Optimal Perturbation Bounds for the Hermitian Eigenvalue Problem

Jesse L. Barlow*

Department of Computer Science and Engineering
The Pennsylvania State University
University Park, PA 16802-6106
e-mail: barlow@cse.psu.edu

Ivan Slapničar†

Faculty of Electrical Engineering, Mechanical Engineering, and Naval Architecture
University of Split
R. Boškovića b.b., 21000 Split, Croatia
e-mail: ivan.slapnicar@fesb.hr

Submitted by Krešimir Veselić

ABSTRACT

There is now a large literature on structured perturbation bounds for eigenvalue problems of the form

$$Hx = \lambda Mx,$$

where H and M are Hermitian. These results give relative error bounds on the ith eigenvalue, λ_i, of the form

$$\frac{|\lambda_i - \tilde{\lambda}_i|}{|\lambda_i|},$$

*The research of Jesse L. Barlow was supported by the National Science Foundation under grants no. CCR-9424435 and no. CCR-9732081.

†The research of Ivan Slapničar was supported by the Croatian Ministry of Science and Technology under grant no. 037012. Part of this work was done while the author was visiting the Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA.

655 Avenue of the Americas, New York, NY 10010 0024- 3795/93/$6.00

and bound the error in the ith eigenvector in terms of the relative gap,

$$\min_{j \neq i} \frac{|\lambda_i - \lambda_j|}{|\lambda_i \lambda_j|^{1/2}}.$$

In general, this theory usually restricts H to be nonsingular and M to be positive definite.

We relax this restriction by allowing H to be singular. For our results on eigenvalues we allow M to be positive semi-definite and for a few results we allow it to be more general. For these problems, for eigenvalues that are not zero or infinity under perturbation, it is possible to obtain local relative error bounds. Thus, a wider class of problems may be characterized by this theory.

Although it is impossible to give meaningful relative error bounds on eigenvalues that are not bounded away from zero, we show that the error in the subspace associated with those eigenvalues can be characterized meaningfully.

---

1. Introduction

We consider the eigenvalue problem

$$Hx = \lambda M x, \quad H, M \in \mathbf{C}^{n \times n}, \quad x \in \mathbf{C}^n, \quad \lambda \in \mathbf{C}, \tag{1.1}$$

where H and M are Hermitian matrices. We assume that there exists a nonsingular matrix $X \in \mathbf{C}^{n \times n}$ such that

$$X^* H X = \Omega, \quad X^* M X = J, \tag{1.2}$$

where

$$\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_n), \quad J = \mathrm{diag}(j_1, \ldots, j_n),$$

and

$$\omega_i \in \mathbf{R}, \quad j_i \in \{e^{i\theta} : \theta \in [0, 2\pi]\} \cup \{0\},\ i = 1, 2, \ldots, n.$$

If we restrict J to be nonsingular, then (1.1) is a statement of eigenproblem for the matrix $A = HM^{-1}$. If we impose the restriction

$$\omega_i \in \mathbf{R}, \quad j_i \in \{0, 1\},$$

then M is positive semi-definite, and (1.1) is the generalized Hermitian eigenvalue problem. Most of the results in this paper concern this class of eigenvalue problems.

We compare (1.1) to the perturbed problem

$$(H + \Delta H)\tilde{x} = \tilde{\lambda}(M + \Delta M)\tilde{x} \tag{1.3}$$

where ΔH and ΔM are "small" Hermitian perturbations. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of the pencil (1.1) and let $\tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_n$ be eigenvalues of the perturbed pencil (1.3). Starting from the theory of Kato [12], we obtain meaningful bounds on

$$\frac{|\lambda_i - \tilde{\lambda}_i|}{|\lambda_i|}. \tag{1.4}$$

Moreover, for the case when M is positive definite, we give conditions under which we can bound the error in the subspaces in terms of a generalization of the relative gap

$$\mathrm{relgap}(\lambda_i) = \min_{j \neq i} \frac{|\lambda_i - \lambda_j|}{|\lambda_i \lambda_j|^{\frac{1}{2}}}.$$

This theory generalizes that in papers by Barlow and Demmel [1], Demmel and Veselić [5], Veselić and Slapničar [18], and some of the results by Gu and Eisenstat [11], Li [13, 14], and Eisenstat and Ipsen [7].

We make the following improvements to the theory given in the above papers:

- The bounds on eigenvalues allow for H and M to be singular. These bounds are used to obtain bounds on the singular value decomposition (SVD).
- The bounds given are local in the sense that each eigenvalue has its own condition number.
- The bounds given are optimal and show clearly the role of structured perturbations.
- The bounds on eigenvectors include bounds on the error in the subspace associated with eigenvalues that are not bounded away from zero.

In §2, we give simple bounds for the relative error of the form (1.4) under weaker assumptions than have been given in previous works [1, 18, 11, 10] and show how this theory can be applied to the singular value decomposition. In §3, we show how this theory accounts for the effect of structured perturbation on the problem (1.1). In §4, we give bounds on error in subspaces for scaled perturbations. Some examples are given in §5 and our conclusions are in §6.

2. Locally optimal perturbation bounds on Hermitian pencils

In this section we first give local condition numbers for eigenvalues. We then derive the perturbation bounds for the singular value decomposition.

2.1. Local condition numbers of eigenvalues

Consider the perturbed pair

$$\tilde{H} \equiv H + \Delta H = H + \delta E, \quad \tilde{M} \equiv M + \Delta M = M + \delta F, \tag{2.1}$$

where δ is a positive real number, $E = \Delta H/\delta$ and $F = \Delta M/\delta$. For $\zeta \in [0, \delta]$ let

$$H(\zeta) = H + \zeta E, \quad M(\zeta) = M + \zeta F. \tag{2.2}$$

Now consider the family of generalized eigenproblems

$$H(\zeta)x(\zeta) = \lambda(\zeta)M(\zeta)x(\zeta), \quad \zeta \in [0, \delta]. \tag{2.3}$$

We assume that (1.2) holds for each $\zeta \in [0, \delta]$ and some $X(\zeta)$, $\Omega(\zeta)$ and $J(\zeta)$. Let $(\lambda_i(\zeta), x_i(\zeta))$ be the ith eigenpair of (2.3). Define $\mathcal{S}$ to be the set of indices given by

$$\mathcal{S} = \{i \colon H(\zeta)x_i(\zeta) \neq 0, M(\zeta)x_i(\zeta) \neq 0 \text{ for all } \zeta \in [0, \delta]\}. \tag{2.4}$$

The set $\mathcal{S}$ is the set of eigenvalues for which relative error bounds can be found. The next theorem gives such a bound. Its proof follows that of Theorem 4 in [1, p.773].

THEOREM 2.1. *Let $(\lambda_i(\zeta), x_i(\zeta))$ be the ith eigenpair of the Hermitian pencil in (2.3). Let $\mathcal{S}$ be defined by (2.4). If $i \in \mathcal{S}$, then*

$$\frac{\lambda_i(\delta) - \lambda_i(0)}{\lambda_i(0)} = \exp\left(\int_0^\delta \rho_i(\zeta) d\zeta\right), \tag{2.5}$$

where

$$\rho_i(\zeta) = \frac{x_i^*(\zeta)Ex_i(\zeta)}{x_i^*(\zeta)H(\zeta)x_i(\zeta)} - \frac{x_i^*(\zeta)Fx_i(\zeta)}{x_i^*(\zeta)M(\zeta)x_i(\zeta)}. \tag{2.6}$$

Proof. Assume that $\lambda_i(\zeta)$ is simple at the point ζ. Then from the classical eigenvalue perturbation theory, we have the first order expansion

$$\lambda_i(\zeta + \xi) = \lambda_i(\zeta) + \xi\dot{\lambda}_i(\zeta) + O(\xi^2). \tag{2.7}$$

Let us compute $\dot{\lambda}(\zeta)$. By differentiating the equation

$$[H + (\zeta + \xi)E]x_i(\zeta + \xi) = \lambda_i(\zeta + \xi)[M + (\zeta + \xi)F]x_i(\zeta + \xi)$$

with respect to ξ, and setting $\xi = 0$ in the result, we obtain

$$Ex_i(\zeta) + H(\zeta)\dot{x}_i(\zeta) = \dot{\lambda}_i(\zeta)M(\zeta)x_i(\zeta) + \lambda_i(\zeta)Fx_i(\zeta) + \lambda_i(\zeta)M(\zeta)\dot{x}_i(\zeta).$$

Applying $x_i^*(\zeta)$ to the left side of this equation, using (2.3), and rearranging gives

$$\dot{\lambda}_i(\zeta) = \frac{x_i^*(\zeta)Ex_i(\zeta)}{x_i^*(\zeta)M(\zeta)x_i(\zeta)} - \frac{\lambda_i(\zeta)x_i^*(\zeta)Fx_i(\zeta)}{x_i^*(\zeta)M(\zeta)x_i(\zeta)}.$$

Since $\lambda_i(\zeta) \neq 0$ for all ζ, dividing by $\lambda_i(\zeta)$ gives

$$\frac{d(\ln \lambda_i(\zeta))}{d\zeta} = \frac{\dot{\lambda}_i(\zeta)}{\lambda_i(\zeta)} = \frac{x_i^*(\zeta)Ex_i(\zeta)}{x_i^*(\zeta)H(\zeta)x_i(\zeta)} - \frac{x_i^*(\zeta)Fx_i(\zeta)}{x_i^*(\zeta)M(\zeta)x_i(\zeta)} = \rho_i(\zeta), \quad (2.8)$$

where ln is the complex version of the natural logarithm function.

