CLASSIFICATION OF ALL NONCOMMUTATIVE POLYNOMIALS WHOSE HESSIAN HAS NEGATIVE SIGNATURE ONE AND A NONCOMMUTATIVE SECOND FUNDAMENTAL FORM

HARRY DYM, JEREMY M. GREENE 1, J. WILLIAM HELTON1, AND SCOTT A. MCCULLOUGH2

1 Partly supported by the NSF DMS-0700758 and the Ford Motor Co.
2 Partly supported by the NSF grants DMS-0140112 and DMS-0457504.
Abstract. Every symmetric polynomial $p = p(x) = p(x_1, \ldots, x_g)$ (with real coefficients) in $g$ noncommuting variables $x_1, \ldots, x_g$ can be written as a sum and difference of squares of noncommutative polynomials:

\[
(SDS) \quad p(x) = \sum_{j=1}^{\sigma_+} f_j^+(x)^T f_j^-(x) - \sum_{\ell=1}^{\sigma_-} f_\ell^+(x)^T f_\ell^-(x),
\]

where $f_j^+, f_\ell^-$ are noncommutative polynomials. Let $\sigma_{\text{min}}(p)$, the negative signature of $p$, denote the minimum number of negative squares used in this representation, and let the Hessian of $p$ be defined by the formula

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

In this paper we classify all symmetric noncommutative polynomials $p(x)$ such that $\sigma_{\text{min}}(p'') \leq 1$.

We also introduce the relaxed Hessian of a symmetric polynomial $p$ of degree $d$ via the formula

\[
p''_{\lambda, \delta}(x)[h] := p''(x)[h] + \delta \sum m(x)^T h^2 m(x) + \lambda p'(x)[h]^T p'(x)[h]
\]

for $\lambda, \delta \in \mathbb{R}$ and show that if this relaxed Hessian is positive semi-definite in a suitable and relatively innocuous way, then $p$ has degree at most 2. Here the sum is over monomials $m(x)$ in $x$ of degree at most $d - 1$ and $1 \leq j \leq g$.

This analysis is motivated by an attempt to develop properties of noncommutative real algebraic varieties pertaining to curvature, since, as will be shown elsewhere,

\[
-\langle p''_{\lambda, \delta}(x)[h]v, v \rangle \quad \text{(appropriately restricted)}
\]

plays the role of of a noncommutative second fundamental form.

Key words: polynomials of real symmetric matrices, noncommutative real algebraic geometry, noncommutative curvature, noncommutative inertia, noncommutative convexity

MSC 2000: 14A22, 14P10, 47A13, 46L07

1. Introduction and Main Results

This paper deals with polynomials $p(x) = p(x_1, \ldots, x_g)$ in noncommuting variables $x_1, \ldots, x_g$. We shall refer to such polynomials as nc polynomials. The first two subsections of this introduction explain the setting and recall some earlier results that will be needed to help justify the classification referred to in the abstract. The main results on nc polynomials are stated in subsection 1.3. In the noncommutative setting, there is a natural and rigid way of representing such polynomials in terms of a matrix called the middle matrix. The results of
subsection 1.3 on nc polynomials are reformulated in subsection 1.4 in terms of the middle matrix. In fact, the strategy implemented in the body of the paper is to establish facts about the middle matrix and then deduce corresponding results about polynomials. Subsection 1.5 concludes the introduction with a discussion of motivation from both geometry and engineering, cf. [HP07].

1.1. **The setting.** The setting of this paper coincides with that of [DHM07a]. The principal definitions are reviewed briefly for the convenience of the reader.

1.1.1. **Polynomials.** Let \( x = \{x_1, \ldots, x_g\} \) denote noncommuting indeterminates and let \( \mathbb{R}\langle x \rangle \) denote the set of polynomials

\[
p(x) = p(x_1, \ldots, x_g)
\]

in the indeterminates \( x \); i.e., the set of finite linear combinations

\[
p = \sum_{|m| \leq d} c_m m \quad \text{with} \quad c_m \in \mathbb{R}
\]

of monomials (words) \( m \) in \( x \). The degree of such a polynomial \( p \) is defined as the the maximum of the lengths \( |m| \) of the monomials \( m \) appearing (non-trivially) in the linear combination (1.1). Thus, for example, if \( g = 3 \), then

\[
p_1 = x_1 x_2^3 + x_2 + x_3 x_1 x_2 \quad \text{and} \quad p_2 = x_1 x_2^3 + x_2^3 x_1 + x_3 x_1 x_2 + x_2 x_1 x_3
\]

are polynomials of degree four in \( \mathbb{R}\langle x \rangle \).

