REFINED PERTURBATION BOUNDS FOR EIGENVALUES OF HERMITIAN AND NON-HERMITIAN MATRICES*

I.C.F. IPSEN† AND B. NADLER‡

Abstract. We present eigenvalue bounds for perturbations of Hermitian matrices, and express the change in eigenvalues in terms of a projection of the perturbation onto a particular eigenspace, rather than in terms of the full perturbation. The perturbations we consider are Hermitian of rank one, and Hermitian or non-Hermitian with norm smaller than the spectral gap of a specific eigenvalue. Applications include principal component analysis under a spiked covariance model, and pseudo arclength continuation methods for the solution of nonlinear systems.

Key words. eigenvalues, Hermitian matrix, eigenvalue gap, perturbation bounds, non-Hermitian perturbations, principal components, numerical continuation.

AMS subject classification. 65F15, 65H10, 65H17, 65H20, 15A18, 15A42

1. Introduction. We present perturbation bounds for eigenvalues of Hermitian matrices that were motivated by two applications: principal component analysis under a spiked covariance model [25], and pseudo arclength continuation methods for the solution of systems of nonlinear equations [7].

Although these applications are very different, they share a common requirement: The change in the eigenvalues of interest should be determined not by the global norm of the full perturbation, which can be quite large, but rather by the norm of a projection of the perturbation on a particular eigenspace. In contrast, most existing eigenvalue bounds are expressed either in terms of the full perturbation or else in terms of a residual, and therefore do not provide sufficient information for our applications.

The paper is organized as follows. We start with the most specific class of perturbations, Hermitian rank-one updates, and then generalize the perturbations first to Hermitian and then to non-Hermitian matrices. In §2 we present bounds for Hermitian rank one updates, and explain why such bounds can be useful in pseudo-arclength continuation methods. In §3 we consider Hermitian perturbations whose norm is smaller than the spectral gap of a specific eigenvalue, and describe their use in principal component analysis. In §4 we extend the bounds to non-Hermitian perturbations.

Notation. The identity matrix of order k is $I_k = \begin{pmatrix} e_1 & \dots & e_k \end{pmatrix}$. The norm $\|\cdot\|$ denotes the two norm. The eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ are numbered so that

$$\lambda_{\min}(A) \equiv \lambda_n(A) \le \ldots \le \lambda_{\max}(A) \equiv \lambda_1(A).$$

The conjugate transpose of a matrix A is denoted by A^*; and an overbar, as in $\overline{A}$, denotes elementwise complex conjugation.

We will use two measures for the separation between adjacent eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n\times n}$: The distance of an eigenvalue $\lambda_i(A)$ to its right neighbour,

$$\mathrm{gap}_i \equiv \lambda_{i-1}(A) - \lambda_i(A), \qquad 2 \le i \le n,$$

*This research was supported in part by NSF grant DMS grant 0112069. BN was supported by a grant from the Lord Sieff of Brimpton memorial fund.

†Department of Mathematics, North Carolina State University, P.O. Box 8205, Raleigh, NC 27695-8205, USA (ipsen@ncsu.edu, http://www4.ncsu.edu/~ipsen/)

‡Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel (boaz.nadler@weizmann.ac.il)

and the minimum of the distance to left and right neighbours,

$$\mathrm{Gap}_i \equiv \min_{j \neq i} |\lambda_i(A) - \lambda_j(A)|.$$

The two measures are related,

$$\mathrm{Gap}_n = \mathrm{gap}_n, \qquad \mathrm{Gap}_1 = \mathrm{gap}_2, \qquad \mathrm{Gap}_i = \min\{\mathrm{gap}_i, \mathrm{gap}_{i+1}\}, \quad 1 < i < n.$$

2. Hermitian Rank One Updates. We present improved perturbation bounds for eigenvalues of Hermitian matrices when the perturbation is Hermitian of rank one.

Before describing an application that requires such bounds, we mention that algorithms for computing eigenvalues and eigenvectors of Hermitian matrices modified by a rank one matrix are well established [3, 10, 14], [11, §8.5.3, §12.5.1]; and that the corresponding inverse eigenvalue problem has also been investigated [21].

2.1. Numerical Continuation. Numerical continuation is the process of solving systems of nonlinear equations $G(u, \lambda) = 0$ for various values of the real parameter λ. Here $G : \mathbb{R}^{N+1} \to \mathbb{R}^N$ is assumed to be sufficiently smooth [12, 20, 22, 29].

Parameter continuation is a method for tracing out a solution path by repeatedly incrementing λ until the desired value of λ is reached. In each iteration, the current solution u serves as an initial iterate for, say, Newton's method to compute a solution for the next value of λ. Although parameter continuation is simple and intuitive, it fails at points (u, λ) where the Jacobian G_u is singular.

One can try to circumvent singularities by reparameterizing the problem and introducing the arclength parameter s. Now both u and λ depend on s, and the original parameter λ is treated as an unknown. The resulting pseudo-arclength continuation method [12, 20, 22, 29] implements parameter continuation on $F(u(s), \lambda(s)) = 0$ with s as the parameter and solves

$$F(x, s) = \begin{pmatrix} G(x) \\ \mathcal{N}(x, s) \end{pmatrix} = 0, \qquad x = \begin{pmatrix} u(s) \\ \lambda(s) \end{pmatrix},$$

where $\mathcal{N}$ represents a normalization equation. Pseudo-arclength continuation requires that the Jacobian

$$F_x = \begin{pmatrix} G_u & G_\lambda \\ \mathcal{N}_u & \mathcal{N}_\lambda \end{pmatrix}$$

be nonsingular. The normalization equation is set up so that at a point where $G(u_0, \lambda_0) = 0$ the row $\begin{pmatrix} \mathcal{N}_u & \mathcal{N}_\lambda \end{pmatrix}$ has unit norm and is almost orthogonal to the rows of $\begin{pmatrix} G_u & G_\lambda \end{pmatrix}$. Hence F_x is nonsingular at (u_0, λ_0) if the rank of $\begin{pmatrix} G_u & G_\lambda \end{pmatrix}$ equals N. In other words, F_x is nonsingular if G_u is nonsingular, or if the nullspace of G_u has dimension 1 and G_λ is not in the range of G_u [20, 29]. The latter singularity is called limit point, fold point, simple fold, or turning point [4, 5, 23, 27, 29].

