University of Kentucky Lexington

Department of Mathematics Technical Report 2000-33

A $\sin 2\Theta$ Theorem for Graded Indefinite Hermitian Matrices[1]

Ninoslav Truhar[2]

Ren-Cang Li[3]

date and revision information go here

ABSTRACT

This paper gives double angle theorems that bound the change in an invariant subspace of an indefinite Hermitian matrix in the graded form $H = D^*AD$ subject to a perturbation $H \to \widetilde{H} = D^*(A + \Delta A)D$. These theorems extend recent results on a definite Hermitian matrix in the graded form (*Linear Algebra Appl.*, **311** (2000), 45–60) but the bounds here are more complicated in that they depend on not only relative gaps and norms of ΔA as in the definite case but also norms of so-called the hyperbolic eigenvector matrices of certain associated matrix pairs. For two special but interest cases, bounds on these hyperbolic eigenvector matrices are obtained to show that their norms are of moderate magnitude.

[1]This report is available on the web at `http://www.ms.uky.edu/~rcli/`.

[2](University Osijek, Croatia) Lehrgebiet Mathematische Physik, Fernuniversität, 58084 Hagen Germany (*Ninoslav.Truhar@Fernuni-Hagen.de*).

[3]Department of Mathematics, University of Kentucky, Lexington, KY 40506 (`rcli@ms.uky.edu`). This work was supported in part by the National Science Foundation under Grant No. ACI-9721388 and by the National Science Foundation CAREER award under Grant No. CCR-9875201.

A $\sin 2\Theta$ Theorem for Graded Indefinite Hermitian Matrices

Ninoslav Truhar* Ren-Cang Li†

September 2000

Abstract

This paper gives double angle theorems that bound the change in an invariant subspace of an indefinite Hermitian matrix in the graded form $H = D^*AD$ subject to a perturbation $H \to \widetilde{H} = D^*(A + \Delta A)D$. These theorems extend recent results on a definite Hermitian matrix in the graded form (*Linear Algebra Appl.*, **311** (2000), 45–60) but the bounds here are more complicated in that they depend on not only relative gaps and norms of ΔA as in the definite case but also norms of so-called the hyperbolic eigenvector matrices of certain associated matrix pairs. For two special but interest cases, bounds on these hyperbolic eigenvector matrices are obtained to show that their norms are of moderate magnitude.

1 Introduction

Let H and $\widetilde{H}$ be two Hermitian matrices whose eigendecompositions are

$$H = [\,U_1 \quad U_2\,] \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}, \quad \widetilde{H} = [\,\widetilde{U}_1 \quad \widetilde{U}_2\,] \begin{bmatrix} \widetilde{\Lambda}_1 & 0 \\ 0 & \widetilde{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{bmatrix}, \tag{1.1}$$

where $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$, $\widetilde{U} = \begin{bmatrix} \widetilde{U}_1 & \widetilde{U}_2 \end{bmatrix}$ are unitary, and

$$\Lambda_1 = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_k), \quad \Lambda_2 = \mathrm{diag}(\lambda_{k+1}, \lambda_{k+2}, \cdots, \lambda_n), \tag{1.2}$$

$$\widetilde{\Lambda}_1 = \mathrm{diag}(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \cdots, \widetilde{\lambda}_k), \quad \widetilde{\Lambda}_2 = \mathrm{diag}(\widetilde{\lambda}_{k+1}, \widetilde{\lambda}_{k+2}, \cdots, \widetilde{\lambda}_n). \tag{1.3}$$

We are interested in bounding the changes in subspace $\mathcal{S} \stackrel{\mathrm{def}}{=} \mathrm{span}(U_1)$, H's invariant subspace spanned by U_1's columns. We shall do this by bounding the sines of the double canonical angles between $\mathcal{S}$ and $\widetilde{\mathcal{S}} \stackrel{\mathrm{def}}{=} \mathrm{span}(\widetilde{U}_1)$. For absolute perturbations, i.e., $H \to H + \Delta H \equiv \widetilde{H}$, this is done by Davis and Kahan [1], and for multiplicative perturbations, i.e., $H \to D^*HD \equiv \widetilde{H}$, as well as for perturbations involving graded *definite* Hermitian matrices, this is done by [4]. The main results of this paper are extensions of a $\sin 2\Theta$ theorem for *positive definite* Hermitian matrices

*(University Osijek, Croatia) Lehrgebiet Mathematische Physik, Fernuniversität, 58084 Hagen Germany (*Ninoslav.Truhar@Fernuni-Hagen.de*)

†Department of Mathematics, University of Kentucky, Lexington, KY 40506 (*rcli@ms.uky.edu*). This work was supported in part by the National Science Foundation under Grant No. ACI-9721388 and by the National Science Foundation CAREER award under Grant No. CCR-9875201.

[4] to *indefinite graded* Hermitian matrices. An advantage of a double angle theorem over a single angle theorem is that a double angle theorem uses a relative gap involving the spectrum of only one matrix, either H or $\widetilde{H}$, in contrast to the gap used by single angle theorems.

Notation. $\|X\|$ and $\|X\|_{\mathrm{F}}$ denote the spectral and Frobenius norms of matrix X, respectively. X^* is the conjugate transpose. $\lambda(X)$ denotes the spectrum of X. I_n denotes the $n \times n$ identity matrix (we may simply write I instead if no confusion).

2 Hyperbolic Singular Value Decomposition (HSVD)

HSVD [5] provides an important tool in our later developments. Throughout this section, Z is $n \times n$ and nonsingular, and J is $n \times n$ and diagonal with ± 1 on its main diagonal.

