Rev. Mat. Iberoam. (Online first)
DOI 10.4171/RMI/1304

Noncommutative partially convex rational functions

Michael Jury, Igor Klep, Mark E. Mancuso, Scott McCullough and James Eldred Pascoe

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 $f\colon [a,b] \to \mathbb{R}$ is convex if

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2}\,(f(x)+f(y))$$

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 [8, 45], wireless communication [31], matrix means [1, 2, 21], perspective functions [15, 16], random matrices and free probability [20] and noncommutative function theory [3, 12, 14, 27, 29]. Often in systems engineering [13], problems have two classes of variables: known unknowns $a = (a_1, \dots, a_{\mathrm{h}})$ and unknown unknowns $x = (x_1, \dots, x_{\mathrm{g}})$. 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. The x variables represent the design variables. A key point is that $p(a,x)$ depends only upon the signal flow diagram. Thus a choice of a value A for a corresponds to a specific problem governed by the given signal flow diagram and in that sense a is a known unknown.

Mathematics Subject Classification (2020): Primary 46N10, 26B25; Secondary 47A63, 52A41, 90C25.
Keywords: Partial convexity, biconvexity, bilinear matrix inequality (BMI), noncommutative rational function, noncommutative polynomial, realization theory.

One then chooses the design variable X to optimize an objective and in that sense x is an unknown unknown. Partial convexity in the unknown unknowns x is then sufficient for reliable numerics and optimization.

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

$$f\Big(\frac{X+Y}{2}\Big)\preceq\frac{1}{2}\,(f(X)+f(Y))$$

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 [7, 37]

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

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 $x=(x_1,\dots,x_g)$. For example,

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

is an 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

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

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 an 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 $r(X)$. An 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 [29] for rational functions and in [38] for more general nc functions. There is a paucity of matrix convex polynomials: as first observed in [27], 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 [23]. Moreover, we also investigate the stronger notion of xy-convexity, modeled on the theory of bilinear matrix inequalities (BMIs) [35].

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

$$X \oplus Y := \big(X_1 \oplus Y_1, \dots, X_{\mathtt{k}} \oplus Y_{\mathtt{k}}\big) = \left(\begin{pmatrix} X_1 & 0 \\ 0 & Y_1 \end{pmatrix}, \dots, \begin{pmatrix} X_{\mathtt{k}} & 0 \\ 0 & Y_{\mathtt{k}} \end{pmatrix}\right) \in \mathcal{D}_{n+m},$$
$$U^*XU := (U^*X_1U, \dots, U^*X_{\mathtt{k}}U) \in \mathcal{D}_n.$$

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, namely $a = (a_1, \dots, a_{\mathtt{h}})$ 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 $X \in \mathbb{S}^{\mathtt{g}}$.

The symmetric version (see Proposition 4.3 in [29]) of the well-known Schützenberger [44] state space similarity theorem implies that 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, the $e \times e$ matrix J is a 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 [33, 47] justify defining the *domain* of r as

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

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

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

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 [5, 11, 17, 18, 22, 24, 26, 28, 34, 39, 40, 42, 43].

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

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

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 $A \in \mathbb{S}^{\mathtt{h}}_k$,

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

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, which is proved as Theorem 2.6, 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 vertebral 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*.

The *vertebral domain* of r is the set

$$\operatorname{dom}_{\mathrm{ver}} r = \{(A,X) \in \operatorname{dom} r : (A,0) \in \operatorname{dom} r\}.$$

Let

$$\operatorname{dom}^+_{\mathrm{ver}} r = \{(A,X) \in \operatorname{dom}^+ r : (A,0) \in \operatorname{dom}^+ r\}$$

Theorem 1.2 gives a realization tailored to partial convexity that provides an algebraic certificate of convexity in x for an $r \in \mathbb{C}\langle\!\langle a,x \rangle\!\rangle$. Given a subset $\mathcal{D} \subseteq \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathtt{g}}$, let

$$\pi_a(\mathcal{D}) = \{A \in \mathbb{S}^{\mathtt{h}} : (A,X) \in \mathcal{D} \text{ for some } X \in \mathbb{S}^{\mathtt{g}}\}. \tag{1.8}$$

Theorem 1.2 (Wurzelschmetterlingrealisierung). *Suppose* $r \in \mathbb{C}\langle\!\langle a,x \rangle\!\rangle$ *is an nc rational function with the SMR as in* (1.3). *Then*

(1) $\operatorname{dom}^+_{\mathrm{ver}} r$ *is a vertebral domain of convexity for* r;

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

(3) *there exist a positive integer* k, *a tuple* $\widehat{T} \in M_k(\mathbb{C})^{\mathrm{g}}$, *and a symmetric rational function* $w(a) \in \mathbb{C}\,\langle\!\!\langle a \rangle\!\!\rangle^{k\times k}$ *defined on* $\pi_a(\mathrm{dom}_{\mathrm{ver}}\, r)$ *such that*

$$\mathrm{dom}^+_{\mathrm{ver}}\, r = \Big\{(A,X)\in \mathrm{dom}_{\mathrm{ver}}\, r : w(A)\succeq 0,\ I-\sqrt{w(A)}\Big[\sum_{i=1}^{\mathrm{g}} \widehat{T}_i\otimes X_i\Big]\sqrt{w(A)}\succ 0\Big\};$$

(4) *there exist a rational function* $\ell(a,x) \in \mathbb{C}\,\langle\!\!\langle a,x \rangle\!\!\rangle^{k\times 1}$, *defined on* $\mathrm{dom}_{\mathrm{ver}}\, r$ *and linear in* x; *and a symmetric rational function* $\mathit{ff}(a,x) \in \mathbb{C}\,\langle\!\!\langle a,x \rangle\!\!\rangle$, *defined on* $\mathrm{dom}_{\mathrm{ver}}\, r$ *and affine linear in* x, *such that* r *admits the following realization, valid on* $\mathrm{dom}^+_{\mathrm{ver}}\, r$:

$$r = \ell(a,x)^* \sqrt{w(a)}\,\Big(I-\sum \sqrt{w(a)}\,\widehat{T}_i x_i \sqrt{w(a)}\Big)^{-1}\sqrt{w(a)}\,\ell(a,x) + \mathit{ff}(a,x).$$

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).

Corollary 1.3 ([23], Proposition 3.1). *Suppose* $\mathcal{D}$ *is a free set that is open in* x, *convex 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 exist* $\ell(a,x) \in \mathbb{C}\langle a,x\rangle$ *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*

$$p = \ell(a,x)^*\, w(a)\, \ell(a,x) + \mathit{ff}(a,x),$$

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

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 $x_1,\dots,x_{\mathrm{g}}, y_1,\dots,y_{\mathrm{h}}$.

An expression of the form

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

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 $(X,Y) \in \mathbb{S}^{\mathrm{g}} \times \mathbb{S}^{\mathrm{h}}$ is a *bilinear matrix inequality* (BMI) [19, 35, 46]. 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 [32].

Symmetric polynomials in two freely noncommuting variables x and y (so $\mathtt{g} = 1 = \mathtt{h}$) that are xy-convex essentially arise from BMIs. Here xy-convex means globally; that is, on all of $\mathbb{S}^1 \times \mathbb{S}^1$.