Denote the partial derivatives at (u_0, λ_0) by $G_u = G_u(u_0, \lambda_0)$ and $y = G_\lambda(u_0, \lambda_0)$. Instead of the singular values of the Jacobian F_x we consider the eigenvalues of

$$F_x F_x^* = \begin{pmatrix} A + yy^* & 0 \\ 0 & 1 \end{pmatrix} + \mathcal{E},$$

where $A = G_u G_u^*$, and we have used the fact that the last row of F_x has unit norm and is almost orthogonal to the others, so that $\|\mathcal{E}\|$ is small. To estimate the condition

number of F_x and the convergence rate of a Newton-GMRES method, it suffices to bound $\|F_x^{-1}\|$ by determining a nontrivial lower bound for the smallest eigenvalue $\lambda_{\min}(A + yy^*)$ [7]. For this positive semi-definite rank one update, Weyl's theorem implies [11, Theorem 8.1.8], [26, Corollary 10.3.1]

$$\lambda_{\min}(A) \leq \lambda_{\min}(A + yy^*).$$

When A is nonsingular, this bound is adequate for our purposes. However it is useless at a fold point, because there A is singular and $0 = \lambda_{\min}(A) = 0 < \lambda_{\min}(A + yy^*)$. We need a lower bound for $\lambda_{\min}(A)$ that takes into account that y is not in the range of A and has a nonzero contribution in the eigenspace of $\lambda_{\min}(A)$.

Our objective is to tighten our previous bound [7, Theorem 3.3] and the bounds in [13]. This is accomplished in Theorem 2.1 below. The results in §2.2 may also be of benefit in the construction of nonsingular bordered matrices.

2.2. Smallest Eigenvalue. For a given Hermitian matrix $A \in \mathbb{C}^{n\times n}$ and a column vector $y \in \mathbb{C}^n$, we improve the inclusion interval from Weyl's theorem for the smallest eigenvalue of Hermitian rank one updates $A \pm yy^*$,

$$\lambda_{\min}(A) \leq \lambda_{\min}(A + yy^*) \leq \lambda_{n-1}(A) \tag{2.1}$$

$$\lambda_{\min}(A) - \|y\|^2 \leq \lambda_{\min}(A - yy^*) \leq \lambda_{\min}(A) \tag{2.2}$$

by taking into account the contribution of y in the eigenspace of $\lambda_{\min}(A)$.

Let $A = V\Lambda V^*$ be an eigenvalue decomposition, where $V = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix}$ is unitary, and

$$\Lambda = \begin{pmatrix} \lambda_1(A) & & \\ & \ddots & \\ & & \lambda_n(A) \end{pmatrix}, \qquad \lambda_{\max}(A) = \lambda_1(A) \geq \ldots \geq \lambda_n(A) = \lambda_{\min}(A).$$

Define the projections of the vector y onto the eigenvectors of A,

$$y_{i:j} \equiv \begin{pmatrix} v_i & \dots & v_j \end{pmatrix}^* y, \qquad 1 \leq i \leq j \leq n.$$

Below we bound the smallest eigenvalues of the rank one updates in terms of rank one updates of projections onto two-dimensional subspaces.

THEOREM 2.1 (Smallest Eigenvalue). *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian, $y \in \mathbb{C}^n$, and*

$$L_\pm \equiv \begin{pmatrix} \lambda_{n-1}(A) & 0 \\ 0 & \lambda_n(A) \end{pmatrix} \pm \begin{pmatrix} \|y_{1:n-1}\| \\ y_n \end{pmatrix} \begin{pmatrix} \|y_{1:n-1}\| & \overline{y}_n \end{pmatrix}$$

$$U_\pm \equiv \begin{pmatrix} \lambda_{n-1}(A) & 0 \\ 0 & \lambda_n(A) \end{pmatrix} \pm \begin{pmatrix} y_{n-1} \\ y_n \end{pmatrix} \begin{pmatrix} \overline{y}_{n-1} & \overline{y}_n \end{pmatrix}.$$

Then $\lambda_{\min}(L_\pm) \leq \lambda_{\min}(A \pm yy^*) \leq \lambda_{\min}(U_\pm)$, where

$$\lambda_{\min}(A) \leq \lambda_{\min}(L_+) \leq \lambda_{\min}(U_+) \leq \lambda_{n-1}(A)$$

$$\lambda_{\min}(A) - \|y\|^2 \leq \lambda_{\min}(L_-) \leq \lambda_{\min}(U_-) \leq \lambda_{\min}(A).$$

Proof. Abbreviate $\alpha_j \equiv \lambda_j(A)$, $1 \leq j \leq n$, and partition the eigenvalue decomposition of A so as to distinguish the smallest eigenvalue $\alpha_n = \lambda_{\min}(A)$.

$$\Lambda = \begin{pmatrix} \Lambda_1 & \\ & \alpha_n \end{pmatrix}, \qquad \Lambda_1 \equiv \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_{n-1} \end{pmatrix},$$

and $V = \begin{pmatrix} V_1 & v_n \end{pmatrix}$ with $V_1 \equiv \begin{pmatrix} v_1 & \dots & v_{n-1} \end{pmatrix}$. We derive the bounds by 'projecting' A onto a 2×2 matrix with eigenvalues α_n and α_{n-1}.

Lower Bounds. We start with the positive semi-definite update. Let z be a unit-norm eigenvector associated with $\lambda_{\min}(A + yy^*)$, i.e. $(A + yy^*)z = \lambda_{\min}(A + yy^*)z$, $\|z\| = 1$. Express z in the V-basis,

$$\begin{pmatrix} z_{1:n-1} \\ z_n \end{pmatrix} = \begin{pmatrix} V_1^* z \\ v_n^* z \end{pmatrix} = V^* z.$$

Then

$$\begin{aligned} \lambda_{\min}(A + yy^*) &= z^*(A + yy^*)z = z_{1:n-1}^* \Lambda_1 z_{1:n-1} + \alpha_n |z_n|^2 + |y^* z|^2 \\ &\geq \alpha_{n-1} \|z_{1:n-1}\|^2 + \alpha_n |z_n|^2 + |z_{1:n-1}^* y_{1:n-1} + \overline{z}_n y_n|^2 \\ &= \begin{pmatrix} z_{1:n-1}^* & \overline{z}_n \end{pmatrix} \left[\begin{pmatrix} \alpha_{n-1} I_{n-1} & 0 \\ 0 & \alpha_n \end{pmatrix} + \begin{pmatrix} y_{1:n-1} \\ y_n \end{pmatrix} \begin{pmatrix} y_{1:n-1}^* & \overline{y}_n \end{pmatrix} \right] \begin{pmatrix} z_{1:n-1} \\ z_n \end{pmatrix}. \end{aligned}$$

