RITZ VALUE BOUNDS THAT EXPLOIT QUASI-SPARSITY

ILSE C.F. IPSEN*

Abstract. Absolute and relative perturbation bounds for Ritz values of complex square matrices are presented. The bounds have several advantages over existing bounds: They target specified eigenvalues, they are tighter for quasi-sparse eigenvectors, and they do not use the entire matrix. They also lead to new a priori bounds for Ritz values of Hermitian matrices.

Two types of bounds are presented: bounds for single eigenvalues, and simultaneous bounds for several eigenvalues. The single eigenvalue bounds are stronger when customized for Hermitian banded and tridiagonal matrices. The multiple eigenvalue bounds are one-to-one matching bounds. They have no condition number when eigenvalues and Ritz values are real.

A bound for the (relative) accuracy of the relative Ritz value separation is also derived.

Key words. a priori bounds, relative error, eigenvector condition number, matching bounds

AMS subject classification. 65F15, 15A18, 15A42

1. Introduction. The perturbation bounds in this paper were motivated by the Quasi-Sparse Eigenvector (QSE) method [14]. The QSE method computes the eigenvalues with (algebraically) smallest real part of extremely large, possibly infinite Hamiltonian matrices in quantum physics.

More specifically, a QSE iteration approximates eigenvalues with smallest real part of a Hamiltonian matrix

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

by the eigenvalues of a 'truncation' H_{11}, whose dimension is small compared to that of H. Ideally, the relative separation of the computed eigenvalues should have 5 percent accuracy.

We derive perturbation bounds to estimate how well the eigenvalues of H_{11}, which are Ritz values of H, approximate a desired eigenvalues of H. There are three reasons why existing bounds are not sufficient for this purpose.

First, numerical experiments show that the QSE method tends to be fast for matrices whose eigenvectors are quasi-sparse, i.e. have many elements of small magnitude. Existing perturbation bounds for Ritz values, both absolute [15, §11], [7, 9, 11, 13, 17, 18] (and many others, such as a priori bounds for Krylov subspace methods) and relative [1, 3], don't exploit quasi-sparsity.

Second, traditional Ritz value bounds don't have control over which eigenvalues they approximate, and may not give a bound for the desired eigenvalue. Suppose one wants to approximate the smallest eigenvalue $\lambda_1(H) \approx 1$ of

$$H = \begin{pmatrix} 100 & \epsilon \\ \epsilon & 1 \end{pmatrix}, \qquad 0 \leq \epsilon < 1$$

by the eigenvalue 100 of $H_{11} = 100$. The Ritz value bound [15, Theorem (11-5-1)] only gives a bound for the large eigenvalue $\lambda_2(H) \approx 100$, $|100 - \lambda_2(H)| \leq \|H_{12}\| = \epsilon$. It does not give information about the accuracy of the small eigenvalue, $|100 - \lambda_1(H)|$.

*Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, P.O. Box 8205, Raleigh, NC 27695-8205, USA (ipsen@math.ncsu.edu, http://www4.ncsu.edu/~ipsen/). This research was supported in part by NSF grants DMS-0209931 and DMS-0209695.

Third, many eigenvalue perturbation bounds depend on the entire matrix. But when a matrix is extremely large or infinite, one can afford to work with only a small piece (in this sense our motivation is similar to that of the Residual Interlace Theorem [15, §10.4]). For instance, Weyl's theorem for Hermitian matrices [15, Fact 1-11] implies

$$\left|\lambda_i(H) - \lambda_i\begin{pmatrix} H_{11} & 0 \\ 0 & H_{22}\end{pmatrix}\right| \leq \|H_{12}\|,$$

where $\lambda_i(\cdot)$ denotes the ith smallest eigenvalue of a matrix and $\|\cdot\|$ the (Euclidean) two-norm. When H is extremely large or infinite, H_{22} may not be available or may not even be known. Instead we need a bound for $|\lambda_i(H) - \lambda_i(H_{11})|$. Fortunately, the matrices in the problems from [14] are often banded, so H_{12} and H_{21} have few non-zero elements and can be part of a bound.

Overview. Perturbation bounds for the approximation of a single eigenvalue by a Ritz value are presented for complex square matrices (§2) and for Hermitian matrices (§3). They are tighter than existing Ritz value bounds when the eigenvector is quasi-sparse. They become even stronger when specialized to Hermitian banded matrices (§4) and Hermitian tridiagonal matrices (§5). Relative bounds for general complex matrices and Hermitian matrices are presented in §6. In §7 simultaneous for several eigenvalues are discussed. They are one-to-one matching bounds, and do not contain a condition number when the truncation is Hermitian and the desired eigenvalues are real. The bounds for Hermitian matrices lead directly to a priori bounds for Ritz values (§8). At last, perturbation bounds for the relative separation of real eigenvalues are derived in in §9.

Notation. A complex matrix V has transpose V^T and conjugate transpose V^*. The identity matrix is I, the *ith* column is e_i. The eigenvalues of a complex square matrix A are denoted by $\lambda_i(A)$. $\|\cdot\|$ denotes the Euclidean two-norm, and $\|\cdot\|_F$ the Frobenius norm.

2. Diagonalizable Matrices. We derive perturbation bounds for the approximation of *any* eigenvalue by an eigenvalue of a leading principal submatrix. The bounds depend on the magnitude of eigenvector components.

Let H be a complex square matrix with eigenvalues λ_j and corresponding eigenvectors v_j, i.e. $Hv_j = \lambda_j v_j$. Partition

$$H = \begin{array}{cc} & m \\ m & \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22}\end{pmatrix} \end{array}, \qquad v_j = \begin{pmatrix} v_j^{(1)} \\ v_j^{(2)}\end{pmatrix}.$$

The eigenvalues of H_{11} are θ_i, $1 \leq i \leq m$. We want to approximate any eigenvalue λ_j of H by an eigenvalue θ_i of H_{11}.

The bound below extends the Ritz value bound for Hermitian matrices [15, Theorem (11-5-1)] to general, complex matrices.

FACT 1. *If H_{11} is diagonalizable with eigenvector matrix W, and λ_j is non-derogatory then*

$$\min_{1\leq i\leq m} |\theta_i - \lambda_j| \leq \kappa(W)\,\|H_{12}\|\,\rho_j, \qquad \rho_j \equiv \|v_j^{(2)}\|/\|v_j^{(1)}\|,$$

assuming $\kappa(W) \equiv \|W\|\,\|W^{-1}\|$ and $\rho_j < \infty$.

