Birkhoff-James approximate orthogonality sets and numerical ranges

Christos Chorianopoulos* and Panayiotis Psarrakos*

December 7, 2010

Abstract

In this paper, the notion of Birkhoff-James approximate orthogonality sets is introduced for rectangular matrices and matrix polynomials. The proposed definition yields a natural generalization of standard numerical range and q-numerical range (and also of recent extensions), sharing with them several geometric properties.

Keywords: Birkhoff-James orthogonality, ϵ-orthogonality, rectangular matrix, numerical range, matrix polynomial, eigenvalue, boundary.

AMS Subject Classifications: 15A60, 47A12.

1 Introduction

The *numerical range* (also known as the *field of values*) of a square matrix $A \in \mathbb{C}^{n\times n}$ is the compact and convex set $F(A) = \{x^*Ax \in \mathbb{C} : x \in \mathbb{C}^n,\, x^*x = 1\}$. The compactness follows readily from the fact that $F(A)$ is the image of the compact unit sphere of $\mathbb{C}^n$ under the continuous mapping $x \longmapsto x^*Ax$, and the convexity of $F(A)$ is the celebrated Hausdorff-Toeplitz Theorem [12, 29]. The numerical range has been studied extensively for many decades, and it is useful in studying and understanding matrices and operators (see [3, 4, 11, 13] and the references therein).

Stampfli and Williams [28, Theorem 4], and later Bonsall and Duncan [4, Lemma 6.22.1], observed that the numerical range of a matrix $A \in \mathbb{C}^{n\times n}$ can be written

$$\begin{aligned} F(A) &= \{\mu \in \mathbb{C} : \|A - \lambda I_n\|_2 \geq |\mu - \lambda|,\, \forall\, \lambda \in \mathbb{C}\} \\ &= \bigcap_{\lambda\in\mathbb{C}} \{\mu \in \mathbb{C} : |\mu - \lambda| \leq \|A - \lambda I_n\|_2\}, \end{aligned}$$

where $\|\cdot\|_2$ denotes the *spectral matrix norm* (i.e., that norm subordinate to the euclidean vector norm) and I_n is the $n\times n$ identity matrix. Hence, $F(A)$ is an infinite intersection of closed (circular) disks $\mathcal{D}(\lambda, \|A - \lambda I_n\|_2) = \{\mu \in \mathbb{C} : |\mu - \lambda| \leq \|A - \lambda I_n\|_2\}$ ($\lambda \in \mathbb{C}$). In this way, it is confirmed once again that $F(A)$ is a compact and convex subset of the complex plane that lies in the closed disk $\mathcal{D}(0, \|A\|_2)$.

*Department of Mathematics, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece (horjoe@yahoo.gr, Ch. Chorianopoulos; ppsarr@math.ntua.gr, P. Psarrakos).

Inspired by the above intersection property, Chorianopoulos, Karanasios and Psarrakos [8] recently introduced a definition of numerical range for rectangular complex matrices. In particular, for any $A, B \in \mathbb{C}^{n\times m}$ with $B \neq 0$, and any matrix norm $\|\cdot\|$, the *numerical range of* A *with respect to* B is defined as

$$\begin{aligned} F_{\|\cdot\|}(A;B) &= \{\mu \in \mathbb{C} : \|A-\lambda B\| \geq |\mu-\lambda|, \forall\, \lambda \in \mathbb{C}\} \qquad (1)\\ &= \bigcap_{\lambda\in\mathbb{C}} \mathcal{D}\left(\lambda, \|A-\lambda B\|\right). \qquad (2)\end{aligned}$$

This set is obviously compact and convex, and satisfies basic properties of the standard numerical range [8]. Moreover, it is nonempty if and only if $\|B\| \geq 1$ [8, Corollary 4].

For a $q \in [0,1]$, the q-numerical range of a square matrix $A \in \mathbb{C}^{n\times n}$ is defined as the compact and convex set $F(A;q) = \{y^*Ax \in \mathbb{C} : x, y \in \mathbb{C}^n,\ x^*x = y^*y = 1,\ y^*x = q\}$. This range was introduced in [22] as a generalization of the standard numerical range $F(A)$ (it is clear that $F(A;1) = F(A)$), and has been systematically investigated in the last two decades [5, 6, 19, 20]. In [1], Aretaki and Maroulas, motivated by the definition of $F_{\|\cdot\|}(A;B)$ in (1) and (2), introduced a definition for the q-numerical range of rectangular complex matrices. Namely, for any $A, B \in \mathbb{C}^{n\times m}$ with $B \neq 0$, any $q \in [0,1]$, and any matrix norm $\|\cdot\|$, they define the *q-numerical range of* A *with respect to* B as the compact and convex set

$$\begin{aligned} F_{\|\cdot\|}(A;B;q) &= \{\mu \in \mathbb{C} : \|A-\lambda B\| \geq |\mu-q\lambda|, \forall\, \lambda \in \mathbb{C}\} \qquad (3)\\ &= \bigcap_{\lambda\in\mathbb{C}} \mathcal{D}\left(q\lambda, \|A-\lambda B\|\right).\end{aligned}$$

They have also obtained that [1]

$$\frac{1}{q_2} F_{\|\cdot\|}(A;B;q_2) \subseteq \frac{1}{q_1} F_{\|\cdot\|}(A;B;q_1)\,; \quad 0 < q_1 < q_2 \leq 1 \qquad (4)$$

(generalizing Theorem 2.5 of [19]), and

$$F_{\|\cdot\|}(A;B;q) = F_{\|\cdot\|}(A;q^{-1}B)\,; \quad 0 < q \leq 1. \qquad (5)$$

By the latter relation and [8, Corollary 4], it follows that $F_{\|\cdot\|}(A;B;q)$ is nonempty if and only if $\|B\| \geq q$ $(0 < q \leq 1)$. Furthermore, it is immediate that for $q = 0$, $F_{\|\cdot\|}(A;B;0) = \mathcal{D}\left(0, \inf_{\lambda\in\mathbb{C}} \|A-\lambda B\|\right)$, extending Proposition 2.11 of [19].

For $n = m$, $\|\cdot\| = \|\cdot\|_2$ and $B = I_n$, we have that $F_{\|\cdot\|_2}(A;I_n) = F(A)$ and $F_{\|\cdot\|_2}(A;I_n;q) = F(A;q)$ $(0 \leq q \leq 1)$ [1, 4, 28], i.e., the ranges $F_{\|\cdot\|}(A;B)$ and $F_{\|\cdot\|}(A;B;q)$ are direct generalizations of the numerical range $F(A)$ and the q-numerical range $F(A;q)$, respectively.

In this article, we introduce a new range of values for rectangular matrices and matrix polynomials, which is based on the notion of Birkhoff-James approximate orthogonality and generalizes the numerical ranges $F_{\|\cdot\|}(A;B)$ and $F_{\|\cdot\|}(A;B;q)$. We also show that it is quite rich in structure by establishing some of its main properties. In the next section, we give the definition together with basic properties of this set for rectangular matrices, and in Section 3, we study the case of matrix polynomials. In Section 4, we obtain necessary and/or sufficient conditions for the boundary points, and finally, in Section 5, we investigate the case of matrix norms induced by inner products of matrices. Simple illustrative examples are also given to verify our results.

2 Approximate orthogonality sets of matrices

The analysis in [8] is based on the properties of matrix norms and the Birkhoff-James orthogonality [2, 14]; namely, for two elements χ and ψ of a complex normed linear space $(\mathcal{X}, \|\cdot\|)$, χ is called *Birkhoff-James orthogonal* to ψ, denoted by $\chi \perp_{BJ} \psi$, if $\|\chi + \lambda\psi\| \geq \|\chi\|$ for all $\lambda \in \mathbb{C}$. This orthogonality is neither symmetric nor additive [14]. However, it is homogeneous, i.e., $\chi \perp_{BJ} \psi$ if and only if $a\chi \perp_{BJ} b\psi$ for any nonzero $a, b \in \mathbb{C}$.

Furthermore, for any $\epsilon \in [0, 1)$, we say[1] that χ is *Birkhoff-James ϵ-orthogonal* to ψ, denoted by $\chi \perp_{BJ}^{\epsilon} \psi$, if $\|\chi + \lambda\psi\| \geq \sqrt{1-\epsilon^2}\,\|\chi\|$ for all $\lambda \in \mathbb{C}$. It is straightforward to see that this relation is also homogeneous. In an inner product space $(\mathcal{X}, \langle\cdot,\cdot\rangle)$, with the standard orthogonality relation $\perp$, a $\chi \in \mathcal{X}$ is called *ϵ-orthogonal* to a $\psi \in \mathcal{X}$, denoted by $\chi \perp^{\epsilon} \psi$, if $|\langle\chi, \psi\rangle| \leq \epsilon\,\|\chi\|\,\|\psi\|$. Moreover, by [7, 9], $\chi \perp \psi$ if and only if $\chi \perp_{BJ} \psi$, and $\chi \perp^{\epsilon} \psi$ if and only if $\chi \perp_{BJ}^{\epsilon} \psi$.

For any $A, B \in \mathbb{C}^{n\times m}$ with $\|B\| \geq 1$, using the Birkhoff-James ϵ-orthogonality for $\epsilon = \epsilon_B = \sqrt{\|B\|^2 - 1}\,/\,\|B\|$, one can verify that (see also Theorem 1 in [8])

$$\begin{aligned} F_{\|\cdot\|}(A;B) &= \{\mu \in \mathbb{C} :\ \|A - (\mu - \lambda)B\| \geq |\lambda|,\ \forall\,\lambda \in \mathbb{C}\} \\ &= \left\{\mu \in \mathbb{C} :\ \left\|\frac{1}{\lambda}(A - \mu B) + B\right\| \geq 1,\ \forall\,\lambda \in \mathbb{C}\backslash\{0\}\right\} \\ &= \left\{\mu \in \mathbb{C} :\ \|\lambda(A - \mu B) + B\| \geq \sqrt{1-\epsilon_B^2}\,\|B\|,\ \forall\,\lambda \in \mathbb{C}\right\} \\ &= \left\{\mu \in \mathbb{C} :\ B \perp_{BJ}^{\epsilon_B} (A - \mu B)\right\}. \end{aligned}$$

In particular, if $||B\| = 1$, then $F_{\|\cdot\|}(A;B) = \{\mu \in \mathbb{C} :\ B \perp_{BJ} (A - \mu B)\}$.

By the above discussion, the next definition arises in a natural way.

