FAST AND ACCURATE CON-EIGENVALUE ALGORITHM FOR OPTIMAL RATIONAL APPROXIMATIONS

T. S. HAUT AND G. BEYLKIN

Abstract. The need to compute small con-eigenvalues and the associated con-eigenvectors of positive-definite Cauchy matrices naturally arises when constructing rational approximations with an optimally small L^∞ error. Specifically, given a rational function with n poles in the unit disk, a rational approximation with $m \ll n$ poles in the unit disk may be obtained from the mth con-eigenvector of an $n \times n$ Cauchy matrix, where the associated con-eigenvalue $\lambda_m > 0$ gives the approximation error in the L^∞ norm. Unfortunately, standard algorithms do not accurately compute small con-eigenvalues (and the associated con-eigenvectors) and, in particular, yield few or no correct digits for con-eigenvalues smaller than the machine roundoff.

We develop a fast and accurate algorithm for computing con-eigenvalues and con-eigenvectors of positive-definite Cauchy matrices, yielding even the tiniest con-eigenvalues with high relative accuracy. The algorithm computes the mth con-eigenvalue in $\mathcal{O}\left(m^2 n\right)$ operations and, since the con-eigenvalues of positive-definite Cauchy matrices decay exponentially fast, we obtain (near) optimal rational approximations in $\mathcal{O}\left(n\left(\log \delta^{-1}\right)^2\right)$ operations, where δ is the approximation error in the L^∞ norm. We derive error bounds demonstrating high relative accuracy of the computed con-eigenvalues and the high accuracy of the unit con-eigenvectors. Finally, numerical tests on random (complex-valued) Cauchy matrices show that the algorithm computes all the con-eigenvalues and con-eigenvectors with nearly full precision.

1. Introduction

We present an algorithm for computing with high relative accuracy the con-eigenvalue decomposition of positive-definite Cauchy matrices,

$$Cu_m = \lambda_m \overline{u_m}, \quad C_{ij} = \frac{\alpha_i \overline{\alpha_j}}{1 - \gamma_i \overline{\gamma_j}}, \; i,j = 1, \ldots, n, \tag{1.1}$$

where γ_i and α_i are complex numbers and $|\gamma_i| < 1$. Although the con-eigenvalue decomposition (see e.g. [21]) is less well-known than the eigenvalue decomposition or the singular value decomposition, it arises naturally in constructing optimal approximations using exponentials or rational functions [1, 2, 3, 10, 28, 5, 6]. For example, for a rational function $f(z)$,

$$f(z) = \sum_{i=1}^{n} \frac{\alpha_i}{z - \gamma_i} + \sum_{i=1}^{n} \frac{\overline{\alpha_i} z}{1 - \overline{\gamma_i} z} + \alpha_0,$$

we may construct a rational approximation $g(z)$ with m poles and with an error,

$$\max_{x \in [0,1]} \left| f\left(e^{2\pi i x}\right) - g\left(e^{2\pi i x}\right) \right| \approx \lambda_m,$$

by solving the con-eigenvalue problem (1.1) (see Section 2.1 for more detail). Ordering the con-eigenvalues, $\lambda_1 \geq \ldots \geq \lambda_n > 0$ (they may all be chosen positive), the number of poles m of the approximant $g(z)$ corresponds to the index of the con-eigenvalue λ_m and leads to a near optimal approximation in the L^∞-norm with the error close to λ_m.

Unfortunately, current algorithms compute an approximate con-eigenvalue $\widehat{\lambda_m}$ with an error no better than $\left|\lambda_m - \widehat{\lambda_m}\right| / |\lambda_1| = \mathcal{O}(\epsilon)$, and an approximate unit con-eigenvector $\widehat{\mathbf{u}_m}$ with an error no

This research was partially supported by NSF grant DMS-100995 and DOE/ORNL grant 4000038129.

better than

$$\|u_m - \widehat{u_m}\|_2 = \mathcal{O}(\epsilon)/\text{absgap}_m, \quad \text{absgap}_m \equiv \min_{p \neq m} |\lambda_m - \lambda_p| / |\lambda_1|,$$

where ϵ denotes the machine roundoff. This implies that a computed con-eigenvalue smaller than $|\lambda_1|\,\epsilon$ will generally have few or no correct digits. Hence, in order to obtain a rational approximation with accuracy $\lambda_m \lesssim 10^{-7}$, we are forced to use at least quadruple precision. Since quadruple precision is typically not supported by the hardware, it slows down the computation by an unpleasant factor (between 30 and 100). Another undesirable feature of current algorithms to solve (1.1) is the $\mathcal{O}\left(n^3\right)$ complexity for finding the $m \ll n$ poles of $g(z)$, where n is the original number of poles of $f(z)$.

Although the construction of optimal rational approximations in the L^∞-norm has a long history (starting with the seminal papers [1, 2, 3]), the difficulties mentioned above limit practical applications of such approximations to situations where the problem size is relatively small and a low accuracy is acceptable. In this regard, we view our results as a stepping stone toward a wider use of optimal L^∞-approximations in numerical analysis.

We develop a fast and accurate algorithm for con-eigenvalue/con-eigenvector computations of positive-definite Cauchy matrices that addresses both of the difficulties mentioned above. Our algorithm computes the mth con-eigenvalue/con-eigenvector in $\mathcal{O}\left(m^2 n\right)$ operations (see Section 5). Since the con-eigenvalues of positive definite Cauchy matrices decay exponentially fast, for a given desired accuracy $\|f\left(e^{2\pi i x}\right) - g\left(e^{2\pi i x}\right)\|_\infty \approx \delta$, the number of poles m in the approximant $g(z)$ is $\mathcal{O}\left(\log \delta^{-1}\right)$. Therefore, the complexity of our algorithm is $\mathcal{O}\left(n\left(\log \delta^{-1}\right)^2\right)$, i.e., it is essentially linear in the number of original poles n and, thus, is mostly controlled by the number of poles of the final optimal approximation.

We also prove that the con-eigenvalue algorithm achieves high relative accuracy, i.e., the computed con-eigenvalue $\widehat{\lambda_m}$ satisfies $\left|\lambda_m - \widehat{\lambda_m}\right| / |\lambda_m| = \mathcal{O}(\epsilon)$, and the computed unit con-eigenvector $\widehat{\mathbf{u}_m}$ satisfies

$$\|u_m - \widehat{u_m}\|_2 = \mathcal{O}(\epsilon)/\text{relgap}_m, \quad \text{relgap}_m \equiv \min_{l \neq m} |\lambda_m - \lambda_l| / (\lambda_l + \lambda_m),$$

(see Theorems 6 and 7 for the exact statement). In contrast to the usual perturbation theory for general matrices, we show that small perturbations of the poles γ_m and residues α_m (determining the Cauchy matrix $C = C(\alpha, \gamma)$ in (1.1)) lead to correspondingly small perturbations in the con-eigenvalues and con-eigenvectors, as long as the poles are well separated in a relative sense and are not too close to the unit circle. Thus, constructing optimal rational approximations using our con-eigenvalue algorithm is a fast and robust procedure.

Our approach is inspired by papers [15, 13, 17, 11, 19], which develop algorithms and theory for highly accurate SVDs of certain structured matrices. Generally speaking, high relative accuracy is achieved when it is possible to avoid catastrophic cancellation resulting from subtracting two close floating point numbers (see [12] for a comprehensive analysis of when efficient and accurate algorithms are possible using floating point arithmetic). Classes of matrices for which highly accurate SVD or eigenvalue algorithms exist include bi-diagonal matrices [14, 9, 18], acyclic matrices [16], graded positive-definite matrices [15], scaled diagonally dominant matrices [4], totally positive matrices [22], symmetric indefinite matrices [27], and Cauchy matrices (as well as, more generally, matrices with displacement rank one) [11].

The con-eigenvalue algorithm considered here is based on computing the eigenvalue decomposition of the product, $\overline{C}C$, of positive-definite Cauchy matrices $\overline{C}$ and C, and is similar to the algorithm in [13] for the generalized eigenvalue decomposition, as well as the algorithm in [17] for the product SVD decomposition. We also rely on the algorithm in [11] for computing, with high relative accuracy, the Cholesky decomposition (with complete pivoting) $C = (PL)\,D^2\,(PL)^*$ of a positive-definite Cauchy matrix C. However, since we are interested in computing only con-eigenvalues of some approximate size δ, we stop Demmel's Cholesky algorithm once the diagonal elements D_{ii} are small with respect to δ and the desired precision. Since the diagonal elements D_{ii} decay exponentially fast, this allows us to accurately compute con-eigenvalues of size δ (and the associated con-eigenvectors)

in $\mathcal{O}\left(n\left(\log\delta^{-1}\right)^2\right)$ operations. We also note that the error bounds developed in [17] are not applicable to our problem (the condition number of a Cauchy matrix cannot be appreciably reduced by scaling the rows and columns). In contrast, the error bounds developed in this paper yield high relative accuracy for all the computed con-eigenvalues larger than δ (and high accuracy for the con-eigenvectors), as long as the unit triangular matrix, L, is well-conditioned, and the relative gap between the con-eigenvalues is not too small (we have always observed this to hold in practice). In particular, if δ is chosen small enough, the full con-eigenvalue decomposition is obtained with high relative accuracy. The derivation of our error bounds makes crucial use of the component-wise perturbation theory developed in [15] for the singular vectors of graded matrices (see also [24]), as well as the component-wise error analysis in [15] and [23] for the one-sided Jacobi method. We also use the error analysis given in [19] for the Householder QR method with complete pivoting.

It has been an established practice, in both numerical analysis and signal processing, to use L^2-type methods for representing functions. On the other hand, it has been understood for some time that nonlinear approximations may be far superior in achieving high accuracy with a minimal number of terms (see e.g., [25]). However, in spite of many interesting theoretical results, the widespread use of nonlinear approximations has been limited by a lack of efficient and accurate algorithms for computing them. Our algorithms provide the necessary tools for computing optimal nonlinear approximations via rational functions, and come with guaranteed accuracy bounds. We believe that these new accurate algorithms may greatly extend the practical use of L^∞ approximations in numerical analysis and signal processing. We note that we have already developed several applications of these algorithms in numerical problems (to be published elsewhere).

We start in Section 2.1 by describing in some detail the reduction problem for rational functions, and connect its solution to a con-eigenvalue problem for positive definite Cauchy matrices. In Section 2.2, we review the results in [15, 13, 17, 11, 19] needed in our derivations (those familiar with highly accurate SVD/eigenvalue algorithms may refer to this section only as needed). Next, in Section 3.1, we formulate the con-eigenvalue problem in terms of an associated eigenvalue problem, and provide an informal description of the algorithms of this paper (we defer proofs of accuracy of these algorithms to Section 5). We then verify the accuracy of the con-eigenvalue algorithm by comparing the con-eigenvalue decomposition of randomly generated Cauchy matrices with that obtained via standard algorithms in extended precision. Finally, in Section 5, we prove that the con-eigenvalue algorithm achieves high relative accuracy and that the con-eigenvalue decomposition is stable with respect to small perturbations of the parameters defining the Cauchy matrix.

2. Preliminaries

We now provide the necessary background for the con-eigenvalue algorithm. In Section 2.1, we explain how the accurate computation of small con-eigenvalues and associated con-eigenvectors allows us to construct optimal rational approximations. Section 2.1 provides important motivation for our algorithm, but is not needed to understand the rest of the paper. In Sections 2.2 - 2.4, we provide necessary background on computing highly accurate SVDs, as well some error bounds that are needed for the analysis of the con-eigenvalue algorithm. Although the results we need in [15, 23, 13, 24, 11, 19] are only stated there for real-valued matrices, they carry over to complex-valued matrices with minor modifications and are formulated as such.

2.1. **Reduction procedure.**

We start with a rational function $f(z)$,

$$f(z) = \sum_{i=1}^{n} \frac{\alpha_i}{z-\gamma_i} + \sum_{i=1}^{n} \frac{\overline{\alpha_i} z}{1-\overline{\gamma_i} z} + \alpha_0, \tag{2.1}$$

where the residues α_j and poles γ_j are complex, and $|\gamma_j| < 1$. We consider an algorithm to find a rational approximation $r(e^{2\pi i x})$ to $f(e^{2\pi i x})$ with a specified number of poles and with a (nearly) optimally small error in the L^∞-norm. The algorithm is based on a theorem of Adamyan, Arov,

and Krein (referred to below as the AAK Theorem) [3]. We note that the formulation given below in terms of a con-eigenvalue problem is similar to the approach taken in [10] and [5].

Given a target accuracy δ for the error in the L^∞-norm, the steps for computing the rational approximant $r(z)$,

$$r(z) = \sum_{i=1}^{m} \frac{\beta_i}{z - \eta_i} + \sum_{i=1}^{m} \frac{\overline{\beta_i} z}{1 - \overline{\eta_i} z} + \alpha_0,$$

are as follows:

(1) Compute a con-eigenvalue $0 < \lambda_m \leq \delta$ and corresponding con-eigenvector u of the Cauchy matrix $C_{ij} = C_{ij}(\gamma_i, \alpha_j)$,

$$Cu = \lambda_m \overline{u}, \text{ where } u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad C_{ij} = \frac{a_i b_j}{x_i + y_j}, \quad i, j = 1, \ldots, n, \tag{2.2}$$

and $a_i = \sqrt{\alpha_i}/\gamma_i$, $b_j = \sqrt{\overline{\alpha}_j}$, $x_i = \gamma_i^{-1}$, $y_j = -\overline{\gamma_j}$. The con-eigenvalues of C are labeled in non-increasing order, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

(2) Find the (exactly) m zeros η_j in the unit disk of the function

$$v(z) = \frac{1}{\lambda_m} \sum_{i=1}^{n} \frac{\sqrt{\overline{\alpha_i}}\, \overline{u_i}}{1 - \overline{\gamma_i} z}. \tag{2.3}$$

The fact that there are exactly m zeros in the unit disk, corresponding to the index m of the con-eigenvalue λ_m, is a consequence of the AAK theorem. The poles of $r(z)$ are given by the zeros η_j of $v(z)$.

