Structure-preserving eigenvalue solvers for robust stability and controllability estimates

Daniel Kressner and Emre Mengi

***Abstract*— Structured eigenvalue problems feature a prominent role in many algorithms for the computation of robust measures for the stability or controllability of a linear control system. Structures that typically arise are Hamiltonian, skew-Hamiltonian, and symplectic. The use of eigenvalue solvers that preserve such structures can enhance the reliability and efficiency of algorithms for robust stability and controllability measures. This aspect is the focus of the present work, which summarizes and extends existing structure-preserving eigenvalue solvers. Also, a new method for estimating the distance to uncontrollability in a cheap manner is presented. The structured eigenvalue algorithms described in this paper are intented to become part of HAPACK, a software package for solving structured eigenvalue problems and applications.**

I. Introduction

Let us consider a discrete linear time-invariant system

$$x_{k+1} = Ax_k + Bu_k, \tag{1}$$

with $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times m}$. This work is concerned with measures that signal (the loss of) stability and controllability of (1) in a robust manner. Existing approaches [1]–[6] for the computation of these measures require the solution of Hamiltonian and symplectic eigenvalue problems, which both have special eigenvalue symmetries. In this work we justify that preserving such symmetries in finite-precision arithmetic is an important ingredient to make these computations more reliable. Another contribution of this paper is to review and extend existing eigenvalue solvers that achieve this goal. Additionally, we introduce an algorithm for a robust controllability measure. First, let us briefly review some of the robust stability and controllability measures that we focus on in this work.

A. Stability

The open loop stability of (1) is determined by the spectral radius $\rho(A) = \max\{|\lambda| : \lambda \in \Lambda(A)\}$, where $\Lambda(A)$ denotes the set of all eigenvalues of A. Specifically, if $\rho(A) < 1$ then regardless of the initial state the system decays exponentially with asymptotic decay rate $\rho(A)$. Nevertheless a more desirable property that is relevant to the transient behavior is to ensure that the norms of the states will not exceed a prespecified amount at all times, see, e.g., [7], [8]. In order to gain insight into the transient behavior of (1), one has to consider additional information besides the eigenvalues of A, unless A is a normal matrix. (When A is normal, the transient behavior is consistent with the asymptotic behavior.) Two useful sets that mimic the transient behavior are the *ϵ-pseudospectrum* and the *field of values*, which are formally defined as

$$\Lambda_\epsilon(A) = \{\lambda : \det(A + E - \lambda I) = 0, \|E\| \le \epsilon\}, \tag{2}$$

and

$$F(A) = \{y^* A y : y \in \mathbb{C}^n, \|y\| = 1\}, \tag{3}$$

respectively. Here and throughout the rest of this paper, $\|\cdot\|$ denotes the matrix 2-norm. Analogous to the spectral radius, the radii of the smallest circles that surround these sets can be considered, amounting to the *ϵ-pseudospectral radius* and the *numerical radius*

$$\rho_\epsilon(A) = \sup_{\lambda\in\Lambda_\epsilon(A)} |\lambda|, \quad \text{and} \quad r(A) = \sup_{z\in F(A)} |z|, \tag{4}$$

respectively.

The relevance of $\rho_\epsilon(A)$ to the transient behavior of (1) is revealed by the inequalities

$$\sup_{\epsilon>0} \frac{\rho_\epsilon(A) - 1}{\epsilon} \le \sup_k \|A^k\| \le e n \sup_{\epsilon>0} \frac{\rho_\epsilon(A) - 1}{\epsilon}, \tag{5}$$

which is an immediate corollary of the Kreiss matrix theorem [8], [9]. Above, e denotes the Euler constant. On the other hand, a tight bound on the norms of the matrix powers is given by

$$\|A^k\| \le 2r(A)^k. \tag{6}$$

The appendix contains some motivating examples for these robust stability measures.

B. Controllability

One (numerically disastrous) way to check whether (1) is controllable is to check whether the controllability matrix

$$C = \begin{bmatrix} B & AB & \dots & A^{n-2}B & A^{n-1}B \end{bmatrix}$$

has full rank. A numerically much better approach is to check whether $S(\lambda) = [A - \lambda I \;\; B]$ is numerically of full rank for all $\lambda \in \Lambda(A)$. Alternatively, one could also compute the controllability staircase form [10] of the matrix pair (A, B). Although usually quite reliable, both approaches are known to occasionally fail to detect nearby uncontrollable systems. In principle, the only numerically reliable way of checking

D. Kressner is with the Department of Computing Science and HPC2N, Umeå University, S-901 87 Umeå, Sweden, kressner@cs.umu.se. The work of the author was supported by the DFG Emmy Noether fellowship KR 2950/1-1 and by the *Swedish Foundation for Strategic Research* under grant A3 02:128.

E. Mengi is with the Department of Mathematics, University of California, San Diego, CA 92093, USA, emengi@math.ucsd.edu. The work of the author was supported in part by the NSF Grant DMS-0412049.

the controllability of the system (1) and nearby systems is provided by the *distance to uncontrollability*

$$\tau(A,B) = \inf\{\|[\Delta A\ \Delta B]\| : [\Delta A\ \Delta B] \in \mathbb{C}^{n\times(n+m)}, \\ x_{k+1} = (A+\Delta A)x_k + (B+\Delta B)u_k \\ \text{is not reachable}\}, \quad (7)$$

introduced by Paige [11].

C. The role of structure-preserving eigenvalue solvers

The importance of using structure-preserving eigenvalue solvers within algorithms for robust stability measures has been observed quite a few times. For example, in [3] it was shown that the distance of a matrix to instability can be computed very accurately, up to the level of machine precision, if strongly backward stable eigenvalue solvers are employed. An analogous result was obtained in [6] for the computation of the ϵ-pseudospectral radius. Moreover, the use of such eigenvalue solvers avoids certain problem-dependent thresholds for deciding, e.g., whether an eigenvalue of a Hamiltonian matrix is on the imaginary axis.

D. Outline

In the next section we first recall existing algorithms [6] for computing the numerical radius and the pseudospectral radius, then we describe a new trisection algorithm for the distance to uncontrollability well-suited for low precision approximation. These algorithms depend on the extraction of the imaginary eigenvalues of Hamiltonian matrices or the unit eigenvalues (eigenvalue with unit modulo) of symplectic pencils. Section III is devoted to the structure-preserving solution of Hamiltonian and symplectic eigenvalue problems. Finally, we reemphasize and substantiate the importance of employing structure-preserving eigenvalue solvers within numerical radius, pseudospectral radius and distance to uncontrollability computations in Section IV.

