Non-commutative Partial Matrix Convexity

DAMON M. HAY, J. WILLIAM HELTON, ADRIAN LIM &
SCOTT MCCULLOUGH

ABSTRACT. Let p be a polynomial in the non-commuting variables $(a,x) = (a_1,\ldots,a_g,x_1,\ldots,x_g)$. If p is convex in the variables x, then p has degree two in x and moreover, p has the form

$$p = L + \Lambda^T\Lambda,$$

where L has degree at most one in x and Λ is a (column) vector which is linear in x, so that $\Lambda^T\Lambda$ is a both sum of squares and homogeneous of degree two. Of course the converse is true also. Further results involving various convexity hypotheses on the x and a variables separately are presented.

1. INTRODUCTION

Fix a positive integer g and let $a = (a_1,\ldots,a_g)$ and $x = (x_1,\ldots,x_g)$ denote two classes of non-commuting variables which are assumed to be symmetric in the sense explained below, and let $\mathbb{R}\langle a,x\rangle$ denote the polynomials in these variables. Thus an element of $\mathbb{R}\langle a,x\rangle$ is an $\mathbb{R}$-linear combination of words built from a and x.

There is a natural involution ^{T} on $\mathbb{R}\langle a,x\rangle$ which reverses the order of a word determined by

$$(fg)^T = g^T f^T$$

for $f, g \in \mathbb{R}\langle a,x\rangle$ and

$$x_j^T = x_j \quad a_j^T = a_j,$$

for $j = 1,2,\ldots,g$.

Indiana University Mathematics Journal ©, Vol. 57, No. 6 (2008), Special Issue

If the polynomial $L(a,x)$ has degree at most one in x and if $\Lambda(a,x)$ is a (column) vector which is linear in x, then the polynomial

$$p = L(a,x) + \Lambda(a,x)^T \Lambda(a,x)$$

is convex in x, since, for each fixed a,

$$\frac{1}{2}(p(a,x) + p(a,y)) - p\left(a, \frac{x+y}{2}\right) = \Lambda\left(a, \frac{x-y}{2}\right)^T \Lambda\left(a, \frac{x-y}{2}\right).$$

The converse is a corollary of the main results of this paper.

We remark that it was already shown in [7] that if a symmetric polynomial with real coefficients is convex in all its variables x, then it must have degree two or less in x.

A notable example where a polynomial which is convex in some of its variables arises is in the Riccati inequality,

$$0 \preceq -xbb^T x + a^T x + xa + c, \quad c = c^T, \ x = x^T$$

where x is a symmetric unknown and a, b, c are not necessarily symmetric knowns. Formulas such as the Riccati inequality are very common in classical linear control theory, where often the problems are 'dimensionfree' in the sense that the resulting matrix inequalities have matrix unknowns such that the form of the inequality does not change with the dimension of the unknowns. It is hoped that for such dimensionless systems problems, convexity will offer more generality than the more restrictive 'linear matrix inequalities' that also arise. However, the results here and in [8] suggest this hope might be unwarranted.

In the remainder of this introduction we introduce the terminology and background necessary to state our main results on the structure of polynomials which satisfy various convexity hypotheses. The exposition is restricted to the case where both the a and x variables are symmetric, but, for the most part, the results go through with the obvious modifications to the situation where some of the variables are symmetric and others are not.

Of course there is no need to assume that the number g_a of a variables is the same as the number g_x of x variables, but it does simplify the exposition. The interested reader should have no problem in refining various estimates which depend upon g in the case where $g_a \neq g_x$.

The authors wish to thank Magdelena Musat for helpful comments and suggestions.

1.1. Non-commutative polynomials. A non-commutative polynomial, or simply polynomial, q in g non-commuting variables $y = (y_1, \dots, y_g)$ is a $\mathbb{R}$-linear combination of words in the letters $y = (y_1, \dots, y_g)$. Thus

$$q = \sum q_w w,$$

where the sum is finite, the w's are words, and $q_w \in \mathbb{R}$.

There is a natural involution ^{T} on words in y which reverses the order of the product. Namely,

$$y_{j_1} y_{j_2} \cdots y_{j_n} = w \mapsto w^T = y_{j_n} \cdots y_{j_2} y_{j_1}. \tag{1.1}$$

This involution naturally extends to polynomials by linearity,

$$q^T = \sum q_w w^T,$$

and a polynomial p is symmetric if $p = p^T$. As noted before, the conventions here mean that $y_j^T = y_j$ and so in this sense the variables themselves are symmetric.

Polynomials are naturally evaluated at g-tuples of symmetric matrices. Let $\mathbb{S}_n$ denote the symmetric $n \times n$ matrices with real entries and let $\mathbb{S}_n(\mathbb{R}^g)$ denote g tuples $Y = (Y_1, \ldots, Y_g)$ where each $Y_j \in \mathbb{S}_n$. There is no requirement that the Y_j commute. Given the word w from Equation (1.1),

$$w(Y) = Y^w = Y_{j_1} Y_{j_2} \cdots Y_{j_n}$$

as expected, and of course,

$$q(Y) = \sum q_w Y^w.$$

Note that the involution on polynomials is compatible with the transpose operation on matrices so that $q(Y)^T = q^T(Y)$. In particular, if p is symmetric, then so is $p(Y)$.

1.2. Convexity. A symmetric polynomial p is convex provided

$$[tp(Y) + (1-t)p(Z)] - p(tY + (1-t)Z) \succeq 0 \tag{1.2}$$

for all n, all pairs Y, $Z \in \mathbb{S}_n(\mathbb{R}^g)$ and $0 < t < 1$. Here $P \succeq 0$ means the square matrix P is positive semi-definite in the sense that $P = P^T$ and all of its eigenvalues are non-negative.

Even in one variable ($g = 1$) convexity in this non-commutative setting is different from ordinary convexity since there is no requirement that Y and Z commute. Indeed, the polynomial $p(y) = y^4$ in the single variable y is not convex (for a simple example where convexity fails for this p see [7]).

The notion of convexity for a polynomial naturally extends to partial convexity; i.e., convexity in a subset of the variables.

1.3. Domains and convexity. It is natural to consider polynomials which are assumed convex only on a subset; i.e., where the inequality of Equation (1.2) is to hold only for some choices of pairs Y and Z. As a preliminary, it is necessary to discuss *non-commutative domains* or *domains* for short.

Let $\mathbb{S}(\mathbb{R}^g)$ denote the sequence $(\mathbb{S}_n(\mathbb{R}^g))_n$. A non-commutative domain $\mathcal{D}$ in $\mathbb{S}(\mathbb{R}^g)$ is a sequence $\mathcal{D} = (\mathcal{D}_n)_n$ where each $\mathcal{D}_n \subset \mathbb{S}_n(\mathbb{R}^g)$ which is *closed under direct sums* in the sense that if $D_j \in S_{n_j}(\mathbb{R}^g)$, then $D_1 \oplus D_2 \in S_{n_1+n_2}(\mathbb{R}^g)$.

1.3.1. Non-commutative domains. The domain $\mathcal{D}$ is *open* if $\mathcal{D}_n$ is open in $\mathbb{S}_n(\mathbb{R}^g)$ for each n; it is *convex* if each $\mathcal{D}_n$ is convex; and it is *matrix convex* if for any isometry $V : \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ and $D = (X_1, \dots, X_g) \in \mathcal{D}_{n_2}$, $V^T DV = (V^T X_1 V, \dots, V^T X_g V) \in \mathcal{D}_{n_1}$; the domain is *semi-algebraic* if there is a finite set $\mathcal{P}$ of symmetric polynomials such that $\mathcal{D}_n = \{X \in \mathbb{S}_n(\mathbb{R}^g) : p(X) \succ 0, \text{ for all } p \in \mathcal{P}\}$.

In general, matrix convex implies convex (this depends upon the closed with respect to direct sums hypothesis). It turns out that if $\mathcal{D}$ is semi-algebraic, then convex implies matrix convex so that these two notions agree.

Below are several examples which are presented to illustrate the ideas or because they will play a role in the sequel.

Example 1.1 (The ε-neighborhood of 0)**.** Given ε the sequence of sets,

$$N_n = \left\{X \in \mathbb{S}_n(\mathbb{R}^g) : \sum X_j^2 \prec \varepsilon I_n\right\}$$

is an open matrix convex semi-algebraic domain.

Example 1.2 (Products)**.** If $\mathcal{U}$, $\mathcal{V} \subset \mathbb{S}(\mathbb{R}^g)$ are domains, then so is the *product* $\mathcal{U} \times \mathcal{V} = (\mathcal{U}_n \times \mathcal{V}_n)_n \subset \mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$.

The conditions, open, convex, matrix-convex, and semi-algebraic, are all preserved under products.

Example 1.3 (Coordinate Projections)**.** Given a domain $\mathcal{W} \subset \mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$, let

$$\pi_a(\mathcal{W}_n) = \left\{A \in \mathbb{S}_n(\mathbb{R}^g) : \text{ there exists } X \in \mathbb{S}_n(\mathbb{R}^g) \text{ such that } (A, X) \in \mathcal{W}_n\right\}.$$

The coordinate projection π_a preserves open, convex, and matrix convex domains.