(3) Find the residues β_m of $r(z)$ by solving the $m \times m$ linear system

$$\sum_{i=1}^{m} \frac{1}{1 - \eta_i \overline{\eta_j}} \beta_i = \sum_{i=1}^{n} \frac{\alpha_i}{1 - \gamma_i \overline{\eta_j}}. \tag{2.4}$$

The L^∞-error of the resulting rational approximation $r(e^{2\pi i x})$ satisfies $\|f - r\|_\infty \approx \lambda_m$, and is close to the best error in the L^∞-norm achievable by rational functions with no more than m poles in the unit disk. Hence, we are led to the problem of computing, to high relative accuracy, small con-eigenvalues and the associated con-eigenvectors of positive-definite Cauchy matrices.

Remark 1. In practice, finding the new poles η_i using the formula for $v(z)$ in (2.3) is ill-advised, since evaluating $v(z)$ in this form could result in loss of significant digits through catastrophic cancellation. It turns out that the con-eigenvector components satisfy $u_i = \sqrt{\alpha_i} v(\gamma_i)$, $i = 1, \ldots, n$, which, along with the n poles $1/\overline{\gamma_i}$ of $v(z)$, completely determines (2.3). Since the poles γ_i of $f(z)$ are often close to the poles η_i of $r(z)$, we have observed that evaluating $v(z)$ by using rational interpolation via continued fractions with the known values $v(\gamma_i)$ allows us to obtain the new poles η_i with nearly full precision.

2.2. **Accurate SVDs of matrices with rank-revealing decompositions.** According to the usual perturbation theory for the SVD (see e.g. [8]), perturbations δA of a matrix A change the ith singular value σ_i by $\delta\sigma_i$ and corresponding unit eigenvector u_i by δu_i, where (assuming for simplicity that σ_i is simple),

$$|\delta\sigma_i| / \sigma_1 \leq \|\delta A\|, \quad \|\delta u_i\| \leq \frac{\|\delta A\|}{\text{absgap}_i}, \quad \text{absgap}_i = \min_{i \neq j} |\sigma_i - \sigma_j| / \sigma_1. \tag{2.5}$$

Therefore, small perturbations in the elements of A may lead to large relative changes in the small singular values and the associated singular vectors. Moreover, since standard algorithms compute an SVD of some nearby matrix $A + \delta A$, where $\|\delta A\| / \|A\| = \mathcal{O}(\epsilon)$, the perturbation bound (2.5) shows that the computed small singular values and corresponding singular vectors will be inaccurate.

In contrast, the authors in [13] show that, for many structured matrices, the ith singular value $\sigma_i \ll \sigma_1$ and the associated singular vector are robust with respect to small perturbations of the matrix that preserve its underlying structure. The sensitivity is instead governed by the ith *relative* gap

$$\text{relgap}_i = \min_{i \neq j} \frac{|\sigma_i - \sigma_j|}{\sigma_i + \sigma_j}.$$

More precisely, let us consider the class of matrices for which a rank-revealing decomposition $A = XDY^*$ is available and may be computed accurately. Here X and Y are $n \times m$ well-conditioned matrices and D is an $m \times m$ diagonal matrix that contains any possible ill-conditioning of A. As is shown in [13], a perturbation of $A = XDY^*$ that is of the form $A + \delta A = (X + \delta X)(D + \delta D)(Y + \delta Y)^*$, where

$$\frac{\|\delta X\|}{\|X\|} = \mathcal{O}(\epsilon), \quad \frac{\|\delta Y\|}{\|Y\|} = \mathcal{O}(\epsilon), \quad \frac{|\delta D_{ii}|}{|D_{ii}|} = \mathcal{O}(\epsilon), \tag{2.6}$$

changes the ith singular value σ_i and associated left (or right) singular vector u_i by amounts $\delta\sigma_i$ and δu_i bounded by

$$\frac{|\delta\sigma_i|}{\sigma_i} \leq \max(\kappa(X), \kappa(Y))\,\mathcal{O}(\epsilon), \quad \|\delta u_i\| \leq \frac{\max(\kappa(X), \kappa(Y))}{\text{relgap}_i}\mathcal{O}(\epsilon), \tag{2.7}$$

where $\kappa(X) = \|X\| \left\|X^\dagger\right\|$ and $X^\dagger$ denotes the pseudo-inverse of A. One reason this class of matrices is so useful is that Gaussian elimination with complete pivoting (GECP) (or simple modifications) computes accurate rank-revealing decompositions of many types of structured matrices (see [13] and [11]). Moreover, small perturbations of such matrices that preserve their underlying structure lead to small perturbations in the rank-revealing factors and, therefore, small relative perturbations of the singular values.

Given the decomposition $A = XDY^*$, it is shown in [13] (see Algorithm 3.1) that an SVD of A may be computed with high relative accuracy, and with about the same cost as standard, less accurate SVD algorithms for dense matrices. The key to this algorithm is the one-sided Jacobi algorithm (briefly reviewed in Section 2.4), which, with an appropriate stopping criterion, accurately computes the SVD of matrices of the form DB, where D is diagonal (and typically highly ill-conditioned) and B is well-conditioned (see [15] and [23]). In the application we are considering, $X = Y$ and D has positive decreasing diagonal elements, and the following simplified version of this algorithm (see [23]) suffices.

Algorithm 1 SVD_RRD(X, D) computes an accurate SVD of $A = XDX^*$. Input: RRD factors X $(n \times m)$ and D $(m \times m)$. Output: SVD factors U and Σ $(XDX^* = U\Sigma^2U^*)$.

$(U, \Sigma, V) \leftarrow$ SVD_RRD(X, D)

`1. Compute QR factors` $(Q, R) \leftarrow$ `Householder_QR`(XD)
`using Householder reflections (may use optional column pivoting)`
`2. Compute SVD factors` $(U, \Sigma, V) \leftarrow$ `Jacobi_SVD`(R) `(`$RD = U\Sigma V^*$`) using the`
`one-sided Jacobi algorithm in [15]`
`4. Output` QU `and` Σ

Algorithm 1 yields computed singular values $\widehat{\sigma_i}$ and left (or right) singular vectors $\widehat{u_i}$ that satisfy

$$\frac{|\sigma_i - \widehat{\sigma_i}|}{\sigma_i} \leq \kappa(X)\,\mathcal{O}(\epsilon), \tag{2.8}$$

$$\|w_i - \widehat{w_i}\| \leq \frac{\kappa(X)}{\text{relgap}_i}\mathcal{O}(\epsilon), \tag{2.9}$$

2.3. **Computing rank-revealing decompositions.** In this section we review how a modification of GECP computes accurate rank-revealing decompositions of Cauchy matrices [11]. We also review how a variant of the QR Householder algorithm with complete pivoting computes accurate rank-revealing decompositions of graded matrices [19].

2.3.1. *LDU factorization of Cauchy matrices.* We describe Demmel's algorithm (see Algorithms 3 and 4 in [11] and Algorithm 2.5 in [7]) for computing an accurate rank-revealing decomposition of a $n \times n$ positive-definite Cauchy matrix $C_{ij} = a_i b_j / (x_i + y_j)$. The algorithm is based on a modification of Gaussian elimination for computing, in $\mathcal{O}\left(n^2\right)$ operations, the Cholesky factorization $C = (PL)\, D\, (PD)^*$ of a positive-definite Cauchy matrix (more generally, the algorithm computes an LDU factorization for an arbitrary Cauchy matrix in $\mathcal{O}\left(n^3\right)$ operations). Here P is a permutation matrix, L is a unit lower triangular matrix, and D is a diagonal matrix with positive diagonal elements. It is shown in [11] that, remarkably, the components of the LDU factors $\widehat{L}$, $\widehat{U}$, and $\widehat{D}$ are computed to high relative accuracy,

$$\left|\widehat{L}_{ij} - L_{ij}\right| \leq |L_{ij}|\, c_n \epsilon, \quad \left|\widehat{U}_{ij} - U_{ij}\right| \leq c_n\, |U_{ij}|\, \epsilon, \quad \left|\widehat{D}_{ii} - D_{ii}\right| \leq c_n\, |D_{ii}|\, \epsilon, \tag{2.10}$$

where c_n is a modest-sized function of n. The basic reason the algorithm achieves high relative accuracy is that the only operations involved are multiplication and division of floating point numbers (additions and subtractions in the algorithm involve only x_i and y_j, which are assumed to be exact).

We now review the basic idea behind the algorithm in [11]. First, ignoring pivoting for a moment, we assume that, after k steps of Gaussian elimination, the Cauchy matrix is transformed to the matrix $G^{(k)}$,

$$G^{(k)} = \begin{pmatrix} G_{11}^{(k)} & G_{12}^{(k)} \\ 0 & G_{22}^{(k)} \end{pmatrix}.$$

The elements of the Schur complement $G_{22}^{(k+1)}$ may be computed from those of $G_{22}^{(k)}$ by using the recursion

$$G_{ij}^{(k)} = \left(\frac{x_i - x_k}{x_i + y_k}\right)\left(\frac{y_j - y_k}{y_j + x_k}\right) G_{ij}^{(k-1)}, \quad i, j = k+1, \ldots, n. \tag{2.11}$$

Introducing pivoting, we observe that the matrix $G^{(k)}$ may be obtained by applying Gaussian elimination to a Cauchy matrix $C^{(k)} = C^{(k)}\left(a^{(k)}, b^{(k)}, x^{(k)}, y^{(k)}\right)$, where $a^{(k)}$, $b^{(k)}$, $x^{(k)}$ and $y^{(k)}$ are permutations of a, b, x and y corresponding to the row and column pivoting of C. As long as the vectors a, b, x and y are permuted according to the pivoting of $G^{(k)}$, the recursive formula (2.11) still holds.

It is observed in [11] that if C is positive-definite (and, therefore, only diagonal pivoting is needed), then the pivot order may be determined in advance in $\mathcal{O}\left(n^2\right)$ operations by computing $\operatorname{diag}\left(G^{(k)}\right)$ from formula (2.11). Once the correct pivot order is known, we do not need to compute the entire Schur complement $G^{(k)}$ to extract the components of L and U, but only its kth row and kth column. Indeed, we may use Algorithm 2.5 in [7], which uses the displacement structure of C, to compute an accurate Cholesky decomposition in $\mathcal{O}\left(n^2\right)$ operations. To see how, note that it easily follows from (2.11) that the Schur complement of a Cauchy matrix is a Cauchy matrix,

$$G^{(k)}(i,j) = \frac{\alpha_i^{(k)} \beta_j^{(k)}}{x_i + y_j}, \quad i, j = k+1, \ldots, n,$$

where the parameters $\alpha_i^{(k)}$ and $\beta_i^{(k)}$ satisfy the recursion

$$\alpha_i^{(k)} = \frac{x_i - x_k}{x_i + y_k} \alpha_i^{(k-1)}, \quad \beta_i^{(k)} = \frac{y_i - y_k}{y_i + x_k} \beta_i^{(k-1)}, \quad i = k+1, \ldots, n.$$

Since the kth column $L(:,k)$ may be extracted from $G^{(k)}(:,k)$, we therefore only require $\mathcal{O}(n)$ operations at each step of Gaussian elimination to compute $L(:,k)$. Updating $\alpha_i^{(k)}$ and $\beta_i^{(k)}$ also requires

only $\mathcal{O}(n)$ operations. In Section 3.2 (see Algorithms 3 and 4), we present an $\mathcal{O}\left(n\left(\log\delta^{-1}\right)^2\right)$ algorithm to compute con-eigenvalues greater than a user specified cutoff δ and, as a result, yielding a fast algorithm for obtaining nearly optimal rational approximations.

Once an accurate LDU factorization $C \approx \left(P\widehat{L}\right)\widehat{D}\left(P\widehat{D}\right)^*$ is available, an accurate SVD of C may be obtained using Algorithm 1.

2.3.2. *Rank-revealing decompositions of graded matrices.* It is shown in [19] that the Householder QR algorithm with complete pivoting may be used to compute a rank-revealing decomposition of a graded matrix of the form $A = D_1BD_2$. Here D_1 and D_2 are diagonal matrices that account for the ill-conditioning of A. Recall that the Householder QR algorithm uses repeated applications of orthogonal matrices to reduce A to an upper-triangular matrix R. On the first step, the parameter β_1 and the vector v_1 of the Householder reflection matrix $Q^{(1)} = I - \beta_1 v_1 v_1^*$ are chosen so that

$$Q^{(1)}\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = \begin{pmatrix} a_{11}^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Consequently, the first application of $Q^{(1)}$ to A results in a matrix of the form

$$A^{(1)} = Q^{(1)}A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}.$$

This process is repeated on the $(n-1)\times(n-1)$ lower block $\left[a_{ij}^{(1)}\right]_{2\le i,j\le n}$ and, after $n-1$ such steps, $A^{(n-1)} = Q^{(n-1)}\dots Q^{(1)}A = R$, where R is upper triangular. In the version considered in [19], the rows of A are first pre-sorted so that so that $\|A(1,:)\|_\infty \ge \dots \ge \|A(n,:)\|_\infty$. The algorithm then proceeds as above, except that at each step, k, column pivoting is performed to ensure that $\left\|A^{(k)}(k:n,k)\right\|_2 \ge \dots \ge \left\|A^{(k)}(k:n,n)\right\|_2$. Letting P_1 denote the row permutation matrix that pre-sorts the rows of A, and letting P_2 denote the column permutation matrix corresponding to the column pivoting, the QR Householder algorithm produces the QR factorization $P_1AP_2 = QR$.

