

# 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 a (near) optimally small  $L_\infty$  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  $m$ th con-eigenvector of an  $n \times n$  Cauchy matrix, where the associated con-eigenvalue  $\lambda_m > 0$  gives the approximation error in the  $L_\infty$  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  $m$ th con-eigenvalue in  $\mathcal{O}(m \log n)$  operations and, since the con-eigenvalues of positive-definite Cauchy matrices decay exponentially fast, we obtain (near) optimal rational approximations in  $\mathcal{O}(n \log \delta^{-1})$  operations, where  $\delta$  is the approximation error in the  $L_\infty$  norm. We provide error bounds demonstrating high relative accuracy of the computed con-eigenvalues and the high accuracy of the unit con-eigenvectors. We also provide examples of using the algorithm to compute (near) optimal rational approximations of functions with singularities and sharp transitions, where approximation errors close to machine roundoff are obtained. Finally, we present numerical tests on random (complex-valued) Cauchy matrices to 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,

$$1.1) \quad \mathbf{C} \mathbf{u}_m = \lambda_m \mathbf{u}_m$$

the error close to  $\mathbf{m}$ . The form (1.2) ensures that  $\mathbf{f}$  (

positive-definite matrices [20], scaled diagonally dominant matrices [4], totally positive matrices [1], certain indefinite matrices [6], and Cauchy matrices (as well as, more generally, matrices with displacement rank one) [15]. For such matrices, recent algorithmic advances (see [24, 25]) make the cost of achieving high relative accuracy comparable to that of alternative (and less accurate) SVD methods.

The con-eigenvalue algorithm considered here is based on computing the eigenvalue decomposition of the product,  $\overline{\mathbf{C}}\mathbf{C}$ , of positive-definite Cauchy matrices  $\overline{\mathbf{C}}$  and  $\mathbf{C}$ , and is similar to the algorithm in [17] for the generalized eigenvalue decomposition, as well as the algorithm in [2] for the product SVD decomposition. We also rely on the algorithm in [15] for computing, with high relative accuracy, the Cholesky decomposition (with complete pivoting)  $\mathbf{C} = (\mathbf{P}\mathbf{L})\mathbf{D}^2(\mathbf{P}\mathbf{L})^*$  of a positive-definite Cauchy matrix  $\mathbf{C}$ . However, since we are interested in computing only con-eigenvalues of some approximate size  $\epsilon$ , we stop

paper with other algorithms in the literature for constructing optimal rational approximations. For the convenience of the reader we also provide relevant background material in Section 7.

2. ACCURATE CON-EIGENVALUE DECOMPOSITION (AN INFORMAL DERIVATION)

2.1. **Constructing optimal rational approximations via a con-eigenvalue problem.** In order to motivate our con-eigenvalue algorithm, let us explain how the accurate computation of small con-eigenvalues and associated con-eigenvectors allows us to construct optimal rational approximations.

We consider an algorithm to find a rational approximation  $r(e^{2ix})$  to  $f(e^{2ix})$  in (1.2) with a specified number of poles and with a (nearly) optimally small error in the  $L^\infty$ -norm. The algorithm is based on a theorem of Adamyan, Arov, and Krein (referred to below as the AAK Theorem) [1]. We note that the formulation given below in terms of a con-eigenvalue problem is similar to the approach taken in [14] and [6].

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

$$r(z) = \sum_{i=1}^m \frac{a_i}{z - \alpha_i} + \sum_{i=1}^m \frac{\overline{a_i}z}{1 - \overline{\alpha_i}z} + o(\epsilon),$$

are as follows:

- 1) Compute a con-eigenvalue  $0 < \lambda_m \leq \dots \leq \lambda_1$  and corresponding con-eigenvector  $u$  of the Cauchy matrix  $C_{ij} = C_{ij}(\alpha_i, \beta_j)$ ,

$$2.1) \quad Cu = \lambda_m u, \text{ where } u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad C_{ij} = \frac{a_i b_j}{x_i + y_j}, \quad i, j = 1, \dots, n,$$

and  $a_i = \sqrt{\lambda_i} / \alpha_i$ ,  $b_j = \sqrt{\lambda_j}$ ,  $x_i = \alpha_i^{-1}$ ,  $y_j = -\beta_j$ . The con-eigenvalues of  $C$  are labeled in non-increasing order,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ .

