NONCOMMUTATIVE POLYNOMIALS WITH ALMOST POSITIVE SECOND DERIVATIVES

An Undergraduate Honors Thesis

Submitted in June 2006

By

Jeremy M. Greene

Advisor: Professor J. William Helton

Department of Mathematics
University of California, San Diego

Contents

1. Acknowledgements

I would like to thank Professor Helton for even giving me the opportunity to do research. This would not have been possible without all of his patience, guidance, encouragement, and expertise. He is a true mentor and a great advisor.

Abstract. Consider symmetric polynomials $p = p(x) = p(x_1, \dots, x_g)$ (with real coefficients) in g noncommuting variables $x_1, \dots, x_g$. Any such polynomial can be written as a sum and difference of squares of noncommutative polynomials. We define the negative signature of a polynomial as the minimum number of negative squares used in the sum and difference of squares representation.

Define the **Hessian** of a polynomial, p, by

$$p''(x)[h] := \frac{d^2 p(x+th)}{dt^2}|_{t=0}.$$

It is a polynomial in x and $h = (h_1, \dots, h_g)$ that is homogeneous of degree 2 in h. In this paper, we classify all symmetric noncommutative polynomials such that the Hessian of the polynomials have negative signature bounded by 1.

This is motivated by an attempt to begin developing geometric properties of noncommutative real algebraic varieties.

2. Introduction

Suppose p is a symmetric polynomial in g symmetric variables. For any such polynomial p we define its first and second directional derivatives:

$$p'(x)[h] := \frac{dp(x+th)}{dt}|_{t=0} \quad and \quad p''(x)[h] := \frac{d^2 p(x+th)}{dt^2}|_{t=0}.$$

The second derivative we call the **Hessian** of p. We now introduce the **modified Hessian**. It is defined for any $\lambda \in \mathbb{R}$ by

$$\mathcal{H}_\lambda(x)[h] := p''(x)[h] + \lambda p'(x)[h]^T p'(x)[h].$$

2.1. The Signature of a Polynomial.

Here we shall establish bounds on the degree of a symmetric polynomial $p = p(x) = p(x_1, \dots, x_g)$ (with real coefficients) in g noncommuting (nc)variables $x_1, \dots, x_g$ in terms of the "signature" of its **Hessian**

$$p''(x)[h] := \frac{d^2 p(x+th)}{dt^2}|_{t=0},$$

which is a polynomial in x and $h = (h_1, \dots, h_g)$ that is homogeneous of degree 2 in h. The bounds are obtained by exploiting the interplay between assorted representations for $p(x)$ and $p''(x)[h]$. In particular, every polynomial $f(x)$ in the class under study admits a representation of the form (a sum and difference of squares)

$$(SDS) \qquad f(x) = \sum_{j=1}^{\sigma_+} f_j^+(x)^T f_j^+(x) - \sum_{\ell=1}^{\sigma_-} f_\ell^-(x)^T f_\ell^-(x)$$

where f_j^+, f_ℓ^- are noncommutative polynomials. Such representations are highly non-unique. However, there is a unique smallest number of positive (resp. negative squares) $\sigma_\pm^{min}(f)$ required in an SDS decomposition of f.

Formal definitions of polynomials and derivatives come in a later section.

From a Corollary and Equation (0.1) in [DHMpreprint], we get the following theorem:

Theorem 2.1. *If $p(x)$ is a symmetric polynomial of degree d in noncommutative symmetric variables $x_1, \dots, x_g$, then*

$$d \leq 2\sigma_\pm(p'') + 2. \tag{2.1}$$

We shall define a **matrix convex noncommutative polynomial** to be one whose Hessian is nonnegative in the sense $\sigma_-^{min}(p'') = 0$. This theorem generalizes an earlier theorem of Helton and McCullough [HM04], which states that every matrix convex noncommutative polynomial has degree $d \leq 2$.

There are refinements which say that if equality holds or is near to holding in Equation (2.1), then the highest degree term, p_d (where $p = \sum_{i=0}^d p_i$), factors or "partially factors" in a certain way.

2.2. **Main Results.** In this thesis we prove the following theorem and make a conjecture based on it.

Consider the special class of degree 4 polynomials of the form

$$\begin{aligned} p = {} & C_0 + K_1x_1 + K_2x_2 \\ & +B_1x_1^2 + B_2x_1x_2 + B_2x_2x_1 + B_4x_2^2 \\ & +(L_1x_1 + L_2x_2)(Q_1x_1^2 + Q_3x_1x_2 + Q_2x_2x_1 + Q_4x_2^2) \\ & +(Q_1x_1^2 + Q_2x_1x_2 + Q_3x_2x_1 + Q_4x_2^2)(L_1x_1 + L_2x_2) \\ & +(L_1x_1 + L_2x_2)(A_1x_1^2 + A_2x_1x_2 + A_2x_2x_1 + A_4x_2^2)(L_1x_1 + L_2x_2). \end{aligned} \tag{2.2}$$

We call polynomials of this form **semifactorable**. Without loss of generality, we assume $L_1^2 + L_2^2 = 1$. We form a matrix of coefficients of the polynomial

$$\begin{pmatrix} B_{11} & B_{12} & Q_1 & Q_2 \\ B_{21} & B_{22} & Q_3 & Q_4 \\ Q_1 & Q_3 & A_{11} & A_{12} \\ Q_2 & Q_4 & A_{21} & A_{22} \end{pmatrix}$$

and abbreviate it to

$$\begin{pmatrix} B & Q \\ Q^T & A \end{pmatrix}.$$

Let $U : \mathbb{R}^3 \to \mathbb{R}^4$ be given by

$$U = \begin{pmatrix} L_2 & 0 & 0 \\ -L_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is an isometry since $L_1^2 + L_2^2 = 1$.

Theorem 2.2. *Let p be a symmetric noncommutative polynomial. The Hessian, p'', of p, having negative signature bounded by 1 (i.e., $\sigma_-^{min}(p'') \leq 1$) is equivalent to p having the semifactorable structure defined in Equation (2.2), where without loss of generality $\begin{pmatrix} L_1 & L_2 \end{pmatrix}$ is a vector of length one, and the matrix*

$$U^T \begin{pmatrix} B & Q \\ Q^T & A \end{pmatrix} U$$

being positive semidefinite.