Theorem 1.4. *Suppose p is a symmetric polynomial in the two freely noncommuting variables x, y. If p is xy-convex, then there exist 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 [32]. 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 $V^* A^2 V = (V^* A V)^2$ it follows that $V^*(A, X)V \in \mathcal{D}$. In [32] 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 $V^* A_j^2 V = (V^* A_j V)^2$ 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 of nc rational functions 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 [44]; see also [4] or Proposition 4.3 in [29].

Proposition 2.1. *If*

$$q(x) = a^*\Big(J - \sum_{j=1}^m A_j x_j\Big)^{-1} a \quad \text{and} \quad q(x) = b^*\Big(K - \sum_{j=1}^m B_j x_j\Big)^{-1} b$$

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

A bit of algebra reveals that $S^*BS = A$. Thus $K - \sum B_j x_j = S^*(J - \sum A_j x_j)S$ and it follows that the definitions of $\operatorname{dom} r$ and $\operatorname{dom}^+ r$ 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 [30]. The *x-partial Hessian* of an SMR as in equation (1.3) is the rational function in $2\mathrm{g} + \mathtt{h}$ freely noncommuting variables,

$$r_{xx}(a,x)[h] = 2c^* R(a,x)\Big(\sum_i T_i h_i\Big) R(a,x)\Big(\sum_i T_i h_i\Big) R(a,x)c \tag{2.1}$$
$$= 2\Big[c^* R(a,x)\Big(\sum_i T_i h_i\Big)\Big] R_T(a,x) \Big[\Big(\sum_i T_i h_i\Big) R(a,x)c\Big],$$

where R is the resolvent

$$R(a,x) := \Big(J - \sum T_j x_j - \sum S_k a_k\Big)^{-1}, \tag{2.2}$$

$\Lambda_T[h] = \sum_{j=1}^{\mathrm{g}} T_j h_j$, and $R_T(a,x) = V_T^* R(a,x) V_T$ is defined as in (1.5). Compare with equation (5.3) in [29], where the *full* Hessian of a SMR is computed in detail. The x-partial Hessian is naturally *evaluated* at a tuple $(A, X, H) \in \mathbb{S}^{\mathtt{h}} \times \mathbb{S}^{\mathrm{g}} \times \mathbb{S}^{\mathrm{g}}$, 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 characterization of convexity in terms of Hessians in [30]. The proof is a straightforward modification of the one in [30], so is only sketched below.

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

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 positive linear functional $\lambda\colon \mathbb{S}_k \to \mathbb{R}$, the function $f_{A,\lambda}\colon S \to \mathbb{R}$ 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,

$$0 \le f''_{A,\lambda}(X)[H] = \lambda \circ r_{xx}(A,X)[H] \quad \text{for all } H.$$

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

2.2. $\operatorname{dom}^+ r$ is open in x and convex in x

In this section we show that $\operatorname{dom}^+ r$ is both open in x and convex in x. Let positive integers m and n, a matrix $D \in \mathbb{S}_n$ and a matrix $B \in M_{m,n}(\mathbb{C})$ be given. Let $V\colon \mathbb{C}^m \to \mathbb{C}^m \oplus \mathbb{C}^n$ denote the inclusion

$$Vx = \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{C}^m \oplus \mathbb{C}^n.$$

Define $L\colon \mathbb{S}_m \to \mathbb{S}_{m+n}$ by

$$L(\mathtt{X}) = \begin{pmatrix} \mathtt{X} & B \\ B^* & D \end{pmatrix}.$$

Let

$$\Omega = \{\mathtt{X} \in \mathbb{S}_n(\mathbb{C}) : \det L(\mathtt{X}) \neq 0\} \quad \text{and} \quad \Omega^+ = \{\mathtt{X} \in \Omega : V^* L(\mathtt{X})^{-1} V \succeq 0\}.$$

Lemma 2.3. *The set Ω^+ is open, convex, and a connected component of Ω.*

Before proving Lemma 2.3, we first establish the following result.

Lemma 2.4. *There exist a subspace $\mathcal{H} \subseteq \mathbb{C}^m$ and a self-adjoint operator F on $\mathcal{H}$ such that, with W equal the inclusion of $\mathcal{H}$ into $\mathbb{C}^m$,*

(1) $\mathtt{X} \in \Omega$ *if and only if* $W^* \mathtt{X} W - F$ *is invertible*; *and*
(2) $\mathtt{X} \in \Omega^+$ *if and only if* $W^* \mathtt{X} W - F \succ 0$.

Proof. The proof is straightforward in the case that D is invertible. Indeed, under the assumption that D is invertible, a standard Schur complement result says $L(\mathtt{X})$ is invertible if and only if the Schur complement of D,

$$S(\mathtt{X}) = \mathtt{X} - BD^{-1}B^*,$$

is invertible and further, in that case,

$$V^* L(\mathtt{X})^{-1} V = S(\mathtt{X})^{-1}.$$

Thus the result holds with $\mathcal{H} = \mathbb{C}^m$ and $F = BD^{-1}B^*$.

The result also holds trivially if $\Omega = \emptyset$ by choosing $\mathcal{H} = \{0\}$. Thus, for the remainder of this proof, assume D is not invertible and $\Omega \neq \emptyset$. In particular, $\ker D \cap \ker B \neq \{0\}$.

With respect to the orthogonal direct sum $\mathbb{C}^n = \ker D \oplus \ker D^\perp$,

$$D = \begin{pmatrix} 0 & 0 \\ 0 & D_0 \end{pmatrix} \quad \text{and} \quad L(\mathtt{X}) = \begin{pmatrix} \mathtt{X} & B_1 & B_2 \\ B_1^* & 0 & 0 \\ B_2^* & 0 & D_0 \end{pmatrix},$$

with D_0 invertible. It follows that $B_1 \colon \ker D \to \mathbb{C}^m$ is one-to-one, as otherwise $L(\mathtt{X})$ is never invertible, violating the assumption $\Omega \neq \emptyset$.

With respect to the orthogonal decomposition $\mathbb{C}^m = \operatorname{rng} B_1 \oplus \operatorname{rng} B_1^\perp$,

$$B_1 = \begin{pmatrix} B_{1,1} \\ 0 \end{pmatrix} : \ker D \to \mathbb{C}^m.$$

In particular, $B_{1,1}$ is invertible. In these coordinates ($\mathbb{C}^m = \operatorname{rng} B_1 \oplus \operatorname{rng} B_1^\perp$ and $\mathbb{C}^n = \ker D \oplus \ker D^\perp$),

$$L(\mathtt{X}) = \begin{pmatrix} \mathtt{X}_{1,1} & \mathtt{X}_{1,2} & B_{1,1} & B_{1,2} \\ \mathtt{X}_{1,2}^* & \mathtt{X}_{2,2} & 0 & B_{2,2} \\ B_{1,1}^* & 0 & 0 & 0 \\ B_{1,2}^* & B_{2,2}^* & 0 & D_0 \end{pmatrix}.$$

