BIANALYTIC FREE MAPS BETWEEN SPECTRAHEDRA AND SPECTRABALLS

J. WILLIAM HELTON[1], IGOR KLEP[2], SCOTT MCCULLOUGH[3], AND JURIJ VOLČIČ[4]

Abstract. Linear matrix inequalities (LMIs) arise in numerous areas such as systems engineering, semidefinite programming and real algebraic geometry. LMIs with (dimension free) matrix unknowns, called free LMIs, are central to the theories of completely positive maps and operator algebras, operator systems and spaces and serve as the paradigm for convex engineering problems prescribed entirely by signal flow diagrams.

The solution set of a free LMI is called a free spectrahedron. In this article, the bianalytic maps from a very general class of *ball-like* free spectrahedra to arbitrary free spectrahedra are explicitly characterized and seen to have an elegant and highly algebraic form. This result depends on a novel free Nullstellensatz, established only after new tools in free analysis are developed and applied to obtain fine detail, geometric in nature locally and algebraic in nature globally, about the boundary of ball-like free spectrahedra. A much cleaner reformulation of existing results on bianalytic mappings between free spectrahedra is a corollary.

1. Introduction

The main results of the article are stated in this introduction. Following a review of basic definitions including that of free spectrahedron, spectraball, free analytic mapping and convexotonic map in Subsection 1.1, the bianalytic mappings from a spectraball to a free spectrahedron are characterized as convexotonic maps in Theorem 1.1 of Subsection 1.2. Essential to this is a concomitant Nullstellensatz (Proposition 1.3), whose proof requires detailed information, both of a local and global nature, about the boundary of a spectraball. This information is collected in Sections 3 and 4 and, as a byproduct, we obtain Theorem 1.4. It is an elegant restatement of the main result from [AHKM18]

Date: April 25, 2018.

2010 *Mathematics Subject Classification.* 47L25, 32H02, 13J30 (Primary); 14P10, 52A05, 46L07 (Secondary).

Key words and phrases. bianalytic map, birational map, linear matrix inequality (LMI), spectrahedron, convex set, Positivstellensatz, free analysis, real algebraic geometry.

[1]Research supported by the NSF grant DMS 1500835.

[2]Supported by the Marsden Fund Council of the Royal Society of New Zealand. Partially supported by the Slovenian Research Agency grants J1-8132, N1-0057 and P1-0222.

[3]Research supported by the NSF grant DMS-1361501.

[4]Research supported by the Deutsche Forschungsgemeinschaft (DFG) Grant No. SCHW 1723/1-1.

characterizing bianalytic maps between free spectrahedra under, what we show here are, canonical irreducibility and minimality hypotheses on the free spectrahedra. With similar irreducibility and minimality hypotheses, triples $(p, \mathcal{B}, \mathcal{D})$, where $\mathcal{B}$ is a spectraball, $\mathcal{D}$ is a free spectrahedron and p is a bianalytic map from $\mathcal{B}$ to $\mathcal{D}$, are completely classified in Theorem 1.6 of Subsection 1.4. The introduction concludes with Subsection 1.5 whose results, Proposition 1.7 and its Corollary 1.8, explain how convexotonic maps naturally arise as maps from a spectraball to a free spectrahedron.

1.1. **Definitions.** Fix a positive integer g. Given a positive integer n, let $M_n(\mathbb{C})^g$ denote the g-tuples $X = (X_1, \dots, X_g)$ of $n \times n$ matrices with entries from $\mathbb{C}$. For $A \in M_d(\mathbb{C})^g$, let $L_A(x, y)$ denote the **monic linear pencil**

$$L_A(x, y) = I + \sum A_j x_j + \sum A_j^* y_j,$$

and let

$$L_A^{\mathrm{re}}(x) = I + \sum A_j x_j + \sum A_j^* x_j^*$$

denote the corresponding **hermitian monic linear pencil**. The set $\mathcal{D}_A(1)$ consisting of $x \in \mathbb{C}^g$ such that $L_A^{\mathrm{re}}(x) \succeq 0$ is a **spectrahedron**. Here $T \succeq 0$ indicates the selfadjoint matrix T is positive semidefinite. Spectrahedra are basic objects in a number of areas of mathematics, e.g. semidefinite programming, convex optimization and in real algebraic geometry [BPR13]. They also figure prominently in determinantal representations [Brä11, GK-VVW16, NT12, Vin93], in the solution of the Kadison-Singer paving conjecture [MSS15], the solution of the Lax conjecture [HV07], and in systems engineering [BGFB94, SIG96].

For $A \in M_{d\times e}(\mathbb{C})^g$, the **homogeneous linear pencil** $\Lambda_A(x) = \sum_j A_j x_j$ evaluates at $X \in M_n(\mathbb{C})^g$ as

$$\Lambda_A(X) = \sum A_j \otimes X_j \in M_{d\times e}(\mathbb{C}) \otimes M_n(\mathbb{C}).$$

In the case A is square ($d = e$), the hermitian monic linear pencil L_A^{re} evaluates at X as

$$L_A^{\mathrm{re}}(X) = I + \Lambda_A(X) + \Lambda_A(X)^* = I + \sum A_j \otimes X_j + \sum A_j^* \otimes X_j^*.$$

If $Y \in M_n(\mathbb{C})^g$, then $L_A(X, Y) = I + \Lambda_A(X) + \Lambda_{A^*}(Y)$.

The **free spectrahedron** determined by A is the sequence of sets $\mathcal{D}_A = (\mathcal{D}_A(n))$, where

$$\mathcal{D}_A(n) = \{X \in M_n(\mathbb{C})^g : L_A^{\mathrm{re}}(X) \succeq 0\}.$$

Free spectrahedra arise naturally in applications such as systems engineering [dOHMP09] and in the theories of matrix convex sets, operator algebras and operator spaces and completely positive maps [EW97, HKM17, Pau02, PSS18]. They also provide tractable useful relaxations for spectrahedral inclusion problems that arise in semidefinite programming and engineering applications such as the matrix cube problem [B-TN02, HKMS, DDOSS17].

Given a tuple $E \in M_{d\times e}(\mathbb{C})^g$, the set

$$\mathcal{B}_E = \{X : \|\Lambda_E(X)\| \leq 1\}$$

is a **spectraball** [EHKM17, BMV]. Spectraballs are special cases of free spectrahedra. Indeed, it is readily seen that

$$\mathcal{B}_E = \mathcal{D}_{\left(\begin{smallmatrix} 0 & E \\ 0 & 0 \end{smallmatrix}\right)}.$$

Further, by [BMV, EHKM17], spectraballs are exactly free circular spectrahedra[1].

1.1.1. *Free analytic functions.* Let $M(\mathbb{C})^g$ denote the sequence $(M_n(\mathbb{C})^g)_n$. A **subset** Γ of $M(\mathbb{C})^g$ is a sequence $(\Gamma_n)_n$ where $\Gamma_n \subseteq M_n(\mathbb{C})^g$. (Sometimes we write $\Gamma(n)$ in place of Γ_n.) The subset Γ is a **free set** if it is closed under direct sums and simultaneous unitary similarity. Examples of such sets include spectraballs and free spectrahedra introduced above. We say the free set $\Gamma = (\Gamma_n)_n$ is **open** if each Γ_n is open. Generally adjectives are applied levelwise to free sets unless noted otherwise.

A **free function** $f : \Gamma \to M(\mathbb{C})$ is a sequence of functions $f_n : \Gamma_n \to M_n(\mathbb{C})$ that **respects intertwining**; that is, if $X \in \Gamma_n$, $Y \in \Gamma_m$, $T : \mathbb{C}^m \to \mathbb{C}^n$, and

$$XT = (X_1T, \dots, X_gT) = (TY_1, \dots, TY_g) = TY,$$

then $f_n(X)T = Tf_m(Y)$. In the case Γ is open, f is **free analytic** if each f_n is analytic in the ordinary sense. We refer the reader to [Voi04, KVV14, AM14] for a fuller discussion of free sets and functions. For further results, not already cited, on free bianalytic and proper free analytic maps see [Pop10, MS08, KŠ17] and the references therein.

Among the results in this article, we characterize free bianalytic maps $p : \mathcal{B}_E \to \mathcal{D}_B$ and, under minimality and irreducibility conditions on A and B, the triples (p, A, B) such that $p : \mathcal{D}_A \to \mathcal{D}_B$ is bianalytic. In both cases the mappings are birational; that is, the inverse of p is also a free rational function in that its components p^j belong to the skew field $\mathbb{C}\langle\!\!\langle x \rangle\!\!\rangle$ of free rational functions. In fact, such maps p form a very restricted class of birational functions we call convexotonic.

Based on the results of [KVV09, Theorem 3.1] and [Vol17, Theorem 3.5] a **free rational function regular at** 0 can, for the purposes of this article, be defined with

[1] A spectrahedron $\mathcal{D}_A$ is free circular if, for each n and $n \times n$ unitary U, $X \in \mathcal{D}_A(n)$ if and only if $UX = (UX_1, \dots, UX_n) \in \mathcal{D}_A(n)$.

minimal overhead as an expression of the form

$$r(x) = c^* \big(I - \Lambda_S(x)\big)^{-1} b,$$

where, for some positive integer s, we have $S \in M_s(\mathbb{C})^g$ and $b, c \in \mathbb{C}^s$. The expression r is known as a realization. Realizations are easy to manipulate and a powerful tool as developed in the series of papers [BGM05, BGM06a, BGM06b] of Ball-Groenewald-Malakorn; see also [Coh95, BR11]. The realization r is evaluated in the obvious fashion on a tuple $X \in M_n(\mathbb{C})^g$ as long as $I - \Lambda_S(X)$ is invertible. Importantly, free rational functions are indeed free analytic functions.

1.1.2. *Convexotonic maps.* A g tuple of $g \times g$ matrices $\Xi = (\Xi_1, \ldots, \Xi_g) \in M_g(\mathbb{C})^g$ satisfying

$$\Xi_k \Xi_j = \sum_{s=1}^{g} (\Xi_j)_{k,s} \Xi_s, \tag{1.1}$$

for each $1 \le j, k \le g$, is **convexotonic**. We say the rational mappings $p = \begin{pmatrix} p^1 & \cdots & p^g \end{pmatrix}$ and $q = \begin{pmatrix} q^1 & \cdots & q^g \end{pmatrix}$ whose entries have the form

$$p^i(x) = \sum_j x_j \left(I - \Lambda_\Xi(x)\right)^{-1}_{j,i} \qquad \text{and} \qquad q^i(x) = \sum_j x_j \left(I + \Lambda_\Xi(x)\right)^{-1}_{j,i},$$

that is, in row form,

$$p(x) = x(I - \Lambda_\Xi(x))^{-1} \qquad \text{and} \qquad q = x(I + \Lambda_\Xi(x))^{-1}$$

are **convexotonic**. It turns out (see [AHKM18, Proposition 6.2]) the mappings p and q are inverses of one another. Hence they are birational maps.

Convexotonic tuples arise naturally as the structure constants of a finite dimensional algebra; cf. [AHKM18, Proposition 6.3]. If $\{J_1, \ldots, J_h\} \subseteq M_r(\mathbb{C})$ is linearly independent and spans an algebra, then, by Lemma 3.7 below, there is a uniquely determined convexotonic tuple $\Xi = (\Xi_1, \ldots, \Xi_h) \in M_h(\mathbb{C})^h$ such that

$$J_k J_j = \sum_{s=1}^{h} (\Xi_j)_{k,s} J_s. \tag{1.2}$$

Finally, note that a convexotonic g-tuple of $g \times g$ matrices Ξ spans an at most g-dimensional algebra in view of equation (1.1). Conversely, if $\Xi \in M_h(\mathbb{C})^h$ is convexotonic, then there exists a positive integer r, an h-dimensional algebra $\mathcal{A} \subseteq M_r(\mathbb{C})$ with (vector space) basis $\{J_1, \ldots, J_h\}$ such that equation (1.2) holds [AHKM18, Section 1.2.1].

1.2. **Main result on maps between spectraballs and free spectrahedra.** Recall a mapping between topological spaces is **proper** if the inverse image of each compact sets is compact. Thus, for free open sets $\mathcal{U} \subseteq M(\mathbb{C})^g$ and $\mathcal{V} \subseteq M(\mathbb{C})^h$, a free mapping $f : \mathcal{U} \to \mathcal{V}$ is proper if each $f_n : \mathcal{U}_n \to \mathcal{V}_n$ is proper. Given subsets $\Omega \subseteq \mathbb{C}^g$ and $\Delta \subseteq \mathbb{C}^h$ (that are not necessarily closed), a map $\psi : \Omega \to \Delta$ **maps boundary to boundary** if $\psi(z^j) \to \partial\Delta$ for each sequence $\Omega \ni z^j \to \partial\Omega$. If ψ is proper, then it maps boundary to boundary. In the case that Ω and Δ are bounded domains, proper is equivalent to mapping boundary to boundary [Kra92].

Let $\mathrm{int}(\mathcal{U})$ denote the interior of the set $\mathcal{U}$. We can now state our principal result on bianalytic mappings from a spectraball onto a free spectrahedron.

Theorem 1.1. *Suppose $E \in M_{d\times e}(\mathbb{C})^g$, $A \in M_r(\mathbb{C})^g$. If $\{E_1, \dots, E_g\}$ and $\{A_1, \dots, A_g\}$ are linearly independent and $f : \mathrm{int}(\mathcal{B}_E) \to \mathrm{int}(\mathcal{D}_A)$ is a proper free analytic map, then, up to affine linear change of variables, f is convexotonic.*

Proof. See Section 3.3. ∎

Remark 1.2. By [HKM11b], if E and A are as in Theorem 1.1 and f is proper from $\mathrm{int}(\mathcal{B}_E)$ to $\mathrm{int}(\mathcal{D}_A)$, then f is bianalytic. The conclusion of Theorem 1.1 is, in part, that f is in fact birational. This phenomenon is also encountered frequently in rigidity theory in several complex variables, cf. [For93]. □

1.2.1. *A Nullstellensatz.* We say $E \in M_{d\times e}(\mathbb{C})^g$ is **ball-minimal** (for $\mathcal{B}_E$) if there does not exist E' of size $d' \times e'$ with $d' + e' < d + e$ such that $\mathcal{B}_E = \mathcal{B}_{E'}$. In fact, in this case, $d \le d'$ and $e \le e'$ for any E' defining the spectraball $\mathcal{B}_E$ by Lemma 3.1(7).[2] A monic linear pencil $L_A = L_A(x, y)$ of size e is **indecomposable** if its coefficients $\{A_1, \dots, A_g, A_1^*, \dots, A_g^*\}$ generate $M_e(\mathbb{C})$ as a $\mathbb{C}$-algebra.[3] We say L_A or L_A^{re} is **minimal** for a free spectrahedron $\mathcal{D}$ if $\mathcal{D} = \mathcal{D}_A$ and if for any other $B \in M_{e'}(\mathbb{C})^g$ satisfying $\mathcal{D} = \mathcal{D}_B$ it follows that $e' \ge e$. In this case, there is a k and **irredundant** indecomposable monic linear pencils L_{A^j} such that

$$L_A = \bigoplus_{j=1}^{k} L_{A^j} = L_{\bigoplus_{j=1}^k A^j},$$

where irredundant means, $\bigcap_{j\ne\ell} \mathcal{D}_{A_j} \not\subseteq \mathcal{D}_{A_\ell}$ for all ℓ. Here the direct sum is in the sense of an orthogonal direct sum decomposition of the space that A acts upon. (Zalar [Zal17] (see also [HKM13]) establishes this result over the reals, but the proofs work (and are easier) over $\mathbb{C}$; it can also be deduced from the results in [KV17] and [HKV].) A minimal L_A for $\mathcal{D}_A$ is unique up to unitary equivalence [HKM13, Zal17]. Theorem 1.1 uses the

[2]See also [HKM11a, Section 5 or Lemma 1.2].