Let Q be a unitary matrix of order $n-1$ so that $Qy_{1:n-1} = \|y_{1:n-1}\| e_{n-1}$ and set $w \equiv \begin{pmatrix} Q z_{1:n-1} \\ z_n \end{pmatrix}$, where $\|w\| = 1$. Then

$$\begin{aligned} \lambda_{\min}(A + yy^*) &\geq w^* \left(\begin{pmatrix} \alpha_{n-1} I_{n-1} & 0 \\ 0 & \alpha_n \end{pmatrix} + \begin{pmatrix} \|y_{1:n-1}\| e_{n-1} \\ y_n \end{pmatrix} \begin{pmatrix} \|y_{1:n-1}\| e_{n-1}^* & \overline{y}_n \end{pmatrix} \right) w \\ &\geq \lambda_{\min} \begin{pmatrix} \alpha_{n-1} I_{n-2} & 0 \\ 0 & L_+ \end{pmatrix} = \min\{\alpha_{n-1}, \lambda_{\min}(L_+)\}. \end{aligned}$$

Applying (2.1) to L_+ gives $\alpha_n \leq \lambda_{\min}(L_+) \leq \alpha_{n-1}$, and

$$\lambda_{\min}(A + yy^*) \geq \min\{\alpha_{n-1}, \lambda_{\min}(L_+)\} = \lambda_{\min}(L_+).$$

Now consider the negative semi-definite update, and let z be a unit-norm eigenvector associated with $\lambda_{\min}(A - yy^*)$, i.e. $(A - yy^*)z = \lambda_{\min}(A - yy^*)z$, $\|z\| = 1$. As above one shows $\lambda_{\min}(A - yy^*) \geq \min\{\alpha_{n-1}, \lambda_{\min}(L_-)\}$. Applying (2.2) to L_- gives $\alpha_n - \|y\|^2 \leq \lambda_{\min}(L_-) \leq \alpha_n$, and

$$\lambda_{\min}(A - yy^*) \geq \min\{\alpha_{n-1}, \lambda_{\min}(L_-)\} = \lambda_{\min}(L_-).$$

Upper Bounds. Since $U_\pm$ are the respective trailing 2×2 principal submatrices of $V^*(A \pm yy^*)V$, Cauchy's interlace theorem [26, §10.1] implies $\lambda_{\min}(A \pm yy^*) \leq \lambda_{\min}(U_\pm)$. Applying (2.1) to U_+ and (2.2) to U_- gives $\lambda_{\min}(U_+) \leq \alpha_{n-1}$ and $\lambda_{\min}(U_-) \leq \alpha_n$. □

Below we give explicit expressions for the bounds in Theorem 2.1 in terms of the absolute gap between the two smallest eigenvalues,

$$\text{gap}_n \equiv \lambda_{n-1}(A) - \lambda_n(A) \geq 0.$$

Corollary 2.2 (Smallest Eigenvalue). *In Theorem 2.1*

$$\lambda_{\min}(L_\pm) = \lambda_{\min}(A) + \frac{1}{2} \left(\text{gap}_n \pm \|y\|^2 - \sqrt{(\text{gap}_n \pm \|y\|^2)^2 \mp 4\text{gap}_n |y_n|^2} \right),$$

and

$$\lambda_{\min}(U_\pm) = \lambda_{\min}(A) + \frac{1}{2} \left(\text{gap}_n \pm \|y_{n-1:n}\|^2 - \sqrt{(\text{gap}_n \pm \|y_{n-1:n}\|^2)^2 \mp 4\text{gap}_n |y_n|^2} \right).$$

Implications for Numerical Continuation. For the application to pseudo arclength continuation in §2.1, it is important to know how $|y_n|$ and gap_n influence $\lambda_{\min}(A + yy^*)$, provided $\lambda_{\min}(A) < \lambda_{n-1}(A)$, $y_n \neq 0$ and $y_{n-1} \neq 0$. This influence becomes very clear in the next bound, which illustrates how much of the increase in the smallest eigenvalue can be due to the contribution of y in the eigenspace of $\lambda_{\min}(A)$.

COROLLARY 2.3. *Under the conditions of Theorem 2.1,*

$$\lambda_{\min}(A + yy^*) \leq \lambda_{\min}(A) + |y_n|\sqrt{\text{gap}_n}.$$

Proof. Abbreviate $\beta = \text{gap}_n + \|y_{n-1:n}\|^2$ and $\gamma = \text{gap}_n|y_n|^2$, and in the expression for $\lambda_{\min}(U_+)$ from Corollary 2.2 write $\lambda_{\min}(U_+) = \lambda_{\min}(A) + \delta$, where

$$\delta = \frac{1}{2}\left(\beta - \sqrt{\beta^2 - 4\gamma}\right) = \frac{2\gamma}{\beta + \sqrt{\beta^2 - 4\gamma}} \leq 2\frac{\gamma}{\beta} \leq \sqrt{\gamma} = |y_n|\sqrt{\text{gap}_n}.$$

The last inequality follows from the fact that the term under the square root is non-negative, i.e. $\beta^2 \geq 4\gamma$. □

2.3. Largest Eigenvalue. We improve the inclusion interval from Weyl's theorem for the largest eigenvalue of $A \pm yy^*$,

$$\begin{aligned} \lambda_{\max}(A) \leq & \lambda_{\max}(A + yy^*) \leq \lambda_{\max}(A) + \|y\|^2 \\ \lambda_2(A) \leq & \lambda_{\max}(A - yy^*) \leq \lambda_{\max}(A) \end{aligned}$$

by taking into account the contribution of y in the eigenspace of $\lambda_{\max}(A)$.

THEOREM 2.4 (Largest Eigenvalue). *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian, $y \in \mathbb{C}^n$, and*

$$\begin{aligned} L_\pm &\equiv \begin{pmatrix} \lambda_1(A) & 0 \\ 0 & \lambda_2(A) \end{pmatrix} \pm \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \begin{pmatrix} \overline{y}_1 & \overline{y}_2 \end{pmatrix} \\ U_\pm &\equiv \begin{pmatrix} \lambda_1(A) & 0 \\ 0 & \lambda_2(A) \end{pmatrix} \pm \begin{pmatrix} y_1 \\ \|y_{2:n}\| \end{pmatrix} \begin{pmatrix} \overline{y}_1 & \|y_{2:n}\| \end{pmatrix}. \end{aligned}$$

Then $\lambda_{\max}(L_\pm) \leq \lambda_{\max}(A \pm yy^*) \leq \lambda_{\max}(U_\pm)$, where

$$\begin{aligned} \lambda_{\max}(A) \leq \lambda_{\max}(L_+) \leq \lambda_{\max}(U_+) \leq \lambda_{\max}(A) + \|y\|^2 \\ \lambda_2(A) \leq \lambda_{\max}(L_-) \leq \lambda_{\max}(U_-) \leq \lambda_{\max}(A). \end{aligned}$$

Proof. Use the fact that $\lambda_{\max}(A) = -\lambda_{\min}(-A)$, and apply Theorem 2.1. □

