On the Shell and the Shell-Extremal Eigenvalues of a Square Matrix

Christos Chorianopoulos* and Panayiotis Psarrakos†

February 13, 2023

Abstract

The shell of a matrix is a cubic curve that provides interesting spectral localization results. We study its geometry in the case where the shell has a closed branch that surrounds a simple extremal eigenvalue of the matrix. Several quantities which are related to the closed branch are introduced and studied as measures of non-normality of that particular eigenvalue.

Key words: shell, cubic curve, numerical range, extremal eigenvalues.

AMS Subject Classifications: 15A18, 15A60.

1 Introduction and preliminaries

It is well known that the real part of each eigenvalue of a square complex matrix A is less than or equal to the largest eigenvalue, say $\delta_1(A)$, of the hermitian part of A. As a consequence, the spectrum of A lies to the left of the vertical (straight) line $\{z \in \mathbb{C} : \operatorname{Re}(z) = \delta_1(A)\}$. Adam and Tsatsomeros [1], extending a methodology of [8, 11], introduced a cubic curve that yields a better localization of the spectrum of A than the above vertical line. The purpose of this paper is to study further this cubic curve in the case where it is not connected and has a closed branch which surrounds one simple eigenvalue of A.

Let $\mathbb{C}^n$ and $\mathbb{C}^{m\times n}$ denote the n-th dimensional complex vector space and the set of $m \times n$ complex matrices, respectively. For a square matrix $A \in \mathbb{C}^{n\times n}$, let $H(A) = \dfrac{A + A^*}{2}$ be the *hermitian part* of A and $K(A) = \dfrac{A - A^*}{2}$ be the *skew-hermitian part* of A. The eigenvalues of $H(A)$ are denoted by $\delta_1(A) \geq \delta_2(A) \geq \cdots \geq \delta_n(A)$, in non-increasing order, and $\mathbf{y}_1 \in \mathbb{C}^n$ is a unit eigenvector of $H(A)$ corresponding to $\delta_1(A)$. The spectrum of A is denoted by $\sigma(A)$, and an eigenvalue $\lambda_0 \in \sigma(A)$ is called *normal* if its geometric multiplicity is equal to its algebraic multiplicity (i.e., λ_0 is a *semi-simple eigenvalue*) and its eigenspace is orthogonal to the eigenspaces of all the rest eigenvalues of A. Moreover, an eigenvalue of A that lies on the boundary of the convex hull of the spectrum $\sigma(A)$ is called an *extremal eigenvalue* of A.

*Department of Electrical and Electronics Engineering, University of West Attica, Egaleo, Greece (cchorian@uniwa.gr).

†Department of Mathematics, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Athens, Greece (ppsarr@math.ntua.gr).

The *numerical range* of a matrix $A \in \mathbb{C}^{n\times n}$ is the compact and convex set [7]

$$F(A) = \{\mathbf{x}^* A\mathbf{x} \in \mathbb{C} : \ \mathbf{x} \in \mathbb{C}^n \ \text{ with } \ \mathbf{x}^*\mathbf{x} = 1\} \subset \mathbb{C}.$$

The region $F(A)$ contains the eigenvalues of A, and for A normal, it reduces to their convex hull. An eigenvalue on the boundary of the numerical range $F(A)$ is a normal eigenvalue of A [7], and apparently, it is an extremal eigenvalue of A. A scalar $\mu \in \mathbb{C}$ on the boundary of the numerical range $F(A)$ is called a *corner* (also called a *sharp point*) of $F(A)$ if there exist real numbers θ_1, θ_2, with $0 \le \theta_1 < \theta_2 < 2\pi$, such that $\mathrm{Re}\left(e^{\mathrm{i}\theta}\mu\right) = \max\left\{\mathrm{Re}(z) : \ z \in F\left(e^{\mathrm{i}\theta}A\right)\right\}$ for all $\theta \in (\theta_1, \theta_2)$. Such corners of $F(A)$ are eigenvalues of A on the boundary of $F(A)$ [7]. The numerical range has stirred the interest of many researchers from the areas of matrix analysis and operator theory for about a century. A standard reference on the topic is [7, Chapter 1].

Generalizing the spectral localization results found in [8, 11], Adam and Tsatsomeros in [1] introduced the cubic curve

$$\Gamma(A) = \{z = x + \mathrm{i}\, y : \ x, y \in \mathbb{R} \ \text{such that} \ f_A(x, y) = 0\},$$

with

$$f_A(x,y) = \left[(\delta_1(A) - x)^2 + (u(A) - y)^2\right](\delta_2(A) - x) + (\delta_1(A) - x)\left(\nu(A) - u(A)^2\right),$$

where $\nu(A) = \|K(A)\mathbf{y}_1\|_2^2$ and $u(A) = \mathrm{Im}\,(\mathbf{y}_1^* A\mathbf{y}_1)$; recall that $\mathbf{y}_1 \in \mathbb{C}^n$ is a unit eigenvector of $H(A)$ corresponding to $\delta_1(A)$, and note that $\|\cdot\|_2$ denotes the 2-norm and $|u(A)| \le |\mathbf{y}_1^* K(A)\mathbf{y}_1| \le \sqrt{\nu(A)}$. The cubic curve $\Gamma(A)$ is called the *shell* of A, it lies in the vertical zone $\mathcal{Z} = \{z \in \mathbb{C} : \ \delta_2(A) \le \mathrm{Re}(z) \le \delta_1(A)\}$ of the complex plane, it is symmetric with respect to the horizontal line

$$\mathcal{L} = \{s + \mathrm{i}\, u(A) : \ s \in \mathbb{R}\},$$

and it has the line $\mathrm{Re}(z) = \delta_2(A)$ as a vertical asymptote. The spectrum of A is contained in the unbounded region defined by $\Gamma_{in}(A) = \{z = x + \mathrm{i}\, y : \ x, y \in \mathbb{R} \ \text{such that} \ f_A(x, y) \ge 0\}$ [1].

Crucial part of the study of the shell $\Gamma(A)$ plays the sign of the quantity (discriminant)

$$D(A) = (\delta_1(A) - \delta_2(A))^2 - 4\left(\nu(A) - u(A)^2\right).$$

Namely, we have the following three cases:

(a) If $D(A) > 0$, then $\Gamma(A)$ is the union of two curves, a closed bounded curve that lies in the vertical zone $\mathcal{Z}_1 = \left\{z \in \mathbb{C} : \ \frac{\delta_1(A) + \delta_2(A) + \sqrt{D(A)}}{2} \le \mathrm{Re}(z) \le \delta_1(A)\right\}$ and surrounds exactly one simple eigenvalue of A [13], and an open unbounded curve which lies in the vertical zone $\mathcal{Z}_2 = \left\{z \in \mathbb{C} : \ \delta_2(A) \le \mathrm{Re}(z) \le \frac{\delta_1(A) + \delta_2(A) - \sqrt{D(A)}}{2}\right\}$ and has the remaining eigenvalues of A to its left. The closed branch of $\Gamma(A)$ intersects the horizontal line $\mathcal{L}$ at the points $\frac{\delta_1(A) + \delta_2(A) + \sqrt{D(A)}}{2} + \mathrm{i}\, u(A)$ and $\delta_1(A) + \mathrm{i}\, u(A)$, and the open branch of $\Gamma(A)$ intersects $\mathcal{L}$ at $\frac{\delta_1(A) + \delta_2(A) - \sqrt{D(A)}}{2} + \mathrm{i}\, u(A)$. It is worth mentioning that if $D(A) > 0$, then the common point $\delta_1(A) + \mathrm{i}\, u(A)$ of the shell and the boundary of the numerical range $F(A)$ cannot belong to a flat portion of the boundary of $F(A)$ because otherwise, $\delta_1(A) = \delta_2(A)$ [3], which is a contradiction.

(b) If $D(A) = 0$, then $\Gamma(A)$ intersects $\mathcal{L}$ at two points, $\frac{\delta_1(A) + \delta_2(A)}{2} + \mathrm{i}\, u(A)$ and $\delta_1(A) + \mathrm{i}\, u(A)$; in particular, the first point is a node point of $\Gamma(A)$.

(c) If $D(A) < 0$, then $\Gamma(A)$ is an unbounded open and simple curve, which has all the eigenvalues of A to its left and intersects $\mathcal{L}$ at the point $\delta_1(A) + \mathrm{i}\, u(A)$.

Figure 1 is illustrative of the aforementioned cases.

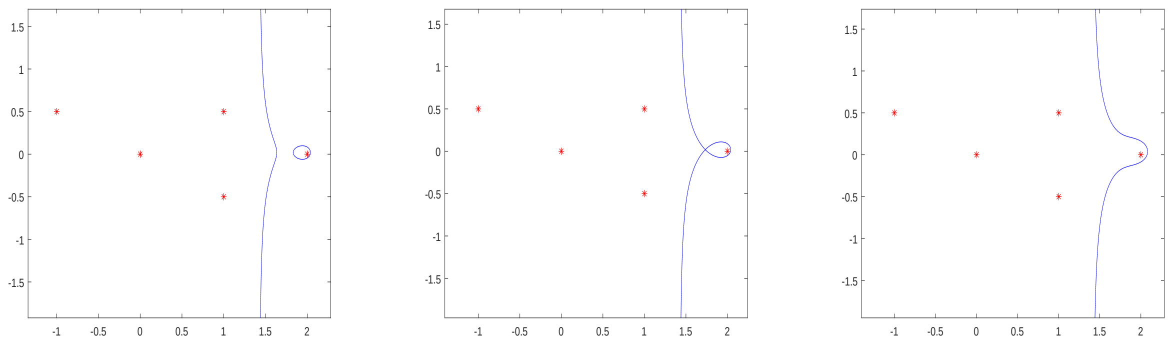

Figure 1: On the left, $D(A) > 0$, and $\Gamma(A)$ consists of a closed branch which surrounds one simple eigenvalue and an unbounded curve leaving all the rest eigenvalues to its left. In the middle, $D(A) = 0$, and $\Gamma(A)$ creates a node. Finally, on the right, $D(A) < 0$, and the shell is a simple open unbounded curve and all eigenvalues lie to its left. All eigenvalues are marked with asterisks.

Remark 1.1. If $\delta_1(A) > \delta_2(A)$ and $\nu(A) - u(A)^2 = 0$, then by the proof of Theorem 3.1 in [1] (see also Proposition 3.15 below), $\delta_1(A) + \mathrm{i}\, u(A)$ is a normal (simple) eigenvalue of A. In this case, the defining function of $\Gamma(A)$ is written in the form $f_A(x, y) = \left[(\delta_1(A) - x)^2 + (u(A) - y)^2\right](\delta_2(A) - x)$, and apparently, the shell $\Gamma(A)$ coincides with the union of the vertical line $\mathrm{Re}(z) = \delta_2(A)$ and the point $\delta_1(A) + \mathrm{i}\, u(A)$ (i.e., the closed branch of $\Gamma(A)$ degenerates to the singleton $\{\delta_1(A) + \mathrm{i}\, u(A)\}$).

Remark 1.2. If $\delta_1(A) = \delta_2(A)$ (i.e., the greatest eigenvalue $\delta_1(A)$ of $H(A)$ is multiple), then the defining function of $\Gamma(A)$ becomes $f_A(x, y) = \left[(\delta_1(A) - x)^2 + (u(A) - y)^2 + \nu(A) - u(A)^2\right](\delta_1(A) - x)$, and the shell $\Gamma(A)$ degenerates to the vertical line $\mathrm{Re}(z) = \delta_1(A)$ $(= \delta_2(A))$.

For reader's convenience, we list below some known basic properties of the shell $\Gamma(A)$ [13].

(P_1) $\Gamma(A^T) = \Gamma(A)$ and $\Gamma(A^*) = \Gamma(\overline{A}) = \overline{\Gamma(A)}$; in particular, if A is real, then $u(A) = 0$ and $\Gamma(A)$ is symmetric with respect to the real axis.

(P_2) For any unitary matrix $U \in \mathbb{C}^{n\times n}$, $\Gamma(U^*AU) = \Gamma(A)$.

(P_3) For any $b \in \mathbb{C}$, $\Gamma(A + bI_n) = \Gamma(A) + b$, where I_n denotes the $n \times n$ identity matrix and adding a scalar to a set means adding this scalar to all elements of the set.

(P_4) For any $r > 0$, $\Gamma(rA) = r\,\Gamma(A)$.

