NONCOMMUTATIVE PARTIALLY CONVEX RATIONAL FUNCTIONS

MICHAEL JURY[1], IGOR KLEP[2], MARK E. MANCUSO[3], SCOTT MCCULLOUGH[4], AND JAMES ELDRED PASCOE[5]

Abstract. Motivated by classical notions of bilinear matrix inequalities (BMIs) and partial convexity, this article investigates partial convexity for noncommutative functions. It is shown that noncommutative rational functions that are partially convex admit novel butterfly-type realizations that necessitate square roots. A strengthening of partial convexity arising in connection with BMIs – xy-convexity – is also considered. A characterization of xy-convex polynomials is given.

1. Introduction

Convexity and its matricial analogs arise naturally in many mathematical and engineering contexts. A function $f : [a, b] \to \mathbb{R}$ is convex if

$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}\left(f(x) + f(y)\right)$$

for all $x, y \in [a, b]$. Convex functions have good optimization properties. For example, local minima are global, making them highly desirable in applications. The dimension-free or scalable matrix analog of convexity appears in many modern applications, such as linear systems engineering [BGFB94, SIG98], wireless communication [JB07], matrix means [And89, And94, Han81], perspective functions [Eff09, ENE11], random matrices and free probability [GS09] and noncommutative function theory [DK+, HMV06, HM04, DHM17, BM14]. Often in systems engineering [dHMP09] problems have two classes of variables: known unknowns $a = (a_1, \dots, a_{\mathtt{h}})$ and unknown unknowns $x = (x_1, \dots, x_{\mathtt{g}})$. Linear system problems specified by a signal flow diagram naturally give rise to matrix inequalities $p(a, x) \succeq 0$, where p is a polynomial, or more generally a rational function, in freely noncommuting variables. The a variables represent system parameters whose size, which can be large, depends upon the specific problem, and the x variables represent the design variables. A key point is that the form of $p(a, x)$ depends only upon the signal flow diagram. Partial convexity in the unknown unknowns x is then sufficient for reliable numerics and optimization.

2010 *Mathematics Subject Classification.* 46N10, 26B25 (Primary); 47A63, 52A41, 90C25 (Secondary).

Key words and phrases. partial convexity, biconvexity, bilinear matrix inequality (BMI), noncommutative rational function, noncommutative polynomial, realization theory.

[1]Research Supported by NSF grant DMS-1900364.

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

[3]Research partially supported by NSF grant DMS-1565243.

[4]Research supported by NSF grants DMS-1361501 and DMS-1764231.

[5]Partially supported by NSF MSPRF DMS 1606260.

A function $f : (-1, 1) \to \mathbb{R}$ is matrix convex if

$$f\left(\frac{X+Y}{2}\right) \preceq \frac{1}{2}\left(f(X) + f(Y)\right)$$

for all hermitian matrices X, Y with spectrum in $(-1, 1)$. Matrix convex functions are automatically real analytic and admit analytic realizations, such as the famous Kraus formula [Kra36, Bha97]

$$f(x) = a + bx + \int_{-1}^{1} \frac{x^2}{1+tx}\, d\mu, \tag{1.1}$$

where $a, b \in \mathbb{R}$ and μ is a finite Borel measure on $[-1, 1]$. Conversely, functions of the form (1.1) are readily seen to be matrix convex on $(-1, 1)$. As an example, the Kraus formula (1.1) in conjunction with the asymptotics at infinity shows that x^2 is matrix convex, but x^4 is not.

In the noncommutative multivariable setting one considers noncommutative (nc) polynomials, rational functions and their generalizations. An nc polynomial is a linear combination of words in the freely noncommuting letters $x = (x_1, \ldots, x_{\mathtt{g}})$. For example,

$$p(x) = x_1x_2 - 17x_2x_1 + 4 \tag{1.2}$$

is a nc (or free) polynomial. Noncommutative polynomials are naturally evaluated at tuples of matrices of any size. For instance, to evaluate $p(x)$ from (1.2) on

$$X_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix},$$

we substitute X_i for the variable x_i, that is,

$$p(X_1, X_2) = X_1X_2 - 17X_2X_1 + 4I_2 = \begin{pmatrix} 69 & 99 \\ 61 & 99 \end{pmatrix}.$$

More generally, an nc rational function is a syntactically valid expression involving $x, +, \cdot, ()^{-1}$ and scalars. Thus

$$r(x) = 1 + (x_1 - x_2(x_1x_2 - x_2x_1)^{-1})^{-1}$$

is an example of a nc rational function. It is evaluated at a tuple $X = (X_1, X_2)$ of $n \times n$ matrices for which $X_1X_2 - X_2X_1$ is invertible and in turn $X_1 - X_2(X_1X_2 - X_2X_1)^{-1}$ is invertible in the natural way to output an $n \times n$ matrix $r(X)$. A nc rational function r is **symmetric** if $r(X) = r(X)^*$ for all hermitian tuples X in its domain.

Matrix convexity for multivariate nc functions is now well understood. Analogs of the Kraus representation, the so-called butterfly realizations, were obtained in [HMV06] for rational functions and in [PTD+] for more general nc functions. There is a paucity of matrix convex polynomials: as first observed in [HM04] they are of degree at most two.

A main result of this paper, Theorem 1.2, is an analog of the Kraus representation for partially convex nc rational functions. Specialized to polynomials, our results extend and generalize results of [HHLM08]. Moreover, we also investigate the stronger notion of xy-convexity, modeled on the theory of bilinear matrix inequalities (BMIs) [KSVS04].

1.1. **Main results.** For positive integers $\mathtt{k}$ and n, let $\mathbb{S}_n^{\mathtt{k}} = \mathbb{S}_n^{\mathtt{k}}(\mathbb{C})$ denote the $\mathtt{k}$-tuples of $n \times n$ hermitian matrices over $\mathbb{C}$. A subset $\mathcal{D} = (\mathcal{D}_n)_n$ of $\mathbb{S}^{\mathtt{k}}$ is a sequence of sets such that $\mathcal{D}_n \subseteq \mathbb{S}_n^{\mathtt{k}}$. This subset is **free**, or a **free set**, if it is **closed under direct sums** and **unitary conjugation**: if $Y \in \mathcal{D}_m$, $X \in \mathcal{D}_n$, and U is an $n \times n$ unitary matrix, then

$$\begin{aligned} X \oplus Y &:= \big(X_1 \oplus Y_1, \quad \cdots, \quad X_{\mathtt{k}} \oplus Y_{\mathtt{k}}\big) = \left(\begin{pmatrix} X_1 & 0 \\ 0 & Y_1 \end{pmatrix}, \cdots, \begin{pmatrix} X_{\mathtt{k}} & 0 \\ 0 & Y_{\mathtt{k}} \end{pmatrix} \right) \in \mathcal{D}_{n+m}, \\ U^* X U &:= (U^* X_1 U, \dots, U^* X_{\mathtt{k}} U) \in \mathcal{D}_n. \end{aligned}$$

It is open if each $\mathcal{D}_n$ is open. (In general adjectives such as open and connected apply term-wise to $\mathcal{D}$.)

Since we are dividing our freely noncommuting variables into two classes $a = (a_1, \dots, a_{\mathtt{h}})$ and $x = (x_1, \dots, x_{\mathtt{g}})$, where $\mathtt{g}$ and $\mathtt{h}$ are positive integers, we take $\mathtt{k} = \mathtt{h} + \mathtt{g}$ and let $\mathbb{S}^{\mathtt{k}} = \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}} = (\mathbb{S}_n^{\mathtt{h}} \times \mathbb{S}_n^{\mathtt{g}})_n$. We express elements of $\mathbb{S}_n^{\mathtt{k}}$ as (A, X), where $A \in \mathbb{S}^{\mathtt{h}}$ and $X \in \mathbb{S}^{\mathtt{g}}$.

A symmetric nc rational function $r(a, x)$ that is **regular** at the origin (has 0 in its domain) admits a symmetric realization

$$r(a,x) = c^* \big(J - \sum_{i=1}^{\mathtt{g}} T_i x_i - \sum_{j=1}^{\mathtt{h}} S_j a_j\big)^{-1} c, \tag{1.3}$$

where, for some positive integer e, J is an $e \times e$ signature matrix ($J^2 = I,\ J^* = J$), the $e \times e$ matrices S_j, T_i are hermitian and $c \in \mathbb{C}^e$. In the case e is the smallest such positive integer the resulting realization is a **symmetric minimal realization (SMR) of size** e. Any two SMRs that determine the same rational function are similar as explained in more detail in Subsection 2.1. In particular, the definitions and results here stated in terms of an SMR do not depend upon the choice of SMR. The results of [Vol17, K-VV09] justify defining the **domain** of r as

$$\operatorname{dom} r = \{(A,X) \in \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}} : \det \Big(J \otimes I - \sum_{i=1}^{\mathtt{g}} T_i \otimes X_i - \sum_{j=1}^{\mathtt{h}} S_j \otimes A_j \Big) \neq 0\}. \tag{1.4}$$

In particular, the domain of a rational function is a free open set. Let $\mathbb{C}\langle\!\langle a, x \rangle\!\rangle$ denote the set of rational functions in the variables a and x.

1.1.1. *The domain of partial convexity.* An nc rational function r is **matrix convex in** x or **partially convex** on $\mathcal{D}$ if

$$r\left(A, \frac{X+Y}{2}\right) \preceq \frac{1}{2}\left(r(A,X) + r(A,Y)\right)$$

whenever $(A,X), (A,Y), (A, \frac{X+Y}{2}) \in \mathcal{D}$. Sublevel sets of such functions have matrix convexity properties, which we do not discuss here save to note that these sublevel sets are very important in real and convex algebraic geometry, polynomial optimization, and the rapidly emerging subject of noncommutative function theory [SSS18, Pop18, PSS18, PS19, K-VV14, HM12, HL18, HKM17, HKM13b, EH19, Eve18, DDSS17, BMV16].

Our first main theorem gives an effective easily computable criterion to determine where r is convex in x. To state this result, let V_T denote the inclusion of the span of the ranges of the T_j into $\mathbb{C}^e$ and let

$$R_T = V_T^* \left(J - \sum_{i=1}^{\mathtt{g}} T_i x_i - \sum_{j=1}^{\mathtt{h}} S_j a_j \right)^{-1} V_T. \tag{1.5}$$

Finally, let

$$\operatorname{dom}^+ r := \{(A,X) \in \operatorname{dom} r : R_T(A,X) \succeq 0\}. \tag{1.6}$$

Given $\mathcal{D} \subseteq \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}}$ and $A \in \mathbb{S}^{\mathtt{h}}_k$, let

$$\mathcal{D}[A] = \{X \in \mathbb{S}^{\mathtt{g}}_k : (A,X) \in \mathcal{D}\}. \tag{1.7}$$

A free set $\mathcal{D}$ is convex (resp. open) in x if $\mathcal{D}[A]$ is convex (resp. open) for each $A \in \mathbb{S}^{\mathtt{h}}$. Theorem 1.1 below says that $\operatorname{dom}^+ r$ deserves the moniker, the *domain of partial convexity of* r. Generally, a free set $\mathcal{D}$ is a **domain of partial convexity** for r if $\mathcal{D}$ is open in x, convex in x, and r is convex in x on $\mathcal{D}$. It is a **full domain of partial convexity** if in addition $\mathcal{D}$ contains a free open set $\mathcal{U}$ with $\mathcal{U}_1 \neq \emptyset$.

Theorem 1.1. *The set* $\operatorname{dom}^+ r$ *is a domain of partial convexity for* r.

Conversely, if $\mathcal{D} \subseteq \operatorname{dom} r$ *is a full domain of partial convexity for* r*, then* $\mathcal{D} \subseteq \operatorname{dom}^+ r$ *and* $\operatorname{dom}^+ r$ *is also a full domain of partial convexity for* r.

1.1.2. *The root butterfly realization: a certificate of partial convexity.* Our second main theorem, the *root butterfly realization,* gives an algebraic certificate for partial convexity near points in the domain of r of the form $(A,0)$. This realization differs from existing realizations in that it contains a square root that appears difficult to avoid. A free set $\mathcal{D}$ is a **vertebral** set if $(A,X) \in \mathcal{D}$ implies $(A,0) \in \mathcal{D}$. We denote the positive (semidefinite) square root of a positive (semidefinite) matrix P by $\sqrt{P}$. A free set $\mathcal{D}$ is a **vertebral domain of convexity for** r provided $\mathcal{D}$ is open in x, convex in x, and if r is convex in x on $\mathcal{D}$. If in addition $\mathcal{D}$ contains a free open set $\mathcal{U}$ with $\mathcal{U}_1 \neq \emptyset$, then $\mathcal{D}$ is a **full vertebral domain of convexity**.

Theorem 1.2 (Wurzelschmetterlingrealisierung)**.** *If* $r \in \mathbb{C}\langle\!\langle a,x \rangle\!\rangle$ *is symmetric and regular at* 0*, then*

(1) *there exists a positive integer* k*, a tuple* $\widehat{T} \in M_k(\mathbb{C})^{\mathtt{g}}$*, a symmetric rational function* $w(a) \in \mathbb{C}\langle\!\langle a,x \rangle\!\rangle^{k\times k}$*, such that*

$$\operatorname{dom}^{\ddagger} r := \left\{(A,X) : w(A) \succeq 0, \quad I - \sqrt{w(A)} \left[\sum_{i=1}^{\mathtt{g}} \widehat{T}_i \otimes X_i\right] \sqrt{w(A)} \succeq 0\right\}.$$

is a vertebral domain of convexity for r;

(2) *if* $\mathcal{D}$ *is a is a full vertebral domain of convexity for* r*, then* $\mathcal{D} \subseteq \operatorname{dom}^{\ddagger} r$ *and* $\operatorname{dom}^{\ddagger} r$ *is a also a full vertebral domain of convexity for* r;

(3) *there exists a rational function* $\ell(a,x) \in \mathbb{C}\langle\!\langle a,x \rangle\!\rangle^{k\times 1}$ *that is linear in* x*, and a symmetric rational function* $f\!\!\!/(a,x) \in \mathbb{C}\langle\!\langle a,x \rangle\!\rangle$ *that is affine linear in* x *such that* r *admits the following realization:*

$$r = \ell(a,x)^* \sqrt{w(a)} \left(I - \sum \sqrt{w(a)} \widehat{T}_i x_i \sqrt{w(a)} \right)^{-1} \sqrt{w(a)}\ \ell(a,x) + f\!\!\!/(a,x). \tag{1.8}$$

As a corollary we obtain the following simple representation for polynomials that are convex in x. We use $\mathbb{C}\langle a,x\rangle$ to denote the set of noncommutative polynomials in (a,x). Given a subset $\mathcal{D} \subseteq \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}}$, let

$$\pi_a(\mathcal{D}) = \{A \in \mathbb{S}^{\mathtt{h}} : (A,X) \in \mathcal{D} \text{ for some } X \in \mathbb{S}^{\mathtt{g}}\}.$$

Corollary 1.3 ([HHLM08, Proposition 3.1])**.** *Suppose $\mathcal{D}$ is a free set that is open in x and contains a free open set $\mathcal{U}$ such that $\mathcal{U}_1 \neq \emptyset$. A polynomial $p(a,x)$ is convex in x on $\mathcal{D}$ if and only if there exists $\ell(a,x) \in \mathbb{C}\langle a,x\rangle$ that is linear in x, and a symmetric $w(a) \in \mathbb{C}\langle a\rangle$ that is positive semidefinite on $\pi_a(\mathcal{D})$ such that*

$$p = \ell(a,x)^* w(a) \ell(a,x) + \mathscr{f}(a,x), \tag{1.9}$$

where $\mathscr{f}(a,x) \in \mathbb{C}\langle a,x\rangle$ is affine linear and symmetric. In particular, if p is convex in x on $\mathcal{D}$, then p is convex in x on $\pi_a(\mathcal{D}) \times \mathbb{S}^{\mathtt{g}}$.

