Relative Residual Bounds for Indefinite Hermitian Matrices *

Ninoslav Truhar† and Ivan Slapničar‡

Abstract

We prove several residual bounds for relative perturbations of the eigenvalues of indefinite Hermitian matrix. The bounds fall into two categories – the Weyl-type bounds and the Hofmann–Wielandt-type bounds. The bounds are expressed in terms of sines of acute principal angles between certain subspaces associated with the indefinite decomposition of the given matrix. The bounds are never worse than the classical residual bounds and can be much sharper in some cases. The bounds generalize the existing relative residual bounds for positive definite matrices to indefinite case.

1 Introduction

Let $H \in \mathbb{C}^{n\times n}$ be a Hermitian matrix, and let $X \in \mathbb{C}^{n\times m}$ where $n \geq m$, be an orthonormal matrix, and

$$M = X^*HX, \qquad R = HX - XM, \qquad \mathcal{X} = \mathcal{R}(X), \tag{1}$$

where $\mathcal{X} = \mathcal{R}(X)$ denote the column space of X. Furthermore, let

$$\lambda_1 \geq \cdots \geq \lambda_n \qquad \text{and} \qquad \mu_1 \geq \cdots \geq \mu_m,$$

*This work was supported by the grant 0023002 from the Croatian Ministry of Science and Technology.

†University Josip Juraj Strossmayer, Faculty of Civil Engineering, Drinska 16 a, 31000 Osijek, Croatia, *truhar@most.gfos.hr*.

‡University of Split, Faculty of Electrical Engineering, Mechanical, Engineering and Naval Architecture, R. Boškovića b.b, 21000 Split, Croatia, *ivan.slapnicar@fesb.hr*.

be the eigenvalues of H and M, respectively.

The eigenvalues of M are sometimes called Ritz values or Rayleigh-Ritz approximations of the eigenvalues of H. Ritz values are optimal in the sense that $\|R\|$ is minimized for $M = X^*HX$, that is, if we replace M by another matrix C we can only increase the spectral norm of R,

$$\|R\| = \|HX - XM\| \leq \|HX - XC\|,$$

for all matrices C of order m (see [10, Theorem 1.15.IV] or [5, Theorem 11-4-5]). Moreover, one can always find m eigenvalues of H that are within absolute distance $\|R\|$ of the Ritz values [5, Theorem 11-5-1]

$$\max_{1 \leq j \leq m} |\lambda_{\tau(j)} - \mu_j| \leq \|R\|, \tag{2}$$

for some permutation τ. There is a similar residual bound given in the Frobenius norm [10, Corollary 4.15.IV]

$$\sqrt{\sum_j (\lambda_{\tau(j)} - \mu_j)^2} \leq \|R\|_F. \tag{3}$$

The above bounds measure absolute distance between eigenvalues, thus they belong to classical or absolute perturbation theory.

Drmač [1, Theorem 6] derived a relative residual error bound for positive definite Hermitian matrix $H = LL^*$ of the following form:

$$\frac{|\lambda_{\tau(j)} - \mu_j|}{|\lambda_{\tau(j)}|} \leq \frac{\sin\psi}{1 - \sin\psi} \quad j = 1, \ldots, m, \tag{4}$$

where ψ is the maximal acute principal angle between $\mathcal{R}(L^*\mathcal{X})$ and $\mathcal{R}(L^{-1}\mathcal{X})$.

We present two relative residual bounds for the eigenvalues of indefinite Hermitian matrices. The first one is similar to (4) and represents the relative version of the Weyl-type residual bound (2). The second one is the relative version of the Hofmann–Wielandt type residual bound (3).

The paper is organized as follows: in the next section we give some preliminary results, in Section 3 we prove our relative residual bounds, in Section 4 we discuss some differences between the positive definite and the indefinite case, and in Section 5 we give a numerical example.

2 Preliminaries

In this section we present some definitions and auxiliary results on the Hermitian eigenvalue problem, Hermitian indefinite decomposition and subspaces and angles between them.