[3]Previously, in [KV17] such pencils were called irreducible.

following Nullstellensatz whose proof depends on Cohn's [Coh95] theory of matrices over the free algebra $\mathbb{C}<x>$ of free (noncommutative) polynomials.

Proposition 1.3. *Suppose $E = (E_1, \dots, E_g) \in M_{d\times e}(\mathbb{C})^g$ is ball-minimal and $V \in \mathbb{C}<x>^{\ell\times e}$ is a (rectangular) matrix polynomial. If for each positive integer n and $(Y, \gamma) \in M_g(\mathbb{C})^n \times (\mathbb{C}^e \otimes \mathbb{C}^n)$ such that $\|\Lambda_E(Y)\| = 1$ and $\|\Lambda_E(Y)\gamma\| = \|\gamma\|$, it follows that $V(Y)\gamma = 0$, then $V = 0$.*

The same conclusion holds for $\mathcal{D}_A$ if L_A is minimal for $\mathcal{D}_A$. Explicitly, if for each (Y, γ) such that $L_A(Y) \succeq 0$ and $L_A(Y)\gamma = 0$ it follows that $V(Y)\gamma = 0$, then $V = 0$.

Proof. See Subsection 3.2. ∎

1.3. **Main result on maps between free spectrahedra.** The article [AHKM18] characterizes the triples (p, A, B) such that $p : \mathcal{D}_A \to \mathcal{D}_B$ is bianalytic under certain awkward irreducibility hypotheses (cf. [AHKM18, §7]). These hypotheses are generic in the sense of algebraic geometry. Here (Section 3.1) we deepen our understanding of these conditions leading to a much cleaner formulation of the main result of [AHKM18] holding with improved hypotheses.

For a tuple of rectangular matrices $E = (E_1, \dots, E_g) \in M_{d\times e}(\mathbb{C})^g$ denote

$$Q_E(x, y) := I - \Lambda_{E^*}(y)\Lambda_E(x), \qquad \mathbb{L}_E(x, y) := \begin{pmatrix} I & \Lambda_E(x) \\ \Lambda_{E^*}(y) & I \end{pmatrix},$$

$$\ker E := \bigcap_{j=1}^{g} \ker E_j.$$

Thus $\mathbb{L}_E(x, y) = L_F(x, y)$ where

$$F = \begin{pmatrix} 0 & E \\ 0 & 0 \end{pmatrix}.$$

We also let $\mathbb{L}_E^{\mathrm{re}}$ denote the hermitian monic linear pencil,

$$\mathbb{L}_E^{\mathrm{re}}(x) := \mathbb{L}_E(x, x^*) = L_F(x, x^*) = L_F^{\mathrm{re}}(x)$$

and likewise let

$$Q_E^{\mathrm{re}}(x) = Q_E(x, x^*).$$

Let $\mathcal{D}_{Q_E} = \{X : Q_E^{\mathrm{re}}(X) \succeq 0\}$ and observe $\mathcal{D}_{Q_E} = \mathcal{B}_E = \mathcal{D}_{\mathbb{L}_E^{\mathrm{re}}} := \{X : \mathbb{L}_E(X, X^*) \succeq 0\} = \mathcal{D}_F$. Finally, for a monic linear pencil L_A, let

$$\mathcal{Z}_{L_A} = \{(X, Y) : \det(L_A(X, Y)) = 0\}, \quad \mathcal{Z}_{L_A}^{\mathrm{re}} = \{X : \det(L_A^{\mathrm{re}}(X)) = 0\}. \tag{1.3}$$

Let us establish some additional terminology. An $F \in \mathbb{C}<x>^{e\times e}$ is an **atom** [Coh95, Chapter 3] if F is not a zero divisor and does not factor, i.e., F cannot be written as $F = F_1F_2$ for some non-invertible $F_1, F_2 \in \mathbb{C}<x>^{e\times e}$. As a consequence of Lemma 3.1(6) below we will see that if Q_E is an atom, $\ker(E) = (0)$ and $\ker(E^*) = (0)$, then E is ball-minimal.

Given a free matrix polynomial Q, the set

$$\mathcal{G}_Q = \{X \in M(\mathbb{C})^g : \|Q(X)\| < 1\} \subseteq M(\mathbb{C})^g$$

is a **free pseudoconvex** set. A free map $p : \mathcal{D}_E \to \mathcal{D}_B \subseteq M(\mathbb{C})^h$ is **analytic** if there is a free open set $\mathscr{U} \supset \mathcal{D}_E$ such that p extends to a free analytic function $\mathscr{U} \to M(\mathbb{C})^h$. (Often we use p to also denote the extension.) Likewise $p : \mathcal{D}_E \to \mathcal{D}_B \subseteq M(\mathbb{C})^g$ is **bianalytic** if it is analytic and has an analytic inverse $q : \mathcal{D}_B \to \mathcal{D}_E$.

Theorem 1.4. *Suppose $E \in M_d(\mathbb{C})^g$, $B \in M_e(\mathbb{C})^g$ and the free map $p : \mathcal{D}_E \to \mathcal{D}_B$ is bianalytic and satisfies $p(0) = 0$ and $p'(0) = I_g$. Assume*

(a) *$\mathcal{D}_E$ is bounded;*
(b) *Q_E and Q_B are atoms, $\ker(B) = (0)$ and E^* is ball-minimal;*
(c) *p analytically extends to a pseudoconvex domain containing $\mathcal{D}_E$ and $q : \mathcal{D}_B \to \mathcal{D}_E$, the inverse of p, also extends analytically to a pseudoconvex domain containing $\mathcal{D}_B$.*

Then p is convexotonic, E and B are of the same size, and there exist highly algebraic relations amongst the triple (p, E, B). Namely, $d = e$ and there exist $d \times d$ unitary matrices Z and M and a convexotonic g-tuple Ξ such that

(1) *p is the convexotonic map $p = x(I - \Lambda_\Xi(x))^{-1}$;*
(2) *for each $1 \leq j, k \leq g$,*

$$E_k(Z - I)E_j = \sum_s (\Xi_j)_{k,s} E_s; \tag{1.4}$$

in particular, the tuple $R = (Z - I)E$ spans an algebra with multiplication table Ξ,

$$R_k R_j = \sum_s (\Xi_j)_{k,s} R_s;$$

(3) *$B_j = M^* Z E_j M$ for $1 \leq j \leq g$.*

Proof. See Section 4.5. ∎

Remark 1.5. The normalization condition $p'(0) = I$ is a matter of convenience. It can be easily enforced by a linear change of variables in the codomain, since linear change of variables preserves the minimality hypotheses. See Section 2.1 and [AHKM18, Section 8]. □

1.4. **Formulas for maps between spectraballs and free spectrahedra.** Theorem 1.1 characterizes bianalytic maps $\mathcal{B}_E \to \mathcal{D}_A$. If we impose a minimality condition on A, then the next theorem, obtained from a slight modification of the proof of Theorem 1.1, produces formulas for the triples (f, E, A) such that $f : \mathcal{B}_E \to \mathcal{D}_A$ is bianalytic.

Theorem 1.6. *Suppose $E \in M_{d\times d'}(\mathbb{C})^g$ and $A \in M_e(\mathbb{C})^g$ are both linearly independent and ball-minimal.*

If $f : \operatorname{int}(\mathcal{B}_E) \to \operatorname{int}(\mathcal{D}_A)$ is a proper free analytic map, $f(0) = 0$ and $f'(0) = I_g$, then A spans an algebra and f is convexotonic. Indeed, $e = \max\{d, d'\}$ and letting Ξ denote the convexotonic tuple associated to A,

$$f = x(I + \Lambda_\Xi(x))^{-1}.$$

Further:

(a) *If $e = d \geq d'$, then there exist unitary $d \times d$ matrices U and V such that $\mathbb{E} = V^*AU$, where*
$$\mathbb{E} = \begin{pmatrix} E \\ 0 \end{pmatrix} \in M_d(\mathbb{C})^g.$$
(b) *If $e = d' \geq d$, then there exist unitary $d' \times d'$ matrices U and V such that $\mathbb{E} = V^*AU$, where*
$$\mathbb{E} = \begin{pmatrix} E & 0 \end{pmatrix} \in M_{d'}(\mathbb{C})^g.$$

As before, the normalization $f'(0) = I_g$ is a matter of convenience. The proof of Theorem 1.6 is given in Section 3.4.

1.5. **Algebras and convexotonic maps.** There is a natural converse to Theorem 1.6 and it figures prominently in the proof of Theorem 1.1.

Proposition 1.7. *If $J \in M_d(\mathbb{C})^g$ is linearly independent (e.g. $\mathcal{D}_J$ is bounded), spans an algebra and Ξ is the resulting convexotonic tuple,*

$$J_k J_j = \sum_{s=1}^{g} (\Xi_j)_{k,s} J_s,$$

then the convexotonic map

$$q(x) = x(I + \Lambda_\Xi(x))^{-1}$$

is bianalytic from $\operatorname{int}(\mathcal{D}_J)$ to $\operatorname{int}(\mathcal{B}_J)$.

Further,

(1) *The domain of q contains $\mathcal{D}_J$;*
(2) *$q : \mathcal{D}_J \to \mathcal{B}_J$ maps the boundary of $\mathcal{D}_J$ into the boundary of $\mathcal{B}_J$;*
(3) *$q : \mathcal{D}_J \to q(\mathcal{D}_J) \subseteq \mathcal{B}_J$ is proper;*
(4) *if, in addition, $\mathcal{D}_J$ is bounded, then q is a bianalytic map between $\mathcal{D}_J$ and $\mathcal{B}_J$. In particular, the domain of $p = x(I - \Lambda_\Xi(x))^{-1}$, the inverse of q, contains $\mathcal{B}_J$.*

Proof. See Section 2. ∎

In the case J does not span an algebra, we have the following corollary of Proposition 1.7.

Corollary 1.8. *Let $A \in M_d(\mathbb{C})^g$ and assume A is linearly independent. Let $\mathcal{A}$ denote the algebra generated by the tuple A. If $C_{g+1}, \dots, C_h \in M_d(\mathbb{C})$ are any matrices such that the tuple $J = (J_1, \dots, J_h) = (A_1, \dots, A_g, C_{g+1}, \dots, C_h)$ is a basis for the vector space $\mathcal{A}$, then there is a rational map $f : \mathcal{D}_A \to \mathcal{B}_J$ with $f(0) = 0$ and $f'(0) = \begin{pmatrix} I_g & 0 \end{pmatrix}$ such that*

(1) *f is an (injective) proper map from $\operatorname{int}(\mathcal{D}_A)$ into $\operatorname{int}(\mathcal{B}_J)$; and*
(2) *f maps the boundary of $\mathcal{D}_A$ into the boundary of $\mathcal{B}_J$.*

Further, a tuple X is in $\operatorname{int}(\mathcal{D}_A) \subseteq M(\mathbb{C})^g$ if and only if $\begin{pmatrix} X & 0 \end{pmatrix} \in \operatorname{int}(\mathcal{D}_J) \subseteq M(\mathbb{C})^h$ and the resulting inclusion mapping $\iota : \operatorname{int}(\mathcal{D}_A) \to \operatorname{int}(\mathcal{D}_J)$ given by $\iota(x) = \begin{pmatrix} x & 0 \end{pmatrix}$ is proper. Letting $\Xi \in M_h(\mathbb{C})^h$ denote the convexotonic tuple associated to the basis J for the algebra $\mathcal{A}$ (equation (1.2)), and q the convexotonic map,

$$q(x) = \begin{pmatrix} x_1 & \cdots & x_h \end{pmatrix} \left(I + \sum_{j=1}^{h} \Xi_j x_j \right)^{-1}, \tag{1.5}$$

$$f(x) = q \circ \iota(x) = \begin{pmatrix} x_1 & \cdots & x_g & 0 & \cdots & 0 \end{pmatrix} \left(I + \sum_{j=1}^{g} \Xi_j x_j \right)^{-1}.$$

2. Convexotonic maps and algebras

This section gives the proofs of Proposition 1.7 and Corollary 1.8, which amount to the easy direction of our main results.

Lemma 2.1. *Suppose $T \in M_d(\mathbb{C})$. If $I + T + T^* \succeq 0$, then $I + T$ is invertible.*

Proof. Arguing the contrapositive, suppose $I + T$ is not invertible. In this case there is a unit vector γ such that

$$T\gamma = -\gamma.$$

Hence,

$$\langle (I + T + T^*)\gamma, \gamma \rangle = \langle T^*\gamma, \gamma \rangle = \langle \gamma, T\gamma \rangle = -1. \qquad \blacksquare$$

Lemma 2.2. *Let $T \in M_d(\mathbb{C})$. Then*

(a) *$I + T + T^* \succeq 0$ if and only if $I + T$ is invertible and $\|(I + T)^{-1}T\| \leq 1$;*
(b) *$I + T + T^* \succ 0$ if and only if $I + T$ is invertible and $\|(I + T)^{-1}T\| < 1$.*

Similarly if $I - T$ is invertible, then $\|T\| \leq 1$ if and only if $I + R + R^* \succ 0$, where $R = T(I - T)^{-1}$.

