

DIFFERENTIAL EQUATIONS FOR REAL-STRUCTURED DEFECTIVITY MEASURES*

P. BUTTÀ[†], N. GUGLIELMI[‡], M. MANETTA[§], AND S. NOSCHESI[†]

Abstract. Let A be a real matrix with all distinct eigenvalues. We propose a new method for the computation of the distance $w_{\mathbb{R}}(A)$ of the matrix A from the set of real defective matrices, i.e., the set of those real matrices with at least one multiple eigenvalue with algebraic multiplicity larger than its geometric multiplicity. For $0 < \varepsilon \leq w_{\mathbb{R}}(A)$, this problem is closely related to the computation of the most ill-conditioned ε -pseudoeigenvalues of A , that is, points in the ε -pseudospectrum of A characterized by the highest condition number. The method we propose couples a system of differential equations on a low-rank manifold which determines the ε -pseudoeigenvalue closest to coalescence, with a fast Newton-like iteration aiming to determine the minimal value ε such that an ε -pseudoeigenvalue becomes defective. The method has a local behavior; this means that, in general, we find upper bounds for $w_{\mathbb{R}}(A)$. However, these bounds usually provide good approximations, in those (simple) cases where we can check this. The methodology can be extended to a structured matrix, where it is required that the distance be computed within some manifold defining the structure of the matrix. In this paper we extensively examine the case of real matrices. As far as we know, there do not exist methods in the literature able to compute such distance.

Key words. pseudospectrum, real-structured pseudospectrum, low-rank differential equations, defective eigenvalue, distance to defectivity, Wilkinson problem

AMS subject classifications. 15A18, 65K05

DOI. 10.1137/140964631

1. Introduction. Let $A \in \mathbb{K}^{n,n}$ be a complex ($\mathbb{K} = \mathbb{C}$) or real ($\mathbb{K} = \mathbb{R}$) matrix with all distinct eigenvalues. We are interested in finding the distance

$$(1.1) \quad w_{\mathbb{K}}(A) = \inf\{\|A - B\|_F : B \in \mathbb{K}^{n,n} \text{ is defective}\},$$

where $\|\cdot\|_F$ is the Frobenius norm (when $\mathbb{K} = \mathbb{C}$, this turns out to be equivalent to the definition in terms of the 2-norm). For a real matrix A , we name $w_{\mathbb{R}}(A)$ as the real-structured distance of A to defectivity. We recall that a matrix is defective if its Jordan canonical form has at least one nondiagonal block associated with an eigenvalue λ . Let us introduce the ε -pseudospectrum of A ,

$$(1.2) \quad \Lambda_{\varepsilon}^{\mathbb{K}}(A) = \{\lambda \in \mathbb{C} : \lambda \in \Lambda(A + E) \text{ for some } E \in \mathbb{K}^{n,n} \text{ with } \|E\|_F \leq \varepsilon\},$$

where we refer to the classical monograph by Trefethen and Embree [TE05] for an extensive exposition. In his well-known paper [Wil65], Wilkinson defined the condition number of a simple eigenvalue as

$$k(\lambda) = \frac{1}{|y^H x|},$$

*Received by the editors April 11, 2014; accepted for publication (in revised form) by D. Kressner March 9, 2015; published electronically May 12, 2015. This research was supported by the Italian MIUR and the INdAM-GNCS.

<http://www.siam.org/journals/simax/36-2/96463.html>

[†]Dipartimento di Matematica, SAPIENZA Università di Roma, 00185 Roma, Italy (butta@mat.uniroma1.it, noschese@mat.uniroma1.it).

[‡]Dipartimento di Ingegneria Scienze Informatiche e Matematica and DEWS, Università degli studi di L'Aquila, and Gran Sasso Science Institute (GSSI), L'Aquila, I-67100 L'Aquila, Italy (guglielm@univaq.it).

[§]Dipartimento di Ingegneria Scienze Informatiche e Matematica, Università degli studi di L'Aquila, I-67100 L'Aquila, Italy (manetta@univaq.it).

where x and y are the right and left eigenvectors such that $\|x\|_2 = \|y\|_2 = 1$. Observing that $k(\lambda) = +\infty$ for a defective eigenvalue since $y^H x = 0$, the search of the closest defective matrix can be pursued by looking for the minimal value ε such that there exists an eigentriple (λ, y, x) satisfying $y^H x = 0$ with $\lambda \in \Lambda(A + E)$ for some E of norm ε . The distance $w_{\mathbb{C}}(A)$ was introduced by Demmel [Dem83] in his well-known Ph.D. thesis and has been studied by several authors, not only for its theoretical interest but also for its practical one (see, e.g., [Ala06] for the bounds on $w_{\mathbb{K}}(A)$ presented in the literature and [AFS13] and references therein for physical applications). An interesting formula for computing $w_{\mathbb{C}}(A)$ has been given by Malyshev [M99]. An appealing algorithm for computing the distance through such a formula has been recently derived by Mengi, Yildirim, and Kiliç [MYK14], based on eigenvalue optimization of Hermitian matrix functions.¹ In the recent and very interesting paper by Alam, Bora, Byers, and Overton [ABBO11], which provides an extensive historical analysis of the problem, the authors have shown that when $\mathbb{K} = \mathbb{C}$, the infimum in (1.1) is actually a minimum. Furthermore, in the same paper, the authors have proposed a computational approach to approximate the nearest defective matrix by a variant of Newton's method. Such a method is well suited to dense problems of moderate size n and—even for a real matrix—computes a nearby defective complex matrix, which is a minimizer for (1.1) in $\mathbb{C}^{n,n}$. A recent fast algorithm has been given in [AFS13], which is based on an extension of the implicit determinant method proposed in [PS05]. This method also applies to the computation of $w_{\mathbb{C}}(A)$ and provides a nearby defective complex matrix.

In this paper we provide a novel approach to the approximation of $w_{\mathbb{K}}(A)$ for both $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$, which may be extended to the approximation of more general structured distances, that is, for example, when restricting the admissible perturbations of A to the set of matrices with a prescribed nonzero pattern. However, a rigorous analysis of general structures is beyond the scope of this paper and we limit the discussion to complex and real perturbations. By the results of Alam and Bora [AB05], we deduce that the distance $w_{\mathbb{C}}(A)$ we compute is the same as that obtained by replacing the Frobenius norm by the 2-norm. Instead, for the real case, the distance we compute is, in general, larger than the corresponding distance in the 2-norm.

1.1. Overview. For a real matrix A we aim to compute the smallest real perturbation which determines a coalescence of two eigenvalues. There are two possible cases: in the first case two real eigenvalues of A coalesce; in the second two genuinely complex eigenvalues coalesce. We need to treat these two cases separately because, as we will show in what follows, the methods we propose present some substantial differences. The first case can be obtained as a by-product of the analogous methodology for the computation of the smallest complex perturbation which determines coalescence of a pair of eigenvalues of a complex matrix. For this reason, we consider this case first, which is interesting in itself. Next we describe our approach. For a given $\varepsilon > 0$ we introduce the quantity

$$(1.3) \quad r(\varepsilon) = \min\{|y^H x| : x \text{ and } y \text{ right/left eigenvectors associated with } z \in \Lambda_{\varepsilon}^{\mathbb{K}}(A)\}$$

and note that every solution to $r(\varepsilon) = 0$ gives an upper bound for $w_{\mathbb{K}}(A)$. We then propose an iterative scheme to compute a minimal solution ε^* to such an equation.

¹The software Eigopt developed by the authors can be downloaded from <http://home.ku.edu.tr/~emengi/software/eigopt.html>.

Every step of the iterative scheme requires the computation of $r(\varepsilon)$. To do this we construct suitable gradient systems of ODEs in the set of the matrices E of norm $\|E\|_F \leq \varepsilon$, whose stationary points \bar{E} are local minimizers and maximizers of $|y^H x|$ as a function of E , where x, y are right and left eigenvectors of $A + E$. A similar approach has been discussed in [GL11, BGN12] for computing pseudospectral abscissas and radii. Next we prove that

$$r(\varepsilon) = \begin{cases} \gamma\sqrt{\varepsilon^* - \varepsilon} + \mathcal{O}((\varepsilon^* - \varepsilon)^{3/2}) & \text{if } \varepsilon < \varepsilon^*, \\ 0 & \text{otherwise.} \end{cases}$$

A root-finding method is applied to $r(\varepsilon)$ to approximate ε^* based on this model.

As $r(\varepsilon)$ is nonsmooth at $\varepsilon = \varepsilon^*$, we consider instead the equation

$$(1.4) \quad r(\varepsilon) = \delta,$$

with a small threshold $\delta \geq 0$. Note that in case $\delta > 0$, we still obtain an estimate for a solution to $r(\varepsilon) = 0$, as a by-product of our method. The substantial differences between the considered two cases are in the system of ODEs; in the first case, the solution evolves within a rank-2 manifold of matrices, while in the second case, within a rank-4 manifold. This implies a different characterization of the stationary points.

The paper is organized as follows.

In section 2 we address the computation of $w_{\mathbb{C}}(A)$ for a complex matrix A (this allows us to treat in the same way the computation of $w_{\mathbb{R}}(A)$ for the case of a real matrix A with two coalescent real eigenvalues). First, we derive a full-rank gradient system of ODEs whose stationary points allow us to compute the solutions of the variational problem (1.3). Then, exploiting the rank-2 property of these stationary points, we project the previous system onto the rank-2 manifold and finally prove that the full-rank and the projected systems have the same stationary points.

In section 3 we consider the second case, where for a real matrix A two genuinely complex eigenvalues coalesce, and we obtain a new full-rank system of ODEs. We first derive a full-rank gradient system of ODEs for this real matrix case, and then we exploit the rank-4 property of the stationary points to project this system onto the rank-4 manifold. Finally, we prove that the full-rank and the projected systems have the same stationary points.

In section 4 we deduce the model for $r(\varepsilon)$ and develop a rapidly convergent root-finding routine to solve (1.4), and then we present the complete algorithm.

Eventually, in section 5 we focus our attention on a few implementation issues, and in section 6 we show some numerical examples.

2. The unstructured case. We denote by $\|A\|_F = \sqrt{\langle A, A \rangle}$ the Frobenius norm of the matrix $A \in \mathbb{C}^{n,n}$, where, for any given $A, B \in \mathbb{C}^{n,n}$, $\langle A, B \rangle = \text{trace}(A^H B)$.

We also need the following definition.

DEFINITION 2.1. *Let M be a singular matrix with a simple zero eigenvalue. The group inverse (reduced resolvent) of M is the unique matrix G satisfying $MG = GM$, $GMG = G$, and $MGM = M$.*

Given a matrix function $M(t) \in \mathbb{C}^{n,n}$, smoothly depending on the real parameter t , we recall results concerning derivatives of right and left eigenvectors $x(t)$ and $y(t)$, respectively, associated with a simple eigenvalue $\lambda(t)$ of $M(t)$. We denote by $G(t)$ the group inverse of $M(t) - \lambda(t)I$ and assume $x(t), y(t)$ smoothly depending on t and such that $\|x(t)\|_2 = \|y(t)\|_2 = 1$. In what follows, we shall often omit the explicit

dependence on t from the notation. The following expressions for the derivatives can be found in [MS88, Theorem 2]:

$$(2.1) \quad \dot{x} = x^H \dot{G} M x x - \dot{G} M x, \quad \dot{y}^H = y^H \dot{M} G y y^H - y^H \dot{M} G.$$

Given $A \in \mathbb{C}^{n,n}$, let λ_0 be a simple eigenvalue of A . We denote by $\mathcal{S}_1 = \{E \in \mathbb{C}^{n,n} : \|E\|_F = 1\}$ the unitary hypersphere in $\mathbb{C}^{n,n}$. By continuity, there exists $\varepsilon_0 > 0$ such that for any $\varepsilon \in [0, \varepsilon_0]$ and $E \in \mathcal{S}_1$ the matrix $A + \varepsilon E$ has a simple eigenvalue λ close to λ_0 .

LEMMA 2.2. *Given $\varepsilon \in (0, \varepsilon_0]$, let*

$$(2.2) \quad S = y y^H G^H + G^H x x^H,$$

where G is the group inverse of $A + \varepsilon E - \lambda I$, whose right and left eigenvectors corresponding to the zero eigenvalue are x and y (recall that we assume $\|x\|_2 = \|y\|_2 = 1$, and note that $y^H x \neq 0$ as λ is simple). Then, for any smooth path $E = E(t) \in \mathcal{S}_1$, we have

$$(2.3) \quad \frac{d}{dt} |y^H x| = \varepsilon |y^H x| \operatorname{Re} \langle \dot{E}, S \rangle.$$

In particular, if $E \not\propto S$, the steepest descent direction is given by

$$(2.4) \quad \arg \min_{\substack{D \in \mathcal{S}_1 \\ \operatorname{Re} \langle D, E \rangle = 0}} \operatorname{Re} \langle D, S \rangle = -\mu (S - \operatorname{Re} \langle E, S \rangle E),$$

where $\mu > 0$ is such that the right-hand side has unit Frobenius norm.

