Integr. equ. oper. theory 46 (2003) 399–454
0378-620X/040399-56, DOI 10.1007/s00020-001-1147-7

Integral Equations and Operator Theory

Matrix Inequalities: A Symbolic Procedure to Determine Convexity Automatically

Juan F. Camino, J. W. Helton, Robert E. Skelton, and Jieping Ye

Abstract. This paper presents a theory of noncommutative functions which results in an algorithm for determining where they are "matrix convex". Of independent interest is a theory of noncommutative quadratic functions and the resulting algorithm which calculates the region where they are "matrix positive". This is accomplished via a theorem (a type of *Positivstellensatz*) on writing noncommutative quadratic functions with noncommutative rational coefficients as a weighted sum of squares. Furthermore the paper gives an LDU algorithm for matrices with noncommutative entries and conditions guaranteeing that the decomposition is successful.

The motivation for the paper comes from systems engineering. Inequalities, involving polynomials in matrices and their inverses, and associated optimization problems have become very important in engineering. When these polynomials are "matrix convex" interior point methods apply directly. A difficulty is that often an engineering problem presents a matrix polynomial whose convexity takes considerable skill, time, and luck to determine. Typically this is done by looking at a formula and recognizing "complicated patterns involving Schur complements"; a tricky hit or miss procedure. Certainly computer assistance in determining convexity would be valuable. This paper, in addition to theory, describes a symbolic algorithm and software which represent a beginning along these lines.

The algorithms described here have been implemented under Mathematica and the noncommutative algebra package NCAlgebra. Examples presented in this article illustrate its use.

Mathematics Subject Classification (2000). Primary 47L99, 93A99; Secondary 15A39, 14P99.

Keywords. Noncommutative algebra, Matrix inequalities, Convexity, Computer algebra.

First author was partly supported by "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES" – Brazil. The second author was partly supported by NSF, the ONR and the Ford Motor Co. The third author was partly supported by DARPA and by AFOSR grant number: F49620–99–1–0331.

1. Introduction

In the last few years, the approaches that have been proposed in the field of optimization and control theory based on linear matrix inequalities and semidefinite programming have become very important and promising, since the same framework can be used for a large set of problems ([SIG97, BEGFB94, EGN99, Roc97, SI95]). Matrix inequalities provide a nice set up for many engineering and related problems, and if they are convex the optimization problem is well behaved and interior point methods provide efficient algorithms ([NN94, AHO98, VB96]) which are effective on moderate sized problems. In practice, many of the problems in engineering and optimization present matrix valued functions that take a large effort to determine their convexity.

1.1. The Idea Behind Our Algorithm

This paper provides a computer algebra algorithm that can be used to find the domain $\mathcal{G}$ of convexity of a noncommutative rational function Γ. This algorithm produces sufficient, and with some weak hypotheses, necessary conditions for convexity.

We now very loosely introduce the idea behind the algorithm even though we have not set down any formal definitions. Let Γ be the noncommutative rational function to be analyzed. Say Γ is a function of the noncommutative variables, $A_1, \dots, A_m, X_1, \dots, X_k$. The main steps of the algorithm are:

1. The second directional derivative with respect to $X_1, \dots, X_k$, the Hessian $\mathcal{H}\Gamma$ of the function Γ, is computed.
2. As the Hessian is always a quadratic function of the update directions, it can be associated with a symmetric matrix $M_{\mathcal{H}\Gamma}$ with noncommutative entries.
3. The noncommutative $\mathrm{LDL}^{\mathrm{T}}$ factorization is applied to the coefficient matrix $M_{\mathcal{H}\Gamma}$.
4. And finally specifying positivity of the resulting diagonal[1] matrix $D(A_1, \dots, A_m, X_1, \dots, X_k)$ gives inequalities describing a region $\mathcal{G}$ of variables on which Γ is matrix convex.

While all this is extremely straightforward in the case where Γ is a conventional commutative function, non-commutativity imposes rather interesting complications. In particular proving that the largest "symbolic inequality domain" on which Γ is "matrix convex" requires a substantial proof, mixing both linear algebra and algebraic representation type arguments.

The implementation of the algorithm has been done completely in NCAlgebra, a noncommutative algebra package that runs under Mathematica. This

[1]If D is not diagonal, it contains 2×2 blocks which are never positive definite. See Section 4.

package with our additions provides a large number of useful commands and functions for symbolic computation. It can be downloaded from

http://math.ucsd.edu/∼ncalg.

As an example, the first and second directional derivatives of noncommutative rational functions, and a noncommutative LDU decomposition are easily computed with NCAlgebra.

1.2. The Theory: Positivity vs Weighted Sums of Squares

To give an idea of what type of mathematics is involved, in proving that our algorithm gives the "largest" domain of convexity, we recall the classic Hilbert 17th problem. The problem is to represent a positive commutative polynomial as a sum of squares, whenever possible. It is shown in [Hel] that noncommutative polynomials which are "matrix positive" are always sums of squares; thus the noncommutative situation behaves more cleanly than the commutative situation. In this paper we shall be interested in noncommutative rational functions $\mathcal{Q}(Z_1, \ldots, Z_v, H_1, \ldots, H_k)$ which are quadratic functions of $\vec{H} = \{H_1, \ldots, H_k\}$. Our concern is describing the region $\mathcal{G}$ of $\vec{Z}= \{Z_1, \ldots, Z_v\}$ on which $\mathcal{Q}$ is "matrix positive" in $\vec{H}$. What we show under reasonable hypotheses is that $\mathcal{Q}$ *has a weighted sum of squares decomposition*

$$\mathcal{Q}(\vec{Z}, \vec{H}) := \sum_{j=1}^{r} \Phi_j(\vec{Z}, \vec{H})^T D_j(\vec{Z}) \Phi_j(\vec{Z}, \vec{H})$$

with $\Phi_j(\vec{Z}, \vec{H})$, $D_j(\vec{Z})$ *rational and* $\Phi_j(\vec{Z}, \vec{H})$ *linear in* $\vec{H}$, *such that formal inequalities involving the* $D_j(\vec{Z})$ *determine a set*

$$\mathcal{G} := \{\vec{Z}\colon D_j(\vec{Z}) > 0, \ j = 1, \ldots, r\}$$

on which $\mathcal{Q}$ *is "matrix positive" in* $\vec{H}$. *Moreover, a certain "closure" of* $\mathcal{G}$ *is the largest such set.* The precise statement of this result is Theorem 8.2 and a weaker more accessible result is Theorem 3.1.

1.3. The Noncommutative LDU Decomposition

In this paper we describe and then analyze the LDU decomposition for general matrices with noncommutative entries. LDU decompositions of small size or for special matrices with noncommutative entries exist scattered through the literature. In our implementation, invertibility of "pivots" is a major issue. Fortunately, one finds that after the LDU algorithm is applied to a matrix M, producing $M = LDU$, the "pivots" appear in the diagonal matrix $D = \text{diag}\{D_1, \ldots, D_r\}$. Our main LDU theorem, Theorem 3.3, says that *if one is willing to assume that the expressions* D_j *are invertible, then the LDU decomposition produced by our algorithm is valid.*

1.4. An Example

We introduce our method for finding the region on which a noncommutative function is convex or concave, with an example of an NCAlgebra command (which embodies it). While this is a bit unusual, since it uses terms which have not been formally introduced, we find most people understand the example anyway and the example eases the many pages of definitions and constructions that the reader must endure before getting to the rewards of the method. The command for finding the region of convexity is

$$\textbf{NCConvexityRegion}[\text{Function } \Gamma,\ \vec{X}\ ].$$

When we input a noncommutative rational function $\Gamma(\vec{Z})$ of $\vec{Z} = \{A_1, \ldots, A_m, X_1, \ldots, X_k\}$ this command outputs a family of inequalities which determine a domain $\mathcal{G}$ of $\vec{Z}$ on which Γ is "matrix convex" in $\vec{X} = \{X_1, \ldots, X_k\}$. This is illustrated by the next two examples.

Example 1.1. Suppose ones wish to determine the domain of convexity (concavity) with respect to X, Y of the following function on matrices $\vec{Z} = \{A, B, R, X, Y\}$:

$$F(\vec{Z}) = -(Y + A^T XB)(R + B^T XB)^{-1}(Y + B^T XA) + A^T XA,$$

where $R = R^T$, $X = X^T$ and $Y = Y^T$. We treat A, B, R, X, Y symbolically as noncommutative indeterminates and apply the command **NCConvexityRegion**[F, $\{X,\ Y\}$] which outputs the list

$$\{-2\,(R + B^T XB)^{-1},\ 0,\ 0,\ 0\}.$$

From this output, we conclude that whenever A, B, R, X, and Y are matrices of compatible dimension, the function F is "matrix concave" in X, Y on the domain $\mathcal{G}$ given by

$$\mathcal{G} := \{(X, Y) : (R + B^T XB)^{-1} > 0\}.$$

The command NCConvexityRegion also has an important feature which for this problem assures us that the "closure" of $\mathcal{G}$, in a certain sense, is the "biggest domain of matrix concavity" for F. □

Example 1.2. Let $X = X^T$ and $Y = Y^T$, and define the function F as

$$F(X, Y) = (X - Y^{-1})^{-1}.$$

The output of **NCConvexityRegion**[F, $\{X,\ Y\}$] is

$$\{(X - Y^{-1})^{-1},\ Y^{-1},\ 0\}.$$

Thus the function F is "matrix convex" on the region

$$\mathcal{G} := \{(X, Y) : Y^{-1} > 0 \text{ and } (X - Y^{-1})^{-1} > 0\},$$

whenever the symbolic elements X and Y are substituted by any matrices of compatible dimension. Of course $\mathcal{G}$ is the same as $\{(X, Y) : Y > 0 \text{ and } X - Y^{-1} > 0\}$. Also in this example our algorithm guarantees that the "closure" of $\mathcal{G}$ in a certain sense is the "biggest domain of matrix convexity" of F. □

1.5. Notation

In order to make an expression symmetric the operator sym, defined as $sym[M] = M + M^T$, is used. The operator $(\cdot)^{-1}$ and $(\cdot)^T$ means the inverse and the transpose respectively. The arrow over a variable is used to indicate that the variable is a list of elements $\overrightarrow{X} = \{X_1, \ldots, X_k\}$. If $\overrightarrow{X}$ contains only one indeterminate, then the notation is $\overrightarrow{X} = X$. Roman upper case letters will commonly represent symbolic elements, and also matrices when it is clear by context. Euler-Script letters are frequently used to indicate the substitution of noncommutative elements by matrices of compatible dimensions. As an example, $\Gamma(X)$ means a noncommutative rational function whose argument X is a symbolic element; on the other hand, the Euler-Script $\mathcal{X}$ is used in $\Gamma(\mathcal{X})$ when $\mathcal{X}$ is a matrix in $\mathbb{R}^{n\times m}$. Another example appears in the definition of the set $\mathcal{R}_L^x := \{(\mathcal{H}\, L\, x) : \text{ all } \mathcal{H} \in \mathbb{R}^{n\times m}\}$ where L is a noncommutative rational function evaluated on certain matrices, $\mathcal{H}$ is a matrix, and x is a vector. Note that we do not use the Euler-Script font for vectors and functions. Even if the argument of the function L is a matrix $\mathcal{Z}$ rather than an indeterminate Z, we would have used $L(\mathcal{Z})$ instead of $\mathcal{L}(\mathcal{Z})$, and often we abbreviate $L(\mathcal{Z})$ to L.

1.6. The Layout of the Paper

This paper is split into two parts. Part I of this paper presents our algorithm, describes its implementation and illustrates its effect on a few examples. We prove in Part II that the region $\mathcal{G}$ of convexity which our algorithm determines is the largest possible in a certain sense. The results in Part II give a satisfying theory of "matrix convexity" and of "matrix positivity" of noncommutative quadratic functions of a certain type. The paper contains a bit of redundancy in order to maximize the reader base. Throughout the presentation of our algorithm we insert actual calls to symbolic routines in NCAlgebra, since this makes clear exactly what can be computed automatically.

Part I should be accessible to readers from many areas, from operator or matrix theory, from symbolic computation, and from engineering who work with matrix inequalities. It is organized as follows. Section 2 gives preliminary definitions about noncommutative rational functions, convexity, positivity, and derivatives. Section 3 concerns quadratic noncommutative functions $\mathcal{Q}$. It gives a representation for $\mathcal{Q}$ in terms of a symmetric matrix $M_{\mathcal{Q}}$ with noncommutative entries and it provides an algorithm to compute the LDU decomposition of $M_{\mathcal{Q}}$. Section 4 gives the convexity algorithm that provides the tools for checking the positivity and presents some examples. Section 5 illustrates how the algorithm when implemented using the noncommutative algebra package NCAlgebra can be used to find the region of convexity of a noncommutative rational function.

Now we describe the organization of Part II. Section 6 formally states and proves a theorem to the effect that our Convexity Algorithm produces a domain $\mathcal{G}$ in which the given function is convex. This is easy and informative. Section 7 gives formal definitions. Section 8 states theorems to the effect that $\mathcal{G}$ is the biggest domain of convexity in a certain sense. Sections 9 and 10 gives proofs of the theorems stated in Section 8 and Section 3.

Section A is an appendix which describes a computer algorithm for representing a noncommutative quadratic function of k variables $H_1, \ldots, H_k$ in terms of a matrix M_Q. This matrix plays an important role in determining the positiveness of a noncommutative rational function.

A brief announcement of the methods in this papers was presented in [CHS00].

Part I. The Algorithm: Its Implementation and Use

We begin with definitions of noncommutative rational functions, of derivatives of noncommutative functions, and of convexity. Next, the procedure to represent a quadratic function together with the noncommutative LDU decomposition is illustrated. Also the idea behind necessary and sufficient conditions for positivity of noncommutative quadratic functions is introduced. Later in Section 4, our Convexity Algorithm is described and then, in Section 5, it is illustrated by some examples.

2. Noncommutative Rational Functions

In this section we present useful definitions and facts about noncommutative rational functions. In fact, the development in this section follows [HM97] and [HM98].

2.1. Noncommutative Symmetric Rational Functions

What occurs in practice are functions Γ which are polynomial or rational in noncommutative variables (often referred to as indeterminates) with coefficient which are real numbers. Noncommutative rational functions of X are polynomials in X and in inverses of polynomials in X. Examples of noncommutative symmetric functions are

$$\Gamma(A, B, X) = AX + XA^T - \frac{3}{4}XBB^TX, \ X = X^T,$$

$$\Gamma(A, D, X, Y) = X^TAX + DYD^T + XYX^T, \ Y = Y^T \text{ and } A = A^T, \tag{2.1}$$

and

$$\Gamma(A, D, E, X, Y) = A(I + DXD^T)^{-1}A^T + E(YXY^T)E^T, \ X = X^T. \tag{2.2}$$

We also assume there is an involution on these rational functions which we denote superscript T, and which will play the role of transpose later when we substitute matrices for the indeterminates.

Often we shall think of some indeterminates as knowns and other indeterminates as unknowns and be concerned primarily about a function's properties with respect to unknowns. For example, in function (2.2) when we are mainly concerned about behavior such as convexity of Γ in X, Y we write $\Gamma(A, D, E, X, Y)$

simply as $\Gamma(X,Y)$. We also use $\vec{Z}$ to abbreviate all indeterminates which appear in the function, for example, in (2.2) we have $\vec{Z} = \{A, D, E, X, Y\}$. Often we distinguish knowns $\vec{A} = \{A_1, \ldots, A_m\}$ from unknowns $\vec{X} = \{X_1, \ldots, X_k\}$ by writing $\vec{Z} = \{\vec{A}, \vec{X}\}$. Throughout this paper, letters near the beginning of the alphabet denote knowns, while the letters X, Y stand for unknowns.

We call a noncommutative function $\Gamma(\vec{A}, \vec{X})$ **symmetric** provided that $\Gamma(\vec{A}, \vec{X})^T = \Gamma(\vec{A}, \vec{X})$. If all $X_1^T, X_2^T, \ldots, X_k^T$ in $\Gamma(\vec{A}, \vec{X})$ appear to the left of every $X_1, X_2, \ldots X_k$ variable, then the noncommutative function $\Gamma(\vec{A}, \vec{X})$ is said to be **hereditary**[2] in $\vec{X}$. Our algorithm when restricted to hereditary noncommutative functions is easier to describe and the theory is easier.

2.2. First Derivatives

Conventional convexity of a function can be characterized by the second derivative being positive. As we shall see in Section 2.4, this is also the case with "noncommutative convex functions" and so we review a notion of second derivative which is suitable for symbolic computation. We begin with first derivatives rather than second derivatives. Later we study convexity tests which are based on derivatives of Γ and their transposes.

Directional derivatives of noncommutative rational $\Gamma(\vec{A}, \vec{X})$ with respect to $\vec{X}$ in the direction $\vec{H}$ are defined in the usual way

$$D\Gamma(\vec{X})[\vec{H}] \quad := \quad \lim_{t\to 0} \frac{1}{t}\left(\Gamma(\vec{X}+t\vec{H}) - \Gamma(\vec{X})\right) = \frac{d}{dt}\,\Gamma(\vec{X}+t\vec{H})\Big|_{t=0}.$$

For example, the derivative of Γ in (2.1) with respect to X is

$$D_X\Gamma(X,Y)[H] = H^TAX + X^TAH + HYX^T + XYH^T.$$

and the derivative of Γ in (2.2) with respect to Y is

$$D_Y\Gamma(X,Y)[K] = E(KXY^T + YXK^T)E^T.$$

It is easy to check that derivatives of symmetric noncommutative rational functions always have the form

$$D\Gamma(X)[H] = sym\left[\sum_{\ell=1}^{k} A_\ell H B_\ell\right].$$

The noncommutative algebra command to generate the directional derivative of $\Gamma(X,Y)$ with respect to X, which is denoted by $D_X\Gamma(X,Y)[H]$, is:

NCAlgebra Command: DirectionalD[Function Γ, X, H].

[2]Note that in our definition of hereditary the variables X_j can not be constrained to be symmetric.

2.3. Second Derivatives

To obtain sufficient conditions for optimization we must use the second order terms of a Taylor expansion of $\Gamma(\vec{X} + t\vec{H})$ about $t = 0 \in \mathbb{R}$:

$$\Gamma(\vec{X} + t\vec{H}) = \Gamma(\vec{X}) + D\Gamma(\vec{X})[\vec{H}]\, t + \vec{\mathcal{H}}\Gamma(\vec{X})[\vec{H}]\, t^2 + \ldots$$

Where the Hessian $\mathcal{H}\Gamma$ of Γ is defined by

$$\mathcal{H}\Gamma(\vec{X})[\vec{H}] := \frac{d^2}{dt^2}\Gamma(\vec{X} + t\vec{H})\Big|_{t=0}.$$

One can easily show that the second derivative of a hereditary symmetric noncommutative rational function Γ with respect to one variable X has the form

$$\mathcal{H}\Gamma(X)[H] = sym\left[\sum_{\ell=1}^{k} A_\ell H^T B_\ell H C_\ell\right].$$

And an analogous more general expression holds for more variables. For example, the second derivative of Γ in (2.2) with respect to X is

$$\mathcal{H}_X\Gamma(X,Y)[H] = \\ 2\,(A(I + DXD^T)^{-1}DHD^T(I + DXD^T)^{-1}DHD^T(I + DXD^T)^{-1}A^T).$$

Once the Hessian $\mathcal{H}\Gamma(\vec{X})[\vec{H}]$ is computed, the only variable of interest is $\vec{H}$. Thus, for convenience, the variables $\vec{X}$ and $\vec{A}$ are gathered in $\vec{Z}$, producing a function $\mathcal{Q}$,

$$\mathcal{Q}(\vec{Z})[\vec{H}] := \mathcal{H}\Gamma(\vec{X})[\vec{H}],$$

which is quadratic in $\vec{H}$. Here of course, a noncommutative polynomial in variables $H_1, H_2, \ldots, H_k$ is said to be **quadratic** if each monomial in the polynomial expression is of order two in the variables $H_1, H_2, \ldots, H_k$.

We emphasize that for our convexity considerations once the Hessian is computed the fact that $\vec{X}$ played a special role has no influence.

NCAlgebra Command: Hessian[function Γ, $\{X_1, H_1\}, \ldots, \{X_k, H_k\}$].

2.4. Matrix Convex Functions

There are several (almost equivalent) notions of noncommutative convexity, and hence we describe two familiar matrix versions. We begin by defining *matrix convex functions* as it is the definition used throughout the paper, and later we define *geometrically matrix convex functions* as it is a common definition for convexity although we do not use it.

We shall be focusing on symmetric noncommutative functions Γ of $\vec{Z}$ defined on a domain $\mathcal{G}$ given by "inequalities" on symmetric noncommutative rational functions ρ_j, $j = 1, \ldots, r$. The tuple $\vec{Z}$ denotes all noncommutative variables

$A, B, C, X, \dots$ which appear in Γ. (Frequently we just denote $\vec{\mathcal{Z}} = \{\mathcal{Z}_1, \dots, \mathcal{Z}_v\}$). We write the formal expression

$$\mathcal{G}_\rho := \{\vec{Z} = \{Z_1, \dots, Z_v\} : \rho_j(\vec{Z}) \geq 0, j = 1, \dots, r\}$$

and call such an expression a **Symbolic Inequality Domain** – **SID**. An example is

$$\mathcal{G} := \{\vec{Z} = \{A, C, X\} : -A^T X - XA - C^T C \geq 0, X \geq 0\}$$

and Example 1.1 and 1.2 of the introduction.

Note that the $\vec{Z}$ are just formal symbols. Since our ultimate interest is matrices we introduce $\mathcal{M}(\mathcal{G}_\rho)$ the set of all matrix tuple $\vec{\mathcal{Z}} = \{\mathcal{Z}_1, \dots, \mathcal{Z}_v\}$ which satisfy

$\rho_j(\vec{\mathcal{Z}})$ *is a positive semidefinite matrix for all* $j = 1, \dots, r$.