Proof. Write the first block row in $(H-\lambda_j I)v_j = 0$ as $(H_{11}-\lambda_j I)v_j^{(1)} = -H_{12}v_j^{(2)}$. Let $H_{11} = W\Theta W^{-1}$, where Θ is a diagonal matrix with diagonal elements θ_i and take norms. □

The above bound does not depend on the eigenvector condition number of H, only on that of the principal submatrix H_{11}. The bound decreases with the 'quasi-sparsity' ρ_j of the eigenvectors v_j. When λ_j is non-derogatory, it has a single eigenspace which has dimension one, and ρ_j is unique.

Definition 2.1. *The quantity*

$$\rho_j \equiv \|v_j^{(2)}\|/\|v_j^{(1)}\|$$

measures the quasi-sparsity of a vector v_j with regard to the partition $v_j = \begin{pmatrix} v_j^{(1)} \\ v_j^{(2)} \end{pmatrix}$. For a given partition, v_j is quasi-sparse if $\rho_j < 1$.

If $v_j^{(2)} = 0$, i.e. $\rho_j = 0$, then λ_j is an eigenvalue of H_{11}. Necessary and sufficient conditions for λ_j to be a Ritz value are discussed in [8].

If $v_j^{(1)} = 0$ then the eigenvalue ordering is 'wrong' and λ_j is represented by the bottom part H_{22}. For instance,

$$H = \begin{pmatrix} 6 & -2 & 0 \\ -2 & 6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

has eigenvalues 1, 4 and 8. The eigenspace for $\lambda_1 = 1$ is spanned by $v_1 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$. The partition

$$H_{11} = 6, \qquad H_{12} = \begin{pmatrix} -2 & 0 \end{pmatrix}, \qquad H_{22} = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

gives $v_1^{(1)} = 0$ and $\rho_1 = \infty$, because $\lambda_1 = 1$ clearly lives in H_{22}.

We are interested in the case $\|v_j^{(2)}\| < \|v_j^{(1)}\|$ where the eigenvector v_j is quasi-sparse. In this case our perturbation bounds for the eigenvalue λ_j are tighter than existing Ritz value bounds.

3. Hermitian Matrices.

We consider the bounds in §2 for Hermitian matrices, where we can say more about the accuracy of the smallest Ritz value. Our bounds resemble those in [12] (a more thorough discussion is given in §xx).

Label the eigenvalues in ascending order, $\lambda_1 \leq \lambda_2 \leq \ldots$, and $\theta_1 \leq \ldots \leq \theta_m$. The traditional Ritz value bound [15, Theorem (11-5-1)] implies that there are m eigenvalues λ_{i_j} of H such that

$$|\theta_j - \lambda_{i_j}| \leq \|H_{12}\|, \qquad 1 \leq j \leq m. \tag{3.1}$$

The bound below is tighter than (3.1) if the eigenvectors are quasi-sparse.

Fact 2 (Hermitian Matrices). *Let H be Hermitian.*

If λ_1 is distinct (and $\rho_1 < \infty$) then

$$0 \leq \theta_1 - \lambda_1 \leq \|H_{12}\|\, \rho_1,$$

If λ_j is distinct (and $\rho_j < \infty$) then

$$\min_{1\leq i\leq m} |\theta_i - \lambda_j| \leq \|H_{12}\|\, \rho_j, \qquad j \geq 2.$$

Proof. This follows from Fact 1 and the bound for λ_1 from the Cauchy interlace theorem [15, §10-1]. □

In contrast to Fact 2, the traditional bound (3.1) may not give any information about the smallest eigenvalue λ_1. For instance, the eigenvalues of

$$H = \begin{pmatrix} 100 & \epsilon \\ \epsilon & 1 \end{pmatrix}, \qquad 0 \le \epsilon < 1$$

satisfy, according to Weyl's theorem [15, Fact 1-11],

$$1 - \epsilon \le \lambda_1 \le 1 + \epsilon, \qquad 100 - \epsilon \le \lambda_2 \le 100 + \epsilon.$$

With $H_{11} = 100$ and $\theta_1 = 100$, (3.1) gives $|\theta_1 - \lambda_2| \le \|H_{12}\| = \epsilon$, but it does not bound $\theta_1 - \lambda_1$. In contrast, Fact 2 yields

$$\rho_1 \le \frac{99 + \epsilon}{\epsilon} \qquad \text{and} \qquad \theta_1 - \lambda_1 \le \|H_{12}\|\, \rho_1 = 99 + \epsilon.$$

The upper bound is the same as the one implied by Weyl's theorem, $99 - \epsilon \le \theta_1 - \lambda_1 \le 99 + \epsilon$.

A bound similar to, but not quite the same as those in Fact 2 was already presented in [12, Theorem 1.1]. In the special case $m = 1$ it is the same as the first bound in Fact 2. For $m > 1$, [12, Theorem 1.1] is more general because it can pair up any Ritz value θ_i with any eigenvalue λ_j, but it is not clear whether it is tighter than Fact 2. In our notation, [12, Theorem 1.1] for $m > 1$ is

$$|\theta_i - \lambda_j|\, |\cos \angle(w_i, v_j^{(1)})| \le \|H_{12}^* w_i\|\, \rho_j, \qquad 1 \le i \le m, \quad 1 \le j,$$

where $H_{11} w_i = \theta_i w_i$, $\|w_i\| = 1$, and $\theta_i \ne \lambda_j$ is required.

Quasi-Sparsity. The bound below shows how the quasi-sparsity depends on the off-diagonal part. The bound also confirms the observation in [14, §3] that eigenvectors are likely to be quasi-sparse if the spacing between eigenvalues is not too small compared to the size of the off-diagonal entries.

FACT 3 (Quasi-Sparsity). *If H is Hermitian, and λ_j distinct and not an eigenvalue of H_{22} (and $\rho_j < \infty$) then*

$$\rho_j \le \frac{\|H_{12}\|}{\min_k |\lambda_k(H_{22}) - \lambda_j|}, \qquad j \ge 1.$$

In particular, $H_{12} = 0$ implies $\rho_j = 0$ (the denominator in the bound is non-zero because λ_j is distinct by assumption).

