

AN OPTIMAL APPROXIMATION PROBLEM FOR FREE POLYNOMIALS

PALAK ARORA, MERIC AUGAT, MICHAEL T. JURY, AND MEREDITH SARGENT

(Communicated by Javad Mashreghi)

ABSTRACT. Motivated by recent work on optimal approximation by polynomials in the unit disk, we consider the following noncommutative approximation problem: for a polynomial f in d freely noncommuting arguments, find a free polynomial p_n , of degree at most n , to minimize $c_n := \|p_n f - 1\|^2$. (Here the norm is the ℓ^2 norm on coefficients.) We show that $c_n \rightarrow 0$ if and only if f is nonsingular in a certain nc domain (the row ball), and prove quantitative bounds. As an application, we obtain a new proof of the characterization of polynomials cyclic for the d -shift.

1. INTRODUCTION

This paper concerns an approximation problem for polynomials in noncommuting indeterminates. To state the problem, let $\mathbf{x} = \{x_1, \dots, x_d\}$ be a set of freely noncommuting indeterminates and write $\mathbf{x} = (x_1, \dots, x_d)$. We write $\mathbb{C}\langle \mathbf{x} \rangle$ for the **free algebra** in the indeterminates x_1, \dots, x_d and its elements are **free polynomials**. We write elements $p \in \mathbb{C}\langle \mathbf{x} \rangle$ as $p(\mathbf{x}) = \sum_{w \in \langle \mathbf{x} \rangle} p_w x^w$ (with $p_w \in \mathbb{C}$ nonzero for all but finitely many w). Here, w is a “word” $w = i_1 i_2 \cdots i_m$ in the “letters” $1, \dots, d$, and $x^w := x_{i_1} x_{i_2} \cdots x_{i_m}$. In this case the length of the word is defined to be m , and the **degree** of p is defined to be the maximum length of words w for which $p_w \neq 0$. For example, the polynomial $p(x_1, x_2, x_3) = x_1 x_3 - 7x_1 x_2 x_3 + 5x_1^2 x_2 x_3$ has degree four since the word $x_1^2 x_2 x_3 = x_1 x_1 x_2 x_3$ has length four, and this is the longest word that appears.

We equip $\mathbb{C}\langle \mathbf{x} \rangle$ with the natural inner product obtained by declaring the “words” in the letters x_i to be an orthonormal system:

$$\langle x_{i_1} \cdots x_{i_n}, x_{j_1} \cdots x_{j_m} \rangle := \begin{cases} 1 & m = n \text{ and } i_\ell = j_\ell \text{ for } 1 \leq \ell \leq n; \\ 0 & \text{otherwise.} \end{cases}$$

Thus for $p = \sum_{w \in \langle \mathbf{x} \rangle} p_w x^w$, $q = \sum_{w \in \langle \mathbf{x} \rangle} q_w x^w$, we have

$$\langle p, q \rangle = \sum_{w \in \langle \mathbf{x} \rangle} p_w \overline{q_w}$$

Received by the editors October 7, 2022, and, in revised form, March 1, 2023.

2020 *Mathematics Subject Classification.* Primary 16S38, 46L52, 47A16; Secondary 47A13.

The first author’s research was supported by NSF grant DMS-2154494.

The second author’s research was supported by NSF grant DMS-2155033.

The third author’s research was supported by NSF grant DMS-2154494.

The fourth author was supported by the Pacific Institute for the Mathematical Sciences (PIMS). The research and findings may not reflect those of the Institute.

and the associated ℓ^2 norm on coefficients

$$\|p\|^2 = \sum_{w \in \langle x \rangle} |p_w|^2.$$

Given a polynomial $f \in \mathbb{C}\langle x \rangle$, for each n there is a unique polynomial $p_n \in \mathbb{C}\langle x \rangle$, of degree at most n , which minimizes the norm

$$c_n := \|p_n f - 1\|^2.$$

We call this p_n the n^{th} *optimal polynomial approximant* (OPA) of f^{-1} . The sequence (c_n) is evidently non-increasing. We can ask two questions about the numbers (c_n) :

Question 1. Under what conditions on f do we have $c_n \rightarrow 0$?

Question 2. If $c_n \rightarrow 0$, how fast? (That is, give bounds on the rate of decay.)

Our main theorem gives answers to Questions 1 and 2, in the more general setting of polynomials with matrix coefficients. (For precise definitions see Section 2)

Theorem 1.1. *Let $F \in M_k(\mathbb{C}) \otimes \mathbb{C}\langle x \rangle$ be a $k \times k$ -matrix valued free polynomial, let (P_n) denote the sequence of optimal polynomial approximants, and put*

$$c_n := \|P_n(x)F(x) - I\|_2^2.$$

Then:

- (1) $c_n \rightarrow 0$ if and only if F is nonsingular in the row ball.
- (2) If F is nonsingular in the row ball, then

$$c_n \lesssim \frac{1}{n^p}$$

for some $p > 0$ depending on F . In particular, if F is a product of ℓ atomic factors, we can take $p = \frac{1}{3\ell-1}$.

Let us describe the motivation for these questions. The first motivation comes from the problem of characterizing the so-called *cyclic* free polynomials. We say f is *cyclic* if every free polynomial q can be approximated arbitrarily well in the ℓ^2 norm by polynomials of the form pf . Further discussion about cyclicity can be found in, for example, [5, 8]. It was proved in [11] that a free polynomial f is cyclic if and only if it is nonsingular in the row ball, but the proof is quite abstract and indirect. It turns out that cyclicity of f is equivalent to the statement that $c_n \rightarrow 0$ (see Section 2). In this light, Question 1 asks about “qualitative cyclicity,” and Question 2 asks about “quantitative cyclicity.” We refer to [11] for further discussion and motivation for the cyclicity question in the noncommutative context.

Thus, one of our motivations was to try to see if there is a more elementary proof of the characterization of cyclic f obtained in [11]. It turns out that such a proof is possible, and in the proof we present in fact allows us to give some answer to Question 2 (namely, item (2) of the theorem), though we do not know if this answer is sharp (see the remarks at the end of Section 3).

Besides this, the problem can be motivated by considering what is already known for polynomials in one complex variable, and asking to what extent this result can be generalized to the free setting. In particular, Bénéteau et al. [2] considered the analogous approximation problem for polynomials in one complex variable: given a

polynomial $f \in \mathbb{C}[z]$, let p_n be the polynomial of degree at most n minimizing the ℓ^2 norm $c_n := \|p_n f - 1\|^2$, and estimate the c_n . Among other things, they proved the following (we have stated their result in a slightly different form):

Theorem 1.2. *Let $f \in \mathbb{C}[z]$ be a polynomial with no zeroes in the disk $|z| < 1$. Then:*

$$c_n := \|p_n f - 1\|^2 \lesssim \frac{1}{n}.$$

In the same paper, the authors have also shown that if f has at least one zero on the boundary of the disk, then the reverse estimate also holds.