1.1.3. *xy-convexity and BMIs.* In this subsection we preview our results on xy-convexity and BMIs. Like partial convexity, here we have two classes of variables. Unlike partial convexity, the roles of the classes of variables appear symmetrically in xy-convexity. With that in mind, we switch notation somewhat and consider freely noncommuting letters $x_1,\dots,x_{\mathtt{g}},y_1,\dots,y_{\mathtt{h}}$.

An expression of the form

$$L(x,y) = A_0 + \sum_{j=1}^{\mathtt{g}} A_j x_j + \sum_{k=1}^{\mathtt{h}} B_k y_k + \sum_{p,q=1}^{\mathtt{g},\mathtt{h}} C_{pq} x_p y_q + \sum_{p,q=1}^{\mathtt{g},\mathtt{h}} D_{pq} y_q x_p, \tag{1.10}$$

where A_j, B_k, C_{pq}, D_{pq} are all matrices of the same size, is an xy**-pencil**. In the case A_j, B_k are hermitian and $D_{pq} = C_{qp}^*$, L is a **hermitian** xy**-pencil**. If $A_0 = I$, then L is **monic**. For a monic hermitian xy-pencil L, the inequality $L(X,Y) \succeq 0$ for $(X,Y) \in \mathbb{S}^{\mathtt{g}} \times \mathbb{S}^{\mathtt{h}}$ is a **bilinear matrix inequality** (BMI) [vAB00, GSL96, KSVS04]. Domains $\mathcal{D}$ defined by BMIs are convex in the x and y variables separately.

We say a function f of two freely noncommuting variables is xy**-convex** on a free set $\mathcal{D}$ if $f(V^*(X,Y)V) \preceq V^* f(X,Y) V$ for all isometries V, and all $X,Y \in \mathcal{D}$ satisfying $V^*(XY)V = (V^*XV)(V^*YV)$. Such a pair $((X,Y),V)$ is called an xy**-pair**. Sublevel sets of xy-convex functions are delineated by (perhaps infinitely many) BMIs as proved in [JKMMP].

Symmetric polynomials in two freely noncommuting variables x and y (so $\mathtt{g} = 1 = \mathtt{h}$) that are xy-convex essentially arise from BMIs.

Theorem 1.4. *Suppose p is a symmetric polynomial in the two freely noncommuting variables x,y. If p is xy-convex, then there exists a hermitian xy-pencil $\lambda \in \mathbb{C}\langle x,y\rangle$, a positive integer k and an xy-pencil $\Lambda \in \mathbb{C}\langle x,y\rangle^{k\times 1}$ such that*

$$p = \lambda(x,y) + \Lambda(x,y)^* \Lambda(x,y).$$

The converse is easily seen to be true.

The notions of partial convexity and xy-convexity are two instantiations of Γ-convexity [JKMMP]. Let $\mathcal{D} \subseteq \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}}$ be a given free open set that is also closed with respect to restrictions to reducing subspaces; that is if $(A,X) \in \mathcal{D}$ and V is an isometry whose range reduces each A_j and X_k, then $V^*(A,X)V \in \mathcal{D}$. The set $\mathcal{D}$ is **convex in** x, or **partially convex**, if for each $A \in \mathbb{S}_k^{\mathtt{h}}$ the slice $\mathcal{D}[A]$ (see (1.7)) is convex. Likewise $\mathcal{D}$ is a^2**-convex** if for each $(A,X) \in \mathcal{D}_n$ and isometry $V : \mathbb{C}^m \to \mathbb{C}^n$ such that $V^*A^2V = (V^*AV)^2$ it follows that $V^*(A,X)V \in \mathcal{D}$. In [JKMMP] it is shown that $\mathcal{D}$ is convex in x if and only if it is a^2-convex. A straightforward variation on the proof of that result establishes Proposition 1.5 below. A rational function $r \in \mathbb{C}\langle\!\!\langle a,x \rangle\!\!\rangle$ is a^2**-convex** on $\mathcal{D}$ if, whenever $(A,X) \in \mathcal{D}$ and $V : \mathbb{C}^m \to \mathbb{C}^n$ is an isometry such that $V^*A_j^2V = (V^*A_jV)^2$ and $V^*(A,X)V \in \mathcal{D}$, we have that

$$V^*r(A,X)V \succeq r(V^*(A,X)V).$$

Proposition 1.5. *If $\mathcal{D} \subseteq \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}}$ is a free set that is closed with respect to reducing subspaces and a^2-convex, then an $r \in \mathbb{C}\langle\!\!\langle a,x \rangle\!\!\rangle$ is a^2-convex on $\mathcal{D}$ if and only if it is convex in x on $\mathcal{D}$.*

2. Partial convexity for nc rational function

In this section we consider partial convexity and establish Theorems 1.1 and 1.2 as well as Corollary 1.3.

2.1. Preliminaries.

Proposition 2.1 below is a version of the well known state space similarity theorem due to Schützenberger [Scü61]; see also [BMG05] or [HMV06, Proposition 4.2].

Proposition 2.1. *If*

$$q(x) = a^* \left(J - \sum_{j=1}^m A_j x_j \right)^{-1} a, \quad s(x) = b^* \left(K - \sum_{j=1}^m B_j x_j \right)^{-1} b$$

are two SMRs for the same rational function, then there is a unique matrix S such that $S^*KS = J$, $SJA_j = KB_jS$ for $1 \le j \le m$ and $SJa = Kb$.

A bit of algebra reveals that $S^*BS = A$. Thus $K - \sum B_jx_j = S^*(J - \sum A_jx_j)S$ and it follows that the definitions of $\operatorname{dom} r$, $\operatorname{dom}^+ r$ and $\operatorname{dom}^\ddagger r$ are independent of the choice of SMR.

Just as in the commutative case, it is well known that convexity properties of a free rational functions can be characterized by positivity of a Hessian. See for instance [HM98]. The x**-partial Hessian** of an SMR as in equation (1.3) is the rational function in $2\mathtt{g} + \mathtt{h}$ freely noncommuting variables,

$$\begin{aligned} r_{xx}(x,a)[h] &= 2c^*R(a,x)(\sum_i T_ih_i)R(a,x)(\sum_i T_ih_i)R(a,x)c \\ &= 2\left[c^*R(a,x)(\sum_i T_ih_i)\right] R_T(a,x) \left[(\sum_i T_ih_i)R(a,x)c\right], \end{aligned} \tag{2.1}$$

where R is the resolvent

$$R(a,x) := (J - \sum T_jx_j - \sum S_ka_k)^{-1}, \tag{2.2}$$

$\Lambda_T[h] = \sum_{j=1}^{\mathtt{g}} T_jh_j$, and $R_T(a,x) = V_T^*R(a,x)V_T$ is defined as in (1.5). Compare with [HMV06, Equation (5.3)] where the *full* Hessian of a SMR is computed in detail. The x-partial Hessian is naturally **evaluated** at a tuple $(A,X,H) \in \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}} \times \mathbb{S}^{\mathtt{g}}$ where $(A,X) \in \operatorname{dom} r$ with output a symmetric $k \times k$ matrix.

Proposition 2.2 is the partial convexity analog of the [HM98] characterization of convexity in terms of Hessians. The proof is a straightforward modification of the one in [HM98] so is only sketched below.

Proposition 2.2. *The rational function r is convex in x on a nonempty open in x set $S \subseteq \operatorname{dom} r \cap (\mathbb{S}_k^{\mathtt{h}} \times \mathbb{S}_k^{\mathtt{g}})$ if and only if $r_{xx}(A,X)[H] \succeq 0$ for all $(A,X) \in S$ and $H \in \mathbb{S}_k^{\mathtt{g}}$.*

Sketch of proof. The rational function r is convex in x on S if and only if for each $A \in \mathbb{S}_k^{\mathtt{h}}$ and each real linear functional $\lambda : \mathbb{S}_k \to \mathbb{R}$ the function $f_{A,\lambda} : S \to \mathbb{R}$ defined by $f_{A,\lambda}(X) = \lambda \circ r(A,X)$ is convex. On the other hand, $f_{A,\lambda}$ is convex if and only if its Hessian is positive; that is

$$0 \leq f''_{A,\lambda}(X)[H] = \lambda \circ r_{xx}(A,X)[H]$$

for all H. Thus $f_{A,\lambda}$ is convex for each A and λ if and only if $r_{xx}(A,X)[H] \succeq 0$. ■

2.2. **Characterization of partial convexity.** Throughout this section we fix an SMR (1.3) for r, and let $R(a,x)$ denote the resolvent of equation (2.2). Recall the definitions of R_T and $\operatorname{dom}^+ r$ of equations (1.5) and (1.6).

Theorem 2.3. *If $r \in \mathbb{C}\langle\!\!\langle a,x \rangle\!\!\rangle$ is a nc rational function with the SMR as in (1.3), then*

1. *if $(B,Z) \in \operatorname{dom}^+ r$ and $(B,Y),(B,Z)$ lie in the same connected component of $\operatorname{dom} r$, then $(B,Y) \in \operatorname{dom}^+ r$;*
2. *$\operatorname{dom}^+ r$ is a domain of partial convexity for r;*
3. *if $\mathcal{D} \subseteq \operatorname{dom} r$ is a full domain of partial convexity for r, then $\mathcal{D} \subseteq \operatorname{dom}^+ r$.*

Corollary 2.4 ([HMV06])**.** *Suppose $r \in \mathbb{C}\langle\!\!\langle x \rangle\!\!\rangle$. If r is convex in a free open set containing 0, then $\operatorname{dom}_0 r$, the component of $\operatorname{dom} r$ containing 0, is convex and r is convex on $\operatorname{dom}_0 r$.*

It is straightforward to verify that $\operatorname{dom}^+ r$ is a free set. Item (1) and that $\operatorname{dom}^+ r$ is open in x and convex in x are proven in Subsection 2.2.3. That r is convex in x on $\operatorname{dom}^+ r$ (and hence $\operatorname{dom}^+ r$ is a domain of partial convexity for r) is a consequence of Proposition 2.5 below. Item (3) is an immediate consequence of the converse portion of Proposition 2.5.

Proposition 2.5. *Let r denote the rational function r of (1.3) and suppose $\mathcal{E} \subseteq \operatorname{dom} r \subseteq \mathbb{S}^{\mathtt{h}+\mathtt{g}}$ is a free set that is open in x.*

If $R_T \succeq 0$ on $\mathcal{E}$, then r is convex in x on $\mathcal{E}$. Conversely, if $\mathcal{E}$ contains a free open set $\mathcal{U}$ with $\mathcal{U}_1 \neq \emptyset$, and if r is convex in x on $\mathcal{E}$, then $R_T \succeq 0$ on $\mathcal{E}$.

2.2.1. *The CHSY lemma.* In this subsection we establish a variant of the CHSY Lemma [CHSY03] (see also [BK13, Vol18]) suitable for a proof of Proposition 2.5, starting with the of independent interest Lemma 2.6 below.

Lemma 2.6. *If $\xi_1,\dots,\xi_K \in \mathbb{C}\langle\!\!\langle x \rangle\!\!\rangle$ are linearly independent rational functions in $\mathtt{g}$ variables, m is a positive integer and $\mathcal{U}$ is a free open subset of $\mathbb{S}^{\mathtt{g}}$ with $\mathcal{U}_1 \neq \emptyset$, then there exists a positive integer M, an $X \in \mathcal{U}(M)$ and a matrix $w \in M_{m,M}(\mathbb{C})$ such that*

$$\{\begin{pmatrix} w\,\xi_1(X)v \\ \vdots \\ w\,\xi_K(X)v \end{pmatrix} : v \in \mathbb{C}^n\} = \mathbb{C}^K \otimes \mathbb{C}^m = \mathbb{C}^{Km}.$$

Proof. Let $\Xi = \operatorname{col}(\xi_1, \dots, \xi_K) \in M_{1,K}(\mathbb{C}\langle\!\langle x \rangle\!\rangle)$. Let $\mathcal{S}$ denote the set of pairs (Y, z), where, for some n, $Y \in \mathcal{U}_n$ and $z \in M_{m,n}(\mathbb{C})$. Given $(Y, z) \in \mathcal{S}_n$, let

$$\mathcal{V}_{(Y,z)} = \{(I_K \otimes z)\Xi(Y)v : v \in \mathbb{C}^n\} \subseteq \mathbb{C}^K \otimes \mathbb{C}^m.$$

Given $A = (Y, z)$ and $\widetilde{A} = (\widetilde{Y}, \widetilde{z})$ both in $\mathcal{S}$, let

$$A \oplus \widetilde{A} = \left(\begin{pmatrix} Y & 0 \\ 0 & \widetilde{Y} \end{pmatrix}, \begin{pmatrix} z \\ \widetilde{z} \end{pmatrix} \right).$$

It is straightforward to verify that $\mathcal{V}_{A\oplus\widetilde{A}} = \mathcal{V}_A + \mathcal{V}_{\widetilde{A}}$. Hence, there exists a (dominating) pair $(X, w) \in \mathcal{S}$ such that

$$\mathcal{V}_{(Y,z)} \subseteq \mathcal{V}_{(X,w)}, \tag{2.3}$$

for all $(Y, z) \in \mathcal{S}$. Suppose $\alpha \in \mathcal{V}^{\perp}_{(X,w)}$. From equation (2.3), it follows that $\alpha \in \mathcal{V}^{\perp}_{(Y,z)}$ for all $(Y, z) \in \mathcal{S}$. Write $\alpha \in \mathbb{C}^K \otimes \mathbb{C}^m$ as $\alpha = \sum \alpha_j \otimes e_j$, where $\{e_1, \dots, e_m\}$ is the standard orthonormal basis for $\mathbb{C}^m$ and $\alpha_j \in \mathbb{C}^K$. We will show, for each j, that $\sum_{s=1}^K \overline{(\alpha_j)_s}\xi_s = 0$, and hence, by the linear independence assumption, that each α_j, and hence α, is zero. Accordingly, fix j and let n and $Y \in \mathcal{U}_n$ be given. Given a vector $f \in \mathbb{C}^n$, let $w_f = e_j f^*$. Since $\alpha \in \mathcal{V}^{\perp}_{(Y,w_f)}$,

$$0 = \alpha^*[I_K \otimes w_f]\Xi(Y) = (\alpha_j^* \otimes f^*)\Xi(Y) = f^* \sum_{s=1}^K \overline{(\alpha_j)_s}\xi_s(Y).$$