1.3.2. Partial convex domains. Many of the notions surrounding domains have partial versions; i.e., versions applied to a subset of the variables. For instance, a domain $\mathcal{W}$ in $\mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$ is *open in x* if for each n and $(A, X) \in \mathcal{W}_n$ there is an open subset U of $\mathbb{S}_n(\mathbb{R}^g)$ containing X such that $\{A\} \times U \subset \mathcal{W}_n$.

1.3.3. Partial convexity. A polynomial $p(a, x)$ is *convex in x* on the open in x domain $\mathcal{W}$ if for each n and $(A, X), (A, Y) \in \mathcal{W}_n$ such that $t(A, X) + (1 - t)(A, Y) \in \mathcal{W}_n$ for each $0 \le t \le 1$, it follows that

$$tp(A, X) + tp(A, Y) \succeq p(A,\ tX + (1 - t)Y).$$

1.3.4. Main convexity results. This subsection contains most of the main results on partially convex polynomials.

Theorem 1.4. *Let $\mathcal{W} = (\mathcal{W}_n)_n$ be an open domain in $\mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$. If $p(a,x)$ is convex in x on $\mathcal{W}$, then*

$$p(a,x) = L(a,x) + V(a,x)^T Z(a) V(a,x),$$

where

(i) *$Z(a)$ is a (square) matrix-valued symmetric polynomial;*
(ii) *$Z(A) \succeq 0$ for each $A \in \pi_a(\mathcal{W})$;*
(iii) *$V(a,x)$ is a (column) vector whose entries are linear in x; and*
(iv) *$L(a,x)$ is a symmetric polynomials with degree at most one in x.*

It follows that $p(a,x)$ is convex in x on the product domain $\pi_a(\mathcal{W}) \times \mathbb{S}(\mathbb{R}^g)$ and has degree at most two in x.

Proof. See Corollary 3.6. □

If $Z(A) \succeq 0$ for all A, then, by a result in [9], Z factors as $Z = R^T R$ for a matrix-valued polynomial R giving the following variant of Theorem 1.4.

Theorem 1.5. *Suppose $p(a,x)$ is a symmetric polynomial and $\mathcal{U} = (\mathcal{U}_n)_n$ is an open domain in $\mathbb{S}(\mathbb{R}^g)$. If $p(a,x)$ is convex on the product domain $\mathbb{S}(\mathbb{R}^g) \times \mathcal{U}$, then*

$$p(a,x) = L(a,x) + \Lambda(a,x)^T \Lambda(a,x),$$

where

(i) *$\Lambda(a,x)$ is a (column) vector and is linear in x; and*
(ii) *$L(a,x)$ has degree at most one in x.*

Consequently, p has degree at most two in x and is globally convex in x.

The previous theorem can be used to deduce the structure of polynomials $p(a,x)$ which are either convex or concave in each variable separately.

Theorem 1.6. *Suppose $p(a,x)$ is symmetric. If p is convex in x and concave in a, then there exists linear (homogeneous of degree one) polynomials $r_j(x)$ and $s_l(a)$ and a polynomial $L(a,x)$ which has degree one in both x and a (so joint degree at most two) such that*

$$p(a,x) = L(a,x) + R(x)^T R(x) - S(a)^T S(a),$$

where $R(x)$ is the (column) vector with entries $r_j(x)$ and likewise for $S(a)$.

Proof. See Theorem 5.1. □

Theorem 1.7. *If $p(a,x)$ is (globally) convex in a and x separately, then there exists a polynomial $L(a,x)$ and and a (column) vector of polynomials $\Lambda(a,x)$ which has degree at most one in x and a separately (thus at most degree two jointly) such that*

$$p(a,x) = L(a,x) + \Lambda(a,x)^T \Lambda(a,x).$$

Here $\Lambda(a,x)$ is a column vector so that $\Lambda(a,x)^T\Lambda(a,x) = \sum \Lambda_j(a,x)^T\Lambda_j(a,x)$ is a sum of squares.

Proof. See Theorem 4.1. □

Remark 1.8. The converse to each of the theorems in this subsubsection is evidently true.

1.4. Positivity of the Hessian. Just as in the classical commutative case, for a (non-commutative) symmetric polynomial, convexity implies that the *Hessian*—a version of the second derivative—is positive semi-definite. Conversely, a variety of fairly weak positivity hypotheses on the Hessian impose very strong restrictions on the polynomial.

1.4.1. Definition of the partial Hessian. Let $p(a,x)$ be a given symmetric polynomial. The *partial Hessian of p* with respect to x in the direction $h = (h_1,\dots,h_g)$ is the formal second derivative of the polynomial in t, $\mathsf{p}(t) = p(a,\, x+th)$. Thus, the partial Hessian,

$$\frac{\partial^2}{\partial x^2}p(a,x)[h] = \mathsf{p}''(0)$$

is a polynomial in $3g$ variables and is homogeneous of degree two in h.

This partial Hessian can also be described algebraically on a monomial $m(a,x)$ by replacing each pair of variables x_j and x_k with h_j and h_k respectively and then multiplying by two. In particular, the partial Hessian of monomials of degree zero or one in x is 0; and the partial Hessian of a monomial of degree two in x is simply that monomial multiplied by 2 with x replaced by h.

A simple observation, which will be used repeatedly without further comment, is that if $p(a,x)$ has degree two in x then $(\partial^2/\partial x^2)p(a,x)[h]$ depends only on a and h, and moreover, there is a polynomial $L(a,x)$ of degree at most one in x so that

$$p(a,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}p(a,x)[x] + L(a,x). \tag{1.3}$$

Moreover, this representation is unique in the sense that the Hessian is the part of $p(a,x)$ which is homogeneous of degree two in x and $L(a,x)$ contains the remaining part of p. In particular, there can be no cancellation between these two terms.

Examples will be given in Section 2.1.

1.4.2. Convexity and the Hessian The following Proposition says that convexity of p on an open in x domain $\mathcal{D}$ implies positivity of the partial Hessian on $\mathcal{D}$.

Proposition 1.9. *Suppose $p(a,x)$ is a symmetric polynomial, $A \in \mathbb{S}_n(\mathbb{R}^g)$, and $\mathcal{U}$ is an open convex set in $\mathbb{S}_n(\mathbb{R}^g)$. If $p(A,X)$ is convex for $X \in \mathcal{U}$, then, for each $X \in \mathcal{U}$ and $H \in \mathbb{S}_n(\mathbb{R}^g)$,*

$$\frac{\partial^2}{\partial x^2} p(A,X)[H] \succeq 0.$$

Explicitly, the hypothesis on p is that for the given A and each $X,\ Y \in \mathcal{U}$, and $0 < t < 1$,

$$p(A,\ tX+(1-t)Y) \preceq tp(A,X) + (1-t)p(A,Y).$$

Proof. Fix $X \in \mathcal{U}$. For a given $H \in \mathbb{S}_n(\mathbb{R}^g)$ and $s \in \mathbb{R}$ small, $X \pm sH \in \mathcal{U}$. By the convexity hypothesis on p,

$$0 \preceq (p(A,\ X+sH) + p(A,\ X-sH)) - 2p(A,X) = s^2 \frac{\partial^2}{\partial x^2} p(A,X)[H] + o(s^4).$$

Dividing by s^2 and letting s tend to 0 gives the result. ◻

Positivity of the Hessian not only implies convexity of p, but also imposes further serious constraints.

Theorem 1.10. *Suppose $p(a,x)$ is a symmetric polynomial of degree d_a and $d_x \geq 2$ in a and x respectively, and $\mathcal{W} \subset \mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$ is a domain. If*

(i) *$\mathcal{W}$ is open in a; and*
(ii) *for each n and each $(A,X) \in \mathcal{W}_n$ and every $H \in \mathbb{S}_n(\mathbb{R}^g)$,*

$$\frac{\partial^2}{\partial x^2} p(A,X)[H] \succeq 0,$$

then, either

(A) *there is a nonzero polynomial q of degree at most d_a and $d_x - 2$ in a and x respectively, so that $q(A,X) = 0$ on $\mathcal{W}$;* or
(B) *p has degree at most two in x, and for each $A \in \pi_a(\mathcal{W}_n)$, the function $p(A,X)$ is (globally) convex in $X \in \mathbb{S}_n(\mathbb{R}^g)$.*

In particular, if $\mathcal{W}$ is the product $\mathcal{W} = \mathbb{S}(\mathbb{R}^g) \times \mathcal{V}$, for some matrix convex set $\mathcal{V} = (\mathcal{V}_n)_n$, then either

(A′) *there is a polynomial q of degree at most d_a and $d_x - 2$ in a and x respectively so that $q(A,X) = 0$ for all $(A,X) \in \mathcal{W}$;* or
(B′) *there exists $\Lambda(a,x)$, linear in x, and $L(a,x)$ a polynomial of degree at most one in x so that*

$$p(a,x) = L(a,x) + \Lambda(a,x)^T \Lambda(a,x).$$

Proof. See Section 3. ◻

Proposition 3.1, which isolates the role of positivity of the Hessian, is a not too technical, but still widely applicable, ingredient in the proof of Theorem 1.10.

1.5. Further results and organization of the paper. The proofs of the main results turn on the *border vector-middle matrix* representation of the partial Hessian. This representation and some examples are discussed in the next section, Section 2.