Denote by $\mathcal{M}_\Delta$ all tuple of matrices $\vec{\mathcal{Z}}$ of size Δ. Denote by $\mathcal{M}_\Delta(\mathcal{G})$ the set of all matrices of size Δ which are in $\mathcal{M}(\mathcal{G})$, that is, $\mathcal{M}_\Delta(\mathcal{G}) = \mathcal{M}_\Delta \bigcap \mathcal{M}(\mathcal{G})$. See section 7.2 for a more complete statement.

Our main definitions of positivity are:

1. A noncommutative rational function $Q(\vec{Z})[\vec{H}]$ which is quadratic in $\vec{H}$ is said to be **matrix positive quadratic** (resp. **matrix strictly positive quadratic**) on a SID $\mathcal{G}_\rho$ provided that $Q(\vec{\mathcal{Z}})[\vec{\mathcal{H}}]$ is a positive semidefinite matrix (resp. positive definite matrix) whenever tuple of matrices $\vec{\mathcal{Z}}$ in $\mathcal{M}(\mathcal{G}_\rho)$ and $\vec{\mathcal{H}}$ are substituted for $\vec{Z}$ and $\vec{H}$.
2. The function $\Gamma(\vec{A}, \vec{X})$ is said to be **matrix convex** with respect to variable $\vec{X}$ on a SID $\mathcal{G}_\rho$ provided its Hessian $\mathcal{H}\Gamma(\vec{\mathcal{X}})[\vec{\mathcal{H}}]$ is a positive semidefinite matrix for all $\vec{\mathcal{A}}, \vec{\mathcal{X}}$ in $\mathcal{M}(\mathcal{G}_\rho)$ and all $\vec{\mathcal{H}}$; in other words, when its Hessian is matrix quadratic.

One Symbolic Inequality Domain $\mathcal{G}_\rho$ **contains another** $\mathcal{G}_{\tilde{\rho}}$, means that whenever tuple of matrices $\vec{\mathcal{Z}}$ of compatible dimension satisfy the inequalities $\tilde{\rho}_j(\vec{\mathcal{Z}}) \geq 0$, for $j = 1, \dots, \tilde{r}$, then they also satisfy the inequalities $\rho_j(\vec{\mathcal{Z}}) \geq 0$, for $j = 1, \dots, r$. In this case we say that

the inequalities $\rho(\vec{\mathcal{Z}}) \geq 0$ are weaker than the inequalities $\tilde{\rho}(\vec{\mathcal{Z}}) \geq 0$.

This condition is the same as $\mathcal{M}(\mathcal{G}_{\tilde{\rho}}) \subseteq \mathcal{M}(\mathcal{G}_\rho)$.

While this looks awkward and elaborate, it is in fact the type of "matrix convexity" which fits reasonably into symbolic processing of the type of matrix inequalities which engineers use. We present a few examples in Section 5 which make this definition clear and natural. Also matrix convexity is strongly connected with usual notions of geometric convexity, as we now discuss.

A noncommutative rational symmetric function Γ of $\vec{X} = \{X_1, \dots, X_k\}$ will be called **geometrically matrix convex** provided that whenever the noncommutative variables $\vec{X}$ are taken to be any matrices of compatible dimension, then for all scalars $0 \leq \alpha \leq 1$ we have that

$$\Gamma(\alpha \vec{\mathcal{X}}^1 + (1-\alpha)\vec{\mathcal{X}}^2) \leq \alpha\Gamma(\vec{\mathcal{X}}^1) + (1-\alpha)\Gamma(\vec{\mathcal{X}}^2).$$

Where $\vec{\mathcal{X}}^1 = \{\mathcal{X}_1^1, \dots, \mathcal{X}_k^1\}$ and $\vec{\mathcal{X}}^2 = \{\mathcal{X}_1^2, \dots, \mathcal{X}_k^2\}$ are tuples of matrices of compatible dimension. The function Γ is **strictly geometrically matrix convex** if the inequality is strict for $0 < \alpha < 1$. The reverse inequality characterizes **geometrically matrix concave**.

Both the definitions, matrix convex and geometrically matrix convex, are equivalent provided that the domain of the function Γ is a convex set; as stated by the following lemma.

Lemma 2.1. *Suppose Γ is a noncommutative rational symmetric function. Then it is geometrically matrix convex (respectively geometrically matrix concave) on a convex region Ω of matrices of fixed sizes if and only if*

$$\mathcal{H}\Gamma(\vec{\mathcal{X}})[\vec{\mathcal{H}}] \geq 0$$

(respectively ≤ 0) for all $\vec{\mathcal{H}}$ and $\vec{\mathcal{X}} \in \Omega$.

Proof. The proof is given in [HM98] where Ω is all matrices of a given size. It extends in a straight forward way to Ω which are convex sets. □

3. Noncommutative Quadratic Functions

An example of a simple quadratic function in $H = H^T$ and $K = K^T$, where the arguments appear outside the expression, is

$$\mathcal{Q}[H, K] := HAH + KBK + HCK + KC^T H.$$

Or yet, a more complicated function, in the sense that the argument H appears inside the monomial is

$$\mathcal{Q}[H] := HAH + G^T HBH + HB^T HG + G^T HDHG.$$

This function can be written in the form

$$\mathcal{Q}[H] = \begin{pmatrix} H & G^T H \end{pmatrix} \begin{pmatrix} A & B^T \\ B & D \end{pmatrix} \begin{pmatrix} H \\ HG \end{pmatrix}. \tag{3.1}$$

This contrasts with the commutative case where (3.1) takes the form

$$Q[H] = H(A + G^T B + B^T G + G^T DG)H.$$

3.1. Representing a Quadratic Function as a Matrix $\mathbf{M}_{\mathcal{Q}}$

As suggested by (3.1), a noncommutative quadratic function $\mathcal{Q}$ which is hereditary in $\overrightarrow{H} = \{H_1, \ldots, H_k\}$ can be always represented as a product of the form $V[\overrightarrow{H}]^T M_{\mathcal{Q}} V[\overrightarrow{H}]$, where $V[\overrightarrow{H}]$ is a "vector" with noncommutative entries and $M_{\mathcal{Q}}$ is a symmetric matrix with noncommutative entries. The "vector" $V[\overrightarrow{H}]$ is called a **border vector of the quadratic function** $\mathcal{Q}$ and the matrix $M_{\mathcal{Q}}$ is the **coefficient matrix of the quadratic function** $\mathcal{Q}$.

The representation $V^T M_{\mathcal{Q}} V$ for a general hereditary quadratic polynomial in $\overrightarrow{H} = \{H, K\}$ is given by $\mathcal{Q}[H, K] :=$

$$\begin{pmatrix} HL_1^1 \\ \vdots \\ HL_{\ell_1}^1 \\ KL_1^2 \\ \vdots \\ KL_{\ell_2}^2 \end{pmatrix}^T \begin{pmatrix} A_{1,1} & \cdots & A_{1,\ell_1} & A_{1,\ell_1+1} & \cdots & A_{1,r} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{1,\ell_1}^T & \cdots & A_{\ell_1,\ell_1} & A_{\ell_1,\ell_1+1} & \cdots & A_{\ell_1,r} \\ A_{1,\ell_1+1}^T & \cdots & A_{\ell_1,\ell_1+1}^T & A_{\ell_1+1,\ell_1+1} & \cdots & A_{\ell_1+1,r} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{1,r}^T & \cdots & A_{\ell_1,r}^T & A_{\ell_1+1,r}^T & \cdots & A_{r,r} \end{pmatrix} \begin{pmatrix} HL_1^1 \\ \vdots \\ HL_{\ell_1}^1 \\ KL_1^2 \\ \vdots \\ KL_{\ell_2}^2 \end{pmatrix}$$

where $r = \ell_1 + \ell_2$. The quantity ℓ_1 is the number of times that a monomial of order two in H appears, and the quantity ℓ_2 is the number of times that a monomial of order two in K appears. The L_j^i, $j = 1, \ldots, \ell_i$ are called the **coefficients of the border vector**. The L_j^1 corresponding to H are distinct and only one may be the identity matrix (equivalently for the L_j^2 corresponding to K). The border vector V is the vector composed of H, K and L_j^i. The coefficient matrix $M_{\mathcal{Q}}$ is the one in the middle with entries $A_{s,t}$, for $s, t = 1, \ldots, r$. See appendix A for an algorithm which compute this decomposition. This general notation illustrated by the example in equation (3.1) is:

$$V[H]^T = \begin{pmatrix} H & G^T H \end{pmatrix} \qquad \text{and} \qquad M_{\mathcal{Q}} = \begin{pmatrix} A & B^T \\ B & D \end{pmatrix}.$$

Noncommutative quadratics even though *not hereditary* have a similar representation (which takes much more space to write) for such a quadratic in H, K. For example, the border vector for a quadratic in H, H^T, K, K^T has the form

$$V[H,K]^T = \Big((L_1^1)^T H^T, \cdots, (L_{\ell_1}^1)^T H^T, (L_1^2)^T K^T, \cdots, (L_{\ell_2}^2)^T K^T, (\tilde{L}_1^1)^T H, \cdots, \\ (\tilde{L}_{\tilde{\ell}_1}^1)^T H, (\tilde{L}_1^2)^T K, \cdots, (\tilde{L}_{\tilde{\ell}_2}^2)^T K \Big).$$

As we shall see from the Example 5.3 in Section 5 the $M_{\mathcal{Q}}$ representation for a quadratic $\mathcal{Q}$ may not be unique. However, this non-uniqueness turns out to produce surprisingly few problems.

We should emphasize that the size of the $M_{\mathcal{Q}}$ representation of a noncommutative quadratic functions $\mathcal{Q}[H_1, \ldots, H_k]$ depends on the particular quadratic and not only on the number of arguments k of the quadratic. For example, there are noncommutative quadratic functions in one variable which have a representation with $M_{\mathcal{Q}}$ a 102×102 matrix.

NCAlgebra Command: NCMatrixOfQuadratic[$\mathcal{Q}$, $\{H_1, \ldots, H_k\}$] generates the list {left border vector, coefficient matrix, right border vector}.

3.2. Positivity of Noncommutative Quadratic Functions

Determining positiveness of the Hessian, which is a quadratic function in $\vec{H}$, is the key to determining the convexity of a rational function of matrices. A critical issue is relating $Q[\vec{\mathcal{H}}]$ being a positive semidefinite matrix for all $\vec{\mathcal{H}}$ to the matrix $M_{\mathcal{Q}}$ being positive semidefinite. In this section we roughly summarize our main result which surprisingly says that under weak hypotheses these two properties are very close to being equivalent. Later, Theorem 3.3 gives a definitive test for the positivity of $M_{\mathcal{Q}}$.

Theorem 3.1 (Positivity: $\mathcal{Q}$ *versus* $M_{\mathcal{Q}}$)**.** *Suppose that the noncommutative rational function $\mathcal{Q}(\vec{Z})[\vec{H}]$ is quadratic in $\vec{H}$. Represent $\mathcal{Q}(\vec{Z})$ with coefficient matrix $M_{\mathcal{Q}(\vec{Z})}$ and border vector $V[\vec{H}]$, that is $\mathcal{Q}(\vec{Z})[\vec{H}] = V[\vec{H}]^T M_{\mathcal{Q}(\vec{Z})} V[\vec{H}]$. Let $\mathcal{G}$ denote the Symbolic Inequality Domain, based on $M_{\mathcal{Q}(\vec{Z})}$, given by*

$$\mathcal{G} := \left\{ \vec{Z} : M_{\mathcal{Q}(\vec{Z})} \geq 0 \right\}.$$

Then $\mathcal{Q}(\vec{Z})[\vec{H}]$ is a matrix positive quadratic function for each $\vec{Z} \in \mathcal{G}$. Conversely, assume:

- i. *the $M_{\mathcal{Q}}$ representation of $\mathcal{Q}$ has a border vector $V[\vec{H}]$ with coefficients $L_1^j(\vec{Z}), \ldots, L_{l_j}^j(\vec{Z})$ for H_j which for each j are linearly independent functions of $\vec{Z}$;*
- ii. *the Symbolic Inequality Domain $\mathcal{G}$ is not thin in the sense that the set $\mathcal{M}_\Delta(\mathcal{G})$ is an open set in $\mathcal{M}_\Delta$, provided that the size Δ is large enough (see the Openness Property in Section 7.2).*

Then the closure of $\mathcal{G}$ in a certain topology is the biggest domain on which $\mathcal{Q}(\vec{Z})[\vec{H}]$ is a matrix positive quadratic function.

Proof. The sufficient side, the symmetric matrix $M_{\mathcal{Q}}$ being positive semidefinite guarantees that the matrix $\mathcal{Q}[\vec{\mathcal{H}}]$ is also positive semidefinite for all tuple of matrices $\vec{\mathcal{H}}$, is trivially proved. To see this, write the quadratic function as $\mathcal{Q}[\mathcal{H}_1, \ldots, \mathcal{H}_k] := V[\mathcal{H}_1, \ldots, \mathcal{H}_k]^T M_{\mathcal{Q}} V[\mathcal{H}_1, \ldots, \mathcal{H}_k]$. Now, let $M_{\mathcal{Q}} \in \mathbb{R}^{r \times r}$ be positive semidefinite. By definition this implies that $x^T M_{\mathcal{Q}} x \geq 0$ for all vectors $x \in \mathbb{R}^r$. So, for any $y \in \mathbb{R}^m$, choose x to be $x = V[\mathcal{H}_1, \ldots, \mathcal{H}_k]y$. Then $x^T M_{\mathcal{Q}} x = y^T V[\mathcal{H}_1, \ldots, \mathcal{H}_k]^T \; M_{\mathcal{Q}} \; V[\mathcal{H}_1, \ldots, \mathcal{H}_k]y = y^T \mathcal{Q}[\mathcal{H}_1, \ldots, \mathcal{H}_k]y \geq 0$.

The necessity side requires involved proof which takes up Part II of this paper. We shall illustrate one of its steps in the simple Example 3.1 below. □

Note that linear dependence of a small set of matrices in a high dimensional space is a rare event. This intuitively speaking is the type of linear dependence in assumption (i) of Theorem 3.1 required to violate the necessity of $M_{\mathcal{Q}}$ being positive. Indeed, this type of linear dependence has never occurred in any experiments we have done, although one could probably make up examples where it occurs.

Even though a quadratic $\mathcal{Q}$ can have two representations $M^1_{\mathcal{Q}}$ and $M^2_{\mathcal{Q}}$ meeting the hypotheses in Theorem 3.1, our result implies that $M^1_{\mathcal{Q}}$ will be positive semidefinite if and only if $M^2_{\mathcal{Q}}$ is also positive semidefinite.

Example 3.1. Consider the noncommutative quadratic function $\mathcal{Q}[H]$ given by

$$\mathcal{Q}[H] := H^T BH + G^T H^T CH + H^T C^T HG + G^T H^T AHG. \tag{3.2}$$

Here, in distinction to most of Part I, we are not forcing H to be symmetric. This is much easier to analyze than the case where H is symmetric. The border vector $V[H]$ and the coefficient matrix $M_{\mathcal{Q}}$ with noncommutative entries are

$$V[H]^T = \begin{pmatrix} H^T & G^T H^T \end{pmatrix} \qquad \text{and} \qquad M_{\mathcal{Q}} = \begin{pmatrix} B & C^T \\ C & A \end{pmatrix},$$

that is, $\mathcal{Q}[H]$ has the form

$$\mathcal{Q}[H] = V[H]^T M_{\mathcal{Q}} V[H] = \begin{pmatrix} H^T & G^T H^T \end{pmatrix} \begin{pmatrix} B & C^T \\ C & A \end{pmatrix} \begin{pmatrix} H \\ HG \end{pmatrix}.$$

Now, if in equation (3.2) the elements A, B, C, G, H are replaced by matrices in $\mathbb{R}^{n\times n}$, then the noncommutative quadratic function $\mathcal{Q}[H]$ becomes a matrix valued function $\mathcal{Q}[\mathcal{H}]$. The matrix valued function $\mathcal{Q}[\mathcal{H}]$ is positive semidefinite if and only if $x^T \mathcal{Q}[\mathcal{H}]x \geq 0$ for all vectors $x \in \mathbb{R}^n$ and all $\mathcal{H} \in \mathbb{R}^{n\times n}$. Or equivalently, the following inequality must hold

$$\begin{pmatrix} x^T \mathcal{H}^T & x^T \mathcal{G}^T \mathcal{H}^T \end{pmatrix} M_{\mathcal{Q}} \begin{pmatrix} \mathcal{H}x \\ \mathcal{H}\mathcal{G}x \end{pmatrix} \geq 0. \tag{3.3}$$

Let

$$y^T := \begin{pmatrix} x^T \mathcal{H}^T & x^T \mathcal{G}^T \mathcal{H}^T \end{pmatrix}. \tag{3.4}$$

Then (3.3) is equivalent to $y^T M_{\mathcal{Q}}\, y \geq 0$. Now it suffices to prove that all vectors of the form y span $\mathbb{R}^{2n}$.

Suppose for a given x, with $n \geq 2$, the vectors x and $\mathcal{G}x$ are linearly independent. Let $y = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ be any vector in $\mathbb{R}^{2n}$, then we can choose $\mathcal{H} \in \mathbb{R}^{n\times n}$ with the property that $v_1 = \mathcal{H}x$ and $v_2 = \mathcal{H}\mathcal{G}x$. It is clear that vectors of the form

$$\mathcal{R}^x := \left\{ \begin{pmatrix} \mathcal{H}x \\ \mathcal{H}\mathcal{G}x \end{pmatrix} : \text{ for all } \mathcal{H} \right\}$$

is all $\mathbb{R}^{2n}$ as required. Thus we are finished unless for all x the vectors x and $\mathcal{G}x$ are linearly dependent. That is for all x, $\lambda_1(x)x + \lambda_2(x)\mathcal{G}x = 0$ for nonzero $\lambda_1(x)$ and

$\lambda_2(x)$. Note $\lambda_2(x) \neq 0$, unless $x = 0$. Set $\tau(x) := \frac{\lambda_1(x)}{\lambda_2(x)}$, then the linear dependence becomes $\tau(x)x + \mathcal{G}x = 0$. This says that every vector x is an eigenvector of $\mathcal{G}$, which implies that $\mathcal{G} = \lambda I$ for some constant λ. This fact can be verified from the Jordan form $\mathcal{G} = \mathcal{M}^{-1}\mathcal{J}\mathcal{M}$ via $\tau(x)\mathcal{M}x + \mathcal{J}\mathcal{M}x = 0$, for all x. Thus the set of all y satisfying (3.4) is all of $\mathbb{R}^{2n}$ unless $\tau I + \mathcal{G} = 0$ for some τ.

Conversely, if $\mathcal{G} = \lambda I$, then the set of y of the form (3.4) is not all of $\mathbb{R}^{2n}$ and has an orthogonal complement $\mathbb{R}^{\perp}$. The function $\mathcal{Q}$ can be positive without $r^T M_{\mathcal{Q}} r$ being positive on vectors $r \in \mathbb{R}^{\perp}$. □

Clearly the method used in the proof above to show that $\mathcal{R}^x$ is all of $\mathbb{R}^{2n}$ is very special. Part II of this paper uses a very different method (there are several parts to this more general proof). In a very vague sense, the main idea behind the proof is that if $\mathcal{R}^x$ is not all of $\mathbb{R}^{2n}$, then the coefficients L^i_j of the border vector form a set of linearly dependent functions. One consequence of this linear dependence property, which is of independent interest, is presented in the following corollary of Theorem 10.10 from Part II.

Corollary 3.2 (Corollary 10.11). *Let $L_1(\vec{Z}), \dots, L_\ell(\vec{Z})$ be noncommutative rational functions of $\vec{Z} = \{Z_1, \dots, Z_v\}$. For each vector x, suppose that the vectors $L_1(\vec{\mathcal{Z}})x$, $\dots$, $L_\ell(\vec{\mathcal{Z}})x$ are linearly dependent whenever matrices $\mathcal{Z}_j$ of compatible dimension are substituted for Z_j for all size Δ bigger than some Δ_0. Then there exist real numbers λ_j for $j = 1, \dots, \ell$ such that*

$$\sum_{j=1}^{\ell} \lambda_j L_j(\vec{Z}) = 0,$$

that is, the functions $L_j(\vec{Z})$ are linearly dependent.

We mention some basic work on positivity of commutative polynomials (not just quadratic polynomials) done in [Par00, PW98]. Our algorithm is somewhat like theirs, in that both use the LDL$^{\mathrm{T}}$ decomposition. While positivity of commutative quadratic functions is easily checked, noncommutative quadratics cause difficulties reminiscent of what happens with non-quadratic higher order commutative polynomials.

3.3. Noncommutative LDU Decomposition

In our approach, the LDU factorization of a matrix with noncommutative entries is the key tool for determination of the matrix positivity of a quadratic function, and hence the region of convexity $\mathcal{G}$ of noncommutative functions.

The LDU factorization applied to a symmetric matrix M of size $r \times r$ with noncommutative entries provides the decomposition $M = LDL^T$, where the $r \times r$ matrix D is diagonal[3] or contains 2×2 blocks with zeros on the diagonal, and

[3]This assumes that at each step of our LDU algorithm a matrix entry called *pivot* is invertible. The case where some pivot may not be invertible will be discussed in details in Theorem 3.3.

the $r \times r$ matrix L is lower triangular and normalized so that each diagonal entry equals the identity. To check the positivity of the symmetric matrix M it suffices to check that D is purely diagonal and to check the positivity of the entries of the diagonal matrix D. It is often very useful (sometimes essential) to perform the LDU decomposition not on a given matrix M but on a matrix PMQ obtained from M by permutation matrices P, Q. When M is symmetric, we shall choose $Q = P^T$ so as to obtain $PMP^T = LDL^T$, or equivalently $M = P^T LDL^T P$.

References on LDU decomposition of matrices with commutative entries are [HJ96, GL83]. The LDU decomposition for noncommutative 2×2 matrices is standard and appears in many places. We do not know a reference on the general $r \times r$ case. However, as we shall see its properties are much like the well understood commuting case. Note that at the k^{th} $\{k := 0, \ldots, r-2\}$ step of the process above, one can choose $(r-k)!$ permutations. The noncommutative $\mathrm{LDL^T}$ decomposition (as implemented in NCAlgebra) is briefly presented here.