Theorem 2.1 (HSVD) *There exist an $n \times n$ unitary matrix Y and an $n \times n$ non-singular matrix X such that*

$$Z = Y\,\Sigma\,JX^*J, \quad X^*JX = J, \tag{2.1}$$

where Σ is $n \times n$ and diagonal with positive real diagonal entries.

We call (2.1) the *HSVD of Z (with respect to J)*. HSVD theorem in its generality allows Z to be rectangular, but the square case is what we will need in this paper.

A matrix X is called *J-unitary* if $X^*JX = J$. It can be proved that *if X is J-unitary, so are X^*, X^{-1}, and X^{-*}*. In fact $X^*JX = J$ implies immediately $X^{-*}JX^{-1} = J$ and thus X^{-1} is J-unitary. Note also $X = JX^{-*}J$ to get

$$(X^*)^*JX^* = XJ\,X^* = JX^{-*}J\,JX^* = JX^{-*}X^* = J,$$

so X^* is J-unitary. Finally $X^{-*} = (X^*)^{-1}$ is J-unitary.

Let M be positive definite, and $J = \mathrm{diag}(\pm 1)$. The *hyperbolic eigenvector matrix* of a matrix pair $\{M, J\}$ is a matrix X which simultaneously diagonalizes the pair:

$$X^*MX = \text{diagonal}, \quad X^*JX = J.$$

Notice that $\{M, J\}$ is a definite matrix pair, and thus it always has a hyperbolic eigenvector matrix [7, p.318].

HSVD of Z (with respect to J) is closely related to the hyperbolic eigenvector matrix and the eigendecomposition of ZJZ^*. This is given in the following theorem.

Theorem 2.2 *In* Theorem 2.1*,*

1. *X is the hyperbolic eigenvector matrix of the pair $\{Z^*Z, J\}$.*

2. *$ZJZ^* = Y\,\Sigma J \Sigma\,Y^*$ is an eigendecomposition.*

On the other hand if $ZJZ^* = Y\Lambda Y^*$ is an eigendecomposition where Λ is diagonal and Y is unitary, there is a J-unitary matrix X such that Z's HSVD is given by (2.1) *where $J = \mathrm{diag}(\pm 1)$ such that $\Lambda = |\Lambda| J$.*

Proof: In Theorem 2.1, by (2.1) we have $X^*JX = J$ and also

$$\begin{aligned} X^*Z^*ZX &= \underbrace{X^*JX}J\Sigma\underbrace{Y^*Y}\Sigma J\underbrace{X^*JX} = JJ\Sigma\Sigma JJ = \Sigma^2, \\ ZJZ^* &= Y\Sigma JX^*J\,J\,JXJ\Sigma Y^* = Y\Sigma J\,X^*JX\,J\Sigma Y^* = Y\Sigma J\,J\,J\Sigma Y^* = Y\,\Sigma J\Sigma\,Y^*. \end{aligned}$$

This proves the first part of the theorem. Now we prove the second part. Suppose we have eigendecomposition $ZJZ^* = Y\Lambda Y^*$. It suffices to find a nonsingular X such that (2.1) holds. Let $\widetilde{X}$ be a hyperbolic eigenvector matrix of the definite matrix pair $\{Z^*Z, J\}$. Using $J = JJJ = J\widetilde{X}^*J\widetilde{X}J = J\widetilde{X}^*J\,J\,J\widetilde{X}J$ we can write

$$ZJZ^* = Y|\Lambda|^{1/2}\,J\,|\Lambda|^{1/2}Y^* = Y|\Lambda|^{1/2}J\widetilde{X}^*J\,J\,J\widetilde{X}J\,|\Lambda|^{1/2}Y^* \stackrel{\text{def}}{=} \widetilde{Z}J\widetilde{Z}, \tag{2.2}$$

where $\widetilde{Z} = Y|\Lambda|^{1/2}J\widetilde{X}^*J$. Set $W = \widetilde{Z}^{-1}Z$. (2.2) yields $WJW^* = J$. This says W^* is J-unitary and so is W, i.e., $W^*JW = J \quad \Rightarrow \quad JW = W^{-*}J$. Now

$$Z = \widetilde{Z}W = Y|\Lambda|^{1/2}J\widetilde{X}^*JW = Y|\Lambda|^{1/2}J\widetilde{X}^*W^{-*}J \equiv Y|\Lambda|^{1/2}JX^*J,$$

where $X = W^{-1}\widetilde{X}$. Since W^{-1} is J-unitary and the product of two J-unitary matrices is still J-unitary, X is J-unitary. ∎

3 Main Results

Let H and $\widetilde{H}$ and their eigen-decompositions be as described in §1. Similarly, as in [4] we shall define a unitary matrix

$$S = [\,U_1 \quad U_2\,]\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix}\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}. \tag{3.1}$$

Note that

$$S^* = S, \quad S^2 = I_n, \quad S^{-1} = S, \quad SHS = H.$$

We now define an auxiliary matrix $\widehat{H}$ as

$$\widehat{H} \equiv S\widetilde{H}S = [\,\widehat{U}_1 \quad \widehat{U}_2\,]\begin{bmatrix} \widetilde{\Lambda}_1 & \\ & \widetilde{\Lambda}_2 \end{bmatrix}\begin{bmatrix} \widehat{U}_1^* \\ \widehat{U}_2^* \end{bmatrix}, \tag{3.2}$$

where $\widehat{U}_i = S\widetilde{U}_i$ for $i = 1, 2$. Geometrically, $S\widetilde{\mathcal{S}}$ is a reflection of $\widetilde{\mathcal{S}}$ where the mirror for S is $\mathcal{S}$ and S reverses $\mathcal{S}^\perp$, the orthogonal complement of $\mathcal{S}$. This explains the following equation due to Davis and Kahan [1] (see also [4]):

$$\|\sin\Theta(\widetilde{U}_1, \widehat{U}_1)\|_{\mathrm{F}} = \|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}}. \tag{3.3}$$

In what follows we shall seek bounds for $\|\sin\Theta(\widetilde{U}_1, \widehat{U}_1)\|_{\mathrm{F}}$.