Let H be indefinite Hermitian matrix and let $H = U\Lambda U^*$ be its eigenvalue decomposition. The spectral absolute value of H is defined as

$$|H|_S = U|\Lambda|U^* = \sqrt{H^2}. \tag{5}$$

Let $H = LJL^*$ be the indefinite Hermitian decomposition of H where L is non-singular and J is diagonal with ± 1 on its diagonal such that $J_{ii} = \operatorname{sign}(\Lambda_{ii})$ (see e.g. [7] for more details). The eigenvalue problem for H is closely related to the hyperbolic eigenvalue problem for the pair (L^*L, J) (see e.g. [9]) – there exists non-singular J-orthogonal matrix V such that

$$V^*L^*LV = |\Lambda|, \qquad V^*JV = J. \tag{6}$$

By inverting $V^*JV = J$ we have $VJV^* = J$. Now, from $V^*J = JV^{-1}$ it follows that[1]

$$\|V\| = \|V^{-1}\|. \tag{7}$$

Thus, for the spectral condition number of V we have

$$\kappa(V) = \|V\|\,\|V^{-1}\| = \|V\|^2.$$

Further,

$$U = LV|\Lambda|^{-1/2}. \tag{8}$$

Indeed,

$$\begin{aligned} U^*HU &= (|\Lambda|^{-1/2}V^*L^*)\,(LJL^*)\,(LV|\Lambda|^{-1/2}) \\ &= |\Lambda|^{-1/2}(V^*L^*LV)J(V^*L^*LV)|\Lambda|^{-1/2} \\ &= |\Lambda|^{-1/2}|\Lambda|J|\Lambda||\Lambda|^{-1/2} = \Lambda. \end{aligned}$$

¿From (8) and (5) we also have

$$|H|_S = U|\Lambda|U^* = LV|\Lambda|^{-1/2}|\Lambda||\Lambda|^{-1/2}V^*L^* = LVV^*L^*. \tag{9}$$

[1] Even more, one can easily show that the singular values of V come in the pairs of reciprocals.

Let $Y = LX$ and $Z = L^{-1}X$, further let

$$Y_L = JL^*X = JY, \qquad Z_L = L^{-1}X = Z. \tag{10}$$

Let $P_{(\cdot)}$ denote the orthogonal projector onto the indicated subspace.

It is easy to show that

$$(P_{\mathcal{Z}} P_{\mathcal{Y}})^\dagger = YZ^*,$$

where † denotes the generalized or pseudo-inverse (for the proof see [1, Proof of Theorem 3]). By using (10), we obtain

$$\begin{aligned}(P_{\mathcal{Z}_L} J P_{\mathcal{Y}_L})^\dagger &= (Z_L Z_L^\dagger J Y_L Y_L^\dagger)^\dagger = (ZZ^\dagger YY^\dagger J)^\dagger = J(ZZ^\dagger YY^\dagger)^\dagger \\ &= JYZ^* = Y_L Z_L^*.\end{aligned} \tag{11}$$

We shall use the angle function $\angle(\mathcal{Y}, \mathcal{Z})$ between arbitrary subspaces $\mathcal{Y}$ and $\mathcal{Z}$ of $\mathbb{C}^n$ defined by (see [14]):

$$\angle(\mathcal{Y}, \mathcal{Z}) = \sin^{-1} \min\{\|(I - P_{\mathcal{Z}})P_{\mathcal{Y}}\|, \|(I - P_{\mathcal{Y}})P_{\mathcal{Z}}\|\}. \tag{12}$$

As we shall see later (see Remark 1) for some perturbation δH, we shall need to derive $\|L^{-1}\delta H L^{-*}\|$. Thus, we shall need the following representation of the pair of orthogonal projectors $P_{\mathcal{Y}_L}$ and $P_{\mathcal{Z}_L}$ in $\mathbb{C}^n$, also due to Wedin [14].

Theorem 1 (Wedin) *Let $\mathcal{Y}_L$, $\mathcal{Z}_L$ be subspaces in $\mathbb{C}^n$. Assume that*

(a) $\operatorname{rank}(P_{\mathcal{Y}_L} P_{\mathcal{Z}_L}) = k + l$

(b) $P_{\mathcal{Y}_L} P_{\mathcal{Z}_L}$ *has k singular values equal to one.*

Then there exist an orthogonal basis of $\mathbb{C}^n$ such that, with respect to this basis, $P_{\mathcal{Y}_L}$ and $P_{\mathcal{Z}_L}$ are represented by block diagonal matrices P_1 and P_2, respectively, where

$$P_1 = \begin{bmatrix} I_k & & \\ & \oplus_i^l \Phi_i(P_{\mathcal{Y}_L}) & \\ & & \Delta_{P_{\mathcal{Y}_L}} \end{bmatrix}, \qquad P_2 = \begin{bmatrix} I_k & & \\ & \oplus_i^l \Psi_i(P_{\mathcal{Z}_L}) & \\ & & \Delta_{P_{\mathcal{Z}_L}} \end{bmatrix},$$

$$\Phi_i(P_{\mathcal{Y}_L}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Psi_i(P_{\mathcal{Z}_L}) = \begin{bmatrix} \cos^2\theta_i & \cos\theta_i \sin\theta_i \\ \cos\theta_i \sin\theta_i & \sin^2\theta_i \end{bmatrix}.$$