Fact 3 implies a quadratic bound similar to [15, Theorem (11-7-1)],

$$\min_{1 \le i \le m} |\theta_i - \lambda_j| \le \frac{\|H_{12}\|^2}{\min_k |\lambda_k(H_{22}) - \lambda_j|}, \qquad j \ge 1, \tag{3.2}$$

provided λ_j is distinct and not an eigenvalue of H_{22}, and $\rho_j < \infty$. This bound is a consequence of Fact 2 and can therefore never be better. In fact, it can be a lot worse.

EXAMPLE 1. *The quadratic bound for θ_1 in (3.2),*

$$0 \le \theta_1 - \lambda_1 \le \frac{\|H_{12}\|^2}{\min_k |\lambda_k(H_{22}) - \lambda_1|},$$

can be arbitrarily worse than the bound in Fact 2.

The Hermitian matrix

$$H = \begin{pmatrix} \theta & \eta & \\ \bar{\eta} & 1 & \\ & & \lambda_1 + \epsilon \end{pmatrix}, \qquad 0 < \epsilon < 1, \quad \theta < |\eta|^2 \text{ real},$$

has eigenvalues $\lambda_1 \equiv \frac{1}{2}(\theta + 1 - \delta) < 1$, *as well as* $\lambda_1 + \epsilon$ *and* $\frac{1}{2}(\theta + 1 + \delta) > 1$, *where* $\delta \equiv \sqrt{4|\eta|^2 + (\theta - 1)^2}$.

Choose $m = 1$, *so* $H_{11} = \theta$ *and* $H_{12} = \begin{pmatrix} \eta & 0 \end{pmatrix}$. *Eigenvectors associated with* λ_1 *are multiples of* $v_1 = \begin{pmatrix} 1 & -\frac{1}{2\eta}(\theta - 1 + \delta) & 0 \end{pmatrix}^T$. *From* $1 - \lambda_1 > 1$ *follows* $\rho_1 = |\eta|/(1-\lambda_1) < \eta$; *and from* $\lambda_1 + \epsilon < \lambda_1 + 1 < 1$ *follows* $\epsilon = \min_k |\lambda_k(H_{22}) - \lambda_1|$. *Fact 2 implies* $\theta - \lambda_1 \leq |\eta|^2$, *but the quadratic bound in Fact 3.2 amounts to* $\theta - \lambda_1 \leq |\eta|^2/\epsilon$, *which is much worse for small* ϵ.

4. Banded Hermitian Matrices. When a Hermitian matrix is banded one can exploit quasi-sparsity and tighten the bounds, especially the one for the smallest eigenvalue. A matrix H with elements h_{ij} has half-bandwidth w if $h_{ij} = 0$ for $i > j+w$.

DEFINITION 4.1. *The quantities*

$$\rho_1^w \equiv \frac{\|v_{m-w+1:m,1}\| \, \|v_{m+1:m+w,1}\|}{\|v_{1:m,1}\|^2}, \qquad \rho_j^w \equiv \frac{\|v_{m+1:m+w,j}\|}{\|v_{1:m,j}\|}, \quad j \geq 2$$

measure the quasi-sparsity of a vector v_j *with regard to the partition*

$$v_j = \begin{pmatrix} v_{1:m,j}^T & v_{m+1:m+w,j}^T & \ldots \end{pmatrix}^T$$

of a matrix with half-bandwidth $w \leq m$. *For a given partition,* v_j *is quasi-sparse if* $\rho_j^w < 1$.

Note that the quasi-sparsity measure for the smallest eigenvalue is stricter than that of the larger eigenvalues. In general $\rho_j^w \leq \rho_j$, because only w rather than all components of $v_j^{(2)}$ participate in the numerator of ρ_j. Hence, the bounds below can be tighter than those for general Hermitian matrices in Fact 2.

FACT 4 (Banded Matrices). *Let* H *be Hermitian with half-bandwidth* $w \leq m$.

If λ_1 *is distinct (and* $\rho_1^w < \infty$*) then*

$$0 \leq \theta_1 - \lambda_1 \leq \|H_{12}\| \, \rho_1^w.$$

If λ_j *is distinct (and* $\rho_j^w < \infty$*) then*

$$\min_{1 \leq i \leq m} |\theta_i - \lambda_j| \leq \|H_{12}\| \, \rho_j^w, \qquad j \geq 2.$$

Proof. Use the fact that $H_{12} = \begin{pmatrix} 0 & 0 \\ L & 0 \end{pmatrix}$ where L is order w. Because $H_{11} - \lambda_1 I$ is positive semi-definite, $(\theta_1 - \lambda_1)\|v_1^{(1)}\|^2 \leq (v_1^{(1)})^*(H_{11} - \lambda_1 I)v_1^{(1)}$. $\square$

5. Hermitian Tridiagonal Matrices. We adapt the bounds for banded matrices to tridiagonal matrices and derive expressions for the Ritz value errors.

Let

$$T \equiv \begin{pmatrix} \alpha_1 & \beta_1 & & \\ \bar{\beta}_1 & \alpha_2 & \beta_2 & \\ & \bar{\beta}_2 & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix}$$

be an unreduced Hermitian tridiagonal matrix, i.e. $\beta_i \neq 0$. The eigenvalues λ_j of T are distinct [15, Lemma (7-7-1)]. Leading and trailing principal submatrices of T are denoted by

$$T_m \equiv \begin{pmatrix} \alpha_1 & \beta_1 & & \\ \bar{\beta}_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{m-1} \\ & & \bar{\beta}_{m-1} & \alpha_m \end{pmatrix}, \qquad \hat{T}_{m+1} \equiv \begin{pmatrix} \alpha_{m+1} & \beta_{m+1} & & \\ \bar{\beta}_{m+1} & \alpha_{m+2} & \beta_{m+2} & \\ & \bar{\beta}_{m+2} & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix},$$

so that

$$T = \begin{pmatrix} T_m & \beta_m e_m e_1^* \\ \bar{\beta}_m e_1 e_m^* & \hat{T}_{m+1} \end{pmatrix},$$

where e_i denotes the ith column of an identity matrix. The leading principal submatrix T_m is also an unreduced tridiagonal with eigenvalues $\theta_1 < \ldots < \theta_m$.

