiguity can arise the subscript  $n$  is omitted.  $\sigma_1(X) \geq \sigma_2(X) \geq \dots \geq 0$  denote  $X$ ’s singular values in descending order.  $\|X\|$  is reserved for a general unitarily invariant norm; two frequently used ones are the spectral norm  $\|X\|_2 \equiv \sigma_1(X)$  and the Frobenius norm  $\|X\|_F \equiv [\sum_i \sigma_i(X)^2]^{1/2}$ .  $X^\dagger$  is the Moore-Penrose inverse of  $X \in \mathbb{C}^{m \times n}$  [15, pp.102-105]. It can be proved that

$$X^\dagger = (X^*X)^{-1}X^* \text{ if } \text{rank}(X) = n \quad \text{and} \quad X^\dagger = X^*(XX^*)^{-1} \text{ if } \text{rank}(X) = m.$$

$\mathcal{P}_X$  denotes the orthogonal projection onto the column space of  $X$ . It can be proved that [15, Theorem 1.3, p.106]

$$\mathcal{P}_X = XX^\dagger, \quad \mathcal{P}_{X^*} = X^\dagger X.$$

The *chordal metric* is defined for  $(\alpha, \beta) \neq (0, 0) \neq (\tilde{\alpha}, \tilde{\beta})$  by

$$(1.2) \quad \rho((\alpha, \beta), (\tilde{\alpha}, \tilde{\beta})) \stackrel{\text{def}}{=} \frac{|\tilde{\beta}\alpha - \tilde{\alpha}\beta|}{\sqrt{|\alpha|^2 + |\beta|^2} \sqrt{|\tilde{\alpha}|^2 + |\tilde{\beta}|^2}}.$$

## 2. PERTURBATION EQUATIONS

We shall develop in this section a unifying framework for creating perturbation equations. Previously used ones which were the cornerstones of the earlier paper [5, 11] can be realized at a particular choice of built-in freedoms. Because of these built-in freedoms, we will be able to pick more suitable perturbation equations

so as to improve the existing bounds on generalized eigenvalues. In the following theorem,  $m$  is an arbitrary integer.

**Theorem 2.1.** *Let  $A - \lambda B \in \mathbb{C}^{n \times n}$  be diagonalizable, and admit the eigendecomposition (1.1), and let  $\tilde{A} - \lambda \tilde{B} \in \mathbb{C}^{n \times n}$  be another pencil. Then for any  $W, \tilde{W} \in \mathbb{C}^{m \times n}$*

$$(2.1) \quad -\tilde{W}(\tilde{A}XY^*B - \tilde{B}XY^*A) = [W(A \ B) - \tilde{W}(\tilde{A} \ \tilde{B})] \begin{pmatrix} XY^* & \\ & XY^* \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}.$$

*Proof.* Noticing that

$$(2.2) \quad AXY^*B - BXY^*A = Y^{-*}(\Lambda\Omega - \Omega\Lambda)X^{-1} = 0,$$

we have

$$\begin{aligned} &-\tilde{W}(\tilde{A}XY^*B - \tilde{B}XY^*A) \\ &= -\tilde{W}(\tilde{A}XY^*B - \tilde{B}XY^*A) + W(AXY^*B - BXY^*A) \\ &= [W(A \ B) - \tilde{W}(\tilde{A} \ \tilde{B})] \begin{pmatrix} XY^* & \\ & XY^* \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}, \end{aligned}$$

as expected. □

*Remark 2.1.* Equation (2.1) can be seen to hold under a weaker condition, namely,  $Y^*AX$  and  $Y^*BX$  commute, where  $X$  and  $Y$  are nonsingular. This condition is fulfilled here by both  $Y^*AX$  and  $Y^*BX$  being diagonal.

Define

$$(2.3) \quad Z \stackrel{\text{def}}{=} (A \ B), \quad \tilde{Z} \stackrel{\text{def}}{=} (\tilde{A} \ \tilde{B}) \in \mathbb{C}^{2n \times n}.$$

The regularity<sup>1</sup> (which will be assumed throughout this paper) of  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  guarantees that  $\text{rank}(Z) = n = \text{rank}(\tilde{Z})$ .

Equation (2.1) is not entirely new. It appeared in [11] for special  $W$  and  $\tilde{W}$ , namely the first two selections in the following list:

1.  $W = \tilde{W} = I_n$ .
2.  $W = Z^\dagger$  and  $\tilde{W} = \tilde{Z}^\dagger$ .
3.  $S \in \mathbb{C}^{n \times n}$  is unitary and to be determined. Either

$$(2.4) \quad W = (ZZ^*)^{-1/2}, \quad \tilde{W} = S(\tilde{Z}\tilde{Z}^*)^{-1/2},$$

or

$$(2.5) \quad W = S^*(ZZ^*)^{-1/2}, \quad \tilde{W} = (\tilde{Z}\tilde{Z}^*)^{-1/2}.$$

4.  $S \in \mathbb{C}^{n \times n}$  is arbitrary and to be determined. Either (2.4) or (2.5).

In the last two selections,  $S$  will be determined by solving one of the following minimization problems:

$$\min_S \left\| (ZZ^*)^{-1/2}Z - S(\tilde{Z}\tilde{Z}^*)^{-1/2}\tilde{Z} \right\|, \quad \min_S \left\| S^*(ZZ^*)^{-1/2}Z - (\tilde{Z}\tilde{Z}^*)^{-1/2}\tilde{Z} \right\|$$

for all or some special unitarily invariant norms.

