Relative Perturbation Theory: (II) Eigenspace and Singular Subspace Variations *

Ren-Cang Li
Mathematical Science Section
Oak Ridge National Laboratory
P.O. Box 2008, Bldg 6012
Oak Ridge, TN 37831-6367
(li@msr.epm.ornl.gov)

LAPACK working notes # 85 (first published July, 1994, revised January, 1996)

Abstract

The classical perturbation theory for Hermitian matrix eigenvalue and singular value problems provides bounds on invariant subspace variations that are proportional to the reciprocals of *absolute gaps* between subsets of spectra or subsets of singular values. These bounds may be bad news for invariant subspaces corresponding to clustered eigenvalues or clustered singular values of much smaller magnitudes than the norms of matrices under considerations when some of these clustered eigenvalues or clustered singular values are perfectly relatively distinguishable from the rest. In this paper, we consider how eigenspaces of a Hermitian matrix A change when it is perturbed to $\widetilde{A} = D^*AD$ and how singular values of a (nonsquare) matrix B change when it is perturbed to $\widetilde{B} = D_1^*BD_2$, where D, D_1 and D_2 are assumed to be close to identity matrices of suitable dimensions, or either D_1 or D_2 close to some unitary matrix. It is proved that under these kinds of perturbations, the change of invariant subspaces are proportional to the reciprocals of *relative gaps* between subsets of spectra or subsets of singular values. We have been able to extend well-known Davis-Kahan $\sin\theta$ theorems and Wedin $\sin\theta$ theorems. As applications, we obtained bounds for perturbations of graded matrices.

*This material is based in part upon work supported, during January, 1992–August, 1995, by Argonne National Laboratory under grant No. 20552402 and by the University of Tennessee through the Advanced Research Projects Agency under contract No. DAAL03-91-C-0047, by the National Science Foundation under grant No. ASC-9005933, and by the National Science Infrastructure grants No. CDA-8722788 and CDA-9401156, and supported, since August, 1995, by a Householder Fellowship in Scientific Computing at Oak Ridge National Laboratory, supported by the Applied Mathematical Sciences Research Program, Office of Energy Research, United States Department of Energy contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. Part of this work was done during summer of 1994 while the author was at Department of Mathematics, University of California at Berkeley.

1 Introduction

Let A and $\widetilde{A}$ be two $n \times n$ Hermitian matrices with eigendecompositions

$$A = U\Lambda U^* \equiv (U_1, U_2) \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} \begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix}, \tag{1.1}$$

$$\widetilde{A} = \widetilde{U}\widetilde{\Lambda}\widetilde{U}^* \equiv (\widetilde{U}_1, \widetilde{U}_2) \begin{pmatrix} \widetilde{\Lambda}_1 & \\ & \widetilde{\Lambda}_2 \end{pmatrix} \begin{pmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{pmatrix}, \tag{1.2}$$

where U, $\widetilde{U} \in \mathbf{U}_n$, U_1, $\widetilde{U}_1 \in \mathbf{C}^{n\times k}$ $(1 \le k < n)$ and

$$\Lambda_1 = \operatorname{diag}(\lambda_1, \cdots, \lambda_k), \qquad \Lambda_2 = \operatorname{diag}(\lambda_{k+1}, \cdots, \lambda_n), \tag{1.3}$$

$$\widetilde{\Lambda}_1 = \operatorname{diag}(\widetilde{\lambda}_1, \cdots, \widetilde{\lambda}_k), \qquad \widetilde{\Lambda}_2 = \operatorname{diag}(\widetilde{\lambda}_{k+1}, \cdots, \widetilde{\lambda}_n). \tag{1.4}$$

Suppose now that A and $\widetilde{A}$ are *close*. The question is: *How close are the eigenspaces spanned by U_i and $\widetilde{U}_i$?* This question has been well answered by four celebrated theorems so-called $\sin\theta$, $\tan\theta$, $\sin 2\theta$ and $\tan 2\theta$ due to Davis and Kahan [2, 1970] for arbitrary additive perturbations in the sense that the perturbations to A can be made arbitrary as long as $\widetilde{A} - A$ is kept small. It is proved that the changes of invariant subspaces are proportional to the reciprocals of *absolute gaps* between subsets of spectra. This paper, however, will address the same question but under multiplicative perturbations: *How close are the eigenspaces spanned by U_i and $\widetilde{U}_i$ under the assumption that $\widetilde{A} = D^*AD$ for some D close to I?* Our bounds suggest that the changes of invariant subspaces be proportional to the reciprocals of *relative gaps* between subsets of spectra. A similar question for singular value decompositions will be answered also. To be specific, we will deal with perturbations of the following kinds:

- **Eigenvalue problems:**
  1. A and $\widetilde{A} = D^*AD$ for the Hermitian case, where D is nonsingular and close to the identity matrix.
  2. $A = S^*HS$ and $\widetilde{A} = S^*\widetilde{H}S$ for the graded nonnegative Hermitian case, where it is assumed that H and $\widetilde{H}$ are nonsingular and often that S is a highly graded diagonal matrix (this assumption is not necessary to our theorems).
- **Singular value problems:**
  1. B and $\widetilde{B} = D_1^*BD_2$, where D_1 and D_2 are nonsingular and close to identity matrices or one of them close to an identity matrix and the other to some unitary matrix.
  2. $B = GS$ and $\widetilde{B} = \widetilde{G}S$ for the graded case, where it is assumed that G and $\widetilde{G}$ are nonsingular and often that S is a highly graded diagonal matrix (this assumption is not necessary to our theorems).

These perturbations cover component-wise relative perturbations to entries of symmetric tridiagonal matrices with zero diagonal [4, 9], entries of bidiagonal and biacyclic matrices [1, 3, 4], and perturbations in graded nonnegative Hermitian matrices [5, 12], in graded matrices of singular value problems [5, 12] and more [6]. Recently, Eisenstat and Ipsen [7, 1994] launched an attack towards the above mentioned perturbations except graded cases. We will give a brief comparison among their results and ours.

This paper is organized as follows. We briefly review Davis-Kahan $\sin\theta$ theorems for Hermitian matrices and their generalizations—Wedin $\sin\theta$ theorems for singular value decompositions in §3. We present in §4.1 our $\sin\theta$ theorems for eigenvalue problems for A and $\widetilde{A} = D^*AD$ and for graded nonnegative Hermitian matrices. Theorems for singular value problems for B and $\widetilde{B} = D_1^*BD_2$ and for graded matrices are given in §4.2. We discuss how to bound from below relative gaps, for example, between Λ_1 and $\widetilde{\Lambda}_2$ by relative gaps between Λ_1 and Λ_2 in §5. A word will be said in §6 regarding Eisenstat-Ipsen's theorems in comparison with ours. Detailed proofs are postponed to §§7, 8, 9, and 10. Finally in §11, we present our conclusions and outline further possible extensions to diagonalizable matrices.

2 Preliminaries

Throughout this paper, we will be following notation we used in the first part of this series Li [11]. Most frequently used are the two kinds of relative distances: ϱ_p and χ defined for $\alpha,\ \widetilde{\alpha} \in \mathbf{C}$ by

$$\varrho_p(\alpha, \widetilde{\alpha}) = \frac{|\alpha - \widetilde{\alpha}|}{\sqrt[p]{|\alpha|^p + |\widetilde{\alpha}|^p}} \quad \text{for } 1 \le p \le \infty, \text{ and} \quad \chi\alpha, \widetilde{\alpha}) = \frac{|\alpha - \widetilde{\alpha}|}{\sqrt{|\alpha\widetilde{\alpha}|}},$$

with convention $0/0 = 0$ for convenience.

Lemma 2.1 (Li) *1. For $\mu,\ \nu \in \mathbf{C}$, $\varrho_p(\mu, \nu) \le 2^{-1/p}\chi(\mu, \nu)$.*

2. For $\mu,\ \nu \in \mathbf{R}$ and $\mu\nu \ge 0$, $\varrho_p(\mu, \nu) \le \varrho_p(\mu^2, \nu^2)$ and $2\chi(\mu, \nu) \le \chi(\mu^2, \nu^2)$.

3. ϱ_p is a metric on $\mathbf{R}$.

4. For $\mu,\ \nu,\ \omega \ge 0$, $\chi(\mu, \nu) \le \chi(\mu, \omega) + \chi(\omega, \nu) + \frac{1}{8}\chi(\mu, \nu)\chi(\mu, \omega)\chi(\omega, \nu)$.

Since this paper concerns with the variations of subspaces, we need some metrics to measure the difference between two subspaces. In this, we follow Davis and Kahan [2, 1970], Stewart and Sun [14]. Let $X,\ \widetilde{X} \in \mathbf{C}^{n\times k}$ $(n > k)$ have full column rank k, and define the angle matrix $\Theta(X, \widetilde{X})$ between X and $\widetilde{X}$ as

$$\Theta(X, \widetilde{X}) \stackrel{\text{def}}{=} \arccos((X^*X)^{-\frac{1}{2}}X^*\widetilde{X}(\widetilde{X}^*\widetilde{X})^{-1}\widetilde{X}^*X(X^*X)^{-\frac{1}{2}})^{-\frac{1}{2}}.$$

The *canonical angles* between the subspace $\mathcal{X} = \mathcal{R}(X)$ and $\widetilde{\mathcal{X}} = \mathcal{R}(\widetilde{X})$ are defined to be the singular values of the Hermitian matrix $\Theta(X, \widetilde{X})$, where $\mathcal{R}(X)$ denotes the subspace spanned by the column vectors of X. The following lemma is well-known. For a proof of it, the reader is referred to, e.g, Li [10, Lemma 2.1].

Lemma 2.2 *Suppose that $(\widetilde{X}, \widetilde{X}_1) \in \mathbf{C}^{n\times n}$ is a nonsingular matrix, and*

$$(\widetilde{X}, \widetilde{X}_1)^{-1} = \begin{pmatrix} \widetilde{Y}^* \\ \widetilde{Y}_1^* \end{pmatrix}, \quad \widetilde{Y} \in \mathbf{C}^{n\times k}.$$

Then for any unitarily invariant norm $|||\cdot|||$

$$\left|\left|\left|\sin\Theta(X, \widetilde{X})\right|\right|\right| = \left|\left|\left|(\widetilde{Y}_1^*\widetilde{Y}_1)^{-1/2}\widetilde{Y}_1^* X (X^*X)^{-1/2}\right|\right|\right|.$$

In this lemma, as well as many other places in the rest of this paper, we talk about the "same" unitarily invariant norm $|||\cdot|||$ that applies to matrices of different dimensions at the same time. Such applications of a unitarily invariant norm are understood in the following sense: First there is a unitarily invariant norm $|||\cdot|||$ on $\mathbf{C}^{M\times N}$ for sufficiently large integers M and N; Then for a matrix $X \in \mathbf{C}^{m\times n}$ ($m \le M$ and $N \le n$), $|||X|||$ is defined by appending X with zero blocks to make it $M \times N$ and then taking the unitarily invariant norm of the enlarged matrix.

Taking $X = U_1$ and $\widetilde{X} = \widetilde{U}_1$ (see (1.1) and (1.2)), with Lemma 2.2 one has

$$\Theta(U_1, \widetilde{U}_1) = \arccos(U_1^*\widetilde{U}_1\widetilde{U}_1^*U_1)^{-1/2} \quad \text{and} \quad \left|\left|\left|\sin\Theta(U_1, \widetilde{U}_1)\right|\right|\right| = \left|\left|\left|\widetilde{U}_2^*U_1\right|\right|\right|. \tag{2.5}$$

For more discussions on angles between subspaces, the reader is referred to Davis and Kahan [2] and Stewart and Sun [14, Chapters I and II].

3 Davis-Kahan $\sin\theta$ Theorems and Wedin $\sin\theta$ Theorems

3.1 Eigenspace Variations

Let A and $\widetilde{A}$ be two Hermitian matrices whose eigendecompositions are given by (1.1) and (1.2):

$$A = U\Lambda U^* = (U_1, U_2)\begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix}\begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix}, \tag{1.1}$$

$$\widetilde{A} = \widetilde{U}\widetilde{\Lambda}\widetilde{U}^* = (\widetilde{U}_1, \widetilde{U}_2)\begin{pmatrix} \widetilde{\Lambda}_1 & \\ & \widetilde{\Lambda}_2 \end{pmatrix}\begin{pmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{pmatrix}, \tag{1.2}$$

where U, $\widetilde{U} \in \mathbf{U}_n$, U_1, $\widetilde{U}_1 \in \mathbf{C}^{n\times k}$ ($1 \le k < n$) and Λ_i's and $\widetilde{\Lambda}_j$'s are defined as in (1.3) and (1.4). Define

$$R \stackrel{\text{def}}{=} \widetilde{A}U_1 - U_1\Lambda_1 = (\widetilde{A} - A)U_1. \tag{3.1}$$

The following two theorems are the matrix versions of two $\sin\theta$ theorems due to Davis and Kahan [2, 1970].

Theorem 3.1 (Davis-Kahan) *Let A and $\widetilde{A}$ be two Hermitian matrices with eigendecompositions (1.1), (1.2), (1.3), and (1.4). If $\delta \stackrel{\text{def}}{=} \min_{1\le i\le k, 1\le j\le n-k} |\lambda_i - \widetilde{\lambda}_{k+j}| > 0$, then*

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} \le \frac{\|R\|_{\mathrm{F}}}{\delta} = \frac{\|(\widetilde{A} - A)U_1\|_{\mathrm{F}}}{\delta}. \tag{3.2}$$

In this theorem, the spectrum of Λ_1 and that of $\widetilde{\Lambda}_2$ are only required to be disjoint. In the next theorem, they are required, more strongly, to be well-separated by intervals.

Theorem 3.2 (Davis-Kahan) *Let A and $\widetilde{A}$ be two Hermitian matrices with eigendecompositions (1.1), (1.2), (1.3), and (1.4). Assume there is an interval $[\alpha,\beta]$ and a $\underline{\delta}>0$ such that the spectrum of Λ_1 lies entirely in $[\alpha,\beta]$ while that of $\widetilde{\Lambda}_2$ lies entirely outside of $(\alpha-\underline{\delta},\beta+\underline{\delta})$ (or such that the spectrum of $\widetilde{\Lambda}_2$ lies entirely in $[\alpha,\beta]$ while that of Λ_1 lies entirely outside of $(\alpha-\underline{\delta},\beta+\underline{\delta})$). Then for any unitarily invariant norm $|||\cdot|||$*

$$\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right| \le \frac{|||R|||}{\underline{\delta}} = \frac{\left|\left|\left|(\widetilde{A}-A)U_1\right|\right|\right|}{\underline{\delta}}. \tag{3.3}$$

3.2 Singular Space Variations

Now, turn to the perturbations of Singular Value Decompositions (SVD). Let B and $\widetilde{B}$ be two $m\times n$ $(m\ge n)$ complex matrices with SVDs

$$B = U\Sigma V^* \equiv (U_1,U_2)\begin{pmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\\ 0 & 0\end{pmatrix}\begin{pmatrix}V_1^*\\ V_2^*\end{pmatrix}, \tag{3.4}$$

$$\widetilde{B} = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^* \equiv (\widetilde{U}_1,\widetilde{U}_2)\begin{pmatrix}\widetilde{\Sigma}_1 & 0\\ 0 & \widetilde{\Sigma}_2\\ 0 & 0\end{pmatrix}\begin{pmatrix}\widetilde{V}_1^*\\ \widetilde{V}_2^*\end{pmatrix}, \tag{3.5}$$

where $U,\widetilde{U}\in\mathbf{U}_m$, $V,\widetilde{V}\in\mathbf{U}_n$, $U_1,\widetilde{U}_1\in\mathbf{C}^{m\times k}$, $V_1,\widetilde{V}_1\in\mathbf{C}^{n\times k}$ $(1\le k<n)$ and

$$\Sigma_1 = \operatorname{diag}(\sigma_1,\cdots,\sigma_k), \quad \Sigma_2 = \operatorname{diag}(\sigma_{k+1},\cdots,\sigma_n), \tag{3.6}$$

$$\widetilde{\Sigma}_1 = \operatorname{diag}(\widetilde{\sigma}_1,\cdots,\widetilde{\sigma}_k), \quad \widetilde{\Sigma}_2 = \operatorname{diag}(\widetilde{\sigma}_{k+1},\cdots,\widetilde{\sigma}_n). \tag{3.7}$$

Define residuals

$$R_{\rm R} \stackrel{\rm def}{=} \widetilde{B}V_1 - U_1\Sigma_1 = (\widetilde{B}-B)V_1 \quad\text{and}\quad R_{\rm L} \stackrel{\rm def}{=} \widetilde{B}^*U_1 - V_1\Sigma_1 = (\widetilde{B}^*-B^*)U_1. \tag{3.8}$$

The following two theorems are due to Wedin [15, 1972].