Proof. By (2.1) and using that $Gx = 0$, $y^H G = 0$, we get

$$\begin{aligned} \frac{d}{dt} |y^H x|^2 &= 2 \operatorname{Re} \{ \overline{y^H x} (\dot{y}^H x + y^H \dot{x}) \} = 2\varepsilon \operatorname{Re} \{ x^H y (y^H \dot{E} G y y^H x + y^H x^H G \dot{E} x x) \} \\ &= 2\varepsilon |y^H x|^2 \operatorname{Re} \{ y^H \dot{E} G y + x^H G \dot{E} x \} \\ &= 2\varepsilon |y^H x|^2 \operatorname{Re} \operatorname{trace} (\dot{E}^H y y^H G^H + \dot{E}^H G^H x x^H); \end{aligned}$$

that is, recalling the definition (2.2), $\frac{d}{dt} |y^H x|^2 = 2\varepsilon |y^H x|^2 \operatorname{Re} \langle \dot{E}, S \rangle$, from which (2.3) follows using that $\frac{d}{dt} |y^H x| = \frac{1}{2} |y^H x|^{-1} \frac{d}{dt} |y^H x|^2$. By (2.3) the steepest descent direction is then given by the variational problem in the left-hand side of (2.4). Finally, to prove (2.4), we have to orthogonally project $-S$ onto the linear manifold $\{D : \operatorname{Re} \langle D, E \rangle = 0\}$ and normalize through the coefficient μ . \square

Remark 2.3. A consequence of Lemma 2.2 is that the gradient $\nabla_E y^H x$ of the function $E \mapsto y^H x$ is proportional to S .

Let \mathcal{M}_2 be the manifold of $n \times n$ matrices of rank 2. Any $E \in \mathcal{M}_2$ can be (nonuniquely) represented in the form $E = U T V^H$, where $U, V \in \mathbb{C}^{n,2}$ have orthonormal columns and $T \in \mathbb{C}^{2,2}$ is nonsingular. We will use instead a unique decomposition in the tangent space. Every tangent matrix $\delta E \in T_E \mathcal{M}_2$ is of the form

$$(2.5) \quad \delta E = \delta U T V^H + U \delta T V^H + U T \delta V^H,$$

where $\delta T \in \mathbb{C}^{2,2}$ and $\delta U, \delta V \in \mathbb{C}^{n,2}$ are such that

$$(2.6) \quad U^H \delta U = 0, \quad V^H \delta V = 0.$$

This representation is discussed in [KL07] for the case of real matrices, but the extension to the case of complex matrices is straightforward.

Under the assumptions above, the orthogonal projection of a matrix $Z \in \mathbb{C}^{n,n}$ onto the tangent space $T_E \mathcal{M}_2$ is given by

$$(2.7) \quad P_E(Z) = Z - (I - U U^H) Z (I - V V^H).$$

2.1. Full-rank system of ODEs. Let A , λ_0 , and ε_0 be defined as before. The following theorem characterizes an evolution onto \mathcal{S}_1 governed by the steepest descent direction of $|y^H x|$, where x and y are unit-norm right and left eigenvectors, respectively, associated with the simple eigenvalue λ .

THEOREM 2.4. *Given $\varepsilon \in (0, \varepsilon_0]$, consider the differential system*

$$(2.8) \quad \dot{E} = -S + \operatorname{Re}\langle E, S \rangle E, \quad E \in \mathcal{S}_1,$$

with S defined in (2.2).

(1) *The right-hand side of (2.8) is antiparallel to the projection onto the tangent space $\mathcal{T}_E \mathcal{S}_1$ of the gradient of $|y^H x|$. More precisely,*

$$(2.9) \quad \frac{d}{dt}|y^H x| = -\varepsilon |y^H x| \|S - \operatorname{Re}\langle E, S \rangle E\|_F^2.$$

In particular, $\frac{d}{dt}|y^H x| = 0$ if and only if $\dot{E} = 0$.

(2) *The matrix S defined in (2.2) has rank 2 if $|y^H x| < 1$, whereas $S = 0$ if $|y^H x| = 1$. As a consequence, the equilibria of system (2.8) for which $|y^H x| < 1$ (in particular, the minimizers of $|y^H x|$) are rank-2 matrices.*

Proof. (1) The assertion is an immediate consequence of Lemma 2.2. In particular, the derivative (2.9) is obtained by plugging the right-hand side of (2.8) into (2.3).

(2) By (2.2) the matrix S has rank not greater than 2. Moreover, it has rank less than 2 if and only if $G^H x = cy$ or $Gy = cx$ for some $c \in \mathbb{C}$. We claim the constant c must vanish in both cases. Indeed, since $Gx = 0$ and $y^H G = 0$, we have $\bar{c}y^H x = x^H Gx = 0$ or $cy^H x = y^H Gy = 0$; hence $c = 0$ as $y^H x \neq 0$. Therefore, the rank of S is less than 2 if and only if $G^H x = 0$ or $Gy = 0$. As G has rank $n - 1$, both conditions are equivalent to having $y = e^{i\theta}x$ for some $\theta \in [0, 2\pi]$, that is to say, $S = 0$. Finally, the assertion concerning the equilibria comes from the fact that $S = \operatorname{Re}\langle E, S \rangle E$ at a stationary point. \square

THEOREM 2.5. *Let x_0 and y_0 be unit-norm right and left eigenvectors, respectively, associated with the simple eigenvalue λ_0 of the matrix A . If $|y_0^H x_0| < 1$, then, for ε small enough, the system (2.8) has only two stationary points, which correspond to the minimum and the maximum of $|y^H x|$ on \mathcal{S}_1 , respectively.*

Proof. As $|y_0^H x_0| < 1$, by item (2) of Theorem 2.4 we have that $S = S_0 + Q(E, \varepsilon)$ with S_0 a nonzero constant matrix and

$$(2.10) \quad \max_{E \in \mathcal{S}_1} \|Q(E, \varepsilon)\|_F = \mathcal{O}(\varepsilon).$$

The equation for the equilibria reads as $F(E, \varepsilon) = 0$, where

$$F(E, \varepsilon) = -S_0 + \operatorname{Re}\langle E, S_0 \rangle E - Q(E, \varepsilon) + \operatorname{Re}\langle E, Q(E, \varepsilon) \rangle E.$$

It is readily seen that $F(E, 0) = 0$ if and only if $E = E_\pm^0 = \pm S_0 / \|S_0\|_F$. Moreover, the Jacobian matrix of $F(E, \varepsilon)$ with respect to E at the point $(E_\pm^0, 0)$ is given by the linear operator $\mathcal{L}_\pm: \mathbb{C}^{n,n} \rightarrow \mathbb{C}^{n,n}$ such that

$$\mathcal{L}_\pm B = \operatorname{Re}\langle B, S_0 \rangle E_\pm^0 + \operatorname{Re}\langle E_\pm^0, S_0 \rangle B, \quad B \in \mathbb{C}^{n,n}.$$

We shall prove below that \mathcal{L}_\pm is invertible. By the implicit function theorem, there are $\varepsilon_1 > 0$, $r > 0$, and E_\pm^ε such that $F(E, \varepsilon) = 0$ for $\varepsilon \in [0, \varepsilon_1]$ and $\|E - E_\pm^0\|_F < r$ if and only if $E = E_\pm^\varepsilon$. On the other hand, by (2.10), if $F(E, \varepsilon) = 0$, then $\|E - E_\pm^0\| = \mathcal{O}(\varepsilon)$

or $\|E - E_-^0\| = \mathcal{O}(\varepsilon)$. We conclude that E_{\pm}^{ε} are the unique equilibria on the whole hypersphere \mathcal{S}_1 for ε small enough. Clearly, E_+^{ε} (resp., E_-^{ε}) is the maximizer (resp., minimizer) of $|y^H x|$.

To prove the invertibility of \mathcal{L}_{\pm} we observe that, recalling $E_{\pm}^0 = \pm S_0 / \|S_0\|_F$, the equation $\mathcal{L}_{\pm} B = 0$ reads as $B = -(\operatorname{Re}\langle B, S_0 \rangle / \|S_0\|_F^2) S_0$, which gives $\operatorname{Re}\langle B, S_0 \rangle = 0$ and therefore $B = 0$. \square

2.2. Projected system of ODEs. By Theorem 2.4, the minimizers of $|y^H x|$ over \mathcal{S}_1 are rank-2 matrices. This suggests using a rank-2 dynamics, obtained as a suitable projection of (2.8) onto \mathcal{M}_2 .

THEOREM 2.6 (the projected system). *Given $\varepsilon \in (0, \varepsilon_0]$, consider the differential system*

$$(2.11) \quad \dot{E} = -P_E(S) + \operatorname{Re}\langle E, S \rangle E, \quad E \in \mathcal{S}_1 \cap \mathcal{M}_2,$$

where the orthogonal projection P_E onto $\mathcal{T}_E \mathcal{M}_2$ is defined in (2.7). Then, the right-hand side of (2.11) is antiparallel to the projection onto the tangent space $\mathcal{T}_E \mathcal{S}_1 \cap \mathcal{T}_E \mathcal{M}_2$ of the gradient of $|y^H x|$. More precisely,

$$(2.12) \quad \frac{d}{dt} |y^H x| = -\varepsilon |y^H x| \|P_E(S) - \operatorname{Re}\langle E, S \rangle E\|_F^2.$$

In particular, $\frac{d}{dt} |y^H x| = 0$ if and only if $\dot{E} = 0$.

Proof. We remark that the definition is well posed, i.e., $\dot{E} \in \mathcal{T}_E \mathcal{S}_1 \cap \mathcal{T}_E \mathcal{M}_2$. Indeed, $\dot{E} \in \mathcal{T}_E \mathcal{M}_2$ as $E = P_E(E)$, and

$$\operatorname{Re}\langle \dot{E}, E \rangle = -\operatorname{Re}\langle E, P_E(S) \rangle + \operatorname{Re}\langle E, S \rangle = -\operatorname{Re}\langle E, P_E(S) \rangle + \operatorname{Re}\langle P_E(E), P_E(S) \rangle = 0,$$

as P_E is an orthogonal projection. We next observe that, by (2.3), the steepest descent direction of $|y^H x|$ is given by the variational problem

$$\arg \min_{\substack{\|D\|_F=1 \\ D \in \mathcal{T}_E \mathcal{S}_1 \cap \mathcal{T}_E \mathcal{M}_2}} \operatorname{Re}\langle D, S \rangle.$$

Since $\operatorname{Re}\langle D, S \rangle = \operatorname{Re}\langle D, P_E(S) \rangle$ for any $D \in \mathcal{T}_E \mathcal{M}_2$, the solution to this problem is

$$D = -\frac{P_E(S) - \operatorname{Re}\langle E, S \rangle E}{\|P_E(S) - \operatorname{Re}\langle E, S \rangle E\|_F}.$$

This proves the first assertion, while the derivative (2.12) is obtained by plugging the right-hand side of (2.11) into (2.3) and using $\frac{d}{dt} |y^H x| = \frac{1}{2} |y^H x|^{-1} \frac{d}{dt} |y^H x|^2$. \square

Remark 2.7. It is worthwhile to note that along the solutions to (2.8) or (2.11) we have $\frac{d}{dt} (\operatorname{Im}(y^H x)) = -\varepsilon |y^H x| \|\dot{E}\|_F^2$, so that the condition $\operatorname{Im}(y^H x) = 0$ is preserved by the dynamics.