In our later development, the norm of  $W(A \ B) - \tilde{W}(\tilde{A} \ \tilde{B})$  measures the error that perturbs  $A - \lambda B$  to  $\tilde{A} - \lambda \tilde{B}$ . There is an obvious reason to choose the first one,  $W = \tilde{W} = I_n$ , because it will explicitly result in the difference  $Z - \tilde{Z}$ , which is

---

<sup>1</sup>It is not necessary to require this in Theorem 2.1.

naturally interpreted as the perturbation that moves the pencil  $A - \lambda B$  to  $\tilde{A} - \lambda \tilde{B}$ . The second selection is attributed to, though not in the context of (2.1), J.-G. Sun, who pioneered the use of the sines of the canonical angles between the column subspace of  $Z^*$  and that of  $\tilde{Z}^*$  to measure the error in a regular pencil [17]. It can be proved, see, e.g., [15, Theorem 5.5, p.43], that if the canonical angles are (arranged in increasing order)

$$(2.6) \quad 0 \leq \theta_n \leq \dots \leq \theta_2 \leq \theta_1 \leq \pi/2,$$

then the singular values of  $Z^\dagger Z - \tilde{Z}^\dagger \tilde{Z} \equiv \mathcal{P}_{Z^*} - \mathcal{P}_{\tilde{Z}^*}$  are

$$\sin \theta_1, \sin \theta_1, \sin \theta_2, \sin \theta_2, \dots, \sin \theta_n, \sin \theta_n.$$

Write the canonical angle matrix<sup>2</sup>

$$(2.7) \quad \Theta(Z^*, \tilde{Z}^*) = \text{diag}(\theta_1, \theta_2, \dots, \theta_n).$$

We have

$$\|\sin \Theta(Z^*, \tilde{Z}^*)\|_2 = \|\mathcal{P}_{Z^*} - \mathcal{P}_{\tilde{Z}^*}\|_2, \quad \|\sin \Theta(Z^*, \tilde{Z}^*)\|_F = \frac{1}{\sqrt{2}} \|\mathcal{P}_{Z^*} - \mathcal{P}_{\tilde{Z}^*}\|_F.$$

Clearly the third selection yields a bigger error than the fourth, so we shall leave it out in our later considerations. Now we turn to the fourth one, which eventually will help us to get rid of the unnatural condition number  $\max\{\|X\|_2, \|X^{-1}\|_2\}^2$  previously appeared in the perturbation theorems for the Frobenius norm [11, Theorem 3.2] and all unitarily invariant norms [5, Theorem 3.1]. For convenience, we introduce

$$(2.8) \quad Z_0 = (ZZ^*)^{-1/2}Z, \quad \tilde{Z}_0 = (\tilde{Z}\tilde{Z}^*)^{-1/2}\tilde{Z}.$$

**Lemma 2.1.** *Let  $S_{\text{opt}} = Z_0 \tilde{Z}_0^*$ .*

1. For  $i = 1, 2, \dots, n$ ,

$$\sigma_i(Z_0 - S_{\text{opt}} \tilde{Z}_0) = \sigma_i(S_{\text{opt}}^* Z_0 - \tilde{Z}_0) = \sin \theta_i,$$

where  $\theta_i$  are the canonical angles between the column subspace of  $Z^*$  and that of  $\tilde{Z}^*$  as in (2.6), and thus for all unitarily invariant norms

$$\| \| Z_0 - S_{\text{opt}} \tilde{Z}_0 \| \| = \| \| S_{\text{opt}}^* Z_0 - \tilde{Z}_0 \| \| = \| \| \sin \Theta(Z^*, \tilde{Z}^*) \| \|.$$

2. Let  $S \in \mathbb{C}^{n \times n}$  be arbitrary. For  $i = 1, 2, \dots, n$ ,

$$\sigma_i(Z_0 - S \tilde{Z}_0) \geq \sigma_i(Z_0 - S_{\text{opt}} \tilde{Z}_0), \quad \sigma_i(S^* Z_0 - \tilde{Z}_0) \geq \sigma_i(S_{\text{opt}}^* Z_0 - \tilde{Z}_0).$$

So  $S_{\text{opt}}$  minimizes both  $\| \| Z_0 - S \tilde{Z}_0 \| \|$  and  $\| \| S^* Z_0 - \tilde{Z}_0 \| \|$ , simultaneously.

*Proof.* We notice that there are unitary matrices  $Q \in \mathbb{C}^{2n \times 2n}$ ,  $U, V \in \mathbb{C}^{n \times n}$  such that [15, Theorem 5.2, p.40]

$$UZ_0Q = (I \ 0), \quad V\tilde{Z}_0Q = (\Gamma \ \Sigma),$$

where  $\Gamma = \text{diag}(\cos \theta_1, \cos \theta_2, \dots, \cos \theta_n)$  and  $\Sigma = \text{diag}(\sin \theta_1, \sin \theta_2, \dots, \sin \theta_n)$ . We have  $S_{\text{opt}} = U^* \Gamma V$ , and

$$U(Z_0 - S_{\text{opt}} \tilde{Z}_0)Q = (I \ 0) - \Gamma(\Gamma \ \Sigma) = (I - \Gamma^2 \ -\Gamma\Sigma) = \Sigma(\Sigma \ -\Gamma);$$

<sup>2</sup>This  $\Theta(Z^*, \tilde{Z}^*)$  is equivalent to  $\Theta_r(Z, \tilde{Z})$  in [11] in the sense that the two share the same singular values.

so  $[U(Z_0 - S_{\text{opt}}\tilde{Z}_0)Q][U(Z_0 - S_{\text{opt}}\tilde{Z}_0)Q]^* = \Sigma^2$ , which implies  $\sigma_i(Z_0 - S_{\text{opt}}\tilde{Z}_0) = \sin \theta_i$ , as desired. On the other hand,

$$V(S_{\text{opt}}^*Z_0 - \tilde{Z}_0)Q = \Gamma(I \ 0) - (\Gamma \ \Sigma) = (0 \ -\Sigma),$$

which implies  $\sigma_i(S_{\text{opt}}^*Z_0 - \tilde{Z}_0) = \sin \theta_i$ . This proves item (1).

Write  $Z_0 - S\tilde{Z}_0 = (Z_0 - S_{\text{opt}}\tilde{Z}_0) + (S - S_{\text{opt}})\tilde{Z}_0$ . Noticing that  $(Z_0 - S_{\text{opt}}\tilde{Z}_0)\tilde{Z}_0^* = 0$ , we have

$$\begin{aligned} (Z_0 - S\tilde{Z}_0)(Z_0 - S\tilde{Z}_0)^* &= (Z_0 - S_{\text{opt}}\tilde{Z}_0)(Z_0 - S_{\text{opt}}\tilde{Z}_0)^* + [(S - S_{\text{opt}})\tilde{Z}_0][(S - S_{\text{opt}})\tilde{Z}_0]^*, \end{aligned}$$

which leads to  $\sigma_i(Z_0 - S\tilde{Z}_0) \geq \sigma_i(Z_0 - S_{\text{opt}}\tilde{Z}_0)$  for all  $i$ . Analogously we have

$$\begin{aligned} (S^*Z_0 - \tilde{Z}_0)(S^*Z_0 - \tilde{Z}_0)^* &= (S_{\text{opt}}^*Z_0 - \tilde{Z}_0)(S_{\text{opt}}^*Z_0 - \tilde{Z}_0)^* + [(S - S_{\text{opt}})^*Z_0][(S - S_{\text{opt}})^*Z_0]^*, \end{aligned}$$

and hence  $\sigma_i(S^*Z_0 - \tilde{Z}_0) \geq \sigma_i(S_{\text{opt}}^*Z_0 - \tilde{Z}_0)$  for all  $i$ . This is item (2). □