If $\lambda_i(\zeta)$ is simple for all $\zeta \in [0, \delta]$, then the bound (2.5) follows by integrating from 0 to δ. In Kato [12, Theorem II.6.1,p.139], it is shown that the eigenvalues of $H(\zeta)$ in $\mathcal{S}$ are real analytic, even when they are multiple. Moreover, Kato goes on to point out that there are only a finite number of ζ where $\lambda_i(\zeta)$ is multiple, so that $\lambda_i(\zeta)$ is continuous and piecewise analytic throughout the interval $[0, \delta]$. Thus we can obtain (2.5) by integrating (2.8) over each of the intervals in which $\lambda_i(\zeta)$ is analytic. ■

Our usual way of interpreting this bound comes from the following corollary.

COROLLARY 2.1. *Assume the hypothesis and terminology of Theorem 2.1. For all $i \in \mathcal{S}$ we have*

$$\left| \frac{\lambda_i(\delta) - \lambda_i(0)}{\lambda_i(0)} \right| \leq |\exp(\delta\kappa_i) - 1|,$$

where

$$\kappa_i = \max_{\zeta\in[0,\delta]} |\mathrm{Re}(\rho_i(\zeta)| + \mathbf{i} \max_{\zeta\in[0,\delta]} |\mathrm{Im}(\rho_i(\zeta))|.$$

Neither the exact expression Theorem 2.1 nor the bound in Corollary 2.1 is computable. We can only hope to compute the function $\rho_i(\zeta)$ at one point, $\zeta = 0$.

Thus we will always use the computable first order approximation[‡]

$$\hat{\kappa}_i = |\mathrm{Re}(\rho_i(0))| + \mathbf{i}|\mathrm{Im}(\rho_i(0))|.$$

If this value is large, then the corresponding eigenvalue is sensitive. Moreover, if δ is sufficiently small, then this approximation will give us appropriate qualitative information.

[‡]This approximation can be computed by switching the roles of the perturbed and unperturbed problem. The backward errors E and F simply change signs, and the "original" vector $x_i(0)$ becomes the computed one.

A very important issue in this paper is whether or not an index i is in the set $\mathcal{S}$. From the proof of Theorem 2.1, it can easily be inferred that $i \in \mathcal{S}$ if and only if $\rho_i(\zeta)$ in (2.6) is bounded for each ζ. That, of course, is not possible to verify since we cannot compute $\rho_i(\zeta)$ for every value of ζ. We will give a heuristic criterion for membership in this set, justified below. The discussion below assumes that $\lambda_i(\zeta)$ is simple and that we can use an argument like that in Theorem 2.1 to consider the case where $\lambda_i(\zeta)$ is not simple.

First, consider the case where $H(\zeta)x_i(\zeta) = 0$ for some ζ, but $M(\zeta)x_i(\zeta) \neq 0$ for all ζ. Then $\lambda_i(\zeta) = 0$ for some ζ. By the mean value theorem for some $\xi \in [0, \zeta]$, we have

$$\lambda_i(\zeta) = \lambda_i(0) + \zeta\dot{\lambda}_i(\xi) = 0.$$

Thus

$$\lambda_i(0) = -\zeta\dot{\lambda}_i(\xi).$$

If

$$|\lambda_i(0)| > \delta \max_{\zeta \in [0,\delta]} |\dot{\lambda}_i(\xi)|,$$

then $i \in \mathcal{S}$. Of course, this is not verifiable either. Thus a good heuristic is

$$|\lambda_i(0)| \gg \delta|\dot{\lambda}_i(0)| \Rightarrow i \in \mathcal{S}.$$

or equivalently,

$$\delta|\hat{\kappa}_i| = \delta|\rho_i(0)| \ll 1 \Rightarrow i \in \mathcal{S}. \tag{2.9}$$

If we assume that $M(\zeta)x_i(\zeta) = 0$ for some ζ, but $H(\zeta)x_i(\zeta) \neq 0$ for all ζ, similar reasoning arrives again at the heuristic (2.9). The only way to know for certain if $i \in \mathcal{S}^c$, the complement of $\mathcal{S}$, is to discover a perturbation in the class of interest that makes either $H(\zeta)x_i(\zeta) = 0$ or $M(\zeta)x_i(\zeta) = 0$.

It is entirely possible that $H(\zeta)x_i(\zeta) = M(\zeta)x_i(\zeta) = 0$ and that $\rho_i(\zeta)$ is defined in the limit and bounded for all ζ. This is the so-called $(0,0)$-eigenvalue case. To exclude this case as well, we may wish to impose the slightly stricter criterion

$$\delta\left(\frac{|x_i^*(0)Ex_i(0)|}{|x_i^*(0)H(0)x_i(0)|} + \frac{|x_i^*(0)Fx_i(0)|}{|x_i^*(0)M(0)x_i(0)|}\right) \ll 1. \tag{2.10}$$

If (2.10) is not true, then it is unlikely that reasonable relative error bound on λ_i is obtainable. Absolute error bounds are unlikely to be improved upon for such eigenvalues. If (2.10) holds, then the estimate

$$\left|\frac{\lambda_i(\delta) - \lambda_i(0)}{\lambda_i(0)}\right| \leq |\exp(\delta\hat{\kappa}_i) - 1| + O(\delta^2)$$

is accurate and computable. We demonstrate this point with an example.

EXAMPLE 2.1. We take $M = I$ and $H = DAD$ where

$$D = \mathrm{diag}(1, 10^8, 10^4, 10^6)$$

and

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

The matrix H has exactly one zero eigenvalue. We then take

$$E = |H|.$$

To the digits displayed, MATLAB computes the eigenvalues

$$\Lambda = \mathrm{diag}(2 \cdot 10^{16}, 1.0001 \cdot 10^{12}, 4.9995 \cdot 10^7, 3.05062 \cdot 10^{-17}).$$

The last eigenvalue should be zero, but is not. The clue here is that the corresponding values of $\hat{\kappa}_i, i = 1, 2, 3, 4$, are

$$\hat{\kappa} = (1, 1, 13, 3.93 \cdot 10^{17}).$$

If we choose $\delta = 2.2204 \cdot 10^{-16}$, the relative distance between two floating point numbers in IEEE double precision, we see that

$$\delta\hat{\kappa}_4 = 87 > 1,$$

thus the last eigenvalue cannot be reliably separated from zero, and the small componentwise perturbation $\tilde{H} = H - \delta|H|$ makes this eigenvalue negative. However, all of the other eigenvalues are well-separated from zero. The non-zero eigenvalues of H are all well-behaved under componentwise perturbations. We will show later (Example 5.5) that the subspace associated with the zero eigenvalue is also well-behaved.

Since for Hermitian eigenvalue problems, the eigenvalues will be real, we can will use the following simplification of Corollary 2.1.

COROLLARY 2.2. *Assume the hypothesis and terminology of Theorem 2.1. If* $\lambda_i(\zeta) \in \mathbf{R}$ *for all* $\zeta \in [0, \delta]$ *and* $i \in \mathcal{S}$ *then*

$$\exp(-\delta\kappa_i) \leq \frac{\lambda_i(\delta)}{\lambda_i(0)} \leq \exp(\delta\kappa_i), \tag{2.11}$$

where

$$\kappa_i = \max_{\zeta \in [0,\delta]} \left| \frac{x_i^*(\zeta) E x_i(\zeta)}{x_i^*(\zeta) H(\zeta) x_i(\zeta)} - \frac{x_i^*(\zeta) F x_i(\zeta)}{x_i^*(\zeta) M(\zeta) x_i(\zeta)} \right|.$$

Our bound improves similar results from [1, 5, 6, 18] since it is applicable to the larger class of matrix pairs, accommodates the case when one or both matrices are singular, and gives nearly optimal local condition number for each non-zero eigenvalue.

If the perturbations E and F are not structured in any particular way with regard to $x_i(\zeta)$, then the bound from Theorem 2.1 will not be much better than classical normwise bounds [19, 17, 9]. For structured perturbations the first order approximation of our bound can be much sharper than the classical bounds and relative bounds from [1, 5, 6, 18], as shown in the examples of §5.

2.2. The singular value decomposition

The singular value decomposition (SVD) of a matrix $A \in \mathbf{C}^{m\times n}$ is given by

$$A = U\Sigma V^*,$$

where U and V are unitary, and $\Sigma = \mathrm{diag}(\sigma_i)$ diagonal and nonnegative. For simplicity, we assume $m \geq n$, for $m < n$, analogous results follow by considering A^*.