The proofs of Proposition 3.1, Theorem 1.10 and Corollaries thereof are in Section 3. The proofs of the results in Subsubsection *1.3.4* are in Sections 4 and 5.

A structure result for the middle matrix of the partial Hessian which is of independent interest and will likely be a valuable tool in further investigations is presented in Section 6 which can be read following Section 2.

2. THE MIDDLE MATRIX OF THE PARTIAL HESSIAN

A polynomial $q(a,x)[h]$ in the non-commuting variables a, x and h which is homogeneous of degree two in h is conveniently represented in terms of a Gram-type representation in the form

$$q(a,x)[h] = V(a,x)[h]^T Z(a,x) V(a,x)[h],$$

where $Z(a,x)$ is a symmetric matrix with polynomial entries, and $V(a,x)[h]$ is a vector, called the *border vector*, whose entries have the form $h_j m(a,x)$ for monomials m of degree at most the degree of q in (a,x). Thus, we can index the entries of $Z := Z(a,x)$ by the monomials $h_j m(a,x)$. The notation $Z(\cdot,\cdot)$ will be used to denote the matrix Z, as well as the entries of Z. For example, $Z(h_j m(a,x), h_k m'(a,x))$ denotes the entry of Z, corresponding to entries $h_j m(a,x)$ and $h_k m'(a,x)$ of the border vector. Context will make clear which of the two uses of the $Z(\cdot,\cdot)$ notation is being employed.

The matrix $Z = Z(a,x)$ is known as the *middle matrix* and is unique up to order and presence of zero rows and columns. When q is symmetric, so is $Z(a,x)$.

Of special interest is the case that $p(a,x)$ is symmetric and $q(a,x)[h] = (\partial^2/\partial x^2)p(a,x)[h]$. A central object in the analysis to follow is the *derived* middle matrix,

$$\mathcal{Z}(a) = Z(a,0).$$

Suppose $p(a,x)$ is a polynomial of degree d_x in x and d_a in a. Let Z be the middle matrix for the partial Hessian of p. We may take the border vector for Z to be of the form

$$V = \begin{bmatrix} V_0 \\ \vdots \\ V_{d_x-2} \end{bmatrix},$$

where V_j is a vector whose entries consist of words of the form

$$h_* m_0 x_{k_1} m_1 \cdots m_{j-1} x_{k_j} m_j,$$

where $m_0, \dots, m_j$ are words in a such that $\deg_a m_0 + \cdots + \deg_a m_j \le d_a$. That is, V_j captures those monomials in x and a of total degree at most d_a in a and degree exactly j in x. Thus V_j is a vector of height

$$N_j := \sum_{n_0+\cdots+n_{j+1} \le d_a} g_a^{n_0+\cdots+n_{j+1}+j+1}$$

and Z has $(d_x - 2)(N_0 + \cdots + N_{d_x-2})$ rows and columns. With respect to this block form of V, we can write Z in block form $Z = [Z_{ij}]$, for $i, j = 1, \dots, d_x - 2$. Thus, Z_{ij} is that part of Z which corresponds to terms the V_i part of the border vector on the left and the V_j part on the right.

2.1. A few simple examples.

Example 2.1. Let $p = ax^3a$. Then $q(a,x)[h] = 2[ahxha + ah^2xa + axh^2a]$, which is equal to

$$\begin{bmatrix} 2h & ah & xh & axh & xah \end{bmatrix} \begin{bmatrix} 0&0&0&0&0 \\ 0&x&0&1&0 \\ 0&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&0&0&0 \end{bmatrix} \begin{bmatrix} h \\ ha \\ hx \\ hxa \\ hax \end{bmatrix}$$

Example 2.2. The polynomial $p = ax^3 + x^3a$ has Hessian $2[ahxh + ah^2x + axh^2 + hxha + h^2xa + xh^2]$ which has the representation

$$\begin{bmatrix} h & ah & xh & axh & xah \end{bmatrix} 2 \begin{bmatrix} 0&x&0&1&0 \\ x&0&1&0&0 \\ 0&1&0&0&0 \\ 1&0&0&0&0 \\ 0&0&0&0&0 \end{bmatrix} \begin{bmatrix} h \\ ha \\ hx \\ hxa \\ hax \end{bmatrix}$$

Example 2.3. Let $p = x^2ax + xax^2$. Then $q(a,x)[h] = 2[hxah + h^2ax + xhah + haxh + xah^2 + hahx]$, which equals

$$\begin{bmatrix} h & ah & xh & axh & xah \end{bmatrix} 2 \begin{bmatrix} xa+ax & 0 & a & 0 & 1 \\ 0 & 0&0&0&0 \\ a & 0&0&0&0 \\ 0 & 0&0&0&0 \\ 1 & 0&0&0&0 \end{bmatrix} \begin{bmatrix} h \\ ha \\ hx \\ hxa \\ hax \end{bmatrix}$$

Recall that Z_{00} is that part of the middle matrix which corresponds to those terms of the border vector with no x variables (both sides). Thus, in this last example we have

$$Z_{00} = 2 \begin{bmatrix} xa+ax & 0 \\ 0 & 0 \end{bmatrix},$$

and likewise

$$Z_{01} = 2\begin{bmatrix} a & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

and the derived matrix is

$$\mathcal{Z} = \mathcal{Z}(0,a) = 2\begin{bmatrix} 0 & 0 & a & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

2.2. Faithfulness. The following standard lemma says more than the totality of matrix evaluations are faithful on $\mathbb{R}\langle a,x\rangle$. Given g and d, let

$$N(g,d) = \sum_0^d g^j. \tag{2.1}$$

Lemma 2.4. *Suppose $p(y)$ is a polynomial of degree d in g variables. If p vanishes on an open subset $U \subset \mathbb{S}_{N(g,d)}(\mathbb{R}^g)$, then $p = 0$.*

Proof. See Lemma 2.2 in [7]. □

3. The Structure of Partially Convex Polynomials

In this section we consider positivity conditions on the partial Hessian of a symmetric polynomial p which impose strong constraints on the form of p.

We begin with Proposition 3.1 below, which isolates the role of direct sums and positivity of the Hessian. It is a not too technical, but still widely applicable, ingredient in the proof of Theorem 1.10. Given $v \in \mathbb{R}^n$ and natural number ℓ, let $[v]_\ell$ denote the vector in $\mathbb{R}^{n\ell} = (\mathbb{R}^n)^\ell$ with each of its ℓ block (of length n) entries equal to v.

Proposition 3.1. *Let $p = p(a,x)$ be a given symmetric polynomial of degree d_a and d_x in a and x respectively. Suppose $n \geq N(g,d_a)$, $v \in \mathbb{R}^n$, $\chi \in \mathbb{S}_n(\mathbb{R}^g)$ and that U is an open set in $\mathbb{S}_n(\mathbb{R}^g)$. If*

(i) *the set $\{m(A,\chi)v : \deg_x m \leq d_x - 2 \text{ and } \deg_a m \leq d_a\}$ is linearly independent in $\mathbb{R}^n$ for each $A \in U$; and*
(ii) *for each natural number ℓ and $A \in U$,*

$$0 \leq \left\langle \frac{\partial^2}{\partial x^2} p(A \otimes I_\ell,\ \chi \otimes I_\ell)[H][v]_\ell, [v]_\ell \right\rangle$$

for all $H \in S_{n\ell}(\mathbb{R}^g)$,

then

(a) *the degree of p in x is at most two;*
(b) *the partial Hessian of p takes the form*

$$\frac{\partial^2}{\partial x^2} p(a,x)[h] = (V(a)[h])^T Z_{0,0}(a) V(a)[h];$$

(c) $Z_{0,0}(A) \succeq 0$ *for* $A \in U$*; and*
(d) *there is a symmetric polynomial* $L(a,x)$ *which has degree at most one in* x *so that*

$$p(a,x) = \frac{1}{2} V(a)[x]^T Z_{0,0}(a) V(a)[x] + L(a,x).$$

In particular, $p(A,X)$ *is (globally) convex in* $X \in \mathbb{S}_n(\mathbb{R}^g)$ *for each fixed* $A \in U$.

Proof. The partial Hessian can be represented in terms of the middle matrix $Z(a,x)$ as

$$\frac{\partial^2}{\partial x^2} p(a,x)[h] = V(a,x)[h]^T Z(a,x) V(a,x)[h].$$

Given ℓ, let $[v]_\ell$ denote the vector in $\mathbb{R}^{n\ell} = (\mathbb{R}^n)^\ell$ with each of its ℓ block (of length n) entries equal to v. Since the set $\{m(A \otimes I_\ell, \chi \otimes I_\ell)[v]_\ell : \deg_x m \leq d_x - 2 \text{ and } \deg_a m \leq d_a\}$ is linearly independent, the [3] Lemma implies that the subspace $W = \{V(A \otimes I_\ell, \chi \otimes I_\ell)[H][v]_\ell : H \in S_{n\ell}(\mathbb{R}^{g_x})\}$ has codimension at most κ in $\mathbb{R}^{nN(g,d)}$ independent of n. (See the appendix, Section 8 for the details.) Choose ℓ so that $\kappa < \ell$. With this fixed ℓ, to simplify the notations, let $\tilde{A} = A \otimes I_\ell$; $\tilde{\chi} = \chi \otimes I_\ell$; and $\tilde{v} = [v]_\ell$.