The study of optimal polynomial approximants has roots tracing back to at least the 1970's where engineering and applied mathematics researchers investigated polynomial least-squares inverses in the context of digital filters in signal processing. A survey for mathematical audiences can be found in [1]. Roughly speaking, one considers the polynomials p_n as approximations to $1/f$, even though $1/f$ will not in general have ℓ^2 coefficients and therefore cannot be approximated directly by Hilbert space methods.

In the mathematical context, OPAs were introduced in [2] as a tool to study cyclicity in a family of Dirichlet-type spaces on the disk \mathbb{D} which includes the classical Hardy space of the disk $H^2(\mathbb{D})$. (This was then extended to a more general reproducing kernel Hilbert space context in [9]. Surveys of OPA results can be found in [1, 17].) A function f is said to be cyclic (for the shift operator $Sf := zf$) in $H^2(\mathbb{D})$ if

$$(1.1) \quad [f] = \overline{\text{span}\{z^k f : k = 0, 1, 2, \dots\}}$$

is the entire space $H^2(\mathbb{D})$. It is not hard to see that to be cyclic, the function f must be zero-free in the disk, and because the polynomials are dense in $H^2(\mathbb{D})$, the function 1 is cyclic. One can then observe that for a cyclic function f , the optimal norms $\|p_n f - 1\|$ converge to zero. Because optimal approximants are unique, the rate of decay of these optimal norms can then be used to quantify and compare “how cyclic” different functions are, e.g. a fast rate of decay would mean a function is strongly cyclic.

The questions about rates of decay and locations of zeros for optimal approximants have been considered in several contexts beyond the spaces mentioned above. The sequence spaces ℓ_A^p are considered in [7, 18], [6] considers L^p and H^p , and the pairs of papers [3, 4] and [15, 16] consider approximations problems adapted to the bidisk and unit ball, respectively.

The role of the hypothesis on the zeroes of f in Theorem 1.2 is explained by the following observation: note that for polynomials p and complex numbers z , the evaluation functional $p \mapsto p(z)$ is continuous for the ℓ^2 norm if and only if $|z| < 1$. Thus, it quickly follows that a necessary condition for $c_n \rightarrow 0$ is that f be nonvanishing in the disk $|z| < 1$. It is a folklore theorem that this condition is also sufficient for $c_n \rightarrow 0$, but as far as we are aware this result from [2] is the first to obtain a quantitative bound. The result can be proved by explicit calculation in the affine linear case $f(z) = 1 - \mu z$ ($|\mu| \leq 1$), and then extended to general f by factoring and induction (though some care is required in handling repeated roots). Thus, in the free case, one may suspect that some sort of nonvanishing condition on f will be necessary for cyclicity, and indeed this is the case, for essentially the same reason that certain evaluation functionals $f \mapsto f(X)$ (now at matrix points X) are bounded for the ℓ^2 norm. The details may be found in Section 2.

Of course, in the free setting we will have no recourse to the fundamental theorem of algebra, so rather different techniques will be required. In particular, while any free polynomial f can be factored into irreducibles, these irreducibles will not in general be linear (and indeed the factorization need not be unique, e.g. $x - xyx = x(1 - yx) = (1 - xy)x$). Nonetheless, it turns out we will be able to reduce the general question to the case of (affine) linear f , though at the cost of introducing matrix coefficients. The technique rests on the (by now) well-understood technique of *linearizations* or *realizations* of free polynomials and rational functions.

1.1. Reader's guide. The next section will give a short tour of our setting in the nc (noncommutative) universe, including the row ball and the Fock space, along with the definition of an nc function (Section 2.1). We also discuss the mechanics of the approximation problem in this universe, including the definitions of the d-shift and cyclicity (Section 2.2), some algebraic notions including linearization, stable associativity, and the free zero locus (Section 2.4), and the outer spectral radius (Section 2.5). In Section 2.4 we also prove Lemma 2.13 which is a crucial reduction step in the proof of Theorem 1.1. Our results hold in the general case of polynomials with matrix coefficients, so Section 2.3 gives background on these.

Section 3 contains the proof of Theorem 1.1, as well as further questions and examples.

2. PRELIMINARIES

2.1. The Fock space \mathcal{F}_d , nc functions and nc domains.

Definition 2.1. Let $\mathfrak{x} = \{x_1, \dots, x_d\}$ be a set of freely noncommuting indeterminates and write $\mathbf{x} = (x_1, \dots, x_d)$. We write $\mathbb{C}\langle \mathfrak{x} \rangle$ for the **free algebra** in the indeterminates x_1, \dots, x_d and its elements are **free polynomials**. We define an inner product on $\mathbb{C}\langle \mathfrak{x} \rangle$ by declaring the “words” in the letters x_i to be an orthonormal system:

$$\langle x_{i_1} \dots x_{i_n}, x_{j_1} \dots x_{j_m} \rangle := \begin{cases} 1 & m = n \text{ and } i_\ell = j_\ell \text{ for } 1 \leq \ell \leq n; \\ 0 & \text{otherwise.} \end{cases}$$

This inner product induces a norm $\|\cdot\|$ and the completion of $\mathbb{C}\langle \mathfrak{x} \rangle$ in terms of this norm is \mathcal{F}_d , the **full Fock space** in d letters. We may identify \mathcal{F}_d with the space of formal power series in the words w , with square-summable coefficients:

$$\mathcal{F}_d = \left\{ \sum_{w \in \langle \mathfrak{x} \rangle} a_w w : \sum_{w \in \langle \mathfrak{x} \rangle} |a_w|^2 < \infty \right\}.$$

When $d = 1$, every power series $\sum a_n z^n$ with square-summable coefficients has radius of convergence at least 1, and hence defines a function $f(z) = \sum a_n z^n$ in the unit disk $|z| < 1$. In our free setting, it turns out that the formal power series we are considering can also be viewed as *noncommutative (nc) function* on an appropriate *nc domain*. To make this precise we begin with some definitions.

The noncommutative analog of the disk $|z| < 1$ will be the *row ball*:

Definition 2.2. Let $d \geq 1$ be an integer. For a d -tuple of $k \times k$ matrices $X = (X_1, \dots, X_d)$ we define the *row norm* of X to be

$$\|X\|_{\text{row}} := \|X_1 X_1^* + \dots + X_d X_d^*\|^{1/2}$$

(that is, the usual operator norm of the $k \times dk$ matrix $(X_1 \ X_2 \ \cdots \ X_d)$). Similarly we define the *column norm*

$$\|X\|_{col} := \|X_1^* X_1 + \cdots + X_d^* X_d\|^{1/2}.$$

We say X is a *row contraction* if $\|X\|_{row} \leq 1$ and a *strict row contraction* if $\|X\|_{row} < 1$. Column contraction and strict column contraction are defined similarly.