As originally observed in [13] (see also [2, 14]), drawing the shells of several rotations $e^{\mathrm{i}\theta}A$ ($\theta \in [0, 2\pi]$) of a square matrix A yields a (not necessarily connected) region that contains the spectrum of A and is contained in $F(A)$; i.e., the shells of the rotated matrices provide a better (tighter) spectral inclusion than the standard numerical range. In this work, we consider the case where $D(A) > 0$, so that the shell $\Gamma(A)$ has a closed branch which surrounds a simple eigenvalue of A; this closed branch offers a nice estimation of the eigenvalue it isolates. In Section 2, we discuss a defining formula and several geometrical properties of the closed branch. In Section 3, we consider the set of extremal eigenvalues of a matrix A that are surrounded by closed branches of shells of some matrices $e^{\mathrm{i}\theta}A$, $\theta \in [0, 2\pi]$, and we treat the size (in various meanings) of these closed branches as a measure of non-normality of the simple extremal eigenvalues that they isolate. In particular, we introduce four such measures and establish some relations among them. In Section 4, we give three examples to illustrate our results. Finally, a technique for the computation of the area enclosed by the closed branch (when it exists) of the shell $\Gamma(A)$ is presented in the appendix at the end of the paper. All calculations were performed in MATLAB 9.10.

Throughout the text, we denote by $\mathcal{C}o\{S\}$ the convex hull of the set S, by ∂S the boundary of the set S, by $\mathcal{U}(n)$ the set of $n \times n$ unitary matrices, and by M^+ the Moore-Penrose pseudoinverse of a matrix $M \in \mathbb{C}^{m\times n}$.

2 The closed branch of the shell

Consider a matrix $A \in \mathbb{C}^{n\times n}$ such that $D(A) = (\delta_1(A) - \delta_2(A))^2 - 4\left(\nu(A) - u(A)^2\right) > 0$, and the horizontal diameter of the closed branch of the shell $\Gamma(A)$

$$d_h(A) = \frac{\delta_1(A) - \delta_2(A) - \sqrt{D(A)}}{2} = \frac{\delta_1(A) - \delta_2(A) - \sqrt{(\delta_1(A) - \delta_2(A))^2 - 4\left(\nu(A) - u(A)^2\right)}}{2}. \tag{1}$$

In this section, for the sake of brevity, we denote

$$\delta_i = \delta_i(A) \quad (i = 1, 2, \ldots, n), \quad D = D(A), \quad \text{and} \quad d_h = d_h(A).$$

Proposition 2.1. *The closed branch of the shell $\Gamma(A)$ is given by $\{x(t) + \mathrm{i}\, y(t) : \ t \in [0, 1]\}$, with*

$$x(t) = \frac{(1+t)\delta_1 + (1-t)\delta_2 + (1-t)\sqrt{D}}{2} = \frac{\delta_1 + \delta_2 + \sqrt{D}}{2} + t\, d_h$$

and

$$y(t) = u(A) \pm d_h \sqrt{t(1-t)\left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}}\right]}.$$

Proof. For any $x \in Z_1 \cap \mathbb{R}$, we need to find a $y \in \mathbb{R}$ such that $x + \mathrm{i}\, y \in \Gamma(A)$. Let $t \in [0, 1]$, and define

$$x(t) = (1-t)\, \frac{\delta_1 + \delta_2 + \sqrt{D}}{2} + t\, \delta_1 = \frac{(1+t)\delta_1 + (1-t)\delta_2 + (1-t)\sqrt{D}}{2}.$$

Substituting $x = x(t)$ in the defining equation of the shell, we have

$$(\delta_2 - x(t))\left[(\delta_1 - x(t))^2 + (u(A) - y)^2\right] + (\delta_1 - x(t))\left(\nu(A) - u(A)^2\right) = 0,$$

or

$$(\delta_2 - x(t))\,(u(A) - y)^2 = (x(t) - \delta_1)\left[(\delta_2 - x(t))\,(\delta_1 - x(t)) + \nu(A) - u(A)^2\right],$$

or

$$(\delta_2 - x(t))\,(u(A) - y)^2 = (x(t) - \delta_1)\left[x(t)^2 - (\delta_1 + \delta_2)\, x(t) + \delta_1\delta_2 + \nu(A) - u(A)^2\right],$$

or

$$(\delta_2 - x(t))\,(u(A) - y)^2 = (x(t) - \delta_1)\left(x(t) - \frac{\delta_1 + \delta_2 + \sqrt{D}}{2}\right)\left(x(t) - \frac{\delta_1 + \delta_2 - \sqrt{D}}{2}\right). \tag{2}$$

Using the definition of $d_h = d_h(A)$ in (1), we simplify each factor of (2) separately:

- $x(t) - \delta_2 = \dfrac{(1+t)\delta_1 + (1-t)\delta_2 + (1-t)\sqrt{D}}{2} - \delta_2 = \dfrac{(1+t)\delta_1 - (1+t)\delta_2 + (1-t)\sqrt{D}}{2} = (1+t)d_h + \sqrt{D}$,
- $x(t) - \delta_1 = \dfrac{(1+t)\delta_1 + (1-t)\delta_2 + (1-t)\sqrt{D}}{2} - \delta_1 = (t-1)\dfrac{\delta_1 - \delta_2 - \sqrt{D}}{2} = (t-1)d_h$,
- $x(t) - \dfrac{\delta_1 + \delta_2 + \sqrt{D}}{2} = \dfrac{(1+t)\delta_1 + (1-t)\delta_2 + (1-t)\sqrt{D}}{2} - \dfrac{\delta_1 + \delta_2 + \sqrt{D}}{2} = t\, d_h$,
- $x(t) - \dfrac{\delta_1 + \delta_2 - \sqrt{D}}{2} = \dfrac{(1+t)\delta_1 + (1-t)\delta_2 + (1-t)\sqrt{D}}{2} - \dfrac{\delta_1 + \delta_2 - \sqrt{D}}{2} = t\, d_h + \sqrt{D}$.

Substituting these expressions in equality (2) yields

$$-\left((1+t)d_h + \sqrt{D}\right)(y - u(A))^2 = t(t-1)\left(t\, d_h + \sqrt{D}\right) d_h^2,$$

or

$$(y - u(A))^2 = t(1-t)\frac{t\, d_h + \sqrt{D}}{(1+t)d_h + \sqrt{D}}\, d_h^2,$$

or

$$(y - u(A))^2 = t(1-t)\left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}}\right] d_h^2,$$

or

$$y = u(A) \pm d_h \sqrt{t(1-t)\left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}}\right]},$$

completing the proof. □

Corollary 2.2. *The closed branch of the shell $\Gamma(A)$ is a convex curve, i.e., it is the boundary of a convex region. Moreover, it cannot be reduced to a line segment unless it is the singleton $\{\delta_1 + \mathrm{i}\, u(A)\}$.*

Proof. The closed branch of the shell is the union of two curves which are symmetric with respect to the horizontal line $\mathcal{L} = \{s + \mathrm{i}\, u(A) :\ s \in \mathbb{R}\}$. The equations of these two curves are

$$y(t) = u(A) \pm d_h \sqrt{t(1-t)\left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}}\right]}, \quad t \in [0,1],$$

and the convexity of the closed branch of the shell follows from the fact that

$$g(t) = d_h \sqrt{t(1-t)\left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}}\right]}$$

is a concave function and $-g(t)$ is a convex function.

For the second part, we assume that the closed branch is a line segment. Then, due to its symmetry with respect to the horizontal line $\mathcal{L} = \{s + \mathrm{i}\, u(A) :\ s \in \mathbb{R}\}$, this can only happen when $d_h = 0$, or $y(t) = u(A)$ for all $t \in [0,1]$. If $d_h \neq 0$, then

$$\frac{|y(t) - u(A)|}{d_h} = \sqrt{t(1-t)\left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}}\right]} = 0 \quad \text{for all } t \in [0,1]. \qquad (3)$$

It is straightforward to verify that (3) holds if and only if $t\, d_h + \sqrt{D} = 0$ for all $t \in (0,1)$. This is a contradiction, and hence, $d_h = 0$ and the closed branch is reduced to the singleton $\{\delta_1 + \mathrm{i}\, u(A)\}$. □

Corollary 2.3. *Let z_1 and z_2 be any two points on the closed branch of the shell $\Gamma(A)$. Then $|z_1 - z_2| \leq d_h$.*

Proof. Consider the midpoint p of the (horizontal) axis of symmetry of the closed branch of the shell, that is, $p = \dfrac{3\delta_1 + \delta_2 + \sqrt{D}}{4} + \mathrm{i}\, u(A)$. Let also $z = x(t) + \mathrm{i}\, y(t)$ ($t \in [0,1]$) be any point on the closed branch of the shell. Then, we have

$$\begin{aligned}
|p - z|^2 &= \left(\frac{3\delta_1 + \delta_2 + \sqrt{D}}{4} - \frac{\delta_1 + \delta_2 + \sqrt{D}}{2} - t\, d_h \right)^2 + (y(t) - u(A))^2 \\
&= \left(\frac{\delta_1 - \delta_2 - \sqrt{D}}{4} - t\, d_h \right)^2 + d_h^2 t(1-t) \left(1 - \frac{d_h}{(1+t)d_h + \sqrt{D}} \right) \\
&\leq d_h^2 \left(\frac{1}{2} - t \right)^2 + d_h^2 t(1-t) \\
&= \frac{d_h^2}{4}.
\end{aligned}$$

The last inequality implies that $|p - z| \leq \dfrac{d_h}{2}$. A simple application of the triangle inequality reveals that for any two boundary points z_1 and z_2 of the closed branch of the shell $\Gamma(A)$, it holds that $|z_1 - z_2| \leq |z_1 - p| + |z_2 - p| \leq d_h$. □

Corollary 2.4. *The area enclosed by the closed branch of the shell $\Gamma(A)$ is smaller than or equal to the quantity*

$$d_h^2 \sqrt{\frac{2}{3} \left(1 - \ln \left(1 + \frac{d_h}{d_h + \sqrt{D}} \right) \right)}$$

Proof. In view of Proposition 2.1, and with the use of the Cauchy-Schwarz inequality, we can verify that the area enclosed by the closed branch of $\Gamma(A)$ is

$$\begin{aligned}
2 \int_{\frac{\delta_1 + \delta_2 + \sqrt{D}}{2}}^{\delta_1} |y(x(t)) - u(A)|\, dx &= 2\, d_h \int_0^1 |y(t) - u(A)|\, dt \\
&= 2\, d_h \int_0^1 d_h \sqrt{t(1-t) \left[1 - \frac{d_h}{(1+t)d_h + \sqrt{D}} \right]}\, dt \qquad (4) \\
&\leq 2\, d_h^2 \sqrt{\int_0^1 t(1-t) dt} \sqrt{\int_0^1 \left[1 - \frac{d_h}{(t+1)d_h + \sqrt{D}} \right] dt} \\
&= 2\, d_h^2 \sqrt{\frac{1}{6}} \sqrt{1 - \left(\ln \left(2d_h + \sqrt{D} \right) - \ln \left(d_h + \sqrt{D} \right) \right)} \\
&= d_h^2 \sqrt{\frac{2}{3} \left(1 - \ln \left(1 + \frac{d_h}{d_h + \sqrt{D}} \right) \right)}. \qquad \square
\end{aligned}$$

Numerical experiments (see Example 4.3 below) show that the closed branch of the shell of matrix A does not lie necessarily (entirely) in the numerical range $F(A)$. A way to further explore this is to compare the radii of curvatures of $F(A)$ and the closed branch of $\Gamma(A)$ at their common right-most point $\delta_1 + \mathrm{i}\, u(A)$.

Proposition 2.5. *For any (general) matrix $A \in \mathbb{C}^{n\times n}$, the radius of curvature of the shell $\Gamma(A)$ at its right-most point $\delta_1 + \mathrm{i}\, u(A)$ is*

$$R_{\Gamma(A)}\left(\delta_1 + \mathrm{i}\, u(A)\right) = \frac{\nu(A) - u(A)^2}{2\left(\delta_1 - \delta_2\right)}. \tag{5}$$

Moreover, $D = (\delta_1 - \delta_2)^2 - 4\left(\nu(A) - u(A)^2\right) \geq 0$ if and only if $R_{\Gamma(A)}\left(\delta_1 + \mathrm{i}\, u(A)\right) \leq \dfrac{\delta_1 - \delta_2}{8}$, where the latter upper bound of $R_{\Gamma(A)}\left(\delta_1 + \mathrm{i}\, u(A)\right)$ is attained if and only if $D = 0$.