By positivity of the partial Hessian, it follows that $Z(\tilde{A}, \tilde{\chi})$ is positive when compressed to W. It follows that $Z(\tilde{A}, \tilde{\chi})$ has at most $\kappa < \ell$ negative eigenvalues for each $A \in U$.

We now partition $V(a,x)[h]$ and $Z(a,x)$ in blocks by the degree of x and let $Z_{jl}(a,x)$ denote the part of $Z(a,x)$ corresponding to the terms of degree j and l in x. In particular, $Z_{0,d_x-2}(a,x)$ represents those terms in the Hessian of the form $m_l(a)h_* s(a,x) h_* m_r(a,x)$, where $m_r(a,x)$ has degree $d_x - 2$ in x and where m_r, m_l, s are monomials. (Here we are assuming the degree in x is d_x, if it is lower, make the obvious adjustments). It follows that $Z_{0,d_x-2}(a,x) = Z_{0,d_x-2}(a)$ is independent of x.

Now $Z_{0,d_x-2}(a)$ is a matrix with entries which are polynomials in a. Since $n \geq N(g, d_a)$ and U is an open subset of $\mathbb{S}_{N(g,d)}(\mathbb{R}^{g_a})$, it follows from Lemma 2.4 that $Z_{0,d_x-2}(A) \neq 0$ for some (and hence most) $A \in U$. Thus, $Z_{0,d_x-2}(\tilde{A})$ has rank at least ℓ. Hence, from the block structure of Z, if $d_x - 2 > 0$, then $Z(\tilde{A}, \tilde{\chi})$ has at least ℓ negative eigenvalues (see Corollary 5.4 in [5]), contradicting $\kappa < \ell$. We conclude that $d_x \leq 2$.

Now that we know the degree of p is two in x, the middle matrix $Z(a,x) = Z_{0,0}(a)$. Which gives the advertised representation.

For the last part of the theorem, we argue as above and find that $Z(\tilde{A}) = Z_{0,0}(\tilde{A}) = Z(A) \otimes I_\ell$ has at most $\kappa < \ell$ negative eigenvalues. It follows that $Z(A)$ can have no negative eigenvalues as otherwise $Z(\tilde{A})$ has at least ℓ. Thus $Z(A) \succeq 0$.

Since $Z(A) \succeq 0$ and (since p has degree at most two in x),

$$p(a,x) = V(a)[x]^T Z(a) V(a)[x] + L(a,x)$$

it follows that, with A fixed, that $p(a,x)$ is convex in $X \in \mathbb{S}_n(\mathbb{R}^g)$. □

Remark 3.2. As seen in the proof, in item (ii) the for every ℓ can be replaced with some $\ell \geq \kappa + 1$ where κ is an integer which depends only upon d and g.

In view of the remark and the following Lemma, hypothesis (ii) of Proposition 3.1 is stronger than needed in that it suffices to tensor with I_ℓ for certain ℓ, thereby weakening the need to tensor with I_ℓ at all.

Lemma 3.3 (R. Guralnick and L. Small)**.** *Let $q \neq 0$ be a given (not necessarily symmetric) polynomial of degree d in g variables. Then there is a sequence of integers $n_k \to \infty$ for which*

$$\{A \in \mathbb{S}_{n_k}(\mathbb{R}^g) : q(A) \text{has full rank } \}$$

is open and dense in $\mathbb{S}_{n_k}(\mathbb{R}^g)$.

See Appendix 9 for a proof.

The proof of Theorem 1.10 combines Proposition 3.1 and the following two lemmas.

Lemma 3.4. *Suppose $\mathcal{W} \subset \mathbb{S}_n(\mathbb{R}^g \times \mathbb{R}^g)$ is a domain and d_a and d_x are natural numbers. Either there exists an n and $(A,X) \in \mathcal{W}_n$ and $v \in \mathbb{R}^n$ such that the set $\{m(A,X)v : m$ is a monomial with $\deg_x m \leq d_x - 2,\ \deg_a m \leq d_a\}$ is linear independent, or there is a polynomial q (not necessarily symmetric) of degree at most d_a and d_x in a and x respectively so that $q(A,X) = 0$ for all $(A,X) \in \mathcal{W}$.*

Proof. A proof can be found in [4][Lemma 4.1]. □

This Lemma naturally combines with the following simple observation.

Lemma 3.5. *Suppose $\mathcal{W} \subset \mathbb{S}_n(\mathbb{R}^g \times \mathbb{R}^g)$ is a domain and and natural numbers d_a and d_x are given. If there exists n and $(A,X) \in \mathcal{W}$ and $v \in \mathbb{R}^n$ such that $\{m(A,X)v : m$ is a monomial with $\deg_x m \leq d_x - 2,\ \deg_a m \leq d_a\}$ is linearly independent, then for each $(B,Y) \in \mathcal{W}$ and vector w (of the correct size), the set $\{m(A \oplus B, X \oplus Y)(v \oplus w) : m$ is a monomial with $\deg_x m \leq d_x - 2,\ \deg_a m \leq d_a\}$ is linearly independent.*

Proof of Theorem 1.10. From Lemma 3.4 either there is a (not necessarily symmetric) polynomial q of degree at most d_a and $d_x - 2$ in a and x respectively so that $q(A,X) = 0$ on $\mathcal{W}$, or there is an n, a pair $(B,Y) \in \mathcal{W}_n$ and vector $u \in \mathbb{R}^n$

so that the set $\{m(B,Y)u : \deg_x m \leq d_x - 2 \text{ and } \deg_a m \leq d_a\}$ is linearly independent.

Given $(C,Z) \in \mathcal{W}$, let $(A',\chi) = (B \oplus C, Y \oplus X) \in \mathcal{W}$ and $v = 0 \oplus u$. By Lemma 3.5, the set $\{m(A',X)v : \deg_x m \leq d_x - 2 \text{ and } \deg_a m \leq d_a\}$ is linearly independent. In particular, by the open in a hypothesis on $\mathcal{W}$, there is a neighborhood U of A' such that for all $A \in U$ the set $\{m(A,X)v : \deg_x m \leq d_x - 2 \text{ and } \deg_a m \leq d_a\}$ is linearly independent. Since $\mathcal{W}$ is also a domain on which the Hessian is non-negative, Theorem 3.1 applies with the conclusion that p has degree at most two in x and

$$p(a,x) = \frac{1}{2}V(a)[x]^T Z(a)V(a)[x] + L(a,x),$$

where $L(a,x)$ has degree at most one in x, and $Z(A) \succeq 0$ for $A \in U$. In particular, $Z(B \oplus C) \succeq 0$. Therefore $Z(C) \succeq 0$ and $p(C,X)$ is convex (globally) in X for each $C \in \pi_a(\mathcal{W})$.

For the second part of the corollary, note that in the absence of the polynomial q, the first part of the corollary implies that $Z(A) \succeq 0$ for all A. Thus, by a Theorem in [9], $Z(A)$ factors as a SoS and the result follows. □

We close this section by pointing out the following simple special cases of Theorem 1.10.

Corollary 3.6. *If $p(a,x)$ is convex in x on an open domain $\mathcal{W} \subset \mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$, then*

$$p(a,x) = L(a,x) + V(a,x)^T Z(a)V(a,x),$$

where

(i) *$Z(A) \succeq 0$ for $A \in \pi_a(\mathcal{W})$;*
(ii) *$V(a,x)$ is the border vector, which is linear in x; and*
(iii) *$L(a,x)$ has degree at most one in x.*

Proof. Note that Lemma 2.4 rules out the possibility that there is a nonzero polynomial q such that $q(A,X) = 0$ for all $(A,X) \in \mathcal{W}$. The convexity implies that $(\partial^2/\partial x^2)p(A,X) \succeq 0$ for all (A,X) in the open domain $\mathcal{W}$. Consequently, the conclusion of the Lemma follows from the argument given for the proof of Theorem 1.10. □

Corollary 3.7. *Suppose $\mathcal{V} \subset \mathbb{S}(\mathbb{R}^g)$ is an open matrix convex set and let $\mathcal{W} = \mathbb{S}(\mathbb{R}^g) \times \mathcal{V}$. If $p(a,x)$ is convex in x on $\mathcal{W}$, then*

$$p(a,x) = L(a,x) + \Lambda(a,x)^T\Lambda(a,x),$$

where

(i) *$\Lambda(a,x)$ is a vector and is linear in x; and*
(ii) *$L(a,x)$ has degree at most one in x.*

Proof. As in the previous corollary, there does not exist a nonzero polynomial q such that $q(A,X) = 0$ for all $(A,X) \in \mathcal{W}$. Thus, option (B') of Proposition 1.10 occurs. □

4. SEPARATE CONVEXITY

Theorem 4.1. *If p is (globally) convex in a and x separately, then there exists an m and polynomials $L(a,x)$ and $\Lambda_j(a,x)$, $j = 1,2,\dots,m$, which are degree (at most) one in x and a separately (thus at most degree two jointly) such that*

$$p(a,x) = L(a,x) + \Lambda(a,x)^T\Lambda(a,x).$$

Here we have used the shorthand, $\Lambda(a,x)^T\Lambda(a,x) = \sum \Lambda_j(a,x)^T\Lambda_j(a,x)$.