Let a symmetric 2×2 matrix with noncommutative entries be given by $M = \begin{pmatrix} A & B^T \\ B & C \end{pmatrix}$ with A and C symmetric elements. Then M has the following $\mathrm{LDL^T}$ decomposition

$$M = \mathrm{LDL^T} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & C - BA^{-1}B^T \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix}, \qquad (3.5)$$

provided that the noncommutative element A is invertible. Our computer algorithm automatically assumes invertibility when it is needed. If the permutation

$$P = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

is applied to both sides of M producing

$$PMP^T = \begin{pmatrix} C & B \\ B^T & A \end{pmatrix},$$

the decomposition is

$$PMP^T = \mathrm{LDL^T} = \begin{pmatrix} I & 0 \\ B^TC^{-1} & I \end{pmatrix} \begin{pmatrix} C & 0 \\ 0 & A - B^TC^{-1}B \end{pmatrix} \begin{pmatrix} I & C^{-1}B \\ 0 & I \end{pmatrix}. \qquad (3.6)$$

Note that matrix D in the two decompositions (3.5) and (3.6) above has the classical Schur complements as its main ingredients.

Now we sketch the computer algebra algorithm for noncommutative symmetric matrices of size $r \times r$. Suppose that matrix M has $r \times r$ noncommutative entries. Then M can be always partitioned as

$$M = \begin{pmatrix} A_{11} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{C} \end{pmatrix} \qquad (3.7)$$

with $\mathbf{C}$ a matrix of size $(r-1) \times (r-1)$ and B a matrix of size $(r-1) \times 1$ with noncommutative entries. Now apply the 2×2 $\mathrm{LDL^T}$ decomposition as in (3.5) to get

$$\begin{pmatrix} I & 0 \\ \mathbf{B}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \mathbf{C} - \mathbf{B}A_{11}^{-1}\mathbf{B}^T \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}\mathbf{B}^T \\ 0 & I \end{pmatrix}.$$

In our symbolic algorithm we assume that if A_{11} is not 0, then it has an inverse denoted A_{11}^{-1}. We call A_{11} the **pivot** for this step of the algorithm. At the next step the $r-1 \times r-1$ matrix $\mathbf{C} - \mathbf{B}A_{11}^{-1}\mathbf{B}^T$ with noncommutative entries, called the **residual matrix**, can also be factored as $\hat{L}\hat{D}\hat{L}^T$ using a partition form analogous to (3.7). In that case M takes the form

$$M = \begin{pmatrix} I & 0 \\ \mathbf{B}A_{11}^{-1} & \hat{L} \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \hat{D} \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}\mathbf{B}^T \\ 0 & \hat{L}^T \end{pmatrix}.$$

The procedure continues until the residual matrix has size 1×1 (in which case we are finished) or the diagonal entry on which we need to pivot is 0. In the later case we find a non-zero diagonal entry A_{kk} and apply a permutation P from right and left to move this diagonal entry A_{kk} to the pivot[4] position. Then we proceed as before. This procedure with permutations stops when the residual matrix R has size 1×1 or all diagonal entries of the residual matrix R of size greater than 1 are identically zero (and no pivot is possible).

The key property of the Noncommutative LDU Algorithm is

Theorem 3.3. *Suppose M is a symmetric matrix of size $r \times r$ with noncommutative rational function entries. The possibly permuted LDU algorithm outputs a matrix D with noncommutative rational entries.* **Either** *D is* **diagonal***,*

i. *in which case, whenever $n \times n$ matrices are substituted for the variables in the function D_j, $j = 1, \dots, r$ in D and produce matrices $\mathcal{D}_j$, which for $j = 1, \dots, r-1$ are invertible, then*

   each $\mathcal{D}_j$ for $j = 1, \dots, r$ is a positive definite (resp. positive semidefinite) matrix if and only if the $rn \times rn$ matrix $\mathcal{M}$ resulting from M is positive definite (resp. positive semidefinite).

or *D can be partitioned as $D = \operatorname{diag}(\bar{D}, R)$,*[5] *where $\bar{D}$ is a diagonal matrix with noncommutative rational entries $\bar{D}_j$, $j = 1, \dots, d$ with $d < r-1$, and R is a* **non-diagonal** *matrix of size $(r-d) \times (r-d)$. We need to distinguish two situations:*

ii. **All entries of the matrix** R **are identically zero***, in which case D is actually diagonal, and the conclusion of case (i) applies.*

iii. **The off diagonal entries of** R **are not identically zero***, in which case some matrices substituted for the variables in M produce $\mathcal{M}$ which is neither a positive semidefinite matrix nor a negative semidefinite matrix.*

---

[4]A appealing way to choose A_{kk} is to observe that each diagonal entry typically will be a rational function of other entries in the matrix. Thus each A_{jj} is given by a formula of some length, and we select A_{kk} to be the nonzero diagonal entry of shortest length. This is a symbolic analog of the common numerical analysis method of picking the pivot of largest size.

[5]$\operatorname{diag}(x_1, \dots, x_r)$ means a diagonal matrix with entries $x_1, \dots, x_r$.

Proof. Prove (i,ii): Suppose D is diagonal with entries $D_j, j = 1, \ldots, d$ not identically zero. Our symbolic algorithm used an expression denoting the inverse of each pivot. Note that the pivots used in the algorithm (and assumed invertible) are exactly the diagonal elements D_j, for $j = 1, \ldots, \min(d, r-1)$. Thus our symbolic formulas are valid when matrices are substituted in, provided that the resulting matrix diagonal entries $\mathcal{D}_j$ for $j = 1, \ldots, \min(d, r-1)$, are invertible. Thus $\mathcal{D}_j$, for $j = 1, \ldots, d$, positive semidefinite (resp. for $j = 1, \ldots, r$, each $\mathcal{D}_j$ a positive definite matrix) implies that $\mathcal{M}$ is positive semidefinite (resp. positive definite). Conversely, if $\mathcal{M}$ is positive semidefinite (resp. positive definite) the $\mathcal{D}_j$, for $j = 1, \ldots, d$, are positive semidefinite (resp. for $j = 1, \ldots, r$, each $\mathcal{D}_j$ is positive definite) since $\mathcal{L}$ is invertible.

Now we prove (iii): If $n \times n$ matrices of any size n are substituted for the variables in M and in R the resulting symmetric residual matrix $\mathcal{R}$ has block diagonal entries equal to the $n \times n$ zero matrix, which implies that $\mathcal{R}$ has trace 0, which implies $\mathcal{R}$ has some positive and some negative eigenvalues. Thus $\mathcal{R}$ and consequently $\mathcal{M}$ can not be either a positive semidefinite matrix or a negative semidefinite matrix. □

While we have presented only enough of the LDL$^{\mathrm{T}}$ decomposition for noncommutative symmetric matrices to determine positivity, in fact the NCAlgebra program can do more. If the user chooses a certain option, NCAlgebra picks a non zero 2×2 block in R and pivots on it. This procedure combined with permutations when needed, ultimately produces a center matrix D which is block diagonal with blocks of size 1×1 or 2×2. This exactly generalizes the standard behavior of the commutative case.

A further feature of our NCAlgebra implementation is that one can retrieve the sequence of permutations which the algorithm selected. Also one can specify exactly which permutations are to be used and thereby override the algorithm's automatic selection of permutations.

A brief summary of a simplified version of the LDL$^{\mathrm{T}}$ algorithm code implemented in the NCAlgebra package follows.

Algorithm 3.4 (Noncommutative LDL$^{\mathrm{T}}$ Decomposition)**.**

Set $k = 0$, $M_k = M$
while $k < r$ **do**
  Apply desired permutation on M_k
  Partition M_k as in (3.7)
  $L_k D_k L_k^T \leftarrow M_k$; as in (3.5)
  Append: $L \leftarrow L_k$; $D \leftarrow D_k$
  Let M_k be the residual $C_k - B_k A_k^{-1} B_k^T$
  $k \leftarrow k + 1$
end

NCAlgebra Command: NCLDUDecomposition[M], gives a permuted LDU decomposition of a symmetric M.

4. Convexity Algorithm

This section presents our main algorithm that provides the region $\mathcal{G}$ in which a given noncommutative symmetric function $\Gamma(\vec{Z})$ is matrix convex in $\vec{X}$.

1. Compute symbolically $\mathcal{Q}(\vec{Z})[\vec{H}] := \mathcal{H}\Gamma(\vec{X})[\vec{H}]$.
2. As $\mathcal{Q}(\vec{Z})[\vec{H}]$ is second order in $\vec{H}$, it can be expressed as $V[\vec{H}]^T M_{\mathcal{Q}(\vec{Z})} V[\vec{H}]$. Extract the matrix $M_{\mathcal{Q}(\vec{Z})}$ from this quadratic expression.
3. Apply the noncommutative $\mathrm{LDL^T}$ decomposition on the matrix $M_{\mathcal{Q}(\vec{Z})}$, i.e., $M_{\mathcal{Q}(\vec{Z})} = LDL^T$, to get matrix D with noncommutative entries.
4. Suppose that matrix D can be partitioned as $D = \mathrm{diag}(\bar{D}, R)$, where $\bar{D}$ is a diagonal matrix with entries $\rho_j(\vec{Z})$, for $j = 1, \ldots, \tilde{d}$ and R is a non-diagonal matrix of size $(r - \tilde{d}) \times (r - \tilde{d})$ containing zeros on the diagonal or 2×2 blocks
$$R_i = \begin{pmatrix} 0 & \rho_i(\vec{Z}) \\ \rho_i(\vec{Z})^T & 0 \end{pmatrix}$$
for $i = \tilde{d} + 1, \ldots, r$. Thus matrix D has the form
$$D = \begin{pmatrix} \rho_1(\vec{Z}) & & & & & & & & \\ & \ddots & & & & & & & \\ & & \rho_{\tilde{d}}(\vec{Z}) & & & & & & \\ & & & 0 & \rho_{\tilde{d}+1}(\vec{Z}) & & & & \\ & & & \rho_{\tilde{d}+1}(\vec{Z})^T & 0 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 0 & \rho_r(\vec{Z}) & \\ & & & & & & \rho_r(\vec{Z})^T & 0 & \\ & & & & & & & & 0 \end{pmatrix}.$$
5. The Hessian $\mathcal{Q}(\vec{\mathcal{Z}})[\vec{\mathcal{H}}]$ is a positive semidefinite matrix for all $\vec{\mathcal{H}}$ whenever the tuple of matrices $\vec{\mathcal{Z}} = \{\mathcal{Z}_1, \ldots, \mathcal{Z}_v\}$ makes the block diagonal matrix D positive semidefinite. Thus a set $\mathcal{G}$ where $\Gamma(\vec{Z})$ is matrix convex is given by
$$\mathcal{G} = \left\{\vec{Z} \;:\; \rho_j(\vec{Z}) > 0,\; j = 1, \ldots, \tilde{d}\right\} \bigcap \left\{\vec{Z} \;:\; \rho_i(\vec{Z}) = 0,\; i = \tilde{d} + 1, \ldots, r\right\}.$$
6. Note that, if $M_{\mathcal{Q}(\vec{Z})}$ is a matrix of size $r \times r$, then there are $\Pi = r!(r - 1)! \cdots 2$ possible $\mathrm{LDL^T}$ decompositions depending on different permutations of the matrix $M_{\mathcal{Q}(\vec{Z})}$. This gives Π different diagonal matrices, D^1,

$D^2, \ldots, D^{\Pi}$. Up to the assumptions that the NCLDUDecomposition algorithm makes about invertibility, each D^i must produce a set $\mathcal{G}$. However, the inequalities produced by the diagonal D^i may be much more elegant and useful than those produced by the diagonal D^j, even though they must produce equivalent sets $\mathcal{G}$.

The main difficulty is the fact that there are Π different permutations for doing the $\mathrm{LDL^T}$ decomposition. Checking them all consumes computer time and leaves the user with many choices. In our experience many permutations work to give the same answer (as will be shown in some examples), so finding a satisfactory one appears not to be time consuming.

The set $\mathcal{G}$ produced by the Convexity Algorithm is the biggest possible in a certain sense. This is the content of Theorem 3.1 and is described precisely in Theorem 8.2 of Part II.

5. Examples

In this section we give several examples of the Convexity Algorithm which vary in complication and which illustrate different points. We begin with a simple example.

Example 5.1. Define the function $\Gamma(X)$ by

$$\Gamma(X) = G^T X^T AXG + X^T BX + G^T X^T CX + X^T C^T XG,$$

where $B = B^T$ and $A = A^T$. The Hessian of $\Gamma(X)$ is given by

$$\mathcal{H}\Gamma(X)[H] = 2(H^T BH + H^T C^T HG + G^T H^T AHG + G^T H^T CH).$$

Equivalently, this quadratic expression takes the form

$$\mathcal{H}\Gamma(X)[H] = V[H]^T M_{\mathcal{H}\Gamma} V[H] = 2(H^T, G^T H^T) \begin{pmatrix} B & C^T \\ C & A \end{pmatrix} \begin{pmatrix} H \\ HG \end{pmatrix}.$$

The $\mathrm{LDL^T}$ decomposition with no permutation applied to $M_{\mathcal{H}\Gamma}$ is

$$\begin{pmatrix} I & 0 \\ CB^{-1} & I \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & A - CB^{-1}C^T \end{pmatrix} \begin{pmatrix} I & B^{-1}C^T \\ 0 & I \end{pmatrix},$$

provided that B is invertible[6].

Therefore, when B is invertible and $G \neq \alpha I$, for any scalar α, the necessary and sufficient conditions for the Hessian to be positive semidefinite are

$$B > 0 \text{ and } A - CB^{-1}C^T \geq 0.$$

[6] The list returned by **NCConvexityRegion** is

$$\{B,\ A - CB^{-1}C^T\}.$$

On the other hand, if A is invertible and a permutation is applied, the $\mathrm{LDL^T}$ decomposition is

$$\begin{pmatrix} I & 0 \\ C^T A^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B - C^T A^{-1} C \end{pmatrix} \begin{pmatrix} I & A^{-1}C \\ 0 & I \end{pmatrix}.$$

For this case, the necessary and sufficient conditions are

$$A > 0 \text{ and } B - C^T A^{-1} C \geq 0.$$

□

5.1. NCAlgebra Examples

Henceforth our examples will use notation which is standard in Mathematica and NCAlgebra. This adds a level of precision and concreteness to the discussion. Also the notation is quite transparent so it causes little reading difficulty. Sometimes for better visualization, TEX notation is employed. In the course of illustrating the Convexity Algorithm we actually show what is inside the command **NCConvexityRegion[]**.

Before going through the examples, it is convenient to explain the basic notation used in NCAlgebra. The transpose of an element x is denoted by $tp[x]$. The identity is denoted 1. The inverse of x is $inv[x]$. The product of the noncommutative elements x and y is $x * * y$. The product of a matrix A with noncommutative entries by another matrix B is provided by the command $MatMult[A, B]$.

The directional derivative, the Hessian, and the LDU decomposition, were already introduced. They are provided from: Hessian[$f(X, Y)$, $\{X, H\}$, $\{Y, K\}$], DirectionalD[$\Gamma(X, Y)$, $\{X, H\}$, $\{Y, K\}$], and NCLDUDecomposition[$Matrix$].

The border vector and the coefficient matrix of a noncommutative quadratic function is given by NCMatrixOfQuadratic[$\mathcal{Q}$, $\{H, K\}$].

The command NCExpand[*expression*] expands out noncommutative multiply's inside an algebraic expression. It is the noncommutative generalization of the Mathematica Expand[ ].

The command NCSimplifyRational[ ], simplifies an expression that includes polynomials and inverses of polynomials. This works by applying a collection of simplifying rules to the expression. The call is NCSimplifyRational[*expression*]. This is in practice an essential command because the expressions obtained by other commands, such as NCLDUDecomposition[ ], Hessian[ ], etc., usually are not in their simplified form. For more details about simplification of noncommutative expressions and symbolic implementation, the reader is referred to [HSW98].

The following examples describe the steps for checking the convexity of a noncommutative function.

Example 5.2. Let the function Γ be given by

$$F := XA + A^T X - (C1^T - X\,B\,D1^T)(Y - D1\,D1^T)^{-1}(C1 - D1\,B^T X) - XBB^T X$$

with $X = X^T$ and $Y = Y^T$. The definition of this function F in Mathematica is:

In[1]:= F := X**A+tp[A]**X - (tp[C1]-X**B**tp[D1])**inv[Y-D1**tp[D1]] ** (C1-D1**tp[B]**X) - X**B**tp[B]**X;

The Hessian of this function is produced by the command

In[2]:= hess = 1/2 **NCHessian**[F, {X, H}, {Y, K}] // **NCSimplifyRational**;

The left (right) border vector and the coefficient matrix Mhess are produced by the command

In[3]:= {LeftBorder, Mhess, RightBorder} = **NCMatrixOfQuadratic**[hess, H, K];

The matrix Mhess from the command above in TEX format is

$$\text{Mhess} = \begin{pmatrix} -BB^T - BD1^T RD1B^T & -BD1^T R & BD1^T R \\ -RD1B^T & -R & R \\ RD1B^T & R & -R \end{pmatrix},$$

where we have made the substitution $R := (Y - D1D1^T)^{-1}$. The LDL^T decomposition of Mhess is obtained by the command

In[4]:= {lu, di, up, P} = **NCLDUDecomposition**[Mhess] // **NCSimplifyRational**;

From the output of this command we obtain the diagonal matrix di, presented below in TEX format

$$\text{di} = \begin{pmatrix} -(Y - D1D1^T)^{-1} & 0 & 0 \\ 0 & -BB^T & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The list returned by **NCConvexityRegion** is the entries of the diagonal matrix di:

$$\{-(Y - D1D1^T)^{-1},\ -BB^T,\ 0\}.$$

Therefore we may conclude that the function F is concave on the region $\mathcal{G} := \{Y : Y - D1D1^T > 0\}$.

To determine that $\bar{\mathcal{G}} := \{Y : Y - D1D1^T \geq 0\}$ is the biggest domain of concavity we need to check if the border vector is linearly independent and if the region $\mathcal{G}$ satisfies the Openness Property[7]. The left border vector "LeftBorder" is

$$\text{LeftBorder} = \{H,\ C1^T(Y - D1D1^T)^{-1}K,\ XBD1^T(Y - D1D1^T)^{-1}K\}.$$

This border vector has linearly independent[8] coefficients for each H and K. To see that, we need to analyze separately the coefficients for the H and K. The H case is trivial as it appears only once. For the K, we need to show that the functions $L_1(Y) := C1^T(Y - D1D1^T)^{-1}$ and $L_2(X,Y) := XBD1^T(Y - D1D1^T)^{-1}$ are linearly independent, which is immediate as L_1 does not depend on X. We remark that the output of the LeftBorder is an option in NCConvexityRegion. Also a sufficient though not necessary test for linear independence of the LeftBorder vector entries is automatically implemented. This test is sketch later in Example 5.3.

---

[7]See the Openness Property in Section 7.2 referred to in item (*ii*) of Theorem 3.1

[8]A rigorous treatment is given in Definition 7.1, where the block linearly independence property is defined.

It is also evident from the strict inequality that for matrices of any compatible dimension the domain $\mathcal{M}(\mathcal{G})$ of matrices is an open set; thus $\mathcal{G}$ satisfy the Openness Property. Therefore we conclude the region $\bar{\mathcal{G}} := \{Y : Y - D1D1^T \geq 0\}$ is the biggest domain of concavity for the function F.

□

An interesting aspect of the next example is that it shows that the $M_{\mathcal{Q}}$ representation may not be unique. This may lead one to conclude that a function is matrix positive instead of being matrix strictly positive.

Example 5.3. Let x, y, h and k be symmetric noncommutative elements. Let's define the noncommutative function $F(x, y)$ to be used in the example as

$$F(x, y) := (x - y^{-1})^{-1}.$$

This function F in Mathematica takes the form:

In[5]:= F := inv[x - inv[y]];

Thus, the Hessian $\mathcal{H}\Gamma(x, y)[h, k]$ of this function is produced by the command

In[6]:= hess = 1/2 **NCHessian**[F, {x, h}, {y, k}] // **NCExpand**

```
inv[x - inv[y]] ** h ** inv[x - inv[y]] ** h ** inv[x - inv[y]] +
inv[x - inv[y]] ** h ** inv[x - inv[y]] ** inv[y] ** k ** inv[y] **
inv[x - inv[y]] + inv[x - inv[y]] ** inv[y] ** k ** inv[y] ** k **
inv[y] ** inv[x - inv[y]] + inv[x - inv[y]] ** inv[y] ** k ** inv[y]
** inv[x - inv[y]] ** h ** inv[x - inv[y]] + inv[x - inv[y]] **
inv[y] ** k ** inv[y] ** inv[x - inv[y]] ** inv[y] ** k ** inv[y] **
inv[x - inv[y]]
```

The left (right) border vector and the coefficient matrix Mhess are produced by the command

In[7]:= {LeftBorder, Mhess, RightBorder} = **NCMatrixOfQuadratic**[hess, {h, k}];

The Hessian of F, denoted by hess, can be rewritten in TEX format as

$$\text{hess} = V^T \text{Mhess}\ V,$$

where V^T = LeftBorder is given by

$$V^T = \begin{bmatrix} h(x - y^{-1})^{-1} \\ ky^{-1}(x - y^{-1})^{-1} \end{bmatrix}^T$$

and the matrix Mhess is given by

$$\text{Mhess} = \begin{bmatrix} (x - y^{-1})^{-1} & (x - y^{-1})^{-1}y^{-1} \\ y^{-1}(x - y^{-1})^{-1} & y^{-1} + y^{-1}(x - y^{-1})^{-1}y^{-1} \end{bmatrix}.$$

The $\mathrm{LDL^T}$ decomposition of the coefficient matrix Mhess is given by the command

In[8]:= {lu, di, up, P} = **NCLDUDecomposition**[Mhess] // **NCSimplifyRational**;

From the output of this command we obtain the following factorization for P Mhess $P^T =$ lu di up

$$\begin{pmatrix} I & 0 \\ y^{-1} & I \end{pmatrix} \begin{pmatrix} (x-y^{-1})^{-1} & 0 \\ 0 & y^{-1} \end{pmatrix} \begin{pmatrix} I & y^{-1} \\ 0 & I \end{pmatrix},$$

where P is a permutation matrix generated automatically by our LDU algorithm. Finally, the list returned by **NCConvexityRegion** is the entries of the diagonal matrix di, i.e.,

$$\{(x-y^{-1})^{-1},\ y^{-1}\}.$$

Therefore the Hessian is matrix strictly positive on the Symbolic Inequality Domain

$$\mathcal{G} := \{(x,y) : y > 0 \text{ and } x - y^{-1} > 0\}. \tag{5.1}$$

Now, Let's analyze the effect of a different representation for the Hessian. Where instead of expanding the expression for the Hessian with the command NCExpand, we apply the command NCSimplifyRational.