Proof. First note that in any case $I+T$ is invertible by Lemma 2.1. (a) We have the following chain of equivalences:

$$\begin{aligned}\|(I+T)^{-1}T\|\le 1 \quad &\iff \quad I-\big((I+T)^{-1}T\big)\big((I+T)^{-1}T\big)^*\succeq 0\\ &\iff \quad I-(I+T)^{-1}TT^*(I+T)^{-*}\succeq 0\\ &\iff \quad (I+T)(I+T)^*-TT^*\succeq 0\\ &\iff \quad I+T+T^*\succeq 0.\end{aligned}$$

The proof of (b) is the same. ■

Proposition 2.3. *For $F\in M_d(\mathbb{C})^g$,*

$$\mathcal{D}_F=\{X\colon \|(1+\Lambda_F(X))^{-1}\Lambda_F(X)\|\le 1\}.$$

Proof. Immediate from Lemma 2.2. ■

Proof of Proposition 1.7. Let q denote the convexotonic map of equation (1.5). Compute

$$\begin{aligned}\Lambda_J(q(x))\,\Lambda_J(x)&=\sum_{s,k=1}^{g}q^s(x)x_kJ_sJ_k=\sum_{j=1}^{g}\sum_{s=1}^{g}q^s(x)\left[\sum_{k=1}^{g}x_k(\Xi_k)_{s,j}\right]J_j\\ &=\sum_{j=1}^{g}\sum_{s=1}^{g}q^s(x)(\Lambda_\Xi(x))_{s,j}J_j=\sum_{j=1}^{g}\sum_{t=1}^{g}x_t\left[\sum_{s=1}^{g}(I+\Lambda_\Xi(x))^{-1}_{t,s}(\Lambda_\Xi(x))_{s,j}\right]J_j\\ &=\sum_{j=1}^{g}\sum_{t=1}^{g}x_t[(I+\Lambda_\Xi(x))^{-1}\Lambda_\Xi(x)]_{t,j}J_j.\end{aligned}$$

Hence,

$$\Lambda_J(q(x))\,(I+\Lambda_J(x))=\sum_{j=1}^{g}\sum_{t}x_t[(I+\Lambda_\Xi(x))^{-1}(I+\Lambda_\Xi(x))]_{t,j}J_j=\Lambda_J(x).$$

Thus, as free matrix rational functions regular at 0,

$$\Lambda_J(q(x))=(I+\Lambda_J(x))^{-1}\,\Lambda_J(x)=:F(x). \tag{2.1}$$

Since J is linearly independent, given $1\le k\le g$, there is a linear functional λ such that $\lambda(J_j)=0$ for $j\ne k$ and $\lambda(J_k)=1$. Applying λ to equation (2.1), gives

$$q^k(x)=\lambda(F(x)). \tag{2.2}$$

Since $\lambda(F(x))$ is a free rational function whose domain contains

$$\mathscr{D}=\{X: I+\Lambda_J(X)\text{ is invertible}\},$$

the same is true for q^k. (As a technical matter, each side of equation (2.2) is a rational expression. Since they are defined and agree on a neighborhood of 0, they determine the same free rational function. It is the domain of this rational function that contains

$\mathscr{D}$. See [Vol17], and also [KVV09], for full details.) By Lemma 2.1, $\mathscr{D}$ contains $\mathcal{D}_A$ (as $X \in \mathcal{D}_A$ implies $I + \Lambda_J(X)$ is invertible). Hence the domain of the free rational mapping q contains $\mathcal{D}_A$ (item (1)). By Lemma 2.2 and equation (2.1), q maps the interior of $\mathcal{D}_J$ into the interior of $\mathcal{B}_J$ and the boundary of $\mathcal{D}_J$ into the boundary of $\mathcal{B}_J$, proving item (2).

Similarly,

$$(I - \Lambda_J(x))^{-1}\, \Lambda_J(x) = \Lambda_J(p(x)), \tag{2.3}$$

where $p(x) = x(I - \Lambda_\Xi(x))^{-1}$. Arguing as above shows the domain of p contains the set

$$\mathscr{E} = \{X : I - \Lambda_J(X) \text{ is invertible}\},$$

which in turn contains $\operatorname{int}(\mathcal{B}_J)$ (since $\|\Lambda_J(X)\| < 1$ allows for an application of Lemma 2.2). By Lemma 2.2 and equation (2.3), p maps the interior of $\mathcal{B}_J$ into the interior of $\mathcal{D}_J$. Hence q is bianalytic between these interiors. Further, if X is in the boundary of $\mathcal{B}_J$, then for $t \in \mathbb{C}$ and $|t| < 1$, we have $p(tX) \in \operatorname{int}(\mathcal{D}_J)$ and

$$\Lambda_J(p(tX)) = (I - \Lambda_J(tX))^{-1}\, \Lambda_J(tX).$$

Assuming $\mathcal{D}_J$ is bounded, it follows that $I - \Lambda_J(X)$ is invertible and thus $X \in \mathscr{E}$ (and is in the domain of p) and $p(X)$ is in the boundary of $\mathcal{D}_J$, proving item (4).

We now turn to item (3). Suppose $Y \in q(\mathcal{D}_J) \cap \partial\mathcal{B}_J$. Thus there is an $X \in \partial\mathcal{D}_J$ such that $q(X) = Y$. Observe that $I - \Lambda_J(Y)$ must be invertible, as otherwise there is a non-zero vector γ such that $\Lambda_J(Y)\gamma = \gamma$, but then, by equation (2.1),

$$(I + \Lambda_J(X))\gamma = \Lambda_J(X)\gamma,$$

a contradiction. It follows that Y is in the domain of p. Thus p restricted to $q(\mathcal{D}_J)$ is a continuous bijection. Therefore, if $K \subseteq q(\mathcal{D}_J)$ is compact, then $q^{-1}(K) = p(K) \subseteq \mathcal{D}_J$ is compact and hence q is proper (as a map from $\mathcal{D}_J$ to $q(\mathcal{D}_J)$). ∎

Proof of Corollary 1.8. Given $X \in M(\mathbb{C})^g$, letting $Y = \begin{pmatrix} X & 0 \end{pmatrix}$,

$$\Lambda_J(Y) = \sum_{j=1}^{h} J_j \otimes Y_j = \sum_{j=1}^{g} A_j \otimes X_j.$$

Hence $L_J(\begin{pmatrix} X & 0 \end{pmatrix}) = L_A(X)$ and it follows that $X \in \operatorname{int}(\mathcal{D}_A)$ if and only if $Y \in \operatorname{int}(\mathcal{D}_J)$. Hence, we obtain a mapping $\iota : \operatorname{int}(cD_A) \mapsto \operatorname{int}(\mathcal{D}_J)$ defined by $\iota(X) = Y$.

Now suppose $K \subseteq \operatorname{int}(\mathcal{D}_J(n))$ is compact and let $L = \iota^{-1}(K) \subseteq \mathcal{D}_A(n)$. In particular, $X \in L$ if and only if $\begin{pmatrix} X & 0 \end{pmatrix} \in K$. Hence if $(X(n))$ is a sequence from L, then $Y(n) = \begin{pmatrix} X(n) & 0 \end{pmatrix}$ is a sequence from K. Since K is compact, $(Y(n))_n$ has a subsequence $(Y(n_j))_j$ that converges to some $Y \in K$. It follows that $Y = \begin{pmatrix} X & 0 \end{pmatrix} \in K \subseteq \operatorname{int}(\mathcal{D}_J)$ for some $X \in \operatorname{int}(\mathcal{D}_A) \cap L$. Hence $(X(n_j))_j$ converges to X and we conclude that L is compact. Hence ι is proper.

Letting $z = (z_1, \dots, z_h)$ denote an h tuple of freely non-commuting indeterminants, and Ξ the convexotonic h tuple as described in the corollary. By Proposition 1.7, the birational mapping

$$q(z) = z(I + \Lambda_\Xi(z))^{-1}$$

is a bianalytic (hence injective and proper) mapping between $\operatorname{int}(\mathcal{D}_J)$ and $\operatorname{int}(\mathcal{B}_J)$ that also maps boundary to boundary. Hence, the composition

$$f(x) = p(\iota(x)) = \begin{pmatrix} x & 0 \end{pmatrix} \, (I - \Lambda_\Xi(x, 0))^{-1}$$

is a proper map from $\operatorname{int}(\mathcal{D}_A)$ into $\operatorname{int}(\mathcal{B}_J)$ that also maps boundary to boundary. ∎

2.1. **Affine linear change of variables.** In this subsection we show that minimality and indecomposability of hermitian monic pencils are preserved under an affine linear change of variables.

For $B \in M_r(\mathbb{C})^g$ and $M \in M_g(\mathbb{C})$, let $M \cdot B$ denote the tuple with entries

$$(M \cdot B)_j = \sum_{k=1}^{g} M_{j,k} B_j. \tag{2.4}$$

Proposition 2.4. *Consider a hermitian monic linear pencil L_A and an affine linear change of variables $\Psi : x \mapsto Mx + b$ for some invertible $g \times g$ matrix M and vector $b \in \mathbb{C}^g$. Assume $L_A(b) \succ 0$. Then $\Psi(\mathcal{D}_A) = \mathcal{D}_F$, where*

$$F = M \cdot (\mathfrak{H} A \mathfrak{H}), \qquad \textit{and} \qquad \mathfrak{H} = L_A(b)^{-1/2}. \tag{2.5}$$

Further,

(1) *L_A is indecomposable if and only if L_F indecomposable;*
(2) *L_A is minimal if and only if L_F is minimal.*

Proof. Equation (2.5) is proved in [AHKM18, §8.2].

Let us first settle the special case $M = I$. If L_A is not indecomposable, then there is a common non-trivial reducing subspace $\mathscr{M}$ for A. It follows that $\mathscr{M}$ is reducing for $L_A(b)$ and hence for $F = \mathfrak{H} A \mathfrak{H}$.

Now suppose L_F is not indecomposable; that is there is a non-trivial reducing subspace $\mathscr{N}$ for $F = \mathfrak{H} A \mathfrak{H}$. We conclude that

$$[I - L_A(b)^{-1}]\mathscr{N} = \mathfrak{H}(L_A(b) - I)\mathfrak{H}\mathscr{N} \subseteq \mathscr{N}$$

Hence $\mathscr{N}$ is invariant for $L_A(b)^{-1}$. Since $\mathscr{N}$ is finite dimensional and $L_A(b)^{-1}$ is invertible, $L_A(b)^{-1}\mathscr{N} = \mathscr{N}$ and consequently $\mathfrak{H}\mathscr{N} = \mathscr{N}$. Because $F = \mathfrak{H} A \mathfrak{H}$ it is now evident that $\mathscr{N}$ is reducing for A.

Now consider the special case $b = 0$. A subspace $\mathscr{M}$ reduces A if and only if it reduces $M \cdot A$. Combining these two special cases proves item (1).

Finally we prove item (2). Since L_A is hermitian, it is unitarily equivalent to $\bigoplus_{j=1}^{\ell} L_{A^j}$, where L_{A^j} are indecomposable hermitian monic linear pencils. Then L_F is unitarily equivalent to $\bigoplus_{j=1}^{\ell} L_{F^j}$, where $F^j = M \cdot (\mathfrak{H} A^j \mathfrak{H})$. By item (1), each of these summands L_{F^j} is indecomposable. Furthermore, since Ψ is bijective it is clear that $\bigcap_{k\neq i} \mathcal{D}_{A^k} \subseteq \mathcal{D}_{A^i}$ if and only if $\bigcap_{k\neq j} \mathcal{D}_{F^k} \subseteq \mathcal{D}_{F^j}$. Therefore free spectrahedra of L_{A^j} are irredundant if and only if free spectrahedra of L_{F^j} are irredundant. Hence L_A is minimal for $\mathcal{D}_A$ if and only if L_F is minimal for $\mathcal{D}_F$. ■

Example 2.5. Even with $M = I$, the property (1) of Proposition 2.4 fails for a general positive definite $\mathfrak{H}$ and F as in (2.5). For example, let

$$A = \begin{pmatrix} 6 & -4 \\ -4 & 3 \end{pmatrix}, \qquad \mathfrak{H} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Then L_A is indecomposable, but since $F = 1 \oplus 2$, the hermitian monic linear pencil L_F is clearly not. ■

3. Characterizing bianalytic maps between spectraballs and free spectrahedra

In this section we prove our main results, Theorems 1.1 and 1.6.

3.1. **Minimality and irreducibility.** Given a g-tuple E of $d \times e$ matrices, let P_E denote the projection onto the span of the ranges of the E_j. Let $\widehat{E} = P_E E$.

Lemma 3.1. *Let E be a g-tuple of $d \times e$ matrices.*

(1) *We have*

$$\begin{pmatrix} I & 0 \\ \Lambda_{E^*} & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & Q_E \end{pmatrix} \begin{pmatrix} I & \Lambda_E \\ 0 & I \end{pmatrix} = \mathbb{L}_E, \quad \begin{pmatrix} I & 0 \\ \Lambda_{\widehat{E}^*} & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & Q_E \end{pmatrix} \begin{pmatrix} I & \Lambda_{\widehat{E}} \\ 0 & I \end{pmatrix} = \mathbb{L}_{\widehat{E}}. \tag{3.1}$$

In particular, the hermitian monic linear pencils $\mathbb{L}^{\mathrm{re}}_{\widehat{E}}$ and $\mathbb{L}^{\mathrm{re}}_E$ have the same spectrahedra, namely $\mathcal{B}_E$.

(2) *The monic linear pencil $\mathbb{L}_{\widehat{E}}$ is indecomposable if and only if Q_E is an atom and* $\ker(E) = (0)$.

(3) *If A is ball-minimal with $\mathcal{D}_A = \mathcal{B}_E$, then L_A is unitarily equivalent to $\mathbb{L}_{E^1} \oplus \cdots \oplus \mathbb{L}_{E^k}$ for some indecomposable monic linear pencils $\mathbb{L}_{E^1}, \ldots, \mathbb{L}_{E^k}$ with irredundant spectraballs $\mathcal{B}_{E^j}$.*

(4) *E is ball-minimal if and only if $\mathbb{L}^{\mathrm{re}}_E$ is minimal.*

(5) *If E is ball-minimal, then, up to unitary equivalence, $Q_E = Q_{E^1} \oplus \cdots \oplus Q_{E^k}$, where $Q_{E^j} \in \mathbb{C}\langle x, y\rangle^{e_j \times e_j}$ are quadratic atoms,* $\ker(E^j) = (0)$ *for all j, and the domains $\{\mathcal{D}_{Q_{E^j}} : 1 \leq j \leq k\}$ are irredundant.*

(6) *If Q_E is an atom,* $\ker(E) = (0)$ *and* $\ker(E^*) = (0)$*, then E is ball-minimal.*

(7) *If E is ball-minimal, $F \in M_{k\times\ell}(\mathbb{C})^g$ and $\mathcal{B}_E = \mathcal{B}_F$, then there exists isometries V, W such that $E = W^* F V$. In particular,*
   (a) *$d \le k$ and $e \le \ell$; and*
   (b) *if $F \in M_{d\times e}(\mathbb{C})^g$ is ball minimal, then V and W are unitary.*

(8) *If E ball-minimal and $Q_F = Q_E$, then there is a tuple R and unitaries U, V such that $\mathcal{B}_E \subseteq \mathcal{B}_R$ and*

$$F = U \begin{pmatrix} E & 0 \\ 0 & R \end{pmatrix} V$$

Proof. (1) Straightforward.

(2) By (3.1), Q_E and $\mathbb{L}_{\widehat{E}}$ are stably associated, cf. [HKV, Section 4]. Hence $\mathbb{L}_{\widehat{E}}$ does not factor in $\mathbb{C}{<}x,y{>}^{(d+e)\times(d+e)}$ if and only if Q_E does not factor in $\mathbb{C}{<}x,y{>}^{e\times e}$ by [HKV, Section 4]. Next, $\mathbb{L}_{\widehat{E}}$ is indecomposable if and only if it does not factor and

$$\ker(\begin{pmatrix} 0 & \widehat{E} \\ 0 & 0 \end{pmatrix}) \cap \ker(\begin{pmatrix} 0 & 0 \\ (\widehat{E})^* & 0 \end{pmatrix}) = (0)$$

([HKV, Section 2.1 and Theorem 3.4]). Thus $\mathbb{L}_{\widehat{E}}$ is indecomposable if and only if Q_E does not factor and $\ker(E) = (0)$.