Proof. The curvature of the shell $\Gamma(A)$ (which is defined implicitly by $f = f_A(x, y) = 0$) at a point $x + \mathrm{i}\, y$ $(x, y \in \mathbb{R})$ is given by [6]

$$\frac{1}{R_{\Gamma(A)}(x + \mathrm{i}\, y)} = \frac{\left|\left(\dfrac{\partial f}{\partial y}\right)^2 \dfrac{\partial^2 f}{\partial x^2} - 2\,\dfrac{\partial f}{\partial x}\,\dfrac{\partial f}{\partial y}\,\dfrac{\partial^2 f}{\partial x \partial y} + \left(\dfrac{\partial f}{\partial x}\right)^2 \dfrac{\partial^2 f}{\partial y^2}\right|}{\sqrt{\left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2}^{\,3}}.$$

It is a matter of straightforward computations to verify the following:

- $\dfrac{\partial f}{\partial x} = -3x^2 + 2x\left(2\delta_1 + \delta_2\right) - (u(A) - y)^2 - \left(\delta_1^2 + 2\delta_1\delta_2 + \nu(A) - u(A)^2\right)$,
- $\dfrac{\partial^2 f}{\partial x^2} = -6x + 2\left(2\delta_1 + \delta_2\right)$,
- $\dfrac{\partial f}{\partial y} = 2(y - u(A))\left(\delta_2 - x\right)$,
- $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x} = -2(y - u(A))$,
- $\dfrac{\partial^2 f}{\partial y^2} = -2\left(x - \delta_2\right)$.

As a consequence, the curvature of the shell at the point $\delta_1 + \mathrm{i}\, u(A)$ is

$$\frac{1}{R_{\Gamma(A)}\left(\delta_1 + \mathrm{i}\, u(A)\right)} = \frac{2\left(\delta_1 - \delta_2\right)}{\nu(A) - u(A)^2},$$

and the first part of the proposition is obtained.

For the second part of the proposition, it is clear that $D \geq 0$ if and only if

$$\left(\delta_1 - \delta_2\right)^2 - 4\left(\nu(A) - u(A)^2\right) \geq 0,$$

or equivalently, if and only if

$$\frac{\nu(A) - u(A)^2}{\delta_1 - \delta_2} \leq \frac{\delta_1 - \delta_2}{4}. \tag{6}$$

Recalling (5), the inequality (6) holds if and only if

$$R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A)) \leq \frac{\delta_1 - \delta_2}{8}. \qquad \square$$

At this point, we need to recall Theorem 1.1 of [5] (see also Theorem 3.3 of [4]), which states that if $A = H_1 + \mathrm{i}\, H_2$, where H_1 and H_2 are hermitian matrices (i.e., $H_1 = H(A)$ and $H_2 = (-\mathrm{i})K(A)$), then the radius of curvature at a boundary point $\mu = \mathbf{x}^* A\mathbf{x}$ ($\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x}^*\mathbf{x} = 1$) of the numerical $F(A)$ which has a supporting line with equation $ax + by + c = 0$, is given by

$$R_{F(A)}(\mu) = -\frac{2}{\sqrt{a^2+b^2}}\, \mathbf{x}^* Q \left(\tilde{A} + cI_n\right)^+ Q\mathbf{x}, \tag{7}$$

where $\tilde{A} = aH_1 + bH_2$ and $Q = bH_1 - aH_2$.

Proposition 2.6. *Let $A \in \mathbb{C}^{n\times n}$ with $D > 0$. Then the following hold:*

(i) $\dfrac{4(\delta_1-\delta_2)}{\delta_1-\delta_n} \le \dfrac{R_F(\delta_1 + \mathrm{i}\, u(A))}{R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))} \le 4.$

(ii) $2R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A)) \le d_h \le 4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A)).$

Proof. (i) Let $A = H(A)+K(A) = H(A)+\mathrm{i}\,(-\mathrm{i})K(A)$, in terms of the hermitian and the skew-hermitian parts of A, and consider a unitary matrix $U \in \mathcal{U}(n)$ such that $U^*H(A)U = \operatorname{diag}\{\delta_1, \delta_2, \ldots, \delta_n\}$. Then, it follows

$$U^*AU = \operatorname{diag}\{\delta_1, \delta_2, \ldots, \delta_n\} + \left[\begin{array}{c|c} \mathrm{i}\, u(A) & \mathbf{z}^* \\ \hline -\mathbf{z} & K_1 \end{array}\right],$$

where $\mathbf{z} \in \mathbb{C}^{n-1}$ and K_1 is an $(n-1)\times(n-1)$ skew-hermitian matrix. Using this decomposition, we will find upper and lower bounds for $R_{F(A)}(\delta_1 + \mathrm{i}\, u(A))$. The supporting line of the numerical range at the point $\delta_1 + \mathrm{i}\, u(A)$ is $\mathrm{Re}(z) = \delta_1$. A unit eigenvector $\mathbf{y}_1$ of $H(A)$ corresponding to δ_1, which can be considered as the first column of U, is such that $\mathbf{y}_1^* A\mathbf{y}_1 = \delta_1 + \mathrm{i}\, u(A)$. Also, note that $\mathbf{z}^*\mathbf{z} = \nu(A) - u(A)^2$ (see the proof of Theorem 3.1 in [1]).

By (7), we have

$$\begin{aligned}
R_{F(A)}(\delta_1 + \mathrm{i}\, u(A)) &= -2\,\mathbf{y}_1^*(\mathrm{i}\, K(A))\,(H(A) - \delta_1 I_n)^+\,(\mathrm{i}\, K(A))\mathbf{y}_1 \\
&= 2\,\mathbf{y}_1^* K(A)\,(H(A) - \delta_1 I_n)^+\, K(A)\mathbf{y}_1 \\
&= 2\,\mathbf{y}_1^* K(A)\,(U\operatorname{diag}\{\delta_1, \delta_2, \ldots, \delta_n\}\,U^* - \delta_1 UU^*)^+\, K(A)\mathbf{y}_1 \\
&= 2\,\mathbf{y}_1^* UU^* K(A)U\,(\operatorname{diag}\{\delta_1, \delta_2, \ldots, \delta_n\} - \delta_1 I_n)^+\, U^*K(A)UU^*\mathbf{y}_1 \\
&= 2\,(U^*\mathbf{y}_1)^* \left[\begin{array}{c|c} \mathrm{i}\, u(A) & \mathbf{z}^* \\ \hline -\mathbf{z} & K_1 \end{array}\right] (\operatorname{diag}\{\delta_1, \delta_2, \ldots, \delta_n\} - \delta_1 I_n)^+ \left[\begin{array}{c|c} \mathrm{i}\, u(A) & \mathbf{z}^* \\ \hline -\mathbf{z} & K_1 \end{array}\right] (U^*\mathbf{y}_1) \\
&= 2\,\mathbf{e}_1^T \left[\begin{array}{c|c} \mathrm{i}\, u(A) & \mathbf{z}^* \\ \hline -\mathbf{z} & K_1 \end{array}\right] \operatorname{diag}\left\{0, (\delta_2-\delta_1)^{-1}, \ldots, (\delta_n-\delta_1)^{-1}\right\} \left[\begin{array}{c|c} \mathrm{i}\, u(A) & \mathbf{z}^* \\ \hline -\mathbf{z} & K_1 \end{array}\right] \mathbf{e}_1 \\
&= -2\,\mathbf{z}^* \operatorname{diag}\left\{(\delta_2-\delta_1)^{-1}, (\delta_3-\delta_1)^{-1}, \ldots, (\delta_n-\delta_1)^{-1}\right\} \mathbf{z},
\end{aligned}$$

where $\mathbf{e}_1$ denotes the first vector of the standard basis. Since

$$\frac{2\,\mathbf{z}^*\mathbf{z}}{\delta_1-\delta_n} \le -2\,\mathbf{z}^* \operatorname{diag}\left\{(\delta_2-\delta_1)^{-1}, (\delta_3-\delta_1)^{-1}, \ldots, (\delta_n-\delta_1)^{-1}\right\} \mathbf{z} \le \frac{2\,\mathbf{z}^*\mathbf{z}}{\delta_1-\delta_2}$$

and $\mathbf{z}^*\mathbf{z} = \nu(A) - u(A)^2$, and keeping in mind (5), we obtain

$$\frac{2\,\dfrac{\nu(A)-u(A)^2}{\delta_1-\delta_n}}{\dfrac{\nu(A)-u(A)^2}{2(\delta_1-\delta_2)}} \le \frac{R_F(\delta_1 + \mathrm{i}\, u(A))}{R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))} \le \frac{2\,\dfrac{\nu(A)-u(A)^2}{\delta_1-\delta_2}}{\dfrac{\nu(A)-u(A)^2}{2(\delta_1-\delta_2)}},$$

and the result follows.

(ii) By the definition of $d_h = d_h(A)$ in (1), it is apparent that

$$(\delta_1 - \delta_2)^2 - 4(\nu(A) - u(A)^2) = (2d_h - (\delta_1 - \delta_2))^2. \tag{8}$$

Recalling (5), equality (8) is equivalent to

$$\begin{aligned} 4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A)) &= \frac{(\delta_1 - \delta_2)^2}{2(\delta_1 - \delta_2)} - \frac{(2d_h - (\delta_1 - \delta_2))^2}{2(\delta_1 - \delta_2)} \\ &= \frac{d_h}{\delta_1 - \delta_2}\left(2(\delta_1 - \delta_2) - 2d_h\right) \\ &= \frac{d_h}{\delta_1 - \delta_2}\left(\delta_1 - \delta_2 + \sqrt{D}\right). \end{aligned}$$

The last equality yields

$$\frac{d_h}{\delta_1 - \delta_2} = \frac{4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))}{\delta_1 - \delta_2 + \sqrt{D}} \le \frac{4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))}{\delta_1 - \delta_2},$$

and hence, $d_h \le 4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))$. Moreover, since $\sqrt{D} \le \delta_1 - \delta_2$, it follows that

$$\frac{d_h}{\delta_1 - \delta_2} = \frac{4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))}{\delta_1 - \delta_2 + \sqrt{D}} \ge \frac{4R_{\Gamma(A)}(\delta_1 + \mathrm{i}\, u(A))}{2(\delta_1 - \delta_2)},$$

which completes the proof. □

3 Shell-extremal eigenvalues and normality

Definition 3.1. Let $A \in \mathbb{C}^{n\times n}$, and let $\lambda_0 \in \sigma(A) \cap \partial\mathcal{C}o\{\sigma(A)\}$ (i.e., λ_0 is an extremal eigenvalue of A). If for some $\theta_0 \in [0, 2\pi]$, it holds that $D\left(e^{\mathrm{i}\theta_0}A\right) > 0$ and $e^{\mathrm{i}\theta_0}\lambda_0$ is surrounded by the closed branch of $\Gamma\left(e^{\mathrm{i}\theta_0}A\right)$, then λ_0 is called a *shell-extremal eigenvalue of A*.

Another way to interpret Definition 3.1 is that λ_0 is the (unique) simple eigenvalue of A surrounded by the closed branch of the rotated curve $e^{-\mathrm{i}\theta_0}\Gamma\left(e^{\mathrm{i}\theta_0}A\right)$ for some $\theta_0 \in [0, 2\pi]$. In this case, it also follows that $e^{\mathrm{i}\theta_0}\lambda_0$ is such that $\mathrm{Re}\left(e^{\mathrm{i}\theta_0}\lambda_0\right) = \max\left\{\mathrm{Re}(\lambda(\theta_0)) : \lambda(\theta_0) \in \sigma\left(e^{\mathrm{i}\theta_0}A\right)\right\}$, and λ_0 is a vertex of the polygon $\partial\mathcal{C}o\{\sigma(A)\}$.

For any eigenvalue λ_0 of a (general) matrix $A \in \mathbb{C}^{n\times n}$, we define the set

$$\mathcal{A}(\lambda_0) = \left\{\theta \in [0, 2\pi] : D\left(e^{\mathrm{i}\theta}A\right) > 0 \text{ and } e^{\mathrm{i}\theta}\lambda_0 \text{ is surrounded by the closed branch of } \Gamma\left(e^{\mathrm{i}\theta}A\right)\right\}.$$

It is clear that if λ_0 is a shell-extremal eigenvalue, then $\mathcal{A}(\lambda_0) \neq \emptyset$ and for every $\theta \in \mathcal{A}(\lambda_0)$, $e^{\mathrm{i}\theta}\lambda_0$ is the (unique) simple eigenvalue of matrix $e^{\mathrm{i}\theta}A$ that lies inside the closed branch of the shell $\Gamma\left(e^{\mathrm{i}\theta_0}A\right)$. If λ_0 is not a shell-extremal eigenvalue, then $\mathcal{A}(\lambda_0) = \emptyset$.