Since the positive semidefinite case has been studied by [4], interesting to us is the case when H is indefinite Hermitian in the graded form $H = D^*AD$, where D is nonsingular and its role here is to scale H so that A is well-conditioned, i.e., $\|A\|\,\|A^{-1}\|$ is of moderate magnitude. Suppose H is perturbed to

$$\widetilde{H} = D^*\widetilde{A}D \equiv D^*(A + \Delta A)D.$$

Let A's eigendecomposition be

$$A = Q\Omega Q^* = Q|\Omega|^{1/2}J|\Omega|^{1/2}Q^*, \tag{3.4}$$

where $J = \mathrm{diag}(J_k, J_{n-k})$ is diagonal and partitioned conformally to (1.1). It can be seen that the diagonal elements of J are the signs of the corresponding eigenvalues of A. Sylvester's theorem [2, Theorem 4.5.8] implies A, H, and J all have the same inertia. Set

$$\begin{aligned} G &= D^*Q|\Omega|^{1/2}, & (3.5)\\ E &= |\Omega|^{-1/2}Q^*\Omega A Q|\Omega|^{-1/2} & (3.6)\end{aligned}$$

to get

$$H = GJG^*, \quad \widetilde{H} = G(J+E)G^*.$$

We shall assume that $\|A^{-1}\|\|\Delta A\| < 1$ which insures $\|E\| \leq \|A^{-1}\|\|\Delta A\| < 1$ and the existence of $(I+EJ)^{1/2}$ defined by the following series [3, Theorem 6.2.8]

$$\begin{aligned} T &\stackrel{\text{def}}{=} (I+EJ)^{1/2} & (3.7)\\ &= I + \sum_{n=1}^{\infty}(-1)^{n-1}\frac{(2n-1)!!}{2^n n!}(EJ)^n, \end{aligned}$$

where $(2n-1)!! = 1 \cdot 1 \cdot 3 \cdot 5 \cdots (2n-1)$. It can be verified that $T = JT^*J$, and therefore

$$J + E = TJT^*,$$

$\|E\| < 1$ implies that $H = GJG^*$, $J+E$ and $\widetilde{H} = G(J+E)G^*$ all have the same inertia as J. From now on we'll *assume that* Λ_1 *and* $\widetilde{\Lambda}_1$ *have the same inertia*, and thus (1.1) can be re-written as

$$H = U|\Lambda|JU^*, \quad \widetilde{H} = \widetilde{U}|\widetilde{\Lambda}|J\widetilde{U}^* \tag{3.8}$$

(if necessary some re-ordering may be needed for the columns of U_i and of $\widetilde{U}_i$, and the diagonal entries of Λ_i and $\widetilde{\Lambda}_i$ *without* affecting the invariant subspaces $\mathrm{span}(U_i)$ and $\mathrm{span}(\widetilde{U}_i)$ in question), where $J = \mathrm{diag}(\pm 1)$ as in (3.4). We have

$$H = GJG^*, \quad \widetilde{H} = GTJT^*G^* \equiv \widetilde{G}J\widetilde{G}^*, \quad \widetilde{G} = GT. \tag{3.9}$$

Bearing (3.8) and (3.9) in mind, we now invoke Theorem 2.2 to get the HSVDs of G and $\widetilde{G}$:

$$\begin{aligned} G &= U|\Lambda|^{1/2}JV^*J, \qquad V^*JV = J, & (3.10)\\ \widetilde{G} &= \widetilde{U}|\widetilde{\Lambda}|^{1/2}J\widetilde{V}^*J, \qquad \widetilde{V}^*J\widetilde{V} = J. & (3.11)\end{aligned}$$

To derive a bound for $\sin\Theta$ between $\widehat{U}_1$ and $\widetilde{U}_1$, we define

$$W \stackrel{\text{def}}{=} G^{-1}SG = V\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k}\end{bmatrix}JV^*J, \tag{3.12}$$

which is J-unitary, indeed

$$WJW^* = G^{-1}SGJG^*SG^{-*} = G^{-1}HG^{-*} = J,$$

where we have used $SHS = H$. Other properties of W useful to our later developments are

$$W^2 = I, \quad W^* J = JW, \quad \|W\| \le \|V\|^2. \tag{3.13}$$

We use $GW = SG$ by (3.12) to get

$$\widehat{H} = SGTJT^*G^*S = GT\underbrace{T^{-1}WT}J\underbrace{T^*W^*T^{-*}}T^*G^* = \widetilde{G}\widetilde{T}J\widetilde{T}^*\widetilde{G}^* \equiv \widehat{G}J\widehat{G}^*,$$

where $\widetilde{G}$ is as in (3.9),

$$\widetilde{T} = T^{-1}WT, \quad \widehat{G} = \widetilde{G}\widetilde{T}. \tag{3.14}$$

Similarly to (3.10) and (3.11) we can write the HSVD of $\widehat{G}$ as

$$\widehat{G} = \widehat{U}|\widehat{\Lambda}|^{1/2}J\widehat{V}^*J, \qquad \widehat{V}^*J\widehat{V} = J. \tag{3.15}$$