Definition 1. For any $A, B \in \mathbb{C}^{n\times m}$ with $B \neq 0$, any matrix norm $\|\cdot\|$, and any $\epsilon \in [0, 1)$, the *Birkhoff-James ϵ-orthogonality set of A with respect to B* is defined and denoted by

$$\begin{aligned} F_{\|\cdot\|}^{\epsilon}(A;B) &= \{\mu \in \mathbb{C} :\ B \perp_{BJ}^{\epsilon} (A - \mu B)\} \\ &= \left\{\mu \in \mathbb{C} :\ \|A - \lambda B\| \geq \sqrt{1-\epsilon^2}\,\|B\|\,|\mu - \lambda|,\ \forall\,\lambda \in \mathbb{C}\right\} \\ &= \bigcap_{\lambda\in\mathbb{C}} \mathcal{D}\left(\lambda, \frac{\|A - \lambda B\|}{\sqrt{1-\epsilon^2}\,\|B\|}\right). \end{aligned}$$

Apparently, the Birkhoff-James ϵ-orthogonality set $F_{\|\cdot\|}^{\epsilon}(A;B)$ is a compact and convex subset of the complex plane that lies in the closed disk $\mathcal{D}(0, \|A\|/(\sqrt{1-\epsilon^2}\,\|B\|))$. By Lemma 3 in [8] (see also [14, Corollary 2.2]), $F_{\|\cdot\|}^{0}(A;B)$ is nonempty. Moreover, for any $\epsilon_1, \epsilon_2 \in [0, 1)$ with $\epsilon_1 < \epsilon_2$, Definition 1 yields $F_{\|\cdot\|}^{\epsilon_1}(A;B) \subseteq F_{\|\cdot\|}^{\epsilon_2}(A;B)$, generalizing (4) (in combination with Theorem 5 below). Hence, the Birkhoff-James ϵ-orthogonality set $F_{\|\cdot\|}^{\epsilon}(A;B)$ is always nonempty.

[1] With regard to the Birkhoff-James orthogonality, various definitions for approximate orthogonality are proposed in [7, 9]. The definition given here seems to be the most suitable to the concept of this work.

The points of the Birkhoff-James ϵ-orthogonality set $F^{\epsilon}_{\|\cdot\|}(A;B)$ have a remarkable geometric interpretation. In particular, a scalar $\mu \in \mathbb{C}$ lies in $F^{\epsilon}_{\|\cdot\|}(A;B)$ if and only if $B \perp^{\epsilon}_{BJ} (A-\mu B)$, or equivalently, if and only if $\|B+\lambda(A-\mu B)\| \geq \sqrt{1-\epsilon^2}\,\|B\|$ for all $\lambda \in \mathbb{C}$. This means that $\mu \in F^{\epsilon}_{\|\cdot\|}(A;B)$ if and only if the one-complex-dimensional affine space $\{B+\lambda(A-\mu B) : \lambda \in \mathbb{C}\}$ does not intersect the open ball $\mathcal{B}^{\circ}(0,\sqrt{1-\epsilon^2}\,\|B\|) = \{M \in \mathbb{C}^{n\times m} : \|M\| < \sqrt{1-\epsilon^2}\,\|B\|\}$, as illustrated in Figure 1.

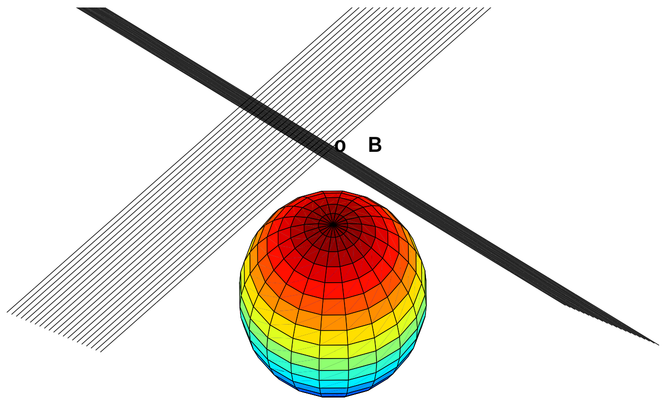

Figure 1: Two affine spaces that contain B and do not intersect $\mathcal{B}^{\circ}(0,\sqrt{1-\epsilon^2}\,\|B\|)$.

We remark that in the sequel, the zero matrix is always considered as a scalar multiple of B.

Proposition 2. *Let $A,B \in \mathbb{C}^{n\times m}$ with $B \neq 0$, and $0 \leq \epsilon_1 < \epsilon_2 < 1$. If the matrix A is not a scalar multiple of B, then $F^{\epsilon_1}_{\|\cdot\|}(A;B) \subset F^{\epsilon_2}_{\|\cdot\|}(A;B)$, and in particular, $F^{\epsilon_1}_{\|\cdot\|}(A;B)$ lies in the interior of $F^{\epsilon_2}_{\|\cdot\|}(A;B)$.*

Proof. For any $\mu \in F^{\epsilon_1}_{\|\cdot\|}(A;B)$, we have

$$\|A-\mu B+(\mu-\lambda)B\| \geq \sqrt{1-\epsilon_1^2}\,\|B\|\,|\mu-\lambda|, \ \ \forall\, \lambda \in \mathbb{C},$$

or

$$\|\lambda B+A-\mu B\| \geq \sqrt{1-\epsilon_1^2}\,\|B\|\,|\lambda| > \sqrt{1-\epsilon_2^2}\,\|B\|\,|\lambda|, \ \ \forall\, \lambda \in \mathbb{C}\backslash\{0\}.$$

Since matrix A is not a scalar multiple of B, there exists a real $\delta > 0$ such that

$$\delta \leq \min\left\{\min_{|\lambda|\leq 1}\left\{\|\lambda B+A-\mu B\| - \sqrt{1-\epsilon_2^2}\,\|B\|\,|\lambda|\right\}, \left(\sqrt{1-\epsilon_1^2}-\sqrt{1-\epsilon_2^2}\right)\|B\|\right\}.$$

As a consequence,

$$\delta \leq \inf_{\lambda\in\mathbb{C}}\left\{\|\lambda B+A-\mu B\| - \sqrt{1-\epsilon_2^2}\,\|B\|\,|\lambda|\right\},$$

and for any $\xi \in \mathcal{D}(0,\delta/\|B\|)$,

$$\|\lambda B+A-(\mu+\xi)B\| \geq \|\lambda B+A-\mu B\| - \|\xi B\| \geq \sqrt{1-\epsilon_2^2}\,\|B\|\,|\lambda|, \ \ \forall\, \lambda \in \mathbb{C}.$$

Hence, every point of the compact set $F^{\epsilon_1}_{\|\cdot\|}(A;B)$ lies in the interior of $F^{\epsilon_2}_{\|\cdot\|}(A;B)$. $\square$

Corollary 3. *Suppose $A, B \in \mathbb{C}^{n\times m}$ such that $B \neq 0$ and A is not a scalar multiple of B. Then for every $\epsilon \in (0,1)$, the ϵ-orthogonality set $F^{\epsilon}_{\|\cdot\|}(A;B)$ has a nonempty interior, and it cannot be degenerated to a singleton or a line segment.*

If we allow the value 1 for the parameter ϵ, then the first two equalities in Definition 1 yield $F^{1}_{\|\cdot\|}(A;B) = \mathbb{C}$. Moreover, if A is not a scalar multiple of B, then the ϵ-orthogonality set $F^{\epsilon}_{\|\cdot\|}(A;B)$ can be arbitrarily large for ϵ sufficiently close to 1.

Proposition 4. *Suppose $A, B \in \mathbb{C}^{n\times m}$ such that $B \neq 0$ and A is not a scalar multiple of B. Then for any bounded region $\Omega \subset \mathbb{C}$, there is an $\epsilon_\Omega \in [0,1)$ such that $\Omega \subseteq F^{\epsilon_\Omega}_{\|\cdot\|}(A;B)$.*

Proof. Without loss of generality, we can assume that Ω is compact. Consider a $\mu \in \Omega$, such that $\mu \notin F^{\epsilon}_{\|\cdot\|}(A;B)$ for every $\epsilon \in [0,1)$. Then for every $\epsilon_k = \sqrt{1-\frac{1}{k^2}}$, $k = 2, 3, \dots$, there exists a $\lambda_k \in \mathbb{C}$ such that

$$\|A - (\mu - \lambda_k)B\| < \sqrt{1 - \left(\sqrt{1-\frac{1}{k^2}}\right)^2}\, \|B\|\, |\lambda_k|,$$

or

$$\|\lambda_k B + A - \mu B\| < \frac{1}{k}\, \|B\|\, |\lambda_k|, \tag{6}$$

or

$$\big|\, \|\lambda_k B\| - \|A - \mu B\|\, \big| < \frac{1}{k}\, \|B\|\, |\lambda_k|,$$

or

$$|\lambda_k|\, \|B\| \left(1 - \frac{1}{k}\right) < \|A - \mu B\| \leq \|A\| + |\mu|\, \|B\|,$$

or

$$|\lambda_k| < \frac{\|A\| + |\mu|\, \|B\|}{\|B\|\,\left(1-\frac{1}{k}\right)} \leq 2\,\frac{\|A\| + |\mu|\, \|B\|}{\|B\|}.$$

Thus, the sequence $\{\lambda_k\}_{k\in\mathbb{N}}$ is always bounded, and hence, it has a converging subsequence $\{\lambda_{k_t}\}_{t\in\mathbb{N}}$. If we assume that $\lambda_{k_t} \to \lambda_0$, then by (6),

$$\|\lambda_{k_t} B + A - \mu B\| < \frac{1}{k_t}\, \|B\|\, |\lambda_{k_t}|, \quad \forall\, t \in \mathbb{N},$$

or

$$\lim_t \|\lambda_{k_t} B + A - \mu B\| = 0,$$

or

$$\|\lambda_0 B + A - \mu B\| = 0,$$

where the latter relation is a contradiction since A is not a scalar multiple of B. As a consequence, there is an $\epsilon_\mu \in [0,1)$ such that $\mu \in F^{\epsilon_\mu}_{\|\cdot\|}(A;B)$. Without loss of generality, we may assume that every $\mu \in \Omega$ lies in the interior of $F^{\epsilon_\mu}_{\|\cdot\|}(A;B)$ (choosing, if necessary, a larger ϵ_μ). Hence, $\Omega \subseteq \bigcup_{\mu\in\Omega} \mathrm{Int}[F^{\epsilon_\mu}_{\|\cdot\|}(A;B)]$, where $\mathrm{Int}[\,\cdot\,]$ denotes the interior of a set. Since Ω is compact, there is a finite number of points $\mu_1, \mu_2, \dots, \mu_s \in \Omega$ such that $\Omega \subseteq \bigcup_{i=1}^{s} \mathrm{Int}[F^{\epsilon_{\mu_i}}_{\|\cdot\|}(A;B)]$. Setting $\epsilon_\Omega = \max\{\epsilon_{\mu_i} : i = 1, 2, \dots, s\}$, Proposition 2 completes the proof. ☐

As mentioned above, for $\|B\| \geq 1$ and $\epsilon_B = \sqrt{\|B\|^2 - 1} \,/\, \|B\|$, the ϵ-orthogonality set $F^{\epsilon_B}_{\|\cdot\|}(A;B)$ coincides with the numerical range $F_{\|\cdot\|}(A;B)$. It is also easy to see that

$$\begin{aligned} F^{\epsilon}_{\|\cdot\|}(A;B) &= \left\{ \mu \in \mathbb{C} : \, \|A - \lambda B\| \geq \sqrt{1-\epsilon^2}\, \|B\|\, |\mu - \lambda|, \, \forall\, \lambda \in \mathbb{C} \right\} \\ &= \left\{ \frac{\mu}{\sqrt{1-\epsilon^2}\, \|B\|} \in \mathbb{C} : \, \|A - \lambda B\| \geq \left| \mu - \sqrt{1-\epsilon^2}\, \|B\| \lambda \right|, \, \forall\, \lambda \in \mathbb{C} \right\} \\ &= \frac{1}{\sqrt{1-\epsilon^2}\, \|B\|} \left\{ \mu \in \mathbb{C} : \, \|A - \lambda B\| \geq \left| \mu - \sqrt{1-\epsilon^2}\, \|B\| \lambda \right|, \, \forall\, \lambda \in \mathbb{C} \right\}. \end{aligned}$$

Thus, keeping in mind [8, Proposition 8] and (5), we have the following results.