Thus, for each j, the rational function $\sum_{s=1}^K \overline{(\alpha_j)_s}\xi_s$ vanishes on the free open set $\mathcal{U}$ and is thus identically zero (since there are not rational identities), and the desired conclusion follows. ■

Lemma 2.7. *If the realization (1.3) is minimal and of size N and $\mathcal{U}$ is a free open subset of* $\operatorname{dom} r$, *then, for each $m \in \mathbb{N}$, there exists an M, $(A, X) \in \mathcal{U}$, a $w \in M_{m,M}(\mathbb{C})$ and an $H \in \mathbb{S}^{\mathtt{g}}_M$ such that*

$$V_{A,X,H,w} := \{(I_N \otimes w)(\sum_i T_i \otimes H_i)R(A, X)(c \otimes I_n)v \mid H \in \mathbb{S}^{\mathtt{g}}_n,\ v \in \mathbb{C}^n\} = (\operatorname{rng} T) \otimes \mathbb{C}^m.$$

Proof. Let K denote the dimension of $\operatorname{rng} T$ and U a unitary matrix mapping $\operatorname{rng} T$ into the first K coordinates of $\mathbb{C}^N$. The entries η_j of the $N \times 1$ matrix $R(a, x)c$ are linearly independent nc rational functions by minimality of (1.3) and hence so are the entries of the $\mathtt{g}\,N \times 1$ matrix

$$Q(a, x, h) := \begin{pmatrix} h_1 R(a, x)c \\ \vdots \\ h_{\mathtt{g}} R(a, x)c \end{pmatrix}.$$

Thus there are $\xi_j \in \mathbb{C}\langle\!\langle h, a, x \rangle\!\rangle$ such that

$$\sum T_i h_i R(a, x)c = \left[\begin{pmatrix} T_1 & \cdots & T_{\mathtt{g}} \end{pmatrix}\right] Q(a, x, h) = U^* \operatorname{col}(\xi_1, \cdots, \xi_K, 0, \cdots, 0).$$

Further, since the entries of Q are linearly independent, the set $\{\xi_1, \dots, \xi_K\}$ is linearly independent. By Lemma 2.6, for each positive integer m, there exists a positive integer M, a tuple $(H, A, X) \in \mathbb{S}^{\mathtt{g}}_M \times \mathcal{U}_M$ and a matrix $w \in M_{M,m}(\mathbb{C})$ such that the conclusion of Lemma 2.6 holds, completing the proof. ■

2.2.2. *Proof of Proposition 2.5.* Observe that, from equation (2.1) it is evident that the inequality $R_T \succeq 0$ on $\mathcal{E}$ implies r_{xx} is positive semidefinite on $\mathcal{E}$, equivalently r is convex in x on $\mathcal{E}$ by Proposition 2.2.

Now suppose r_{xx} is positive semidefinite on $\mathcal{E}$. To prove that the inequality $R_T \succeq 0$ holds on $\mathcal{E}$, disaggregate the variables. That is, let

$$x_i = \begin{pmatrix} x_i^1 & 0 \\ 0 & x_i^2 \end{pmatrix}, \quad h_i = \begin{pmatrix} 0 & k_i \\ k_i^* & 0 \end{pmatrix}, \quad a_i = \begin{pmatrix} a_i^1 & 0 \\ 0 & a_i^2 \end{pmatrix}.$$

In these coordinates the $(1,1)$ entry of r_{xx} in (2.1) equals

$$2\left[c^* R(a^1,x^1)(\sum_i T_i k_i)\right] R(a^2,x^2) \left[(\sum_i T_i(k_i)^*)R(a^1,x^1)c\right]. \tag{2.4}$$

We next apply Lemma 2.7. Given a positive integer m and $(A^2,X^2) \in \mathcal{E}(m)$, choose M and $(A^1,X^1) \in \mathcal{U}(M)$, $w \in M_{m,M}(\mathbb{C})$ and $H \in \mathbb{S}_M^{\mathtt{g}}$ satisfying the conclusion of Lemma 2.7. Thus $(A,X) = (A^1 \oplus A^2, X^1 \oplus X^2) \in \mathcal{E}(m+M)$ and hence $r_{xx}(A,X)[H] \succeq 0$. Choose $K = wH \in M_{m,M}(\mathbb{C})$. Substituting into (2.4) and observing that $\{[\sum T_j \otimes K_j]R(A^1,X^1)(c \otimes I) : v \in \mathbb{C}^n\}$ spans $\operatorname{rng} T \oplus \mathbb{C}^m$, it now follows that $R_T(A^2,X^2) \succeq 0$. ■

2.2.3. *Convexity in x and inverses of structured pencils.* Items (1) and (2) are established in this subsection, completing the proof of Theorem 2.3.

Let N denote the size of realization. Thus $J \in M_N(\mathbb{C})$. Express J, S, T with respect to the orthogonal decomposition $\mathbb{C}^N = \operatorname{rng} T \oplus (\operatorname{rng} T)^\perp$ as

$$J = \begin{pmatrix} J_0 & J_1 \\ J_1^* & J_2 \end{pmatrix}, \quad S_k = \begin{pmatrix} S_{k,0} & S_{k,1} \\ S_{k,1}^* & S_{k,2} \end{pmatrix}, \quad T_j = \begin{pmatrix} T_{j,0} & 0 \\ 0 & 0 \end{pmatrix}.$$

Let

$$\begin{aligned} L(a,x) = J - \sum T_j x_j - \sum S_k a_k &= \begin{pmatrix} J_0 - \sum T_{j,0}x_j - \sum S_{k,0}a_k & J_1 - \sum S_{k,1}a_k \\ J_1^* - \sum S_{k,1}^* a_k & J_2 - \sum S_{k,2}a_k \end{pmatrix} \\ &= \begin{pmatrix} L_0(a,x) & L_1(a) \\ L_1(a)^* & L_2(a) \end{pmatrix}. \end{aligned}$$

Proof of Theorem 2.3(1). Let $\operatorname{dom} r[n] = \operatorname{dom} r \cap (\mathbb{S}_n^{\mathtt{h}} \times \mathbb{S}_n^{\mathtt{g}})$.

Given a real number τ, let

$$L^\tau(a,x) = \begin{pmatrix} L_0(a,x) & L_1(a) \\ L_1(a)^* & L_2(a) - \tau I \end{pmatrix}, \quad R^\tau(a,x) = L^\tau(a,x)^{-1} = \begin{pmatrix} R_0^\tau(a,x) & R_1^\tau(a,x) \\ R_1^\tau(a,x)^* & R_2^\tau(a,x) \end{pmatrix},$$

when this inverse exists. Note that $R_0^0(a,x) = R_T(a,x)$.

For a given $(A,X) \in \mathbb{S}_n^{\mathtt{h}+\mathtt{g}}$, it is well known that if $L^\tau(A,X)$ and $L_2(A) - \tau I$ are invertible, then so is the Schur complement [LM00]

$$S_\tau(A,X) := L_0(A,X) - L_1(A)(L_2(A) - \tau I)^{-1} L_1(A)^*$$

and $R_0^\tau(A,X) = S_\tau(A,X)^{-1}$. Further that the signature of L^τ equals the signature of R^τ equals the amalgam of the signatures of $L_2(A) - \tau I$ and $S_\tau(A,X)$; that is, letting $\varepsilon_\pm(C)$

denote the number of positive and negative eigenvalues of a hermitian matrix C,

$$\varepsilon_\pm(L^\tau(A,X)) = \varepsilon_\pm(L_2(A)-\tau) + \varepsilon_\pm(S_\tau(A,X)).$$

Moreover, for a given n the signature of $L^\tau(A,X)$ is constant on any component of $\operatorname{dom} r[n]$ since such components are precisely the components of the domain of R as a function on $\mathbb{S}_n^{\mathtt{h}+\mathtt{g}}$.

Now suppose $\mathcal{C}$ is a connected component of $\operatorname{dom} r[n]$, $(B,Y),(B,Z)\in\mathcal{C}$ and $R_T(B,Z)\succeq 0$. Let $\gamma:[0,1]\to\mathcal{C}$ be a continuous function with $\gamma(0)=(B,Z)$ and $\gamma(1)=(B,Y)$. Because the range of γ is a compact subset of $\operatorname{dom} r[n]$, there exists a $\delta>0$ such that for each $0<\tau<\delta$,

(a) $L^\tau(\gamma(s))$ is invertible for each $s\in[0,1]$ and each $L^\tau(\gamma(s))$ has the same signature as L does on $\mathcal{C}$; and
(b) $L_2^\tau(B) = L_2(B)-\tau I$ is invertible.

Since the Schur complement is matrix monotone [LM00] and $L^\tau(B,W)\preceq L^d(B,W)$,

$$R_0^\tau(B,W)\succeq R_0^d(B,W)$$

for $0<d<\tau<\delta$ and $W=Y,Z$. Since also $\lim_{d\to 0^+} R_0^d(B,Z) = R_T(B,Z)\succeq 0$, it follows that $R_0^\tau(B,Z)\succeq 0$. Further, since $L^\tau(B,Z)$ and $L_2^\tau(B)$ are invertible, $R_0^\tau(B,Z)\succ 0$. On the other hand, $R^\tau(B,Y)$ and $R^\tau(B,Z)$ have the same signature and these are the amalgams of the signatures of $R_0^\tau(B,Y)$ and $L_2(B)$ and of $R_0^\tau(B,Z)$ and $L_2(B)$ respectively, it follows that the signatures of $R_0^\tau(B,Y)$ and $R_0^\tau(B,Z)$ are the same. Thus $R_0^\tau(B,Y)\succ 0$. Letting τ tend to 0 (through positive values), it follows that $R_T(B,Y)\succeq 0$ and thus $(B,Y)\in\operatorname{dom}^+ r$. ■

Since $\operatorname{dom}^+ r$ is open in x follows from item (1) of Theorem 2.3, Lemma 2.8 below certifies item (2) of Theorem 2.3.

Lemma 2.8. *Suppose m,n are positive integers, $\phi_0,\phi_1\in\mathbb{S}_n$, $\Psi\in M_{n,m}(\mathbb{C})$ and $\Gamma\in\mathbb{S}_m$. For λ real, let $\Phi(\lambda)=\phi_0+\lambda\phi_1$,*

$$L(\lambda) = \begin{pmatrix}\Phi(\lambda) & \Psi\\ \Psi^* & \Gamma\end{pmatrix} \qquad \text{and} \qquad L(\lambda)^{-1} = \begin{pmatrix} M_0(\lambda) & M_1(\lambda)\\ M_1(\lambda)^* & M_2(\lambda)\end{pmatrix},$$

when this inverse exists.

If $L(0)$ and $L(1)$ are invertible and $M_0(0),M_0(1)\succeq 0$, then $L(\lambda)$ is invertible and $M_0(\lambda)\succeq 0$ for $0\le\lambda\le 1$.

Proof. Let $\mathcal{K}$ the kernel of Γ and $\mathcal{R}=\Psi(\mathcal{K})$. Express $L(\lambda)$ as a 4×4 block matrix with respect to the decomposition $\mathbb{C}^n\oplus\mathbb{C}^m = [\mathcal{R}\oplus\mathcal{R}^\perp]\oplus[\mathcal{K}\oplus\mathcal{K}^\perp]$,

$$L(\lambda) = \begin{pmatrix} \Phi_{11}(\lambda) & \Phi_{12}(\lambda) & \Psi_{1,1} & \Psi_{1,2}\\ \Phi_{12}(\lambda)^* & \Phi_{22}(\lambda) & 0 & \Psi_{2,2}\\ \Psi_{1,1}^* & 0 & 0 & 0\\ \Psi_{1,2}^* & \Psi_{2,2}^* & 0 & D\end{pmatrix},$$

where D is invertible. Thus $L(\lambda)$ is invertible if and only if the Schur complement of the upper 3×3 block,

$$S(\lambda) = \left(\begin{array}{cc|c} \Omega_{1,1}(\lambda) & \Omega_{1,2}(\lambda) & \Psi_{1,1}\\ \Omega_{1,2}(\lambda)^* & \Omega_{2,2}(\lambda) & 0\\ \hline \Psi_{1,1}^* & 0 & 0\end{array}\right)$$

is invertible, where $\Omega_{j,k} = \Phi_{j,k} - \Psi_{j,2} D^{-1} \Psi_{k,2}^*$. By construction $\Psi_{1,1}$ is onto. On the other hand, if $S(\lambda)$ is invertible, then $\Psi_{1,1}$ is one-one. Thus $S(\lambda)$ is invertible if and only if $\Psi_{1,1}$ and $\Omega_{2,2}(\lambda)$ are both invertible. Since, by hypothesis, $L(0)$ is invertible, it follows that $\Psi_{1,1}$ is invertible (and in particular the dimensions of $\mathcal{K}$ and $\mathcal{R}$ are the same). Hence $L(\lambda)$ is invertible if and only if $\Omega_{2,2}(\lambda)$ is invertible. Further, in that case,

$$L(\lambda)^{-1} = \left(\begin{array}{c|c} S(\lambda)^{-1} & * \\ \hline * & * \end{array}\right) = \left(\begin{array}{ccc|c} 0 & 0 & \Psi_{1,1}^{-*} & * \\ 0 & \Omega_{2,2}(\lambda)^{-1} & * & * \\ \Psi_{1,1}^{-1} & * & * & * \\ \hline * & * & * & * \end{array}\right).$$

In particular, when $L(\lambda)$ is invertible,

$$M_0(\lambda) = \begin{pmatrix} 0 & 0 \\ 0 & \Omega_{2,2}(\lambda)^{-1} \end{pmatrix}.$$

Since, by assumption, $M_0(0) \succeq 0$ and $L(0)$ is invertible, it follows that $\Omega_{2,2}(0) \succ 0$. Similarly, $\Omega_{2,2}(1) \succ 0$. Since $\Omega(\lambda)$ is affine linear, it follows that $\Phi_{2,2}(\lambda) \succ 0$ for $0 \le \lambda \le 1$. Thus $\Omega_{2,2}(\lambda)$, and therefore $L(\lambda)$ is invertible and $M_0(\lambda) \succeq 0$ for $0 \le \lambda \le 1$. ∎

2.3. Realizations for partial convexity.

Proposition 2.9. *The rational function $r \in \mathbb{C}\langle\!\!\langle a, x \rangle\!\!\rangle$ of equation (1.3) admits the realization*

(2.5)
$$\begin{aligned} r = {} & c^*(J - \sum S_i a_i)^{-1} c + c^*(J - \sum S_i a_i)^{-1} \sum T_i x_i \, (J - \sum S_i a_i)^{-1} c \\ & + c^*(J - \sum S_j a_j)^{-1} \sum T_i x_i \left(J - \sum T_j x_j - \sum S_k a_k\right)^{-1} \sum T_i x_i (J - \sum S_i a_i)^{-1} c. \end{aligned}$$

We will refer to a realization of the form (2.5) as a **caterpillar realization**.