For fixed d and each $k \geq 1$, we denote

$$\mathfrak{B}_k^d = \{X \in M_k(\mathbb{C})^d : \|X\|_{row} < 1\}$$

and put

$$\mathfrak{B}^d := \bigsqcup_{k=1}^{\infty} \mathfrak{B}_k^d.$$

The set \mathfrak{B}^d is called the *row ball*, and we refer to \mathfrak{B}_k^d as the k^{th} “level” of the row ball.

We note that each \mathfrak{B}_k^d is an open set in the usual topology on $M_k(\mathbb{C})^d$, and \mathfrak{B}^d is closed under direct sums: if $X \in \mathfrak{B}_{k_1}^d$ and $Y \in \mathfrak{B}_{k_2}^d$ then $X \oplus Y := (X_1 \oplus Y_1, \dots, X_d \oplus Y_d) \in \mathfrak{B}_{k_1+k_2}^d$. The row ball is thus an *nc domain*.

With these definitions, we have:

Theorem 2.3 ([12, Theorem 1.1]). *If $\sum_w |a_w|^2 < \infty$ and X is a d -tuple of $k \times k$ matrices with $\|X\|_{row} < 1$, then the series*

$$f(X) := \sum_{w \in \langle \mathfrak{z} \rangle} a_w X^w$$

converges in the usual topology of $M_k(\mathbb{C})$.

Thus, every element of the Fock space \mathcal{F}_d determines a graded function on the row ball $f : \mathfrak{B}_k^d \rightarrow M_k(\mathbb{C})$. This function respects direct sums: $f(X \oplus Y) = f(X) \oplus f(Y)$ and similarities: if $S \in GL_k(\mathbb{C})$ and both $\|X\|_{row} < 1$, $\|S^{-1}XS\|_{row} < 1$, then $f(S^{-1}XS) = S^{-1}f(X)S$. This f is then an *nc function* on the row ball. (We remark that the theorem and subsequent remarks also hold with the column ball in place of the row ball).

In particular, for any X in the row ball at level k , the map from \mathcal{F}_d to $M_k(\mathbb{C})$ given by

$$f \mapsto f(X)$$

is continuous for the norm topologies on each space. (One thinks of this as “bounded point evaluation” analogous to the point evaluation $f \mapsto f(z)$ for f in the Hardy space and $|z| < 1$.) As a consequence we obtain the following proposition, relevant to our approximation problem:

Theorem 2.4. *If f is an nc polynomial and p_n is a sequence of nc polynomials such that*

$$\lim_{n \rightarrow \infty} \|p_n f - 1\|_2 = 0,$$

then f is nonsingular in the row ball.

Proof. Suppose X is in the row ball and $\det f(X) = 0$. Then also $\det(p_n(X)f(X)) = 0$ for all n , so each of the matrices $p_n(X)f(X)$ is singular, and hence these cannot converge to I . \square

2.2. The d -shift and cyclicity. From now on we think of the Fock space \mathcal{F}_d in this way, as a space of nc functions represented as convergent power series in the row ball, in analogy with spaces of holomorphic functions in the disk. Let us introduce an important class of operators acting in the space \mathcal{F}_d :

Definition 2.5. The **left d -shift** is the tuple of operators $L = (L_1, \dots, L_d)$ where each $L_i : \mathcal{F}_d \rightarrow \mathcal{F}_d$ is given by $L_i : f \mapsto x_i f$. We similarly define the **right d -shift**.

In the nc function picture, for each $f \in \mathcal{F}_d$ and each $X \in \mathfrak{B}^d$

$$(L_i f)(X) = X_i f(X).$$

One may observe that each L_i is a bounded operator for the ℓ^2 norm on coefficients. In fact, it follows quickly from definitions that each of the operators L_i is an isometry for the ℓ^2 norm, and their ranges are mutually orthogonal. These facts are summarized algebraically in the relations

$$L_i^* L_j = \delta_{ij} I.$$

For any polynomial $p \in \mathbb{C}\langle x \rangle$, the “left multiplication” operator

$$f(X) \mapsto p(X)f(X)$$

is bounded, since this operator is just $p(L_1, \dots, L_d)$. Similarly the “right multiplication operators” $f(X) \mapsto f(X)p(X)$ are bounded.

Analogous to the notion of cyclicity in Hardy space, we make the following definition:

Definition 2.6. An element $f \in \mathcal{F}_d$ is said to be **cyclic** for the d -shift if the set $\{pf : p \in \mathbb{C}\langle x \rangle\}$ is dense in \mathcal{F}_d .

For a given free polynomial f , it is evident that f is cyclic if and only if $c_n := \|p_n f - 1\|^2 \rightarrow 0$. Indeed, if f is cyclic then there is some sequence of polynomials q_n so that $\|q_n f - 1\|^2 \rightarrow 0$, so by the optimality of the p_n we must have $\|p_n f - 1\|^2 \rightarrow 0$ as well. On the other hand, if $\|p_n f - 1\|^2 \rightarrow 0$ and q is any polynomial, then since q is a bounded left multiplier, we have $\|qp_n f - q\|^2 \rightarrow 0$ as well, so that f is cyclic.

2.3. Polynomials with matrix coefficients. We will also have cause to deal with matrix-valued versions of \mathcal{F}_d . For fixed m, n we view $M_{m \times n}(\mathbb{C})$ as a Hilbert space with the tracial inner product $\langle A, B \rangle = \text{tr}(B^* A)$, so that $M_{m \times n}(\mathbb{C})$ is a Hilbert space with the Hilbert-Schmidt or *Frobenius* norm. Elements of the Hilbert space tensor product $M_{m \times n}(\mathbb{C}) \otimes \mathcal{F}_d$ may be identified with formal power series with $m \times n$ matrix coefficients

$$\sum_{w \in \langle x \rangle} A_w x^w$$

with $\sum_{w \in \langle x \rangle} \text{tr}(A_w^* A_w) < \infty$. These may in turn be identified with “matrix-valued” functions on the row ball, where at each level k for $\|X\|_{\text{row}} < 1$ we have a convergent power series in $M_{m \times n} \otimes M_k$

$$F(X) := \sum_{w \in \langle x \rangle} A_w \otimes X^w.$$

We again obtain bounded point evaluations $f \mapsto f(X)$, and the analog of Theorem 2.4 holds in the (square) matrix-valued case as well. When the series contains only finitely many terms, we of course obtain a polynomial with matrix coefficients,

with the degree of the polynomial defined the same way as in the scalar case. Observe also that a polynomial with matrix coefficients is naturally identified with a matrix having polynomial entries; in which case the degree is equal to the maximum of the degrees of the entries.