Theorem 3.3 (Wedin) *Let B and $\widetilde{B}$ be two $m\times n$ $(m\ge n)$ complex matrices with SVDs (3.4), (3.5), (3.6), and (3.7). If $\delta \stackrel{\rm def}{=} \min\left\{\min_{1\le i\le k,1\le j\le n-k}|\sigma_i-\widetilde{\sigma}_{k+j}|, \min_{1\le i\le k}\sigma_i\right\}>0$, then*

$$\begin{aligned}&\sqrt{\|\sin\Theta(U_1,\widetilde{U}_1)\|_{\rm F}^2+\|\sin\Theta(V_1,\widetilde{V}_1)\|_{\rm F}^2}\\ &\qquad\le \frac{\sqrt{\|R_{\rm R}\|_{\rm F}^2+\|R_{\rm L}\|_{\rm F}^2}}{\delta} = \frac{\sqrt{\|(\widetilde{B}-B)V_1\|_{\rm F}^2+\|(\widetilde{B}^*-B^*)U_1\|_{\rm F}^2}}{\delta}.\end{aligned} \tag{3.9}$$

Theorem 3.4 (Wedin) *Let B and $\widetilde{B}$ be two $m\times n$ $(m\geq n)$ complex matrices with SVDs (3.4), (3.5), (3.6), and (3.7). If there exist $\alpha>0$ and $\underline{\delta}>0$ such that*

$$\min_{1\leq i\leq k}\sigma_i\geq\alpha+\underline{\delta}\quad \textit{and}\quad \max_{1\leq j\leq n-k}\widetilde{\sigma}_{k+j}\leq\alpha,$$

then for any unitarily invariant norm $|||\cdot|||$

$$\max\left\{\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right|,\left|\left|\left|\sin\Theta(V_1,\widetilde{V}_1)\right|\right|\right|\right\} \tag{3.10}$$
$$\leq\frac{\max\{|||R_{\rm R}|||,|||R_{\rm L}|||\}}{\underline{\delta}}=\frac{\max\left\{\left|\left|\left|(\widetilde{B}-B)V_1\right|\right|\right|,\left|\left|\left|(\widetilde{B}^*-B^*)U_1\right|\right|\right|\right\}}{\underline{\delta}}.$$

4 Relative Perturbation Theorems

4.1 Eigenspace Variations

The following theorem is an extension of Theorem 3.1.

Theorem 4.1 *Let A and $\widetilde{A}=D^*AD$ be two $n\times n$ Hermitian matrices with eigendecompositions (1.1), (1.2), (1.3), and (1.4), where D is nonsingular and close to I. If $\eta_2\stackrel{\rm def}{=}\min_{1\leq i\leq k,\,1\leq j\leq n-k}\varrho_2(\lambda_i,\widetilde{\lambda}_{k+j})>0$, then*

$$\|\sin\Theta(U_1,\widetilde{U}_1)\|_{\rm F}\leq\frac{\sqrt{\|(I-D^{-1})U_1\|_{\rm F}^2+\|(I-D^*)U_1\|_{\rm F}^2}}{\eta_2}. \tag{4.1}$$

By imposing a stronger condition on the separation between the spectra of $\widetilde{\Lambda}_2$ and Λ_1, we have the following bound on any unitarily invariant norm of $\sin\Theta(U_1,\widetilde{U}_1)$.

Theorem 4.2 *Let A and $\widetilde{A}=D^*AD$ be two $n\times n$ Hermitian matrices with eigendecompositions (1.1), (1.2), (1.3), and (1.4), where D is nonsingular and close to I. Assume that there exist $\alpha>0$ and $\delta>0$ such that the spectrum of Λ_1 lies entirely in $[-\alpha,\alpha]$ while that of $\widetilde{\Lambda}_2$ lies entirely outside $(-\alpha-\delta,\alpha+\delta)$ (or such that the spectrum of Λ_1 lies entirely outside $(-\alpha-\delta,\alpha+\delta)$ while that of $\widetilde{\Lambda}_2$ lies entirely in $[-\alpha,\alpha]$). Then for any unitarily invariant norm $|||\cdot|||$*

$$\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right|\leq\frac{\sqrt[q]{|||(I-D^{-1})U_1|||^q+|||(I-D^*)U_1|||^q}}{\underline{\eta}_p}, \tag{4.2}$$

where $\underline{\eta}_p\stackrel{\rm def}{=}\varrho_p(\alpha,\alpha+\delta)$.

Now we consider eigenspace variations for a graded Hermitian matrix $A=S^*HS\in\mathbf{C}^{n\times n}$ perturbed to $\widetilde{A}=S^*\widetilde{H}S$. Set $\Delta H\stackrel{\rm def}{=}\widetilde{H}-H$.

Theorem 4.3 *Let $A = S^*HS$ and $\widetilde{A} = S^*\widetilde{H}S$ be two $n\times n$ Hermitian matrices with eigen-decompositions (1.1), (1.2), (1.3), and (1.4). H is positive definite and $\|H^{-1}\|_2\|\Delta H\|_2 < 1$. If $\eta_\chi \stackrel{\rm def}{=} \min_{1\le i\le k,\,1\le j\le n-k} \chi(\lambda_i, \widetilde{\lambda}_{k+j}) > 0$, then*

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_{\rm F} \le \frac{\|D - D^{-1}\|_{\rm F}}{\eta_\chi} \le \frac{\|H^{-1}\|_2}{\sqrt{1 - \|H^{-1}\|_2\|\Delta H\|_2}} \frac{\|\Delta H\|_{\rm F}}{\eta_\chi}. \tag{4.3}$$

where $D = (I + H^{-1/2}(\Delta H)H^{-1/2})^{1/2} = D^*$.

Theorem 4.4 *Let $A = S^*HS$ and $\widetilde{A} = S^*\widetilde{H}S$ be two $n\times n$ Hermitian matrices with eigen-decompositions (1.1), (1.2), (1.3), and (1.4). H is positive definite and $\|H^{-1}\|_2\|\Delta H\|_2 < 1$. Assume that there exist $\alpha > 0$ and $\delta > 0$ such that*

$$\max_{1\le i\le k} \lambda_i \le \alpha \quad \text{and} \quad \min_{1\le j\le n-k} \widetilde{\lambda}_{k+j} \ge \alpha + \delta$$

or

$$\min_{1\le i\le k} \lambda_i \ge \alpha + \delta \quad \text{and} \quad \max_{1\le j\le n-k} \widetilde{\lambda}_{k+j} \le \alpha.$$

Then for any unitarily invariant norm $|||\cdot|||$

$$\left|\left|\left|\sin\Theta(U_1, \widetilde{U}_1)\right|\right|\right| \le \frac{|||D - D^{-1}|||}{\underline{\eta}_\chi} \le \frac{\|H^{-1}\|_2}{\sqrt{1 - \|H^{-1}\|_2\|\Delta H\|_2}} \frac{|||\Delta H|||}{\underline{\eta}_\chi}, \tag{4.4}$$

where $\underline{\eta}_\chi \stackrel{\rm def}{=} \chi(\alpha, \alpha + \delta)$ and $D = (I + H^{-1/2}(\Delta H)H^{-1/2})^{1/2} = D^*$.

4.2 Singular Space Variations

The following two theorems concern singular space perturbations.

Theorem 4.5 *Let B and $\widetilde{B} = D_1^*BD_2$ be two $m\times n$ $(m \ge n)$ (complex) matrices with SVDs (3.4), (3.5), (3.6), and (3.7), where D_1 and D_2 are nonsingular and close to identities. Let*

$$\eta_2 \stackrel{\rm def}{=} \begin{cases} \min\left\{\min_{1\le i\le k, 1\le j\le n-k} \varrho_2(\sigma_i, \widetilde{\sigma}_{k+j}),\ \min_{1\le i\le k} \varrho_2(\sigma_i, 0)\right\}, & \text{if } m > n, \\ \min\left\{\min_{1\le i\le k, 1\le j\le n-k} \varrho_2(\sigma_i, \widetilde{\sigma}_{k+j})\right\}, & \text{otherwise.} \end{cases} \tag{4.5}$$

If $\eta_2 > 0$, then

$$\sqrt{\|\sin\Theta(U_1, \widetilde{U}_1)\|_{\rm F}^2 + \|\sin\Theta(V_1, \widetilde{V}_1)\|_{\rm F}^2} \tag{4.6}$$
$$\le \frac{\sqrt{\|(I - D_1^*)U_1\|_{\rm F}^2 + \|(I - D_1^{-1})U_1\|_{\rm F}^2 + \|(I - D_2^*)V_1\|_{\rm F}^2 + \|(I - D_2^{-1})V_1\|_{\rm F}^2}}{\eta_2}.$$

Theorem 4.6 *Let B and $\widetilde{B} = D_1^* B D_2$ be two $m \times n$ $(m \geq n)$ (complex) matrices with SVDs (3.4), (3.5), (3.6), and (3.7), where D_1 and D_2 are nonsingular and close to identities. If there exist $\alpha > 0$ and $\delta > 0$ such that*

$$\min_{1 \leq i \leq k} \sigma_i \geq \alpha + \delta \quad \text{and} \quad \max_{1 \leq j \leq n-k} \widetilde{\sigma}_{k+j} \leq \alpha,$$

then for any unitarily invariant norm $|||\cdot|||$

$$\max\left\{ \left|\left|\left| \sin\Theta(U_1, \widetilde{U}_1) \right|\right|\right|, \left|\left|\left| \sin\Theta(V_1, \widetilde{V}_1) \right|\right|\right| \right\} \leq \frac{1}{\underline{\eta}_p} \max\left\{ \sqrt[q]{|||(I - D_2^{-1})V_1|||^q + |||(D_1^* - I)U_1|||^q}, \sqrt[q]{|||(I - D_1^{-1})U_1|||^q + |||(D_2^* - I)V_1|||^q} \right\}, \tag{4.7}$$

$$\left|\left|\left| \begin{pmatrix} \sin\Theta(U_1, \widetilde{U}_1) & \\ & \sin\Theta(V_1, \widetilde{V}_1) \end{pmatrix} \right|\right|\right| \leq \frac{\sqrt[q]{\left|\left|\left| \begin{pmatrix} (I - D_1^{-1})U_1 & \\ & (I - D_2^{-1})V_1 \end{pmatrix} \right|\right|\right|^q + \left|\left|\left| \begin{pmatrix} (I - D_1^*)U_1 & \\ & (I - D_2^*)V_1 \end{pmatrix} \right|\right|\right|^q}}{\underline{\eta}_p}, \tag{4.8}$$

where $\underline{\eta}_p \stackrel{\text{def}}{=} \varrho_p(\alpha, \alpha + \delta)$.

In Theorems 4.5 and 4.6, we assumed that both D_1 and D_2 are close to identity matrices. But, intuitively D_2 should not affect U_1 much as long as it is close to a unitary matrix. In fact, if $D_2 \in \mathbf{U}_n$, it does not affect U_1 at all. The following theorems indeed confirm this observation.

Theorem 4.7 *Let B and $\widetilde{B} = D_1^* B D_2$ be two $m \times n$ $(m \geq n)$ (complex) matrices with SVDs (3.4), (3.5), (3.6), and (3.7), where D_1 and D_2 are close to some unitary matrices. Let η_2 be defined by (4.5) and set*

$$\eta_\chi \stackrel{\text{def}}{=} \begin{cases} \min\left\{ \min_{1 \leq i \leq k, 1 \leq j \leq n-k} \chi(\sigma_i, \widetilde{\sigma}_{k+j}), \min_{1 \leq i \leq k} \chi(\sigma_i, 0) \right\}, & \text{if } m > n, \\ \min\left\{ \min_{1 \leq i \leq k, 1 \leq j \leq n-k} \chi(\sigma_i, \widetilde{\sigma}_{k+j}) \right\}, & \text{otherwise.} \end{cases} \tag{4.9}$$

Assume[1]

$$\eta_2 > \frac{1}{2\sqrt{2}} \max\{\|D_1^* - D_1^{-1}\|_2, \|D_2^* - D_2^{-1}\|_2\}. \tag{4.10}$$

If D_1 is close to identity, then

$$\left\| \sin\Theta(U_1, \widetilde{U}_1) \right\|_{\mathrm{F}} \leq \frac{\sqrt{\|(I - D_1^{-1})U_1\|_{\mathrm{F}}^2 + \|(I - D_1^*)U_1\|_{\mathrm{F}}^2}}{\eta_2 - 2^{-3/2}\epsilon_2} + \left(1 + \frac{\epsilon_1}{16}\eta_\chi\right) \frac{\|D_2^* - D_2^{-1}\|_{\mathrm{F}}}{2\eta_\chi - \epsilon_1}, \tag{4.11}$$

$$\left\| \sin\Theta(U_1, \widetilde{U}_1) \right\|_{\mathrm{F}} \leq \frac{\sqrt{\|(I - D_1^{-1})U_1\|_{\mathrm{F}}^2 + \|(I - D_1^*)U_1\|_{\mathrm{F}}^2}}{\eta_2 - 2^{-3/2}\epsilon_2} + \frac{\|D_2^* - D_2^{-1}\|_{\mathrm{F}}}{2^{3/2}\eta_2 - \epsilon_1}; \tag{4.12}$$

[1] This implies, by Lemma 2.1, $\eta_\chi > \frac{1}{2} \max\{\|D_1^* - D_1^{-1}\|_2, \|D_2^* - D_2^{-1}\|_2\}$.

If D_2 is close to identity, then

$$\left\|\sin\Theta(V_1,\widetilde{V}_1)\right\|_{\rm F} \le \left(1+\frac{\epsilon_2}{16}\eta_\chi\right)\frac{\|D_1^*-D_1^{-1}\|_{\rm F}}{2\eta_\chi-\epsilon_2}+\frac{\sqrt{\|(I-D_2^{-1})V_1\|_{\rm F}^2+\|(I-D_2^*)V_1\|_{\rm F}^2}}{\eta_2-2^{-3/2}\epsilon_1}, \tag{4.13}$$

$$\left\|\sin\Theta(V_1,\widetilde{V}_1)\right\|_{\rm F} \le \frac{\|D_1^*-D_1^{-1}\|_{\rm F}}{2^{3/2}\eta_2-\epsilon_2}+\frac{\sqrt{\|(I-D_2^{-1})V_1\|_{\rm F}^2+\|(I-D_2^*)V_1\|_{\rm F}^2}}{\eta_2-2^{-3/2}\epsilon_1}, \tag{4.14}$$

where $\epsilon_1=\|D_1^*-D_1^{-1}\|_2$ *and* $\epsilon_2=\|D_2^*-D_2^{-1}\|_2$.

Inequalities (4.11) and (4.12) which differ slightly in their last term clearly says that D_2 contributes to $\sin\Theta(U_1,\widetilde{U}_1)$ with its departure from *some* unitary matrix, and similarly do (4.13) and (4.14).

Remark. When one of the D_1 and D_2 is I, assumption (4.10) can actually be weakened to $\eta_2>0$, as shall be seen from our proofs.

Theorem 4.8 *Let B and $\widetilde{B}=D_1^*BD_2$ be two $m\times n$ $(m\ge n)$ (complex) matrices with SVDs (3.4), (3.5), (3.6), and (3.7), where D_1 and D_2 are close to some unitary matrices. Suppose that there exist $\alpha>0$ and $\delta>0$ such that*

$$\min_{1\le i\le k}\sigma_i\ge\alpha+\delta \quad \text{and} \quad \max_{1\le j\le n-k}\widetilde{\sigma}_{k+j}\le\alpha.$$

Assume[2]

$$\varrho_p(\alpha,\alpha+\delta)>\frac{1}{2^{1+1/p}}\max\{\|D_1^*-D_1^{-1}\|_2,\,\|D_2^*-D_2^{-1}\|_2\}. \tag{4.15}$$

If D_1 is close to identity, then for any unitarily invariant norm $|||\cdot|||$

$$\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right| \le \frac{\sqrt[q]{\left|\left|\left|(I-D_1^{-1})U_1\right|\right|\right|^q+|||(I-D_1^*)U_1|||^q}}{\underline{\eta}_p-2^{-1-1/p}\epsilon_2}+\left(1+\frac{\epsilon_1}{16}\underline{\eta}_\chi\right)\frac{\left|\left|\left|D_2^*-D_2^{-1}\right|\right|\right|}{2\underline{\eta}_\chi-\epsilon_1}, \tag{4.16}$$

$$\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right| \le \frac{\sqrt[q]{\left|\left|\left|(I-D_1^{-1})U_1\right|\right|\right|^q+|||(I-D_1^*)U_1|||^q}}{\underline{\eta}_p-2^{-1-1/p}\epsilon_2}+\frac{\left|\left|\left|D_2^*-D_2^{-1}\right|\right|\right|}{2^{1+1/p}\underline{\eta}_2-\epsilon_1}; \tag{4.17}$$

If D_2 is close to identity, then

$$\left|\left|\left|\sin\Theta(V_1,\widetilde{V}_1)\right|\right|\right| \le \left(1+\frac{\epsilon_2}{16}\underline{\eta}_\chi\right)\frac{\left|\left|\left|D_1^*-D_1^{-1}\right|\right|\right|}{2\underline{\eta}_\chi-\epsilon_2}+\frac{\sqrt[q]{\left|\left|\left|(I-D_2^{-1})V_1\right|\right|\right|^q+|||(I-D_2^*)V_1|||^q}}{\underline{\eta}_p-2^{-1-1/p}\epsilon_1}, \tag{4.18}$$

$$\left|\left|\left|\sin\Theta(V_1,\widetilde{V}_1)\right|\right|\right| \le \frac{\left|\left|\left|D_1^*-D_1^{-1}\right|\right|\right|}{2^{1+1/p}\underline{\eta}_p-\epsilon_2}+\frac{\sqrt[q]{\left|\left|\left|(I-D_2^{-1})V_1\right|\right|\right|^q+|||(I-D_2^*)V_1|||^q}}{\underline{\eta}_p-2^{-1-1/p}\epsilon_1}. \tag{4.19}$$

---

[2] This implies, by Lemma 2.1, $\underline{\eta}_\chi>\frac{1}{2}\max\{\|D_1^*-D_1^{-1}\|_2,\,\|D_2^*-D_2^{-1}\|_2\}$.

where $\underline{\eta}_p \stackrel{\rm def}{=} \varrho_p(\alpha, \alpha+\delta)$, $\underline{\eta}_\chi \stackrel{\rm def}{=} \chi(\alpha, \alpha+\delta)$, *and* $\epsilon_1 = \|D_1^* - D_1^{-1}\|_2$ *and* $\epsilon_2 = \|D_2^* - D_2^{-1}\|_2$.