II. Algorithms for robust stability and controllability measures

The remarkable similarity in the computation of the numerical radius, the pseudospectral radius and the distance to uncontrollability is that these quantities can be expressed as global optima of optimization problems involving eigenvalues or singular values. These problems are nonsmooth, due to ties in the eigenvalues or singular values, and nonconvex usually with more than one local minimum. Nonsmoothness is mainly of theoretical concern, since generically the problems are smooth around the global optima. The algorithms described in the following are devised to avoid stagnation at local minima and converge to global minima.

A. Numerical Radius

In [6], the Boyd-Balakrishnan algorithm [12], [13] for the H_∞-norm is modified for the numerical radius. The algorithm exploits the equivalent eigenvalue maximization characterization over the unit circle

$$r(A) = \sup_{\theta\in[0,2\pi)} \lambda_{\max}(H(Ae^{i\theta}))$$

where $H(A) = (A + A^*)/2$ is the Hermitian part of A and $\lambda_{\max}$ denotes the largest eigenvalue. In [6, Theorem 3.1] it has been proved that μ is an eigenvalue of the matrix $H(Ae^{i\theta})$ if and only if $e^{i\theta}$ is contained in the spectrum of the pencil

$$R(\mu) - \lambda S = \begin{bmatrix} 2\mu I & -A^* \\ I & 0 \end{bmatrix} - \lambda \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}. \quad (8)$$

Therefore for a fixed μ we are capable of finding the intersection points of the horizontal line $y = \mu$ and the function $f(\theta) = \lambda_{\max}(H(Ae^{i\theta}))$ by means of the two step procedure:

1) extract all unit eigenvalues of $R(\mu) - \lambda S$;
2) for each unit eigenvalue $e^{i\theta}$, check whether the equality $\lambda_{\max}(H(Ae^{i\theta})) = \mu$ holds.

Given an estimate μ for the numerical radius, using the procedure above the intervals (of values of θ) in which $\mu > f(\theta)$ can be determined. The estimate μ for the numerical radius is refined to the maximum value attained by f at the midpoints of the intervals. The convergence of the algorithm follows from the fact that the length of the largest interval in which $\mu > f(\theta)$ is at least halved at each iteration and therefore approaches zero in the limit. In practice, quadratic convergence is observed.

B. Pseudospectral Radius

In [6], an iterative algorithm for computing the ϵ-pseudospectral radius with generically quadratic convergence rate is described. The algorithm benefits from the singular value characterization

$$\Lambda_\epsilon(A) = \{\lambda \in \mathbb{C} : \sigma_{\min}(A - \lambda I) \le \epsilon\} \quad (9)$$

of the ϵ-pseudospectrum with $\sigma_{\min}$ denoting the nth largest singular value of a $n\times m$ matrix assuming $n \le m$ throughout this work. See for instance [8] for a proof of equivalence of (2) and (9). Therefore the ϵ-pseudospectral radius amounts to an optimization problem with the singular values appearing in the constraints.

The intersection points of a circle of radius r centered at the origin with the boundary of the ϵ-pseudospectrum can be retrieved efficiently by solving the eigenvalue problem

$$P(r,\epsilon) - \lambda Q(r,\epsilon) = \begin{bmatrix} -\epsilon I & A \\ rI & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & rI \\ A^* & -\epsilon I \end{bmatrix}. \quad (10)$$

Each unit eigenvalue corresponds to a θ satisfying $\epsilon \in \sigma(A - re^{i\theta}I)$ [3, Theorem 4], where $\sigma(\cdot)$ denotes the singular values of a matrix. This gives the *circular search with radius* r:

1) extract all unit eigenvalues of $P(r,\epsilon) - \lambda Q(r,\epsilon)$ and
2) for each unit eigenvalue $e^{i\theta}$ check whether $\sigma_{\min}(A - re^{i\theta}I) = \epsilon$ holds.

Additionally, there is an even more efficient way to compute the largest modulus intersection point of the line passing through the origin with slope θ and the ϵ-pseudospectrum. In [6, Theorem 2.6] it was proved that this quantity corresponds

to the imaginary part of the largest modulus imaginary eigenvalue of the matrix

$$K(\theta) = \begin{bmatrix} ie^{i\theta}A^* & \epsilon I \\ -\epsilon I & ie^{-i\theta}A \end{bmatrix}. \quad (11)$$

This procedure is called *the radial search in the direction* θ.

Given an estimate η of the ϵ-pseudospectral radius, at each iteration the algorithm for the ϵ-pseudospectral radius performs a circular search with radius η. From the set of intersection points the set of arcs on the circle of radius η lying inside the ϵ-pseudospectrum are inferred. For each such arc a radial search with slope equal to the angle of the midpoint of the arc yields a better estimate than the current estimate. The estimate η is refined to the maximum of the values returned by the radial searches. The initial estimate for the algorithm is generated by the radial search with slope equal to the angle of an eigenvalue with modulus ρ.

C. Distance to uncontrollability

Eising [14], [15] has shown that the distance to uncontrollability defined in (7) is equivalent to

$$\tau(A,B) = \inf_{\lambda\in\mathbb{C}} \sigma_{\min}([A - \lambda I \;\; B]). \quad (12)$$

The need to search over the whole complex plane makes this optimization harder than the optimization problems above, which are defined over lines or curves.

In the following, we present a trisection algorithm that works on the δ-level set of the function $g(\lambda) = \sigma_{\min}([A - \lambda I \;\; B])$ for successive values of δ that are estimates of the distance to uncontrollability. Let $\tau(A,B) = g(\alpha_* + \beta_* i)$. We assume the *a priori* knowledge of a positive real ν satisfying $\nu > |\alpha_*|$. A well known bound is $\nu = 2(\|A\| + \|B\|)$, but it may be possible to come up with tighter bounds for special cases. We define the vertical cross section at α of the δ-level sets of $g(\lambda)$ as

$$\mathcal{S}_\delta(\alpha) = \{\beta \in \mathbb{R} : g(\alpha + \beta i) = \delta\}. \quad (13)$$

The next theorem states that for all $\alpha \in [\alpha_* - (\delta - \tau(A,B)), \alpha_* + (\delta - \tau(A,B))]$ the set $\mathcal{S}_\delta(\alpha)$ is nonempty.

Theorem 1: Let $\tau(A,B) = g(\alpha_* + \beta_* i)$ *and assume* $\delta > \tau(A,B)$ *is given. For any* $\alpha \in [\alpha_* - (\delta - \tau(A,B)), \alpha_* + (\delta - \tau(A,B))]$ *there exists a real number* $\beta_\alpha \in \mathcal{S}_\delta(\alpha)$.