Also as before, each matrix-valued polynomial $P \in M_{m \times n}(\mathbb{C}\langle x \rangle)$ will define bounded left and right multipliers $F \mapsto PF$, $F \mapsto FP$ between matrix-valued \mathcal{F}_d spaces of appropriate sizes. We will need the following lemma about multiplication by linear polynomials:

Lemma 2.7. *Let $A = (A_1, \dots, A_d) \in M_m(\mathbb{C})^d$ and let $P(x) := A_1x_1 + \dots + A_dx_d$. Then the left multiplication operator $F \mapsto PF$ has norm equal to $\|A\|_{col}$.*

Proof. Writing L_i for the left shifts as above, the operator $F \mapsto PF$ acting in the Hilbert space tensor product $M_m \otimes \mathcal{F}_d$ is given by $\sum_{j=1}^d A_j \otimes L_j$. (Here we identify A_j with the operator $B \mapsto A_jB$ acting in the Hilbert space $M_m(\mathbb{C})$.) From definitions, the operator $B \mapsto CB$ in the tracial Hilbert space $M_m(\mathbb{C})$ has norm equal to $\|C\|_{op}$. Since the L_j are isometries with orthogonal ranges, the relations $L_i^*L_j = \delta_{ij}$ entail

$$\begin{aligned} \left\| \left(\sum_{j=1}^d A_j \otimes L_j \right) \right\|^2 &= \left\| \left(\sum_{j=1}^d A_j \otimes L_j \right)^* \left(\sum_{j=1}^d A_j \otimes L_j \right) \right\| \\ &= \left\| \sum_{j=1}^d (A_j^* A_j) \otimes I \right\| = \left\| \sum_{j=1}^d (A_j^* A_j) \right\|_{op} \\ &= \|A\|_{col}^2. \end{aligned}$$

□

2.4. Stable equivalence and linearization.

Definition 2.8. An $m \times \ell$ nc linear pencil (in d indeterminates) is an expression of the form

$$L_A(x) = A_0 + Ax,$$

where $Ax = A_1x_1 + \dots + A_dx_d$ and $A_i \in M_{m \times \ell}(\mathbb{C})$ for $i = 1, \dots, d$. If $A_0 = I$ then we call $L_A(x)$ a monic linear pencil. A matrix tuple $A = (A_1, \dots, A_d) \in M_m(\mathbb{C})^d$ is **irreducible** if A generates $M_m(\mathbb{C})$ as a (unital) algebra, i.e. $M_m(\mathbb{C}) = \{p(A) : p \in \mathbb{C}\langle x \rangle\}$. The monic linear pencil L_A is **irreducible** if A is irreducible.

The protagonist of our proof is stable associativity. Let us define that as well:

Definition 2.9. Given $A \in M_{k \times k}(\mathbb{C}\langle x \rangle)$ and $B \in M_{\ell \times \ell}(\mathbb{C}\langle x \rangle)$. We say A and B are **stably associated** if there exists $N \in \mathbb{Z}^+$ and $P, Q \in \text{GL}_N(\mathbb{C}\langle x \rangle)$ such that

$$P \begin{pmatrix} A & \\ & I \end{pmatrix} Q = \begin{pmatrix} B & \\ & I \end{pmatrix}.$$

We use the notation $A \sim_{\text{sta}} B$ for A being stably associated to B ; this is an equivalence relation.

Definition 2.10. If $f \in M_{k \times k}(\mathbb{C}\langle x \rangle)$, then

$$\mathcal{X}_n(f) = \{X \in M_n(\mathbb{C})^d : \det(f(X)) = 0\},$$

and

$$\mathcal{Z}(f) := \bigcup_{n \geq 1} \mathcal{Z}_n(f).$$

The set $\mathcal{Z}(f)$ is the **free zero locus** of f .

It also follows from the definitions that if $f \sim_{\text{sta}} g$ then $\mathcal{Z}(f) = \mathcal{Z}(g)$. The concept of stable associativity is useful to us because of the following classical fact (we refer to [10] for a discussion of this in the context of factorization of free polynomials):

Lemma 2.11 (Linearization trick). *Suppose $F \in M_k(\mathbb{C}\langle x \rangle)$ and $F(0) = I$. Then F is stably associated to a monic linear pencil. Moreover, every monic pencil is similar to a pencil of the form*

$$L_A(x) = \begin{bmatrix} L_1(x) & * & * & \dots & * \\ 0 & L_2(x) & * & \dots & * \\ 0 & 0 & L_3(x) & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & L_l(x) \end{bmatrix}$$

where for every k , $L_k(x) = I - \sum_{j=1}^d \bar{A}_j^{(k)} x_j = I - \bar{A}^{(k)} x$ for $\bar{A}^{(k)} = (\bar{A}_1^{(k)} \bar{A}_2^{(k)} \dots \bar{A}_d^{(k)})$ and either $\bar{A}^{(k)} = 0$ or irreducible.

For example, the 1×1 polynomial $F(x, y) = 1 - xy$ is stably associated to the 2×2 monic linear pencil $\begin{pmatrix} 1 & x \\ y & 1 \end{pmatrix}$ via

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 - xy & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} = \begin{pmatrix} 1 & x \\ y & 1 \end{pmatrix}.$$

(This calculation is sometimes known as “Higman’s trick.”) In fact the full lemma can be proved by iteratively applying this trick (increasing the size of the matrices at each step) to gradually reduce the degree of each monomial appearing in the entries of F . This shows F is stably associated to linear pencil, which can then be put into block upper triangular form using Burnside’s theorem.

We say $F \in M_k(\mathbb{C}\langle x \rangle)$ is **regular** if it is not a zero divisor. In particular, if $F(0) = I$ then F is regular. A regular non-invertible matrix polynomial F is an **atom** if it is not the product of two non-invertible matrices in $M_k(\mathbb{C}\langle x \rangle)$. When $k = 1$, every nonzero polynomial is regular and a nonconstant polynomial p is an atom if it is not the product of two nonconstant polynomials. If F_1, F_2 are regular matrices over $\mathbb{C}\langle x \rangle$ (not necessarily of the same size) and $F_1 \sim_{\text{sta}} F_2$, then F_1 is an atom if and only if F_2 is an atom. The following lemma from [10] relates irreducibility of the polynomial F to irreducibility of the pencil L_A .