Proof. Formula (2.5) follows from a routine calculation. ∎

The free set $\mathcal{N}_\epsilon = \{X \in \mathbb{S}^g \mid \|X\| \preceq \epsilon\}$ is a **free ball** about the origin. Recall the definitions of V_T and $\pi_a(\mathcal{D})$ from Theorem 1.2.

Theorem 2.10 (Wurzelschmetterlingrealisierung)**.** *Suppose $r \in \mathbb{C}\langle\!\!\langle a, x \rangle\!\!\rangle$ is symmetric with SMR of size N as in equation (1.3). Let $\widehat{T}_j = V_T^* T_j V_T$ and let k denote the dimension of $\operatorname{rng} T$.*

There exists a rational function $w(a) \in M_k(\mathbb{C}\langle\!\!\langle a, x \rangle\!\!\rangle)$, and rational functions $\ell_j(a) \in \mathbb{C}\langle\!\!\langle a, x \rangle\!\!\rangle^k$ for $1 \le j \le g$, and an affine linear in x rational function $\not{f}(a, x)$ such that, with

$$\ell(a, x) = \sum x_j \ell_j(a), \tag{2.6}$$

1. *if $(B, 0), (B, Y) \in \operatorname{dom} r$, then $I - (\sum T_j \otimes Y_j) w(B)$ is invertible and*

$$r(B, Y) = \ell(B, Y)^* w(B) \left(I - (\sum \widehat{T}_i \otimes Y_i) w(B)\right)^{-1} \ell(B, Y) + \not{f}(B, Y);$$

(2) *if* $(B,Y) \in \operatorname{dom} r$ *and* $w(B)$, $I - \sqrt{w(B)}\,[\sum \widehat{T}_j \otimes Y_j]\,\sqrt{w(B)} \succeq 0$, *then*

$$I - \sqrt{w(B)}\,[\sum \widehat{T}_j \otimes Y_j]\,\sqrt{w(B)} \succ 0;$$

(3) *the set*

$$\operatorname{dom}^\ddagger r = \{(A,X) \in \operatorname{dom} r : w(A) \succeq 0 \text{ and } I - \sqrt{w(A)}\,[\sum \widehat{T}_j \otimes X_j]\,\sqrt{w(A)} \succeq 0\}$$

is a vertebral domain of convexity for r *and*

$$r|_{\operatorname{dom}^\ddagger r}(a,x) = \ell(a,x)^* \sqrt{w(a)}\,\Big(I - \sqrt{w(a)}\,\sum \widehat{T}_i x_i\,\sqrt{w(a)}\Big)^{-1}\sqrt{w(a)}\,\ell(a,x) + \mathscr{f}(a,x); \tag{2.7}$$

(4) *If* $\mathcal{D} \subseteq \operatorname{dom} r$ *is a full vertebral domain of convexity for* r, *then* $\mathcal{D} \subseteq \operatorname{dom}^\ddagger r$*;*

(5) *If* r *is a polynomial and* $\mathcal{D}$ *is a full vertebral domain of convexity for* r, *then*
   (a) $\mathscr{f}$, w, ℓ *are also polynomials;*
   (b) r *has the representation,*

$$r(a,x) = \ell(a,x)^* w(a)\ell(a,x) + \mathscr{f}(a,x), \tag{2.8}$$

and hence r *is convex in* x *on* $\pi_a(\mathcal{D}) \times \mathbb{S}^{\mathrm{g}}$ *and has degree at most two in* x.

Conversely, any (rational) function of the form (2.7) *is convex in* x *on the set* $\operatorname{dom}^\ddagger r$ *and any polynomial of the form of equation* (2.8) *is convex on the free strip* $\{A \in \mathbb{S}^{\mathrm{h}} : w(A) \succeq 0\} \times \mathbb{S}^{\mathrm{g}}$.

Given the symmetric realization (1.3), express the matrices T_i, S_j as block 2×2 matrices with respect to the orthogonal decomposition $\operatorname{rng} T \oplus \operatorname{rng} T^\perp$ as

$$T_i = \begin{pmatrix} \widehat{T}_j & 0 \\ 0 & 0 \end{pmatrix}, \quad S_i = \begin{pmatrix} S^i_{11} & S^i_{12} \\ S^{i*}_{12} & S^i_{22} \end{pmatrix}, \quad J = \begin{pmatrix} J_{11} & J_{12} \\ J^*_{12} & J_{22} \end{pmatrix}. \tag{2.9}$$

The proof of Theorem 2.10 will also use the following elementary fact.

Lemma 2.11. *If* $P, Q \in \mathbb{S}_k$ *and* $I - QP^2$ *is invertible and* $I - PQP \succeq 0$, *then* $I - PQP \succ 0$.

Proof. Since $I - QP^2$ is invertible and P and Q are symmetric, $I - P^2Q$ is also invertible. Suppose $PQPx = x$. It follows that $P^2QPx = Px$ and hence $(I - P^2Q)Px = 0$. Since $I - P^2Q$ is invertible, $Px = 0$ and therefore $x = PQPx = 0$. Hence $I - PQP$ is both invertible and positive semidefinite. Thus $I - PQP \succ 0$. ∎

Proof of Theorem 2.10. By Proposition 2.9, r admits the caterpillar realization (2.5) and the resolvent

$$R(a,x) = \begin{pmatrix} J_{11} - \sum T^i_{11} x_i - \sum S^i_{11} a_i & J_{12} - \sum S^i_{12} a_i \\ J^*_{12} - \sum S^{i*}_{12} a_i & J_{22} - \sum S^i_{22} a_i \end{pmatrix}^{-1} \tag{2.10}$$

is defined on the domain of r, which includes $(0,0)$. We obtain a free rational function $W(a) = R(a,0) \in \mathbb{C}\langle\!\langle a,x\rangle\!\rangle$. Let $w(a) = V_T^* R(a,0) V_T$ denote the (block) $(1,1)$-entry of $W(a)$. Thus $\operatorname{dom} w \supseteq \operatorname{dom} W \supseteq \operatorname{dom} r$. Likewise the domain of the rational function

$$\ell(a,x) = V_T^* \sum T_i x_i W(a) c$$

contains $\operatorname{dom} W$.

If $(A,0),(A,X) \in \operatorname{dom} r$, then $A \in \operatorname{dom} W$, and hence

$$\begin{aligned}\Big(J - \sum T_j \otimes X_j - \sum S_k \otimes A_k\Big) W(A) &= I - \Big(\sum T_j \otimes X_j\Big) W(A) \\ &= \begin{pmatrix} I - \sum \widehat{T}_j x_j w(a) & * \\ 0 & I \end{pmatrix}.\end{aligned}$$

It follows that $I - (\sum \widehat{T}_j \otimes X_j) w(A)$ is invertible whenever $(A,0),(A,X) \in \operatorname{dom} r$, establishing the first half of item (1). Moreover, in that case,

$$R_T(a,x) = V_T^* R(a,x) V_T = w(a)\Big(I - (\sum \widehat{T}_i x_i) w(a)\Big)^{-1}. \tag{2.11}$$

Thus,

$$r(A,X) = \ell(A,X)^* w(A)\Big(I - (\sum \widehat{T}_i \otimes X_i) w(A)\Big)^{-1} \ell(A,X) + \mathfrak{f}(A,X),$$

when $(A,0),(A,X) \in \operatorname{dom} r$, proving item (1).

Item (2) is an immediate consequence of Lemma 2.11. That $\operatorname{dom}^\ddagger r$ is a free vertebral set follows from item (2). To prove the rest of this item, suppose $(A,X) \in \operatorname{dom}^\ddagger r$. By item (1) $I - (\sum T_j \otimes X_j) w(A)$ is invertible and by item (2) $I - \sqrt{w(A)}\Big(\sum \widehat{T}_i \otimes X_i\Big)\sqrt{w(A)}$ is invertible. For t small a power series argument shows

$$w(A)\Big(I - t(\sum \widehat{T}_i \otimes X_i) w(A)\Big)^{-1} = \sqrt{w(A)}\Big(I - t\sqrt{w(A)}(\sum \widehat{T}_i \otimes X_i)\sqrt{w(A)}\Big)^{-1}\sqrt{w(A)}.$$

Since both sides are rational functions in t, it follows that they agree where both are defined and thus, in particular, at $t = 1$ and the realization of equation (2.7) is established. Since $\operatorname{dom}^\ddagger r \subseteq \operatorname{dom}^+ r$, the convexity of r on $\operatorname{dom}^\ddagger r$ follows from Theorem 2.3 and the proof of item (3) is complete.

Suppose $\mathcal{D}$ is a full vertebral domain of convexity for r. In this case, given $(A,X) \in \mathcal{D}$ it follows that $(A,0) \in \mathcal{D}$. Since r is convex in x on $\mathcal{D}$, Proposition 2.5 and equation (2.11) together imply both $R_T(A,0) = w(A) \succeq 0$ and $R_T(A,X) \succeq 0$. It follows that equation (2.11) that $\Big(I - w(A)\sum \widehat{T}_i \otimes X_i\Big) w(A) \succeq 0$ and thus $(A,X) \in \operatorname{dom}^\ddagger r$ proving item (4).

In the case r is a polynomial, $R(a,x)$ is globally defined (has no singularities) and is therefore a (matrix-valued) polynomial by [KV17, Corollary 3.4]. Hence both $w(a)$ and $\ell(a,x)$ are polynomials. By hypothesis, there is a free open set $\mathcal{U} \subseteq \mathcal{D}$ with $\mathcal{U}_1 \neq \emptyset$. Choose a point $(\mathtt{a},\mathtt{x}) \in \mathcal{U}_1 \subseteq \mathbb{R}^{\mathtt{h}} \times \mathbb{R}^{\mathtt{g}}$ and consider the polynomial $q(a,x) = r(a - \mathtt{a}, x)$. Let $\mathcal{D}' = \{(A - \mathtt{a}I, X) : (A,X) \in \mathcal{D}\}$. If $(A,X) \in \mathcal{D}'$, then $(A - \mathtt{a}I, X) \in \mathcal{D}$ and hence $(A - \mathtt{a}, 0) \in \mathcal{D}$ and finally $(A,0) \in \mathcal{D}'$. Thus $\mathcal{D}'$ is a vertebral domain of partial convexity for q. Hence, without loss of generality, we assume from the outset that $(0,0) \in \mathcal{D}$. Next $w(0) = V_T^* R(0,0) V_T$ is positive semidefinite by Theorem 2.3 since we have now convexity in x in a neighborhood of 0. Next $R(0,0) = J^{-1} = J$ and so $w(0) = J_0 \succeq 0$. Since r is a polynomial (and the realization is minimal), TJ is (jointly) nilpotent by [KV17, Corollary 3.4]. But

$$TJ = \begin{pmatrix} \hat{T} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} J_0 & J_1 \\ J_1^* & J_2 \end{pmatrix} = \begin{pmatrix} \hat{T} J_0 & \hat{T} J_1 \\ 0 & 0 \end{pmatrix},$$

whence $\hat{T}J_1$ is (jointly) nilpotent. Thus $Y = \sqrt{J_1}T_j\sqrt{J_1}$ is self-adjoint and nilpotent and hence 0. Thus, from equation (2.7), r has the representation of equation (2.8). From this representation it is immediate that r has degree (at most) two and is convex in x on the set $\{(A,X) : w(A) \succeq 0\}$, which includes $\pi_a(\mathcal{D}) \times \mathbb{S}^{\mathrm{g}}$. ∎

Corollary 2.12. *Let $\mathcal{D}$ be a vertebral set. Let $r \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle$ be a nc rational function in two classes of variables $x = (x_1, \dots, x_{\mathrm{g}})$ and $a = (a_1, \dots, a_{\mathrm{h}})$. Let r have a SMR (1.3). Consider the matrices in block form based on* $\operatorname{rng} T$ *in equation (2.9) and let k denote the dimension of* $\operatorname{rng} T$.

If J_2 is invertible, then the function r is convex in x on $\mathcal{D}$ if and only if there exists a rational function $\ell(a,x) \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle^{k\times 1}$ that is linear in x, and a rational function $m(a) \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle^{k\times k}$ such that

$$r = \ell(a,x)^* \left(m(a) - \sum \widehat{T}_i x_i\right)^{-1} \ell(a,x) + \mathbf{\not{f}}(a,x),$$

where $\mathbf{\not{f}}(a,x) \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle$ is affine linear in x, and the resolvent $(m(a) - \sum \widehat{T}_i x_i)^{-1}$ is positive on a dense subset of $\mathcal{D}_n$ for large n.

Proof. This result follows by using the Schur complement form for the inverse of a block matrix in Proposition 2.9, the positivity condition follows from Proposition 2.5. ∎

3. A polynomial factorization

In this section we introduce an auxiliary operation $\mathscr{E}$ on both matrices and polynomials and in Theorem 3.3 provide a decomposition of symmetric polynomials $\rho \in M_2(\mathbb{C}\langle x, y\rangle)$ for which $\mathscr{E}\rho$ is (matrix) positive. This result is a key ingredient in the proof of Theorem 1.4 characterizing xy-convex polynomials in Section 4.