In addition, p has degree 3 if and only if p_3 has the form $L\Lambda L$ where Λ and L are linear, with $L = L_1 x_1 + L_2 x_2$, and

$$(L_2 \ \ -L_1) B \begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} \geq 0.$$

2.3. **Conjectures.** Related in character to our theorem is the next conjecture.

Conjecture 2.3. *If there is a $\lambda \in \mathbb{R}^+$ such that $\sigma_-^{min}(\mathcal{H}_\lambda(x)[h]) = 0$ then p has degree 2 or less.*

This is a strengthening of the case $d = 2$ in Theorem 2.1, which we recall says

Theorem 2.4. *If the $\sigma_-^{min}(p''(x)[h]) = 0$ then p has degree 2 or less.*

The importance of the strengthening is that this may well be a key step in developing a geometric theory of noncommutative real algebraic varieties with "negative curvature". See Section 4 for this motivation. The next lemma shows that Theorem 2.1 considerably narrows down the class of polynomials, p, we must analyze to prove the conjecture.

Lemma 2.5. *If the modified Hessian of a polynomial p is positive semidefinite then $\sigma_-^{min}(p''(x)[h]) \leq 1$.*

Proof. The modified Hessian, $\mathcal{H}_\lambda(x)[h]$, of p can be represented in SDS form as

$$\sum_{j=1}^{\sigma_+} f_j^+(x)^T f_j^+(x) - \sum_{\ell=1}^{\sigma_-} f_\ell^-(x)^T f_\ell^-(x).$$

Then, by definition of the modified Hessian, we see that

$$\begin{aligned} p''(x)[h] &= \sum_{j=1}^{\sigma_+} f_j^+(x)^T f_j^+(x) - \sum_{\ell=1}^{\sigma_-} f_\ell^-(x)^T f_\ell^-(x) - \lambda p'(x)[h]^T p'(x)[h] \\ &= \sum_{j=1}^{\sigma_+} f_j^+(x)^T f_j^+(x) - \sum_{\ell=1}^{1+\sigma_-} f_\ell^-(x)^T f_\ell^-(x). \end{aligned}$$

The modified Hessian positive semidefinite means $\sigma_-^{min}(\mathcal{H}_\lambda) = 0$. It follows that

$$\sigma_-^{min}(p''(x)[h]) \leq 1 + \sigma_-^{min}(\mathcal{H}_\lambda(x)[h]) \leq 1. \tag{2.3}$$

Note that if $\lambda = 0$, we have $\sigma_-^{min}(p'') = 0$ because $p''(x)[h] = \mathcal{H}_\lambda(x)[h]$ and if $\lambda > 0$, we get equality in Equation (2.3). ■

3. Details of the Setup

We introduce our notation and terminology in detail. Our layout follows [DHMpreprint].

3.1. **The Setup.** Let $x = \{x_1, \dots, x_g\}$ denote noncommuting indeterminates and let $\mathbb{R}\langle x\rangle$ denote the set of polynomials $p(x) = p(x_1, \dots, x_g)$ in the indeterminates x. Thus a polynomial $p \in \mathbb{R}\langle x\rangle$ is a finite (real) linear combination of monomials (words) in x whose degree is, by definition, the maximum of the lengths of the monomials appearing (non-trivially) in the linear combination. For example, if $g = 3$, then

$$p_1 = x_1x_2^3 + x_2 + x_3x_1x_2 \quad \text{and} \quad p_2 = x_1x_2^3 + x_2^3x_1 + x_3x_1x_2 + x_2x_1x_3$$

are polynomials of degree 4 in $\mathbb{R}\langle x\rangle$.

More generally, with $|m|$ denoting the length of the monomial m, a polynomial of degree at most d is an expression of the form

$$p = \sum_{|m|\leq d} c_m m\,, \tag{3.1}$$

where each $c_m \in \mathbb{R}$.

There is a natural **involution** m^T on monomials given by the rule

$$\text{if} \quad m = x_{i_1}x_{i_2}\cdots x_{i_k}\,, \quad \text{then} \quad m^T = x_{i_k}\cdots x_{i_2}x_{i_1}$$

which of course extends to polynomials as in Equation (3.1) by

$$p^T = \sum_{|m|\leq d} c_m m^T.$$

A polynomial $p \in \mathbb{R}\langle x\rangle$ is said to be **symmetric** if $p = p^T$. The second polynomial p_2 listed above is symmetric, the first is not.

3.1.1. *Substituting Matrices for Indeterminates.* Let $(\mathbb{R}^{n\times n}_{sym})^g$ denote the set of g-tuples $(X_1,\dots,X_g)$ of real symmetric $n\times n$ matrices. We shall be interested in evaluating a polynomial $p(x) = p(x_1,\dots,x_g)$ that belongs to $\mathbb{R}\langle x\rangle$ at a tuple $X = (X_1,\dots,X_g) \in (\mathbb{R}^{n\times n}_{sym})^g$. In this case $p(X)$ is also an $n\times n$ matrix and the involution on $\mathbb{R}\langle x\rangle$ that was introduced earlier is compatible with matrix transposition, i.e.,

$$p^T(X) = p(X)^T,$$

where $p(X)^T$ denotes the transpose of the matrix $p(X)$. When $X \in (\mathbb{R}^{n\times n}_{sym})^g$ is substituted into p the constant term $p(0)$ of $p(x)$ becomes $p(0)I_n$. Thus, for example,

$$p(x) = 3 + x^2 \quad \Longrightarrow \quad p(X) = 3I_n + X^2.$$

A symmetric polynomial $p \in \mathbb{R}\langle x\rangle$ is **matrix positive** if $p(X)$ is a positive semidefinite matrix for each tuple $X = (X_1,\dots,X_g) \in (\mathbb{R}^{n\times n}_{sym})^g$.

3.2. **Derivatives.** Now we formally define the derivatives on noncommutative polynomials. We follow the exposition in [HPpreprint].

For a polynomial $p \in \mathbb{R}\langle x\rangle$ we define the *directional derivative*:

$$p'(x)[h] := \frac{d}{dt} p(x+th)_{|t=0}.$$

It is a linear form in h. Similarly, the k^{th} derivative

$$p^{(k)}(x)[h] := \frac{d^k}{dt^k} p(x+th)_{|t=0}$$

is homogeneous of degree k in h.