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

\[
x_j^T = x_j \quad \text{and if} \quad m = x_{i_1} x_{i_2} \cdots x_{i_k} \quad \text{then} \quad m^T = x_{i_k} \cdots x_{i_2} x_{i_1},
\]

which of course extends to polynomials \( p = \sum c_m m \) by linearity:

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

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. Because of the assumption \( x_j^T = x_j \) (which will be in force throughout this paper) the variables are said to be symmetric too.
1.1.2. **Substituting Matrices for Indeterminates.** Let \((\mathbb{R}_{sym}^{n \times n})^g\) denote the set of \(g\)-tuples \((X_1, \ldots, X_g)\) of real symmetric \(n \times n\) matrices. We shall be interested in evaluating a polynomial \(p(x) = p(x_1, \ldots, x_g)\) that belongs to \(\mathbb{R}\langle x \rangle\) at a tuple \(X = (X_1, \ldots, X_g) \in (\mathbb{R}_{sym}^{n \times n})^g\). 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 \(X \in (\mathbb{R}_{sym}^{n \times n})^g\) is substituted into \(p\) the constant term \(p(0)\) of \(p(x)\) becomes \(p(0)I_n\). For example, if \(p(x) = 3 + x^2\), then

\[
p(X) = 3I_n + X^2.
\]

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

\[
p(tX + (1-t)Y) \preceq tp(X) + (1-t)p(Y)
\]

for every pair of tuples \(X, Y \in (\mathbb{R}_{sym}^{n \times n})^g\) and \(0 \leq t \leq 1\).

1.1.3. **Derivatives.** We define the directional derivative of the monomial \(m = x_{j_1} x_{j_2} \cdots x_{j_n}\) as the linear form:

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

and extend the definition to polynomials \(p = \sum c_m m\) by linearity; i.e.,

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

Thus, \(p'(x)[h] \in \mathbb{R}\langle x, h \rangle\) is the coefficient of \(t\) in the expression \(p(x + th) - p(x)\);

it is an nc polynomial in \(2g\) (symmetric) variables \((x_1, \ldots, x_g, h_1, \ldots, h_g)\). Higher order derivatives are computed in the same way: If \(q(x)[h] = h_{j_1} x_{j_2} \cdots x_{j_n}\), then

\[
q'(x)[h] = h_{j_1} h_{j_2} x_{j_3} \cdots x_{j_n} + h_{j_1} x_{j_2} h_{j_3} \cdots x_{j_n} + \cdots + h_{j_1} x_{j_2} \cdots x_{j_{n-1}} h_{j_n},
\]

and the definition is extended to finite linear combinations of such terms by linearity. If \(p\) is symmetric, then so is \(p'\). For \(g\)-tuples of symmetric matrices of a fixed size \(X, H\), the evaluation formula

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

holds, and if \(q(t) = p(X + tH)\), then

\[
p'(X)[H] = q'(0) \quad \text{and} \quad p''(X)[H] = q''(0).
\]
The second formula in (1.3) is the evaluation of the Hessian, $p''(x)[h]$ of a polynomial $p \in \mathbb{R}\langle x \rangle$; it 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 $(x_1, \ldots, x_g)$ and is homogeneous of degree two in $h$. The same conclusion holds for the $k^{th}$ directional derivative $p^{(k)}(x)[h]$ of $p$ if $k \leq d$, the degree of $p$. The expositions in [HMV06] and in [HP07] give more detail on the derivative and Hessian of a polynomial in noncommuting variables.

**Example 1.1.** A few concrete examples are listed for practice with the definitions, if the reader is so inclined.

1. If $p(x) = x^4$, then $p'(x)[h] = hxxx + xhxx + xhx + xhx$, $p''(x)[h] = 2hhxx + 2hxhx + 2xhx + 2xxh + 2xhh$,

   $p''(3)(x)[h] = 6(hhxx + hhx + hxx + hh)$,
   $p''(4)(x)[h] = 24hhhh$ and $p''(5)(x)[h] = 0$.

2. If $p(x) = x_2x_1x_2$, then $p'(x)[h] = h_2x_1x_2 + x_2h_1x_2 + x_2x_1h_2$.

3. If $p(x) = x_2^2x_2$, then $p''(x)[h] = 2(h_2^2x_2 + h_1x_2h_2 + x_1h_1h_2)$.

1.1.4. **The Signature of a Polynomial.** Every symmetric polynomial $p(x)$ admits a representation of the form (a sum and difference of squares)

$$
(SDS) \quad p(x) = \sum_{j=1}^{\sigma_+} f^+_j(x)^T f^+_j(x) - \sum_{\ell=1}^{\sigma_-} f^-_\ell(x)^T f^-_\ell(x)
$$

where $f^+_j, f^-_\ell$ are noncommutative polynomials; see e.g., Lemmas 4.7 and 4.8 of [DHM07] for details. Such representations are highly non-unique. However, there is a unique smallest number of positive (resp. negative squares) $\sigma_{\min}^+(f)$ required in an SDS decomposition of $f$. These numbers will be identified with $\mu_+(Z)$, the number of positive (negative) eigenvalues of an appropriately chosen symmetric matrix $Z$ in subsection 1.3.