Given a pair of block 2×2 matrices $A = (A_{i,j})$ and $B = (B_{i,j})$ define

$$A \circledast B = \left(A_{i,j} \otimes B_{i,j}\right).$$

Thus $A \circledast B$ is a mix of Schur product $(*)$ and tensor product $(\otimes)$. Let $V_1 = \begin{pmatrix} I \\ 0 \end{pmatrix}$ and $V_2 = \begin{pmatrix} 0 \\ I \end{pmatrix}$ with respect to the block decomposition of A and define W_1, W_2 similarly with respect to the block decomposition of B. Let

$$E = \begin{pmatrix} V_1 \otimes W_1 & V_2 \otimes W_2 \end{pmatrix}.$$

Lemma 3.1. *With notation as above, $A \circledast B = E^*[A \otimes B]E$.*

Proof. Note that

$$E^*[A \otimes B]E = \left((V_j^* \otimes W_j^*)[A \otimes B](V_k \otimes W_k)\right)_{j,k=1}^2$$

and $(V_j^* \otimes W_j^*)[A \otimes B](V_k \otimes W_k) = A_{jk} \otimes B_{jk}$. ∎

Let, for $j = 1, 2$,

$$s_j = \begin{pmatrix} s_{j,0} & s_{j,1} \\ s_{j,1}^* & s_{j,2} \end{pmatrix},$$

where $\{s_{j,k} : 1 \le j \le 2,\ 0 \le k \le 2\}$ are freely noncommuting variables with $s_{j,0}$ and $s_{j,2}$ symmetric; that is $s_{j,k}^* = s_{j,k}$ for $k = 0, 2$. For notational purposes, let

$$s_0 = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Suppose $p = \sum_{j,k=0} p_{j,k}\, x_j x_k$, is a 2×2 symmetric matrix polynomial of degree (at most) two in two symmetric variables $x = (x_1, x_2)$, where, for notation purposes, $x_0 = 1$ (the unit in $\mathbb{C}\langle x \rangle$), each $p_{j,k} \in M_2(\mathbb{C})$ and $p_{j,k}^* = p_{k,j}$. Let $\mathscr{E}p$ denote the matrix polynomial in the six variables $\{s_{j,0}, s_{j,1}, s_{j,2} : 1 \le j \le 2\}$ defined by

$$\mathscr{E}p(s) = \sum_{j,k=0}^{2} p_{j,k} \circledast s_j s_k.$$

Such a polynomial is naturally evaluated at a pair of block 2×2 symmetric matrices,

$$S_j = \begin{pmatrix} S_{j,0} & S_{j,1} \\ S_{j,1}^* & S_{j,2} \end{pmatrix} \in M_{n+m}(\mathbb{C}) \tag{3.1}$$

using $\circledast$ via

$$\mathscr{E}p(S) = \sum_{j,k=0}^{2} p_{j,k} \circledast S_j S_k \in M_{m+n}(\mathbb{C}).$$

By contrast,

$$p(S) = \sum_{j,k=0}^{2} p_{j,k} \otimes S_j S_k \in M_2(\mathbb{C}) \otimes M_{m+n}(\mathbb{C}).$$

However, p and $\mathscr{E}p$ are closely related, as the following lemma describes. Its proof is similar to that of Lemma 3.1.

Lemma 3.2. *With notations as above,*

$$\mathscr{E}p(S) = E^* \left(\sum_{j,k=0}^{2} p_{j,k} \otimes S_j S_k \right) E = E^* p(S) E.$$

In particular, if $p(S) \succeq 0$, then $\mathscr{E}p(S) \succeq 0$ too.

Theorem 3.3 is the main result of this section.

Theorem 3.3. *Suppose $\rho(x)$ is a symmetric 2×2 polynomial of degree at most two in the symmetric variables $x = (x_1, x_2)$. If $\mathscr{E}\rho(S) \succeq 0$ for all positive integers m, n and pairs $S = (S_1, S_2) \in \mathbb{S}_{n+m}^2$ of 2×2 block symmetric matrices, then there exists an $N \le 12$ and $q_0, q_1, q_2 \in M_{N,2}(\mathbb{C})$ such that*

$$\begin{aligned} q_j^* q_k &= \rho_{j,k}, \quad 1 \le j, k \le 2, \\ q_0^* q_k + q_k^* q_0 &= \rho_{k,0} + \rho_{0,k}, \quad k = 1, 2, \end{aligned} \tag{3.2}$$

$$(q_0^* q_0)_{1,1} = (\rho_{0,0})_{1,1}, \quad (q_0^* q_0)_{2,2} = (\rho_{0,0})_{2,2}. \tag{3.3}$$

In particular, letting q denote the affine linear polynomial $q = \sum_{j=0}^{2} q_j x_j \in \mathbb{C}\langle x\rangle^{N\times 2}$, there is an $r_1 \in \mathbb{C}$ such that

$$\rho = q^* q + r, \quad \textit{where } r = \begin{pmatrix} 0 & r_1 \\ r_1^* & 0 \end{pmatrix}.$$

The remainder of this section is devoted to the proof of Theorem 3.3. Let $\{e_1, e_2\}$ denote the standard orthonormal basis for $\mathbb{C}^2$ with resulting matrix units $e_a e_b^*$ for $1 \le a, b \le 2$. Let $\langle x_1, x_2\rangle_k$ denote the words in x_1, x_2 of length at most k. Thus $\langle x_1, x_2\rangle_1 = \{x_0, x_1, x_2\}$, where, as above, $x_0 = 1$. We will view $\mathbb{C}^3$ as the span of $\langle x_1, x_2\rangle_1$ with $\langle x_1, x_2\rangle_1$ as an orthonormal basis and $M_3(\mathbb{C})$ as matrices indexed by $\langle x_1, x_2\rangle_1 \times \langle x_1, x_2\rangle_1$. In this case $x_j x_k^*$ are the matrix units.

Let $\mathscr{S}$ denote the subspace of $M_2(\mathbb{C}) \otimes M_3(\mathbb{C})$ consisting of matrices

$$T = \left(T_{\alpha,\beta}\right)_{\alpha,\beta \in \langle x_1, x_2\rangle_1},$$

where $T_{\alpha,\beta} \in M_2(\mathbb{C})$ satisfy, for $\beta \in \langle x_1, x_2\rangle_1$,

$$T_{\beta,x_0} = T_{x_0,\beta}, \quad T_{x_0,x_0} \in \operatorname{span}\{e_1 e_1^*, e_2 e_2^*\}.$$

Thus T_{x_0,x_0} is diagonal and $\mathscr{S}$ is an **operator space**; that is, a self-adjoint subspace of $M_2(\mathbb{C}) \otimes M_3(\mathbb{C})$ that contains the identity.

Define $\psi : \mathscr{S} \to M_2(\mathbb{C})$ by

$$\psi\left(T_{\alpha,\beta}\right) = \sum_{\alpha,\beta \in \langle x_1,x_2\rangle_1} \rho_{\alpha,\beta} * T_{\alpha,\beta} = \sum_{\alpha,\beta \in \langle x_1,x_2\rangle_1} \rho_{\alpha,\beta} \circledast T_{\alpha,\beta}. \tag{3.4}$$

Proposition 3.4. *The mapping ψ of equation (3.4) is completely positive (cp).*

Proof. To prove that ψ is cp, let a positive integer n and positive definite $Z \in M_n(\mathbb{C}) \otimes \mathscr{S}$ be given. In particular,

$$Z = \left(Z_{\alpha,\beta}\right)_{\alpha,\beta \in \langle x_1,x_2\rangle_1},$$

where $Z_{\alpha,\beta} = \left((Z_{\alpha,\beta})_{a,b}\right)_{a,b=1}^{2} \in M_n(\mathbb{C}) \otimes M_2(\mathbb{C})$, $(Z_{\alpha,\beta})_{a,b} \in M_n(\mathbb{C})$ and

$$Z_{x_0,\beta} = Z_{\beta,x_0}, \quad Z_{x_0,x_0} = \sum_{a=1}^{2} (Z_{x_0,x_0})_{a,a} \otimes e_a e_a^*.$$

Since Z is positive definite, $Z_{x_0,\alpha}^* = Z_{x_0,\alpha}$ and letting $\Theta = Z_{x_0,x_0}^{-1}$,

$$0 \preceq \left(Z_{\alpha,\beta} - Z_{\alpha,x_0} \Theta Z_{x_0,\beta}\right)_{|\alpha|=|\beta|=1} = GG^* = \left(G_\alpha G_\beta^*\right)_{|\alpha|=|\beta|=1},$$

for some m and matrices

$$G_\alpha = \left((G_\alpha)_{a,j}\right)_{a,j=1}^{2} \in M_{n,m}(\mathbb{C}) \otimes M_2(\mathbb{C}).$$

In particular, for $1 \le a, b \le 2$,

$$(Z_{\alpha,\beta})_{a,b} - \left[Z_{\alpha,x_0} \begin{pmatrix} \Theta_{1,1} & 0 \\ 0 & \Theta_{2,2} \end{pmatrix} Z_{x_0,\beta} \right]_{a,b} = \sum_{j=1}^{2} (G_\alpha)_{a,j} \, (G_\beta)_{b,j}^*,$$

where $\Theta_{j,j} = (Z_{x_0,x_0})^{-1}_{j,j}$. Thus, for $|\alpha| = 1 = |\beta|$,

$$\sum_{j=1}^{2} (Z_{\alpha,x_0})_{a,j}\Theta_{j,j}(Z_{x_0,\beta})_{j,b} + \sum_{j=1}^{2} (G_\alpha)_{a,j}\,(G_\beta)^*_{b,j} = (Z_{\alpha,\beta})_{a,b}.$$

Let

$$\Psi = \begin{pmatrix} \Psi_{1,1} & 0 \\ 0 & \Psi_{2,2} \end{pmatrix} \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C}), \quad \text{where} \quad \Psi_{a,a} = \begin{pmatrix} (Z_{x_0,x_0})_{a,a} & 0 \\ 0 & I_m \end{pmatrix} \in M_{n+m}(\mathbb{C}).$$

Let, for $j = 1, 2$,

$$W_j = \big((W_j)_{a,b}\big) \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C}), \quad \text{where} \quad (W_j)_{a,b} = \begin{pmatrix} (Z_{x_0,x_j})_{a,b} & (G_{x_j})_{a,b} \\ (G_{x_j})^*_{a,b} & 0 \end{pmatrix} \in M_{n+m}(\mathbb{C}).$$

Since Z_{α,x_0} is self-adjoint, so is W_j. By construction,

$$(W_j\Psi^{-1}W_k)_{a,b} = \begin{pmatrix} (Z_{x_j,x_k})_{a,b} & * \\ * & * \end{pmatrix} \in M_{n+m}(\mathbb{C}).$$

Thus, letting $V \in M_{2(n+m),2n}(\mathbb{C})$ denote the isometry whose adjoint is

$$V^* = \begin{pmatrix} I_n & 0 & 0 & 0 \\ 0 & 0 & I_n & 0 \end{pmatrix} \in M_{2n,2(n+m)}(\mathbb{C}),$$

we have

$$V^*W_j\Psi^{-1}W_kV = Z_{x_j,x_k}, \quad V^*W_jV = Z_{x_0,x_j}, \quad V^*\Psi V = Z_{x_0,x_0}. \tag{3.5}$$

Let

$$W = \begin{pmatrix} \Psi & W_1 & W_2 \\ W_1 & W_1\Psi^{-1}W_1 & W_1\Psi^{-1}W_2 \\ W_2 & W_2\Psi^{-1}W_1 & W_2\Psi^{-1}W_2 \end{pmatrix} \in M_{n+m}(\mathbb{C}) \otimes \mathscr{S}.$$

It follows from (3.5) that $\psi(Z) = V^*\psi(W)V$. Thus to prove $\psi(Z) \succeq 0$ it suffices to show $\psi(W) \succeq 0$.

Viewing $S_j = \Psi^{-\frac{1}{2}}W_j\Psi^{-\frac{1}{2}}$ as block 2×2 matrices with respect to the natural block decomposition, the S_j are self-adjoint and

$$\Psi^{-\frac{1}{2}}\psi(W)\Psi^{-\frac{1}{2}} = \sum_{j,k} \rho_{j,k} \circledast S_jS_k = \mathcal{E}\rho(S).$$

By hypothesis $\mathcal{E}\rho(S) \succeq 0$ and hence $\psi(W) \succeq 0$. A limiting argument now shows if $Z \in M_n(\mathbb{C}) \otimes \mathscr{S}$ is positive semidefinite, then $\psi(Z) \succeq 0$ and hence ψ is completely positive. ∎

Proof of Theorem 3.3. Since, by Proposition 3.4, ψ is cp it extends, by the Arveson Extension Theorem [Pau02, Theorem 7.5], to a cp map $\varphi : M_2(\mathbb{C}) \otimes M_3(\mathbb{C}) \to M_2(\mathbb{C})$. By a well-known result of Choi [Pau02, Theorem 3.14], its Choi matrix

$$C_\varphi = \sum_{j,k=0}^{2} \sum_{a,b=1}^{2} [e_ae_b^* \otimes x_jx_k^*] \otimes [\varphi(e_ae_b^* \otimes x_jx_k^*)] \in M_2(\mathbb{C}) \otimes M_3(\mathbb{C}) \otimes M_2(\mathbb{C})$$

is positive semidefinite. In particular, C_φ factors as F^*F where,

$$F = \sum_{a=1}^{2}\sum_{j=1}^{3} e_a^* \otimes x_j^* \otimes F_{j,a}$$

for some N (≤ 12) and $N \times 2$ matrices $F_{j,a}$ and

$$F_{j,a}^* F_{k,b} = \varphi(e_a e_b^* \otimes x_j x_k^*). \tag{3.6}$$

For $q_j = \begin{pmatrix} F_{j,1}e_1 & F_{j,2}e_2 \end{pmatrix} \in M_{N,2}(\mathbb{C})$, we have $q_j^* q_k = \left(e_a^* F_{j,a}^* F_{k,b} e_b\right)_{a,b=1}^2 \in M_2(\mathbb{C})$. So, using (3.6), for $a = 1, 2$,

$$(\rho_{0,0})_{a,a} = (\rho_{0,0} \circledast e_a e_a^*)_{a,a} = \psi(e_a e_a^* \otimes x_0 x_0^*)_{a,a} = \varphi(e_a e_a^* \otimes x_0 x_0^*)_{a,a} = e_a^* F_{0,a}^* F_{0,a} e_a = (q_0^* q_0)_{a,a}.$$

Hence equation (3.3) holds. Next, for $\ell = 1, 2$ and $1 \leq a, b \leq 2$,

$$\begin{aligned}(\rho_{0,\ell} + \rho_{\ell,0})_{a,b} &= e_a^* \left[(\rho_{0,\ell} + \rho_{\ell,0}) \circledast e_a e_b^*\right] e_b = e_a^* \psi\left(e_a e_b^* \otimes (x_0 x_\ell^* + x_\ell x_0^*)\right) e_b \\ &= e_a^* \varphi\left(e_a e_b^* \otimes (x_0 x_\ell^* + x_\ell x_0^*)\right) e_b = e_a^* [F_{0,a}^* F_{\ell,b} + F_{\ell,a}^* F_{0,b}] e_b \\ &= (q_0^* q_\ell + q_\ell^* q_0)_{a,b}.\end{aligned}$$

Thus $q_0^* q_\ell + q_\ell^* q_0 = \rho_{0,\ell} + \rho_{\ell,0}$.