COROLLARY 2.3. *Let $\tilde{A} = A + \Delta A = A + \delta E$, where $E = \Delta A/\delta$. Define $A(\zeta) = A + \zeta E$ for $\zeta \in [0,\delta]$. Let $A(\zeta)$ have the singular value decomposition*

$$A(\zeta) = U(\zeta)\Sigma(\zeta)V(\zeta)^*, \quad \zeta \in [0,\delta],$$

where $U(\zeta) \in \mathbf{C}^{m\times m}$ and $V(\zeta) \in \mathbf{C}^{n\times n}$ are unitary and

$$\begin{aligned}
\Sigma(\zeta) &= \mathrm{diag}(\sigma_1(\zeta), \ldots, \sigma_n(\zeta)), \\
U(\zeta) &= (u_1(\zeta), \ldots, u_m(\zeta)), \\
V(\zeta) &= (v_1(\zeta), \ldots, v_n(\zeta)).
\end{aligned}$$

Then for each index i such that $\sigma_i(\zeta) \neq 0$ for all $\zeta \in [0,\delta]$ we have

$$\exp(-\delta\kappa_i) \leq \frac{\sigma_i(\delta)}{\sigma_i(0)} \leq \exp(\delta\kappa_i), \tag{2.12}$$

where

$$\kappa_i = \max_{\zeta\in[0,\delta]} \left| \frac{\mathrm{Re}(u_i^*(\zeta)Ev_i(\zeta))}{u_i^*(\zeta)A(\zeta)v_i(\zeta)} \right|. \tag{2.13}$$

Proof. We note that $\sigma_i(\zeta)$ is the ith eigenvalue of (1.1) with $M = I$ and $H(\zeta)$ given by

$$H(\zeta) = \begin{pmatrix} 0 & A(\zeta) \\ A(\zeta)^* & 0 \end{pmatrix}.$$

The corresponding eigenvector is

$$x_i(\zeta) = \frac{1}{\sqrt{2}} \begin{pmatrix} u_i(\zeta) \\ v_i(\zeta) \end{pmatrix}.$$

If we evaluate $\rho_i(\zeta)$ in (2.6) we obtain

$$\rho_i(\zeta) = \frac{x_i^*(\zeta) \begin{pmatrix} 0 & E \\ E^* & 0 \end{pmatrix} x(\zeta)}{x_i^*(\zeta) H(\zeta) x_i(\zeta)} = \frac{\mathrm{Re}(u_i^*(\zeta) E v_i(\zeta))}{u_i^*(\zeta) A(\zeta) v_i(\zeta)}.$$

If we then apply Corollary 2.2, we have (2.13). ■

3. Effect of structured perturbations

In this section, we discuss the effect of common structured errors. For this part of the theory we state the results for the SVD and the Hermitian pencil (H, M). Similar bounds which can be derived for generalizations of the SVD are given in [2].

If we had exact expressions for the perturbation matrices E and F from (2.1), then Theorem 2.1 and Corollary 2.1 could always be used to give a good bound. However, it is extremely rare that a perturbation resulting from either data or an algorithm is known exactly. Usually, we just have a bound for it, and with some luck, we have a structured bound. In this section, we discuss what can be determined from such a bound.

Two structures of perturbations are discussed, although there are others. The first structure discussed are scaled perturbations, that is, when E has the form (and analogously F)

$$E = D^* E_S D, \quad \|E_S\| \leq 1.$$

Such perturbations are common in the discussion of Jacobi-type methods [5]. When discussing the SVD, we may discuss the two-sided perturbation

$$E = D_L^* E_S D_R, \quad \|E_S\| \leq 1.$$

However, for simplicity, we will just simply discuss the perturbation

$$E = E_S D, \quad \|E_S\| \leq 1.$$

Our results are easy to generalize to the two-sided perturbations.

The other structure discussed here are componentwise perturbations. Here we assume that

$$|E| \leq |H|, \quad |F| \leq |M|,$$

where both the inequality and the absolute value are componentwise. Such perturbations are common for highly structured eigenvalue problems and for data perturbations.

3.1. Structured perturbations of the SVD

We suppose that

$$\tilde{A} = A + \Delta A = A + \delta E, \quad E = E_S D, \tag{3.1}$$

where D is some right grading matrix and $\|E_S\| \leq 1$. An important context for this class of perturbations is the analysis of Jacobi-type methods [5] where

$$D = \mathrm{diag}(\|A(:,1)\|, \|A(:,2)\|, \ldots, \|A(:,n)\|).$$

However, the results here allow D to be any matrix.

For $\zeta \in [0, \delta]$ let $A(\zeta)$ and its SVD be defined as in Corollary 2.3. As before, let $\mathcal{S}$ be the set of indices for which $\sigma_i(\zeta) \neq 0$ for all $\zeta \in [0, \delta]$. We now introduce the notion of a truncated SVD. In this case, we truncate with respect to the index set $\mathcal{S}$.

Definition 3.1. Let $k = |\mathcal{S}|$, and let the singular values of $A(\zeta)$ whose indices are in $\mathcal{S}$ correspond to singular values $\sigma_1(\zeta), \ldots, \sigma_k(\zeta)$ in non-increasing order. The truncated SVD of $A(\zeta)$ with respect to $\mathcal{S}$ is defined as

$$A(\zeta; \mathcal{S}) = U(\zeta)\Sigma(\zeta; \mathcal{S})V^*(\zeta),$$

where

$$\Sigma(\zeta; \mathcal{S}) = \mathrm{diag}(\sigma_1(\zeta), \sigma_2(\zeta), \ldots, \sigma_k(\zeta), 0, \ldots, 0).$$

It is also appropriate to define the Moore-Penrose pseudoinverse of $A(\zeta; \mathcal{S})$. For a fixed matrix $A \in \mathbf{C}^{m \times n}$, the Moore-Penrose pseudoinverse is the unique matrix $A^\dagger \in \mathbf{C}^{n \times m}$ satisfying the four Penrose conditions

$$\begin{array}{ll} 1.\ AA^\dagger A = A, & 3.\ (AA^\dagger)^* = AA^\dagger, \\ 2.\ A^\dagger AA^\dagger = A^\dagger, & 4.\ (A^\dagger A)^* = A^\dagger A. \end{array}$$

It is easily verified that the Moore-Penrose pseudoinverse of $A(\zeta; \mathcal{S})$ is given by

$$A^\dagger(\zeta; \mathcal{S}) = V(\zeta)\Sigma^\dagger(\zeta; \mathcal{S})U^*(\zeta),$$

where

$$\Sigma^\dagger(\zeta; \mathcal{S}) = \mathrm{diag}(\sigma_1^{-1}(\zeta), \ldots, \sigma_k^{-1}(\zeta), 0, \ldots, 0),$$

and k is as specified in Definition 3.1. We now use this form to establish global error bounds for all $\sigma_i, i \in \mathcal{S}$.

PROPOSITION 3.1. *For $\zeta \in [0, \delta]$ let $A(\zeta) = A + \zeta E_S D$, where E_S and D are defined by (3.1), have the singular value decomposition assumed in Corollary 2.3. Then (2.12) holds for each $i \in \mathcal{S}$ with κ_i bounded by*

$$\kappa_i \leq \max_{\zeta \in [0,\delta]} \|DA^{\dagger}(\zeta; \mathcal{S}) u_i(\zeta)\| \leq \max_{\zeta \in [0,\delta]} \|DA^{\dagger}(\zeta; \mathcal{S})\|.$$

Proof. From (2.13), for each $i \in \mathcal{S}$ we have

$$\kappa_i = \max_{\zeta \in [0,\delta]} \frac{|\mathrm{Re}(u_i^*(\zeta) E_S D v_i(\zeta))|}{|u_i^*(\zeta) A(\zeta) v_i(\zeta)|}. \tag{3.2}$$

Using the fact that $\|u_i^*(\zeta) E_S\| \leq 1$ with (3.2) yields

$$\kappa_i \leq \max_{\zeta \in [0,\delta]} \frac{\|D v_i(\zeta)\|}{\sigma_i(\zeta)}. \tag{3.3}$$

By the definition of $A^{\dagger}(\zeta; \mathcal{S})$ we have

$$v_i(\zeta) = A^{\dagger}(\zeta; \mathcal{S}) u_i(\zeta) \sigma_i(\zeta). \tag{3.4}$$

Combining (3.3) with (3.4) yields the desired result. ■

The following corollary is a componentwise error bound that we might expect from singular value improvement procedures. Its proof is very similar to the scaled case.

COROLLARY 3.1. *Let*

$$\tilde{A} = A + \Delta A = A + \delta E, \qquad |E| \leq |A|.$$

Here both the inequality and the absolute value are componentwise. For $\zeta \in [0, \delta]$ let $A(\zeta) = A + \zeta E$, have the singular value decomposition assumed in Corollary 2.3. Then (2.12) holds for each $i \in \mathcal{S}$ with κ_i bounded by

$$\kappa_i \leq \max_{\zeta \in [0,\delta]} \|\,|A|\,|A^{\dagger}(\zeta; \mathcal{S}) u_i(\zeta)|\,\| \leq \max_{\zeta \in [0,\delta]} \|\,|A|\,|A^{\dagger}(\zeta; \mathcal{S})|\,\|.$$

Proposition 3.1 and Corollary 3.1 generalize the corresponding results by Demmel and Veselić [5] and Li [13] to matrices which do not necessarily have full rank. Corollary 3.1 is illustrated by Example 5.1 in §5.

3.2. Structured perturbations for the Hermitian generalized eigenvalue problem

We consider the Hermitian pencil $H - \lambda M$, where M is positive semidefinite, and the perturbed pencil $\tilde{H} - \lambda \tilde{M}$, where $\tilde{H}$ and $\tilde{M}$ are defined by

(2.1). Suppose that the family of pencils $H(\zeta)-\lambda M(\zeta), \zeta \in [0,\delta]$, defined by (2.2), has the form

$$H(\zeta) = X^{-*}(\zeta)\Lambda(\zeta)X^{-1}(\zeta), \quad M(\zeta) = X^{-*}(\zeta)J(\zeta)X^{-1}(\zeta), \tag{3.5}$$

where

$$\begin{aligned} X(\zeta) &= (x_1(\zeta),\ldots,x_n(\zeta)), \\ \Lambda(\zeta) &= \operatorname{diag}(\lambda_1(\zeta),\ldots,\lambda_n(\zeta)), \\ J(\zeta) &= \operatorname{diag}(j_1(\zeta),\ldots,j_n(\zeta)). \end{aligned}$$

Here

$$j_i(\zeta) = \begin{cases} 0 & \text{if } x_i(\zeta) \in \operatorname{Null}(M(\zeta)) \ , \\ 1 & \text{otherwise.} \end{cases}$$

As done in Veselić and Slapničar [18], we relate our problem to a positive definite eigenvalue problem. To do so, we first define the spectral absolute value of the matrix $H(\zeta)$ with respect to $M(\zeta)$.