Proof: The function $\sigma_{\min}([A - (\alpha' + \beta_* i)I \quad B])$ approaches ∞ as $\alpha' \to \infty$. By the continuity of $\sigma_{\min}$ as a function of α', there exists $\mu' > 0$ so that

$$\sigma_{\min}([A - (\alpha_* + \mu' + \beta_* i)I \;\; B]) = \delta.$$

Let μ_1 be the smallest such μ'. Similarly, let μ_2 be the smallest positive μ' satisfying $\sigma_{\min}([A - (\alpha_* - \mu' + \beta_* i)I \;\; B]) = \delta$. Note that for all $\alpha' \in [\alpha_* - \mu_2, \alpha_* + \mu_1]$, the inequality

$$\sigma_{\min}([A - (\alpha' + \beta_* i)I \;\; B]) \le \delta \quad (14)$$

holds. Furthermore from Weyl's theorem [16, Theorem 4.3.1], the singular values are globally Lipschitz with Lipschitz constant one. Therefore we can deduce $\mu_1 \ge \delta - \tau(A,B)$ and $\mu_2 \ge \delta - \tau(A,B)$.

Now choose any α such that $\alpha_* - (\delta - \tau(A,B)) \le \alpha \le \alpha_* + (\delta - \tau(A,B))$. Since α lies in the interval $[\alpha_* - \mu_2, \alpha_* + \mu_1]$, it follows from (14) that

$$\sigma_{\min}([A - (\alpha + \beta_* i)I \;\; B]) \le \delta. \quad (15)$$

As $\lim_{\beta\to\infty} \sigma_{\min}([A - (\alpha + \beta i)I \;\; B]) = \infty$, the continuity of the minimum singular value as a function of β together with (15) imply that for some $\beta_\alpha \ge \beta_*$, $\sigma_{\min}([A - (\alpha + \beta_\alpha i)I \;\; B]) = \delta$, as desired. ∎

Whether the set $\mathcal{S}_\delta(\alpha)$ is non-empty can be verified by checking whether

$$D(\alpha,\delta) = \begin{bmatrix} -(A^* - \alpha I) & \delta I \\ \frac{BB^*}{\delta} - \delta I & A - \alpha I \end{bmatrix}. \quad (16)$$

has an imaginary eigenvalue [17]. We call this verification *the vertical search at* α.

The algorithm maintains an interval $[L,U]$ containing the distance to uncontrollability. Putting together all the tools presented results in a technique that refines either the lower bound or the upper bound reducing the length of the interval by two-third at each iteration. Let $\delta_1 = L + 2(U-L)/3$, $\delta_2 = L + (U-L)/3$, $\delta = \delta_1$ and $\eta = 2(\delta_1 - \delta_2)$. We apply the vertical search at $-\nu, -\nu + \eta, \ldots, -\nu + \lceil \frac{2\nu}{\eta} \rceil \eta$. (Recall that ν is an upper bound on α_* in absolute value.) If any of the vertical searches returns an intersection point, then using definition (12) we can deduce the upper bound

$$\delta_1 = \delta \ge \tau(A,B).$$

If none of the vertical searches returns an intersection point, then suppose that the closest vertical line among $\alpha = -\nu + j\eta,\ j = 0, \ldots, \lceil \frac{2\nu}{\eta} \rceil$ to α_* is $\alpha = \alpha'$. Clearly

$$|\alpha' - \alpha_*| > \delta - \tau(A,B) \;\text{ and }\; |\alpha' - \alpha_*| \le \eta/2.$$

The first inequality holds, because otherwise according to Theorem 1 the set $\mathcal{S}_\delta(\alpha')$ would not be empty as verified. Meanwhile, since α' is the closest vertical line to α_*, we have the second inequality. Combining these inequalities yields

$$\delta - \tau(A,B) < \eta/2 \Longleftrightarrow \delta_2 < \tau(A,B).$$

Each iteration of the algorithm is of complexity $O(\frac{\nu}{\eta} n^3)$, since we perform the vertical search at $\lceil \frac{2\nu}{\eta} \rceil + 1$ different positions.

III. Structure-exploiting algorithms

The algorithms presented in the previous section require the extraction of imaginary or unit eigenvalues from particularly structured eigenvalue problems. In the presence of roundoff error, an eigenvalue solver that does not take special care of the eigenvalue symmetries would introduce real parts for the imaginary eigenvalues and perturb the unit eigenvalues off the unit circle. In contrast, structure-preserving algorithms are capable of returning exact imaginary or unit eigenvalues. Therefore, the need for deciding whether a nearby eigenvalue is actually on the imaginary axis or the unit circle largely disappears, making the algorithms described in the previous section more reliable.

A. Hamiltonian eigenvalue problems

The matrix $K(\theta)$ in (11) belongs to the class of Hamiltonian matrices which take the form

$$H = \begin{bmatrix} A & G \\ Q & -A^* \end{bmatrix}, \quad G = G^*, \quad Q = Q^*$$

where $A, G, Q \in \mathbb{C}^{n\times n}$. Eigenvalues of such matrices are always symmetric with respect to the imaginary axis, i.e., if λ is an eigenvalue of H then $-\bar{\lambda}$ is also an eigenvalue. Hence, if λ is a simple purely imaginary eigenvalue it will stay on the imaginary axis when computed by structure-preserving algorithms; otherwise the eigenvalue symmetry would be broken. MATLAB functions providing such algorithms are readily available in HAPACK [18].

B. Symplectic eigenvalue problems

Considering the matrix pencil $P(r,\epsilon) - \lambda Q(r,\epsilon)$ in (10) upon which we must decide whether it has eigenvalues on the unit circle, we can see that its conjugate transpose, the pencil

$$\lambda E_S - A_S = \lambda \begin{bmatrix} 0 & A \\ rI & -\epsilon I \end{bmatrix} - \begin{bmatrix} -\epsilon I & rI \\ A^* & 0 \end{bmatrix}$$

satisfies $E_S J E_S^* = A_S J A_S^*$ with the matrix $J = \left[\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right]$. This relationship makes $\lambda E_S - A_S$ a *symplectic matrix pencil*. The same property holds for the pencil $R(\mu) - \lambda S$ in (8). Eigenvalues of such matrix pencils come in pairs $\{\lambda, 1/\bar{\lambda}\}$ (a zero eigenvalue pairs with an infinite eigenvalue). Similarly as for the Hamiltonian case, eigenvalue solvers that are capable to preserve these eigenvalue pairings in finite-precision arithmetic can be expected to produce more reliable results for eigenvalues on or close to the unit circle. A number of existing methods are based on symplectic but non-orthogonal transformations, see [19] for an overview. Patel's algorithm [20], which covers *real* symplectic pencils having a particular block structure, employs orthogonal transformations after a so called $S + S^{-1}$ transform. In this section, we focus on a method based on the generalized Cayley transform [21] and propose a new extension to the complex case.