Finally, we see that $q_j^* q_k = \rho_{j,k}$ (for $1 \leq j, k \leq 2$) by computing, for $1 \leq a, b \leq 2$,

$$\begin{aligned}(\rho_{j,k})_{a,b} &= e_a^* [\rho_{j,k} \circledast e_a e_b^*] e_b = e_a^* \psi(e_a e_b^* \otimes x_j x_k^*) e_b \\ &= e_a^* \varphi(e_a e_b^* \otimes x_j x_k^*) e_b = e_a^* F_{j,a}^* F_{k,b} e_b = (q_j^* q_k)_{a,b}.\end{aligned}$$

■

4. The characterization of xy-convex polynomials

In this section we prove Theorem 1.4. In Subsection 4.1 it is established that xy-convex polynomials are biconvex (convex in x and y separately). Two applications of equation (2.8) of Theorem 2.10 then significantly reduce the complexity of the problem of characterizing xy-convex polynomials. The notion of the xy-Hessian of a polynomial is introduced in Subsection 4.2 where a **border vector-middle matrix** (see for instance [HKM13a]) representation for this Hessian is established. Further, it is shown that this middle matrix is positive for xy-convex polynomials. The proof of Theorem 1.4 concludes in Subsection 4.3 by combining positivity of the middle matrix and Theorem 3.3.

4.1. xy-convexity implies biconvexity.

The notion of xy-convexity for polynomials has a convenient concrete reformulation.

Proposition 4.1. *A triple $((X, Y), V)$ is an xy-pair if and only if, up to unitary equivalence, it has the block form*

$$X = \begin{pmatrix} X_0 & A & 0 \\ A^* & * & * \\ 0 & * & * \end{pmatrix}, \quad Y = \begin{pmatrix} Y_0 & 0 & C \\ 0 & * & * \\ C^* & * & * \end{pmatrix}, \quad V = \begin{pmatrix} I & 0 & 0 \end{pmatrix}^*. \tag{4.1}$$

Thus, a polynomial $p(x, y) \in M_\mu(\mathbb{C}\langle x, y\rangle)$ is xy-convex if and only if for each xy-pair $((X, Y), V)$ of the form of equation (4.1), we have

$$(I_\mu \otimes V)^* p(X, Y)(I_\mu \otimes V) - p(X_0, Y_0) \succeq 0.$$

Proof. Observe that $(X_0, Y_0) = V^*(X,Y)V$ and $((X,Y),V)$ is an xy-pair; that is $V^*YXV = V^*YVV^*XV$. Thus, if p is xy-convex on $\mathcal{K}$, then

$$0 \preceq (I_\mu \otimes V)^* p(X,Y)(I_\mu \otimes V) - p(V^*(X,Y)V) = (I_\mu \otimes V)^* p(X,Y)(I_\mu \otimes V) - p(X_0,Y_0).$$

To establish the reverse implication, given an xy-pair $((X,Y),V)$ decompose the space (X,Y) act upon as $\operatorname{rng} V \oplus (\operatorname{rng} V)^\perp$ and note that, with respect to this orthogonal decomposition, X and Y have the block form

$$X = \begin{pmatrix} X_0 & \alpha \\ \alpha^* & \beta \end{pmatrix}, \quad Y = \begin{pmatrix} Y_0 & \gamma \\ \gamma^* & \delta \end{pmatrix},$$

where X_0, Y_0, β, δ are hermitian. The relation $V^*YXV = V^*YVV^*XV$ implies $\alpha\gamma^* = 0$. But then, α and γ are, up to unitary equivalence, of the form $\begin{pmatrix} A & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & C \end{pmatrix}$, respectively. ∎

Consider the following list of monomials:

$$\mathscr{L} = \{1, x, y, x^2, y^2, xy, yx, xy^2, y^2x, x^2y, yx^2, xyx, yxy, xyxy, yxyx, xy^2x, yx^2y\}. \tag{4.2}$$

Proposition 4.2. *If $p \in \mathbb{C}\langle x,y\rangle$ is convex in both x and y (separately), then p has degree at most two in both x and y (separately) and p contains no monomials of the form x^2y^2 or y^2x^2, only the monomials in the set $\mathscr{L}$.*

Proof. The degree bounds follow from Theorem 2.10. The representation of p in (2.8) and that of ℓ in (2.6) imply p does not contain the monomials x^2y^2 and y^2x^2. ∎

Let $[\mathscr{L}]$ denote the $\mathbb{C}$-vector space with basis $\mathscr{L}$ of equation (4.2).

Lemma 4.3. *If $p \in \mathbb{C}\langle x,y\rangle$ is xy-convex, then p is convex in both x and y. Hence $p \in [\mathscr{L}]$.*

Proof. Given (X_1, Y) and (X_2, Y), let $V = \frac{1}{\sqrt{2}}\begin{pmatrix} I & I \end{pmatrix}^T$ and note $((X_1 \oplus X_2, Y \oplus Y), V)$ is an xy-pair. Since p is xy-convex,

$$p\Big(\frac{X_1+X_2}{2}, Y\Big) = p(V^*(X,Y)V) \preceq V^*p(X,Y)V = \frac{1}{2}\big(p(X_1,Y) + p(X_2,Y)\big)$$

Thus p is convex in x. By symmetry p is convex in y. The conclusion of the lemma now follows from Proposition 4.2. ∎

4.2. **The xy-Hessian.** In view of Lemma 4.3 we now consider only symmetric polynomials $p \in [\mathscr{L}]$. Let $\{s_0, t_0, \alpha, \beta_j, \gamma, \delta_j : 0 \le j \le 2\}$ denote freely noncommuting variables with $s_0, t_0, \beta_0, \beta_2, \delta_0, \delta_2$ symmetric. Let, in view of Proposition 4.1,

$$s = \begin{pmatrix} s_0 & \begin{pmatrix} \alpha & 0 \end{pmatrix} \\ \begin{pmatrix} \alpha^* \\ 0 \end{pmatrix} & \begin{pmatrix} \beta_0 & \beta_1 \\ \beta_1^* & \beta_2 \end{pmatrix} \end{pmatrix}, \quad t = \begin{pmatrix} t_0 & \begin{pmatrix} 0 & \gamma \end{pmatrix} \\ \begin{pmatrix} 0 \\ \gamma^* \end{pmatrix} & \begin{pmatrix} \delta_0 & \delta_1 \\ \delta_1^* & \delta_2 \end{pmatrix} \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^*.$$

The xy**-Hessian** of $p \in \mathbb{C}\langle x,y\rangle$, denoted $H^{xy}p$, is the quadratic in α, γ part of $V^*p(s,t)V - p(V^*(s,t)V) = V^*p(s,t)V - p(s_0,t_0)$. In particular, for $p \in [\mathscr{L}]$,

$$H^{xy}p := V^*p(s,t)V - p(V^*(s,t)V) = V^*p(s,t)V - p(s_0,t_0).$$

The proof of the following lemma is routine.

Lemma 4.4. *If $p = \sum_{u\in\mathscr{L}} p_u u \in [\mathscr{L}]$, then $H^{xy}p$ is a function of $\{\alpha, \gamma, s_0, t_0, \delta_0, \delta_1, \beta_1, \beta_2\}$ with the explicit form*

$$\begin{aligned}
H^{xy}p = {} & [p_{x^2}\alpha\alpha^* + p_{y^2}\gamma\gamma^*] + [p_{xyx}\alpha\delta_0\alpha^* + p_{yxy}\gamma\beta_2\gamma^* + p_{xy^2}(s_0\gamma\gamma^* + \alpha\delta_1\gamma^*) \\
& + p_{y^2x}(\gamma\gamma^* s_0 + \gamma\delta_1^*\alpha^*) + p_{x^2y}(\alpha\alpha^* t_0 + \alpha\beta_1\gamma^*) + p_{yx^2}(t_0\alpha\alpha^* + \gamma\beta_1^*\alpha^*)] \\
& + [p_{xy^2x}(s_0\gamma\gamma^* s_0 + \alpha\delta_1\gamma^* s_0 + s_0\gamma\delta_1^*\alpha^* + \alpha(\delta_0^2 + \delta_1\delta_1^*)\alpha^*) \\
& + p_{xyxy}(\alpha\delta_0\alpha^* t_0 + \alpha\delta_0\beta_1\gamma^* + s_0\gamma\beta_2\gamma^* + \alpha\delta_1\beta_2\gamma^*) \\
& + p_{yxyx}(t_0\alpha\delta_0\alpha^* + \gamma\beta_1^*\delta_0\alpha^* + \gamma\beta_2\gamma^* s_0 + \gamma\beta_2\delta_1^*\alpha^*) \\
& + p_{yx^2y}(t_0\alpha\alpha^* t_0 + \gamma\beta_1^*\alpha^* t_0 + t_0\alpha\beta_1\gamma^* + \gamma(\beta_1^*\beta_1 + \beta_2^2)\gamma^*)] \\
= {} & \alpha\,[p_{x^2} + p_{xy}\delta_0 + p_{xy^2x}(\delta_0 + \delta_1\delta_1^*)]\,\alpha^* + \alpha\,[p_{xy^2} + p_{xyxy}\delta_0]\,\alpha^* t_0 + t_0\alpha\,[p_{yx^2} + p_{yxyx}\delta_0]\,\alpha^* \\
& + \alpha\,[p_{xy^2}\delta_1 + p_{x^2y}\beta_1 + p_{xyxy}(\delta_0\beta_1 + \delta_1\beta_2)]\,\gamma^* \\
& + \gamma\,[p_{y^2x}\delta_1^* + p_{yx^2}\beta_1^* + p_{yxyx}(\beta_1^*\delta_0 + \beta_2^*\delta_1)]\,\alpha^* \\
& + \alpha\,[p_{xy^2x}\delta_1]\,\gamma^* s_0 + s_0\gamma\,[p_{xy^2x}\delta_1^*]\,\alpha^* + t_0\alpha\,[p_{yx^2y}]\,\alpha^* t_0 + t_0\alpha\,[p_{yx^2y}\beta_1]\,\gamma^* \\
& + \gamma\,[p_{yx^2y}\beta_1^*]\,\alpha^* t_0 + \gamma\,\left[p_{y^2} + p_{yxy}\beta_2 + p_{yx^2y}(\beta_1^*\beta_1 + \beta_2^2)\right]\gamma^* + \gamma\,[p_{y^2x} + p_{yxyx}\beta_2]\,\gamma^* s_0 \\
& + s_0\gamma\,[p_{xy^2} + p_{xyxy}\beta_2]\,\gamma^* + s_0\gamma\,[p_{yx^2y}]\,\gamma^* s_0.
\end{aligned}$$

Lemma 4.5. *If $p \in [\mathscr{L}]$ and $H^{xy}p = 0$, then p is an xy-pencil. If $p, q \in [\mathscr{L}]$ satisfy $H^{xy}p = H^{xy}q$, then there is an xy-pencil $\lambda \in \mathbb{C}\langle x, y\rangle$ such that $p = q + \lambda$.*

Proof. Since H^{xy} is a linear mapping, it suffices to show, if $p = \sum_{w\in\mathscr{L}} p_w w$ satisfies $H^{xy}p = 0$, then p is an xy-pencil. To this end, observe, if $H^{xy}p = 0$, then, in view of Lemma 4.4, $p_w = 0$ for w in the set

$$\{x^2, y^2, xyx, yxy, xy^2, y^2x, x^2y, yx^2, xy^2x, xyxy, yxyx, yx^2y\}.$$

Hence the only possible nonzero coefficients of p are $p_1, p_x, p_y, p_{xy}, p_{yx}$ and the result follows. ■

The Hessian of an xy-convex p has a border vector-middle matrix representation that we now describe. Since $p \in [\mathscr{L}]$,

$$p(x, y) = \lambda(x, y) + \sum_{w\in\mathscr{L}_*} p_w w,$$

where $\lambda(x, y)$ is an xy-pencil and

$$\mathscr{L}_* = \{x^2, y^2, xyx, yxy, xy^2, y^2x, x^2y, yx^2, xy^2x, xyxy, yxyx, yx^2y\} = \mathscr{L} \setminus \{1, x, y, xy, yx\}.$$

Since p is symmetric, there are relations among its coefficients. For instance, $p_{xyx}, p_{yxy} \in \mathbb{R}$ and $p_{yx^2} = \overline{p_{x^2y}}$.

Let $\mathfrak{B} = \mathfrak{B}(s_0, t_0, \alpha, \gamma)$ denote the row vector-valued free polynomial,

$$\mathfrak{B}(s_0, t_0)[\alpha, \gamma] = \begin{pmatrix} \alpha & t_0\alpha & \gamma & s_0\gamma \end{pmatrix}.$$

We call $\mathfrak{B}$ the xy-**border vector**, or simply the border vector.

For $1 \le j,k \le 2$, let $\mathfrak{M}_{j,k}$ denote the 2×2 matrix polynomial,

$$\mathfrak{M}_{11} = \begin{pmatrix} p_{x^2} + p_{xyx}\delta_0 + p_{xy^2x}(\delta_0^2 + \delta_1\delta_1^*) & p_{x^2y} + p_{xyxy}\delta_0 \\ p_{yx^2} + p_{yxyx}\delta_0 & p_{yx^2y} \end{pmatrix},$$
$$\mathfrak{M}_{12} = \begin{pmatrix} p_{x^2y}\beta_1 + p_{xy^2}\delta_1 + p_{xyxy}(\delta_0\beta_1 + \delta_1\beta_2) & p_{xy^2x}\delta_1 \\ p_{yx^2y}\beta_1 & 0 \end{pmatrix},$$
$$\mathfrak{M}_{21} = \begin{pmatrix} p_{yx^2}\beta_1^* + p_{y^2x}\delta_1^* + p_{yxyx}(\beta_1^*\delta_0 + \beta_2\delta_1^*) & p_{yx^2y}\beta_1^* \\ p_{xy^2x}\delta_1^* & 0 \end{pmatrix},$$
$$\mathfrak{M}_{22} = \begin{pmatrix} p_{y^2} + p_{yxy}\beta_2 + p_{yx^2y}(\beta_2^2 + \beta_1^*\beta_1) & p_{y^2x} + p_{yxyx}\beta_2 \\ p_{xy^2} + p_{xyxy}\beta_2 & p_{xy^2x} \end{pmatrix}.$$

Let $\mathfrak{M} = (\mathfrak{M}_{j,k})_{j,k=1}^2$ denote the resulting 4×4 (2×2 block matrix with 2×2 entries) matrix polynomial. The matrix $\mathfrak{M}$ is the xy**-middle matrix**, or simply the middle matrix, of p.

Lemma 4.6. *If $p \in [\mathscr{L}]$ is symmetric, then then*

$$H^{xy}p = \mathfrak{B}\mathfrak{M}\mathfrak{B}^*.$$

Proposition 4.7 shows xy-convexity of p is equivalent to positivity of its middle matrix.