It can be verified that

$$\widehat{H} - \widetilde{H} = \widehat{G}\left(J\widetilde{T}^* - \widetilde{T}^{-1}J\right)\widetilde{G}^*.$$

Pre- and post-multiply the equation by $\widehat{U}_2$ and $\widetilde{U}_1$, respectively, to get

$$\widehat{\Lambda}_2\widehat{U}_2^*\widetilde{U}_1 - \widehat{U}_2^*\widetilde{U}_1\widetilde{\Lambda}_1 = |\widehat{\Lambda}_2|^{1/2}J_{n-k}\widehat{V}_2^*J\left(J\widetilde{T}^* - \widetilde{T}^{-1}J\right)J\widetilde{V}_1J_k|\widetilde{\Lambda}_1|^{1/2},$$

where $\widehat{V} = \begin{bmatrix} \widehat{V}_1 & \widehat{V}_2 \end{bmatrix}$ is partitioned accordingly. Now, we can state our main theorem, which generalizes [4, Theorem 2.3] to indefinite Hermitian matrices.

Theorem 3.1 *Let $H = D^*AD$ and $\widetilde{H} = D^*\widetilde{A}D$ be two $n \times n$ Hermitian matrices with eigendecompositions* (1.1)*, and let A be non-singular and $\|A^{-1}\|\,\|\Delta A\| < 1$. Suppose Λ_1 and $\widetilde{\Lambda}_1$ have the same inertia. If*

$$\widetilde{\eta}_\chi \stackrel{\text{def}}{=} \min_{\mu\in\lambda(\widetilde{\Lambda}_1),\,\nu\in\lambda(\widetilde{\Lambda}_2)} \frac{|\mu-\nu|}{\sqrt{|\mu\nu|}} > 0,$$

then

$$\|\sin 2\Theta(U_1,\widetilde{U}_1)\|_{\mathrm{F}} \le \|\widehat{V}_2\|\|\widetilde{V}_1\|\frac{\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, \tag{3.16}$$

where T and $\widetilde{T}$ are as in (3.7) *and* (3.14).

If A is positive definite, $J = I$ and $\|\widehat{V}_2\| = \|\widetilde{V}\| = 1$, and thus Theorem 3.1 becomes [4, Theorem 2.4]. In general when A is indefinite, it's not clear at all how big $\|\widehat{V}_2\|$ and $\|\widetilde{V}_1\|$ may get. Also we would like to make $\|\widehat{V}_2\|$ disappear from the bound since it corresponds to the intermediate $\widehat{H}$. For practical purpose, $\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}$ is not immediately available and will likely be bounded in terms of norms of E (and thus of $\|\Delta A\|_{\mathrm{F}}$). We shall now deal with these issues. For the ease of presentation, define

$$\delta = \|A^{-1}\|\,\|\Delta A\|, \quad \delta_{\mathrm{F}} = \|A^{-1}\|\,\|\Delta A\|_{\mathrm{F}}. \tag{3.17}$$

It can be seen that $\|E\| \le \delta$ and $\|E\|_{\mathrm{F}} \le \delta_{\mathrm{F}}$.

3.1 Bounding $\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}$

We shall present a couple of lemmas in which the factor $\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}}$ in the right-hand side of (3.16) will be bounded in terms of $T - T^{-1}$, ΔA, and A. Using (3.14), $T^* = JTJ$, $W^{-1} = W$, and $W^* = JWJ$, we have

$$\begin{aligned} J\widetilde{T}^* - \widetilde{T}^{-1}J &= TWT^{-1}J - T^{-1}WTJ \\ &= \left(TW(T^{-1} - T) + (T - T^{-1})WT\right) J. \end{aligned}$$

Thus we have the following bound

$$\|J\widetilde{T}^* - \widetilde{T}^{-1}J\|_{\mathrm{F}} \le 2\|W\|\|T\|\|T^{-1} - T\|_{\mathrm{F}}. \tag{3.18}$$

Write $T = I + \Gamma$. We have by (3.7)

$$\begin{aligned} \|\Gamma\| &\le \sum_{n=1}^{\infty} \frac{(2n-1)!!}{2^n n!}\|E\|^n \le \frac{1}{2}\|E\| \sum_{n=1}^{\infty} \|E\|^{n-1} \\ &= \frac{1}{2} \cdot \frac{\|E\|}{1 - \|E\|} \le \frac{1}{2}\frac{\delta}{1-\delta}, &(3.19) \\ \|\Gamma\|_{\mathrm{F}} &\le \frac{1}{2} \cdot \frac{\|E\|_{\mathrm{F}}}{1 - \|E\|} \le \frac{1}{2}\frac{\delta_{\mathrm{F}}}{1-\delta}. &(3.20) \end{aligned}$$

So if $\delta < 2/3$, $\|\Gamma\| < 1$ which implies $T^{-1} = I - \Gamma + \Gamma^2 - \Gamma^3 + \ldots$ to yield

$$\begin{aligned} \|T^{-1} - T\|_{\mathrm{F}} &\le \|\Gamma\|_{\mathrm{F}} + \frac{\|\Gamma\|_{\mathrm{F}}}{1 - \|\Gamma\|} \\ &= \|\Gamma\|_{\mathrm{F}} \frac{2 - \|\Gamma\|}{1 - \|\Gamma\|} \\ &\le \frac{1}{2}\frac{4 - 5\delta}{(1-\delta)(2-3\delta)}\delta_{\mathrm{F}} \\ &\le \frac{2}{2 - 3\delta}\delta_{\mathrm{F}}. &(3.21) \end{aligned}$$

An immediate consequence of Theorem 3.1, (3.13), (3.18), and (3.21) is the following corollary.