Since D_0 is invertible, $L(\mathtt{X})$ is invertible if and only if the Schur complement of D_0,

$$T(\mathtt{X}) = \begin{pmatrix} \mathtt{X}_{1,1} & \mathtt{X}_{1,2} & B_{1,1} \\ \mathtt{X}_{1,2}^* & \mathtt{X}_{2,2} & 0 \\ B_{1,1}^* & 0 & 0 \end{pmatrix} - \begin{pmatrix} B_{1,2} \\ B_{2,2} \\ 0 \end{pmatrix} D_0^{-1} \begin{pmatrix} B_{1,2}^* & B_{2,2}^* & 0 \end{pmatrix},$$

is invertible. Writing $T(\mathtt{X})$ as

$$\begin{pmatrix} \mathtt{X}_{1,1}-C_{1,1} & \mathtt{X}_{1,2}-C_{1,2} & B_{1,1} \\ \mathtt{X}_{1,2}^*-C_{1,2}^* & \mathtt{X}_{2,2}-C_{2,2} & 0 \\ B_{1,1}^* & 0 & 0 \end{pmatrix},$$

we observe that $T(\mathtt{X})$ is invertible if and only if $\mathtt{X}_{2,2}-C_{2,2}$ is invertible, proving item (1) with $\mathscr{H} = \operatorname{rng} B_1^{\perp}$ and $F = C_{2,2}$. Moreover,

$$T(\mathtt{X})^{-1} = \begin{pmatrix} 0 & 0 & B_{1,1}^{-1} \\ 0 & (\mathtt{X}_{2,2}-C_{2,2})^{-1} & * \\ B_{1,1}^{-*} & * & * \end{pmatrix}.$$

Since the upper 3×3 block of $L(\mathtt{X})^{-1}$ is $T(\mathtt{X})^{-1}$, it follows that

$$V^*L(\mathtt{X})^{-1}V = \begin{pmatrix} 0 & 0 \\ 0 & (\mathtt{X}_{2,2}-C_{2,2})^{-1} \end{pmatrix}.$$

Hence $\mathtt{X}\in\Omega^+$ if and only if $\mathtt{X}_{2,2}-C_{2,2}\succ 0$, proving item (2) again with $\mathscr{H} = \operatorname{rng} B_1^{\perp}$ and $F = C_{2,2}$. ■

Proof of Lemma 2.3. Since, by Lemma 2.4, $\mathtt{X}\in\Omega^+$ if and only if $W^*\mathtt{X}W - F\succ 0$, the set Ω^+ is both open and convex. Since Ω^+ is convex, to prove Ω^+ is a connected component of Ω, it suffices to prove Ω^+ is closed in Ω. To this end, suppose $(\mathtt{X}_n)_n$ is a sequence from Ω^+ that converges to $\mathtt{X}\in\Omega$. It follows from Lemma 2.4 that $W^*\mathtt{X}_nW - F\succ 0$ for each n and hence, after taking a limit, $W^*\mathtt{X}W - F\succeq 0$. On the other hand, $\mathtt{X}\in\Omega$ implies $W^*\mathtt{X}W - F$ is invertible by Lemma 2.4. Hence $W^*\mathtt{X}W - F\succ 0$ and therefore $\mathtt{X}\in\Omega^+$ by yet another application of Lemma 2.4. ■

Proposition 2.5. *Suppose $r\in\mathbb{C}\langle\!\langle a,x\rangle\!\rangle$ is an nc rational function with the SMR as in* (1.3) *and $A\in\mathbb{S}_n^{\mathtt{h}}$. The set*

$$\Omega[A]^+ = \{X\in\mathbb{S}_n^{\mathtt{g}} : (A,X)\in\operatorname{dom}^+ r\}$$

is open, convex and a connected component of the set

$$\Omega[A] = \{X\in\mathbb{S}_n^{\mathtt{g}} : (A,X)\in\operatorname{dom} r\}\subseteq\mathbb{S}_n^{\mathtt{g}}.$$

Proof. Let N denote the size of realization. Thus $J\in M_N(\mathbb{C})$. Without loss of generality, we assume that $\operatorname{rng} T\oplus\operatorname{rng} T^{\perp}$ decomposes $\mathbb{C}^N$ as $\mathbb{C}^a\oplus\mathbb{C}^b$. Express J,S,T with respect to this orthogonal decomposition as

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

Let $B = J_{1,2}\otimes I - \sum S_{k,1}\otimes A_k\in M_{a,b}(\mathbb{C})\otimes\mathbb{S}_n$ and $D = J_{2,2}\otimes I - \sum_k S_{k,2}\otimes A_k\in\mathbb{S}_b\otimes\mathbb{S}_n\subseteq\mathbb{S}_{bn}$ and define $L\colon\mathbb{S}_{am}\to\mathbb{S}_{am+bn}$ by

$$L(\mathtt{X}) = \begin{pmatrix} \mathtt{X} & B \\ B^* & D \end{pmatrix}$$

and let V denote the inclusion of $\mathbb{C}^a \otimes \mathbb{C}^n$ into $(\mathbb{C}^a \otimes \mathbb{C}^n) \oplus \mathbb{C}^b \otimes \mathbb{C}^n$. Let $\Omega = \{\mathtt{X} \in \mathbb{S}_{am} : \det L(\mathtt{X}) \neq 0\}$ and let

$$\Omega^+ = \{\mathtt{X} \in \Omega : V^* L(\mathtt{X})^{-1} V \succeq 0\}.$$

By Lemma 2.3, Ω^+ is open, convex and a connected component of Ω. In particular, Ω^+ is closed in Ω.

Define $\Lambda\colon \mathbb{S}_n^g \to \mathbb{S}_{an}$ by

$$\Lambda(X) = \Big(J_{1,1} \otimes I - \sum_k S_{k,0} \otimes A_k\Big) - \sum_j T_{j,0} \otimes X_j.$$

Observe that Λ is affine linear, $\Omega[A] = \Lambda^{-1}(\Omega)$ and $\Omega[A]^+ = \Lambda^{-1}(\Omega^+)$. Thus, since Λ is continuous and Ω^+ is open, $\Omega[A]^+$ is open. Likewise, since Λ is affine linear and Ω^+ is convex, $\Omega[A]^+$ is convex and thus connected. Finally, since $\Omega[A]^+$ is connected, to show it is a component of $\Omega[A]$, it suffices to observe that it is closed since it is the inverse image under the continuous map $\Lambda|_{\Omega[A]} : \Omega[A] \to \Omega$ of the closed (in Ω) set Ω^+. ∎

2.3. 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 2.6. *If $r \in \mathbb{C}\langle a,x \rangle$ is an nc rational function with the SMR as in (1.3), then*

(1) $\operatorname{dom}^+ r$ *is a domain of partial convexity for* r;

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

Corollary 2.7 ([29]). *Suppose $r \in \mathbb{C}\langle x \rangle$. 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. That $\operatorname{dom}^+ r$ is open in x and convex in x was established in Proposition 2.5. Thus to prove $\operatorname{dom}^+ r$ is a domain of partial convexity for r, it remains to prove that r is convex in x on $\operatorname{dom}^+ r$, a statement that follows from Proposition 2.8 below. Item (2) of Theorem 2.6 is an immediate consequence of the converse portion of Proposition 2.8.

Proposition 2.8. *Let r denote the rational function of* (1.3) *and suppose* $\mathcal{E} \subseteq \operatorname{dom} r$ *is a free set that is open in x and convex 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 $\mathcal{E}$.

2.3.1. The CHSY lemma. In this section we establish a variant of the CHSY lemma [10] (see also [9, 48]) suitable for a proof of Proposition 2.8, starting with the of independent interest Lemma 2.9 below.