With our own best interests in mind we shall require that  $\tilde{W}$  in (2.1) have full column rank and thus  $\tilde{W}^\dagger \tilde{W} = I$ ; so

$$(2.9) \quad \tilde{A}XY^*B - \tilde{B}XY^*A = -\tilde{W}^\dagger [W(A \ B) - \tilde{W}(\tilde{A} \ \tilde{B})] \begin{pmatrix} XY^* & \\ & XY^* \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}.$$

This is valid for the cases of selections (1), (2), and (4) with (2.5).

Also to our interest in this paper,  $\tilde{A} - \lambda\tilde{B}$  is diagonalizable and admits the eigendecomposition as in (1.1) except with all symbols having tildes, i.e.,

$$(2.10) \quad \tilde{Y}^*\tilde{A}\tilde{X} = \tilde{\Lambda} \equiv \text{diag}(\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_n), \quad \tilde{Y}^*\tilde{B}\tilde{X} = \tilde{\Omega} \equiv \text{diag}(\tilde{\beta}_1, \tilde{\beta}_2, \dots, \tilde{\beta}_n),$$

where  $\tilde{X}$  and  $\tilde{Y}$  are nonsingular.

**Theorem 2.2.** *Let  $A - \lambda B, \tilde{A} - \lambda\tilde{B} \in \mathbb{C}^{n \times n}$  be diagonalizable, and admit the eigendecompositions (1.1) and (2.10), respectively, and let  $W, \tilde{W} \in \mathbb{C}^{m \times n}$ .*

1. *If  $\tilde{W}$  has full column rank, then*

$$(2.11) \quad \tilde{\Lambda}\tilde{X}^{-1}X\Omega - \tilde{\Omega}\tilde{X}^{-1}X\Lambda = -\tilde{Y}^*\tilde{W}^\dagger(WZ - \tilde{W}\tilde{Z}) \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}.$$

2. *If  $W$  has full column rank, then*

$$(2.12) \quad \Lambda X^{-1}\tilde{X}\tilde{\Omega} - \Omega X^{-1}\tilde{X}\tilde{\Lambda} = -Y^*W^\dagger(\tilde{W}\tilde{Z} - WZ) \begin{pmatrix} \tilde{X} & \\ & \tilde{X} \end{pmatrix} \begin{pmatrix} \tilde{\Omega} \\ -\tilde{\Lambda} \end{pmatrix}.$$

*Proof.* Equation (2.11) follows from (2.9), and the eigendecompositions (1.1) and (2.10). Equation (2.12) is a consequence of (2.11) with the positions of  $A - \lambda B$  and  $\tilde{A} - \lambda\tilde{B}$  swapped. □

**Lemma 2.2.** *Let  $A - \lambda B \in \mathbb{C}^{n \times n}$  be diagonalizable, and admit the eigendecomposition (1.1). If  $|\alpha_j|^2 + |\beta_j|^2 = 1$  for all  $j$  and  $W = (ZZ^*)^{-1/2}$ , then*

$$\|Y^*W^\dagger\|_2 \leq \|X^{-1}\|_2.$$

*Proof.* Since  $W$  is invertible,  $W^\dagger = W^{-1} = (ZZ^*)^{1/2}$ . Then

$$\begin{aligned} (Y^*W^\dagger)(Y^*W^\dagger)^* &= Y^*(ZZ^*)Y = \Lambda X^{-1}X^{-*}\Lambda^* + \Omega X^{-1}X^{-*}\Omega^* \\ &\leq \|X^{-1}\|_2^2(\Lambda\Lambda^* + \Omega\Omega^*) = \|X^{-1}\|_2^2 I, \end{aligned}$$

which implies the desired inequality. Here for Hermitian matrices  $M$  and  $N$ ,  $M \leq N$  means that  $N - M$  is positive semidefinite.  $\square$

**Lemma 2.3** ([11, Lemma 3.4, p.248]). *Let  $A - \lambda B \in \mathbb{C}^{n \times n}$  be diagonalizable, and admit the eigendecomposition (1.1). If  $|\alpha_j|^2 + |\beta_j|^2 = 1$  for all  $j$ , then*

$$\|Y^*\|_2 \leq \|X^{-1}\|_2 \|Z^\dagger\|_2.$$

*Remark 2.2.* Although this paper mainly deals with matrix pencils with real spectra, this section does not require that all  $\alpha_j$ 's,  $\beta_j$ 's,  $\tilde{\alpha}_j$ 's, and  $\tilde{\beta}_j$ 's be real.

### 3. NORMS OF CERTAIN MATRICES

Throughout this section,  $\Lambda$  and  $\Omega$  denote

$$(3.1) \quad \Lambda = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad \Omega = \text{diag}(\beta_1, \beta_2, \dots, \beta_n),$$

and similarly for  $\tilde{\Lambda}$  and  $\tilde{\Omega}$ , except with tildes over the associated symbols. We shall assume

$$(3.2) \quad \alpha_j \text{'s, } \beta_j \text{'s, } \tilde{\alpha}_j \text{'s, and } \tilde{\beta}_j \text{'s are all real and } \alpha_j^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1 \text{ for all } j.$$

We shall look for bounds from below on

$$\left\| \tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda \right\| \left\| \Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda} \right\|$$

for special or all unitarily invariant norms, where  $T$  is nonsingular. We accomplish this by first transforming both  $\tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda$  and  $\Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda}$  into generalized commutators involving two unitary matrices as in the proofs of [9, p.220] and [5, p.642], and then use recent results due to Bhatia, Kittaneh, and Li [4].