Corollary 3.1 *To the conditions of Theorem 3.1 add this: $\|A^{-1}\|\|\Delta A\| < 2/3$. Then*

$$\begin{aligned} \frac{1}{2}\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} &\le \|V\|^2\|\widetilde{V}_1\|\|\widehat{V}_2\| \frac{\|T\|_2\|T - T^{-1}\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, &(3.22) \\ &\le \|V\|^2\|\widetilde{V}_1\|\|\widehat{V}_2\| \frac{\varepsilon}{\widetilde{\eta}_\chi}, &(3.23) \end{aligned}$$

where

$$\varepsilon = \frac{2 - \delta}{(1-\delta)(2-3\delta)}\delta_{\mathrm{F}} = \delta_{\mathrm{F}} + \mathcal{O}(\delta_{\mathrm{F}}^2). \tag{3.24}$$

Note that bounds (3.22) and (3.23) are proper generalizations of [4, (2.35)], since in the positive definite case V, $\widetilde{V}$ and $\widehat{V}$ are unitary and thus disappear from these inequalities altogether.

3.2 Bounding $\|\widetilde{V}\|$ and $\|\widehat{V}\|$ in terms of $\|V\|$

We shall now bound $\|\widetilde{V}\|$ and $\|\widehat{V}\|$ in terms of $\|V\|$. For this we will need the following lemma.

Lemma 3.1 *Let V and $\widetilde{V}$ be the hyperbolic eigenvector matrices of the pairs $\{G^*G, J\}$ and $\{T^*G^*GT, J\}$, respectively. Write $\Gamma = T - I$ and define $\gamma = \|\Gamma\|_{\mathrm{F}}/(1 - \|\Gamma\|)$.*

$$\text{If } \|\Gamma\| < 1 \text{ and } \gamma\|V\|^2 < \tfrac{1}{4}, \text{ then } \quad \|\widetilde{V}\| \leq \frac{\|V\|}{\sqrt{1 - 4\gamma\|V\|^2}}. \tag{3.25}$$

Proof: (3.25) follows from [8, Lemma 5] (see [8, Lemma 4] or [6, Lemma 1]). ∎

It is reasonable to expect that $\|\Gamma\|_{\mathrm{F}}$ in defining γ in this lemma should be replaced by $\|\Gamma\|$, the spectral norm of Γ. But we are unable to prove this.

To use Lemma 3.1, we shall interpret that V, $\widetilde{V}$, and $\widehat{V}$ assigned above are the hyperbolic eigenvector matrices of the pairs $\{G^*G, J\}$, $\{\widetilde{G}^*\widetilde{G}, J\}$, and $\{\widehat{G}^*\widehat{G}, J\}$, respectively (see (3.5), (3.9), and (3.14) for the assignments of G, $\widetilde{G}$, and $\widehat{G}$). By (3.20),

$$\frac{\|\Gamma\|_{\mathrm{F}}}{1 - \|\Gamma\|} \leq \frac{1}{2 - 3\delta}\delta_{\mathrm{F}} \equiv \alpha. \tag{3.26}$$

Lemma 3.1 applied to $\{G^*G, J\}$ and $\{\widetilde{G}^*\widetilde{G}, J\}$ yields

$$\|\widetilde{V}\| \leq \frac{\|V\|}{\sqrt{1 - 4\alpha\|V\|^2}} \quad \text{if } 4\alpha\|V\|^2 < 1 \text{ and } \delta < 2/3. \tag{3.27}$$

Note by (3.9) and (3.14) and $\widehat{G} = GWT$, thus $\widehat{G}^*\widehat{G} = T^*W^*G^*GWT$ and

$$GW = U|\Lambda|^{1/2}JV^*JW = U|\Lambda|^{1/2}JV^*W^*J, \tag{3.28}$$

by (3.13). Since both W and V are J-unitary, WV is also J-unitary. (3.28) is the HSVD of GW and consequently WV is the hyperbolic eigenvector matrix of the pair $\{W^*G^*GW, J\}$. It follows from (3.12) that

$$WV = V \begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix} J,$$

and thus $\|WV\| = \|V\|$. Lemma 3.1 applied to the pairs $\{W^*G^*GW, J\}$ and $\{\widehat{G}^*\widehat{G}, J\}$ (recall $\widehat{G} = GWT$) implies

$$\|\widehat{V}\| \leq \frac{\|V\|}{\sqrt{1 - 4\alpha\|V\|^2}} \quad \text{if } 4\alpha\|V\|^2 < 1 \text{ and } \delta < 2/3. \tag{3.29}$$

Corollary 3.2 *To the conditions of Theorem 3.1 add these:*

$$\delta < 2/3, \quad \alpha \equiv \frac{1}{2 - 3\delta}\delta_{\mathrm{F}} \leq \frac{1}{4\|V\|^2}.$$

Then

$$\begin{aligned} \frac{1}{2}\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} &\leq \frac{\|V\|^4}{1 - 4\alpha\|V\|^2} \frac{\|T\|_2\|T - T^{-1}\|_{\mathrm{F}}}{\widetilde{\eta}_\chi}, && (3.30) \\ &\leq \frac{\|V\|^4}{1 - 4\alpha\|V\|^2} \frac{\varepsilon}{\widetilde{\eta}_\chi}, && (3.31) \end{aligned}$$

where ε is defined by (3.24).