More formally, we regard the directional derivative $p'(x)[h] \in \mathbb{R}\langle x, h\rangle$ as a polynomial in $2g$ free symmetric (i.e. invariant under *) variables $(x_1,\dots,x_g,h_1,\dots,h_g)$. In the case of a monomial $m = x_{j_1}x_{j_2}\cdots x_{j_n}$ the derivative is:

$$m'[h] = h_{j_1}x_{j_2}\cdots x_{j_n} + x_{j_1}h_{j_2}x_{j_3}\cdots x_{j_n} + \ \dots\ + x_{j_1}\cdots x_{j_{n-1}}h_{j_n}$$

and for a polynomial $p = \sum p_m m$ the derivative is

$$p'(x)[h] = \sum p_m m'[h].$$

If p is symmetric, then so is p'. For g-tuples of symmetric matrices of a fixed size X, H, observe that the evaluation formula

$$p'(X)[H] = \lim_{t\to 0} \frac{p(X+tH) - p(X)}{t}$$

holds. Alternately, with $q(t) = p(X+tH)$, we find

$$p'(X)[H] = q'(0).$$

Likewise for a polynomial $p \in \mathbb{R}\langle x\rangle$, the *Hessian*, $p''(x)[h]$ of $p(x)$, can be thought of as the formal second directional derivative of p in the "direction" h.

If $p'' \neq 0$, that is, if degree $p \geq 2$, then the degree of $p''(x)[h]$ as a polynomial in the $2g$ variables $x_1, \ldots, x_g, h_1 \ldots, h_g$ is equal to the degree of $p(x)$ as a polynomial in $x_1, \ldots, x_g$.

Likewise for k^{th} derivatives.

Example 3.1. 1. $p(x) = x_2x_1x_2$

$$p'(x)[h] = \frac{d}{dt}[(x_2+th_2)(x_1+th_1)(x_2+h_2)]_{|t=0} = h_2x_1x_2+x_2h_1x_2+x_2x_1h_2.$$

2. One variable $p(x) = x^4$. Then

$$p'(x)[h] = hxxx + xhxx + xxhx + xxxh$$

Note each term is linear in h and h replaces each occurrence of x once and only once:

$$p''(x)[h] =$$
$$hhxx + hhxx + hxhx + hxxh+$$
$$hxhx + xhhx + xhhx + xhxh+$$
$$hxxh + xhxh + xxhh + xxhh,$$

which yields

$$p''(x)[h] = 2hhxx + 2hxhx + 2hxxh + 2xhhx + 2xhxh + 2xxhh.$$

Note each term is degree two in h and h replaces each pair of x's exactly once. Likewise

$$p^{(3)}(x)[h] = 6(hhhx + hhxh + hxhh + xhhh)$$

and $p^{(4)}(x)[h] = 24hhhh$ and $p^{(5)}(x)[h] = 0$.

3. $p = x_1^2x_2$

$$p''(x)[h] = h_1^2x_2 + h_1x_1h_2 + x_1h_1h_2.$$

4. Motivation

Many linear systems engineering problems lead to Matrix Inequalities, MIs. These take the form of a polynomial or a matrix of polynomials with matrix variables taking a positive semidefinite value. One very much desires these polynomials to be "convex" or "quasiconvex" in the unknown matrix variables, since if they are, then numerical calculations are reliable. Many systems problems have the property that they are "dimension free" and in this case the form of the polynomials remains the same for matrices of all size. In other words, we have noncommutative polynomials as we have studied in this thesis.

As mentioned in Theorem 2.4 it was shown that matrix convex noncommutative polynomials have degree two or less. This thesis builds machinery we expect to be pertinent to quasiconvex polynomials, to be discussed in a moment. Also, more generally, there are classes of engineering problems where convexity is relaxed and the $\sigma_-^{min}(p'') \leq 1$ class is an example.

For motivation we discuss smooth quasiconvex functions f on $\mathbb{R}^g$. These are functions all of whose sublevel sets

$$\mathcal{C}_c := \{X : \ f(X) \leq c\}$$

are convex. It can be shown that a set $\mathcal{C}_c$ with boundary denoted $\partial\mathcal{C}_c$ with gradient $\nabla f(X) \neq 0$ for $X \in \partial\mathcal{C}_c$ is convex if and only if *the Hessian of f restricted to the tangent plane of f at X is positive semidefinite.*

This is the X, H having dimension 1 case of the statement $p''(X)[H] \geq 0$ for all H satisfying $p'(X)[H] = 0$ in the notation of this thesis. One can show (Professor Helton's unpublished notes) that this is equivalent to *there is a $\lambda > 0$ making the modified Hessian, $\mathcal{H}_\lambda(X)[H]$, nonnegative for all H.* Thus a smooth function f with nowhere vanishing gradient (except at 0) are quasiconvex on a bounded domain B if and only if for λ large enough $\mathcal{H}_\lambda(X)[H]$ is nonnegative for all $X \in B$ and all H.

Details of the noncommutative version of this are not complete, but are encouraging enough that we think it is very worthwhile to classify noncommutative polynomials p having nonnegative modified Hessian and to classify the bigger class consisting of p having $\sigma_-^{min}(p'') \leq 1$.

5. Background for Proofs

In this section, we provide some of the tools needed to prove Theorem 2.2. We first explain what the middle matrix is and how we compute it. We then provide key matrix lemmas that are heavily needed in order to prove Theorem 2.2.

5.1. A Representation for Quadratics: Defining $Z(0)$ and $\mathcal{Z}$.

The following representation for symmetric noncommutative polynomials $q(x)[h]$ that are of degree ℓ in x and homogeneous of degree 2 in h was proved in [CHSY03]. This representation is exploited extensively in this paper.

(5.1)
$$q(x)[h] = [V_0^T, V_1^T, \ldots, V_\ell^T] \begin{bmatrix} Z_{00} & Z_{01} & \cdots & Z_{0,\ell-1} & Z_{0\ell} \\ Z_{10} & Z_{11} & \cdots & Z_{1,\ell-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ Z_{\ell 0} & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_\ell \end{bmatrix},$$

where:

(1) $\ell = d - 2$.