- 2) Find the (exactly)  $m$  smallest con-eigenvalues  $\lambda_1, \dots, \lambda_m$ .

*Remark 1.* In practice, finding the new poles  $\zeta_i$  using the formula for  $\mathbf{v}(\mathbf{z})$  in (2.2) is ill-advised, since evaluating  $\mathbf{v}(\mathbf{z})$  in this form could result in loss of significant digits through catastrophic cancellation. Indeed, it turns out (see [6, Section 6] and [27]) that the values of the con-eigenvector components satisfy  $\mathbf{u}_i = \sqrt{\lambda_i} \mathbf{v}(\zeta_i)$ ,  $i = 1, \dots, \mathbf{n}$ . It then follows that the sum (2.2) must suffer cancellation of about  $\log_{10}(\lambda_i^{-1})$  digits if  $\mathbf{v}(\zeta_i)$  and  $\mathbf{v}(\mathbf{z})$  are of comparable size (note that  $\lambda_i$  controls the approximation error and, thus, is necessarily small). On the other hand, the function values  $\mathbf{v}(\zeta_i) = \mathbf{u}_i / \sqrt{\lambda_i}$ ,  $i = 1, \dots, \mathbf{n}$ , along with the  $\mathbf{n}$  poles  $1/\sqrt{\lambda_i}$  of  $\mathbf{v}(\mathbf{z})$ , completely determine (2.2). Since the poles  $\zeta_i$  of  $\mathbf{f}(\mathbf{z})$  are often close to the poles  $\zeta_i$  of  $\mathbf{r}(\mathbf{z})$ , we have observed that evaluating  $\mathbf{v}(\mathbf{z})$  by using rational interpolation via continued fractions with the known values  $\mathbf{v}(\zeta_i)$  allows us to obtain the new poles  $\zeta_i$  with nearly full precision. In particular, an approximation  $\tilde{\mathbf{v}}(\mathbf{z})$  to  $\mathbf{v}(\mathbf{z})$  is computed via continued fractions,

$$(2.4) \quad \tilde{\mathbf{v}}(\mathbf{z}) = \frac{\mathbf{a}_1}{1 + \mathbf{a}_2(\mathbf{z} - \zeta_1) / (1 + \mathbf{a}_3(\mathbf{z} - \zeta_2) / (1 + \dots))},$$

where the coefficients  $\mathbf{a}_j$  are determined from the interpolation conditions  $\tilde{\mathbf{v}}(\zeta_i) = \mathbf{v}(\zeta_i)$ . If the poles  $\zeta_i$  are given in the form  $\zeta_i = \exp(-\lambda_i)$ , we find that Newton's method on  $\tilde{\mathbf{v}}(\exp(-\lambda))$  yields the new poles  $\zeta_i = \exp(-\lambda_i)$  with nearly full relative accuracy even when  $\text{Re}(\lambda_i)$

---

**Algorithm 1** `ConEig_RRD` ( $\mathbf{X}, \mathbf{D}$ ) computes accurate con-eigenvalue decomposition of  $\mathbf{X}\mathbf{D}^2\mathbf{X}^*$ . Input: rank-revealing factors  $\mathbf{X}$  and  $\mathbf{D}$  of dimensions  $\mathbf{n} \times \mathbf{m}$  and  $\mathbf{m} \times \mathbf{m}$ , where the diagonal of  $\mathbf{D} > 0$  is decreasing. Output:  $\mathbf{m}$  con-eigenvalues/con-eigenvectors of  $\mathbf{X}\mathbf{D}\mathbf{X}^*$ , contained in  $\lambda$  and  $\mathbf{T}$ .