1.2. **Previous results.** The number (or rather dearth) of negative squares in an SDS decomposition of the Hessian places serious restrictions on the degree of a symmetric nc polynomial. A theorem of Helton and McCullough [HM04] states that a symmetric nc polynomial that
is matrix convex has degree at most two. Since a symmetric nc polynomial
\begin{equation}
    p \text{ is matrix convex if and only if } \sigma_{\min}^{\prime\prime}(p'') = 0,
\end{equation}
(see e.g., [HMe98]) the result in [HM04] is a special case of the following more general result in [DHM07a].

**Theorem 1.2.** If $p(x)$ is a symmetric nc polynomial of degree $d$ in noncommutative symmetric variables $x_1, \ldots, x_g$, then
\begin{equation}
    d \leq 2\sigma_{\min}^{\prime\prime}(p'') + 2.
\end{equation}

There are refinements which say that if equality holds or is near to holding in Equation (1.5), then the highest degree term, $p_d$, factors or “partially factors” in a certain way. Here, and often in what follows, we write
\[ p = \sum_{i=0}^{d} p_i, \]
where $p_i$ is either a homogeneous polynomial of degree $i$, or the zero polynomial if there are no terms of degree $i$ in $p$.

### 1.3. Main Results

In this paper we determine the structure of a symmetric nc polynomial $p$ when either
\begin{enumerate}
    \item $\sigma_{\min}^{\prime\prime}(p'') \leq 1$; or
    \item a variant of the Hessian, described below in subsubsection 1.3.2 is positive.
\end{enumerate}

The first follows the investigation in [DHM07a] and the second extends efforts in [DHM07b], [HM04] and [HMV06]. The two are linked together via questions about the convexity of sublevel sets of symmetric polynomials.

The main conclusions on the structure of $p(x)$ are summarized in Theorem 1.3 and Theorem 1.4 below.

The proofs of these theorems are based on a representation for $p''$ in terms of a matrix $Z$ and very detailed analysis of the structure of $Z$. The main structural results on $Z$ are discussed in subsection 1.4.

### 1.3.1. The case of at most one negative square

An nc polynomial $\varphi$ in $g$ variables that is homogeneous of degree one may be represented by a vector from $\mathbb{R}^g$; i.e., associate the vector $u$ with entries $u_j$ to the polynomial $\varphi = \sum_{j=1}^{g} u_j x_j$.

Similarly, an nc polynomial $q$ in $g$ variables that is homogeneous of degree two (i.e., a noncommutative quadratic form) may be conveniently represented by a $g \times g$ matrix; i.e., associate the matrix $Q(q) = [q_{ij}]$ to the polynomial $q = \sum_{i,j} q_{ij} x_i x_j$. 
Theorem 1.3. Suppose \( p(x) \) is a symmetric nc polynomial of degree \( d \) in \( g \) symmetric variables \( x_1, \ldots, x_g \). Then \( \sigma_{\text{min}}(p'') \leq 1 \) if and only if

\[
p(x) = p_0 + p_1(x) + p_2(x) + \varphi(x)q(x) + q^T(x)\varphi(x) + \varphi(x)f_0(x)\varphi(x),
\]

where

1. Each nonzero \( p_j(x) \) is a homogeneous symmetric polynomial of degree \( j \);
2. \( \varphi(x) = \sum u_j x_j \) is a homogeneous polynomial of degree one and \( u \) is a unit vector with components \( u_1, \ldots, u_g \);
3. Each of the polynomials \( p_2(x) \), \( q(x) \) and \( f_0(x) \) is either 0 or a homogeneous polynomial of degree two; and the matrix

\[
E_2 = \begin{bmatrix} P Q(p_2)' & P Q(q)' \\ Q(q)' & Q(f_0) \end{bmatrix}
\]

is positive semidefinite, where \( P = I - uu^T \) is the orthogonal projection onto the orthogonal complement of \( u \).

If in addition \( p \) has degree three, then \( f_0 = 0 \) and \( q(x) = f_1(x)\varphi(x) \) for some homogeneous polynomial \( f_1(x) \) of degree one.