(2) V_j abbreviates $V_j(x)[h]$, $j = 0, \ldots, \ell$, is a vector of height g^{j+1} with entries of the form $h_i m, i = 1, \ldots g$, and m runs through all the g^j monomials in the noncommuting variables $x_1, \ldots, x_g$ of degree j. Thus, e.g., if $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}$ for short, then

$$V_0 = \begin{bmatrix} h_1 \\ \vdots \\ h_g \end{bmatrix}, \quad V_1 = \begin{bmatrix} h_1\mathbf{x} \\ \vdots \\ h_g\mathbf{x} \end{bmatrix}.$$

In particular, V_j, alias $V_j(x)[h]$, depends linearly on h. The height of V we denote by $ht(V)$, it is $ht(V) := \sum_{j=0}^{d-2} g^{j+1}$.

(3) Z_{ij} which abbreviates $Z_{ij}(x)$ is a matrix of size $g^{i+1} \times g^{j+1}$ and the entries in Z_{ij} are polynomials in the noncommuting variables $x_1, \ldots, x_g$ of degree $\leq d - 2 - (i + j) = \ell - (i + j)$. In particular, $Z_{i,\ell-i} = Z_{i,\ell-i}(x)$ is a constant matrix for $i = 0, \ldots, \ell$.

(4) $Z_{ij}^T = Z_{ji}$.

The matrix Z in Equation (5.1) will be referred to as the **middle matrix of the polynomial** $q(x)[h]$ and the vectors $V_j = V_j(x)[h]$ with monomials as entries will be referred to as **border vectors**. There is a fair amount of flexibility in the choice of the border vector. The particular choice made above has the advantage of keeping track of all the monomials that come into play in the representation of arbitrary $q(x)[h]$. However, for high degree polynomials with few terms, this will result in a middle matrix with most of its entries equal to zero.

We say Z and M are **polynomially congruent** and write $Z \sim_p M$ if

(5.2)
$$Z(x) = W(x)^T M(x) W(x)\,,$$

where $W(x)$ is an invertible polynomial matrix with an inverse $W(x)^{-1}$ that is also a matrix with polynomial entries. Theorem 3.5 in [DHMpreprint] says that $Z(x) \sim_p Z(0)$ and we define $\mathcal{Z} := Z(0)$.

5.2. **Examples.** Here are some examples that take a polynomial, compute the modified Hessian, factor it as $v^T Z v$, and perform a LDL^T decomposition on the middle matrix with $x = 0$. It is important to note that if we set $\lambda = 0$ we are dealing with the Hessian.

Example 5.1. Take $p(x) = x^2$. Then we take the modified Hessian to get

$$\mathcal{H}_\lambda(x)[h] = 2h^2 + \lambda hxhx + \lambda hx^2h + \lambda xh^2x + \lambda xhxh.$$

Since $\mathcal{H}_\lambda(x)[h]$ is quadratic in h, we can factor this as in Equation (5.1):

$$\mathcal{H}_\lambda(x)[h] = v^T Z_\lambda(x) v = \begin{pmatrix} h & xh \end{pmatrix} \begin{pmatrix} 2+\lambda x^2 & \lambda x \\ \lambda x & \lambda \end{pmatrix} \begin{pmatrix} h \\ hx \end{pmatrix}.$$

So we note that $Z_\lambda(x) = \begin{pmatrix} 2+\lambda x^2 & \lambda x \\ \lambda x & \lambda \end{pmatrix}$ and thus if we substitute $x = 0$ we obtain the matrix:

$$\mathcal{Z}_\lambda = Z_\lambda(0) = \begin{pmatrix} 2 & 0 \\ 0 & \lambda \end{pmatrix}.$$

Next we perform a noncommutative LDL^T decomposition to get

$$\mathcal{Z}_\lambda = \begin{pmatrix} 2 & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = LDL^T.$$

One should note that the LDL^T algorithm also gives a permutation matrix and in this example, the permutation matrix is $P = I_2$ (the 2×2 identity matrix).

Example 5.2. Take $p(x) = x_1x_2 + x_2x_1$. Then we get the modified Hessian to be:

$$\begin{aligned} \mathcal{H}_\lambda(x)[h] = \quad & 2h_1h_2 + 2h_2h_1 + \lambda h_1x_2h_1x_2 + \lambda h_1x_2h_2x_1 \\ + \quad & \lambda h_1x_2x_1h_2 + \lambda h_1x_2x_2h_1 + \lambda h_2x_1h_1x_2 + \lambda h_2x_1h_2x_1 \\ + \quad & \lambda h_2x_1x_1h_2 + \lambda h_2x_1x_2h_1 + \lambda x_1h_2h_1x_2 + \lambda x_1h_2h_2x_1 \\ + \quad & \lambda x_1h_2x_1h_2 + \lambda x_1h_2x_2h_1 + \lambda x_2h_1h_1x_2 + \lambda x_2h_1h_2x_1 \\ + \quad & \lambda x_2h_1x_1h_2 + \lambda x_2h_1x_2h_1. \end{aligned}$$

Since $\mathcal{H}_\lambda(x)[h]$ is quadratic in h (where $h = \{h_1, h_2\}$), we can factor this as in Equation (5.1):

$$\mathcal{H}_\lambda(x)[h] = v^T Z_\lambda(x) v =$$

$$\begin{pmatrix} h_1 & h_2 & x_1h_2 & x_2h_1 \end{pmatrix} \begin{pmatrix} \lambda x_2^2 & 2+\lambda x_2x_1 & \lambda x_2 & \lambda x_2 \\ 2+\lambda x_1x_2 & \lambda x_1^2 & \lambda x_1 & \lambda x_1 \\ \lambda x_2 & \lambda x_1 & \lambda & \lambda \\ \lambda x_2 & \lambda x_1 & \lambda & \lambda \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_2x_1 \\ h_1x_2 \end{pmatrix}.$$

So we note that $Z_\lambda(x) = \begin{pmatrix} \lambda x_2^2 & 2+\lambda x_2x_1 & \lambda x_2 & \lambda x_2 \\ 2+\lambda x_1x_2 & \lambda x_1^2 & \lambda x_1 & \lambda x_1 \\ \lambda x_2 & \lambda x_1 & \lambda & \lambda \\ \lambda x_2 & \lambda x_1 & \lambda & \lambda \end{pmatrix}$ and thus if we substitute $x = 0$ we obtain the matrix:

$$\mathcal{Z}_\lambda = Z_\lambda(0) = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & \lambda & \lambda \\ 0 & 0 & \lambda & \lambda \end{pmatrix}.$$

Next we perform a noncommutative LDL^T decomposition to get

$$\mathcal{Z}_\lambda = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & \lambda & \lambda \\ 0 & 0 & \lambda & \lambda \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = LDL^T.$$

In this example the permutation matrix is

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

5.3. **A Lower Bound for Negative and Positive Eigenvalues.** Given a matrix M, let $\mu_\pm(M)$ denote the number of strictly positive (resp. strictly negative) eigenvalues of M. Note that $\mu_\pm(M)$ does not say anything about the number of zero eigenvalues. Lemma 5.4 below, which is Lemma 4.3 in [DHMpreprint], gives a lower bound on the number of positive (resp. negative) eigenvalues of a 3×3 block matrix. We prepare for its proof with the following lemma which gives basic properties of matrices.