(3) If A is ball-minimal with $\mathcal{D}_A = \mathcal{B}_E$, then L_A is unitarily equivalent to $L_{A^1} \oplus \cdots \oplus L_{A^k}$ for some indecomposable irredundant monic linear pencils $L_{A^1}, \dots, L_{A^k}$. It suffices to see that the L_{A^j} are of the desired form.

Let $U \in U_n(\mathbb{C})$ be arbitrary and denote $\phi_U(X) = UX$. The structure of $\mathbb{L}_E$ implies

$$\phi_U(\mathcal{Z}^{\mathrm{re}}_{\mathbb{L}_E}(n)) \subseteq \mathcal{Z}^{\mathrm{re}}_{\mathbb{L}_E}(n).$$

Moreover,

$$\mathcal{Z}^{\mathrm{re}}_{\mathbb{L}_E}(n) = \mathcal{Z}^{\mathrm{re}}_{L_{A^1}}(n) \cup \cdots \cup \mathcal{Z}^{\mathrm{re}}_{L_{A^k}}(n)$$

is a union of pairwise distinct real hypersurfaces for all n large enough by [HKV, Remark 3.9]. Note ϕ_U is an analytic map that analytically depends on its parameter U; since ϕ_I is the identity map, it follows that

$$\phi_U(\mathcal{Z}^{\mathrm{re}}_{L_{A^i}}(n)) \subseteq \mathcal{Z}^{\mathrm{re}}_{L_{A^i}}(n)$$

for $i = 1, \dots, \ell$ and large n. By continuity of ϕ_U it further follows that

$$\phi_U(\mathcal{D}_{L_{A^i}}(n)) \subseteq \mathcal{D}_{L_{A^i}}(n).$$

Since n was arbitrary large enough, and U was arbitrary, it follows that $\mathcal{D}_{L_{A^i}}$ is a free circular free spectrahedron ([BMV] or [EHKM17, Section 1.1]), so A^i is after a unitary equivalence of the form

$$\begin{pmatrix} 0 & E^i \\ 0 & 0 \end{pmatrix}$$

for some tuple of rectangular matrices E^i by [EHKM17, Theorem 1.1(2)].

(4) That minimality of $\mathbb{L}_E$ implies ball-minimality of E is immediate.

Conversely, suppose $\mathbb{L}_E$ is not minimal amongst monic linear pencils L_A that define $\mathcal{D}_{\mathbb{L}_E} = \mathcal{B}_E$. By [EHKM17, Theorem 1.1(2)] there is a proper reducing subspace $\mathscr{M}$ for the tuple

$$B = \begin{pmatrix} 0 & E \\ 0 & 0 \end{pmatrix}$$

such that, with $A = B|_{\mathscr{M}}$, we have L_A is minimal for $\mathcal{D}_{\mathbb{L}_E}$. Hence, by item (3), there is an F such that that L_A is unitarily equivalent to $\mathbb{L}_F$. In particular, $\mathcal{B}_F = \mathcal{D}_{\mathbb{L}_E} = \mathcal{B}_E$. Since $\mathscr{M}$ is proper, the size of F is strictly smaller than that of E it follows that E is not ball-minimal.

(5) Combine items (4), (3) and (2) of Lemma 3.1 in that order.

(6) The hypothesis implies $\widehat{E} = E$. It follows that $\mathbb{L}_E$ is indecomposable by item (2). For a pencil L, indecomposability of L implies minimality L^{re}. Thus $\mathbb{L}_E^{\mathrm{re}}$ is minimal and hence E is ball-minimal by item (4).

(7) By item (4), $\mathbb{L}_E^{\mathrm{re}}$ is minimal. Since $\mathbb{L}_F$ defines $\mathcal{B}_E$, it follows that there is a reducing subspace $\mathscr{M}$ for

$$B = \begin{pmatrix} 0 & F \\ 0 & 0 \end{pmatrix}$$

and, with $A = B|_{\mathscr{M}}$ the pencil L_A is unitarily equivalent to $\mathbb{L}_E$. Hence there is an isometry,

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} : \mathbb{C}^d \oplus \mathbb{C}^e \to \mathbb{C}^k \oplus \mathbb{C}^\ell$$

and

$$BV = V \begin{pmatrix} 0 & E \\ 0 & 0 \end{pmatrix}, \quad B^*V = V \begin{pmatrix} 0 & 0 \\ E^* & 0 \end{pmatrix}.$$

From here routine calculations using $\ker(E) = (0)$ and $\ker(E^*) = (0)$ (since E is ball-minimal) show $V_{12} = V_{21} = 0$, V_{11} and V_{22} are isometries and $V_{11}E = FV_{22}$. In particular, $d \le k$ and $e \le \ell$. Moreover, if F is also ball-minimal, then reversing the roles of E and F shows $d = k$ and $e = \ell$ and hence both V_{11} and V_{22} are unitaries.

(8) For notational ease, let $V_j = V_{jj}$ denote the isometries from item (7) viewed as maps onto their ranges. Thus $V_1 : \mathbb{C}^d \to \mathrm{ran}(V_{11})$ and similarly for V_2. In particular, $F(\mathrm{ran}(V_2)) \subseteq \mathrm{ran}(V_1)$ and $F(\mathrm{ran}(V_2)^\perp) \subseteq \mathrm{ran}(V_1)^\perp$. Let R denote the mapping from $\mathrm{ran}(V_2)^\perp \to \mathrm{ran}(V_1)^\perp$ induced by F. Similarly, let S denote the mapping from $\mathrm{ran}(V_2) \to \mathrm{ran}(V_1)$ induced by F. In particular, $V_1^* S V_2 = E$. Choose any unitary mapping $V_1^\perp : \mathbb{C}^{k-d} \to \mathrm{ran}(V_{11})^\perp$ and similarly for $V_2^\perp$. Thus

$$U_1 = \begin{pmatrix} V_1 & 0 \\ 0 & V_1^\perp \end{pmatrix} : \mathbb{C}^d \oplus \mathbb{C}^{k-d} \to \mathrm{ran}(V_{11}) \oplus \mathrm{ran}(V_{11})^\perp$$

is unitary and U_2 defined similarly is also. We have,

$$U_1^* F U_2 = \begin{pmatrix} V_1^* & 0 \\ 0 & (V_1^\perp)^* \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} V_2 & 0 \\ 0 & V_2^\perp \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & (V_1^\perp)^* R V_2^\perp \end{pmatrix}. \qquad \blacksquare$$

Proposition 3.2. *Suppose $E \in M_{d\times e}(\mathbb{C})^g$ and $C \in M_g(\mathbb{C})$ is invertible. If E is ball-minimal, then $C \cdot E$ (see equation (2.4)) is ball-minimal.*

Proof. E is ball-minimal if and only if $\mathbb{L}_E = L_{\left(\begin{smallmatrix} 0 & E \\ 0 & 0 \end{smallmatrix}\right)}$ is minimal for $\mathcal{B}_E$ (Lemma 3.1(4)) if and only if $L_{C\cdot\left(\begin{smallmatrix} 0 & E \\ 0 & 0 \end{smallmatrix}\right)} = L_{\left(\begin{smallmatrix} 0 & E \\ 0 & 0 \end{smallmatrix}\right)}$ is minimal for $\mathcal{B}_{CE}$ (Proposition 2.4) if and only if $C \cdot E$ is ball-minimal (Lemma 3.1(4) again). $\blacksquare$

3.2. **Proof of Proposition 1.3.** We continue to let $x = (x_1, \dots, x_g)$ and $y = (y_1, \dots, y_g)$ denote tuples of freely non-commuting variables such that the x_j and y_k are also freely non-commuting.

The proof of Proposition 1.3 depends crucially on Cohn's theory of projective modules and matrices over the free algebra $\mathbb{C}{<}x{>}$ [Coh95]. An alternative reference is [Sco85]. Let $\mathscr{R}$ denote either $\mathbb{C}{<}x{>}$ or $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle$. The **inner rank**, $\rho_{\mathscr{R}}(V)$ of a $V \in \mathscr{R}^{\ell\times e}$ is the smallest nonnegative integer r for which there exists $V_1 \in \mathscr{R}^{\ell\times r}$ and $V_2 \in \mathscr{R}^{r\times e}$ such that $V = V_1 V_2$. Given $V \in \mathbb{C}{<}x{>}^{\ell\times e}$ it is evident that $\rho_{\mathbb{C}<x>}(V) \geq \rho_{\mathbb{C}\langle\!\langle x \rangle\!\rangle}(V)$. Cohn ([Coh95, Proposition 4.6.13]) proves equality holds; that is, if $V \in \mathbb{C}{<}x{>}^{\ell\times e}$, then

$$\rho_{\mathbb{C}<x>}(V) = \rho_{\mathbb{C}\langle\!\langle x \rangle\!\rangle}(V),$$

justifying writing $\rho(V)$ and calling it the inner rank of V. We note that the analogous statement in the commutative case, where $\mathbb{C}{<}x{>}$ is replaced by $\mathbb{C}[x]$ and $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle$ is replaced by $\mathbb{C}(x)$, is false.

Let $\operatorname{rk} T$ denote the rank of the matrix T.

Lemma 3.3. *Let $0 \neq V \in \mathbb{C}{<}x{>}^{\ell\times e}$ and assume $\ell \geq e$. If $\rho(V) = e$ then there exist infinitely many $n \in \mathbb{N}$ for which there exists a nonempty Zariski open subset $\mathcal{O} \subseteq M_n(\mathbb{C})^g$ such that*

$$\operatorname{rk} V(X) = en \quad \text{for all} \quad X \in \mathcal{O}.$$

A similar statement holds if $\ell \leq e$.

Proof. From the discussion preceding the statement of the lemma, V is of inner rank e as a $\ell \times e$ matrix over $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle$. It follows that the set of columns $\{v_1, \dots, v_e\}$ form a linearly independent subset of $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle^{e\times 1}$ as a (left) vector space over $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle$. Since a linearly independent set over a skew field can be extended to a basis, there is a $\ell \times (\ell - e)$ matrix V' over $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle$ so that $\tilde{V} = \begin{pmatrix} V & V' \end{pmatrix}$ is invertible over $\mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle$; that is there is a $\tilde{W} \in \mathbb{C}\,\langle\!\!\langle x \rangle\!\!\rangle^{\ell\times\ell}$ such that $\tilde{V}\tilde{W} = I_\ell$. By Amitsur's theorem [Ami66] (cf. also [KVV, Proposition 3.8]), there is an $n \in \mathbb{N}$ and a tuple $X^0 \in M_n(\mathbb{C})^g$ so that $\tilde{V}$ and $\tilde{W}$ are both defined at X^0; in particular, $\det \tilde{V}(X^0) \neq 0$. Therefore $\tilde{V}$ is defined and invertible on a

Zariski open subset $\mathcal{O} \subseteq M_n(\mathbb{C})^g$. Clearly, the same conclusion holds for every multiple of n. ■

For $n \in \mathbb{N}$ let $\Omega^{(n)} = (\Omega_1^{(n)}, \dots, \Omega_g^{(n)})$ be a g-tuple of $n \times n$ **generic matrices** [Pro76], i.e.,

$$\Omega_j^{(n)} = (\omega_{j\imath\jmath})_{\imath\jmath},$$

where $\omega_{j\imath\jmath}$ for $1 \leq j \leq g$ and $1 \leq \imath, \jmath \leq n$ are commuting indeterminates. Further, we let $\Upsilon^{(n)}$ and $\Theta^{(n)}$ be further tuples of $n \times n$ generic matrices, with

$$\Upsilon_j^{(n)} = (\upsilon_{j\imath\jmath})_{\imath\jmath}, \qquad \Theta_j^{(n)} = (\theta_{j\imath\jmath})_{\imath\jmath}.$$

Proposition 3.4. *Suppose $F \in \mathbb{C}\langle x, y\rangle^{p\times p}$ is an atom, $F(0) = I$, $\det F(\Omega^{(n)}, \Upsilon^{(n)})$ depends on $\Upsilon^{(n)}$ for large enough n, and $V \in \mathbb{C}\langle x\rangle^{\ell\times e}$. If*

$$\det F(X, Y) = 0 \quad \Rightarrow \quad \operatorname{rk} V(X) < n\,\rho(V)$$

for all $n \in \mathbb{N}$ and all tuples X, Y of $n \times n$ matrices, then $V = 0$.

Proof. Arguing by contradiction, suppose $V \neq 0$. Let $r = \rho(V)$. By the definition of inner rank, $r \leq \min\{\ell, e\}$ and $V = V_1V_2$ for some $V_1 \in \mathbb{C}\langle x\rangle^{\ell\times r}$ and $V_2 \in \mathbb{C}\langle x\rangle^{r\times e}$. Clearly, $\rho(V_1) = \rho(V_2) = r$.

Let $n \in \mathbb{N}$ and X, Y be tuples of $n \times n$ matrices. If $\det F(X, Y) = 0$ then either $\operatorname{rk} V_1(X) < rn$ or $\operatorname{rk} V_2(X) < rn$. The sets

$$\{(X, Y) \colon \operatorname{rk} V_i(X) < rn\}$$

are Zariski closed and we have just seen that their union covers the singularity set $\mathscr{Z}_F$ of F. Since $\det F(\Omega^{(n)}, \Upsilon^{(n)})$ is an irreducible polynomial for large n by [HKV, Theorem 4.3], one of two cases occurs. Namely, either V_1 or V_2 is rank deficient on $\mathscr{Z}_F$. In the first case,

$$\det F(X, Y) = 0 \quad \Rightarrow \quad \operatorname{rk} V_1(X) < nr.$$

Let z_{ij} be a new $\ell \times (\ell - r)$ tuple of free noncommuting variables, let W denote the rectangular matrix polynomial $W = W(z) = \big(z_{ij}\big)_{i,j=1}^{\ell,\ell-r}$ and set

$$\tilde{V} = \begin{pmatrix} V_1 & W \end{pmatrix} \in \mathbb{C}\langle x, z\rangle^{\ell\times\ell}.$$

Observe that if $X \in M_n(\mathbb{C})^g$ and $V_1(X)$ has full rank, then there is a $Z = (Z_{ij})_{i,j=1}^{\ell,\ell-r} \in M_n(\mathbb{C})^{\ell(\ell-r)}$ such that $\tilde{V}(X, Z)$ is invertible. Thus, the polynomial $\det \tilde{V}(\Omega^{(n)}, \Theta^{(n)})$ is not (identically) zero. Hence, by Lemma 3.3, $\det \tilde{V}(\Omega^{(n)}, \Theta^{(n)})$ is not (identically) zero for infinitely many n.