Here I_k is the $k \times k$ identity matrix, $\Delta_{P_{\mathcal{Y}_L}}$, $\Delta_{P_{\mathcal{Z}_L}}$ are diagonal matrices with entries from $\{0, 1\}$ and $\Delta_{P_{\mathcal{Y}_L}} \Delta_{P_{\mathcal{Z}_L}} = 0$. The numbers

$$0 < \theta_1 \le \cdots \le \theta_l < \pi/2$$

are the acute principal angles between $\mathcal{Y}_L$ and $\mathcal{Z}_L$.

3 Residual bounds

First we prove the Weyl-type relative residual bound for the eigenvalues of non-singular indefinite Hermitian matrix H.

Theorem 2 *Let $H = LJL^*$, where L and J are non-singular and J is diagonal with ± 1 on its diagonal and let*

$$\delta H = RX^* + XR^*, \tag{13}$$

where X is an $n \times m$ orthonormal matrix. Then there are at least m eigenvalues λ_{i_k}, $k = 1, \ldots, m$, of H for which

$$\frac{|\lambda_{i_k} - \mu_k|}{|\lambda_{i_k}|} \leq \kappa(V) \|L^{-1} \delta H L^{-*}\|, \qquad k = 1, \ldots, m. \tag{14}$$

Here, μ_k are eigenvalues of matrix M defined by (1) and V is J-unitary matrix which diagonalizes the pair (L^*L, J) as in (6).

Proof. First, notice that the Hermitian matrix $\widetilde{H} = H - \delta H$ has $\mathcal{X}$ as an invariant subspace. Indeed, using the fact that $X^*R = R^*X = 0$ we can write

$$\begin{aligned}
(H - \delta H)X &= HX - RX^*X + XR^*X \\
&= HX - (HX - XM) - XR^*X = XM.
\end{aligned}$$

This means that the eigenvalues of M coincide with at least m eigenvalues of $\widetilde{H}$. By applying a result of Veselić and Slapničar [13, Theorem 2.1] we know that if

$$|x^* \delta H x| \leq \eta x^* |H|_S x, \quad \forall\ x, \ \ \eta < 1,$$

then

$$1 - \eta \leq \frac{\widetilde{\lambda}_i}{\lambda_i} \leq 1 + \eta.$$

By (7) we have $\kappa(V) = \|V\|\,\|V^{-1}\| = \|V^{-1}\|^2$. Using (9) it can be written

$$\begin{aligned}
|x^* \delta H x| &= |x^* L L^{-1} \delta H L^{-*} L^* x| \leq \|L^{-1} \delta H L^{-*}\|\, x^* L L^* x \\
&\leq \|L^{-1} \delta H L^{-*}\|\, x^* L V V^{-1} V^{-*} V^* L^* x \\
&\leq \|L^{-1} \delta H L^{-*}\|\, \kappa(V)\, x^* |H|_S x,
\end{aligned}$$

which means that there are at least m eigenvalues λ_{i_k} of H such that (14) holds. ∎

Now we need to bound $\|L^{-1}\delta H L^{-*}\|$. Notice that

$$\begin{aligned} L^{-1}\delta H L^{-*} &= L^{-1}\left[(HX - XM)X^* + X(X^*H - MX^*)\right]L^{-*} \\ &= JL^*XX^*L^{-*} - L^{-1}XX^*L\,JL^*XX^*L^{-*} \\ &\quad + L^{-1}XX^*LJ - L^{-1}XX^*LJL^*XX^*L^{-*} \\ &= (I - Z_L Y_L^* J)Y_L Z_L^* + Z_L Y_L^*(I - JY_L Z_L^*). \end{aligned}$$

Using (11) we have

$$\begin{aligned} L^{-1}\delta H L^{-*} &= \left[I - (P_{\mathcal{Y}_L} J P_{\mathcal{Z}_L})^\dagger J\right](P_{\mathcal{Z}_L} J P_{\mathcal{Y}_L})^\dagger \\ &\quad + (P_{\mathcal{Y}_L} J P_{\mathcal{Z}_L})^\dagger \left[I - J(P_{\mathcal{Z}_L} J P_{\mathcal{Y}_L})^\dagger\right]. \end{aligned}$$

The following remark is due to Drmač [1, Remark 8].