Following [19], we consider the error analysis of the Householder algorithm (without pivoting) applied to P_1AP_2, where P_1 and P_2 are chosen so that no column or row exchanges are necessary (e.g. the matrix A is pre-pivoted). Assume that the matrix P_1AP_2 may be factored as $P_1AP_2 = D_1BD_2$, where D_1 and D_2 are diagonal matrices, and that the Householder algorithm, applied to the row-scaled matrix $C = D_1B$, produces intermediate matrices $C^{(k)}$ with columns $c_j^{(k)}$. Finally, define the quantities ρ, μ, and ψ by

$$\rho = \max_i \frac{\max_{j,k}\left|c_{ij}^{(k)}\right|}{\max_j|c_{ij}|}, \quad \mu = \max_k \max_{j\ge k} \frac{\left\|c_j^{(k)}(k:m)\right\|}{\left\|c_k^{(k)}(k:m)\right\|}, \quad \psi = \max_{\substack{1\le i\le n\\ i\le k\le n}} \frac{\max_j|c_{kj}|}{\max_j|c_{ij}|}. \tag{2.12}$$

The above quantities measure the extent to which the Householder algorithm preserves the scaling in the intermediate matrices $A^{(k)}$, and are almost always small (this is analogous to the pivot growth factor in Gaussian elimination with row pivoting). It is shown in [19] that

Theorem 2. *Suppose that A is pre-pivoted, and the Householder algorithm is used to compute the upper triangular matrix $\widehat{R}$ of the QR decomposition. Then there is an orthogonal matrix Q such that $Q\widehat{R} = D_1(B+\delta B)D_2$, where δB satisfies*

$$\|\delta B\| \le \rho\psi\mu\|B\|\mathcal{O}(\epsilon),$$

and ρ, μ, and ψ are defined in (2.12).

In [19] Theorem 2 is combined with the theory developed in [13] (e.g., see Theorems 4.1 and 4.2 in [13]) to show that the QR algorithm with complete pivoting produces accurate rank revealing decompositions of graded matrices of the form $A = D_1 B D_2$, as long as the principal minors of B are well-conditioned and the diagonal elements of D_1 and D_2 are approximately decreasing in magnitude. In our specific application, we are able to obtain stronger error bounds (see Proposition 10).

Remark. Instead of pre-sorting the rows of A and applying the Householder algorithm with column pivoting, one may also use a version of the Householder algorithm in which both row and column pivoting is employed (see [19] for more details). Gaussian elimination with complete pivoting may also be used to obtain accurate rank-revealing decompositions of graded matrices [13].

2.4. **Modified one-sided Jacobi algorithm .** The heart of Algorithm 1 is the modified one-sided Jacobi algorithm, which accurately computes the SVD of matrices of the form DB and BD, where D is diagonal and typically highly graded, and B is well-conditioned (see [15] and [23]). Although we focus on the one-sided Jacobi algorithm as applied to $G = BD$, analogous considerations apply to $G = DB$ by replacing G by G^*. The one-sided Jacobi algorithm works by applying a sequence of Jacobi matrices $J_1, \ldots, J_M$ to G from the right (i.e., the same side as the scaling, which ensures that components of the right singular vectors are computed with high relative accuracy). Each Jacobi matrix J is chosen to orthogonalize two selected columns, and one sweep consists of orthogonalizing columns in the order $(1,1),(1,2),\ldots,(1,n)$, followed by columns $(2,3),(2,4),\ldots,(2,n)$, and so on. Sweeps are repeated until all the columns are orthogonal to each other to within the bound

$$G\left(J_1 \cdots J_M\right) = W, \;\; \frac{|w_i^* w_j|}{|w_i^* w_i|^{1/2} \, |w_i^* w_i|^{1/2}} \leq n\epsilon, \;\; \text{if } i \neq j.$$

This stopping criterion is used to ensure that even the smallest singular values are computed with high relative accuracy. The SVD of $G = U\Sigma V^*$ immediately follows by taking $\Sigma_{ii} = W(:,i)$, $V = W/\Sigma$, and $U = \left(J_1 J_2 \cdots J_M\right)^*$.

It will be crucial for the error bounds developed in this paper that the components of the right singular vectors of DB (or the left singular vectors of BD) scale in a way similar to D, and are computed accurately relative to this scaling. At each step m of the Jacobi algorithm, we write $\left(J_0 \cdots J_m\right) G = B_m D_m$, where the columns of B_m have unit l^2-norm and the matrix D_m is diagonal. Defining $\nu = \max_{1 \leq m \leq M} \kappa_2 \left(B_m\right)^2$, we then have the following result from [23] (an improvement on results in [15]).

Theorem 3. *Let $G = DB$ be a $n \times n$ full-rank, complex-valued matrix, where the diagonal matrix D is chosen so that the l^2-norm of each column of B is unity. Suppose that one-sided Jacobi algorithm is used to compute an approximation $\widehat{v_i}$ to the ith left singular vector v_i of G (normalized so that $v_i(i) = 1$), and the iteration converges after M sweeps. Then the following error bound holds on the computed components of v_i:*

$$|v_i(j) - \widehat{v_i}(j)| \leq \min\left(\left|\frac{D_{ii}}{D_{jj}}\right|, \left|\frac{D_{jj}}{D_{ii}}\right|\right)\left(\frac{\rho(M,n)\,\nu}{relgap_i}\epsilon + \mathcal{O}\left(\epsilon^2\right)\right), \tag{2.13}$$

where

$$\text{relgap}_i = \frac{|\sigma_i - \sigma_j|}{\sigma_i + \sigma_j},$$

and $\rho(M,n)$ *is proportional to* $M \cdot n^{3/2}$.

3. Accurate con-eigenvalue decomposition (an informal derivation)

3.1. **Accurate con-eigenvalue decompositions of positive-definite matrices with RRDs.** Since in the con-eigenvalues are determined only up to an arbitrary phase factor $e^{i\phi}$, we may assume

they are real without loss of generality. Moreover, the con-eigenvalue problem for a positive-definite Cauchy matrix $C_{ij} = a_i b_j / (x_i + y_j)$ reduces to an eigenvalue problem,

$$\overline{C} C u = \lambda \overline{C} \bar{u} = |\lambda|^2 u. \tag{3.1}$$

Remark 4. For applications to computing optional rational approximations (see Section 2.1), we need to compute the con-eigenvalues/con-eigenvectors of Cauchy matrices of the slightly different form, $C_{ij} = \alpha_i \overline{\alpha_j} / (1 - \gamma_i \overline{\gamma_j})$, i.e., with $a_i = \sqrt{\alpha_i}/\gamma_i$, $b_j = \sqrt{\overline{\alpha_j}}$, $x_i = \gamma_i^{-1}$, and $y_j = -\overline{\gamma_j}$. The same reasoning as in [11] shows that the Cholesky computation of C (see Section 2.3.1) is performed with high relative accuracy, as long as the differences $\gamma_j^{-1} - \overline{\gamma_i}$ are computed with high relative accurately. As noted in [11], $\gamma_j^{-1} - \overline{\gamma_i}$ may be accurately computed using techniques from [26]. Alternatively, if needed, we may compute some of these differences using extended precision without impacting the overall speed of the algorithm.

We now discuss how to compute accurate eigenvectors and eigenvalues of matrices $\overline{A}A$, where A is of the form $A = XD^2X^*$, with X a (well-conditioned) $n \times m$ matrix ($m \leq n$) and D an $m \times m$ diagonal matrix with positive diagonal entries. To do so, let us define the $m \times m$ matrix $G = D\left(X^{\mathrm{T}}X\right)D$, and consider its SVD, $G = W\Lambda V^*$. Then $G^*G = V\Lambda^2 V^*$, and the ith right singular vector ($1 \leq i \leq m$), $v_i = V(:,i)$, satisfies $\left(DX^*\overline{X}D\right)\left(DX^{\mathrm{T}}XD\right)v_i = \lambda_i^2 v_i$. It then follows that $z_i = XDv_i$ is an eigenvector of $A\overline{A}$ with eigenvalue λ_i^2, since

$$\begin{aligned} A\overline{A}z_i &= \left(XD^2X^*\right)\left(\overline{X}D^2X^{\mathrm{T}}\right)z_i = \\ &= XD\left(DX^*\overline{X}D\right)\left(DX^{\mathrm{T}}XD\right)v_i = \lambda_i^2 XDv_i = \lambda_i^2 z_i. \end{aligned}$$

and, thus, $\overline{z_i} = \overline{X}D\overline{v_i}$ is an eigenvector of $\overline{A}A$. To summarize: given the decomposition $A = XD^2X^*$, an eigenvector z_i ($i \leq m$) of $\overline{A}A$ is given by $\overline{z_i} = \overline{X}D\overline{v_i}/D_{ii}$, where v_i is the ith right singular vector of the $m \times m$ matrix $G = D\left(X^{\mathrm{T}}X\right)D$.

To compute eigenvectors and eigenvalues of $\overline{A}A$, we first use the Householder QR algorithm from Section 2.3.2 to obtain an accurate rank-revealing decomposition of $G = QR$. Algorithm 1 then computes an accurate SVD $G = W\Lambda^2 V^*$ (see Section 2.2), where the diagonal elements of Λ contain m con-eigenvalues of A, and the columns of the matrix $T = \overline{X}\left(D\overline{V}D^{-1}\right)$ contain m con-eigenvectors. The main steps are shown in Algorithm 2.

Algorithm 2 ConEig_RRD(X, D) computes accurate con-eigenvalue decomposition of XDX^*. Input: rank-revealing factors X and D (of dimensions $n \times m$ and $m \times m$). Output: m con-eigenvalues/con-eigenvectors of XDX^*, contained in Λ and T.

$(\Lambda, T) \leftarrow$ ConEig_RRD(X, D)

1. Form $G = D\left(X^{\mathrm{T}}X\right)D$
2. Compute QR factors $(Q, R, P_1, P_2) \leftarrow$ Householder_QR of G ($G = P_1 Q R P_2$), using complete pivoting (see Section 2.3.2)
3. Compute the SVD factors $(W, \Lambda, V) \leftarrow$ Jacobi(RP_2) of RP_2 ($RP_2 = W\Lambda^2 V^*$), using one-sided Jacobi, applied from the right (see Section 2.4)
4. Form matrix of con-eigenvectors $T = \overline{X}D\overline{V}/D^{-1}$, and output con-eigenvalues Λ and con-eigenvectors T

In our application, the elements of D decay exponentially fast and it would appear that computing the con-eigenvectors $\overline{z_i} = \overline{X}D\overline{v_i}/D_{ii}$ might lead to wildly inaccurate results even if v_i is computed accurately. The basic reason Algorithm 2 achieves high accuracy is that the matrix containing the right singular vectors scales like $|V_{ij}| \leq c_V \min\left(D_{ii}/D_{jj}, D_{jj}/D_{ii}\right)$, and the computed components $\widehat{V}_{ij}$ are also accurate relative to the scaling in D, i.e.,

$$\left|V_{ij} - \widehat{V}_{ij}\right| \leq \min\left(D_{ii}/D_{jj}, D_{jj}/D_{ii}\right)\mathcal{O}(\epsilon).$$

3.2. **Accurate con-eigenvalue decompositions of positive-definite Cauchy matrices.** If $A = C$ is a positive-definite Cauchy matrix, then we know from Section 2.2 that the modified GECP algorithm in [11] computes the Cholesky decomposition $C = (PL)\,D^2\,(PL)^*$ with high relative accuracy. Therefore, Algorithm 2 for the eigenvalue problem of $\overline{C}C$ may be used, with $X = PL$, to compute all the eigenvalues and eigenvectors (and, therefore, the con-eigenvectors and con-eigenvalues of C). Alternatively, from the decomposition $C = (PL)\,D^2\,(PL)^*$, we may first use the one-sided Jacobi algorithm of Section 2.4 to compute an SVD of $PLD = U\Sigma\,(U')*$ (where U and U' are unitary matrices), yielding $C = U\Sigma^2U^*$. Then, with $X = U$ and $D = \Sigma$, Algorithm 2 computes the eigenvalues and eigenvectors of $\overline{C}C$. In our analysis of this version of the con-eigenvalue algorithm, the error bounds for the computed con-eigenvectors are better but, in practice, we have not observed a significant difference in accuracy between the two.

For our purposes, we are only interested in computing a single con-eigenvector with associated con-eigenvalue of approximate size δ (see Section 2.1). However, the diagonal elements of D may be many orders of magnitude smaller than δ, and it is then natural to expect that, by computing a partial Cholesky decomposition of C, we may obtain the ith con-eigenvector in much fewer than $\mathcal{O}\left(n^3\right)$ operations. In this case, we stop Demmel's algorithm for the Cholesky decomposition of C once the diagonal elements D_{ii} are small with respect to the product of δ and the machine round-off ϵ, that is, as soon as $D_{mm} \leq \delta\epsilon$ for some m. We then obtain $C \approx \widetilde{C} = \left(\widetilde{P}\widetilde{L}\right)\widetilde{D}^2\left(\widetilde{P}\widetilde{L}\right)^*$, where $\widetilde{L}$ is an $n \times m$ matrix and $\widetilde{D}$ is a diagonal $m \times m$ matrix. Algorithms 3 and 4 contain pseudo-code for computing $\widetilde{L}$, $\widetilde{D}$, and $\widetilde{P}$. In the pseudo-code $I\,(n, m)$ denotes the first $m \leq n$ columns of the $n \times n$ identity matrix.