Theorem 5. *For any $A, B \in \mathbb{C}^{n\times m}$ with $B \neq 0$, and $\epsilon \in [0,1)$, it holds that*

$$F^{\epsilon}_{\|\cdot\|}(A;B) = F_{\|\cdot\|}(q_\epsilon^{-1}A; B; q_\epsilon) = F_{\|\cdot\|}(q_\epsilon^{-1}A; q_\epsilon^{-1}B),$$

where $q_\epsilon = \sqrt{1-\epsilon^2}\, \|B\|$. Equivalently, for any $A, B \in \mathbb{C}^{n\times m}$ and $q \in (0,1]$, with $\|B\| \geq q$,

$$F^{\epsilon_q}_{\|\cdot\|}(A;B) = F_{\|\cdot\|}(q^{-1}A; B; q) = F_{\|\cdot\|}(q^{-1}A; q^{-1}B),$$

where $\epsilon_q = \sqrt{\|B\|^2 - q^2} \,/\, \|B\|$.

Corollary 6. *For any $A, B \in \mathbb{C}^{n\times m}$ and $q \in (0,1]$, with $\|B\| = q$, it holds that*

$$F_{\|\cdot\|}(A;B;q) = F^{0}_{\|\cdot\|}(A;B) = \{\mu \in \mathbb{C} : \, B \perp_{BJ} (A - \mu B)\}.$$

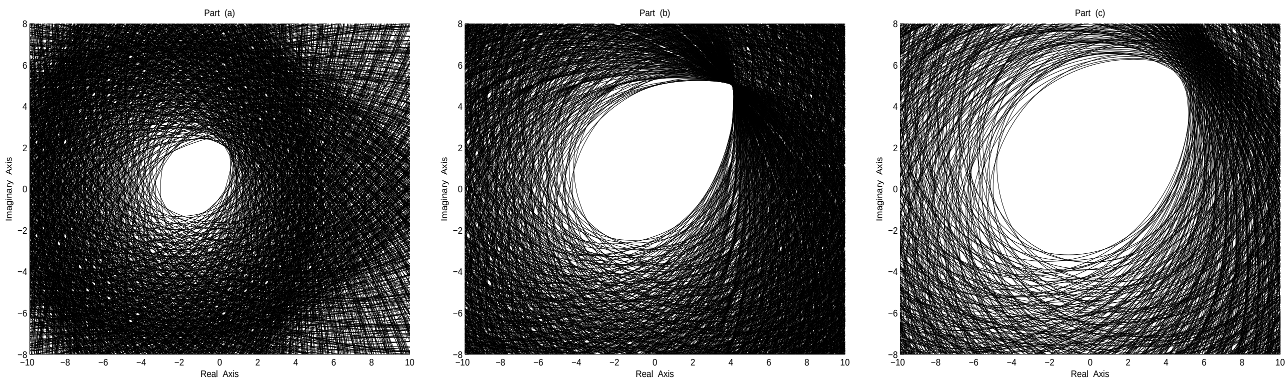

Figure 2: The sets $F^{0.5}_{\|\cdot\|_2}(A;B)$, $F^{\sqrt{0.5}}_{\|\cdot\|_2}(A;B)$ and $F^{\sqrt{0.6}}_{\|\cdot\|_2}(A;B)$.

As an example, we consider the 3×4 complex matrices

$$A = \begin{bmatrix} 4 + \mathrm{i}\,5 & 0 & \mathrm{i} & 0 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & 0 & -\mathrm{i}\,2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & -\mathrm{i} & 0 \end{bmatrix}$$

with $\|B\|_2 = \sqrt{2}$. The approximate orthogonality sets $F^{0.5}_{\|\cdot\|_2}(A;B)$, $F^{\sqrt{0.5}}_{\|\cdot\|_2}(A;B)$ and $F^{\sqrt{0.6}}_{\|\cdot\|_2}(A;B)$ are illustrated (as the unshaded regions) in parts (a), (b) and (c) of Figure

2, respectively, confirming Proposition 2. Note also that $\sqrt{0.5} = \sqrt{\|B\|_2^2 - 1}\,/\,\|B\|_2$ and $\sqrt{0.6} = \sqrt{\|B\|_2^2 - \sqrt{0.8}^{\,2}}\,/\,\|B\|_2$, and thus, Theorem 5 yields $F^{\sqrt{0.5}}_{\|\cdot\|_2}(A;B) = F_{\|\cdot\|_2}(A;B)$ and $F^{\sqrt{0.6}}_{\|\cdot\|_2}(A;B) = F_{\|\cdot\|_2}(\sqrt{0.8}^{\,-1}A;B;\sqrt{0.8}) = F_{\|\cdot\|_2}(\sqrt{0.8}^{\,-1}A;\sqrt{0.8}^{\,-1}B)$.

By the above discussion and Theorem 5, it is apparent that the Birkhoff-James ϵ-orthogonality set $F^{\epsilon}_{\|\cdot\|}(A;B)$ is a generalization[2] of the numerical ranges $F_{\|\cdot\|}(A;B)$ and $F_{\|\cdot\|}(A;B;q)$, in the sense that it does not require any condition for the norm of matrix $B \neq 0$ and it coincides with $F_{\|\cdot\|}(A;B)$ and $F_{\|\cdot\|}(A;B;q)$ for certain values of ϵ. Furthermore, basic properties of the numerical range $F_{\|\cdot\|}(A;B)$ obtained in [8] are extended readily to $F^{\epsilon}_{\|\cdot\|}(A;B)$.

(P$_1$) If $A = bB$ for some $b \in \mathbb{C}$, then $F^{\epsilon}_{\|\cdot\|}(bB;B) = \{b\}$. The converse is not true in general; for example, if the matrix norm $\|\cdot\|$ is induced by an inner product of matrices, then $F^{0}_{\|\cdot\|}(A;B)$ is always a singleton (see Property (P$_7$) below).

(P$_2$) For any scalars $a, b \in \mathbb{C}$, it holds that $F^{\epsilon}_{\|\cdot\|}(aA + bB;B) = aF^{\epsilon}_{\|\cdot\|}(A;B) + b$.

(P$_3$) Suppose the matrix norm $\|\cdot\|$ is induced by a vector norm (acting on $\mathbb{C}^n$ and $\mathbb{C}^m$) and $n \geq m$, and let $\mu_0 \in \mathbb{C}$ be an *eigenvalue of* A *with respect to* B, with an associate unit *eigenvector* $x_0 \in \mathbb{C}^m$; that is, $(A - \mu_0 B)x_0 = 0$. Then for every $\epsilon \in \left[\sqrt{\|B\|^2 - \|Bx_0\|^2}\,/\,\|B\|, 1\right)$, μ_0 lies in $F^{\epsilon}_{\|\cdot\|}(A;B)$. (In combination with Theorem 5, this property is a direct generalization of Theorem 2.7 in [19].)

(P$_4$) If $A \neq 0$, then $\left\{\mu^{-1} \in \mathbb{C} : \mu \in F^{\epsilon}_{\|\cdot\|}(A;B),\ |\mu| \geq \|A\|/\|B\|\right\} \subseteq F^{\epsilon}_{\|\cdot\|}(B;A)$.

(P$_5$) $\mathrm{Int}[F^{\epsilon}_{\|\cdot\|}(A;B)] \subseteq \left\{\mu \in \mathbb{C} : \|A - \lambda B\| > \sqrt{1-\epsilon^2}\,\|B\|\,|\mu - \lambda|,\ \forall\, \lambda \in \mathbb{C}\right\}$.

(P$_6$) Suppose that $f : (\mathbb{C}^{n_1 \times m_1}, \|\cdot\|) \to (\mathbb{C}^{n_2 \times m_2}, |||\cdot|||)$ is a linear map such that $|||f(M)||| = (\,\geq\,,\,\leq\,)\ \|M\|$ for every $M \in \mathbb{C}^{n_1 \times m_1}$. Then for any $A, B \in \mathbb{C}^{n_1 \times m_1}$, $F^{\epsilon}_{|||\cdot|||}(f(A);f(B)) = (\,\supseteq\,,\,\subseteq\,)\ F^{\epsilon}_{\|\cdot\|}(A;B)$.

(P$_7$) If the matrix norm $\|\cdot\|$ is induced by the inner product of matrices $\langle\cdot,\cdot\rangle$ (this is the case of the *Frobenius norm* $\|\cdot\|_F$), then

$$F^{\epsilon}_{\|\cdot\|}(A;B) = \mathcal{D}\left(\frac{\langle A,B\rangle}{\|B\|^2}\,,\ \left\|A - \frac{\langle A,B\rangle}{\|B\|^2}\,B\right\|\frac{\epsilon}{\sqrt{1-\epsilon^2}\,\|B\|}\right).$$

If the matrix norm $\|\cdot\|$ is induced by a vector norm, then by Property (P$_3$) (see also [8, Proposition 17]), an eigenvalue μ_0 of A with respect to B lies in $F^{\epsilon}_{\|\cdot\|}(A;B)$ if there is an associated unit eigenvector x_0 (i.e., $(A - B\mu_0)x_0 = 0$) such that $\|Bx_0\| \geq \sqrt{1-\epsilon^2}\,\|B\|$. As a consequence, if the matrices A and B are square, say $n \times n$, and B is invertible with $\|B^{-1}\|^{-1} \geq \sqrt{1-\epsilon^2}\,\|B\|$, then all the eigenvalues of A with respect to B lie in $F^{\epsilon}_{\|\cdot\|}(A;B)$. In this case, we have the following result on the *generalized resolvent* $(A - zB)^{-1}$, $z \in \mathbb{C}$ (for the standard resolvent of operators, see [15, Theorem V.3.2] and [28, Lemma 1]).

[2] We remark that the definitions of $F_{\|\cdot\|}(A;B)$, $F_{\|\cdot\|}(A;B;q)$ and $F^{\epsilon}_{\|\cdot\|}(A;B)$ are applicable to the elements of any normed linear space.