Remark. If either D_1 or D_2 is I, (4.15) can actually be removed.

Now, Let's briefly mention a possible application of Theorems 4.5, 4.6, 4.7 and 4.8. It has something to do with *deflation* in computing the singular value systems of a bidiagonal matrix. Taking account of the remarks we have made, we get

Corollary 4.1 *Assume* $D_1 = I$ *and* D_2 *takes the form*

$$D_2 = \begin{pmatrix} I & X \\ & I \end{pmatrix},$$

where X *is a matrix of suitable dimensions. Let* η_2 *be defined by (4.5), and* η_χ *by (4.9). If* $\eta_2 > 0$, *then*

$$\left\|\sin\Theta(U_1, \widetilde{U}_1)\right\|_{\rm F} \leq \frac{\|X\|_{\rm F}}{\sqrt{2}\,\eta_\chi}, \tag{4.20}$$

$$\sqrt{\left\|\sin\Theta(U_1, \widetilde{U}_1)\right\|_{\rm F}^2 + \left\|\sin\Theta(V_1, \widetilde{V}_1)\right\|_{\rm F}^2} \leq \frac{\sqrt{2}\|X\|_{\rm F}}{\eta_2}. \tag{4.21}$$

Proof: Inequality (4.21) follows from (4.6). Inequality (4.20) follows from (4.11) and

$$D_2^* - D_2^{-1} = \begin{pmatrix} & X \\ X^* & \end{pmatrix} \Rightarrow \|D_2^* - D_2^{-1}\|_{\rm F} = \sqrt{2}\|X\|_{\rm F}.$$ ∎

Corollary 4.2 *Assume* $D_2 = I$ *and* D_1 *takes the form*

$$D_1 = \begin{pmatrix} I & X \\ & I \end{pmatrix},$$

where X *is a matrix of suitable dimensions. Let* η_2 *be defined by (4.5), and* η_χ *by (4.9). If* $\eta_2 > 0$, *then*

$$\left\|\sin\Theta(V_1, \widetilde{V}_1)\right\|_{\rm F} \leq \frac{\|X\|_{\rm F}}{\sqrt{2}\,\eta_\chi},$$

$$\sqrt{\left\|\sin\Theta(U_1, \widetilde{U}_1)\right\|_{\rm F}^2 + \left\|\sin\Theta(V_1, \widetilde{V}_1)\right\|_{\rm F}^2} \leq \frac{\sqrt{2}\|X\|_{\rm F}}{\eta_2}.$$

Corollary 4.3 *Assume* $D_1 = I$ *and* D_2 *takes the form*

$$D_2 = \begin{pmatrix} I & X \\ & I \end{pmatrix},$$

where X *is a matrix of suitable dimensions. Suppose that there exist* $\alpha > 0$ *and* $\delta > 0$ *such that*

$$\min_{1\leq i\leq k} \sigma_i \geq \alpha + \delta \quad \textit{and} \quad \max_{1\leq j\leq n-k} \widetilde{\sigma}_{k+j} \leq \alpha.$$

Then

$$\left\|\sin\Theta(U_1,\widetilde{U}_1)\right\| \le \frac{1}{2\underline{\eta}_\chi}\left\|\begin{pmatrix} & X \\ X^* & \end{pmatrix}\right\|, \quad \left\|\sin\Theta(V_1,\widetilde{V}_1)\right\| \le \frac{\|X\|}{\underline{\eta}_p}, \tag{4.22}$$

where $\underline{\eta}_p \overset{\rm def}{=} \varrho_p(\alpha,\alpha+\delta)$, $\underline{\eta}_\chi \overset{\rm def}{=} \chi(\alpha,\alpha+\delta)$.

Proof: The first inequality in (4.22) follows from (4.16), and the second from (4.7). ∎

Corollary 4.4 *Assume* $D_2 = I$ *and* D_1 *takes the form*

$$D_1 = \begin{pmatrix} I & X \\ & I \end{pmatrix},$$

where X *is a matrix of suitable dimensions. Suppose that there exist* $\alpha > 0$ *and* $\delta > 0$ *such that*

$$\min_{1\le i\le k}\sigma_i \ge \alpha+\delta \quad \textit{and} \quad \max_{1\le j\le n-k}\widetilde{\sigma}_{k+j} \le \alpha.$$

Then

$$\left\|\sin\Theta(U_1,\widetilde{U}_1)\right\| \le \frac{\|X\|}{\underline{\eta}_p}, \quad \left\|\sin\Theta(V_1,\widetilde{V}_1)\right\| \le \frac{1}{2\underline{\eta}_\chi}\left\|\begin{pmatrix} & X \\ X^* & \end{pmatrix}\right\|,$$

where $\underline{\eta}_p \overset{\rm def}{=} \varrho_p(\alpha,\alpha+\delta)$, $\underline{\eta}_\chi \overset{\rm def}{=} \chi(\alpha,\alpha+\delta)$.

Now, we consider singular space variations for a graded matrix $B = GS \in \mathbf{C}^{n\times n}$ perturbed to $\widetilde{B} = \widetilde{G}S \in \mathbf{C}^{n\times n}$, where G is nonsingular. Set $\Delta G \overset{\rm def}{=} \widetilde{G} - G$. If $\|(\Delta G)G^{-1}\|_2 < 1$, then $\widetilde{G} = G + \Delta G = [I + (\Delta G)G^{-1}]G$ is nonsingular also.

Theorem 4.9 *Let* $B = GS \in \mathbf{C}^{n\times n}$ *and* $\widetilde{B} = \widetilde{G}S \in \mathbf{C}^{n\times n}$ *with SVDs (3.4), (3.5), (3.6), and (3.7), where* G *is nonsingular. Assume* $\|(\Delta G)G^{-1}\|_2 < 1$. *If*

$$\eta_2 \overset{\rm def}{=} \min_{1\le i\le k, 1\le j\le n-k} \varrho_2(\sigma_i, \widetilde{\sigma}_{k+j}) > 0,$$

then

$$\begin{aligned} &\sqrt{\left\|\sin\Theta(U_1,\widetilde{U}_1)\right\|_{\rm F}^2 + \left\|\sin\Theta(V_1,\widetilde{V}_1)\right\|_{\rm F}^2} \\ &\quad \le \frac{\sqrt{\|(\Delta G)G^{-1}U_1\|_{\rm F}^2 + \|[I + G^{-*}(\Delta G)^*]^{-1}G^{-*}(\Delta G)^*U_1\|_{\rm F}^2}}{\eta_2} \\ &\quad \le \|G^{-1}\|_2\sqrt{1 + \frac{1}{(1-\|G^{-1}\|_2\|\Delta G\|_2)^2}}\,\frac{\|\Delta G\|_{\rm F}}{\eta_2}, \end{aligned} \tag{4.23}$$

and

$$\left\|\sin\Theta(V_1,\widetilde{V}_1)\right\|_{\rm F} \le \frac{\|I+(\Delta G)G^{-1}-(I+(\Delta G)G^{-1})^{-*}\|_{\rm F}}{2\,\eta_\chi} \tag{4.24}$$
$$\le \left(\frac{\|(\Delta G)G^{-1}+G^{-*}(\Delta G)^*\|_{\rm F}}{\|(\Delta G)G^{-1}\|_{\rm F}}+\frac{\|(\Delta G)G^{-1}\|_2}{1-\|(\Delta G)G^{-1}\|_2}\right)\frac{\|(\Delta G)G^{-1}\|_{\rm F}}{2\,\eta_\chi}$$
$$\le \left(1+\frac{1}{1-\|G^{-1}\|_2\|\Delta G\|_2}\right)\frac{\|G^{-1}\|_2\|\Delta G\|_{\rm F}}{2\,\eta_\chi},$$

where $\eta_\chi \stackrel{\rm def}{=} \min_{1\le i\le k, 1\le j\le n-k} \chi(\sigma_i,\widetilde{\sigma}_{k+j})$.

Proof: Write $\widetilde{B}=\widetilde{G}S=[I+(\Delta G)G^{-1}]GS=D_1^*BD_2$, where $D_1^*=I+(\Delta G)G^{-1}$ and $D_2=I$. Then apply Theorem 4.5 to get (4.23), and apply Theorem 4.7 to get (4.24). ∎

Theorem 4.10 *Let* $B=GS\in\mathbf{C}^{n\times n}$ *and* $\widetilde{B}=\widetilde{G}S\in\mathbf{C}^{n\times n}$ *with SVDs (3.4), (3.5), (3.6), and (3.7), where G is nonsingular. Assume* $\|(\Delta G)G^{-1}\|_2<1$. *If there exist* $\alpha>0$ *and* $\delta>0$ *such that*

$$\min_{1\le i\le k}\sigma_i\ge\alpha+\delta \quad \textit{and} \quad \max_{1\le j\le n-k}\widetilde{\sigma}_{k+j}\le\alpha$$

or, the other way around, i.e.,

$$\max_{1\le i\le k}\sigma_i\le\alpha \quad \textit{and} \quad \min_{1\le j\le n-k}\widetilde{\sigma}_{k+j}\ge\alpha+\delta,$$

then for any unitarily invariant norm $|||\cdot|||$

$$\max\left\{\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right|,\left|\left|\left|\sin\Theta(V_1,\widetilde{V}_1)\right|\right|\right|\right\} \tag{4.25}$$
$$\le \frac{\max\left\{|||(\Delta G)G^{-1}U_1|||,|||[I+G^{-*}(\Delta G)^*]^{-1}G^{-*}(\Delta G)^*U_1|||\right\}}{\eta_\infty}$$
$$\le \frac{\|G^{-1}\|_2}{1-\|G^{-1}\|_2\|\Delta G\|_2}\,\frac{|||\Delta G|||}{\eta_\infty},$$

$$\left|\left|\left|\begin{pmatrix}\sin\Theta(U_1,\widetilde{U}_1) & \\ & \sin\Theta(V_1,\widetilde{V}_1)\end{pmatrix}\right|\right|\right| \tag{4.26}$$
$$\le \frac{\sqrt[q]{|||(\Delta G)G^{-1}U_1|||^q+|||[I+G^{-*}(\Delta G)^*]^{-1}G^{-*}(\Delta G)^*U_1|||^q}}{\underline{\eta}_p}$$
$$\le \|G^{-1}\|_2\sqrt[q]{1+\frac{1}{(1-\|G^{-1}\|_2\|\Delta G\|_2)^q}}\;\frac{|||\Delta G|||}{\underline{\eta}_p},$$

and

$$\left|\left|\left|\sin\Theta(V_1,\widetilde{V}_1)\right|\right|\right| \le \frac{|||I+(\Delta G)G^{-1}-(I+(\Delta G)G^{-1})^{-1}|||}{2\,\underline{\eta}_\chi} \tag{4.27}$$

$$\leq \left(\frac{|||(\Delta G)G^{-1} + G^{-*}(\Delta G)^*|||}{|||(\Delta G)G^{-1}|||} + \frac{\|(\Delta G)G^{-1}\|_2}{1 - \|(\Delta G)G^{-1}\|_2} \right) \frac{|||(\Delta G)G^{-1}|||}{2\,\underline{\eta}_\chi}$$

$$\leq \left(1 + \frac{1}{1 - \|G^{-1}\|_2 \|\Delta G\|_2} \right) \frac{\|G^{-1}\|_2\, |||\Delta G|||}{2\,\underline{\eta}_\chi}.$$

where $\underline{\eta}_p \stackrel{\rm def}{=} \varrho_p(\alpha, \alpha+\delta)$*, and* $\underline{\eta}_\chi \stackrel{\rm def}{=} \chi(\alpha, \alpha+\delta)$*.*

Proof: Again write $\widetilde{B} = \widetilde{G}S = [I + (\Delta G)G^{-1}]GS = D_1^* B D_2$, where $D_1^* = I + (\Delta G)G^{-1}$ and $D_2 = I$. Then apply Theorem 4.6 to get (4.25) and (4.26), and apply Theorem 4.8 to get (4.27). ∎

Remark. Inequalities (4.24) and (4.27) may provide much tighter bounds than (4.23) and (4.25), especially when $(\Delta G)G^{-1}$ is very close to a skew Hermitian matrix.

5 More on Relative Gaps

In the theorems of §4, various relative gaps play an indispensable role. Those gaps are imposed on either between Λ_1 and $\widetilde{\Lambda}_2$ or between Σ_1 and $\widetilde{\Sigma}_2$. In some applications, it may be more convenient to have theorems where only positive relative gaps between Λ_1 and Λ_2 or between Σ_1 and Σ_2 are assumed. Based on results of Ostrowski [13, 1959] (see also [8, pp.224–225]), Barlow and Demmel [1, 1990], Demmel and Veselić [5, 1992], Eisenstat and Ipsen [6, 1993], Mathias [12, 1994], and Li [11, 1994], theorems in §4 can be modified to accommodate this need. In what follows, we list inequalities for how to bound relative gaps between Λ_1 and $\widetilde{\Lambda}_2$ or between Σ_1 and $\widetilde{\Sigma}_2$ from below for each theorem by relative gaps between Λ_1 and Λ_2 or between Σ_1 and Σ_2. The derivations of these inequalities depends on the fact listed in Lemma 2.1.

For Theorem 4.1: $\eta_2 \geq \min_{1\leq i\leq k,\, 1\leq j\leq n-k} \varrho_2(\lambda_i, \lambda_{k+j}) - \|I - D^*D\|_2$.

For Theorem 4.2: Assume that there exist $\hat{\alpha} > 0$ and $\hat{\delta} > 0$ such that the spectrum of Λ_1 lies entirely in $[-\hat{\alpha}, \hat{\alpha}]$ while that of Λ_2 lies entirely outside $(-\hat{\alpha}-\hat{\delta}, \hat{\alpha}+\hat{\delta})$ (or such that the spectrum of Λ_1 lies entirely outside $(-\hat{\alpha}-\hat{\delta}, \hat{\alpha}+\hat{\delta})$ while that of Λ_2 lies entirely in $[-\hat{\alpha}, \hat{\alpha}]$). If $\varrho_p(\hat{\alpha}, \hat{\alpha}+\hat{\delta}) > \|I - D^*D\|_2$, then there are $\alpha > 0$ and $\delta > 0$ as the theorem requires, and

$$\underline{\eta}_p \geq \varrho_p(\hat{\alpha}, \hat{\alpha}+\hat{\delta}) - \|I - D^*D\|_2.$$

For Theorem 4.3: $\eta_\chi > \min_{1\leq i\leq k,\, 1\leq j\leq n-k} \chi(\lambda_i, \lambda_{k+j}) - \frac{1}{2}\|D - D^{-1}\|_2$.