Remark 1 Using canonical representation of the pair $P_{\mathcal{Y}_L}$ and $P_{\mathcal{Z}_L}$ we can see that, in a suitably chosen orthonormal basis, the matrix $L^{-1}\delta H L^{-*}$ is block diagonal with diagonal blocks of the form

$$\Gamma_i = \pm \begin{bmatrix} 0 & \tan\psi_i \\ \tan\psi_i & 2\tan^2\psi_i \end{bmatrix}, \qquad \|\Gamma_i\| = \frac{\sin\psi_i}{1 - \sin\psi_i},$$

with acute principal angles ψ_i between $\mathcal{Y}_L$ and $\mathcal{Z}_L$.

Now we can formulate Theorem 2 in terms of acute principal angles.

Theorem 3 *Let $H = LJL^*$, where L and J are non-singular and J is diagonal with ± 1 on its diagonal. Let*

$$\mathcal{Y}_L = JL^*\mathcal{X}, \qquad \mathcal{Z}_L = L^{-1}\mathcal{X},$$

and let ψ be the maximal acute principal angle between $\mathcal{Y}_L$ and $\mathcal{Z}_L$. Then there are at least m eigenvalues λ_{i_k}, $k = 1, \ldots, m$, of H for which

$$\frac{|\lambda_{i_k} - \mu_k|}{|\lambda_{i_k}|} \leq \kappa(V)\frac{\sin\psi}{1 - \sin\psi}, \qquad k = 1, \ldots, m, \tag{15}$$

provided that right hand side in (15) is less than one. Here V is a J-unitary matrix which diagonalizes the pair (L^*L, J) as in (6).

Proof. From Theorem 2 it follows that there are at least m eigenvalues λ_{i_k} of H such that (14) holds. This, together with Remark 1, yields (15). ■

In order to prove our second result, a Hofmann–Wielandt type relative residual bound, we need some results on doubly stochastic matrices. A real $n \times n$ matrix Y is doubly stochastic if $Y_{ij} \geq 0$ and $\sum_{i=1}^{n} Y_{ik} = \sum_{i=1}^{n} Y_{ki} = 1$ for $k = 1, 2, \cdots, n$. By Birkhoff's theorem [3, Theorem 8.7.1], a matrix is doubly stochastic if and only if it lies in the convex hull of all permutation matrices. This result has lead to the following lemma by R.-C. Li [6, Lemma 5.1].

Lemma 1 (Li) *Let Y be a $n \times n$ doubly stochastic matrix, and let M be a $n \times n$ complex matrix. Then there exists a permutation τ of $\{1, 2, \ldots, n\}$ such that*

$$\sum_{i,j=1}^{n} |M_{ij}|^2 Y_{ij} \geq \sum_{i=1}^{n} |M_{i\tau(i)}|^2.$$

The following theorem gives Hofmann–Wielandt type relative residual bound for nonsingular indefinite Hermitian matrix.

Theorem 4 *Let $H = LJL^*$, where L and J are non-singular and J is diagonal with ± 1 on its diagonal. Let $\widetilde{H} = H - \delta H$, where δH is defined by (13), and let $\widetilde{H}$ be decomposed as $\widetilde{H} = \widetilde{L} J \widetilde{L}^*$. Set $N = -L^{-1} \delta H L^{-*}$. Then there are at least m eigenvalues λ_{i_k}, $k = 1, \ldots, m$, of H for which*

$$\sqrt{\sum_{k=1}^{m} \left(\frac{|\lambda_{i_k} - \mu_k|}{\sqrt{|\lambda_{i_k}||\mu_k|}} \right)^2} \leq \|V\|_F \, \|\widetilde{V}\| \frac{\|N\|}{\sqrt{1 - \|N\|}}, \tag{16}$$

provided that the right hand side in (16) is positive. Here V and $\widetilde{V}$ are J-unitary matrices which simultaneously diagonalize the pairs (L^*L, J) and $(\widetilde{L}^*\widetilde{L}, J)$ as in (6), respectively.

Proof. Positivity of the right hand side in (16) implies $\|N\| < 1$. Therefore, J, $J + N$, $H = LJL^*$ and $\widetilde{H} = L(J+N)L^*$ have the same inertia. Therefore, the eigenvalue decompositions of H and $\widetilde{H}$ can be written as

$$H = U|\Lambda|^{1/2} J |\Lambda|^{1/2} U^*, \qquad \widetilde{H} = \widetilde{U} |\widetilde{\Lambda}|^{1/2} J |\widetilde{\Lambda}|^{1/2} \widetilde{U}^*. \tag{17}$$

The assumption $\|N\| < 1$ implies $\|NJ\| < 1$ which means that $I + NJ$ is nonsingular. According to [4, Theorem 6.4.12 (a)], the square roots of $I+NJ$ are well defined, and can be expressed as a polynomial in NJ.