Proposition 7. *Suppose the matrix norm $\|\cdot\|$ is induced by a vector norm. Let $\epsilon \in [0,1)$, and let A, B be two $n \times n$ matrices with B invertible and $\|B^{-1}\|^{-1} \geq \sqrt{1-\epsilon^2}\,\|B\|$. Then for any point $\xi \notin F^{\epsilon}_{\|\cdot\|}(A;B)$, the distance $d(\xi, F^{\epsilon}_{\|\cdot\|}(A;B))$ from ξ to $F^{\epsilon}_{\|\cdot\|}(A;B)$ satisfies*

$$d(\xi, F^{\epsilon}_{\|\cdot\|}(A;B)) \leq \frac{1}{\sqrt{1-\epsilon^2}\,\|B\|\,\|(A-\xi B)^{-1}\|}.$$

Proof. For any $\mu \in F_{\|\cdot\|}(AB^{-1}; I_n)$,

$$|\mu - \lambda| \leq \|AB^{-1} - \lambda I_n\| \leq \|(A-\lambda B)\|\,\|B^{-1}\| \leq \frac{\|A-\lambda B\|}{\sqrt{1-\epsilon^2}\,\|B\|}, \quad \forall\, \lambda \in \mathbb{C},$$

and thus, $F_{\|\cdot\|}(AB^{-1}; I_n) \subseteq F^{\epsilon}_{\|\cdot\|}(A;B)$. By [18], for any convex set V that contains $F_{\|\cdot\|}(AB^{-1}; I_n)$ and any $\xi \notin V$, we have $d(\xi, V) \leq \|(AB^{-1} - \xi I_n)^{-1}\|^{-1}$. Setting $V = F^{\epsilon}_{\|\cdot\|}(A;B)$ yields

$$d(\xi, F^{\epsilon}_{\|\cdot\|}(A;B)) \leq \frac{1}{\|(AB^{-1} - \xi I_n)^{-1}\|} \leq \frac{(\sqrt{1-\epsilon^2}\,\|B\|)^{-1}}{\|B^{-1}\|\,\|(AB^{-1} - \xi I_n)^{-1}\|},$$

and the proof is complete. □

3 Approximate orthogonality sets of matrix polynomials

Consider an $n \times m$ matrix polynomial

$$P(z) \;=\; A_l z^l + A_{l-1} z^{l-1} + \cdots + A_1 z + A_0, \tag{7}$$

where z is a complex variable and $A_j \in \mathbb{C}^{n\times m}$ $(j = 0, 1, \ldots, l)$ with $A_l \neq 0$. The study of matrix polynomials has a long history, especially with regard to their applications on higher order linear systems of differential equations (see [10, 17, 23] and the references therein).

If $n \geq m$, then a scalar $\mu_0 \in \mathbb{C}$ is said to be an *eigenvalue of* $P(z)$ in (7) if $P(\mu_0)x_0 = 0$ for some nonzero vector $x_0 \in \mathbb{C}^m$. This vector x_0 is called an *eigenvector of* $P(z)$ corresponding to μ_0.

For an $n \times n$ matrix polynomial $P(z)$, the *(standard) numerical range of* $P(z)$ is defined as

$$\begin{aligned} W(P(z)) &= \{\mu \in \mathbb{C} :\, x^* P(\mu) x = 0,\, x \in \mathbb{C}^n,\, x \neq 0\} \\ &= \{\mu \in \mathbb{C} :\, 0 \in F(P(\mu))\}. \end{aligned} \tag{8}$$

This range and its properties have been studied extensively in [21, 24, 25, 26, 27].

Motivated by (8), and recalling (1), (3) and Definition 1, for an $n \times m$ matrix polynomial $P(z)$ as in (7), any nonzero matrix $B \in \mathbb{C}^{n\times m}$, and any matrix norm $\|\cdot\|$, we define the *numerical range of* $P(z)$ *with respect to* B ($\|B\| \geq 1$)

$$\begin{aligned} W_{\|\cdot\|}(P(z);B) &= \left\{\mu \in \mathbb{C} :\, 0 \in F_{\|\cdot\|}(P(\mu);B)\right\} \qquad (9) \\ &= \{\mu \in \mathbb{C} :\, \|P(\mu) - \lambda B\| \geq |\lambda|,\, \forall\, \lambda \in \mathbb{C}\} \\ &= \left\{\mu \in \mathbb{C} :\, B \perp^{\epsilon_B}_{BJ} P(\mu),\, \epsilon_B = \frac{\sqrt{\|B\|^2 - 1}}{\|B\|}\right\}, \end{aligned}$$

the *q-numerical range of* $P(z)$ *with respect to* B $(0 \le q \le 1,\ \|B\| \ge q)$

$$\begin{aligned} W_{\|\cdot\|}(P(z);B;q) &= \left\{\mu \in \mathbb{C} :\ 0 \in F_{\|\cdot\|}(P(\mu);B;q)\right\} \\ &= \{\mu \in \mathbb{C} :\ \|P(\mu) - \lambda B\| \ge q|\lambda|,\ \forall\, \lambda \in \mathbb{C}\} \\ &= \left\{\mu \in \mathbb{C} :\ B \perp_{BJ}^{\epsilon_q} P(\mu),\ \epsilon_q = \frac{\sqrt{\|B\|^2 - q^2}}{\|B\|}\right\}, \end{aligned} \tag{10}$$

and the *Birkhoff-James* ϵ*-orthogonality set of* $P(z)$ *with respect to* B $(0 \le \epsilon < 1)$

$$\begin{aligned} W_{\|\cdot\|}^{\epsilon}(P(z);B) &= \left\{\mu \in \mathbb{C} :\ 0 \in F_{\|\cdot\|}^{\epsilon}(P(\mu);B)\right\} \qquad (11) \\ &= \left\{\mu \in \mathbb{C} :\ \|P(\mu) - \lambda B\| \ge \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|,\ \forall\, \lambda \in \mathbb{C}\right\} \\ &= \left\{\mu \in \mathbb{C} :\ B \perp_{BJ}^{\epsilon} P(\mu)\right\}. \qquad (12) \end{aligned}$$

The closeness of these sets follows from the continuity of matrix norms, and for $q = 0$, $W_{\|\cdot\|}(P(z);B;0) = \mathbb{C}$. Furthermore, by definitions (9), (10) and (11), Theorem 5 and Corollary 6 are extended readily to the case of matrix polynomials.

Theorem 8. *Let* $P(z)$ *be an* $n \times m$ *matrix polynomial as in (7). For any nonzero* $B \in \mathbb{C}^{n\times m}$ *and* $\epsilon \in [0,1)$,

$$W_{\|\cdot\|}^{\epsilon}(P(z);B) = W_{\|\cdot\|}(P(z);B;q_\epsilon) = W_{\|\cdot\|}(P(z);q_\epsilon^{-1}B),$$

where $q_\epsilon = \sqrt{1-\epsilon^2}\,\|B\|$. *Equivalently, for any* $B \in \mathbb{C}^{n\times m}$ *and* $q \in (0,1]$, *with* $\|B\| \ge q$,

$$W_{\|\cdot\|}^{\epsilon_q}(P(z);B) = W_{\|\cdot\|}(P(z);B;q) = W_{\|\cdot\|}(P(z);q^{-1}B),$$

where $\epsilon_q = \sqrt{\|B\|^2 - q^2}\,/\,\|B\|$.

Corollary 9. *For any* $B \in \mathbb{C}^{n\times m}$ *and* $q \in (0,1]$, *with* $\|B\| = q$, *it holds that*

$$W_{\|\cdot\|}(P(z);B;q) = W_{\|\cdot\|}^{0}(P(z);B) = \{\mu \in \mathbb{C} :\ B \perp_{BJ} P(z)\}.$$

It is worth noting that for the linear pencil $P(z) = Bz - A$, the first equality of Definition 1 and (12) yield $W_{\|\cdot\|}^{\epsilon}(Bz - A;B) = F_{\|\cdot\|}^{\epsilon}(A;B)$. Furthermore, if a $\mu_0 \in \mathbb{C}$ satisfies $P(\mu_0) = 0$, then it is immediate that $\mu_0 \in W_{\|\cdot\|}^{\epsilon}(P(z);B)$.

If all the coefficient matrices of $P(z)$ are scalar multiples of B, then the matrix polynomial is written in the form $P(z) = p(z)B$ for some scalar polynomial $p(z)$. Thus, for any $\epsilon \in [0,1)$, the Birkhoff-James ϵ-orthogonality set

$$W_{\|\cdot\|}^{\epsilon}(p(z)B;B) = \left\{\mu \in \mathbb{C} :\ |p(\mu) - \lambda|\,\|B\| \ge \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|,\ \forall\, \lambda \in \mathbb{C}\right\}$$

contains all zeros of $p(z)$.

As in the case of constant matrices, the ϵ-orthogonality set $W_{\|\cdot\|}^{\epsilon}(P(z);B)$ is a natural generalization of the numerical ranges $W_{\|\cdot\|}(P(z);B)$ and $W_{\|\cdot\|}(P(z);B;q)$, and hence, in the remainder of the paper, we focus our interest on this set. In the special case where $n = m$, $B = I_n$ and $\|\cdot\| = \|\cdot\|_2$, it is clear that $W_{\|\cdot\|_2}(P(z);I_n) =$

$\{\mu \in \mathbb{C} : 0 \in F_{\|\cdot\|_2}(P(\mu); I_n)\} = \{\mu \in \mathbb{C} : 0 \in F(P(\mu))\} = W(P(z))$, i.e., the definition of $W_{\|\cdot\|}(P(z); B)$ introduced above is a direct extension of the definition of the standard numerical range $W(P(z))$.

Consider an $n \times m$ matrix polynomial $P(z) = \sum_{l=0}^{l} A_l z^l$ as in (7), a nonzero matrix $B \in \mathbb{C}^{n\times m}$, a matrix norm $\|\cdot\|$, and an $\epsilon \in [0, 1)$.

Proposition 10. *The following hold:*

(i) *For any scalar $\alpha \in \mathbb{C}\backslash\{0\}$, $W^{\epsilon}_{\|\cdot\|}(\alpha P(z); B) = W^{\epsilon}_{\|\cdot\|}(P(z); B)$, $W^{\epsilon}_{\|\cdot\|}(P(\alpha z); B) = \alpha^{-1} W^{\epsilon}_{\|\cdot\|}(P(z); B)$ and $W^{\epsilon}_{\|\cdot\|}(P(z+\alpha); B) = W^{\epsilon}_{\|\cdot\|}(P(z); B) - \alpha$.*

(ii) *If $R(z) = A_0 z^l + \cdots + A_{l-1} z + A_l = z^l P(z^{-1})$ is the reverse matrix polynomial of $P(z)$, then $W^{\epsilon}_{\|\cdot\|}(R(z); B)\backslash\{0\} = \left\{\mu \in \mathbb{C} : \mu^{-1} \in W^{\epsilon}_{\|\cdot\|}(P(z); B)\backslash\{0\}\right\}$.*

(iii) *If the norm $\|\cdot\|$ is invariant under the conjugate operation $\overline{\cdot}$, and the coefficients of $P(z)$ and B are all real matrices, then $W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is symmetric with respect to the real axis.*

(iv) *Suppose the matrix norm $\|\cdot\|$ is induced by a vector norm. If there exist two unit vectors $x_0 \in \mathbb{C}^n$ and $y_0 \in \mathbb{C}^m$ such that $|x_0^* B y_0| \geq \sqrt{1-\epsilon^2}\, \|B\|$, and $x_0^* A_j y_0 = 0$ for every $j = 0, 1, \dots, l$, then $W^{\epsilon}_{\|\cdot\|}(P(z); B) = \mathbb{C}$.*

Proof. (i) It is easy to see that

$$W^{\epsilon}_{\|\cdot\|}(\alpha P(z); B) = \left\{\mu \in \mathbb{C} : \left\|P(\mu) - \frac{\lambda}{\alpha} B\right\| \geq \sqrt{1-\epsilon^2}\, \|B\| \left|\frac{\lambda}{\alpha}\right|, \forall\, \frac{\lambda}{\alpha} \in \mathbb{C}\right\},$$

$$W^{\epsilon}_{\|\cdot\|}(P(\alpha z); B) = \left\{\alpha^{-1}\mu \in \mathbb{C} : \|P(\mu) - \lambda B\| \geq \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda|, \forall\, \lambda \in \mathbb{C}\right\},$$

and

$$W^{\epsilon}_{\|\cdot\|}(P(z+\alpha); B) = \left\{\mu - \alpha \in \mathbb{C} : \|P(\mu) - \lambda B\| \geq \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda|, \forall\, \lambda \in \mathbb{C}\right\}.$$

(ii) A nonzero $\mu \in \mathbb{C}$ lies in $W^{\epsilon}_{\|\cdot\|}(R(z); B)$ if and only if

$$\|\mu^l P(\mu^{-1}) - \lambda B\| \geq \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

or equivalently, if and only if

$$\left\|P(\mu^{-1}) - \frac{\lambda}{\mu^l} B\right\| \geq \sqrt{1-\epsilon^2}\, \|B\| \left|\frac{\lambda}{\mu^l}\right|, \quad \forall\, \frac{\lambda}{\mu^l} \in \mathbb{C}.$$

(iii) It follows from the equalities $\|P(\mu) - \lambda B\| = \|\overline{P(\mu) - \lambda B}\| = \left\|P(\overline{\mu}) - \overline{\lambda} B\right\|$ and $|\overline{\lambda}| = |\lambda|$ $(\mu, \lambda \in \mathbb{C})$.