For our purposes, we rewrite the projected system in terms of the representations of \mathcal{M}_2 and $\mathcal{T}_E \mathcal{M}_2$ discussed right before section 2.1. In terms of the decomposition $E = UTV^H$ the system (2.11) can be expressed as

$$\dot{U}TV^H + U\dot{T}V^H + UT\dot{V}^H = -SVV^H - UU^H S + UU^H SVV^H + \operatorname{Re}\langle E, S \rangle E.$$

By using that $U^H \dot{U} = 0$, $\dot{V}^H V = 0$, we obtain

$$\begin{cases} \dot{T} = -U^H S V + \operatorname{Re}\langle E, S \rangle T, \\ \dot{U}T + U\dot{T} = -S V + \operatorname{Re}\langle E, S \rangle U T, \\ \dot{T}V^H + T\dot{V}^H = -U^H S + \operatorname{Re}\langle E, S \rangle T V^H. \end{cases}$$

By introducing the notation

$$(2.13) \quad p = U^H y, \quad q = V^H x, \quad r = U^H G^H x, \quad s = V^H G y,$$

noting that $U^H S V = p s^H + r q^H$ and that $\text{Re}\langle E, S \rangle = \text{Re}(s^H T^H p + q^H T^H r)$, by simple algebraic manipulation we finally obtain the following system of ODEs:

$$(2.14) \quad \begin{cases} \dot{T} = -(ps^H + rq^H) + \text{Re}(s^H T^H p + q^H T^H r) T, \\ \dot{U} = -((y - Up)s^H + (G^H x - Ur)q^H) T^{-1}, \\ \dot{V} = -((Gy - Vs)p^H + (x - Vq)r^H) T^{-H}. \end{cases}$$

2.3. Stationary points of the projected ODEs. We start by providing a characterizing result for stationary points of system (2.11). By (2.14) and nonsingularity of T , these are given by the solutions to

$$(2.15) \quad \begin{cases} -(ps^H + rq^H) + \text{Re}(s^H T^H p + q^H T^H r) T = 0, \\ (y - Up)s^H + (G^H x - Ur)q^H = 0, \\ (Gy - Vs)p^H + (x - Vq)r^H = 0. \end{cases}$$

LEMMA 2.8. *Given $\varepsilon \in (0, \varepsilon_0]$, assume $E \in \mathcal{S}_1 \cap \mathcal{M}_2$ is a stationary point of (2.11) (or, equivalently, of (2.14)) such that λ is not an eigenvalue of A and $|y^H x| < 1$. Then $P_E(S) \neq 0$.*

Proof. We argue by contradiction and assume that $P_E(S) = 0$. Then $S = (I - UU^H) S (I - VV^H)$, which implies that

$$\begin{aligned} U^H S = 0 & \implies U^H S y = r(x^H y) = 0, \\ S V = 0 & \implies V^H S^H x = s(y^H x) = 0, \end{aligned}$$

whence, as $|y^H x| \neq 0$ onto $\mathcal{S}_1 \cap \mathcal{M}_2$, $r = 0$, $s = 0$. By inserting these formulas into (2.15) we would obtain

$$(2.16) \quad \begin{cases} 0 = 0, \\ G^H x x^H V = 0, \\ G y y^H U = 0. \end{cases}$$

Since $|y^H x| < 1$ implies $G^H x \neq 0$ and $G y \neq 0$, the previous relations imply that $V^H x = 0$ and $y^H U = 0$, so that $E x = 0$ and $y^H E = 0$. As a consequence, λ would be an eigenvalue of A , which contradicts the assumptions. This means that $P_E(S) \neq 0$. The proof is complete. \square

At a stationary point the equation $P_E(S) = \text{Re}\langle E, S \rangle E$ reads as

$$(2.17) \quad E = \mu P_E(S)$$

for some nonzero $\mu \in \mathbb{R}$. Indeed, now we prove that at stationary point $E = \mu S$.

THEOREM 2.9. *Given $\varepsilon \in (0, \varepsilon_0]$, assume $E = U T V^H \in \mathcal{S}_1 \cap \mathcal{M}_2$ is a stationary point of (2.11) (or, equivalently, of (2.14)) such that λ is not an eigenvalue of A and $|y^H x| < 1$. Then $E = \mu S$ holds for some real μ .*

Proof. Since, by (2.17),

$$\text{Re}(y^H E G y + x^H G E x) = \text{Re}\langle S, E \rangle = \text{Re}\langle \mu P_E(S), P_E(S) \rangle = \frac{1}{\mu} \langle E, E \rangle \neq 0,$$

assuming $E = UTV^H$, we get

$$(2.18) \quad \operatorname{Re} (y^H UTV^H Gy + x^H GUTV^H x) \neq 0 .$$

Recalling (2.7), in order to prove the theorem we have to show that the matrix

$$B = (I - UU^H) S (I - VV^H) = (I - UU^H) (yy^H G^H + G^H xx^H) (I - VV^H)$$

is zero. Note that B is different from S , as we have shown in the proof of Lemma 2.8. By (2.15),

$$(2.19) \quad (y - Up)s^H + (G^H x - Ur)q^H = 0 ,$$

$$(2.20) \quad (Gy - Vs)p^H + (x - Vq)r^H = 0 .$$

The assumption that λ is not an eigenvalue of A implies $Ex \neq 0$ and $y^H E \neq 0$, which means $p \neq 0$ and $q \neq 0$.

If (2.19) is fulfilled, then at least one of the following three cases holds true:

$$(1a) \quad s = 0, \quad G^H x - Ur = 0.$$

$$(1b) \quad y = Up, \quad G^H x - Ur = 0.$$

$$(1c) \quad s \propto q \quad (\implies V^H Gy \propto V^H x).$$

Analogously, if (2.20) is fulfilled, then at least one of the following three cases holds true:

$$(2a) \quad r = 0, \quad Gy - Vs = 0.$$

$$(2b) \quad x = Vq, \quad Gy - Vs = 0.$$

$$(2c) \quad p \propto r \quad (\implies U^H G^H x \propto U^H y).$$

If case (1b) holds, then

$$y = UU^H y \quad \text{and} \quad G^H x = UU^H G^H x \quad \implies \quad S = UU^H S ,$$

and thus $B = 0$.

If case (2b) holds, then

$$x = VV^H x \quad \text{and} \quad Gy = VV^H Gy \quad \implies \quad S = SVV^H ,$$

and thus $B = 0$.

Assume now that $B \neq 0$, which excludes (1b) and (2b), so that four cases may occur:

(1) If (1a) and (2a) hold, then $s = V^H Gy = 0$ and $r = U^H G^H x = 0$, which contradicts (2.18).

(2) If (1c) and (2c) hold, then $ps^H + rq^H \propto U^H yx^H V$, and from the first equation of (2.15), it follows that T has rank 1, which contradicts the invertibility of T .

(3) If (1a) and (2c) hold, then, as $s = 0$, again by (2.15), $T \propto rq^H$ has rank 1.

(4) If (1c) and (2a) hold, then, as $r = 0$, again by (2.15), $T \propto ps^H$ has rank 1.

This completes the proof. \square

To summarize, if λ is not an eigenvalue of A , the stationary points of the projected and the unprojected ODEs coincide (recall that if $|y^H x| = 1$, this is obvious as $S = 0$; see Theorem 2.4). Moreover, since E is proportional to S at such points, by the same arguments leading to Theorem 2.5 we obtain the following result.

COROLLARY 2.10. *Under the same assumptions of Theorem 2.9, for sufficiently small ε the projected ODE (2.11) has only two stationary points.*

3. The real-structured case. We now assume that the matrix A is real, and we restrict the perturbations E to be real as well. To our knowledge there are no methods to compute the most ill-conditioned eigenvalue in the real ε -pseudospectrum,

$$(3.1) \quad \Lambda_\varepsilon^{\mathbb{R}}(A) = \{ \lambda \in \mathbb{C} : \lambda \in \Lambda(A + E) \text{ for some } E \in \mathbb{R}^{n,n} \text{ with } \|E\|_F \leq \varepsilon \},$$

that is, the eigenvalue of $A + \varepsilon E$ which corresponds to $\min y^H x$.

We denote by \mathcal{R}_1 the unitary hypersphere in $\mathbb{R}^{n,n}$. We next fix $\tilde{\varepsilon}_0 > 0$ such that for any $\varepsilon \in [0, \tilde{\varepsilon}_0]$ and $E \in \mathcal{R}_1$ the matrix $A + \varepsilon E$ has a simple eigenvalue λ close to λ_0 . The same reasoning as in the proof of Lemma 2.2 gives, for a smooth path $E(t) \in \mathcal{R}_1$,

$$(3.2) \quad \frac{d}{dt} |y^H x| = \varepsilon |y^H x| \langle \dot{E}, \text{Re}(S) \rangle,$$

where S is defined as in (2.2), from which the steepest descent direction is given by

$$\arg \min_{\substack{D \in \mathcal{R}_1 \\ \langle D, E \rangle = 0}} \langle D, \text{Re}(S) \rangle = -\mu (\text{Re}(S) - \langle E, \text{Re}(S) \rangle E),$$

with μ the normalization constant. Note that the matrix $\text{Re}(S)$ has rank at most 4, being the real part of the complex matrix S , whose rank is at most 2. Clearly, for a real eigenvalue, S is real and the situation is identical to the one considered in the unstructured case $\mathbb{K} = \mathbb{C}$; thus the peculiar difference arises when we consider nonreal eigenvalues.

Let \mathcal{M}_4 be the manifold of the real matrices of rank 4. The matrix representations both in \mathcal{M}_4 and in the tangent space $T_E \mathcal{M}_4$ are analogous to (2.5), (2.6), provided that $U, V \in \mathbb{R}^{n,4}$ have orthonormal columns and $T \in \mathbb{R}^{4,4}$ is nonsingular; see [KL07, sect. 2.1]. More precisely, any rank-4 matrix of order n can be written in the form

$$(3.3) \quad E = UTV^T,$$

with, now, U and V such that $U^T U = I_4$ and $V^T V = I_4$, where I_4 is the identity matrix of order 4, and $T \in \mathbb{R}^{4,4}$ is nonsingular. As before, since this decomposition is not unique, we use a unique decomposition on the tangent space. For a given choice of U, V, T any matrix $\delta E \in T_E \mathcal{M}_4$ can be uniquely written as

$$\delta E = \delta UTV^T + U\delta TV^T + UT\delta V^T,$$

with $U^T \delta U = 0$, $V^T \delta V = 0$. Accordingly, the orthogonal projection of a matrix $Z \in \mathbb{R}^{n,n}$ onto the tangent space $T_E \mathcal{M}_4$ is defined by

$$(3.4) \quad \tilde{P}_E(Z) = Z - (I - UU^T)Z(I - VV^T),$$

which differs from (2.7) since U and V are now real matrices.

3.1. Full-rank system of ODEs. Given $\varepsilon \in (0, \tilde{\varepsilon}_0]$, the role of the differential system in (2.8) is now played by

$$(3.5) \quad \dot{E} = -\text{Re}(S) + \langle E, \text{Re}(S) \rangle E, \quad E \in \mathcal{R}_1.$$

More precisely, the right-hand side of (3.5) is antiparallel to the projection onto the tangent space $T_E \mathcal{R}_1$ of the gradient of $|y^H x|$ and

$$(3.6) \quad \frac{d}{dt} |y^H x| = -\varepsilon |y^H x| \|\text{Re}(S) - \langle E, \text{Re}(S) \rangle E\|_F^2.$$

The proof of Theorem 2.5 is easily adapted to the real case. Therefore, under the same hypothesis, the system (3.5) has only two stationary points, which correspond to the minimum and the maximum of $|y^H x|$ on \mathcal{R}_1 , respectively.

A natural question is concerned with the possibility of $\text{Re}(S)$ vanishing. We have the following result concerning the matrix $\text{Re}(S)$.

THEOREM 3.1. *Assume that the matrix B is real and has a pair of simple complex conjugate eigenvalues λ and $\bar{\lambda}$. Let x and y be its right and left eigenvectors associated with λ such that $y^H x < 1$. Let G be the group inverse of $B - \lambda I$ and S be given by (2.2). Then $\text{Re}(S)$ is different from zero.*

Proof. First observe that the eigenvectors x and y are necessarily genuinely complex vectors, that is, $\text{Re}(x) \neq 0$, $\text{Im}(x) \neq 0$, $\text{Re}(y) \neq 0$, and $\text{Im}(y) \neq 0$. Let us denote the range of a matrix M as $\mathcal{R}(M)$. By definition of S (see (2.2)) we have $\mathcal{R}(S) = \text{span}(y, G^H x)$. Assume that S is purely imaginary, that is, $\text{Re}(S) = 0$. Under this assumption, recalling that S is a rank-2 matrix, we have that $v \in \mathcal{R}(S)$ implies $\bar{v} \in \mathcal{R}(S)$, and therefore

$$\mathcal{R}(S) = \text{span}(y, \bar{y}, G^H x, \overline{G^H x})$$

has dimension 2. Since y and $G^H x$ are linearly independent, we get $\bar{y} = \alpha y + \beta G^H x$. A left premultiplication by x^H gives

$$x^H \bar{y} = \alpha x^H y + \beta x^H G^H x \implies \alpha = 0,$$

which is due to the fact that (i) $x^H \bar{y} = 0$ by well-known biorthogonality of right and left eigenvectors, (ii) $Gx = 0$, a property of the group inverse G , and (iii) $x^H y \neq 0$ by simplicity of λ . This implies $\bar{y} \propto G^H x$. Analogously, we have that

$$\mathcal{R}(S^T) = \text{span}(x, \bar{x}, Gy, \overline{Gy})$$

has dimension 2. Proceeding in the same way, we obtain that $\bar{x} \propto Gy$.