Lemma 5.3. *Let A, C, W, and U be matrices where W and U are invertible. If the range of C contains the range of A then*

1. *the range of U^TC contains the range of U^TA, and*
2. *the range of U^TA is equal to the range of U^TAW.*

Proof. To prove (1), suppose x is in the range of $U^T A$. This means that $x = U^T Ay_1$ for some vector y_1. Since the range of C contains the range of A we can write this as

$$x = U^T Ay_1 = U^T Cw_1$$

for some vector w_1 which means that x is in the range of $U^T C$. Hence the range of $U^T C$ contains the range of $U^T A$.

To prove (2), let x be in the range of $U^T A$. This means that $x = U^T Ay$ for some vector y. Since W is invertible, there exists some vector q such that $y = Wq$ so that

$$x = U^T Ay = U^T AWq$$

which means that x is in the range of $U^T AW$.

Now suppose x is in the range of $U^T AW$. Then $x = U^T AWy$ for some vector y. But then

$$x = U^T AWy = U^T Av$$

where $v = Wy$ which means x is in the range of $U^T A$. ■

Lemma 5.4. *If*

$$E = \begin{bmatrix} A & B & C \\ B^T & D & 0 \\ C^T & 0 & 0 \end{bmatrix}$$

is a real symmetric matrix, then

$$\mu_\pm(E) \geq \mu_\pm(D) + rank C\,. \tag{5.3}$$

If the range of C contains the range of A and the range of B then Equation (5.3) holds with equality.

Proof. Let $C = USV^T$ be the singular value decomposition of C. Then

$$E \sim \begin{bmatrix} \widetilde{A} & \widetilde{B} & S \\ \widetilde{B}^T & D & 0 \\ S^T & 0 & 0 \end{bmatrix},$$

where $\widetilde{A} = U^T AU$ and $\widetilde{B} = U^T B$. Thus, if

$$S = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix},$$

where F is a positive definite diagonal matrix and $\widetilde{A}$ and $\widetilde{B}$ are written in compatible block form as

$$\widetilde{A} = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \widetilde{B}_1 \\ \widetilde{B}_2 \end{bmatrix},$$

respectively, then

$$E \sim \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} & \widetilde{B}_1 & F & 0 \\ \widetilde{A}_{21} & \widetilde{A}_{22} & \widetilde{B}_2 & 0 & 0 \\ \widetilde{B}_1^T & \widetilde{B}_2^T & D & 0 & 0 \\ F & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since F is invertible, there exists matrices R_j so that $\widetilde{A}_{11} = FR_1$, $\widetilde{A}_{12} = FR_2$ and $\widetilde{B}_1 = FR_3$.

Let $\mathcal{E}$ denote the elementary matrix

$$\mathcal{E} = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ -\frac{1}{2}R_1 & -R_2 & -R_3 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix},$$

and verify that

$$E \sim \mathcal{E}^T E \mathcal{E} = \begin{bmatrix} 0 & 0 & 0 & F & 0 \\ \widetilde{0} & \widetilde{A}_{22} & \widetilde{B}_2 & 0 & 0 \\ \widetilde{0} & \widetilde{B}_2^T & D & 0 & 0 \\ F & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let G denote the matrix obtained by removing the second and fifth rows and columns of the matrix on the right hand side above. By the principle of eigenvalue interlacing

$$\mu_\pm(E) \geq \mu_\pm(G).$$

By Lemma 4.2 in [DHMpreprint]

$$\mu_\pm(G) \geq \operatorname{rank}(F) + \mu_\pm(D).$$

The last two displayed equations combine to complete the proof, since $\operatorname{rank} F = \operatorname{rank} C$.

To prove equality in Equation (5.3), we first show that the range of S contains the range of $\widetilde{A}$ and the range of $\widetilde{B}$. To show this, we use

Lemma 5.3 to get

$$\begin{aligned} range(\widetilde{A}) &= range(U^T AU) \\ &= range(U^T A) \\ &\subset range(U^T C) \\ &= range(U^T CV) \\ &= range(S) \end{aligned}$$

and for $\widetilde{B}$ we use Lemma 5.3 again to get

$$\begin{aligned} range(\widetilde{B}) &= range(U^T B) \\ &\subset range(U^T C) \\ &= range(U^T CV) \\ &= range(S). \end{aligned}$$

If we apply S to any (nonzero) vector we get that the range of S is

$$S \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} Fv \\ 0 \end{bmatrix}.$$

Since the range of S contains the range of $\widetilde{A}$ we must have that

$$\widetilde{A} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} \star \\ 0 \end{bmatrix}.$$

It then follows that $\widetilde{A}_{21} = \widetilde{A}_{22} = 0$. Similarly, since the range of S contains the range of $\widetilde{B}$, it follows that $\widetilde{B}_2 = 0$ so $\widetilde{B} = \begin{bmatrix} \widetilde{B}_1 \\ 0 \end{bmatrix}$. Then

$$E \sim \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} & \widetilde{B}_1 & F & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \widetilde{B}_1^T & 0 & D & 0 & 0 \\ F & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since F is invertible, there exist matrices R_j so that $\widetilde{A}_{11} = FR_1$, $\widetilde{A}_{12} = FR_2$ and $\widetilde{B}_1 = FR_3$. Let $\mathcal{E}$ denote the elementary matrix as before

$$\mathcal{E} = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ -\frac{1}{2}R_1 & -R_2 & -R_3 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}.$$

One can verify that

$$E \sim \mathcal{E}^T E \mathcal{E} = \begin{bmatrix} 0 & 0 & 0 & F & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & D & 0 & 0 \\ F & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Recall that $\mu_\pm(M)$ is defined as the number of strictly positive (resp. strictly negative) eigenvalues for some matrix M. Let $\hat{G}$ denote the matrix obtained by removing the second and fifth rows and columns of $\mathcal{E}^T E \mathcal{E}$. Since we are removing rows and columns of only zeros, $\mu_\pm(\hat{G}) = \mu_\pm(\mathcal{E}^T E \mathcal{E}) = \mu_\pm(E)$. Then by Lemma 4.2 in [DHMpreprint], we get equality in Equation (5.3). ■

6. Properties of the Hessian

The following lemma is Lemma 3.8 in [DHMpreprint] and relates $\sigma_\pm^{min}(q)$ for a special $q(x)[h]$, which is quadratic in h, to the number of positive (resp. negative) eigenvalues of $Z(0)$, denoted $\mu_\pm(Z(0))$, where $Z(0)$ is the middle matrix of q evaluated at 0.