**Lemma 3.1** (Bhatia, Kittaneh, and Li [4]). *Let  $\Xi$  be positive definite, and let  $M$  and  $\tilde{M}$  be unitary. Then, for any unitarily invariant norms,*

$$\left\| \tilde{M}\Xi - \Xi M \right\| \left\| M\Xi^{-1} - \Xi^{-1}\tilde{M} \right\| \geq \left\| \tilde{M} - M \right\|^2.$$

Let  $i$  be the imaginary unit. Define

$$\begin{pmatrix} \Lambda_0 \\ \Omega_0 \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \Lambda \\ \Omega \end{pmatrix}, \quad \begin{pmatrix} \tilde{\Lambda}_0 \\ \tilde{\Omega}_0 \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \tilde{\Lambda} \\ \tilde{\Omega} \end{pmatrix},$$

$$D \stackrel{\text{def}}{=} \Omega_0^* \Lambda_0, \quad \tilde{D} \stackrel{\text{def}}{=} \tilde{\Omega}_0^* \tilde{\Lambda}_0.$$

It is easy to see that these six matrices just defined are diagonal and have unimodular diagonal entries, and thus are all unitary. It can be verified that

$$\begin{aligned} \tilde{\Lambda}_0 T \Omega_0 - \tilde{\Omega}_0 T \Lambda_0 &= 2 \left( \tilde{\Lambda} T \Omega - \tilde{\Omega} T \Lambda \right), \\ \Lambda_0 T^{-1} \tilde{\Omega}_0 - \Omega_0 T^{-1} \tilde{\Lambda}_0 &= 2 \left( \Lambda T^{-1} \tilde{\Omega} - \Omega T^{-1} \tilde{\Lambda} \right). \end{aligned}$$

Since  $\Omega_0$  and  $\tilde{\Omega}_0$  are unitary and diagonal, we then have

$$(3.3) \quad \tilde{D}T - TD = 2\tilde{\Omega}_0^* (\tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda) \Omega_0^*,$$

$$(3.4) \quad DT^{-1} - T^{-1}\tilde{D} = 2\Omega_0^* (\Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda}) \tilde{\Omega}_0^*.$$

Let  $T = U\Xi V^*$  be the SVD of  $T$ , and set

$$M = U^*\tilde{D}U, \quad \tilde{M} = V^*DV, \quad \text{both unitary.}$$

Pre-multiply and post-multiply (3.3) by  $U^*$  and  $V$ , respectively, to get

$$(3.5) \quad \tilde{M}\Xi - \Xi M = 2U^*\tilde{\Omega}_0^* (\tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda) \Omega_0^*V,$$

and pre-multiply and post-multiply (3.4) by  $V^*$  and  $U$ , respectively, to get

$$(3.6) \quad M\Xi^{-1} - \Xi^{-1}\tilde{M} = 2V^*\Omega_0^* (\Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda}) \tilde{\Omega}_0^*U.$$

The conditions of Lemma 3.1 are satisfied with  $M, \tilde{M}$  and  $\Xi$  here; so we have

**Theorem 3.1.** *For any nonsingular  $T \in \mathbb{C}^{n \times n}$ , there are a unitary matrix  $M$  whose eigenvalues are  $\{(\alpha_j + i\beta_j)/(i\alpha_j + \beta_j)\}_{j=1}^n$  and a unitary matrix  $\tilde{M}$  whose eigenvalues are  $\{(\tilde{\alpha}_j + i\tilde{\beta}_j)/(i\tilde{\alpha}_j + \tilde{\beta}_j)\}_{j=1}^n$  such that*

$$\left\| \tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda \right\| \left\| \Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda} \right\| \geq \frac{1}{4} \left\| \tilde{M} - M \right\|^2,$$

where  $\alpha_j$ 's,  $\beta_j$ 's,  $\tilde{\alpha}_j$ 's, and  $\tilde{\beta}_j$ 's are assumed to satisfy (3.2), and  $\Lambda$  and  $\Omega$  are defined by (3.1).

Notice that

$$(3.7) \quad \left| \frac{\alpha_j + i\beta_j}{i\alpha_j + \beta_j} - \frac{\tilde{\alpha}_k + i\tilde{\beta}_k}{i\tilde{\alpha}_k + \tilde{\beta}_k} \right| = 2\rho((\alpha_j, \beta_j), (\tilde{\alpha}_k, \tilde{\beta}_k)).$$

Next we shall employ the following existing perturbation results on unitary matrices to deduce a corollary to Theorem 3.1. Let  $M$  and  $\tilde{M}$  be assigned as above. Then there is a permutation matrix  $P$  such that ([3], [1, p.178])

$$(3.8) \quad \left\| D - P\tilde{D}P^{-1} \right\| \leq \frac{\pi}{2} \left\| \tilde{M} - M \right\|.$$

The factor  $\frac{\pi}{2}$  in inequality (3.8) can be replaced by 1 in two special cases: the spectral norm and the Frobenius norm. For the spectral norm  $\|\cdot\|_2$  this is a consequence of results in [2]. For the Frobenius norm  $\|\cdot\|_F$  this is subsumed in the more general Hoffman-Wielandt theorem [8]. Combine these with Theorem 3.1 and equation (3.7) to get

**Corollary 3.1.** *Under the conditions of Theorem 3.1, there exists a permutation  $\{j_1, j_2, \dots, j_n\}$  of  $\{1, 2, \dots, n\}$  such that*

$$\left\| \Delta \right\|^2 \leq \frac{\pi^2}{4} \left\| \tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda \right\| \left\| \Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda} \right\|,$$

where

$$(3.9) \quad \Delta \equiv \text{diag} \left( \rho((\alpha_1, \beta_1), (\tilde{\alpha}_{j_1}, \tilde{\beta}_{j_1})), \dots, \rho((\alpha_n, \beta_n), (\tilde{\alpha}_{j_n}, \tilde{\beta}_{j_n})) \right).$$

The factor  $\frac{\pi^2}{4}$  can be replaced by 1 in two special cases: the spectral norm and the Frobenius norm.

## 4. PERTURBATION BOUNDS—DIAGONALIZABLE PENCILS WITH REAL SPECTRA

We shall, in this section, develop perturbation bounds on the changes of generalized eigenvalues of a diagonalizable pencil with real spectra when the pencil is perturbed to another one which is also diagonalizable and has real spectra. These bounds always improve the existing ones in [5, 11], and the improvements are significant when the condition numbers of the eigenvector matrices are much bigger than 1.