For Theorem 4.4: Assume that there exist $\hat{\alpha} > 0$ and $\hat{\delta} > 0$ such that

$$\max_{1\leq i\leq k} \lambda_i \leq \hat{\alpha} \quad \text{and} \quad \min_{1\leq j\leq n-k} \lambda_{k+j} \geq \hat{\alpha} + \hat{\delta}$$

or

$$\min_{1\le i\le k}\lambda_i \ge \hat\alpha+\hat\delta \quad\text{and}\quad \max_{1\le j\le n-k}\lambda_{k+j}\le\hat\alpha.$$

If $\chi(\hat\alpha,\hat\alpha+\hat\delta) > \frac{1}{2}\|D-D^{-1}\|_2$, then there are $\alpha>0$ and $\delta>0$ as the theorem requires, and

$$\underline{\eta}_\chi \ge \frac{\chi(\hat\alpha,\hat\alpha+\hat\delta)-\frac{1}{2}\|D-D^{-1}\|_2}{1+\frac{\|D-D^{-1}\|_2}{16}\chi(\hat\alpha,\hat\alpha+\hat\delta)}.$$

For Theorem 4.5: $\eta_2\ge\hat\eta_2-\epsilon$, where $\epsilon=\frac{1}{2\sqrt{2}}(\|D_1^*-D_1^{-1}\|_2+\|D_2^*-D_2^{-1}\|_2)$ (or $\epsilon=\max\{|1-\sigma_{\min}(D_1)\sigma_{\min}(D_2)|,\,|1-\sigma_{\max}(D_1)\sigma_{\max}(D_2)|\}$) and

$$\hat\eta_2 \stackrel{\rm def}{=} \begin{cases} \min\left\{\min\limits_{1\le i\le k,1\le j\le n-k}\varrho_2(\sigma_i,\widetilde\sigma_{k+j}),\ \min\limits_{1\le i\le k}\varrho_2(\sigma_i,0)\right\}, & \text{if } m>n,\\ \min\left\{\min\limits_{1\le i\le k,1\le j\le n-k}\varrho_2(\sigma_i,\widetilde\sigma_{k+j})\right\}, & \text{otherwise.}\end{cases} \tag{5.1}$$

For Theorems 4.6: Assume there exist $\hat\alpha>0$ and $\hat\delta>0$ such that

$$\min_{1\le i\le k}\sigma_i \ge \hat\alpha+\hat\delta \quad\text{and}\quad \max_{1\le j\le n-k}\sigma_{k+j}\le\hat\alpha.$$

If $\varrho_p(\hat\alpha,\hat\alpha+\hat\delta)>\epsilon$, then there are $\alpha>0$ and $\delta>0$ as the theorem requires, and

$$\underline{\eta}_p \ge \varrho_p(\hat\alpha,\hat\alpha+\hat\delta)-\epsilon,$$

where $\epsilon=\frac{1}{2^{1+1/p}}(\|D_1^*-D_1^{-1}\|_2+\|D_2^*-D_2^{-1}\|_2)$ (or $\epsilon=\max\{|1-\sigma_{\min}(D_1)\sigma_{\min}(D_2)|,\,|1-\sigma_{\max}(D_1)\sigma_{\max}(D_2)|\}$).

For Theorems 4.7: Let $\hat\eta_2$ be defined by (5.1) and set

$$\hat\eta_\chi \stackrel{\rm def}{=} \begin{cases} \min\left\{\min\limits_{1\le i\le k,1\le j\le n-k}\chi(\sigma_i,\sigma_{k+j}),\ \min\limits_{1\le i\le k}\chi(\sigma_i,0)\right\}, & \text{if } m>n,\\ \min\left\{\min\limits_{1\le i\le k,1\le j\le n-k}\chi(\sigma_i,\sigma_{k+j})\right\}, & \text{otherwise.}\end{cases} \tag{5.2}$$

If $\hat\eta_2>\frac{1}{2\sqrt{2}}(\epsilon_1+\epsilon_2+\max\{\epsilon_1,\epsilon_2\})$, then

$$\eta_2\ge\hat\eta_2-\frac{1}{2\sqrt{2}}(\epsilon_1+\epsilon_2)\quad\text{and}\quad \eta_\chi\ge\frac{\hat\eta_\chi-\frac{1}{2}(\epsilon_1+\epsilon_2)/\left(1+\frac{1}{32}\epsilon_1\epsilon_2\right)}{1+\frac{\epsilon_1+\epsilon_2}{16}\hat\eta_\chi/\left(1+\frac{1}{32}\epsilon_1\epsilon_2\right)}.$$

For Theorems 4.8: Suppose that there exist $\hat\alpha>0$ and $\hat\delta>0$ such that

$$\min_{1\le i\le k}\sigma_i \ge \hat\alpha+\hat\delta \quad\text{and}\quad \max_{1\le j\le n-k}\sigma_{k+j}\le\hat\alpha.$$

If $\varrho_p(\hat\alpha,\hat\alpha+\hat\delta)>\frac{1}{2^{1+1/p}}(\epsilon_1+\epsilon_2+\max\{\epsilon_1,\epsilon_2\})$, then there are $\alpha>0$ and $\delta>0$ as the theorem requires, and

$$\underline{\eta}_p\ge\varrho_p(\hat\alpha,\hat\alpha+\hat\delta)-\frac{1}{2^{1+1/p}}(\epsilon_1+\epsilon_2)\quad\text{and}\quad \underline{\eta}_\chi\ge\frac{\chi(\hat\alpha,\hat\alpha+\hat\delta)-\frac{1}{2}(\epsilon_1+\epsilon_2)/\left(1+\frac{1}{32}\epsilon_1\epsilon_2\right)}{1+\frac{\epsilon_1+\epsilon_2}{16}\chi(\hat\alpha,\hat\alpha+\hat\delta)/\left(1+\frac{1}{32}\epsilon_1\epsilon_2\right)}.$$

For Theorem 4.9:

$$\eta_2 \geq \hat{\eta}_2 - \frac{1}{2\sqrt{2}}\epsilon \quad \text{and} \quad \eta_\chi \geq \frac{\hat{\eta}_\chi - \epsilon/2}{1 + \frac{\epsilon}{16}\hat{\eta}_\chi},$$

where $\hat{\eta}_2 \stackrel{\text{def}}{=} \min_{1\leq i\leq k, 1\leq j\leq n-k} \varrho_2(\sigma_i, \sigma_{k+j})$, $\hat{\eta}_\chi \stackrel{\text{def}}{=} \min_{1\leq i\leq k, 1\leq j\leq n-k} \chi(\sigma_i, \sigma_{k+j})$, $\epsilon = \|D^* - D^{-1}\|_2$, and $D = I + (\Delta G)G^{-1}$.

For Theorem 4.10: Suppose there exist $\hat{\alpha} > 0$ and $\hat{\delta} > 0$ such that

$$\min_{1\leq i\leq k} \sigma_i \geq \hat{\alpha} + \hat{\delta} \quad \text{and} \quad \max_{1\leq j\leq n-k} \sigma_{k+j} \leq \hat{\alpha}$$

or, the other way around, i.e.,

$$\max_{1\leq i\leq k} \sigma_i \leq \hat{\alpha} \quad \text{and} \quad \min_{1\leq j\leq n-k} \sigma_{k+j} \geq \hat{\alpha} + \hat{\delta},$$

If $\varrho_p(\hat{\alpha}, \hat{\alpha} + \hat{\delta}) > \frac{1}{2^{1+1/p}}\epsilon$, then there are $\alpha > 0$ and $\delta > 0$ as the theorem requires, and

$$\underline{\eta}_p \geq \varrho_p(\hat{\alpha}, \hat{\alpha} + \hat{\delta}) - \frac{1}{2^{1+1/p}}\epsilon \quad \text{and} \quad \underline{\eta}_\chi \geq \frac{\chi(\hat{\alpha}, \hat{\alpha} + \hat{\delta}) - \epsilon/2}{1 + \frac{\epsilon}{16}\chi(\hat{\alpha}, \hat{\alpha} + \hat{\delta})},$$

where $\epsilon = \|D^* - D^{-1}\|_2$ and $D = I + (\Delta G)G^{-1}$.

6 A Word on Eisenstat-Ipsen's Theorems

Eisenstat and Ipsen [7, 1994] have obtained a few bounds on eigenspace variations for A and $\widetilde{A} = D^*AD$ and on singular space variations for B and $= \widetilde{B} = D_1^*BD_2$. Their bounds for subspaces of dimension $k > 1$ contain a factor $\sqrt{k}$ which makes their results less competitive to ours. For this reason we will not compare their bounds for subspaces of dimension $k > 1$ with ours.

For the problem studied in Theorem 4.1, Eisenstat and Ipsen [7, 1994] tried to bound the angle θ_j between $\widetilde{u}_j$ and $\mathcal{R}(U_1)$, where $\widetilde{u}_1, \widetilde{u}_2, \cdots, \widetilde{u}_k$ are the columns of $\widetilde{U}_1$. They showed

$$\sin\theta_j \leq \frac{\|I - D^{-*}D^{-1}\|_2}{\delta_j} + \|I - D\|_2, \tag{6.1}$$

where

$$\delta_j \stackrel{\text{def}}{=} \begin{cases} \min_{k+1\leq i\leq n} \frac{|\lambda_i - \widetilde{\lambda}_j|}{|\widetilde{\lambda}_j|}, & \text{if } \widetilde{\lambda}_j \neq 0, \\ \infty & \text{otherwise.} \end{cases}$$

Inequality (6.1) does provide a nice bound. Unfortunately it does not bound straightforwardly $\|\sin\Theta(U_1, \widetilde{U}_1)\|_2$, and generally,

$$\sin\theta_j \leq \|\sin\Theta(U_1, \widetilde{U}_1)\|_2 \quad \text{for } j = 1, 2, \ldots, k$$

and all of them may be strict. In the worst case, $\|\sin\Theta(U_1,\widetilde{U}_1)\|_2$ could be as large as $\sqrt{k}\max_{1\le j\le k}\sin\theta_j$. To make a fair comparison to our inequality (4.1), we consider the case $k=1$. One infers from inequality (4.1) that[3]

$$\sin\theta_1 \le \frac{\sqrt{\|I-D^{-1}\|_2^2+\|I-D^*\|_2^2}}{\min_{2\le i\le n}\varrho_2(\lambda_i,\widetilde{\lambda}_1)}. \tag{6.2}$$

It looks that (6.1) may be potentially sharper because δ_1 may be much larger than $\min_{2\le i\le n}\varrho_2(\lambda_i,\widetilde{\lambda}_1)$. But this is not quite true because of the extra term $\|I-D\|_2$ in (6.1) which stays no matter how large δ_1 is. We present our arguments as follows.

1. These relative perturbation bounds are most likely to be used when the closest eigenvalue λ_ℓ among all $\{\lambda_i\}_{i=2}^n$ to $\widetilde{\lambda}_1$ has about the same magnitude as $\widetilde{\lambda}_1$, i.e., when $|\lambda_\ell|\approx|\widetilde{\lambda}_1|$ and

$$\delta_1\approx|\lambda_\ell-\widetilde{\lambda}_1|/|\widetilde{\lambda}_1|\approx\sqrt{2}\varrho_2(\lambda_\ell,\widetilde{\lambda}_1)\approx\sqrt{2}\min_{2\le i\le n}\varrho_2(\lambda_i,\widetilde{\lambda}_1).$$

   In such a situation, if $I-D$ is very tiny, then

$$\begin{aligned}\sqrt{\|I-D^{-1}\|_2^2+\|I-D^*\|_2^2} &\approx \sqrt{2}\|I-D\|_2+O(\|I-D\|_2^2),\\ \|I-D^{-*}D^{-1}\|_2 &\approx 2\|I-D\|_2+O(\|I-D\|_2^2).\end{aligned}$$

   Hence asymptotically, inequalities (6.2) and (6.1) read, respectively

$$\begin{aligned}\sin\theta_1 &\le \frac{\sqrt{2}\|I-D\|_2}{\varrho_2(\lambda_\ell,\widetilde{\lambda}_1)}+O(\|I-D\|_2^2),\\ \sin\theta_1 &\le \frac{\sqrt{2}\|I-D\|_2}{\varrho_2(\lambda_\ell,\widetilde{\lambda}_1)}+\|I-D\|_2+O(\|I-D\|_2^2).\end{aligned}$$

   So our bound (6.2) is asymptotically sharper by amount $\|I-D\|_2$.

2. On the other hand, if the closest eigenvalue λ_ℓ among all $\{\lambda_i\}_{i=2}^n$ to $\widetilde{\lambda}_1$ has much bigger magnitude than $|\widetilde{\lambda}_1|$, i.e., $|\lambda_\ell|\gg|\widetilde{\lambda}_1|$, then

$$\min_{2\le i\le n}\varrho_2(\lambda_i,\widetilde{\lambda}_1)\approx 1.$$

   Thus asymptotically, inequality (6.2) read

$$\sin\theta_1\le\sqrt{2}\|I-D\|_2+O(\|I-D\|_2^2)$$

   which cannot be much worse than (6.1).

[3]By treating A and $\widetilde{A}$ symmetrically, one can see Theorem 4.1 remains valid with the relative gap η_2 between Λ_1 and $\widetilde{\Lambda}_2$ replaced by the relative gap $\min_{1\le i\le k,\,1\le j\le n-k}\varrho_2(\widetilde{\lambda}_i,\lambda_{k+j})$ between $\widetilde{\Lambda}_1$ and Λ_2.

To be short, we have shown that it is possible for our inequality (6.2) to be less sharp than (6.1) by a constant factor and in the case when these bounds are most likely to be used in estimating errors (6.2) is sharper.

Another point we like to make is that we had given up some sharpness for elegantness in the derivation of (6.2). Recall that we have, besides (7.1), also

$$\widetilde{\Lambda}_1 \widetilde{U}_1^* U_2 - \widetilde{U}_1^* U_2 \Lambda_2 = \widetilde{\Lambda}_1 \widetilde{U}_1^* (I - D^{-1}) U_2 + \widetilde{U}_1^* (D^* - I) U_2 \Lambda_2. \tag{6.3}$$

When $\widetilde{\Lambda}_1 = 0$, this equation reduces to $-\widetilde{U}_1^* U_2 = \widetilde{U}_1^*(I - D^*) U_2$ which implies

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_2 \le \|D^* - I\|_2,$$

a better bound than (6.2), provided $\lambda_j \neq 0$ for $j = k+1, \cdots, n$. Equation (6.3) implies

$$\widetilde{u}_1^* U_2 = \widetilde{u}_1^* (D^{-1} - I) U_2 \widetilde{\lambda}_1 (\widetilde{\lambda}_1 I - \Lambda_2)^{-1} + \widetilde{u}_1^* (I - D^*) U_2 \Lambda_2 (\widetilde{\lambda}_1 I - \Lambda_2)^{-1},$$

which can be used to obtain an identity for $\sin\theta_1 = \|\widetilde{u}_1^* U_2\|_2$! Generally (6.3) produces

$$\begin{aligned} &\|\sin\Theta(U_1, \widetilde{U}_1)\|_2 \\ &\quad \le \max_{1\le i\le k,\, 1\le j\le n-k} \frac{|\widetilde{\lambda}_i|}{|\widetilde{\lambda}_i - \lambda_{k+j}|} \|I - D^{-1}\|_2 + \max_{1\le i\le k,\, 1\le j\le n-k} \frac{|\lambda_{k+j}|}{|\widetilde{\lambda}_i - \lambda_{k+j}|} \|D^* - I\|_2 \end{aligned}$$

which would be a better bound than (4.1) in the case when

$$\text{either } \min_{1\le i\le k} |\widetilde{\lambda}_i| \gg \max_{1\le j\le n-k} |\lambda_{k+j}| \quad \text{or} \quad \max_{1\le i\le k} |\widetilde{\lambda}_i| \ll \min_{1\le j\le n-k} |\lambda_{k+j}|.$$

Eisenstat and Ipsen [7, 1994] treated singular value problems in a very similar way as they did to eigenspaces. This makes our arguments above apply to our bounds and their bounds for singular value problems.

Eisenstat and Ipsen [7, 1994] did not study bounds in other matrix norms.

7 Proofs of Theorems 4.1 and 4.2

Let $R = \widetilde{A} U_1 - U_1 \Lambda_1 = (\widetilde{A} - A) U_1$ as defined in (3.1). Notice that

$$\begin{aligned} \widetilde{U}_2^* R &= \widetilde{U}_2^* \widetilde{A} U_1 - \widetilde{U}_2^* U_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{U}_2^* U_1 - \widetilde{U}_2^* U_1 \Lambda_1, \\ \widetilde{U}_2^* R &= \widetilde{U}_2^* (\widetilde{A} - A) U_1 = \widetilde{U}_2^* \left[D^* A D - D^* A + D^* A - A\right] U_1 \\ &= \widetilde{U}_2^* \left[D^* A D (I - D^{-1}) + (D^* - I) A\right] U_1 \\ &= \widetilde{\Lambda}_2 \widetilde{U}_2^* (I - D^{-1}) U_1 + \widetilde{U}_2^* (D^* - I) U_1 \Lambda_1. \end{aligned}$$

Thus, we have

$$\widetilde{\Lambda}_2 \widetilde{U}_2^* U_1 - \widetilde{U}_2^* U_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{U}_2^* (I - D^{-1}) U_1 + \widetilde{U}_2^* (D^* - I) U_1 \Lambda_1. \tag{7.1}$$

Lemma 7.1 *Let $\Omega \in \mathbf{C}^{s\times s}$ and $\Gamma \in \mathbf{C}^{t\times t}$ be two Hermitian matrices, and let E, $F \in \mathbf{C}^{s\times t}$. If $\lambda(\Omega)\bigcap\lambda(\Gamma) = \emptyset$, then matrix equation $\Omega X - X\Gamma = \Omega E + F\Gamma$ has a unique solution, and moreover $\|X\|_{\rm F} \le \sqrt{\|E\|_{\rm F}^2 + \|F\|_{\rm F}^2}\Big/ \eta_2$, where $\eta_2 \stackrel{\rm def}{=} \min_{\omega\in\lambda(\Omega),\gamma\in\lambda(\Gamma)} \varrho_2(\omega,\gamma)$.*

Proof: For any unitary matrices $P \in \mathbf{U}_s$ and $Q \in \mathbf{U}_t$, the substitutions

$$\Omega \leftarrow P^*\Omega P, \quad \Gamma \leftarrow Q^*\Gamma Q, \quad X \leftarrow P^*XQ, \quad E \leftarrow P^*EQ, \quad \text{and} \quad F \leftarrow P^*FQ$$

leave the lemma unchanged, so we may assume without loss of generality that $\Omega = \operatorname{diag}(\omega_1, \omega_2, \ldots, \omega_s)$ and $\Gamma = \operatorname{diag}(\gamma_1, \gamma_2, \ldots, \gamma_t)$.