In[9]:= hess = 1/2 **NCHessian**[F, {x, h}, {y, k}] // **NCSimplifyRational**

```
k ** h ** inv[x - inv[y]] + inv[x - inv[y]] ** h ** k - k ** x **
inv[x - inv[y]] ** h ** inv[x - inv[y]] + k ** x ** inv[x -
inv[y]] ** inv[y] ** k - inv[x - inv[y]] ** h ** k ** x ** inv[x -
inv[y]] + inv[x - inv[y]] ** h ** inv[x - inv[y]] ** h ** inv[x -
inv[y]] - inv[x - inv[y]] ** h ** inv[x - inv[y]] ** x ** k -
inv[x - inv[y]] ** x ** k ** h ** inv[x - inv[y]] - k ** x **
inv[x - inv[y]] ** inv[y] ** k ** x ** inv[x - inv[y]] + inv[x -
inv[y]] ** h ** inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]]
+ inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] ** h ** inv[x
- inv[y]] - inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] **
inv[y] ** k + inv[x - inv[y]] ** x ** k ** x ** inv[x - inv[y]] **
inv[y] ** k ** x ** inv[x - inv[y]]
```

The LeftBorder (RightBorder) vector and the coefficient matrix Mhess are produced by the command

In[10]:= {LeftBorder, Mhess, RightBorder} = **NCMatrixOfQuadratic**[hess, {h, k}];

The Hessian of F can be rewritten in TEX format as hess = V^T Mhess V, where $V^T =$ LeftBorder, given by

$$V^T = \left(k,\ (x-y^{-1})^{-1}h,\ (x-y^{-1})^{-1}xk\right),$$

has linearly independent coefficients, and the matrix Mhess is

$$\text{Mhess} = \begin{pmatrix} x(x-y^{-1})^{-1}y^{-1} & 1-x(x-y^{-1})^{-1} & -x(x-y^{-1})^{-1}y^{-1} \\ 1-(x-y^{-1})^{-1}x & (x-y^{-1})^{-1} & -1+(x-y^{-1})^{-1}x \\ -x(x-y^{-1})^{-1}y^{-1} & -1+x(x-y^{-1})^{-1} & x(x-y^{-1})^{-1}y^{-1} \end{pmatrix}.$$

The LDL^T decomposition of the coefficient matrix Mhess is given by the command

In[11]:= {lu, di, up, P} = **NCLDUDecomposition**[Mhess];

From the output of this command we obtain the following factorization for P Mhess P^T = lu di up

$$\begin{pmatrix} I & 0 & 0 \\ y^{-1} & I & 0 \\ -y^{-1} & -I & I \end{pmatrix} \begin{pmatrix} (x-y^{-1})^{-1} & 0 & 0 \\ 0 & y^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} I & y^{-1} & -y^{-1} \\ 0 & I & -I \\ 0 & 0 & I \end{pmatrix}. \tag{5.2}$$

Finally, the list returned by **NCConvexityRegion** is

$$\{(x-y^{-1})^{-1},\ y^{-1},\ 0\}.$$

Thus the region of convexity for F contains

$$\mathcal{G} := \{(x,y) : y > 0 \text{ and } x - y^{-1} > 0\}. \tag{5.3}$$

Naturally, this is the same domain that was already determined in (5.1).

To insure that $\bar{\mathcal{G}} := \{(x,y) : y > 0 \text{ and } x - y^{-1} \geq 0\}$ contains the biggest region of convexity of F, we must verify hypotheses (i) and (ii) of Theorem 3.1. The linear dependence of the coefficients of the border vector states, as in hypothesis (i), that there exist λ_1, λ_2 scalars such that $\lambda_1 I + (x-y^{-1})^{-1}x\lambda_2 = 0$ for all symmetric x, y. It follows that the coefficients of the border vector are linearly independent. Now we say a few words about a practical test guaranteeing linear independence of the border vector, that is guaranteeing hypotheses (i) of Theorem 3.1. This test is implemented in the command NCConvexityRegion. The idea is to declare all variables to commute; then compute a linear combination of the coefficient functions of the border vector which is 0. If the only linear combination is 0, then this insures that condition (i) holds. This is a conservative test and our example passes it.

To check condition (ii) of Theorem 3.1, without going into the topology involved, we just say that because the inequalities in 5.3 are strict, the set of $n \times n$ symmetric matrices which satisfy them (for each large n) contains an open set. This suffices to satisfy (ii).

We should emphasize the fact that if we conclude that a function is matrix convex, it could be quite possible that the function actually is matrix "strictly" convex. This happens because we do not have a way to guarantee a unique representation for the matrix M_Q. However, the biggest possible domain of convexity of F, the "closure" of $\mathcal{G}$, is uniquely determined whatever representation is used.

Now we discuss permutations. One can observe that for this example (the 3×3 case) there are 12 $\mathrm{LDL^T}$ factorizations, related to all possible permutations. We computed them and found that four permutations provide identical decompositions to the one in (5.2), four permutations give division[9] by 0, and the other four give

[9]NCLDUDecomposition[ ] contains (automatic) logical rules for permutations to bypass division by 0. Using this automatic permutation, which is the default, the four decompositions provide diagonal matrices identical to the one in (5.4).

the following diagonal matrix:

$$\begin{pmatrix} -x + x(x-y^{-1})^{-1}x & 0 & 0 \\ 0 & \left\{ \begin{array}{l} -y + x^{-1} - (x-y^{-1})^{-1} + yx(x-y^{-1})^{-1} \\ + (x-y^{-1})^{-1}xy + (x-y^{-1})^{-1}x(x-y^{-1})^{-1} \\ - (x-y^{-1})^{-1}xyx(x-y^{-1})^{-1} \end{array} \right\} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.4}$$

□

Example 5.4. Define the function Γ as

$$F := -X + Y - (Y + A^T XB)(R + B^T XB)^{-1}(Y + B^T XA) + A^T XA, \tag{5.5}$$

with $X = X^T$, $Y = Y^T$ and $R = R^T$. In Mathematica it takes the form

In[12]:= F := - X + Y - (Y+tp[A]**X**B) ** inv[R+tp[B]**X**B] ** (Y+tp[B]**X**A) + tp[A]**X**A;

For that function the Hessian and the coefficient matrix are obtained from the commands:

In[13]:= hess = **NCHessian**[F, {X,H}, {Y,K}] // **NCSimplifyRational**;

In[14]:= {LeftBorder, Mhess, RightBorder} = **NCMatrixOfQuadratic**[hess, H, K];

The $\mathrm{LDL^T}$ decomposition of Mhess is obtained by

In[15]:= {lu, di, up, P} = **NCLDUDecomposition**[Mhess] // **NCSimplifyRational**;

From the output of this command we obtain the diagonal matrix di, presented below

$$\mathrm{di} = \begin{pmatrix} -2\, inv[R + tp[B] **X**B] & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The list returned by **NCConvexityRegion** is the entries of the diagonal matrix di above. The corresponding lower triangular matrix lu is

$$\mathrm{lu} = \begin{pmatrix} I & 0 & 0 & 0 \\ B & I & 0 & 0 \\ -B & 0 & I & 0 \\ -B & 0 & 0 & I \end{pmatrix}.$$

The coefficient matrix is

$$\mathrm{Mhess} = -2 \begin{pmatrix} I \\ B \\ -B \\ -B \end{pmatrix} inv[R + tp[B] **X**B] \begin{pmatrix} I & tp[B] & -tp[B] & -tp[B] \end{pmatrix}.$$

Therefore the condition for negative semi-definiteness of Mhess is $R + tp[B] **X**B > 0$. In which, one concludes that the function F in (5.5) is concave on the region $\{X : R + tp[B] **X**B > 0\}$. □

Part II. Theoretical Results and Proofs

Earlier in Section 3.2, we saw that positivity of the matrix $M_{\mathcal{Q}}$ implies matrix positivity of the associated quadratic function $\mathcal{Q}$. Also, Example 3.1 in Section 3.1 gives a glimpse of the main linear independence idea behind the converse. Part II fully addresses the converse; we know that the quadratic function $\mathcal{Q}$ is matrix positive in some sense and we wish to conclude that the matrix $M_{\mathcal{Q}}$ is also matrix positive. Our main results show a substantial class of cases in which this is true. From these results we obtain under weak hypotheses that our Convexity Algorithm determines exactly the correct Symbolic Inequality Domain up to its "closure".

Part II of this paper is a bit redundant with Part I, so that it can be read without constantly flipping back to Part I.

6. Main Theorem on Sufficient Condition for Convexity

As we now see, it is easy to prove that our Convexity Algorithm in Section 4 produces a Symbolic Inequality Domain $\mathcal{G}$ on which a noncommutative symmetric rational function Γ is matrix convex on $\mathcal{G}$.

Remark 6.1. We do not analyze the full Convexity Algorithm, but we shall treat only the case where the residual matrix R in the LDU decomposition is identically zero. The reason we do little work on this case is that matrix D can be partitioned as

$$D = \begin{pmatrix} \rho_1(\vec{Z}) & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \rho_{\tilde{d}}(\vec{Z}) & & & & & & & \\ & & & 0 & \rho_{\tilde{d}+1}(\vec{Z}) & & & & & \\ & & & \rho_{\tilde{d}+1}(\vec{Z})^T & 0 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 0 & \rho_r(\vec{Z}) & & \\ & & & & & & \rho_r(\vec{Z})^T & 0 & & \\ & & & & & & & & & 0 \end{pmatrix}.$$

This matrix D is positive semidefinite for $\vec{\mathcal{Z}}$ only if $\vec{\mathcal{Z}}$ makes $\rho_j(\vec{\mathcal{Z}}) \geq 0$ for $j = 1, \ldots, \tilde{d}$ and $\rho_i(\vec{\mathcal{Z}}) = 0$ for $i = \tilde{d}+1, \ldots, r$. The constraint $\rho_i(\vec{\mathcal{Z}}) = 0$ is very demanding and typically will force the Symbolic Inequality Domain $\mathcal{G}$ to violate the Openness Property. We have not analyzed this situation carefully, since we felt confident that it would not cause difficulties in our Convexity Algorithm. The NCConvexityRegion command lists the domain of convexity $\mathcal{G}$ for $\Gamma(\vec{Z})$ as those

$\vec{Z}$ such that

$$\mathcal{G} = \left\{ \vec{Z} \ : \ \rho_j(\vec{Z}) > 0, \ j = 1, \ldots, \tilde{d} \right\} \bigcap \left\{ \vec{Z} \ : \ \rho_i(\vec{Z}) = 0, \ i = \tilde{d} + 1, \ldots, r \right\}.$$

The strict inequality $\rho_j(\vec{Z}) > 0$ reflects the fact that the LDU algorithm requires invertibility of the ρ_j for $j = 1, \ldots, \tilde{d}$.

Theorem 6.2 (Sufficient Condition for Convexity)**.** *Let $\Gamma(\vec{Z})$ with $\vec{Z} = \{\vec{A}, \vec{X}\}$ be a noncommutative symmetric rational function. The function $\Gamma(\vec{Z})$ may be or may not be hereditary*[10]*. Suppose that the coefficient matrix $M_{\mathcal{H}\Gamma}$ of the Hessian $\mathcal{H}\Gamma(\vec{X})[\vec{H}]$ has a noncommutative $L(\vec{Z})D(\vec{Z})L(\vec{Z})^T$ decomposition with diagonal $D(\vec{Z})$ whose entries are all matrix positive on a Symbolic Inequality Domain*[11] *$\mathcal{G}$. Then $\Gamma(\vec{Z})$ is matrix convex on $\mathcal{G}$.*

Proof. It suffices to prove that the Hessian $\mathcal{H}\Gamma(\vec{X})[\vec{H}]$ is a matrix positive quadratic function for $\vec{Z} = \{\vec{A}, \vec{X}\}$ in the Symbolic Inequality Domain $\mathcal{G}$. Let $\mathcal{H}\Gamma(\vec{X})$ $[\vec{H}]$ be in the form $V[\vec{H}]^T M_{\mathcal{H}\Gamma} V[\vec{H}]$, where $M_{\mathcal{H}\Gamma} = L(\vec{Z})D(\vec{Z})L(\vec{Z})^T$. Thus

$$\mathcal{H}\Gamma(\vec{X})[\vec{H}] = V[\vec{H}]^T L(\vec{Z})D(\vec{Z})L(\vec{Z})^T V[\vec{H}]. \tag{6.1}$$

Now, substitute for $\vec{Z}$ and $\vec{H}$ in (6.1) any tuple of matrices $\vec{\mathcal{H}}$ and $\vec{\mathcal{Z}} = \{\vec{\mathcal{A}}, \vec{\mathcal{X}}\}$ in $\mathcal{M}(\mathcal{G})$[12] of compatible dimension. Since $D(\vec{\mathcal{Z}})$ has positive semidefinite entries, formula (6.1) implies that $\mathcal{H}\Gamma(\vec{\mathcal{X}})[\vec{\mathcal{H}}]$ is positive semidefinite. This says that $\Gamma(\vec{Z})$ is matrix convex on $\mathcal{G}$. □

7. Key Definitions

This section presents the definitions essential for the statement of our most general theorem, which shows that no "bigger" Symbolic Inequality Domain than the $\mathcal{G}$ produced by our Convexity Algorithm yields a function Γ which is matrix convex on $\mathcal{G}$. We start with a simple illustrative case and then we present the general case.

7.1. Definitions of Linearly Dependent Functions and Borders

To make sure there is no confusion in understanding our results and discussion of borders we include notational discussion which looks at the border of a quadratic function $\mathcal{Q}$ carefully.

---

[10]Defined in Section 2.1, Part I.
[11]Defined in Section 2.4, Part I.
[12]Defined in Section 2.4, Part I.

7.1.1. The Basic Idea Now we illustrate what we mean by linearly independent border vector. For simplicity of exposition, the hereditary function $\mathcal{Q}$ is limited to be quadratic in two noncommutative variables H_1 and H_2 ($\vec{H} := \{H_1, H_2\}$). In the next section, we will extend the idea to the case of several variables. Let the hereditary quadratic function $\mathcal{Q}(\vec{Z})[\vec{H}]$ take the form

$$\begin{aligned}\mathcal{Q}(\vec{Z})[\vec{H}] =& \sum_{s=1}^{\ell_1}\sum_{t=1}^{\ell_1} {L_s^1}^T(\vec{Z}) H_1^T A_{s,t}(\vec{Z}) H_1 L_t^1(\vec{Z}) \\ &+ sym \sum_{s=1}^{\ell_1}\sum_{t=1}^{\ell_2} {L_s^1}^T(\vec{Z}) H_1^T A_{s,t+\ell_1}(\vec{Z}) H_2 L_t^2(\vec{Z}) \\ &+ \sum_{s=1}^{\ell_2}\sum_{t=1}^{\ell_2} {L_s^2}^T(\vec{Z}) H_2^T A_{s+\ell_1,\,t+\ell_1}(\vec{Z}) H_2 L_t^2(\vec{Z}).\end{aligned}$$

Where each $L_j^i(\vec{Z})$ is a rational function not necessarily distinct; may even be the identity matrix. The quantity ℓ_i is the number of times that the monomial of order two in H_i appears. For the case above, the border of the matrix valued function $\mathcal{Q}(\vec{Z})[\vec{H}]$ has the form

$$V(\vec{Z})[\vec{H}] := \begin{pmatrix} H_1 L_1^1(\vec{Z}) \\ H_1 L_2^1(\vec{Z}) \\ \vdots \\ H_1 L_{\ell_1}^1(\vec{Z}) \\ H_2 L_1^2(\vec{Z}) \\ \vdots \\ H_2 L_{\ell_2}^2(\vec{Z}) \end{pmatrix}. \tag{7.1}$$

In this border, the H_1 and H_2 parts operate independently, so we shall consider separately the polynomials, which are the coefficients of H_1 and H_2:

$$\vec{L}^1(\vec{Z}) := \{L_1^1(\vec{Z}), \ldots, L_{\ell_1}^1(\vec{Z})\} \tag{7.2}$$

and

$$\vec{L}^2(\vec{Z}) := \{L_1^2(\vec{Z}), \ldots, L_{\ell_2}^2(\vec{Z})\}. \tag{7.3}$$

Definition 7.1 (Linearly Independent Functions Property)**.** For a given i, the noncommutative rational functions $L_j^i(\vec{Z})$ for $j = 1, \ldots, \ell_i$ are said to be **linearly independent functions** if the only scalars λ_j, such that

$$\sum_{j=1}^{\ell_i} \lambda_j L_j^i(\vec{Z}) = 0$$

are $\lambda_1 = \lambda_2 = \cdots = \lambda_{\ell_i} = 0$. We emphasize that the scalars λ_j do not depend on $\vec{Z}$. If there exists such nonzero scalars, the functions $L^i_j(\vec{Z})$ are said to be **linearly dependent functions**.

As we shall see what is critical for our Convexity Algorithm is when either $\vec{L}^1(\vec{Z})$ or $\vec{L}^2(\vec{Z})$ is a **linearly dependent set of functions**. We say that the border vector $V(\vec{Z})[\vec{H}]$ in (7.1) has block linearly independent coefficients, if neither the functions $\vec{L}^1(\vec{Z})$ in (7.2) nor the functions $\vec{L}^2(\vec{Z})$ in (7.3) are linearly dependent. In the next section, we repeat all of these definitions for the most general case.

7.1.2. The General Case In the most general case, the quadratic function $\mathcal{Q}(\vec{Z})[\vec{H}]$ is not constrained to be hereditary. Let's define $\vec{H}$ as

$$\vec{H} := \{H_{-h}, \ldots, H_{-1}, H_1, \ldots, H_h, H_{h+1}, \ldots, H_g, H_{g+1}, \ldots, H_k\}, \tag{7.4}$$

where $\{H_j\}_{j=g+1}^k$ are constrained to be symmetric and $H_j = H_{-j}^T$, for $j = 1, \ldots, h$. That is, we can separate $\vec{H}$ into three different parts as follows: the first part[13] $\{H_j\}_{j=-h}^h$ has the pairwise restriction that $H_{-j} = H_j^T$, for $j = 1, \ldots, h$, the second part $\{H_j\}_{j=h+1}^g$ has no restriction, the third part $\{H_j\}_{j=g+1}^h$ has each H_j constrained to be symmetric. Let $\mathcal{I}$ denote the integers between $-h$ and k except for 0. This is the index set for the H_j which are the entries of $\vec{H}$.

Any noncommutative symmetric quadratic $\mathcal{Q}(\vec{Z})[\vec{H}]$ can be put in the form $V(\vec{Z})[\vec{H}]^T M_{\mathcal{Q}(\vec{Z})} V(\vec{Z})[\vec{H}]$, where the border $V(\vec{Z})[\vec{H}]$ has the form

$$V(\vec{Z})[\vec{H}] := \begin{pmatrix} V^{mix}(\vec{Z})[\vec{H}] \\ V^{pure}(\vec{Z})[\vec{H}] \\ V^{sym}(\vec{Z})[\vec{H}] \end{pmatrix}, \tag{7.5}$$

[13]The integer 0 is not included in the index set $j = -h, \ldots, h$ of the first part, but for simplicity of notation we do not make this explicit, since it is clear from context.

with $V^{mix}(\overrightarrow{Z})[\overrightarrow{H}]$, $V^{pure}(\overrightarrow{Z})[\overrightarrow{H}]$, and $V^{sym}(\overrightarrow{Z})[\overrightarrow{H}]$ defined as follows:

$$V^{mix}(\overrightarrow{Z})[\overrightarrow{H}] = \begin{pmatrix} H_{-h}L_1^{-h}(\overrightarrow{Z}) \\ \vdots \\ H_{-h}L_{\ell_{-h}}^{-h}(\overrightarrow{Z}) \\ \vdots \\ H_{-1}L_1^{-1}(\overrightarrow{Z}) \\ \vdots \\ H_{-1}L_{\ell_{-1}}^{-1}(\overrightarrow{Z}) \\ H_1L_1^1(\overrightarrow{Z}) \\ \vdots \\ H_1L_{\ell_1}^1(\overrightarrow{Z}) \\ \vdots \\ H_hL_1^h(\overrightarrow{Z}) \\ \vdots \\ H_hL_{\ell_h}^h(\overrightarrow{Z}) \end{pmatrix} \qquad V^{pure}(\overrightarrow{Z})[\overrightarrow{H}] = \begin{pmatrix} H_{h+1}L_1^{h+1}(\overrightarrow{Z}) \\ \vdots \\ H_{h+1}L_{\ell_{h+1}}^{h+1}(\overrightarrow{Z}) \\ \vdots \\ H_gL_1^g(\overrightarrow{Z}) \\ \vdots \\ H_gL_{\ell_g}^g(\overrightarrow{Z}) \end{pmatrix}$$

$$V^{sym}(\overrightarrow{Z})[\overrightarrow{H}] = \begin{pmatrix} H_{g+1}L_1^{g+1}(\overrightarrow{Z}) \\ \vdots \\ H_{g+1}L_{\ell_{g+1}}^{g+1}(\overrightarrow{Z}) \\ \vdots \\ H_kL_1^k(\overrightarrow{Z}) \\ \vdots \\ H_kL_{\ell_k}^k(\overrightarrow{Z}) \end{pmatrix}$$

In order to illustrate the above definitions, we give a simple example of a quadratic function and its border vector representation. Let the quadratic function $\mathcal{Q}(\overrightarrow{Z})[\overrightarrow{H}]$ be given by $\mathcal{Q}(\overrightarrow{Z})[\overrightarrow{H}] = H_1^T * H_1 + H_1 * H_1^T + H_2 * H_2^T + H_3^T * H_3 + H_4 * H_4$, where H_1, H_2, and H_3 are not symmetric and $H_4 = H_4^T$. The symbol $*$ means any expression that does not contain H_i. For this quadratic, the border vector has the following structure:

$$V[\overrightarrow{H}] = \begin{pmatrix} \left.\begin{matrix} H_1 \\ H_1^T \end{matrix}\right\} & \text{Mixed} \\ \\ \left.\begin{matrix} H_2^T \\ H_3 \end{matrix}\right\} & \text{Pure} \\ \\ \left.H_4\right\} & \text{Symmetric} \end{pmatrix}$$

Note that this representation of $\mathcal{Q}(\overrightarrow{Z})[\overrightarrow{H}]$ might require simple relabeling of variables. For example, if $\mathcal{Q}[\{H, K\}] = H^TAH + KBK^T$, then $H_1 = H$, $H_2 = K^T$

and

$$V[\vec{H}] = V^{pure}[\vec{H}] = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}. \tag{7.6}$$

Indeed, the representations with only $V^{pure}[\vec{H}]$ give precisely the hereditary[14] $\mathcal{Q}$. Allowing simple relabeling of variables increases the scope of such representations to include all cases like those in example (7.6).