Our goal will be accomplished with the help of Theorem 2.2 and Corollary 3.1. To use Theorem 2.2, we have to make selections for  $W$  and  $\tilde{W}$ . In §2, we have discussed four possibilities. There we excluded selection (3) from further consideration because it will yield weaker results than selection (4). We now argue that selection (2) is not so good as selection (4) either when the underlying matrix norm is a unitarily invariant norm other than the spectral norm  $\|\cdot\|_2$  for which the reappearance of each  $\sin \theta_i$  as singular values of  $\mathcal{P}_{Z^*} - \mathcal{P}_{\tilde{Z}^*}$  does not matter. It is precisely this reappearance that makes each chordal distance between generalized eigenvalues, properly matched, be listed twice in the inequalities in [5, Theorems 3.1, 3.4], see Remark 4.1 below. That does not seem to be natural. In what follows we shall see that with selection (4) this unnatural inclusion, among others, will disappear.

**Theorem 4.1.** *Let  $A - \lambda B, \tilde{A} - \lambda \tilde{B} \in \mathbb{C}^{n \times n}$  be diagonalizable, and admit the eigendecompositions (1.1) and (2.10), respectively. Assume that both pencils have only real spectra, i.e.,  $\alpha_j$ 's,  $\beta_j$ 's,  $\tilde{\alpha}_j$ 's, and  $\tilde{\beta}_j$ 's are all real. Then there exists a permutation  $\{j_1, j_2, \dots, j_n\}$  of  $\{1, 2, \dots, n\}$  such that*

$$(4.1) \quad \|\Delta\| \leq \frac{\pi}{2} \sqrt{\kappa(X)\kappa(\tilde{X})} \left\| \left\| \sin \Theta(Z^*, \tilde{Z}^*) \right\| \right\|,$$

where  $\Delta$  is as defined in (3.9). The factor  $\frac{\pi}{2}$  can be replaced by 1 in two special cases: the spectral norm and the Frobenius norm.

*Proof.* Without loss of generality we may assume  $\alpha_j^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1$  for all  $j$ , for otherwise we may rescale  $Y, \Lambda, \Omega$  and  $\tilde{Y}, \tilde{\Lambda}, \tilde{\Omega}$ . Write  $T = \tilde{X}^{-1}X$ . Taking  $W = S^*(ZZ^*)^{-1/2}$  and  $\tilde{W} = (\tilde{Z}\tilde{Z}^*)^{-1/2}$  with  $S$  being optimal as determined by Lemma 2.1, we have from (2.11) that

$$\left\| \tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda \right\| \leq \|\tilde{Y}^*(\tilde{Z}\tilde{Z}^*)^{1/2}\|_2 \left\| \left\| \sin \Theta(Z^*, \tilde{Z}^*) \right\| \right\| \|X\|_2.$$

Taking  $W = (ZZ^*)^{-1/2}$  and  $\tilde{W} = S(\tilde{Z}\tilde{Z}^*)^{-1/2}$  with  $S$  again optimal as determined by Lemma 2.1, we have from (2.12) that

$$\left\| \Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda} \right\| \leq \|Y^*(ZZ^*)^{1/2}\|_2 \left\| \left\| \sin \Theta(Z^*, \tilde{Z}^*) \right\| \right\| \|\tilde{X}\|_2.$$

Combine the above two inequalities, together with Lemma 2.2, to get

$$\left\| \tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda \right\| \left\| \Lambda T^{-1}\tilde{\Omega} - \Omega T^{-1}\tilde{\Lambda} \right\| \leq \kappa(X)\kappa(\tilde{X}) \left\| \left\| \sin \Theta(Z^*, \tilde{Z}^*) \right\| \right\|^2.$$

Now apply Corollary 3.1 to conclude the proof.  $\square$

*Remark 4.1.* Theorem 4.1 improves [5, Theorem 3.1]:

$$\left\| \begin{pmatrix} \Delta & \\ & \Delta \end{pmatrix} \right\| \leq \frac{\pi}{2} \hat{\kappa}(X)\hat{\kappa}(\tilde{X}) \|\mathcal{P}_{Z^*} - \mathcal{P}_{\tilde{Z}^*}\|$$

in two ways, where  $\widehat{\kappa}(X) \equiv \max\{\|X\|_2, \|X^{-1}\|_2\}^2$ . First, (4.1) is more elegant in that there is no repetition of each chordal difference of generalized eigenvalues, and second, (4.1) is sharper since

$$\sqrt{\kappa(X)\kappa(\widetilde{X})} \leq \kappa(X)\kappa(\widetilde{X}) \leq \widehat{\kappa}(X)\widehat{\kappa}(\widetilde{X})$$

and the improvement can be significant if  $\kappa(X), \kappa(\widetilde{X}) \gg 1$ . For the same reason, Theorem 4.1 improves

$$\|\Delta\|_2 \leq \kappa(X)\kappa(\widetilde{X})\|\sin \Theta(Z^*, \widetilde{Z}^*)\|_2 \quad ([11, \text{Theorem 3.1}] ),$$

$$\|\Delta\|_F \leq \widehat{\kappa}(X)\widehat{\kappa}(\widetilde{X})\|\sin \Theta(Z^*, \widetilde{Z}^*)\|_F \quad ([11, \text{Theorem 3.2}] ).$$

Roughly speaking, when the two pencils are sufficiently close, as one might expect for those bounds to be useful,  $\kappa(X) \approx \kappa(\widetilde{X})$  and then the new bound, and those new ones below, are proportional to  $\kappa(X)$ , while the old ones are proportional to  $\kappa(X)^2$ .

**Theorem 4.2.** *To the conditions of Theorem 4.1 add  $\alpha_j^2 + \beta_j^2 = \widetilde{\alpha}_j^2 + \widetilde{\beta}_j^2 = 1$  for all  $j$ . Then*

$$(4.2) \quad \|\Delta\| \leq \frac{\pi}{2} \sqrt{\|X\|_2 \|Y^*\|_2 \|\widetilde{X}\|_2 \|\widetilde{Y}^*\|_2} \left\| \|Z - \widetilde{Z}\| \right\|,$$

where  $\Delta$  is as defined in (3.9). The factor  $\frac{\pi}{2}$  can be replaced by 1 in two special cases: the spectral norm and the Frobenius norm.