Algorithm 3 Pivot_Order (a, b, x, y, δ) pre-computes pivot order for Cholesky factorization of $n \times n$ positive-definite Cauchy matrix $C_{ij} = a_ib_j/\,(x_i + y_j)$. Input: a, b, x, and y defining $C_{ij} = a_ib_j/\,(x_i + y_j)$, and target size δ of con-eigenvalue. Output: correctly pivoted vectors a, b, x, and y, truncation size m, and $m \times n$ permutation matrix $\widetilde{P}$

$$\left(a, b, x, y, \widetilde{P}, m\right) \leftarrow \text{Pivot\_Order}\,(a, b, x, y, \delta)$$

Form vector $g_i := a_ib_i/(x_i + y_i)$, $i = 1, \ldots, n$
Set cutoff for GECP termination: $\eta := \epsilon\delta$
Initialize permutation matrix ($n \times n$ identity): $\widetilde{P} = I\,(n, n)$
Compute correctly pivoted vectors:
$m := 1$
  while $|g\,(m)| \geq \eta$ or $m = n - 1$
    Find $m \leq l \leq n$ such that $|g(l)| = \max|g\,(m : n)|$
    Swap elements:
      $g(l) \leftrightarrow g(m)$, $x(l) \leftrightarrow x(m)$, $y(l) \leftrightarrow y(m)$
      $a(l) \leftrightarrow a(m)$,$b(l) \leftrightarrow b(m)$
    Swap rows of permutation matrix:
      $\widetilde{P}(l, :) \leftrightarrow \widetilde{P}(m, :)$
    Update diagonal of Schur complement:
      $g(m + 1 : n) :=$
   $(x\,(m + 1 : n) - x(m))\,/\,(y\,(m + 1 : n) - y(m))\,g(m + 1 : n)$
    Increment iteration count:
      $m := m + 1$
Output $a, b, x, y, \widetilde{P}\,(1 : m, n)\,, m$

Algorithm 4 Cholesky_Cauchy (x, y, a, b, δ) computes partial Cholesky factorization of positive-definite Cauchy matrix $C_{ij} = a_i b_j / (x_i + y_j)$. Input: a, b, x, and y defining $C_{ij} = a_i b_j / (x_i + y_j)$, and target size δ of con-eigenvalue. Output: $n \times m$ matrix $\widetilde{L}$, $m \times m$ matrix $\widetilde{D}$, and permutation $m \times n$ matrix P in partial Cholesky factorization.

$(L, D, P) \leftarrow$ Cholesky_Cauchy (a, b, x, y, δ)

`Compute pivoted vectors and matrix size` m `(Algorithm 3):`

$\left(a, b, x, y, \widetilde{P}, m\right) \leftarrow$ `Pivot_Order` (a, b, x, y, δ)

`Initialize generators:`

$\alpha := a$, $\beta := b$

`Compute first column of Schur complement:`

$G(:, 1) := \alpha * \beta / (x + y)$

`for` $k = 2, m$

`Update generators:`

$\alpha(k:n) := \alpha(k:n) * (x(k:n) - x(k-1)) / (x(k:n) + y(k-1))$

$\beta(k:n) := \beta(k:n) * (y(k:n) - y(k-1)) / (y(k:n) + x(k-1))$

`Extract` k`th column for Cholesky factors:`

$G(k:n, k) := \alpha(k:n) * \beta(k:n) / (x(k:n) + y(k:n))$

`Output partial Cholesky factors:`

$\widetilde{D} =$ `diag` $(G(1:n, 1:m)^{1/2}$, $\widetilde{L} =$ `tril` $(G(1:n, 1:m))\, \widetilde{D}^{-2} + I(n, m)$, $\widetilde{P}$

Once the partial Cholesky decomposition $C \approx \widetilde{C} = \left(\widetilde{P}\widetilde{L}\right) \widetilde{D}^2 \left(\widetilde{P}\widetilde{L}\right)^*$ is computed, Algorithm 2 for the eigenvalue problem of $\overline{\widetilde{C}}\widetilde{C}$ may then be used, with $X = \widetilde{P}\widetilde{L}$ and $D = \widetilde{D}$, to compute accurate con-eigenvalues and con-eigenvectors of $\widetilde{C}$ (see Theorem 7). As before, we may optionally use Algorithm 1 to compute an SVD of $\widetilde{P}\widetilde{L}\widetilde{D} = \widetilde{U}\widetilde{\Sigma}\left(\widetilde{U'}\right)^*$, yielding $\widetilde{C} = \widetilde{U}\left(\widetilde{\Sigma}\right)^2 \widetilde{U}^*$. Since the con-eigenvalues decay exponentially fast, the complexity of this algorithm is $\mathcal{O}\left(n\left(\log(\delta\epsilon)^{-1}\right)^2\right)$ operations. Therefore, when used in the reduction procedure outlined in Section 2.1, the near optimal rational approximation may be obtained by computing the SVD of a matrix that is roughly the size of the optimal number of poles. The pseudo-code is given in Algorithm 5.

Algorithm 5 Con_Eigvector (a, b, x, y, δ) computes accurate con-eigenvalue decomposition of positive-definite Cauchy matrix $C_{ij} = a_i b_j / (x_i + y_j)$. Input: a, b, x, and y defining $C_{ij} = a_i b_j / (x_i + y_j)$, and target size δ of con-eigenvalue. Output: con-eigenvalues lager than δ, and associated con-eigenvectors.

$(\Lambda, T) \leftarrow$ Con_Eigvector (a, b, x, y, δ)

`1. Compute partial Cholesky factors` $\left(\widetilde{L}, \widetilde{D}, \widetilde{P}\right) \leftarrow$

`Cholesky_Cauchy` (a, b, x, y, δ) `(Algorithm 4) and set` $X = \widetilde{P}\widetilde{L}$

`2. Optionally, compute the SVD factors` $\left(\widetilde{U}, \widetilde{\Sigma}, \widetilde{U'}\right) \leftarrow$

`SVD_RRD` $\left(\widetilde{P}\widetilde{L}, \widetilde{D}\right)$ `(Algorithm 1) and set` $X = \widetilde{U}$, $D = \widetilde{\Sigma}$

`3. Compute con-eigenvalues and con-eigenvectors` $(\Lambda, T) \leftarrow$

`ConEig_RRD` (X, D) `(Algorithm 2) using Algorithm 2`

`4. Select largest` l `such that` $\Lambda_{ll} \geq$

δ `and output` $\Lambda(1:l, 1:l)$, $T(1:n, 1:l)$

4. Experimental verification

We test the accuracy of Algorithm 5 on 300 random Cauchy matrices, $C_{ij} = (\alpha_i \overline{\alpha_j}) / (1 - \gamma_i \overline{\gamma_j})$, $i, j = 1, \ldots, 120$. The complex poles $\gamma_j = \rho_j e^{2\pi i \phi_j}$ and residues $w_j = \zeta_j e^{2\pi i \psi_j}$ are generated by taking ρ_j, ϕ_j, and ψ_j from the uniform distribution on $(0, 1)$, and taking ζ_j from the uniform distribution on $(0, 10)$. For each randomly generated matrix, we first compute, as a gauge, $\overline{C}C = T\Lambda^2 T^{-1}$ using the in-built Mathematica^{TM} eigenvalue solver with 300 digits of precision, and compare the result with $\widehat{T}$ and $\widehat{\Lambda}$ computed via Algorithm 5 using standard double precision. We then evaluate the maximum relative error in the con-eigenvalues $\lambda_j = \Lambda_{jj}$, $\max_j \left|\lambda_j - \widehat{\lambda_j}\right| / |\lambda_j|$, and the maximum error in the computed con-eigenvectors, $\max_j \left\| T(:,j) - \widehat{T}(:,j) \right\|_2 / \|T(:,j)\|_2$. We first scale $\widehat{T}(:,j)$ by the complex-valued constant $T(i_0,j) / \widehat{T}(i_0,j)$, $i_0 = \max_{1 \le i \le n} |T(i,j)|$, since $T(:,j)$ and $\widehat{T}(:,j)$ are defined only up to an arbitrary complex-valued factor.

Figures 4.1 and 4.2 summarize the result of a typical run. Figure 4.1(a) shows the distribution of the poles γ_j inside the unit disk and Figure 4.1(b) displays $\log_{10} \lambda_j^2$ as a function of the index j. Figure 4.2(a) shows the relative error in the con-eigenvalues $\left|\lambda_j - \widehat{\lambda_j}\right| / |\lambda_j|$, and Figure 4.2(b) displays the error in the normalized con-eigenvectors $\|z_j - \widehat{z_j}\|_2 / \|z_j\|_2$, as functions of the index j.

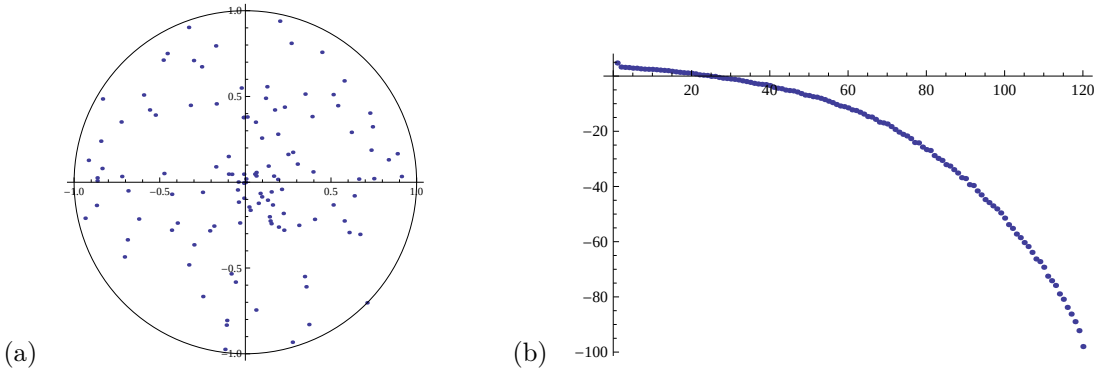

Figure 4.1. (a) Distribution of poles γ_j determining Cauchy matrix C in a typical run. (b) Exponential decay of the eigenvalues λ_j^2 of $\overline{C}C$ as a function of the index j using $\log_{10}$ scale.

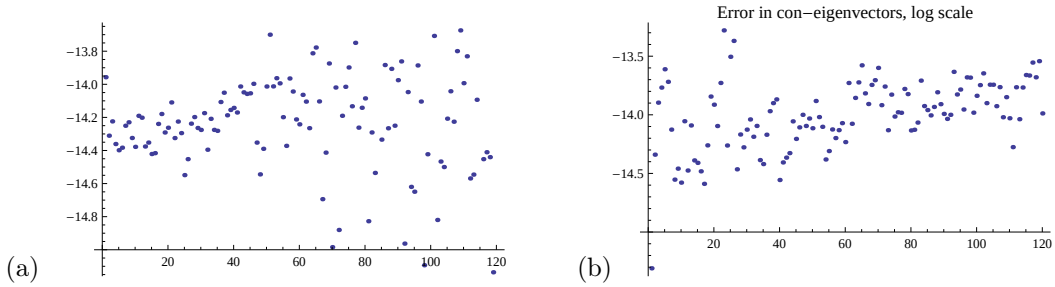

Figure 4.2. (a) Relative error in the jth con-eigenvalue, $\left|\lambda_j - \widehat{\lambda_j}\right| / |\lambda_j|$, as a function of the index j. (b) The error in the jth con-eigenvector, $\|z_j - \widehat{z_j}\|_2 / \|z_j\|_2$, $z_j = T(:,j)$, as a function of the index j.

In Figures 4.3 and 4.4 for each of the 300 random Cauchy matrices, we plot the error in the computed con-eigenvalues $\left|\widehat{\lambda_j} - \lambda_j\right| / |\lambda_j|$ and con-eigenvectors $\|\widehat{u_j} - u_j\|_2 / \|u_j\|_2$for $j = 1, 40, 80, 120$

(note the exponential decay of λ_j). We see that the con-eigenvalues and the con-eigenvectors are computed with nearly full precision for all the Cauchy matrices. In fact, the largest errors $\left|\widehat{\lambda_j} - \lambda_j\right| / |\lambda_j|$ and $\|\widehat{u_j} - u_j\|_2 / \|u_j\|_2$ in the computed con-eigenvalues and con-eigenvectors, for any of the 300 Cauchy matrices and any $1 \le j \le n$, are 5.3×10^{-13} and 6.4×10^{-13}.