Since $(I+NJ)^{1/2} = J[(I+NJ)^{1/2}]^* J$, the matrix $J+N$ can be decomposed as

$$J + N = (I + NJ)^{1/2} J [(I + NJ)^{1/2}]^*. \tag{18}$$

Thus, we can write $\widetilde{H}$ as

$$\widetilde{H} = L(I + NJ)^{1/2} J [(I + NJ)^{1/2}]^* L^* \equiv \widetilde{L} J \widetilde{L}^*\,, \tag{19}$$

where $\widetilde{L} = L(I + NJ)^{1/2}$

Further, (8) implies

$$L = U|\Lambda|^{1/2} V^{-1}. \tag{20}$$

Similarly, (17), (19) and (8) imply

$$\widetilde{L} = \widetilde{U}|\widetilde{\Lambda}|^{1/2} \widetilde{V}^{-1}. \tag{21}$$

Now (19) and (18) imply that

$$\begin{aligned} \widetilde{H} - H &= L(I + NJ)^{1/2} J [(I + NJ)^{1/2}]^* L^* - LJL^* \\ &= L(I + NJ)^{1/2}\, \Xi\, L^* = \widetilde{L}\, \Xi\, L^*, \end{aligned} \tag{22}$$

where

$$\Xi = J[(I + NJ)^{1/2}]^* - (I + NJ)^{-1/2} J = (I + NJ)^{-1/2} N. \tag{23}$$

Pre- and post-multiplication of (22) by $\widetilde{U}^*$ and U, respectively, together with eigenvalue decompositions (17), and relations (20) and (21), gives

$$\widetilde{\Lambda}\widetilde{U}^* U - \widetilde{U}^* U \Lambda = |\widetilde{\Lambda}|^{1/2} \widetilde{V}^{-1}\, \Xi\, V^{-*} |\Lambda|^{1/2}.$$

By interpreting this equality component-wise we have

$$\frac{\widetilde{\lambda}_p - \lambda_q}{\sqrt{|\widetilde{\lambda}_p||\lambda_q|}}\; S_{pq} = [\widetilde{V}^{-1}\, \Xi\, V^{-*}]_{pq},$$

where $S = \widetilde{U}^* U$ and $p, q \in \{1, \dots, n\}$. By taking the Frobenius norm we have

$$\sum_{p,q=1}^{n} \left(\frac{|\widetilde{\lambda}_p - \lambda_q|}{\sqrt{|\widetilde{\lambda}_p||\lambda_q|}} \right)^2 |S_{pq}|^2 = \|\widetilde{V}^{-1} \, \Xi \, V^{-*}\|_F^2.$$

Since $(|S_{pq}|^2)$ is a doubly stochastic matrix, by applying Lemma 1 we obtain

$$\sum_{p=1}^{n} \left(\frac{|\widetilde{\lambda}_p - \lambda_{\tau(p)}|}{\sqrt{|\widetilde{\lambda}_p||\lambda_{\tau(p)}|}} \right)^2 \le \|\widetilde{V}^{-1} \, \Xi \, V^{-*}\|_F^2 \tag{24}$$

for some permutation τ of $\{1, 2, \dots, n\}$. Further,

$$\|\widetilde{V}^{-1} \, \Xi \, V^{-*}\|_F \le \|\widetilde{V}\| \, \|V\|_F \, \|\Xi\|. \tag{25}$$

Relation (23) implies

$$\|\Xi\| \le \|(I + JN)^{-1/2}\| \, \|N\| \le \frac{1}{\sqrt{1 - \|N\|}} \|N\| \, .$$

In the proof of Theorem 2 we have shown that $\widetilde{H}$ has $\mathcal{X}$ as its invariant subspace. Thus, there are there are at least m eigenvalues λ_{i_k}, $k = 1, \dots, m$, of H such that

$$\sum_{k=1}^{m} \left(\frac{|\lambda_{i_k} - \mu_k|}{\sqrt{|\lambda_{i_k}||\mu_k|}} \right)^2 \le \sum_{p=1}^{n} \left(\frac{|\widetilde{\lambda}_p - \lambda_{\tau(p)}|}{\sqrt{|\widetilde{\lambda}_p||\lambda_{\tau(p)}|}} \right)^2 .$$

The theorem follows by combining this with (24), (25) and (23). ■

Now we can formulate Theorem 4 in terms of acute principal angles.