A tridiagonal matrix has half-bandwidth $w = 1$, and the measures for quasi-sparsity are

$$\tau_1 \equiv \rho_1^1 = \frac{|v_{m,1}|\,|v_{m+1,1}|}{\|v_{1:m,1}\|^2}, \qquad \tau_j \equiv \rho_j^1 = \frac{|v_{m+1,j}|}{\|v_{1:m,j}\|}, \quad j \geq 2.$$

Since an unreduced Hermitian tridiagonal has distinct eigenvalues, all eigenspaces are one-dimensional, and the leading component of each eigenvector is non-zero [15, Theorem (7-9-5)], i.e. $v_{1:m,j} \neq 0$. Therefore τ_j is always well-defined. Moreover, all elements of an eigenvector v_1 for the smallest eigenvalue are nonzero [15, Theorem (7-9-5)], hence $\tau_1 > 0$.

Fact 5 (Tridiagonal Matrices). *Let T be unreduced Hermitian tridiagonal. Then*

$$0 \leq \theta_1 - \lambda_1 = c_1\, |\beta_m|\, \tau_1, \qquad \min_{1\leq i\leq m} |\theta_i - \lambda_j| = c_j\, |\beta_m|\, \tau_j, \quad j \geq 2,$$

where $0 \leq c_j \leq 1$, $j \geq 1$, and

$$c_1 \equiv \frac{\sum_{i=1}^m |\gamma_{i1}|^2}{|\gamma_{m1}|\, \prod_{i=2}^m |\theta_i - \lambda_1|}, \qquad c_j \equiv \frac{(\sum_{i=1}^m |\gamma_{ij}|^2)^{1/2}}{\prod_{i=1, i\neq k}^m |\theta_i - \lambda_j|}, \quad j \geq 2,$$

$\gamma_{1j} \equiv \beta_1 \cdots \beta_{m-1}$,

$$\gamma_{ij} \equiv \beta_i \cdots \beta_{m-1}\, \det(\lambda_j I - T_{i-1}), \qquad 2 \leq i \leq m-1,$$

and $\gamma_{mj} = \det(\lambda_j I - T_{m-1})$.

Proof. An eigenvector v_j is a multiple of [15, §7-10], [20, §5.48]

$$\begin{pmatrix} \gamma_{1j} & \ldots & \gamma_{m-1,j} & \gamma_{mj} & \frac{\det(\lambda_j I - T_m)}{\beta_m} & \frac{\det(\lambda_j I - T_{m+1})}{\beta_m \beta_{m+1}} & \ldots \end{pmatrix}^T.$$

If λ_j is an eigenvalue of T_m then $\tau_j = 0$, and the desired equalities hold. Now assume that λ_j is not an eigenvalue of T_m.

Using the above expression in τ_1 yields

$$\tau_1 = \frac{|v_{m,1}|\,|v_{m+1,1}|}{\|v_{1:m,1}\|^2} = \frac{|\gamma_{m1}|\,|\det(\lambda_1 I - T_m)|}{|\beta_m|\, \sum_{i=1}^m |\gamma_{i1}|^2} = \frac{|\gamma_{mj}|\,|\theta_1 - \lambda_1| \cdots |\theta_m - \lambda_1|}{|\beta_m|\, \sum_{i=1}^m |\gamma_{ij}|^2}.$$

Solving for $\theta_1 - \lambda_1$ gives $\theta_1 - \lambda_1 = c_1 |\beta_m| \tau_1$, where $c_1 \geq 0$. Since all elements of v_1 are non-zero [15, Theorem (7-9-5)], $\gamma_{m1} \neq 0$ and c_1 is well-defined. The proof for $j \geq 2$ is similar.

Fact 4 implies $\min_{1\leq i\leq m} |\theta_i - \lambda_j| \leq |\beta_m|\, \tau_j$, which means $c_j \leq 1$. ◻

If T is almost decoupled, i.e. $|\beta_m|$ is small, and if v_j is quasi-sparse then some Ritz value θ_i is close to λ_j. The quantity c_j indicates the tightness of the bound in Fact 4 for tridiagonal matrices,

$$\min_{1\leq i\leq m} |\theta_i - \lambda_j| \leq |\beta_m|\, \tau_j.$$

The bound can be loose if λ_j is well separated from all but one eigenvalue of T_m.

As in Fact 3, one can bound the quasi-sparsity.

FACT 6. *If T is an unreduced Hermitian tridiagonal and λ_j is not an eigenvalue of $\hat{T}_{m+1}$ then*

$$\tau_1 = |\beta_m|\, \frac{|v_{m,1}|^2}{\|v_{1:m,1}\|^2}\, e_1^*(\hat{T}_{m+1} - \lambda_1 I)^{-1} e_1$$

and

$$\tau_j = |\beta_m|\, \frac{|v_{m,j}|}{\|v_{1:m,j}\|}\, |e_1^*(\hat{T}_{m+1} - \lambda_j I)^{-1} e_1|, \qquad j \geq 2.$$

Thus $\tau_j \leq |\beta_m| |e_1^*(\hat{T}_{m+1} - \lambda_j I)^{-1} e_1|$. This means, an eigenvector v_j is quasi-sparse if the off-diagonal part β_m and the leading diagonal element of $(\hat{T}_{m+1} - \lambda_j I)^{-1}$ are small in magnitude.

Example 1 illustrates that the quadratic bounds (3.2) for general Hermitian matrices can be much worse than the quasi-sparsity bounds in Fact 2. This is not true for tridiagonal matrices: the quadratic bounds below are equal to the quasi-sparse bounds in Fact 5 because the expression for τ_j in Fact 6 holds with equality,

$$\theta_1 - \lambda_1 \leq |\beta_m|^2\, \frac{|v_{m,1}|^2}{\|v_{1:m,1}\|^2}\, e_1^*(\hat{T}_{m+1} - \lambda_1 I)^{-1} e_1 \tag{5.1}$$

$$\min_{1\leq i\leq m} |\theta_i - \lambda_j| \leq |\beta_m|^2\, \frac{|v_{m,j}|}{\|v_{1:m,j}\|}\, |e_1^*(\hat{T}_{m+1} - \lambda_j I)^{-1} e_1|, \qquad j \geq 2, \tag{5.2}$$

provided λ_j is not an eigenvalue of $\hat{T}_{m+1}$.