Proposition 4.7. *If $p(x,y)$ is xy-convex, then $\mathfrak{M}(B_1,B_2,D_0,D_1) \succeq 0$ for all matrices (B_1,B_2,D_0,D_1) of compatible sizes.*

Proof. Since p is xy-convex, $H^{xy}p \succeq 0$. Let positive integers M,N and matrices $D_0 \in M_M(\mathbb{C})$, $B_2 \in M_N(\mathbb{C})$ and $B_1, D_1 \in M_{N,M}(\mathbb{C})$ be given. Choose a vector $h \in \mathbb{C}^2$ and $X_0, Y_0 \in M_2(\mathbb{C})$ such that $\{h, X_0h\}$ and $\{h, Y_0h\}$ are linearly independent. Positivity of the Hessian gives

$$\begin{aligned} 0 &\le h^* H^{xy}p(X_0,A,B_1,B_2,Y_0,C,D_0,D_1)h \\ &= [h^*\mathfrak{B}(X_0,A,Y_0,C)]\,\mathfrak{M}(B_1,B_2,D_0,D_1)\,[h^*\mathfrak{B}(X_0,A,Y_0,C)]^*. \end{aligned}$$

On the other hand, given vectors $f_1,\dots,f_4 \in \mathbb{C}^M$, there exists $A \in M_{2,M}(\mathbb{C})$ and $C \in M_{2,N}(\mathbb{C})$ such that

$$\mathfrak{B}(X_0,A,Y_0,C)^*h = \begin{pmatrix} A^*h \\ A^*Y_0h \\ C^*h \\ C^*X_0h \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix}.$$

It follows that $\mathfrak{M}(B_1,B_2,D_0,D_1) \succeq 0$. ∎

4.3. **Proof of Theorem 1.4.** The convexity assumption on p implies the middle matrix $\mathfrak{M}$ of its Hessian takes positive semidefinite values by Proposition 4.7.

Let

$$\sigma = \left(\begin{pmatrix} \delta_0 & \delta_1 \\ \delta_1^* & \delta_2 \end{pmatrix}, \begin{pmatrix} \beta_0 & \beta_1 \\ \beta_1^* & \beta_2 \end{pmatrix} \right).$$

Let Q denote the 2×2 matrix polynomial obtained from the first and third rows and columns of $\mathfrak{M}$. Thus,

$$Q = \begin{pmatrix} p_{x^2} + p_{xyx}\delta_0 + p_{xy^2x}(\delta_0^2 + \delta_1\delta_1^*) & p_{x^2y}\beta_1 + p_{xy^2}\delta_1 + p_{xyxy}(\delta_0\beta_1 + \delta_1\beta_2) \\ p_{yx^2}\beta_1^* + p_{y^2x}\delta_1^* + p_{yxyx}(\beta_1^*\delta_0 + \beta_2\delta_1^*) & p_{y^2} + p_{yxy}\beta_2 + p_{yx^2y}(\beta_2^2 + \beta_1^*\beta_1) \end{pmatrix}. \tag{4.3}$$

Define a 2×2 polynomial $P(x_1,x_2)=\sum P_{j,k}x_jx_k$ (with $x_0=1$ as usual) by setting

$$\begin{aligned}
&P_{0,0}=\begin{pmatrix} p_{x^2} & 0\\ 0 & p_{y^2}\end{pmatrix},\quad P_{0,1}=P_{1,0}=\frac12\begin{pmatrix} p_{xyx} & p_{xy^2}\\ p_{y^2x} & 0\end{pmatrix},\quad P_{0,2}=P_{2,0}=\frac12\begin{pmatrix} 0 & p_{x^2y}\\ p_{yx^2} & p_{yxy}\end{pmatrix}\\
&P_{1,2}=\begin{pmatrix} 0 & p_{xyxy}\\ 0 & 0\end{pmatrix},\quad P_{2,1}=\begin{pmatrix} 0 & 0\\ p_{yxyx} & 0\end{pmatrix},\quad P_{1,1}=\begin{pmatrix} p_{xy^2x} & 0\\ 0 & 0\end{pmatrix},\quad P_{2,2}=\begin{pmatrix} 0 & 0\\ 0 & p_{yx^2y}\end{pmatrix}
\end{aligned}\tag{4.4}$$

and observe $\mathscr{E}P(\sigma)=Q$.

Since $\mathfrak{M}$ takes positive semidefinite values, $\mathscr{E}P(S)\succeq 0$ for all tuples of hermitian matrices of the form (3.1). Hence Theorem 3.3 produces an N and $F=\sum F_js_j$, where $F_j\in M_{N,2}(\mathbb{C})$, and an $R=\begin{pmatrix} 0 & r\\ r^* & 0\end{pmatrix}$ such that $F^*F+R=P$, where $r\in\mathbb{C}$. In particular,

$$\begin{gathered}
F_j^*F_k=P_{j,k},\ 1\le j,k\le 2\\
F_0^*F_k+F_k^*F_0=P_{k,0}+P_{0,k},\ k=1,2\\
F_0^*F_0=P_{0,0}+R,\\
F_1^*F_1=P_{1,1}=\begin{pmatrix} p_{xy^2x} & 0\\ 0 & 0\end{pmatrix},\quad F_2^*F_2=P_{2,2}=\begin{pmatrix} 0 & 0\\ 0 & p_{yx^2y}\end{pmatrix}.
\end{gathered}$$

Hence, letting $\{e_1,e_2\}$ denote the standard orthonormal basis for $\mathbb{C}^2$, $F_1e_2=0=F_2e_1$. In particular, $e_1^*F_2^*F_0=0$. Now set $\Lambda_x=F_0e_1$, $\Lambda_y=F_0e_2$, $\Lambda_{yx}=F_1e_1$ and $\Lambda_{xy}=F_2e_2$ and verify,

$$\begin{aligned}
\Lambda_x^*\Lambda_x&=e_1^*F_0^*F_0e_1=e_1^*P_{0,0}e_1=p_{x^2}\\
\Lambda_y^*\Lambda_y&=e_2^*F_0^*F_0e_2=e_2^*P_{0,0}e_2=p_{y^2}\\
\Lambda_{yx}^*\Lambda_x+\Lambda_x^*\Lambda_{yx}&=e_1^*F_1^*F_0e_1+e_1^*F_0^*F_1e_1=e_1^*(F_1^*F_0+F_0^*F_1)e_1=(2P_{1,0})_{1,1}=p_{xyx}\\
\Lambda_{xy}^*\Lambda_y+\Lambda_y^*\Lambda_{xy}&=e_2^*F_2^*F_0e_2+e_2^*F_0^*F_2e_2=e_2^*(F_2^*F_0+F_0^*F_2)e_2=e_2^*(2P_{2,0})e_2=p_{yxy}\\
\Lambda_x^*\Lambda_{xy}&=e_1^*F_0^*F_2e_2=e_1^*(F_0^*F_2+F_2^*F_0)e_2=e_1^*(2P_{2,0})e_2=p_{x^2y}\\
\Lambda_y^*\Lambda_{yx}&=e_2^*F_0^*F_1e_1=e_2^*(F_0^*F_1+F_1^*F_0)e_1=e_2^*(2P_{1,0})e_1=p_{y^2x}\\
\Lambda_{xy}^*\Lambda_x&=e_2^*F_2^*F_0e_1=e_2^*(F_2^*F_0+F_0^*F_2)e_1=e_2^*(2P_{2,0})e_1=p_{yx^2}\\
\Lambda_{yx}^*\Lambda_y&=e_1^*F_1^*F_0e_2=e_1^*(F_1^*F_0+F_0F_1^*)e_2=e_1^*(2P_{1,0})e_2=p_{xy^2}\\
\Lambda_{yx}^*\Lambda_{yx}&=e_1^*F_1^*F_1e_1=e_1^*P_{1,1}e_1=p_{xy^2x}\\
\Lambda_{xy}^*\Lambda_{xy}&=e_2^*F_2^*F_2e_2=e_2^*P_{2,2}e_2=p_{yx^2y}\\
\Lambda_{xy}^*\Lambda_{yx}&=e_2^*F_2^*F_1e_1=e_2^*P_{2,1}e_1=p_{yxyx}\\
\Lambda_{yx}^*\Lambda_{xy}&=e_1^*F_1^*F_2e_2=e_1^*P_{1,2}e_2=p_{xyxy}.
\end{aligned}\tag{4.5}$$

Let

$$q=\Lambda(x,y,xy)^*\Lambda(x,y,xy),$$

where Λ denotes the xy-pencil

$$\Lambda=\Lambda_xx+\Lambda_yy+\Lambda_{xy}xy+\Lambda_{yx}yx.$$

A straightforward calculation, based on the identities of equation (4.5) and an appeal to the formula for the xy-Hessian in Lemma 4.4, shows $H^{xy}q = H^{xy}p$. Hence, by Lemma 4.5, there is a hermitian xy-pencil λ such that $p = q + \lambda = \Lambda^*\Lambda + \lambda$, completing the proof. ■

Remark 4.8. Note that $\Lambda_x^*\Lambda_y + \Lambda_y^*\Lambda_x = R = \begin{pmatrix} 0 & r \\ r^* & 0 \end{pmatrix}$.

References

[And89] T. Ando: *Majorization, doubly stochastic matrices, and comparison of eigenvalues*, Linear Algebra Appl. 118 (1989) 163–248. 1

[And94] T. Ando: *Majorizations and inequalities in matrix theory*, Linear Algebra Appl. 199 (1994) 17–67. 1

[vAB00] J.G. VanAntwerp, R.D. Braatz: *A tutorial on linear and bilinear matrix inequalities*, J. Process Control 10 (2000) 363–385. 5

[BM14] S. Balasubramanian, S. McCullough: *Quasi-convex free polynomials*, Proc. Amer. Math. Soc. 142 (2014) 2581–2591. 1

[BMG05] J.A. Ball, G. Groenewald, T. Malakorn: *Structured noncommutative multidimensional linear systems*, SIAM J. Control Optim. 44 (2005) 1474–1528. 6

[BMV16] J.A. Ball, G. Marx, V. Vinnikov: *Noncommutative reproducing kernel Hilbert spaces.* J. Funct. Anal. 271 (2016) 1844–1920. 3

[Bha97] R. Bhatia: *Matrix analysis*, Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997. 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. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. 1

[BK13] M. Brešar, I. Klep: *A local-global principle for linear dependence of noncommutative polynomials*, Israel J. Math. 193 (2013) 71–82. 7

[CHSY03] J.F. Camino, J.W. Helton, R.E. Skelton, J. Ye: *Matrix inequalities: A symbolic procedure to determine convexity automatically*, Integral Equations Operator Theory 46 (2003) 399–454. 7

[DDSS17] 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

[DK+] K. Davidson, M. Kennedy: *Noncommutative Choquet theory*, preprint https://arxiv.org/abs/1905.08436 1

[dHMP09] M.C. de Oliveira, J.W. Helton, S.A. McCullough, M. Putinar: *Engineering systems and free semi-algebraic geometry.* Emerging applications of algebraic geometry, 17–61, IMA Vol. Math. Appl., 149, Springer, New York, 2009. 1

[DHM17] H. Dym, J.W. Helton, S. McCullough: *Non-commutative polynomials with convex level slices*, Indiana Univ. Math. J. 66 (2017) 2071–2135. 1

[ENE11] A. Ebadian, I. Nikoufar, M. Eshaghi Gordji: *Perspectives of matrix convex functions*, Proc. Natl. Acad. Sci. USA 108 (2011) 7313–7314. 1

[Eff09] E.G. Effros: *A matrix convexity approach to some celebrated quantum inequalities* Proc. Natl. Acad. Sci. USA 106 (2009) 1006–1008. 1

[Eve18] E. Evert: *Matrix convex sets without absolute extreme points,* Linear Algebra Appl. 537 (2018) 287–301. 3

[EH19] E. Evert, J.W. Helton: *Arveson extreme points span free spectrahedra*, Math. Ann. 375 (2019) 629–653. 3

[GSL96] K.-C. Goh, M.G. Safonov, J.H. Ly: *Robust synthesis via bilinear matrix inequalities*, Internat. J. Robust Nonlinear Control 6 (1996) 1079–1095. 5

[GS09] A. Guionnet, D. Shlyakhtenko: *Free diffusions and matrix models with strictly convex interaction*, Geom. Funct. Anal. 18 (2009) 1875–1916. 1

[Han81] F. Hansen: *Selfadjoint means and operator monotone functions*, Math. Ann. 256 (1981) 29–35. 1

[HHLM08] D.M. Hay, J.W. Helton, A. Lim, S. McCullough: *Non-commutative partial matrix convexity*, Indiana Univ. Math. J. 57 (2008) 2815–2842. 2, 5

[HKM13a] J.W. Helton, I. Klep, S. McCullough: *Free convex algebraic geometry.* In: Semidefinite optimization and convex algebraic geometry, 341–405, MOS-SIAM Ser. Optim., 13, SIAM, Philadelphia, PA, 2013. 18

[HKM13b] J.W. Helton, I. Klep, S. McCullough: *The matricial relaxation of a linear matrix inequality,* Math. Program. Ser. A 138 (2013) 401–445. 3

[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. (JEMS) 19 (2017) 1845–1897. 3

[HM04] J.W. Helton, S.A. McCullough: *Convex noncommutative polynomials have degree two or less*, SIAM J. Matrix Anal. Appl. 25 (2004) 1124–1139. 1, 2

[HMV06] J.W. Helton, S.A. McCullough, V. Vinnikov: *Noncommutative convexity arises from linear matrix inequalities*, J. Funct. Anal. 240 (2006) 105–191. 1, 2, 6, 7

[HL18] M. Hartz, M. Lupini, *The classification problem for operator algebraic varieties and their multiplier algebras.* Trans. Amer. Math. Soc. 370 (2018) 2161–2180. 3

[HM12] J.W. Helton, S. McCullough: *Every convex free basic semi-algebraic set has an LMI representation,* Ann. of Math. (2) 176 (2012) 979–1013. 3

[HM98] J.W. Helton, O. Merino: *Sufficient conditions for optimization of matrix functions,* in: Proceedings of the Conference on Decision and Control, 1998, vol 3, 3361–3365. 6, 7

[JB07] E. Jorswieck, H. Boche: *Majorization and matrix-monotone functions in wireless communications*, Foundations and Trends® in Communications and Information Theory 3.6 (2007) 553–701. 1

[JKMMP] M.T. Jury, I. Klep, M.E. Mancuso, S. McCullough, J.E. Pascoe: *Noncommutative partial convexity via* Γ*-convexity*, to appear in J. Geom. Anal., DOI 10.1007/s12220-020-00387-1. 5, 6

[K-VV09] D.S. 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

[K-VV14] D.S. Kaliuzhnyi-Verbovetskyi, V. Vinnikov: *Foundations of free noncommutative function theory,* Mathematical Surveys and Monographs, 199. American Mathematical Society, Providence, RI, 2014. 3