Definition 7.2 (Block Linearly Dependent Coefficients)**.** The border $V(\vec{Z})[\vec{H}]$ in (7.5) has **block linearly dependent coefficients** if for some i the functions $L^i_j(\vec{Z})$ for $j = 1, \ldots, \ell_i$ are linearly dependent, otherwise the border vector $V(\vec{Z})[\vec{H}]$ has **block linearly independent coefficients**.

The "block" nature of the definition above is because we shall often consider separately the set

$$\vec{L}^i(\vec{Z}) := \{L^i_1(\vec{Z}), \ldots, L^i_{\ell_i}(\vec{Z})\}$$

for each $i \in \mathcal{I}$.

7.2. Substituting Matrices for Indeterminates

In this section we discuss the substitution of matrices for indeterminates and give some definitions. Let $\vec{Z} = \{Z_1, \ldots, Z_v\}$ be all indeterminates (variables) occurring in whatever noncommutative rational functions $\Gamma(\vec{Z})$ and constraints $\mathcal{G}$ we are studying. If these indeterminates are replaced by matrices we must be careful to replace them by tuple of matrices $\vec{\mathcal{Z}} := \{\mathcal{Z}_1, \ldots, \mathcal{Z}_v\}$ of **sizes**

$$\vec{\mathcal{Z}}^{\#} := \{m_1 \times n_1, \ldots, m_v \times n_v\}$$

compatible with the function $\Gamma(\vec{Z})$ and the constraints $\mathcal{G}$. Let C^{dim} denote the set of all compatible dimensions. A partial order $\succeq$ on C^{dim}, denoted by $\vec{\mathcal{Z}}^{\#} \succeq \vec{\mathcal{Z}}^{a\#}$, is given by

$$\{m_1 \geq m^a_1, n_1 \geq n^a_1, \ldots, m_v \geq m^a_v, n_v \geq n^a_v\},$$

and if strict inequality holds in every entry we write $\vec{\mathcal{Z}}^{\#} \succ \vec{\mathcal{Z}}^{a\#}$. Once a size $\Delta \in C^{dim}$ has been selected we let $\mathcal{M}_\Delta$ denote the set of all v tuples of matrices of size Δ. Moreover, if $\mathcal{G}$ is a Symbolic Inequality Domain, then let $\mathcal{M}(\mathcal{G})$ (resp. $\mathcal{M}_\Delta(\mathcal{G})$) denote the set of all matrices meeting the constraints defining $\mathcal{G}$ (resp. and lying in $\mathcal{M}_\Delta$). Often we suppress the subscript Δ because its presence is clear from context.

[14]Note that in our definition of hereditary the variables H_j can not be constrained to be symmetric.

Definition 7.3 (Openness Property)**.** The domain $\mathcal{G}$ has the Openness Property provided that there is a size Δ_0 in C^{dim} with the property that when indeterminates are replaced by matrices with size $\Delta \succeq \Delta_0$, then the set of matrices $\mathcal{M}_\Delta(\mathcal{G})$ is contained in the closure of the interior of $\mathcal{M}_\Delta(\mathcal{G})$.

8. Theorems on Convexity and Positivity

8.1. Main Result on Convexity: Theorem 8.2

Theorem 8.2, which follows, gives a test which can in fact be implemented with a noncommutative Gröbner basis algorithm ([Mor86, Mor94, Frö97]). The linear dependence check is purely algebraic and can be performed automatically by computer (software willing). We have not considered seriously the practicality of the Openness Property. However, in all the examples we have done, it is obvious that the set $\mathcal{G}$ obtained satisfy it. Now we set down a class of quadratic functions for which the theory in this paper works. The definition also serves as a reminder of Theorem 3.3 on LDL^T decompositions.

Definition 8.1 (Nice Quadratic on a Symbolic Inequality Domain $\mathcal{G}$)**.** A noncommutative symmetric function $\mathcal{Q}(\vec{Z})[\vec{H}]$, which is rational in $\vec{Z}$ and quadratic in $\vec{H}$, can be always put in the form $V(\vec{Z})[\vec{H}]^T\ M_\mathcal{Q}(\vec{Z})\ V(\vec{Z})[\vec{H}]$ with $V(\vec{Z})[\vec{H}]$ as in (7.5). Suppose that the coefficient matrix $M_\mathcal{Q}(\vec{Z})$ has a noncommutative $L(\vec{Z})$ $D(\vec{Z})\ L(\vec{Z})^T$ decomposition (we may have applied some permutation) with $D(\vec{Z})$ a diagonal matrix (no matrix R in Theorem 3.3, unless all entries of the matrix R are identically zero) having entries $D_j(\vec{\mathcal{Z}})$, for $j = 1, \dots, r-1$, each of which are zero or invertible matrices whenever tuple of matrices $\vec{\mathcal{Z}}$ of compatible dimension in $\mathcal{M}_\Delta(\mathcal{G})$ for large enough Δ are substituted for $\vec{Z}$, then we call $\mathcal{Q}(\vec{Z})[\vec{H}]$ a **nice quadratic**.

Theorem 8.2 (A Checkable Necessary and Sufficient Condition for Convexity)**.**
Assumptions: *Define $\vec{Z} = \{\vec{A}, \vec{X}\}$ where X_j may or may not be constrained to be symmetric. Let $\Gamma(\vec{Z})$ be any noncommutative symmetric rational function, whose Hessian $\mathcal{H}\Gamma(\vec{Z})[\vec{H}]$ is a nice quadratic, satisfying the following two conditions:*

- i. *the function $\Gamma(\vec{Z})$ is matrix convex for $\vec{Z}$ on a Symbolic Inequality Domain $\mathcal{G}$ satisfying the Openness Property for some big enough Δ_0;*
- ii. *the border vector $V(\vec{Z})[\vec{H}]$ of the Hessian $\mathcal{H}\Gamma(\vec{Z})[\vec{H}]$ has block linearly independent coefficients.*

Conclusion: *The following statements are equivalent:*

- a. *when tuple of matrices $\vec{\mathcal{Z}}$ in $\mathcal{M}_\Delta(\mathcal{G})$ of compatible dimension $\Delta \succeq \Delta_0$ are substituted into the Hessian $\mathcal{H}\Gamma$, we obtain $\mathcal{H}\Gamma(\vec{\mathcal{Z}})[\vec{\mathcal{H}}] \geq 0$ for all $\vec{\mathcal{H}}$.*

b. *for all tuple of matrices* $\vec{\mathcal{Z}}$ *in the closure of* $\mathcal{M}_\Delta(\mathcal{G})$ *the diagonal entries of the* $L(\vec{\mathcal{Z}})\ D(\vec{\mathcal{Z}})\ L(\vec{\mathcal{Z}})^T$ *decomposition are positive semidefinite matrices (that is* $D(\vec{\mathcal{Z}}) \geq 0$*) provided that* $D(\vec{\mathcal{Z}})$ *is defined.*

Proof. That (b) implies (a) is easy to prove and follows from Theorem 6.2. That (a) implies (b) is difficult to prove and follows from:

- the next Theorem 8.3 which applies only to quadratic functions and proves under appropriate hypotheses that $\mathcal{H}\Gamma(\vec{\mathcal{Z}})[\vec{\mathcal{H}}] \geq 0$ implies $M_{\mathcal{H}\Gamma}(\vec{\mathcal{Z}}) \geq 0$ for $\vec{\mathcal{Z}}$ defined as in (a) above;
- and that $M_{\mathcal{H}\Gamma}(\vec{\mathcal{Z}}) \geq 0$ implies $D(\vec{\mathcal{Z}}) \geq 0$, which is true since $M_{\mathcal{H}\Gamma}(\vec{\mathcal{Z}}) = L(\vec{\mathcal{Z}})D(\vec{\mathcal{Z}})L(\vec{\mathcal{Z}})^T$ with $L(\vec{\mathcal{Z}})$ an invertible matrix.[15]

□

8.2. Main Result on Quadratic Functions: Theorem 8.3

This section gives results about quadratic functions. The main result is Theorem 8.3 that concerns positivity of a noncommutative rational function $\mathcal{Q}(\vec{Z})[\vec{H}]$ which is quadratic in $\vec{H}$. The statement of this theorem is presented in this section and its proof is finished in Section 10.

Theorem 8.3 (Main Result on Quadratic Functions)**.**

Assumptions*: Let* $\vec{H} := \{H_{-h}, \ldots, H_k\}$ *be defined as in (7.4). Consider a noncommutative rational function* $\mathcal{Q}(\vec{Z})[\vec{H}]$ *which is a quadratic*[16] *in the variables* $\vec{H}$ *on a Symbolic Inequality Domain* $\mathcal{G}$*. Write* $\mathcal{Q}(\vec{Z})[\vec{H}]$ *in the form* $\mathcal{Q}(\vec{Z})[\vec{H}] = V(\vec{Z})[\vec{H}]^T M_{\mathcal{Q}(\vec{Z})} V(\vec{Z})[\vec{H}]$*. Suppose that the following two conditions hold:*

i. *the Symbolic Inequality Domain* $\mathcal{G}$ *satisfies the Openness Property for some big enough* Δ_0*;*

ii. *the border vector* $V(\vec{Z})[\vec{H}]$ *of the quadratic function* $\mathcal{Q}(\vec{Z})[\vec{H}]$ *has block linearly independent coefficients.*

Conclusion*: The following statements are equivalent:*

a. *when tuple of matrices* $\vec{\mathcal{Z}}$ *in* $\mathcal{M}_\Delta(\mathcal{G})$ *of compatible dimension* $\Delta \succeq \Delta_0$ *are substituted into* $\mathcal{Q}$*, we obtain* $\mathcal{Q}(\vec{\mathcal{Z}})[\vec{\mathcal{H}}]$ *is a positive semidefinite matrix for each tuple of matrices* $\vec{\mathcal{H}}$*;*

b. *we have* $M_{\mathcal{Q}(\vec{\mathcal{Z}})} \geq 0$ *for all* $\vec{\mathcal{Z}}$ *in the closure of* $\mathcal{M}_\Delta(\mathcal{G})$ *on which* $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ *is defined.*

---

[15] $L(\vec{\mathcal{Z}})$ is an invertible matrix since it is lower triangular with ones on its diagonal.

[16] We emphasize that $\mathcal{Q}(\vec{Z})[\vec{H}]$ is not restricted to be a **nice quadratic**.

Proof. Clearly (b) implies (a). The hard part is (a) implies (b). The proof of this result consumes the following Section 9 and is finalized in Section 10. □

9. Theorems Concerning Quadratic Functions

Before beginning the proof of Theorem 8.3 in earnest, we sketch some of the ideas for the simplest type of quadratic functions. Section 9, which consist of Section 9.1 and Section 9.2, concerns primarily a matrix valued quadratic function $\mathcal{Q}[\vec{\mathcal{H}}]$ of tuple $\vec{\mathcal{H}}$ of $n \times n$ matrices; there is no dependence on symbolic variables or on variables $\vec{Z}$. In Section 9.1, we treat quadratic functions which are hereditary in the variables $\vec{\mathcal{H}}$.

Later, in Section 10, we begin to combine the matrix results of Section 9.1 with symbolic variables, and also we study quadratic functions of $\vec{H}$ which also depend on $\vec{Z}$. We reemphasize that the function $\mathcal{Q}(\vec{Z})[\vec{H}]$ is quadratic in $\vec{H}$, but it need not be quadratic in $\vec{Z}$.

9.1. Some Ideas of the Proof

This section gives a very special case of Theorem 8.3 in order to illustrate a few of the ideas involved and expose the readers to easy cases of the notation. This tutorial proof takes up Section 9.1 and then after that the fully general proof begins.

The special case we consider is that of a hereditary quadratic function $\mathcal{Q}[\vec{\mathcal{H}}]$. To assume that $\mathcal{Q}[\vec{\mathcal{H}}]$ is a hereditary function is equivalent to imposing that $\vec{\mathcal{H}}$ has the special form $\vec{\mathcal{H}} := \{\mathcal{H}_{h+1}, \dots, \mathcal{H}_g\}$, which in our notation says that $\{\mathcal{H}_i\}_{i=-h}^{h}$ and $\{\mathcal{H}_j\}_{j=g+1}^{k}$ are missing in $\vec{\mathcal{H}} := \{\mathcal{H}_{-h}, \dots, \mathcal{H}_{-1}, \mathcal{H}_1, \dots, \mathcal{H}_h, \mathcal{H}_{h+1}, \dots, \mathcal{H}_g, \mathcal{H}_{g+1}, \dots, \mathcal{H}_k\}$. Note that we are treating a purely quadratic function $\mathcal{Q}[\vec{\mathcal{H}}]$, in other words, $\mathcal{Q}(\vec{\mathcal{Z}})[\vec{\mathcal{H}}]$ has no $\vec{\mathcal{Z}}$ dependence. This special type of $\mathcal{Q}[\vec{\mathcal{H}}]$ has the following representation

$$\mathcal{Q}[\vec{\mathcal{H}}] = V^{pure}[\vec{\mathcal{H}}]^T M_{\mathcal{Q}} V^{pure}[\vec{\mathcal{H}}],$$

where $V^{pure}[\vec{\mathcal{H}}]$ is defined as follows

$$V^{pure}[\vec{\mathcal{H}}] = \begin{pmatrix} \mathcal{H}_{h+1}L_1^{h+1} \\ \vdots \\ \mathcal{H}_{h+1}L_{\ell_{h+1}}^{h+1} \\ \vdots \\ \mathcal{H}_g L_1^g \\ \vdots \\ \mathcal{H}_g L_{\ell_g}^g \end{pmatrix}, \tag{9.1}$$

with each L_j^i being a fixed matrix, that is, they do not depend on matrices $\vec{\mathcal{Z}}$.

The main result of this section, Proposition 9.1, is easy to prove, and serves as an introduction to the ideas of the proof of the main Theorem 8.3.

Proposition 9.1 (Necessary Condition for Positivity). *Let $\mathcal{Q}[\vec{\mathcal{H}}]$ be a hereditary quadratic function of tuple $\vec{\mathcal{H}} = \{\mathcal{H}_j\}_{j=h+1}^g$, where each matrix $\mathcal{H}_j$ has dimension $n \times n$. Also assume that this quadratic has a border vector of the type defined in* (9.1). *Suppose that $\mathcal{Q}[\vec{\mathcal{H}}]$ is a positive semidefinite matrix for each tuple $\vec{\mathcal{H}}$, then either*

i. *the matrix $M_{\mathcal{Q}}$ is positive semidefinite*

or

ii. *there is an integer $d \in [h+1, g]$ and real valued functions*

$$\lambda_j : \mathbb{R}^n \to \mathbb{R}, \quad j = 1, \ldots, \ell_d,$$

such that

$$\sum_{j=1}^{\ell_d} \lambda_j(x) L_j^d x = 0, \quad \text{for } x \in \mathbb{R}^n.$$

We now define some sets that will be used throughout the paper, and especially in the proof of Proposition 9.1 above. Let each L_j^i be fixed matrices of dimension $n \times n$. For a given $x \in \mathbb{R}^n$, define the set $\mathcal{R}_{\vec{L}^i}^{pure,x}$ to be

$$\mathcal{R}_{\vec{L}^i}^{pure,x} := \left\{ \begin{pmatrix} \mathcal{H}_i L_1^i x \\ \vdots \\ \mathcal{H}_i L_{\ell_i}^i x \end{pmatrix} : \text{all } \mathcal{H}_i \in \mathbb{R}^{n \times n} \right\}, \tag{9.2}$$

and the set $\mathcal{R}_{\vec{L}^i}^{pure}$ to be

$$\mathcal{R}_{\vec{L}^i}^{pure} := \left\{ \mathcal{R}_{\vec{L}^i}^{pure,x} : \text{all } x \in \mathbb{R}^n \right\}.$$

Define also the set $\mathcal{S}_{\vec{L}}$ to be

$$\mathcal{S}_{\vec{L}} := \begin{pmatrix} \mathcal{R}^{pure}_{\vec{L}^{h+1}} \\ \vdots \\ \mathcal{R}^{pure}_{\vec{L}^{g}} \end{pmatrix} = \left\{ \begin{pmatrix} \mathcal{H}_{h+1} L_1^{h+1} x \\ \vdots \\ \mathcal{H}_{h+1} L_{\ell_{h+1}}^{h+1} x \\ \vdots \\ \mathcal{H}_g L_1^g x \\ \vdots \\ \mathcal{H}_g L_{\ell_g}^g x \end{pmatrix} : \begin{array}{l} \text{all } \mathcal{H}_{h+1}, \ \ldots, \ \mathcal{H}_g \ \in \ \mathbb{R}^{n\times n} \\ \text{and all } \ x \in \mathbb{R}^n \end{array} \right\}.$$

The Proof of Proposition 9.1 follows immediately from Lemma 9.2 and Proposition 9.3, which we now present.

Lemma 9.2. *Let $\mathcal{Q}[\vec{\mathcal{H}}]$ be a hereditary quadratic function of tuple $\vec{\mathcal{H}}$ of matrices of dimension $n \times n$. Also assume that this quadratic has a border vector of the type defined in (9.1). The function $\mathcal{Q}[\vec{\mathcal{H}}]$ is positive semidefinite for all $\vec{\mathcal{H}}$ implies $M_{\mathcal{Q}} \geq 0$, provided that for some y the space $\mathcal{R}^{pure,y}_{\vec{L}^i}$ fills out the whole space $\mathbb{R}^{n\ell_i}$ for all $i = h+1, \ldots, g$.*

Proof. Let $\mathcal{Q}[\vec{\mathcal{H}}]$ be positive semidefinite. By definition this implies that $y^T \mathcal{Q}[\vec{\mathcal{H}}] y \geq 0$ for all $y \in \mathbb{R}^n$ and all $\{\mathcal{H}_j\}_{j=h+1}^g \in \mathbb{R}^{n\times n}$. Therefore

$$y^T \mathcal{Q}[\vec{\mathcal{H}}] y = y^T V[\vec{\mathcal{H}}]^T M_{\mathcal{Q}} V[\vec{\mathcal{H}}] y = w^T M_{\mathcal{Q}} w \geq 0$$

for all $w = V[\vec{\mathcal{H}}] y \in \mathbb{R}^{n(\ell_{h+1}+\cdots+\ell_g)}$ and all $\{\mathcal{H}_j\}_{j=h+1}^g \in \mathbb{R}^{n\times n}$. Now it suffices to prove that for some y all vectors of the form w equals $\mathbb{R}^{n(\ell_{h+1}+\cdots+\ell_g)}$. But this condition is directly satisfied from the assumption that the space $\mathcal{R}^{pure,y}_{\vec{L}^i}$ fills out the whole space $\mathbb{R}^{n\ell_i}$ for all $i = h+1, \ldots, g$. □

Proposition 9.3. *For a given $x \in \mathbb{R}^n$, let $\mathcal{R}^{pure,x}_{\vec{L}^i}$ be defined as in* (9.2). *The following holds:*

i. *If $\mathcal{R}^{pure,x}_{\vec{L}^i}$* **is** *all of $\mathbb{R}^{n\ell_i}$, then $L_1^i x, L_2^i x, \ldots, L_{\ell_i}^i x$ are linearly independent vectors.*

ii. *If $\mathcal{R}^{pure,x}_{\vec{L}^i}$* **is not** *all of $\mathbb{R}^{n\ell_i}$, then $L_1^i x$, $L_2^i x$, $\ldots$, $L_{\ell_i}^i x$ are linearly dependent vectors, and consequently there exist nontrivial scalar functions $\lambda_j(x)$, that may depend on x, such that*

$$\lambda_1(x) L_1^i x + \lambda_2(x) L_2^i x + \cdots + \lambda_{\ell_i}(x) L_{\ell_i}^i x = 0. \tag{9.3}$$

An obvious consequence of the above fact is that if $\mathcal{R}^{pure}_{\vec{L}^i} = \{\mathcal{R}^{pure,x}_{\vec{L}^i} :$ all $x \in \mathbb{R}^n\}$, is not all of $\mathbb{R}^{n\ell_i}$, then for each x, $\mathcal{R}^{pure,x}_{\vec{L}^i}$ is not all of $\mathbb{R}^{n\ell_i}$, and thus equation (9.3) *holds for all x.*

Proof. For a given $x \in \mathbb{R}^n$, let $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ be all of $\mathbb{R}^{n\ell_i}$. Suppose $L^i_1 x, L^i_2 x, \ldots, L^i_{\ell_i} x$ are linearly dependent vectors. Without loss of generality, let $L^i_1 x = \sum_{j=2}^{\ell_i} \lambda_j(x) L^i_j x$, where $\lambda_j(x)$ are scalar functions. Define $s_j = \mathcal{H}_i L^i_j x$, then $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ becomes

$$\mathcal{R}^{pure,x}_{\overrightarrow{L}^i} = \left\{ \begin{pmatrix} \lambda_2(x)s_2 + \cdots + \lambda_s(x)s_{\ell_i} \\ s_2 \\ \vdots \\ s_{\ell_i} \end{pmatrix} : \text{ some } s_j \in \mathbb{R}^n \right\}$$

which can not possibly be $\mathbb{R}^{n\ell_i}$. This fact contradicts our assumption on $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ being all of $\mathbb{R}^{n\ell_i}$, thus $L^i_1 x, L^i_2 x, \ldots, L^i_{\ell_i} x$ must be a linearly independent set of vectors.