As in §2.2, we give explicit expressions for the bounds in Theorem 2.4 in terms of the absolute gap between the two largest eigenvalues,

$$\text{gap}_2 \equiv \lambda_{\max}(A) - \lambda_2(A) \geq 0.$$

COROLLARY 2.5 (Largest Eigenvalue). *In Theorem 2.4*

$$\lambda_{\max}(L_\pm) = \lambda_{\max}(A) + \frac{1}{2}\left(-\text{gap}_2 \pm \|y_{1:2}\|^2 + \sqrt{\left(\text{gap}_2 \pm \|y_{1:2}\|^2\right)^2 \mp 4\text{gap}_2|y_2|^2}\right),$$

and

$$\lambda_{\max}(U_\pm) = \lambda_{\max}(A) + \frac{1}{2}\left(-\text{gap}_2 \pm \|y\|^2 + \sqrt{\left(\text{gap}_2 \pm \|y\|^2\right)^2 \mp 4\text{gap}_2\|y_{2:n}\|^2}\right).$$

2.4. Interior Eigenvalues. We improve the inclusion intervals from Weyl's theorem for the interior eigenvalues of $A \pm yy^*$,

$$\lambda_i(A) \le \lambda_i(A + yy^*) \le \lambda_{i-1}(A), \qquad 2 \le i \le n-1 \tag{2.3}$$

$$\lambda_{i+1}(A) \le \lambda_i(A - yy^*) \le \lambda_i(A), \tag{2.4}$$

by using the bounds for the extreme eigenvalues in Theorems 2.1 and 2.4 on principal submatrices.

THEOREM 2.6 (Interior Eigenvalues). *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian, $y \in \mathbb{C}^n$ and*

$$L_{\pm}^{(i)} \equiv \begin{pmatrix} \lambda_{i-1}(A) & 0 \\ 0 & \lambda_i(A) \end{pmatrix} \pm \begin{pmatrix} \|y_{1:i-1}\| \\ y_i \end{pmatrix} \begin{pmatrix} \|y_{1:i-1}\| & \overline{y}_i \end{pmatrix}$$

$$U_{\pm}^{(i)} \equiv \begin{pmatrix} \lambda_i(A) & 0 \\ 0 & \lambda_{i+1}(A) \end{pmatrix} \pm \begin{pmatrix} y_i \\ \|y_{i+1:n}\| \end{pmatrix} \begin{pmatrix} \overline{y}_i & \|y_{i+1:n}\| \end{pmatrix}.$$

Then

$$\lambda_{\min}(L_+^{(i)}) \le \lambda_{\min}(A \pm yy^*) \le \min\{\lambda_{\max}(U_+^{(i)}), \lambda_{i-1}(A)\}, \qquad 2 \le i \le n-1,$$

where $\lambda_i(A) \le \lambda_{\min}(L_+^{(i)}) \le \lambda_{\max}(U_+^{(i)}) \le \lambda_i(A) + \|y_{i:n}\|^2$. Moreover

$$\max\{\lambda_{i+1}(A), \lambda_{\min}(L_-^{(i)})\} \le \lambda_i(A - yy^*) \le \lambda_{\max}(U_-^{(i)}), \qquad 2 \le i \le n-1,$$

where $\lambda_i(A) - \|y_{1:i}\|^2 \le \lambda_{\min}(L_-^{(i)}) \le \lambda_{\max}(U_-^{(i)}) \le \lambda_i(A)$.

Proof. As before, abbreviate $\alpha_j \equiv \lambda_j(A)$, $1 \le j \le n$.

Lower Bounds. Partition the eigenvalue decomposition so that

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}, \qquad \Lambda_1 \equiv \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_i \end{pmatrix}, \qquad \Lambda_2 \equiv \begin{pmatrix} \alpha_{i+1} & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix},$$

and $V = \begin{pmatrix} V_1 & V_2 \end{pmatrix}$ with $V_1 \equiv \begin{pmatrix} v_1 & \dots & v_i \end{pmatrix}$ and $V_2 \equiv \begin{pmatrix} v_{i+1} & \dots & v_n \end{pmatrix}$. Since $V_1^*(A \pm yy^*)V_1$ is a principal submatrix of order i of $V^*(A \pm yy^*)V$, the Cauchy interlace theorem [26, §10.1] implies

$$\lambda_i(A \pm yy^*) = \lambda_i(V^*(A \pm yy^*)V) \ge \lambda_i(V_1^*(A \pm yy^*)V_1) = \lambda_{\min}\left(\Lambda_1 \pm y_{1:i}y_{1:i}^*\right).$$

Apply the lower bounds in Theorem 2.1. The second term in the maximum follows from (2.4).

Upper Bounds. Partition

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}, \qquad \Lambda_1 \equiv \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_{i-1} \end{pmatrix}, \qquad \Lambda_2 \equiv \begin{pmatrix} \alpha_i & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix},$$

and $V = \begin{pmatrix} V_1 & V_2 \end{pmatrix}$ with $V_1 \equiv \begin{pmatrix} v_1 & \dots & v_{i-1} \end{pmatrix}$ and $V_2 \equiv \begin{pmatrix} v_i & \dots & v_n \end{pmatrix}$. Since $V_2^*(A + yy^*)V_2$ is a principal submatrix of order $n-(i-1)$ of $V^*(A + yy^*)V$, the Cauchy interlace theorem implies

$$\lambda_i(A + yy^*) = \lambda_i\left(V^*(A + yy^*)V\right) \le \lambda_1(V_2^*(A + yy^*)V_2) = \lambda_{\max}\left(\Lambda_2 + y_{i:n}y_{i:n}^*\right).$$

Applying the upper bound in Theorem 2.4 yields $\lambda_{\max}(\Lambda_2 + y_{i:n}y_{i:n}^*) \leq \lambda_{\max}(U_+^{(i)})$. The second term in the bound follows from (2.3). □

We use the absolute gap between the ith eigenvalue and its right neighbour,

$$\text{gap}_i \equiv \lambda_{i-1}(A) - \lambda_i(A) \geq 0, \qquad 2 \leq i \leq n$$

to determine explicit expressions for the bounds in Theorem 2.6.

COROLLARY 2.7. *In Theorem 2.6*

$$\lambda_{\min}(L_\pm^{(i)}) = \lambda_i(A) + \frac{1}{2}\left(\text{gap}_i \pm \|y_{1:i}\|^2 - \sqrt{(\text{gap}_i \pm \|y_{1:i}\|^2)^2 \mp 4\text{gap}_i|y_i|^2}\right),$$

and

$$\lambda_{\max}(U_+^{(i)}) = \lambda_i(A) + \frac{1}{2}\left(-\text{gap}_{i+1} \pm \|y_{i:n}\|^2 + \sqrt{\left(\text{gap}_{i+1} \pm \|y_{i:n}\|^2\right)^2 \mp 4\text{gap}_{i+1}\|y_{i+1:n}\|^2}\right).$$

3. Hermitian Perturbations. We present improved perturbation bounds for well-separated eigenvalues of Hermitian matrices. As in the previous section, we start by presenting an application that motivates these bounds.