Proof: By (3.26), γ in Lemma 3.1 satisfies $\gamma \le \alpha$. The inequality (3.30) follows by inserting (3.25) and (3.29) into (3.22) and (3.31) follows by inserting (3.25) and (3.29) into (3.23). ∎

3.3 Bounding V

The bounds (3.30) and (3.31) contain an additional factor which depends on J-unitary matrix V whose norm may be big. Here we shall show that $\|V\|$ is of modest magnitudes for special but interesting matrices. In [8] two classes of so-called *Well-Behaved Matrices* for which $\kappa(V) \equiv \|V\|\,\|V^{-1}\|$ can be bounded in a satisfactory way are defined. The first class consists of *scaled diagonal dominant (SDD) matrices*, and the second class consists of *quasi-definite matrices*. We shall now review bounds in [8], as well as derive new and improved ones, for both classes of matrices.

A matrix H is *scaled diagonal dominant*, if it can be written as $H = D(J+N)D$ with diagonal positive definite D, $J = \mathrm{diag}(\pm 1)$, and $\|N\| < 1$. The following theorem was proved in [8].

Theorem 3.2 *Let $H = D(J+N)D$ be SDD. Then $\|V\| \le \sqrt{n(1+\|N\|)/(1-\|N\|)}$.*

This bound depends on the square root of the dimension. Note that

$$H = D(I+NJ)^{1/2} J \left[(I+NJ)^{1/2}\right]^* D = \widetilde{B} J \widetilde{B}^*,$$

where $\widetilde{B} = D(I+NJ)^{1/2}$. Thus, we can interpret $\widetilde{B}$ as obtained by multiplicatively perturbing D, and this way V (the hyperbolic eigenvector matrix of $\left\{\left[(I+NJ)^{1/2}\right]^* D^2 (I+NJ)^{1/2}, J\right\}$) is resulted from perturbing I (the hyperbolic eigenvector matrix of $\{D^2, J\}$). Lemma 3.1 applied to the two pairs yields a theorem as follows.

Theorem 3.3 *Let $H = D(J+N)D$ be SDD. If $\|N\|_{\mathrm{F}} < 2/7$, we have*

$$\|V\|^2 \le \frac{1}{1-4\gamma}, \tag{3.32}$$

where $\gamma = \|N\|_{\mathrm{F}}/(2-3\|N\|)$.

Proof: Write $I + \Gamma = (I+NJ)^{1/2}$. Analogously to (3.20), we have $\|\Gamma\|_{\mathrm{F}} \le \frac{1}{2}\|N\|_{\mathrm{F}}/(1-\|N\|)$. Thus if $\|N\|_{\mathrm{F}} < 2/7$, $\|\Gamma\|_{\mathrm{F}}/(1-\|\Gamma\|) \le \gamma < 1/4$, and then (3.25) implies (3.32). ∎

The bound in Theorem 3.3 does not explicitly depend on n while the bound in Theorem 3.2 does.

Next we make an extension of the bound (3.32) to a larger class of matrices, containing SDD matrices. We say a Hermitian matrix H is *block scaled diagonally dominant (BSDD)* if it admits

$$H = D_{\mathrm{b}}^*(J_{\mathrm{b}} + N_{\mathrm{b}})D_{\mathrm{b}}, \tag{3.33}$$

where $D_{\mathrm{b}} = D_1 \oplus D_2 \oplus \ldots \oplus D_k$ and D_i is non-singular, $J_{\mathrm{b}} = J_1 \oplus J_2 \oplus \ldots \oplus J_k$, with $J_i = I$ or $J_i = -I$, and $\|N_{\mathrm{b}}\|_{\mathrm{F}} \le 1$. A BSDD H as just described can be rewritten as

$$H \equiv B J_{\mathrm{b}} B^*, \quad \text{where } B = D_{\mathrm{b}}^*(I + N_{\mathrm{b}} J_{\mathrm{b}})^{1/2}. \tag{3.34}$$

Using similar approach as above for SDD matrices we have the following bound.

Theorem 3.4 *Let $H = D_{\rm b}^*(J_{\rm b} + N_{\rm b})D_{\rm b}$ be BSDD. If $\|N_{\rm b}\|_{\rm F} < 2/7$, we have*

$$\|V\|^2 \leq \frac{1}{1 - 4\gamma_{\rm b}}, \tag{3.35}$$

where $\gamma_{\rm b} = \|N_{\rm b}\|_{\rm F}/(2 - 3\|N_{\rm b}\|)$.

Lastly we consider so-called quasi-definite matrices. A Hermitian matrix H is said to be a *quasi-definite* if there exists a permutation matrix P such that

$$H_{\rm q} \equiv P^T HP = \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^* & -H_{22} \end{bmatrix},$$

where H_{11} and H_{22} are positive definite. The following theorem was proved in [8].

Theorem 3.5 *Let H be quasi-definite as just described. Then*

$$\kappa(V) \leq n \max\{\|A_{11}\| + \|A_{12}A_{22}^{-1}A_{12}^*\|, \|A_{22}\| + \|A_{12}^*A_{11}^{-1}A_{12}\|\}, \tag{3.36}$$

where

$$H_{\rm q} \equiv DAD = D \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & -A_{22} \end{bmatrix} D,$$

$D = \mathrm{diag}(H_{11}^{1/2}, H_{22}^{1/2})$, and V is the hyperbolic eigenvector matrix of $\{G^*G, J\}$, $J_{ii} = \mathrm{sign}(H_{ii})$ and $G = DF$ such that $H_{\rm q} \equiv GJG = DFJF^*D$ (that is, $A = FJF^*$).