The generalized Cayley transform transforms a matrix pencil $\lambda E_S - A_S$ into

$$\lambda E_H - A_H := \lambda(E_S - A_S) - (E_S + A_S).$$

The generalized Cayley transform preserves the eigenvalues in the sense that λ is an eigenvalue of $\lambda E_H - A_H$ if and only if $(\lambda+1)/(\lambda-1)$ is an eigenvalue of $\lambda E_S - A_S$. The fact that $\lambda E_S - A_S$ is symplectic implies $E_H J A_H^* = -A_H J E_H^*$, i.e., $\lambda E_H - A_H$ is a *Hamiltonian matrix pencil*.

1) The real case: Given a *real* pencil $\lambda E_H - A_H$, we can apply the numerically backward stable, structured algorithm for Hamiltonian matrix pencils described in [21] to compute the eigenvalues of $\lambda E_H - A_H$.

This algorithm proceeds as follows. First, orthogonal symplectic matrices U_1, U_2 (an orthogonal symplectic matrix U satisfies $U^T U = I$ *and* $U^T J U = J$), and an orthogonal matrix U_2 are computed such that

$$U_3^T E_H U_1 = \begin{bmatrix} E_{11} & E_{12} \\ 0 & E_{22}^T \end{bmatrix} = \begin{bmatrix} \ddots & \square \\ & \ddots \end{bmatrix},$$

$$U_3^T A_H U_2 = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}^T \end{bmatrix} = \begin{bmatrix} \ddots & \square \\ & \ddots \end{bmatrix},$$

i.e., A_{11}, E_{11}, E_{22} are upper triangular while A_{22} is upper Hessenberg. Since

$$(\lambda E_H - A_H) J (\lambda E_H^T - A_H^T) = \lambda^2 E_H J E_H^T - A_H J A_H^T,$$

and

$$U_3^T E_H J E_H^T U_3 J = \begin{bmatrix} -E_{11}E_{22} & * \\ 0 & -E_{22}^T E_{11}^T \end{bmatrix},$$

$$U_3^T A_H J A_H^T U_3 J = \begin{bmatrix} A_{11}A_{22} & * \\ 0 & A_{22}^T A_{11}^T \end{bmatrix},$$

it follows that the eigenvalues of $\lambda E_H - A_H$ can be obtained as the square roots of the eigenvalues of the matrix pencil $\lambda E_{11}E_{22} + A_{11}A_{22}$. The periodic QZ algorithm [22], [23] can be used to compute the eigenvalues of the latter pencil in a numerically backward stable manner.

2) The complex case: In this section, we will present a structure-preserving algorithm for addressing a *complex* Hamiltonian matrix pencil $\lambda E_H - A_H$, by combining ideas from [21] and [24]. It will be necessary to assume that E_H is not (close to) a singular matrix. This can be guaranteed by applying structured staircase algorithms [25] beforehand. Let us define an embedded real $4n \times 4n$ pencil $\lambda E_W - A_W$ by

$$\lambda \left[\begin{array}{cc|cc} E_{11,R} & -E_{11,I} & E_{12,R} & -E_{12,I} \\ -E_{11,I} & -E_{11,R} & -E_{12,I} & -E_{12,R} \\ \hline E_{21,R} & -E_{21,I} & E_{22,R} & -E_{22,I} \\ -E_{21,I} & -E_{21,R} & -E_{22,I} & -E_{22,R} \end{array}\right] - \left[\begin{array}{cc|cc} A_{11,I} & A_{11,R} & A_{12,I} & A_{12,R} \\ A_{11,R} & -A_{11,I} & A_{12,R} & -A_{12,I} \\ \hline A_{21,I} & A_{21,R} & A_{22,I} & A_{22,R} \\ A_{21,R} & -A_{21,I} & A_{22,R} & -A_{22,I} \end{array}\right],$$

where $E_H = \left[\begin{smallmatrix} E_{11,R} & E_{12,R} \\ E_{21,R} & E_{22,R} \end{smallmatrix}\right] + \imath \left[\begin{smallmatrix} E_{11,I} & E_{12,I} \\ E_{21,I} & E_{22,I} \end{smallmatrix}\right]$ and $A_H = \left[\begin{smallmatrix} A_{11,R} & A_{12,R} \\ A_{21,R} & A_{22,R} \end{smallmatrix}\right] + \imath \left[\begin{smallmatrix} A_{11,I} & A_{12,I} \\ A_{21,I} & A_{22,I} \end{smallmatrix}\right]$ with all submatrices $E_{ij,*}$, $A_{ij,*}$ being real and $n \times n$. It can be shown that $E_H J A_H^* = -A_H J E_H^*$ implies $E_W J A_W^T = A_W J E_W^T$, which tempts us to call $\lambda E_W - A_W$ a *skew-Hamiltonian matrix pencil*. Moreover, λ is an eigenvalue of $\lambda E_H - A_H$ if and only if $\{-\imath\lambda, \imath\bar{\lambda}\}$ is an eigenvalue pair of $\lambda E_W - A_W$.

It remains to develop a structure-preserving algorithm for skew-Hamiltonian pencils. First, an orthogonal matrix U_1 and an orthogonal symplectic matrix U_2 are computed such that

$$U_1^T E_W U_2 = \begin{bmatrix} E_{11} & E_{12} \\ 0 & E_{22}^T \end{bmatrix} = \begin{bmatrix} \ddots & \square \\ & \ddots \end{bmatrix}, \quad (17)$$

$$U_1^T A_W U_2 \quad = \quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}^T \end{bmatrix} = \begin{bmatrix} \diagdown & \square \\ \diagdown & \square \end{bmatrix}, \quad (18)$$

i.e., A_{12}, E_{11}, E_{22} are upper triangular while A_{11} is upper Hessenberg. A modification of the standard Hessenberg-triangular reduction algorithm [26] which achieves the form (17) and (18) is described in the appendix. An important observation, which makes the form (17) and (18) useful for eigenvalue computations, is that the relationship $E_W J A_W^T = A_W J E_W^T$ is preserved under the reduction to the form (17) and (18). This implies $A_{21} E_{22} = 0$ (which, together with the nonsingularity of E_{22}, yields $A_{21} = 0$) and $E_{11} A_{22} = A_{11} E_{22}$. Hence, the eigenvalues of $\lambda E_{11} - A_{11}$, which can be computed by the QZ algorithm [26], are the eigenvalues of $\lambda E_W - A_W$ with halved multiplicities. Since E_{11} and A_{11} are real, these eigenvalues come in pairs $\{\mu, \bar{\mu}\}$, each of which corresponds to an eigenvalue pair $\{\imath\mu, \imath\bar{\mu}\}$ of the Hamiltonian pencil $\lambda E_H - A_H$.