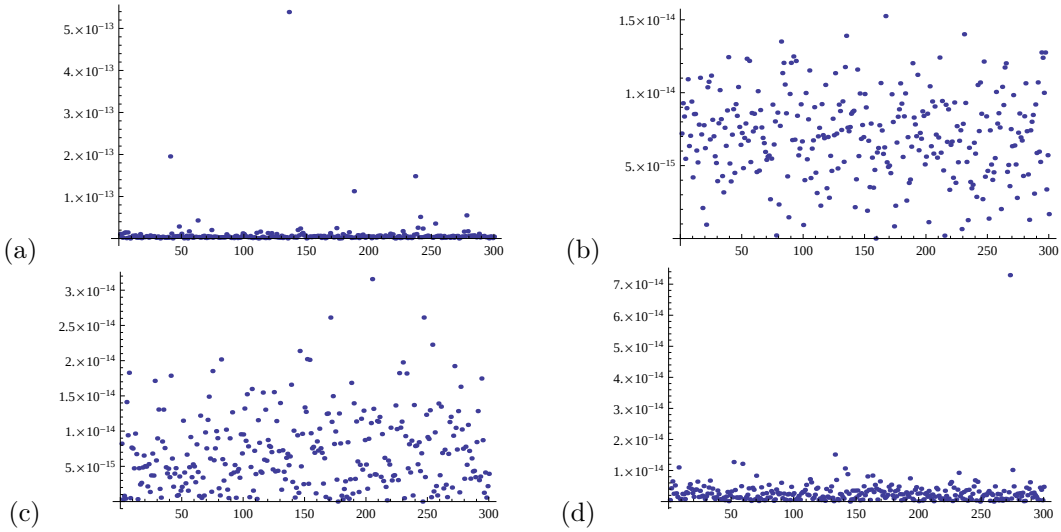

FIGURE 4.3. Relative error in the computed con-eigenvalues, $\left|\widehat{\lambda_j} - \lambda_j\right| / |\lambda_j|$, for $j = 1, 40, 80, 120$ ((a), (b), (c), and (d), respectively), plotted for each of the 300 random Cauchy matrices.

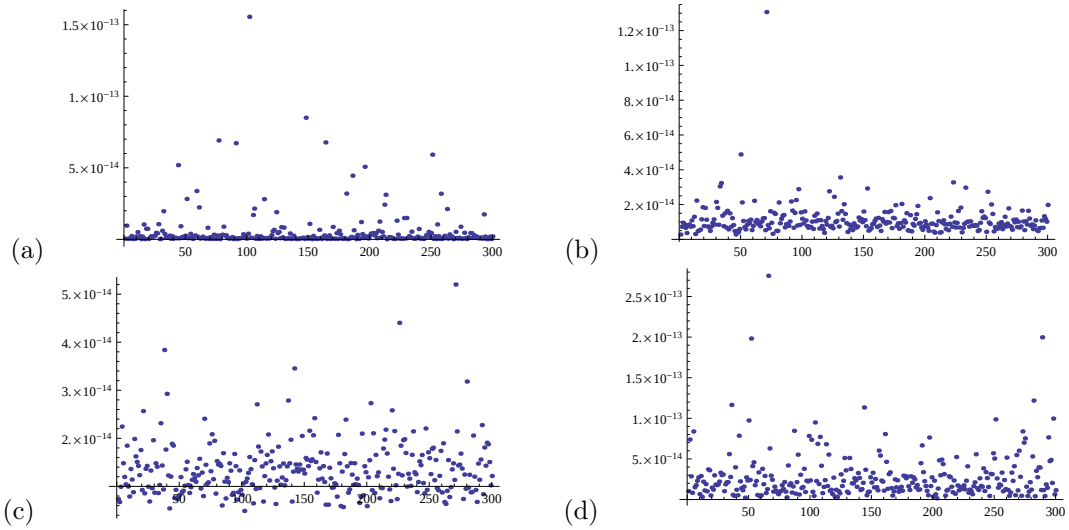

FIGURE 4.4. Relative error in the computed con-eigenvectors, $\|\widehat{u_j} - u_j\|_2 / \|u_j\|_2$, for $j = 1, 40, 80, 120$ ((a), (b), (c), and (d), respectively), plotted for each of the 300 random Cauchy matrices.

5. Accuracy and perturbation theory

We show that Algorithm 5 of the previous section (with the optional SVD step included) achieves high relative accuracy. We also demonstrate that small perturbations of a_i, b_j, x_i, and y_j determining C lead to small relative perturbations of the con-eigenvalues and small perturbations of the angles between subspaces spanned by the con-eigenvectors, as long as the parameters x_i and y_j are not too close in a relative sense. We mention that omitting Step 2 in Algorithm 5 also yields high relative accuracy, but with a more pessimistic error bound (in practice we have observed essentially no difference in accuracy between the two algorithms).

For two (complex) floating point numbers x and y, we denote by $\mathrm{fl}\,(x \odot y)$ the result of applying the operation $x \odot y$ in floating point arithmetic, where $\odot$ is one of the four basic operations, $\odot \in \{+, -, \times, \div\}$. We use that $\mathrm{fl}\,(x \odot y) = (x \odot y)\,(1+\delta)$, where $|\delta| \leq c\epsilon + \mathcal{O}\left(\epsilon^2\right)$, ϵ denote the machine round-off, and c is a small constant (cf. [20]). We will also abuse notation by letting $\mathrm{fl}\,(XY)$ denote the result of multiplying matrices X and Y in floating point arithmetic.

Let us define a quantity that will appear often in our estimates. For a given diagonal matrix D', we define

$$d_{ij}\,(D') = \min\left(\frac{|D'_{ii}|}{|D'_{jj}|}, \frac{|D'_{jj}|}{|D'_{ii}|}\right), \quad i,j = 1,\ldots,n. \tag{5.1}$$

For the diagonal matrices considered in this paper, the elements of the matrix $d_{ij}\,(D')$ decrease exponentially fast away from its diagonal $i = j$.

In Theorems 5-7 below we always assume that the con-eigenvalues are simple, although this is not a crucial restriction. In the statements and proofs of these theorems, the implicit constant factor implied by the notation $\mathcal{O}\,(\eta)$ and $\mathcal{O}\,(\epsilon)$ (here $\epsilon, \eta \ll 1$) depends only on the size n of the matrix C. We also use the notation $\mathcal{O}\,(1)$ to denote a quantity that depends only on the size n. We note that all these implicit constants may be tracked more carefully and are modest-sized functions of n.

Theorem 5. *Suppose that the parameters defining the positive-definite Cauchy matrix $C = C(a,b,x,y)$ are perturbed to $\tilde{a} = a + \delta a$, $\tilde{b} = b + \delta b$, $x = x + \delta x$, and $y = y + \delta y$. Let us define*

$$\eta = (1/\eta_1 + 1/\eta_2 + 1/\eta_3) \max\left\{\|\delta a\|_\infty, \|\delta b\|_\infty, \|\delta x\|_\infty, \|\delta y\|_\infty\right\},$$

where

$$\eta_1 = \min_{i\neq j} \frac{|x_i - x_j|}{|x_j| + |x_i|}, \quad \eta_2 = \min_{i\neq j} \frac{|y_i - y_j|}{|y_j| + |y_i|}, \quad \eta_3 = \min_{i\neq j} \frac{|x_i + y_j|}{|x_i| + |y_j|}.$$

Let $C = LDL^*$ denote the Cholesky factorization of C, and let $\widetilde{C} = C(\widetilde{a}, \widetilde{b}, \widetilde{x}, \widetilde{y})$ denote the Cauchy matrix corresponding to the perturbed parameters. Finally, let λ_i, $\widetilde{\lambda}_i$ and z_i, $\widetilde{z}_i$ denote the con-eigenvalues and con-eigenvectors of C and $\widetilde{C}$.

Then the con-eigenvalues λ_i and $\widetilde{\lambda}_i$ satisfy

$$\frac{\left|\lambda_i - \widetilde{\lambda}_i\right|}{|\lambda_i|} \leq \kappa\,(L)\,\mathcal{O}\,(\eta)\,,$$

and the acute angle between the con-eigenvectors z_i and $\widetilde{z}_i$ is bounded by

$$\sin\left(\angle z_i, \widetilde{z}_i\right) \leq \frac{\kappa\,(L)}{relgap_i}\mathcal{O}\,(\eta)\,, \quad relgap_i = \min_{j\neq i} \frac{|\lambda_i - \lambda_j|}{|\lambda_i| + |\lambda_j|},$$

where $\kappa(L) = \left\|L^{-1}\right\| \|L\|$ is the condition number.

Next we state

Theorem 6. *Suppose that Algorithm 5 (with the optional SVD step included) is used to compute the full con-eigenvalue decomposition of a positive-definite Cauchy matrix C. Suppose also that C has the Cholesky factorization $C = (PL)\,D^2\,(PL)^*$, where P is the permutation matrix that encodes*

complete pivoting. Then, letting λ_i, $\widehat{\lambda_i}$, *and* z_i, $\widehat{z_i}$ *denote exact and computed con-eigenvalues and con-eigenvectors of* C, *we bound the relative error*

$$\frac{\left|\lambda_i - \widehat{\lambda_i}\right|}{|\lambda_i|} \leq (\rho\mu\psi + \nu + \kappa(L))\,\mathcal{O}(\eta),$$

and the acute angle between z_i *and* $\widehat{z_i}$

$$\sin(\angle\widehat{z_i}, z_i) \leq \frac{\rho\mu\psi + \nu + \kappa(L)}{relgap_i}\mathcal{O}(\epsilon), \quad relgap_i = \min_{j\neq i}\frac{|\lambda_i - \lambda_j|}{|\lambda_i| + |\lambda_j|}.$$

The value of ν *comes from the error analysis of the modified Jacobi algorithm described in Section 2.4. The values of* ρ, μ, *and* ψ *come from the error analysis of the Householder QR algorithm described in Section 2.3.2.*

Theorem 7. *Suppose Algorithm 5 (with the optional SVD step included) is used to compute* m *approximate con-eigenvalues* $\widehat{\lambda_i}$ *and associated con-eigenvectors* $\widehat{z_i}$ *of a positive-definite Cauchy matrix* C. *Suppose also that* C *has the Cholesky factorization* $C = (PL)\,D^2\,(PL)^*$, *where* P *is the permutation matrix that encodes complete pivoting. Let* $\widetilde{L} = L(1:n, m)$ *and* $\widetilde{D} = D(1:m, 1:1m)$, *where* m *is chosen so that* $D_{mm} \leq \delta\epsilon$ *for some* $\delta > 0$, *and* $\widetilde{L}\widetilde{D} = Q\widetilde{R}\widetilde{P}$ *is the QR factorization of* $\widetilde{L}\widetilde{D}$ *with column pivoting.*

Then the acute angle between z_m *and* $\widehat{z_m}$ *may be bounded by*

$$\sin(\angle\widehat{z_i}, z_i) \leq \left(\frac{\rho\mu\psi + \nu + \kappa(L)}{relgap_i} + \frac{\|C\|}{relgap_i}\mathcal{O}(\epsilon)\right),$$

where, as in Theorem 6, the value of ν *comes from the Jacobi algorithm and values of* ρ, μ, *and* ψ *from the Householder QR algorithm.*

Although estimates (2.8) and (2.9) imply that the SVD of the matrix $G = \Sigma\left(U^{\mathrm{T}}U\right)\Sigma = (QW)\,SV$ in Algorithm 5 is computed accurately (and, therefore, the con-eigenvalues are computed accurately), the con-eigenvector z_i is obtained from the right singular vector v_i of G by the transformation $z_i = \left(\overline{U}\Sigma\right)v_i / \left\|\left(\overline{U}\Sigma\right)v_i\right\|$. Since the elements of Σ decay exponentially fast, the matrix $U\Sigma$ has a very large condition number, and it would at first appear that the computed con-eigenvector $\widehat{z_i} = \mathrm{fl}\left(\left(\widehat{\overline{U}}\widehat{\Sigma}\right)\widehat{v_i} / \left\|\left(\widehat{\overline{U}}\widehat{\Sigma}\right)\widehat{v_i}\right\|\right)$ could be a wildly inaccurate approximation of z_i, even if $\widehat{v_i}$, $\widehat{U}$, and $\widehat{\Sigma}$ are known accurately. The reason Algorithm 5 achieves high accuracy is that the singular vectors v_i are bounded component-wise by $|v_i(j)| \leq d_{ij}(\Sigma)\,c_0$, and the computed singular vector components $\widehat{v_i}(j)$ of G are accurate relative to the scaling of Σ in the sense that $|v_i(j) - \widehat{v_i}(j)| \leq c_0 d_{ij}(\Sigma)\,\epsilon$. The constant c_0 is moderate sized if L, the unit triangular matrix from the Cholesky factorization of C, is well-conditioned, and the relative gap, relgap_i, is not too small. Recall that the quantities $d_{ij}(\Sigma) = \min(\Sigma_{ii}/\Sigma_{jj}, \Sigma_{jj}/\Sigma_{ii})$ may be many orders of magnitude smaller than ϵ. Then it follows that

$$\left|\frac{\Sigma_{jj}}{\Sigma_{ii}}v_i(j) - \mathrm{fl}\left(\frac{\widehat{\Sigma_{jj}}}{\widehat{\Sigma_{ii}}}\widehat{v_i}(j)\right)\right| \leq c_1\epsilon, \tag{5.2}$$

where c_1 is a constant (small if c_0 is small). It also turns out that $\|\Sigma v_i/\Sigma_{ii}\| \geq 1$. Therefore, since U is orthogonal, and the ith exact and computed con-eigenvectors z_i and $\widehat{z_i}$ are given by $z_i = U(\Sigma w_i/\Sigma_{ii})$ and $\widehat{z_i} = \mathrm{fl}\left(\widehat{U}\left(\widehat{\Sigma}\widehat{w_i}/\widehat{\Sigma_{ii}}\right)\right)$, it follows from the above inequality that $\sin(\angle z_i, \widehat{z_i}) = c_1\mathcal{O}(\epsilon)$. The details of the proof may be found in Section 5.2.

5.1. **Perturbation theorem.** In this section, we prove Theorem 5. We start by formulating several preliminary results. Lemma 8 describes how perturbations of the vectors a, b, x, and y defining the Cauchy matrix $C = C(a,b,x,y)$ change the factors L and D in the Cholesky decomposition $C = LDL^*$ (see [11] for a proof).