EXAMPLE 2 (Toeplitz Matrices). *The real symmetric tridiagonal Toeplitz matrix*

$$T \equiv \begin{pmatrix} \alpha & \beta & & \\ \beta & \alpha & \ddots & \\ & \ddots & \ddots & \beta \\ & & \beta & \alpha \end{pmatrix}, \qquad \beta > 0,$$

of order n has as smallest eigenvalue [16, §2.6.2] $\lambda_1 = \alpha + 2\beta\cos\left(\frac{n\pi}{n+1}\right)$ and eigenvector

$$v_1 = \frac{2}{\sqrt{n+1}} \left(\sin\left(\tfrac{1\pi}{n+1}\right) \quad \sin\left(\tfrac{2\pi}{n+1}\right) \quad \dots \quad \sin\left(\tfrac{n\pi}{n+1}\right) \right)^T.$$

For $m \ll n$ approximate $\sin x \approx x$ and $\cos x \approx 1 - \frac{x^2}{2}$. Then the error in the smallest Ritzvalue is

$$\theta_1 - \lambda_1 = 2\beta \left(\cos \frac{\pi}{n+1} - \cos \frac{\pi}{m+1} \right) \approx \beta \, \frac{\pi^2}{(m+1)^2},$$

while Fact 5 gives the bound

$$\beta \, \tau_1 \approx \frac{6\beta}{2m+1}.$$

That is, the error is proportional to β/m^2 and the bound is proportional to β/m. Therefore the error bound predicts correctly that the error is proportional to the magnitude β of the offdiagonal elements.

6. Relative Bounds. We derive perturbation bounds on the relative error for eigenvalues of a leading principal submatrix of H. Relative eigenvalue bounds are surveyed in [6].

The relative error bound below corresponds to the absolute bound in Fact 1. It resembles the relative bounds in [4, §5] but exploits quasi-sparsity.

FACT 7. *If H_{11} is non-singular and diagonalizable with eigenvector matrix W, and λ_j is non-derogatory (and $\rho_j < \infty$) then*

$$\min_{1 \leq i \leq m} \left| \frac{\theta_i - \lambda_j}{\theta_i} \right| \leq \kappa(W) \, \|H_{11}^{-1} H_{12}\| \, \rho_j,$$

where $\kappa(W) \equiv \|W\| \, \|W^{-1}\|$.

Proof. Write the first block row of $(H - \lambda_j I)v_j = 0$ as $(I - \lambda_j H_{11}^{-1})v_j^{(1)} = -H_{11}^{-1} H_{12} v_j^{(2)}$. □

Like the absolute bound in Fact 1, the relative bound decreases with the quasi-sparsity. The bound itself is also relative in the sense that the off-diagonal part H_{12} is 'normalized' by H_{11}.

When H is Hermitian one can bound the relative error between jth eigenvalue and Ritz value for the m smallest eigenvalues of H (m is the dimension of H_{11}), provided the error is sufficiently small compared to the eigenvalue separation. To prove the relative bounds we define the eigenvalue separation as

$$\Delta_k \equiv \frac{\lambda_{k+1} - \lambda_k}{\max\{|\lambda_k|, |\lambda_{k+1}|\}}, \qquad k \geq 1.$$

First we derive a bound that holds without regard to quasi-sparsity.

FACT 8 (m Smallest Eigenvalues). *Let H be Hermitian and H_{11} be nonsingular; and let the m smallest eigenvalues $\lambda_1 < \ldots < \lambda_m$ of H be distinct and non-zero. Let $\lambda_i = \theta_i(1 + \epsilon_i)$.*

If $|\epsilon_1| < 1$ then

$$|\epsilon_1| = \min_{1 \leq i \leq m} \left| \frac{\theta_i - \lambda_1}{\theta_i} \right|.$$

If for some $2 \leq i \leq m$,

$$|\epsilon_k| < \min \left\{ \frac{1}{2}\Delta_{k-1}, \, \frac{1}{2}\Delta_k, \, 1 \right\}, \qquad 1 \leq k \leq i,$$

then[1]

$$|\epsilon_i| = \min_{1\le l\le m} \left|\frac{\theta_l - \lambda_i}{\theta_l}\right|.$$

Proof. The Cauchy interlace theorem [15, §10-1] implies for the m smallest eigenvalues of H, $\lambda_i \le \theta_i$, $1 \le i \le m$. The case $\lambda_i < 0 < \theta_i$ cannot occur because $1 < 1-\frac{\lambda_i}{\theta_i}$ contradicts the assumption $|\epsilon_i| < 1$.

$i = 1$. For $\lambda_1 > 0$ or $\theta_1 < 0$ one gets, respectively,

$$0 \le 1 - \frac{\lambda_1}{\theta_1} \le 1 - \frac{\lambda_1}{\theta_i} \qquad \text{or} \qquad 0 \ge 1 - \frac{\lambda_1}{\theta_1} \ge 1 - \frac{\lambda_1}{\theta_i}, \qquad i \ge 2.$$

Thus

$$\left|\frac{\theta_1 - \lambda_1}{\theta_1}\right| = \min_{1\le i\le m} \left|\frac{\theta_i - \lambda_1}{\theta_i}\right|.$$

$i = 2$. As above one shows

$$\left|\frac{\theta_2 - \lambda_2}{\theta_2}\right| = \min_{2\le i\le m} \left|\frac{\theta_i - \lambda_2}{\theta_i}\right|.$$

It remains to show that θ_2 is closer to λ_2 than θ_1.