Lemma 2.12 ([10, Lemma 4.2]). *If $F \in M_k(\mathbb{C}\langle x \rangle)$ is an atom and $F(0) = I$, then F is stably associated to an irreducible monic linear pencil L . That is, if F is an atom, then there exists an irreducible tuple $A \in M_m(\mathbb{C})^d$ such that F is stably associated to $L_A = I_m - \sum_{i=1}^d A_i x_i$.*

The following lemma shows that our rate-of-decay problem is essentially invariant under stable associativity. Precisely:

Lemma 2.13. *Suppose F and G are matrix polynomials. If F and G are stably associated, then there exists $C_1, C_2 \in \mathbb{R}$ and $D_1, D_2 \in \mathbb{N}$ such that for P_n, Q_n , the n th degree OPAs of F and G respectively, the following holds:*

$$C_1 \|P_{n+D_1}F - I\|_F^2 \leq \|Q_nG - I\|_F^2 \leq C_2 \|P_{n-D_2}F - I\|_F^2.$$

Proof. Since F and G are stably associated, there exists $N \in \mathbb{Z}^+$ and $A, B \in \text{GL}_N(\mathbb{C}\langle x \rangle)$ such that

$$(F \oplus I) = A(G \oplus I)B.$$

Since B and B^{-1} are both polynomials, it follows that the map $H \mapsto BHB^{-1}$ is bounded in \mathcal{F}_d , so there exists $C_1 \in \mathbb{R}$ such that $\|BHB^{-1}\|_F \leq C_1 \|H\|_F$ for all $H \in M_N(\mathbb{C}\langle x \rangle)$. Next, set $D_1 = \deg(A) + \deg(B)$.

Observe that

$$\begin{aligned} B \begin{pmatrix} P_n & \\ & I \end{pmatrix} A \begin{pmatrix} G & \\ & I \end{pmatrix} - \begin{pmatrix} I & \\ & I \end{pmatrix} &= B \left(\begin{pmatrix} P_n & \\ & I \end{pmatrix} A \begin{pmatrix} G & \\ & I \end{pmatrix} B - \begin{pmatrix} I & \\ & I \end{pmatrix} \right) B^{-1} \\ &= B \left(\begin{pmatrix} P_n & \\ & I \end{pmatrix} \begin{pmatrix} F & \\ & I \end{pmatrix} - \begin{pmatrix} I & \\ & I \end{pmatrix} \right) B^{-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \left\| B \begin{pmatrix} P_n & \\ & I \end{pmatrix} A \begin{pmatrix} G & \\ & I \end{pmatrix} - \begin{pmatrix} I & \\ & I \end{pmatrix} \right\|_F^2 &\leq C \left\| \begin{pmatrix} P_n & \\ & I \end{pmatrix} \begin{pmatrix} F & \\ & I \end{pmatrix} - \begin{pmatrix} I & \\ & I \end{pmatrix} \right\|_F^2 \\ &= C \|P_nF - I\|_F^2. \end{aligned}$$

The matrix $B(P_n \oplus I)A$ has a block structure of $\begin{bmatrix} Q & * \\ * & * \end{bmatrix}$, and we note that the nature of the Frobenius norm implies that

$$\|QG - I\|_F^2 \leq \left\| \begin{pmatrix} Q & * \\ * & * \end{pmatrix} \begin{pmatrix} G & \\ & I \end{pmatrix} - \begin{pmatrix} I & \\ & I \end{pmatrix} \right\|_F^2.$$

Thus,

$$\|QG - I\|_F^2 \leq C \|P_nF - I\|_F^2.$$

Finally, note that

$$\deg(Q) \leq \deg(B(P_n \oplus I)A) \leq \deg(B) + \deg(P_n) + \deg(A) = \deg(P_n) + D_1.$$

For n th degree OPA, Q_n of G , $\|Q_nG - I\|$ decreases with n , thus, we can write

$$\|Q_nG - I\|_F^2 \leq C_1 \|P_{n-D_1}F - I\|_F^2.$$

The other inequality follows by interchanging the roles of F and G , and chasing through the proof, we find that in this case we can take $D_1 = \deg(A^{-1}) + \deg(B^{-1})$. \square

2.5. The Outer Spectral Radius. The last technical tool we will need is the *outer spectral radius*, which is one of several possible notions of a “joint spectral radius” for a system of matrices $A = (A_1, \dots, A_d)$.

Definition 2.14. Let $n \in \mathbb{N}$ and $X \in M_{n \times n}(\mathbb{C})^d$. We associate to X the completely positive map on $M_{n \times n}$ defined by

$$\Psi_X(T) = \sum_{j=1}^d X_j T X_j^*.$$

The **outer spectral radius** is defined to be $\rho(X) := \lim_{k \rightarrow \infty} \|\Psi_X^k(I)\|^{1/2k}$ (that is, the square root of the spectral radius of Ψ_X viewed as a linear transformation on M_n).

Remark 2.15. Observe that if we equip $M_n(\mathbb{C})$ with tracial inner product then $\Psi_X^* = \Psi_{X^*}$, so $\rho(X) = \rho(X^*)$.

Several properties of the outer spectral radius are pointed out in [14] Section 4;

- (1) $\rho(X) = \rho(S^{-1}XS)$ for any $S \in \text{GL}_n(\mathbb{C})$, follows from Definition 2.14;
- (2) $\rho(X) = \rho(X^*)$, follows from Remark 2.15;
- (3) $\rho(X) < 1$ if and only if X is similar to an element of the row ball, this follows from Theorem 3.8, [13] and Lemma 2.5, [14];
- (4) if X is irreducible, then $\rho(X) = \min \{\|S^{-1}XS\| : S \in \text{GL}_n(\mathbb{C})\}$, follows from Lemma 2.4, [14];
- (5) if X is irreducible and $\rho(X) = 1$, then X is similar to a row coisometry, from Lemma 2.9. [14].

Since X is a row contraction if and only if X^* is a column contraction, by combining (2), (3), and (5) we get:

Lemma 2.16. *If $A = (A_1, \dots, A_d)$ is irreducible with outer spectral radius $\rho(A) \leq 1$ then A is similar to a column contraction.*

We also require the following lemma from [11]:

Lemma 2.17 ([11] Proposition 4.1). *A monic linear pencil $I - Ax$ is nonsingular in the row ball if and only if $\rho(A) \leq 1$.*

3. PROOF OF THEOREM [11]

The proof of the theorem consists of first proving it in the special case where F is a contractive, monic linear pencil (accomplished by the following lemma), then reducing the general case to that one by means of the algebraic and analytic machinery of the previous section, and finishing by induction.

Throughout the proof, for a matrix polynomial $P(x) = \sum_{w \in \langle x \rangle} P_w x^w$ we will write $\|P\|_2$ for the Hilbert space norm $(\sum_{w \in \langle x \rangle} \text{tr}(P_w^* P_w))^{1/2}$ and $\|P\|_{\text{mult}}$ for the norm of P as a left multiplier $F \mapsto PF$.