Write $X = (x_{ij})$, $E = (e_{ij})$, and $F = (f_{ij})$. Entrywise, equation $\Omega X - X\Gamma = \Omega E + F\Gamma$ reads $\omega_i x_{ij} - x_{ij}\gamma_j = \omega_i e_{ij} + f_{ij}\gamma_j$. Thus x_{ij} exists uniquely provided $\omega_i \ne \gamma_j$ which is guaranteed by the assumption $\lambda(\Omega)\bigcap\lambda(\Gamma) = \emptyset$, and moreover

$$|(\omega_i - \gamma_j)x_{ij}|^2 = |\omega_i x_{ij} - x_{ij}\gamma_j|^2 = |\omega_i e_{ij} + f_{ij}\gamma_j|^2 \le (|\omega_i|^2 + |\gamma|^2)(|e_{ij}|^2 + |f_{ij}|^2)$$

by the Cauchy-Schwarz inequality. This implies

$$\begin{aligned} |x_{ij}|^2 &\le \frac{|e_{ij}|^2 + |f_{ij}|^2}{[\varrho_2(\omega_i,\gamma_j)]^2} \le \frac{|e_{ij}|^2 + |f_{ij}|^2}{\eta_2^2} \\ \Rightarrow \quad \|X\|_{\rm F}^2 = \sum_{i,j}|x_{ij}|^2 &\le \frac{\sum_{i,j}|e_{ij}|^2 + \sum_{i,j}|f_{ij}|^2}{\eta_2^2} = \frac{\|E\|_{\rm F}^2 + \|F\|_{\rm F}^2}{\eta_2^2}, \end{aligned}$$

as was to be shown. ∎

Proof of Theorem 4.1: By Lemma 7.1 and equation (7.1), we have

$$\begin{aligned} \|\widetilde{U}_2^* U_1\|_{\rm F}^2 &\le \frac{\|\widetilde{U}_2^*(I - D^{-1})U_1\|_{\rm F}^2 + \|\widetilde{U}_2^*(D^* - I)U_1\|_{\rm F}^2}{\eta_2^2} \\ &\le \frac{1}{\eta_2^2}\left(\|(I - D^{-1})U_1\|_{\rm F}^2 + \|(D^* - I)U_1\|_{\rm F}^2\right). \end{aligned}$$

This completes the proof of Theorem 4.1, since $\|\sin\Theta(U_1, \widetilde{U}_1)\|_{\rm F} = \|\widetilde{U}_2^* U_1\|_{\rm F}$. ∎

Lemma 7.2 *Let $\Omega \in \mathbf{C}^{s\times s}$ and $\Gamma \in \mathbf{C}^{t\times t}$ be two Hermitian matrices, and let E, $F \in \mathbf{C}^{s\times t}$. If there exist $\alpha > 0$ and $\delta > 0$ such that*

$$\|\Omega\|_2 \le \alpha \quad \textit{and} \quad \|\Gamma^{-1}\|_2^{-1} \ge \alpha + \delta$$

or

$$\|\Omega^{-1}\|_2^{-1} \ge \alpha + \delta \quad \textit{and} \quad \|\Gamma\|_2 \le \alpha,$$

then matrix equation $\Omega X - X\Gamma = \Omega E + F\Gamma$ has a unique solution, and moreover for any unitarily invariant norm $|||\cdot|||$, $|||X||| \le \sqrt[q]{|||E|||^q + |||F|||^q}\Big/ \underline{\eta}_p$, where $\underline{\eta}_p \stackrel{\rm def}{=} \varrho_p(\alpha, \alpha+\delta)$.

Proof: First of all, the conditions of this lemma imply $\lambda(\Omega)\bigcap\lambda(\Gamma)=\emptyset$, thus X exists uniquely by Lemma 7.1. In what follows, we present a proof of the bound for $|||X|||$ for the case $\|\Omega\|_2\le\alpha$ and $\|\Gamma^{-1}\|_2^{-1}\ge\alpha+\delta$. A proof for the other case is analogous. Post-multiply equation $\Omega X - X\Gamma = \Omega E + F$, by Γ^{-1} to get

$$\Omega X\Gamma^{-1} - X = \Omega E\Gamma^{-1} + F. \tag{7.2}$$

Under the assumptions $\|\Omega\|_2\le\alpha$ and $\|\Gamma^{-1}\|_2^{-1}\ge\alpha+\delta \Rightarrow \|\Gamma^{-1}\|_2\le\frac{1}{\alpha+\delta}$, we have

$$\begin{aligned}\left|\left|\left|\Omega X\Gamma^{-1} - X\right|\right|\right| &\ge |||X||| - \left|\left|\left|\Omega X\Gamma^{-1}\right|\right|\right| \ge |||X||| - \|\Omega\|_2\,|||X|||\,\|\Gamma^{-1}\|_2 \\ &\ge |||X||| - \alpha\,|||X|||\,\frac{1}{\alpha+\delta} = \left(1-\frac{\alpha}{\alpha+\delta}\right)|||X|||,\end{aligned}$$

and

$$\begin{aligned}\left|\left|\left|\Omega E\Gamma^{-1} + F\right|\right|\right| &\le \left|\left|\left|\Omega E\Gamma^{-1}\right|\right|\right| + |||F||| \le \|\Omega\|_2\,|||E|||\,\|\Gamma^{-1}\|_2 + |||F||| \\ &\le \alpha\,|||E|||\,\frac{1}{\alpha+\delta} + |||F||| \le \sqrt[p]{1+\frac{\alpha^p}{(\alpha+\delta)^p}}\sqrt[q]{|||E|||^q + |||F|||^q}.\end{aligned}$$

By equation (7.2), we deduce that

$$\left(1-\frac{\alpha}{\alpha+\delta}\right)|||X||| \le \sqrt[p]{1+\frac{\alpha^p}{(\alpha+\delta)^p}}\sqrt[q]{|||E|||^q + |||F|||^q}$$

from which the desired inequality follows. ∎

Proof of Theorem 4.2: By Lemma 7.2 and equation (7.1), we have

$$\begin{aligned}\left|\left|\left|\widetilde{U}_2^*U_1\right|\right|\right| &\le \sqrt[q]{\left|\left|\left|\widetilde{U}_2^*(I-D^{-1})U_1\right|\right|\right|^q + \left|\left|\left|\widetilde{U}_2^*(D^*-I)U_1\right|\right|\right|^q}\Big/\underline{\eta}_p \\ &\le \sqrt[q]{|||(I-D^{-1}I)U_1|||^q + |||(D^*-I)U_1|||^q}\Big/\underline{\eta}_p,\end{aligned}$$

as required since $\left|\left|\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\right|\right| = \left|\left|\left|\widetilde{U}_2^*U_1\right|\right|\right|$. ∎

8 Proofs of Theorems 4.3 and 4.4

Notice that

$$\begin{aligned}A &= S^*HS = (H^{1/2}S)^*H^{1/2}S, \\ \widetilde{A} &= S^*H^{1/2}(I+H^{-1/2}(\Delta H)H^{-1/2})H^{1/2}S \\ &= \left((I+H^{-1/2}(\Delta H)H^{-1/2})^{1/2}H^{1/2}S\right)^*(I+H^{-1/2}(\Delta H)H^{-1/2})^{1/2}H^{1/2}S.\end{aligned}$$

Set $B = S^*H^{1/2}$ and $\widetilde{B} = S^*H^{1/2}(I + H^{-1/2}(\Delta H)H^{-1/2})^{1/2} \stackrel{\text{def}}{=} BD$, where $D = \left(I + H^{-1/2}(\Delta H)H^{-1/2}\right)^{1/2}$. Given the eigendecompositions of A and $\widetilde{A}$ as in (1.1), (1.2), (1.3), and (1.4), one easily see that B and $\widetilde{B}$ admit the following SVDs.

$$\begin{aligned} B &= U\Lambda^{1/2}V^* \equiv (U_1, U_2)\begin{pmatrix} \Lambda_1^{1/2} & \\ & \Lambda_2^{1/2} \end{pmatrix}\begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix}, \\ \widetilde{B} &= \widetilde{U}\widetilde{\Lambda}^{1/2}\widetilde{V}^* \equiv (\widetilde{U}_1, \widetilde{U}_2)\begin{pmatrix} \widetilde{\Lambda}_1^{1/2} & \\ & \widetilde{\Lambda}_2^{1/2} \end{pmatrix}\begin{pmatrix} \widetilde{V}_1^* \\ \widetilde{V}_2^* \end{pmatrix}, \end{aligned}$$

where U, $\widetilde{U}$ are the same as in (1.1) and (1.2), V_1, $\widetilde{V}_1 \in \mathbf{C}^{n\times k}$. Notice that

$$\widetilde{A} - A = \widetilde{B}\widetilde{B}^* - BB^* = \widetilde{B}D^*B^* - \widetilde{B}D^{-1}B^* = \widetilde{B}(D^* - D^{-1})B^*.$$

Pre- and post-multiply the equations by $\widetilde{U}^*$ and U, respectively, to get $\widetilde{\Lambda}\widetilde{U}^*U - \widetilde{U}^*U\Lambda = \widetilde{\Lambda}^{1/2}\widetilde{V}^*(D^* - D^{-1})V\Lambda^{1/2}$ which yields

$$\widetilde{\Lambda}_2\widetilde{U}_2^*U_1 - \widetilde{U}_2^*U_1\Lambda_1 = \widetilde{\Lambda}_2^{1/2}\widetilde{V}_2^*(D^* - D^{-1})V_1\Lambda_1^{1/2}. \tag{8.1}$$

The following inequality will be very useful in the rest of our proofs.

$$\begin{aligned} \left|\!\left|\!\left| \widetilde{V}_2^*(D^* - D^{-1})V_1 \right|\!\right|\!\right| &\le \left|\!\left|\!\left| D^* - D^{-1} \right|\!\right|\!\right| \\ &= \left|\!\left|\!\left| \left(I + H^{-1/2}(\Delta H)H^{-1/2}\right)^{1/2} - \left(I + H^{-1/2}(\Delta H)H^{-1/2}\right)^{-1/2} \right|\!\right|\!\right| \\ &\le \left\| \left(I + H^{-1/2}(\Delta H)H^{-1/2}\right)^{-1/2} \right\|_2 \left|\!\left|\!\left| H^{-1/2}(\Delta H)H^{-1/2} \right|\!\right|\!\right| \\ &\le \frac{\|H^{-1}\|_2 \, |\!|\!|\Delta H|\!|\!|}{\sqrt{1 - \|H^{-1}\|_2\|\Delta H\|_2}}. \end{aligned}$$

Lemma 8.1 *Let* $\Omega \in \mathbf{C}^{s\times s}$ *and* $\Gamma \in \mathbf{C}^{t\times t}$ *be two nonnegative definite Hermitian matrices, and let* $E \in \mathbf{C}^{s\times t}$. *If* $\lambda(\Omega)\bigcap\lambda(\Gamma) = \emptyset$, *then matrix equation* $\Omega X - X\Gamma = \Omega^{1/2}E\Gamma^{1/2}$ *has a unique solution* $X \in \mathbf{C}^{s\times t}$, *and moreover* $\|X\|_{\rm F} \le \|E\|_{\rm F}/\eta_\chi$, *where* $\eta_\chi \stackrel{\text{def}}{=} \min_{\omega\in\lambda(\Omega), \gamma\in\lambda(\Gamma)} \chi(\omega, \gamma)$.

Proof: For any unitary matrices $P \in \mathbf{U}_s$ and $Q \in \mathbf{U}_t$, the substitutions

$$\Omega \leftarrow P^*\Omega P, \quad \Omega^{1/2} \leftarrow (P^*\Omega P)^{1/2}, \quad \Gamma \leftarrow Q^*\Gamma Q, \quad \Gamma^{1/2} \leftarrow (Q^*\Gamma Q)^{1/2},$$

$$X \leftarrow P^*XQ, \quad \text{and} \quad E \leftarrow P^*EQ$$

leave the lemma unchanged, so we may assume without loss of generality that $\Omega = \mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_s)$ and $\Gamma = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_t)$.

Write $X = (x_{ij})$, $E = (e_{ij})$. Entrywise, equation $\Omega X - X\Gamma = \Omega^{1/2}E\Gamma^{1/2}$ reads $\omega_i x_{ij} - x_{ij}\gamma_j = \sqrt{\omega_i}e_{ij}\sqrt{\gamma_j}$. As long as $\omega_i \ne \gamma_j$, x_{ij} exists uniquely, and

$$|x_{ij}|^2 = |e_{ij}|^2/\chi(\omega_i, \gamma_j) \le |e_{ij}|^2/\eta_\chi$$

summing which over $1 \le i \le s$ and $1 \le j \le t$ leads to the desired inequality. ∎

Proof of Theorem 4.3: Equation (8.1) and Lemma 8.1 imply

$$\left\|\sin\Theta(U_1,\widetilde{U}_1)\right\|_{\mathrm{F}} = \left\|\widetilde{U}_2^* U_1\right\|_{\mathrm{F}} \le \frac{\|\widetilde{V}_2^*(D^* - D^{-1})V_1\|_{\mathrm{F}}}{\eta_\chi} \le \frac{\|D^* - D^{-1}\|_{\mathrm{F}}}{\eta_\chi},$$

as required. ∎

Lemma 8.2 *Let $\Omega \in \mathbf{C}^{s\times s}$ and , $\in \mathbf{C}^{t\times t}$ be two nonnegative definite Hermitian matrices, and let $E \in \mathbf{C}^{s\times t}$. If there exist $\alpha > 0$ and $\delta > 0$ such that*

$$\|\Omega\|_2 \le \alpha \quad \textit{and} \quad \|,^{-1}\|_2^{-1} \ge \alpha + \delta$$

or

$$\|\Omega^{-1}\|_2^{-1} \ge \alpha + \delta \quad \textit{and} \quad \|,\|_2 \le \alpha,$$

then matrix equation $\Omega X - X, = \Omega^{1/2} E ,^{1/2}$ has a unique solution $X \in \mathbf{C}^{s\times t}$, and moreover $\|\!|X|\!\| \le \|\!|E|\!\| / \underline{\eta}_\chi$, where $\underline{\eta}_\chi \stackrel{\mathrm{def}}{=} \chi(\alpha, \alpha+\delta)$.

Proof: The existence and uniqueness of X are easy to see because the conditions of this lemma imply $\lambda(\Omega) \bigcap \lambda(,) = \emptyset$. To bound $\|\!|X|\!\|$, we present a proof for the case $\|\Omega\|_2 \le \alpha$ and $\|,^{-1}\|_2^{-1} \ge \alpha + \delta$. A proof for the other case is analogous. Post-multiply equation $\Omega X - X, = \Omega^{1/2} E ,^{1/2}$ by $,^{-1}$ to get

$$\Omega X ,^{-1} - X = \Omega^{1/2} E ,^{-1/2}. \tag{8.2}$$

Under the assumptions $\|\Omega\|_2 \le \alpha$ and $\|,^{-1}\|_2^{-1} \ge \alpha + \delta \Rightarrow \|,^{-1}\|_2 \le \frac{1}{\alpha+\delta}$, we have $\|\!|\Omega X ,^{-1} - X|\!\| \ge \left(1 - \frac{\alpha}{\alpha+\delta}\right) \|\!|X|\!\|$ as in the proof of Lemma 7.2 and

$$\left\|\!\left|\Omega^{1/2} E ,^{-1/2}\right|\!\right\| \le \|\Omega^{1/2}\|_2 \,\|\!|E|\!\|\, \|,^{-1/2}\|_2 \le \sqrt{\alpha}\, \|\!|E|\!\| \,\frac{1}{\sqrt{\alpha+\delta}}.$$

By equation (8.2), we deduce that

$$\left(1 - \frac{\alpha}{\alpha+\delta}\right) \|\!|X|\!\| \le \sqrt{\frac{\alpha}{\alpha+\delta}}\, \|\!|E|\!\|$$

from which the desired inequality follows. ∎ ∎

Proof of Theorem 4.4: Equation (8.1) and Lemma 8.2 imply

$$\left\|\!\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\!\right\| = \left\|\!\left|\widetilde{U}_2^* U_1\right|\!\right\| \le \frac{\left\|\!\left|\widetilde{V}_2^*(D^* - D^{-1})V_1\right|\!\right\|}{\underline{\eta}_\chi} \le \frac{\|\!|D^* - D^{-1}|\!\|}{\underline{\eta}_\chi},$$

as required. ∎

9 Proofs of Theorems 4.5 and 4.6

Let $R_{\rm R} = \widetilde{B}V_1 - U_1\Sigma_1 = (\widetilde{B} - B)V_1$ and $R_{\rm L} = \widetilde{B}^*U_1 - V_1\Sigma_1 = (\widetilde{B}^* - B^*)U_1$ as defined in (3.8).