[KSVS04] S. Kanev, C. Scherer, M. Verhaegen, B. De Schutter: *Robust output-feedback controller design via local BMI optimization*, Automatica J. IFAC 40 (2004) 1115–1127. 2, 5

[KV17] I. Klep, J. Volčič: *Free loci of matrix pencils and domains of noncommutative rational functions.* Comment. Math. Helv. 92 (2017) 105–130. 13

[Kra36] F. Kraus: *Über konvexe Matrixfunktionen*, Math. Z. 41 (1936) 18–42. 2

[LM00] C.-K. Li, R. Mathias: *Extremal characterizations of the Schur complement and resulting inequalities*, SIAM Review 42 (2000) 233–246. 9, 10

[PS19] B. Passer, O.M. Shalit: *Compressions of compact tuples,* Linear Algebra Appl. 564 (2019) 264–283. 3

[PTD+] J.E. Pascoe, R. Tully-Doyle: *The royal road to automatic noncommutative real analyticity, monotonicity, and convexity*, preprint https://arxiv.org/abs/1907.05875 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 Studies in Advanced Mathematics, 78. Cambridge University Press, Cambridge, 2002. 17

[Pop18] G. Popescu: *Invariant subspaces and operator model theory on noncommutative varieties,* Math. Ann. 372 (2018) 611–650. 3

[SSS18] G. Salomon, O.M. Shalit, E. Shamovich: *Algebras of bounded noncommutative analytic functions on subvarieties of the noncommutative unit ball,* Trans. Amer. Math. Soc. 370 (2018) 8639–8690 3

[Scü61] M.P. Schützenberger: *On the definition of a family of automata*, Inform. Control 4 (1961) 245–270. 6

[SIG98] R.E. Skelton, T. Iwasaki, K.M. Grigoriadis: *A unified algebraic approach to linear control design*, The Taylor & Francis Systems and Control Book Series. Taylor & Francis, Ltd., London, 1998. 1

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

[Vol18] J. Volčič: *Matrix coefficient realization theory of noncommutative rational functions*, J. Algebra 499 (2018) 397–437. 7

Appendix A. Not for publication

A.1. **Proof of Proposition 1.5.** First suppose $r \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle$ is a^2-convex on $\mathcal{D} \subseteq \mathbb{S}^{\mathrm{h}} \times \mathbb{S}^{\mathrm{g}}$. To prove r is convex in x on $\mathcal{D}$, suppose $(A, X), (A, Y)$. Consider the matrices

$$B = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}, \quad Z = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}, \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} I \\ I \end{pmatrix}.$$

Note that V reduces B. Equivalently $V^*B^2V = (V^*BV)^2$. Since $V^*(B, Z)V = (A, \frac{X+Y}{2}) \in \mathcal{D}$ (by the convexity hypothesis on $\mathcal{D}$) and of course $(B, Z) \in \mathcal{D}$ too,

$$\frac{1}{2}(r(A, X) + r(A, Y)) = V^*r(B, Z)V \succeq r(V^*(B, Z)V) = r\Big(A, \frac{X+Y}{2}\Big).$$

Hence r is convex in x on $\mathcal{D}$.

Now suppose r is convex in x on $\mathcal{D}$ and $(B, Z) \in \mathcal{D}_n$, and $V : \mathbb{C}^m \to \mathbb{C}^n$ is an isometry such that $V^*B^2V = (V^*BV)^2$. Thus the range of V reduces B and up to unitary equivalence,

$$B = \begin{pmatrix} A & 0 \\ 0 & \alpha \end{pmatrix}, \quad Z = \begin{pmatrix} X & \beta \\ \beta^* & \delta \end{pmatrix}, \quad V = \begin{pmatrix} I \\ 0 \end{pmatrix}.$$

Let

$$U = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

Since U is unitary, $\mathcal{D}$ is a free set and $(B, Z) \in \mathcal{D}$, we have $U^*(B, Z)U \in \mathcal{D}$. Since $\mathcal{D}$ is, by hypothesis, convex in x,

$$\mathcal{D} \ni \frac{(B, Z) + U^*(B, Z)U}{2} = \left(B, \begin{pmatrix} X & 0 \\ 0 & \delta \end{pmatrix}\right).$$

Because r is convex in x,

$$\begin{aligned}\begin{pmatrix} r(V^*(B, Z)V) & 0 \\ 0 & r(\alpha, \beta) \end{pmatrix} = \begin{pmatrix} r(A, X) & 0 \\ 0 & r(\alpha, \beta) \end{pmatrix} &= r\left(B, \begin{pmatrix} X & 0 \\ 0 & \delta \end{pmatrix}\right) \\ &\preceq \frac{1}{2}V^*\left(r(B, Z) + r(U^*(B, Z)U\right)V \\ &= \frac{1}{2}V^*\left(r(B, Z) + U^*r(B, Z)U\right)V = V^*r(B, Z)V.\end{aligned}$$

Thus r is a^2-convex. ■

A.2. **Proof of Lemma 4.4.** We provide the routine verification of the formula for $H^{xy}w$ for words $w \in \mathscr{L}$. The result then follows by linearity of H^{xy}.

It is clear that the xy-pencil terms ($1, x, y, xy$ and yx) vanish under H^{xy}. The $(1, 1)$ entry of t^2 is $t_0^2 + \gamma\gamma^*$. Thus,

$$H^{xy}y^2 = t_0^2 + \gamma\gamma^* - t_0^2 = \gamma\gamma^*.$$

Similarly, the $(1, 1)$ entry of st^2 is $s_0t_0^2 + s_0\gamma\gamma^* + \alpha\delta_1\gamma^*$. Hence,

$$H^{xy}xy^2 = (s_0t_0^2 + s_0\gamma\gamma^* + \alpha\delta_1\gamma^*) - s_0t_0^2 = s_0\gamma\gamma^* + \alpha\delta_1\gamma^*.$$

The $(1, 1)$ entry of sts is $(s_0t_0s_0 + \alpha\delta_0\alpha^*) - s_0t_0s_0$. Hence,

$$H^{xy}xyx = \alpha\delta_0\alpha^*.$$

The $(1,1)$ entry of $stst$ is $(s_0t_0s_0 + \alpha\delta_0\alpha^*)t_0 + (\alpha\delta_0\beta_1 + (s_0\gamma + \alpha\delta_1)\beta_2)\gamma^* - s_0t_0s_0t_0$. Hence

$$H^{xy}xyxy = \alpha\delta_0\alpha^*t_0 + \alpha(\delta_0\beta_1 + \delta_1\beta_2)\gamma^* + s_0\gamma\beta_2\gamma^*.$$

The $(1,1)$ entry of st^2s is $s_0t_0^2s_0 + s_0\gamma\gamma^*s_0 + \alpha\delta_1\gamma^*s_0 + s_0\gamma\delta_1^*\alpha^* + \alpha(\delta_0^2 + \delta_1\delta_1^*)\alpha^*$. Thus,

$$\begin{aligned} H^{xy}xy^2x &= (s_0t_0^2s_0 + s_0\gamma\gamma^*s_0 + \alpha\delta_1\gamma^*s_0 + s_0\gamma\delta_1^*\alpha^* + \alpha(\delta_0^2 + \delta_1\delta_1^*)\alpha^*) - s_0t_0^2s_0 \\ &= s_0\gamma\gamma^*s_0 + \alpha\delta_1\gamma^*s_0 + s_0\gamma\delta_1^*\alpha^* + \alpha(\delta_0^2 + \delta_1\delta_1^*)\alpha^*. \end{aligned}$$

The remainder follow by symmetry in x and y. ■

A.3. **Examples.**

Example A.1. Consider the polynomial

$$\begin{aligned} p(x,y) &= x^2 + y^2 + xy^2x + 2(xyxy + yxyx) + yx^2y \\ &= \begin{pmatrix} 1 & y \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \begin{pmatrix} 1+y^2 & 2y \\ 2y & 1 \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \begin{pmatrix} 1 \\ y \end{pmatrix} + y^2 \\ &= \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} \begin{pmatrix} 1+x^2 & 2x \\ 2x & 1 \end{pmatrix} \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} + x^2. \end{aligned}$$

It is, by its very form, convex in x on the free set $\{(X,Y) : I - 3Y^2 \succeq 0\}$ and it is convex in y on the set $\{(X,Y) : I - 3X^2 \succeq 0\}$. Thus p is biconvex on $\mathscr{D} = \{(X,Y) : I - 3X^2,\ I - 3Y^2 \succeq 0\}$. That the set $\mathscr{E} = \{(X,Y) : I - 3X^2,\ I - 3Y^2 \succ 0\}$ is the largest open free set on which p is biconvex follows from Theorem 2.10. ■

Example A.2. Consider the polynomial p from Example A.1 and recall $\mathscr{D} = \{(X,Y) : I - 3X^2,\ I - 3Y^2 \succeq 0\}$ contains any free set on which p is biconvex. The middle matrix of the xy-Hessian of p is given by

$$\mathfrak{M}(x,y) = \begin{pmatrix} \begin{pmatrix} I + \delta_0^2 + \delta_1\delta_1^* & 2\delta_0 \\ 2\delta_0 & I \end{pmatrix} & \begin{pmatrix} 2(\delta_0\beta_1 + \delta_1\beta_2) & \delta_1 \\ \beta_1 & 0 \end{pmatrix} \\ \begin{pmatrix} 2(\beta_1^*\delta_0 + \beta_2\delta_1^*) & \beta_1^* \\ \delta_1^* & 0 \end{pmatrix} & \begin{pmatrix} I + \beta_2^2 + \beta_1^*\beta_1 & 2\beta_2 \\ 2\beta_2 & I \end{pmatrix} \end{pmatrix}.$$

Evidently $\mathfrak{M} \succeq 0$ in a neighborhood of 0 and thus p is xy-convex in a neighborhood of 0. On the other hand, $\mathfrak{M}$ is not positive semidefinite on all of $\mathscr{D}$ and thus, arguing as in the proof of Theorem 1.4, p is not xy-convex on the interior of $\mathscr{D}$. ■

A.4. **Equivalence of positivity of $\mathfrak{M}$ and $\mathscr{E}P$.** In this subsection we show directly that positivity of $\mathfrak{M}$ is equivalent to positivity of $\mathscr{E}P = Q$, where P and Q are defined in equations (4.4) and (4.3). Of course that positivity of $\mathfrak{M}$ implies positivity of Q is immediate. On the other hand, the proof of Theorem 1.4 only required the, a priori weaker, condition $\mathscr{E}P \succeq 0$ and hence this latter condition implies positivity of the middle matrix $\mathfrak{M}$.

Assume Q takes positive semidefinite values. Given $\delta_1, \beta_1 \in M_{n,m}(\mathbb{C})$, $\delta_0 \in M_n(\mathbb{C})$ and $\beta_2 \in M_m(\mathbb{C})$, make the replacements,

$$\widehat{\delta}_0 = \begin{pmatrix} \delta_0 & 0_{n\times m} \\ 0_{m\times n} & 0_{m\times m} \end{pmatrix}, \quad \widehat{\delta}_1 = \begin{pmatrix} \delta_1 & 0_{n\times n} \\ tI_m & 0_{m\times n} \end{pmatrix}, \quad \widehat{\beta}_1 = \begin{pmatrix} \beta_1 & tI_n \\ 0_{m\times m} & 0_{m\times n} \end{pmatrix}, \quad \widehat{\beta}_2 = \begin{pmatrix} \beta_2 & 0_{m\times n} \\ 0_{n\times m} & 0_{m\times m} \end{pmatrix}.$$

Substituing into Q gives, $Q(\widehat{\beta},\widehat{\delta}) = \big(Q_{j,k}\big)_{j,k=1}^{2}$, where

$$Q_{1,1} = \begin{pmatrix} p_{x^2} + p_{xyx}\delta_0 + p_{xy^2x}(\delta_0^2 + \delta_1\delta_1^*) & tp_{xy^2x}\delta_1 \\ tp_{xy^2x}\delta_1^* & p_{x^2} + t^2 p_{xy^2x} \end{pmatrix}$$

$$Q_{1,2} = \begin{pmatrix} p_{x^2y}\beta_1 + p_{xy^2}\delta_1 + p_{xyxy}(\delta_0\beta_1 + \delta_1\beta_2) & t(p_{x^2y} + p_{xyxy}\delta_0) \\ t(p_{xy^2} + p_{xyxy}\beta_2) & 0 \end{pmatrix} = Q_{2,1}^*$$

$$Q_{2,2} = \begin{pmatrix} p_{y^2} + p_{yxy}\beta_2 + p_{yx^2y}(\beta_2^2 + \beta_1^*\beta_1) & tp_{yx^2y}\beta_1^* \\ tp_{yx^2y}\beta_1 & p_{y^2} + t^2 p_{yx^2y} \end{pmatrix}.$$

Now conjugate each block with $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{t} \end{pmatrix}$ and let t tend to infinity to deduce that $Q' = \big(Q'_{j,k}\big)_{j,k=1}^{2} \succeq 0$, where

$$Q'_{1,1} = \begin{pmatrix} p_{x^2} + p_{xyx}\delta_0 + p_{xy^2x}(\delta_0^2 + \delta_1\delta_1^*) & p_{xy^2x}\delta_1 \\ p_{xy^2x}\delta_1^* & p_{xy^2x} \end{pmatrix}$$

$$Q'_{1,2} = \begin{pmatrix} p_{x^2y}\beta_1 + p_{xy^2}\delta_1 + p_{xyxy}(\delta_0\beta_1 + \delta_1\beta_2) & p_{x^2y} + p_{xyxy}\delta_0 \\ p_{xy^2} + p_{xyxy}\beta_2 & 0 \end{pmatrix} = (Q'_{1,2})^*$$

$$Q'_{2,2} = \begin{pmatrix} p_{y^2} + p_{yxy}\beta_2 + p_{yx^2y}(\beta_2^2 + \beta_1^*\beta_1) & p_{yx^2y}\beta_1^* \\ p_{yx^2y}\beta_1 & p_{yx^2y} \end{pmatrix}.$$

Finally, Q' is unitarily equivalent to $\mathfrak{M}$ via the permutation that interchanges the second and fourth rows and columns.

Michael Jury, Department of Mathematics, University of Florida, Gainesville

Email address: `mjury@ufl.edu`

Igor Klep, Faculty of Mathematics and Physics, University of Ljubljana, Slovenia

Email address: `igor.klep@fmf.uni-lj.si`

Mark E. Mancuso, Department of Mathematics and Statistics, Washington University in St. Louis

Email address: `mark.mancuso@wustl.edu`

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

Email address: `sam@math.ufl.edu`

James E. Pascoe, Department of Mathematics, University of Florida, Gainesville

Email address: `pascoej.ufl.edu`

NOT FOR PUBLICATION

Contents