For $\lambda_1 > 0$ or $\lambda_2 < 0$ the assumption $|\epsilon_1| < \Delta_1$ implies $\theta_1 < \lambda_2$. For $\theta_1 < 0$ and $\lambda_2 > 0$ this is true automatically. Therefore $\lambda_1 \le \theta_1 < \lambda_2 \le \theta_2$. Hence

$$\frac{\lambda_2 - \theta_1}{\theta_1} = z + \frac{\theta_2 - \lambda_2}{\theta_2},$$

where

$$z \equiv \frac{\lambda_2 - \theta_1}{\theta_1} - \frac{\theta_2 - \lambda_2}{\theta_2} = \frac{\lambda_2 - \lambda_1}{\lambda_1} + \frac{\lambda_2}{\lambda_1}\,\epsilon_1 + \epsilon_2.$$

If $\theta_2 < 0$ then $z < 0$. When $\lambda_1 > 0$ write

$$z = \frac{\lambda_2}{\lambda_1}\left(\frac{\lambda_2 - \lambda_1}{\lambda_2} + \epsilon_1 + \frac{\lambda_1}{\lambda_2}\,\epsilon_2\right)$$

which shows $z > 0$. If $\theta_1 < 0 < \lambda_2$ then

$$\frac{\lambda_2 - \theta_1}{|\theta_1|} > 1 > |\epsilon_2|.$$

Therefore

$$\left|\frac{\theta_2 - \lambda_2}{\theta_2}\right| = \min_{1\le i\le m} \left|\frac{\theta_i - \lambda_2}{\theta_i}\right|.$$

[1] The boundary conditions are $|\epsilon_1| \le \min\{\frac{1}{2}\Delta_1, 1\}$ and $|\epsilon_m| \le \min\{\frac{1}{2}\Delta_{m-1}, 1\}$.

$i \geq 3$. The proof proceeds by induction and is similar to the case $i = 2$. □

Therefore, if the relative distances between the first $i-1$ eigenvalues and Ritz values are sufficiently small (compared to the separation of the adjacent eigenvalues) then θ_i is the Ritz value closest to λ_i in the relative sense. As in the case of absolute bounds, the bound for the smallest eigenvalue requires the fewest assumptions. If H is Hermitian positive-definite the condition on ϵ_k simplifies to

$$|\epsilon_k| \leq \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}.$$

Now we add quasi-sparsity.

COROLLARY 6.1. *Let H be Hermitian and H_{11} be nonsingular; and let the m smallest eigenvalues $\lambda_1 < \ldots < \lambda_m$ of H be distinct and non-zero.*

If $\|H_{11}^{-1} H_{12}\| \rho_1 < 1$ then

$$\left| \frac{\theta_1 - \lambda_1}{\theta_1} \right| \leq \|H_{11}^{-1} H_{12}\| \rho_1.$$

If for some $2 \leq i \leq m$,

$$\|H_{11}^{-1} H_{12}\| \rho_k \leq \min\left\{ \frac{1}{2}\Delta_{k-1},\, \frac{1}{2}\Delta_k,\, 1 \right\}, \qquad 1 \leq k \leq i,$$

then[2]

$$\left| \frac{\theta_i - \lambda_i}{\theta_i} \right| \leq \|H_{11}^{-1} H_{12}\| \rho_i.$$

Proof. Follows from Facts 7 and 8. □

Therefore, if the bound in Fact 7 is small compared to the eigenvalue separation, then Fact 7 bounds the relative distance between ith Ritzvalue and eigenvalue.

7. Several Eigenvalues. We present simultaneous bounds for all eigenvalues of H_{11}. They are one-to-one matching bounds, and do not contain a condition number when the truncation is Hermitian and the desired eigenvalues are real.

Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of a complex square matrix H. Set

$$\Lambda \equiv \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}, \qquad V \equiv (\, v_1 \quad \ldots \quad v_m \,),$$

where v_j is an eigenvector for λ_j, so $HV = V\Lambda$. Partition V conformally with H, $V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$.

Below we use the quantities $\|V_2 V_1^{-1}\|$ and $\|V_2\| \|V_1^{-1}\|$ to measure the block quasi-sparsity of the vectors V with regard to the partition $(\, V_1^T \quad V_2^T \,)^T$. They appear to represent a more stringent measure of quasi-sparsity then ρ_j from Definition 2.1 because

$$\min_{1 \leq i \leq m} \rho_i \leq \sqrt{m} \|V_2 V_1^{-1}\| \leq \sqrt{m} \|V_2\| \, \|V_1^{-1}\|.$$

[2] The boundary conditions are $\|H_{11}^{-1} H_{12}\| \rho_1 \leq \min\{\frac{1}{2}\Delta_1, 1\}$ and $\|H_{11}^{-1} H_{12}\| \rho_m \leq \min\{\frac{1}{2}\Delta_{m-1}, 1\}$.

The bound below extends [15, Theorem (11-5-7)] from Hermitian to general matrices.

FACT 9. *If $\lambda_1, \ldots, \lambda_m$ are distinct, V_1 is non-singular, and H_{11} is diagonalizable with eigenvector matrix W then there exists a permutation $\sigma(\cdot)$ so that*

$$\left(\sum_{i=1}^{m} |\theta_{\sigma(i)} - \lambda_i|^2\right)^{1/2} \leq \sqrt{m}\, \kappa(W)\, \|H_{12}\|_F\, \|V_2 V_1^{-1}\|,$$

where $\kappa(W) \equiv \|W\|\, \|W^{-1}\|$.

Moreover,

$$H_{11} + E = V_1 \Lambda V_1^{-1}, \qquad \|E\| \leq \|H_{12}\|\, \|V_2 V_1^{-1}\|.$$

Proof. Write the first block row of $H \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \Lambda$ as

$$H_{11} - V_1 \Lambda V_1^{-1} = -H_{12} V_2 V_1^{-1},$$

and decompose $H_{11} = W \Theta W^{-1}$. Since Θ is normal, [19, Theorem 1.1], [2, Problem VI.8.11] imply that there is a permutation $\sigma(\cdot)$ so that

$$\left(\sum_{i=1}^{m} |\theta_{\sigma(i)} - \lambda_i|^2\right)^{1/2} \leq \sqrt{m}\, \|W^{-1}\, H_{12} V_2 V_1^{-1}\, W\|_F.$$

□

Fact 9 provides a bound for a one-to-one matching of eigenvalues and Ritz values. The quasi-sparsity measure $\|V_2 V_1^{-1}\|$ does not depend on the conditioning of the columns of V. When V is block quasi-sparse, i.e. $\|V_2 V_1^{-1}\| \leq 1$, then the Ritz value errors are determined by the (computable) condition number $\kappa(W)$ of the Ritz vectors. Moreover, quasi-sparsity also implies that H_{11} is close to a matrix with eigenvectors V_1 and eigenvalues λ_i.