Now recall that (see [MS88]) a vector v is an eigenvector of $B - \lambda I$ corresponding to the eigenvalue μ if and only if v is an eigenvector of G corresponding to the eigenvalue $\mu^\#$, where $\mu^\# = 1/\mu$ if $\mu \neq 0$ and $\mu^\# = 0$ if $\mu = 0$. Since $\bar{y} \propto G^H x$ we have

$$(3.7) \quad Gy = \gamma \bar{x}, \quad G\bar{x} = i \frac{1}{2\text{Im}(\lambda)} \bar{x},$$

$$(3.8) \quad G^H x = \eta \bar{y}, \quad G^H \bar{y} = -i \frac{1}{2\text{Im}(\lambda)} \bar{y},$$

with $\gamma \neq 0$ and $\eta \neq 0$.

Since x and y are the only vectors in the kernels of G and G^H , respectively, we deduce by (3.7) and (3.8) that

$$x \propto \frac{1}{\gamma} y + 2i \text{Im}(\lambda) \bar{x}, \quad y \propto \frac{1}{\eta} x - 2i \text{Im}(\lambda) \bar{y},$$

implying that $\text{span}(x, \bar{x}) = \text{span}(y, \bar{y})$. This implies, by biorthogonality of right and left eigenvectors, that the set $X = \{x, \bar{x}\}$ is orthogonal to the set \widehat{X} consisting of the remaining $n - 2$ right eigenvectors of B , and similarly the set $Y = \{y, \bar{y}\}$ is orthogonal to the set \widehat{Y} of the remaining $n - 2$ left eigenvectors of B . Note that $\text{span}(X)$ and $\text{span}(\widehat{X})$ are right-invariant subspaces of both B and B^T .

Denote by $\text{Orth}(C)$ a real orthonormal basis for the range of C , and define $U = \text{Orth}(X) \in \mathbb{R}^{n,2}$ (as determined by the procedure to get the Schur canonical form of B) and $V = \text{Orth}(\widehat{X}) \in \mathbb{R}^{n,n-2}$. Set $Q = (U, V) \in \mathbb{R}^{n,n}$, which implies Q is an orthogonal matrix. Now consider the similarity transformation associated with Q ,

$$\tilde{B} = Q^T B Q = \begin{pmatrix} B_1 & O^T \\ O & B_2 \end{pmatrix},$$

where O stands for the $((n - 2) \times 2)$ -dimensional zero matrix, $B_1 \in \mathbb{R}^{2,2}$, and $B_2 \in \mathbb{R}^{n-2,n-2}$. This means that \tilde{B} is block-diagonal and the matrix

$$(3.9) \quad B_1 = \begin{pmatrix} \varrho & \sigma \\ -\tau & \varrho \end{pmatrix}$$

is such that $\varrho = \text{Re}(\lambda)$ and $\sigma > 0, \tau > 0$ with $\sigma\tau = \text{Im}(\lambda)^2 > 0$ so that B_1 has eigenvalues λ and $\bar{\lambda}$. If $\sigma = \tau$, then B_1 is normal, which implies that the pair of right and left eigenvectors associated with λ , say \tilde{x}, \tilde{y} (scaled to have unit 2-norm and real and positive Hermitian scalar product), is such that $\tilde{y}^H \tilde{x} = 1$. Since $x = Q\tilde{x}$ and $y = Q\tilde{y}$, the orthogonality of Q implies $y^H x = 1$, which gives a contradiction. As a consequence we can assume $\tau \neq \sigma$. By the properties of the group inverse we have

$$\tilde{G} = Q^T G Q = \begin{pmatrix} G_1 & O^T \\ O & G_2 \end{pmatrix},$$

where G_1 is the group inverse of $B_1 - \lambda I$ and G_2 is the inverse of $B_2 - \lambda I$, which is nonsingular. It is easy to verify the following formula for the group inverse by simply checking the three conditions in Definition 2.1:

$$G_1 = \begin{pmatrix} \frac{i}{4\sqrt{\sigma\tau}} & -\frac{1}{4\tau} \\ \frac{1}{4\sigma} & \frac{i}{4\sqrt{\sigma\tau}} \end{pmatrix}.$$

It follows that also $Q^T S Q$ is block triangular such that we can write

$$(3.10) \quad \tilde{S} = Q^T S Q = \begin{pmatrix} S_1 & O^T \\ O & S_2 \end{pmatrix},$$

with

$$(3.11) \quad S_1 = \tilde{y}_1 \tilde{y}_1^H G_1^H + G_1^H \tilde{x}_1 \tilde{x}_1^H,$$

where, denoting by e_1, e_2 the first two vectors of the canonical basis, $\tilde{y}_1 \in \mathbb{C}^2$ and $\tilde{x}_1 \in \mathbb{C}^2$ are the projections on $\text{span}(e_1, e_2)$ of the eigenvectors \tilde{y}, \tilde{x} of B associated with λ , that is, $\tilde{x} = Q^T x$ and $\tilde{y} = Q^T y$,

$$\tilde{x} = \nu_x^{-1} \begin{pmatrix} i\frac{\sqrt{\sigma}}{\sqrt{\tau}} & 1 & 0 & \dots & 0 \end{pmatrix}^T, \quad \tilde{y} = \nu_y^{-1} \begin{pmatrix} -i\frac{\sqrt{\tau}}{\sqrt{\sigma}} & 1 & 0 & \dots & 0 \end{pmatrix}^T,$$

where $\nu_x = \sqrt{\frac{\sigma}{\tau} + 1}$ and $\nu_y = \sqrt{\frac{\tau}{\sigma} + 1}$ are such that $\|\tilde{x}\| = \|\tilde{y}\| = 1$ and $\tilde{x}^H \tilde{y} \in \mathbb{R}^+$. Finally, we obtain

$$S_1 = \begin{pmatrix} 0 & \frac{\tau - \sigma}{2\sigma(\sigma + \tau)} \\ \frac{\tau - \sigma}{2\tau(\sigma + \tau)} & 0 \end{pmatrix},$$

which is real and cannot vanish due to the fact that $\sigma \neq \tau$.

Recalling that Q is real, if S were purely imaginary, then S_1 would be purely imaginary as well, which gives a contradiction.

We remark that it can also be shown that $S_2 = 0$ in (3.10). \square

Remark 3.2. Note that when A is real and we compute S for a complex eigenvalue, it can occur that S is real. The simplest example is given by the matrix (3.9).

According to Theorem 3.1 we have that $\operatorname{Re}(S) \neq 0$ for every path of genuinely complex eigenvalues. Based on this result we can characterize stationary points of (3.5) to have rank (at most) 4.

This suggests projecting the ODE onto the rank-4 manifold of real matrices.

3.2. Projected system of ODEs. The following theorem characterizes the projected system onto the tangent space $\mathcal{T}_E\mathcal{M}_4$.

THEOREM 3.3 (the real projected system). *Given $\varepsilon \in (0, \tilde{\varepsilon}_0]$, consider the differential system*

$$(3.12) \quad \dot{E} = -\tilde{P}_E(\operatorname{Re}(S)) + \langle E, \operatorname{Re}(S) \rangle E, \quad E \in \mathcal{R}_1 \cap \mathcal{M}_4,$$

where the orthogonal projection \tilde{P}_E onto $\mathcal{T}_E\mathcal{M}_4$ is defined in (3.4). Then, the right-hand side of (3.12) is antiparallel to the projection onto the tangent space $\mathcal{T}_E\mathcal{R}_1 \cap \mathcal{T}_E\mathcal{M}_4$ of the gradient of $|y^H x|$. More precisely,

$$(3.13) \quad \frac{d}{dt}|y^H x| = -\varepsilon |y^H x| \|\tilde{P}_E(S) - \langle E, \operatorname{Re}(S) \rangle E\|_F^2.$$

Proof. By (3.2), the steepest descent direction is given by the variational problem

$$\arg \min_{\substack{\|D\|_F=1 \\ D \in \mathcal{T}_E\mathcal{R}_1 \cap \mathcal{T}_E\mathcal{M}_4}} \langle D, \operatorname{Re}(S) \rangle.$$

Since $\langle D, \operatorname{Re}(S) \rangle = \langle D, \tilde{P}_E(\operatorname{Re}(S)) \rangle$ for any $D \in \mathcal{T}_E\mathcal{M}_4$, this problem has solution

$$D = -\frac{\tilde{P}_E(\operatorname{Re}(S)) - \langle E, \operatorname{Re}(S) \rangle E}{\|\tilde{P}_E(\operatorname{Re}(S)) - \langle E, \operatorname{Re}(S) \rangle E\|_F}. \quad \square$$

Based on the previous theorem, our aim is to write a system of differential equations on the manifold of rank-4 matrices for the projected system. Given $\varepsilon \in (0, \tilde{\varepsilon}_0]$, consider the differential system (3.12). The reasoning leading to the derivation of system (2.14) can be adapted to the real case to obtain the differential equation in a form that uses the factors in the decomposition (3.3) rather than the full $n \times n$ matrix E . More precisely, by plugging $E = UTV^T$ into (3.12) and using the fact that $U^T \dot{U} = 0$ and $\dot{V}^T V = 0$, we obtain

$$\begin{cases} \dot{T} = -U^T \operatorname{Re}(S)V + \langle E, \operatorname{Re}(S) \rangle T, \\ \dot{U}T + U\dot{T} = -\operatorname{Re}(S)V + \langle E, \operatorname{Re}(S) \rangle UT, \\ \dot{T}V^T + T\dot{V}^T = -U^T \operatorname{Re}(S) + \langle E, \operatorname{Re}(S) \rangle TV^T. \end{cases}$$

Now, writing

$$(3.14) \quad \operatorname{Re}(S) = \operatorname{Re}(yy^H G^H + G^H xx^H) = YW^T + ZX^T,$$

where

$$(3.15) \quad \begin{cases} X = (\operatorname{Re}(x), \operatorname{Im}(x)), & Y = (\operatorname{Re}(y), \operatorname{Im}(y)), \\ W = (\operatorname{Re}(Gy), \operatorname{Im}(Gy)), & Z = (\operatorname{Re}(G^H x), \operatorname{Im}(G^H x)) \end{cases}$$

are matrices in $\mathbb{R}^{n,2}$, we obtain the following system of ODEs:

$$(3.16) \quad \begin{cases} \dot{T} = -(PN^T + RQ^T) + \operatorname{trace}(N^T T^T P + Q^T T^T R) T, \\ \dot{U} = -((Y - UP)N^T + (Z - UR)Q^T)T^{-1}, \\ \dot{V} = -((W - VN)P^T + (X - VQ)R^T)T^{-T}, \end{cases}$$

where

$$(3.17) \quad P = U^T Y, \quad Q = V^T X, \quad R = U^T Z, \quad N = V^T W$$

are matrices in $\mathbb{R}^{4,2}$.

We now provide a characterizing result for stationary points of (3.12).