Remark 3.2. Let $A \in \mathbb{C}^{n\times n}$ with $D(A) > 0$, and let λ_0 be the simple eigenvalue of A surrounded by the closed branch of the shell $\Gamma(A)$. Then, by the continuity of $\nu(e^{\mathrm{i}\theta}A)$ and $u(e^{\mathrm{i}\theta}A)$ in $\theta \in [0, 2\pi]$, there exist a $\theta_1 \in [0, \pi)$ and a $\theta_2 \in (\pi, 2\pi]$ such that for every $\theta \in [0, \theta_1) \cup (\theta_2, 2\pi]$, $D\left(e^{\mathrm{i}\theta}A\right) > 0$ and $e^{\mathrm{i}\theta}\lambda_0$ is the simple eigenvalue of $e^{\mathrm{i}\theta}A$ that lies inside the closed branch of $\Gamma\left(e^{\mathrm{i}\theta}A\right)$, i.e., $[0, \theta_1) \cup (\theta_2, 2\pi] \subseteq \mathcal{A}(\lambda_0)$.

Remark 3.3. Let $A \in \mathbb{C}^{n\times n}$, and let λ_0 be an eigenvalue of A. If λ_0 is not a vertex of the polygon $\partial\mathcal{C}o\{\sigma(A)\}$ or λ_0 is a vertex of $\partial\mathcal{C}o\{\sigma(A)\}$ and its algebraic multiplicity is greater than 1 (i.e., it is a multiple extremal eigenvalue of A), then it is apparent that $\mathcal{A}(\lambda_0)=\emptyset$ and λ_0 cannot be a shell-extremal eigenvalue of A. If λ_0 is a simple eigenvalue of A which is a vertex of the polygon $\partial\mathcal{C}o\{\sigma(A)\}$, then there exist some $\phi\in[0,2\pi]$ such that $e^{\mathrm{i}\phi}\lambda_0$ has real part less than or equal to the real part of some other eigenvalue of $e^{\mathrm{i}\phi}A$, and consequently, $\phi\notin\mathcal{A}(\lambda_0)$.

Remark 3.4. Let $A \in \mathbb{C}^{n\times n}$, and let λ_1 and λ_2 be two distinct (simple) shell-extremal eigenvalues of A. Then $\mathcal{A}(\lambda_1)\cap\mathcal{A}(\lambda_2)=\emptyset$. If not, and $\theta_0\in\mathcal{A}(\lambda_1)\cap\mathcal{A}(\lambda_2)$, then there are two eigenvalues of $e^{\mathrm{i}\theta_0}A$, namely, $e^{\mathrm{i}\theta_0}\lambda_1$ and $e^{\mathrm{i}\theta_0}\lambda_2$, in the vertical complex zone defined by $\dfrac{\delta_1(e^{\mathrm{i}\theta_0}A)+\delta_2(e^{\mathrm{i}\theta_0}A)+\sqrt{D(e^{\mathrm{i}\theta_0}A)}}{2}$ and $\delta_1(e^{\mathrm{i}\theta_0}A)$, which contradicts [1, 13].

Proposition 3.5. *Let A be an $n\times n$ real matrix. Then the following hold:*

(i) The function $D(e^{\mathrm{i}\theta}A)$, $\theta\in[0,2\pi]$, is "symmetric" with respect to 0, in the sense that $D\left(e^{\mathrm{i}\theta}A\right)=D\left(e^{\mathrm{i}(2\pi-\theta)}A\right)$ for every $\theta\in[0,2\pi]$.

(ii) The function $D(e^{\mathrm{i}\theta}A)$, $\theta\in[0,2\pi]$, is symmetric with respect to π.

Proof. (i) We want to prove that for every $\theta\in[0,2\pi]$, the quantities (discriminants)

$$D\left(e^{\mathrm{i}\theta}A\right)=\left(\delta_1(e^{\mathrm{i}\theta}A)-\delta_2(e^{\mathrm{i}\theta}A)\right)^2-4\left(\nu(e^{\mathrm{i}\theta}A)-u(e^{\mathrm{i}\theta}A)^2\right)$$

and

$$D\left(e^{\mathrm{i}(2\pi-\theta)}A\right)=\left(\delta_1(e^{\mathrm{i}(2\pi-\theta)}A)-\delta_2(e^{\mathrm{i}(2\pi-\theta)}A)\right)^2-4\left(\nu(e^{\mathrm{i}(2\pi-\theta)}A)-u(e^{\mathrm{i}(2\pi-\theta)}A)^2\right)$$

are equal. By the symmetry of the numerical range

$$\begin{aligned} F(A) &= \{\mathbf{x}^*A\mathbf{x}\in\mathbb{C}:\ \mathbf{x}\in\mathbb{C}^n \text{ with } \mathbf{x}^*\mathbf{x}=1\} \\ &= \bigcap_{\theta\in[0,2\pi]}\left\{e^{-\mathrm{i}\theta}(x+\mathrm{i}\,y):\ x,y\in\mathbb{R} \text{ with } x\le\delta_1(e^{\mathrm{i}\theta}A)\right\} \end{aligned}$$

with respect to the real axis [7], it follows that for every $\theta\in[0,2\pi]$, the right-most point of $F(e^{\mathrm{i}\theta}A)$ and the right-most point of $F(e^{\mathrm{i}(2\pi-\theta)}A)$ have the same real part, which means that $\delta_1(e^{\mathrm{i}\theta}A)=\delta_1(e^{\mathrm{i}(2\pi-\theta)}A)$. Similarly, by the symmetry of the *rank-2 numerical range*

$$\begin{aligned} \Lambda_2(A) &= \left\{z\in\mathbb{C}:\ X^*AX=zI_2 \text{ for some } X\in\mathbb{C}^{n\times 2} \text{ such that } X^*X=I_2\right\} \\ &= \bigcap_{\theta\in[0,2\pi]}\left\{e^{-\mathrm{i}\theta}(x+\mathrm{i}\,y):\ x,y\in\mathbb{R} \text{ with } x\le\delta_2(e^{\mathrm{i}\theta}A)\right\} \end{aligned}$$

with respect to the real axis [10], it follows that for every $\theta\in[0,2\pi]$, the right-most point of $\Lambda_2(e^{\mathrm{i}\theta}A)$ and the right-most point of $\Lambda_2(e^{\mathrm{i}(2\pi-\theta)}A)$ have the same real part, which means that $\delta_2(e^{\mathrm{i}\theta}A)=\delta_2(e^{\mathrm{i}(2\pi-\theta)}A)$. Moreover, since A is real, one can verify that for any $\theta\in[0,2\pi]$, $\mathbf{y}_\theta$ is a unit eigenvector of $H\left(e^{\mathrm{i}\theta}A\right)$ corresponding to $\delta_1\left(e^{\mathrm{i}\theta}A\right)$ if and only if (its conjugate) $\overline{\mathbf{y}}_\theta$ is a unit eigenvector of $H\left(e^{\mathrm{i}(2\pi-\theta)}A\right)$ corresponding to $\delta_1\left(e^{\mathrm{i}(2\pi-\theta)}A\right)$. As a consequence,

$$\begin{aligned} \nu(e^{\mathrm{i}\theta}A)-u(e^{\mathrm{i}\theta}A)^2 &= \left\|K(e^{\mathrm{i}\theta}A)\mathbf{y}_\theta\right\|_2^2-\mathrm{Im}\left(\mathbf{y}_\theta^*(e^{\mathrm{i}\theta}A)\mathbf{y}_\theta\right)^2 \\ &= \left\|\overline{K(e^{\mathrm{i}\theta}A)\mathbf{y}_\theta}\right\|_2^2-\mathrm{Im}\left(\overline{\mathbf{y}_\theta^*(e^{\mathrm{i}\theta}A)\mathbf{y}_\theta}\right)^2 \\ &= \left\|K(e^{\mathrm{i}(2\pi-\theta)}A)\overline{\mathbf{y}}_\theta\right\|_2^2-\mathrm{Im}\left(\overline{\mathbf{y}}_\theta^*(e^{\mathrm{i}(2\pi-\theta)}A)\overline{\mathbf{y}}_\theta\right)^2 \\ &= \nu(e^{\mathrm{i}(2\pi-\theta)}A)-u(e^{\mathrm{i}(2\pi-\theta)}A)^2, \end{aligned}$$

and the "symmetry" of the function $D(e^{\mathrm{i}\theta}A)$ ($\theta \in [0, 2\pi]$) with respect to 0, is obtained.

(ii) Applying part (i) to the real matrix $-A = e^{\mathrm{i}\pi}A$, the symmetry of $D(e^{\mathrm{i}\theta}A)$ ($\theta \in [0, 2\pi]$) with respect to π follows readily. □

Proposition 3.6. *Let A be an $n \times n$ real matrix with $D(A) > 0$, and let λ_0 be the simple eigenvalue of A surrounded by the closed branch of the shell $\Gamma(A)$. Then the set $\mathcal{A}(\lambda_0)$ is "symmetric" with respect to 0, in the sense that $\theta \in \mathcal{A}(\lambda_0)$ if and only if $2\pi - \theta \in \mathcal{A}(\lambda_0)$, and it is a subset of $\left[0, \frac{\pi}{2}\right) \cup \left(\frac{3\pi}{2}, 2\pi\right]$.*

Proof. The first part of the proposition follows directly from Proposition 3.5 (i).

For the second part of the proposition, by the symmetry of the spectrum $\sigma(A)$ with respect to the real axis, it is apparent that for every $\phi \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$, $e^{\mathrm{i}\phi}\lambda_0$ has real part less than or equal to the real part of some other eigenvalue of matrix $e^{\mathrm{i}\phi}A$, and thus, $\phi \notin \mathcal{A}(\lambda_0)$. □

Corollary 3.7. *If A is an $n \times n$ real matrix with $D(-A) > 0$, and λ_0 is the simple eigenvalue of A surrounded by the closed branch of $-\Gamma(-A)$, then $\mathcal{A}(\lambda_0)$ is symmetric with respect to π, and it is a subset of $\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$.*

Remark 3.8. In this work, we are mainly interested in matrices which have shell-extremal eigenvalues, because these (simple and extremal) eigenvalues appear to have relatively low non-normality. On the other hand, it is worth mentioning that not all (square) matrices need to have shell-extremal eigenvalues. For instance, it is clear that a matrix with all its extremal eigenvalues multiple cannot have shell-extremal eigenvalues. Moreover, a matrix may not have shell-extremal eigenvalues when its non-normality is relatively high and/or every simple extremal eigenvalue of it is close enough to some other eigenvalue of the same matrix. For example, the upper triangular matrices

$$A = \begin{bmatrix} -2 & 6 & -5 & 4 \\ 0 & -3 & 8 & -7 \\ 0 & 0 & 4 & 9 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & -7 & 9 & 6 & -8 \\ 0 & -1 & -6 & 8 & -8 \\ 0 & 0 & -6 & 7 & -9 \\ 0 & 0 & 0 & -7 & 5 \\ 0 & 0 & 0 & 0 & 6 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 5 & 1 & 4 & \mathrm{i} & \mathrm{i} \\ 0 & 6 & 2 & 5 & \mathrm{i}\,4 \\ 0 & 0 & \mathrm{i}\,4 & 3 & \mathrm{i}\,4 \\ 0 & 0 & 0 & 7 & -\mathrm{i}\,6 \\ 0 & 0 & 0 & 0 & \mathrm{i}\,5 \end{bmatrix}$$

do not have shell-extremal eigenvalues. This is confirmed by Figure 2, where the functions $D(e^{\mathrm{i}\theta}A)$, $D(e^{\mathrm{i}\theta}B)$ and $D(e^{\mathrm{i}\theta}C)$ are evaluated for $\theta \in [0, 2\pi]$; the "symmetry" with respect to 0 and the symmetry with respect to π of Proposition 3.5 are apparently confirmed for the real matrices A and B.