FACT 10 (Real Ritz Values). *If, in addition to the conditions of Fact 9, $\theta_1 < \cdots < \theta_m$ are also real then*

$$\left(\sum_{i=1}^{m} |\theta_i - \lambda_i|^2\right)^{1/2} \leq \sqrt{2}\, \kappa(W)\, \|H_{12}\|_F\, \|V_2 V_1^{-1}\|,$$

where $\Re(\lambda_1) \leq \ldots \leq \Re(\lambda_m)$.

Proof. In the proof of Fact 9 use the fact that Θ is real, hence Hermitian, and apply [10, §0, (ii)], [2, Problem VI.8.7]. □

Fact 10 differs from Fact 9 in that the ordering of eigenvalues and Ritz values is now known: they are ordered with regard to increasing real parts. In the special case $m = 1$, when Fact 10 bounds the error in a single Ritz value, it is by a factor $\sqrt{2}$ worse than Fact 1.

The next bound, for a Hermitian leading principal submatrix and desired real eigenvalues, contains no condition number.

COROLLARY 7.1 (Hermitian Truncation). *If H_{11} is Hermitian, $\lambda_1 < \ldots < \lambda_m$, and V_1 non-singular then*

$$\left(\sum_{i=1}^{m} |\theta_i - \lambda_i|^2\right)^{1/2} \leq \sqrt{2}\, \|H_{12}\|_F\, \|V_2 V_1^{-1}\|.$$

In the bounds below we remove the condition number $\kappa(W)$ of the Ritz vectors, at the expense of making the quasi-sparsity measure more stringent.

FACT 11. *If $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of H, and V_1 is non-singular, then there is a permutation $\sigma(\cdot)$ so that*

$$\left(\sum_{i=1}^{m} |\theta_{\sigma(i)} - \lambda_i|^2\right)^{1/2} \leq \sqrt{m}\, \|H_{12}\|_F\, \|V_2\|\, \|V_1^{-1}\|.$$

Proof. Write $V_1^{-1} H_{11} V_1 - \Lambda = -V_1^{-1} H_{12} V_2$. Since Λ is normal, [19, Theorem 1.1], [2, Problem VI.8.11] imply that there is a permutation $\sigma(\cdot)$ so that

$$\begin{aligned} \left(\sum_{i=1}^{m} |\theta_{\sigma(i)} - \lambda_i|^2\right)^{1/2} &= \left(\sum_{j=1}^{m} |\lambda_{\sigma(i)}(V_1^{-1} H_{11} V_1) - \lambda_i|^2\right)^{1/2} \\ &\leq \sqrt{m}\, \|H_{12}\|_F\, \|V_2\|\, \|V_1^{-1}\|. \end{aligned}$$

□

For a block of vectors to be quasi-sparse, V_1 must be well-conditioned with respect to inversion and $\|V_2\|$ must be small. Unfortunately, unlike the quasi-sparsity measure in Facts 9 and 10, $\|V_2\|\, \|V_1^{-1}\|$ does depend on the conditioning of the columns of V. The bound can be improved when the desired eigenvalues are real.

FACT 12 (Real Eigenvalues). *If, in addition to the conditions of Fact 11, $\lambda_1 < \ldots < \lambda_m$ are also real, then*

$$\left(\sum_{i=1}^{m} |\theta_i - \lambda_i|^2\right)^{1/2} \leq \sqrt{2}\, \|H_{12}\|_F\, \|V_2\|\, \|V_1^{-1}\|,$$

where $\Re(\theta_1) \leq \ldots \leq \Re(\theta_m)$.

Proof. In the proof of Fact 11, use the fact Λ is Hermitian and apply [10, §0, (ii)], [2, Problem VI.8.7]. □

8. A Priori Bounds for Hermitian Matrices. We present two a priori bounds for Ritz values of Hermitian matrices, one for a single Ritz value and on for several Ritz values.

Let A be a Hermitian matrix with eigenvalues λ_j and corresponding eigenvectors v_j, i.e. $Av_j = \lambda_j v_j$. Let $(\,Q \quad Q_\perp\,)$ be unitary. The eigenvalues of Q^*AQ are θ_i. The bound below can be viewed as a spcial case of [12, Theorem 3.2] with $\tilde{R} = \tilde{Q}$, and it is tighter than the last bound in [18, (2.1)].

FACT 13. *Let A be Hermitian, and $S \equiv \mathrm{range}(Q)$, where Q has orthonormal columns.*

If λ_1 is distinct (and $Q^*v_1 \neq 0$) then

$$0 \leq \theta_1 - \lambda_1 \leq \|Q^*AQ_\perp\|\, \tan\angle(S, v_1).$$

If λ_j is distinct (and $Q^*v_j \neq 0$) then

$$\min_i |\theta_i - \lambda_j| \leq \|Q^*AQ_\perp\|\, \tan\angle(S, v_j), \qquad j \geq 2.$$

Proof. Follows from Fact 2 after the basis transformation $H = (\,Q \quad Q_\perp\,)^*\, A\, (\,Q \quad Q_\perp\,)$.
$\square$

Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of a complex square matrix A. Set

$$\Lambda \equiv \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m s \end{pmatrix}, \qquad V \equiv (\,v_1 \quad \ldots \quad v_m\,),$$

where v_j is an eigenvector for λ_j, so $HV = V\Lambda$.

FACT 14. *Let A be Hermitian, and $S = \text{range}(Q)$ with $\dim(S) = m$. If $\lambda_1 < \ldots < \lambda_m$, and Q^*V is non-singular then*

$$\left(\sum_{i=1}^{m} |\theta_i - \lambda_i|^2\right)^{1/2} \leq \sqrt{2}\, \|Q^* A Q_\perp\|_F \, \tan \angle(S, V).$$

Proof. Similar to the proof of Fact 13, this follows from Corollary 7.1, and from the CS decomposition [5, Theorem 2.6.2]. $\square$

9. Relative Separation. One of the requirements for the QSE method [14] is that the computed eigenvalues have a relative separation that is accurate to at least 5 percent. We present a perturbation bound for the relative separation of the Ritz values, when eigenvalues and Ritz values are real.