Lemma 6.1. *Let $q(x)[h]$ be a symmetric polynomial in nc variables $(x, h) = (x_1, \dots, x_g, h_1, \dots, h_g)$ that is of degree ℓ in x and homogeneous of degree 2 in h with middle matrix $Z(x)$. Then*

$$\mu_\pm(Z(0)) \leq \sigma_\pm^{min}(q). \tag{6.1}$$

If $q(x)[h] = p''(x)[h]$ is the Hessian of a symmetric polynomial $p(x)$ of degree $\ell + 2$ in g nc variables $(x_1, \dots, x_g)$, then equality prevails in (6.1),i.e.,

$$\mu_\pm(Z(0)) = \sigma_\pm^{min}(q). \tag{6.2}$$

7. Proof of Theorem 2.2

Now we prove Theorem 2.2. We first use Lemma 6.1 to see that $\sigma_-^{min}(p'') \leq 1$ is equivalent to the middle matrix $Z(0)$ for p'' satisfying $\mu_-(Z(0)) \leq 1$. Thus the whole weight of our proof falls on the following proposition.

Proposition 7.1. *Let p be a symmetric noncommutative polynomial. The middle matrix $Z(x)$, or equivalently $Z(0)$, of the Hessian of p having at most one negative eigenvalue is equivalent to p being semi-factorable, where without loss of generality $\begin{pmatrix} L_1 & L_2 \end{pmatrix}$ is a vector of* *length one, and the matrix*

$$U^T \begin{pmatrix} B & Q \\ Q^T & A \end{pmatrix} U \tag{7.1}$$

being positive semidefinite.

In addition, p has degree 3 if and only if p_3 has the form $L\Lambda L$ where Λ and L are linear, with $L = L_1x_1 + L_2x_2$, and

$$(L_2 \ \ -L_1)B\begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} \geq 0.$$

7.1. **Proof of Proposition 7.1.** First, we assume that $Z(0)$ of the Hessian of p has at most one nc negative eigenvalue. Lemma 6.1 implies that $Z(0)$ having at most one nc negative eigenvalue is equivalent to the Hessian of p having negative signature bounded by 1 (i.e., $\sigma_-^{min}(p'') \leq 1$). Theorem 2.1 says that if p is a symmetric polynomial of degree d in noncommutative symmetric variables $x_1, \ldots, x_n$ then

$$d \leq 2\sigma_{\pm}(p'') + 2$$

where p'' denotes the Hessian of p. From this we can see that if $\sigma_-^{min}(p'') \leq 1$ then $d \leq 4$. Moreover, when this happens, any polynomial, p, with $\sigma_-^{min}(p'') \leq 1$ has a special form which we now describe.

First, we write p as

$$p = p_0 + p_1 + p_2 + p_3 + p_4$$

where p_i is homogeneous of degree i for $i = 0, 1, 2, 3, 4$. From Theorem 1.4 in [DHMpreprint], as illustrated by Example 1.4 in [DHMpreprint], we get the decomposition of the degree 4 term, p_4, as

$$p_4(x) = L(x)A(x)L(x) \tag{7.2}$$

where L is a linear function of x and A is a quadratic function of x; that is,

$$\begin{aligned} L(x) &= L_1x_1 + L_2x_2 \\ A(x) &= A_1x_1^2 + A_2x_1x_2 + A_2x_2x_1 + A_4x_2^2. \end{aligned}$$

We also have that $p_2 = B_1x_1^2 + B_2x_1x_2 + B_2x_2x_1 + B_4x_2^2$ is a general quadratic function of x, $p_1 = K_1x_1 + K_2x_2$ is linear in x, and $p_0 = C_0$ is a constant. This contains all of the structure which can be read off from [DHMpreprint].

Next we show the polynomial has semifactorable form.

7.1.1. *Obtaining Semifactorable Form.* At this point we have a polynomial of the form

$$p = p_0 + p_1 + p_2 \tag{7.3}$$

$$+T_1x_1^3+T_2x_1^2x_2+T_2x_2x_1^2+T_3x_1x_2^2+T_3x_2^2x_1+T_4x_1x_2x_1+T_5x_2x_1x_2+T_6x_2^3$$

$$+p_4.$$

The following calculations were greatly assisted by running the noncommutative symbolic algebra package, NCAlgebra [HMS02].

The middle matrix of the Hessian of p given by Equation (7.3) evaluated at 0 is a 14×14 denoted $Z(0)$. We conjugate it with the matrix M defined by

$$M = \begin{bmatrix} L_1 & L_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ L_2 & -L_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & L_1 & L_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & L_1 & L_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & L_2 & -L_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & L_2 & -L_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

to obtain the matrix $R = M^T Z(0) M$ of the form

$$R = \begin{pmatrix} \alpha & \beta_1 & \beta_2 \\ \beta_1^T & \tau & 0 \\ \beta_2^T & 0 & 0 \end{pmatrix}$$

where α is in $\mathbb{R}^{1\times 1}, \beta \in \mathbb{R}^{1\times 13}$, the matrix $\tau \in \mathbb{R}^{5\times 5}$ is given by

$$\tau = 2 \begin{bmatrix} f_0 & f_1 & L_2^2T_1 - 2L_1L_2T_2 + L_1^2T_5 & f_2 & L_2^2T_4 - 2L_1L_2T_3 + L_1^2T_6 \\ f_1 & A_1 & 0 & A_2 & 0 \\ L_2^2T_1 - 2L_1L_2T_2 + L_1^2T_5 & 0 & 0 & 0 & 0 \\ f_2 & A_2 & 0 & A_4 & 0 \\ L_2^2T_4 - 2L_1L_2T_3 + L_1^2T_6 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where $L_1^2 + L_2^2 = 1$,