*Proof.* A proof can be given along the same lines as the proof of Theorem 4.1 but with  $W = \widetilde{W} = I_n$ . □

*Remark 4.2.* We now compare Theorem 4.2 to [5, Theorem 3.2]:

$$\|\Delta\| \leq \frac{\pi}{2} \kappa(X) \|\widetilde{X}\|_2 \|\widetilde{Y}^*\|_2 \left\| \|Z - \widetilde{Z}\| \right\|,$$

which holds under the same condition of Theorem 4.2. The ratio of the right hand side of this inequality to that of (4.2) is

$$\frac{\|X^{-1}\|_2 \sqrt{\|X\|_2 \|\widetilde{X}\|_2 \|\widetilde{Y}^*\|_2}}{\sqrt{\|Y^*\|_2}}.$$

In general, there is not much we can say about this ratio. Nevertheless we may argue that the cases for which  $\|\widetilde{X}\|_2 \approx \|X\|_2$  and  $\|\widetilde{Y}^*\|_2 \approx \|Y^*\|_2$  may be interesting, and if we agree on this, then for such cases the ratio is approximately equal to

$$\|X^{-1}\|_2 \|X\|_2 = \kappa(X),$$

which is always bigger than 1 and can be much bigger, and thus Theorem 4.2 does potentially improve [5, Theorem 3.2]. For the same reason, Theorem 4.2 can be regarded as improvements to [11, Theorem 3.3]:

$$\|\Delta\|_2 \leq \kappa(X) \|\widetilde{X}\|_2 \|\widetilde{Y}^*\|_2 \|Z - \widetilde{Z}\|_2,$$

$$\|\Delta\|_F \leq \kappa(X) \|\widetilde{X}\|_2 \|\widetilde{Y}^*\|_2 \|Z - \widetilde{Z}\|_F.$$

**Corollary 4.1.** *Assume the conditions of Theorem 4.1. Then*

$$(4.3) \quad \|\Delta\| \leq \frac{\pi}{2} \sqrt{\kappa(X)\kappa(\widetilde{X})} \|Z^\dagger\|_2 \|\widetilde{Z}^\dagger\|_2 \left\| \|Z - \widetilde{Z}\| \right\|,$$

where  $\Delta$  is as defined in (3.9). The factor  $\frac{\pi}{2}$  can be replaced by 1 in two special cases: the spectral norm and the Frobenius norm.

*Proof.* Since  $Y$  and  $\tilde{Y}$  do not appear in the inequality, we may assume, without loss of generality, that  $\alpha_j^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1$  for all  $j$ , for otherwise we may rescale  $Y, \Lambda, \Omega$  and  $\tilde{Y}, \tilde{\Lambda}, \tilde{\Omega}$ . We then have (4.2). Apply Lemma 2.3 to conclude the proof.  $\square$

*Remark 4.3.* The use of  $\sin \Theta(Z^*, \tilde{Z}^*)$  to measure errors in  $\tilde{A} - \lambda \tilde{B}$  has theoretical implications. For example, if  $\tilde{A} - \lambda \tilde{B} = R(A - \lambda B)$  for some nonsingular  $R$ , then  $\tilde{Z} = RZ$ , and thus  $Z^*$  and  $\tilde{Z}^*$  have the same column space and all canonical angles are zeros, which implies the underlying perturbation does not affect any of the eigenvalues, as it should. However the use of  $Z - \tilde{Z}$  won't yield such perfect error bounds for the example, but it is perhaps more convenient in actual applications since it explicitly measures the perturbation in a straightforward way. Finally, we point out that we can always bound  $\sin \Theta(Z^*, \tilde{Z}^*)$  by  $Z - \tilde{Z}$  from above:

$$(4.4) \quad \left\| \sin \Theta(Z^*, \tilde{Z}^*) \right\| \leq \min \{ \|(ZZ^*)^{-1/2}\|_2, \|(\tilde{Z}\tilde{Z}^*)^{-1/2}\|_2 \} \left\| Z - \tilde{Z} \right\|,$$

due to J.-G. Sun (see, e.g., [11, pp.250-251]), but not the other way around.

## 5. PERTURBATION BOUNDS—DEFINITE PENCILS

A matrix pencil  $A - \lambda B$  is said to be a *definite pencil* [6, 14, 15] if  $A$  and  $B$  are Hermitian and

$$\gamma(A, B) \stackrel{\text{def}}{=} \min_{\|x\|_2=1} \sqrt{(x^*Ax)^2 + (x^*Bx)^2} > 0.$$

It is well-known [14, 15] that if  $A - \lambda B \in \mathbb{C}^{n \times n}$  is definite, there is a nonsingular matrix  $X$  such that

$$(5.1) \quad X^*AX = \Lambda \equiv \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad X^*BX = \Omega \equiv \text{diag}(\beta_1, \beta_2, \dots, \beta_n).$$

So necessarily all  $\alpha_j$ 's and  $\beta_j$ 's are real.

**Lemma 5.1** (Elsner and Sun [7]). *Let  $A - \lambda B \in \mathbb{C}^{n \times n}$  be a definite pencil and admit the eigendecomposition (5.1). If  $\alpha_j^2 + \beta_j^2 = 1$  for all  $j$ , then*

$$\|X\|_2 \leq \frac{1}{\sqrt{\gamma(A, B)}}, \quad \|X^{-1}\|_2 \leq \frac{\|Z\|_2}{\sqrt{\gamma(A, B)}}.$$

Now let  $\tilde{A} - \lambda \tilde{B} \in \mathbb{C}^{n \times n}$  be another definite pencil that admits the following decomposition:

$$(5.2) \quad \tilde{X}^* \tilde{A} \tilde{X} = \tilde{\Lambda} \equiv \text{diag}(\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_n), \quad \tilde{X}^* \tilde{B} \tilde{X} = \tilde{\Omega} \equiv \text{diag}(\tilde{\beta}_1, \tilde{\beta}_2, \dots, \tilde{\beta}_n),$$

where  $\tilde{X}$  is nonsingular. Also necessarily all  $\tilde{\alpha}_j$ 's and  $\tilde{\beta}_j$ 's are real.