Lemma 2.9. *If $\xi_1, \dots, \xi_K \in \mathbb{C}\langle x \rangle$ are linearly independent rational functions in g variables, m is a positive integer and $\mathcal{U}$ is a free open subset of $\mathbb{S}^g$ with $\mathcal{U}_1 \neq \emptyset$, then there* *exist a positive integer* M*, an* $X \in \mathfrak{U}_M$ *and a matrix* $w \in M_{m,M}(\mathbb{C})$ *such that*

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

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

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

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

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

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 $(w, X) \in \mathcal{S}$ such that

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

for all $(z, Y) \in \mathcal{S}$. Suppose $\alpha \in \mathcal{V}_{(w,X)}^{\perp}$. From equation (2.3), it follows that $\alpha \in \mathcal{V}_{(z,Y)}^{\perp}$ for all $(z, Y) \in \mathcal{S}$. Write $\alpha \in \mathbb{C}^K \otimes \mathbb{C}^m$ as $\alpha = \sum \alpha_j \otimes e_j$, where $\{e_1, \ldots, e_m\}$ 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 $Y \in \mathfrak{U}_n$ be given. Given a vector $f \in \mathbb{C}^n$, let $w_f = e_j f^*$. Since $\alpha \in \mathcal{V}_{(Y,w_f)}^{\perp}$,

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

Thus, for each j, the rational function $\xi = \sum_{s=1}^K \overline{(\alpha_j)_s}\, \xi_s$ vanishes on $\mathfrak{U}$. By hypothesis, $\mathfrak{U}_1 \neq \emptyset$ and $\mathfrak{U}$ is an open free set. Hence, for each n, the set $\mathfrak{U}_n$ is nonempty and open and ξ vanishes identically on $\mathfrak{U}$. Hence ξ is identically zero since there are no rational identities [6]; cf. the definition of nc rational functions via matrix evaluations in [29]. The desired conclusion follows. ∎

Lemma 2.10. *If the realization* (1.3) *is minimal and of size* N *and* $\mathfrak{U}$ *is a free open subset of* dom r*, then, for each* $m \in \mathbb{N}$*, there exist an* M*,* $(A, X) \in \mathfrak{U}$*, a* $w \in M_{m,M}(\mathbb{C})$ *and an* $H \in \mathbb{S}_M^{\mathrm{g}}$ *such that*

$$V_{A,X,H,w} := \Big\{ (I_N \otimes w)\Big(\sum_i T_i \otimes H_i\Big) R(A, X)(c \otimes I_M)v \mid v \in \mathbb{C}^M \Big\} = (\operatorname{rng} T) \otimes \mathbb{C}^m.$$

Proof. Let K denote the dimension of $\operatorname{rng} T$ and let U be 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 g $N \times 1$ matrix

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

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

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

Further, since the entries of Q are linearly independent, the set $\{\xi_1,\dots,\xi_K\}$ is linearly independent. By Lemma 2.9, for each positive integer m, there exist a positive integer M, a tuple $(H,A,X) \in \mathbb{S}_M^{\mathrm{g}} \times \mathcal{U}_M$ and a matrix $w \in M_{M,m}(\mathbb{C})$ such that the conclusion of Lemma 2.9 holds, completing the proof. ■

2.3.2. Proof of Proposition 2.8. 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, in the following way. Let

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

where the x_i^j, k_i and a_i^j form a $2(2\mathrm{g}+\mathtt{h})$ collection of freely noncommuting variables. In these coordinates the $(1,1)$ entry of r_{xx} in (2.1) equals

$$(2.4) \qquad 2\Big[c^* R(a^1,x^1)\Big(\sum_i T_i k_i\Big)\Big]\, R(a^2,x^2)\,\Big[\Big(\sum_i T_i (k_i)^*\Big) R(a^1,x^1)c\Big].$$

We next apply Lemma 2.10. 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 $H \in \mathbb{S}_M^{\mathrm{g}}$ satisfying the conclusion of Lemma 2.10. 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

$$\left\{\Big[\sum T_j \otimes K_j\Big] R(A^1,X^1)(c \otimes I) : v \in \mathbb{C}^n\right\}$$

spans $\mathrm{rng}\, T \oplus \mathbb{C}^m$, it now follows that $R_T(A^2,X^2) \succeq 0$. ■

2.3.3. Proof of Theorem 2.6. For item (1), Proposition 2.5 says that $\mathrm{dom}^+ r$ is open in x and convex in x. The forward direction of Proposition 2.8 says that r is convex in x on $\mathrm{dom}^+ r$.

The converse direction of Proposition 2.8 says that, if $\mathcal{D}$ is a full domain of convexity for r, then $R_T \succeq 0$ on $\mathcal{E}$. Thus $\mathcal{E} \subseteq \mathrm{dom}^+ r$. ■

2.4. Realizations for partial convexity

Proposition 2.11. *The rational function $r \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle$ of equation (1.3) admits the realization*

$$r = c^*\Big(J - \sum S_i a_i\Big)^{-1} c + c^*\Big(J - \sum S_i a_i\Big)^{-1} \sum T_i x_i \Big(J - \sum S_i a_i\Big)^{-1} c$$
$$+ c^*\Big(J - \sum S_j a_j\Big)^{-1} \sum T_i x_i \Big(J - \sum T_j x_j - \sum S_k a_k\Big)^{-1}$$
$$\cdot \sum T_i x_i \Big(J - \sum S_i a_i\Big)^{-1} c. \tag{2.5}$$

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

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

Recall the definitions of V_T and $\pi_a(\mathcal{D})$ from equations (1.5) and (1.8), respectively.

Theorem 2.12 (Wurzelschmetterlingrealisierung). *Suppose $r \in \mathbb{C}\langle\!\langle a, x\rangle\!\rangle$ is symmetric with SMR as in equation (1.3).*

(1) *The set $\operatorname{dom}^+_{\mathrm{ver}} r$ is a vertebral domain of convexity for r.*
(2) *If $\mathcal{D} \subseteq \operatorname{dom} r$ is a full vertebral domain of convexity for r, then $\mathcal{D} \subseteq \operatorname{dom}^+_{\mathrm{ver}} r$.*

Let $\widehat{T}_j = V_T^* T_j V_T$ and let k be the dimension of $\operatorname{rng} T$. There exist a rational function $w(a) \in M_k(\mathbb{C}\langle\!\langle a\rangle\!\rangle)$, defined on $\pi_a(\operatorname{dom}_{\mathrm{ver}} r)$ and positive semidefinite on $\pi_a(\operatorname{dom}^+_{\mathrm{ver}} r)$; rational functions $\ell_j(a) \in \mathbb{C}\langle\!\langle a\rangle\!\rangle^k$ for $1 \le j \le g$, that are defined on $\operatorname{dom}_{\mathrm{ver}} r$; and a rational function $f\!f(a, x)$ that is affine linear in x and defined on $\operatorname{dom}_{\mathrm{ver}} r$ such that, with

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

(3) *if $(B, Y) \in \operatorname{dom}_{\mathrm{ver}} r$; then $I - (\sum T_j \otimes Y_j) w(B)$ is invertible and*

$$r(B, Y) = \ell(B, Y)^* w(B)\Big(I - \Big(\sum \widehat{T}_i \otimes Y_i\Big) w(B)\Big)^{-1} \ell(B, Y) + f\!f(B, Y);$$

(4) *and we have*

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

and

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

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

(b) *r has the representation*

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

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 $\mathrm{dom}^+_{\mathrm{ver}}\, r$ *and any polynomial of the form of equation (2.8) is convex in x 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_j, S_j as block 2×2 matrices with respect to the orthogonal decomposition $\mathrm{rng}\, T \oplus \mathrm{rng}\, T^{\perp}$ as

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

Proof of Theorem 2.12. By definition, $\mathrm{dom}^+_{\mathrm{ver}}\, r$ is convex in x and a subset of $\mathrm{dom}^+\, r$. Thus, since r is convex in x on $\mathrm{dom}^+\, r$, it is also convex in x on $\mathrm{dom}^+_{\mathrm{ver}}\, r$. Thus item (1) holds.