LEMMA 3.4. *Given $\varepsilon \in (0, \tilde{\varepsilon}_0]$, assume $E \in \mathcal{R}_1 \cap \mathcal{M}_4$ is a stationary point of (3.12) such that λ is not an eigenvalue of A and $|x^H y| < 1$. Then $\tilde{P}_E(\operatorname{Re}(S)) \neq 0$.*

Proof. Assume by contradiction that $\tilde{P}_E(\operatorname{Re}(S)) = 0$; then we have

$$(3.18) \quad \operatorname{Re}(S) = (I - UU^T) \operatorname{Re}(S) (I - VV^T).$$

By using (3.14), (3.18) can be rewritten as

$$(3.19) \quad YW^T + ZX^T = (I - UU^T)(YW^T + ZX^T)(I - VV^T).$$

Multiplying by U^T to the left and V to the right we obtain $U^T YW^T V + U^T Z X^T V = 0$. By Theorem 3.1 we have that $YW^T + ZX^T$ cannot be zero; therefore we have the following possibilities:

$$(3.20) \quad U^T Y = 0 \quad \text{and} \quad U^T Z = 0,$$

$$(3.21) \quad U^T Y = 0 \quad \text{and} \quad X^T V = 0,$$

$$(3.22) \quad W^T V = 0 \quad \text{and} \quad X^T V = 0,$$

$$(3.23) \quad W^T V = 0 \quad \text{and} \quad U^T Z = 0.$$

Each of the conditions (3.20), (3.21), and (3.22) implies that λ is an eigenvalue of A , and this contradicts the hypothesis. Instead, by (3.17), condition (3.23) means $N = 0$ and $R = 0$, whence, replacing these values in (3.16), we get

$$(3.24) \quad \begin{cases} \dot{T} = 0, \\ \dot{U} = -ZQ^T T^{-1}, \\ \dot{V} = -WP^T T^{-T}. \end{cases}$$

In order to have $\dot{E} = 0$ we need $\dot{U} = 0$ and $\dot{V} = 0$. Since $|x^H y| < 1$ implies $G^H x \neq 0$ and $Gy \neq 0$, we have $Z \neq 0$ and $W \neq 0$. Moreover, by a rotation $(x, y) \rightarrow (e^{i\phi} x, e^{i\phi} y)$ we can assume both $G^H x$ and Gy are genuine complex vectors. As a consequence, the previous relations imply $Q = V^T X = 0$ and $P^T = Y^T U = 0$, so that $Ex = 0$ and $y^H E = 0$. Therefore λ would be an eigenvalue of A , which contradicts the assumptions. This means that $\tilde{P}_E(\operatorname{Re}(S)) \neq 0$. \square

Consequently, at a stationary point of (3.12) we have $E = \mu \tilde{P}_E(\operatorname{Re}(S))$ for some real $\mu \neq 0$.

3.3. Stationary points of the projected system of ODEs. In order to study stationary points of (3.16) we define

$$(3.25) \quad B = (I - UU^T) \operatorname{Re}(S) (I - VV^T) = (I - UU^T) (YW^T + ZX^T) (I - VV^T) .$$

THEOREM 3.5. *Given $\varepsilon \in (0, \tilde{\varepsilon}_0)$, assume $E = UTV^T \in \mathcal{R}_1 \cap \mathcal{M}_4$ is a stationary point of (3.12) (or, equivalently, of (3.16)) such that λ is not an eigenvalue of A and $|y^H x| < 1$. Then it holds that $E = \mu \operatorname{Re}(S)$ for some real μ .*

Proof. To prove that $\tilde{P}_E(\operatorname{Re}(S)) = \operatorname{Re}(S)$ we have to show that the matrix B in (3.25) is zero. Assume (T, U, V) is a stationary point of (3.16). The first equation yields

$$(3.26) \quad PN^T + RQ^T = cT \quad \implies \quad U^T \operatorname{Re}(S)V = cT ,$$

where c is a nonzero constant.

By the assumptions, we have that $Q \neq 0$ and $P \neq 0$; otherwise we would have either $Ex = 0$ or $y^H E = 0$, which would imply that λ is an eigenvalue of A . From the nonsingularity of T , the last two equations in (3.16) imply that at the stationary point we have

$$(3.27) \quad (Y - UP)N^T + (Z - UR)Q^T = 0 ,$$

$$(3.28) \quad (W - VN)P^T + (X - VQ)R^T = 0 .$$

If (3.27) is fulfilled, at least one of the following three cases holds true:

- (1a) $N = 0$ and $Z - UR = 0$.
- (1b) $Y - UP = 0$ and $Z - UR = 0$.
- (1c) $Y - UP \neq 0$, $N \neq 0$, $Z - UR \neq 0$.

Similarly, if (3.28) is fulfilled, at least one of the following three cases holds true:

- (2a) $R = 0$ and $W - VN = 0$.
- (2b) $V - WN = 0$ and $X - VQ = 0$.
- (2c) $W - VN \neq 0$, $X - VQ \neq 0$, $R \neq 0$.

Case (1c) implies $\operatorname{span}(N) = \operatorname{span}(Q)$. From (3.26) we gather that T has rank 2, which means the 4×4 matrix T is singular, which can be excluded.

Case (2c) implies $\operatorname{span}(P) = \operatorname{span}(R)$. From (3.26) we gather that T has rank 2, which means the 4×4 matrix T is singular, which can be excluded.

Case (1b): the first equality $Y - UP = 0$ implies $Y = UU^T Y$, and then $B = 0$.

Case (2b): the second equality $X - VQ = 0$ implies $X = VV^T X$, and then $B = 0$.

Finally, if (1a) and (2a) hold true simultaneously, $Z - UR = 0$ implies $Z = UU^T Z$, and $W - VN = 0$ implies $W = VV^T W$. Replacing in (3.25), we get again $B = 0$.

This proves that $\tilde{P}_E(\operatorname{Re}(S)) = \operatorname{Re}(S)$. \square

An immediate consequence of Theorem 3.5 is that the rank-4 stationary points of (3.5) and (3.16) coincide.

4. Approximating the distance to defectivity. For a given ε recall that

$$(4.1) \quad r(\varepsilon) = \min_{E: \|E\|_F=1} y^H x ,$$

where x and y are respectively the right and left eigenvectors of $A + \varepsilon E$ with $y^H x \geq 0$. Given a matrix A with all distinct eigenvalues and $\delta \geq 0$, we look for

$$\varepsilon^{\delta,*} = \arg \min_{\varepsilon > 0} \{ \varepsilon : r(\varepsilon) \leq \delta \} .$$

Starting from $\varepsilon > 0$ such that $r(\varepsilon) > \delta$, which can be computed by integrating the ODEs in previous sections, we want to compute a root $\varepsilon^{\delta,*}$ of the equation $r(\varepsilon) = \delta$. We expect generically that $r(\varepsilon)$ is not smooth at zero, when two eigenvalues coalesce to form a Jordan block. Nevertheless, the eigenvalue $\lambda(\varepsilon)$ and the eigenvectors $x(\varepsilon)$ and $y(\varepsilon)$ are smooth if $\varepsilon < \varepsilon^{0,*}$, so that $r(\varepsilon)$ is smooth as well.

More precisely, by Theorem 2.5 and Corollary 2.10 we know that at least for ε small enough there exists a unique branch of minimizers $E(\varepsilon)$, which are nondegenerate solutions to the stationary equation $P_E(S) = \text{Re}\langle E, S \rangle E$ and $\tilde{P}_E(\text{Re}(S)) = \langle E, \text{Re}(S) \rangle E$. These equations have the form $F(E, \varepsilon) = 0$, with F a (complex or real) analytic function of E and ε ; hence the solution $E(\varepsilon)$, as well as the corresponding simple eigenvalue $\lambda(\varepsilon)$ and eigenvectors $x(\varepsilon)$ and $y(\varepsilon)$, are (complex or real) analytic functions of ε . In order to derive an equation for ε to approximate $\varepsilon^{\delta,*}$, the δ -distance to defectivity, we need to compute the derivative of $r(\varepsilon)$ with respect to ε , which we denote by $r'(\varepsilon)$.

THEOREM 4.1. *Let $r(\varepsilon)$, $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$, be a branch of minima such that $r(\varepsilon) > 0$ for all ε , so that $\lambda(\varepsilon)$ is a simple eigenvalue of $A + \varepsilon E(\varepsilon)$, where $E(\varepsilon)$ is a smooth minimizer of (4.1) with either $E \in \mathbb{C}^{n,n}$ or $E \in \mathbb{R}^{n,n}$ (in this case A is assumed real); i.e., $\lambda(\varepsilon) \in \Lambda(A + \varepsilon E(\varepsilon))$ with $\|E(\varepsilon)\|_F = 1$ for all $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$. For all ε , let $x(\varepsilon)$ and $y(\varepsilon)$ be smooth vector-valued functions determining right and left eigenvectors of $A + \varepsilon E(\varepsilon)$ of unit norm and such that $y(\varepsilon)^H x(\varepsilon) > 0$, and let $G(\varepsilon)$ be the group inverse of $A + \varepsilon E(\varepsilon) - \lambda(\varepsilon)I$. For the function $r(\varepsilon)$ we have*

$$(4.2) \quad \frac{dr(\varepsilon)}{d\varepsilon} = r(\varepsilon) \text{Re} \left(x(\varepsilon)^H G(\varepsilon) E(\varepsilon) x(\varepsilon) + y(\varepsilon)^H E(\varepsilon) G(\varepsilon) y(\varepsilon) \right) \leq 0 .$$

Proof. We get

$$(4.3) \quad r'(\varepsilon) = y'(\varepsilon)^H x(\varepsilon) + y(\varepsilon)^H x'(\varepsilon) .$$

Making use of (2.1) and recalling that $G(\varepsilon)x(\varepsilon) \equiv 0$ and $y(\varepsilon)^H G(\varepsilon) \equiv 0$, we obtain

$$(4.4) \quad \begin{aligned} r'(\varepsilon) &= r(\varepsilon) \text{Re} \left(x(\varepsilon)^H G(\varepsilon) E(\varepsilon) x(\varepsilon) + y(\varepsilon)^H E(\varepsilon) G(\varepsilon) y(\varepsilon) \right) \\ &+ \varepsilon r(\varepsilon) \text{Re} \left(x(\varepsilon)^H G(\varepsilon) E'(\varepsilon) x(\varepsilon) + y(\varepsilon)^H E'(\varepsilon) G(\varepsilon) y(\varepsilon) \right) . \end{aligned}$$

In order to prove the theorem we have to show that the second term in (4.4) vanishes. By Theorem 2.4 we have that the extremizer (stationary point of the corresponding ODE) of (4.1) with $E \in \mathbb{C}^{n,n}$ is

$$(4.5) \quad E(\varepsilon) = c S(\varepsilon), \quad \text{where } S(\varepsilon) = y(\varepsilon)y(\varepsilon)^H G(\varepsilon)^H + G(\varepsilon)^H x(\varepsilon)x(\varepsilon)^H ,$$

for a suitable constant c . Furthermore, by norm conservation of $E(\varepsilon)$ as ε varies,

$$(4.6) \quad \text{Re} \langle E(\varepsilon), E'(\varepsilon) \rangle = 0 .$$

Similarly, by (3.5), in the case of (4.1) with $E \in \mathbb{R}^{n,n}$,

$$(4.7) \quad E(\varepsilon) = c \text{Re}(S(\varepsilon)), \quad \langle E(\varepsilon), E'(\varepsilon) \rangle = 0 .$$

Equations (4.5), (4.6), and (4.7) imply that in both cases (see (4.4)),

$$r'(\varepsilon) = r(\varepsilon) \text{Re} \left(\langle S(\varepsilon), E(\varepsilon) \rangle + \varepsilon \langle S(\varepsilon), E'(\varepsilon) \rangle \right) = r(\varepsilon) \text{Re} \langle S(\varepsilon), E(\varepsilon) \rangle ,$$

which is (4.2), which completes the first part of the proof. The nonpositivity is due to monotonicity of $r(\varepsilon)$ with respect to ε . This property follows immediately by noticing that the minimum in (4.1) can be equivalently computed on the closed ball $\{E: \|E\|_F \leq 1\}$. Indeed, given a matrix $E \in \mathbb{C}^{n,n}$ of norm $\|E\|_F < 1$, the matrix $E_\alpha = E + \alpha(I - xy^H/y^Hx)$ is such that $A + \varepsilon E_\alpha$ has the same right and left eigenvectors x and y , and the parameter α can be chosen to have $\|E_\alpha\|_F = 1$. Similarly, for A real and given a matrix $E \in \mathbb{R}^{n,n}$ of norm $\|E\|_F < 1$, we can repeat the argument with the real matrix $E_\alpha = E + \alpha(I - 2\text{Re}(xy^H)/y^Hx)$ (recall $y^Hx > 0$). Indeed, as \bar{x}, \bar{y} are eigenvectors of $A + \varepsilon E$, by biorthogonality it follows that also in this case $A + \varepsilon E_\alpha$ has the same right and left eigenvectors x and y . \square

Remark 4.2. It is worthwhile to note that the monotonicity property of $r(\varepsilon)$ implies that the constant c in (4.5) is negative. More precisely, we get $c = -\|S(\varepsilon)\|_F^{-1}$ in the complex case and $c = -\|\text{Re}(S(\varepsilon))\|_F^{-1}$ in the real case. For the complex case, consider the Cauchy problem

$$\begin{cases} \dot{\mathcal{E}}(t) = -S(t), \\ \mathcal{E}(0) = E(\varepsilon), \end{cases}$$

where $S(t) = y(t)y(t)^H G(t)^H + G(t)^H x(t)x(t)^H$ is the maximal descent direction associated with $A + \varepsilon \mathcal{E}(t)$, so that $y(t)^H x(t) < r(\varepsilon)$ for t positive and small. On the other hand, as $E(\varepsilon) = cS(\varepsilon)$,

$$\left. \frac{d}{dt} \|\mathcal{E}(t)\|_F^2 \right|_{t=0} = 2c \|S(\varepsilon)\|_F^2.$$

Hence, $c > 0$ would imply both $\|\mathcal{E}(t)\|_F > \|E(\varepsilon)\|_F$ and $y(t)^H x(t) < r(\varepsilon)$ for t positive and small, in contradiction with the nonincreasing property of $r(\varepsilon)$. For the real case, the argument is analogous, and we omit the details.