IV. Advantages from preserving structure

In the presence of rounding errors algorithms that respect the symmetry of eigenvalues such as the ones in the previous section move simple imaginary eigenvalues of a Hamiltonian matrix only along the imaginary axis and simple unit eigenvalues of a symplectic pencil only along the unit circle. This property for the eigenvalue solvers used inside the algorithms in Section II is essential to compute the robust stability and controllability measures in a reliable fashion. In Figure 1.a)[1] we see that for $\epsilon = 0.1$ the circle of radius 2 intersects the boundary of the ϵ-pseudospectrum of the twisted Toeplitz matrix available in *EigTool*'s demo menu [27] at 8 distinct points. The twisted Toeplitz is a tridiagonal matrix with fixed subdiagonal and superdiagonal entries and the main diagonal allowed to vary smoothly. For this example our implementation of the symplectic eigenvalue solver returns 8 eigenvalues (each with algebraic multiplicity 2) with unit moduli. These values multiplied by $r = 2$ correspond to the intersection points as displayed by asterisks in Figure 1.a). The moduli of the eigenvalues corresponding to the intersection points computed by `eig` differ from one by about 4×10^{-15} . To compute the distance to uncontrollability for A chosen to be the 5×5 matrix that can be obtained in MATLAB by typing `gallery(5)`, B a 5×1 matrix with entries selected from a normal distribution with zero mean and variance one, we need the imaginary eigenvalues of the Hamiltonian matrix $D(\alpha, \delta)$ for various α and δ. For $\delta = 10^{-0.8421}$ and $\alpha = 0.963980$, the MATLAB routine `eig` introduces significant real parts of order 10^{-7} for the imaginary eigenvalues due to rounding. The vertical line at α intersects the δ-level set of $\sigma_{\min}([A - \lambda I \quad B])$ at four distinct points. The real Hamiltonian eigensolver in HAPACK [18] returns the pairs $\pm 0.47623i$ and $\pm 0.47621i$. The intersection points are marked by asterisks in Figure 1.b). When a general-purpose eigenvalue solver is used, a problem-dependent tolerance is needed to decide whether the departure of an eigenvalue from the imaginary axis or unit circle is solely caused by rounding errors.

[1] All of the figures in this section are generated by *EigTool*.

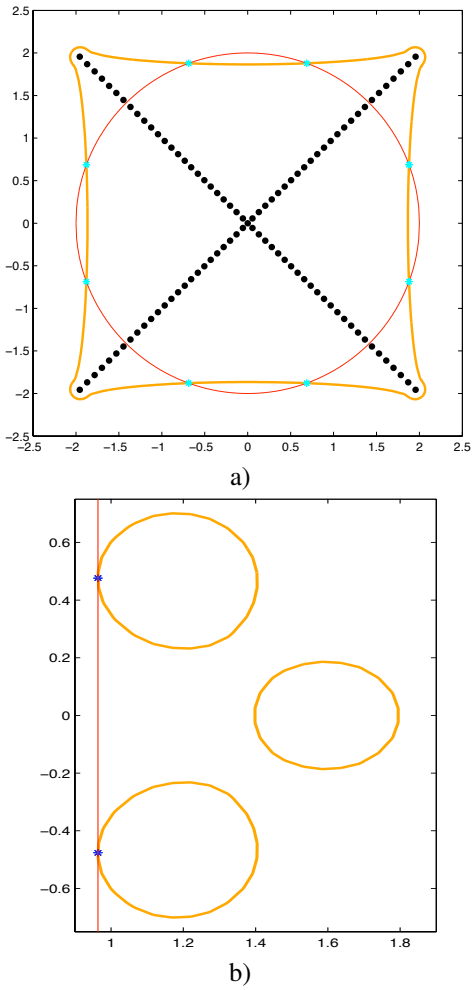

Fig. 1. The closed curves consist of the boundaries of the ϵ-pseudospectrum for the twisted Toeplitz matrix ($\epsilon = 0.1$) and the δ-level curve of $\sigma_{\min}([A - \lambda I \quad B])$ for the `Gallery5`-normally distributed matrix pair ($\delta = 10^{-0.8421}$), respectively. a) The circle of radius 2 intersects the ϵ-pseudospectrum boundary at eight points. b) The intersection points of the vertical line at $\alpha = 963980$ and the level curve are marked with asterisks.

An eigenvalue solver is called numerically *strongly backward stable* if it returns the exact eigenvalues of a nearby problem having the *same* structure. Although observed in practice, it is not yet clear whether the algorithms presented in Section III share this very desirable property.

In [6], the ϵ-pseudospectral radius computation was proved to be backward stable when strongly backward stable eigenvalue solvers are used. Here we analyze the accuracy of the presented algorithm for computing the distance to uncontrollability. When applied to the Hamiltonian matrix $D(\alpha, \delta)$ defined in (16), we assume that the eigenvalue solver returns the eigenvalues of $\hat{D}(\alpha, \delta) = D(\alpha, \delta) + E$, for some Hamiltonian matrix E with norm of the order of machine precision. For the proofs of the results below see the appendix. The first result is related to the error of the computed upper bound. For notational convenience, we reveal the dependence of D

on B by writing $D_B(\alpha,\delta)$.

Theorem 2: Suppose iy is an imaginary eigenvalue of $\hat{D}_B(\alpha,\delta)$. Then iy is also an eigenvalue of $D_{\left(\sqrt{\frac{\delta+\zeta}{\delta}}\right)B}(\alpha,\delta+\zeta)$ for some ζ with $|\zeta|\le\|E\|$.

An immediate implication of the theorem above is that there exists a β such that $\sigma_{\min}\left(\left[A-(\alpha+\beta)i\ \left(\sqrt{\delta+\zeta}/\sqrt{\delta}\right)B\right]\right) = \delta+\zeta$ or by applying Weyl's Theorem once again, we obtain

$$\begin{aligned}\tau(A,B) &\le \sigma_{\min}([A-(\alpha+\beta)i\ B]) \le \\ \delta+\zeta+\frac{|\sqrt{\delta+\zeta}-\sqrt{\delta}|}{\sqrt{\delta}}\|B\| &= \delta+O(|\zeta|(1+\|B\|/(2\delta))).\end{aligned}$$

where $|\zeta|\le\|E\|$. Thus, the error in the computed upper bound δ is $O(\|E\|(1+\|B\|/(2\delta)))$.

The next result concerns the case when none of the vertical lines at $\alpha=-\nu+\eta j,\ \ j=0,\ldots,\lceil\frac{2\nu}{\eta}\rceil$ intersects the δ-level sets of $\sigma_{\min}([A-\lambda I\ B])$ and therefore the lower bound.