If $\mathcal{D} \subseteq \mathrm{dom}\, r$ is full vertebral domain of convexity for r, then $\mathcal{D}$ is a full domain of partial convexity for r. Hence, by Theorem 1.1, $\mathcal{D} \subseteq \mathrm{dom}^+\, r$. If $(A,X) \in \mathcal{D}$, then $(A,0) \in \mathcal{D}$, since $\mathcal{D}$ is a vertebral set. Thus both (A,X) and $(A,0) \in \mathrm{dom}^+\, r$ and hence $(A,X) \in \mathrm{dom}^+_{\mathrm{ver}}\, r$, proving item (2).

By Proposition 2.11, r admits the caterpillar realization (2.5), whose resolvent

$$R(a,x) = \begin{pmatrix} J_{11} - \sum \widehat{T}_j x_j - \sum S^j_{11} a_j & J_{12} - \sum S^j_{12} a_j \\ J^*_{12} - \sum S^{j*}_{12} a_j & J_{22} - \sum S^j_{22} a_j \end{pmatrix}^{-1}$$

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

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

contains $\mathrm{dom}\, W$.

Suppose $(A,X) \in \mathrm{dom}_{\mathrm{ver}}\, r$. Thus $(A,0), (A,X) \in \mathrm{dom}\, r$, and hence

$$\begin{aligned} R^{-1}(A,X)W(A) &= \Big(J - \sum T_j \otimes X_j - \sum S_k \otimes A_k\Big) W(A) \\ &= I - \Big(\sum T_j \otimes X_j\Big) W(A) = \begin{pmatrix} \big(I - \sum \widehat{T}_j \otimes X_j\big) w(A) & * \\ 0 & I \end{pmatrix}. \end{aligned} \tag{2.10}$$

It follows that $I - (\sum \widehat{T}_j \otimes X_j) w(A)$ is invertible whenever $(A,0), (A,X) \in \mathrm{dom}\, r$, establishing the first half of item (3). Moreover, in that case, from equation (2.10),

$$\begin{aligned} R(A,X) &= W(A) \begin{pmatrix} \big(I - \sum \widehat{T}_j \otimes X_j\big) w(A) & * \\ 0 & I \end{pmatrix}^{-1} \\ &= \begin{pmatrix} w(A)\big(I - \sum \widehat{T}_j \otimes X_j\big) w(A) & * \\ 0 & I \end{pmatrix}^{-1} \end{aligned}$$

and thus

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

Letting ff denote the affine linear in x term from the caterpillar realization of equation (2.5),

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

when $(A,X) \in \operatorname{dom}_{\mathrm{ver}} r$, proving item (3).

Given square matrices P and Q of the same size, the eigenvalues of PQ and QP are the same. Now suppose $(X,A) \in \operatorname{dom}_{\mathrm{ver}} r$ and $w(A) \succeq 0$ and let

$$\mathtt{T} = \sum \widehat{T}_i \otimes X_i.$$

Choosing $P = \mathtt{T}\sqrt{w(A)}$ and $Q = \sqrt{w(A)}$, it follows that $\mathtt{T}w(A)$ and $\sqrt{w(A)}\mathtt{T}\sqrt{w(A)}$ have the same eigenvalues. Thus, in view of item (3), if $I - \sqrt{w(A)}\mathtt{T}\sqrt{w(A)} \succeq 0$, then $I - \sqrt{w(A)}\mathtt{T}\sqrt{w(A)} \succ 0$. Hence

$$R_T(A,X) = w(A)(I - \mathtt{T}w(A))^{-1} = \sqrt{w(A)}\left(I - \sqrt{w(A)}\mathtt{T}\sqrt{w(A)}\right)^{-1}\sqrt{w(A)} \succeq 0$$

and therefore $(A,X) \in \operatorname{dom}^+ r$. The assumption $R_T(A,0) = w(A) \succeq 0$ is equivalent to $(A,0) \in \operatorname{dom}^+ r$. Hence $(A,X) \in \operatorname{dom}^+_{\mathrm{ver}} r$.

Conversely, if $(A,X) \in \operatorname{dom}^+_{\mathrm{ver}} r$, then $w(A) \succeq 0$ and, since $\operatorname{dom}^+_{\mathrm{ver}} r$ is convex in x and $(A,0) \in \operatorname{dom}^+_{\mathrm{ver}} r$, for each $0 \le t \le 1$, the matrix $I - t\mathtt{T}w(A)$ is invertible and hence so is $M(t) = I - \sqrt{w(A)}\mathtt{T}\sqrt{w(A)}$. Since $M(0)$ is positive and $M(t)$ is invertible and self-adjoint for $0 \le t \le 1$, it follows that $M(1) \succ 0$ and the proof of item (4) is complete.

In the case r is a polynomial, $R(a,x)$ is globally defined (has no singularities) and is therefore a (matrix-valued) polynomial by Corollary 3.4 in [36]. 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 $(A - \mathtt{a}I, 0) \in \mathcal{D}$ 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 $(0,0) \in \mathcal{D}$. Then $w(0) = V_T^* R(0,0) V_T$ is positive semidefinite by Theorem 2.6 since we have now convexity in x in a neighborhood of 0. Next $R(0,0) = J^{-1} = J$ and so $w(0) = J_{1,1} \succeq 0$. Since r is a polynomial (and the realization is minimal), TJ is (jointly) nilpotent by Corollary 3.4 in [36]. But

$$TJ = \begin{pmatrix} \hat{T} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} J_{11} & J_{12} \\ J_{12}^* & J_{22} \end{pmatrix} = \begin{pmatrix} \hat{T} J_{11} & \hat{T} J_{12} \\ 0 & 0 \end{pmatrix},$$

whence $\hat{T} J_{12}$ is (jointly) nilpotent. Thus $Y = \sqrt{J_{11}} T_j \sqrt{J_{11}}$ 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 in x and is convex in x on the set $\{(A,X) : w(A) \succeq 0\}$, which includes $\pi_a(\mathcal{D}) \times \mathbb{S}^{\mathtt{g}}$. ∎

Corollary 2.13. *Let $\mathfrak{D}$ be a vertebral set. Let $r \in \mathbb{C}\langle\!\langle a, x \rangle\!\rangle$ be an nc rational function in two classes of variables $x = (x_1, \ldots, x_g)$ and $a = (a_1, \ldots, a_h)$. Let r have a SMR (1.3). Consider the matrices in block form based on* rng T *in equation (2.9) and let k denote the dimension of* rng T.

If J_{22} is invertible, then the function r is convex in x on $\mathfrak{D}$ if and only if there exist 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

$$r = \ell(a, x)^* \left(m(a) - \sum \widehat{T}_i x_i\right)^{-1} \ell(a, x) + \mathit{ff}(a, x),$$

where $\mathit{ff}(a, x) \in \mathbb{C}\langle\!\langle a, x \rangle\!\rangle$ is affine linear in x, and the resolvent $(m(a) - \sum \widehat{T}_i x_i)^{-1}$ is positive on a dense subset of $\mathfrak{D}_n$ for large n.