(iv) For any $\mu \in \mathbb{C}$, it holds that

$$\|P(\mu) - \lambda B\| = \|x_0^*\|\, \|P(\mu) - \lambda B\|\, \|y_0\| \geq \|x_0^* P(\mu) y_0 - \lambda(x_0^* B y_0)\| \geq \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda|$$

for every $\lambda \in \mathbb{C}$. □

Proposition 11. *Suppose the matrix norm $\|\cdot\|$ is induced by a vector norm and $n \geq m$, and let μ_0 be an eigenvalue of $P(z)$ with an associated unit eigenvector $x_0 \in \mathbb{C}^n$. Then for every $\epsilon \in \left[\sqrt{\|B\|^2 - \|Bx_0\|^2} \,/\, \|B\|, 1\right)$, μ_0 lies in $W^{\epsilon}_{\|\cdot\|}(P(z); B)$.*

Proof. Since $\|Bx_0\| \geq \sqrt{1-\epsilon^2}\,\|B\|$, it follows

$$\|P(\mu_0) - \lambda B\|\,\|x_0\| \geq \|P(\mu_0)x_0 - \lambda Bx_0\| = \|\lambda Bx_0\| \geq \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|, \quad \forall\, \lambda \in \mathbb{C}. \ \square$$

For a square matrix polynomial $P(z) = \sum_{j=0}^{l} A_j z^j$, it is known that the numerical range $W(P)$ is unbounded if and only if $0 \in F(A_l)$ [21].

Theorem 12. *Let $P(z)$ be an $n \times m$ matrix polynomial as in (7), $B \in \mathbb{C}^{n\times m}$ be nonzero, and $\epsilon \in [0, 1)$.*

(i) *If $W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is unbounded, then $0 \in F^{\epsilon}_{\|\cdot\|}(A_l; B)$.*

(ii) *Suppose $0 \in F^{\epsilon}_{\|\cdot\|}(A_l; B)$ and 0 is not an isolated point of $W^{\epsilon}_{\|\cdot\|}(R(z); B)$, where $R(z) = \sum_{j=0}^{l} A_{l-j} z^j = z^l P(z^{-1})$. Then the ϵ-orthogonality set $W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is unbounded.*

Proof. (i) Suppose that the ϵ-orthogonality set $W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is unbounded, and let $\mu \in W^{\epsilon}_{\|\cdot\|}(P(z); B)\backslash\{0\}$. Then it holds that

$$\left\|A_l\mu^l + A_{l-1}\mu^{l-1} + \cdots + A_1\mu + A_0 - \lambda B\right\| \geq \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

or

$$|\mu^l| \left\|A_l + A_{l-1}\frac{1}{\mu} + \cdots + A_1\frac{1}{\mu^{l-1}} + A_0\frac{1}{\mu^l} - \frac{\lambda}{\mu^l}B\right\| \geq \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

or

$$\left\|A_l - \frac{\lambda}{\mu^l}B\right\| + \left\|A_{l-1}\frac{1}{\mu} + \cdots + A_1\frac{1}{\mu^{l-1}} + A_0\frac{1}{\mu^l}\right\| \geq \sqrt{1-\epsilon^2}\,\|B\|\left|\frac{\lambda}{\mu^l}\right|, \quad \forall\, \lambda \in \mathbb{C}.$$

For the sake of contradiction, we assume that $0 \notin F^{\epsilon}_{\|\cdot\|}(A_l; B)$, or equivalently, that there exists a $\lambda_0 \in \mathbb{C}$ such that $\|A_l - \lambda_0 B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda_0|$. Since the set $W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is unbounded, for sufficiently large $\mu \in W^{\epsilon}_{\|\cdot\|}(P(z); B)$, the quantity $\left\|A_{l-1}\frac{1}{\mu} + \cdots + A_1\frac{1}{\mu^{l-1}} + A_0\frac{1}{\mu^l}\right\|$ becomes smaller than the difference $\sqrt{1-\epsilon^2}\,\|B\|\,|\lambda_0| - \|A_l - \lambda_0 B\|$. Then setting $\lambda = \lambda_0\mu^l$ yields

$$\left\|A_l - \frac{\lambda}{\mu^l}B\right\| + \left\|A_{l-1}\frac{1}{\mu} + \cdots + A_1\frac{1}{\mu^{l-1}} + A_0\frac{1}{\mu^l}\right\| < \sqrt{1-\epsilon^2}\,\|B\|\left|\frac{\lambda}{\mu^l}\right|.$$

This is a contradiction.

(ii) Consider the reverse matrix polynomial $R(z) = \sum_{j=0}^{l} A_{l-j} z^j = z^l P(z^{-1})$. By Proposition 10 (ii), $W^{\epsilon}_{\|\cdot\|}(R(z); B)\backslash\{0\} = \left\{\mu \in \mathbb{C} : \mu^{-1} \in W^{\epsilon}_{\|\cdot\|}(P(z); B)\backslash\{0\}\right\}$. Since $0 \in F^{\epsilon}_{\|\cdot\|}(A_l; B)$, it follows that $0 \in W^{\epsilon}_{\|\cdot\|}(R(z); B)$. Moreover, since 0 is not an isolated point of $W^{\epsilon}_{\|\cdot\|}(R(z); B)$, there is a sequence $\{\mu_k\}_{k\in\mathbb{N}} \subset W^{\epsilon}_{\|\cdot\|}(R(z); B)\backslash\{0\}$ that converges to the origin. This means that the sequence $\{\mu_k^{-1}\}_{k\in\mathbb{N}} \subset W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is unbounded. Hence, $W^{\epsilon}_{\|\cdot\|}(P(z); B)$ is also unbounded. $\square$

The condition that the origin is not an isolated point of the set $W^{\epsilon}_{\|\cdot\|}(R(z);B)$ is always satisfied in the case of standard numerical range $W(P(z))$. This can be verified by the second part of the proof of [21, Theorem 2.3], the second part of [24, Lemma in page 103], and the fact that the leading coefficient of $P(z)$ is nonzero.

As in the case of constant matrices, the ϵ-orthogonality set $W^{\epsilon}_{\|\cdot\|}(P(z);B)$ can be arbitrarily large for ϵ sufficiently close to 1.

Proposition 13. *Let $P(z)$ be an $n \times m$ matrix polynomial as in (7), $B \in \mathbb{C}^{n\times m}$ be nonzero, and $\epsilon \in [0,1)$. Suppose also that $\Omega \subset \mathbb{C}$ is a compact region such that that for every $\mu \in \Omega$, $P(\mu)$ is not a nonzero scalar multiple of B. Then there is an $\epsilon_\Omega \in [0,1)$ such that $\Omega \subseteq W^{\epsilon_\Omega}_{\|\cdot\|}(P(z);B)$.*

Proof. By Proposition 4, we have that for any $\mu \in \Omega$, there is an $\epsilon_\mu \in [0,1)$ such that $0 \in F^{\epsilon_\mu}_{\|\cdot\|}(P(\mu);B)$, or equivalently, $\mu \in W^{\epsilon_\mu}_{\|\cdot\|}(P(z);B)$. The last part of the proof of Proposition 4 implies the desired conclusion. ☐

Finally, we consider an $n \times n$ matrix polynomial $P(z) = \sum_{j=0}^{l} A_j z^j$ and the norm $\|\cdot\|_2$.

Proposition 14. *Suppose $\mu_0 \in W(P(z))$, and let $x_0 \in \mathbb{C}^n$ such that $\|x_0\|_2 = 1$ and $x_0^* P(\mu_0) x_0 = 0$. Then for every $\epsilon \in \left[\sqrt{\|B\|_2^2 - |x_0^* B x_0|^2}\,/\,\|B\|_2, 1\right)$, μ_0 lies in $W^{\epsilon}_{\|\cdot\|_2}(P(z);B)$.*

Proof. Since $|x_0^* B x_0| \geq \sqrt{1-\epsilon^2}\,\|B\|_2$, it is straightforward to verify that for every $\lambda \in \mathbb{C}$, $\|P(\mu_0) - \lambda B\|_2 \geq \|x_0^*(P(\mu_0) - \lambda B)x_0\|_2 = |\lambda|\,|x_0^* B x_0| \geq \sqrt{1-\epsilon^2}\,\|B\|_2 |\lambda|$. ☐

Corollary 15. *For every $\epsilon \in [0,1)$ such that the interior of disk $\mathcal{D}(0, \sqrt{1-\epsilon^2}\,\|B\|_2)$ has an empty intersection with the standard numerical range $F(B)$, it holds that $W(P(z)) \subseteq W^{\epsilon}_{\|\cdot\|_2}(P(z);B)$.*

4 The boundary

By Definition 1, it is apparent that a $\mu_0 \in \mathbb{C}$ lies in the Birkhoff-James ϵ-orthogonality set $F^{\epsilon}_{\|\cdot\|}(A;B)$ if and only if $\inf_{\lambda\in\mathbb{C}}\{\|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda|\} \geq 0$. Motivated by the last part of the proof of Proposition 2, we specialize this characterization to the boundary of $F^{\epsilon}_{\|\cdot\|}(A;B)$, $\partial F^{\epsilon}_{\|\cdot\|}(A;B)$.

Proposition 16. *Let $A, B \in \mathbb{C}^{m\times n}$ with $B \neq 0$, $\epsilon \in [0,1)$, and $\mu_0 \in F^{\epsilon}_{\|\cdot\|}(A;B)$.*

(i) *The point μ_0 lies on the boundary $\partial F^{\epsilon}_{\|\cdot\|}(A;B)$ if and only if*

$$\inf_{\lambda\in\mathbb{C}} \left\{ \|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda| \right\} = 0.$$

(ii) *If $\epsilon > 0$, then $\mu_0 \in \partial F^{\epsilon}_{\|\cdot\|}(A;B)$ if and only if*

$$\min_{\lambda\in\mathbb{C}} \left\{ \|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda| \right\} = 0,$$

i.e., if and only if there is a $\lambda_0 \in \mathbb{C}$ such that $\|A - \lambda_0 B\| = \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_0|$.