3.1. Principal Component Analysis Under the Spiked Covariance Model. Principal component analysis (PCA) is a common tool in the analysis of high dimensional data [15, 17]. Given m samples $x_i \in \mathbb{R}^n$, stored in a (mean centered) $m \times n$ matrix X, PCA proceeds in three steps: It computes the *empirical covariance matrix* $C = \frac{1}{m} X^*X$; finds orthonormal directions with maximal variance of the data, represented by the largest eigenvalues and eigenvectors of the matrix C; and at last determines a low dimensional representation of the data from linear projections onto these directions associated with maximal variance.

A common model for the analysis of PCA on high dimensional data is a small rank linear mixture or "spiked covariance model" [8, 16, 25]. Under this model, each data sample x_i is an independent identically distributed random vector of the form

$$x = \sum_{j=1}^{k} u_j v_j + \sigma\xi$$

where u_j are random variables, also referred to as *components* or *latent variables*, the vectors $v_j \in \mathbb{R}^n$ are the *responses*, $\xi \in \mathbb{R}^n$ is a multivariate Gaussian noise vector with identity covariance matrix, and the scalar σ is the level of noise.

If all k random variables u_j are uncorrelated with zero mean and unit variance, and all eigenvectors v_j are orthogonal, then the first k eigenvalues and eigenvector pairs of the population covariance matrix are $(\|v_j\|^2+\sigma^2, v_j)$. Given that we have only a finite dataset $\{x_i\}_{i=1}^m$, the question is how close are the eigenvalues and eigenvectors of the empirical noisy covariance matrix C to their limiting values?

We formulate this problem in terms of matrix perturbation theory by working in an orthonormal basis whose first k vectors are $v_j/\|v_j\|$ and by writing the empirical covariance matrix as

$$C = A + E, \qquad \text{where} \quad A = \begin{pmatrix} \|v_1\|^2 + \sigma^2 & & & \\ & \ddots & & \\ & & \|v_k\|^2 + \sigma^2 & \\ & & & \sigma^2 I_{n-k} \end{pmatrix}.$$

We want to determine under which conditions the first few (largest) eigenvalues and eigenvectors of C correspond to the first few latent variables and characteristic responses of A, and how close these noisy estimates are to the unperturbed eigenvalues and eigenvectors of A. Many papers in statistics have studied the asymptotic distribution of the eigenvalues and eigenvectors of C in the limit as $m \to \infty$, see [1, 2, 9, 15, 24] and references therein. However, in our application we are interested in answers to these questions for a finite number of samples m.

In the context of matrix perturbation theory, we look for absolute normwise perturbation bounds for eigenvalues of Hermitian matrices A and $A + E$. In particular, assuming that the signals u_j have a significant signal to noise ratio, we want bounds for eigenvalues $\lambda_j(A)$ that are well separated from all others, in the sense that $\text{gap}_j > \|E\|$. Moreover, to obtain sharp bounds we cannot afford to deal with the global norm of E, but rather we need to restrict ourselves to the contribution of E in the relevant eigenspace of A. In the present paper, we derive such bounds that depend on the projection of E onto a space spanned by an eigenvector v_j. The analysis is completed in a second paper [25], where we derive probabilistic bounds of the type "$\|Ev_j\| \leq f(m, n)$ with probability $1 - \delta$".

3.2. Perturbation Bounds. Two types of existing two-norm results could potentially be applicable for the application in §3.1: Two-norm bounds that hold for all eigenvalues, and residual bounds that hold for a few eigenvalues. The best known example of a two-norm bound for Hermitian matrices $A, A + E \in \mathbb{C}^{n\times n}$ is Weyl's theorem [11, Theorem 8.1.6], [26, Theorem 10.3.1],

$$|\lambda_j(A) - \lambda_j(A + E)| \leq \|E\|, \qquad 1 \leq j \leq n. \tag{3.1}$$

The advantage of (3.1) is that it applies to all eigenvalues of A and $A + E$. The disadvantage is that the bound is the same for all eigenvalues, and depends on the global norm of E which can be quite large, specifically in high dimensions, $n \gg 1$, [16].

For a single perturbed eigenvalue, one can either tighten the Bauer-Fike theorem [6, §4.6.1], [28, Corollary 3.3] or derive a residual bound from scratch [26, Theorem 4.5.1] as follows. If w_j is a unit norm eigenvector associated with an eigenvalue $\lambda_j(A + E)$, i.e. $(A + E)w_j = \lambda_j(A + E)w_j$, $\|w_j\| = 1$, then

$$\min_i |\lambda_i(A) - \lambda_j(A + E)| \leq \|Ew_j\|. \tag{3.2}$$

The problem is that this bound depends on the a priori unknown projection of E onto the perturbed eigenvector w_j. However, by switching the roles of A and $A + E$, we obtain for each eigenvalue $\lambda_i(A)$,

$$\min_j |\lambda_i(A) - \lambda_j(A + E)| \leq \|Ev_i\|, \qquad 1 \leq i \leq n. \tag{3.3}$$

While this bound depends on the projection of E onto an eigenspace of A, it doesn't pair up $\lambda_i(A)$ with the corresponding perturbed eigenvalue $\lambda_i(A + E)$.

Below we show that such a pairing is possible for eigenvalues $\lambda_i(A)$ that are well separated from the other eigenvalues of A, and that the distance between $\lambda_i(A)$ and $\lambda_i(A+E)$ is bounded only by the projection of E onto the eigenspace of $\lambda_i(A)$, rather than by the full perturbation E. Now we use the two-sided eigenvalue separation,

$$\text{Gap}_i \equiv \min_{j\neq i} |\lambda_i(A) - \lambda_j(A)|, \qquad 1 \leq i \leq n.$$

In the following lemma we present a bound that is probably known, but we were not able to find it in the literature.