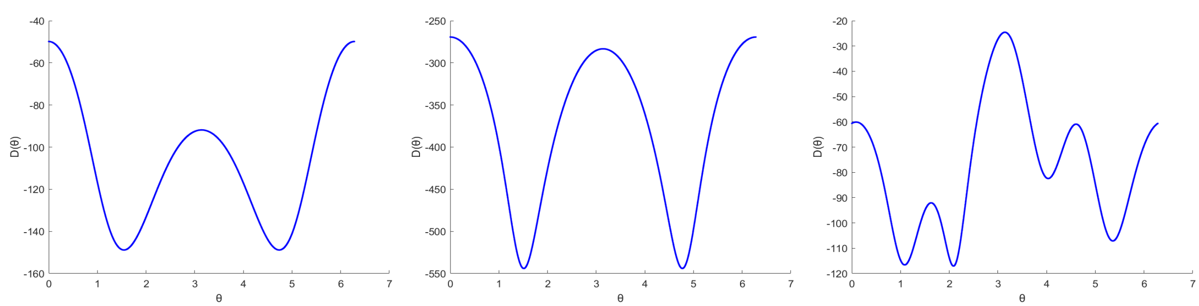

Figure 2: The functions $D(e^{\mathrm{i}\theta}A)$ (left), $D(e^{\mathrm{i}\theta}B)$ (middle), and $D(e^{\mathrm{i}\theta}C)$ (right) take negative values for all $\theta \in [0, 2\pi]$.

The next lemma is a direct implication of Schur's triangularization.

Lemma 3.9. *Let $A \in \mathbb{C}^{n\times n}$, and let λ_0 be an eigenvalue of A. Then λ_0 is a normal eigenvalue of A of multiplicity k if and only if there is a unitary matrix $U \in \mathcal{U}(n)$ such that $A = UTU^*$, where the matrix T is of the form $T = \lambda_0 I_k \oplus T_1$ with $T_1 \in \mathbb{C}^{(n-k)\times(n-k)}$ a triangular matrix that does not have λ_0 as an eigenvalue.*

The following theorem is crucial for our analysis. In particular, motivated by Remark 1.1, we prove that normal shell-extremal eigenvalues appear as isolated points of shells; that is, the closed branch of any shell that surrounds a normal eigenvalue is reduced to a singleton.

Theorem 3.10. *Let $A \in \mathbb{C}^{n\times n}$, and let λ_0 be a shell-extremal eigenvalue of A. Then λ_0 is a normal eigenvalue of A if and only if $e^{\mathrm{i}\theta}\lambda_0$ is an isolated point of the curve $\Gamma(e^{\mathrm{i}\theta}A)$ for every $\theta \in \mathcal{A}(\lambda_0)$. When this is the case, λ_0 is a corner of $F(A)$.*

Proof. Since λ_0 is a (simple) shell-extremal eigenvalue of A, it is obvious that $\mathcal{A}(\lambda_0) \neq \emptyset$. For any $\theta \in \mathcal{A}(\lambda_0)$, we have that $D\left(e^{\mathrm{i}\theta}A\right) > 0$, and $e^{\mathrm{i}\theta}\lambda_0$ is an isolated point of the curve $\Gamma(e^{\mathrm{i}\theta}A)$ if and only if

$$d_h\left(e^{\mathrm{i}\theta}A\right) = \frac{\delta_1(e^{\mathrm{i}\theta}A) - \delta_2(e^{\mathrm{i}\theta}A) - \sqrt{D\left(e^{\mathrm{i}\theta}A\right)}}{2} = 0$$

(see the proof of Corollary 2.2), or equivalently, if and only if

$$\delta_1(e^{\mathrm{i}\theta}A) - \delta_2(e^{\mathrm{i}\theta}A) = \sqrt{\left(\delta_1(e^{\mathrm{i}\theta}A) - \delta_2(e^{\mathrm{i}\theta}A)\right)^2 - 4\left(\nu(e^{\mathrm{i}\theta}A) - u(e^{\mathrm{i}\theta}A)^2\right)},$$

or equivalently, if and only if

$$\nu(e^{\mathrm{i}\theta}A) - u(e^{\mathrm{i}\theta}A)^2 = 0. \tag{9}$$

By Remark 1.1 and the proof of Theorem 3.1 in [1] (see also Proposition 3.15 below), equality (9) holds if and only if $e^{\mathrm{i}\theta}\lambda_0$ is a normal eigenvalue of $e^{\mathrm{i}\theta}A$. Hence, $e^{\mathrm{i}\theta}\lambda_0$ is an isolated point of the curve $\Gamma(e^{\mathrm{i}\theta}A)$ if and only if λ_0 is a normal eigenvalue of A.

Suppose now that λ_0 is a normal shell-extremal eigenvalue of A. By the first part of the theorem and Remark 3.2, there exist $\theta_1, \theta_2 \in [0, 2\pi]$, with $\theta_1 < \theta_2$, such that for every $\theta \in (\theta_1, \theta_2)$, $D\left(e^{\mathrm{i}\theta}A\right) > 0$ and $e^{\mathrm{i}\theta}\lambda_0$ is an isolated point of $\Gamma\left(e^{\mathrm{i}\theta}A\right)$. This means that for every $\theta \in (\theta_1, \theta_2)$, $e^{\mathrm{i}\theta}\lambda_0$ is the right-most point of $\Gamma\left(e^{\mathrm{i}\theta}A\right)$ and the right-most point of $F\left(e^{\mathrm{i}\theta}A\right)$. As a consequence, λ_0 is a corner of $F(A)$. □

Corollary 3.11. *Let $A \in \mathbb{C}^{n\times n}$, and let λ_0 be a simple eigenvalue of A. Then the following are equivalent:*

(a) *For some $\theta \in [0, 2\pi]$, $e^{\mathrm{i}\theta}\lambda_0$ is an isolated point of $\Gamma(e^{\mathrm{i}\theta}A)$.*

(b) *For every $\theta \in \mathcal{A}(\lambda_0) \neq \emptyset$, $e^{\mathrm{i}\theta}\lambda_0$ is an isolated point of $\Gamma(e^{\mathrm{i}\theta}A)$.*

(c) *λ_0 is a normal shell-extremal eigenvalue of A.*

(d) *λ_0 is a corner of $F(A)$.*

Proof. By Theorem 3.10 and its proof, the statements (a), (b) and (c) are equivalent and yield (d).

Suppose now that λ_0 is a corner of $F(A)$. Then, by [7], λ_0 is a normal eigenvalue of A and there is a unitary matrix $U \in \mathcal{U}(n)$ such that $A = UTU^*$, where the matrix T is of the form $T = [\lambda_0] \oplus T_1$ with $T_1 \in \mathbb{C}^{(n-1)\times(n-1)}$ a triangular matrix that does not have λ_0 as an eigenvalue (see Lemma 3.9).

Moreover, without loss of generality, we may assume that λ_0 is the right-most point of $F(A)$ and $\Gamma(A)$ (otherwise, we can work with an appropriate rotation of A). Since λ_0 is a simple normal eigenvalue of A, $\delta_1(A) = \mathrm{Re}\,(\lambda_0)$ is a simple eigenvalue of $H(A)$ [9, Lemma 1], and thus, $\delta_1(A) > \delta_2(A)$ and $D(A) = (\delta_1(A) - \delta_2(A))^2 > 0$. As a consequence, the shell $\Gamma(A)$ has a closed branch that is reduced to the singleton $\{\lambda_0\}$, and λ_0 is a normal shell-extremal eigenvalue of A. □

Let $A \in \mathbb{C}^{n\times n}$, and let λ_0 be a (simple) shell-extremal eigenvalue of A. Let also $\mathbf{x}_0 \in \mathbb{C}^n$ be a unit eigenvector of A corresponding to λ_0, recall that $\mathbf{y}_1 \in \mathbb{C}^n$ is a unit eigenvector of $H(A)$ corresponding to $\delta_1(A)$, and denote by $\mathbf{y}_1\left(e^{\mathrm{i}\theta}A\right)$ a unit eigenvector of $H\left(e^{\mathrm{i}\theta}A\right)$ corresponding to $\delta_1\left(e^{\mathrm{i}\theta}A\right)$ for any $\theta \in [0, 2\pi]$ (clearly, $\mathbf{y}_1 = \mathbf{y}_1(A)$). Moreover, recall that the (real nonnegative) *cosine* and *sine* of the angle between two nonzero vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ are defined by

$$\cos\left(\widehat{\mathbf{x},\mathbf{y}}\right) = \frac{|\mathbf{y}^*\mathbf{x}|}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2} \quad \text{and} \quad \sin\left(\widehat{\mathbf{x},\mathbf{y}}\right) = \sqrt{1 - \cos\left(\widehat{\mathbf{x},\mathbf{y}}\right)^2}.$$

Definition 3.12. For a shell-extremal eigenvalue λ_0 of a (square) matrix A, the following quantities can be considered as measures of non-normality:

- The infimum of the horizontal diameters

$$\eta_{A,1}(\lambda_0) = \inf\left\{ d_h\left(e^{\mathrm{i}\theta}A\right) = \frac{\delta_1\left(e^{\mathrm{i}\theta}A\right) - \delta_2\left(e^{\mathrm{i}\theta}A\right) - \sqrt{D\left(e^{\mathrm{i}\theta}A\right)}}{2} \,:\, \theta \in \mathcal{A}(\lambda_0) \right\};$$

  this is a direct implication of Theorem 3.10 and Corollaries 2.2 and 2.3.

- The infimum

$$\eta_{A,2}(\lambda_0) = \inf\left\{ \nu\left(e^{\mathrm{i}\theta}A\right) - u\left(e^{\mathrm{i}\theta}A\right)^2 \,:\, \theta \in \mathcal{A}(\lambda_0) \right\};$$

  this quantity is equal to zero if and only if span$\{\mathbf{x}_0\}$ coincides with the one-dimensional eigenspace of $H\left(e^{\mathrm{i}\theta}A\right)$ corresponding to $\delta_1\left(e^{\mathrm{i}\theta}A\right)$ for some $\theta \in \mathcal{A}(\lambda_0)$, or equivalently, if and only if the shell-extremal eigenvalue λ_0 is a normal eigenvalue of A on the boundary of the numerical range $F(A)$ (see also Remark 1.1 and the proof of Theorem 3.10).

- The infimum of the sines

$$\eta_{A,3}(\lambda_0) = \inf\left\{ \sin\left(\widehat{\mathbf{x}_0, \mathbf{y}_1\left(e^{\mathrm{i}\theta}A\right)}\right) \,:\, \theta \in \mathcal{A}(\lambda_0) \right\};$$

  as both λ_0 and $\delta_1\left(e^{\mathrm{i}\theta}A\right)$ ($\theta \in \mathcal{A}(\lambda_0)$) are simple, this quantity is equal to zero if and only if $\mathbf{x}_0 \in \mathrm{span}\left\{\mathbf{y}_1\left(e^{\mathrm{i}\theta}A\right)\right\}$ (which means that $\mathbf{y}_1\left(e^{\mathrm{i}\theta}A\right)$ is an eigenvector of both $H\left(e^{\mathrm{i}\theta}A\right)$ and A), or equivalently, if and only if the shell-extremal eigenvalue λ_0 is a normal eigenvalue of A [9, Lemma 1].

- The infimum of the radii of curvature

$$\eta_{A,4}(\lambda_0) = \inf\left\{ R_{\Gamma(e^{\mathrm{i}\theta}A)}(\delta_1(e^{\mathrm{i}\theta}A) + \mathrm{i}u(e^{\mathrm{i}\theta}A)) \,:\, \theta \in \mathcal{A}(\lambda_0) \right\};$$

  this is a direct implication of Theorem 3.10.

The next result will help us to obtain a connection between the measures $\eta_{A,1}(\lambda_0)$, $\eta_{A,2}(\lambda_0)$ and $\eta_{A,3}(\lambda_0)$.