$$\begin{aligned} f_0 &= B_4L_1^2 - 2B_2L_1L_2 + B_1L_2^2 \\ f_1 &= L_1L_2T_1 - L_1^2T_2 + L_2^2T_2 - L_1L_2T_5 \\ f_2 &= L_1L_2T_4 - L_1^2T_3 + L_2^2T_3 - L_1L_2T_6, \end{aligned}$$

and in particular β_2 is the 1×8 vector

(7.4) $$\beta_2 =$$

$$[2A_1L_1^3{+}2A_1L_1L_2^2\ ,\ 2A_2L_1^3{+}2A_2L_1L_2^2\ ,\ 2A_2L_1^3{+}2A_2L_1L_2^2\ ,\ 2A_4L_1^3{+}2A_4L_1L_2^2,$$
$$2A_1L_1^2L_2{+}2A_1L_2^3\ ,\ 2A_2L_1^2L_2{+}2A_2L_2^3\ ,\ 2A_2L_1^2L_2{+}2A_2L_2^3\ ,\ 2A_4L_1^2L_2{+}2A_4L_2^3].$$

If $\beta_2 \neq 0$ then the rank of β_2 is 1 and by Equation (5.3) in Lemma 5.4 we need τ positive semidefinite. Next we use the permutation matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

to multiply τ on the left and right to get

$$\tau_P = \left(\frac{1}{2}\right) P\tau P$$

$$= \begin{bmatrix} f_0 & f_1 & f_2 & L_2^2T_1 - 2L_1L_2T_2 + L_1^2T_5 & L_2^2T_4 - 2L_1L_2T_3 + L_1^2T_6 \\ f_1 & A_1 & A_2 & 0 & 0 \\ f_2 & A_2 & A_4 & 0 & 0 \\ L_2^2T_1 - 2L_1L_2T_2 + L_1^2T_5 & 0 & 0 & 0 & 0 \\ L_2^2T_4 - 2L_1L_2T_3 + L_1^2T_6 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

At this point we invoke that τ_P is positive semidefinite. Indeed the form of τ_P via Lemma 5.4 gives the necessary conditions

$$\begin{aligned} L_2^2T_1 - 2L_1L_2T_2 + L_1^2T_5 &= 0 \\ L_2^2T_4 - 2L_1L_2T_3 + L_1^2T_6 &= 0. \end{aligned}$$

The general solutions are

$$T_1 = 2c_2L_1 \quad \textit{and} \quad T_2 = L_1c_1 + L_2c_2 \quad \textit{and} \quad T_5 = 2c_1L_2$$

$$T_3 = L_1c_3 + L_2c_4 \quad \textit{and} \quad T_4 = 2c_4L_1 \quad \textit{and} \quad T_6 = 2c_3L_2$$

where c_1, c_2, c_3, and c_4 are constants in $\mathbb{R}$. We then take these values and substitute them into the polynomial in Equation (7.3) and we factor it to get
(7.5)

$$p_3 = (L_1x_1+L_2x_2)(c_2x_1^2+c_1x_1x_2+c_4x_2x_1+c_3x_2^2)+(c_2x_1^2+c_4x_1x_2+c_1x_2x_1+c_3x_2^2)(L_1x+L_2y).$$

This shows p_3 to be of the semifactorable form

$$p_3(x) = L(x)Q^T(x) + Q(x)L(x)$$

where Q is a quadratic function of x. Thus, we have proved that if $Z(0)$ of the Hessian of p has at most one negative eigenvalue then p has semifactorable form as stated in the proposition.

7.1.2. *Positivity of Matrix (7.1).* Next we will show that $Z(0)$ of the Hessian of p having one negative eigenvalue is equivalent to Expression (7.1) being positive semidefinite, assuming now that p has the semifactorable form as in Equation (2.2). The middle matrix of the Hessian of p evaluated at 0 is a 14×14 denoted $Z(0)$. We conjugate it with the same matrix M as before to obtain the matrix $\hat{R} = M^T Z(0) M$ of the form

$$\hat{R} = \begin{pmatrix} \hat{\alpha} & \hat{\beta}_1 & \hat{\beta}_2 \\ \hat{\beta}_1^T & \hat{\tau} & 0 \\ \hat{\beta}_2^T & 0 & 0 \end{pmatrix}$$

where $\hat{\alpha}$ is in $\mathbb{R}^{1\times 1}$, $\hat{\beta} \in \mathbb{R}^{1\times 13}$, the matrix $\hat{\tau} \in \mathbb{R}^{5\times 5}$ is given by

$$\hat{\tau} = \begin{bmatrix} B_4L_1^2 - 2B_2L_2L_1 + B_1L_2^2 & L_2Q_1 - L_1Q_3 & 0 & L_2Q_2 - L_1Q_4 & 0 \\ L_2Q_1 - L_1Q_3 & A_1 & 0 & A_2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ L_2Q_2 - L_1Q_4 & A_2 & 0 & A_4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and $\hat{\beta}_2 = \beta_2$ as given by Equation (7.4).

Because $\hat{R}$ and $Z(0)$ are conjugate, $Z(0)$ having at most one negative eigenvalue is equivalent to $\hat{R}$ having at most one negative eigenvalue as well. Since $\hat{\beta} \in \mathbb{R}^{1\times 13}$, the range of $\hat{\beta}$ is contained in $\mathbb{R}$. The range of $\hat{\beta}_2$ is either 0 or $\mathbb{R}$. We shall later see that $\hat{\beta}_2 = 0$ is equivalent to p having degree 3.

First, assume $\hat{\beta}_2 \neq 0$. Then the range of $\hat{\beta}_2$ equals $\mathbb{R}$. Note that the range of $\hat{\beta}_1$ and the range of $\hat{\alpha}$ are also contained in $\mathbb{R}$. Putting this all together, we have that the range of $\hat{\beta}_2$ contains the range of $\hat{\beta}_1$ and the range of $\hat{\alpha}$. Also note that since $\hat{\beta}_2 \neq 0$ we have that $\operatorname{rank}(\hat{\beta}_2) = 1$.