In particular, (4.2) can be written in the more concise form

$$(4.8) \quad r'(\varepsilon) = \begin{cases} -r(\varepsilon)\|S(\varepsilon)\|_F & \text{in the complex case,} \\ -r(\varepsilon)\|\text{Re}(S(\varepsilon))\|_F & \text{in the real case.} \end{cases}$$

We may use Theorem 4.1 to devise a Newton-bisection iteration to solve the equation $r(\varepsilon) = \delta$. However, since $r(\varepsilon)$ is not differentiable at $\varepsilon = \varepsilon^{0,*}$, this is not recommended for small δ (as our experiments show). We now analyze the behavior of $r(\varepsilon)$ as ε approaches $\varepsilon^{0,*}$. For the sake of simplicity we consider only the complex case, but similar results can be deduced in the real-structured case. To state and prove the main result some extra remarks are needed.

A first remark is that, for the generic case under consideration in which only two eigenvalues coalesce to form a Jordan block, the matrix $A + \varepsilon^{0,*}E(\varepsilon^{0,*})$ is non-derogatory, and hence it is a continuity point for its invariant subspaces [CH80]. In particular (the limits below are intended from the left),

$$(4.9) \quad x_* = x(\varepsilon^{0,*}) = \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} x(\varepsilon), \quad y_* = y(\varepsilon^{0,*}) = \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} y(\varepsilon),$$

where, obviously, $y_*^H x_* = r(\varepsilon^{0,*}) = 0$. For the same reason, with

$$(4.10) \quad M(\varepsilon) = A + \varepsilon E(\varepsilon),$$

the rank of the matrix $M(\varepsilon) - \lambda(\varepsilon)I$ remains equal to $n - 1$ also at $\varepsilon = \varepsilon^{0,*}$. Consequently, for Q^\dagger being the Moore–Penrose pseudoinverse of the matrix Q (see, e.g., [GVL13]) and setting

$$(4.11) \quad B(\varepsilon) = (M(\varepsilon) - \lambda(\varepsilon)I)^\dagger, \quad C(\varepsilon) = y(\varepsilon)^H B(\varepsilon)x(\varepsilon),$$

the following limits exist and are finite:

$$(4.12) \quad B_* = B(\varepsilon^{0,*}) = \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} B(\varepsilon), \quad C_* = C(\varepsilon^{0,*}) = \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} C(\varepsilon).$$

PROPOSITION 4.3. *Under the conditions of Theorem 4.1, assume that the branch of minima extends to the whole interval $[\underline{\varepsilon}, \varepsilon^{0,*})$. Then,*

$$(4.13) \quad \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} r(\varepsilon)r'(\varepsilon) = \begin{cases} 2|C_*| & \text{in the complex case,} \\ 2|C_*| \|\operatorname{Re}(x_*y_*^H)\|_F & \text{in the real case.} \end{cases}$$

Proof. We recall that the group inverse G of a matrix Q which has nullity one admits the following representation [GO11]:

$$G = \left(I - \frac{xy^H}{y^Hx} \right) Q^\dagger \left(I - \frac{xy^H}{y^Hx} \right),$$

where x, y are right and left null vectors of Q , respectively. In our context this gives

$$(4.14) \quad \begin{aligned} G(\varepsilon) &= \left(I - \frac{x(\varepsilon)y(\varepsilon)^H}{r(\varepsilon)} \right) B(\varepsilon) \left(I - \frac{x(\varepsilon)y(\varepsilon)^H}{r(\varepsilon)} \right) \\ &= \frac{C(\varepsilon)x(\varepsilon)y(\varepsilon)^H}{r(\varepsilon)^2} - \frac{B(\varepsilon)x(\varepsilon)y(\varepsilon)^H + x(\varepsilon)y(\varepsilon)^H B(\varepsilon)}{r(\varepsilon)} + B(\varepsilon), \end{aligned}$$

where $B(\varepsilon)$ and $C(\varepsilon)$ are defined in (4.11). By (4.14) and using properties of the Moore–Penrose pseudoinverse,

$$(4.15) \quad x(\varepsilon)^H B(\varepsilon) = 0, \quad B(\varepsilon)y(\varepsilon) = 0,$$

we have

$$(4.16) \quad S(\varepsilon)^H = \frac{2C(\varepsilon)x(\varepsilon)y(\varepsilon)^H}{r(\varepsilon)^2} - \frac{B(\varepsilon)x(\varepsilon)y(\varepsilon)^H + x(\varepsilon)y(\varepsilon)^H B(\varepsilon)}{r(\varepsilon)}.$$

Now, by applying (4.12) we get

$$(4.17) \quad \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} r(\varepsilon)^2 \|S(\varepsilon)\|_F = 2|C_*|, \quad \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} r(\varepsilon)^2 \|\operatorname{Re}(S(\varepsilon))\|_F = 2|C_*| \|\operatorname{Re}(x_*y_*^H)\|_F.$$

The proof is completed with (4.8). \square

As $2rr' = (r^2)'$, the above proposition shows that $r^2(\varepsilon)$ is at least of order $\varepsilon^{0,*} - \varepsilon$ as $\varepsilon \rightarrow \varepsilon^{0,*}$; this implies, in the generic case in which the limits in (4.13) are not zero, that the function $r(\varepsilon)$ approaches zero like $\sqrt{\varepsilon^{0,*} - \varepsilon}$ as $\varepsilon \rightarrow \varepsilon^{0,*}$. Indeed, under some further generic assumptions, such behavior holds true also for the higher-order term in the Puiseux expansion of $r(\varepsilon)$ at $\varepsilon^{0,*}$. This is the content of the next proposition, where for brevity we treat the complex case.

PROPOSITION 4.4. *Under the conditions of Theorem 4.1, let $E(\varepsilon)$ be a smooth minimizer of (4.1) with $E \in \mathbb{C}^{n,n}$, and assume that the branch of minima extends to the whole interval $[\underline{\varepsilon}, \varepsilon^{0,*})$. Then, if $\varepsilon^{0,*} \|B_*\|_F$ is sufficiently small,*

$$(4.18) \quad \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} r(\varepsilon)^3 r''(\varepsilon) = -2|C_*|(1 + 4|C_*|).$$

Proof. For the proof, see the supplementary materials, accessible at the online version of this paper. \square

Algorithm 1: Basic algorithm for computing the δ -distance to defectivity.

Data: δ , tol, starting eigenvalue λ_0 (candidate to coalesce), ε_ℓ (lower bound, default 0), ε_r (upper bound; see section 5.1), θ (bisection parameter), $\varepsilon_0 \in (\varepsilon_\ell, \varepsilon_r)$, $E_0 = -S/\|S\|$ ($E_0 = -\text{Re}(S)/\|\text{Re}(S)\|_F$ in the real case), where $S = yy^H G^H + G^H xx^H$, with G the group inverse of $A - \lambda_0 I$, and x (resp., y) the right (resp., left) eigenvector associated to λ_0 .

Result: $\varepsilon^{\delta,*}$

begin

```

1  Set  $k = 0$  and  $r(\varepsilon_0) = 2\delta$ .
2  while  $|r(\varepsilon_k) - \delta| \geq \text{tol}$  do
3    Compute  $E_k$  by integrating (2.11) in the complex case (or (3.12) in the
4    real case), with given initial datum  $E(\varepsilon_{k-1})$  (if  $k \geq 1$ ).
5    Compute right/left eigenvectors  $x_k/y_k$  of  $A + \varepsilon_k E_k$  and  $r(\varepsilon_k) = y_k^H x_k$ .
6    if  $r(\varepsilon_k) > \text{tol}$  then
7      | Set  $\varepsilon_\ell = \varepsilon_k$  and Bisect = False.
8    else
9      | Set  $\varepsilon_r = \varepsilon_k$  and Bisect = True.
10   if Bisect = True then
11     | Set  $\varepsilon_{k+1} = (1 - \theta)\varepsilon_\ell + \theta\varepsilon_r$ .
12   else
13     | Compute  $r'(\varepsilon_k)$  by (4.2) and next  $\gamma_k$  and  $\varepsilon_k^{0,*}$  by (4.20).
14     | Set  $\widehat{\varepsilon}_{k+1} = \varepsilon_k^{0,*} - \delta^2/\gamma_k^2$ .
15     | if  $\widehat{\varepsilon}_{k+1} \in (\varepsilon_\ell, \varepsilon_r)$  then
16       | | Set  $\varepsilon_{k+1} = \widehat{\varepsilon}_{k+1}$ .
17     | else
18       | | Set  $\varepsilon_{k+1} = (1 - \theta)\varepsilon_\ell + \theta\varepsilon_r$ .
19   | Set  $k = k + 1$ .
20 Return  $\varepsilon^{\delta,*} \approx \varepsilon_k$ 

```

4.1. The algorithm. For ε close to $\varepsilon^{0,*}$, $\varepsilon < \varepsilon^{0,*}$, we have generically

$$(4.19) \quad \begin{cases} r(\varepsilon) &= \gamma\sqrt{\varepsilon^{0,*} - \varepsilon} + \mathcal{O}((\varepsilon^{0,*} - \varepsilon)^{3/2}), \\ r'(\varepsilon) &= -\frac{\gamma}{2\sqrt{\varepsilon^{0,*} - \varepsilon}} + \mathcal{O}((\varepsilon^{0,*} - \varepsilon)^{1/2}), \end{cases}$$

which corresponds to the coalescence of two eigenvalues. In order to set up an iterative process, given ε_k , we use Theorem 4.1 to compute $r'(\varepsilon)$ and estimate γ and $\varepsilon_k^{0,*}$ by solving (4.19) with respect to γ and $\varepsilon^{0,*}$. We denote the solution as γ_k and $\varepsilon_k^{0,*}$, i.e.,

$$(4.20) \quad \gamma_k = \sqrt{2r(\varepsilon_k)|r'(\varepsilon_k)|}, \quad \varepsilon_k^{0,*} = \varepsilon_k + r(\varepsilon_k)/(2|r'(\varepsilon_k)|),$$

and then compute $\varepsilon_{k+1} = \varepsilon_k^{0,*} - \delta^2/\gamma_k^2$. Algorithm 1 is a realization of this idea.

The correction of ε_{k+1} at line 8 is due to the fact that if ε_{k+1} is larger than $\varepsilon^{0,*}$, the value is corrected by a bisection step. The test at line 2—when negative—means that for $\varepsilon = \varepsilon_k$ there are coalescing eigenvalues (up to a tolerance tol) in the

ε -pseudospectrum. Numerically, the algorithm shows quadratic convergence; that is (for small δ) close to $\varepsilon^{\delta,*}$,

$$|\varepsilon_{k+1} - \varepsilon^{\delta,*}| = \mathcal{O}(|\varepsilon_k - \varepsilon^{\delta,*}|^2),$$

as we will show in the numerical illustrations in section 6. Clearly, if δ is not too small, a classical Newton iteration might also be alternatively used.

Remark 4.5. If we set δ sufficiently small, by exploiting the Puiseux expansion (4.19), we find at the same time an accurate solution $\varepsilon^{\delta,*}$ of the equation $r(\varepsilon) = \delta$ and of $\varepsilon^{0,*}$, that is, an upper bound for the distance to defectivity.

5. Implementation issues. In this section we discuss some aspects relevant to the numerical computation of the δ -distance to defectivity.

Here we consider the cases where a decomposition $A = VDV^{-1}$ is available (D is the diagonal matrix of eigenvalues of A). This case is interesting when the size of A is not too large. The situation where such a decomposition of A is not available would also be interesting, but it needs a guess for the candidate eigenvalues to coalesce, and for this reason we do not discuss it here.