Theorem 5 *Assume the notation of Theorem 4. Let*

$$\mathcal{Y}_L = J L^* \mathcal{X}, \qquad \mathcal{Z}_L = L^{-1} \mathcal{X},$$

and let ψ be the maximal acute principal angle between $\mathcal{Y}_L$ and $\mathcal{Z}_L$. Then there are at least m eigenvalues λ_{i_k}, $k = 1, \dots, m$, of H for which

$$\sqrt{\sum_{k=1}^{m} \left(\frac{|\lambda_{i_k} - \mu_k|}{\sqrt{|\lambda_{i_k}||\mu_k|}} \right)^2} \le \|V\|_F \, \|\widetilde{V}\| \frac{|\sin \psi|}{\sqrt{(1 - \sin \psi)\,(1 - 2 \sin \psi)}}, \tag{26}$$

provided that the right hand side in (26) is positive.

Proof. The proof follows from $\|N\| = \|L^{-1}\delta H L^{-*}\|$, Remark 1 and (16). ■

If we wish to avoid existence of unperturbed and perturbed quantities V and $\widetilde{V}$, respectively, on the right-hand side of (16) and (26), we need to bound $\|\widetilde{V}\|$ in terms of $\|V\|$. For this purpose we need the following theorem by the authors [9, Theorem 5].

Theorem 6 *Let $\widetilde{L} = L(I+E)$ let V and $\widetilde{V}$ be nonsingular J-unitary matrices which simultaneously diagonalize the pairs (L^*L, J) and $(\widetilde{L}^*\widetilde{L}, J)$ as in (6), respectively. If*

$$\alpha \equiv \frac{\|E\|_F}{1-\|E\|} < \frac{1}{4\|V\|^2}, \tag{27}$$

then

$$\|\widetilde{V}\| \leq \frac{\|V\|}{\sqrt{1-4\alpha\|V\|^2}}. \tag{28}$$

Combining previous results gives our final theorem.

Theorem 7 *Assume the notation of Theorem 4. Let*

$$\mathcal{Y}_L = JL^*\mathcal{X}, \qquad \mathcal{Z}_L = L^{-1}\mathcal{X},$$

and let ψ be the maximal acute principal angle between $\mathcal{Y}_L$ and $\mathcal{Z}_L$. Let $N = -L^{-1}\delta H L^{-*}$ be such that

$$\beta \equiv \frac{\|N\|_F}{2\sqrt{1-\|N\|}-\|N\|} < \frac{1}{4\|V\|^2}. \tag{29}$$

Then there are at least m eigenvalues λ_{i_k}, $k = 1, \ldots, m$, of H for which

$$\sqrt{\sum_{k=1}^{m}\left(\frac{|\lambda_{i_k}-\mu_k|}{\sqrt{|\lambda_{i_k}||\mu_k|}}\right)^2} \leq \frac{\|V\|_F\,\|V\|}{\sqrt{1-4\beta\|V\|^2}}\frac{\|\sin\psi\|}{\sqrt{(1-\sin\psi)\,(1-2\sin\psi)}}. \tag{30}$$

Proof. In (19) we have defined

$$\widetilde{L} = L(I+NJ)^{1/2}.$$

The assumption $\|N\| < 1$ ensures the existence of $(I+NJ)^{1/2}$ defined by the following series [4, Theorem 6.2.8]

$$\begin{aligned} I + E &\equiv (I+NJ)^{1/2} \\ &= I + \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(2n-1)!!}{2^n n!} (NJ)^n. \end{aligned}$$

Here $(2n-1)!! = 1 \cdot 1 \cdot 3 \cdot 5 \cdots (2n-1)$. It is easy to see that

$$\|E\| \leq \frac{1}{2} \frac{\|N\|}{\sqrt{1-\|N\|}}\,.$$

Using the fact that β from (29) is the upper bound for α from (27), we can apply bound (28) for $\|\widetilde{V}\|$ with β in role of α. This together with (26) gives (30). ■

4 Comparison of the positive definite and the indefinite case

First, notice that our bound (15) is a proper generalization of the bound (4) to indefinite Hermitian matrices. Indeed, in the positive definite case the matrix V is unitary and the bound (15) is equal to (4).