Lemma 3.1. *Assume that M is a column contraction: $\|M\|_{\text{col}} \leq 1$ and let π_n be the one variable optimal polynomial approximants for $f = 1 - z$. Then*

- (a) $\|\pi_n(Mx)\|_{\text{mult}} \lesssim n$, and
- (b) $\|\pi_n(Mx)(I - Mx) - I\|_2^2 \lesssim \frac{1}{n}$.

Proof. We begin by recalling the one variable commutative case $f(z) = 1 - z$ where the $H^2(\mathbb{D})$ optimal polynomial approximants are given in [9] by

$$(3.1) \quad \pi_n(z) = \sum_{k=0}^n \left(1 - \frac{k+1}{n+2}\right) z^k.$$

To show (a), observe that

$$(3.2) \quad \|\pi_n\|_{\infty} := \sup\{|\pi_n(z)| : |z| < 1\} = \sum_{k=0}^n \left(1 - \frac{k+1}{n+2}\right) = \frac{n+1}{2}.$$

Because M is a column contraction, by Lemma 2.7 we have that multiplication by Mx is a contraction, so we can apply von Neumann's inequality to see that

$$\|\pi_n(Mx)\|_{mult} \leq \|\pi_n\|_\infty \lesssim n$$

as needed.

For part (b), again consider the one variable case and use algebra to see that

$$\begin{aligned} \pi_n(z)(1-z) - 1 &= (1-z) \sum_{k=0}^n \left(1 - \frac{k+1}{n+2}\right) z^k - 1 \\ (3.3) \qquad \qquad &= -\frac{1}{n+2} \sum_{k=1}^{n+1} z^k. \end{aligned}$$

Again, since multiplication by Mx is contractive, we have $\|(Mx)^k\|_2 \leq \|Mx\|_2 := c$ for all $k \geq 1$. Moreover, since the $(Mx)^k$ are orthogonal (they are homogeneous polynomials of different degrees), by the Pythagorean theorem we have

$$\begin{aligned} \|\pi_n(Mx)(I - Mx) - I\|_2^2 &= \left\| -\frac{1}{n+2} \sum_{k=1}^{n+1} (Mx)^k \right\|_2^2 \\ (3.4) \qquad \qquad &= \frac{1}{(n+2)^2} \sum_{k=1}^{n+1} \|(Mx)^k\|_2^2 \leq \frac{c^2(n+1)}{(n+2)^2}, \end{aligned}$$

which gives $\|\pi_n(Mx)(I - Mx) - I\|_2^2 \lesssim \frac{1}{n}$. \square

Let us now prove Theorem 1.1. We recall the statement:

Theorem (Theorem 1.1). *Let $F \in M_k(\mathbb{C}) \otimes \mathbb{C}\langle x \rangle$ be a $k \times k$ -matrix valued free polynomial, let (P_n) denote the sequence of optimal polynomial approximants, and put*

$$c_n := \|P_n(x)F(x) - I\|_2^2.$$

Then:

- (1) $c_n \rightarrow 0$ if and only if F is nonsingular in the row ball.
- (2) If F is nonsingular in the row ball, then

$$c_n \lesssim \frac{1}{n^p}$$

for some $p > 0$ depending on F . In particular, if F is a product of ℓ atomic factors, we can take $p = \frac{1}{3\ell-1}$.

Proof of Theorem 1.1. By Theorem 2.4 (and the remarks in Section 2.3), if $c_n \rightarrow 0$ then necessarily F is nonsingular in the row ball. The other direction of item (1) will evidently follow from item (2). Let $F \in M_k(\mathbb{C}\langle x \rangle)$ and assume it is nonsingular on \mathfrak{B}^d . We can assume, without loss of generality, that $F(0) = I$. Then by lemma 2.11 we have that F is stably associated to a monic linear pencil, say, $L_A(x)$ and

$$L_A(x) \sim \begin{bmatrix} L_1(x) & * & * & \dots & * \\ 0 & L_2(x) & * & \dots & * \\ 0 & 0 & L_3(x) & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & L_\ell(x) \end{bmatrix}$$

where for each $k = 1, \dots, \ell$ the linear pencil $L_k(x) = I - \bar{A}^{(k)}x$ has $\bar{A}^{(k)}$ either 0 or irreducible. By Remark 5.2 and Section 6.2 of [10], the number of blocks ℓ will be equal to the number of atomic factors of F .

By hypothesis, our original F was nonsingular in the row ball, and since stable associativity preserves the zero locus, the pencil L_A , and hence also each of the $L_{\bar{A}^{(k)}}$, is also nonsingular in the row ball. Thus by Lemma 2.17, each of the pencils $\bar{A}^{(k)}$ has outer spectral radius at most 1. It follows, in turn, by Lemma 2.16, that each $\bar{A}^{(k)}$ is similar to a column contractive pencil $M^{(k)}$.

In summary, we conclude that our F is stably associated to a monic linear pencil of the following block upper-triangular form:

$$(3.5) \quad F \sim_{\text{sta}} \begin{bmatrix} I - M^{(1)}(x) & * & * & \dots & * \\ 0 & I - M^{(2)}(x) & * & \dots & * \\ 0 & 0 & I - M^{(3)}(x) & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I - M^{(\ell)}(x) \end{bmatrix}$$

where each $M^{(k)}$ is a column contraction. Note that the off-diagonal entries are arbitrary linear pencils $Y_{ij}(x)$ (genuinely linear, with no constant term).

We now consider a monic linear pencil $L_A(x)$ of the form appearing in the right-hand side of (3.5), in $\ell \times \ell$ block upper triangular form, with contractive pencils down the diagonal. We prove the following

Claim 1. For each such pencil $L_A(x)$, there exists a sequence of matrix polynomials $\sigma_{n,\ell}(x)$, $n = 1, 2, \dots$, satisfying the following conditions:

- (1) $\deg \sigma_{n,\ell} \leq (\ell - 1) + n + n^3 + \dots + n^{3^{\ell-1}} \lesssim n^{3^{\ell-1}}$,
- (2) $\|\sigma_{n,\ell}\|_{\text{mult}} \lesssim n^{(1+3+3^2+\dots+3^{\ell-1})}$, and
- (3) $\|\sigma_{n,\ell}(x)(I - Ax) - I\|_2^2 \lesssim \frac{1}{n}$

for all n , where the implied constants are allowed to depend on ℓ but not on n .