It follows by Lemma 5.4 that we have equality in Equation (5.3):

$$\mu_\pm(\hat{R}) = \mu_\pm(\hat{\tau}) + 1.$$

To this point we have $\mu_-(R) \le 1$ if and only if the matrix $\tilde{\tau}$ given by

$$\tilde{\tau} = \begin{bmatrix} B_4L_1^2 - 2B_2L_1L_2 + B_1L_2^2 & L_2Q_1 - L_1Q_3 & L_2Q_2 - L_1Q_4 \\ L_2Q_1 - L_1Q_3 & A_1 & A_2 \\ L_2Q_2 - L_1Q_4 & A_2 & A_4 \end{bmatrix}$$

is positive semidefinite. We can write this matrix in the appealing form

$$U^T \begin{pmatrix} B & Q \\ Q^T & A \end{pmatrix} U$$

used in the proposition.

Next we turn to the $\hat{\beta}_2 = 0$ case.

7.1.3. *Classification of Degree 3 Polynomials.* Because the entries in $\hat{\beta}_2$ are the coefficients in the degree 4 term, $\hat{\beta}_2 = 0$ is equivalent to p having degree 3. Recall that $\hat{\beta}_2 = \beta_2$ in Equation (7.4) which means $A = 0$.

Consider $\hat{R}$. The third and fifth columns of the submatrix $\hat{\tau}$ are identically zero, therefore $\hat{R}_{1,4}$ and $\hat{R}_{1,6}$ are both zero or $\hat{\tau}$ is positive semidefinite. This is because Lemma 5.4 tells us that if either statement fails then $\hat{R}$ has at least two negative eigenvalues. First we analyze the condition $\hat{\tau}$ positive semidefinite.

The condition $\hat{\tau}$ positive semidefinite was exhausted previously and was shown to be equivalent to $\tilde{\tau}$ positive semidefinite. The difference now is that $A = 0$ since there is no degree 4 term.

Now $\tilde{\tau}$ becomes

$$\begin{aligned} \tilde{\tau}^\star &:= \tilde{\tau}|_{A=0} = \begin{bmatrix} B_4L_1^2 - 2B_2L_1L_2 + B_1L_2^2 & L_2Q_1 - L_1Q_3 & L_2Q_2 - L_1Q_4 \\ L_2Q_1 - L_1Q_3 & 0 & 0 \\ L_2Q_2 - L_1Q_4 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \begin{pmatrix} L_2 & -L_1 \end{pmatrix} B \begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} & \begin{pmatrix} L_2 & -L_1 \end{pmatrix} Q \\ Q^T \begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} & 0 \end{bmatrix}. \end{aligned}$$

We need $\tilde{\tau}^\star$ positive semidefinite which is equivalent to

$$(L_2 \ \ -L_1) B \begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} \ge 0 \tag{7.6}$$

positive semidefinite and

$$Q^T \begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} = 0. \tag{7.7}$$

We write down all solutions to Equation (7.7) and substitute back into the original polynomial Q to obtain

$$c_1 L_1 x_1^2 + c_2 L_1 x_1 x_2 + c_1 L_2 x_2 x_1 + c_2 L_2 x_2^2$$

where c_1 and c_2 are free real constants. This factors as

$$(L_1 x_1 + L_2 x_2)(c_1 x_1 + c_2 x_2).$$

This means the degree 3 term looks like

$$p_3 = L \Lambda L \tag{7.8}$$

where Λ is linear in x.

Now we analyze the case

$$\hat{R}_{1,4} = 2L_1^2 L_2 Q_1 + 2L_2^3 Q_1 - 2L_1^3 Q_3 - 2L_1 L_2^2 Q_3 = 0$$

and

$$\hat{R}_{1,6} = 2L_1^2 L_2 Q_2 + 2L_2^3 Q_2 - 2L_1^3 Q_4 - 2L_1 L_2^2 Q_4 = 0.$$

Simple algebra shows

$$\hat{R}_{1,4} = -2(L_1^2 + L_2^2)(-L_2 Q_1 + L_1 Q_3) = 0$$

$$\hat{R}_{1,6} = -2(L_1^2 + L_2^2)(-L_2 Q_2 + L_1 Q_4) = 0.$$

This gives two equations and four unknowns which the Q's must satisfy. They are the same as Equation (7.7); therefore, we again conclude $p_3 = L \Lambda L$. ■

8. Future Work

Recall the modified Hessian conjecture, Conjecture 2.3, stating that if there is a $\lambda \in \mathbb{R}^+$ such that $\sigma_-^{min}(\mathcal{H}_\lambda(x)[h]) = 0$ then p has degree 2 or less. The conjecture below, Conjecture 8.1, is a more refined version that we believe is true.

Conjecture 8.1. *If there is a $\lambda \in \mathbb{R}$, $\lambda > 0$ and an $\epsilon > 0$, such that the modified Hessian of p takes positive semidefinite values whenever it is evaluated on tuples of matrices $X = \{X_1, X_2\}$ in the ϵ-ball, $X_1^2 + X_2^2 \leq \epsilon^2 I$, and all $H = \{H_1, H_2\}$ where X_j and H_j are symmetric $n \times n$ matrices, then p has degree 2 or less.*

We have done many computer experiments on the modified Hessian conjectures. We are currently in the process of writing a symbolic algebra program to systematically check all cases.

References

[CHSY03] Camino, J. F, Helton, J.W., Skelton, R.E., and Ye, J., Matrix inequalities: A Symbolic Procedure to Determine Convexity Automatically, Integral Equations and Operator Theory Vol 46, issue 4, August 2003 on pp. 399-454.

[DHMpreprint] Dym, H., Helton, J.W., and McCullough, S., The hessian of a noncommutative polynomial has numerous negative eigenvalues, preprint, pp. 1-50.

[HM04] Helton, J. W. and McCullough, S.A., Convex polynomials Degree Two or Less, SIAM Journal of Matrix Analysis, **25**(2004) no. 4, pp. 1124-1139.

[HMS02] Helton, J.W., Miller, R.L., and Stankus, M., NCAlgebra Package for Mathematica Version 3.8, 2002.

[HPpreprint] Helton, J.W. and Putinar, M., Positive Polynomials in Scalar and Matrix Variables, the Spectral Theorem and Optimization, preprint, pp. 1-77.