Further, it is easy to show that in the positive definite case the angle function $\angle(\mathcal{Y}_L, \mathcal{Z}_L)$ defined by (12) does not depend on L but only on H (see [2]). However, in indefinite case this is not true in general. If we decompose matrix H as

$$H = L_1 J L_1 = L_2 J L_2 \tag{31}$$

then we can write

$$\begin{aligned} \mathcal{Y}_{L_1} &= J L_1^* \mathcal{X}, \qquad \mathcal{Z}_{L_1} = L_1^{-1} \mathcal{X} \\ \mathcal{Y}_{L_2} &= J L_2^* \mathcal{X}, \qquad \mathcal{Z}_{L_2} = L_2^{-1} \mathcal{X}\,. \end{aligned}$$

If $\psi_i = \angle(\mathcal{Y}_{L_i}, \mathcal{Z}_{L_i})$, $i = 1, 2$, it is of our interest to find out is there any connection between the angles ψ_1 and ψ_2.

¿From (31) it follows that there exists nonsingular J-unitary matrix W such that $L_2 = L_1 W$. ¿From Remark 1 it follows that

$$\|L_i^{-1}\delta H L_i^{-*}\| = \frac{\sin\psi_i}{1-\sin\psi_i}, \qquad i=1,2\,.$$

This, together with fact that $L_2 = L_1 W$, gives

$$\frac{\sin\psi_2}{1-\sin\psi_2} \le \|W\|^2 \frac{\sin\psi_1}{1-\sin\psi_1}\,.$$

Therefore, for ψ_1 and ψ_2 small enough, we have

$$\sin\psi_2 \lesssim \|W\|^2 \sin\psi_1.$$

We conclude that if the matrix W has moderate norm and if the subspaces ψ_1 and ψ_2 are sufficiently close, then the angle functions will be close, too.

The bounds of Section 3 depend on the spectral condition number or the norm of the J-unitary matrix V which diagonalizes the pair (L^*L, J). Although these quantities can be large, $\kappa(V)$ is bounded by [8, Theorem 3]:

$$\kappa(V) \le \min\sqrt{\kappa(\Delta^* L^* L \Delta)},$$

where the minimum is taken over all matrices which commute with J. Appropriate bounds for $\kappa(V)$ exist for some other classes of "well-behaved matrices" such as *scaled diagonal dominant matrices*, *block scaled diagonally dominant (BSDD) matrices* and *quasi-definite matrices*. Details of these bounds can be found in e.g. [12, Section 3.1] and [11].

5 Numerical example

Let $H = D^*(J+N)D$ be the nonsingular Hermitian matrix with

$$D = \begin{bmatrix} 2\cdot 10^4 & 8\cdot 10^4 & 0 & 0 \\ 2\cdot 10^3 & 4\cdot 10^4 & 0 & 0 \\ 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0.6 & 0.8 \end{bmatrix}, \qquad N = \begin{bmatrix} 0 & 0.08 & 0.01 & 0.03 \\ 0.08 & 0.06 & 0.05 & 0.03 \\ 0.01 & 0.05 & 0.08 & 0.04 \\ 0.03 & 0.03 & 0.04 & 0.04 \end{bmatrix},$$

and $J = \operatorname{diag}(1,1,-1,-1)$. All subsequent quantities are displayed properly rounded to the given number of decimal places. The spectrum of H is

$$\lambda(H) = \{8.9705\cdot 10^9, 4.8108\cdot 10^7, -1.9227, -0.11514\}.$$

Let

$$X = \begin{bmatrix} 0.19962 & 0 \\ 0.97987 & 0 \\ 0 & 0.78274 \\ 0 & 0.62235 \end{bmatrix}$$

be the orthonormal matrix. From (1) it follows that

$$M = \begin{bmatrix} 8.9705 \cdot 10^9 & 6.6073 \cdot 10^3 \\ 6.6073 \cdot 10^3 & -1.9149 \end{bmatrix},$$

with the spectrum

$$\lambda(M) = \{8.9704 \cdot 10^9, -1.9198\}.$$

The residual R is

$$R = HX - XM = \begin{bmatrix} 1.7432 \cdot 10^7 & -352.2 \\ -3.4906 \cdot 10^6 & 71.795 \\ -172.67 & 1.3242 \cdot 10^{-4} \\ 217.16 & -1.6214 \cdot 10^{-4} \end{bmatrix}$$

with $\|R\| \sim 1.8 \cdot 10^7$. Therefore, the residual bounds from the classical perturbation theory (2) and (3) are useless. Further, for δH from (13) we have $\|\delta H H^{-1}\| = 2.3 \cdot 10^3$, so the relative perturbation bounds which use the factor $\|\delta H H^{-1}\|$, like those from [1, Theorem 3], are also useless.