5.1. Upper bound. In order to compute an upper bound ε_r to $\varepsilon^{\delta,*}$, we consider

$$A = \sum_{i=1}^n \lambda_i \frac{1}{y_i^H x_i} x_i y_i^H,$$

the canonical decomposition of A , where x_i and y_i are the right and left eigenvectors of A associated with λ_i , normalized as $\|x_i\| = \|y_i\| = 1$ and $y_i^H x_i > 0$ for all i . In the case of complex perturbations ($\mathbb{K} = \mathbb{C}$) we have the following natural upper bound:

$$(5.1) \quad \varepsilon_r^{\mathbb{C}} = \min_{1 \leq i \leq n} \min_{1 \leq j \leq n, j \neq i} \frac{|\lambda_i - \lambda_j|}{y_i^H x_i}.$$

For a real matrix, in the case of real admissible perturbations, we have the same bound (5.1) if we consider coalescence of two real eigenvalues or two complex conjugate eigenvalues. Otherwise, if two complex eigenvalues (not conjugate to each other) coalesce, we have to consider a double coalescence (due to the fact that the conjugate pair also coalesces). This gives the following value:

$$(5.2) \quad \varepsilon_r^{\mathbb{R}} = \min_{1 \leq i \leq m} \min_{1 \leq j \leq m, j \neq i} 2 \frac{\|\operatorname{Re}((\lambda_i - \lambda_j)x_i y_i^H)\|_F}{y_i^H x_i},$$

where the set of the first $m \leq n/2$ eigenvalues has nonzero imaginary part and does not contain any conjugate pair. Finally, for complex conjugate eigenvalues coalescing on the real axis, we obtain the bound

$$(5.3) \quad \widehat{\varepsilon}_r^{\mathbb{R}} = \min_{1 \leq i \leq n, \lambda_i \notin \mathbb{R}} \frac{\|\operatorname{Re}((\lambda_i - \bar{\lambda}_i)x_i y_i^H)\|_F}{y_i^H x_i}.$$

5.2. Choice of the initial eigenvalue. We consider two possibilities: a simple choice is to choose one among the eigenvalues of largest condition number; a more sophisticated choice is obtained by extending the method proposed in [ABBO11] to the case of real perturbations. Letting p_j be the condition number for λ_j of A , we define the candidate coalescence point, say z_0 , by the following combination of two

eigenvalues: $z_0 = (p_{j'} \lambda_{k'} + p_k \lambda_{j'}) / (p_{j'} + p_{k'})$, where the index pair (j', k') minimizes $|\lambda_j - \lambda_k| / (p_j + p_k)$ over all distinct pairs of eigenvalues. The starting guess is derived from first-order perturbation bounds for simple eigenvalues. If A is perturbed by εE with ε small, then for an eigenvalue λ_j of A there exists an eigenvalue $\tilde{\lambda}_j$ of $A + \varepsilon E$ such that $|\tilde{\lambda}_j - \lambda_j| \leq p_j \varepsilon + \mathcal{O}(\varepsilon^2)$. Thus, for sufficiently small ε , the component of $\Lambda_\varepsilon^{\mathbb{K}}(A)$ containing λ_j is approximately a disk of radius $p_j \varepsilon$ centered at λ_j . Therefore, a point of coalescence of two components of $\Lambda_\varepsilon^{\mathbb{K}}(A)$ containing eigenvalues, say, λ_j and λ_k , is expected to be approximated by the point of coalescence of the disks $|z - \lambda_j| \leq p_j \varepsilon$ and $|z - \lambda_k| \leq p_k \varepsilon$. This gives the guess z_0 which is the point of coalescence of the disks for the eigenvalues $\lambda_{j'}$ and $\lambda_{k'}$. Note that when $\mathbb{K} = \mathbb{C}$, the rate p_j of λ_j is given by the traditional eigenvalue condition number $\kappa(\lambda_j)$, whereas when $\mathbb{K} = \mathbb{R}$, the role of p_j is played by the first-order measure in the Frobenius norm of the worst-case effect on λ_j of real perturbations, in detail, $p_j = \|\text{Re}(y_j x_j^H)\|_F / y_j^H x_j$; see, e.g., [BK04, KKT06].

5.3. Euler discretization. The stepsize control in the numerical integration of the ODEs by the Euler method is simply driven by the monotonicity requirement $r(t_{n+1}) < r(t_n)$. The stepsize h can be selected by Algorithm 2.

Algorithm 2: Stepsize selection.

Data: the eigenvalue λ_n and the normalized eigenvectors x_n, y_n of the matrix $A + \varepsilon E_n$ (with $E_n \approx E(t_n)$, $t_n = t_{n-1} + h_{n-1}$), $\sigma > 1$ (stepsize ratio), h (predicted stepsize).

Result: $h_{n+1}, E_{n+1}, \lambda_{n+1}$

begin

- 1 | Apply the Euler method to (2.11) (or (3.12)) to get $\tilde{E}_{n+1} \approx E(t_n + h)$.
- 2 | Compute the eigenvalue $\tilde{\lambda}$ of $A + \varepsilon \tilde{E}_{n+1}$, closest to λ_n .
- 3 | Compute the normalized eigenvectors \tilde{x}, \tilde{y} associated with $\tilde{\lambda}$.
- 4 | **if** $\tilde{y}^H \tilde{x} \geq y_n^H x_n$ **then**
 | Reduce the stepsize $h = h/\sigma$; go to 1.
- 5 | **if** $h \geq h_{n-1}$ **then**
 | Increase the stepsize $h = \sigma h$.
- 6 | **Return** $\lambda_{n+1} = \tilde{\lambda}(h)$, $E_{n+1} = \tilde{E}_{n+1}$

5.4. Rescaling and computational issues. When $\varepsilon = \varepsilon^{0,*}$ the group inverse of $A + \varepsilon E(\varepsilon) - \lambda(\varepsilon)I$ is not defined, and, in fact, in the considered ODEs it becomes singular as time approaches infinity, i.e., close to coalescence. This means that when $\varepsilon^{0,*} - \varepsilon$ is very small, the group inverse gets very large at the stationary points. To overcome this problem we may premultiply the right-hand side of the ODEs by the real scalar $r^2 = (y^H x)^2$. Indeed, this is equivalent to replacing G by $(y^H x)^2 G$, which has a finite limit as $\varepsilon \rightarrow \varepsilon^{0,*}$; that is, by (4.11), (4.12), and (4.14),

$$\lim_{\varepsilon \rightarrow \varepsilon^{0,*}} r(\varepsilon)^2 G(\varepsilon) = \lim_{\varepsilon \rightarrow \varepsilon^{0,*}} C(\varepsilon) x(\varepsilon) y(\varepsilon)^H = C_* x_* y_*^H,$$

where x_*, y_* are defined in (4.9). Indeed, using (5.4) below, the constant C_* defined in (4.12) is equivalently given by $C_* = y_*^H (B(\varepsilon^{0,*}) + y_* x_*^H)^{-1} x_*$, where $B(\varepsilon) = A + \varepsilon E(\varepsilon) - \lambda(\varepsilon)I$. We get this by showing the existence of the limit of $(B(\varepsilon) + y(\varepsilon)x(\varepsilon)^H)^{-1}$

as $\varepsilon \rightarrow \varepsilon^{0,*}$. Let us consider the 2×2 block

$$\begin{pmatrix} a & 1 \\ -a^2 + \delta a & -a + \delta \end{pmatrix},$$

where $a, \delta \in \mathbb{C}$ are dependent on ε , such that $\lim_{\varepsilon \rightarrow \varepsilon^{0,*}} \delta = 0$ and $\lim_{\varepsilon \rightarrow \varepsilon^{0,*}} a = a_* \in \mathbb{R}$. This is a generic matrix with a simple zero eigenvalue for $\varepsilon < \varepsilon^{0,*}$ which becomes double and defective at $\varepsilon = \varepsilon^{0,*}$. An explicit computation yields

$$B(\varepsilon^{0,*}) + y_* x_*^H = \begin{pmatrix} 2a_* & 1 - a_*^2 \\ 1 - a_*^2 & -2a_* \end{pmatrix},$$

which is clearly invertible. For problems of small dimension we need to compute the eigenvalues and the eigenvectors of the matrices $A + \varepsilon E$ by using the MATLAB routine *eig* . In particular, once we have the right eigenvector, we compute the left one by making a second call to *eig* to compute the right eigenvector of $(A + \varepsilon E)^H$; see [GO11] for more details. For problems of large dimension (and possibly sparse structure) we use the routine *eigs* , which is an interface for ARPACK [LSY98], the implicitly restarted Arnoldi implementation. This choice enables us to compute eigenpairs without forming the dense matrix $A + \varepsilon E$ explicitly. The main computational problem when computing the right-hand side of the differential equations (2.11) and (3.12) is the application of the group inverse G to a vector. In order to compute it efficiently we make use of the following result from [GOS14].

THEOREM 5.1. *Suppose that B has index one with $x \in \ker(B)$ and $y \in \ker(B^H)$ of unit norm and such that $y^H x > 0$. Let $B^\#$ be the group inverse of B . Then*

$$(5.4) \quad B^\# = \Pi B^\dagger \Pi = \Pi (B + yx^H)^{-1} \Pi,$$

where Π is the projection $(I - xz^H)$ with $z = y/(y^H x)$.

In the case considered here we can apply Theorem 5.1 with $B = A - \lambda I + \varepsilon E$, so that—using the second equation of (5.4)— $G = \Pi (A - \lambda I + F)^{-1} \Pi$, where $F = \varepsilon E + \mu xz^H = U_1 \Sigma_1 V_1^H$ is a rank- ℓ matrix, with $\ell = 3$ in the complex case and $\ell = 5$ in the real case. Thus U_1 and V_1 are $n \times \ell$ matrices, and Σ_1 is an $\ell \times \ell$ diagonal matrix. Since $(A - \lambda I)^{-1} = V(D - \lambda I)^{-1} V^{-1}$ and we have V , D , and V^{-1} (which we have already computed to detect a candidate eigenvalue of A to coalesce), we can make use of the well-known *Sherman–Morrison–Woodbury* formula [W50] to exploit the low-rank structure of F , given by (with $H = (A - \lambda I)^{-1}$)

$$(A - \lambda I + F)^{-1} = H - HU_1 \Sigma_1 (\Sigma_1 + \Sigma_1 V_1 H U_1 \Sigma_1)^{-1} \Sigma_1 V_1 H.$$

This allows for an efficient computation of the matrix vector product Gw , w being an arbitrary vector. When the matrix A is sparse, we can make use of a sparse linear solver. For example, we can use GMRES as the computation of matrix products of the form $A + F$, where A is sparse and F of low rank is inexpensive. Additionally, making use of the routine *ilu* [Sa96], we have the sparse incomplete LU factorization of A , i.e., $\tilde{L}\tilde{U}$, and we can compute as before the vector product $\tilde{G}w$, w being an arbitrary vector and \tilde{G} the group inverse of $\tilde{L}\tilde{U} + F$, as a preconditioner for GMRES.

Remark 5.2. For ε not too close to the critical value $\varepsilon^{0,*}$, the computation of the expressions in (5.4) is well-conditioned. Since we compute $\varepsilon^{0,*}$ by means of (4.19), we do not need to get too close to $\varepsilon^{0,*}$. Whenever $\varepsilon \approx \varepsilon^{0,*}$, the formulas contain sensitive computations, i.e., the approximation of the eigenvectors y and x associated to a quasi-defective eigenvalue, and the solution of a possibly ill-conditioned linear system. However, since we are not interested in computing the exact solution but only in decreasing $r(\varepsilon)$, we usually also observe good behavior in such a situation.

6. Numerical illustrations. First we consider two illustrative examples of small dimension, of both complex and real matrices. For the second real example we compute both the complex and the real distance to defectivity. All computations were carried out in MATLAB (Version R2014a).

Example 1. Consider the complex matrix

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

By using the method proposed in [ABBO11] to identify a pair of eigenvalues of A candidate to coalesce, which we have discussed in section 5.2, we compute

$$\lambda_1 = 1.416177710 + 1.260523165i, \quad \lambda_2 = 0.338991381 + 0.455810180i.$$

Figure 1 shows the ε -pseudospectrum for $\varepsilon = \varepsilon^{0,*}$, which reveals that this value corresponds to the exact distance to defectivity of the matrix A .