---

$(\lambda, \mathbf{T}) \leftarrow \text{ConEig\_RRD}(\mathbf{X}, \mathbf{D})$

1. Form  $\mathbf{G} = \mathbf{D}(\mathbf{X}^T\mathbf{X})\mathbf{D}$
  2. Compute QR factors  $(\mathbf{Q}, \mathbf{R}) \leftarrow \text{Householder\_QR}$  of  $\mathbf{G}$  ( $\mathbf{G} = \mathbf{Q}\mathbf{R}$ ), with optional pivoting (see Section 7.3)
  3. Compute the SVD factors  $(\mathbf{U}_l, \sigma, \mathbf{U}_r) \leftarrow \text{Jacobi}(\mathbf{R})$  of  $\mathbf{R}$  ( $\mathbf{R} = \mathbf{U}_l \mathbf{U}_r^*$ ), using one-sided Jacobi, applied from the left (see Section 7.4)
  4. Compute  $\mathbf{R}_1 = \mathbf{D}^{-1}\mathbf{R}\mathbf{D}^{-1}$ ,  $\mathbf{X}_1 = \mathbf{D}^{-1}\mathbf{U}_l$ ,  $\mathbf{Y}_1 = \mathbf{R}_1^{-1}\mathbf{X}_1$  (see (2.6) below)
  5. Form the matrix of con-eigenvectors  $\mathbf{T} = \overline{\mathbf{X}}\mathbf{Y}_1$ , and output con-eigenvalues  $\lambda$  and con-eigenvectors  $\mathbf{T}$
- 

Importantly, for Cauchy matrices  $\mathbf{A} = \mathbf{C}$ ) the elements of  $\mathbf{D}$  decay exponentially fast, and it would appear that computing the con-eigenvectors  $\mathbf{z}_i = \overline{\mathbf{X}}\mathbf{D}\mathbf{v}_i / \sqrt{\mathbf{D}_{ii}}$  might lead to wildly inaccurate results even if the right singular vector of  $\mathbf{G}$ ,  $\mathbf{v}_i$ , is computed accurately. However, as we show in Section 5, Algorithm 1 achieves high accuracy despite the extreme ill-conditioning of  $\mathbf{D}$ . The key reason is that the right singular vector  $\mathbf{v}_i$ , corresponding to the singular value  $\sigma_i$ , scales like  $|\mathbf{v}_i(\mathbf{j})| \leq c_v \min(\mathbf{D}_{jj} / \sqrt{\mathbf{D}_{ii}}, \sqrt{\mathbf{D}_{ii}} / \mathbf{D}_{jj})$ , and the computed singular vector  $\widehat{\mathbf{v}}_i$  is accurate relative to the scaling in  $\mathbf{D}$  and  $\sigma_i$  in the sense that

$$|\mathbf{v}_i(\mathbf{j}) - \widehat{\mathbf{v}}_i(\mathbf{j})| \leq \min\left\{\frac{\mathbf{D}_{jj}}{\sqrt{\mathbf{D}_{ii}}}, \frac{\sqrt{\mathbf{D}_{ii}}}{\mathbf{D}_{jj}}\right\} \mathcal{O}(\epsilon).$$

For Cauchy matrices, the quantity  $\min(\mathbf{D}_{jj} / \sqrt{\mathbf{D}_{ii}}, \sqrt{\mathbf{D}_{ii}} / \mathbf{D}_{jj})$  decreases exponentially fast away from the diagonal  $\mathbf{i} = \mathbf{j}$ .