We use the same stringent concept of separation as in the previous section,

$$\Delta_k(\lambda) \equiv \frac{\lambda_{k+1} - \lambda_k}{\max\{|\lambda_k|,\, |\lambda_{k+1}|\}}, \qquad \Delta_k(\theta) \equiv \frac{\theta_{k+1} - \theta_k}{\max\{|\theta_k|,\, |\theta_{k+1}|\}},$$

where $\lambda_k \leq \lambda_{k+1}$ and $\theta_k \leq \theta_{k+1}$. The relative accuracy of $\Delta_k(\theta)$ is

$$\left| \frac{\Delta_k(\lambda) - \Delta_k(\theta)}{\Delta_k(\lambda)} \right|.$$

FACT 15. *Let $\lambda_1 < \lambda_2$ and $\theta_1 < \theta_2$ be real and non-zero with $\lambda_1 = \theta_1(1+\epsilon_1)$ and $\lambda_2 = \theta_2(1+\epsilon_2)$ where $|\epsilon_1|, |\epsilon_2| < \epsilon$ for some $0 \leq \epsilon < 1$.*
Then

$$\left| \frac{\Delta_1(\lambda) - \Delta_1(\theta)}{\Delta_1(\lambda)} \right| \leq \frac{1}{1 - \frac{\mu}{M}} \, \frac{2\epsilon}{1-\epsilon},$$

where $\mu \equiv \min\{|\lambda_1|,\, |\lambda_2|\}$ and $M \equiv \max\{|\lambda_1|,\, |\lambda_2|\}$.

Proof. The assumption $|\epsilon_i| < 1$ assures that λ_i and θ_i have the same sign. $\square$

The factor $1/(1 - \frac{\mu}{M})$ is a condition number for the relative separation. It's basically the same as the condition number for subtraction. The condition number is close to one, if the relative separation between λ_1 and λ_2 is large. The accuracy requirement of 5 percent for the QSE method is not so hard to achieve, as the following example illustrates. Suppose λ_1 and λ_2 are accurate to 8 digits, $\epsilon \leq 10^{-8}$. To obtain a relative accuracy of at least .05 for the relative separation, it suffices to have $\lambda_1 \leq (1 - 10^{-6})\lambda_2$.

COROLLARY 9.1. *Let H be Hermitian and H_{11} be nonsingular; and let the m smallest eigenvalues $\lambda_1 < \ldots < \lambda_m$ of H be distinct and non-zero.*

If $\|H_{11}^{-1}H_{12}\|\rho_i < 1$, $1 \le i \le m$, *then*

$$\left|\frac{\Delta_i(\lambda) - \Delta_i(\theta)}{\Delta_i(\lambda)}\right| \le \frac{1}{1 - \frac{\mu_i}{M_i}} \, \frac{2\epsilon_i}{1 - \epsilon_i}, \qquad 1 \le i \le m-1,$$

where $\epsilon_i \equiv \|H_{11}^{-1}H_{12}\| \, \max\{\rho_i, \rho_{i+1}\}$, *and*

$$\mu_i \equiv \min\{|\lambda_i|, \, |\lambda_{i+1}|\}, \qquad M_i \equiv \max\{|\lambda_i|, \, |\lambda_{i+1}|\}.$$

Therefore, if the Ritz values are sufficiently accurate then the accuracy of the Ritz value separation is comparable to the accuracy of the Ritz values. Note that the conditions for an accurate Ritz value separation are less stringent than the ones in Corollary 6.1 that guarantee the pairing up of a Ritz value with the corresponding eigenvalue.

Acknowledgements. I thank Andrew Knyazev for helpful discussions on a priori bounds, which lead to the bounds in §8; and Beresford Parlett for suggesting not to separate the off-diagonal part from the eigenvector, which motivated the results on tridiagonal matrices.

REFERENCES

[1] C. Beattie and I. Ipsen, *Inclusion regions for matrix eigenvalues*, Linear Algebra Appl., 358 (2003), pp. 281–91.
[2] R. Bhatia, *Matrix Analysis*, Springer Verlag, New York, 1997.
[3] Z. Drmač and V. Hari, *Relative residual bounds for the eigenvalues of a Hermitian semidefinite matrix*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 21–9.
[4] S. Eisenstat and I. Ipsen, *Three absolute perturbation bounds for matrix eigenvalues imply relative bounds*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 149–58.
[5] G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, third ed., 1996.
[6] I. Ipsen, *Relative perturbation results for matrix eigenvalues and singular values*, in Acta Numerica 1998, vol. 7, Cambridge University Press, Cambridge, 1998, pp. 151–201.
[7] Z. Jia and G. Stewart, *An analysis of the Rayleigh-Ritz method for approximating eigenspaces*, Math. Comp., 70 (2001), pp. 637–47.
[8] C. Johnson and B. Kroschel, *Principal submatrices, geometric multiplicities, and structured eigenvectors*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1004–12.
[9] W. Kahan, *Inclusion theorems for clusters of eigenvalues of Hermitian matrices*, Tech. Rep. CS42, Computer Science Department, University of Toronto, 1967.
[10] ———, *Spectra of nearly Hermitian matrices*, Proc. Amer. Math. Soc., 48 (1975), pp. 11–7.
[11] S. Kaniel, *Estimates for some computational techniques in linear algebra*, Math. Comp., 20 (1966), pp. 369–78.
[12] A. Knyazev, *New estimates for Ritz values*, Math. Comp., 66 (1997), pp. 985–95.
[13] A. Kuijlaars, *Which eigenvalues are found by the Lanczos method?*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 306–21.
[14] D. Lee, N. Salwen, and D. Lee, *The diagonalization of quantum field Hamiltonians*, Phys. Lett. B, 503 (2001), pp. 223–35.
[15] B. Parlett, *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, 1980.
[16] P. Roebuck and S. Barnett, *A survey of Toeplitz and related matrices*, Int. J. Systems Sci., 9 (1978), pp. 921–34.
[17] Y. Saad, *Error bounds on the interior Rayleigh-Ritz approximations from Krylov subspaces*, SIAM J. Numer. Anal., 17 (1980), pp. 687–706.
[18] G. Sleijpen, J. Van Den Eshof, and P. Smit, *Optimal a priori bounds for the Rayleigh-Ritz method*, Math. Comp., 72 (2002), pp. 677–84.
[19] J. Sun, *On the variation of the spectrum of a normal matrix*, Linear Algebra Appl., 246 (1996), pp. 215–23.
[20] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.