Definition 3.2. Let For $\zeta \in [0,\delta]$ let the pair $(H(\zeta), M(\zeta))$ have the generalized eigendecomposition in (3.5). The spectral absolute value of $H(\zeta)$ with respect to $M(\zeta)$ is the matrix $\| H(\zeta) \|_M$ given by

$$\| H(\zeta) \|_M = X^{-*}(\zeta)|\Lambda(\zeta)|X^{-1}(\zeta),$$

where $|\Lambda(\zeta)| = \operatorname{diag}(|\lambda_1(\zeta)|,\ldots,|\lambda_n(\zeta)|)$. If $M(\zeta) = I$ for all ζ, then we define $\| H(\zeta) \| = \| H(\zeta) \|_I$.

If we let $X^{-1}(\zeta)$ have the factorization

$$X^{-1}(\zeta) = Q(\zeta)R(\zeta),$$

where $Q(\zeta)$ is unitary, then it is easily seen that

$$\| H(\zeta) \|_M = R^*(\zeta) \| R^{-*}(\zeta)H(\zeta)R^{-1}(\zeta) \| R(\zeta).$$

This is the definition given by Veselić and Slapničar [18] for the case where M is nonsingular. We also note that for the case $M = I$, we have

$$\| H(\zeta) \| = \sqrt{H^2(\zeta)},$$

where $\sqrt{\cdot}$ denotes matrix square root, that is, $\| H(\zeta) \|$ is the positive semi-definite polar factor of $H(\zeta)$.

We will now define a truncated version of $\| H(\zeta) \|_M$. Define $\mathcal{S}$ as in (2.4).

DEFINITION 3.3. The truncated spectral absolute value of $H(\zeta)$ with respect to $M(\zeta)$ is the matrix $|H(\zeta;\mathcal{S})|_M$ such that

$$|H(\zeta;\mathcal{S})|_M = X^{-*}(\zeta)|\Lambda(\zeta;\mathcal{S})|X^{-1}(\zeta),$$

where

$$\Lambda(\zeta;\mathcal{S}) = \mathrm{diag}(\lambda_1(\zeta;\mathcal{S}),\ldots,\lambda_n(\zeta;\mathcal{S}))$$

and

$$\lambda_i(\zeta;\mathcal{S}) = \begin{cases} \lambda_i(\zeta), & i \in \mathcal{S}, \\ 0, & \text{otherwise.} \end{cases}$$

We define $M(\zeta;\mathcal{S})$ and $J(\zeta;\mathcal{S})$ conformally.

Clearly, $|H(\zeta;\mathcal{S})|_M$ is positive semi-definite. We can factor both $|H(\zeta;\mathcal{S})|_M$ and $M(\zeta;\mathcal{S})$ into the form

$$\begin{aligned} |H(\zeta;\mathcal{S})|_M &= C^*(\zeta;\mathcal{S})C(\zeta;\mathcal{S}), \\ M(\zeta;\mathcal{S}) &= G^*(\zeta;\mathcal{S})G(\zeta;\mathcal{S}), \end{aligned} \tag{3.6}$$

respectively, where

$$\begin{aligned} C(\zeta;\mathcal{S}) &= U(\zeta)\Phi(\zeta;\mathcal{S})X^{-1}(\zeta) \in \mathbf{C}^{m\times n}, \quad m \le n, \\ G(\zeta;\mathcal{S}) &= V(\zeta)J(\zeta;\mathcal{S})X^{-1}(\zeta) \in \mathbf{C}^{p\times n}, \quad p \le n. \end{aligned} \tag{3.7}$$

Here

$$\Phi(\zeta;\mathcal{S}) = \mathrm{diag}(\phi_i(\zeta;\mathcal{S})) = \mathrm{diag}(\sqrt{|\lambda_i(\zeta;\mathcal{S})|}),$$

and $U(\zeta)$ and $V(\zeta)$ are matrices with orthonormal rows and orthonormal nontrivial columns. That is, columns of $U(\zeta)$ which correspond to $i \in \mathcal{S}$ are orthonormal, and columns of $V(\zeta)$ for which $j_i(\zeta) = 1$ are orthonormal. Note that the form (3.7) describes the quotient singular value decomposition (QSVD) of the pair $(C(\zeta;\mathcal{S}), G(\zeta;\mathcal{S}))$ [15, 16]. The $G(\zeta;\mathcal{S})$-weighted pseudoinverse of $C(\zeta;\mathcal{S})$ [3, 8] is given by

$$C_G^\dagger(\zeta;\mathcal{S}) \equiv X(\zeta)\Phi^\dagger(\zeta;\mathcal{S})U^*(\zeta). \tag{3.8}$$

Likewise, the $C(\zeta;\mathcal{S})$-weighted psuedoinverse of $G(\zeta;\mathcal{S})$ is

$$G_C^\dagger(\zeta;\mathcal{S}) \equiv X(\zeta)J^\dagger(\zeta;\mathcal{S})V^*(\zeta).$$

Using this structure, we can establish bounds on all of the eigenvalues that do not change sign under the perturbation.

Theorem 3.1. *Let the pair $(\tilde{H}, \tilde{M})$ be defined by (2.1). For $\zeta \in [0, \delta]$ let $(\lambda_i(\zeta), x_i(\zeta))$ be the ith eigenpair of the pair $(H(\zeta), M(\zeta))$, defined by (2.2). Let $C(\zeta; \mathcal{S})$ and $G(\zeta; \mathcal{S})$ be defined by (3.6), and let the QSVD of $(C(\zeta; \mathcal{S}), G(\zeta; \mathcal{S}))$ be given by (3.7). Define $\mathcal{S}$ as in (2.4). Then each $\lambda_i(\zeta), i \in \mathcal{S}$, satisfies (2.11), where*

$$
\begin{aligned}
\kappa_i \quad &\leq \quad \max_{\zeta \in [0,\delta]} \frac{|x_i^*(\zeta) E x_i(\zeta)|}{x_i^*(\zeta)\, |H(\zeta;\mathcal{S})|_M x_i(\zeta)} + \max_{\zeta \in [0,\delta]} \frac{|x_i^*(\zeta) F x_i(\zeta)|}{x_i^*(\zeta)\ M(\zeta) x_i(\zeta)} \qquad (3.9)\\
&= \quad \max_{\zeta \in [0,\delta]} |u_i^*(\zeta)[C_G^\dagger(\zeta;\mathcal{S})]^* E C_G^\dagger(\zeta;\mathcal{S}) u_i(\zeta)| \\
&\quad + \max_{\zeta \in [0,\delta]} |v_i^*(\zeta)[G_C^\dagger(\zeta;\mathcal{S})]^* F G_C^\dagger(\zeta;\mathcal{S}) v_i(\zeta)| \\
&\leq \quad \max_{\zeta \in [0,\delta]} \|[C_G^\dagger(\zeta;\mathcal{S})]^* E C_G^\dagger(\zeta;\mathcal{S})\| + \max_{\zeta \in [0,\delta]} \|[G_C^\dagger(\zeta;\mathcal{S})]^* F G_C^\dagger(\zeta;\mathcal{S})\|.
\end{aligned}
$$

Proof. For each $i \in \mathcal{S}$ considering (2.11) as in Corollary 2.2 yields

$$
\begin{aligned}
\kappa_i \quad &= \quad \max_{\zeta \in [0,\delta]} \frac{|x_i^*(\zeta) E x_i(\zeta)|}{|x_i^*(\zeta) H(\zeta) x_i(\zeta)|} + \max_{\zeta \in [0,\delta]} \frac{|x_i^*(\zeta) F x_i(\zeta)|}{|x_i^*(\zeta) M(\zeta) x_i(\zeta)|} \\
&= \quad \max_{\zeta \in [0,\delta]} \frac{|x_i^*(\zeta) E x_i(\zeta)|}{\phi_i(\zeta)^2} + \max_{\zeta \in [0,\delta]} \frac{|x_i^*(\zeta) F x_i(\zeta)|}{j_i(\zeta)^2}.
\end{aligned}
$$

By (3.6) and (3.7) this is just the first equality in (3.9). Since

$$x_i(\zeta) = \phi_i(\zeta) C_G^\dagger(\zeta;\mathcal{S}) u_i(\zeta) = j_i(\zeta) G_C^\dagger(\zeta;\mathcal{S}) v_i(\zeta),$$

we have

$$
\begin{aligned}
\kappa_i \quad &= \quad \max_{\zeta \in [0,\delta]} |u_i^*(\zeta)[C^\dagger(\zeta;\mathcal{S})]^* E C^\dagger(\zeta;\mathcal{S}) u_i(\zeta)| \\
&\qquad + \max_{\zeta \in [0,\delta]} |v_i^*(\zeta)[G_C^\dagger(\zeta;\mathcal{S})]^* F G_C^\dagger(\zeta;\mathcal{S}) v_i(\zeta)|.
\end{aligned}
$$

which is the second equality in (3.9). Classical norm inequalities yield the inequality in (3.9). ∎

The following corollary yields a bound for the case of scaled perturbations discussed by Barlow and Demmel [1].

Corollary 3.2. *Assume the hypothesis and terminology of Theorem 3.1. Assume that E and F have the form*

$$E = D_H^* E_S D_H, \quad \|E_S\| \leq 1, \quad F = D_M^* F_S D_M, \quad \|F_S\| \leq 1.$$

Then for each $i \in \mathcal{S}$

$$\begin{aligned}\kappa_i &\leq \max_{\zeta\in[0,\delta]} \|D_H C_G^{\dagger}(\zeta;\mathcal{S})u_i(\zeta)\|^2 + \max_{\zeta\in[0,\delta]} \|D_M G_C^{\dagger}(\zeta;\mathcal{S})v_i(\zeta)\|^2. \\ &\leq \max_{\zeta\in[0,\delta]} \|D_H C_G^{\dagger}(\zeta;\mathcal{S})\|^2 + \max_{\zeta\in[0,\delta]} \|D_M G_C^{\dagger}(\zeta;\mathcal{S})\|^2.\end{aligned}$$

The componentwise version of Theorem 3.1 is obtained similarly as in §3.1 and §3.2.