LEMMA 3.1. *If $A, A+E \in \mathbb{C}^{n\times n}$ are Hermitian, then for every eigenvalue $\lambda_i(A)$ with $Gap_i > 2\|E\|$*

$$|\lambda_i(A+E) - \lambda_i(A)| \le \|Ev_i\|.$$

Proof. According to (3.3) for every eigenvalue $\lambda_i(A)$ there exists an eigenvalue $\lambda_j(A+E)$ such that

$$|\lambda_i(A) - \lambda_j(A+E)| \le \|Ev_i\|.$$

We now prove that under the gap condition, $j = i$. Weyl's theorem implies

$$|\lambda_j(A) - \lambda_j(A+E)| \le \|E\|.$$

Moreover, all other eigenvalues of A are further from $\lambda_i(A)$, because for $j \ne i$,

$$|\lambda_i(A) - \lambda_j(A+E)| \ge |\lambda_i(A) - \lambda_j(A)| - |\lambda_j(A) - \lambda_j(A+E)| \ge Gap_i - \|E\| > \|E\|.$$

This means, for each eigenvalue $\lambda_i(A)$ satisfying the gap condition there is exactly one eigenvalue of $A+E$ at distance less than $\|E\|$, and this eigenvalue must be $\lambda_i(A+E)$. Therefore, $j = i$. □

The condition $\mathrm{Gap}_i > 2\|E\|$ appears in many other contexts, because it is a sufficient condition that prevents the eigenvalue $\lambda_i(A+\epsilon E)$ from crossing other eigenvalues for $|\epsilon| < 1$ [19, Theorem II.3.9]. Now we improve the gap condition in Lemma 3.1 by a factor of 2, but at the expense of a multiplicative factor of $\sqrt{2}$ in the perturbation bound.

THEOREM 3.2. *If $A, A+E \in \mathbb{C}^{n\times n}$ are Hermitian then for every eigenvalue $\lambda_i(A)$ with $\mathrm{Gap}_i > \|E\|$*

$$|\lambda_i(A) - \lambda_i(A+E)| \le \sqrt{2}\|Ev_i\|.$$

Proof. Fix an index i, and let

$$V^*AV = \begin{pmatrix} \lambda_i(A) & \\ & \Lambda_i \end{pmatrix}$$

be an eigenvalue decomposition of A, where V is unitary with leading column $Ve_1 = v_i$. Partition

$$V^*EV = \begin{pmatrix} f_{ii} & f^* \\ f & E_i \end{pmatrix}$$

conformally with V^*AV. Then we can write $V^*(A+E)V = M + F$ where

$$M \equiv \begin{pmatrix} \lambda_i(A) & \\ & \Lambda_i + E_i \end{pmatrix}, \qquad F \equiv \begin{pmatrix} f_{ii} & f^* \\ f & \end{pmatrix}.$$

From $\sqrt{|f_{ii}|^2 + 4\|f\|^2} \le |f_{ii}| + 2\|f\|$ follows

$$\|F\| = \frac{1}{2}\left(|f_{ii}| + \sqrt{|f_{ii}|^2 + 4\|f\|^2}\right) \le |f_{ii}| + \|f\| \le \sqrt{2}\|Ev_i\|.$$

Weyl's theorem (3.1) implies

$$|\lambda_j(M) - \lambda_j(M+F)| \le \|F\| \le \sqrt{2}\|Ev_i\|, \qquad 1 \le j \le n.$$

Let $\lambda_i(A)$ be the kth eigenvalue of M so that $\lambda_k(M) = \lambda_i(A)$ and $|\lambda_i(A) - \lambda_k(A+E)| \le \sqrt{2}\|Ev_i\|$. We now prove that if $\mathrm{Gap}_i > \|E\|$, then $\lambda_i(A)$ is indeed the ith eigenvalue of M.

Assume $\mathrm{Gap}_i > \|E\|$. Since $\lambda_j(A)$ for $j \ne i$ are the eigenvalues of Λ_i, we can write the eigenvalues of $\Lambda_i + E_i$ as $\lambda_j(A) + \gamma_j$ for $j \ne i$. Weyl's theorem (3.1) implies $|\gamma_j| \le \|E_i\| \le \|E\|$. For $i+1 \le j \le n$ we have

$$\lambda_i(A) - (\lambda_j(A) + \gamma_j) \ge \mathrm{Gap}_i - \|E\| > 0,$$

and for $1 \le j \le i-1$

$$\lambda_j(A) + \gamma_j - \lambda_i(A) \ge \mathrm{Gap}_i - \|E\| > 0.$$

This means, there are exactly $n-i$ eigenvalues of M that are smaller than $\lambda_i(A)$, and $i-1$ eigenvalues that are larger. Thus $\lambda_i(A) = \lambda_i(M)$. □

For every eigenvalue $\lambda_i(A)$, Theorem 3.2 bounds the distance to a perturbed eigenvalue in terms of $\|Ev_i\|$. Since E is Hermitian, $\|Ev_i\| = \|v_i v_i^* E\|$ is the orthogonal projection of E onto the eigenspace of $\lambda_i(A)$.

The bound in Theorem 3.2 is tighter than (3.1) for a particular eigenvalue $\lambda_i(A)$, if the projection of E on the eigenspace of $\lambda_i(A)$ is small compared to $\|E\|$, i.e. if $\sqrt{2}\|Ev_i\| \le \|E\|$. This is typically the case for random perturbations in high dimensions, as in principal component analysis. In contrast to (3.1), which matches up all eigenvalues, Theorem 3.2 bounds the distance between corresponding eigenvalues of A and $A+E$ only for those eigenvalues of A that are sufficiently well separated from all other eigenvalues of A.

4. Non-Hermitian Perturbations. In §3 we showed that a small Hermitian perturbation E of a Hermitian matrix A changes a well separated eigenvalue $\lambda_i(A)$ by at most $\|Ev_i\|$ rather than by the full norm $\|E\|$. We extend this approach to general non-Hermitian perturbations E and obtain bounds that are comparable to those for Hermitian perturbations.

Since a non-Hermitian perturbation of a Hermitian matrix may lead to a non-diagonalizable matrix, there is relatively little work on eigenvalue bounds for non-Hermitian perturbations. A notable exception is the work by Kahan [18, 30], who proved that all eigenvalues of $A+E$ are included in the union of the regions

$$\{z \in \mathbb{C} \,:\, |z - \lambda_k| \le \|E\| \ \text{ and } \ |\Im(z)| \le \|(E - E^*)/2\|\}.$$

If one of these regions is isolated from the others, then it contains exactly one eigenvalue, and if both matrices A and E are real then this eigenvalue must also be real. Another type of eigenvalue bound for general matrices is a Gershgorin theorem [31], which in the simplest form states that for a diagonal matrix $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, all eigenvalues of $A+E$ are in the union of the disks

$$\left\{ z \in \mathbb{C} \,:\, |z - \lambda_k - E_{kk}| \le \sum_{j \ne k} |E_{kj}| \right\}.$$

Below we derive a bound that is sharper whenever a perturbed eigenvalue is close to a well separated eigenvalue, where the separation condition involves the two-sided gap

$$\mathrm{Gap}_i \equiv \min_{j \neq i} |\lambda_i(A) - \lambda_j(A)|.$$

The bound is almost, but not quite, the same as the one for Hermitian perturbations in Theorem 3.2.