Proof. This result follows by using the Schur complement form for the inverse of a block matrix in Proposition 2.11; the positivity condition follows from Proposition 2.8. ∎

3. A polynomial factorization

In this section we introduce an auxiliary operation $\mathcal{E}$ 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 $\mathcal{E}\rho$ is (matrix) positive. This result is a key ingredient in the proof of Theorem 1.4, which appears in Section 4, characterizing xy-convex polynomials.

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 $(\otimes)$. It is known as the Khatri–Rao product. Let $V_1 = \binom{I}{0}$ and $V_2 = \binom{0}{I}$ with respect to the block decomposition of A and define W_1 and 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 $\mathcal{E}p$ denote the matrix polynomial in the six variables $\{s_{j,0}, s_{j,1}, s_{j,2} : 1 \le j \le 2\}$ defined by

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

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

$$S_j = \begin{pmatrix} S_{j,0} & S_{j,1} \\ S_{j,1}^* & S_{j,2} \end{pmatrix} \in M_\mu(\mathbb{C}) \otimes M_2(\mathbb{C}), \tag{3.1}$$

using $\circledast$ via

$$\mathcal{E}p(S) = \sum_{j,k=0}^{2} p_{j,k} \circledast S_j S_k \in M_\mu(\mathbb{C}) \otimes M_2(\mathbb{C}).$$

By contrast,

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

However, p and $\mathcal{E}p$ 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,*

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

In particular, if $p(S) \succeq 0$, then $\mathcal{E}p(S) \succeq 0$ 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 $\mathcal{E}\rho(S) \succeq 0$ 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 exist an $N \le 12$ and $q_0, q_1, q_2 \in M_{N,2}(\mathbb{C})$ such that*

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

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 $\mathcal{S}$ 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 $\beta \in \langle x_1, x_2\rangle_1$,

$$T_{\beta,x_0} = T_{x_0,\beta}, \quad T_{x_0,x_0} \in \operatorname{span}\{e_1e_1^*, e_2e_2^*\}.$$

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

Define $\psi\colon \mathcal{S} \to M_2(\mathbb{C})$ by

$$\psi\left(T_{\alpha,\beta}\right) = \sum_{\alpha,\beta\in\langle x_1,x_2\rangle_1} \rho_{\alpha,\beta} * T_{\alpha,\beta} = \sum_{\alpha,\beta\in\langle x_1,x_2\rangle_1} \rho_{\alpha,\beta} \circledast T_{\alpha,\beta}. \tag{3.3}$$

Proposition 3.4. *The mapping* ψ *of equation* (3.3) *is completely positive* (*cp*).

Proof. To prove that ψ is cp, let a positive integer n and positive definite $Z \in M_n(\mathbb{C}) \otimes \mathcal{S}$ 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 $\Theta_{j,j} = (Z_{x_0,x_0})_{j,j}^{-1}$. 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

$$\Psi = \begin{pmatrix} \Psi_{1,1} & 0 \\ 0 & \Psi_{2,2} \end{pmatrix} \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C}),$$

where

$$\Psi_{a,a} = \begin{pmatrix} (Z_{x_0,x_0})_{a,a} & 0 \\ 0 & I_m \end{pmatrix} \in M_{n+m}(\mathbb{C}).$$

Let, for $j = 1, 2$,

$$W_j = \big((W_j)_{a,b}\big) \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C}), \tag{3.4}$$

where

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

Since $Z_{\alpha,x_0} = Z_{x_0,\alpha}$ is self-adjoint, so is W_j. By construction,

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

Let

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

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

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

From the definition (3.3) of ψ (and letting ψ also denote its ampliations $\psi \otimes I_\ell$, where I_ℓ is the identity on $M_\ell(\mathbb{C})$),

$$\psi(W) = \rho_{x_0,x_0} \circledast \Psi + \rho_{x_0,x_1} \circledast W_1 + \rho_{x_0,x_2} \circledast W_2 + \sum_{j,k=1}^{2} \rho_{x_j,x_k} \circledast W_j\Psi^{-1}W_k. \tag{3.5}$$

By definition of the $\circledast$ operation, given

$$R = \begin{pmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \end{pmatrix} \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C}), \tag{3.6}$$

$$R_{i,j} = \begin{pmatrix} R^{1,1}_{i,j} & R^{1,2}_{i,j} \\ R^{2,1}_{i,j} & R^{2,2}_{i,j} \end{pmatrix} \in M_n(\mathbb{C}) \oplus M_m(\mathbb{C}),$$

$$\tau = \begin{pmatrix} \tau_{1,1} & \tau_{1,2} \\ \tau_{2,1} & t_{2,2} \end{pmatrix} \in M_2(\mathbb{C})$$

we have $\tau \circledast R = (\tau_{i,j} R_{i,j})$ and hence

$$V^* [\tau \circledast R]\, V = \big(\tau_{i,j} R^{1,1}_{i,j}\big) = \tau \circledast \widetilde{R},$$

where $\widetilde{R} = \left(R^{1,1}_{i,j}\right)^2_{i,j=1}$. Hence,

$$\begin{aligned} V^* \left[\rho_{x_0,x_0} \circledast \Psi\right] V &= \rho_{x_0,x_0} \circledast Z_{x_0,x_0}, \\ V^* \left[\rho_{x_0,x_j} \circledast W_j\right] V &= \rho_{x_j,x_k} \circledast Z_{x_j,x_k}, \\ V^* \left[\rho_{x_j,x_k} \circledast W_j \Psi^{-1} W_k\right] V &= \rho_{x_j,x_k} \circledast Z_{x_j,x_k}. \end{aligned}$$

Thus, from equation (3.5),

$$V^* \psi(W) V = \psi(Z).$$

Hence, to prove $\psi(Z) \succeq 0$ it suffices to show $\psi(W) \succeq 0$.

With R and τ as in equation (3.6), given a block diagonal matrix

$$D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C}),$$

we have

$$\begin{aligned} D\,[\tau \circledast R]\, D &= \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \left(\tau_{i,j} R_{i,j}\right) \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \\ &= \left(\tau_{i,j} D_i R_{i,j} D_j\right) = \tau \circledast (DRD). \end{aligned}$$

Hence, $S_j = \Psi^{-1/2} W_j \Psi^{-1/2} \in M_{n+m}(\mathbb{C}) \otimes M_2(\mathbb{C})$ are self-adjoint and

$$\Psi^{-1/2}\, \psi(W)\, \Psi^{-1/2} = \sum_{j,k} \Psi^{-1/2} \left[\rho_{j,k} \circledast W_{j,k}\right] \Psi^{-1/2} = \sum_{j,k} \rho_{j,k} \circledast S_j S_k = \mathcal{E}\rho(S).$$

By hypothesis $\mathcal{E}\rho(S) \succeq 0$ and hence $\psi(W) \succeq 0$. Thus $\psi(Z) \succeq 0$ under the extra assumption that $Z \succ 0$.