Lemma 8. *Suppose the data defining the Cauchy matrix $C = C(a,b,x,y)$ is perturbed to $\tilde{a} = a+\delta a$, $\tilde{b} = b+\delta b$, $x = x+\delta x$, and $y = y+\delta y$. Let us define*

$$\eta = (1/\eta_1 + 1/\eta_2 + 1/\eta_3) \max\left\{ \|\delta a\|_\infty, \|\delta b\|_\infty, \|\delta x\|_\infty, \|\delta y\|_\infty \right\},$$

where

$$\eta_1 = \min_{i\neq j} \frac{|x_i - x_j|}{|x_j| + |x_i|}, \quad \eta_2 = \min_{i\neq j} \frac{|y_i - y_j|}{|y_j| + |y_i|}, \quad \eta_3 = \min_{i\neq j} \frac{|x_i + y_j|}{|x_i| + |y_j|}.$$

Then $C = C(a,b,x,y)$ and $\widetilde{C} = \widetilde{C}(\widetilde{a},\widetilde{b},\widetilde{x},\widetilde{y})$ have Cholesky factorizations $C = LDL^*$ and $\widetilde{C} = \widetilde{L}\widetilde{D}\widetilde{L}^*$, where L, $\widetilde{L}$ are unit lower triangular matrices, D, $\widetilde{D}$ are diagonal matrices with positive entries, and

$$\left| L_{ij} - \widehat{L}_{ij} \right| = |L_{ij}|\, \mathcal{O}(\eta), \quad \left| D_{ii} - \widehat{D_{ii}} \right| = |D_{ii}|\, \mathcal{O}(\eta).$$

We also need the following proposition, the proof of which may be obtained via techniques developed in [4] and [15] (see also [24]). The proof is given in the Appendix (it is a straightforward modification of that found in [15, Proposition 2.12], and is provided for completeness).

Proposition 9. *Suppose $G = DBD$ and $G+\delta G = D(B+\delta B)D$, where D is a diagonal matrix with positive diagonal entries, and B, $B+\delta B$ are non-singular Hermitian matrices. Then, letting Σ denote the diagonal matrix of eigenvalues of G, unit eigenvectors x_i and $\widetilde{x}_i$ of DBD and $D(B+\delta B)D$, may be chosen so that*

$$|x_i(j) - \widetilde{x}_i(j)| \quad \leq \quad d_{ij}(D) \left(\frac{\kappa(B)^{1/2}}{\sigma_{\min}(B)} \right) \frac{\mathcal{O}(\|\delta B\|)}{relgap_i}.$$

Also, the following norm-wise and component-wise bounds hold:

$$\kappa^{-1/2}(B) \leq \left\| \frac{Dx_i}{D_{ii}} \right\|, \quad |x_i(j)| \leq \kappa(B)^{1/2} d_{ij}(D).$$

Finally, we state the main result needed to prove Theorem 5.

Proposition 10. *Suppose $G = DBD$ and $G+\delta G = D(B+\delta B)D$, where B is a non-singular complex-valued matrix, and D is a diagonal matrix with positive diagonal elements. Then the ith (left or right) singular vectors v_i and $\widetilde{v}_i$ of G and $G+\delta G$ may be chosen so that*

$$|v_i(j) - \widetilde{v}_i(j)| \leq d_{ij}(D) \frac{\kappa(B)^2}{relgap_i} \mathcal{O}\left(\frac{\|\delta B\|}{\|B\|} \right), \quad |v_i(j)| \leq \kappa(B)^{1/2} d_{ij}(D).$$

If, in addition, the matrix B is complex symmetric ($B^T = B$), then

$$\left\| \frac{Dv_i}{D_{ii}} \right\| \geq \kappa^{-1/2}(B).$$

Proof. Note that the SVD of DBD may be obtained from the eigenvalue decomposition of the Hermitian matrix

$$F = \begin{pmatrix} 0 & DBD \\ (DBD)^* & 0 \end{pmatrix} = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix}.$$

Similarly, the SVD of $D(B+\delta B)D$ may be obtained from the eigenvalue decomposition of the Hermitian matrix

$$F + \delta F = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} 0 & B+\delta B \\ (B+\delta B)^* & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix}.$$

Indeed, given the SVD, $DBD = USV^*$, the matrix F has the eigenvalue decomposition

$$F = \left[\frac{1}{\sqrt{2}}\begin{pmatrix} V & V \\ U & -U \end{pmatrix}\right]\begin{pmatrix} S & 0 \\ 0 & -S \end{pmatrix}\left[\frac{1}{\sqrt{2}}\begin{pmatrix} V & V \\ U & -U \end{pmatrix}\right]^*.$$

A similar statement applies to $F + \delta F$. The component-wise bounds now follow from applying Proposition 10 to F and $F + \delta F$, and using the equalities

$$\left\|\begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}^{-1}\right\| = \left\|\begin{pmatrix} 0 & B^{-1} \\ (B^{-1})^* & 0 \end{pmatrix}\right\| = 2\left\|B^{-1}\right\|, \text{ and } \left\|\begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}\right\| = 2\|B\|.$$

If B is complex symmetric, then we may take $U = \overline{V}$ and apply Proposition 10, yielding

$$\left\|\begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix}\begin{pmatrix} v_i \\ \overline{v_i} \end{pmatrix}/D_{ii}\right\| = 2\left\|\frac{Dv_i}{D_{ii}}\right\| \geq \kappa^{1/2}\left(\begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}\right) = 2\kappa^{1/2}(B).$$

We now prove Theorem 5.

Recall that the matrix T of con-eigenvectors satisfies $T = \overline{U}\Sigma\overline{V}$, where $C = U\Sigma^2U^*$ is the SVD of C and V is the matrix of right singular vectors of $G = \Sigma\left(U^{\mathrm{T}}U\right)\Sigma$. Let $\widetilde{C} = \widetilde{U}\widetilde{\Sigma}^2\widetilde{U}^*$ denote the SVD of $\widetilde{C}$. From Lemma 8 and the discussion in Section 2.2 (see also [11]),

$$\left\|U - \widetilde{U}\right\| = \kappa(L)\,\mathcal{O}(\eta), \quad \left|\frac{\Sigma_{ii} - \widetilde{\Sigma_{ii}}}{\Sigma_{ii}}\right| = \kappa(L)\,\mathcal{O}(\eta). \tag{5.3}$$

Defining $\widetilde{G} = \widetilde{\Sigma}\left(\widetilde{U}^{\mathrm{T}}\widetilde{U}\right)\widetilde{\Sigma}$, the above bounds yield $\widetilde{G} = \Sigma\left(U^{\mathrm{T}}U + E\right)\Sigma$, where $\|E\| = \kappa(L)\,\mathcal{O}(\eta)$.

Proposition 10 states that unit singular vectors $\widetilde{v}_i$ of $\widetilde{G}$ and v_i of G may be chosen so that

$$|v_i(j)| \leq \kappa^{1/2}\left(U^{\mathrm{T}}U\right)d_{ij}(\Sigma) = d_{ij}(\Sigma), \tag{5.4}$$

and

$$|v_i(j) - \widetilde{v}_i(j)| \leq d_{ij}(\Sigma)\left(\frac{\kappa^2\left(U^{\mathrm{T}}U\right)}{\mathrm{relgap}_i}\right)O(\|E\|) \leq d_{ij}(\Sigma)\frac{\kappa(L)}{\mathrm{relgap}_i}\mathcal{O}(\eta). \tag{5.5}$$

Therefore, defining $w_i = \Sigma v_i/\Sigma_{ii}$ and $\widetilde{w}_i = \widetilde{\Sigma}\widetilde{v}_i/\widetilde{\Sigma_{ii}}$, we have

$$\begin{aligned} |w_i(j) - \widetilde{w}_i(j)| &= \frac{\Sigma_{jj}}{\Sigma_{ii}}\left|v_i(j) - \frac{\widetilde{\Sigma_{jj}}}{\Sigma_{jj}}\frac{\Sigma_{ii}}{\widetilde{\Sigma_{ii}}}\widetilde{v}_i(j)\right| \\ &\leq \frac{\Sigma_{jj}}{\Sigma_{ii}}\left|v_i(j) - (1 + \kappa(L)\,\mathcal{O}(\eta))\,\widetilde{v}_i(j)\right| \\ &\leq \frac{\Sigma_{jj}}{\Sigma_{ii}}\left(|v_i(j) - \widetilde{v}_i(j)| + |v_i(j)|\,\kappa(L)\,\mathcal{O}(\eta)\right) \\ &\leq \left(\frac{\kappa(L)}{\mathrm{relgap}_i} + \kappa(L)\right)\mathcal{O}(\eta), \end{aligned}$$

where we used (5.3) in the first inequality, and (5.4)-(5.5) in the last one. Proposition 10 also implies that $1/\|w_i\| \leq \kappa^{1/2}\left(U^{\mathrm{T}}U\right) = 1$ and $1/\|\widehat{w_i}\| \leq \mathcal{O}(1)$.

The proof now follows upon noting that the con-eigenvectors z_i and $\widetilde{z_i}$ satisfy $z_i = Uw_i$, $\widetilde{z_i} = \widetilde{U}\widetilde{w}_i$, and using (5.3),

$$\begin{aligned}
\sin\angle\left(\widehat{z_i}, z_i\right) &= \min_{\alpha} \left\| \frac{\widetilde{U}\widetilde{w}_i}{\left\|\widetilde{U}\widetilde{w}_i\right\|} - \alpha\frac{Uw_i}{\|Uw_i\|} \right\| \leq \left\| \frac{\widetilde{U}\widetilde{w}_i}{\left\|\widetilde{U}\widetilde{w}_i\right\|} - \frac{Uw_i}{\|U\widetilde{w_i}\|} \right\| \\
&\leq \left\| \left(\frac{\widetilde{U}\widetilde{w}_i}{\left\|\widetilde{U}\widetilde{w}_i\right\|} - \frac{U\widetilde{w}_i}{\left\|\widetilde{U}\widetilde{w}_i\right\|} \right) + \left(\frac{U\widetilde{w}_i}{\left\|\widetilde{U}\widetilde{w}_i\right\|} - \frac{Uw_i}{\left\|\widetilde{U}\widetilde{w}_i\right\|} \right) \right\| \\
&\leq \frac{\left\|\left(\widetilde{U} - U\right)\widetilde{w}_i\right\|}{\left\|\widetilde{U}\widetilde{w}_i\right\|} + \frac{\|U\|}{\left\|\widetilde{U}\widetilde{w}_i\right\|}\left\|(\widetilde{w}_i - w_i)\right\| \leq \left(\frac{\kappa(L)}{\mathrm{relgap}_i} + \kappa(L) \right) \mathcal{O}(\eta).
\end{aligned}$$

□

5.2. **Proof of Theorem 6 (high relative accuracy of Algorithm 5).** First consider Steps 1 and 2 of Algorithm 5. From [11], the computed SVD factors $\widehat{\Sigma}$ and $\widehat{U}$ of C satisfy

$$\left| D_{ij} - \widehat{D_{ij}} \right| \leq |D_{ij}|\,\kappa(L)\,\mathcal{O}(\epsilon), \quad \left\| U - \widehat{U} \right\| \leq \kappa(L)\,\mathcal{O}(\epsilon). \tag{5.6}$$

Next, Algorithm 2 is used to compute the con-eigenvalue decomposition of $\widehat{\Sigma}\left(\widehat{U}^{\mathrm{T}}\widehat{U}\right)\widehat{\Sigma}$. In Step 1, the computed matrix $\widehat{G}$ satisfies $\widehat{G} = \mathrm{fl}\left(\widehat{\Sigma}\left(\widehat{U}^{\mathrm{T}}\widehat{U}\right)\widehat{\Sigma}\right) = \Sigma\left(U^{\mathrm{T}}U + E_0\right)\Sigma$, where $\|E_0\| \leq \kappa(L)\,\mathcal{O}(\epsilon)$.

In Step 2 of Algorithm 2, a computed upper triangular factor $\widehat{R}$ of $\widehat{G}$ is obtained using the Householder QR algorithm with complete pivoting. By Theorem 2, there is an orthogonal matrix Q and permutation matrices P_1 and P_2 such that the computed $\widehat{R}$ satisfies

$$P_1 Q \widehat{R} P_2 = \Sigma\left(U^{\mathrm{T}}U + E_2\right)\Sigma, \tag{5.7}$$

where $E_2 = E_0 + E_1$, $\|E_1\| \leq \rho\mu\psi\mathcal{O}(\epsilon)$, and ρ, μ, and ψ are "pivot growth factors" described in Section 2.3.2.

Step 3 of Algorithm 2 involves computing an approximate SVD $\widehat{R}P_2 \approx \widehat{U_l}\widehat{\Lambda}\widehat{U_r}^*$ using the modified one-sided Jacobi algorithm, applied from the right. Note that, from (5.7), if $\widehat{R}P_2$ has the (exact) SVD $\widehat{R}P_2 = U_l \Lambda U_r^*$, then U_r is the matrix of right singular vectors of $\Sigma\left(U^{\mathrm{T}}U + E_2\right)\Sigma$. Therefore, defining $U_r' = \Sigma U_r \Sigma^{-1}$, Proposition 10 implies that $\left|(U_r')_{ij}\right| \leq \kappa\left(U^{\mathrm{T}}U + E_2\right) \leq \mathcal{O}(1)$. Also,

$$\left|(U_r')^{-1}_{ij}\right| = \left|\left(\Sigma U_r^{-1}\Sigma^{-1}\right)_{ij}\right| = \left|\left(\Sigma U_r^{*}\Sigma^{-1}\right)_{ij}\right| = \left|(U_r')_{ji}\right| \leq \mathcal{O}(1).$$

Therefore, we have $\kappa(U_r') \leq \mathcal{O}(1)$. Then it follows that we may write $\widehat{R}P_2 = R'\Sigma$, where $R' = U_l\left(\Sigma U_r \Sigma^{-1}\right)$ and $\kappa(R') \leq \mathcal{O}(1)$.