9.1 The Square Case: $m = n$

When $m = n$, the SVDs (3.4) and (3.5) read

$$\begin{aligned} B &= U\Sigma V^* \equiv (U_1, U_2)\begin{pmatrix} \Sigma_1 & \\ & \Sigma_2 \end{pmatrix}\begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix}, \\ \widetilde{B} &= \widetilde{U}\widetilde{\Sigma}\widetilde{V}^* \equiv (\widetilde{U}_1, \widetilde{U}_2)\begin{pmatrix} \widetilde{\Sigma}_1 & \\ & \widetilde{\Sigma}_2 \end{pmatrix}\begin{pmatrix} \widetilde{V}_1^* \\ \widetilde{V}_2^* \end{pmatrix}. \end{aligned}$$

Notice that

$$\begin{aligned} \widetilde{U}_2^* R_{\rm R} &= \widetilde{U}_2^*\widetilde{B}V_1 - \widetilde{U}_2^*U_1\Sigma_1 = \widetilde{\Sigma}_2\widetilde{V}_2^*V_1 - \widetilde{U}_2^*U_1\Sigma_1, \\ \widetilde{U}_2^* R_{\rm R} &= \widetilde{U}_2^*(\widetilde{B} - B)V_1 = \widetilde{U}_2^*(D_1^*BD_2 - D_1^*B + D_1^*B - B)V_1 \\ &= \widetilde{U}_2^*\left[\widetilde{B}(I - D_2^{-1}) + (D_1^* - I)B\right]V_1 \\ &= \widetilde{\Sigma}_2\widetilde{V}_2^*(I - D_2^{-1})V_1 + \widetilde{U}_2^*(D_1^* - I)U_1\Sigma_1 \end{aligned}$$

to get

$$\widetilde{\Sigma}_2\widetilde{V}_2^*V_1 - \widetilde{U}_2^*U_1\Sigma_1 = \widetilde{\Sigma}_2\widetilde{V}_2^*(I - D_2^{-1})V_1 + \widetilde{U}_2^*(D_1^* - I)U_1\Sigma_1. \tag{9.1}$$

On the other hand,

$$\begin{aligned} \widetilde{V}_2^* R_{\rm L} &= \widetilde{V}_2^*\widetilde{B}^*U_1 - \widetilde{V}_2^*V_1\Sigma_1 = \widetilde{\Sigma}_2\widetilde{U}_2^*U_1 - \widetilde{V}_2^*V_1\Sigma_1. \\ \widetilde{V}_2^* R_{\rm L} &= \widetilde{V}_2^*(\widetilde{B}^* - B^*)U_1 = \widetilde{V}_2^*(D_2^*B^*D_1 - D_2^*B^* + D_2^*B^* - B^*)U_1 \\ &= \widetilde{V}_2^*\left[\widetilde{B}^*(I - D_1^{-1}) + (D_2^* - I)B^*\right]U_1 \\ &= \widetilde{\Sigma}_2\widetilde{U}_2^*(I - D_1^{-1})U_1 + \widetilde{V}_2^*(D_2^* - I)V_1\Sigma_1 \end{aligned}$$

which produce

$$\widetilde{\Sigma}_2\widetilde{U}_2^*U_1 - \widetilde{V}_2^*V_1\Sigma_1 = \widetilde{\Sigma}_2\widetilde{U}_2^*(I - D_1^{-1})U_1 + \widetilde{V}_2^*(D_2^* - I)V_1\Sigma_1. \tag{9.2}$$

Equations (9.1) and (9.2) take an equivalent forms as a single matrix equation with dimensions doubled.

$$\begin{aligned} &\begin{pmatrix} & \widetilde{\Sigma}_2 \\ \widetilde{\Sigma}_2 & \end{pmatrix}\begin{pmatrix} \widetilde{U}_2^*U_1 & \\ & \widetilde{V}_2^*V_1 \end{pmatrix} - \begin{pmatrix} \widetilde{U}_2^*U_1 & \\ & \widetilde{V}_2^*V_1 \end{pmatrix}\begin{pmatrix} & \Sigma_1 \\ \Sigma_1 & \end{pmatrix} \\ &\qquad = \begin{pmatrix} & \widetilde{\Sigma}_2 \\ \widetilde{\Sigma}_2 & \end{pmatrix}\begin{pmatrix} \widetilde{U}_2^*(I - D_1^{-1})U_1 & \\ & \widetilde{V}_2^*(I - D_2^{-1})V_1 \end{pmatrix} \\ &\qquad\quad + \begin{pmatrix} \widetilde{U}_2^*(D_1^* - I)U_1 & \\ & \widetilde{V}_2^*(D_2^* - I)V_1 \end{pmatrix}\begin{pmatrix} & \Sigma_1 \\ \Sigma_1 & \end{pmatrix}. \end{aligned} \tag{9.3}$$

Proof of Theorem 4.5: Notice that the eigenvalues of $\begin{pmatrix} & \widetilde{\Sigma}_2 \\ \widetilde{\Sigma}_2 & \end{pmatrix}$ are $\pm\widetilde{\sigma}_{k+j}$ and these of $\begin{pmatrix} & \Sigma_1 \\ \Sigma_1 & \end{pmatrix}$ are $\pm\sigma_i$, and that

$$\varrho_2(\sigma_i, -\widetilde{\sigma}_{k+j}) \geq \varrho_2(\sigma_i, \widetilde{\sigma}_{k+j}) \quad \text{and} \quad \varrho_2(-\sigma_i, \widetilde{\sigma}_{k+j}) \geq \varrho_2(\sigma_i, \widetilde{\sigma}_{k+j}).$$

By Lemma 7.1 and equation (9.3), we have

$$\begin{aligned} &\|\widetilde{U}_2^* U_1\|_{\mathrm{F}}^2 + \|\widetilde{V}_2^* V_1\|_{\mathrm{F}}^2 \\ &\quad \leq \frac{1}{\eta_2^2}\left[\|\widetilde{U}_2^*(I - D_1^{-1})U_1\|_{\mathrm{F}}^2 + \|\widetilde{U}_2^*(D_1^* - I)U_1\|_{\mathrm{F}}^2 + \|\widetilde{V}_2^*(I - D_2^{-1})V_1\|_{\mathrm{F}}^2 + \|\widetilde{V}_2^*(D_2^* - I)V_1\|_{\mathrm{F}}^2\right] \\ &\quad \leq \frac{1}{\eta_2^2}\left[\|(I - D_1^{-1})U_1\|_{\mathrm{F}}^2 + \|(D_1^* - I)U_1\|_{\mathrm{F}}^2 + \|(I - D_2^{-1})V_1\|_{\mathrm{F}}^2 + \|(D_2^* - I)V_1\|_{\mathrm{F}}^2\right] \end{aligned}$$

which completes the proof. ∎

Lemma 9.1 *Let* $\Omega \in \mathbf{C}^{s\times s}$ *and* , $\in \mathbf{C}^{t\times t}$ *be two Hermitian matrices, and let* X, Y, E, $\widetilde{E}$, F, $\widetilde{F}, \in \mathbf{C}^{s\times t}$ *satisfying*

$$\Omega X - Y, = \Omega E + F, \quad \textit{and} \quad \Omega Y - X, = \Omega \widetilde{E} + \widetilde{F}, .$$

If there exist $\alpha > 0$ *and* $\delta > 0$ *such that*

$$\|\Omega\|_2 \leq \alpha \quad \textit{and} \quad \|,^{-1}\|_2^{-1} \geq \alpha + \delta$$

or

$$\|\Omega^{-1}\|_2^{-1} \geq \alpha + \delta \quad \textit{and} \quad \|,\|_2 \leq \alpha,$$

then for any unitarily invariant norm $|||\cdot|||$,

$$\max\{|||X|||, |||Y|||\} \leq \frac{1}{\underline{\eta}_p}\max\left\{\sqrt[q]{|||E|||^q + |||F|||^q}, \sqrt[q]{\left|\left|\left|\widetilde{E}\right|\right|\right|^q + \left|\left|\left|\widetilde{F}\right|\right|\right|^q}\right\}, \tag{9.4}$$

where $\underline{\eta}_p \stackrel{\mathrm{def}}{=} \varrho_p(\alpha, \alpha + \delta)$.

Proof: We present a proof for the case $\|\Omega\|_2 \leq \alpha$ and $\|,^{-1}\|_2^{-1} \geq \alpha + \delta$. A proof for the other case is analogous. Consider first the subcase $|||X||| \geq |||Y|||$. Post-multiply equation $\Omega Y - X, = \Omega\widetilde{E} + \widetilde{F},$ by $,^{-1}$ to get

$$\Omega Y,^{-1} - X = \Omega \widetilde{E},^{-1} + \widetilde{F}. \tag{9.5}$$

Then we have, by $\|\Omega\|_2 \leq \alpha$ and $\|,^{-1}\|_2^{-1} \geq \alpha + \delta \Rightarrow \|,^{-1}\|_2 \leq \frac{1}{\alpha+\delta}$, that

$$\begin{aligned} \left|\left|\left|\Omega Y,^{-1} - X\right|\right|\right| &\geq |||X||| + -\left|\left|\left|\Omega Y,^{-1}\right|\right|\right| \geq |||X||| - \|\Omega\|_2\, |||Y|||\, \|,^{-1}\|_2 \\ &\geq |||X||| - \alpha\, |||Y|||\, \frac{1}{\alpha+\delta} \geq |||X||| - \alpha\, |||X|||\, \frac{1}{\alpha+\delta} \\ &= \left(1 - \frac{\alpha}{\alpha+\delta}\right)|||X||| \end{aligned}$$

and

$$\begin{aligned}\left\|\!\left|\Omega\widetilde{E}, ^{-1}+\widetilde{F}\right|\!\right\| &\le \left\|\!\left|\Omega\widetilde{E}, ^{-1}\right|\!\right\| + \left\|\!\left|\widetilde{F}\right|\!\right\| \le \|\Omega\|_2 \left\|\!\left|\widetilde{E}\right|\!\right\| \, \|, ^{-1}\|_2 + \left\|\!\left|\widetilde{F}\right|\!\right\| \\ &\le \alpha \left\|\!\left|\widetilde{E}\right|\!\right\| \frac{1}{\alpha+\delta} + \left\|\!\left|\widetilde{F}\right|\!\right\| \le \sqrt[p]{1+\frac{\alpha^p}{(\alpha+\delta)^p}} \sqrt[q]{\left\|\!\left|\widetilde{E}\right|\!\right\|^q + \left\|\!\left|\widetilde{F}\right|\!\right\|^q}.\end{aligned}$$

By equation (9.5), we deduce that

$$\left(1-\frac{\alpha}{\alpha+\delta}\right) |\!|\!| X |\!|\!| \le \sqrt[p]{1+\frac{\alpha^p}{(\alpha+\delta)^p}} \sqrt[q]{\left\|\!\left|\widetilde{E}\right|\!\right\|^q + \left\|\!\left|\widetilde{F}\right|\!\right\|^q}$$

which produce that if $|\!|\!| X |\!|\!| \ge |\!|\!| Y |\!|\!|$, $|\!|\!| X |\!|\!| \le \frac{1}{\underline{\eta}_p} \sqrt[q]{\left\|\!\left|\widetilde{E}\right|\!\right\|^q + \left\|\!\left|\widetilde{F}\right|\!\right\|^q}$. Similarly if $|\!|\!| X |\!|\!| < |\!|\!| Y |\!|\!|$, from $\Omega X - Y, \; = \Omega E + F$, we can obtain $|\!|\!| Y |\!|\!| \le \frac{1}{\underline{\eta}_p} \sqrt[q]{|\!|\!| E |\!|\!|^q + |\!|\!| F |\!|\!|^q}$. Inequality (9.4) now follows. ∎

Proof of Theorem 4.6: By equations (9.1) and (9.2) and Lemma 9.1, we have

$$\begin{aligned} &\max\left\{ \left\|\!\left|\widetilde{U}_2^* U_1\right|\!\right\|, \left\|\!\left|\widetilde{V}_2^* V_1\right|\!\right\| \right\} \\ &\le \frac{1}{\underline{\eta}_p} \max\left\{ \sqrt[q]{\left\|\!\left|\widetilde{V}_2^*(I-D_2^{-1})V_1\right|\!\right\|^q + \left\|\!\left|\widetilde{U}_2^*(D_1^*-I)U_1\right|\!\right\|^q}, \right. \\ &\qquad\qquad \left. \sqrt[q]{\left\|\!\left|\widetilde{U}_2^*(I-D_1^{-1})U_1\right|\!\right\|^q + \left\|\!\left|\widetilde{V}_2^*(D_2^*-I)V_1\right|\!\right\|^q} \right\} \\ &\le \frac{1}{\underline{\eta}_p} \max\left\{ \sqrt[q]{|\!|\!|(I-D_2^{-1})V_1|\!|\!|^q + |\!|\!|(D_1^*-I)U_1|\!|\!|^q}, \sqrt[q]{|\!|\!|(I-D_1^{-1})U_1|\!|\!|^q + |\!|\!|(D_2^*-I)V_1|\!|\!|^q} \right\}, \end{aligned}$$

as required. Turning to inequality (4.8), we have by equation (9.3) and Lemma 7.2 that

$$\begin{aligned} \left\|\!\left|\begin{pmatrix} \widetilde{U}_2^* U_1 & \\ & \widetilde{V}_2^* V_1 \end{pmatrix}\right|\!\right\| \le \frac{1}{\underline{\eta}_p} \left(\left\|\!\left|\begin{pmatrix} \widetilde{U}_2^*(I-D_1^{-1})U_1 & \\ & \widetilde{V}_2^*(I-D_2^{-1})V_1 \end{pmatrix}\right|\!\right\|^q \right. \\ \left. + \left\|\!\left|\begin{pmatrix} \widetilde{U}_2^*(D_1^*-I)U_1 & \\ & \widetilde{V}_2^*(D_2^*-I)V_1 \end{pmatrix}\right|\!\right\|^q \right)^{1/q}, \end{aligned} \tag{9.6}$$

since the conditions of Theorem 4.6 imply

$$\left\|\begin{pmatrix} & \widetilde{\Sigma}_2 \\ \widetilde{\Sigma}_2 & \end{pmatrix}\right\|_2 \le \alpha, \qquad \left\|\begin{pmatrix} & \Sigma_1 \\ \Sigma_1 & \end{pmatrix}^{-1}\right\|_2 \le \frac{1}{\alpha+\delta}.$$

Since $\widetilde{U}_2^* U_1$ and $\sin\Theta(U_1,\widetilde{U}_1)$ have the same nonzero singular values and so do $\widetilde{V}_2^* V_1$ and $\sin\Theta(V_1,\widetilde{V}_1)$,

$$\left\|\!\left|\begin{pmatrix} \widetilde{U}_2^* U_1 & \\ & \widetilde{V}_2^* V_1 \end{pmatrix}\right|\!\right\| = \left\|\!\left|\begin{pmatrix} \sin\Theta(U_1,\widetilde{U}_1) & \\ & \sin\Theta(V_1,\widetilde{V}_1) \end{pmatrix}\right|\!\right\|. \tag{9.7}$$

Note also

$$\begin{pmatrix} \widetilde{U}_2^*(I-D_1^{-1})U_1 & \\ & \widetilde{V}_2^*(I-D_2^{-1})V_1 \end{pmatrix} = \begin{pmatrix} \widetilde{U}_2^* & \\ & \widetilde{V}_2^* \end{pmatrix} \begin{pmatrix} (I-D_1^{-1})U_1 & \\ & (I-D_2^{-1})V_1 \end{pmatrix},$$

$$\begin{pmatrix} \widetilde{U}_2^*(D_1^*-I)U_1 & \\ & \widetilde{V}_2^*(D_2^*-I)V_1 \end{pmatrix} = \begin{pmatrix} \widetilde{U}_2^* & \\ & \widetilde{V}_2^* \end{pmatrix} \begin{pmatrix} (D_1^*-I)U_1 & \\ & (D_2^*-I)V_1 \end{pmatrix}.$$

Thus, one has

$$\left|\!\left|\!\left| \begin{pmatrix} \widetilde{U}_2^*(I-D_1^{-1})U_1 & \\ & \widetilde{V}_2^*(I-D_2^{-1})V_1 \end{pmatrix} \right|\!\right|\!\right| \le \left|\!\left|\!\left| \begin{pmatrix} (I-D_1^{-1})U_1 & \\ & (I-D_2^{-1})V_1 \end{pmatrix} \right|\!\right|\!\right|, \tag{9.8}$$

$$\left|\!\left|\!\left| \begin{pmatrix} \widetilde{U}_2^*(D_1^*-I)U_1 & \\ & \widetilde{V}_2^*(D_2^*-I)V_1 \end{pmatrix} \right|\!\right|\!\right| \le \left|\!\left|\!\left| \begin{pmatrix} (D_1^*-I)U_1 & \\ & (D_2^*-I)V_1 \end{pmatrix} \right|\!\right|\!\right|. \tag{9.9}$$