Let us give an informal explanation of the reasons why Algorithm 1 yields accurate results. As discussed in Section 7, the QR Householder algorithm computes an accurate rank-revealing decomposition of  $\mathbf{G} = \mathbf{Q}\mathbf{R}$ . It turns out (see the online version [28, Lemma 11]) that  $\mathbf{R}$  may be factored as  $\mathbf{R} = \mathbf{D}^2\mathbf{R}_0$ , where  $\mathbf{R}_0$  is graded relative to  $\mathbf{D}$  in the sense that  $\|\mathbf{D}\mathbf{R}_0\mathbf{D}^{-1}\|$  and  $\|\mathbf{D}\mathbf{R}_0^{-1}\mathbf{D}^{-1}\|$  are not too large, as long as the  $\mathbf{n}$  leading principal minors of  $\mathbf{X}^T\mathbf{X}$  are well-conditioned. Therefore, from the discussion in Section 7.4 (see in particular Theorem 10), the one-sided Jacobi algorithm computes the  $i$ th left singular vector  $\mathbf{u}_i$  of  $\mathbf{R}$  accurately relative to the scaling  $\min\{\mathbf{D}_{jj} / \sqrt{\mathbf{D}_{ii}}, \sqrt{\mathbf{D}_{ii}} / \mathbf{D}_{jj}\}$ . It follows that  $\mathbf{D}^{-1}\mathbf{u}_i$  may also be computed accurately. Finally, since the  $i$ th right singular vector  $\mathbf{v}_i$  of  $\mathbf{R}$  (and  $\mathbf{G}$ ) satisfies

$$\begin{aligned} \mathbf{D}\mathbf{v}_i &= \mathbf{D}\mathbf{R}^{-1}\mathbf{u}_i \\ \mathbf{D}\mathbf{v}_i / \sqrt{\mathbf{D}_{ii}} &= \mathbf{D}\mathbf{R}^{-1}\mathbf{u}_i / \sqrt{\mathbf{D}_{ii}} \\ &= (\mathbf{D}\mathbf{R}_0\mathbf{D}^{-1})^{-1} (\mathbf{D}^{-1}\mathbf{u}_i / \sqrt{\mathbf{D}_{ii}}), \end{aligned} \tag{2.6}$$

the con-eigenvector  $\mathbf{z}_i = \overline{\mathbf{X}}(\mathbf{D}\mathbf{v}_i / \sqrt{\mathbf{D}_{ii}})$  may be computed accurately, as long as  $\mathbf{D}\mathbf{R}_0\mathbf{D}^{-1}$  is computed accurately and is well-conditioned (we show this is the case if  $\mathbf{n}$  leading principal minors of  $\mathbf{X}^T\mathbf{X}$  are well-conditioned).

accuracy. As explained in the next section,  $\mathbf{j}^{-1} - \overline{\mathbf{i}}$  may be accurately computed

---

**Algorithm 2** `Pivot_Order`( $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \epsilon$ ) pre-computes pivot order for Cholesky factorization of  $\mathbf{n} \times \mathbf{n}$  positive-definite Cauchy matrix  $\mathbf{C}_{ij} = \mathbf{a}_i \mathbf{b}_j / (\mathbf{x}_i + \mathbf{y}_j)$ . Input:  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{x}$ , and  $\mathbf{y}$  defining  $\mathbf{C}_{ij} = \mathbf{a}_i \mathbf{b}_j / (\mathbf{x}_i + \mathbf{y}_j)$ , and target size  $\epsilon$  of con-eigenvalue. Output: correctly pivoted vectors  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{x}$ , and  $\mathbf{y}$ , truncation size  $\mathbf{m}$ , and  $\mathbf{m} \times \mathbf{n}$  permutation matrix  $\tilde{\mathbf{P}}$

---

$$(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{P}}, \mathbf{m}) \leftarrow \text{Pivot\_Order}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \epsilon)$$

Form vector



CON-EIGENVALUE

$$\widehat{\mathbf{f}}_{\mathbf{n}} = \mathbf{h}_{\mathbf{n}} + \frac{(-1)^{\mathbf{L}}}{(2 \mathbf{i}\mathbf{n})^{\mathbf{L}}} \int_0^{\mathbf{x}_0} \mathbf{f}^{(\mathbf{L})}(\mathbf{x}) e^{2 \mathbf{i}\mathbf{n}\mathbf{x}} d\mathbf{x} + \frac{(-1)^{\mathbf{L}}}{(2 \mathbf{i}\mathbf{n})^{\mathbf{L}}} \int_{\mathbf{x}_0}^1 \mathbf{f}^{(\mathbf{L})}(\mathbf{x}) e^{2 \mathbf{i}\mathbf{n}\mathbf{x}} d\mathbf{x},$$