We begin with a lemma.

Lemma 4.2. *Suppose $p(a,x) = q(a,x) + \sum r_j(a,x)^T r_j(a,x)$. If*

(i) *each $r_j(a,x)$ is homogeneous of degree one in x;*
(ii) *$q(a,x)$ is degree one in x; and*
(iii) *$p(a,x)$ has degree at most two in a,*

then each $r_j(a,x)$ has degree at most one in a and $q(a,x)$ has degree at most two in a.

Proof. Terms from $q(a,x)$ cannot cancel those from

$$s(a,x) = \sum r_j(a,x)^T r_j(a,x),$$

since the former are of at most degree one in x and the later homogeneous of degree two in x. Since p has degree two in a and there can't be cancellation of the highest degree term in a in the sum of squares term s, each $r_j(a,x)$ has degree at most one in a. Likewise, $q(a,x)$ has degree at most two in a. □

Let $\mathcal{J}$ denote those monomials in x and a which are linear in each of x and a separately. Thus $\mathcal{J}$ has $2g^2$ elements. Let $V(a,x)$ denote the tautological vector whose entries are the monomials from $\mathcal{J}$. Given a $\mathcal{J} \times \mathcal{J}$ matrix M, the expression,

$$s(a,x) = V(a,x)^T M V(a,x) = \sum_{m,\ell} M_{m,\ell}\, m(a,x)\ell(a,x) \tag{4.1}$$

is then a polynomial which is homogeneous of degree two in each of x and a separately.

If $M = R^T R$, then

$$s(a,x) = \sum r_j(a,x)^T r_j(a,x) \tag{4.2}$$

where $r_j(a,x) = R_j V(a,x)$ and R_j is the j-th row of R.

Conversely, if s has the form in equation (4.1), and $r_j = \sum r_j(m)m(a,x)$ (where the sum is over $m \in \mathcal{J}$), then s has the form in equation (4.1 with $M = R^T R$, where R is the matrix whose j-th row has entries $R_j(m) = r_j(m)$.

Lemma 4.3. *Suppose $s(a,x)$ is homogeneous of degree two in each of x and a separately. If A and A' are $n \times \mathcal{J}$ and $n' \times \mathcal{J}$ matrices respectively such that*

$$s(a,x) = V(a,x)^T A^T AV(a,x) = V(a,x)^T (A')^T A' V(a,x),$$

then there is a partial isometry $U : \mathbb{R}^n \to \mathbb{R}^{n'}$ such that $A = UA'$.

Proof. It is readily verified that

$$A^T A = (A')^T A'. \tag{4.3}$$

The existence of U now follows from the Douglas Range Inclusion Theorem. For convenience of the reader we give the argument in this case. Define U on the range of A into the range of A' by $UAm = A'm$ (here m is the vector with a 1 in the m-th place and 0 elsewhere). The equality (4.3) implies U unitary and thus extends to a partial isometry (by defining it to be 0 on the orthogonal complement of the range of A). ◻

Similarly, a polynomial $p(x)$ which homogeneous of degree two in x alone (no a) can be represented as

$$p(x) = V(x)^T PV(x),$$

where $V(x)$ is now the vector with entries x_j. The polynomial p is a sum of squares if and only if $P = R^T R$ for some R (so if and only if P is positive semi-definite). In particular, we are using V in two different ways which should cause no confusion.

Proof of Theorem 4.1. Since p is convex in x, Corollary 1.5 implies p can be written in the form,

$$p(a,x) = \mathsf{L}(a,x) + \mathsf{H}(a,x)^T \mathsf{H}(a,x),$$

where $\mathsf{H}(a,x)$ is linear in x and $\mathsf{L}(a,x)$ has degree at most one in x. Here H is a vector with entries H_j.

Since p is convex in a it has degree at most two in a. Thus, Lemma 4.2 says $\mathsf{H}(a,x)$ has degree at most one in a and $\mathsf{L}(a,x)$ has degree at most two in a, in addition to the degree restrictions relative to x above.

Write $\mathsf{H}_\mathsf{j}(a,x) = a_j(a,x) + b_j(x)$ with $a_j(a,x)$ homogeneous of degree one in a and $b_j(x)$ a polynomial in x alone. Similarly, since L has degree at most one in x and two in a, it can be written as

$$\mathsf{L}(a,x) = C(a,x) + D(a) + E(a,x)$$

where $C(a,x)$ is homogeneous of degree two in a and one in x; $D(a,x)$ is homogeneous of degree two in a (and has no x); and E has degree at most one in each of x and a.

Now let A and B respectively denote matrices which produce the representations

$$\begin{aligned}\sum a_j^T a_j &= V(a,x)^T A^T A V(a,x)\\ \sum b_j^T b_j &= V(x) B^T B V(x).\end{aligned}$$

We have,

$$\begin{aligned}p(a,x) = V(a,x)^T A^T A V(a,x) + [V(a,x)A^T B V(x) + V(x)^T B^T A V(a,x)]\\ + C(a,x) + V(x) B^T B V(x) + D(a) + E(a,x).\end{aligned}$$

Note that the term $[\cdot\,]$ is the part of $p(a,x)$ which is homogeneous of degree two in x and one in a; whereas $C(a,x)$ is the part homogeneous of degree one in x and two in a.

Reversing the roles of x and a, p can also be written as

$$\begin{aligned}p(a,x) &= V(a,x)^T (A')^T A' V(a,x)\\ &\quad + [V(a,x)(A')^T B' V(a) + V(a)^T (B')^T A' V(a,x)]\\ &\quad\quad + C'(a,x) + V(a)(B')^T B' V(a) + D'(x) + E'(a,x).\end{aligned}$$

Note that the term $[\cdot]$ is the part of $p(a,x)$ which is homogeneous of degree two in a and one in x; whereas $C'(a,x)$ is the part homogeneous of degree one in a and two in x.

Comparing these last two representations we find,

$$\begin{aligned}V(a,x)A^T A V(a,x) &= V(a,x)(A')^T A' V(a,x)\\ V(a,x)A^T B V(x) + V(x)^T B^T A V(a,x) &= C'(a,x)\\ V(a,x)(A')^T B' V(a) + V(a)^T (B')^T A' V(a,x) &= C(a,x)\\ V(a)^T (B')^T B' V(a) &= D(a).\end{aligned}$$

From Lemma 4.3 there is a partial isometry U so that $A' = UA$. Choose W so that $I - UU^T = WW^T$. Consider,

$$\begin{aligned}&(AV(a,x) + BV(x) + U^T B' V(a))^T (AV(a,x) + BV(x) + U^T B'(a) V(a))\\ &\quad + (WB'V(a))^T (WB'V(a))\\ &\quad - [V(a)^T (B')^T U B V(x) + V(x)^T B^T U^T B' V(a)] + E(a,x) \quad =\end{aligned}$$

$$
\begin{aligned}
&= (AV(a,x)+BV(x))^T(AV(a,x)+BV(x)) \\
&\quad + V(a,x)^T A^T U^T B' V(a) + V(a)^T (B')^T U A V(a,x) \\
&\quad + V(a)^T B^T U U^T B V(a) + V(a)^T (B')^T W^T W B' V(a) \\
&= (AV(a,x)+BV(x))^T(AV(a,x)+BV(x)) \\
&\quad + [V(a,x)^T (A')^T B' V(a) + V(a)^T (B')^T A' V(a,x)] \\
&\quad + V(a)^T B^T B V(a) + E(a,x) \\
&= (AV(a,x)+BV(x))^T(AV(a,x)+BV(x)) + C(a,x) + D(a) + E(a,x) \\
&= p(a,x).
\end{aligned}
$$

□

5. Convex in x and Concave in a

Theorem 5.1. *Suppose $p(a,x)$ is symmetric. If p is convex in x and concave in a, then there exists linear (homogeneous of degree one) polynomials $r_j(x)$ and $s_l(a)$ and a polynomial $L(a,x)$ which has degree one in both x and a (so joint degree at most two) such that*

$$p(a,x) = L(a,x) + R(x)^T R(x) - S(a)^T S(a),$$

where $R(x)^T R(x)$ and $S(a)^T S(a)$ are shorthand for the sums $\sum r_j(x)^T r_j(x)$ and $\sum s_j(a)^T s_j(a)$ respectively.

Proof. Since p is (globally) convex in x, it can be written in the form,

$$p(a,x) = \Lambda(a,x)^T \Lambda(a,x) + L(a,x),$$

where Λ is linear and L has degree at most one in x. Thus, $\Lambda(a,x)^T\Lambda(a,x)$ is homogeneous of degree two in x; whereas $L(a,x)$ had degree at most one in x. In particular, there can be no cancellation between these terms.

Since p is (globally) concave in a, it has degree at most two in a and since the terms in $\Lambda(a,x)^T\Lambda(a,x)$ can not cancel with those in $L(a,x)$, it thus follows that $\Lambda(a,x)$ has degree at most one in a and likewise $L(a,x)$ has degree at most two in a.

Write,

$$L(a,x) = L_0(x) + L_1(a,x) + L_2(a,x),$$

where L_j is homogeneous of degree j in a (and degree at most one in x). Similarly, write $\Lambda(a,x) = \Lambda_0(x) + \Lambda_1(a,x)$, with Λ_j homogeneous of degree j in a.