**Theorem 5.1.** *Let  $A - \lambda B, \tilde{A} - \lambda \tilde{B} \in \mathbb{C}^{n \times n}$  be definite pencils and admit the eigendecompositions (5.1) and (5.2), respectively. Then there exists a permutation*

$\{j_1, j_2, \dots, j_n\}$  of  $\{1, 2, \dots, n\}$  such that

$$(5.3) \quad \|\Delta\| \leq \frac{\pi}{2} \sqrt{\frac{\|Z\|_2 \|\tilde{Z}\|_2}{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})}} \|\sin \Theta(Z^*, \tilde{Z}^*)\|,$$

$$(5.4) \quad \|\Delta\| \leq \frac{\pi}{2} \frac{\|Z - \tilde{Z}\|}{\sqrt{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})}},$$

where  $\Delta$  is as defined in (3.9). The factor  $\frac{\pi}{2}$  can be replaced by 1 in two special cases: the spectral norm and the Frobenius norm.

*Proof.* Since  $X$  and  $\tilde{X}$  do not appear in the desired inequalities, we may assume, without loss of generality, that  $\alpha_j^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1$  for all  $j$ , for otherwise we may rescale  $X, \Lambda, \Omega$  and  $\tilde{X}, \tilde{\Lambda}, \tilde{\Omega}$ . Now inequality (5.3) is a consequence of Theorem 4.1 and Lemma 5.1, and inequality (5.4) follows from Theorem 4.2 and Lemma 5.1.  $\square$

It has been proved that [11, pp.250-251]

$$\|(ZZ^*)^{-1/2}\|_2 \leq [\gamma(A, B)]^{-1}, \quad \|Z\|_2 \geq \gamma(A, B).$$

Together with (4.4), we can see that the bounds in Theorem 5.1 improve all their earlier versions:<sup>3</sup>

1. [11, Theorem 4.1].

$$(5.5) \quad \begin{aligned} \|\Delta\|_2 &\leq \frac{\|Z\|_2 \|\tilde{Z}\|_2}{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})} \|\sin \Theta(Z^*, \tilde{Z}^*)\|_2, \\ \|\Delta\|_2 &\leq \frac{\min\{\|Z\|_2, \|\tilde{Z}\|_2\}}{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})} \|Z - \tilde{Z}\|_2. \end{aligned}$$

2. [11, Theorem 4.2].

$$(5.6) \quad \begin{aligned} \|\Delta\|_F &\leq \frac{\max\{1, \|Z\|_2^2\} \max\{1, \|\tilde{Z}\|_2^2\}}{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})} \|\sin \Theta(Z^*, \tilde{Z}^*)\|_F, \\ \|\Delta\|_F &\leq \frac{\min\{\|Z\|_2, \|\tilde{Z}\|_2\}}{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})} \|Z - \tilde{Z}\|_F. \end{aligned}$$

3. [5, Theorems 3.4 and 3.5].

$$(5.7) \quad \begin{aligned} \left\| \begin{pmatrix} \Delta & \\ & \Delta \end{pmatrix} \right\| &\leq \frac{\pi \max\{1, \|Z\|_2^2\} \max\{1, \|\tilde{Z}\|_2^2\}}{2 \gamma(A, B)\gamma(\tilde{A}, \tilde{B})} \|\mathcal{P}_{Z^*} - \mathcal{P}_{\tilde{Z}^*}\|, \\ \|\Delta\| &\leq \frac{\pi \min\{\|Z\|_2, \|\tilde{Z}\|_2\}}{2 \gamma(A, B)\gamma(\tilde{A}, \tilde{B})} \|Z - \tilde{Z}\|. \end{aligned}$$

<sup>3</sup>Rigorously speaking, we cannot say that (5.4) always improves each of (5.5), (5.6), and (5.7) because we only have  $\|Z\|_2 \geq \gamma(A, B)$  and  $\|\tilde{Z}\|_2 \geq \gamma(\tilde{A}, \tilde{B})$  but not  $\min\{\|Z\|_2, \|\tilde{Z}\|_2\} \geq \sqrt{\gamma(A, B)\gamma(\tilde{A}, \tilde{B})}$  always. We argue that when  $\|Z\|_2 \gg \|\tilde{Z}\|_2$  or the other way around, none of the four inequalities would yield meaningful bounds. (Notice  $\sin \theta_i \leq 1$  always.) For this reason the interesting case may be those for which  $\|Z\|_2 \approx \|\tilde{Z}\|_2$ , and then (5.4) will be the sharpest.

Inequality (5.4) for the spectral norm (and thus with  $\frac{\pi}{2}$  replaced by 1) is comparable to a theorem of Stewart [14]: If  $\|\tilde{Z} - Z\|_2 < \gamma(A, B)$ , then

$$\|\Delta\|_2 \leq \|Z - \tilde{Z}\|_2 / \gamma(A, B).$$

(This is a slightly improved version, due to [10], of Stewart's.) This is because when the two definite pencils are close enough, it is safe to assume  $\gamma(A, B) \approx \gamma(\tilde{A}, \tilde{B})$ . Another bound for the same purpose is due to Sun [16]:

$$\|\Delta\|_2 \leq \max_{\|x\|_2=1} \rho((x^*Ax, x^*Bx), (x^*\tilde{A}x, x^*\tilde{B}x)).$$

This bound does not lend itself to an immediate comparison, and it is not clear how to actually compute the bound. However it implies Stewart's bound above.

## 6. CONCLUSIONS

We have presented a unifying framework for perturbation equations which may be of independent interest. Although we mainly use them in this paper under the real spectrum assumption, these perturbation equations hold for all diagonalizable pencils regardless of the nature of their spectra (and for an even wider range of pencils, see Remark 2.1). The existing perturbation equations which were the cornerstones of the earlier paper [11, 5] fall into the framework, and what's more the framework offers room to pick more suitable equations as needed for better bounds. With the help of the unifying framework and the recent result [4] on generalized commutators, not only have we obtained perturbation bounds in all unitarily invariant norms by a more concise argument than before, but also these new bounds represent significant improvements over the previous ones.

Our new bounds are established for diagonalizable pencils with only real eigenvalues. It is natural at this point to ask what kind of uniform bounds may hold when the pencils have complex eigenvalues. In this regard, we are able to establish Theorem 6.1 below, which provides a bound for the Frobenius norm, but unfortunately the bound is, roughly speaking, proportional to the square of the condition number of the eigenvector matrix, as happened to the old bounds, see Remark 4.1.