THEOREM 4.1. *Let $A, E \in \mathbb{C}^{n\times n}$ where A is Hermitian, let μ be a (possibly complex) eigenvalue of $A + E$, and let $\lambda_i(A)$ be an eigenvalue of A closest to μ, i.e.*

$$|\lambda_i(A) - \mu| = \min_{1\le l\le n} |\lambda_l(A) - \mu|.$$

If $\mathrm{Gap}_i > 3\|E\|$ *then*

$$|\lambda_i(A) - \mu| \le \sqrt{5}\|Ev_i\|.$$

Proof. Abbreviate $\lambda_i \equiv \lambda_i(A)$, and let w be a unit norm eigenvector of μ, i.e. $(A + E)w = \mu w$, $\|w\| = 1$. By assumption λ_i is, among all eigenvalues of A, an eigenvalue that is closest to μ. Thus the Bauer Fike-Theorem (3.2) applied to the Hermitian matrix A and the perturbed eigenvalue μ of the matrix $A + E$ yields

$$|\lambda_i - \mu| \le \|Ew\|. \tag{4.1}$$

We now perform a similarity transformation of A that makes it possible to express the perturbed eigenvector w in terms of the exact eigenvector v_i.

Let $W = \begin{pmatrix} w & W_2 \end{pmatrix}$ be a unitary matrix and perform the similarity transformation

$$W^*AW = \begin{pmatrix} \mu - w^*Ew & b^* \\ b & M \end{pmatrix},$$

where $\|b\| = \|W_2^*Ew\| \le \|Ew\|$. Isolate the block diagonal part, $W^*AW = D + F$, where

$$D \equiv \begin{pmatrix} \mu - w^*Ew & \\ & M \end{pmatrix}, \qquad F \equiv \begin{pmatrix} & b^* \\ b & \end{pmatrix}.$$

The matrix D is Hermitian because it consists of principal submatrices of the Hermitian matrix W^*AW; in particular the scalar $\mu - w^*Ew$ is real. We show in two steps that $\mu - w^*Ew$ is the ith eigenvalue of D.

1. Under the gap condition $\mathrm{Gap}_i > 3\|E\|$, among all eigenvalues of A, λ_i is the only eigenvalue closest to $\mu - w^*Ew$, with all other eigenvalues at a distance of at least $\|E\|$.
   To show this, apply the Bauer Fike-Theorem (3.2) to the leading diagonal element of D to conclude that there exists an eigenvalue $\lambda_k(A)$ so that

   $$|\lambda_k(A) - (\mu - w^*Ew)| \le \|F\| \le \|Ew\|.$$

   We show that $k = i$ by showing that $\mu - w^*Ew$ is too far away from all eigenvalues of A but λ_i. The gap condition implies for $j \neq i$

   $$\begin{aligned} |\lambda_j(A) - (\mu - w^*Ew)| &= |\lambda_j(A) - \lambda_i + \lambda_i - \mu + w^*Ew| \\ &\ge |\lambda_j(A) - \lambda_i| - \|Ew\| - |w^*Ew| \\ &\ge \mathrm{Gap}_i - 2\|Ew\| > \|E\|. \end{aligned}$$

Therefore λ_i is the only eigenvalue of A that is close to $\mu - w^*Ew$. Hence $k = i$, and

$$|\lambda_i - (\mu - w^*Ew)| \leq \|Ew\|. \tag{4.2}$$

2. $\mu - w^*Ew$ is the ith eigenvalue of D.
As in the proof of Theorem 3.2 we show that this follows from the gap condition. Weyl's theorem (3.1),

$$|\lambda_l(A) - \lambda_l(D)| \leq \|F\| \leq \|Ew\|, \qquad 1 \leq l \leq n. \tag{4.3}$$

allow us to write the eigenvalues of D as $\lambda_j(A) + \gamma_j$, where $|\gamma_j| \leq \|Ew\|$. Assuming $\mathrm{Gap}_i > 3\|E\|$, we have for $i + 1 \leq j \leq n$

$$\begin{aligned}\mu - w^*Ew - (\lambda_j(A) + \gamma_j) &= (\mu - w^*Ew - \lambda_i) + (\lambda_i - \lambda_j(A)) + \gamma_j \\ &\geq -|\mu - w^*Ew - \lambda_i| + (\lambda_i - \lambda_j(A)) - |\gamma_j| \\ &\geq \mathrm{Gap}_i - 2\|Ew\| > 0,\end{aligned}$$

where the last inequality follows from (4.2). Similarly for $1 \leq j \leq i - 1$

$$\lambda_j(A) + \gamma_j - (\mu - w^*Ew) \geq \mathrm{Gap}_i - 2\|Ew\| > 0.$$

This means, there are exactly $n - i$ eigenvalues of D that are smaller than $\lambda_i(A)$, and $i - 1$ eigenvalues that are larger. Thus $\mu - w^*Ew = \lambda_i(D)$.

Because $\mu - w^*Ew$ is the ith eigenvalue of D, no eigenvalue of M can be the ith eigenvalue of D. We use this fact to express w in terms of v_i. The partitioning of W^*AW provides a 2×2 system from which one can solve for v_i. Abbreviate

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} w^*v_i \\ W_2^*v_i \end{pmatrix} = W^*v_i.$$

From $Av_i = \lambda_i v_i$ follows

$$0 = (W^*AW - \lambda_i I)W^*v_i = \begin{pmatrix} \mu - w^*Ew - \lambda_i & b^* \\ b & M - \lambda_i I \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}. \tag{4.4}$$

We show that $M - \lambda_i I$ is nonsingular, by showing that λ_i cannot be an eigenvalue of M. Above we established that no eigenvalue of M can be the ith eigenvalue of D. Hence $\lambda_j(M) = \lambda_{j_k}(D)$ for some $j_k \neq i$, and (4.3) and the gap condition imply

$$\begin{aligned}|\lambda_j(M) - \lambda_i| &= |\lambda_{j_k}(D) - \lambda_i| = |(\lambda_{j_k}(A) - \lambda_i) + (\lambda_{j_k}(D) - \lambda_{j_k}(A))| \\ &\geq \mathrm{Gap}_i - \|E\| > 2\|E\| > 0.\end{aligned}$$

Thus λ_i is not an eigenvalue of M, and $M - \lambda_i I$ is nonsingular.