Proof. (i) If μ_0 is a boundary point of $F^{\epsilon}_{\|\cdot\|}(A;B)$, then for every $\delta > 0$, there is a $\lambda_\delta \in \mathbb{C}$ such that

$$\|A - \lambda_\delta B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_\delta| + \delta. \tag{13}$$

Since the difference $\|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda|$ is nonnegative for every $\lambda \in \mathbb{C}$, it follows that $\inf_{\lambda\in\mathbb{C}}\{\|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda|\} = 0$.

For the converse, suppose $\inf_{\lambda\in\mathbb{C}}\{\|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda|\} = 0$, and for the sake of contradiction, assume that $\mu_0 \in \mathrm{Int}[F^{\epsilon}_{\|\cdot\|}(A;B)]$. Then there is a real $r > 0$ such that $\mathcal{D}(\mu_0, r) \subset \mathrm{Int}[F^{\epsilon}_{\|\cdot\|}(A;B)]$, and hence,

$$\mathcal{D}(\mu_0, r) \subset \mathrm{Int}\left[\mathcal{D}\left(\lambda, \frac{\|A - \lambda B\|}{\sqrt{1-\epsilon^2}\,\|B\|}\right)\right], \quad \forall\, \lambda \in \mathbb{C},$$

or

$$\|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda| > \sqrt{1-\epsilon^2}\,\|B\|\,r, \quad \forall\, \lambda \in \mathbb{C},$$

or

$$\inf_{\lambda\in\mathbb{C}}\left\{\|A - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda|\right\} \geq \sqrt{1-\epsilon^2}\,\|B\|\,r > 0.$$

This is a contradiction.

(ii) Suppose $\epsilon > 0$ and $\mu_0 \in \partial F^{\epsilon}_{\|\cdot\|}(A;B)$. Setting $\delta = 1/k$ and $\lambda_\delta = \lambda_k$ ($k = 1, 2, \dots$) in (13) yields

$$\|A - \lambda_k B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_k| + \frac{1}{k},$$

or

$$|\,\|A\| - \|\lambda_k B\|\,| < \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_k| + \frac{1}{k}.$$

Next, we adapt arguments from the proof of Proposition 4. It is clear that $|\lambda_k|\,\|B\| - \|A\| < \sqrt{1-\epsilon^2}\,\|B\|\,(|\mu_0| + |\lambda_k|) + 1/k$, and since $\epsilon > 0$, we have

$$|\lambda_k| < \frac{\|A\| + \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0| + 1}{\|B\|\,(1 - \sqrt{1-\epsilon^2})}.$$

Hence, the sequence $\{\lambda_k\}_{k\in\mathbb{N}}$ is always bounded, and thus, it has a converging subsequence $\{\lambda_{k_t}\}_{t\in\mathbb{N}}$. If we assume that $\lambda_{k_t} \to \lambda_0$, then

$$\|A - \lambda_{k_t} B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_{k_t}| + \frac{1}{k_t}, \quad \forall\, t \in \mathbb{N},$$

or

$$\lim_t \left(\|A - \lambda_{k_t} B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_{k_t}| - \frac{1}{k_t}\right) \leq 0,$$

or

$$\|A - \lambda_0 B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda_0| \leq 0,$$

where the latter relation is possible only as an equality.

The converse follows readily from (i). □

If $\epsilon > 0$, then the above proposition implies that for any $\mu_0 \in \partial F^{\epsilon}_{\|\cdot\|}(A;B)$, there is a generating disk $\mathcal{D}\left(\lambda_0, \frac{\|A-\lambda_0 B\|}{\sqrt{1-\epsilon^2}\,\|B\|}\right)$ such that $\mu_0 \in \partial\mathcal{D}\left(\lambda_0, \frac{\|A-\lambda_0 B\|}{\sqrt{1-\epsilon^2}\,\|B\|}\right)$. As a consequence, since $F^{\epsilon}_{\|\cdot\|}(A;B)$ is convex and lies in $\mathcal{D}\left(\lambda_0, \frac{\|A-\lambda_0 B\|}{\sqrt{1-\epsilon^2}\,\|B\|}\right)$, we have the following corollaries (see also Corollary 3).

Corollary 17. *If $0 < \epsilon < 1$, then the boundary $\partial F^{\epsilon}_{\|\cdot\|}(A;B)$ does not have any flat portions.*

Corollary 18. *Suppose $0 < q < 1$ and $\|B\| \geq q$. Then for any $\mu_0 \in \partial F_{\|\cdot\|}(A;B;q)$, there is a $\lambda_0 \in \mathbb{C}$ such that $\|A - \lambda_0 B\| = |\mu_0 - q\lambda_0|$. In particular, the boundary of the q-numerical range $F_{\|\cdot\|}(A;B;q)$ does not have any flat portions.*

On the other hand, if $\|B\| = 1$, then $F^{0}_{\|\cdot\|}(A;B) = F_{\|\cdot\|}(A;B)$ might have flat portions; see, for example, Proposition 20 in [8]. Hence, in Proposition 16 (ii), the condition $\epsilon > 0$ cannot be omitted.

The properties of a point μ of the standard numerical range $W(P(z))$ are strongly related to the properties of the origin as a point of $F(P(\mu))$ [16, 24, 26]. Parts (i) and (ii) of the following theorem are generalizations of Theorem 1.1 in [24] and Theorem 2 in [16], respectively. (We denote the derivative of $P(z)$ by $P'(z)$.)

Theorem 19. *Suppose $P(z)$ is an $n \times m$ matrix polynomial as in (7), $B \in \mathbb{C}^{n\times m}$ is nonzero, $\epsilon \in [0,1)$, and $\mu_0 \in W^{\epsilon}_{\|\cdot\|}(P(z);B)$.*

(i) *If $\mu_0 \in \partial W^{\epsilon}_{\|\cdot\|}(P(z);B)$, then $0 \in \partial F^{\epsilon}_{\|\cdot\|}(P(\mu_0);B)$.*

(ii) *If $0 \in \partial F^{\epsilon}_{\|\cdot\|}(P(\mu_0);B)\backslash F^{\epsilon}_{\|\cdot\|}(P'(\mu_0);B)$ and $P(\mu_0) \neq 0$, then μ_0 lies on the boundary $\partial W^{\epsilon}_{\|\cdot\|}(P(z);B)$.*

Proof. (i) Since $\mu_0 \in W^{\epsilon}_{\|\cdot\|}(P(z);B)$, it is clear that $0 \in F^{\epsilon}_{\|\cdot\|}(P(\mu_0);B)$. For the sake of contradiction, we assume that the origin lies in the interior of $F^{\epsilon}_{\|\cdot\|}(P(\mu_0);B)$, $\mathrm{Int}[F^{\epsilon}_{\|\cdot\|}(P(\mu_0);B)]$. Then by Proposition 16 (i), there exists a $\delta > 0$ such that

$$\inf_{\lambda\in\mathbb{C}} \left\{ \|P(\mu_0) - \lambda B\| - \sqrt{1-\epsilon^2}\,\|B\|\,|\mu_0 - \lambda| \right\} > \delta,$$

and hence,

$$\|P(\mu_0) - \lambda B\| - \delta > \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|, \quad \forall\, \lambda \in \mathbb{C}.$$

Also, we have

$$P(z) = P(\mu_0) + (z-\mu_0)P'(\mu_0) + (z-\mu_0)E(z,\mu_0), \tag{14}$$

where $\|E(z,\mu_0)\| = o(1)$ as $|z - \mu_0| \to 0$. As a consequence, there is a real $r > 0$ such that for every $\mu \in \mathcal{D}(\mu_0, r)$, $|\mu - \mu_0|\,\|P'(\mu_0) + E(\mu,\mu_0)\| \leq \delta$. Thus, for every $\mu \in \mathcal{D}(\mu_0, r)$, it holds that

$$\|P(\mu_0) - \lambda B\| - |\mu - \mu_0|\,\|P'(\mu_0) + E(\mu,\mu_0)\| > \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

or

$$\|P(\mu_0) + (\mu - \mu_0)P'(\mu_0) + (\mu - \mu_0)E(\mu, \mu_0) - \lambda B\| > \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

or

$$\|P(\mu) - \lambda B\| > \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

and hence, μ_0 is an interior point of $W^{\epsilon}_{\|\cdot\|}(P(z); B)$; this is a contradiction. Thus, the origin is a boundary point of $F^{\epsilon}_{\|\cdot\|}(P(\mu); B)$.

(ii) For the sake of contradiction, assume that $\mu_0 \in \mathrm{Int}[W^{\epsilon}_{\|\cdot\|}(P(z); B)]$. Hence, there is a $\delta > 0$ such that $\mathcal{D}(\mu_0, \delta) \subset \mathrm{Int}[W^{\epsilon}_{\|\cdot\|}(P(z); B)]$. Recall (14), and observe that since $0 \notin F^{\epsilon}_{\|\cdot\|}(P'(\mu_0); B)$, there is a $\lambda_1 \in \mathbb{C}$ such that $\|P'(\mu_0) - \lambda_1 B\| < \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_1|$. By choosing δ sufficiently small, we may assume that for every μ in the (closed) circular annulus $\mathcal{D}(\mu_0, \delta, \delta/2) = \{\mu \in \mathbb{C} : \delta/2 \leq |\mu - \mu_0| \leq \delta\}$, it holds that

$$\|E(\mu, \mu_0)\| + \|P'(\mu_0) - \lambda_1 B\| < \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_1|,$$

or

$$\|(\mu - \mu_0)P'(\mu_0) + (\mu - \mu_0)E(\mu, \mu_0) - (\mu - \mu_0)\lambda_1 B\| < \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_1|\, |\mu - \mu_0|.$$

Hence, we can define

$$\xi = \min_{\mu \in \mathcal{D}(\mu_0, \delta, \delta/2)} \left\{ |\mu - \mu_0| \left(\sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_1| - \|P'(\mu_0) + E(\mu, \mu_0) - \lambda_1 B\| \right) \right\} > 0.$$

Since $0 \in \partial F^{\epsilon}_{\|\cdot\|}(P(\mu_0); B)$, Proposition 16 (i) implies that there is a $\lambda_0 \in \mathbb{C}$ such that

$$\|P(\mu_0) - \lambda_0 B\| < \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_0| + \xi.$$

Consequently, for every $\mu \in \mathcal{D}(\mu_0, \delta, \delta/2)$,

$$\|P(\mu_0) - \lambda_0 B\| < \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_0| + |\mu - \mu_0| \left(\sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_1| - \|P'(\mu_0) + E(\mu, \mu_0) - \lambda_1 B\| \right),$$

or

$$\|P(\mu) - (\lambda_0 + \lambda_1(\mu - \mu_0))B\| < \sqrt{1-\epsilon^2}\, \|B\|\, (|\lambda_0| + |\lambda_1(\mu - \mu_0)|).$$

Observe now that λ_0 and λ_1 do not depend on μ, and thus, we can choose a $\hat{\mu} \in \mathcal{D}(\mu_0, \delta, \delta/2)$ such that $\arg(\lambda_1(\hat{\mu} - \mu_0)) = \arg(\lambda_0)$. Then it follows

$$\|P(\hat{\mu}) - (\lambda_0 + \lambda_1(\hat{\mu} - \mu_0))B\| < \sqrt{1-\epsilon^2}\, \|B\|\, |\lambda_0 + \lambda_1(\hat{\mu} - \mu_0)|\,,$$

and hence, $\hat{\mu} \notin W^{\epsilon}_{\|\cdot\|}(P(z); B)$; this is a contradiction. □

Definition (11), Proposition 2 and Theorem 19 (i) yield the following.