Now suppose $Z \in M_n(\mathbb{C}) \otimes \mathcal{S}$ is positive semidefinite. Since the identity is contained in $M_n(\mathbb{C}) \otimes \mathcal{S}$, for each $\epsilon > 0$, the matrix $Z + \epsilon I$ is positive definite and in $M_n(\mathbb{C}) \otimes \mathcal{S}$. Thus, by what has already been proved, $\psi(Z + \epsilon I) \succeq 0$ and hence, by letting ϵ tend to 0, it follows that $\psi(Z) \succeq 0$ and the proof is complete. ∎

Proof of Theorem 3.3. Since, by Proposition 3.4, ψ is cp it extends, by the Arveson extension theorem ([41], Theorem 7.5), to a cp map $\varphi\colon M_2(\mathbb{C}) \otimes M_3(\mathbb{C}) \to M_2(\mathbb{C})$. By a well-known result of Choi ([41], Theorem 3.14), its Choi matrix

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

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, in particular,

$$F_{j,a}^* F_{k,b} = \varphi(e_a e_b^* \otimes x_j x_k^*). \tag{3.7}$$

For $q_j = (F_{j,1}e_1 \;\; F_{j,2}e_2) \in M_{N,2}(\mathbb{C})$, we have $q_j^* q_k = (e_a^* F_{j,a}^* F_{k,b} e_b)_{a,b=1}^2 \in M_2(\mathbb{C})$. So, using (3.7), for $a = 1, 2$,

$$\begin{aligned}(\rho_{0,0})_{a,a} &= \left(\rho_{0,0} \circledast e_a e_a^*\right)_{a,a} = \psi(e_a e_a^* \otimes x_0 x_0^*)_{a,a} \\ &= \varphi(e_a e_a^* \otimes x_0 x_0^*)_{a,a} = e_a^* F_{0,a}^* F_{0,a} e_a = (q_0^* q_0)_{a,a}.\end{aligned}$$

Hence equation (3.2) holds. Next, for $\ell = 1, 2$ and $1 \le a, b \le 2$,

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

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 $1 \le j, k \le 2$) by computing, for $1 \le a, b \le 2$,

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

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 2.12 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 [25]) 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 that $((X, Y), V)$ is an xy-pair; that is, $V^*YXV = V^*YVV^*XV$. Thus, if p is xy-convex, then

$$\begin{aligned} 0 &\preceq (I_\mu \otimes V)^* p(X, Y)(I_\mu \otimes V) - p(V^*(X, Y)V) \\ &= (I_\mu \otimes V)^* p(X, Y)(I_\mu \otimes V) - p(X_0, Y_0). \end{aligned}$$

To establish the reverse implication, given an xy-pair $((X, Y), V)$, decompose the space (X, Y) act upon as rng $V \oplus (\text{rng } V)^\perp$ 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 and 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 $(A\ 0)$ and $(0\ C)$, respectively. ∎

Consider the following list of monomials:

$$\begin{aligned} \mathcal{L} = \{&1, x, y, x^2, y^2, xy, yx, xy^2, y^2x, x^2y, yx^2, xyx, \\ &yxy, xyxy, yxyx, xy^2x, yx^2y\}. \end{aligned} \tag{4.2}$$

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 $\mathcal{L}$.*

Proof. The degree bounds follow from Theorem 2.12. 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 $[\mathcal{L}]$ denote the $\mathbb{C}$-vector space with basis $\mathcal{L}$ 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 $p \in [\mathcal{L}]$.*

Proof. Given (X_1, Y) and (X_2, Y), let $V = \frac{1}{\sqrt{2}}(I\ I)^T$ 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 $p \in [\mathcal{L}]$. Denote by $\{s_0, t_0, \alpha, \beta_j, \gamma, \delta_j : 0 \le j \le 2\}$ freely noncommuting variables with s_0, t_0, β_0, β_2, δ_0 and δ_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 and 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 $p \in [\mathcal{L}]$,

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

The proof of the following lemma is routine.

Lemma 4.4. *If $p = \sum_{u \in \mathcal{L}} p_u u \in [\mathcal{L}]$, then $H^{xy}p$ is a function of α, γ, s_0, t_0, δ_0, δ_1, β_1 and β_2 with the explicit form*

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

Lemma 4.5. *If $p \in [\mathcal{L}]$ and $H^{xy}p = 0$, then p is an xy-pencil. If $p, q \in [\mathcal{L}]$ 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 $p = \sum_{w \in \mathcal{L}} p_w w$ 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 a $p \in [\mathcal{L}]$ has a border vector-middle matrix representation that we now describe. Since $p \in [\mathcal{L}]$,

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

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

$$\begin{aligned}
\mathcal{L}_* &= \{x^2, y^2, xyx, yxy, xy^2, y^2x, x^2y, yx^2, xy^2x, xyxy, yxyx, yx^2y\} \\
&= \mathcal{L} \setminus \{1, x, y, xy, yx\}.
\end{aligned}$$

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 $\mathfrak{M}_{j,k}(\beta_1, \beta_2, \delta_0, \delta_1)$ denote the 2×2 matrix polynomials

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

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 $p \in [\mathcal{L}]$ is symmetric, then*

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

Proposition 4.7 shows that 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 exist $A \in M_{2,M}(\mathbb{C})$ and $C \in M_{2,N}(\mathbb{C})$ such that

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

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 that 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,

$$Q = Q(\delta_{a,b}, \beta_{a,b}) \tag{4.3}$$

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

and, given $S = (S_1, S_2)$ of the block form of equation (3.1), we have $Q(S) \succeq 0$ since $\mathfrak{M}(S_{2,1}, S_{2,2}, S_{1,0}, S_{1,1}) \succeq 0$ by Proposition 4.7.

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

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

and observe $\mathcal{E}P(\sigma) = Q(\sigma)$. Thus $\mathcal{E}P(S) \succeq 0$ for all tuples of Hermitian matrices of the form (3.1). Hence Theorem 3.3 produces an N and $F = \sum F_j s_j$, 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,

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

Hence, letting $\{e_1, e_2\}$ denote the standard orthonormal basis for $\mathbb{C}^2$, $F_1 e_2 = 0 = F_2 e_1$. In particular, $e_1^* F_2^* F_0 = 0$. Now set $\Lambda_x = F_0 e_1$, $\Lambda_y = F_0 e_2$, $\Lambda_{yx} = F_1 e_1$ and $\Lambda_{xy} = F_2 e_2$

and verify

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

Let

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

where Λ denotes the xy-pencil

$$\Lambda = \Lambda_x x + \Lambda_y y + \Lambda_{xy} xy + \Lambda_{yx} yx.$$

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

Funding. MJ: Research Supported by NSF grant DMS-1900364. IK: Supported by the Slovenian Research Agency grants J1-8132, N1-0217 and P1-0222. Partially supported by the Marsden Fund Council of the Royal Society of New Zealand. MM: Research partially supported by NSF grant DMS-1565243. SM: Research supported by NSF grants DMS-1361501 and DMS-1764231. JP: Partially supported by NSF MSPRF DMS 1606260.

References

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

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

[3] Balasubramanian, S. and McCullough, S.: Quasi-convex free polynomials. *Proc. Amer. Math. Soc.* **142** (2014), no. 8, 2581–2591.

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

[5] Ball, J. A., Marx, G. and Vinnikov, V.: Noncommutative reproducing kernel Hilbert spaces. *J. Funct. Anal.* **271** (2016), no. 7, 1844–1920.

[6] Bergman, G.: Rational relations and rational identities in division ring I, II. *J. Algebra* **43** (1976), no. 1, 252–266 and 267–297.