As a consequence we can solve for z_2 in (4.4) and obtain $z_2 = -z_1(M - \lambda_i I)^{-1}b$. Since $z_1 = 0$ would imply $z_2 = 0$, and then $v_i = 0$ we must have $z_1 \neq 0$. From the definition of z_1 and z_2 follows

$$W^*v_i = z_1 \begin{pmatrix} 1 \\ (M - \lambda_i I)^{-1}b \end{pmatrix}.$$

Multiplying on the left by the unitary matrix $W = \begin{pmatrix} w & W_2 \end{pmatrix}$ yields

$$v_i = z_1 \left(w + W_2(M - \lambda_i I)^{-1}b\right), \qquad z_1 = 1/\left\| \begin{pmatrix} 1 \\ (M - \lambda_i I)^{-1}b \end{pmatrix} \right\|.$$

Solving for w gives

$$w = \sqrt{1 + \|(M - \lambda_i I)^{-1} b\|^2}\, v_i - W_2 (M - \lambda_i I)^{-1} b,$$

and a subsequent multiplication by E yields

$$Ew = \sqrt{1 + \|(M - \lambda_i I)^{-1} b\|^2}\, E v_i - E W_2 (M - \lambda_i I)^{-1} b.$$

Thus

$$\|Ew\| \le \sqrt{1 + \|(M - \lambda_i I)^{-1} b\|^2}\, \|E v_i\| + \|E\| \|(M - \lambda_i I)^{-1} b\|.$$

To bound $\|(M - \lambda_i I)^{-1}\|$ from above, we use the fact from item 2 that $\mu - w^* E w$ is the ith eigenvalue of D. As a consequence the eigenvalues of M correspond to $\lambda_j(D)$ for $j \neq i$. This means there is a $k \neq i$ so that

$$\begin{aligned} 1/\|(M - \lambda_i I)^{-1}\| &\ge \min_j |\lambda_j(M) - \lambda_i| = |\lambda_k(D) - \lambda_i| \ge |\lambda_i - \lambda_k(A)| - |\lambda_k(D) - \lambda_k(A)| \\ &\ge \mathrm{Gap}_i - \|E\| > 2\|E\|, \end{aligned}$$

where the next to last inequality follows from (4.3). Hence

$$\|(M - \lambda_i I)^{-1} b\| \le \frac{\|Ew\|}{2\,\|E\|} \le \frac{1}{2}.$$

At last, we substitute this into the above bound for $\|Ew\|$ to obtain

$$\|Ew\| \le \frac{\sqrt{5}}{2} \|E v_i\| + \|E\| \frac{\|Ew\|}{2\,\|E\|} \le \frac{\sqrt{5}}{2} \|E v_i\| + \frac{\|Ew\|}{2},$$

so $\|Ew\|| \le \sqrt{5} \|E v_i\|$. The result follows by substituting this bound for $\|Ew\|$ in (4.1). □

Acknowledgments. We thank Tim Kelley and Alastair Spence for helpful discussions; and we are grateful to Jan Brandts whose valuable suggestions greatly improved the presentation of the paper.

REFERENCES

[1] T. W. Anderson, *Asymptotic theory for principal component analysis*, Ann. Math. Stat., 34 (1963), pp. 122–148.

[2] ———, *An introduction to multivariate statistical analysis*, Wiley, New York, 1984.

[3] J. R. Bunch, C. P. Nielsen, and D. C. Sorensen, *Rank-one modification of the symmetric eigenvalue problem*, Numer. Math., 31 (1978), pp. 31–48.

[4] T. F. C. Chan, *Newton-like pseudo-arclength methods for computing simple turning points*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 135–148.

[5] T. F. C. Chan and H. B. Keller, *Arc-length continuation and multigrid techniques for nonlinear elliptic eigenvalue problems*, SIAM J. Sci. Stat. Comput., 3 (1982), pp. 173–194.

[6] F. Chatelin, *Valeurs Propres de Matrices*, Masson, Paris, 1986.

[7] K. I. Dickson, C. T. Kelley, I. C. F. Ipsen, and I. G. Kevrekidis, *Condition estimation for pseudo-arclength continuation*, SIAM J. Num. Anal., 45 (2007), pp. 263–276.

[8] R. B. Dozier and J. W. Silverstein, *On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices*, Journal of Multivariate Analysis, to appear, (2007).

[9] M. Eaton and D. E. Tyler, *On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix*, Ann. Stat., 19 (1991), pp. 260–271.

[10] G. H. Golub, *Some modified matrix eigenvalue problems*, SIAM Rev., 15 (1973), pp. 318–334.
[11] G. H. Golub and C. F. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, third ed., 1996.
[12] W. Govaerts, *Numerical Methods for Bifurcations of Dynamic Equilibria*, SIAM, Philadelphia, 2000.
[13] W. Govaerts and J. D. Pryce, *A singular value inequality for block matrices*, Linear Algebra Appl., 125 (1989), pp. 141–148.
[14] M. Gu and S. C. Eisenstat, *A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1266–1276.
[15] J. E. Jackson, *A User's Guide to Principal Components*, Wiley, New York, 1991.
[16] I. M. Johnstone, *On the distribution of the largest eigenvalue in principal components analysis*, Ann. Stat., 29 (2001), pp. 295–327.
[17] I. T. Jolliffe, *Principal Component Analysis*, Springer, New York, 2002.
[18] W. Kahan, *Spectra of nearly Hermitian matrices*, Proc. AMS, 48 (1975).
[19] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag, Berlin, 1995.
[20] H. B. Keller, *Lectures on Numerical Methods in Bifurcation Theory*, Tata Institute of Fundamental Research, Lectures on Mathematics and Physics, Springer Verlag, New York, 1987.
[21] M. Krupnik, *Changing the spectrum of an operator by a perturbation*, Linear Algebra Appl., 167 (1992), pp. 113–118.
[22] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer Verlag, New York, 1998.
[23] R. Menzel and H. Schwetlick, *Zur Lösung parameterabhängiger nichtlinearer Gleichungen mit singulären Jacobi-Matrizen*, Numer. Math., (1978), pp. 65–79.
[24] R. J. Muirhead, *Aspects of multivariate statistical theory*, Wiley, New York, 1982.
[25] B. Nadler, *Finite sample approximation results for principal component analysis: A matrix perturbation approach*, (2007). Submitted for publication.
[26] B. N. Parlett, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.
[27] W. C. Rheinboldt, *Solution fields of nonlinear equations and continuation methods*, SIAM J. Numer. Anal., 17 (1980), pp. 221–237.
[28] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, New York, 1992.
[29] A. Spence, *Numerical methods for bifurcation problems*, in The Graduate's Guide to Numerical Analysis, Series in Computational Mathematics, Springer, 1999.
[30] G. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.
[31] R. S. Varga, *Geršgorin and His Circles*, Springer, Heidelberg, 2004.