On the other hand,

$$\det F(X, Y) = 0 \quad \Rightarrow \quad \det \tilde{V}(X, Z) = 0$$

for all n and tuples $X,Y \in M_n(\mathbb{C})^g$ and $Z \in M_n(\mathbb{C})^{\ell(\ell-r)}$. Since $\det F(\Omega^{(n)},\Upsilon^{(n)})$ is irreducible for n large enough, it divides $\det \tilde{V}(\Omega^{(n)},\Theta^{(n)})$ for all n large enough. However, there are no $\upsilon_{j\imath\jmath}$ in the non-zero polynomial $\det \tilde{V}(\Omega^{(n)},\Theta^{(n)})$. On the other hand for sufficiently large n, there are some $\upsilon_{j,\imath,\jmath}$ in $\det F(\Omega^{(n)},\Upsilon^{(n)})$ by assumption and we have arrived at a contradiction. Hence $V=0$.

In the second case, where

$$\det F(X,Y)=0 \quad\Rightarrow\quad \operatorname{rk} V_2(X) < nr,$$

we proceed as above, but use

$$\tilde{V} = \begin{pmatrix} V_1 \\ W \end{pmatrix} \in \mathbb{C}\langle x,z\rangle^{e\times e}.$$

Here W is the rectangular matrix polynomial $W=W(z)=\big(z_{ij}\big)_{i,j=1}^{e-r,e}$ consisting of free noncommuting variables z_{ij}. Then apply the $\ell\le e$ case of Lemma 3.3. ∎

Proof of Proposition 1.3. Since E is ball-minimal, by Lemma 3.1(5), after a unitary change of basis we can assume that $Q_E = Q_{E^1}\oplus\cdots\oplus Q_{E^k}$, where $Q_{E^j}\in\mathbb{C}\langle x,y\rangle^{e_j\times e_j}$ are quadratics atoms and the domains $\{\mathcal{D}_{Q_{E^j}} : 1\le j\le k\}$ are irredundant. Let

$$V = \begin{pmatrix} V^1 & \dots & V^k\end{pmatrix}$$

be the decomposition with respect to the above block structure of Q_E. Thus $V^j \in \mathbb{C}\langle x\rangle^{\ell\times e_j}$. Observe that

$$Q^{\mathrm{re}}_{E^j}(X)v=0 \quad\Rightarrow\quad V^j(X)v=0$$

holds for all X,v and $j=1,\dots,k$. Thus, by Lemma 3.1(5), we may (and do) assume that Q_E is an atom.

Let X,Y be tuples of $n\times n$ matrices. By assumption

$$\{(X,X^*)\colon Q^{\mathrm{re}}_E(X)\succeq 0 \,\&\, \det Q^{\mathrm{re}}_E(X)=0\}\subseteq\{(X,X^*)\colon \operatorname{rk} V(X) < n\,\rho(V)\}.$$

Hence,

$$\{(X,X^*)\colon \mathbb{L}^{\mathrm{re}}_{\widehat{E}}(X)\succeq 0 \,\&\, \det \mathbb{L}^{\mathrm{re}}_{\widehat{E}}(X)=0\}\subseteq\{(X,X^*)\colon \operatorname{rk} V(X) < n\,\rho(V)\}$$

by Lemma 3.1(1). Therefore

$$\{(X,Y)\colon \det \mathbb{L}_{\widehat{E}}(X,Y)=0\}\subseteq\{(X,Y)\colon \operatorname{rk} V(X) < n\,\rho(V)\}$$

by [HKV, Proposition 8.3]. Further, $\mathbb{L}_{\widehat{E}}(0)=I$ and $\det(\mathbb{L}_{\widehat{E}}(\Omega^{(n)},\Upsilon^{(n)}))$ depends on some $\upsilon_{j,\imath,\jmath}$ since $\det(\mathbb{L}_{\widehat{E}}(\Omega^{(n)},0))=1$. Finally, $\det\mathbb{L}_{\widehat{E}}=\det Q_E$ and hence $V=0$ by Proposition 3.4.

It is clear that the same proof works for $\mathcal{D}_A$ with L_A minimal. ∎

3.3. **Proof of Theorem 1.1.** In this subsection we prove Theorem 1.1. By [HKM11b], f is bianalytic from $\operatorname{int}(\mathcal{B}_E)$ to $\operatorname{int}(\mathcal{D}_A)$; that is, proper implies bianalytic. The proper condition on f is used to establish the following lemma.

A free analytic function f defined in a neighborhood of 0 has a power series expansion,

$$f(x) = \sum_{j=0}^{\infty} G_j = \sum_{j=0}^{\infty} \sum_{|\alpha|=j} f_\alpha x^\alpha, \tag{3.2}$$

where the α are words in x and $|\alpha|$ is the length of the word α. The term G_j is the **homogeneous of degree** j part of f.

Lemma 3.5. *Suppose $E \in M_{d\times e}(\mathbb{C})^g$ is linearly independent and $B \in M_r(\mathbb{C})^h$. Suppose $f : \operatorname{int}(\mathcal{B}_E) \to \operatorname{int}(\mathcal{D}_B)$ is proper. If $X \in \mathcal{B}_E$ is nilpotent of order N, then there is a (free) polynomial $p(x)$ of degree at most N such that $f_X(z) = f(zX) = p(zX)$ for $z \in \mathbb{C}$ with $|z| < 1$. Further, if $X \in \partial\mathcal{B}_E$ (equivalently $\|\Lambda_E(X)\| = 1$), then $p(X) \in \partial\mathcal{D}_B$.*

Proof. The series expansion of equation (3.2) converges as written on $\mathcal{N}_\epsilon = \{X \in M(\mathbb{C})^g : \sum X_j X_j^* \prec \epsilon^2\}$ for any $\epsilon > 0$ such that $N_\epsilon \subseteq \operatorname{int}(\mathcal{B}_E)$ [HKM12b, Proposition 2.24]. In particular, if $X \in \mathcal{B}_E$ is nilpotent of order N and $|z|$ is small, then

$$f_X(z) := f(zX) = \sum_{j=1}^{N} G_j(zX) = \sum_{j=1}^{N} \left[\sum_{|\alpha|=j} f_\alpha X^\alpha \right] z^j =: p(zX).$$

It now follows that $f_X(z) = p(zX)$ for $|z| < 1$ (since $zX \in \operatorname{int}(\mathcal{B}_E)$ for such z and both sides are analytic in z and agree on a neighborhood of 0).

Now suppose $X \in \partial\mathcal{B}_E(n)$ (still nilpotent of order N). Since $f : \operatorname{int}(\mathcal{B}_E) \to \operatorname{int}(\mathcal{D}_B)$, it follows that $L_A(p(tX)) \succ 0$ for $0 < t < 1$. Arguing by contradiction, suppose $L_B(p(X)) \succ 0$; that is $p(X) \in \operatorname{int}(\mathcal{D}_B(n))$. Hence there is an η such that

$$B_\eta(p(X)) := \{Y \in M_n(\mathbb{C})^g : \|Y - p(X)\| \le \eta\} \subseteq \operatorname{int}(\mathcal{D}_B(n)).$$

Since $K = B_\eta(p(x))$ is compact $L = f_n^{-1}(K) \subseteq \operatorname{int}(\mathcal{B}_E)$ is also compact by the proper hypothesis on f (and hence on each $f_n : \operatorname{int}(\mathcal{B}_E(n)) \to \operatorname{int}(\mathcal{D}_J(n))$). On the other hand, for $t < 1$ sufficiently large, $tX \in L$, but $X \notin \operatorname{int}(\mathcal{B}_E)$, and we have arrived at the contradiction that L cannot be compact. ∎

Remark 3.6. In view of Lemma 3.5, for $X \in \partial\mathcal{B}_E$ nilpotent we let $f(X)$ denote $f_X(1)$. Observe also, if $g = h$, $f(0) = 0$, $f'(0) = I_g$ and $X \in \mathcal{B}_E$ is nilpotent of order two, then $f(X) = X$. □

Lemma 3.7. *Suppose $G \in M_{d\times e}(\mathbb{C})^g$ and $\{G_1,\dots,G_g\}$ is linearly independent, $C \in M_{e\times d}(\mathbb{C})$ and $\Gamma \in M_g(\mathbb{C})^g$. If*

$$G_\ell C G_j = \sum_{s=1}^{g} (\Gamma_j)_{\ell,s} G_s,$$

then the tuple Γ is convexotonic.

Proof. For notational ease let $T = CG \in M_e(\mathbb{C})^g$. The hypothesis implies T spans an algebra (but not that T is linearly independent). Routine calculations give

$$[G_\ell T_j]T_k = \sum_{t=1}^{g} (\Gamma_j)_{\ell,t} G_t\, T_k = \sum_{s,t=1} (\Gamma_j)_{\ell,t} (\Gamma_k)_{t,s} G_s = \sum_{s} (\Gamma_j\, \Gamma_k)_{\ell,s} G_s.$$

On the other hand

$$G_\ell [T_j T_k] = \sum_t G_\ell (\Gamma_k)_{j,t} T_t = \sum_{s,t} (\Gamma_t)_{\ell,s} (\Gamma_k)_{j,t} G_s.$$

By independence of G,

$$(\Gamma_j \Gamma_k)_{\ell,s} = \sum_t (\Gamma_k)_{j,t} (\Gamma_t)_{\ell,s}$$

and therefore

$$\Gamma_j \Gamma_k = \sum_t (\Gamma_k)_{j,t} \Gamma_t$$

and the proof is complete. ∎

Lemma 3.8. *Suppose $B \in M_r(\mathbb{C})^g$ and $Q \in M_{r\times u}(\mathbb{C})$ and let $\mathscr{B}$ denote the algebra generated by B. Let h denote the dimension of $\mathscr{B}$ as a vector space. If $\{B_1 Q,\dots,B_g Q\}$ is linearly independent, then there exists a $g \le t \le h$ and a basis $\{J_1,\dots,J_h\}$ of $\mathscr{B}$ such that*

1. *$J_j = B_j$ for $1 \le j \le g$;*
2. *$\{J_1 Q,\dots,J_t Q\}$ is linearly independent; and*
3. *$J_j Q = 0$ for $t < j \le h$.*

Letting $\Xi \in M_h(\mathbb{C})^h$ denote the convexotonic tuple associated to J,

$$(\Xi_j)_{\ell,k} = 0 \quad \text{for } j > t \text{ and } k \le t.$$

Proof. The set $\mathscr{N} = \{T \in \mathscr{B} : TQ = 0\} \subseteq \mathscr{B}$ is subspace (in fact a left ideal). Choose t so that $h - t$ is the dimension of $\mathscr{N}$ and a basis $\{J_{t+1},\dots,J_h\}$ for $\mathscr{N}$. By item (2), the span $\mathscr{M}$ of $\{B_1,\dots,B_g\}$ satisfies $\mathscr{M} \cap \mathscr{N} = (0)$. Thus $\{B_1,\dots,B_g,J_{t+1},\dots,J_h\}$ is a linearly independent set. Extend this set to a basis $\{J_1,\dots,J_h\}$. It only remains to see item (2) holds. Arguing by contradiction, if $\{J_1 Q,\dots,J_t Q\}$ is linearly dependent, then some linear combination of $\{J_1,\dots,J_t\}$ lies in $\mathscr{N}$.

To prove the last statement, the tuple Ξ satisfies,

$$J_\ell J_j = \sum_{k=1}^{h} (\Xi_j)_{\ell,k} J_k$$

for $1 \le j, \ell \le h$. Thus, for $j > t$ and $1 \le \ell \le h$,

$$0 = J_\ell J_j Q = \sum_{k=1}^{h} (\Xi_j)_{\ell,k} J_k Q = \sum_{s=1}^{t} (\Xi_j)_{\ell,k} J_k Q.$$

By independence of $\{J_k Q : 1 \le k \le t\}$, it follows that $(\Xi_j)_{\ell,k} = 0$ for $k \le t$. ■

Proof of Theorem 1.1. We assume, without loss of generality, that E is ball-minimal. By an affine linear change of variables in the codomain, we assume $f(0) = 0$ and $f'(0) = I_g$. We will now show f is convexotonic.

We perform the off diagonal trick. Given a tuple X, let

$$S_X = \begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix}.$$

Suppose $X \in M_n(\mathbb{C})^g$ and $\|\Lambda_E(X)\| = 1$. It follows that $\|\Lambda_E(S_X)\| = 1$. Thus $S_X \in \partial\mathcal{B}_E$. Since $f : \operatorname{int}(\mathcal{B}_E) \to \operatorname{int}(\mathcal{D}_A)$ is proper (and S_X is nilpotent), $f(S_X) = S_X$ (see Remark 3.6), $S_X \in \partial\mathcal{D}_A$. Thus $I + \Lambda_A(S_X) + \Lambda_A(S_X)$ is positive semidefinite and has a (non-trivial) kernel. Equivalently,

$$1 = \|\Lambda_A(S_X)\| = \|\Lambda_A(X)\|.$$

Thus, by homogeneity, $\|\Lambda_E(X)\| = \|\Lambda_A(X)\|$ for all n and $X \in M_n(\mathbb{C})^g$. Thus $\mathcal{B}_E = \mathcal{B}_A$.

Apply Lemma 3.1(8) and in the notation of that lemma, we assume that $V = I$ (since $\mathcal{D}_A$ and $\mathcal{B}_A$ are unchanged when A is replaced by a unitarily equivalent tuple) and write,

$$A = U \begin{pmatrix} E & 0 \\ 0 & F \end{pmatrix}, \tag{3.3}$$

where $\mathcal{B}_E \subseteq \mathcal{B}_F$. With respect to the orthogonal decomposition in equation (3.3), let

$$R = \begin{pmatrix} I_e \\ 0_{k-e,e} \end{pmatrix}.$$

We will use later the fact that if $Q_E^{\mathrm{re}}(X) \succeq 0$ and $Q_E^{\mathrm{re}}(X)\gamma = 0$, then $Q_A^{\mathrm{re}}(X)R\gamma = 0$. For now observe

$$A_j R = U \begin{pmatrix} E_j \\ 0 \end{pmatrix}. \tag{3.4}$$

Thus, since $\{E_1, \dots, E_g\}$ is linearly independent, the set $\{A_1 R, \dots, A_g R\}$ is linearly independent.