To prove (ii), suppose for a given $x \in \mathbb{R}^n$ the vectors $L^i_1 x, L^i_2 x, \ldots, L^i_{\ell_i} x$ are linearly independent. Let

$$y = \begin{pmatrix} w_1 \\ \vdots \\ w_{\ell_i} \end{pmatrix}$$

be any vector in $\mathbb{R}^{n\ell_i}$. Then we can choose $\mathcal{H}_i \in \mathbb{R}^{n\times n}$ with the property that $w_1 = \mathcal{H}_i L^i_1 x$, $w_2 = \mathcal{H}_i L^i_2 x$, $\ldots$, $w_{\ell_i} = \mathcal{H}_i L^i_{\ell_i} x$. Thus $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ is all of $\mathbb{R}^{n\ell_i}$. $\square$

What we have demonstrated is only the beginning of the proof of Theorem 8.3 for a hereditary quadratic function. Next, we must show that the λ_j do not depend on x. For the particular case we have been treating, there are several ways to do this, but they do not all work for the general case of interest. The method we use later to prove that the λ_j are independent of x uses the fact that the quadratic function depends on the variables $\overrightarrow{Z}$ (see Theorem 10.10 in Section 10). Another difficulty is that the sets analogous to $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ never equal the whole space for the case where $\mathcal{Q}$ is non-hereditary or $\overrightarrow{\mathcal{H}}$ contains symmetric elements. Fortunately these sets have co-dimension which depends only on the dimension of the coefficient matrix $M_{\mathcal{Q}}$ and does not depend on the dimension of the matrices contained in the tuple $\overrightarrow{\mathcal{Z}}$ substituted for $\overrightarrow{Z}$ (See Proposition 9.8). We combine this fact about co-dimension with the algebraic dependence of the functions $\mathcal{Q}(\overrightarrow{Z})$ and $L^i_j(\overrightarrow{Z})$ on $\overrightarrow{Z}$ to complete the proof of Theorem 8.3 in Section 10.

9.2. The Range of the Border Vector of a Matrix Quadratic Function

Earlier in Section 9.1, a necessary condition for positivity was presented in Proposition 9.3 for a particular type of quadratic function. The key was a linear independence property guaranteeing that the space $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ is all $\mathbb{R}^{n\ell_i}$, that means, the co-dimension of the space $\mathcal{R}^{pure,x}_{\overrightarrow{L}^{i}}$ equals zero. Unfortunately, this only characterizes the unconstrained part (the second part) of $\overrightarrow{H}$ defined in (7.4). Section 9.2 gives similar conditions on the other two parts of $\overrightarrow{H}$, the pairwise symmetric part (the first part) and the symmetric part (the third part). General quadratic functions are treated in Proposition 9.4, and the key property is a uniform bound on certain co-dimensions. Again, as in Section 9.1, we study quadratic functions $\mathcal{Q}(\overrightarrow{Z})[\overrightarrow{H}]$ with no $\overrightarrow{Z}$ dependence.

First define $\mathcal{R}^{sym,x}_{\overrightarrow{L}^{i}}$ and $\mathcal{R}^{mix,x}_{\overrightarrow{L}^{i}}$ to be

$$\mathcal{R}^{sym,x}_{\overrightarrow{L}^{i}} := \left\{ \begin{pmatrix} \mathcal{H}_i L_1^i x \\ \vdots \\ \mathcal{H}_i L_{\ell_i}^i x \end{pmatrix} : \text{all } \mathcal{H}_i = \mathcal{H}_i^T \in \mathbb{R}^{n\times n} \right\}, \tag{9.4}$$

$$\mathcal{R}^{mix,x}_{\overrightarrow{L}^{i}} := \left\{ \begin{pmatrix} \mathcal{H}_{-i} L_1^{-i} x \\ \vdots \\ \mathcal{H}_{-i} L_{\ell_{-i}}^{-i} \\ \mathcal{H}_i L_1^i x \\ \vdots \\ \mathcal{H}_i L_{\ell_i}^i x \end{pmatrix} : \text{all } \mathcal{H}_{-i} = \mathcal{H}_i^T \in \mathbb{R}^{n\times n} \right\}. \tag{9.5}$$

Define also $\mathcal{R}^{sym}_{\overrightarrow{L}^{i}}$ and $\mathcal{R}^{mix}_{\overrightarrow{L}^{i}}$ to be

$$\mathcal{R}^{sym}_{\overrightarrow{L}^{i}} := \left\{ \mathcal{R}^{sym,x}_{\overrightarrow{L}^{i}} : \text{all } x \in \mathbb{R}^n \right\},$$

$$\mathcal{R}^{mix}_{\overrightarrow{L}^{i}} := \left\{ \mathcal{R}^{mix,x}_{\overrightarrow{L}^{i}} : \text{all } x \in \mathbb{R}^n \right\}.$$

The following Proposition 9.4 introduces our main results concerning $\mathcal{R}^{sym,x}_{\overrightarrow{L}^{i}}$ and $\mathcal{R}^{mix,x}_{\overrightarrow{L}^{i}}$, and also summarizes similar results concerning $\mathcal{R}^{pure,x}_{\overrightarrow{L}^{i}}$ given in Proposition 9.3.

Proposition 9.4. *For a given $x \in \mathbb{R}^n$, let $\mathcal{R}^{pure,x}_{\overrightarrow{L}^{i}}$, $\mathcal{R}^{sym,x}_{\overrightarrow{L}^{i}}$ and $\mathcal{R}^{mix,x}_{\overrightarrow{L}^{i}}$ be defined as in* (9.2) *and* (9.4–9.5)*. The following holds:*

i. *If $\mathcal{R}^{pure,x}_{\overrightarrow{L}^{i}}$* **is** *all of $\mathbb{R}^{n\ell_i}$, then $L_1^i x, L_2^i x, \ldots, L_{\ell_i}^i x$ are linearly independent vectors.*

ii. *If $\mathcal{R}^{pure,x}_{\overrightarrow{L}^{i}}$* **is not** *all of $\mathbb{R}^{n\ell_i}$ (resp. If $\mathcal{R}^{sym,x}_{\overrightarrow{L}^{i}}$ has co-dimension in $\mathbb{R}^{n\ell_i}$ greater than $\ell_i[\ell_i - 1]/2$), then $L_1^i x$, $L_2^i x$, $\ldots$, $L_{\ell_i}^i x$ are linearly dependent vectors, and consequently there exist nontrivial scalar functions $\lambda_j(x)$, that may depend on x, such that*

$$\lambda_1(x) L_1^i x + \lambda_2(x) L_2^i x + \cdots + \lambda_{\ell_i}(x) L_{\ell_i}^i x = 0. \tag{9.6}$$

iii. *If* $\mathcal{R}^{mix,x}_{\overrightarrow{L}^i}$ *has co-dimension in* $\mathbb{R}^{n(\ell_i+\ell_{-i})}$ *greater than* $\ell_i\ell_{-i}$*, then either* $L^i_1x,\ L^i_2x,\ \ldots,\ L^i_{\ell_i}x$ *or* $L^{-i}_1x,\ L^{-i}_2x,\ \ldots,\ L^{-i}_{\ell_{-i}}x$ *are linearly dependent vectors, and consequently there exist nontrivial scalar functions* $\lambda_j(x)$*, that may depend on* x*, such that either*

$$\lambda_1(x)L^i_1x + \lambda_2(x)L^i_2x + \cdots + \lambda_{\ell_i}(x)L^i_{\ell_i}x = 0 \tag{9.7}$$

or

$$\lambda_1(x)L^{-i}_1x + \lambda_2(x)L^{-i}_2x + \cdots + \lambda_{\ell_{-i}}(x)L^{-i}_{\ell_{-i}}x = 0. \tag{9.8}$$

Proof. The results concerning $\mathcal{R}^{pure,x}_{\overrightarrow{L}^i}$ were proved in Proposition 9.3.

First we treat the case where the $\mathcal{H}_i$ are constrained to be symmetric. If (9.6) fails, then $L^i_1x, \ldots, L^i_{\ell_i}x$ are linearly independent; thus we may use Lemma 9.5 below to obtain that $\mathcal{R}^{sym,x}_{\overrightarrow{L}^i}$ is a space of co-dimension equal to $\ell_i(\ell_i-1)/2$. This contradicts the assumption that $\mathcal{R}^{sym,x}_{\overrightarrow{L}^i}$ has co-dimension in $\mathbb{R}^{n\ell_i}$ greater than $\ell_i(\ell_i-1)/2$. This proves part (ii) of Proposition 9.4.

The proof of part (iii) follows the same line. If both (9.7) and (9.8) fail, then both $L^i_1x, \ldots, L^i_{\ell_i}x$ and $L^{-i}_1x,\ L^{-i}_2x,\ \ldots,\ L^{-i}_{\ell_{-i}}x$ are linearly independent vectors; thus Lemma 9.6 below implies that $\mathcal{R}^{mix,x}_{\overrightarrow{L}^i}$ is a space of co-dimension equal to $\ell_i\ell_{-i}$, contradicting the assumption that $\mathcal{R}^{mix,x}_{\overrightarrow{L}^i}$ has co-dimension greater than $\ell_i\ell_{-i}$. This completes the proof of Proposition 9.4. □

Now we present the Lemmas required in the proof of Proposition 9.4. We use H instead of $\mathcal{H}$ to stand for a matrix in $\mathbb{R}^{n\times n}$ in Lemma 9.5 and Lemma 9.6. This makes the rather involved formulas easier to read.

Lemma 9.5. *For linearly independent vectors* $v_1, \ldots, v_\ell \in \mathbb{R}^n$ *the space* S *defined by*

$$S = \left\{ \begin{pmatrix} Hv_1 \\ \vdots \\ Hv_\ell \end{pmatrix} :\ \textit{all } H = H^T \in \mathbb{R}^{n\times n} \right\}$$

is a subspace in $\mathbb{R}^{n\ell}$ *with co-dimension* $\ell(\ell-1)/2$.

Proof. Define invertible matrices $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{\ell\times\ell}$ by

$$\begin{pmatrix} v_1| & \cdots & |v_\ell \end{pmatrix} = P \begin{pmatrix} I \\ 0 \end{pmatrix} Q,$$

where I is the identity matrix with dimension ℓ and $\begin{pmatrix} v_1| & \cdots & |v_\ell \end{pmatrix}$ denotes the matrix whose columns are $v_1, \ldots, v_\ell$. (Note that the hypotheses of this theorem imply $n > \ell$.) The dimension of the space S is

$$\begin{aligned}
\dim(S) &= \dim\left(\left\{\begin{pmatrix} Hv_1 \\ \vdots \\ Hv_\ell \end{pmatrix} : \text{ all } H = H^T \in \mathbb{R}^{n\times n}\right\}\right) \\
&= \dim\left(\left\{H\begin{pmatrix} v_1 | & \dots & |v_\ell \end{pmatrix} : \text{ all } H = H^T \in \mathbb{R}^{n\times n}\right\}\right) \\
&= \dim\left(\left\{HP\begin{pmatrix} I \\ 0 \end{pmatrix} Q : \text{ all } H = H^T \in \mathbb{R}^{n\times n}\right\}\right) \\
&= \dim\left(\left\{HP\begin{pmatrix} I \\ 0 \end{pmatrix} : \text{ all } H = H^T \in \mathbb{R}^{n\times n}\right\}\right) \\
&= \dim\left(\left\{P^T HP\begin{pmatrix} I \\ 0 \end{pmatrix} : \text{ all } H = H^T \in \mathbb{R}^{n\times n}\right\}\right) \\
&= \dim\left(\left\{\tilde{H}\begin{pmatrix} I \\ 0 \end{pmatrix} : \text{ all } \tilde{H} = \tilde{H}^T \in \mathbb{R}^{n\times n}\right\}\right) \\
&= n\ell - \ell(\ell-1)/2.
\end{aligned}$$

Thus the co-dimension equals $\ell(\ell-1)/2$. The last step above was a consequence of the following argument. Partition

$$\tilde{H} = \begin{array}{c} \\ \ell \\ n-\ell \end{array}\begin{array}{c} \begin{array}{cc} \ell & n-\ell \end{array} \\ \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \end{array}.$$

Then

$$\begin{aligned}
&\dim\left(\left\{\tilde{H}\begin{pmatrix} I \\ 0 \end{pmatrix} : \text{for all } \tilde{H} = \tilde{H}^T\right\}\right) \\
&= \dim\left(\left\{\begin{pmatrix} H_{11} \\ H_{21} \end{pmatrix} : \text{for all } H_{11} = H_{11}^T \in \mathbb{R}^{\ell\times\ell} \text{and } H_{21} \in \mathbb{R}^{(n-\ell)\times\ell}\right\}\right) \\
&= \dim\left(\left\{H_{11} : \text{for all } H_{11} = H_{11}^T \in \mathbb{R}^{\ell\times\ell}\right\}\right) + \\
&\qquad\qquad \dim\left(\left\{H_{21} : \text{for all } H_{21} \in \mathbb{R}^{(n-\ell)\times\ell)}\right\}\right) \\
&= \tfrac{\ell(\ell+1)}{2} + (n-\ell)\ell \\
&= n\ell - \ell(\ell-1)/2.
\end{aligned}$$

□

Lemma 9.6. *Suppose that $\{u_i\}_{i=1}^r$ and $\{v_j\}_{j=1}^s$ are two sets of linearly independent vectors in $\mathbb{R}^n$. (The set $\{u_i, v_j\}_{i,j}$ need not consist of linearly independent vectors.)*

Then the space S defined by

$$S = \left\{ \begin{pmatrix} Hu_1 \\ \vdots \\ Hu_r \\ H^T v_1 \\ \vdots \\ H^T v_s \end{pmatrix} : \textit{ for all } H \in \mathbb{R}^{n\times n} \right\}$$

is a subspace in $\mathbb{R}^{n(r+s)}$ with co-dimension rs.

Proof. Define invertible matrices $P_1 \in \mathbb{R}^{n\times n}$, $Q_1 \in \mathbb{R}^{r\times r}$, $P_2 \in \mathbb{R}^{n\times n}$, and $Q_2 \in \mathbb{R}^{s\times s}$ by

$$\begin{pmatrix} u_1| & \cdots & |u_r \end{pmatrix} = P_1 \begin{pmatrix} I_r \\ 0 \end{pmatrix} Q_1,$$

$$\begin{pmatrix} v_1| & \cdots & |v_s \end{pmatrix} = P_2 \begin{pmatrix} I_s \\ 0 \end{pmatrix} Q_2,$$

where I_r and I_s are the identity matrices with dimension r and s respectively. (Note that the hypotheses of this theorem imply $n > r$ and $n > s$.) The dimension of the space S is

$$\begin{aligned}
\dim(S) &= \dim \left(\left\{ \begin{pmatrix} Hu_1 \\ \vdots \\ Hu_r \\ H^T v_1 \\ \vdots \\ H^T v_s \end{pmatrix} : \text{ for all } H \in \mathbb{R}^{n\times n} \right\} \right) \\
&= \dim \big(\big\{ \begin{pmatrix} H \begin{pmatrix} u_1| & \ldots & |u_r \end{pmatrix} & | & H^T \begin{pmatrix} v_1| & \ldots & |v_s \end{pmatrix} \end{pmatrix} : \\
&\qquad \text{for all } H \in \mathbb{R}^{n\times n} \big\} \big) \\
&= \dim \left(\left\{ \left(HP_1 \begin{pmatrix} I_r \\ 0 \end{pmatrix} Q_1 \;\middle|\; H^T P_2 \begin{pmatrix} I_s \\ 0 \end{pmatrix} Q_2 \right) : \right.\right. \\
&\qquad \left.\left. \text{for all } H \in \mathbb{R}^{n\times n} \right\} \right) \\
&= \dim \left(\left\{ \left(HP_1 \begin{pmatrix} I_r \\ 0 \end{pmatrix} Q_1 \;\middle|\; H^T P_2 \begin{pmatrix} I_s \\ 0 \end{pmatrix} Q_2 \right) \begin{pmatrix} Q_1^{-1} & 0 \\ 0 & Q_2^{-1} \end{pmatrix} : \right.\right. \\
&\qquad \left.\left. \text{for all } H \in \mathbb{R}^{n\times n} \right\} \right) \\
&= \dim \left(\left\{ \left(HP_1 \begin{pmatrix} I_r \\ 0 \end{pmatrix} \;\middle|\; H^T P_2 \begin{pmatrix} I_s \\ 0 \end{pmatrix} \right) : \right.\right. \\
&\qquad \left.\left. \text{for all } H \in \mathbb{R}^{n\times n} \right\} \right)
\end{aligned}$$

$$= \dim\left(\left\{\left(P_2^T H P_1 \begin{pmatrix} I_r \\ 0 \end{pmatrix} \quad | \quad P_1^T H^T P_2 \begin{pmatrix} I_s \\ 0 \end{pmatrix} \right) : \text{for all } H \in \mathbb{R}^{n\times n}\right\}\right)$$

$$= \dim\left(\left\{\left(\tilde{H} \begin{pmatrix} I_r \\ 0 \end{pmatrix} \quad | \quad \tilde{H}^T \begin{pmatrix} I_s \\ 0 \end{pmatrix} \right) : \text{ for all } \tilde{H} \in \mathbb{R}^{n\times n}\right\}\right)$$

$$= n(r+s) - rs.$$

Thus the co-dimension equals to rs. The last step above follows from the following argument. Partition (assume $r < s$)

$$\tilde{H} = \begin{array}{r} \\ r \\ s-r \\ n-s \end{array} \begin{array}{c} \quad r \quad\; s-r \quad n-s \\ \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \end{array} . \tag{9.9}$$

Then

$$\dim\left(\left\{\left(\tilde{H} \begin{pmatrix} I_r \\ 0 \end{pmatrix} \middle| \tilde{H}^T \begin{pmatrix} I_s \\ 0 \end{pmatrix}\right) : \text{for all } \tilde{H} \in \mathbb{R}^{n\times n}\right\}\right)$$

$$= \dim\left(\left\{\begin{pmatrix} H_{11} & H_{11}^T & H_{21}^T \\ H_{21} & H_{12}^T & H_{22}^T \\ H_{31} & H_{13}^T & H_{23}^T \end{pmatrix} : \text{for all } \tilde{H} \in \mathbb{R}^{n\times n} \text{ partitioned as in (9.9)}\right\}\right)$$

$$= \dim(\{\tilde{H} : \text{for all } \tilde{H} \in \mathbb{R}^{n\times n}\})$$

$$- \dim\left(\left\{(H_{32} \quad H_{33}) : \text{for all } H_{32} \in \mathbb{R}^{(n-s)\times(s-r)} \text{ and } H_{33} \in \mathbb{R}^{(n-s)\times(n-s)}\right\}\right)$$

$$= n^2 - [(n-s)(s-r) + (n-s)(n-s)]$$

$$= n(r+s) - rs$$

□

We now present a lemma concerning co-dimensions, which will be used in the proof of Proposition 9.8.

Lemma 9.7. *Suppose that each S_i for $i = 1, \ldots, k$ is a subspace in $\mathbb{R}^{n_i}$ with co-dimension m_i, then the space $S = \begin{pmatrix} S_1 \\ \vdots \\ S_k \end{pmatrix}$ is a subspace in $\mathbb{R}^{n_1+\cdots+n_k}$ with co-dimension $m_1 + \cdots + m_k$.*

Proof. The space S is the direct sum of the spaces $\begin{pmatrix} 0 \\ \vdots \\ S_i \\ \vdots \\ 0 \end{pmatrix}$, each of which has dimension $n_i - m_i$. The dimension of S equals to the sum of the dimensions of S_i, or equivalently the co-dimension of S equals to the sum of the co-dimensions of S_i, which is $m_1 + \cdots + m_k$. □

Finally, we present Proposition 9.8, which introduces our main result concerning the co-dimension of the range of a border vector.

Proposition 9.8. *If there is an $x \in \mathbb{R}^n$ such that $L_1^i x, \ldots, L_{\ell_i}^i x$ are linearly independent vectors for every $i \in \mathcal{I}$, then the following space $\mathcal{R}_{\vec{L}}^{all,x}$ has co-dimension less than or equal to $t := t_1 + \cdots + t_k$, where*

$$t_i = \begin{cases} \ell_{-i}\ell_i & \text{for } i = 1, \ldots, h \\ 0 & \text{for } i = h+1, \ldots, g \\ \ell_i(\ell_i - 1)/2 & \text{for } i = g+1, \ldots, k \end{cases}$$

and $\mathcal{R}_{\vec{L}}^{all,x}$ is defined as

$$\begin{pmatrix} \mathcal{R}_{\vec{L}_1}^{mix,x} \\ \vdots \\ \mathcal{R}_{\vec{L}_h}^{mix,x} \\ \mathcal{R}_{\vec{L}_{h+1}}^{pure,x} \\ \vdots \\ \mathcal{R}_{\vec{L}_g}^{pure,x} \\ \mathcal{R}_{\vec{L}_{g+1}}^{sym,x} \\ \vdots \\ \mathcal{R}_{\vec{L}_k}^{sym,x} \end{pmatrix} = \left\{ \begin{pmatrix} \mathcal{H}_{-h} L_1^{-h} x \\ \vdots \\ \mathcal{H}_h L_{\ell_h}^h x \\ \mathcal{H}_{h+1} L_1^{h+1} x \\ \vdots \\ \mathcal{H}_g L_{\ell_g}^g x \\ \mathcal{H}_{g+1} L_1^{g+1} x \\ \vdots \\ \mathcal{H}_k L_{\ell_k}^k x \end{pmatrix} : \begin{array}{l} \text{for all } \mathcal{H}_i \in \mathbb{R}^{n \times n} \ (i \in \mathcal{I}), \text{ satisfying the constraints } \mathcal{H}_{-j} = \mathcal{H}_j^T \\ \text{for } j = 1, \ldots, h \text{ and } \mathcal{H}_j = \mathcal{H}_j^T \text{ for } j = g+1, \ldots, k \end{array} \right\}$$

Proof. It follows directly from Lemma 9.7, Lemma 9.6, and Lemma 9.5. □

10. Linear Dependence of Symbolic Functions

Let Δ_0 be a size sufficiently large that the domain $\mathcal{G}$ posses the Openness Property[17]. Let $\mathcal{N}_{\Delta_0}(\mathcal{G})$ be the subset of the set of all matrices meeting the inequality

[17]See definition 7.3 in Section 7.2.

constraints $\mathcal{M}(\mathcal{G})$ defined by $\mathcal{N}_{\Delta_0}(\mathcal{G}) := \bigcup_{\Delta \geq \Delta_0} \mathcal{M}_\Delta(\mathcal{G})$. Define also three subsets of $\mathcal{N}_{\Delta_0}(\mathcal{G})$, namely $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, by

$\mathcal{A} := \{\overrightarrow{\mathcal{Z}} \in \mathcal{N}_{\Delta_0}(\mathcal{G})$: the matrix $M_{\mathcal{Q}(\overrightarrow{\mathcal{Z}})}$ has less than or equal to t negative eigenvalues$\}$, where t is defined in Proposition 9.8.