Proposition 20. *Let $P(z)$ be an $n \times m$ matrix polynomial as in (7), $B \in \mathbb{C}^{n \times m}$ be nonzero, and $0 \leq \epsilon_1 < \epsilon_2 < 1$. Then $W^{\epsilon_1}_{\|\cdot\|}(P(z); B) \subseteq W^{\epsilon_2}_{\|\cdot\|}(P(z); B)$, and for any $\mu \in W^{\epsilon_1}_{\|\cdot\|}(P(z); B)$ such that $P(\mu) \neq 0$, μ lies in the interior of $W^{\epsilon_2}_{\|\cdot\|}(P(z); B)$.*

Proof. For any $\mu \in W^{\epsilon_1}_{\|\cdot\|}(P(z);B)$, the origin lies in $F^{\epsilon_1}_{\|\cdot\|}(P(\mu);B) \subseteq F^{\epsilon_2}_{\|\cdot\|}(P(\mu);B)$, and thus, $\mu \in W^{\epsilon_2}_{\|\cdot\|}(P(z);B)$. Moreover, if $\mu \in W^{\epsilon_1}_{\|\cdot\|}(P(z);B)$ with $P(\mu) \neq 0$, then the matrix $P(\mu)$ cannot be a scalar multiple of B and Proposition 2 implies that the origin lies in the interior of $F^{\epsilon_2}_{\|\cdot\|}(P(\mu);B)$. Hence, by Theorem 19 (i), μ is an interior point of $W^{\epsilon_2}_{\|\cdot\|}(P(z);B)$. □

Corollary 21. *For any scalar $b \in \mathbb{C}$, $\epsilon \in [0,1)$ and $q \in (0,1]$, we have that $W^{\epsilon}_{\|\cdot\|}(P(z);bB) = W^{\epsilon}_{\|\cdot\|}(P(z);B)$, and $W_{\|\cdot\|}(P(z);bB;q) = (\supseteq, \subseteq)\ W_{\|\cdot\|}(P(z);B;q)$ if $|b| = (>, <)\ 1$.*

Proof. Since the Birkhoff-James ϵ-orthogonality is homogeneous,

$$W^{\epsilon}_{\|\cdot\|}(P(z);bB) = \{\mu \in \mathbb{C} :\ bB \perp^{\epsilon}_{BJ} P(\mu)\} = \{\mu \in \mathbb{C} :\ B \perp^{\epsilon}_{BJ} P(\mu)\}.$$

Moreover,

$$\begin{aligned} W_{\|\cdot\|}(P(z);bB;q) &= \left\{\mu \in \mathbb{C} :\ bB \perp^{\epsilon}_{BJ} P(\mu),\, \epsilon = \frac{\sqrt{|b|^2\|B\|^2 - q^2}}{|b|\,\|B\|}\right\} \\ &= \left\{\mu \in \mathbb{C} :\ B \perp^{\epsilon}_{BJ} P(\mu),\, \epsilon = \frac{\sqrt{\|B\|^2 - q^2/|b|^2}}{\|B\|}\right\}. \end{aligned}$$

The proof is completed by Proposition 20. □

Corollary 22. *Suppose the matrix norm $\|\cdot\|$ is induced by a vector norm, and let $x_0 \in \mathbb{C}^n$ and $y_0 \in \mathbb{C}^m$ be two unit vectors such that $|x_0^* B y_0| = \|B\|$. Then for any $\epsilon \in [0,1)$, the Birkhoff-James ϵ-orthogonality set $W^{\epsilon}_{\|\cdot\|}(P(z);B)$ contains all zeros of the scalar polynomial $x_0^* P(z) y_0 = x_0^* A_l y_0 z^l + \cdots + x_0^* A_1 y_0 z + x_0^* A_0 y_0$. Moreover, for any $\mu \in \mathbb{C}$ such that $P(\mu) \neq 0$ and $x_0^* P(\mu) y_0 = 0$, it holds that $\mu \in \mathrm{Int}[W^{\epsilon}_{\|\cdot\|}(P(z);B)]$ for every $\epsilon \in (0,1)$.*

Proof. Let $\mu_0 \in \mathbb{C}$ be a zero of the scalar polynomial $x_0^* P(z) y_0$. Then for every $\lambda \in \mathbb{C}$, $\|P(\mu_0) - \lambda B\| \geq \|x_0^* [P(\mu_0) - \lambda B]\, y_0\| \geq \|B\|\,|\lambda|$. Thus, μ_0 lies in $W^{0}_{\|\cdot\|}(P(z);B)$, and Proposition 20 completes the proof. □

The last result of the section is partially complementary to Proposition 20 and gives a sufficient condition for the appearance of isolated points.

Proposition 23. *Let $P(z)$ be an $n \times m$ matrix polynomial as in (7), $B \in \mathbb{C}^{n\times m}$ be nonzero, and $0 \leq \epsilon < 1$. If there is a $\mu_0 \in \mathbb{C}$ such that $P(\mu_0) = 0$ and $0 \notin F^{\epsilon}_{\|\cdot\|}(P'(\mu_0);B)$, then μ_0 is an isolated point of $W^{\epsilon}_{\|\cdot\|}(P(z);B)$.*

Proof. As in the proof of Theorem 19 (see (14)), we have

$$P(z) = P(\mu_0) + (z-\mu_0)P'(\mu_0) + (z-\mu_0)E(z,\mu_0),$$

where $P(\mu_0) = 0$ and $\|E(z,\mu_0)\| = o(1)$ as $|z-\mu_0| \to 0$. Since $0 \notin F^{\epsilon}_{\|\cdot\|}(P'(\mu_0);B)$, there is a λ_0 such that $\|P'(\mu_0) - \lambda_0 B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda_0|$, and by choosing a sufficiently small $\delta > 0$, we may assume that for every $\mu \in \mathcal{D}(\mu_0,\delta)\backslash\{\mu_0\}$,

$$\|E(\mu,\mu_0)\| + \|P'(\mu_0) - \lambda_0 B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda_0|,$$

or

$$\|(\mu-\mu_0)P'(\mu_0)+(\mu-\mu_0)E(\mu,\mu_0)-(\mu-\mu_0)\lambda_0 B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda_0|\,|\mu-\mu_0|,$$

or

$$\|P(\mu)-\lambda_0(\mu-\mu_0)B\| < \sqrt{1-\epsilon^2}\,\|B\|\,|\lambda_0(\mu-\mu_0)|.$$

As a consequence, $W^{\epsilon}_{\|\cdot\|}(P(z);B)\cap\mathcal{D}(\mu_0,\delta)=\{\mu_0\}$, and μ_0 is an isolated point of $W^{\epsilon}_{\|\cdot\|}(P(z);B)$. □

5 The case of norms induced by inner products

Let $A,B\in\mathbb{C}^{n\times m}$ with $B\neq 0$, and $\epsilon\in[0,1)$, and suppose that the matrix norm $\|\cdot\|$ is induced by the inner product of matrices $\langle\cdot,\cdot\rangle$. Then by Property ($P_7$) (see also Proposition 13 in [8]), the Birkhoff-James ϵ-orthogonality set of A with respect to B is a closed disk, namely,

$$F^{\epsilon}_{\|\cdot\|}(A;B)=\mathcal{D}\left(\frac{\langle A,B\rangle}{\|B\|^2}\,,\,\left\|A-\frac{\langle A,B\rangle}{\|B\|^2}\,B\right\|\frac{\epsilon}{\sqrt{1-\epsilon^2}\,\|B\|}\right).$$

It is worth mentioning that this relation (independently from the proof of [8, Proposition 13]) can be confirmed by the observation that the Birkhoff-James ϵ-orthogonality coincides with the inner product ϵ-orthogonality [7, 9]. In particular, a scalar $\mu\in\mathbb{C}$ lies in $F^{\epsilon}_{\|\cdot\|}(A;B)$ if and only if

$$B\perp^{\epsilon}(A-\mu B),$$

or equivalently, if and only if

$$|\langle B,A-\mu B\rangle|\leq\epsilon\,\|B\|\,\|A-\mu B\|,$$

or equivalently, if and only if

$$\langle B,A-\mu B\rangle\,\langle A-\mu B,B\rangle\leq\epsilon^2\|B\|^2\langle A-\mu B,A-\mu B\rangle,$$

or equivalently, if and only if

$$\frac{|\langle A,B\rangle|^2}{\|B\|^4}-\mu\,\frac{\langle B,A\rangle}{\|B\|^2}-\overline{\mu}\,\frac{\langle A,B\rangle}{\|B\|^2}+|\mu|^2\leq\epsilon^2\left(\frac{\|A\|^2}{\|B\|^2}-\mu\,\frac{\langle B,A\rangle}{\|B\|^2}-\overline{\mu}\,\frac{\langle A,B\rangle}{\|B\|^2}+|\mu|^2\right),$$

or equivalently, if and only if

$$\left|\mu-\frac{\langle A,B\rangle}{\|B\|^2}\right|^2(1-\epsilon^2)\leq\frac{\epsilon^2}{\|B\|^2}\left\|A-\frac{\langle A,B\rangle}{\|B\|^2}\,B\right\|^2.$$

Consider now an $n\times m$ matrix polynomial $P(z)=\sum_{j=0}^{l}A_jz^j$ as in (7). Then by (12), we have

$$\begin{aligned}W^{\epsilon}_{\|\cdot\|}(P(z);B) &= \{\mu\in\mathbb{C}:\,B\perp^{\epsilon}_{BJ}P(\mu)\}\\ &= \{\mu\in\mathbb{C}:\,B\perp^{\epsilon}P(\mu)\}\\ &= \{\mu\in\mathbb{C}:\,|\langle P(\mu),B\rangle|\leq\epsilon\,\|B\|\,\|P(\mu)\|\}\,. \qquad (15)\end{aligned}$$

As a consequence,

$$\begin{aligned} W^{\epsilon}_{\|\cdot\|}(P(z);B) &= \left\{\mu\in\mathbb{C}:\, |\langle P(\mu),B\rangle|^2 \le \epsilon^2\|B\|^2\|P(\mu)\|^2\right\} \\ &= \left\{\mu\in\mathbb{C}:\, \langle P(\mu),B\rangle\langle B,P(\mu)\rangle \le \epsilon^2\|B\|^2\langle P(\mu),P(\mu)\rangle\right\} \\ &= \left\{\mu\in\mathbb{C}:\, \langle\sum_{j=0}^{l}A_j\mu^j,B\rangle\langle B,\sum_{j=0}^{l}A_j\mu^j\rangle \le \epsilon^2\|B\|^2\langle\sum_{j=0}^{l}A_j\mu^j,\sum_{j=0}^{l}A_j\mu^j\rangle\right\} \\ &= \left\{\mu\in\mathbb{C}:\, \sum_{i,j=0}^{l}\langle A_i,B\rangle\langle B,A_j\rangle\mu^i\overline{\mu}^j - \epsilon^2\|B\|^2\sum_{i,j=0}^{l}\langle A_i,A_j\rangle\mu^i\overline{\mu}^j \le 0\right\}. \end{aligned}$$

Writing $\mu = u + \mathrm{i}v$ $(u, v \in \mathbb{R})$, the function

$$p_\epsilon(u,v) = \sum_{i,j=0}^{l}\langle A_i,B\rangle\langle B,A_j\rangle(u+\mathrm{i}v)^i(u-\mathrm{i}v)^j - \epsilon^2\|B\|^2\sum_{i,j=0}^{l}\langle A_i,A_j\rangle(u+\mathrm{i}v)^i(u-\mathrm{i}v)^j$$

is a scalar polynomial in $u, v \in \mathbb{R}$ of total degree $2l$, with real coefficients. Thus, the boundary $\partial W^{\epsilon}_{\|\cdot\|}(P(z);B)$ lies on the algebraic curve

$$\{u+\mathrm{i}v\in\mathbb{C}:\, p_\epsilon(u,v)=0,\, u,v\in\mathbb{R}\}.$$

Furthermore, for $\epsilon = 0$, we have

$$W^{0}_{\|\cdot\|}(P(z);B) = \left\{\mu\in\mathbb{C}:\, \langle A_l,B\rangle\mu^l+\cdots+\langle A_1,B\rangle\mu+\langle A_0,B\rangle=0\right\}. \qquad (16)$$

As a consequence, Proposition 20 yields the following result which is similar to Corollary 22.