COROLLARY 3.3. *Assume the hypothesis and terminology of Theorem 3.1. Assume that $|E| \leq |H|$ and $|F| \leq |M|$. Then*

$$\begin{aligned}\kappa_i &\leq \max_{\zeta\in[0,\delta]} |u_i^*(\zeta)[C_G^{\dagger}(\zeta;\mathcal{S})]^*|\,|H|\,|C^{\dagger}(\zeta;\mathcal{S})u_i(\zeta)| \\ &\quad + \max_{\zeta\in[0,\delta]} |v_i^*(\zeta)[G_C^{\dagger}(\zeta;\mathcal{S})]^*|\,|M|\,|G_G^{\dagger}(\zeta;\mathcal{S})v_i(\zeta)| \\ &\leq \max_{\zeta\in[0,\delta]} \|\,|[C_G^{\dagger}(\zeta;\mathcal{S})]^*|\,|H|\,|C_G^{\dagger}(\zeta;\mathcal{S})|\,\| \\ &\quad + \max_{\zeta\in[0,\delta]} \|\,|[G_C^{\dagger}(\zeta;\mathcal{S})]^*|\,|M|\,|G_C^{\dagger}(\zeta;\mathcal{S})|\,\|.\end{aligned}$$

Corollary 3.3 is illustrated by Examples 5.2, 5.3 and 5.4 in §5. It generalizes bounds in Veselić and Slapničar [18] to pencils that are singular and as shown in the examples, often allows us to improve upon them.

4. Error bounds on subspaces

We now consider the effect of structured perturbations on the eigenvectors of H. We confine our attention to the perturbed problem

$$(H + \Delta H)\tilde{x} = \tilde{\lambda}\tilde{x},$$

where $\Delta H = \delta E$. Consider the family of Hermitian eigenproblems

$$H(\zeta)x(\zeta) = \lambda(\zeta)x(\zeta), \quad H(\zeta) = H + \zeta E, \quad \zeta \in [0,\delta]. \tag{4.1}$$

Define the set $\mathcal{S}$ by

$$\mathcal{S} = \{i\colon \lambda_i(\zeta) \neq 0, \zeta \in [0,\delta]\}, \tag{4.2}$$

in which case its set complement is

$$\mathcal{S}^c = \{i\colon \lambda_i(\zeta) = 0, \text{ for some } \zeta \in [0,\delta]\}. \tag{4.3}$$

Suppose that $\mathcal{S}$ has k elements and that $\mathcal{S}^c$ has $n-k$ elements. Let $X_1, \tilde{X}_1 \in \mathbf{C}^{n\times k}$ be the eigenvectors of H and $H+\Delta H$ associated with $\mathcal{S}$ and let $X_2, \tilde{X}_2 \in \mathbf{C}^{n\times(n-k)}$ be the matrix of eigenvectors associated with $\mathcal{S}^c$. Thus we have

$$\begin{aligned} HX_j &= X_j\Lambda_j, \quad j=1,2, \\ (H+\Delta H)\tilde{X}_j &= \tilde{X}_j\tilde{\Lambda}_j, \quad j=1,2. \end{aligned} \tag{4.4}$$

We now define two separate types of *relative gaps*:

$$\begin{aligned} \mathrm{relgap}(\lambda,\mu) &= \frac{|\lambda-\mu|}{|\lambda\mu|^{1/2}}, \quad \mu,\lambda \neq 0, \\ \mathrm{relgap}_0(\lambda,\mu) &= \frac{|\lambda-\mu|}{|\lambda|}. \end{aligned}$$

The first definition is just that from Barlow and Demmel [1], and the second definition allows μ (but not λ) to be zero. This allows to bound the error in the "zero" subspaces, that is, the eigenvectors in $\mathcal{S}^c$. These straightforward extensions to sets of eigenvalues will be used: for $\Lambda = \mathrm{diag}(\lambda_i)$ and $\Gamma = \mathrm{diag}(\gamma_j)$

$$\begin{aligned} \mathrm{relgap}(\Lambda,\Gamma) &= \min_{(i,j)} \mathrm{relgap}(\lambda_\mathrm{i},\gamma_\mathrm{j}), \\ \mathrm{relgap}_0(\Lambda,\Gamma) &= \min_{(i,j)} \mathrm{relgap}_0(\lambda_\mathrm{i},\gamma_\mathrm{j}). \end{aligned}$$

We will bound

$$\|X_1^*\tilde{X}_2\|_F$$

which is "sin Θ" bound for the error in the zero subspace X_2 from Davis and Kahan [4].

We will also partition X_1 (conformally $\tilde{X}_1$) into

$$X_1 = \begin{pmatrix} X_{11} & X_{12} \end{pmatrix},$$

where X_{11} and X_{12} are the eigenvectors associated with the eigenvalue matrices Λ_{11} and Λ_{12} such that

$$\Lambda_1 = \mathrm{diag}(\Lambda_{11}, \Lambda_{12}).$$

We also bound

$$\|X_{11}^*\tilde{X}_{12}\|_F$$

which is "sin Θ" bound for the error in two subspaces within the non-zero subspace X_1. There can be no meaningful bound within the zero subspace X_2.

We can write down the following theorem on the perturbations of these subspaces.

THEOREM 4.1. *Let $H, \Delta H \in \mathbf{C}^{n\times n}$ be Hermitian and let $\mathcal{S}$ and $\mathcal{S}^c$ be defined by (4.2) and (4.3), respectively. Let $X_j, \tilde{X}_j$ and Λ_j, $\tilde{\Lambda}_j$, $j = 1,2$ satisfy (4.4). Furthermore, let Λ_1 be partitioned into $\Lambda_1 = \mathrm{diag}(\Lambda_{11}, \Lambda_{12})$, and let $\tilde{\Lambda}_1$ be partitioned conformally. Define X_{1i}, $i = 1,2$ such that*

$$HX_{1i} = X_{1i}\Lambda_{1i},$$

and define $\tilde{X}_{1i}, i = 1,2$ conformally as well. For $\zeta \in [0,\delta]$ let $X(\zeta)$ be the eigenvector matrix of the $H(\zeta)$ from (4.1), and let $\Lambda(\zeta;\mathcal{S})$ be defined as in Definition 3.3 with $M = I$ and $X^{-1}(\zeta) = X^*(\zeta)$. Let

$$H(\zeta;\mathcal{S}) = X(\zeta)\Lambda(\zeta;\mathcal{S})X^*(\zeta), \quad C(\zeta;\mathcal{S}) = |\Lambda(\zeta;\mathcal{S})|^{1/2}X^*(\zeta).$$

Then

$$\|X_{11}^*\tilde{X}_{12}\|_F \le \frac{\|[C^\dagger(0;\mathcal{S})]^*\Delta H C^\dagger(\delta;\mathcal{S})\|_F}{\mathrm{relgap}(\Lambda_{11},\tilde{\Lambda}_{12})} \tag{4.5}$$

and

$$\|X_1^*\tilde{X}_2\|_F \le \frac{\|H^\dagger(0;\mathcal{S})\Delta H X_2\|_F}{\mathrm{relgap}_0(\Lambda_1,\tilde{\Lambda}_2)}. \tag{4.6}$$

Proof. For any $\lambda_i \ne \tilde{\lambda}_j$ and $i \ne j$, we have

$$x_i^*\tilde{x}_j = \frac{x_i^*\Delta H\tilde{x}_j}{\tilde{\lambda}_j - \lambda_i}. \tag{4.7}$$

If x_i is a column of X_{11} and $\tilde{x}_j$ is a column of X_{12}, then $i,j \in \mathcal{S}$, thus

$$\begin{aligned} |x_i^*\tilde{x}_j| &\le \frac{|x_i^*\Delta H\tilde{x}_j|\,|\lambda_i\tilde{\lambda}_j|^{1/2}}{|\tilde{\lambda}_j-\lambda_i|\,|\lambda_i\tilde{\lambda}_j|^{1/2}} = \frac{|x_i^*\Delta H\tilde{x}_j|\,|\lambda_i\tilde{\lambda}_j|^{1/2}}{|\tilde{\lambda}_j-\lambda_i|\,\|C(0;\mathcal{S})x_i\|\,\|C(\delta;\mathcal{S})\tilde{x}_j\|} \\ &= \frac{|u_i^*[C^\dagger(0;\mathcal{S})]^*\Delta H C^\dagger(\delta;\mathcal{S})\tilde{u}_j|}{\mathrm{relgap}(\lambda_i,\tilde{\lambda}_j)}. \end{aligned}$$

Summing these up and using standard norm inequalities yields

$$\|X_{11}^*\tilde{X}_{12}\|_F \le \frac{\|U_{11}^*[C^\dagger(0;\mathcal{S})]^*\Delta H C^\dagger(\delta;\mathcal{S})\tilde{U}_{12}\|_F}{\mathrm{relgap}(\Lambda_{11},\tilde{\Lambda}_{12})},$$

where $U_{11} = C(0;\mathcal{S})X_{11}|\Lambda_{11}|^{1/2}$ and $\tilde{U}_{12} = C(\delta;\mathcal{S})X_{12}|\Lambda_{12}|^{1/2}$. Since U_{11} and $\tilde{U}_{12}$ have orthonormal columns, the bound (4.5) follows.