$\mathcal{B} := \{\overrightarrow{\mathcal{Z}} \in \mathcal{N}_{\Delta_0}(\mathcal{G})$: for every x with compatible dimension, there exists $i \in \mathcal{I}$ such that the vectors $L_1^i(\overrightarrow{\mathcal{Z}})x, \ldots, L_{\ell_i}^i(\overrightarrow{\mathcal{Z}})x$ are linearly dependent, that is, for each $\overrightarrow{\mathcal{Z}}$ and x, there exists $\lambda_j(\overrightarrow{\mathcal{Z}}, x)$, such that $\sum_{j=1}^{\ell_i} \lambda_j(\overrightarrow{\mathcal{Z}}, x) L_j^i(\overrightarrow{\mathcal{Z}})x = 0\}$. We emphasize that i also depends on $\overrightarrow{Z}$ and x, that is $i = i(\overrightarrow{\mathcal{Z}}, x)$.

$\mathcal{C} := \mathcal{B} \bigcap \mathcal{A}^c$, where $\mathcal{A}^c$ denotes the theoretic complement of set $\mathcal{A}$.

We will show later that the set $\mathcal{N}_{\Delta_0}(\mathcal{G})$ is the disjoint union of the two sets $\mathcal{A}$ and $\mathcal{C}$. Let $\mathcal{A}_\Delta$ be the set of tuples in $\mathcal{A}$ with size Δ. Similarly, $\mathcal{C}_\Delta$ is the set of tuples in $\mathcal{C}$ with size Δ. The next three lemmas give basic properties of the sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$.

Lemma 10.1. *Let the sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be defined as above. Suppose that the quadratic function $\mathcal{Q}(\overrightarrow{\mathcal{Z}})[\overrightarrow{\mathcal{H}}]$ is positive semidefinite for all $\overrightarrow{\mathcal{H}}$ provided that the variables $\overrightarrow{\mathcal{Z}}$, having compatible dimension, are in $\mathcal{N}_{\Delta_0}(\mathcal{G})$. Then the set $\mathcal{N}_{\Delta_0}(\mathcal{G})$ is the union of the sets $\mathcal{A}$ and $\mathcal{B}$, that is, $\mathcal{N}_{\Delta_0}(\mathcal{G}) = \mathcal{A} \bigcup \mathcal{B}$, furthermore, $\mathcal{N}_{\Delta_0}(\mathcal{G})$ is the disjoint union of the sets $\mathcal{A}$ and $\mathcal{C}$.*

Proof. Observe what happens when we replace $\overrightarrow{Z}$ by tuple of matrices $\overrightarrow{\mathcal{Z}}$ of compatible dimension. Fix a vector x. Suppose that $x^T \mathcal{Q}(\overrightarrow{\mathcal{Z}})[\overrightarrow{\mathcal{H}}]x \geq 0$ for all $\overrightarrow{\mathcal{H}}$. This implies, that $\overrightarrow{w}^T M_{\mathcal{Q}(\overrightarrow{\mathcal{Z}})} \overrightarrow{w} \geq 0$ for all $\overrightarrow{w}$ in $\mathcal{R}^{all,x}_{\overrightarrow{L}(\overrightarrow{\mathcal{Z}})}$. Thus the number of negative eigenvalues of $M_{\mathcal{Q}(\overrightarrow{\mathcal{Z}})}$ is less than or equal to the co-dimension of the space $\mathcal{R}^{all,x}_{\overrightarrow{L}(\overrightarrow{\mathcal{Z}})}$, which by Proposition 9.8 either is bounded by t or there is a $d \in \mathcal{I}$, which depends on $\overrightarrow{\mathcal{Z}}$ and x, such that $L_1^d(\overrightarrow{\mathcal{Z}})x, \ldots, L_{\ell_d}^d(\overrightarrow{\mathcal{Z}})x$ are linearly dependent for every vector x with compatible dimension.

As a consequence of the above result, the set $\mathcal{N}_{\Delta_0}(\mathcal{G})$ is the union of the sets $\mathcal{A}$ and $\mathcal{B}$, and consequently the disjoint union of the sets $\mathcal{A}$ and $\mathcal{C}$. In particular, the set $\mathcal{M}_\Delta(\mathcal{G})$ is the disjoint union of $\mathcal{A}_\Delta$ and $\mathcal{C}_\Delta$ for each $\Delta \succeq \Delta_0$. □

Lemma 10.2. *For every $\Delta \succeq \Delta_0$, suppose the closure of $\mathcal{A}_\Delta$, denoted by $\overline{\mathcal{A}_\Delta}$, contains $\mathcal{M}_\Delta(\mathcal{G})$, in other words, $\mathcal{A}_\Delta$ is dense in $\mathcal{M}_\Delta(\mathcal{G})$. Then $\mathcal{A}_\Delta$ actually equals the whole set $\mathcal{M}_\Delta(\mathcal{G})$.*

Proof. The lemma follows directly from the fact that the eigenvalues of a symmetric matrix continuously depend on the norm of the matrix, c.f. Appendix D of [GL83]. □

We present some definitions about direct sum and sets which respect direct sums, since they are important tools for proving linear dependence of the coefficient of the border vector.

Definition 10.3 (Direct Sum)**.** Our definition of the direct sum is the usual one, which for two matrices $\mathcal{Z}_1$ and $\mathcal{Z}_2$ is given by

$$\mathcal{Z}_1 \oplus \mathcal{Z}_2 := \begin{pmatrix} \mathcal{Z}_1 & 0 \\ 0 & \mathcal{Z}_2 \end{pmatrix}.$$

Now, we extend this definition for v tuples of matrices $\vec{\mathcal{Z}} := \{\mathcal{Z}_1, \dots, \mathcal{Z}_v\}$. For any positive integer J, we denote by $\vec{\mathcal{Z}}^J$ the direct sum $\vec{\mathcal{Z}} \oplus \cdots \oplus \vec{\mathcal{Z}}$ of J copies of $\vec{\mathcal{Z}}$. For instance, the direct sum of three v tuples of matrices $\vec{\mathcal{Z}}_1 := \{\mathcal{Z}_{11}, \dots, \mathcal{Z}_{1v}\}$, $\vec{\mathcal{Z}}_2 := \{\mathcal{Z}_{21}, \dots, \mathcal{Z}_{2v}\}$, and $\vec{\mathcal{Z}}_3 := \{\mathcal{Z}_{31}, \dots, \mathcal{Z}_{3v}\}$ is given by

$$\vec{\mathcal{Z}}_1 \oplus \vec{\mathcal{Z}}_2 \oplus \vec{\mathcal{Z}}_3 := \{\mathcal{Z}_{11} \oplus \mathcal{Z}_{21} \oplus \mathcal{Z}_{31}, \dots, \mathcal{Z}_{1v} \oplus \mathcal{Z}_{2v} \oplus \mathcal{Z}_{3v}\}.$$

Note that from the above definition, if noncommutative functions L_j^i applied to a v tuples of matrices $\vec{\mathcal{Z}}$ produce matrices $L_j^i(\vec{\mathcal{Z}}) \in \mathbb{R}^{n \times n}$, then these functions L_j^i applied to the direct sum $\vec{\mathcal{Z}}^J$ produce matrices $L_j^i(\vec{\mathcal{Z}}^J) \in \mathbb{R}^{Jn \times Jn}$.

Definition 10.4 (A Set Respects Direct Sums)**.** A set $\mathcal{P}$ is said to respect direct sums if $\vec{\mathcal{Z}}_i$ for $i = 1, \dots, \mu$ is contained in the set $\mathcal{P}$ implies that the direct sum $\vec{\mathcal{Z}}_i^J$ is also contained in $\mathcal{P}$ for each positive integer J. Furthermore, the direct sum $\vec{\mathcal{Z}}_1^J \oplus \cdots \oplus \vec{\mathcal{Z}}_\mu^J$ is also contained in $\mathcal{P}$.

We present Proposition 10.5 below because it foreshadow a key idea in the proof of Theorem 8.3.

Lemma 10.5. *Under the same assumptions as Lemma* 10.1*, the set* $\mathcal{C}$ *(a subset of* $\mathcal{B}$*) respects direct sums.*

Proof. The proof is by contradiction. Pick $\vec{\mathcal{Z}}_i \in \mathcal{C}$, thus $\vec{\mathcal{Z}}_i \in \mathcal{A}^c$, which means $M_{\mathcal{Q}(\vec{\mathcal{Z}}_i)}$ has at least $t+1$ negative eigenvalues. Next suppose that $\vec{\mathcal{Z}}_i^J$ is not contained in $\mathcal{C}$ for some integer J. Then by Lemma 10.1, $\mathcal{Z}_i^J$ is contained in $\mathcal{A}$, which by the definition of the set $\mathcal{A}$ implies that $M_{\mathcal{Q}(\vec{\mathcal{Z}}_i^J)}$ has less than or equal to t negative eigenvalues. On the other hand, by the property of direct sum, the number of negative eigenvalues of $M_{\mathcal{Q}(\vec{\mathcal{Z}}_i^J)}$ equals J times the number of the negative eigenvalues of $M_{\mathcal{Q}(\vec{\mathcal{Z}}_i)}$. Thus, $M_{\mathcal{Q}(\vec{\mathcal{Z}}_i)}$ also has less than or equal to t negative eigenvalues, which is a contradiction. Hence, $\vec{\mathcal{Z}}_i^J$ is contained in $\mathcal{C}$ for all integers J.

Similarly, we can further prove that the direct sum $\vec{\mathcal{Z}}_1^J \oplus \cdots \oplus \vec{\mathcal{Z}}_\mu^J$ is also contained in $\mathcal{C}$. □

10.1. Subsets of $\mathcal{B}$ Which Respect Direct Sums

The following few lemmas pertain to a subset $\mathcal{P}$ of $\mathcal{B}$ which respects direct sums. The next lemma shows that for a finite set denoted by $\mathcal{S}$, consisting of different elements in $\mathcal{P}$, we can find a linear combination of the coefficients of the border vector which equals zero for any $\vec{\mathcal{Z}} \in \mathcal{S}$. We actually prove something a little more general. That is,

Lemma 10.6. *Let $\mathcal{P}$ be a subset of $\mathcal{B}$ which respects direct sums. Suppose that $\mathcal{S}$ is a finite subset of $\mathcal{P}$. Then, there are scalars $\lambda_j(\mathcal{S})$ and an integer $d(\mathcal{S}) \in \mathcal{I}$ (which depend upon the choice of the set $\mathcal{S}$) such that*

$$\sum_{j=1}^{\ell_{d(\mathcal{S})}} \lambda_j(\mathcal{S}) L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}}) = 0, \tag{10.1}$$

for every $\vec{\mathcal{Z}} \in \mathcal{S}$.

Proof. The proof relies on taking direct sums of matrices. Write the set $\mathcal{S}$ as $\mathcal{S} = \{\vec{\mathcal{Z}}_1, \ldots, \vec{\mathcal{Z}}_\mu\}$, where each $\vec{\mathcal{Z}}_i \in \mathcal{P}$ for $i = 1, \ldots, \mu$. For this proof, it suffices to take each $L_j^d(\vec{\mathcal{Z}}_i)$ to be in $\mathbb{R}^{n\times n}$. Choose $\vec{\mathcal{Z}}^*$ to be the direct sum $\vec{\mathcal{Z}}_1^n \oplus \cdots \oplus \vec{\mathcal{Z}}_\mu^n$, where each $\vec{\mathcal{Z}}_i^n$ for $i = 1, \ldots, \mu$ is the direct sum of n copies of $\vec{\mathcal{Z}}_i$. Define the vector e^* to be

$$e^* := \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \in \mathbb{R}^{n^2},$$

where the e_k for $k = 1, \ldots, n$ are the standard basis elements for $\mathbb{R}^n$. Also let x^* be a vector that contains μ copies of e^*, that is,

$$x^* = \begin{pmatrix} e^* \\ \vdots \\ e^* \end{pmatrix} \in \mathbb{R}^{\mu n^2}.$$

Since (by assumption) the set $\mathcal{P}$ respects direct sum, $\vec{\mathcal{Z}}^*$ is also contained in $\mathcal{P}$. Then, by the definition of the set $\mathcal{B}$, there exist scalars $\lambda_j(\vec{\mathcal{Z}}^*, x^*)$ and an integer

$d \in \mathcal{I}$ (we reemphasize that $d = d(\vec{\mathcal{Z}}^*, x^*)$), such that

$$\sum_{j=1}^{\ell_d} \lambda_j(\vec{\mathcal{Z}}^*, x^*) L_j^d(\vec{\mathcal{Z}}^*) x^* = 0.$$

It follows that

$$\sum_{j=1}^{\ell_d} \lambda_j(\vec{\mathcal{Z}}^*, x^*) L_j^d(\vec{\mathcal{Z}}_i^n) e^* = 0, \text{ for } i = 1, \dots, \mu.$$

This implies that for $i = 1, \dots, \mu$,

$$\sum_{j=1}^{\ell_d} \lambda_j(\vec{\mathcal{Z}}^*, x^*) L_j^d(\vec{\mathcal{Z}}_i) e_k = 0, \text{ for } k = 1, \dots, n.$$

Since the $\{e_k\}_{k=1}^n$, is a basis for $\mathbb{R}^n$, we obtain that

$$\sum_{j=1}^{\ell_d} \lambda_j(\vec{\mathcal{Z}}^*, x^*) L_j^d(\vec{\mathcal{Z}}_i) = 0, \text{ for } i = 1, \dots, \mu.$$

Since $(\vec{\mathcal{Z}}^*, x^*)$ are determined by the choice of the set $\mathcal{S}$, we conclude that

$$\sum_{j=1}^{\ell_{d(\mathcal{S})}} \lambda_j(\mathcal{S}) L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}}_i) = 0,$$

for each $\vec{\mathcal{Z}}_i \in \mathcal{S}$, with $\lambda_j(\mathcal{S}) := \lambda_j(\vec{\mathcal{Z}}^*, x^*)$ and $d(\mathcal{S}) := d(\vec{\mathcal{Z}}^*, x^*)$. Thus we obtain equation (10.1) required for the lemma. □

The next Lemma 10.7 extends this result from the finite set $\mathcal{S}$ to the bigger set $\mathcal{M}_\Delta(\mathcal{G})$.

Lemma 10.7. *Let $\mathcal{P}$ be a subset of $\mathcal{B}$ which respects direct sums. For $\Delta \succeq \Delta_0$, if there is an open set $\mathcal{U}_\Delta$ contained in $\mathcal{P}_\Delta := \mathcal{P} \bigcap \mathcal{M}_\Delta(\mathcal{G})$, then there exist scalars $\lambda_j(\Delta)$ and an integer $d(\Delta) \in \mathcal{I}$, such that*

$$\sum_{j=1}^{\ell_{d(\Delta)}} \lambda_j(\Delta) L_j^{d(\Delta)}(\vec{\mathcal{Z}}) = 0,$$

for every $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$.

Proof. Fix a size $\Delta \succeq \Delta_0$. Denote by **vec** the map which sends a tuple of matrices $\vec{\mathcal{Z}}$ in $\mathcal{P}_\Delta$ to their entries arranged as a vector $(y_1, \dots, y_K) \in \mathbb{R}^K$ as follows

$$\mathbf{vec} : \mathcal{P}_\Delta \to \mathbb{R}^K,$$

where K is total number of entries in the matrices in $\vec{\mathcal{Z}}$. The order of the arrangement does not matter, but the same order must be used consistently. Denote

$\mathbf{vec}^-$ the inverse map of $\mathbf{vec}$. Then each entry of the matrix $L^i_j(\vec{\mathcal{Z}})$ is a rational function of the elements $y_1, \ldots, y_K$. By multiplying through by some polynomials if necessary, we can assume without loss of generality that each entry of $L^i_j(\vec{\mathcal{Z}})$ is a polynomial in the K variables $y_1, \ldots, y_K$. Let D_r be the maximum degree of y_r among all of the polynomials which are entries of $L^i_j(\vec{\mathcal{Z}})$, for all i and j.

Since $\mathcal{P}_\Delta$ contains an open set $\mathcal{U}_\Delta$, we can choose a finite set

$$\tilde{\mathcal{S}} := \{(y_1^{v_1}, \ldots, y_K^{v_K}) \in \mathbb{R}^K, \text{ here } v_r = 1, \ldots, D_r + 1 \text{ for all } r = 1, \ldots, K\},$$

such that for every $r = 1, \ldots, K$, the elements $y_r^1, \ldots, y_r^{D_r+1}$ are distinct. That is, the values in each coordinate of $\tilde{\mathcal{S}}$ are distinct. The set $\tilde{\mathcal{S}}$ is a subset of the space $\mathbb{R}^K$. As a consequence, the cardinality Π of the set $\tilde{\mathcal{S}}$ (the number of elements in $\tilde{\mathcal{S}}$) equals $\Pi = \prod_{r=1}^{K} (D_r + 1)$.

Define $\mathcal{S} = \mathbf{vec}^-(\tilde{\mathcal{S}}) \in \mathcal{P}_\Delta$. By Lemma 10.6, for each tuple $\vec{\mathcal{Z}} \in \mathcal{S}$, there are constants $\lambda_j(\mathcal{S})$ and an integer $d(\mathcal{S}) \in \mathcal{I}$, both depending on $\mathcal{S}$ such that

$$\sum_{j=1}^{\ell_{d(\mathcal{S})}} \lambda_j(\mathcal{S}) L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}}) = 0, \tag{10.2}$$

for every tuple of matrices $\vec{\mathcal{Z}} \in \mathcal{S}$.

Now we show that (10.2) actually holds for every $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$. Note that (10.2) can be equivalently written as

$$\sum_{j=1}^{\ell_{d(\mathcal{S})}} \lambda_j(\mathcal{S}) \left[L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}})\right]_{(p,q)} = 0, \tag{10.3}$$

for every tuple of matrices $\vec{\mathcal{Z}} \in \mathcal{S}$, where $\left[L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}})\right]_{(p,q)}$ denotes the (p, q)th entry of $L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}})$. By the previous argument, $\left[L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}})\right]_{(p,q)}$ is a polynomial in the K variables $y_1, \ldots, y_K$, and also the maximum degree on each indeterminate y_r is no greater than D_r. Clearly all the elements in $\tilde{\mathcal{S}}$ give rise to matrix tuple $\vec{\mathcal{Z}}$ that satisfy the polynomial equation (10.3) for all p and q. By the elementary theorem of algebra which says that every nonzero polynomial in one complex variable with degree D_r has at most D_r zeros, we conclude by the construction (cardinality Π) of the set $\tilde{\mathcal{S}}$ that for every $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$

$$\sum_{j=1}^{\ell_{d(\mathcal{S})}} \lambda_j(\mathcal{S}) \left[L_j^{d(\mathcal{S})}(\vec{\mathcal{Z}})\right]_{(p,q)} = 0, \text{ for each } p \text{ and } q,$$

Thus it follows that

$$\sum_{j=1}^{\ell_{d(\Delta)}} \lambda_j(\Delta) L_j^{d(\Delta)}(\vec{\mathfrak{Z}}) = 0,$$

for every $\vec{\mathfrak{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$, by choosing constants $\lambda_j(\Delta) := \lambda_j(\mathcal{S})$ and integer $d(\Delta) = d(\mathcal{S})$. □

Now we have obtained the linear combination $\lambda_j(\Delta)$ of $L_j^{d(\Delta)}(\vec{\mathfrak{Z}})$, which is zero for all elements $\vec{\mathfrak{Z}}$ in $\mathcal{M}_\Delta(\mathcal{G})$ for one fixed size Δ. The following lemma connects the coefficients $\lambda_j(\Delta)$ of the linear combinations between different size. It says that if we have an annihilating linear combination for $\mathcal{M}_\Delta(\mathcal{G})$, then this same combination will also be annihilated for all size Δ' with $\Delta \succeq \Delta'$.

Lemma 10.8. *Fix a size Δ. Suppose there are scalars $\lambda_j(\Delta)$ and an integer $i(\Delta) \in \mathcal{I}$ such that*

$$\sum_{j=1}^{\ell_i(\Delta)} \lambda_j(\Delta) L_j^{i(\Delta)}(\vec{\mathfrak{Z}}) = 0,$$

for every $\vec{\mathfrak{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$. Then

$$\sum_{j=1}^{\ell_i(\Delta)} \lambda_j(\Delta) L_j^{i(\Delta)}(\vec{\mathfrak{Z}}) = 0,$$

for every $\vec{\mathfrak{Z}} \in \mathcal{M}_{\Delta'}(\mathcal{G})$, with $\Delta \succeq \Delta'$.