Assuming the claim for the moment, we conclude that for this fixed A (which fixes ℓ) there exist, for each n , matrix polynomials $\sigma_N := \sigma_{n,\ell}$ of degree $N \lesssim n^{3^{\ell-1}}$ such that

$$\|\sigma_N(x)(I - Ax) - I\|_2^2 \lesssim \frac{1}{n}.$$

It follows that for the optimal approximants Q_N at this same degree N , we will also have $\|Q_N(x)(I - Ax) - I\|_2^2 \lesssim \frac{1}{n} \lesssim \frac{1}{N^p}$ (where $p = \frac{1}{3^{\ell-1}}$), and thus (since these quantities decrease as the degree N increases), we conclude that for all degrees n

$$\|Q_n(x)(I - Ax) - I\|_2^2 \lesssim \frac{1}{n^p},$$

where $p = \frac{1}{3^{\ell-1}}$. Finally, since $I - Ax$ was stably equivalent to our original matrix polynomial F , we conclude by Lemma 2.13 that the optimal approximants P_n of the original F achieve the same rate of decay.

To complete the proof, it remains to prove the claim, which we will do via induction on ℓ .

In the case $\ell = 1$, our pencil $L_A(x)$ has the form $1 - M^{(1)}x$ for some column contractive irreducible pencil $M^{(1)}$, so we can take $\sigma_{n,1}$ to be the polynomials provided by Lemma 3.1, namely $\sigma_{n,1}(x) = \pi_n(M^{(1)}x)$ where π_n are the 1-variable OPAs for the polynomial $f(z) = 1 - z$.

Now suppose the claim is proved for ℓ , let us prove it for $\ell + 1$. We write our pencil in the form

$$L_A(x) = \begin{bmatrix} I - \tilde{A}x & Yx \\ 0 & I - M^{(\ell+1)}x \end{bmatrix},$$

where \tilde{A} has at most ℓ irreducible blocks. We will choose $\sigma_{n,\ell+1}$ of the form

$$\sigma_{n,\ell+1}(x) = \begin{bmatrix} \sigma_{n,\ell}(x) & r(x) \\ 0 & q(x) \end{bmatrix},$$

where $\sigma_{n,\ell}$ is chosen by applying the induction hypothesis to the block $\ell \times \ell$ pencil $I - \tilde{A}x$.

We choose $q(x)$ to be $\pi_n(M^{(\ell+1)}(x))$, as in Lemma 3.1, so $\deg(q) \leq n$ and $\|q\|_{mult} \lesssim n$ by the lemma.

Finally we define

$$r(x) = -\sigma_{n,\ell}(x) \cdot (Yx) \cdot \pi_N(M^{(\ell+1)}x),$$

where again π_N is the one-variable OPA for $f(z) = 1 - z$, at degree $N = n^{3^\ell}$. Let us now verify the claims (1)–(3). First, the degree of $\sigma_{n,\ell+1}$ will be the maximum of the degrees of $\sigma_{n,\ell}$, $r(x)$, and $q(x)$, and we see by inspection that $r(x)$ has the largest degree, which is

$$(3.6) \quad \deg r(x) = \deg(\sigma_{n,\ell}) + 1 + n^{3^\ell}$$

$$(3.7) \quad \leq \ell + n + n^3 + \cdots + n^{3^\ell},$$

which proves (1). For (2), observe that the $\|\cdot\|_{mult}$ norm of $\sigma_{n,\ell+1}$ is comparable to the maximum of the $\|\cdot\|_{mult}$ norms of its entries; by the induction hypothesis, Lemma 3.1, and the definition of $r(x)$, we have

$$\begin{aligned} \|r(x)\|_{mult} &\leq \|\sigma_{n,\ell}\|_{mult} \|Yx\|_{mult} \|\pi_N(M^{(\ell+1)}x)\|_{mult} \\ &\lesssim n^{1+3+\cdots+3^\ell}. \end{aligned}$$

Since this is larger than $\|\sigma_{n,\ell}\|_{mult}$ and $\|q(x)\|_{mult}$, (2) is proved.

Finally for (3), we have

$$\sigma_{n,\ell+1}(x)(I - A^{(\ell+1)}x) = \begin{bmatrix} \sigma_{n,\ell}(x)(I - A^{(\ell)}x) & \sigma_{n,\ell}(x) \cdot Yx + r(x)(I - M^{(\ell+1)}x) \\ 0 & q(x)(I - M^{(\ell+1)}x) \end{bmatrix}.$$

We subtract $\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ from this block matrix. Applying the induction hypothesis we can estimate the $(1, 1)$ block as

$$\|\sigma_{n,\ell}(x)(I - A^{(\ell)}x) - I\|_2^2 \lesssim \frac{1}{n}.$$

Since $M^{(\ell+1)}$ is a column contraction, we can apply Lemma 3.1 to estimate the $(2, 2)$ block as

$$\|q(x)(I - M^{(\ell+1)}x) - I\|_2^2 = \|\pi_n(M^{(\ell+1)}x)(I - M^{(\ell+1)}x) - I\|_2^2 \lesssim \frac{1}{n}.$$

In the off-diagonal block, we have by the definition of $r(x)$, the induction hypothesis, and Lemma 3.1 (applied at degree $N = n^{3^\ell}$),

$$\begin{aligned} \|\sigma_{n,\ell}(x) \cdot Yx + r(x)(I - M^{(\ell+1)}x)\|_2^2 &\leq \\ &\leq \|\sigma_{n,\ell}\|_{mult}^2 \|Yx\|_{mult}^2 \|I - \pi_N(M^{(\ell+1)}x)(I - M^{(\ell+1)}x)\|_2^2 \\ &\lesssim n^{2(1+3+\dots+3^{(\ell-1)})} \cdot \frac{1}{n^{3^\ell}} \\ &= n^{3^\ell-1} \cdot \frac{1}{n^{3^\ell}} \\ &= \frac{1}{n}. \end{aligned}$$

This completes the proof of the claim. \square

As an immediate consequence of the theorem (and the remarks following Definition 2.6), we obtain a new proof of the following fact, first proved in [11]:

Corollary 3.2. *Let f be an nc polynomial with scalar coefficients. If f is nonvanishing in the row ball, then f is cyclic for the d -shift.*

3.1. Remarks and questions. In general, the exponent $p = \frac{1}{3^\ell-1}$ appearing in Theorem 3.1 (where ℓ is the number of irreducible factors of F) is not sharp. Already in one variable, the bound $c_n \lesssim \frac{1}{n}$ holds regardless of the number of factors, as was already mentioned in the introduction.

It is possible to run the proof presented here in the one-variable case, making more careful choices of the polynomials constructed in the induction step, to recover the uniform $\frac{1}{n}$ bound in that case, but this relies crucially on the commutativity of the polynomial ring $\mathbb{C}[z]$.

At the moment, we do not know if the same is true in the free setting, or even if there is a uniform power p so that $\|P_n F - I\|_2^2 \lesssim \frac{1}{n^p}$ independently of F . Also, observe that our proof does not actually construct the optimal approximants for any given f —we only construct a sequence p_n obeying the claimed bounds, so that the optimal approximants must be at least as good. We have computed the optimal approximants (with computer assistance) in some simple examples, but the results do not appear enlightening and we have not included them.