Taking the partial Hessian of p with respect to a gives,

$$\frac{\partial^2}{\partial a^2} p(a,x)[k] = 2[\Lambda_1(k,x)^T \Lambda_1(k,x) + L_2(k,x).$$

Since $(\partial^2/\partial a^2)p(a,x)[k]$ is negative semi-definite (in k for each (a,x)), and since the first term above is homogeneous of degree two in k while the second has degree at most one in k, it follows that $\Lambda_1(k,x) = 0$ and $L_2(k,x)$ is negative semi-definite. Since $L_2(k,x)$ has degree at most one in x and is negative semi-definite, it does not depend on x; $L_2(a,x) = L_2(a)$. Further, $L_2(a)$ is negative semi-definite and hence can be written as $-S(a)^T S(a)$. □

We anticipate that many of the results in this section and the last section will *localize* as the following theorem illustrates. By $\|a\| < 1$ we mean the open matrix convex domain with $(\{A \in \mathbb{S}_n(\mathbb{R}^g) : \sum A_j^2 < I_n\})_n \subset \mathbb{S}(\mathbb{R}^g)$.

Theorem 5.2. *Suppose $p(a,x)$ is symmetric. If*

(i) *$p(a,x)$ is convex in x for $\|a\| < 1$;*
(ii) *$p(a,x) \succeq 0$ for $\|A\| < 1$ and all X; and*
(iii) *p is concave in a,*

then p has the form,

$$p(a,x) = W(x)^T (R(a) - Q(a))W(x),$$

where

(a) *$W(x)$ is the vector with $g+1$ entries the monomials $\{\varnothing, x_1, \dots, x_g\}$;*
(b) *$R(a)$ and $Q(a)$ are symmetric matrix polynomials in a;*
(c) *$R(a)$ has degree at most one in a;*
(d) *$Q(a)$ is homogeneous of degree two and $Q(A) \succeq 0$ for every A (and so $Q(a)$ is a sum of squares); and*
(e) *$R(A) - Q(A) \succeq 0$ for $\|A\| < 1$.*

Note that the converse is true too; i.e., if p has the form above and (a)–(e) are satisfied, then p satisfies (i)–(iii).

Proof. The convexity in x and concavity in a are enough to imply that p has degree at most two in each of x and a (so at most degree four). Further, by convexity in X for $\|A\| < 1$, it follows from Corollary 3.6 that

$$p(a,x) = \frac{1}{2}V(a)[x]^T Z(a)V(a)[x] + L(a,x),$$

where $Z(A) \succeq 0$ for $\|A\| < 1$ and $V(x)[h]$ is the border vector (with respect to x) for $p(x,a)$, which is linear in x (homogeneous of degree one) and $L(a,x)$ has degree at most one in x. In particular, by considering the differing degrees of the x terms, there can be no cancellation between the two terms. Because of this lack of cancellation, both terms have degree at most two in a.

Represent $V(x)[a] = V_0(x) \oplus V_1(x)[a]$ and decompose Z with respect to this direct sum as

$$Z(a) = \begin{pmatrix} M_{00}(a) & M_{01}(a) \\ M_{01}^T(a) & M_{11} \end{pmatrix}.$$

Note that M_{11} is constant, $M_{01}(a)$ has degree at most one, and $M_{00}(a)$ has degree at most two. Since $Z(A) \succeq 0$ for $\|A\| < 1$, it follows that $M_{11} \succeq 0$.

We now take the Hessian of p with respect to a,

$$\frac{1}{2}\frac{\partial^2}{\partial a^2} p(a,x)[k] = V(x)[k]^T \begin{pmatrix} 2M''_{00}[k] & 2M'_{01}[k] \\ M'_{10}[k] & 2M_{11} \end{pmatrix} V(x)[k] + \frac{\partial^2}{\partial a^2} L(x)[k].$$

Replacing x with tx, choosing t large, and using the hypothesis that $p(a,x)$ is concave in a, so that $(\partial^2/\partial a^2)p(A,X)[K] \preceq 0$ for all choices of X, A, K it follows that the first term above is negative semidefinite; i.e.,

$$V(X)[K]^T \begin{pmatrix} M''_{00}[K] & M'_{01}[K] \\ M'_{10}[K] & M_{11} \end{pmatrix} V(X)[K] \preceq 0$$

for all X, A, K. Thus, by Lemma 8.3 (really it may first be necessary to take direct sums), the matrix

$$\begin{pmatrix} M''_{00}[K] & M'_{01}[K] \\ M'_{10}[K] & M_{11} \end{pmatrix}$$

is negative definite and, therefore, $M_{11} \preceq 0$. However, as noted above, $M_{11} \succeq 0$, and thus $M_{11} = 0$.

Since $Z(A) \succeq 0$ for $\|A\| < 1$ and $M_{11} = 0$, it follows that $M_{01} = 0$, too. We conclude that p has the form,

$$p(a,x) = V(x)^T Z(a) V(x) + L(a,x), \tag{5.1}$$

where $V(x)$ is the vector with entries $\{x_1, \dots, x_g\}$; $Z(a)$ (is a matrix and has) degree at most two; $Z(A) \succeq 0$ for $\|A\| < 1$; and $L(a,x)$ has degree at most two in a and at most one in x.

Coming at things from the other way, the fact that $p(a,x)$ is concave (globally)in a (for each x) implies that

$$p(a,x) = -\Lambda(a,x)^T \Lambda(a,x) + M(a,x), \tag{5.2}$$

where $\Lambda(a,x)$ is linear in a and $M(a,x)$ has degree at most one in a. As usual, the fact that p also has degree at most two in x implies that $\Lambda(a,x)$ has degree at most one in x and $M(a,x)$ has degree at most two in x.

Comparing the representations of Equations (5.1) and (5.2) with an eye toward the terms which are homogeneous of degree two in each of a and x, it follows that $\Lambda(a,x)$ is a linear combination of terms of the form a_j and $a_j x_\ell$. Consequently, there is a matrix-valued $Q(a)$ so that

$$\Lambda(a,x)^T \Lambda(a,x) = W(x)^T Q(a) W(x),$$

where $Q(a)$ is a sum of squares, and $W(x)$ is the vector (of polynomials) defined in the statement of the Theorem with entries $\{\emptyset, x_1, \dots, x_g\}$.

From Equation (5.1), $p(a,x)$ does not contain monomials of the form $x_j x_k a_\ell$ (or $a_\ell x_k x_j$). Hence, $M(a,x)$ can be written as

$$M(a,x) = W(x)^T R(a) W(x),$$

where $R(a)$ has degree at most one. Thus,

$$p(a,x) = W(x)^T (R(a) - Q(a)) W(x),$$

where R, Q satisfy the conditions (a) - (d).

To complete the proof we need to show $R(a) - Q(a) \succeq 0$ for $\|a\| < 1$.

For a given n, the set

$$\Gamma_n = \left\{ W(X)v = \begin{pmatrix} v \\ X_1 v \\ \vdots \\ X_g v \end{pmatrix} : v \in \mathbb{R}^n, \;\; X \in \mathbb{S}_n(\mathbb{R}^g), \;\; v \in \mathbb{R}^n \right\}$$

is dense in $\mathbb{R}^{(g+1)n}$. To prove this claim we follow the route of the CHSY-Lemma (see Appendix 8). Suppose first that $v = e_1$ and let

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_g \end{pmatrix} \in \mathbb{R}^{gn}$$

be given. Letting

$$X_j = \begin{pmatrix} w_{j1} & w_{j2} & \dots \\ w_{j2} & 0 & \dots \\ \vdots & 0 & 0 \end{pmatrix},$$

so that $X_j = X_j^T$ and $X_j e_1 = w_j$, it follows that

$$\Gamma_n \supset W(X)e_1 = \begin{pmatrix} e_1 \\ w \end{pmatrix}.$$

It follows that Γ_n contains all vectors of the form,

$$\begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_g \end{pmatrix} \in \mathbb{R}^{(g+1)n}$$

for which $w_0 \neq 0$ and the claim follows.

Finally, since $p(A,X) \succeq 0$ for $\|A\| < 1$ and all X, if $\|A\| < 1$ and $X \in \mathbb{S}_n(\mathbb{R}^g)$ and $v \in \mathbb{R}^n$, then

$$0 \leq p(A,X)v, v\rangle = \langle (Q(A) - P(A))W(X)v, W(X)v\rangle.$$

From the density of Γ_n it follows that $Q(A) - P(A) \succeq 0$ ☐

6. The Polynomial Congruence

In this section we establish a polynomial congruence between $Z(a,x)$ and $\mathcal{Z}(a)$ (the derived matrix, $\mathcal{Z}(a) = Z(a,0)$) in the case that $Z(a,x)$ is the middle matrix of the partial Hessian of a symmetric polynomial $p(a,x)$. The relation is analogous to that found in [5]

Suppose $p(a,x)$ is a polynomial of degree d_x in x and d_a in a. The partial Hessian of p can be written as a sum of terms of the form $m_L h_L r h_R m_R$, where $r = Z(m_L h_L, h_R m_R)$, and m_L and m_R are monomials in x and a. We use $(m_L h_L, h_R m_R)$ to index these terms. We will further write m_L (and similarly m_R) in the form

$$x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i},$$

where each m_{ℓ_s} is a monomial in a alone. So in particular, here m_L has degree i in x.