To obtain (4.6), simply assume that $i \in \mathcal{S}$ and $j \in \mathcal{S}^c$ in (4.7). Then

$$\begin{aligned} |x_i^*\tilde{x}_j| &\le \frac{|x_i^*\Delta H\tilde{x}_j|\,|\lambda_i|}{|\tilde{\lambda}_j-\lambda_i|\,|\lambda_i|} = \frac{|x_i^*\Delta H\tilde{x}_j|}{\mathrm{relgap}_0(\lambda_i,\tilde{\lambda}_j)\,\|H(0;\mathcal{S})x_i\|} \\ &= \frac{|x_i^*H^\dagger(0;\mathcal{S})\Delta H\tilde{x}_j|}{\mathrm{relgap}_0(\lambda_i,\tilde{\lambda}_j)}. \end{aligned}$$

A similar argument to that above produces (4.6). ■

This theorem can be generalized to (1.3) with M positive definite if we substitute an M-weighted norm for the Euclidean and Frobenius norms.

Bounds on the perturbed subspaces for structured perturbations are easy to derive from Theorem 4.1. For instance, let

$$\Delta H = D^* \Delta A D, \quad \|\Delta A\|_F = \delta_F.$$

Then short arguments from (4.5) and (4.6) lead to

$$\begin{aligned} \|X_{11}^* \tilde{X}_{12}\|_F &\leq \delta_F \max_{0 \leq \zeta \leq \delta} \frac{\|DC^\dagger(\zeta;\mathcal{S})\|^2}{\operatorname{relgap}(\Lambda_{11}, \tilde{\Lambda}_{12})}, \\ \|X_1^* \tilde{X}_2\|_F &\leq \delta_F \frac{\|DH^\dagger(0;\mathcal{S})\| \, \|D\tilde{X}_2\|}{\operatorname{relgap}_0(\Lambda_1, \tilde{\Lambda}_2)}. \end{aligned}$$

If we instead assume that

$$|\Delta H| \leq \delta |H|,$$

then

$$\begin{aligned} \|X_{11}^* \tilde{X}_{12}\|_F &\leq \delta \frac{\| \, |C^\dagger(0;\mathcal{S})|^T |H| \, |C^\dagger(\delta;\mathcal{S})| \, \|_F}{\operatorname{relgap}(\Lambda_{11}, \tilde{\Lambda}_{12})}, \\ \|X_1^* \tilde{X}_2\|_F &\leq \delta \frac{\| \, |H^\dagger(0;\mathcal{S})| \, |H| \, \|_F}{\operatorname{relgap}_0(\Lambda_1, \tilde{\Lambda}_2)}. \end{aligned}$$

We note that the error in the zero subspace X_2, given by $\|X_1^* \tilde{X}_2\|_F$ is modest if $\| \, |H^\dagger(0;\mathcal{S})| \, |H| \, \|_F$ or $\|DH^\dagger(0;\mathcal{S})\| \, \|D\tilde{X}_2\|$ is modest and Λ_1 has a good relative separation from the near zero eigenvalues.

5. Examples

In this section we illustrate our results on several examples. We give examples for structured perturbations of §3, in particular for the relative componentwise perturbations of the type

$$\Delta H = \delta E, \quad |E| \leq |H|.$$

Such perturbations are highly interesting since they appear during various numerical algorithms for eigenvalue and singular value problems [1, 5, 9, 17, 18, 19]. Such perturbations are sometimes called *floating-point* perturbations [18]. In all examples we compute the first order approximations of our bounds, thus we cannot expect optimality in all cases.

The first example deals with the singular value decomposition and illustrates Corollary 3.1.

EXAMPLE 5.1. Let

$$A = \begin{pmatrix} -2 \cdot 10^{40} & 7 \cdot 10^{20} & 7 \\ -8 \cdot 10^{40} & -6.0001 \cdot 10^{20} & -6 \\ -7 \cdot 10^{40} & 2 \cdot 10^{20} & 2 \end{pmatrix}.$$

Note that the last two column vectors of A are nearly parallel. Let $\Delta A = \delta E$ where $\delta = 10^{-6}$ and

$$E = \begin{pmatrix} 7 \cdot 10^{39} & -1 \cdot 10^{20} & 3 \\ -3 \cdot 10^{40} & 1 \cdot 10^{20} & -1 \\ -9 \cdot 10^{39} & 3 \cdot 10^{19} & 0.4 \end{pmatrix}.$$

Also, both A and D are strongly scaled from the right. Let $\delta\sigma_i = \sigma_i(A + \Delta A) - \sigma_i(A)$. The singular values of A are (properly rounded)

$$(\sigma_1, \sigma_2, \sigma_3) = (1.08 \cdot 10^{41}, 9.25 \cdot 10^{20}, 0.45),$$

and the relative changes in the singular values are

$$\left(\frac{|\delta\sigma_1|}{\sigma_1}, \frac{|\delta\sigma_2|}{\sigma_2}, \frac{|\delta\sigma_3|}{\sigma_3} \right) = (2.5 \cdot 10^{-7}, 1.6 \cdot 10^{-7}, 3.0 \cdot 10^{-2}).$$

Both singular value decompositions, of A and $A + \Delta A$, are computed by the one-sided Jacobi method whose sufficiently high accuracy is guaranteed by the analysis of Demmel and Veselić [5].

Since $|E| \leq 0.42857|A|$, we can apply Corollary 3.1. We compute the first order approximations of the corresponding bounds, that is,

$$\frac{|\delta\sigma_i|}{\sigma_i} \leq \delta_A \| \, |A| \, |A^\dagger u_i| \, \| \leq \delta_A \| \, |A| \, |A^\dagger| \, \|, \quad \delta_A = 0.42857. \tag{5.1}$$

Note that we can use the fact that A is strongly scaled from the right to compute inverse much more accurately. The bounds obtained by the first inequality in (5.1) are

$$\left(\frac{|\delta\sigma_1|}{\sigma_1}, \frac{|\delta\sigma_2|}{\sigma_2}, \frac{|\delta\sigma_3|}{\sigma_3} \right) \leq (4.3 \cdot 10^{-7}, 5 \cdot 10^{-7}, 1.8 \cdot 10^{-1}).$$

This shows that our bounds are local and even the first order approximations can be nearly optimal. Note that our relative bound for σ_1 is slightly worse than the the bound $\|\delta E\|/\sigma_1 = 2.97 \cdot 10^{-7}$ which is derived from the classical normwise perturbation theory. This is to be expected for the largest singular value since it is always perfectly conditioned in the relative sense (unless it is 0) and our bounds have an extra condition number. However, the classical bound is meaningless for other singular values.

The simplified bounds, that is, the second inequality in (5.1)

$$\max_{i=1,2,3} \frac{|\delta\sigma_i|}{\sigma_i} \leq 1.8 \cdot 10^{-1}$$

and the Demmel-Veselić bound,

$$\max_{i=1,2,3} \frac{|\delta\sigma_i|}{\sigma_i} \leq n\delta_A \| \, [A \overset{-1}{\operatorname{diag}}(\|A_{:i}\|)]^{-1} \| \leq 3.8 \cdot 10^{-1},$$

respectively, both cover only the worst case.

The following two examples illustrate Corollary 3.3. Both examples were also analyzed in [18].

Example 5.2. Let

$$H = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 10^{-8} \end{pmatrix}.$$

Let $\Delta H = \delta E$, where $\delta = 0.5 \cdot 10^{-5}$ and

$$E = \begin{pmatrix} 0.6 & -1 & 0.8 \\ -1 & 0 & 0 \\ 0.8 & 0 & -1.2 \cdot 10^{-11} \end{pmatrix}.$$

Thus, $|E| \leq |H|$. Let $\delta\lambda_i = \lambda_i(H + \Delta H) - \lambda_i(H)$. The eigenvalues of H are (properly rounded)

$$(\lambda_1, \lambda_2, \lambda_3) = (2, -1, 5 \cdot 10^{-9}),$$

and the relative changes in the eigenvalues are

$$\left(\frac{|\delta\lambda_1|}{|\lambda_1|}, \frac{|\delta\lambda_2|}{|\lambda_2|}, \frac{|\delta\lambda_3|}{|\lambda_3|} \right) = (6.7 \cdot 10^{-7}, 1.7 \cdot 10^{-6}, 9.0 \cdot 10^{-6}).$$

We want to apply Corollary 3.3 with $M = I$. Since the eigenvector matrix $X(\zeta)$ is itself unitary, we can take $U(\zeta) = V(\zeta) = X^{-*}(\zeta) = X(\zeta)$ in (3.7), which implies $C(\zeta; \mathcal{S}) = \mathbf{|} H(\zeta; \mathcal{S}) \mathbf{|}^{1/2}, G(\zeta) = I$. The first order approximations of the bounds from Corollary 3.3 are

$$\begin{aligned} \frac{|\delta\lambda_i|}{|\lambda_i|} &\leq \delta |u_i^* \mathbf{|} H \mathbf{|}^{-1/2}| \, |H| \, |\mathbf{|} H \mathbf{|}^{-1/2} u_i| \\ &\leq \delta \| \, |\mathbf{|} H \mathbf{|}^{-1/2}| \, |H| \, |\mathbf{|} H \mathbf{|}^{-1/2}| \, \|. \end{aligned} \tag{5.2}$$

The bounds obtained by the first inequality in (5.2) are

$$\left(\frac{|\delta\lambda_1|}{|\lambda_1|}, \frac{|\delta\lambda_2|}{|\lambda_2|}, \frac{|\delta\lambda_3|}{|\lambda_3|}\right) \leq (5 \cdot 10^{-6}, 8.3 \cdot 10^{-6}, 1.5 \cdot 10^{-5}),$$

which is again nearly optimal. The heuristic (2.9) implies that even the tiniest eigenvalue λ_3 does not cross zero for any $\zeta \in [0, \delta]$, even though λ_3 is in magnitude much less that δ. On the other hand, the simplified bound, that is, the bound obtained from the second inequality in (5.2), is

$$\max_{i=1,2,3} \frac{|\delta\lambda_i|}{|\lambda_i|} \leq 0.095.$$

The bound from Veselić and Slapničar [18],

$$\max_{i=1,2,3} \frac{|\delta\lambda_i|}{|\lambda_i|} \leq n\delta \|(D^{-1} | H | D^{-1})^{-1}\| \leq 2 \cdot 10^3, \tag{5.3}$$

where $D = \operatorname{diag}(\sqrt{|H|_{ii}})$, is useless.