Inequality (4.8) is a consequence of (9.6), (9.7), (9.8) and (9.9). ∎

9.2 The Non-Square Case: $m > n$

Augment B and $\widetilde{B}$ by a zero block $0_{m,m-n}$ to $B_{\rm a} = (B, 0_{m,m-n})$ and $\widetilde{B}_{\rm a} = (\widetilde{B}, 0_{m,m-n})$. From $\widetilde{B} = D_1^* B D_2$, we get

$$\widetilde{B}_{\rm a} = D_1^* B_{\rm a} \begin{pmatrix} D_2 & \\ & I_{m-n} \end{pmatrix} \stackrel{\rm def}{=} D_1^* B_{\rm a} D_{\rm a2}.$$

From the SVDs (3.4) and (3.5) of B and $\widetilde{B}$, one can calculate the SVDs of $B_{\rm a}$ and $\widetilde{B}_{\rm a}$:

$$B_{\rm a} = U\Sigma_{\rm a} V_{\rm a}^* = (U_1, U_2) \begin{pmatrix} \Sigma_1 & \\ & \Sigma_{\rm a2} \end{pmatrix} \begin{pmatrix} V_{\rm a1}^* \\ V_{\rm a2}^* \end{pmatrix}, \tag{9.10}$$

$$\widetilde{B}_{\rm a} = \widetilde{U}\widetilde{\Sigma}_{\rm a} \widetilde{V}_{\rm a}^* = (\widetilde{U}_1, \widetilde{U}_2) \begin{pmatrix} \widetilde{\Sigma}_1 & \\ & \widetilde{\Sigma}_{\rm a2} \end{pmatrix} \begin{pmatrix} \widetilde{V}_{\rm a1}^* \\ \widetilde{V}_{\rm a2}^* \end{pmatrix}, \tag{9.11}$$

where

$$\Sigma_{\rm a2} = \begin{pmatrix} \Sigma_2 & \\ & 0_{m-n,m-n} \end{pmatrix}, V_{\rm a1} = \begin{pmatrix} V_1 \\ 0_{m-n,k} \end{pmatrix}, V_{\rm a2} = \begin{pmatrix} V_2 & \\ & I_{m-n} \end{pmatrix},$$

similarly for $\widetilde{\Sigma}_{\rm a2}$, $\widetilde{V}_{\rm a1}$ and $\widetilde{V}_{\rm a2}$. The following fact is easy to establish

$$\left|\!\left|\!\left| \sin\Theta(V_{\rm a1}, \widetilde{V}_{\rm a1}) \right|\!\right|\!\right| = \left|\!\left|\!\left| \sin\Theta(V_1, \widetilde{V}_1) \right|\!\right|\!\right|.$$

Applying the square case of Theorems 4.5 and 4.6 to $m \times m$ matrices $B_{\rm a}$ and $\widetilde{B}_{\rm a}$ just defined will complete the proofs.

10 Proofs of Theorems 4.7 and 4.8

We have seen in §9 how to deal with the nonsquare case by transforming it to the square case. So here we will only give proofs for the square case: $m = n$. Let $\hat{B} = D_1^* B$ and $\check{B} = BD_2$ and their SVDs be

$$\hat{B} \;=\; \hat{U}\hat{\Sigma}\hat{V}^* \equiv (\hat{U}_1, \hat{U}_2)\begin{pmatrix} \hat{\Sigma}_1 & \\ & \hat{\Sigma}_2 \end{pmatrix}\begin{pmatrix} \hat{V}_1^* \\ \hat{V}_2^* \end{pmatrix}, \tag{10.1}$$

$$\check{B} \;=\; \check{U}\check{\Sigma}\check{V}^* \equiv (\check{U}_1, \check{U}_2)\begin{pmatrix} \check{\Sigma}_1 & \\ & \check{\Sigma}_2 \end{pmatrix}\begin{pmatrix} \check{V}_1^* \\ \check{V}_2^* \end{pmatrix}, \tag{10.2}$$

where $\widetilde{U}$, $\check{U} \in \mathbf{U}_n$, $\widetilde{V}$, $\check{V} \in \mathbf{U}_n$, U_1, $\check{U}_1 \in \mathbf{C}^{n\times k}$, V_1, $\check{V}_1 \in \mathbf{C}^{n\times k}$ and

$$\hat{\Sigma}_1 = \mathrm{diag}(\hat{\sigma}_1, \cdots, \hat{\sigma}_k), \quad \hat{\Sigma}_2 = \mathrm{diag}(\hat{\sigma}_{k+1}, \cdots, \hat{\sigma}_n),$$
$$\check{\Sigma}_1 = \mathrm{diag}(\check{\sigma}_1, \cdots, \check{\sigma}_k), \quad \check{\Sigma}_2 = \mathrm{diag}(\check{\sigma}_{k+1}, \cdots, \check{\sigma}_n).$$

Partitionings in (10.1) and (10.2) shall be done in such a way that

$$\begin{aligned} &\max_{1\le i\le n} \chi(\sigma_i, \hat{\sigma}_i) \le \tfrac{1}{2}\|D_1^* - D_1^{-1}\|_2, \qquad \max_{1\le i\le n} \chi(\hat{\sigma}_i, \widetilde{\sigma}_i) \le \tfrac{1}{2}\|D_2^* - D_2^{-1}\|_2, \\ &\max_{1\le i\le n} \chi(\sigma_i, \check{\sigma}_i) \le \tfrac{1}{2}\|D_2^* - D_2^{-1}\|_2, \qquad \max_{1\le i\le n} \chi(\check{\sigma}_i, \widetilde{\sigma}_i) \le \tfrac{1}{2}\|D_1^* - D_1^{-1}\|_2. \end{aligned} \tag{10.3}$$

Such partitionings are possible because of the relative perturbation theorems proved in Li [11]. By the fact $\varrho_p(\xi, \zeta) \le 2^{-1/p}\chi(\xi, \zeta)$ (see Lemma 2.1 below), these inequalities imply

$$\begin{aligned} &\max_{1\le i\le n} \varrho_p(\sigma_i, \hat{\sigma}_i) \le \tfrac{1}{2^{1+1/p}}\|D_1^* - D_1^{-1}\|_2, \quad \max_{1\le i\le n} \varrho_p(\hat{\sigma}_i, \widetilde{\sigma}_i) \le \tfrac{1}{2^{1+1/p}}\|D_2^* - D_2^{-1}\|_2, \\ &\max_{1\le i\le n} \varrho_p(\sigma_i, \check{\sigma}_i) \le \tfrac{1}{2^{1+1/p}}\|D_2^* - D_2^{-1}\|_2, \quad \max_{1\le i\le n} \varrho_p(\check{\sigma}_i, \widetilde{\sigma}_i) \le \tfrac{1}{2^{1+1/p}}\|D_1^* - D_1^{-1}\|_2. \end{aligned} \tag{10.4}$$

Consider $\hat{B}$ and $\widetilde{B} = \hat{B}D_2$. We have

$$\hat{B}\hat{B}^* \;=\; \hat{U}\hat{\Sigma}\hat{\Sigma}^*\hat{U}^* \equiv (\hat{U}_1, \hat{U}_2)\begin{pmatrix} \hat{\Sigma}_1^2 & \\ & \hat{\Sigma}_2^2 \end{pmatrix}\begin{pmatrix} \hat{U}_1^* \\ \hat{U}_2^* \end{pmatrix}, \tag{10.5}$$

$$\widetilde{B}\widetilde{B}^* \;=\; \widetilde{U}\widetilde{\Sigma}\widetilde{\Sigma}^*\widetilde{U}^* \equiv (\widetilde{U}_1, \widetilde{U}_2)\begin{pmatrix} \widetilde{\Sigma}_1^2 & \\ & \widetilde{\Sigma}_2^2 \end{pmatrix}\begin{pmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{pmatrix}. \tag{10.6}$$

Notice that

$$\widetilde{B}\widetilde{B}^* - \hat{B}\hat{B}^* = \widetilde{B}D_2^*\hat{B}^* - \widetilde{B}D_2^{-1}\hat{B}^* = \widetilde{B}(D_2^* - D_2^{-1})\hat{B}^*,$$

Pre- and post-multiply the equations by $\widetilde{U}^*$ and $\hat{U}$, respectively, to get $\widetilde{\Sigma}^2\widetilde{U}^*\hat{U} - \widetilde{U}^*\hat{U}\hat{\Sigma}^2 = \widetilde{\Sigma}\widetilde{V}^*(D_2^* - D_2^{-1})\hat{V}\hat{\Sigma}$ which gives

$$\widetilde{\Sigma}_2^2\widetilde{U}_2^*\hat{U}_1 - \widetilde{U}_2^*\hat{U}_1\hat{\Sigma}_1^2 = \widetilde{\Sigma}_2\widetilde{V}_2^*(D_2^* - D_2^{-1})\hat{V}_1\hat{\Sigma}_1. \tag{10.7}$$

Consider now $\check{B}$ and $\widetilde{B} = D_1^* \check{B}$. We have

$$\check{B}^* \check{B} = \check{V} \check{\Sigma}^* \check{\Sigma} \check{V}^* \equiv (\check{V}_1, \check{V}_2) \begin{pmatrix} \check{\Sigma}_1^2 & \\ & \check{\Sigma}_2^2 \end{pmatrix} \begin{pmatrix} \check{V}_1^* \\ \check{V}_2^* \end{pmatrix}, \tag{10.8}$$

$$\widetilde{B}^* \widetilde{B} = \widetilde{V} \widetilde{\Sigma}^* \widetilde{\Sigma} \widetilde{V}^* \equiv (\widetilde{V}_1, \widetilde{V}_2) \begin{pmatrix} \widetilde{\Sigma}_1^2 & \\ & \widetilde{\Sigma}_2^2 \end{pmatrix} \begin{pmatrix} \widetilde{V}_1^* \\ \widetilde{V}_2^* \end{pmatrix}. \tag{10.9}$$

Notice that

$$\widetilde{B}^* \widetilde{B} - \check{B}^* \check{B} = \widetilde{B}^* D_1^* \check{B} - \widetilde{B}^* D_1^{-1} \check{B} = \widetilde{B}^* (D_1^* - D_1^{-1}) \check{B}.$$

Pre- and post-multiply the equations by $\widetilde{V}^*$ and $\check{V}$, respectively, to get $\widetilde{\Sigma}^2 \widetilde{V}^* \check{V} - \widetilde{V}^* \check{V} \check{\Sigma}^2 = \widetilde{\Sigma} \widetilde{U}^* (D_1^* - D_1^{-1}) \check{U} \check{\Sigma}$ which gives

$$\widetilde{\Sigma}_2^2 \widetilde{V}_2^* \check{V}_1 - \widetilde{V}_2^* \check{V}_1 \check{\Sigma}_1^2 = \widetilde{\Sigma}_2 \widetilde{U}_2^* (D_1^* - D_1^{-1}) \check{U}_1 \check{\Sigma}_1. \tag{10.10}$$

Two other eigendecompositions that will be used later in the proofs are

$$B B^* = U \Sigma \Sigma^* U^* \equiv (U_1, U_2) \begin{pmatrix} \Sigma_1^2 & \\ & \Sigma_2^2 \end{pmatrix} \begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix}, \tag{10.11}$$

$$B^* B = V \Sigma^* \Sigma V^* \equiv (V_1, V_2) \begin{pmatrix} \Sigma_1^2 & \\ & \Sigma_2^2 \end{pmatrix} \begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix}. \tag{10.12}$$

Proof of Theorem 4.7: Equations (10.7) and (10.10) and Lemma 8.1 produce

$$\left\| \sin \Theta(\hat{U}_1, \widetilde{U}_1) \right\|_{\mathrm{F}} = \left\| \widetilde{U}_2^* \hat{U}_1 \right\|_{\mathrm{F}} \leq \frac{\left\| \widetilde{V}_2^* (D_2^* - D_2^{-1}) \hat{V}_1 \right\|_{\mathrm{F}}}{\eta_\chi (\hat{\Sigma}_1^2, \widetilde{\Sigma}_2^2)},$$

$$\left\| \sin \Theta(\check{V}_1, \widetilde{V}_1) \right\|_{\mathrm{F}} = \left\| \widetilde{V}_2^* \check{V}_1 \right\|_{\mathrm{F}} \leq \frac{\left\| \widetilde{U}_2^* (D_1^* - D_1^{-1}) \check{U}_1 \right\|_{\mathrm{F}}}{\eta_\chi (\check{\Sigma}_1^2, \widetilde{\Sigma}_2^2)},$$

where[4] $\eta_\chi(\hat{\Sigma}_1^2, \widetilde{\Sigma}_2^2) \stackrel{\mathrm{def}}{=} \min_{1 \leq i \leq k, 1 \leq j \leq n-k} \chi(\hat{\sigma}_i^2, \widetilde{\sigma}_{k+j}^2)$ and $\eta_\chi(\check{\Sigma}_1^2, \widetilde{\Sigma}_2^2) \stackrel{\mathrm{def}}{=} \min_{1 \leq i \leq k, 1 \leq j \leq n-k} \chi(\check{\sigma}_i^2, \widetilde{\sigma}_{k+j}^2)$. On the other hand, applying Theorem 4.1 to BB^* and $\hat{B}\hat{B}^* = D_1^* B B^* D_1$ leads to (see (10.5) and (10.11))

$$\left\| \sin \Theta(U_1, \hat{U}_1) \right\|_{\mathrm{F}} \leq \frac{\sqrt{\|(I - D_1^{-1}) U_1\|_{\mathrm{F}}^2 + \|(I - D_1^*) U_1\|_{\mathrm{F}}^2}}{\eta_2(\Sigma_1^2, \hat{\Sigma}_2^2)},$$

where $\eta_2(\Sigma_1^2, \hat{\Sigma}_2^2) \stackrel{\mathrm{def}}{=} \min_{1 \leq i \leq k, 1 \leq j \leq n-k} \varrho_2(\sigma_i^2, \hat{\sigma}_{k+j}^2)$; Applying Theorem 4.1 to $B^* B$ and $\check{B}^* \check{B} = D_2^* B^* B D_2$ leads to (see (10.8) and (10.12))

$$\left\| \sin \Theta(V_1, \check{V}_1) \right\|_{\mathrm{F}} \leq \frac{\sqrt{\|(I - D_2^{-1}) V_1\|_{\mathrm{F}}^2 + \|(I - D_2^*) V_1\|_{\mathrm{F}}^2}}{\eta_2(\Sigma_1^2, \check{\Sigma}_2^2)},$$

[4]We abuse notation η_χ here for convenience. As we recall, η_χ has its own assignment in the statement of Theorem 4.7. However, it is re-defined as a function in this proof. Hopefully, this would not cause any confusion.

where $\eta_2(\Sigma_1^2, \check{\Sigma}_2^2) \stackrel{\rm def}{=} \min_{1\le i\le k, 1\le j\le n-k} \varrho_2(\sigma_i^2, \check{\sigma}_{k+j}^2)$; Let $\epsilon_1 = \|D_1^* - D_1^{-1}\|_2$ and $\epsilon_2 = \|D_2^* - D_2^{-1}\|_2$. We claim

$$\eta_\chi(\hat{\Sigma}_1^2, \widetilde{\Sigma}_2^2) \ge 2\,\eta_\chi(\hat{\Sigma}_1, \widetilde{\Sigma}_2) \ge \begin{cases} \frac{2\eta_\chi(\Sigma_1,\widetilde{\Sigma}_2)-\epsilon_1}{1+\frac{\epsilon_1}{16}\eta_\chi(\Sigma_1,\widetilde{\Sigma}_2)}, \\ 2^{3/2}\eta_2(\Sigma_1, \widetilde{\Sigma}_2) - \epsilon_1. \end{cases} \tag{10.13}$$

This is because

$$\begin{aligned} \chi(\hat{\sigma}_i^2, \widetilde{\sigma}_{k+j}^2) \ge 2\chi(\hat{\sigma}_i, \widetilde{\sigma}_{k+j}) &\ge 2^{1+1/2}\varrho_2(\hat{\sigma}_i, \widetilde{\sigma}_{k+j}) && \text{(by Lemma 2.1)} \\ &\ge 2^{3/2}\left[\varrho_2(\sigma_i, \widetilde{\sigma}_{k+j}) - \varrho_2(\hat{\sigma}_i, \sigma_i)\right] && (\varrho_2 \text{ is a metric on } \mathbf{R}) \\ &\ge 2^{3/2}\left[\varrho_2(\sigma_i, \widetilde{\sigma}_{k+j}) - 2^{-3/2}\epsilon_1\right] && \text{(by (10.4)} \end{aligned}$$

and because

$$\chi(\sigma_i, \widetilde{\sigma}_{k+j}) \le \chi(\sigma_i, \hat{\sigma}_i) + \chi(\hat{\sigma}_i, \widetilde{\sigma}_{k+j}) + \frac{1}{8}\chi(\sigma_i, \hat{\sigma}_i)\chi(\hat{\sigma}_i, \widetilde{\sigma}_{k+j})\chi(\sigma_i, \widetilde{\sigma}_{k+j}).$$

$$\Rightarrow \quad \chi(\hat{\sigma}_i, \widetilde{\sigma}_{k+j}) \ge \frac{\chi(\sigma_i, \widetilde{\sigma}_{k+j}) - \chi(\sigma_i, \hat{\sigma}_i)}{1 + \frac{1}{8}\chi(\sigma_i, \hat{\sigma}_i)\chi(\sigma_i, \widetilde{\sigma}_{k+j})} \ge \frac{\chi(\sigma_i, \widetilde{\sigma}_{k+j}) - \epsilon_1/2}{1 + \frac{\epsilon_1}{16}\chi(\sigma_i, \widetilde{\sigma}_{k+j})}. \qquad \text{(by (10.4)}$$