Let $Z = [Z_{ij}]$ be the middle matrix for the partial Hessian of p. Here the block structure indicated by Z_{ij} is determined by the the degree in x of the monomials in the middle matrix per the usual convention.

For each $j = 0, \ldots, d_x - 2$, let K_j be the $N_j \times N_{j-1}$ matrix with entries

$$K_j(h_{k_{j+2}} m_{k_{j+2}} x_{k_{j+1}} m_{k_{j+1}} \cdots x_{k_1} m_{k_1}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1}) = x_{k_{j+2}} m_{k_{j+2}},$$

and all others being 0.

Lemma 6.1. *Let p be a polynomial of degree $d_x \geq 2$ in x and of arbitrary degree in a. Let $Z = [Z_{ij}]$ be the middle matrix for the Hessian of p. Then $Z_{i,j+1}K_j + Z_{i,j}(a,0) = Z_{i,j}$ for $i = 0, \ldots, d_x - 2$, and $j = 0, \ldots, d_x - 3$, where $i + j \leq d_x - 2$.*

Proof. It suffices to prove the result for monomials, since it is evidently linear. Thus, suppose $i + j < d_x - 2$ and the monomial

$$x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}} r h_{k_{j+2}} m_{k_{j+2}} x_{k_{j+1}} m_{k_{j+1}} \cdots x_{k_1} m_{k_1}$$

is the $(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+2}} m_{k_{j+2}} x_{k_{j+1}} m_{k_{j+1}} \cdots x_{k_1} m_{k_1})$-term of the Hessian (so this is an entry of $Z_{i,j+1}$). Then the

$$(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1})$$

term will be

$$x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}} r x_{k_{j+2}} m_{k_{j+2}} h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1}.$$

In other words, $Z_{i,j}(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1})$ equals

$$Z_{i,j+1}(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+2}} m_{k_{j+2}} x_{k_{j+1}} m_{k_{j+1}} \cdots x_{k_1} m_{k_1}) x_{k_{j+2}} m_{k_{j+2}}.$$

Now, the $(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1})$-entry of the matrix $Z_{i,j+1} K_j$ is the product of

$$\text{row} \quad x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}} \quad \text{of } Z_{i,j+1},$$

and

$$\text{column} \quad h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1} \;\text{ of } K_j.$$

The only nonzero entry of column $h_{k_{j+1}} m_{k_{j+1}} x_{k_j} m_{k_j} \cdots x_{k_1} m_{k_1}$ of K_j is matrix

$$K_j(h_{k_{j+2}} m_{k_{j+2}} x_{k_{j+1}} m_{k_{j+1}} \cdots x_{k_1} m_{k_1}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1}) = x_{k_{j+2}} m_{k_{j+2}}.$$

Hence, the $(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1})$-entry of $Z_{i,j+1} K_j$ is

$$Z_{i,j+1}(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}} m_{\ell_{i+1}}, h_{k_{j+2}} m_{k_{j+2}} x_{k_{j+1}} m_{k_{j+1}} \cdots x_{k_1} m_{k_1}) x_{k_{j+2}} m_{k_{j+2}},$$

which equals

$$Z_{i,j}(x_{\ell_1} m_{\ell_1} \cdots x_{\ell_i} m_{\ell_i} h_{\ell_{i+1}}, h_{k_{j+1}} m_{k_{j+1}} x_{k_j} \cdots x_{k_1} m_{k_1}).$$

We conclude that $Z_{i,j+1} K_j + Z_{i,j}(0) = Z_{i,j}$ whenever $i + j < d_x - 2$.

If $i + j = d_x - 2$, then $Z_{i,j+1} = 0$ and $Z_{i,j} = Z_{i,j}(a, 0)$, so that

$$Z_{i,j+1} K_j + Z_{i,j}(a, 0) = Z_{i,j}.$$

Clearly the result also holds when $i + j > d_x - 2$. □

To illustrate, in Example 2.3, note that $Z_{00}(a, x) = Z_{01}(a, x) K_1 + Z_{00}(0, a)$, where

$$K_1 = \begin{bmatrix} x & 0 \\ 0 & 0 \\ xa & 0 \end{bmatrix}.$$

Theorem 6.2. *There is a matrix polynomial $A(a,x)$ so that*

$$Z(a,x)A(a,x) = Z(a,0).$$

Further, A has a square root B, so that $B^2 = A$, which is also a polynomial for which

$$B^T(a,x)Z(a,x)B(a,x) = Z(a,0).$$

Further, B is invertible (and its inverse is a polynomial). In particular, for any $(A,X) \in \mathbb{S}(\mathbb{R}^g \times \mathbb{R}^g)$, $B(A,X)$ is invertible.

With Lemma 6.1 in place, the proof of Theorem 6.2 follows along the lines of the proof of Theorem 7.3 in [5]. Let $A(a,x)$ denote the matrix

$$A(a,x) := \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ -K_0 & I & \cdots & 0 & 0 \\ 0 & -K_1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \\ 0 & 0 & \cdots & -K_{d_x-3} & I \end{bmatrix}. \tag{6.1}$$

and observe $Z(a,x)A(a,x) = Z(a,0)$ follows from Lemma 6.1. For $B(a,x)$, one takes

$$B(a,x) := \sum_{j=0}^{d_x-2} \binom{\frac{1}{2}}{j} (A(a,x)-I)^j.$$

7. Signatures

Since the partial Hessian, $(\partial^2/\partial x^2)p(a,x)[h]$, of the symmetric polynomial p is symmetric, it has a sum of difference of squares (SDS) decomposition,

$$\frac{\partial^2}{\partial x^2} p(a,x)[h] = \sum_{j=1}^{n} q_j^T(a,x)[h]q_j(a,x)[h] - \sum_{\ell=1}^{m} r_\ell^T(a,x)[h]r_\ell(a,x)[h],$$

where q_j, r_ℓ are linear in h.

The minimum number of positive (resp. negative) squares needed in such a decomposition is the *positive* (resp. *negative*) signature of $(\partial^2/\partial x^2)p(a,x)[h]$, denoted $\sigma_\pm((\partial^2/\partial x^2)p(a,x)[h])$. More generally, any symmetric polynomial $q(a,x)[h]$ which is homogeneous of degree two in h has a SDS decomposition (again with factors q_j, r_ℓ linear in h). Accordingly, we may define the signature $\sigma_\pm(q)$.

For a symmetric matrix polynomial $Z(y)$ in y and $Y \in \mathbb{S}_n(\mathbb{R}^g)$ we let $\mu_\pm(Z(Y))$ denote the number of positive or negative eigenvalues of the symmetric matrix $Z(Y)$.

The following proposition is a generalization of a result from [5].

Proposition 7.1. *Let $q(a,x)[h]$ be a symmetric polynomial in nc variables $(a,x,h) = (a_1,\dots,a_g,x_1,\dots,x_g,h_1,\dots,h_g)$ that is of degree ℓ in x and homogeneous of degree two in h with middle matrix $Z(a,x)$. Then*

$$\mu_\pm(Z(A,X)) \le n\sigma_\pm(q)$$

for each n and (A,X) in $\mathbb{S}_n(\mathbb{R}^g \times \mathbb{R}^g)$.

Proof. Suppose

$$q(a,x)[h] = \sum_{j=1}^{\sigma_+} f_j^T f_j - \sum_{j=1}^{\sigma_-} g_j^T g_j,$$

where the f_j and g_ℓ are linear in h, is a SDS decomposition with the minimum number of positive squares. Then

$$f_j^T(a,x)[h] f_j(a,x)[h] = V(a,x)[h]^T F_j(a,x) V(a,x)[h]$$

and

$$g_j^T(a,x)[h] g_j(a,x)[h] = V(a,x)[h]^T G_j(a,x) V(a,x)[h],$$

where each of the matrix polynomials $F_j(a,x)$ and $G_j(a,x)$ on the right is of the form

$$F_j(a,x) = \Phi_j(a,x)\Phi_j(a,x)^T \quad \text{and} \quad G_j(a,x) = \Psi_j(a,x)\Psi(a,x)^T,$$

for vectors Φ_j and Ψ_j whose entries are polynomials of degree less than or equal to ℓ and the border vector $V(a,x)[h]$ is linear in h.