Proof. Let $\vec{\emptyset} = \{\emptyset_1, \ldots, \emptyset_v\}$ be a tuple of zero matrices of compatible dimension. For every $\vec{\mathfrak{Z}}_0 \in \mathcal{M}_{\Delta'}(\mathcal{G})$ let $\vec{\mathfrak{Z}}$ be

$$\vec{\mathfrak{Z}} = \vec{\mathfrak{Z}}_0 \oplus \vec{\emptyset}$$

to get $\vec{\mathfrak{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$ with $\Delta \succeq \Delta'$. By assumption, there are scalars $\lambda_j(\Delta)$ and an integer $i(\Delta)$ such that

$$\sum_{j=1}^{\ell_i(\Delta)} \lambda_j(\Delta) L_j^{i(\Delta)}(\vec{\mathfrak{Z}}) = 0,$$

for every $\vec{\mathfrak{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$. Then plug in the decomposition of $\vec{\mathfrak{Z}}$ given above, together with the fact that

$$L_j^{i(\Delta)}(\vec{\mathfrak{Z}}_0 \oplus \vec{\emptyset}) = L_j^{i(\Delta)}(\vec{\mathfrak{Z}}_0) \oplus L_j^{i(\Delta)}(\vec{\emptyset}) = \begin{pmatrix} L_j^{i(\Delta)}(\vec{\mathfrak{Z}}_0) & 0 \\ 0 & 0 \end{pmatrix},$$

to obtain

$$\sum_{j=1}^{\ell_i(\Delta)} \lambda_j(\Delta) L_j^{i(\Delta)}(\vec{\mathfrak{Z}}_0) = 0$$

for every $\vec{\mathcal{Z}}_0 \in \mathcal{M}_{\Delta'}(\mathcal{G})$. □

So far we have shown that for every fixed size Δ, there exists an annihilating linear combination (that may depend on the size Δ), which also holds for any size Δ' with $\Delta \succeq \Delta'$. Now we show that actually there exists an annihilating linear combination for all $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$ that does not depend on the size Δ.

Lemma 10.9. *Let $\mathcal{P}$ be a subset of $\mathcal{B}$ which respect direct sums. Suppose there is a size $\Delta_1 \succeq \Delta_0$, such that for every size $\Delta \succeq \Delta_1$ there is an open set $\mathcal{U}_\Delta$ contained in $\mathcal{P}_\Delta := \mathcal{P} \bigcap \mathcal{M}_\Delta(\mathcal{G})$. Then, there are constants λ_j and an integer $d \in \mathcal{I}$ (we emphasize that λ_j and the integer d do not depend on the size Δ) such that*

$$\sum_{j=1}^{\ell_d} \lambda_j L_j^d(\vec{\mathcal{Z}}) = 0,$$

for every $\vec{\mathcal{Z}} \in \mathcal{M}(\mathcal{G})$.

Proof. Define the set Λ_*^Δ as

$$\Lambda_*^\Delta := \Big\{(d(\Delta), \lambda_1(\Delta), \ldots, \lambda_{\ell_{d(\Delta)}}(\Delta)) : \sum_{j=1}^{\ell_{d(\Delta)}} \lambda_j(\Delta) L_j^{d(\Delta)}(\vec{\mathcal{Z}}) = 0,$$
$$\text{for every } \vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G}) \text{ and an integer } d(\Delta) \in \mathcal{I}\Big\}.$$

Since for every $\Delta \succeq \Delta_1$, the set $\mathcal{P}_\Delta$ contains an open set $\mathcal{U}_\Delta$, we have from Lemma 10.7, that the set Λ_*^Δ is nonempty. Thus there exists a point

$$(\tilde{d}(\Delta), \tilde{\lambda_1}(\Delta), \ldots, \tilde{\lambda}_{\ell_{d(\Delta)}}(\Delta)) \in \Lambda_*^\Delta$$

for every $\Delta \succeq \Delta_1$. We can define a collection of sets for every $\Delta \succeq \Delta_1$ and every integer $\tilde{d}(\Delta)$ as

$$\Lambda_*^\Delta(\tilde{d}(\Delta)) := \{(\lambda_1(\Delta), \ldots, \lambda_{\ell_d(\Delta)}(\Delta)) : (\tilde{d}(\Delta), \lambda_1(\Delta), \ldots, \lambda_{\ell_d(\Delta)}(\Delta)) \in \Lambda_*^\Delta\}.$$

It is clear by the construction that $\Lambda_*^\Delta(d(\Delta))$ is a linear space, which is nontrivial since $(\tilde{\lambda_1}(\Delta), \ldots, \tilde{\lambda}_{\ell_d(\Delta)}(\Delta)) \in \Lambda_*^\Delta(\tilde{d}(\Delta))$ for every $\Delta \succeq \Delta_1$. Since the integer $d(\Delta)$ only has finitely many possibilities in $\mathcal{I}$ there exists an infinite increasing sequence $\{j_i\}_{i=1}^\infty$ and an integer d in $\mathcal{I}$, such that $\Lambda_*^{\Delta_{j_i}}(d)$ is nonempty for any i and such that

$$\Delta_{j_{i_1}} \succ \Delta_{j_{i_2}}, \text{ for any } i_1 > i_2.$$

By Lemma 10.8, the dimension of the space $\Lambda_*^{\Delta_{j_i}}(d)$ is a nonincreasing sequence, which is bounded below by 1. Thus

$$\min_{i \geq 1} \dim(\Lambda_*^{\Delta_{j_i}}(d)) \geq 1.$$

Hence $\bigcap_{i\geq 1} \Lambda_*^{\Delta_{j_i}}(d) \neq \emptyset$, and consequently there is an integer d (that does not depend on Δ) and scalars $\lambda_j(\Delta)$, such that

$$\sum_{j=1}^{\ell_d} \lambda_j(\Delta) L_j^d(\vec{\mathcal{Z}}) = 0,$$

for every $\vec{\mathcal{Z}} \in \bigcup_{i=1}^{\infty} \mathcal{M}_{\Delta_{j_i}}(\mathcal{G})$.

So far we have shown that the integer d does not depend on the size Δ. The next step is to show that the scalars λ_j are also independent of Δ. This is accomplished by applying Lemma 10.8 successively. Thus, we conclude that

$$\sum_{j=1}^{\ell_d} \lambda_j L_j^d(\vec{\mathcal{Z}}) = 0,$$

for every $\vec{\mathcal{Z}} \in \mathcal{M}(\mathcal{G})$. □

From all of this we obtain the following result which is interesting in areas independent of this paper.

Theorem 10.10. *Let $L_1(\vec{Z}), \ldots, L_\ell(\vec{Z})$ be noncommutative rational functions of $\vec{Z} = \{Z_1, \ldots, Z_v\}$. Let $\mathcal{G}$ be a Symbolic Inequality Domain satisfying the Openness Property. Suppose for all $\Delta \succeq \Delta_0$ we have for each $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$ of compatible dimension and each vector x that the vectors*

$$L_1(\vec{\mathcal{Z}})x, \ldots, L_\ell(\vec{\mathcal{Z}})x$$

are linearly dependent. Then the functions $L_1(\vec{Z}), \ldots, L_\ell(\vec{Z})$ are linearly dependent, that is, there are scalars λ_j (that do not depend on $\vec{Z}$) such that

$$\sum_{j=1}^{\ell} \lambda_j L_j(\vec{Z}) = 0$$

Proof. Form a subset of $\mathcal{B}$ denoted by $\mathcal{P}$ associated with $L_1(\vec{\mathcal{Z}}), \ldots, L_\ell(\vec{\mathcal{Z}})$ by

$$\mathcal{P} = \left\{ \vec{\mathcal{Z}} \in \mathcal{N}_{\Delta_0}(\mathcal{G}) : \text{ for each } \vec{\mathcal{Z}}, x \text{ there exist } \lambda(\vec{\mathcal{Z}}, x), \text{ such that } \sum_{j=1}^{\ell} \lambda_j L_j(\vec{\mathcal{Z}})x = 0 \right\}.$$

Now, we show that this set $\mathcal{P}$ respects direct sums. For $t = 1, \ldots, \mu$ let $\vec{\mathcal{Z}}_t$ be contained in $\mathcal{P}$. By definition of the set $\mathcal{P}$, for each $\vec{\mathcal{Z}}_t, x$ there exist $\lambda(\vec{\mathcal{Z}}_t, x)$ such that

$$\sum_{j=1}^{\ell} \lambda_j(\vec{\mathcal{Z}}_t, x) L_j(\vec{\mathcal{Z}}_t) x = 0, \qquad t = 1, \ldots, \mu.$$

Let x^* be a vector that contains J copies of x. Since $L_j(\vec{\mathcal{Z}}_t^{\,J}) = L_j(\vec{\mathcal{Z}}_t) \oplus \cdots \oplus L_j(\vec{\mathcal{Z}}_t)$, we have that, for $t = 1, \ldots, \mu$,

$$\sum_{j=1}^{\ell} \lambda_j(\vec{\mathcal{Z}}_t, x) L_j(\vec{\mathcal{Z}}_t^{\,J}) x^* = \begin{pmatrix} \sum_{j=1}^{\ell} \lambda_j(\vec{\mathcal{Z}}_t, x) L_j(\vec{\mathcal{Z}}_t) x & & 0 \\ & \ddots & \\ 0 & & \sum_{j=1}^{\ell} \lambda_j(\vec{\mathcal{Z}}_t, x) L_j(\vec{\mathcal{Z}}_t) x \end{pmatrix} = 0,$$

and consequently $\vec{\mathcal{Z}}_t^{\,J} \in \mathcal{P}$ for each $t = 1, \ldots, \mu$.

Thus Lemma 10.6, 10.7, 10.8 and 10.9 apply to $\mathcal{P}$. In particular Lemma 10.9 implies Theorem 10.10. □

Also Theorem 10.10 lays behind Corollary 10.11, which is here repeated.

Corollary 10.11. *Let $L_1(\vec{Z}), \ldots, L_\ell(\vec{Z})$ be noncommutative rational functions of $\vec{Z} = \{Z_1, \ldots, Z_v\}$. For each vector x, suppose that the vectors $L_1(\vec{\mathcal{Z}})x, \ldots, L_\ell(\vec{\mathcal{Z}})x$ are linearly dependent whenever matrices $\mathcal{Z}_j$ of compatible dimension are substituted for Z_j for all size Δ bigger than some Δ_0. Then there exist real numbers λ_j for $j = 1, \ldots, \ell$ such that*

$$\sum_{j=1}^{\ell} \lambda_j L_j(\vec{Z}) = 0,$$

that is, the functions $L_j(\vec{Z})$ are linearly dependent.

Proof. In Theorem 10.10 take $\mathcal{G}$ to be everything. That is, $\mathcal{G}$ contains no inequality constraints. Thus $\mathcal{G}$ has the Openness Property, since $\mathcal{M}_\Delta(\mathcal{G}) = \mathcal{M}_\Delta$. □

We need the following lemmas to complete the proof of the main Theorem.

Lemma 10.12. *Let Δ_0 be any size. Assume that T is a symmetric matrix with noncommutative rational functions $t_{ij}(\vec{Z})$ as entries. Suppose there is an integer r such that whenever tuple of matrices $\vec{\mathcal{Z}} \in \mathcal{N}_{\Delta_0}(\mathcal{G})$ of compatible dimension are substituted for $\vec{Z}$, the resulting matrix $T(\vec{\mathcal{Z}})$ has at most r negative eigenvalues. Then $T(\vec{\mathcal{Z}})$ is positive semidefinite (that is, $r = 0$) for each $\vec{\mathcal{Z}} \in \mathcal{M}(\mathcal{G})$.*

Proof. The key fact is

$$T(\vec{\mathcal{Z}} \oplus \vec{\mathcal{Z}}) = T(\vec{\mathcal{Z}}) \oplus T(\vec{\mathcal{Z}}).$$

This implies that if $T(\vec{\mathcal{Z}})$ has η negative eigenvalues, then T applied to the $2r$-fold direct sum $\vec{\mathcal{Z}} \oplus \cdots \oplus \vec{\mathcal{Z}}$ has $2r\eta$ negative eigenvalues. Consequently the hypothesis $2r\eta \leq r$ implies that $\eta = 0$. □

Lemma 10.13. *Suppose that $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ is positive semidefinite for every $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$, then $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ is also positive semidefinite for every $\vec{\mathcal{Z}} \in \mathcal{M}_{\Delta'}(\mathcal{G})$ with $\Delta \succeq \Delta'$.*

Proof. Use an idea similar to the one in the proof of Lemma 10.8. □

10.2. Proof of Theorem 8.3

Proof. For any $\Delta \succeq \Delta_0$, if $\mathcal{A}_\Delta$ is dense in $\mathcal{M}_\Delta(\mathcal{G})$, that is, $\overline{\mathcal{A}_\Delta} \supseteq \mathcal{M}_\Delta(\mathcal{G})$, then by Lemma 10.2, we have $\mathcal{A}_\Delta = \mathcal{M}_\Delta(\mathcal{G})$. Hence, the number of negative eigenvalues of $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ is uniformly bounded by t for all $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$. Now we apply Lemma 10.12 with $r = t$ to obtain that, for each tuple of matrices $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$ substituted for $\vec{Z}$, the matrix $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ is positive semidefinite. On the other hand, if $\mathcal{A}_\Delta$ is not dense in $\mathcal{M}_\Delta(\mathcal{G})$, then by Lemma 10.1 there exists an open set $\mathcal{U}_\Delta$ contained in $\mathcal{C}_\Delta \subset \mathcal{M}_\Delta(\mathcal{G})$.

So far we have shown that for any $\Delta \succeq \Delta_0$ one of the following must be satisfied, either

a. the matrix $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ is positive semidefinite for each $\vec{\mathcal{Z}} \in \mathcal{M}_\Delta(\mathcal{G})$,

or

b. there exists an open set $\mathcal{U}_\Delta$ contained in $\mathcal{C}_\Delta \subset \mathcal{M}_\Delta(\mathcal{G})$.

The final step is to show that if positivity of $M_{\mathcal{Q}(\vec{\mathcal{Z}})}$ fails, then the block linear independence of the border vector (in assumption (*ii*) of Theorem 8.3) of the quadratic function $\mathcal{Q}$ also fails. Assume there is a size Δ^* such that (a) is not satisfied. Then by Lemma 10.13, (a) is not satisfied for every $\Delta \succeq \Delta^*$. Hence (b) is true for every $\Delta \succeq \Delta^*$, which by Lemma 10.9 (with $\mathcal{P}_\Delta = \mathcal{C}_\Delta$) and Lemma 10.5 implies that there are constants λ_j and an integer d such that $\sum_{j=1}^{\ell_d} \lambda_j L_j^d(\vec{\mathcal{Z}}) = 0$ for every $\vec{\mathcal{Z}} \in \mathcal{M}(\mathcal{G})$. Thus, by Corollary 10.11, the noncommutative rational functions $L_j^d(\vec{Z})$ are linearly dependent for $j = 1, \ldots, \ell_d$ and consequently the border $V(\vec{Z})[\vec{H}]$ has block linearly dependent coefficients. But this contradicts assumption (*ii*) of Theorem 8.3, finalizing in this way the proof of the main Theorem 8.3. □

Remark 10.14. It is enough (a weaker hypotheses) to consider square matrices of dimension $n \times n$ (when substituting matrices for indeterminate) to prove the theorems concerning convexity and matrix positive of noncommutative rational functions.

Appendix A. Computer Algorithm For Representing $\mathcal{Q}(\vec{\mathbf{Z}})[\vec{\mathbf{H}}]$ With $\mathbf{M}_{\mathcal{Q}}(\vec{\mathbf{Z}})$ and $\mathbf{V}(\vec{\mathbf{Z}})[\vec{\mathbf{H}}]$

In our approach, we are given a noncommutative rational function $\mathcal{Q}(\vec{Z})[\vec{H}]$, which is quadratic and hereditary in $\vec{H}$ but usually not quadratic in $\vec{Z}$, and we need to express this function as $V(\vec{Z})[\vec{H}]^T\ M_{\mathcal{Q}}(\vec{Z})\ V(\vec{Z})[\vec{H}]$. That means we have to construct the border vector $V(\vec{Z})[\vec{H}]$ and the coefficient matrix $M_{\mathcal{Q}}$. *This representation of $\mathcal{Q}(\vec{Z})[\vec{H}]$ may not be unique.*

This section describes a simplified version of the algorithm used. The algorithm is based on a simple pattern match, that is illustrated here for the case were $\vec{H} := \{H_1, H_2\}$. It can be easily expanded for the more general case where $\vec{H}$ has k entries. The algorithm explained here does not assume $\vec{H}$ necessarily symmetric. For the symmetric case, just let $\vec{H} = \vec{H}^T$ and the steps are the same.

1. Expand the quadratic function in H_1 and H_2.
2. In that case, there are four types of quadratic terms involving the H_i :
$$*H_1^T * H_1* \ , \qquad *H_1^T * H_2* \ , \qquad *H_2^T * H_1* \ , \qquad \text{and} \qquad * H_2^T * H_2 * .$$
   The pattern matching symbol $*$ means any expression that does not contain H_i.
3. We work on each one of these quadratic terms $*H_i^T * H_j*$ individually. Let $i = j = 1$. Then find all pattern of the form $*H_1^T * H_1*$. Before the pattern matching is processed, it is important that all the terms of the expression to be found are collected. That means, if there is an expression like
$$L_1^{1^T} H_1^T B_1 H_1 L_1^1 + \cdots + L_1^{1^T} H_1^T B_m H_1 L_1^1$$
   then collect all of the B_i in $A_{1,1} = \sum_{i=1}^{m} B_i$. Follows this procedure, then at the end we may have a sum of terms like:
$$L_1^{1^T} H_1^T A_{1,1} H_1 L_1^1 + L_1^{1^T} H_1^T A_{1,2} H_1 L_2 + \cdots + L_{\ell_1}^{1}{}^T H_1^T A_{\ell_1,\ell_1} H_1 L_{\ell_1}^1$$
   Where the $A_{i,j}$ for $i, j = 1, \ldots, \ell_1$ collect all the terms that match the expression $L_i^{1^T} H_1^T * H_1 L_j^1$. This step was illustrated in the example above, where all the terms that match the expression $L_1^{1^T} H_1^T * H_1 L_1^1$ are collected in the coefficient $A_{1,1}$.
4. The same procedure applies for the terms $*H_1^T * H_2*$, $*H_2^T * H_1*$, and $*H_2^T * H_2*$.
5. Once the finding of all the patterns is finished, the $A_{t,s}$ are the entries of the coefficient matrix $M_{\mathcal{Q}}$, and the $H_i L_j^i$ are the entries of the border vector $V(\vec{Z})[\vec{H}]$.

References

[AHO98] Farid Alizadeh, Jean-Pierre A. Haeberly, and Michael L Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM J. Optim.*, 8(3):746–768, 1998.

[BEGFB94] S. Boyd, L. El-Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[CHS00] Juan F. Camino, J. W. Helton, and R. E. Skelton. A symbolic algorithm for determining convexity of a matrix function: How to get Schur complements out of your life. In 39^{th} *IEEE Conference on Decision and Control*, 2000.

[EGN99] L. El-Ghaoui and S. Niculescu. *Advances in Linear Matrix Inequality Methods in Control*. SIAM, 1999.

[Frö97] R. Fröberg. *An Introduction to Gröbner Bases*. Pure and Applied Mathematics. John Wiley & Sons, 1997.

[GL83] G. Golub and C. Van Loan. *Matrix Computation*. Johns Hopkins University Press, 1983.

[Hel] J. W. Helton. "Positive" noncommutative polynomial are sums of squares. To appear in *Annals of Mathematics*.

[HJ96] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1996.

[HM97] J. W. Helton and O. Merino. Coordinate optimization for bi-convex matrix inequalities. In 36^{th} *IEEE Conference on Decision and Control*, volume 4, pages 3609–3613, Dec. 1997.

[HM98] J. W. Helton and O. Merino. Sufficient conditions for optimization of matrix functions. In 37^{th} *IEEE Conference on Decision and Control*, volume 3, pages 3361–3365, Dec. 1998.

[HSW98] J. W. Helton, M. Stankus, and J. J. Wavrik. Computer simplification of formulas in linear system theory. *IEEE transaction on Automatic Control*, 4(3):302–314, March 1998.

[Mor86] Teo Mora. Gröebner bases for noncommutative polynomial rings. *Lecture Notes in Computer Sci.*, 1(229):353–362, 1986.

[Mor94] Teo Mora. An introduction to commutative and noncommutative Gröbner bases. *Theoretical Computer Science*, 134(1):131–173, 7 November 1994.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13. SIAM studies in applied mathematics, 1994.

[Par00] P. A. Parrilo. On a decomposition of multivariable forms via LMI methods. In *American Control Conference*, 2000.

[PW98] V. Powers and T. Wörmann. An algorithm for sums of squares of real polynomials. *Journal of Pure and Applied Algebra*, 127:99–104, 1998.

[Roc97] R. T. Rockafellar. *Convex Analysis*. Princeton Press, 1997.

[SI95] R. E. Skelton and T. Iwasaki. Increased roles of linear algebra in control education. *IEEE Control Systems*, 8:76–90, 1995.

[SIG97] R. E. Skelton, T. Iwasaki, and K. M. Grigoriadis. *A Unified Algebraic Approach to Linear Control Design*. Taylor & Francis, 1997.

[VB96] L. Vandeberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, March 1996.

Acknowledgment

The authors would like to thank Joshua Griffin who helped considerably with our software; particularly with the LDU algorithm and its implementation.

Juan F. Camino
Department of Mechanical and Aerospace Engineering, UC San Diego
Mail code 0411, 9500 Gilman Drive, La Jolla, CA, 93092-0411, USA
E-mail: camino@ucsd.edu

J. W. Helton
Department of Mathematics, UC San Diego
Mailcode 0112, 9500 Gilman Drive, La Jolla, CA, 93093-0112,USA
E-mail: helton@euclid.ucsd.edu

Robert E. Skelton
Department of Mechanical and Aerospace Engineering, UC San Diego
Mail code 0411, 9500 Gilman Drive, La Jolla, CA, 93092-0411, USA
E-mail: bobskelton@ucsd.edu

Jieping Ye
Department of Mathematics, UC San Diego
Mailcode 0112, 9500 Gilman Drive, La Jolla, CA, 93093-0112, USA
E-mail: jye@euclid.ucsd.edu

Submitted: May 20, 2001