[7] Bhatia, R.: *Matrix analysis*. Graduate Texts in Mathematics 169, Springer-Verlag, New York, 1997.

[8] Boyd, S., El Ghaoui, L., Feron, E. and Balakrishnan, V.: *Linear matrix inequalities in system and control theory*. SIAM Studies in Applied Mathematics 15, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[9] Brešar, M. and Klep, I.: A local-global principle for linear dependence of noncommutative polynomials. *Israel J. Math.* **193** (2013), no. 1, 71–82.

[10] Camino, J. F., Helton, J. W., Skelton, R. E. and Ye, J.: Matrix inequalities: a symbolic procedure to determine convexity automatically. *Integral Equations Operator Theory* **46** (2003), no. 4, 399–454.

[11] Davidson, K. R., Dor-On, A., Shalit, O. M. and Solel, B.: Dilations, inclusions of matrix convex sets, and completely positive maps. *Int. Math. Res. Not. IMRN* (2017), no. 13, 4069–4130.

[12] Davidson, K. and Kennedy, M.: *Noncommutative Choquet theory*. Preprint 2019, arXiv: 1905.08436.

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

[14] Dym, H., Helton, J. W. and McCullough, S.: Non-commutative polynomials with convex level slices. *Indiana Univ. Math. J.* **66** (2017), no. 6, 2071–2135.

[15] Ebadian, A., Nikoufar, I. and Eshaghi Gordji, M.: Perspectives of matrix convex functions. *Proc. Natl. Acad. Sci. USA* **108** (2011), no. 18, 7313–7314.

[16] Effros, E. G.: A matrix convexity approach to some celebrated quantum inequalities. *Proc. Natl. Acad. Sci. USA* **106** (2009), no. 4, 1006–1008.

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

[18] Evert, E. and Helton, J. W.: Arveson extreme points span free spectrahedra. *Math. Ann.* **375** (2019), no. 1-2, 629–653.

[19] Goh, K.-C., Safonov, M. G. and Ly, J. H.: Robust synthesis via bilinear matrix inequalities. Linear matrix inequalities in control theory and applications. *Internat. J. Robust Nonlinear Control* **6** (1996), no. 9-10, 1079–1095.

[20] Guionnet, A. and Shlyakhtenko, D.: Free diffusions and matrix models with strictly convex interaction. *Geom. Funct. Anal.* **18** (2009), no. 6, 1875–1916.

[21] Hansen, F.: Selfadjoint means and operator monotone functions. *Math. Ann.* **256** (1981), no. 1, 29–35.

[22] Hartz, M. and Lupini, M.: The classification problem for operator algebraic varieties and their multiplier algebras. *Trans. Amer. Math. Soc.* **370** (2018), no. 3, 2161–2180.

[23] Hay, D. M., Helton, J. W., Lim, A. and McCullough, S.: Non-commutative partial matrix convexity. *Indiana Univ. Math. J.* **57** (2008), no. 6, 2815–2842.

[24] Helton, J. W., Klep, I. and McCullough, S.: The matricial relaxation of a linear matrix inequality. *Math. Program. Ser. A* **138** (2013), no. 1-2, 401–445.

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

[26] Helton, J. W., Klep, I. and McCullough, S.: The tracial Hahn–Banach theorem, polar duals, matrix convex sets, and projections of free spectrahedra. *J. Eur. Math. Soc. (JEMS)* **19** (2017), no. 6, 1845–1897.

[27] Helton, J. W. and McCullough, S.: Convex noncommutative polynomials have degree two or less. *SIAM J. Matrix Anal. Appl.* **25** (2004), no. 4, 1124–1139.

[28] Helton, J. W. and McCullough, S.: Every convex free basic semi-algebraic set has an LMI representation. *Ann. of Math. (2)* **176** (2012), no. 2, 979–1013.

[29] Helton, J. W., McCullough, S. A. and Vinnikov, V.: Noncommutative convexity arises from linear matrix inequalities. *J. Funct. Anal.* **240** (2006), no. 1, 105–191.

[30] Helton, J. W. and Merino, O.: Sufficient conditions for optimization of matrix functions. In *Proceedings of the 37th IEEE Conference on decision and control, vol. 3*, 3361–3365. IEEE Control Systems Society, 1998.

[31] Jorswieck, E. and Boche, H.: *Majorization and matrix-monotone functions in wireless communications*. Foundations and Trends® in Communications and Information Theory 3.6, Now Foundations and Trends, 2007.

[32] Jury, M., Klep, I., Mancuso, M. E., McCullough, S. and Pascoe, J. E.: Noncommutative partial convexity via Γ-convexity. *J. Geom. Anal.* **31** (2021), no. 3, 3137–3160.

[33] Kaliuzhnyi-Verbovetskyi, D. S. and Vinnikov, V.: Singularities of rational functions and minimal factorizations: the noncommutative and the commutative setting. *Linear Algebra Appl.* **430** (2009), no. 4, 869–889.

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

[35] Kanev, S., Scherer, C., Verhaegen, M. and De Schutter, B.: Robust output-feedback controller design via local BMI optimization. *Automatica J. IFAC* **40** (2004), no. 7, 1115–1127.

[36] Klep, I. and Volčič, J.: Free loci of matrix pencils and domains of noncommutative rational functions. *Comment. Math. Helv.* **92** (2017), no. 1, 105–130.

[37] Kraus, F.: Über konvexe Matrixfunktionen. *Math. Z.* **41** (1936), no. 1, 18–42.

[38] Pascoe, J. E. and Tully-Doyle, R.: The royal road to automatic noncommutative real analyticity, monotonicity, and convexity. Preprint 2019, arXiv: 1907.05875.

[39] Passer, B., Shalit, O. M. and Solel, B.: Minimal and maximal matrix convex sets. *J. Funct. Anal.* **274** (2018), no. 11, 3197–3253.

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

[41] Paulsen, V.: *Completely bounded maps and operator algebras*. Cambridge Studies in Advanced Mathematics 78, Cambridge University Press, Cambridge, 2002.

[42] Popescu, G.: Invariant subspaces and operator model theory on noncommutative varieties. *Math. Ann.* **372** (2018), no. 1-2, 611–650.

[43] Salomon, G., Shalit, O. M. and Shamovich, E.: Algebras of bounded noncommutative analytic functions on subvarieties of the noncommutative unit ball. *Trans. Amer. Math. Soc.* **370** (2018), no. 12, 8639–8690.

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

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

[46] VanAntwerp, J. G. and Braatz, R. D.: A tutorial on linear and bilinear matrix inequalities. *J. Process Control* **10** (2000), no. 4, 363–385.

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

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

Received May 21, 2020; revised July 26, 2021. Published online November 4, 2021.

Michael Jury
Department of Mathematics, University of Florida, Gainesville, FL 32601, USA;
mjury@ufl.edu

Igor Klep
Department of Mathematics, University of Ljubljana, Jadranska 21, 1000 Ljubljana, Slovenia;
igor.klep@fmf.uni-lj.si

Mark E. Mancuso
Department of Mathematics, Lafayette College, 730 High St., Easton, PA 18042, USA;
mancusom@lafayette.edu

Scott McCullough
Department of Mathematics, University of Florida, Gainesville, FL 32601, USA;
sam@math.ufl.edu

James Eldred Pascoe
Department of Mathematics, University of Florida, Gainesville, FL 32601, USA;
pascoej@ufl.edu