Proposition 3.13. *Let $A \in \mathbb{C}^{n\times n}$ with $D(A) > 0$, and let $\mathbf{x}_0 \in \mathbb{C}^n$ be a unit eigenvector of A corresponding to the eigenvalue of A that is surrounded by the closed branch of $\Gamma(A)$. Then, it holds that*

$$\sin\left(\widehat{\mathbf{x}_0, \mathbf{y}_1}\right) \le \sqrt{\frac{d_h(A)}{\delta_1(A) - \delta_2(A)}} = \sqrt{\frac{1}{2}\left(1 - \sqrt{1 - 8\,\frac{R_{\Gamma(A)}\left(\delta_1(A) + \mathrm{i}\,u(A)\right)}{\delta_1(A) - \delta_2(A)}}\right)}.$$

Proof. Theorem 5.1 of [14] states that $\cos\left(\widehat{\mathbf{x}_0, \mathbf{y}_1}\right) \ge \sqrt{\frac{1}{2} + \frac{\sqrt{D(A)}}{2\left(\delta_1(A) - \delta_2(A)\right)}}$. As a consequence,

$$\sin\left(\widehat{\mathbf{x}_0, \mathbf{y}_1}\right)^2 = 1 - \cos\left(\widehat{\mathbf{x}_0, \mathbf{y}_1}\right)^2 \le \frac{1}{2} - \frac{\sqrt{D(A)}}{2\left(\delta_1(A) - \delta_2(A)\right)} = \frac{\delta_1(A) - \delta_2(A) - \sqrt{D(A)}}{2\left(\delta_1(A) - \delta_2(A)\right)}. \tag{10}$$

Recalling the definition of $d_h(A)$ in (1), the inequality in (10) holds if and only if

$$\sin\left(\widehat{\mathbf{x}_0, \mathbf{y}_1}\right) \le \sqrt{\frac{d_h(A)}{\delta_1(A) - \delta_2(A)}}.$$

Observing also that (5) yields

$$\begin{aligned} d_h(A) &= \frac{\delta_1(A) - \delta_2(A) - \sqrt{(\delta_1(A) - \delta_2(A))^2 - 4(\nu(A) - u(A)^2)}}{2} \\ &= \frac{\delta_1(A) - \delta_2(A)}{2}\left(1 - \sqrt{1 - 8\,\frac{R_{\Gamma(A)}(\delta_1(A) + \mathrm{i}\,u(A))}{\delta_1(A) - \delta_2(A)}}\right), \end{aligned}$$

the result follows. □

Corollary 3.14. *Let $A \in \mathbb{C}^{n\times n}$, and let λ_0 be a shell-extremal eigenvalue of A. Then the following hold:*

(i) $2\,\eta_{A,4}(\lambda_0) \le \eta_{A,1}(\lambda_0) \le 4\,\eta_{A,4}(\lambda_0)$.

(ii) If λ_0 is not a normal eigenvalue of A, then

$$\eta_{A,1}(\lambda_0) \ge 2\sqrt{\eta_{A,2}(\lambda_0)}\;\eta_{A,3}(\lambda_0)^2.$$

Proof. (i) The double inequality $2\,\eta_{A,4}(\lambda_0) \le \eta_{A,1}(\lambda_0) \le 4\,\eta_{A,4}(\lambda_0)$ is derived immediately from Proposition 2.6 (ii).

(ii) Since λ_0 is a shell-extremal eigenvalue of matrix A, $\mathcal{A}(\lambda_0) \ne \emptyset$ and $\delta_1(e^{\mathrm{i}\theta}A) > \delta_2(e^{\mathrm{i}\theta}A)$ for every $\theta \in \mathcal{A}(\lambda_0)$. If, in addition, λ_0 is not a normal eigenvalue of A, then for every $\theta \in \mathcal{A}(\lambda_0)$,

$$\delta_1(e^{\mathrm{i}\theta}A) - \delta_2(e^{\mathrm{i}\theta}A) > 2\sqrt{\nu(e^{\mathrm{i}\theta}A) - u(e^{\mathrm{i}\theta}A)^2} \ge 2\sqrt{\eta_{A,2}(\lambda_0)} > 0.$$

As a consequence, Proposition 3.13 implies that for every $\theta \in \mathcal{A}(\lambda_0)$,

$$\eta_{A,3}(\lambda_0) \le \sin\left(\widehat{\mathbf{x}_0, \mathbf{y}_1\,(e^{\mathrm{i}\theta}A)}\right) \le \sqrt{\frac{d_h(e^{\mathrm{i}\theta}A)}{\delta_1(e^{\mathrm{i}\theta}A) - \delta_2(e^{\mathrm{i}\theta}A)}} < \sqrt{\frac{d_h(e^{\mathrm{i}\theta}A)}{2\sqrt{\eta_{A,2}(\lambda_0)}}},$$

and hence,

$$\eta_{A,1}(\lambda_0) \ge 2\sqrt{\eta_{A,2}(\lambda_0)}\;\eta_{A,3}(\lambda_0)^2. \quad \square$$

As in the proof of Proposition 2.6, consider an $n\times n$ unitary matrix $U=\begin{bmatrix}\mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_n\end{bmatrix}$, with (orthonormal) columns $\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n$, such that $H(A)=U\mathrm{diag}\{\delta_1(A),\delta_2(A),\ldots,\delta_n(A)\}U^*$. Then

$$A=U\left[\begin{array}{c|c}\delta_1(A)+\mathrm{i}\,u(A) & \mathbf{z}^* \\ \hline -\mathbf{z} & \mathrm{diag}\{\delta_2(A),\ldots,\delta_n(A)\}+K_1\end{array}\right]U^*, \tag{11}$$

where $\mathbf{z}\in\mathbb{C}^{n-1}$, with $\|\mathbf{z}\|_2=\sqrt{\nu(A)-u(A)^2}$, and K_1 is an $(n-1)\times(n-1)$ skew-hermitian matrix (see also the proof of Theorem 3.1 in [1]). The equality (11) yields the definition of $\eta_{A,2}(\lambda_0)$ and can be refined in the following form.

Proposition 3.15. *For any $A\in\mathbb{C}^{n\times n}$, there exists a unitary matrix $W\in\mathcal{U}(n)$ such that*

$$A=W\left[\begin{array}{c|cccc}\delta_1(A)+\mathrm{i}\,u(A) & -\sqrt{\nu(A)-u(A)^2} & 0 & \cdots & 0 \\ \hline \sqrt{\nu(A)-u(A)^2} & & & & \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{array}\right]W^*,$$

where A_1 is an $(n-1)\times(n-1)$ matrix.

Proof. Consider the matrix $\hat{A}=A-(\delta_1(A)+\mathrm{i}\,u(A))\,I_n$, and observe that $\delta_1(\hat{A})=0$, $u(\hat{A})=0$, and we can choose a common unit eigenvector $\mathbf{y}_1=\mathbf{y}_1(A)=\mathbf{y}_1(\hat{A})$ of the hermitian matrices $H(A)$ and $H(\hat{A})$, corresponding to $\delta_1(A)$ and $\delta_1(\hat{A})=0$, respectively. As a consequence,

$$\begin{aligned}\nu(\hat{A}) &= \|K(\hat{A})\mathbf{y}_1\|_2^2 \;=\; \mathbf{y}_1^*K(\hat{A})^*K(\hat{A})\mathbf{y}_1 \;=\; -\mathbf{y}_1^*K(\hat{A})^2\mathbf{y}_1 \\ &= -\mathbf{y}_1^*\left(K(A)-\mathrm{i}\,u(A)I_n\right)^2\mathbf{y}_1 \;=\; -\mathbf{y}_1^*\left(K(A)^2-2\,\mathrm{i}\,u(A)K(A)-u(A)^2I_n\right)\mathbf{y}_1 \\ &= -\mathbf{y}_1^*K(A)^2\mathbf{y}_1-2u(A)^2+u(A)^2 \;=\; \mathbf{y}_1^*K(A)^*K(A)\mathbf{y}_1-u(A)^2 \;=\; \nu(A)-u(A)^2.\end{aligned}$$

Moreover, we can see that

$$\begin{aligned}\|\hat{A}\mathbf{y}_1\|_2^2 &= \mathbf{y}_1^*\hat{A}^*\hat{A}\mathbf{y}_1 \;=\; \mathbf{y}_1^*(H(\hat{A})+K(\hat{A}))^*(H(\hat{A})+K(\hat{A}))\mathbf{y}_1 \\ &= \mathbf{y}_1^*(H(\hat{A})^2+H(\hat{A})K(\hat{A})+K(\hat{A})^*H(\hat{A})+K(\hat{A})^*K(\hat{A}))\mathbf{y}_1 \\ &= \mathbf{y}_1^*K(\hat{A})^*K(\hat{A})\mathbf{y}_1 \;=\; \nu(\hat{A}) \;=\; \nu(A)-u(A)^2,\end{aligned}$$

i.e., $\|\hat{A}\mathbf{y}_1\|_2=\sqrt{\nu(A)-u(A)^2}$.

If $\hat{A}\mathbf{y}_1=\mathbf{0}$, then it is obvious that $\nu(A)-u^2(A)=0$, and thus, (11) holds with $\mathbf{z}=\mathbf{0}$.

If $\hat{A}\mathbf{y}_1\neq\mathbf{0}$, then consider the unit vector $\mathbf{w}_1=\dfrac{\hat{A}\mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2}$ and observe that $\mathbf{y}_1^*\mathbf{w}_1=\dfrac{\mathbf{y}_1^*\hat{A}\mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2}=0$. Furthermore, extend the pair $\{\mathbf{y}_1,\mathbf{w}_1\}$ to an orthonormal basis $\{\mathbf{y}_1,\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_{n-1}\}$, consider the unitary matrix $W=\begin{bmatrix}\mathbf{y}_1 & \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_{n-1}\end{bmatrix}$, and define the matrix $B=[b_{i,j}]=W^*\hat{A}W$. Keeping in mind that

$$H(\hat{A})\mathbf{y}_1=\mathbf{0},\quad \mathbf{y}_1^*H(\hat{A})=\mathbf{0}^T,\quad \mathbf{y}_1^*K(\hat{A})\mathbf{y}_1=0\quad\text{and}\quad \|\hat{A}\mathbf{y}_1\|_2=\sqrt{\nu(A)-u(A)^2},$$

it is straightforward to verify that

$$b_{1,1}=\mathbf{y}_1^*\hat{A}\mathbf{y}_1=\mathbf{y}_1^*(H(\hat{A})+K(\hat{A}))\mathbf{y}_1=\mathbf{y}_1^*H(\hat{A})\mathbf{y}_1+\mathbf{y}_1^*K(\hat{A})\mathbf{y}_1=0,$$

$$
\begin{aligned}
b_{1,2} &= \mathbf{y}_1^* \hat{A} \mathbf{w}_1 = \frac{\mathbf{y}_1^* \hat{A}^2 \mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2} = \frac{\mathbf{y}_1^*(H(\hat{A}) + K(\hat{A}))^2 \mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2} \\
&= \frac{\mathbf{y}_1^*(H(\hat{A})^2 + H(\hat{A})K(\hat{A}) + K(\hat{A})H(\hat{A}) + K(\hat{A})^2)\mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2} \\
&= \frac{\mathbf{y}_1^* K(\hat{A})^2 \mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2} = -\frac{\mathbf{y}_1^* K(\hat{A})^* K(\hat{A}) \mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2} = -\frac{\nu(\hat{A})}{\|\hat{A}\mathbf{y}_1\|_2} \\
&= -\frac{\nu(A) - u(A)^2}{\|\hat{A}\mathbf{y}_1\|_2} = -\sqrt{\nu(A) - u(A)^2}
\end{aligned}
$$

and

$$
b_{2,1} = \mathbf{w}_1^* \hat{A} \mathbf{y}_1 = \frac{\mathbf{y}_1^* \hat{A}^* \hat{A} \mathbf{y}_1}{\|\hat{A}\mathbf{y}_1\|_2} = \|\hat{A}\mathbf{y}_1\|_2 = \sqrt{\nu(A) - u(A)^2}.
$$

The rest of the entries of the first column of B, namely, $b_{3,1}, b_{4,1}, \ldots, b_{n,1}$, are

$$
b_{k+1,1} = \mathbf{w}_k^* \hat{A} \mathbf{y}_1 = \|\hat{A}\mathbf{y}_1\|_2 \mathbf{w}_k^* \mathbf{w}_1 = 0, \quad k = 2, 3, \ldots, n-1.
$$

The rest of the entries of the first row of B, namely, $b_{1,3}, b_{1,4}, \ldots, b_{1,n}$, are

$$
\begin{aligned}
b_{1,k+1} &= \mathbf{y}_1^* \hat{A} \mathbf{w}_k = \mathbf{y}_1^* K(\hat{A}) \mathbf{w}_k = \overline{\mathbf{w}_k^* K(\hat{A})^* \mathbf{y}_1} \\
&= -\overline{\mathbf{w}_k^* K(\hat{A}) \mathbf{y}_1} = -\overline{\mathbf{w}_k^* \hat{A} \mathbf{y}_1} = -\bar{b}_{k+1,1} = 0, \quad k = 2, 3, \ldots, n-1,
\end{aligned}
$$

and the proof is complete. □

4 Examples

Example 4.1. Consider the 4×4 real *almost skew-symmetric* (i.e., with rank-1 symmetric part) matrix

$$
A = K + \mathbf{x}\mathbf{x}^* = \begin{bmatrix} 4.0000 & -0.7523 & -0.7050 & -0.1972 \\ 0.7523 & 0 & 0.5568 & -0.6441 \\ 0.7050 & -0.5568 & 0 & -0.6782 \\ 0.1972 & 0.6441 & 0.6782 & 0 \end{bmatrix},
$$

where K is a randomly generated skew-symmetric matrix and $\mathbf{x} = [2, 0, 0, 0]^T$. The eigenvalues of A are

$$
\lambda_1 = 3.7276, \quad \lambda_2 = 0.1359 + \mathrm{i}\,1.1192, \quad \lambda_3 = 0.1359 - \mathrm{i}\,1.1192 \quad \text{and} \quad \lambda_4 = 0.0006,
$$

and the only shell-extremal eigenvalue is λ_1. We evaluate the rotated shells $e^{-\mathrm{i}\,\theta}\Gamma(e^{\mathrm{i}\theta}A)$ for some values of $\theta \in [0, 2\pi]$ in Figure 3. In the same figure, we sketch the numerical range $F(A)$ and mark with asterisks the eigenvalues of A. In Table 1, we notice that the inequalities in both parts of Proposition 2.6 are confirmed, and the ratio $\frac{R_{F(e^{\mathrm{i}\theta}A)}(\delta_1(e^{\mathrm{i}\theta}A) + \mathrm{i}\,u(e^{\mathrm{i}\theta}A))}{R_{\Gamma(e^{\mathrm{i}\theta}A)}(\delta_1(e^{\mathrm{i}\theta}A) + \mathrm{i}\,u(e^{\mathrm{i}\theta}A))}$ attains the maximum bound 4 provided by Proposition 2.6 (i).