Since

$$q(a,x)[h] = V(a,x)[h]^T \Big(\sum_{j=1}^{\sigma_+} F_j(a,x) - \sum_{j=1}^{\sigma_-} G_j(a,x) \Big) V(a,x)[h],$$

and for a given q the middle matrix is unique, once the border vector is fixed, it follows that

$$Z(a,x) = \sum_{j=1}^{\sigma_+} F_j(a,x) - \sum_{j=1}^{\sigma_-} G_j(a,x).$$

Since the rank of $F_j(A,X)$ is at most n, it follows that $Z(A,X)$ has at most $n\sigma_+^{min}$ positive eigenvalues. □

Corollary 7.2. *Suppose q is the partial Hessian of a symmetric polynomial of degree $d_x \geq 3$ in x. If $\sigma_\pm(q) \leq 1$, then*

(i) $$\sigma_\pm = 1;$$

(ii) $$\sup_n \left\{ \frac{\mu_\pm(Z(A,0))}{n} : A \in \mathbb{S}_n(\mathbb{R}^g) \right\} = \sigma_\pm(q).$$

Proof. If $\sigma_+(q) = 0$, then q is the negative of a sum of squares, so that earlier results then imply that the degree of p is two. Thus, $\sigma_+(q) = 1$. In this case, in view of the previous proposition, it suffices to prove that there is an n and an $A \in \mathbb{S}_n(\mathbb{R}^g)$ such that $\mu_+(Z(A,0)) \geq n$. The $Z_{0,d_x-2}(a,x) = Z_{0,d_x-2}(a,0)$ block of the middle matrix of the Hessian of p is not zero. It therefore has an entry (NC polynomial) which is not zero and therefore, by the Guralnick-Small Lemma (Lemma 3.3), there is an n and an $A \in \mathbb{S}_n(\mathbb{R}^g)$ so that $Z_{0,d-2}(A,0)$ has rank at least n. Since the middle matrix is symmetric and *upper anti-diagonal* and the upper right has rank n, it has at least n positive eigenvalues (Corollary 5.4 in [5]). This gives item (ii). ◻

Conjecture 7.3. *Suppose q is the partial Hessian of a symmetric polynomial. Then*

$$\sup_n \left\{ \frac{\mu_\pm(Z(A,0))}{n} : A \in \mathbb{S}_n(\mathbb{R}^g) \right\} = \sigma_\pm(q).$$

8. Appendix A. The [CHSY] Lemma

At the root of the [3] Lemma is the following

Lemma 8.1. *Fix $n > d$. If $\{z_1, \dots, z_d\}$ is a linearly independent set in $\mathbb{R}^n$, then the codimension of*

$$\left\{ \begin{pmatrix} Hz_1 \\ Hz_2 \\ \vdots \\ Hz_d \end{pmatrix} : H \in S_n(\mathbb{R}) \right\} \subset \mathbb{R}^{nd}$$

is $(d(d-1))/2$. In particular, this codimension is independent of n.

Proof. Consider the mapping Φ given by

$$S_n(\mathbb{R}) \ni H \mapsto \begin{pmatrix} Hz_1 \\ Hz_2 \\ \vdots \\ Hz_d \end{pmatrix}.$$

Since the span of $\{z_1, \dots, z_d\}$ has dimension d, it follows that the kernel of Φ has dimension $\kappa = ((n-d)(n-d+1))/2$ and hence the range has dimension

$(n(n+1))/2-\kappa$. To see this assertion, it suffices to assume that the span of $\{z_1,\dots,z_d\}$ is the span of $\{e_1,\dots,e_d\}\subset\mathbb{R}^n$ (the first d standard basis vectors in $\mathbb{R}^n$) in which case H is symmetric and $Hz_j=0$ for all j if and only if

$$H=\begin{pmatrix}0 & 0\\ 0 & H'\end{pmatrix},$$

where H' is symmetric and $(n-d)\times(n-d)$.

Finally, we conclude that the codimension of the range is

$$nd-\left(\frac{n(n+1)}{2}-\kappa\right)=\frac{d(d-1)}{2}.$$

$\square$

Lemma 8.2. *[3] If $n>d$ and $\{z_1,\dots,z_d\}$ is a linearly independent subset of $\mathbb{R}^n$, then the codimension of*

$$\left\{\bigoplus_{j=1}^{g}\begin{pmatrix}H_jz_1\\ H_jz_2\\ \vdots\\ H_jz_d\end{pmatrix} : H=(H_1,\dots,H_g)\in\mathbb{S}_n(\mathbb{R}^g)\right\}\subset\mathbb{R}^{gnd}$$

is $g(d(d-1))/2$ (independent of n).

Finally, the form in which we generally apply the lemma is the following.

Lemma 8.3. *Fix n, $v\in\mathbb{R}^n$, and $A\in\mathbb{S}_n(\mathbb{R}^g)$ and $X\in\mathbb{S}_n(\mathbb{R}^g)$. If the set $\{m(a,x)v : |m|\le d\}$ (bi-degree of m is at most d, meaning degree at most d_a in a and d_x in x) is linearly independent, then the codimension of*

$$\left\{V(A,X)[H]v : H\in\mathbb{S}_n(\mathbb{R}^g)\right\}$$

is at most $g(\kappa(\kappa-1))/2$, where $\kappa=\sum^{d_a+d_x} g^j$ and where V is the border vector associated to the given collection of monomials. Again, this codimension is independent of n.

Proof. Let $z_m=m(A,X)v$ for the given collection of monomials m. There are at most κ of these. Now apply the previous lemma. $\square$

9. APPENDIX: GENERIC INVERTIBILITY

In this Appendix we give the proof of Lemma 3.3 supplied to us by R. Guralnick and L. Small.

Proof. (1) By a result of Amitsur [1, Theorem 1], if q vanishes on all symmetric matrices of size m and $deg\ q = d$, then all matrices of size m satisfy the standard identity S_{2d}, but this bounds the size of the matrices. So taking n sufficiently large, p does not vanish on all $n \times n$ symmetric matrices.

(2) Now consider $n \times n$ matrices $Y_1, \ldots, Y_g$ with distinct commuting variables as entries (often called "generic matrices") e.g.., $Y_1 = (y_{ij})_{i,j=1,\cdots,n}$ and their transposes $Y_1^T, \ldots, Y_g^T$. Consider the algebra of polynomials in these matrices. If n is a power of 2, then this is a domain with a quotient division ring D contained in the $n \times n$ matrices over the field of rational functions. See for example, [2].

(3) To finish the argument: Take q and consider $q(Y_1 + Y_1^T, \ldots, Y_g + Y_g^T)$ for n a sufficiently large power of 2 (depending only the degree of q), so q is nonzero (note that since we are in characteristic not 2, the $Y_i + Y_i^T$ are generic symmetric matrices, so q not zero means that q does not vanish with those arguments) and q is an element of the ring generated by $Y_1, \ldots Y_g, Y_1^T, \ldots Y_g^T$ and this is a domain with a quotient division ring, it follows that $\det q(Y_1 + Y_1^T, \ldots, Y_g + Y_g^T)$ is a polynomial which is not identically zero. Thus it is nonzero on an open dense set, so the result follows. ◻

There are n for which the set of A (of size $n \times n$) such that $p(A)$ is invertible is empty. Suppose n is odd and let p denote the polynomial (in symmetric variables), $p(x, y) = (xy - yx)^2$. When n is odd, and X, Y are symmetric matrices, $XY - YX$ is skew self-adjoint and thus does not have full rank. Hence $p(X, Y)$ (is self-adjoint) and does not have full rank.

REFERENCES

[1] S.A. AMITSUR, *Identities in rings with involutions*, Israel J. Math **7** (1969), 99–106.

[2] ALLAN BERELE and DAVID SALTMAN, *The centers of generic division algebras with involution*, Israel J. Math. **63** (1988), 98–118.

[3] J.F. CAMINO, J.W. HELTON, R.E. SKELTON, and J. YE, *Matrix inequalities: A symbolic procedure to determine convexity automatically*, Integral Equations and Operator Theory **46** (2003), 399–454.

[4] H. DYM, J.W. HELTON, and S. MCCULLOUGH, *Irreducible non-commutative defining polynomials for convex sets have degree four or less*, Indiana Univ. Math. J. **56** (2007), 1189–1232.

[5] ______, *The Hessian of a non-commutative polynomial has numerous negative eigenvalues*, Journal D'Analyse Mathematique **102** (2007), 29–76.

[6] J.W. HELTON and A. LIM, *Non-commutative polynomials which are matrix convex in some of its variables have degree two or less* (2003) (Preprint).

[7] J.W. HELTON and S. MCCULLOUGH, *Convex non-commutative polynomials have degree two or less*, SIAM J. Matrix Anal. Appl. **25** (2004), 1124–1139.

[8] J.W. HELTON, S. MCCULLOUGH, and V. VINNIKOV, *Noncommutative convexity arises from linear matrix inequalities*, J. Funct. Anal. **240** (2006), 105–191.

[9] S. MCCULLOUGH, *Factorization of operator-valued polynomials in several non-commuting variables*, Linear Algebra and its Applications **326** (2001), 193–203.

DAMON M. HAY:
Department of Mathematics and Statistics
University of North Florida
1 UNF Drive
Jacksonville, FL 32224, U.S.A.
E-MAIL: damon.hay@unf.edu

J. WILLIAM HELTON:
Department of Mathematics
University of California
La Jolla, CA 92093-0112, U.S.A.
E-MAIL: helton@math.ucsd.edu

ADRIAN LIM:
Department of Mathematics
Vanderbilt University
Nashville, TN 37240, U.S.A.
E-MAIL: adrian.lim@vanderbilt.edu

SCOTT MCCULLOUGH:
Department of Mathematics
University of Florida
Box 118105
Gainesville, FL 32611-8105, U.S.A.
E-MAIL: sam@math.ufl.edu

KEY WORDS AND PHRASES: Non-commutative convex polynomials; linear matrix inequalities.

2000 MATHEMATICS SUBJECT CLASSIFICATION: 47Axx (47A63, 14P10).

Received: *April 13th, 2008; revised: November 1st, 2008.*

Article electronically published on November 14th, 2008.