We now apply Lemma 3.8 to A and R in place of B and Q and obtain a basis $\{J_1,\dots,J_h\}$ for $\mathscr{A}$, the algebra generated by $\{A_1,\dots,A_g\}$, and a $g\le t\le h$ such that $J_j=A_j$ for $1\le j\le g$, the set $\{J_jR : 1\le j\le t\}$ is linearly independent and $J_jR=0$ for $t<j\le h$. Let $\xi\in M_h(\mathbb{C})^h$ denote the resulting convexotonic tuple, let $\Xi=-\xi$ and let $\varphi:\operatorname{int}(\mathcal{D}_J)\to\operatorname{int}(\mathcal{B}_J)$ denote the convexotonic map

$$\varphi(x)=x(I-\Lambda_\Xi(x))^{-1}.$$

By Corollary 1.8 the composition $\varphi\circ\iota$ is proper from $\operatorname{int}(\mathcal{D}_E)$ to $\operatorname{int}(\mathcal{B}_J)$. Hence, $\mathscr{F}=\varphi\circ\iota\circ f$ is proper from $\operatorname{int}(\mathcal{B}_E)$ to $\operatorname{int}(\mathcal{B}_J)$. Further $\mathscr{F}(0)=0$ and $\mathscr{F}'(0)=\begin{pmatrix}I_g & 0\end{pmatrix}$ because essentially the same is true for each of the components f,ι,φ. Thus $\mathscr{F}(x)=\begin{pmatrix}x & 0\end{pmatrix}+\rho(x)$, where $\rho(x)$ consists of terms of degree at least two.

Write

$$\mathscr{F}=\begin{pmatrix}\mathscr{F}^1 & \dots & \mathscr{F}^h\end{pmatrix}.$$

Expand $\mathscr{F}$ as a power series,

$$\mathscr{F}=\sum H_j=\sum_{j=1}^{\infty}\sum_{|\alpha|=j}\mathscr{F}_\alpha\,\alpha,$$

where H_j is the homogeneous of degree j part of $\mathscr{F}$. Thus,

$$H_j=\begin{pmatrix}H_j^1 & \dots & H_j^h\end{pmatrix}$$

and $H_1(x)=\begin{pmatrix}x & 0\end{pmatrix}$. Likewise,

$$\mathscr{F}_{x_j}(x)=\begin{pmatrix}0 & \dots & 0 & x_j & 0 & \dots & 0\end{pmatrix}$$

for $1\le j\le g$ and $\mathscr{F}_{x_j}=0$ for $j>g$.

The next objective is to show $H_m^s=0$ for $m\ge 2$ and $s\le t$. Given a positive integer m and a tuple $Y\in M_n(\mathbb{C})^g$, let $Z=Z^Y$ denote the g-tuple of block $(m+1)\times(m+1)$ matrices $Z_j=\big((Z_j)_{a,b}\big)_{a,b=0}^m$ with entries $(Z_j)_{a,a+1}=Y_j$ and $Z_{a,b}=0$ otherwise. In particular, $Z^Y\in\mathcal{B}_E(n(m+1))$ if and only if $Y\in\mathcal{B}_E(n)$. If $Y\in\mathcal{B}_E$, then since Z^Y is also nilpotent, Lemma 3.5 (and Remark 3.6) imply $\mathscr{F}(Z^Y)\in\mathcal{B}_J$; that is if $\|\Lambda_E(Y)\|\le 1$, then $\|\Lambda_J(\mathscr{F}(Z^Y))\|\le 1$. Let $\mathscr{Z}^j=\mathscr{F}^j(Z^Y)$ and note that $\mathscr{Z}^j_{0,m}=H^j_m(Y)$ and $\mathscr{Z}^j_{m-1,m}=H^j_1(Y)$. thus $\mathscr{Z}^j_{m-1,m}=Y_j$ for $1\le j\le g$ and $\mathscr{Z}^j_{m-1,m}=H^j_1(Y)=0$ for $j>g$. Now $\|\Lambda_J(\mathscr{Z})\|\le 1$ is equivalent to $I-\Lambda_J(\mathscr{Z})^*\Lambda_J(\mathscr{Z})\succeq 0$. Thus,

$$I-\Lambda_A(Y)^*\Lambda_A(Y)-\Lambda_J(H_m(Y))^*\Lambda_J(H_m(Y))\succeq 0.$$

Multiplying on the right by $R\otimes I$ and on the left by $R^*\otimes I$,

$$I-\Lambda_{AR}(Y)^*\Lambda_{AR}(Y)-\Lambda_{JR}(H_m(Y))^*\Lambda_{JR}(H_m(Y))\succeq 0.$$

By equation (3.4) and hence, $\Lambda_{AR}(Y)^*\Lambda_{AR}(Y)=\Lambda_E(Y)^*\Lambda_E(Y)$,

$$I-\Lambda_E(Y)^*\Lambda_E(Y)-\Lambda_{JR}(H_m(Y))^*\Lambda_{JR}(H_m(Y))\succeq 0. \tag{3.5}$$

Now suppose $Q_E^{\rm re}(Y) \succeq 0$ and $Q_E^{\rm re}(Y)\gamma = 0$. From the inequality of equation (3.5), it follows that

$$\Lambda_{JR}(H_m(Y))\gamma = 0.$$

Letting $V(y) = \Lambda_{JR}(H_m(y))$ we conclude if (Y,γ) is in the detailed boundary of $\mathcal{B}_E$, then $V(Y)\gamma = 0$. By Proposition 1.3 it follows that $V = 0$; that is,

$$0 = V(y) = \sum_{j=1}^{h} H_m^j(y) J_j R = \sum_{j=1}^{t} H_m^j(y) J_j R.$$

Since $\{J_1R,\dots,J_tR\}$ is linearly independent, it follows that $H_m^j(y) = 0$ for all $1 \le j \le t$ and all $m \ge 2$. Hence,

$$\mathscr{F}(x) = \begin{pmatrix} x & 0 & \Psi(x) \end{pmatrix}$$

where the 0 has length $t-g$ and Ψ has length $h-t$, where Ψ consists of terms of degree at least two.

Let ψ denote the inverse of φ,

$$\psi(x) = x(I + \Lambda_\Xi(x))^{-1}.$$

Thus, $\psi \circ \mathscr{F} = \begin{pmatrix} f(x) & 0 & 0 \end{pmatrix}$ and consequently,

$$\begin{pmatrix} f(x) & 0 & 0 \end{pmatrix} = \begin{pmatrix} x & 0 & \Psi(x) \end{pmatrix} \left((I + \Lambda_\Xi(\begin{pmatrix} x & 0 & \Psi(x) \end{pmatrix})) \right)^{-1}.$$

Rearranging gives,

$$\begin{aligned} \begin{pmatrix} x & 0 & \Psi(x) \end{pmatrix} &= \begin{pmatrix} f(x) & 0 & 0 \end{pmatrix} (I + \Lambda_\Xi(\begin{pmatrix} x & 0 & \Psi(x) \end{pmatrix})) \\ &= \begin{pmatrix} f(x) & 0 & 0 \end{pmatrix} (I + \sum_{j=1}^{g} \Xi_j x_j + \sum_{j=t+1}^{h} \Xi_j \Psi_{j-t}). \end{aligned} \tag{3.6}$$

We now examine, for $1 \le k \le t$, the k-th entry on the right hand side of equation (3.6). Since $(\Xi_j)_{\ell,k} = 0$ for $j > t$ and $k \le t$ (see Lemma 3.8),

$$\begin{aligned} (I + \Lambda_\Xi(\begin{pmatrix} x & 0 & \Psi(x) \end{pmatrix}))_{\ell,k} &= (I + \sum_{j=1}^{g} \Xi_j x_j + \sum_{j=t+1}^{h} \Xi_j \Psi_{j-t})_{\ell,k} \\ &= I_{\ell,k} + \sum_{j=1}^{g} (\Xi_j)_{\ell,k} x_j + \sum_{j=t+1}^{h} (\Xi_j)_{\ell,k} \Psi_{j-t}, \end{aligned}$$

for $1 \le k \le t$. Hence, from equation (3.6), for $g < k \le t$ (so that for $\ell \le g$ we have $I_{\ell,k} = 0$),

$$\sum_{\ell=1}^{g} f^\ell(x) \left(I + \sum_{j=1}^{g} \Xi_j x_j + \sum_{j=t+1}^{h} \Xi_j \Psi_{j-t} \right)_{\ell,k} = \sum_{j,\ell=1}^{g} (\Xi_j)_{\ell,k} f^\ell(x)\, x_j \tag{3.7}$$

and similarly, for $1 \le k \le g$,

$$\sum_{\ell=1}^{g} f^{\ell}(x)\,(I+\sum_{j=1}^{g}\Xi_j x_j+\sum_{j=t+1}^{h}\Xi_j\Psi_{j-t})_{\ell,k}=f^{k}(x)+\sum_{j,\ell=1}^{g}(\widehat{\Xi})_{\ell,k}f^{\ell}(x)\,x_j. \tag{3.8}$$

Combining equations (3.7) and (3.6), for $1 \le j \le g$ and $g < k \le t$,

$$\sum_{\ell=1}^{g}(\Xi_j)_{\ell,k}f^{\ell}(x)=0.$$

Since $\{f^1,\dots,f^g\}$ is linearly independent, it follows that

$$(\Xi_j)_{\ell,k}=0$$

for $1 \le j,\ell \le g$ and $g < k \le t$.

We next show $\widehat{\Xi} \in M_g(\mathbb{C})^g$ defined by

$$(\widehat{\Xi}_j)_{\ell,k}=(\Xi_j)_{\ell,k},\quad 1\le j,\ell,k\le g$$

is convexotonic. For $1 \le j,\ell \le g$,

$$A_\ell A_j R = J_\ell J_r R=\sum_{s=1}^{h}(\Xi_j)_{\ell,s}J_sR=\sum_{s=1}^{t}(\Xi_j)_{\ell,s}J_sR=\sum_{s=1}^{g}(\Xi_j)_{\ell,s}J_sR=\sum_{s=1}^{g}(\Xi_j)_{\ell,s}A_sR. \tag{3.9}$$

Multiplying equation (3.9) on the right by U^* and using equation (3.4) gives

$$\begin{pmatrix}E_\ell & 0\\ 0 & F_\ell\end{pmatrix} U \begin{pmatrix}E_j\\ 0\end{pmatrix}=\begin{pmatrix}\sum_{s=1}^{g}(\Xi_j)_{\ell,s}E_s\\ 0\end{pmatrix}.$$

Writing

$$U=\begin{array}{cc} \begin{array}{cc} d & k-d \end{array} & \\ \begin{pmatrix}U_{11} & U_{12}\\ U_{21} & U_{22}\end{pmatrix} & \begin{array}{c} e \\ k-e \end{array}\end{array} \tag{3.10}$$

it follows that

$$E_\ell U_{11}E_j=\sum_{s=1}^{g}(\Xi_j)_{\ell,s}E_s=\sum_{s=1}^{g}(\widehat{\Xi}_j)_{\ell,s}E_s. \tag{3.11}$$

By Lemma 3.7, the tuple $\widehat{\Xi}$ is convexotonic.

Combining equation (3.6) and equation (3.8), if $1 \le k \le g$, then

$$\begin{aligned} x_k &= \sum_{\ell=1}^{g} f^{\ell}(x)\,(I+\Lambda_{\Xi}(\begin{pmatrix}x & 0 & \Psi(x)\end{pmatrix}))_{\ell,k}\\ &= f^k+\sum_{j,\ell=1}^{g}(\Xi_j)_{\ell,k}f^{\ell}(x)\,x_j = f^k+\sum_{j,\ell=1}^{g}(\widehat{\Xi})_{\ell,k}f^{\ell}(x)\,x_j.\end{aligned}$$

Thus,

$$x = f(x)(I + \Lambda_{\hat{\Xi}}(x))$$

and consequently

$$f(x) = x(I + \Lambda_{\hat{\Xi}}(x))^{-1}$$

is convexotonic. ∎

3.4. **Proof of Theorem 1.6.** We merely indicate the modifications of the proof of Theorem 1.1 that are needed to prove Theorem 1.6. In this case we assume that $f'(0) = I_g$. Observe that, at equation (3.3), the minimal square size hypothesis on A implies F is not there; that is, either

$$A = U \begin{pmatrix} E \\ 0 \end{pmatrix}$$

or

$$A = U \begin{pmatrix} E & 0 \end{pmatrix}.$$

Thus $e = \max\{d, d'\}$ and further, using equation (3.10), $A_j = U_{11} E_j$. By equation (3.11), $\{A_1, \dots, A_g\}$ spans an algebra. Thus $g = h$, $\mathscr{F}(x) = x$ and finally, $f(x) = x(I + \Lambda_{\Xi}(x))^{-1}$. ∎

4. Characterizing bianalytic maps between spectrahedra

In this section we prove Theorem 1.4. We first investigate polynomials defining spectrahedra and relate minimality of these polynomials to certain geometric properties of the boundaries of the corresponding spectrahedra. The main results here are Propositions 4.2 and 4.4. A major accomplishment, Subsection 4.4, is to reduce the eig-generic type hypotheses of [AHKM18] to various natural and cleaner algebraic conditions on the corresponding pencils defining spectrahedra.

Lemma 4.1. *Let L_A be a linear pencil. The set $\{(X, X^*) \mid X \in \mathcal{Z}^{\mathrm{re}}_{L_A}(n)\}$ is Zariski dense in the set $\mathcal{Z}_{L_A}(n)$ for every n. Likewise, $\{(X, X^*) \mid X \in \mathcal{Z}^{\mathrm{re}}_{Q_A}(n)\}$ is Zariski dense in $\mathcal{Z}_{Q_A}(n)$.*

Proof. The first part holds by [KV17, Proposition 5.2]. The second follows immediately from the first. ∎

4.1. **The detailed boundary.** Let ρ be a hermitian $d \times d$ free matrix polynomial with $\rho(0) = I_d$. Thus $\rho \in \mathbb{C}\langle x, y\rangle^{d\times d}$ and $\rho(X, X^*)^* = \rho(X, X^*)$ for all $X \in M(\mathbb{C})^g$. The **detailed boundary** of $\mathcal{D}_\rho$ is the union of sets

$$\widehat{\partial \mathcal{D}_\rho}(n) := \left\{(X, v) \in M_n(\mathbb{C})^g \times \mathbb{C}^{dn} \mid X \in \partial \mathcal{D}_\rho,\ \rho(X, X^*) v = 0\right\}$$

for $n \in \mathbb{N}$. The nomenclature and notation are somewhat misleading in that $\widehat{\partial D_\rho}$ is not determined by the set $\mathcal{D}_\rho$ but by its defining polynomial ρ. Denote also

$$\widehat{\partial^1 \mathcal{D}_\rho}(n) := \left\{ (X, v) \in \widehat{\partial \mathcal{D}_\rho}(n) \mid \dim \ker \rho(X, X^*) = 1 \right\}.$$

For $(X, v) \in \widehat{\partial^1 \mathcal{D}_\rho(n)}$, we call v the **hair** at X and we call X the **follicle**. Letting

$$\pi_1 : M_n(\mathbb{C})^g \times \mathbb{C}^{dn} \to M_n(\mathbb{C})^g \quad \text{and} \quad \pi_2 : M_n(\mathbb{C})^g \times \mathbb{C}^{dn} \to \mathbb{C}^{dn}$$

denote the canonical projections, set

$$\partial^1 \mathcal{D}_\rho(n) = \pi_1 \left(\widehat{\partial^1 \mathcal{D}_\rho}(n) \right), \qquad \operatorname{hair} \mathcal{D}_\rho(n) = \pi_2 \left(\widehat{\partial^1 \mathcal{D}_\rho}(n) \right).$$

We will also abbreviate $\widehat{\partial \mathcal{B}_E}(n) := \widehat{\partial \mathcal{D}_{Q_E}}(n)$, etc.