Theorem 3.4 also applies to the current case if H_{12} is "small" enough. To do so, we let $H_{11} = L_1L_1^*$ and $H_{22} = L_2L_2^*$ be, e.g., Cholesky factorizations, and write

$$H_{\rm q} = L_{\rm q}(J + N_{\rm q})L_{\rm q}^*,$$

where

$$L_{\rm q} = \begin{bmatrix} L_1 & \\ & L_2 \end{bmatrix}, \quad N_{\rm q} = \begin{bmatrix} 0 & L_1^{-1}H_{12}L_2^{-*} \\ L_2^{-1}H_{12}^*L_2^{-*} & 0 \end{bmatrix}, \quad J = \begin{bmatrix} I & \\ & -I \end{bmatrix}.$$

Note that if $\|N_{\rm q}\|$ is small enough we can apply (3.25). Write $\gamma_{\rm q} = \|N_{\rm q}\|_{\rm F}/(2 - 3\|N_{\rm q}\|)$. If $\|L_1^{-1}H_{12}L_2^{-*}\|_{\rm F} < 2/(4\sqrt{2} + 3)$ which implies $\gamma_{\rm q} < 1/4$, from (3.35) it follows

$$\|V\|^2 \leq \frac{1}{1 - 4\gamma_{\rm q}}. \tag{3.37}$$

We have so far considered 3 related classes, and all fall into the category of BSDD matrices. So we shall present a theorem as a corollary to Theorem 3.1 for BSDD matrices.

Theorem 3.6 *Let $H = D_{\rm b}^*AD_{\rm b}$ be BSDD matrix where $A = J_{\rm b} + N_{\rm b}$, $D_{\rm b} = D_1 \oplus \ldots \oplus D_k$ and $J_{\rm b} = J_1 \oplus \ldots \oplus J_k$ are non-singular with $J_i = I$ or $J_i = -I$, for $i = 1, \ldots, k$. Suppose H is perturbed to $\widetilde{H} = D^*(A + \Delta A)D$. Set*

$$\alpha_{\rm b} \equiv \frac{\delta_{\rm F}}{2 - 3\delta} = \frac{\|(J_{\rm b} + N_{\rm b})^{-1}\|\|\Delta A\|_{\rm F}}{2 - 3\|(J_{\rm b} + N_{\rm b})^{-1}\|\|\Delta A\|} \quad \text{and} \quad \gamma_{\rm b} \equiv \frac{\|N_{\rm b}\|_{\rm F}}{2 - 3\|N_{\rm b}\|}.$$

If $\|N_{\rm b}\|_{\rm F} < 2/7$ *and if* $4\alpha_{\rm b}/(1-4\gamma_{\rm b}) < 1$, *then*

$$\frac{1}{2}\|\sin 2\Theta(U_1,\widetilde{U}_1)\|_{\rm F} \le \frac{1}{(1-4\gamma_{\rm b})\cdot(1-4(\gamma_{\rm b}+\alpha_{\rm b}))}\frac{\|T\|_2\|T-T^{-1}\|_{\rm F}}{\widetilde{\eta}_\chi}, \tag{3.38}$$

$$\le \frac{1}{(1-4\gamma_{\rm b})\cdot(1-4(\gamma_{\rm b}+\alpha_{\rm b}))}\frac{\varepsilon}{\widetilde{\eta}_\chi}, \tag{3.39}$$

where ε *is defined by* (3.24).

Proof: The upper bounds are obtained by bounding α and $\|V\|$ as in Corollary 3.2. For BSDD matrices we have $A^{-1} = (J+N_{\rm b})^{-1}$. Further since $\|\alpha_{\rm b}\|_{\rm F} < 2/7$, from (3.35) it follows that $\|V\|^2 \le 1/(1-4\gamma_{\rm b})$, and thus $4\alpha_{\rm b}\|V\|^2 < 1$ which allow us to use the bounds (3.30) and (3.31). It can be seen that

$$\frac{\|V\|^4}{1-4\alpha_{\rm b}\|V\|^2} \le \frac{1}{(1-4\gamma_{\rm b})\cdot(1-4(\gamma_{\rm b}+\alpha_{\rm b}))},$$

which together with (3.30) and (3.31) yields (3.38) and (3.39), respectively. ∎

4 A Numerical Example

Consider $H = D^*AD$:

$$D = \begin{bmatrix} L_1 & \\ & L_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21}^* & A_{22} \end{bmatrix},$$

$$L_1 = \begin{bmatrix} 6.7676e+2 & 4.4692e+2 & 7.6564e+2 \\ 0 & 2.3767e-1 & 8.5502e-2 \\ 0 & 0 & 1.2205e-1 \end{bmatrix}, \; L_2 = \begin{bmatrix} 1.4078e+3 & -4.4582e+2 & -2.1765e+3 \\ 0 & 3.4030e+3 & -5.5373e+3 \\ 0 & 0 & 4.8642e+1 \end{bmatrix},$$

$$A_{11} = I, \quad A_{22} = -I,$$

$$A_{12} = \begin{bmatrix} 4.6675e-3 & 3.9688e-3 & 1.0076e-3 \\ 9.7469e-2 & 4.7937e-2 & 3.4591e-2 \\ 2.7080e-5 & -2.9377e-5 & -5.8637e-4 \end{bmatrix}.$$