Theorem 3: Suppose that $\hat{D}_B(\alpha,\delta)$ does not have any imaginary eigenvalue and the matrix A does not have an eigenvalue with real part equal to α. For all ζ such that $\|E\|\le\zeta\le\delta$ the matrix $D_{\left(\sqrt{\frac{\delta-\zeta}{\delta}}\right)B}(\alpha,\delta-\zeta)$ does not have any imaginary eigenvalue.

In particular let us apply the theorem above with $\zeta=\|E\|$ when for all $\alpha=-\nu+\eta j,\ \ j=0,\ldots,\lceil\frac{2\nu}{\eta}\rceil$ the matrix $\hat{D}_B(\alpha,\delta)$ does not have any imaginary eigenvalue. The theorem implies that for A and $\left(\sqrt{\delta-\|E\|}/\sqrt{\delta}\right)B$ the vertical cross sections at all α of the $(\delta-\|E\|)$-level sets are empty. We can conclude from Theorem 1 that

$$\begin{aligned}2(\delta-\|E\|-\tau(A,\left(\sqrt{\frac{\delta-\|E\|}{\delta}}\right)B)) \le \ \ \eta \Longrightarrow \\ \delta-\|E\|-\eta/2 \le \tau(A,\left(\sqrt{\frac{\delta-\|E\|}{\delta}}\right)B)) \\ \le \tau(A,B)+\frac{\sqrt{\delta}-\sqrt{\delta-\|E\|}}{\sqrt{\delta}}\|B\|,\end{aligned}$$

where the last inequality follows from Weyl's Theorem. In this case we update the lower bound to $\delta-\eta/2$ which differs from an exact lower bound by $O(\|E\|(1+\|B\|/(2\delta)))$.

V. Conclusions

We have reviewed methods for computing the pseudospectral and numerical radius of a matrix. Also, a new method for estimating the distance to uncontrollability of a matrix pair has been introduced and analyzed. It has been shown how existing structure-exploiting algorithms for the eigenvalue problems arising from these methods can be extended to cover complex symplectic pencils.

References

[1] S. Boyd, V. Balakrishnan, and P. Kabamba, "A bisection method for computing the $\mathcal{H}_\infty$ norm of a transfer matrix and related problems," *Math. Control, Signals, Sys.*, vol. 2, pp. 207–219, 1989.

[2] J. V. Burke, A. S. Lewis, and M. L. Overton, "Robust stability and a criss-cross algorithm for pseudospectra," *IMA J. Numer. Anal.*, vol. 23, no. 3, pp. 359–375, 2003.

[3] R. Byers, "A bisection method for measuring the distance of a stable to unstable matrices," *SIAM J. Sci. Statist. Comput.*, vol. 9, pp. 875–881, 1988.

[4] M. Gu, "New methods for estimating the distance to uncontrollability," *SIAM J. Matrix Analysis Appl.*, vol. 21, no. 3, pp. 989–1003, 2000.

[5] D. Hinrichsen and A. J. Pritchard, "Stability radii of linear systems," *Systems Control Lett.*, vol. 7, no. 1, pp. 1–10, 1986.

[6] E. Mengi and M. L. Overton, "Algorithms for the computation of the pseudospectral radius and the numerical radius of a matrix," *IMA J. of Numerical Anal.*, vol. 25, no. 4, pp. 648–669, 2005.

[7] D. Hinrichsen and A. J. Pritchard, *Mathematical Systems Theory I*, ser. Texts in Applied Mathematics. Springer-Verlag, 2005, vol. 48.

[8] L. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices*. Princeton Univ. Press, 2005.

[9] H.-O. Kreiss, "Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differenzialgleichungen approximieren," *BIT*, vol. 2, pp. 153–181, 1962.

[10] P. H. Petkov, N. D. Christov, and M. M. Konstantinov, *Computational Methods for Linear Control Systems*. Hertfordshire, UK: Prentice-Hall, 1991.

[11] C. Paige, "Properties of numerical algorithms relating to controllability," *IEEE Trans. Automat. Control*, vol. AC-26, pp. 130–138, 1981.

[12] S. Boyd and V. Balakrishnan, "A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm," *Sys. Control Lett.*, vol. 15, pp. 1–7, 1990.

[13] N. A. Bruinsma and M. Steinbuch, "A fast algorithm to compute the H_∞-norm of a transfer function matrix," *Sys. Control Lett.*, vol. 14, no. 4, pp. 287–293, 1990.

[14] R. Eising, "The distance between a system and the set of uncontrollable systems," in *Memo COSOR 82-19, Eindhoven Univ. Technol., Eindhoven, The Netherlands*, 1982.

[15] ——, "Between controllable and uncontrollable," *Sys. Control Lett.*, vol. 4, pp. 263–264, 1984.

[16] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[17] R. Byers, "Detecting nearly uncontrollable pairs," in *Numerical Methods Proceedings of the International Symposium MTNS-89*, vol. III. M.A. Kaashoek, J.H. van Schuppen, and A.C.M. Ran, eds., Springer-Verlag, 1990, pp. 447–457.

[18] P. Benner and D. Kressner, "Algorithm 854: Fortran 77 subroutines for computing the eigen values of Hamiltonian matrices II," *ACM Trans. Math. Software*, vol. 32, no. 2, pp. 352–373, 2006, see also http://www.tu-chemnitz.de/mathematik/hapack/.

[19] H. Faßbender, *Symplectic methods for the symplectic eigenproblem*. Kluwer Academic/Plenum Publishers, New York, 2000.

[20] R. V. Patel, "On computing the eigenvalues of a symplectic pencil," *Linear Algebra Appl.*, vol. 188/189, pp. 591–611, 1993.

[21] P. Benner, V. Mehrmann, and H. Xu, "A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils," *Numerische Mathematik*, vol. 78, no. 3, pp. 329–358, 1998.

[22] A. Bojanczyk, G. H. Golub, and P. Van Dooren, "The periodic Schur decomposition; algorithm and applications," in *Proc. SPIE Conference*, vol. 1770, 1992, pp. 31–42.

[23] J. J. Hench and A. J. Laub, "Numerical solution of the discrete-time periodic Riccati equation," *IEEE Trans. Automat. Control*, vol. 39, no. 6, pp. 1197–1210, 1994.

[24] P. Benner, V. Mehrmann, and H. Xu, "A note on the numerical solution of complex Hamiltonian and skew-Hamiltonian eigenvalue problems," *Electron. Trans. Numer. Anal.*, vol. 8, pp. 115–126, 1999.

[25] R. Byers, V. Mehrmann, and H. Xu, "A structured staircase algorithm for skew-symmetric/symmetric pencils," MATHEON, DFG Research Center "Mathematics for key technologies" in Berlin, TU Berlin," Technical Report, 2005.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins University Press, 1996.