Similarly, we have

$$\begin{aligned} \eta_\chi(\check{\Sigma}_1^2, \widetilde{\Sigma}_2^2) &\ge 2\,\eta_\chi(\check{\Sigma}_1, \widetilde{\Sigma}_2) \ge \begin{cases} \frac{2\eta_\chi(\Sigma_1,\widetilde{\Sigma}_2)-\epsilon_2}{1+\frac{\epsilon_2}{16}\eta_\chi(\Sigma_1,\widetilde{\Sigma}_2)}, \\ 2^{3/2}\eta_2(\Sigma_1, \widetilde{\Sigma}_2) - \epsilon_2, \end{cases} \\ \eta_2(\Sigma_1^2, \hat{\Sigma}_2^2) &\ge \eta_2(\Sigma_1, \hat{\Sigma}_2) \ge \eta_2(\Sigma_1, \widetilde{\Sigma}_2) - 2^{-3/2}\epsilon_2, \\ \eta_2(\Sigma_1^2, \check{\Sigma}_2^2) &\ge \eta_2(\Sigma_1, \check{\Sigma}_2) \ge \eta_2(\Sigma_1, \widetilde{\Sigma}_2) - 2^{-3/2}\epsilon_1. \end{aligned}$$

The proof will be completed by employing

$$\begin{aligned} \left\|\sin\Theta(U_1, \widetilde{U}_1)\right\|_{\rm F} &\le \left\|\sin\Theta(U_1, \hat{U}_1)\right\|_{\rm F} + \left\|\sin\Theta(\hat{U}_1, \widetilde{U}_1)\right\|_{\rm F}, \\ \left\|\sin\Theta(V_1, \widetilde{V}_1)\right\|_{\rm F} &\le \left\|\sin\Theta(V_1, \check{V}_1)\right\|_{\rm F} + \left\|\sin\Theta(\check{V}_1, \widetilde{V}_1)\right\|_{\rm F}, \end{aligned}$$

since $\|\sin\Theta(\cdot\,,\cdot)\|_{\rm F}$ is a metric on the space of k-dimensional subspaces [14]. ∎

Proof of Theorems 4.8: Denote $\beta = \alpha + \delta$. Let $\hat{\alpha}$ and $\check{\alpha}$ be the largest positive numbers such that

$$\chi(\alpha, \hat{\alpha}) \le \frac{1}{2}\|D_2^* - D_2^{-1}\|_2 \quad \text{and} \quad \chi(\alpha, \check{\alpha}) \le \frac{1}{2}\|D_1^* - D_1^{-1}\|_2$$

which guarantee that $\|\hat{\Sigma}_2\|_2 \le \hat{\alpha}$, $\|\check{\Sigma}_2\|_2 \le \check{\alpha}$, and

$$\varrho_p(\alpha, \hat{\alpha}) \le \frac{1}{2^{1+1/p}}\|D_2^* - D_2^{-1}\|_2 \quad \text{and} \quad \varrho_p(\alpha, \check{\alpha}) \le \frac{1}{2^{1+1/p}}\|D_1^* - D_1^{-1}\|_2;$$

and let $\hat{\beta}$ and $\check{\beta}$ be the smallest numbers such that

$$\chi(\beta,\hat{\beta}) \le \frac{1}{2}\|D_1^* - D_1^{-1}\|_2 \quad \text{and} \quad \chi(\beta,\check{\beta}) \le \frac{1}{2}\|D_2^* - D_2^{-1}\|_2$$

which guarantee that $\|\hat{\Sigma}_1^{-1}\|_2^{-1} \ge \hat{\beta}$, $\|\check{\Sigma}_1^{-1}\|_2^{-1} \ge \check{\beta}$, and

$$\varrho_p(\beta,\hat{\beta}) \le \frac{1}{2^{1+1/p}}\|D_1^* - D_1^{-1}\|_2 \quad \text{and} \quad \varrho_p(\beta,\check{\beta}) \le \frac{1}{2^{1+1/p}}\|D_2^* - D_2^{-1}\|_2.$$

(4.15) implies $\min\{\hat{\beta},\check{\beta}\} > \alpha$ and $\beta > \max\{\hat{\alpha},\check{\alpha}\}$.

Equations (10.7) and (10.10) and Lemma 8.2 produce

$$\begin{aligned} \left\|\!\left|\sin\Theta(\hat{U}_1,\widetilde{U}_1)\right|\!\right\| = \left\|\!\left|\widetilde{U}_2^*\hat{U}_1\right|\!\right\| &\le \frac{\left\|\!\left|\widetilde{V}_2^*(D_2^* - D_2^{-1})\hat{V}_1\right|\!\right\|}{\chi(\alpha^2,\hat{\beta}^2)}, \\ \left\|\!\left|\sin\Theta(\check{V}_1,\widetilde{V}_1)\right|\!\right\| = \left\|\!\left|\widetilde{V}_2^*\check{V}_1\right|\!\right\| &\le \frac{\left\|\!\left|\widetilde{U}_2^*(D_1^* - D_1^{-1})\check{U}_1\right|\!\right\|}{\chi(\alpha^2,\check{\beta}^2)}. \end{aligned}$$

On the other hand, applying Theorem 4.2 to BB^* and $\hat{B}\hat{B}^* = D_1^*BB^*D_1$ leads to (see (10.5) and (10.11))

$$\left\|\!\left|\sin\Theta(U_1,\hat{U}_1)\right|\!\right\| \le \frac{\sqrt[q]{\left\|\!\left|(I - D_1^{-1})U_1\right|\!\right\|^q + \|\!|(I - D_1^*)U_1|\!\|^q}}{\chi(\hat{\alpha}^2,\beta^2)};$$

Applying Theorem 4.1 to B^*B and $\check{B}^*\check{B} = D_2^*B^*BD_2$ leads to (see (10.8) and (10.12))

$$\left\|\!\left|\sin\Theta(V_1,\check{V}_1)\right|\!\right\| \le \frac{\sqrt[q]{\left\|\!\left|(I - D_2^{-1})V_1\right|\!\right\|^q + \|\!|(I - D_2^*)V_1|\!\|^q}}{\chi(\check{\alpha}^2,\beta^2)}.$$

Notice that

$$\begin{aligned} \chi(\alpha^2,\hat{\beta}^2) &\ge 2\chi(\alpha,\hat{\beta}) \ge \begin{cases} \frac{2\chi(\alpha,\beta)-\epsilon_1}{1+\frac{\epsilon_1}{16}\chi(\alpha,\beta)}, \\ 2^{1+1/p}\varrho_p(\alpha,\beta) - \epsilon_1, \end{cases} \\ \chi(\alpha^2,\check{\beta}^2) &\ge 2\chi(\alpha,\check{\beta}) \ge \begin{cases} \frac{2\chi(\alpha,\beta)-\epsilon_2}{1+\frac{\epsilon_2}{16}\chi(\alpha,\beta)}, \\ 2^{1+1/p}\varrho_p(\alpha,\beta) - \epsilon_2, \end{cases} \\ \varrho_p(\hat{\alpha}^2,\beta^2) &\ge \varrho_p(\hat{\alpha},\beta) \ge \varrho_p(\alpha,\beta) - 2^{1+1/p}\epsilon_2, \\ \varrho_p(\check{\alpha}^2,\beta^2) &\ge \varrho_p(\check{\alpha},\beta) \ge \varrho_p(\alpha,\beta) - 2^{1+1/p}\epsilon_1, \end{aligned}$$

where $\epsilon_1 = \|D_1^* - D_1^{-1}\|_2$ and $\epsilon_2 = \|D_2^* - D_2^{-1}\|_2$. The proof will be completed by employing

$$\begin{aligned} \left\|\!\left|\sin\Theta(U_1,\widetilde{U}_1)\right|\!\right\| &\le \left\|\!\left|\sin\Theta(U_1,\hat{U}_1)\right|\!\right\| + \left\|\!\left|\sin\Theta(\hat{U}_1,\widetilde{U}_1)\right|\!\right\|, \\ \left\|\!\left|\sin\Theta(V_1,\widetilde{V}_1)\right|\!\right\| &\le \left\|\!\left|\sin\Theta(V_1,\check{V}_1)\right|\!\right\| + \left\|\!\left|\sin\Theta(\check{V}_1,\widetilde{V}_1)\right|\!\right\|. \end{aligned}$$

since $\|\!|\sin\Theta(\,\cdot\,,\,\cdot\,)|\!\|$ is a metric on the space of k-dimensional subspaces [14]. ∎

11 Conclusions and Further Extensions to Diagonalizable Matrices

We have developed a relative perturbation theory for eigenspace and singular space variations under multiplicative perturbations. In the theory, extensions of Davis-Kahan $\sin\theta$ theorems and Wedin $\sin\theta$ theorems from the classical perturbation theory are made. Our unifying treatment covers almost all previously studied cases over the last six years or so.

Using the similar technique in this paper, one can also develop a relative perturbation theory for eigenspaces of diagonalizable matrices: A and $\widetilde{A} = D_1^* A D_2$ are diagonalizable, where D_1 and D_2 are close to the identity matrix. We outline a way of doing this. Let eigendecompositions of A and $\widetilde{A}$ be

$$AX = X\Lambda \equiv (X_1, X_2)\begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} \quad \text{and} \quad \widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{\Lambda} \equiv (\widetilde{X}_1, \widetilde{X}_2)\begin{pmatrix} \widetilde{\Lambda}_1 & \\ & \widetilde{\Lambda}_2 \end{pmatrix},$$

where X, $\widetilde{X} \in \mathbf{C}^{n\times n}$ are nonsingular, and X_1, $\widetilde{X}_1 \in \mathbf{C}^{n\times k}$ $(1 \le k < n)$ and

$$\begin{aligned} \Lambda_1 &= \mathrm{diag}(\lambda_1, \cdots, \lambda_k), & \Lambda_2 &= \mathrm{diag}(\lambda_{k+1}, \cdots, \lambda_n), \\ \widetilde{\Lambda}_1 &= \mathrm{diag}(\widetilde{\lambda}_1, \cdots, \widetilde{\lambda}_k), & \widetilde{\Lambda}_2 &= \mathrm{diag}(\widetilde{\lambda}_{k+1}, \cdots, \widetilde{\lambda}_n). \end{aligned}$$

λ_i's and $\widetilde{\lambda}_j$'s may be complex. Partition

$$X^{-1} = \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix} \quad \text{and} \quad \widetilde{X}^{-1} = \begin{pmatrix} \widetilde{Y}_1^* \\ \widetilde{Y}_2^* \end{pmatrix},$$

where Y_1, $\widetilde{Y}_1 \in \mathbf{C}^{n\times k}$. Define $R \stackrel{\text{def}}{=} \widetilde{A}X_1 - X_1\Lambda_1 = (\widetilde{A} - A)X_1$. We have

$$\begin{aligned} \widetilde{Y}_2^* R &= \widetilde{Y}_2^* \widetilde{A} X_1 - \widetilde{Y}_2^* X_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{Y}_2^* X_1 - \widetilde{Y}_2^* X_1 \Lambda_1, \\ \widetilde{Y}_2^* R &= \widetilde{Y}_2^* (\widetilde{A} - A) X_1 = \widetilde{Y}_2^* (D_1^* A D_2 - D_1^* A + D_1^* A - A) X_1 \\ &= \widetilde{Y}_2^* \left[\widetilde{A}(I - D_2^{-1}) + (D_1^* - I)A \right] X_1 \\ &= \widetilde{\Lambda}_2 \widetilde{Y}_2^* (I - D_2^{-1}) X_1 + \widetilde{Y}_2^* (D_1^* - I) X_1 \Lambda_1. \end{aligned}$$

Thus we have the following perturbation equation

$$\widetilde{\Lambda}_2 \widetilde{Y}_2^* X_1 - \widetilde{Y}_2^* X_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{Y}_2^* (I - D_2^{-1}) X_1 + \widetilde{Y}_2^* (D_1^* - I) X_1 \Lambda_1$$

from which various bounds on $\sin\Theta(X_1, \widetilde{X}_1)$ can be derived under certain conditions. For example, let $\eta_2 \stackrel{\text{def}}{=} \min_{1\le i\le k,\, 1\le j\le n-k} \varrho_2(\lambda_i, \widetilde{\lambda}_{k+j})$. If $\eta_2 > 0$, then by Lemma 7.1 we have

$$\begin{aligned} \|\widetilde{Y}_2^* X_1\|_{\mathrm{F}} &\le \frac{1}{\eta_2}\sqrt{\|\widetilde{Y}_2^*(I - D_2^{-1})X_1\|_{\mathrm{F}}^2 + \|\widetilde{Y}_2^*(D_1^* - I)X_1\|_{\mathrm{F}}^2} \\ &\le \frac{1}{\eta_2}\|\widetilde{Y}_2^*\|_2 \|X_1\|_2 \sqrt{\|I - D_2^{-1}\|_{\mathrm{F}}^2 + \|D_1^* - I\|_{\mathrm{F}}^2}. \end{aligned}$$

Notice that by Lemma 2.2

$$\begin{aligned}\|\sin\Theta(X_1,\widetilde{X}_1)\|_{\rm F} &= \|(\widetilde{Y}_2^*\widetilde{Y}_2)^{-1/2}\widetilde{Y}_2^*X_1(X_1^*X_1)^{-1/2}\|_{\rm F}\\ &\le \|(\widetilde{Y}_2^*\widetilde{Y}_2)^{-1/2}\|_2\|\widetilde{Y}_2^*X_1\|_{\rm F}\|(X_1^*X_1)^{-1/2}\|_2.\end{aligned}$$

Then a bound on $\|\sin\Theta(X_1,\widetilde{X}_1)\|_{\rm F}$ is immediately available.

Acknowledgement: I thank Professor W. Kahan for his consistent encouragement and support, Professor J. Demmel and Professor B. N. Parlett for helpful discussions.

References

[1] J. Barlow and J. Demmel. Computing accurate eigensystems of scaled diagonally dominant matrices. *SIAM Journal on Numerical Analysis*, 27:762–791, 1990.

[2] C. Davis and W. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7:1–46, 1970.

[3] J. Demmel and W. Gragg. On computing accurate singular values and eigenvalues of matrices with acyclic graphs. *Linear Algebra and its Application*, 185:203–217, 1993.

[4] J. Demmel and W. Kahan. Accurate singular values of bidiagonal matrices. *SIAM Journal on Scientific and Statistical Computing*, 11:873–912, 1990.

[5] J. Demmel and K. Veselić. Jacobi's method is more accurate than QR. *SIAM Journal on Matrix Analysis and Applications*, 13(4):1204–1245, 1992.

[6] S. C. Eisenstat and I. C. F. Ipsen. Relative perturbation techniques for singular value problems. Research Report YALEU/DCS/RR–942, Department of Computer Science, Yale University, 1993.

[7] S. C. Eisenstat and I. C. F. Ipsen. Relative perturbation bounds for eigenspaces and singular vector subspaces. In J. G. Lewis, editor, *Proceedings of the Fifth SIAM Conference on Applied Linear Algebra*, pages 62–66, Philadelphia, 1994. SIAM Publications.

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

[9] W. Kahan. Accurate eigenvalues of a symmetric tridiagonal matrix. Technical Report CS41, Computer Science Department, Stanford University, Stanford, CA, 1966. (revised June 1968).

[10] R.-C. Li. On perturbations of matrix pencils with real spectra. *Mathematics of Computation*, 62:231–265, 1994.

[11] R.-C. Li. Relative perturbation theory: (i) eigenvalue and singular value variations. Technical Report UCB//CSD-94-855, Computer Science Division, Department of EECS, University of California at Berkeley, 1994. (revised January 1996).

[12] R. Mathias. Spectral perturbation bounds for graded positive definite matrices. Manuscript, Department of Mathematics, College of William & Mary, 1994.

[13] A. M. Ostrowski. A quantitative formulation of Sylvester's law of inertia. *Proc. National Acad. Sciences (USA)*, 45:740–744, 1959.

[14] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.

[15] P.-Å. Wedin. Perturbation bounds in connection with singular value decomposition. *BIT*, 12:99–111, 1972.