Let

$$\Delta A = \begin{bmatrix} 2.2e-4 & 1.1e-5 & 2.4e-5 & 3.8e-6 & 3.2e-6 & 1.3e-5 \\ 1.1e-5 & 8.4e-6 & 3.4e-4 & 3.0e-6 & 3.5e-5 & 9.9e-6 \\ 2.4e-5 & 3.4e-4 & 6.7e-5 & 1.3e-8 & -2.1e-6 & -2.2e-7 \\ 3.8e-6 & 3.0e-6 & 1.3e-8 & -9.2e-5 & 3.4e-4 & 4.7e-5 \\ 3.2e-6 & 3.5e-5 & -2.1e-6 & 3.4e-4 & -4.7e-5 & 1.9e-5 \\ 1.3e-5 & 9.9e-6 & -2.2e-7 & 4.7e-5 & 1.9e-5 & -2.5e-4 \end{bmatrix}.$$

Spectrum of matrix $\widetilde{H} = D^*(A+\Delta A)D$ is

$$\lambda(\widetilde{H}) = \{-4.5142e+7,\ -4.0159e+6,\ -2.9967e+2,\ 7.6909e-3,\ 4.0809e-2,\ 1.2443e+6\}.$$

If we write $A = Q|\Omega|^{1/2}J|\Omega|^{1/2}Q^*$ then $H = GJG^*$ (similarly for perturbed quantities). Now applying the HSVD[1] to $\widetilde{G}$ we obtain $\widetilde{G} = \widetilde{U}|\widetilde{\Lambda}|^{1/2}J\widetilde{V}^*J$ where

$$\widetilde{U} = \begin{bmatrix} 6.0678e-1 & -4.4527e-1 & -6.5844e-1 & -4.9637e-4 & 8.8632e-4 & 1.8055e-4 \\ 4.0071e-1 & 8.8676e-1 & -2.3040e-1 & -3.3034e-4 & 5.9491e-4 & -3.7053e-4 \\ 6.8647e-1 & -1.2404e-1 & 7.1650e-1 & -5.6247e-4 & 1.0062e-3 & 3.2199e-5 \\ 1.0515e-3 & 3.0853e-4 & 7.8402e-6 & 5.5603e-2 & -6.7732e-1 & 7.3358e-1 \\ -5.8103e-4 & 2.3473e-4 & 5.9004e-6 & 4.7068e-1 & 6.6574e-1 & 5.7901e-1 \\ -1.1751e-3 & 1.4252e-4 & 3.5703e-6 & -8.8055e-1 & 3.1309e-1 & 3.5582e-1 \end{bmatrix},$$

$$\widetilde{V} = \begin{bmatrix} 9.9786e-1 & -6.0033e-2 & 2.6174e-2 & -3.2594e-4 & -1.4212e-3 & 1.5757e-3 \\ 5.6316e-2 & 9.9210e-1 & 1.2569e-1 & -3.5942e-2 & 4.0352e-2 & -1.7707e-2 \\ -3.3464e-2 & -1.2391e-1 & 9.9173e-1 & 1.8614e-3 & -2.0045e-3 & 1.1523e-3 \\ -1.9456e-3 & -5.6730e-2 & -4.1672e-3 & 6.0271e-1 & -7.3069e-1 & 3.2568e-1 \\ -4.4352e-4 & -1.3248e-3 & 6.3785e-5 & 5.5096e-1 & 6.7381e-1 & 4.9236e-1 \\ 1.7198e-3 & 1.6388e-3 & 3.5202e-4 & -5.7834e-1 & -1.1706e-1 & 8.0736e-1 \end{bmatrix}.$$

We derive a bound for $\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}}$, where U_1 and $\widetilde{U}_1$ contain eigenvectors corresponding to eigenvalues $\lambda_1 = 1.2440e+6, \lambda_2 = 7.6900e-3$, and $\widetilde{\lambda}_1 = 1.2443e+6$, $\widetilde{\lambda}_2 = 7.6909e-3$, respectively. We have

$$\alpha_{\mathrm{b}} = 8.7366e-004, \quad \gamma_{\mathrm{b}} = 9.1307e-002, \quad \widetilde{\eta}_{\chi} = 1.8694.$$

(3.39) of Theorem 3.6 gives

$$\frac{1}{2}\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} \leq 1.0333e-003,$$

in comparison to $\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} = 3.5304e-004$.

Note that the absolute gap is $3.3118e-2$ and $\|\Delta H\| = 7.2828e+3$, and thus the classical Davis-Kahan $\sin 2\Theta$ theorem produces

$$\frac{1}{2}\|\sin 2\Theta(U_1, \widetilde{U}_1)\|_{\mathrm{F}} \leq 2.2e+5$$

which is too big to be useful.

Also note, since $\|N_{\mathrm{b}}\| = 0.11415$ the bound (3.37) yields that $\|V\| \leq 1.2551$, a quite accurate estimation of $\|V\| = 1.059$.

[1] A MATLAB function to compute HSVDs is available from the first author upon request.

References

[1] C. Davis and W. M. Kahan, The rotation of eigenvectors by a perturbation. III, *SIAM J. Numer. Anal.*, 7:1–46 (1970).

[2] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.

[3] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[4] R.-C. Li, Relative perturbation theory: (IV) $\sin 2\Theta$ theorems, *Linear Algebra Appl.*, 311:45–60 (2000).

[5] R. Onn, A. O. Steinhardt and A. Bojanczyk, Hyperbolic singular value decompositions and applications, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 1575–1588 (1991).

[6] I. Slapničar and N. Truhar, Relative perturbation theory for hyperbolic eigenvalue problem, *Linear Algebra Appl.*, 309:57–72 (2000).

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

[8] N. Truhar and I. Slapničar, Relative perturbation bound for invariant subspaces of Hermitian matrices, *Linear Algebra Appl.*, 301:171–185 (1999).