[27] T. Wright, "Eigtool: a graphical tool for nonsymmetric eigenproblems," Oxford University Computing Laboratory. See http://www.comlab.ox.ac.uk/pseudospectra/eigtool/.

Appendix

A motivating example for the robust stability estimates To illustrate the relevance of the robust stability measures in this work with the transient behavior of a discrete system , let us consider two matrices that have similar spectral properties but different transient behaviors.

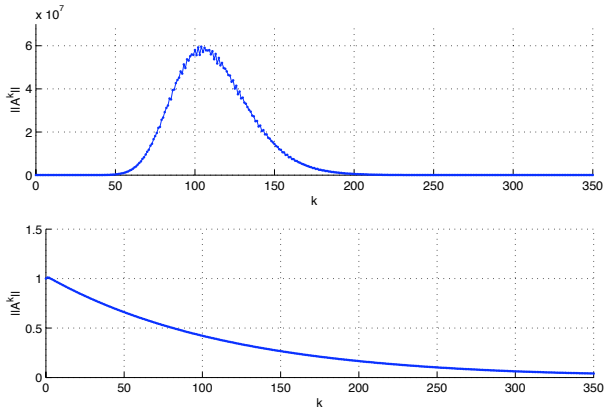

Fig. 2. The norms of A^k for the scaled Grcar matrix at the top and the scaled Airy matrix at the bottom.

The 100×100 "Grcar" matrix is a Toeplitz matrix with -1 on the subdiagonal, 1 on the diagonal, first, second and third superdiagonals and all the other entries 0. The 50×50 Airy matrix is obtained by discretizing the differential operator $L(u) = \frac{d^2u}{dx^2} + ixu$. Both of the matrices are available in *EigTool*, a graphical MATLAB tool to compute pseudospectra and related measures [27]. We multiply both of the matrices by constants so that their eigenvalues are contained inside the unit circle with the spectral radii for the scaled Grcar matrix and the scaled Airy matrix equal to 0.9051 and 0.9907, respectively. This in turn means that the eventual decay for the scaled Grcar matrix is faster than the Airy matrix. From Figure 2, it is apparent that the norms of the powers of the Grcar matrix reach 10^7 before any decay takes place, while the powers of the Airy matrix never exceed 1.5 in norm. The poor transient behavior of the Grcar matrix can be estimated by the ϵ-pseudospectral radius of the Grcar matrix as $\rho_\epsilon(A) \approx 1.0321$ and the ratio $\frac{\rho_\epsilon(A)-1}{\epsilon} \approx 3.2138 \times 10^6$ for $\epsilon = 10^{-8}$. Therefore the initial growth on the top in Figure 2 agrees with the bounds in (5). Another indicator of the the poor transient behavior of the Grcar matrix is its numerical radius (≈ 1.2941), which is considerably larger than the numerical radius of the Airy matrix (≈ 1.0025).

REDUCTION OF SKEW-HAMILTONIAN MATRIX PENCILS

A reduction of skew-Hamiltonian pencil to the form displayed in (17) and (18) can be achieved by Algorithm 4. This algorithm makes use of Givens rotation matrices which have the form

$$G_{ij}(\theta) = \begin{bmatrix} I_{i-1} & & & & \\ & \cos\theta & & \sin\theta & \\ & & I_{j-i-1} & & \\ & -\sin\theta & & \cos\theta & \\ & & & & I_{m-j} \end{bmatrix},$$

for a given i, j such that $1 \le i < j \le m$ and for some angle $\theta \in [-\pi/2, \pi/2)$. The angle can always be chosen so that the jth component of $G_{ij}(\theta)x$ is zero for a fixed vector $x \in \mathbb{R}^n$. In this case, we identify $G_{ij}(\theta) \equiv G_{ij}(x)$. For $i < j$, we use the notation $G_{ij}(\theta) \equiv G_{ji}(x)$ to identify the Givens rotation which maps the ith component of $G_{ij}(\theta)^T \cdot x$ to zero. Note that the direct sum of two identical Givens rotations, $G_{ij}(x) \oplus G_{ij}(x)$ as well as the $2m \times 2m$ Givens rotation $G_{i,m+i}(x)$ are orthogonal symplectic matrices.

Algorithm 4:
Hessenberg-triangular-like reduction

Input: A matrix $E \in \mathbb{R}^{2m \times 2m}$ having the block triangular form displayed in (17), a matrix $A \in \mathbb{R}^{2m \times 2m}$.

Output: An orthogonal symplectic matrix $U_1 \in \mathbb{R}^{2m \times 2m}$ and an orthogonal symplectic matrix $U_2 \in \mathbb{R}^{2m \times 2m}$ such that $U_1^T E U_2$ and $U_1^T A U_2$ have the reduced form displayed in (17) and (18) respectively. The matrices E and A are overwritten by the reduced matrices.

[1.] Compute an orthogonal matrix U_1 such that $E \leftarrow U_1^T E$ has the form displayed in (17) (can be achieved by a slightly modified QR decomposition [21]).

[2.] $U_2 \leftarrow I_{2m}$
for $j \leftarrow 1, 2, \ldots, m-1$ **do**
  % Annihilate entries in jth column of $(2,1)$ block of A.
  for $i \leftarrow j, j+1, \ldots, m-1$ **do**
    $G \leftarrow I_m \oplus G_{i+1,i}(A(m+1:2m, j))$
    $A \leftarrow GA,\ E \leftarrow GE,\ U_1 \leftarrow U_1 G^T$
  end for
  % Propagate transformations through E while preserving its triangular shape.
  for $i \leftarrow j, j+1, \ldots, m-1$ **do**
    $G \leftarrow G_{i,i+1}(E(i, m+1:2m)) \oplus G_{i,i+1}(E(i, m+1:2m))$
    $A \leftarrow AG^T,\ E \leftarrow EG^T,\ U_2 \leftarrow U_2 G^T$
    $G \leftarrow G_{i,i+1}(E(1:m, i)) \oplus I_m$
    $A \leftarrow GA,\ E \leftarrow GE,\ U_1 \leftarrow U_1 G^T$
  end for
  % Annihilate $(2m, j)$ entry of A.
  $G \leftarrow G_{m,2m}(A(:, j))$
  $A \leftarrow GA,\ E \leftarrow GE,\ U_1 \leftarrow U_1 G^T$
  $G \leftarrow G_{2m,m}(E(2m, :))$
  $A \leftarrow AG^T,\ E \leftarrow EG^T,\ U_2 \leftarrow U_2 G^T$
  % Annihilate entries in jth column of $(1,1)$ block of A.
  for $i \leftarrow m-1, m-2, \ldots, j+1$ **do**
    $G \leftarrow I_m \oplus G_{i,i+1}(A(1:m, j))$
    $A \leftarrow GA,\ E \leftarrow GE,\ U_1 \leftarrow U_1 G^T$
  end for
  % Propagate transformations through E while preserving its triangular shape.
  for $i \leftarrow m-1, m-2, \ldots, j+1$ **do**