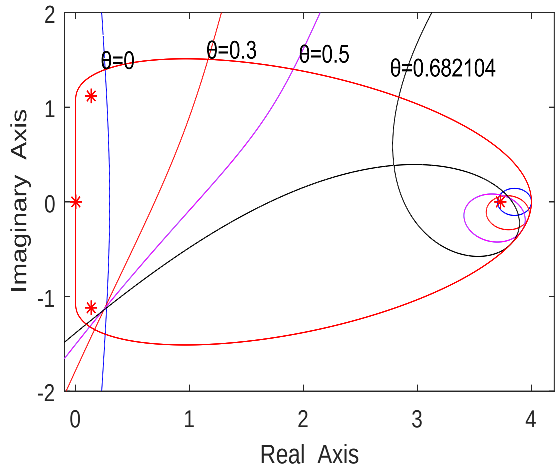

Figure 3: For the right-most point $\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A)$ of every shell $\Gamma(e^{-i\theta}A)$ ($\theta\in[0,2\pi]$), the scalar $e^{-i\theta}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))$ is always a boundary point of the numerical range $F(A)$.

θ	0	0.3	0.5	0.682104
$\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A)$	4	$3.9736-i\,0.1706$	$3.9176-i\,0.3015$	$3.8184-i\,0.4476$
$D(e^{i\theta}A)$	11.5926	7.7161	3.9972	$8.93\cdot 10^{-5}$
$d_h(e^{i\theta}A)$	0.2976	0.3802	0.5507	1.3162
$\nu(e^{i\theta}A)-u(e^{i\theta}A)^2$	1.1019	1.2008	1.4043	1.7447
$R_{\Gamma(e^{i\theta}A)}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))$	0.1377	0.1697	0.2265	0.3302
$\dfrac{R_{F(e^{i\theta}A)}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))}{R_{\Gamma(e^{i\theta}A)}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))}$	4	3.6775	3.4594	3.2519

Table 1: The quantities $\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A)$, $D(e^{i\theta}A)$, $d_h(e^{i\theta}A)$, $\nu(e^{i\theta}A)-u(e^{i\theta}A)^2$, $R_{\Gamma(e^{i\theta}A)}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))$ and $\dfrac{R_{F(e^{i\theta}A)}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))}{R_{\Gamma(e^{i\theta}A)}(\delta_1(e^{i\theta}A)+i\,u(e^{i\theta}A))}$ for $\theta=0,0.3,0.5,0.682104$.

Example 4.2. Let

$$A=\begin{bmatrix} 5 & 1 & 0 & 0.1\\ 0 & 3 & 0 & 2\\ 0 & 0 & -4 & 0.4\\ 0 & 0 & 0 & 0\end{bmatrix}$$

with $\sigma(A) = \{-4, 0, 3, 5\}$. The eigenvalues $\lambda_1 = -4$ and $\lambda_2 = 5$ are shell-extremal eigenvalues, with

$$\mathcal{A}(\lambda_1) = [0, 0.7165) \cup (2\pi - 0.7165, 2\pi] \quad \text{and} \quad \mathcal{A}(\lambda_2) = (2.0255,\ 4.2576)$$

(see Figure 4). Notice that $D(e^{\mathrm{i}\theta}A)$ ($\theta \in [0, 2\pi]$) and $\mathcal{A}(\lambda_1)$ are "symmetric" with respect to 0, and $D(e^{\mathrm{i}\theta}A)$ ($\theta \in [0, 2\pi]$) and $\mathcal{A}(\lambda_2)$ are symmetric with respect to π, verifying Propositions 3.5 and 3.6, and Corollary 3.7.

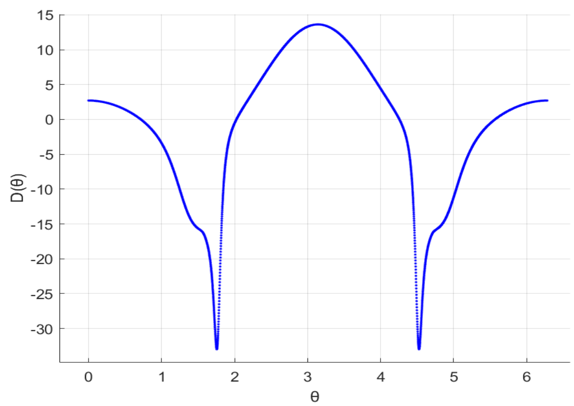

Figure 4: The function $D(e^{\mathrm{i}\theta}A)$, $\theta \in [0, 2\pi]$.

Example 4.3. Consider the 7×7 upper triangular matrix

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 + \mathrm{i}\,3 & 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & \mathrm{i}\,4 & 3 & -2 & \mathrm{i}\,3 \\ 0 & 0 & 0 & 0 & 1 + \mathrm{i}\,3 & -1.5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1.5 + \mathrm{i}\,3 & -0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathrm{i}\,3 \end{bmatrix},$$

which has two shell-extremal eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 0$, with

$$\mathcal{A}(\lambda_1) = [0, 1.4655) \cup (2\pi - 0.1585, 2\pi] \quad \text{and} \quad \mathcal{A}(\lambda_2) = (1.5721, 2.4262).$$

In Figure 5, we sketch the shell $\Gamma(A)$ and the numerical range $F(A)$, and mark with asterisks the eigenvalues of A. The eigenvalue $\lambda_2 = 0$ is a normal eigenvalue of A and a corner of $F(A)$. It is worth noting that in this example, the closed branch of the shell $\Gamma(A)$ that surrounds $\lambda_1 = 5$ does not lie (entirely) in the numerical range $F(A)$. Moreover,

$$\nu(A) - u(A)^2 = 0.0313 \quad \text{and} \quad d_h(A) = 0.0327,$$

which is consistent with the fact that the only nonzero off-diagonal entry of A in the row and the column of the diagonal entry 5 is equal to 0.2 and the eigenvalue $\lambda_1 = 5$ can be considered as "nearly" normal.

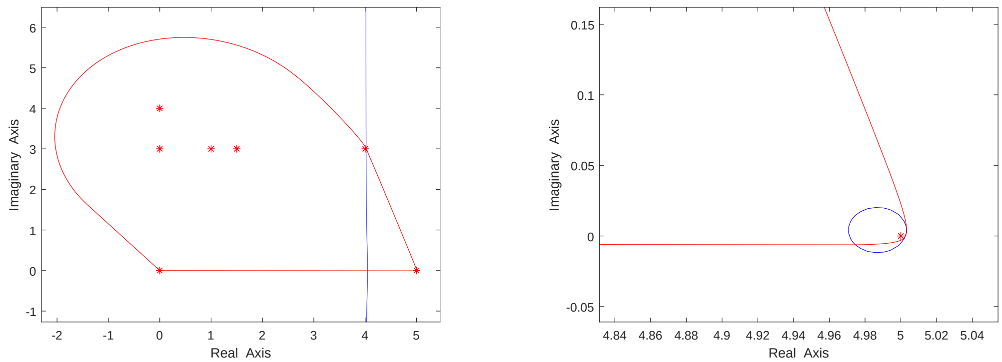

Figure 5: In the left part, we have the numerical range and the shell of A, and in the right part, we have a magnified picture of the closed branch of the shell that surrounds the shell-extremal eigenvalue $\lambda_1 = 5$.

Appendix A The area enclosed by the closed branch of $\Gamma(A)$

Let A be an $n \times n$ matrix such that $\delta_1(A) > \delta_2(A)$, $\nu(A) - u(A)^2 > 0$ and $D(A) = (\delta_1(A) - \delta_2(A))^2 - 4\left(\nu(A) - u(A)^2\right) > 0$. Denote also by $\mathcal{E}$ the (nonzero) area enclosed by the closed branch of the shell $\Gamma(A)$, and by $\mathcal{E}_{ub}$ the upper bound of this area, given by Corollary 2.4, i.e.,

$$\mathcal{E}_{ub} = d_h(A)^2 \sqrt{\frac{2}{3}\left(1 - \ln\left(1 + \frac{d_h(A)}{d_h(A) + \sqrt{D(A)}}\right)\right)}. \tag{12}$$

In this appendix, we describe a method for the computation of the area $\mathcal{E}$, using elliptic integrals, suggested by an anonymous referee.

By (4) (in the proof of Corollary 2.4), we know that

$$\mathcal{E} = 2\, d_h(A)^2 \int_0^1 \sqrt{t(1-t)\left[1 - \frac{d_h(A)}{(1+t)d_h(A) + \sqrt{D(A)}}\right]}\, dt.$$

If we set

$$w = \sqrt{\frac{\delta_1(A) - \delta_2(A) + 2\sqrt{\nu(A) - u(A)^2}}{\delta_1(A) - \delta_2(A) - 2\sqrt{\nu(A) - u(A)^2}}} \in (1, +\infty),$$

then

$$2\sqrt{\nu(A) - u(A)^2} = (\delta_1(A) - \delta_2(A))\,\frac{w^2 - 1}{w^2 + 1}, \quad D(A) = (\delta_1(A) - \delta_2(A))^2\,\frac{4w^2}{(w^2+1)^2}$$

and (recalling the definition of $d_h(A)$ in (1))

$$d_h(A) = \frac{\delta_1(A) - \delta_2(A) - \sqrt{(\delta_1(A) - \delta_2(A))^2\,\dfrac{4w^2}{(w^2+1)^2}}}{2} = (\delta_1(A) - \delta_2(A))\,\frac{(w-1)^2}{2(w^2+1)}.$$

As a consequence, straightforward calculations imply that

$$\begin{aligned}
\mathcal{E} &= (\delta_1(A)-\delta_2(A))^2\,\frac{(w-1)^4}{2(w^2+1)^2}\int_0^1\sqrt{t(1-t)\left[1-\frac{(w-1)^2}{(1+t)(w-1)^2+4w}\right]}\,dt\\
&= (\delta_1(A)-\delta_2(A))^2\,\frac{(w-1)^4}{2(w^2+1)^2}\int_0^1\sqrt{t(1-t)\,\frac{t(w-1)^2+4w}{t(w-1)^2+(w+1)^2}}\,dt\\
&= (\delta_1(A)-\delta_2(A))^2\,\frac{(w-1)^6}{2(w^2+1)^2}\int_0^1\frac{\sqrt{-t\left[t+\frac{4w}{(w-1)^2}\right](t-1)\left[t+1+\frac{4w}{(w-1)^2}\right]}}{(w+1)^2+t(w-1)^2}\,dt.
\end{aligned}$$

Setting $s=t+\dfrac{2w}{(w-1)^2}$ directly yields

$$\mathcal{E}=(\delta_1(A)-\delta_2(A))^2\,\frac{(w-1)^6}{2(w^2+1)^2}\int_{\frac{2w}{(w-1)^2}}^{\frac{w^2+1}{(w-1)^2}}\frac{\sqrt{-\left[s^2-\frac{4w^2}{(w-1)^4}\right]\left[s^2-\frac{(w^2+1)^2}{(w-1)^4}\right]}}{(w-1)^2s+w^2+1}\,ds.$$