On the other hand, consider two decompositions of H: $H = L_1 J L_1^*$, where L_1 is obtained using Gaussian elimination, and $H = L_2 J L_2^*$, where L_2 is obtained from the eigenvalue decomposition of the matrix $J + N$. More precisely,

$$L_1 = \begin{bmatrix} 2.0264 \cdot 10^4 & 0 & 0 & 0 \\ 8.6931 \cdot 10^4 & 3.2418 \cdot 10^4 & 0 & 0 \\ 0.03435 & 0.0609 & -1.1057 & 0 \\ 0.03346 & 0.04231 & -0.79638 & -0.42555 \end{bmatrix}$$

and

$$L_2 = \begin{bmatrix} 1.3743 \cdot 10^4 & 1.4896 \cdot 10^4 & 2.7393 \cdot 10^2 & -2.0233 \cdot 10^2 \\ 8.2797 \cdot 10^4 & 4.1918 \cdot 10^4 & 1.9839 \cdot 10^3 & -7.8674 \cdot 10^2 \\ 3.7731 \cdot 10^{-2} & -8.2197 \cdot 10^{-3} & -1.1041 & -1.2542 \cdot 10^{-2} \\ 2.9716 \cdot 10^{-2} & -2.2666 \cdot 10^{-3} & -0.79993 & 0.41642 \end{bmatrix},$$

where $L_2 = D^* U_a |\Lambda_a|^{1/2}$ and $J + N = U_a |\Lambda_a| J U_a^*$ is the eigenvalue decomposition of $J + N$. We have $L_1 = L_2 W$, where $W^* J W = J$ and $\|W\| = 1.029$.

For the matrices V_1 and V_2 which diagonalize the pairs $(L_1 J L_1^*)$ and $(L_2 J L_2^*)$ as in (6) we have

$$\kappa(V_1) \approx 1, \qquad \kappa(V_2) = 1.059,$$

respectively. For the matrix δH defined by (13) we have

$$\|L_1^{-1} \delta H L_1^{-*}\| = 0.0498, \quad \|L_2^{-1} \delta H L_2^{-*}\| = 0.0475\,.$$

Therefore, Theorems 2 and 3 bound well the relative perturbation

$$\max \left\{ \frac{|\lambda_1(H) - \lambda_1(M)|}{|\lambda_1(H)|}, \frac{|\lambda_3(H) - \lambda_2(M)|}{|\lambda_3(H)|} \right\} = 0.0015.$$

Finally, Theorems 4, 5 and 7 bound equally well the relative perturbation

$$\sqrt{\frac{(\lambda_1(H) - \lambda_1(M))^2}{|\lambda_1(H)\lambda_1(M)|} + \frac{(\lambda_3(H) - \lambda_2(M))^2}{|\lambda_3(H)\lambda_2(M)|}} = 0.0015.$$

References

[1] Z. Drmač, On relative residual bounds for the eigenvalues of a Hermitian matrix, *Linear Algebra Appl.*, 244:155-163 (1996)

[2] Z. Drmač, V. Hari, Relative residual bounds for the eigenvalues of a Hermitian semidefinite matrix, *SIAM Journal on Matrix Analysis and Appl.*, 18:1:21–29 (1997).

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

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

[5] B. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

[6] R.-C. Li, Relative perturbation theory: (i) eigenvalue and singular value variations, *SIAM J. Matrix Anal. Appl.*, 19:956–982 (1998).

[7] I. Slapničar, Componentwise analysis of direct factorization of real symmetric and hermitian matrix, *Linear Algebra Appl.*, 272:227–275 (1998).

[8] I. Slapničar and K. Veselić, A bound for the condition of a hyperbolic eigenvector matrix, *Linear Algebra Appl.*, 290:247–255 (1999).

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

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

[11] N. Truhar and R.-C. Li, A $\sin 2\Theta$ theorem for graded indefinite Hermitian matrices, *Linear Algebra Appl.*, 359:263–276 (2003).

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

[13] K.Veselić and I. Slapničar, Floating–point perturbations of Hermitian matrices, *Linear Algebra Appl.*, 195:81–116 (1993).

[14] P.-Å. Wedin, On angles between subspaces of a finite dimensional inner product space, in: *Matrix Pencils*, B. Kågström and A. Ruhe, eds, Springer Lecture Notes in Mathematics 973, Springer, Berlin, (1983).