where

$$\mathbf{h}_{\mathbf{n}} = \sum_{\mathbf{p}=1}^{\mathbf{L}} \frac{(-1)^{\mathbf{p}}}{(2 \mathbf{i}\mathbf{n})^{\mathbf{p}}} \left( e^{2 \mathbf{i}\mathbf{n}\mathbf{x}_0} \mathbf{F}^{(\mathbf{p}-1)}(\mathbf{x}_0) + \mathbf{F}^{(\mathbf{p}-1)}(0) \right),$$

$\mathbf{F}^{(\mathbf{p})}(\mathbf{x}) = \mathbf{f}^{(\mathbf{p})}(\mathbf{x}^+) - \mathbf{f}^{(\mathbf{p})}(\mathbf{x}^-)$  and  $\mathbf{x}^+$ ,  $\mathbf{x}^-$  indicate directional limits. As the first step in constructing a (near) optimal rational approximation to  $\mathbf{f}$ , we subtract the leading  $\mathbf{L}$  terms of the asymptotic expansion of  $\widehat{\mathbf{f}}_{\mathbf{n}}$  and consider  $\mathbf{g}$





CON-EIGENVALUE

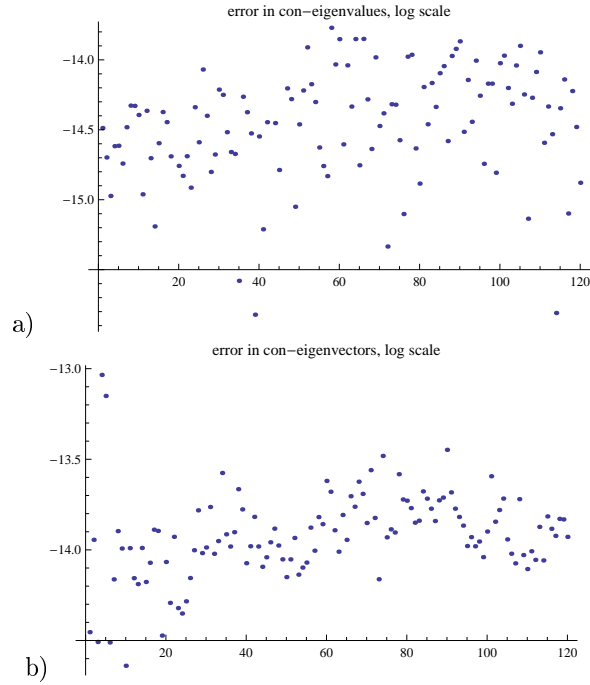


FIGURE 4.2. a) Relative error in the  $\mathbf{j}$ th con-eigenvalue,  $\left| \lambda_{\mathbf{j}} - \widehat{\lambda}_{\mathbf{j}} \right| / \left| \lambda_{\mathbf{j}} \right|$ , as a function of the index  $\mathbf{j}$ . b) The error in the  $\mathbf{j}$ th con-eigenvector,  $\| \mathbf{z}_{\mathbf{j}} - \widehat{\mathbf{z}}_{\mathbf{j}} \|_2 / \| \mathbf{z}_{\mathbf{j}} \|_2$ ,  $\mathbf{z}_{\mathbf{j}} = \mathbf{Z}(:, \mathbf{j})$ , as a function of the index  $\mathbf{j}$ .

where the condition number  $\kappa(\mathbf{L}) = \| \mathbf{L} \| \| \mathbf{L}^{-1} \|$  is typically small. The estimates in Theorems 6-7 also depend on

$$5.2) \quad \mu_3(\mathbf{L}) = \| \mathbf{L}^{-1} \| \left( \mu_1(\mathbf{L}) + \mu_2(\mathbf{L}) + \kappa^3(\mathbf{L}) \right),$$

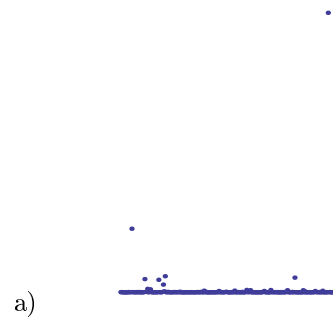
where  $\mu_1$ ,  $\mu_2$ , and  $\mu_3$  are “pivot growth” factors associated with the QR factorization (see Section 7.1), and the factor  $\kappa$  is associated with the one-sided Jacobi algorithm (see 7.12)).