$G \leftarrow G_{i+1,i}(E(i+1,1:m)) \oplus G_{i+1,i}(E(i+1,1:m))$
$A \leftarrow AG^T,\ E \leftarrow EG^T,\ U_2 \leftarrow U_2G^T$
$G \leftarrow G_{i+1,i}(E(1:m,i+1)) \oplus I_m$
$A \leftarrow GA,\ E \leftarrow GE,\ U_1 \leftarrow U_1G^T$
end for
end for

PROOFS OF THE THEOREMS FOR THE ACCURACY OF THE ALGORITHM FOR THE DISTANCE TO UNCONTROLLABILITY

When a vertical search applied by the algorithm succeeds in finite precision, the vertical search in exact arithmetic succeeds for a slightly perturbed level set as shown next.

Theorem 5: Suppose iy is an imaginary eigenvalue of $\hat{D}_B(\alpha,\delta)$. Then iy is also an eigenvalue of $D_{\left(\sqrt{\frac{\delta+\zeta}{\delta}}\right)B}(\alpha,\delta+\zeta)$ for some ζ with$|\zeta| \leq \|E\|$.

Proof: Since iy is an eigenvalue of $\hat{D}_B(\alpha,\delta)$,

$$0 = \det(D_B(\alpha,\delta)+E-iy) = \det(JD_B(\alpha,\delta)+JE-Jiy).$$

The matrices JE and $JD_B(\alpha,\delta) - Jiy$ are Hermitian. Therefore Weyl's Theorem [16, Theorem 4.3.1] implies that the matrix $JD_B(\alpha,\delta) - Jiy$ has an eigenvalue less than or equal to $\|E\|$ in absolute value, *i.e.*

$$\begin{aligned} 0 &= \det(JD_B(\alpha,\delta) - Jiy - \zeta I) \\ &= \det(JD_{\left(\sqrt{\frac{\delta+\zeta}{\delta}}\right)B}(\alpha,\delta+\zeta) - Jiy). \end{aligned}$$

holds for some ζ such that $|\zeta| \leq \|E\|$. Since J is unitary, it follows that iy is an eigenvalue of $D_{\left(\sqrt{\frac{\delta+\zeta}{\delta}}\right)B}(\alpha,\delta+\zeta)$. ■

Similarly when a vertical search fails in finite precision, the vertical seach in exact arithmetic fails on a nearby level set which is proved below.

Theorem 6: Suppose $\hat{D}_B(\alpha,\delta)$ does not have any imaginary eigenvalue and the matrix A does not have an eigenvalue with real part equal to α. For all ζ such that $\|E\| \leq \zeta \leq \delta$ the matrix $D_{\left(\sqrt{\frac{\delta-\zeta}{\delta}}\right)B}(\alpha,\delta-\zeta)$ does not have any imaginary eigenvalue.

Proof: Consider the eigenvalues of $JD_B(\alpha,\delta)+\delta I - iyJ$ and $J\hat{D}_B(\alpha,\delta)+\delta I-iyJ$ as functions of $y \in \mathbb{R}$ which we denote in ascending order by $\varepsilon_j(y),\ \ j=1,\ldots,2n$ and $\hat{\varepsilon}_j(y),\ \ j=1,\ldots,2n$, respectively. Since the Hermitian matrix $J\hat{D}_B(\alpha,\delta) + \delta I - iyJ$ is obtained by perturbing $JD_B(\alpha,\delta) + \delta I - iyJ$ by JE, an immediate consequence of Weyl's Theorem [16, Theorem 4.3.1] is that for all j and y

$$|\varepsilon_j(y) - \hat{\varepsilon}_j(y)| \leq \|E\|. \tag{19}$$

The fact that for all $y \in \mathbb{R}$, the imaginary number iy is not an eigenvalue of $\hat{D}_B(\alpha,\delta)$ implies

$$0 \neq \det(J\hat{D}_B(\alpha,\delta)-iyJ) = \det(J\hat{D}_B(\alpha,\delta)+\delta I-iyJ-\delta I)$$

that is for all y and j, $\hat{\varepsilon}_j(y) \neq \delta$. Indeed as $\hat{\varepsilon}_j(y)$ depends on y continuously, either $\forall y,\ \ \hat{\varepsilon}_j(y) > \delta$ or $\forall y,\ \ \hat{\varepsilon}_j(y) < \delta$. If the first case is satisfied, from (19) we can deduce that $\forall y,\ \ \varepsilon_j(y) > \delta - \|E\|$. Otherwise, $\forall y,\ \ \varepsilon_j(y) < \delta + \|E\|$.

By assumption A does not have an eigenvalue with the real part equal to α, so there cannot exist a y such that

$$\begin{aligned} 0 &= \det\left(\begin{bmatrix} BB^*/\delta & A-(\alpha+iy)I \\ A^*-(\alpha-iyI) & 0 \end{bmatrix}\right) \\ &= \det(JD(\alpha,\delta)+\delta I - iyJ) \end{aligned}$$

which means $\varepsilon_j(y) \neq 0$ for all y as well. Therefore by the continuity of ε either $\forall y,\ \ \varepsilon_j(y) > 0$ or otherwise $\forall y,\ \ \varepsilon_j(y) < 0$ holds. Regarding where the function $\varepsilon_j(y)$ lies, so far we have shown the following three possibilities

- $\forall y, \varepsilon_j(y) > \delta - \|E\|$,
- $\forall y, \varepsilon_j(y) < 0$,
- $\forall y, \delta + \|E\| > \varepsilon_j(y) > 0$.

But we can also eliminate the third possibility, because clearly the function $\varepsilon_j(y)$ approaches ∞ as y goes to ∞. Therefore the function $\varepsilon_j(y)$ cannot take values in the range $[0, \delta-\|E\|]$. Let ζ be contained in $[\|E\|, \delta]$, by the definition of $\varepsilon_j(y)$ for all y we have

$$\begin{aligned} 0 &\neq \det(JD_B(\alpha,\delta)+\delta I - iyJ - (\delta-\zeta)I) \\ &= \det(JD_{\left(\sqrt{\frac{\delta-\zeta}{\delta}}\right)B}(\alpha,\delta-\zeta) - iyJ). \end{aligned}$$

By multiplying the matrix inside the last determinent by $-J$ from left, we obtain the result. ■