Corollary 24. *For any $\epsilon \in [0, 1)$, all zeros of the scalar polynomial $\langle P(z), B\rangle = \langle A_l, B\rangle z^l + \cdots + \langle A_1, B\rangle z + \langle A_0, B\rangle$ lie in the ϵ-orthogonality set $W^{\epsilon}_{\|\cdot\|}(P(z); B)$. Moreover, for any $\mu \in \mathbb{C}$ such that $P(\mu) \neq 0$ and $\langle P(\mu), B\rangle = 0$, it holds that $\mu \in \mathrm{Int}[W^{\epsilon}_{\|\cdot\|}(P(z); B)]$ for every $\epsilon \in (0, 1)$.*

The above discussion allows the construction of empty approximate orthogonality sets for matrix polynomials. In particular, if $\langle A_j, B\rangle = 0$, $j = 1, 2, \ldots, l$, and $\langle A_0, B\rangle \neq 0$, then the polynomial $\langle P(z), B\rangle = \langle A_0, B\rangle$ is constant and nonzero, and by (16), $W^{0}_{\|\cdot\|}(P(z); B) = \emptyset$. In this special case, we may also say that $W^{0}_{\|\cdot\|}(P(z); B) = \{\infty\}$, since the Birkhoff-James ϵ-orthogonality set of the reverse matrix polynomial $R(z) = \sum_{j=0}^{l} A_{l-j}z^j$, $W^{0}_{\|\cdot\|}(R(z); B) = \left\{\mu \in \mathbb{C} :\, \langle A_0, B\rangle\mu^l = 0\right\}$, coincides with the origin. This is compatible to Proposition 10 (ii), Theorem 12, and the fact that for any $\epsilon \in (0, 1)$, $W^{\epsilon}_{\|\cdot\|}(P(z); B) = \{\mu \in \mathbb{C} :\, |\langle A_0, B\rangle| \le \epsilon\,\|B\|\,\|P(\mu)\|\}$ is unbounded and contains a set of the form $\{z \in \mathbb{C} :\, |z| \ge r\}$ for some real $r > 0$.

In our last example, we consider the 3×2 quadratic matrix polynomial

$$P(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0.8 \end{bmatrix} z^2 + \begin{bmatrix} 1 & \mathrm{i} \\ 0 & -1 \\ 0.5 & 0.1 \end{bmatrix} z + \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ -0.1 & 0 \end{bmatrix},$$

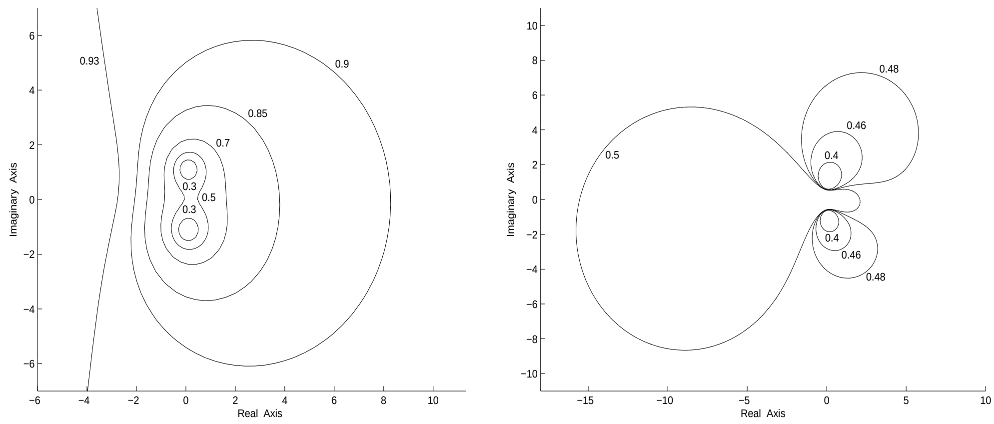

Figure 3: Birkhoff-James ϵ-orthogonality sets of $P(z)$ (left part) and $R(z)$ (right part).

its reverse matrix polynomial

$$R(z) = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ -0.1 & 0 \end{bmatrix} z^2 + \begin{bmatrix} 1 & \mathrm{i} \\ 0 & -1 \\ 0.5 & 0.1 \end{bmatrix} z + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0.8 \end{bmatrix},$$

and the matrix

$$B = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.9 \\ 0 & 0.2 \end{bmatrix}.$$

For the Frobenius norm, and by applying (15), we have drawn the boundaries of the ϵ-orthogonality sets $W^{\epsilon}_{\|\cdot\|_F}(P(z);B)$, $\epsilon = 0.3, 0.5, 0.7, 0.85, 0.9, 0.93$, and $W^{\epsilon}_{\|\cdot\|_F}(R(z);B)$, $\epsilon = 0.4, 0.46, 0.48, 0.5$, in the left and right parts of Figure 3, respectively. Note that $W^{0.5}_{\|\cdot\|_F}(R(z);B)$ coincides with the complex plane excluded the lemniscus containing the origin. The sets $W^{\epsilon}_{\|\cdot\|_F}(P(z);B)$ and $W^{\epsilon}_{\|\cdot\|_F}(R(z);B)$ become unbounded when $\epsilon = 0.9288$ and $\epsilon = 0.4928$, respectively, and the origin meets the Birkhoff-James ϵ-orthogonality sets of the corresponding leading coefficients, confirming Theorem 12. Propositions 10 (ii), 13 and 20 are also apparently verified. Furthermore, the zeros $0.0843 \pm \mathrm{i}\,1.1216$ of the polynomial $\langle P(z), B\rangle = 1.66z^2 - 0.28z + 2.1$ lie in $W^{\epsilon}_{\|\cdot\|_F}(P(z);B)$, and the zeros $0.0667 \pm \mathrm{i}\,0.8866$ of the polynomial $\langle R(z), B\rangle = 2.1z^2 - 0.28z + 1.66$ lie in $W^{\epsilon}_{\|\cdot\|_F}(R(z);B)$, confirming Corollary 24.

Acknowledgment. The authors wish to thank an anonymous referee for his valuable comments.

References

[1] A. Aretaki and I. Maroulas, Investigating the numerical range and q-numerical range of nonsquare matrices, *preprint*, 2010.

[2] G. Birkhoff, Orthogonality in linear metric spaces, *Duke Math. J.*, **1** (1935), 169–172.

[3] F.F. Bonsall and J. Duncan, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, London Mathematical Society Lecture Note Series, Cambridge University Press, New York, 1971.

[4] F.F. Bonsall and J. Duncan, *Numerical Ranges II*, London Mathematical Society Lecture Notes Series, Cambridge University Press, New York, 1973.

[5] M.-T. Chien and H. Nakazato, The boundary of the q-numerical range of a reducible matrix, *Linear and Multilinear Algebra*, **55** (2007), 275–292.

[6] M.-T. Chien, H. Nakazato and P. Psarrakos, On the q-numerical range of matrices and matrix polynomials, *Linear and Multilinear Algebra*, **53** (2005), 357–374.

[7] J. Chmielinski, On an ε-Birkhoff orthogonality, *J. Ineq. Pure and Appl. Math.*, **6** (2005), Article 79.

[8] Ch. Chorianopoulos, S. Karanasios and P. Psarrakos, A definition of numerical range of rectangular matrices, *Linear and Multilinear Algebra*, **57** (2009), 459–475.

[9] S.S. Dragomir, On approximation of continuous linear functionals in normed linear spaces, *An. Univ. Timişoara Ser. Ştiinţ. Mat.*, **29** (1991), 51–58.

[10] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

[11] K.E. Gustafson and D.K.M. Rao, *Numerical Range: The Field of Values of Linear Operators and Matrices*, Springer-Verlag, New York, 1997.

[12] F. Hausdorff, Der Wertvorrat einer Bilinearform, *Math. Z.*, **3** (1919), 314–316.

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

[14] R.C. James, Orthogonality and linear functionals in normed linear spaces, *Trans. Amer. Math. Soc.*, **61** (1947), 265–292.

[15] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1976.

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

[17] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd edition, Academic Press, Orlando, 1985.

[18] H.W.J. Lenferink and M.N. Spijker, On generalization of the numerical range of a matrix, *Linear Algebra Appl.*, **140** (1990), 251–266.

[19] C.-K. Li, P.P. Mehta and L. Rodman, A generalized numerical range: the range of a constrained sesqulinear form, *Linear and Multilinear Algebra*, **37** (1994), 25–49.

[20] C.-K. Li and H. Nakazato, Some results on the q-numerical range, *Linear and Multilinear Algebra*, **43** (1998), 385–409.

[21] C.-K. Li and L. Rodman, Numerical range of matrix polynomials, *SIAM J. Matrix Anal. Appl.*, **15** (1994), 1256–1265.

[22] M. Marcus and P. Andresen, Constrained extrema of bilinear functionals, *Monatsh. Math.*, **84** (1977), 219–235.

[23] A.S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.

[24] J. Maroulas and P. Psarrakos, The boundary of numerical range of matrix polynomials, *Linear Algebra Appl.*, **267** (1997), 101–111.

[25] J. Maroulas and P. Psarrakos, On the connectedness of numerical range of matrix polynomials, *Linear Algebra Appl.*, **280** (1998), 97–108.

[26] H. Nakazato and P. Psarrakos, On the shape of numerical range of matrix polynomials, *Linear Algebra Appl.*, **338** (2001), 105–123.

[27] P. Psarrakos and M. Tsatsomeros, On the relation between the numerical range and the joint numerical range of matrix polynomials, *Electr. J. Linear Algebra*, **6** (2000), 20–30.

[28] J.G. Stampfli and J.P. Williams, Growth conditions and the numerical range in a Banach algebra, *Tôhoku Math. Journ.*, **20** (1968), 417–424.

[29] O. Toeplitz, Das algebraische Analogon zu einern Satze von Fejér, *Math. Z.*, **2** (1918), 187–197.