A potential note we could add: In Bénéteau et al [3], they consider the case of separable functions $f(z_1, z_2) = g(z_1)h(z_2)$ on the bidisk, and show in Theorem 2.6 that the rate of decay for OPAs of a separable function is approximately $1/n$ conditional on the locations of the zeros of the one variable g, h . However, this factorization relies on the structure of the bidisk, which we unfortunately do not have access to when working on the row ball.

The sharp value of p seems difficult to calculate even for relatively simple cases. For example, for the free polynomial in two variables $f(x, y) = (1-x)(1-y)$ one can construct polynomials p_n (by essentially the method used in the induction step of the proof, but making more careful choices of σ_n, r_n) for which

$$\|p_n(x, y)(1-x)(1-y) - 1\|_2^2 \lesssim \frac{1}{n^{1/2}},$$

which improves on the bound $\frac{1}{n^{1/3}}$ provided by the theorem, but we do not know if this value $p = 1/2$ is sharp. (Computing sharp values seems to be similarly difficult

in the commutative multivariable case, for example the polynomial $2 - z - w$ on the bidisk, see [16, Section 6].)

Finally, it would be of interest to know (1) whether the methods used here can be modified to handle the case of nc rational functions F , and (2) what can be said in the case of non-square polynomial matrix functions F .

REFERENCES

- [1] Catherine Bénéteau and Raymond Centner, *A survey of optimal polynomial approximants, applications to digital filter design, and related open problems*, Complex Anal. Synerg. **7** (2021), no. 2, Paper No. 16, 12, DOI 10.1007/s40627-021-00068-x. MR4244844
- [2] Catherine Bénéteau, Alberto A. Condori, Constanze Liaw, Daniel Seco, and Alan A. Sola, *Cyclicity in Dirichlet-type spaces and extremal polynomials*, J. Anal. Math. **126** (2015), 259–286, DOI 10.1007/s11854-015-0017-1. MR3358033
- [3] Catherine Bénéteau, Alberto A. Condori, Constanze Liaw, Daniel Seco, and Alan A. Sola, *Cyclicity in Dirichlet-type spaces and extremal polynomials II: Functions on the bidisk*, Pacific J. Math. **276** (2015), no. 1, 35–58, DOI 10.2140/pjm.2015.276.35. MR3366027
- [4] Catherine Bénéteau, Greg Knese, Łukasz Kosiński, Constanze Liaw, Daniel Seco, and Alan Sola, *Cyclic polynomials in two variables*, Trans. Amer. Math. Soc. **368** (2016), no. 12, 8737–8754, DOI 10.1090/tran6689. MR3551587
- [5] Leon Brown and Allen L. Shields, *Cyclic vectors in the Dirichlet space*, Trans. Amer. Math. Soc. **285** (1984), no. 1, 269–303, DOI 10.2307/1999483. MR748841
- [6] Raymond Centner, *Optimal polynomial approximants in L^p* , Concr. Oper. **9** (2022), no. 1, 96–113, DOI 10.1515/conop-2022-0131. MR4448804
- [7] Raymond Cheng, William T. Ross, and Daniel Seco, *Zeros of optimal polynomial approximants in ℓ_A^p* . part A, Adv. Math. **404** (2022), no. part A, Paper No. 108396, 39, DOI 10.1016/j.aim.2022.108396. MR4406262
- [8] Omar El-Fallah, Karim Kellay, Javad Mashreghi, and Thomas Ransford, *A primer on the Dirichlet space*, Cambridge Tracts in Mathematics, vol. 203, Cambridge University Press, Cambridge, 2014. MR3185375
- [9] Emmanuel Fricain, Javad Mashreghi, and Daniel Seco, *Cyclicity in reproducing kernel Hilbert spaces of analytic functions*, Comput. Methods Funct. Theory **14** (2014), no. 4, 665–680, DOI 10.1007/s40315-014-0073-z. MR3274894
- [10] J. William Helton, Igor Klep, and Jurij Volčič, *Geometry of free loci and factorization of non-commutative polynomials*, Adv. Math. **331** (2018), 589–626, DOI 10.1016/j.aim.2018.04.007. MR3804686
- [11] Michael T. Jury, Robert T. W. Martin, and Eli Shamovich, *Non-commutative rational functions in the full Fock space*, Trans. Amer. Math. Soc. **374** (2021), no. 9, 6727–6749, DOI 10.1090/tran/8418. MR4302175
- [12] Gelu Popescu, *Free holomorphic functions on the unit ball of $B(\mathcal{H})^n$* , J. Funct. Anal. **241** (2006), no. 1, 268–333, DOI 10.1016/j.jfa.2006.07.004. MR2264252
- [13] Gelu Popescu, *Similarity problems in noncommutative polydomains*, J. Funct. Anal. **267** (2014), no. 11, 4446–4498, DOI 10.1016/j.jfa.2014.09.023. MR3269883
- [14] Guy Salomon, Orr M. Shalit, and Eli Shamovich, *Algebras of noncommutative functions on subvarieties of the noncommutative ball: the bounded and completely bounded isomorphism problem*, J. Funct. Anal. **278** (2020), no. 7, 108427, 54, DOI 10.1016/j.jfa.2019.108427. MR4053626
- [15] Meredith Sargent and Alan A. Sola, *Optimal approximants and orthogonal polynomials in several variables II: Families of polynomials in the unit ball*, Proc. Amer. Math. Soc. **149** (2021), no. 12, 5321–5330, DOI 10.1090/proc/15641. MR4327435
- [16] Meredith Sargent and Alan A. Sola, *Optimal approximants and orthogonal polynomials in several variables*, Canad. J. Math. **74** (2022), no. 2, 428–456, DOI 10.4153/S0008414X20000826. MR4410997
- [17] Daniel Seco, *Some problems on optimal approximates*, Recent progress on operator theory and approximation in spaces of analytic functions, Contemp. Math., vol. 679, Amer. Math. Soc., Providence, RI, 2016, pp. 193–205, DOI 10.1090/conm/679. MR3589676

- [18] Daniel Seco and Roberto Téllez, *Polynomial approach to cyclicity for weighted ℓ_A^p* , Banach J. Math. Anal. **15** (2021), no. 1, Paper No. 1, 16, DOI 10.1007/s43037-020-00085-8. MR4159768

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA 32611
Email address: pa8@williams.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH FLORIDA, TAMPA, FLORIDA 33620
Email address: mla026@bucknell.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA 32611
Email address: mjury@ufl.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MANITOBA, WINNIPEG, MANITOBA R3T 2N2,
CANADA
Email address: meredithsargent@gmail.com