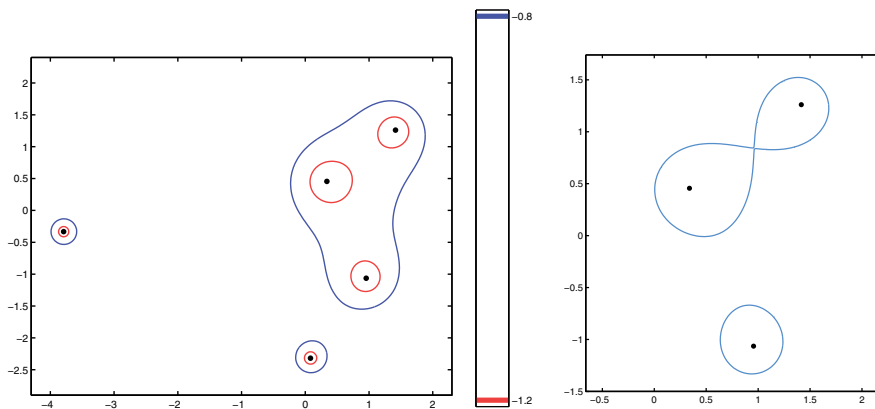


FIG. 1. *Left: The complex ε -pseudospectrum for $\varepsilon_1 = 10^{-1.2}$ and $\varepsilon_2 = 10^{-0.8}$ indicates that the distance to defectivity $\varepsilon^{0,*} \in (\varepsilon_1, \varepsilon_2)$. Right: The ε -pseudospectrum for $\varepsilon = \varepsilon^{0,*}$. The point of coalescence of the eigenvalues, $\lambda = 0.961516149290911 + 0.840702239813292i$, is localized at the double point of the lemniscate in the upper part of the picture.*

We apply Algorithm 1, setting $\lambda_0 = \lambda_1$, $\varepsilon_0 = 10^{-1.2}$, $\delta = 10^{-3}$, $\theta = 0.8$, and $\text{tol} = 10^{-6}$. As a result we obtain

$$\varepsilon^{\delta,*} = 0.082876706760826.$$

Table 1 illustrates the behavior of Algorithm 1. The computed distance agrees with that computed by the code Eigopt by Mengi, Yildirim, and Kiliç [MYK14] and the code Neardefmat by Alam et al. [ABBO11].

TABLE 1
Computed values of ε and $r(\varepsilon)$ for Example 1.

k	ε_k	$r(\varepsilon_k)$	k	ε_k	$r(\varepsilon_k)$
0	0.0630957344480	0.1052552110776	5	0.0828797867212	< tol
1	0.0860136668542	< tol	6	0.0828287046155	0.0059191631349
2	0.0814300803729	0.0312807047835	7	0.0828769469626	0.0009101061019
3	0.0829229501475	< tol	8	0.0828767067896	0.0009999896898
4	0.0826243761926	0.0133249418460	9	0.0828767067608	0.0009999999997

Example 2. Consider the real matrix A ,

$$(6.1) \quad A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix},$$

known as *Grcar*, of dimension 6. A pair of initial eigenvalues candidate to coalesce is

$$\lambda_1 = 0.358489183 - 1.950114681i, \quad \lambda_2 = 1.139108055 - 1.230297560i.$$

TABLE 2
Computed values of ε and $r(\varepsilon)$ for Example 2 (real distance).

k	ε_k	$r(\varepsilon_k)$	k	ε_k	$r(\varepsilon_k)$
0	0.1000000000000	0.5896332470947	3	0.3007166340568	0.0009775689197
1	0.3507801853487	0.0000663072831	4	0.3007166135743	0.0009999998497
2	0.3006241482790	0.0141840323629	5	0.3007166135741	0.0010000000000

We apply Algorithms 1 and 2, setting $\lambda_0 = \lambda_1$, $\varepsilon_0 = 10^{-1}$, $\delta = 10^{-3}$, $\theta = 0.8$, and $\text{tol} = 10^{-7}$. As results we obtain

$$\begin{aligned} \varepsilon^{\delta,*} &= 0.215185436319885 && \text{(complex distance)}, \\ \varepsilon^{\delta,*} &= 0.300716613574112 && \text{(real distance)}. \end{aligned}$$

Table 2 illustrates the behavior of Algorithm 1 for computing the real δ -distance. The coalescence point is $\lambda = 0.755369495776934 - 1.590698410975971i$.

The estimated real distance to defectivity is $\varepsilon^{0,*} \approx 0.300717075321045$.

Example 3: West0067 matrix, dimension 67. This real matrix is part of the Chemwest collection [BPR97], which models chemical engineering plants.

Complex distance. We set $\delta = 10^{-3}$ and obtain the following values:

$$\varepsilon^{0,*} = 0.00551675100, \quad \gamma = 2.01525876223, \quad \varepsilon^{\delta,*} = 0.00551650477.$$

The computed eigenvalues, which are expected to coalesce for $\varepsilon = \varepsilon^{0,*}$, are

$$\lambda_1 = -0.252310074 + 0.853692119i, \quad \lambda_2 = -0.252142643 + 0.853721574i,$$

having a distance $|\lambda_1 - \lambda_2| \approx 1.7 \cdot 10^{-4}$. These values agree with those computed by the code in [ABBO11], i.e., $\varepsilon^{0,*} = 0.00551675$.

Real distance. We set $\delta = 10^{-3}$ and obtain the results in Table 3; as a by-product we compute the following values:

$$\varepsilon^{0,*} = 0.0078798150, \quad \gamma = 0.2864276818, \quad \varepsilon^{\delta,*} = 0.0078676260.$$

The computed four eigenvalues, expected to coalesce pairwise for $\varepsilon = \varepsilon^{0,*}$, are

$$\lambda_1 = -0.251808570 + 0.853267804i, \quad \lambda_2 = -0.251792424 + 0.853170539i$$

and their conjugates, having a distance $|\lambda_1 - \lambda_2| \approx 9.8 \cdot 10^{-5}$.

TABLE 3
Computed values of ε and $r(\varepsilon)$ for Example 3 (real-structured distance).

k	ε_k	$r(\varepsilon_k)$	k	ε_k	$r(\varepsilon_k)$
0	0.0100000000000000	< tol	4	0.0078678769846	0.0005348142974
1	0.0075002500000	0.0321588584749	5	0.0078591741703	0.0050035775278
2	0.0078716932627	< tol	6	0.0078676283725	0.0009966139068
3	0.0077788324471	0.0159016209587	7	0.0078676259959	0.0009999979770

Example 4: Orr-Sommerfeld matrix, dimension 99. This matrix is taken from EigTool [Wri02] and arises in the discretization of the Orr-Sommerfeld operator at large Reynolds numbers. Since the matrix is complex, we consider its real part.

Real distance. We set $\delta = 10^{-4}$. Since two real eigenvalues coalesce, the real and complex distances give the same estimate,

$$\varepsilon^{0,*} = 7.56829388628 \cdot 10^{-4}, \quad \gamma = 55.720252607, \quad \varepsilon^{\delta,*} = 7.56829385407 \cdot 10^{-4}.$$

The computed eigenvalues, which are expected to coalesce pairwise for $\varepsilon = \varepsilon^{0,*}$, are

$$\lambda_1 = -0.2675984616 \cdot 10^{-5}, \quad \lambda_2 = -0.2676113453 \cdot 10^{-5},$$

whose distance is $1.3 \cdot 10^{-10}$.

Interestingly, the value computed by the code in [ABBO11] is equal to $\varepsilon^{0,*} = 7.364562414280245 \cdot 10^{-4}$, and the two closest computed eigenvalues have distance $|\lambda_1 - \lambda_2| = 0.00032389$ and condition numbers $k_1 = 3.96107$ and $k_2 = 4.18569$.

Example 5: str_600 matrix, dimension 363. This real matrix is part of the SMTAPE collection [BPR97] and arises from an application of the simplex method.

The complex distance to nearest defective matrix found by the code in [ABBO11] is $6.80846 \cdot 10^{-5}$.

TABLE 4
Computed values of ε and $r(\varepsilon)$ for Example 5 (real-structured distance).

k	ε_k	$r(\varepsilon_k)$
0	0.0001000000000000	0.000800966492236
1	0.000095636329862	0.001000272860068
2	0.000095645234943	0.001000000062242

Real distance. We set $\delta = 10^{-3}$ and obtain the results in Table 4; as a by-product we compute the following values:

$$\varepsilon^{0,*} = 1.133466479 \cdot 10^{-4}, \quad \gamma = 0.237681872, \quad \varepsilon^{\delta,*} = 9.5645234943 \cdot 10^{-5}.$$

The computed eigenvalues, which are expected to coalesce pairwise for $\varepsilon = \varepsilon^{0,*}$, are complex conjugate in this case,

$$\lambda_1 = 0.008426530 + 0.0021636406i, \quad \lambda_2 = 0.008426530 - 0.0021636406i,$$

having a distance $|\lambda_1 - \lambda_2| \approx 4.3 \cdot 10^{-3}$.

Acknowledgments. The authors are grateful to the anonymous referees for their valuable comments and suggestions. MATLAB codes implementing the described algorithms are available from the website <http://univaq.it/~guglielm/>.

REFERENCES

- [AFS13] R. O. AKINOLA, M. A. FREITAG, AND A. SPENCE, *The calculation of the distance to a nearby defective matrix*, Numer. Linear Algebra Appl., 21 (2014), pp. 403–414.
- [Ala06] R. ALAM, *Wilkinson’s problem revisited*, J. Anal., 14 (2006), pp. 175–205.
- [AB05] R. ALAM AND S. BORA, *On sensitivity of eigenvalues and eigendecompositions of matrices*, Linear Algebra Appl., 396 (2005), pp. 273–301.
- [ABBO11] R. ALAM, S. BORA, R. BYERS, AND M. L. OVERTON, *Characterization and construction of the nearest defective matrix via coalescence of pseudospectral components*, Linear Algebra Appl., 435 (2011), pp. 494–513.
- [BPR97] R. F. BOISVERT, R. POZO, K. REMINGTON, R. F. BARRETT, AND J. J. DONGARRA, *Matrix Market: A web resource for test matrix collections*, in The Quality of Numerical Software: Assessment and Enhancement, R. F. Boisvert, ed., Chapman & Hall, London, 1997, pp. 125–137; also available online from <http://math.nist.gov/MatrixMarket/>.
- [BGN12] P. BUTTÀ, N. GUGLIELMI, AND S. NOSCHESI, *Computing the structured pseudospectrum of a Toeplitz matrix and its extreme points*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1300–1319.
- [BK04] R. BYERS AND D. KRESSNER, *On the condition of a complex eigenvalue under real perturbations*, BIT, 44 (2004), pp. 209–214.
- [CH80] J. B. CONWAY AND P. R. HALMOS, *Finite-dimensional points of continuity of Lat*, Linear Algebra Appl., 31 (1980), pp. 93–102.
- [Dem83] J. W. DEMMEL, *A Numerical Analyst’s Jordan Canonical Form*, Ph.D. thesis, Computer Science Division, University of California, Berkeley, 1983.
- [GVL13] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, 2013.
- [GL11] N. GUGLIELMI AND C. LUBICH, *Differential equations for roaming pseudospectra: Paths to extremal points and boundary tracking*, SIAM J. Numer. Anal., 49 (2011), pp. 1194–1209.
- [GO11] N. GUGLIELMI AND M. L. OVERTON, *Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1166–1192.
- [GOS14] N. GUGLIELMI, M. L. OVERTON, AND G. W. STEWART, *An efficient algorithm for computing the generalized null space decomposition*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 38–54.
- [KKT06] M. KAROW, D. KRESSNER, AND F. TISSEUR, *Structured eigenvalue condition numbers*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1052–1068.
- [KL07] O. KOCH AND C. LUBICH, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 434–454.
- [LSY98] R. B. LEHOUCQ, D. C. SORENSEN, AND C. YANG, *ARPACK Users’ Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, Software Environ. Tools 6, SIAM, Philadelphia, 1998.
- [M99] A. N. MALYSHEV, *A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues*, Numer. Math., 83 (1999), pp. 443–454.
- [MYK14] E. MENGI, E. A. YILDIRIM, AND M. KILIÇ, *Numerical optimization of eigenvalues of Hermitian matrix functions*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 699–724.
- [MS88] C. D. MEYER AND G. W. STEWART, *Derivatives and perturbations of eigenvectors*, SIAM J. Numer. Anal., 25 (1988), pp. 679–691.
- [Sa96] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, Boston, 1996.
- [PS05] A. SPENCE AND C. POULTON, *Photonic band structure calculations using nonlinear eigenvalue techniques*, J. Comput. Phys., 204 (2005), pp. 65–81.
- [TE05] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra. The Behavior of Non-normal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.

- [Wil65] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, UK, 1965.
- [W50] M. A. WOODBURY, *Inverting Modified Matrices*, Memo. Rep. no. 42, Statistical Research Group, Princeton University, Princeton, NJ, 1950.
- [Wri02] T. G. WRIGHT, *EigTool: A Graphical Tool for Nonsymmetric Eigenproblems*, Oxford University Computing Laboratory, <http://www.comlab.ox.ac.uk/pseudospectra/eigtool/> (2002).