If we consider now the variable $z\in[0,1]$ such that $s=s(z)=\dfrac{1}{(w-1)^2}\sqrt{(w^2+1)^2-(w^2-1)^2z^2}$, then it follows that $s(1)=\dfrac{2w}{(w-1)^2}$, $s(0)=\dfrac{w^2+1}{(w-1)^2}$, $ds=-\dfrac{(w+1)^2z}{\sqrt{(w^2+1)^2-(w^2-1)^2z^2}}\,dz$ and

$$\begin{aligned}
\mathcal{E} &= (\delta_1(A)-\delta_2(A))^2\,\frac{(w^2-1)^4}{2(w^2+1)^2}\int_1^0\frac{-z^2\sqrt{1-z^2}}{\left[w^2+1+\sqrt{(w^2+1)^2-(w^2-1)^2z^2}\right]\sqrt{(w^2+1)^2-(w^2-1)^2z^2}}\,dz\\
&= (\delta_1(A)-\delta_2(A))^2\,\frac{(w^2-1)^2}{2(w^2+1)^2}\int_0^1\frac{\sqrt{1-z^2}\left[w^2+1-\sqrt{(w^2+1)^2-(w^2-1)^2z^2}\right]}{\sqrt{(w^2+1)^2-(w^2-1)^2z^2}}\,dz\\
&= (\delta_1(A)-\delta_2(A))^2\,\frac{(w^2-1)^2}{2(w^2+1)^2}\left[\int_0^1\frac{(w^2+1)\sqrt{1-z^2}}{\sqrt{(w^2+1)^2-(w^2-1)^2z^2}}\,dz-\int_0^1\sqrt{1-z^2}\,dz\right]\\
&= (\delta_1(A)-\delta_2(A))^2\left[\frac{(w^2-1)^2}{2(w^2+1)}\int_0^1\frac{\sqrt{1-z^2}}{\sqrt{(w^2+1)^2-(w^2-1)^2z^2}}\,dz-\frac{(w^2-1)^2}{2(w^2+1)^2}\,\frac{\pi}{4}\right]\\
&= (\delta_1(A)-\delta_2(A))^2\left[\frac{1}{2}\int_0^1\frac{\sqrt{1-\frac{(w^2-1)^2}{(w^2+1)^2}z^2}}{\sqrt{1-z^2}}\,dz-\frac{2w^2}{(w^2+1)^2}\int_0^1\frac{1}{\sqrt{1-z^2}\sqrt{1-\frac{(w^2-1)^2}{(w^2+1)^2}z^2}}\,dz\right]\\
&\quad-\frac{(\delta_1(A)-\delta_2(A))^2}{8}\,\frac{(w^2-1)^2}{(w^2+1)^2}\,\pi. \qquad (13)
\end{aligned}$$

Setting the elliptic integral of first kind $E_1(\lambda)=\int_0^1\dfrac{1}{\sqrt{1-x^2}\sqrt{1-\lambda^2x^2}}\,dx$ and the elliptic integral of second kind $E_2(\lambda)=\int_0^1\dfrac{\sqrt{1-\lambda^2x^2}}{\sqrt{1-x^2}}\,dx$ (see [12]), and keeping in mind that $\dfrac{w^2-1}{w^2+1}=\dfrac{2\sqrt{\nu(A)-u(A)^2}}{\delta_1(A)-\delta_2(A)}$

and $(\delta_1(A)-\delta_2(A))^2\,\dfrac{2w^2}{(w^2+1)^2}=\dfrac{D(A)}{2}$, (13) is written in the form

$$\mathcal{E}=\frac{(\delta_1(A)-\delta_2(A))^2}{2}\,E_2\!\left(\frac{2\sqrt{\nu(A)-u(A)^2}}{\delta_1(A)-\delta_2(A)}\right)-\frac{D(A)}{2}\,E_1\!\left(\frac{2\sqrt{\nu(A)-u(A)^2}}{\delta_1(A)-\delta_2(A)}\right)-\frac{\nu(A)-u(A)^2}{2}\,\pi. \quad (14)$$

By (12) and (14), we can calculate the area $\mathcal{E}$, the upper bound $\mathcal{E}_{ub}$ and the ratio $\dfrac{\mathcal{E}_{ub}}{\mathcal{E}}$ for several values of $\delta_1(A)-\delta_2(A)$ and $\nu(A)-u(A)^2$ such that $D(A)=(\delta_1(A)-\delta_2(A))^2-4\left(\nu(A)-u(A)^2\right)>0$ (see Table 2). Our numerical experiments verify that the size of the closed branch of $\Gamma(A)$ is increasing in the measure of non-normality $\nu(A)-u(A)^2$ and decreasing in the distance $\delta_1(A)-\delta_2(A)$, and remarkably suggest that the ratio $\dfrac{\mathcal{E}_{ub}}{\mathcal{E}}$ is always close to 1.04 (the value 1.0396 appears in most cases).

	$\nu(A)-u(A)^2=0.2499$	$\nu(A)-u(A)^2=1$	$\nu(A)-u(A)^2=2.2499$	$\nu(A)-u(A)^2=5.4$
$\delta_1(A)-\delta_2(A)=1$	$\mathcal{E}=0.1069$ $\mathcal{E}_{ub}=0.1120$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0483$			
$\delta_1(A)-\delta_2(A)=3$	$\mathcal{E}=0.0057$ $\mathcal{E}_{ub}=0.0059$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.1065$ $\mathcal{E}_{ub}=0.1107$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.9652$ $\mathcal{E}_{ub}=1.0150$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0520$	
$\delta_1(A)-\delta_2(A)=6$	$\mathcal{E}=13.7680\cdot 10^{-4}$ $\mathcal{E}_{ub}=14.3130\cdot 10^{-4}$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.0228$ $\mathcal{E}_{ub}=0.0237$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.1224$ $\mathcal{E}_{ub}=0.1272$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.8525$ $\mathcal{E}_{ub}=0.8862$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0395$
$\delta_1(A)-\delta_2(A)=10$	$\mathcal{E}=4.9232\cdot 10^{-4}$ $\mathcal{E}_{ub}=5.1182\cdot 10^{-4}$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.0080$ $\mathcal{E}_{ub}=0.0083$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.0412$ $\mathcal{E}_{ub}=0.0428$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.2500$ $\mathcal{E}_{ub}=0.2599$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$
$\delta_1(A)-\delta_2(A)=15$	$\mathcal{E}=2.1835\cdot 10^{-4}$ $\mathcal{E}_{ub}=2.2700\cdot 10^{-4}$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.0035$ $\mathcal{E}_{ub}=0.0037$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.0179$ $\mathcal{E}_{ub}=0.0187$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$	$\mathcal{E}=0.1056$ $\mathcal{E}_{ub}=0.1098$ $\frac{\mathcal{E}_{ub}}{\mathcal{E}}=1.0396$

Table 2: The area $\mathcal{E}$, the upper bound $\mathcal{E}_{ub}$ and the ratio $\dfrac{\mathcal{E}_{ub}}{\mathcal{E}}$ for $\delta_1(A)-\delta_2(A)=1,3,6,10,15$ and $\nu(A)-u(A)^2=0.2499,1,2.2499,5.4$.

Finally, if we set $\rho(A)=\dfrac{2\sqrt{\nu(A)-u(A)^2}}{\delta_1(A)-\delta_2(A)}\in(0,1)$, then $\nu(A)-u(A)^2=\dfrac{1}{4}(\delta_1(A)-\delta_2(A))^2\rho(A)^2$, $D(A)=(\delta_1(A)-\delta_2(A))^2\left(1-\rho(A)^2\right)$ and $d_h(A)=\dfrac{1}{2}(\delta_1(A)-\delta_2(A))\left(1-\sqrt{1-\rho(A)^2}\right)$, and (12) and (14) yield

$$\mathcal{E}_{ub}=\frac{(\delta_1(A)-\delta_2(A))^2}{2\sqrt{6}}\left(1-\sqrt{1-\rho(A)^2}\right)^2\sqrt{1+\ln\left(\frac{1+\sqrt{1-\rho(A)^2}}{2}\right)}$$

and

$$\mathcal{E}=\frac{(\delta_1(A)-\delta_2(A))^2}{2}\left[E_2(\rho(A))-\left(1-\rho(A)^2\right)E_1(\rho(A))-\frac{\rho(A)^2}{4}\,\pi\right],$$

respectively. Thus, we can have the graph of $\frac{\mathcal{E}_{ub}}{\mathcal{E}}$ as a function of $\rho(A) \in (0,1)$; see Figure 6. This graph suggests that the ratio $\frac{\mathcal{E}_{ub}}{\mathcal{E}}$ takes all its values in the open interval $(1.0385, 1.0540)$, and it is almost equal to 1.0396 for all $\rho(A) \in (0, 0.7)$. This means that $0.9619\,\mathcal{E}_{ub}$ is a satisfactory approximation of the area $\mathcal{E}$.

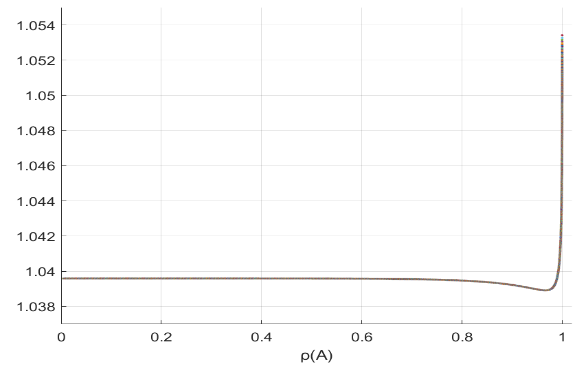

Figure 6: The ratio $\frac{\mathcal{E}_{ub}}{\mathcal{E}}$ as a function of $\rho(A) \in (0,1)$.

Acknowledgments. The authors would like to thank the anonymous referees for their valuable suggestions. They also want to express their gratitude to one of the referees for bringing to their attention the technique described in the appendix for the computation of the area enclosed by the closed branch (when it exists) of the shell $\Gamma(A)$.

References

[1] M. Adam and M. Tsatsomeros, An eigenvalue inequality and spectrum localization for complex matrices, *Electron. J. Linear Algebra*, **15** (2006) 239–250.

[2] Aik. Aretaki, P.J. Psarrakos, and M.J. Tsatsomeros, The envelope of tridiagonal Toeplitz and block-shift matrices, *Linear Algebra Appl.*, **532** (2017) 60–85.

[3] M.-T. Chien and H. Nakazato, Flat portions on the boundary of the numerical ranges of certain Toeplitz matrices, *Linear Multilinear Algebra*, **56** (2008) 143–162.

[4] M. Fiedler, Geometry of the numerical range of matrices, *Linear Algebra Appl.*, **37** (1981) 81–96.

[5] M. Fiedler, Numerical range of matrices and Levinger's theorem, *Linear Algebra Appl.*, **220** (1995) 171–180.

[6] R. Goldman, Curvature formulas for implicit curves and surfaces, *Comput. Aided Geom. Des.*, **22** (2005) 632–658.

[7] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[8] S. Kirkland, P. Psarrakos, and M. Tsatsomeros, On the location of the spectrum of hypertournament matrices, *Linear Algebra Appl.*, **323** (2001) 37–49.

[9] P. Lancaster and P. Psarrakos, Normal and seminormal eigenvalues of matrix functions, *Integral Equations Operator Theory*, **41** (2001) 331–342.

[10] C.K. Li and N.S. Sze, Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations, *Proc. Amer. Math. Soc.*, **136** (2008) 3013–3023.

[11] J. McDonald, P. Psarrakos, and M. Tsatsomeros, Almost skew-symmetric matrices, *Rocky Mountain J. Math.*, **34** (1) (2004) 269–288.

[12] L.M. Milne-Thomson, Elliptic integrals, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (M. Abramowitz and I.A. Stegun, eds.), Dover, New York, 1972, 587–626.

[13] P.J. Psarrakos and M.J. Tsatsomeros, An envelope for the spectrum of a matrix, *Cent. Eur. J. Math.*, **10** (1) (2012) 292–302.

[14] P.J. Psarrakos and M.J. Tsatsomeros, On the geometry of the envelope of a matrix, *Appl. Math. Comput.*, **244** (2014) 132–141.