Example 5.3. Another interesting example is the following: let $H = DAD$, where

$$A = \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 10^8 & & & \\ & 1 & & \\ & & 1 & \\ & & & 10^8 \end{pmatrix}.$$

The eigenvector matrix of H is

$$X = \begin{pmatrix} 1/\sqrt{2} & 1/2 & 1/2 & 0 \\ 0 & -1/2 & 1/2 & 1/\sqrt{2} \\ 0 & -1/2 & 1/2 & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/2 & 1/2 & 0 \end{pmatrix}.$$

Let $\Delta H = \delta E$, where $\delta = 0.5 \cdot 10^{-6}$, $E = DE_SD$, $E_S = ww^T$, $w = (1 \quad 1 \quad -1 \quad 1)^T$. The eigenvalues of H are

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (2 \cdot 10^{16}, 2 \cdot 10^8, -2 \cdot 10^8, 2),$$

and the relative changes in the eigenvalues are

$$\left(\frac{|\delta\lambda_1|}{|\lambda_1|}, \frac{|\delta\lambda_2|}{|\lambda_2|}, \frac{|\delta\lambda_3|}{|\lambda_3|}, \frac{|\delta\lambda_4|}{|\lambda_4|}\right) = (0, 49, 0.98, 5 \cdot 10^{-7}).$$

There is no change in λ_1 because its eigenvector satisfies $Ex_1 = 0$. Unless we use the exact perturbation E with Theorem 2.1, or incorporate this into the structure of our bound, we will not detect this. We see that λ_2 and λ_3 are very sensitive. The bounds obtained by the first inequality in (5.2) are

$$\left(\frac{|\delta\lambda_1|}{|\lambda_1|}, \frac{|\delta\lambda_2|}{|\lambda_2|}, \frac{|\delta\lambda_3|}{|\lambda_3|}, \frac{|\delta\lambda_4|}{|\lambda_4|}\right) \leq (5\cdot 10^{-7}, 25, 25, 5\cdot 10^{-7}),$$

and clearly show the sensitivity of λ_2 and λ_3. The bound on λ_2 is too optimistic because this is a first order theory and value of $\hat{\kappa}_i$ is now too large (> 1) for it to be relevant. All eigenvalues are in the set $\mathcal{S}$ since all of the eigenvectors retain their sign pattern under the perturbation, but two of them are not well-behaved and can only be meaningfully bounded by exactly computing the integral in Theorem 2.1 or in absolute error terms. For λ_2 the absolute bound

$$|\lambda_2 - \tilde{\lambda}_2| \leq \delta \|E\| = 10^{10}$$

is a good estimate, but does not tell us whether the eigenvalue crosses zero or not. Theorem 2.1 would tell us that, but at great expense.

The bounds (the first order approximations) for λ_1 and λ_4 are good in the sense that they show that these eigenvalues are well-behaved. The bound for λ_1 is not optimal since it only uses the information that $|E| \leq |H|$.

The bounds obtained from the second inequality in (5.2) and (5.3),

$$\max_{i=1,2,3,4} \frac{|\delta\lambda_i|}{|\lambda_i|} \leq 50, \quad \text{and} \quad \max_{i=1,2,3,4} \frac{|\delta\lambda_i|}{|\lambda_i|} \leq 100,$$

respectively, as well as the bound for λ_4 obtained by the classical normwise perturbation theory, are useless.

The next example illustrates Corollary 3.3 on a matrix pair (H, M).

Example 5.4. Let $H = D_H A^T \Sigma A D_H$ and $M = D_M B^T B D_M$, where

$$\begin{aligned} D_H &= \operatorname{diag}(10^8, 10^4, 10, 10, 1), \quad \Sigma = \operatorname{diag}(-1, -1, 1, 1), \\ D_M &= \operatorname{diag}(10^{-4}, 10^{-2}, 10^{-2}, 10^{-1}, 1), \end{aligned}$$

and

$$A = \begin{pmatrix} -3 & -5 & -5 & 0 & 2 \\ 4 & 2 & -2 & -4 & -5 \\ -1 & -1 & 1 & 1 & 1 \\ 1 & 5 & 3 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -2 & -4 & -4 & -2 \\ -3 & 4 & 4 & 4 & 0 \\ -1 & 1 & 2 & 0 & -2 \end{pmatrix}.$$

Thus, H is indefinite singular of rank four, M is semi-definite of rank three, and H and M are scaled in opposite directions. Altogether,

$$H = \begin{pmatrix} -2.3\cdot 10^{17} & -1.7\cdot 10^{13} & -5.0\cdot 10^{9} & 1.6\cdot 10^{10} & 2.8\cdot 10^{9} \\ -1.7\cdot 10^{13} & -3.0\cdot 10^{8} & -7.0\cdot 10^{5} & 1.2\cdot 10^{6} & 3.4\cdot 10^{5} \\ -5.0\cdot 10^{9} & -7.0\cdot 10^{5} & -1.9\cdot 10^{3} & -4.0\cdot 10^{2} & 1.0\cdot 10^{2} \\ 1.6\cdot 10^{10} & 1.2\cdot 10^{6} & -4.0\cdot 10^{2} & -1.4\cdot 10^{3} & -1.6\cdot 10^{2} \\ 2.8\cdot 10^{9} & 3.4\cdot 10^{5} & 1.0\cdot 10^{2} & -1.6\cdot 10^{2} & -1.9\cdot 10^{1} \end{pmatrix},$$

$$M = \begin{pmatrix} 1.9\cdot 10^{-7} & -1.9\cdot 10^{-5} & -2.6\cdot 10^{-5} & -2.4\cdot 10^{-4} & -4.0\cdot 10^{-4} \\ -1.9\cdot 10^{-5} & 2.1\cdot 10^{-3} & 2.6\cdot 10^{-3} & 2.4\cdot 10^{-2} & 2.0\cdot 10^{-2} \\ -2.6\cdot 10^{-5} & 2.6\cdot 10^{-3} & 3.6\cdot 10^{-3} & 3.2\cdot 10^{-2} & 4.0\cdot 10^{-2} \\ -2.4\cdot 10^{-4} & 2.4\cdot 10^{-2} & 3.2\cdot 10^{-2} & 3.2\cdot 10^{-1} & 8.0\cdot 10^{-1} \\ -4.0\cdot 10^{-4} & 2.0\cdot 10^{-2} & 4.0\cdot 10^{-2} & 8.0\cdot 10^{-1} & 8.0\cdot 10^{0} \end{pmatrix}.$$

The eigenvector matrix of the pair (H, M) is (properly rounded)

$$X = \begin{pmatrix} 1.00 & -7.40\cdot 10^{-5} & -4.80\cdot 10^{-6} & 1.86\cdot 10^{-7} & 2.18\cdot 10^{-8} \\ -2.52\cdot 10^{-3} & 1.00 & 3.01\cdot 10^{-2} & -2.96\cdot 10^{-4} & -6.72\cdot 10^{-5} \\ 3.74\cdot 10^{-3} & 5.00\cdot 10^{-1} & 8.66\cdot 10^{1} & -8.41\cdot 10^{-1} & -5.04\cdot 10^{-2} \\ 6.29\cdot 10^{-4} & -1.50\cdot 10^{-1} & -1.01\cdot 10^{1} & 2.13 & 1.68\cdot 10^{-1} \\ -2.53\cdot 10^{-5} & 1.00\cdot 10^{-2} & 5.76\cdot 10^{-1} & -2.37\cdot 10^{-1} & 3.36\cdot 10^{-1} \end{pmatrix}.$$

We have

$$\begin{aligned} X^*HX &= \mathrm{diag}(-2.3\cdot 10^{17}, 9.57\cdot 10^{8}, -1.3019\cdot 10^{7}, 7.7388, 2.6\cdot 10^{-15}), \\ X^*MX &= \mathrm{diag}(-3.6\cdot 10^{-23}, 1.1\cdot 10^{-18}, 1, 1, 1). \end{aligned}$$

We conclude that $\mathcal{S} = \{3, 4\}$,

$$\Lambda(0;\mathcal{S}) = \mathrm{diag}(0, 0, -1.3019\cdot 10^{7}, 7.7388, 0), \quad J(\zeta) = \mathrm{diag}(0, 0, 1, 1, 1),$$

where $\Lambda(0;\mathcal{S})$ and $J(\zeta)$ are defined by Definition 3.3 and (3.5), respectively. We can arrive to this conclusion in two ways, by using heuristic (2.9), or by observing that the null subspaces of of H and M have only the trivial intersection.