Now, let $\widehat{v_i}$ denote the computed right singular vector of $\widehat{R}P_2$. Then, by Theorem 3 and the equality $\widehat{R}P_2 = R'\Sigma$, there is an exact right singular vector $v_i^{(1)}$ of $\widehat{R}$ such that

$$\left|\widehat{v_i}(j) - v_i^{(1)}(j)\right| \leq d_{ij}(\Sigma)\,\frac{\nu}{\mathrm{relgap}_i}\mathcal{O}(\epsilon). \tag{5.8}$$

Also, since $P_1 Q \widehat{R} P_2 = (P_1 Q)\,\widehat{U_l}\widehat{\Sigma}\widehat{U_r} = D\left(U^{\mathrm{T}}U + E_2\right)D$, Theorem 10 ensures that there is an exact left singular vector v_i of $G = D\left(U^{\mathrm{T}}U\right)D$ that satisfies

$$\left|v_i(j) - v_i^{(1)}(j)\right| \leq d_{ij}(\Sigma)\,\frac{\kappa^2\left(U^{\mathrm{T}}U\right)}{\mathrm{relgap}_i}\mathcal{O}\left(\|E_2\|\right) \leq d_{ij}(\Sigma)\,\frac{\kappa(L) + \rho\mu\psi}{\mathrm{relgap}_i}\mathcal{O}(\epsilon). \tag{5.9}$$

Therefore, from (5.8) and (5.9) we obtain

$$|v_i(j) - \widehat{v_i}(j)| \leq d_{ij}(\Sigma)\,\frac{\rho\mu\psi + \nu + \kappa(L)}{\mathrm{relgap}_i}\mathcal{O}(\epsilon). \tag{5.10}$$

We will also need the following component-wise bounds,

$$|v_i(j)|, |\widehat{v_i}(j)| \leq d_{ij}(\Sigma)\,\mathcal{O}(1), \tag{5.11}$$

which follow from Proposition 10.

Finally, Step 4 of Algorithm 2 involves computing the con-eigenvector $\widehat{z_i} = \mathrm{fl}\left(\widehat{U}\left(\widehat{\Sigma}\widehat{v_i}/\widehat{\Sigma_{ii}}\right)\right)$. Now let $\widehat{w_i} = \mathrm{fl}\left(\widehat{\Sigma}\widehat{v_i}/\widehat{\Sigma_{ii}}\right) = \widehat{\Sigma}\widehat{v_i}/\widehat{\Sigma_{ii}} + E_3$, $w_i = \Sigma v_i/\Sigma_{ii}$ and note that

$$|E_4| \leq \left|\widehat{\Sigma}\right| |\widehat{v_i}| / \left|\widehat{\Sigma_{ii}}\right| \mathcal{O}(\epsilon) = (1+\mathcal{O}(\epsilon))\,|\Sigma|\,|\widehat{v_i}|\,/\,|\Sigma_{ii}|\,\mathcal{O}(\epsilon) = \mathcal{O}(\epsilon),$$

where (5.6) is used in the first equality and (5.11) is used in the second equality. Therefore, we obtain

$$\begin{aligned}
|w_i(j) - \widehat{w_i}(j)| &= \frac{\Sigma_{jj}}{\Sigma_{ii}}\left|v_i(j) - \frac{\widehat{\Sigma_{jj}}}{\Sigma_{jj}}\frac{\Sigma_{ii}}{\widehat{\Sigma_{ii}}}\widehat{v_i}(j)\right| \\
&= \frac{\Sigma_{jj}}{\Sigma_{ii}}\left|v_i(j) - (1+\kappa(L)\,\mathcal{O}(\epsilon))\,\widehat{v_i}(j)\right| \\
&\leq \frac{\Sigma_{jj}}{\Sigma_{ii}}\left(|v_i(j) - \widehat{v_i}(j)| + |\widehat{v_i}(j)|\,\kappa(L)\,\mathcal{O}(\epsilon)\right) \\
&\leq \frac{\Sigma_{jj}}{\Sigma_{ii}} d_{ij}(\Sigma)\left(\frac{\rho\mu\psi + \nu + \kappa(L)}{\mathrm{relgap}_i} + \kappa(L)\right)\mathcal{O}(\epsilon) \\
&\leq \left(\frac{\rho\mu\psi + \nu + \kappa(L)}{\mathrm{relgap}_i} + \kappa(L)\right)\mathcal{O}(\epsilon),
\end{aligned}$$

where we used (5.6) in the second equality, and (5.11) and (5.10) in the second inequality. We also have from Proposition 10 that $1/\|w_i\| \leq \kappa^{1/2}\left(U^{\mathrm{T}}U\right) = 1$ and $1/\|\widehat{w_i}\| \leq \kappa^{1/2}\left(U^{\mathrm{T}}U + E\right) = \mathcal{O}(1)$. Therefore, recalling that U is orthogonal and $\left\|\widehat{U} - U\right\| = \kappa(L)\,\mathcal{O}(\epsilon)$, a similar calculation as in the proof of Theorem 5 shows that the acute angle between the computed and exact con-eigenvectors $\widehat{z_i} = \mathrm{fl}\left(\widehat{U}\widehat{w_i}\right)$ and $z_i = Uw_i$ is bounded by the quantity in the last inequality above, thus completing the proof.

5.3. **Proof of Theorem 7.** We need the following well-known result describing the sensitivity of the eigenvalue problem for diagonalizable matrices.

Lemma 11. *Assuming that A has simple eigenvalues, we consider its perturbation $A + E$. Let $X = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ denote the matrix of unit eigenvectors of A, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Then the acute angle between the ith unit eigenvectors x_i and $\widetilde{x_i}$ of A and $\widetilde{A}$ is bounded by*

$$\sin(\angle x_i, \widetilde{x_i}) \leq \kappa(X)\,\frac{\mathcal{O}(\|E\|)}{absgap_i}, \quad \textit{where} \quad absgap_i = \min_{j\neq i}|\lambda_i - \lambda_j|.$$

The next result shows that the matrix of eigenvectors of $\overline{C}C$ is well-conditioned.

Lemma 12. *Let C denote a positive-definite Cauchy matrix, and let T denote the matrix of unit eigenvectors of $\overline{C}C$. Then we have $\kappa(T) = \mathcal{O}(1)$, where $\mathcal{O}(1)$ denotes a modest-sized function of n.*

Proof. From Section 3.1, we know that the matrix of (unnormalized) eigenvectors of $\overline{C}C$ is given by $Z = \overline{U}\left(\Sigma\overline{V}\Sigma^{-1}\right)$, where V is the matrix of right singular vectors of $\Sigma\left(U^{\mathrm{T}}U\right)\Sigma$ and U is the unitary matrix in $C = U\Sigma U^*$. Now, by Proposition 10, we have that

$$\left|\left(\Sigma\overline{V}\Sigma^{-1}\right)_{ij}\right| = \left|\left(\Sigma V\Sigma^{-1}\right)_{ij}\right| \leq \kappa^{1/2}\left(U^{\mathrm{T}}U\right) = 1. \tag{5.12}$$

Also,

$$\left|\left(\Sigma\overline{V}\Sigma^{-1}\right)_{ij}^{-1}\right| = \left|\left(\Sigma\overline{V}^{-1}\Sigma^{-1}\right)_{ij}\right| = \left|\left(\Sigma V\Sigma^{-1}\right)_{ji}\right| \leq 1.$$

Therefore, $\kappa\left(\Sigma\overline{V}\Sigma^{-1}\right) = \mathcal{O}(1)$, and it follows that $\kappa(Z) \leq \kappa\left(\Sigma\overline{V}\Sigma^{-1}\right) = \mathcal{O}(1)$.

Now define $\Omega_{ii} = \|z_i\|$, and let $T = Z\Omega^{-1}$ denote the matrix of unit eigenvectors. From (5.12), $\Omega_{ii} = \mathcal{O}(1)$. Also, again using Proposition 10, we have

$$\Omega_{ii} = \left\|\overline{U}\Sigma\overline{v_i}\Sigma_{ii}^{-1}\right\| = \left\|\Sigma\overline{v_i}\Sigma_{ii}^{-1}\right\| \geq \kappa^{1/2}\left(U^{\mathrm{T}}U\right) = 1.$$

Therefore, we finally obtain $\kappa(T) \leq \kappa(Z)\,\kappa(\Omega) = \mathcal{O}(1)$. □

The next lemma is the key to proving Theorem 7.

Proposition 13. *Suppose that PCP has the Cholesky factorization $C = (PL)\,D^2\,(PL)$ (where complete pivoting is used). For a given $m \leq n$, define $\widetilde{C} = L_1 D_1^2 L_1^*$, where $L_1 = (PL)(1:n, 1:m)$, $D_1 = D(1:m, 1:m)$, and where $1 \leq m \leq n$ is such that $D_{mm} \leq \epsilon\lambda_i$. Assume that the eigenvalues of $\overline{C}C$ and $\overline{\widetilde{C}}\widetilde{C}$ are simple. Then if $\overline{C}Cz_i = \lambda_i z_i$ and $\overline{\widetilde{C}}\widetilde{C}\widetilde{z}_i = \widetilde{\lambda}_i\widetilde{z}_i$, we may bound the acute angle between z_i and $\widetilde{z}_i$ by*

$$\sin(\angle z_i, \widetilde{z}_i) \leq \frac{\|C\|}{relgap_i}\mathcal{O}(\epsilon), \quad where\ relgap_i = \min_{j\neq i}\frac{|\lambda_i - \lambda_j|}{\lambda_i + \lambda_j}.$$

Proof. Let

$$PL = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \quad D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}.$$

Then, since complete pivoting is used and the components of PL are bounded by 1,

$$\begin{aligned} C = (PL)\,D^2\,(PL)^* &= \begin{pmatrix} L_{11}D_1 & L_{12}D_2 \\ L_{21}D_1 & L_{22}D_2 \end{pmatrix}\begin{pmatrix} D_1L_{11}^* & D_1L_{21}^* \\ D_2L_{12}^* & D_2L_{22}^* \end{pmatrix}, \\ &= \begin{pmatrix} L_{11}D_1 & 0 \\ L_{21}D_1 & 0 \end{pmatrix}\begin{pmatrix} D_1L_{11}^* & D_1L_{21}^* \\ 0 & 0 \end{pmatrix} + E_1 \\ &= L_1D_1^2L_1^* + E_1 = \widetilde{C} + E_1, \end{aligned}$$

where $\|E_1\| \leq \mathcal{O}(1)D_{mm}$. Therefore, $\overline{\widetilde{C}}\widetilde{C} = \overline{C}C + E_2$, where $E_2 = \overline{\widetilde{C}}E_1 + \overline{E}_1\widetilde{C} + \overline{E}_1E_1$ and $\|E_2\| \leq \mathcal{O}(1)\,\|C\|\,D_{mm}$.

Now, let T denote the matrix of unit eigenvectors of $\overline{C}C$. Then by Lemmas 11 and 12, we have

$$\begin{aligned} \sin(\angle z_i, \widetilde{z}_i) &\leq \kappa(T)\,\frac{\mathcal{O}(\|E_2\|)}{\mathrm{absgap}_i} \leq \frac{\mathcal{O}(1)\,\|C\|}{\mathrm{absgap}_i}D_{mm}. \\ &\leq \frac{\mathcal{O}(1)}{\mathrm{relgap}_i}\,\|C\|\,\frac{D_{mm}}{\lambda_i} \leq \frac{\|C\|}{\mathrm{relgap}_i}\mathcal{O}(\epsilon). \end{aligned}$$

□

Finally, we are ready to prove Theorem 7.

Proof. The proof of Theorem 6 shows that the m computed con-eigenvectors $\widehat{\widetilde{z}}_i$ of $\widetilde{C} = \widetilde{L}\widetilde{D}^2\widetilde{L}$ are close to the exact eigenvectors $\widetilde{z}_i$, i.e.,

$$\sin\left(\angle\widehat{\widetilde{z}}_i, \widetilde{z}_i\right) \leq \frac{\rho\mu\psi + \nu + \kappa(L)}{\mathrm{relgap}_i}\mathcal{O}(\epsilon),$$

Since Proposition 13 implies that

$$\sin(\angle z_i, \widetilde{z}_i) \leq \frac{\|C\|}{\mathrm{relgap}_i}\mathcal{O}(\epsilon),$$

the claim follows. □

6. Appendix

We prove Proposition 10. For the sake of clarity, we first prove several lemmas. Throughout this section, unless stated otherwise B will denote a non-singular Hermitian matrix and D will denote a real-valued matrix with positive diagonal elements.