4.1.1. *Boundary hair spans.* In this subsection we connect the notion of boundary hair to the notion of ball-minimal. Letting $\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ denote the standard basis for $\mathbb{C}^n$, expand $u \in \mathbb{C}^e \otimes \mathbb{C}^n$ as

$$u = \sum_{k=1}^{n} u_k \otimes \mathrm{e}_k$$

Let $\pi : \mathbb{C}^{en} \to \mathbb{C}^e$ denote the projection of u onto u_1.

Proposition 4.2. *Let $E \in M_{d\times e}(\mathbb{C})^g$. Then E is ball-minimal if and only if $\pi(\operatorname{hair} \mathcal{B}_E)$ spans $\mathbb{C}^e$.*

Proof. ($\Rightarrow$) First we prove that if E is ball-minimal, Q_E is an atom and $\mathcal{O}$ is a Zariski dense subset of $\partial^1 \mathcal{D}_{Q_E}$, then

$$\mathcal{S} = (\pi \circ \pi_2) \left(\pi_1^{-1}(\mathcal{O}) \cap \widehat{\partial^1 \mathcal{B}_E} \right)$$

spans $\mathbb{C}^e$.

Assume $\mathcal{S}$ spans a subspace V of dimension $e' < e$. Let P denote the projection of $\mathbb{C}^e$ onto V. Observe that

$$PQ_E(x, y)P^* = PP^* - P\Lambda_{E^*}(y)\Lambda_E(x)P^* = Q_{EP^*}(x, y).$$

Then $(X, v) \in \widehat{\partial^1 \mathcal{D}_{Q_E}}$ and $X \in \mathcal{O}$ implies

$$Q^{\mathrm{re}}_{EP^*}(X)(P \otimes I)v = (P \otimes I)Q^{\mathrm{re}}_E(X)v = 0,$$

so $\mathcal{O} \subseteq \mathcal{Z}_{\mathbb{L}_{EP^*}}$.

By equation (3.1), $\partial^1 \mathcal{D}_{\mathbb{L}_E} = \partial^1 \mathcal{D}_{Q_E}$, and

$$\{(X, X^*) \colon X \in \partial^1 \mathcal{D}_{Q_E}\}$$

is Zariski dense in $\mathcal{Z}_{\mathbb{L}_E}$ by [HKV, Corollary 8.5]. Since $\mathcal{O}$ is Zariski dense in $\partial^1 \mathcal{D}_{Q_E}$, we have $\mathcal{Z}_{\mathbb{L}_E} \subseteq \mathcal{Z}_{\mathbb{L}_{EP^*}}$. By [KV17, Theorem 3.6], there exists a surjective homomorphism from the algebra generated by the coefficients of $\mathbb{L}_{EP^*}$ to the algebra generated by the coefficients of $\mathbb{L}_E$, which equals $M_{d+e}(\mathbb{C})$ since $\mathbb{L}_E$ is indecomposable by Lemma 3.1 items (6) and (2). However, since the first algebra lies in $M_{d+e'}(\mathbb{C})$, we have arrived at a contradiction.

If E is ball-minimal, then $Q_E = Q_{E^1} \oplus \cdots \oplus Q_{E^k}$ for some ball-minimal E^i where the Q_{E^j} are atoms and the spectraballs $\{\mathcal{D}_{Q_E^j} : 1 \le j \le k\}$ are irredundant by Lemma 3.1(5). Note that $\partial^1 \mathcal{D}_{Q_E}$ is Zariski dense in

$$\partial^1 \mathcal{D}_{Q_{E^1}} \cup \cdots \cup \partial^1 \mathcal{D}_{Q_{E^k}},$$

since it is precisely the union of these hypersurfaces minus their intersections. Now the previous paragraph yields the desired conclusion.

($\Leftarrow$) Decompose $\mathbb{L}_E$ into a direct sum of indecomposable hermitian monic linear pencils

$$\mathbb{L}_E = \mathbb{L}_{E^1} \oplus \mathbb{L}_{E^2} \oplus \cdots \oplus \mathbb{L}_{E^k},$$

which corresponds to decomposing Q_E as

$$Q_E = Q_{E^1} \oplus Q_{E^2} \oplus \cdots \oplus Q_{E^k}. \tag{4.1}$$

If E is not ball-minimal, then, without loss of generality $\bigcap_{j>1} \mathcal{D}_{Q_{E^j}} \subseteq \mathcal{D}_{Q_{E^1}}$. In particular, $Q^{\mathrm{re}}_{E^2}(X) \succeq 0$ implies $Q^{\mathrm{re}}_{E^1}(X) \succeq 0$. Suppose $X \in \mathcal{D}_{Q_{E^1}}$ and there is a non-zero vector γ_1 such that $Q^{\mathrm{re}}_{E^1}(X)\gamma_1 = 0$. It follows that $Q^{\mathrm{re}}_{E^j}(X)$ has a non-trivial kernel for some $j > 1$ and therefore the the kernel of $Q^{\mathrm{re}}_E(X)$ is at least two dimensional. We conclude that $P_1 \pi(\operatorname{hair} \mathcal{B}_E) = (0)$, where $\oplus H_j$ denote the decomposition of $\mathbb{C}^e$ with respect to the decomposition implicit in equation (4.1) and P_1 is the projection onto the first coordinate H_1. ■

4.2. **From basis to hyperbasis.** Call a set $\{u^1, \ldots, u^{d+1}\}$ a **hyperbasis** for $\mathbb{C}^d$ if each d element subset is a basis. This notion critically enters the genericity conditions considered in [AHKM18].

Lemma 4.3. *For $E \in M_{d\times e}(\mathbb{C})^g$ and $n \in \mathbb{N}$ assume that $\mathcal{Z}_{Q_E}(n)$ is an irreducible hypersurface in $M_n(\mathbb{C})^{2g}$,*

$$\{(X, X^*) \colon X \in \partial^1 \mathcal{B}_E(n)\}$$

is Zariski dense in $\mathcal{Z}_{Q_E}(n)$, and $\mathcal{S} := \pi(\operatorname{hair} \mathcal{B}_E)$ spans $\mathbb{C}^e$. Then $\mathcal{S}$ contains a hyperbasis for $\mathbb{C}^e$.

Proof. By the spanning assumption, there exist $X^1, \ldots, X^e \in \partial^1 \mathcal{B}_E(n)$ such that

$$\mathbb{C}^e = \bigoplus_{k=1}^{e} \ker Q^{\mathrm{re}}_E(X^k). \tag{4.2}$$

If $X \in \partial^1\mathcal{B}_E(n)$, then $\operatorname{adj} Q_E^{\mathrm{re}}(X)$ is of rank one, and its range lies in $\ker Q_E^{\mathrm{re}}(X)$. Let M_i denote the i-th column of a matrix M. Then for every $k = 1,\dots,e$ there exists $1 \le i_k \le en$ such that $\ker Q_E^{\mathrm{re}}(X^k) = \mathbb{C}\cdot(\operatorname{adj} Q_E^{\mathrm{re}}(X^k))_{i_k}$. Now consider

$$v(t,X,Y) := \sum_{k=1}^{e} t_k(\operatorname{adj} Q_E(X,Y))_{i_k} \tag{4.3}$$

as a vector of polynomials in indeterminates $t=(t_1,\dots,t_e)$ and entries of (X,Y) (i.e., coordinates of $M_n(\mathbb{C})^{2g}$). Let $\{\mathrm{e}_1,\dots,\mathrm{e}_e\}$ denote the standard basis for $\mathbb{C}^e$. For every k we have $v(\mathrm{e}_k,X^k,X^{k*}) \neq 0$ by the construction of v. Since the complements of zero sets are Zariski open and dense in the affine space, for each k the set $U_k = \{t \in \mathbb{C}^g : v(t, X^kX^{k*}) \neq 0\} \subseteq \mathbb{C}^g$ is open and dense and thus so is $\bigcap_{k=1}^e U_k$. Hence there exists $\lambda \in \mathbb{C}^e$ such that $v(\lambda, X^k, X^{k*}) \neq 0$ for every k. Now define a map

$$u : \mathcal{Z}_{Q_E}(n) \to \mathbb{C}^e, \qquad u(X,Y) := v(\lambda, X, Y).$$

Note that u is a polynomial map by (4.3) and $u(X^1,X^{1*}),\dots,u(X^e,X^{e*})$ form a basis of $\mathbb{C}^e$ by (4.2). Therefore

$$u(X,Y) = \sum_{k=1}^{d} r_k(X,Y)u(X^k,X^{k*})$$

for every $(X,Y) \in \mathcal{Z}_{Q_E}(n)$, where r_k are rational functions defined on $\mathcal{Z}_{Q_E}(n)$. In particular, $r_k(X^j,X^{j*}) = \delta_{j,k}$.

Suppose that the product $r_1\cdots r_e \equiv 0$ on

$$\{(X,X^*) \colon X \in \partial^1\mathcal{B}_E(n)\}.$$

Then $r_1\cdots r_e \equiv 0$ on $\mathcal{Z}_{Q_E}(n)$ by the Zariski denseness hypothesis. Therefore $r_k \equiv 0$ on $\mathcal{Z}_{Q_E}(n)$ for some k by the irreducibility hypothesis, contradicting $r_k(X^k,X^{k*}) = 1$.

Consequently there exists $X^0 \in \partial^1\mathcal{B}_E(n)$ such that $r_1(X^0,X^{0*})\cdots r_e(X^0,X^{0*}) \neq 0$. By the construction it follows that $u(X^0,X^{0*}), u(X^1,X^{1*}),\dots,u(X^e,X^{e*}) \in \mathcal{S}$ form a hyperbasis of $\mathbb{C}^e$. ∎

Proposition 4.4. *Let $E \in M_{d\times e}(\mathbb{C})^g$. Then Q_E is an atom and* $\ker(E) = (0)$ *if and only if $\pi(\operatorname{hair}\mathcal{B}_E)$ contains a hyperbasis of $\mathbb{C}^e$.*

Proof. ($\Rightarrow$) If $Q_{\widehat{E}} = Q_E$ is an atom and $\ker(E) = (0)$, then $\widehat{E}$ is ball-minimal by Lemma 3.1(6), so $\pi(\operatorname{hair}\mathcal{B}_E)$ spans $\mathbb{C}^e$ by Proposition 4.2. By [HKV, Corollaries 3.6 and 8.5] the assumptions of Lemma 4.3 are satisfied for some $n \in \mathbb{N}$, so $\pi(\operatorname{hair}\mathcal{B}_E)$ contains a hyperbasis for $\mathbb{C}^e$.

($\Leftarrow$) If E is not ball-minimal, then $\pi(\operatorname{hair}\mathcal{B}_E)$ does not span $\mathbb{C}^e$ by Proposition 4.2. If E is ball-minimal but Q_E is not at atom, then $\mathbb{L}_E$ is minimal but not indecomposable, so $\mathbb{L}_E$ decomposes as $\mathbb{L}_{E^1} \oplus \mathbb{L}_{E^2}$ by Lemma 3.1(3). Hence Q_E decomposes as $Q_{E^1} \oplus Q_{E^2}$. If e_i is the size of Q_{E^i}, then

$$\pi(\operatorname{hair} \mathcal{B}_E) \subseteq (\mathbb{C}^{e_1} \oplus \{0\}^{e_2}) \cup (\{0\}^{e_1} \oplus \mathbb{C}^{e_2}),$$

so $\pi(\operatorname{hair} \mathcal{B}_E)$ cannot contain a hyperbasis for $\mathbb{C}^e = \mathbb{C}^{e_1} \oplus \mathbb{C}^{e_2}$. ■

4.3. **Additional remarks on irreducibility.**

Remark 4.5. Note that Q_E is an atom, $\ker(E) = (0)$ and $\ker(E^*) = (0)$ (or equivalently, $\mathbb{L}_E$ is indecomposable) if and only if the centralizer of

$$\begin{pmatrix} 0 & E_1 \\ 0 & 0 \end{pmatrix}, \dots \begin{pmatrix} 0 & E_g \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ E_1^* & 0 \end{pmatrix}, \dots \begin{pmatrix} 0 & 0 \\ E_g^* & 0 \end{pmatrix},$$

is trivial. This amounts to checking whether a system of linear equations has a solution. □

Remark 4.6. If $\mathbb{L}_E$ is indecomposable, then so is L_E. Indeed, if $L_E = L_{E^1} \oplus L_{E^2}$, then $\mathbb{L}_E$ equals $\mathbb{L}_{E^1} \oplus \mathbb{L}_{E^2}$ up to a canonical shuffle.

However, the converse is not true. For example, with

$$\Lambda(x) = \begin{pmatrix} 0 & x_2 \\ x_1 & 0 \end{pmatrix},$$

$$I + \Lambda(x) + \Lambda^*(y) = \begin{pmatrix} 1 & x_2 + y_1 \\ x_1 + y_2 & 1 \end{pmatrix}$$

is an indecomposable hermitian monic linear pencil, but

$$I - \Lambda\Lambda^* = \begin{pmatrix} 1 - x_1 y_1 & 0 \\ 0 & 1 - x_2 y_2 \end{pmatrix}$$

factors. □

4.4. **The eig-generic condition.** In this subsection we connect the various genericity assumptions used in [AHKM18] to clean, purely algebraic conditions of the corresponding hermitian monic linear pencils, see Proposition 4.9. We begin by recalling these assumptions precisely.

Definition 4.7 ([AHKM18, §7.1.2])**.** A tuple $A \in M_d(\mathbb{C})^g$ is **weakly eig-generic** if there exists an $\ell \leq d+1$ and, for $1 \leq j \leq \ell$, positive integers n_j and tuples $\alpha^j \in (\mathbb{M}_{n_j}(\mathbb{C})^g$ such that

(a) for each $1 \leq j \leq \ell$, the eigenspace corresponding to the largest eigenvalue of $\Lambda_A(\alpha^j)^* \Lambda_A(\alpha^j)$ has dimension one and hence is spanned by a vector $u^j = \sum_{a=1}^{n_j} u_a^j \otimes e_a$; and
(b) the set $\mathscr{U} = \{u_a^j : 1 \leq j \leq \ell, 1 \leq a \leq n_j\}$ contains a hyperbasis for $\ker(A)^\perp = \operatorname{rg}(A^*)$.

The tuple is **eig-generic** if it is weakly eig-generic and $\ker(A) = (0)$ (equivalently, $\mathrm{rg}(A^*) = \mathbb{C}^d$).

Finally, a tuple A is **$*$-generic** (resp. **weakly $*$-generic**) if there exists an $\ell \le d$ and tuples β^j such that the kernels of $I - \Lambda_A(\beta^j)\Lambda_A(\beta^j)^*$ have dimension one and are spanned by vectors $\mu^j = \sum \mu_a^j \otimes e_a$ for which the set $\{\mu_a^j : j, a\}$ spans $\mathbb{C}^d$ (resp. $\mathrm{rg}(A) = \ker(A^*)^\perp$).