*Remark.* There are simple formulas for  $\mathbf{L}_{ij}$  and  $(\mathbf{L}^{-1})_{ij}$  [10] in terms of the parameters  $\mathbf{a}_i$ ,  $\mathbf{b}_j$ ,  $\mathbf{x}_i$  and  $\mathbf{y}_j$  defining the Cauchy matrix  $\mathbf{C}$ , and it is possible that the bounds below may be improved by using this additional structure.

**Theorem 5.** Suppose  $i2Q\text{Td}(t(J2.01a(i20\text{Td}(t(J27\text{Td}(t(J27\text{Tp})3\text{TJ2Td})3\text{TJ2Tdcobi})\text{T}31.2m247.4243.5lS2722.71425.5m$

(

$\kappa$



a) 



depend only on the well-conditioned matrix  $\mathbf{L}$  and, in particular, are independent of the exponentially decaying diagonal matrix  $\mathbf{D}$ , they still scale like  $\|\mathbf{L}\|$ ; the bounds on the con-eigenvalues are better—they scale like  $\|\mathbf{L}\|^{-3}$ . However, in practice Algorithm 4 achieves nearly full precision for all the con-eigenvalues and con-eigenvectors. While it is likely that better estimates can be obtained, those presented here elucidate the basic mechanism behind the high accuracy that we observe in our experiments.

## 6. DISCUSSION: COMPARISON WITH RELATED APPROACHES FOR CONSTRUCTING OPTIMAL RATIONAL APPROXIMATIONS

Numerical approaches for finding near optimal rational approximations originate in theoretical results of Adamyan, Arov, and Krein [1, 2, 3]. In particular, given a periodic function  $\mathbf{f}(e^{2i\mathbf{x}}) \in \mathbf{L}^\infty(0, 1)$ , AAK theory yields an optimal “rational-like” approximation  $\mathbf{r}_M(e^{2i\mathbf{x}})$ ,

$$6.1) \quad \mathbf{r}_M(\mathbf{z}) = \frac{\mathbf{a}_0 + \mathbf{a}_1\mathbf{z} + \mathbf{a}_2\mathbf{z}^2 + \dots}{(\mathbf{z} - \mathbf{z}_1)\dots(\mathbf{z} - \mathbf{z}_M)}, \quad |\mathbf{z}_j| < 1,$$

constructed from the left and right singular vectors corresponding to the  $M$ th singular value,  $\sigma_M$ , of the infinite Hankel matrix  $\mathbf{H}_{ij} = \hat{\mathbf{f}}(\mathbf{i} + \mathbf{j} - 1)$ ,  $\mathbf{i}, \mathbf{j} = 1, 2, \dots$ . The numerator of  $\mathbf{r}_M(\mathbf{z})$  in 6.1) is analytic in the unit disk. The approximation error satisfies

$$\max_{\mathbf{x}}$$

high degree polynomials (determined at the SVD step) may be sensitive to perturbations in their coefficients. However, when limited to approximating smooth functions, these “truncated Hankel” methods can yield surprisingly high accuracy since the errors in the poles may be compensated by the residues. As far as we are aware, truncated Hankel methods for constructing optimal rational approximations for functions with singularities generally do not achieve approximation errors better than  $\approx 10^{-4}$ . In contrast, in Section 3.1 we show that the reduction algorithm approximates piecewise smooth functions with errors close to machine precision.