Lemma 14. *The eigenvalues λ_i of DBD may be arranged so that*

$$\lambda_{\min}(B) D_{ii}^2 \leq \lambda_i \leq \lambda_{\max}(B) D_{ii}^2.$$

Proof. It suffices to show that $\lambda_{min}(B)\left(D^2x, x\right) \leq (DBDx, x) \leq \lambda_{\max}(B)\left(D^2x, x\right)$, since D^2 and DBD are Hermitian. These inequalities follows from the fact that $(DBDx, x) = (BDx, Dx)$, since

$$\lambda_{\max}(B)\left(D^2x, x\right) = \lambda_{\max}(B)(Dx, Dx) \leq (BDx, Dx) \leq \lambda_{\max}(B)(Dx, Dx) = \lambda_{\max}(B)\left(D^2x, x\right).$$

□

Lemma 15. *The components of the ith eigenvector x_i of DBD, normalized so that $\|x_i\| = 1$, satisfy the estimate*

$$|x_i(j)| \leq \kappa(B)^{1/2} d_{ij}(D).$$

It also holds that

$$\left\|\frac{Dx_i}{D_{ii}}\right\| \geq \kappa^{-1/2}(B).$$

Proof. Let $G = DBD$, and suppose that $Gx_i = \lambda_i x_i$. Then $y_i = Dx_i$ satisfies, $By_i = \lambda_i D^{-2} y_i$. Therefore,

$$|\lambda_i| = |\lambda_i| \left\|D^{-1}y_i\right\|^2 = |\lambda_i| \left|\left(y_i, D^{-2}y_i\right)\right| = |\lambda_i| |(y_i, By_i)| \geq \sigma_{\min}(B) \|y_i\|^2.$$

Similarly,

$$|\lambda_i| = |\lambda_i| \left\|D^{-1}y_i\right\|^2 = |\lambda_i| \left|\left(y_i, D^{-2}y_i\right)\right| = |\lambda_i| |(y_i, By_i)| \leq \sigma_{\max}(B) \|y_i\|^2.$$

It follows that

$$\left(\frac{|\lambda_i|}{\sigma_{\max}(B)}\right)^{1/2} \leq \|y_i\| \leq \left(\frac{|\lambda_i|}{\sigma_{\min}(B)}\right)^{1/2}. \tag{6.1}$$

The estimate on the norm of Dx_i/D_{ii} follows from the first inequality above and Lemma 14, since

$$\left\|\frac{Dx_i}{D_{ii}}\right\| = \frac{\|y_i\|}{D_{ii}} \geq \left(\frac{|\lambda_i|}{D_{ii}^2}\right)^{1/2} \frac{1}{\sigma_{\max}(B)^{1/2}} \geq \left(\frac{\sigma_{\min}(B)}{\sigma_{\min}(B)}\right)^{1/2}.$$

To prove the inequality on the components of x_i, first suppose that $D_{ii} \leq D_{jj}$. Then

$$\begin{aligned}
|x_i(j)| &= \frac{|y_i(j)|}{D_{jj}} \leq \frac{1}{D_{jj}} \left(\frac{|\lambda_i|}{\sigma_{\min}(B)}\right)^{1/2} \\
&= \left(\frac{D_{ii}}{D_{jj}}\right) \left(\frac{|\lambda_i|^{1/2}}{D_{ii}}\right) \left(\frac{1}{\sigma_{\min}(B)}\right)^{1/2} \\
&\leq \left(\frac{\sigma_{\max}(B)}{\sigma_{\min}(B)}\right)^{1/2} \left(\frac{D_{ii}}{D_{jj}}\right) \\
&\leq \kappa(B)^{1/2} \left(\frac{D_{ii}}{D_{jj}}\right),
\end{aligned}$$

where we used (6.1) in the first inequality and Lemma 14 in the second inequality.

Now suppose that $D_{jj} > D_{ii}$. Then x_i is an eigenvector of G^{-1},

$$G^{-1}x_i = \left(D^{-1}B^{-1}D^{-1}\right)x_i = \frac{1}{\lambda_i}x_i.$$

Since $D_{jj}^{-1} < D_{ii}^{-1}$, we may apply the first case to G^{-1},

$$\begin{aligned} |x_i(j)| &\leq \kappa(B^{-1})^{1/2} \frac{D_{ii}^{-1}}{D_{jj}^{-1}} \\ &= \kappa(B)^{1/2} \frac{D_{jj}}{D_{ii}}. \end{aligned}$$

The inequalities on the components of x_i follow. □

Lemma 16. *Given a (Hermitian) perturbation δB, the eigenvalues λ_i and $\widetilde{\lambda}_i$ of DBD and $D(B+\delta B)D$ may be arranged so that*

$$1 - \frac{\|\delta B\|}{\sigma_{\min}(B)} \leq \left| \frac{\widetilde{\lambda}_i}{\lambda_i} \right| \leq 1 + \frac{\|\delta B\|}{\sigma_{\min}(B)}.$$

Proof. Use Lemma 2.1 of [15] with $H = B$, $K = D^{-2}$, $\delta H = \delta B$, and $\delta K = 0$.

□

We now prove Proposition 10.

Proof. From standard perturbation theory, if $x_i(t)$ denotes the ith eigenvector of $D(B + t\delta B)D$, $0 \leq t \leq 1$, then

$$x_i'(t) = t \sum_{k \neq j} \frac{x_k(t)^* (D\delta BD) x_i(t)}{\lambda_i(t) - \lambda_k(t)} x_k(t) = t \sum_{k \neq j} \frac{(Dx_k(t))^* \delta B (Dx_i(t))}{\lambda_i(t) - \lambda_k(t)} x_k(t).$$

Now let $y_k(t) = Dx_k(t)$, $z_k(t) = y_k(t)/\|y_k(t)\|$. Then

$$|x_i'(t)(j)| \leq t \sum_{k \neq j} \xi_{kj} \frac{|z_k(t)^* \delta B z_i(t)|}{|\lambda_i(t) - \lambda_k(t)| / |\lambda_i(t)\lambda_k(t)|^{1/2}} |x_k(t)(j)|, \tag{6.2}$$

where

$$|\xi_{ki}| = \frac{\|y_k(t)\|}{|\lambda_k(t)|^{1/2}} \frac{\|y_i(t)\|}{|\lambda_i(t)|^{1/2}}.$$

Since $y_k(t)$ satisfies

$$(B + t\delta B) y_k(t) = \lambda_k(t) D^{-2} y_k(t),$$

we may estimate

$$\begin{aligned} |\lambda_k(t)| &= |\lambda_k(t)| \left| \left(y_k(t), D^{-2} y_k(t) \right) \right| \\ &= |(y_k(t), (B + t\delta B) y_k(t))| \\ &\geq \|y_k(t)\|^2 \sigma_{\min}(B) (1 - t\|\delta B\|). \end{aligned}$$

From the above bound,

$$\begin{aligned} |\xi_{kj}| &\leq \left(\frac{1}{\sigma_{\min}(B)(1 - t\|\delta B\|)} \right) \\ &= \frac{1}{\sigma_{\min}(B)} + \mathcal{O}(\|\delta B\|). \end{aligned} \tag{6.3}$$

Using (6.3) and Lemma 15 in equation (6.2), we obtain

$$\begin{aligned}
|x_i'(t)(j)| &\leq t\sum_{k\neq j}\xi_{kj}\frac{|z_k(t)^*\,(\delta B)\,z_i(t)|}{|\lambda_i(t)-\lambda_k(t)|\,/\,|\lambda_i(t)\lambda_k(t)|^{1/2}}\,|x_k(t)(j)| \\
&\leq \frac{(n-1)\,t}{\sigma_{\min}(B)}\frac{\|\delta B\|}{|\lambda_i(t)-\lambda_k(t)|\,/\,|\lambda_i(t)\lambda_k(t)|^{1/2}}\,|x_k(t)(j)| \\
&\leq \frac{(n-1)\,t\kappa(B)^{1/2}}{\sigma_{\min}(B)}\min\left(\frac{|\lambda_i(t)\lambda_k(t)|^{1/2}}{|\lambda_i(t)-\lambda_k(t)|}\left|\frac{\lambda_k(t)}{\lambda_j(t)}\right|^{1/2},\frac{|\lambda_i(t)\lambda_k(t)|^{1/2}}{|\lambda_i(t)-\lambda_k(t)|}\left|\frac{\lambda_j(t)}{\lambda_k(t)}\right|^{1/2}\right) \\
&\leq \|\delta B\|\frac{(n-1)\,t\kappa(B)^{1/2}}{\sigma_{\min}(B)}\min\left(\frac{|\lambda_k(t)|}{|\lambda_i(t)-\lambda_k(t)|}\left|\frac{\lambda_i(t)}{\lambda_j(t)}\right|^{1/2},\frac{|\lambda_i(t)|}{|\lambda_i(t)-\lambda_k(t)|}\left|\frac{\lambda_j(t)}{\lambda_i(t)}\right|^{1/2}\right) \\
&\leq \|\delta B\|\left(\frac{(n-1)\,t\kappa(B)^{1/2}}{\sigma_{\min}(B)}\frac{|\lambda_k(t)|+|\lambda_i(t)|}{|\lambda_i(t)-\lambda_k(t)|}\right)\min\left(\left|\frac{\lambda_i(t)}{\lambda_j(t)}\right|^{1/2},\left|\frac{\lambda_j(t)}{\lambda_i(t)}\right|^{1/2}\right).
\end{aligned}$$

The result follows from the equality $|\widetilde{x_i}(j)-x_i(j)|=\left|\int_0^1 x_i'(t)(j)\,dt\right|$. □

References

[1] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Infinite Hankel matrices and generalized Carathéodory-Fejér and I. Schur problems. *Funkcional. Anal. i Priložen.*, 2(4):1–17, 1968.

[2] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz. *Funkcional. Anal. i Priložen.*, 2(1):1–19, 1968.

[3] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem. *Mat. Sb. (N.S.)*, 86(128):34–75, 1971.

[4] J. Barlow and J. Demmel. Computing accurate eigensystems of scaled diagonally dominant matrices. Technical report, New York University, Technical Report 421, 1988.

[5] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Appl. Comput. Harmon. Anal.*, 19(1):17–48, 2005.

[6] G. Beylkin and L. Monzón. Nonlinear inversion of a band-limited Fourier transform. *Appl. Comput. Harmon. Anal.*, 27(3):351–366, 2009. doi: 10.1016/j.acha.2009.04.003.

[7] T. Boros, T. Kailath, and V. Olshevsky. Pivoting and backward stability of fast algorithms for solving Cauchy linear equations. *Linear Algebra Appl.*, 343/344:63–99, 2002. Special issue on structured and infinite systems of linear equations.

[8] C. Davis and W. M. Kahan. Some new bounds on perturbation of subspaces. *Bulletin of the American Mathematical Society*, 75(4):863–&, 1969.

[9] P. Deift, J. Demmel, C. Li, and C. Tomei. The bidiagonal singular value decomposition and hamiltonian mechanics. *SIAM J. Num. Anal*, 28:1463–1516, 1991.

[10] Ph. Delsarte, Y. Genin, and Y. Kamp. On the role of the Nevanlinna-Pick problem in circuit and system theory. *International Journal of Circuit Theory and Applications*, 9(2):177–187, 1981.

[11] J. Demmel. Accurate singular value decompositions of structured matrices. *SIAM J. Matrix Anal. Appl.*, 21(2):562–580 (electronic), 1999.

[12] J. Demmel, I. Dumitriu, O. Holtz, and P. Koev. Accurate and efficient expression evaluation and linear algebra. *Acta Numer.*, 17:87–145, 2008.

[13] J. Demmel, M. Gu, S. Eisenstat, I. Slapnicar, K. Veselic, and Z. Drmac. Computing the singular value decomposition with high relative accuracy. *LAPACK Working Note*, 119(CS-97-348), 1997.

[14] J. Demmel and W. Kahan. Accurate singular values of bidiagonal matrices. *SIAM J. Sci. Stat. Comput*, 11:873–912, 1990.

[15] J. Demmel and K. Veselic. Jacobi's method is more accurate than QR. *SIAM. J. Matrix Anal. and Appl.*, 13(4):1204–1245, 1992.

[16] J. W. Demmel and W. Gragg. On computing accurate singular values and eigenvalues of matrices with acyclic graphs. *Linear Algebra Appl.*, 185:203–217, 1993.

[17] Z. Drmač. Accurate computation of the product-induced singular value decomposition with applications. *SIAM Journal of Numerical Analysis*, 35(5):1969–1994, 1998.

[18] K. V. Fernando and B. N. Parlett. Accurate singular values and differential qd algorithms. *Numer. Math.*, 67(2):191–229, 1994.

[19] N. J. Higham. QR factorization with complete pivoting and accurate computation of the SVD. In *Proceedings of the International Workshop on Accurate Solution of Eigenvalue Problems (University Park, PA, 1998)*, volume 309, pages 153–174, 2000.
[20] N. J. Higham. *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2002.
[21] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990.
[22] P. Koev. Accurate eigenvalues and SVDs of totally nonnegative matrices. *SIAM J. Matrix Anal. Appl*, 27:1–23, 2005.
[23] R. Mathias. Accurate eigensystem computations by Jacobi methods. *SIAM J. Matrix Anal. Appl*, 16:977–1003, 1996.
[24] R. Mathias. Spectral perturbation bounds for positive definite matrices. *SIAM J. Matrix Anal. Appl*, 18:959–80, 1997.
[25] D. J. Newman. Rational approximation of |x|. *Michigan Math. J.*, 11:11–14, 1964.
[26] D. M. Priest. Algorithms for arbitrary precision floating point arithmetic. In *Proceedings of the 10th Symposium on Computer Arithmetic*, pages 132–145. IEEE Computer Society Press, 1991.
[27] I. Slapnicar. Highly accurate symmetric eigenvalue decomposition and hyperbolic SVD. *Linear Algebra and its Applications*, 358:387–424, 2003.
[28] N. J. Young. The singular-value decomposition of an infinite Hankel matrix. *Linear Algebra and its Applications*, 50:639–656, 1983.

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526, United States