Remark 4.8. One can replace n_j with $\sum_{j=1}^{\ell} n_j$ in Definition 4.7, so we can without loss of generality assume $n_1 = \cdots = n_g =: n$. □

Mixtures of these generic conditions were critical assumptions in the main theorems of [AHKM18]. The next proposition gives elegant and much more familiar replacements for them.

Proposition 4.9. *Let $A \in M_d(\mathbb{C})^g$.*

(1) *A is eig-generic if and only if Q_A is an atom and* $\ker(A) = (0)$.
(2) *A is $*$-generic if and only if A^* is ball-minimal.*
(3) *Let P be the projection onto* $\mathrm{rg}(A^*)$. *Then A is weakly eig-generic if and only if Q_{AP^*} is an atom and* $\ker(AP^*) = (0)$.
(4) *Let P be the projection onto* $\mathrm{rg}(A)$. *Then A is weakly $*$-generic if and only if A^*P^* is ball-minimal.*

Proof. (1) Follows from Proposition 4.4 and Remark 4.8.

(2) Follows from the $*$-analog of Proposition 4.2 and Remark 4.8.

(3) Follows from (1).

(4) Follows from (2). ■

4.5. **Proof of Theorem 1.4.** We use Proposition 4.9. In the terminology of [AHKM18], assumptions (a) and (b) imply that E is eig-generic and $*$-generic, and B is eig-generic. Theorem 1.4 thus follows from [AHKM18, Theorem 7.10]. ■

References

[AM14] J. Agler, J. McCarthy: *Global holomorphic functions in several non-commuting variables*, Canad. J. Math. 67 (2015) 241–285. 3

[Ami66] S.A. Amitsur: *Rational identities and applications to algebra and geometry*, J. Algebra 3 (1966) 304–359. 16

[AHKM18] M. Augat, J.W. Helton, I. Klep, S. McCullough: *Bianalytic Maps Between Free Spectrahedra*, Math. Ann. 371 (2018) 883–959. 1, 4, 6, 7, 12, 25, 27, 29, 30

[BGM05] J.A. Ball, G. Groenewald, T. Malakorn: Structured noncommutative multidimensional linear systems. SIAM J. Control Optim. 44 (2005), no. 4, 1474–1528. 4

[BGM06a] J.A. Ball, G. Groenewald, T. Malakorn: *Bounded Real Lemma for Structured Non-Commutative Multidimensional Linear Systems and Robust Control*, Multidimens. Syst. Signal Process. 17 (2006) 119–150. 4

[BGM06b] J.A. Ball, G. Groenewald, T. Malakorn: Conservative structured noncommutative multidimensional linear systems. The state space method generalizations and applications, 179–223, Oper. Theory Adv. Appl., 161, Linear Oper. Linear Syst., Birkhuser, Basel, 2006. 4

[BMV] J.A. Ball, G. Marx, V. Vinnikov: *Interpolation and transfer-function realization for the noncommutative Schur-Agler class*, preprint, `http://arxiv.org/abs/1602.00762` 3, 14

[B-TN02] A. Ben-Tal, A. Nemirovski: *On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty*, SIAM J. Optim. 12 (2002) 811–833. 3

[BR11] J. Berstel, C. Reutenauer: *Noncommutative rational series with applications*, Encyclopedia of Mathematics and its Applications, 137. Cambridge University Press, 2011. 4

[BPR13] G. Blekherman, P.A. Parrilo, R.R. Thomas (editors): *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization 13, SIAM, 2013. 2

[BGFB94] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan: *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics 15, SIAM, 1994. 2

[Brä11] P. Brändén: *Obstructions to determinantal representability*, Adv. Math. 226 (2011) 1202–1212. 2

[Coh95] P.M. Cohn: *Skew fields. Theory of general division rings*, Encyclopedia of Mathematics and its Applications, 57. Cambridge University Press, 1995. 4, 6, 16

[DDOSS17] K.R. Davidson, A. Dor-On, O.M. Shalit, B. Solel: *Dilations, Inclusions of Matrix Convex Sets, and Completely Positive Maps*, Int. Math. Res. Not. IMRN 2017, 4069–4130. 3

[dOHMP09] M. de Oliveira, J.W. Helton, S. McCullough, M. Putinar: *Engineering systems and free semi-algebraic geometry*, in: Emerging applications of algebraic geometry (edited by M. Putinar, S. Sullivant), 17–61, Springer-Verlag, 2009. 3

[EW97] E.G. Effros, S. Winkler: *Matrix convexity: operator analogues of the bipolar and Hahn-Banach theorems*, J. Funct. Anal. 144 (1997) 117–152. 3

[EHKM17] E. Evert, J.W. Helton, I. Klep, S. McCullough: *Circular Free Spectrahedra*, J. Math. Anal. Appl. 445 (2017) 1047–1070. 3, 14, 15

[For93] F. Forstnerič: *Proper holomorphic mappings: a survey*, in: Several complex variables (Stockholm, 1987/1988) 297–363, Math. Notes 38, Princeton Univ. Press, 1993. 5

[GK-VVW16] A. Grinshpan, D.S. Kaliuzhnyi-Verbovetskyi, V. Vinnikov, H.J. Woerdeman: *Matrix-valued Hermitian Positivstellensatz, lurking contractions, and contractive determinantal representations of stable polynomials*, in Oper. Theory Adv. Appl 255, 123–136, Birkhäuser/Springer, 2016. 2

[HKM11a] J.W. Helton, I. Klep, S. McCullough: *Analytic mappings between noncommutative pencil balls,* J. Math. Anal. Appl. 376 (2011), no. 2, 407–428. 5

[HKM11b] J.W. Helton, I. Klep, S. McCullough: *Proper Analytic Free Maps*, J. Funct. Anal. 260 (2011) 1476–1490. 5, 19

[HKM12a] J.W. Helton, I. Klep, S. McCullough: *The convex Positivstellensatz in a free algebra*, Adv. Math. 231 (2012) 516–534.

[HKM12b] J.W. Helton, I. Klep, S. McCullough: *Free analysis, convexity and LMI domains,* Mathematical methods in systems, optimization, and control, 195–219, Oper. Theory Adv. Appl., 222, Birkhäuser/Springer, Basel, 2012. 19

[HKM13] J.W. Helton, I. Klep, S. McCullough: *The matricial relaxation of a linear matrix inequality*, Math. Program. 138 (2013) 401–445. 5

[HKM17] J.W. Helton, I. Klep, S. McCullough: *The tracial Hahn-Banach theorem, polar duals, matrix convex sets, and projections of free spectrahedra*, J. Eur. Math. Soc. 19 (2017) 1845–1897. 3

[HKMS] J.W. Helton, I. Klep, S. McCullough, M. Schweighofer: *Dilations, Linear Matrix Inequalities, the Matrix Cube Problem and Beta Distributions*, to appear in Mem. Amer. Math. Soc., `https://arxiv.org/abs/1412.1481` 3

[HKV] J.W. Helton, I. Klep, J. Volčič: *Geometry of free loci and factorization of noncommutative polynomials*, to appear in Adv. Math., `https://arxiv.org/abs/1708.05378` 5, 14, 17, 18, 26, 28

[HV07] J.W. Helton, V. Vinnikov: *Linear matrix inequality representation of sets*, Comm. Pure Appl. Math. 60 (2007) 654–674. 2

[KVV09] D. Kaliuzhnyi-Verbovetskyi, V. Vinnikov: *Singularities of rational functions and minimal factorizations: the noncommutative and the commutative setting*, Linear Algebra Appl. 430 (2009) 869–889. 3, 11

[KVV14] D. Kaliuzhnyi-Verbovetskyi, V. Vinnikov: *Foundations of Free Noncommutative Function Theory,* Mathematical Surveys and Monographs 199, AMS, 2014. 3

[KŠ17] I. Klep, Š. Špenko: *Free function theory through matrix invariants*, Canad. J. Math. 69 (2017) 408–433. 3

[KVV] I. Klep, V. Vinnikov, J. Volčič: *Multipartite rational functions*, preprint `https://arxiv.org/abs/1509.03316` 16

[KV17] I. Klep, J. Volčič: *Free loci of matrix pencils and domains of noncommutative rational functions*, Comment. Math. Helv. 92 (2017) 105–130. 5, 25, 26

[Kra92] S.G. Krantz: *Function theory of several complex variables*, Second edition. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992. 5

[MSS15] A.W. Marcus, D.A. Spielman, N. Srivastava: *Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem*, Ann. of Math. 182 (2015) 327–350. 2

[MS08] P.S. Muhly, B. Solel: *Schur class functions and automorphism of Hardy algebras*, Doc. Math. 13 (2008) 365–411. 3

[NT12] T. Netzer, A. Thom: *Polynomials with and without determinantal representations*, Linear Algebra Appl. 437 (2012) 1579–1595. 2

[PSS18] B. Passer, O.M. Shalit, B. Solel: *Minimal and maximal matrix convex sets*, J. Funct. Anal. 274 (2018) 3197–3253. 3

[Pau02] V. Paulsen: *Completely bounded maps and operator algebras*, Cambridge Univ. Press, 2002. 3

[Pop10] G. Popescu: *Free holomorphic automorphisms of the unit ball of $B(H)^n$*, J. reine angew. Math. 638 (2010) 119–168. 3

[Pro76] C. Procesi: *The invariant theory of $n \times n$ matrices*, Adv. Math. 19 (1976) 306–381. 17

[Sco85] A.H. Schofield: *Representation of rings over skew fields*, London Mathematical Society Lecture Note Series, 92. Cambridge University Press, 1985. 16

[SIG96] R.E. Skelton, T. Iwasaki, K.M. Grigoriadis: *A Unified Algebraic Approach to Linear Control Design,* Taylor and Francis, 1996. 2

[Vin93] V. Vinnikov: *Self-adjoint determinantal representations of real plane curves*, Math. Ann. 296 (1993) 453–479. 2

[Voi04] D.-V. Voiculescu: *Free analysis questions I: Duality transform for the coalgebra of $\partial_{X:B}$*, Int. Math. Res. Not. 16 (2004) 793–822. 3

[Vol17] J. Volčič: *On domains of noncommutative rational functions*, Linear Algebra Appl. 516 (2017) 69–81. 3, 11

[Zal17] A. Zalar: *Operator Positivstellensätze for noncommutative polynomials positive on matrix convex sets*, J. Math. Anal. Appl. 445 (2017) 32–80. 5

Appendix A. More on the proof of Theorem 1.1 in matrix form

Here we illustrate the argument that $H_m^s = 0$ for the special case $H_3^s = 0$ for $1 \le s \le t$ in $g = 2$ variables. We do the same argument as in Section 3.3 but in block matrix notation because some may find it more intuitive.

Given a tuple $Y = (Y_1, Y_2)$, let

$$Z_j^Y = \begin{pmatrix} 0 & Y_j & 0 & 0 \\ 0 & 0 & Y_j & 0 \\ 0 & 0 & 0 & Y_j \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{for} \quad j = 1, \dots g,$$

and observe

$$\Lambda_E(Z^Y) = \begin{pmatrix} 0 & \Lambda_E(Y) & 0 & 0 \\ 0 & 0 & \Lambda_E(Y) & 0 \\ 0 & 0 & 0 & \Lambda_E(Y) \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

In particular, $Y \in \mathcal{B}_E(n)$ if and only if $Z^Y \in \mathcal{B}_E(3n)$. Since

$$Z_j^Y Z_k^Y = \begin{pmatrix} 0 & 0 & Y_jY_k & 0 \\ 0 & 0 & 0 & Y_jY_k \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad Z_j^Y Z_k^Y Z_\ell^Y = \begin{pmatrix} 0 & 0 & 0 & Y_jY_kY_\ell \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

it follows that

$$\mathscr{F}^j(Z^Y) = \begin{pmatrix} 0 & H_1^j(Y) & H_2^j(Y) & H_3^j(Y) \\ 0 & 0 & H_1^j(Y) & H_2^j(Y) \\ 0 & 0 & 0 & H_1^j(Y) \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence, for $1 \le j \le g$,

$$\mathscr{F}^j(Z^Y) = \begin{pmatrix} 0 & Y_j & H_2^j(Y) & H_3^j(Y) \\ 0 & 0 & Y_j & H_2^j(Y) \\ 0 & 0 & 0 & Y_j \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and for $j > g$,

$$\mathscr{F}^j(Z^Y) = \begin{pmatrix} 0 & 0 & H_2^j(Y) & H_3^j(Y) \\ 0 & 0 & 0 & H_2^j(Y) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Consequently, after the canonical shuffle

$$\Lambda_J(\mathscr{F}(Z^Y)) \stackrel{\text{c.s.}}{=} \begin{pmatrix} 0 & \Lambda_A(Y) & \Lambda_J(H_2(Y)) & \Lambda_J(H_3(Y)) \\ 0 & 0 & \Lambda_A(Y) & \Lambda_A(H_2(Y)) \\ 0 & 0 & 0 & \Lambda_J(Y) \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{A.1}$$

Now if $Y \in \mathcal{B}_E$, then $\mathscr{F}(Z^Y) \in \mathcal{B}_E$ and thus $\|\Lambda_J(\mathscr{F}(Z^Y))\| \le 1$ from which it follows, by considering the last column in the matrix representation of equation (A.1), that

$$I - \Lambda_A(Y)^*\Lambda_A(Y) - \Lambda_J(H_3(Y))^*\Lambda_J(H_3(Y)) \succeq 0.$$

Multiplying by $R \otimes I$ on the right and $R^* \otimes I$ on the left and using $R^*A_k^*A_jR = E_k^*E_j$, produces

$$I - \Lambda_E(Y)^*\Lambda_E(Y) - \Lambda_{JR}(H_3(Y))^*\Lambda_{JR}(H_3(Y)) \succeq 0.$$

In particular, if also there is a vector γ such that $Q_E(X)\gamma = 0$, then $\Lambda_{JR}(H^3(Y))\gamma = 0$. Hence, with $V(y) = \Lambda_{JR}(H^3(y))$, if $Q_E(Y) \succeq 0$ and $Q_E(Y)\gamma = 0$, then $V(Y)\gamma = 0$. An application of Proposition 1.3 gives $V = 0$. Therefore,

$$0 = \sum_{j=1}^{h} J_jR \otimes H_3^j(y) = \sum_{j=1}^{t} J_jR \otimes H_3^j(y).$$

Since $\{J_1R, \dots, J_tR\}$ is linearly independent, it follows that $H_3^j(y) = 0$ for $1 \le j \le t$. ■

J. William Helton, Department of Mathematics, University of California, San Diego

Email address: helton@math.ucsd.edu

Igor Klep, Department of Mathematics, The University of Auckland, New Zealand

Email address: igor.klep@auckland.ac.nz

Scott McCullough, Department of Mathematics, University of Florida, Gainesville

Email address: sam@math.ufl.edu

Jurij Volčič, Department of Mathematics, Ben-Gurion University of the Negev, Israel

Email address: volcic@post.bgu.ac.il

Contents