We also note that the results in [27] (illustrated in Section 3.2) demonstrate an effective numerical calculus based on the reduction algorithm, capable of computing highly accurate solutions to viscous Burgers’ equation for viscosity as small as  $10^{-5}$ . These solutions exhibit moving transition regions of width  $\approx 10^{-5}$ , and computing them with high accuracy over long time intervals is a nontrivial task for any numerical method. The con-eigenvalue algorithm of this paper is critical to the high accuracy and efficiency of this numerical calculus.

## 7. APPENDIX: BACKGROUND ON ALGORITHMS FOR HIGH RELATIVE ACCURACY

Here we provide necessary background on computing highly accurate SVDs. Although the results we need in [20, 17, 4, 15, 29] are only stated there for real-valued matrices, they carry over to complex-valued matrices with minor modifications and are formulated as such.

**7.1. Accurate SVDs of matrices with rank-revealing decompositions.** According to the usual perturbation theory for the SVD (see e.g. [12]), perturbations  $\mathbf{A}$  of a matrix  $\mathbf{A}$  change the  $i$ th singular value  $\sigma_i$  by  $\delta\sigma_i$  and corresponding unit eigenvector  $\mathbf{u}_i$  by  $\delta\mathbf{u}_i$ , where (assuming for simplicity that  $\sigma_i$  is simple),

$$(7.1) \quad |\delta\sigma_i|/\sigma_i \leq \|\mathbf{A}\|, \quad \|\delta\mathbf{u}_i\| \leq \frac{\|\mathbf{A}\|}{\text{absgap}_i}, \quad \text{absgap}_i = \min_{i \neq j} |\sigma_i - \sigma_j|/\sigma_i.$$

Therefore, small perturbations in the elements of  $\mathbf{A}$  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  $\mathbf{A} + \delta\mathbf{A}$ , where  $\|\delta\mathbf{A}\|/\|\mathbf{A}\| = \mathcal{O}(\epsilon)$ , the perturbation bound (7.1) shows that the computed small singular values and corresponding singular vectors will be inaccurate.

In contrast, the authors in [17] show that, for many structured matrices, the  $i$ th 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  $i$ th *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  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  holds

[15]). 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.



the matrix  $\mathbf{A}$  is pre-pivoted). Assume that the matrix  $\mathbf{P}_1\mathbf{A}\mathbf{P}_2$  may be factored as  $\mathbf{P}_1\mathbf{A}\mathbf{P}_2 = \mathbf{D}_1\mathbf{B}\mathbf{D}_2$ , where  $\mathbf{D}_1$  and  $\mathbf{D}_2$  are diagonal matrices, and that the Householder algorithm, applied to the row-scaled matrix  $\mathbf{C} = \mathbf{D}_1\mathbf{B}$ , produces intermediate matrices  $\mathbf{C}^{(k)}$  with columns  $\mathbf{c}_j^{(k)}$ . Finally, define the quantities  $\mu$ ,  $\nu$ , and  $\kappa$  by

$$7.10) \quad \nu = \max_i \frac{\max_{j,k} |\mathbf{c}_{ij}^{(k)}|}{\max_j |\mathbf{c}_{ij}|}, \quad \mu = \max_k \max_{j \geq k} \frac{\|\mathbf{c}_j^{(k)}(\mathbf{k} : \mathbf{m})\|}{\|\mathbf{c}_k^{(k)}(\mathbf{k} : \mathbf{m})\|}, \quad \kappa = \max_{\substack{1 \leq i \leq n \\ i \leq k \leq n}} \frac{\max_j |\mathbf{c}_{kj}|}{\max_j |\mathbf{c}_{ij}|}.$$

The above quantities measure the extent to which the Householder algorithm preserves the scaling in the intermediate matrices  $\mathbf{A}^{(k)}$ , and are

and

$$7.12) \quad \quad \quad = (\mathbf{M}, \mathbf{n}) \frac{2}{0},$$

where  $(\mathbf{M}, \mathbf{n})$  is proportional to  $\mathbf{M} \cdot \mathbf{n}^{3/2}$ , and  $\frac{2}{0}$  in defined in 7.11). Then we have the following result from [ ]