**Theorem 6.1.** *Let  $A - \lambda B, \tilde{A} - \lambda \tilde{B} \in \mathbb{C}^{n \times n}$  be diagonalizable, and admit the eigendecompositions (1.1) and (2.10), respectively. Then there exists a permutation  $\{j_1, j_2, \dots, j_n\}$  of  $\{1, 2, \dots, n\}$  such that*

$$(6.1) \quad \|\Delta\|_F \leq \kappa(X)\kappa(\tilde{X}) \|\sin \Theta(Z^*, \tilde{Z}^*)\|_F,$$

where  $\Delta$  is as defined in (3.9).

*Proof.* Without loss of generality we may assume  $\alpha_j^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1$  for all  $j$ . Set  $T = \tilde{X}^{-1}X$ . Taking  $W = S^*(ZZ^*)^{-1/2}$  and  $\tilde{W} = (\tilde{Z}\tilde{Z}^*)^{-1/2}$  with  $S$  being optimal as determined by Lemma 2.1, we have from (2.11) that

$$\|\tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda\|_F \leq \|\tilde{Y}^*(\tilde{Z}\tilde{Z}^*)^{1/2}\|_2 \|\sin \Theta(Z^*, \tilde{Z}^*)\|_F \|X\|_2.$$

Notice that  $\tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda = T \circ G$ , where  $G = (\tilde{\alpha}_i\beta_j - \tilde{\beta}_i\alpha_j) \in \mathbb{C}^{n \times n}$  and  $T \circ G$  is the Hadamard product. Now [12, Proposition 3.1] implies that

$$(6.2) \quad \|T^{-1}\|_2 \|\tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda\|_F \equiv \|T^{-1}\|_2 \|T \circ G\|_F \geq \min_{P \text{ permutation}} \|P \circ G\|_F.$$

Combine the above two inequalities with Lemma 2.2 to get (6.1).  $\square$

This proof does not seem to go well with any other unitarily invariant norm. Part of the conjecture in [12] was proposed precisely for the possibility of extending the inequality in (6.2) when the Frobenius norm is replaced by a general unitarily invariant norm; namely, it was asked whether there is a universal constant  $c$ , independent of the dimension  $n$  and the matrices, such that

$$c\|T^{-1}\|_2\|T \circ G\|_F \geq \min_{P \text{ permutation}} \|P \circ G\|_F?$$

Recent work by Romeo and Tilli [13, 18] in connection with the conjecture shows that this is in general not possible unless the unitarily invariant norm is equivalent to the Frobenius norm in the sense that  $c^{-1}\|\cdot\|_F \leq \|\cdot\| \leq c\|\cdot\|_F$ . This implies that we will have to invent a different approach to extend Theorem 6.1 to general unitarily invariant norms. Also whether the factor  $\kappa(X)\kappa(\tilde{X})$  in Theorem 6.1 could be replaced by its square root is not known.

We shall also point out that our unifying framework for perturbation equations will lead to various perturbation bounds on the changes of eigenspaces of a diagonalizable pencil  $A - \lambda B$  (even with complex spectra) in the Frobenius norm or any unitarily invariant norm, depending upon eigenvalue separation assumptions. However, such developments will be very much parallel to Part II of [11] in techniques. We omit details.

## REFERENCES

1. R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, vol. 169, Springer, New York, 1996. MR **98i**:15003
2. R. Bhatia and C. Davis, *A bound for the spectral variation of a unitary operator*, Linear and Multilinear Algebra **15** (1984), 71–76. MR **85b**:15020
3. R. Bhatia, C. Davis, and A. McIntosh, *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra and Its Applications **52–53** (1983), 45–67. MR **85a**:47020
4. R. Bhatia, F. Kittaneh, and R.-C. Li, *Some inequalities for commutators and an application to spectral variation. II*, Linear and Multilinear Algebra **43** (1997), no. 1–3, 207–220. MR **99f**:47014
5. R. Bhatia and R.-C. Li, *On perturbations of matrix pencils with real spectra. II*, Mathematics of Computation **65** (1996), no. 214, 637–645. MR **96g**:15019
6. C. R. Crawford, *A stable generalized eigenvalue problem*, SIAM Journal on Numerical Analysis **13** (1976), 854–860. MR **55**:4628; MR **80c**:65085
7. L. Elsner and J.-G. Sun, *Perturbation theorems for the generalized eigenvalue problem*, Linear Algebra and Its Applications **48** (1982), 341–357. MR **84f**:15012
8. A. J. Hoffman and H. W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Mathematical Journal **20** (1953), 37–39. MR **14**:611b
9. R.-C. Li, *Norms of certain matrices with applications to variations of the spectra of matrices and matrix pencils*, Linear Algebra and its Application **182** (1993), 199–234. MR **94c**:15040
10. ———, *A perturbation bound for definite pencils*, Linear Algebra and Its Applications **179** (1993), 191–202. MR **94h**:15007
11. ———, *On perturbations of matrix pencils with real spectra*, Mathematics of Computation **62** (1994), 231–265. MR **94f**:15011
12. ———, *Spectral variations and Hadamard products: Some problems*, Linear Algebra and its Application **278** (1998), 317–326. MR **99j**:15035
13. M. Romeo and P. Tilli, *On some variational problems in the theory of unitarily invariant norms and Hadamard products*, Linear Algebra and Its Applications **326** (2001), 161–172. MR **2001m**:15066
14. G. W. Stewart, *Perturbation bounds for the definite generalized eigenvalue problem*, Linear Algebra and Its Application **23** (1979), 69–86. MR **80c**:15007
15. G. W. Stewart and J.-G. Sun, *Matrix perturbation theory*, Academic Press, Boston, 1990. MR **97a**:65017

16. J.-G. Sun, *A note on Stewart's theorem for definite matrix pairs*, *Linear Algebra and Its Applications* **48** (1982), 331–339. MR **84f**:15013
17. ———, *Perturbation analysis for the generalized eigenvalue problem and the generalized singular value problem*, *Matrix Pencils* (New York) (B. Kågström and A. Ruhe, eds.), *Lecture Notes in Math.*, vol. 973, Springer Verlag, 1983, pp. 221–244. MR **84c**:65009
18. P. Tilli, *On some problems involving invariant norms and Hadamard products*, *Structured Matrices in Mathematics, Computer Science, and Engineering. II* (Providence, RI) (Vadim Olshevsky, ed.), *Contemporary Mathematics*, American Mathematical Society, 2001, vol. 281, pp. 319–324.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF KENTUCKY, LEXINGTON, KY 40506

*E-mail address*: `rcli@ms.uky.edu`

*URL*: `http://www.ms.uky.edu/~rcli`