Let us perturb H to $H + \Delta H$ with $\Delta H = \delta E$, where $\delta = 10^{-6}$ and

$$E = \begin{pmatrix} -1\cdot 10^{17} & 3\cdot 10^{12} & 9\cdot 10^{8} & 7\cdot 10^{9} & -3\cdot 10^{8} \\ 3\cdot 10^{12} & -4\cdot 10^{6} & 2\cdot 10^{5} & -3\cdot 10^{5} & 1\cdot 10^{5} \\ 9\cdot 10^{8} & 2\cdot 10^{5} & 9\cdot 10^{2} & 8\cdot 10^{1} & -4\cdot 10^{1} \\ 7\cdot 10^{9} & -3\cdot 10^{5} & 8\cdot 10^{1} & 4\cdot 10^{2} & 2\cdot 10^{1} \\ -3\cdot 10^{8} & 1\cdot 10^{5} & -5\cdot 10^{1} & 2\cdot 10^{1} & -6 \end{pmatrix}.$$

Thus, $|E_H| \leq |H|$. The relative changes in the eigenvalues λ_i, $i \in \mathcal{S}$, are

$$\left(\frac{|\delta\lambda_3|}{|\lambda_3|}, \frac{|\delta\lambda_4|}{|\lambda_4|}\right) = (3.5\cdot 10^{-7}, 5.3\cdot 10^{-4}).$$

The first order approximations of the bounds from Corollary 3.3 are

$$\begin{aligned} \frac{|\delta\lambda_i|}{|\lambda_i|} &\leq \delta |u_i^*(0)[C^\dagger(0;\mathcal{S})]^*|\,|H|\,|C^\dagger(0;\mathcal{S})u_i(0)| \qquad (5.4) \\ &\leq \delta \|\,|[C^\dagger(0;\mathcal{S})]^*|\,|H|\,|C^\dagger(0;\mathcal{S})|\,\|, \end{aligned}$$

where $C^\dagger(0;\mathcal{S})$ and $U(0)$ are defined by (3.6) and (3.7), respectively. We can take

$$U = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

in which case

$$C^\dagger(0;\mathcal{S}) = \begin{pmatrix} -1.331 \cdot 10^{-9} & 6.675 \cdot 10^{-8} \\ 8.341 \cdot 10^{-6} & -1.065 \cdot 10^{-4} \\ 2.400 \cdot 10^{-2} & -3.024 \cdot 10^{-1} \\ -2.796 \cdot 10^{-3} & 7.667 \cdot 10^{-1} \\ 1.595 \cdot 10^{-4} & -8.525 \cdot 10^{-2} \end{pmatrix}.$$

The bounds obtained from the first inequality in (5.4) are

$$\left(\frac{|\delta\lambda_3|}{|\lambda_3|}, \frac{|\delta\lambda_4|}{|\lambda_4|} \right) \leq (2.7 \cdot 10^{-6}, 4.6 \cdot 10^{-3}),$$

and the bound obtained from the second inequality in (5.4) is

$$\max_{i=3,4} \frac{|\delta\lambda_i|}{|\lambda_i|} \leq 4.6 \cdot 10^{-3}.$$

Note that choosing $C^\dagger(0;\mathcal{S})$ with another $U(0)$ in (3.8) would yield the same bounds.

Our last example deals with subspace bounds of §5.

EXAMPLE 5.5. Let us reconsider the matrix in Example 2.1. If we choose $\delta = 2.2204 \cdot 10^{-16}$ then the bound in the perturbation of its zero subspace is

$$\|X_1^* \tilde{X}_2\|_F \leq \delta \frac{\|\,|H^\dagger(0;\mathcal{S})|\,|H|\,\|_F}{\mathrm{relgap}_0(\Lambda_1, \tilde{\Lambda}_2)}.$$

The truncated psuedoinverse of H is

$$H^\dagger(0,\mathcal{S}) = \begin{pmatrix} 2.0004 \cdot 10^{-16} & -1.0001 \cdot 10^{-16} & -2.0002 \cdot 10^{-12} & -2.0003 \cdot 10^{-14} \\ -1.0001 \cdot 10^{-16} & 1 \cdot 10^{-16} & 1 \cdot 10^{-12} & 1 \cdot 10^{-14} \\ -2.0002 \cdot 10^{-12} & 1 \cdot 10^{-12} & 2 \cdot 10^{-8} & 2 \cdot 10^{-10} \\ -2.0003 \cdot 10^{-14} & 1 \cdot 10^{-14} & 2 \cdot 10^{-10} & 3 \cdot 10^{-12} \end{pmatrix},$$

the associated condition number is

$$\| \, |H^{\dagger}(0, \mathcal{S})| \, |H| \, \|_F = 4.0004 \cdot 10^4,$$

and we have that $\mathrm{relgap}_0(\Lambda_1, \tilde{\Lambda}_2) \approx 1$. Thus

$$\|X_1^* \tilde{X}_2\|_F \leq \delta \frac{\| \, |H^{\dagger}(0; \mathcal{S})| \, |H| \, \|_F}{\mathrm{relgap}_0(\Lambda_1, \tilde{\Lambda}_2)} = 8.8827 \cdot 10^{-12}$$

Actually, this bound is very pessimistic. The scaled bound is much better. If we have

$$|E| \leq |H|,$$

then

$$\delta_F \leq \delta \|A\|_F = 8.8818 \cdot 10^{-16}.$$

We have that

$$\|DH^{\dagger}(0; \mathcal{S})\| = 3.0002 \cdot 10^{-4}, \quad \|D\tilde{X}_2\| = 2,$$

thus

$$\|X_1^* \tilde{X}_2\|_F \leq \delta_F \frac{\|DH^{\dagger}(0; \mathcal{S})\| \, \|D\tilde{X}_2\|}{\mathrm{relgap}_0(\Lambda_1, \tilde{\Lambda}_2)} = 5.3295 \cdot 10^{-19}.$$

Thus it is reasonable to expect the zero subspace of this matrix to be computed accurately.

The standard absolute gap bound is

$$\|X_1^* X_2\|_F \leq \delta \frac{\|E\|_F}{\mathrm{gap}(\Lambda_1, \tilde{\Lambda}_2)} = 8.8827 \cdot 10^{-8}.$$

This is far too pessimistic.

6. Conclusion

As a general conclusion based on our bounds and the above examples we note that the error bounds on individual eigenvalues and vectors tend to be tighter, sometimes much tighter, than the global error bounds for all of the eigenvalues of the matrix given in [1] or [18]. Moreover, they are easier to generalize to large classes of eigenvalue problems. We also note that we obtain structured perturbation results on Hermitian pencils when one or both of the matrices are singular (see Proposition 3.1, Corollaries 3.1 and 3.3).

The above examples also lead us to observation that we can often obtain meaningful relative error bounds on eigenvalues of numerically singular matrices as long as those eigenvalues have good local condition numbers. If these "non-zero" eigenvalues are well-behaved, it is possible that the subspace associated with the "zero" eigenvalues is also well-behaved (see comments after Theorem 4.1 and Example 5.5).

Thus, we expand the definition of well-behaved matrices to include matrices whose non-zero eigenvalues have modest local condition numbers and whose zero subspace is well-behaved. This definition includes the matrices in Examples 5.2 and 5.5, but does not include Example 5.3 because of its two badly behaved eigenvalues. Examples 5.1 and 5.2 have eigenvalues (and singular values) that are much better behaved than the normwise theory would tell us, but we would not expect any numerical method to compute all of the digits of the smallest eigenvalues (or singular values) correctly.

We would like to thank the referee for very helpful comments.

REFERENCES

1 J.L. Barlow and J.W. Demmel, Computing accurate eigensystems of scaled diagonally dominant matrices, *SIAM J. Numer. Anal.*, 27:762–791, 1990.

2 J.L. Barlow and I. Slapničar, Optimal perturbation bounds for the Hermitian eigenvalue problem, Technical report, The Pennsylvania State University, 1999.

3 S.L. Campbell and Jr. C.D. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, London, 1979. Reprint by Dover, New York, 1991.

4 C. Davis and W.M. Kahan, The rotation of eigenvectors by a perturbation III, *SIAM J. Num. Anal.*, 7:1–46, 1970.

5 J.W. Demmel and K. Veselić, Jacobi's method is more accurate than QR, *SIAM J. Matrix Anal. Appl.*, 13:1204–1243, 1992.

6 F.M. Dopico, J. Moro, and J.M. Molera, Weyl-type perturbation bounds for eigenvalues of Hermitian matrices, preprint, 1998.

7 S.C. Eisenstat and I.C.F. Ipsen, Relative perturbation techniques for singular value problems, *SIAM J. Numer. Anal.*, 32:1972–1988, 1995.

8 L. Eldén, Error analysis of direct method of matrix inversion, *J. Assoc. Comput. Mach.*, 8:281–330, 1961.

9 G.H. Golub and C.F. Van Loan, *Matrix Computations, Second Edition*, The Johns Hopkins Press, Baltimore, 1989.

10 M. Gu, *Studies in Numerical Linear Algebra*, PhD thesis, Yale University, New Haven, CT, 1994.

11 M. Gu and S.C. Eisenstat, Relative perturbation theory for eigenproblems, Research Report YALEU/DCS/RR-934, Department of Computer Science, Yale University, New Haven, CT, February 1993.

12 T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1982.

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

14 R.-C. Li, Relative perturbation theory: (ii) eigenspace ans singular subspace variations, *SIAM J. Matrix Anal. Appl.*, 20:471–492, 1998.

15 C.F. Van Loan, Generalizing the singular value decomposition, *SIAM J. Num. Anal.*, 13:76–83, 1976.

16 C.C. Paige and M.A. Saunders, Towards a generalized singular value decomposition, *SIAM J. Num. Anal.*, 18:398–405, 1981.

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

18 K. Veselić and I. Slapničar, Floating-point perturbations of Hermitian matrices, *Linear Algebra and Its Applications*, 195:81–116, 1993.

19 J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, London, 1965.