Finally, \( \sigma_{\text{min}}(p'') = 0 \) if and only if \( d \leq 2 \) and \( p_2(X) \geq 0 \) for every \( X \in \mathbb{R}^{n \times n}_{\text{sym}} \).

Proof. The proof is in \[1.4\] \( \square \)

1.3.2. The Relaxed Hessian. In applications to convex sublevel sets and the curvature of noncommutative real varieties it is useful to introduce a variant of the Hessian for symmetric polynomials \( p \) of degree \( d \) in \( g \) noncommuting variables. Let \( V_k(x)[h] \) denote the vector of polynomials with entries \( h_j m(x) \), where \( m(x) \) runs through the set of \( g^k \) monomials of length \( k \), \( j = 1, \ldots, g \). The order of the entries (which is irrelevant for the moment) is fixed in Section \[1.5.2\] via formula \[2.4\]. Thus, \( V_k = V_k(x)[h] \) is a vector of height \( g^{k+1} \), and the vectors

\[
V(x)[h] = \text{col}(V_0, \ldots, V_{d-2}) \quad \text{and} \quad \tilde{V}(x)[h] = \text{col}(V_0, \ldots, V_{d-1})
\]

are vectors of height \( g\nu \) and \( g\tilde{\nu} \), respectively, where

\[
\nu = 1 + g + \cdots + g^{d-2} \quad \text{and} \quad \tilde{\nu} = 1 + g + \cdots + g^{d-1}.
\]

Note that

\[
\tilde{V}(x)[h]^T \tilde{V}(x)[h] = \sum_{j=1}^g \sum_{|m| \leq d-1} m(x)^T h_j^2 m(x),
\]

where \( |m| \) denotes the degree (length) of the monomial \( m \).
The relaxed Hessian of the symmetric nc polynomial \( p \) of degree \( d \) is the polynomial
\[
    p''_{\lambda, \delta} := p''(x)[h] + \delta \tilde{V}(x)[h]^T \tilde{V}(x)[h] + \lambda p'(x)[h]^T p'(x)[h],
\]
in which \( \lambda, \delta \in \mathbb{R} \). We say that the relaxed Hessian is positive at \((X, v) \in (\mathbb{R}^{n \times n}_{\text{sym}})^g \times \mathbb{R}^n\) if for each \( \delta > 0 \) there is a \( \lambda > 0 \) so that
\[
    0 \leq \langle p''_{\lambda, \delta}(X)[H]v, v \rangle
\]
for all \( H \in (\mathbb{R}^{n \times n}_{\text{sym}})^g \). Note that
\[
    \lambda_1 \geq \lambda \quad \text{and} \quad \delta_1 \geq \delta \implies \langle p''_{\lambda_1, \delta_1}(X)[H]v, v \rangle \geq \langle p''_{\lambda, \delta}(X)[H]v, v \rangle.
\]

The next theorem says that positivity of the relaxed Hessian plus a type of irreducibility implies \( p \) has degree at most two.

**Theorem 1.4.** Let \( p \) be a symmetric nc polynomial of degree \( d \) in \( g \) symmetric variables. Suppose \( n > \frac{1}{2} q \tilde{\nu}(\tilde{\nu} - 1) \),
\[
    X \in (\mathbb{R}^{n \times n}_{\text{sym}})^g, \ v \in \mathbb{R}^n, \ \text{and there is no nonzero polynomial} \ q \ (\text{not necessarily symmetric}) \ \text{of degree less than} \ d \ \text{such that} \ q(X)v = 0. \ \text{If the relaxed Hessian is positive at} \ (X, v), \ \text{then} \ p \ \text{is convex and has degree at most two.}
\]

Theorem 1.4 in turn follows from the following theorem together with an application of the CHSY-Lemma (see Lemma 6.1).

**Theorem 1.5.** Let \( p \) be a symmetric nc polynomial of degree \( d \) in \( g \) symmetric variables. Suppose \( X \in (\mathbb{R}^{n \times n}_{\text{sym}})^g \) and \( v \in \mathbb{R}^n \). If the set of vectors \( \{\tilde{V}(X)[h]v : H \in (\mathbb{R}^{n \times n}_{\text{sym}})^g\} \) has codimension at most \( n - 1 \) in \( \mathbb{R}^{nq \tilde{\nu}} \) and if the relaxed Hessian is positive at \((X, v)\), then \( \sigma^m_{\min}(p'') = 0 \) and hence \( p \) is convex and has degree at most two.

**Proof.** Theorems 1.4 and 1.5 are established in section 6. \( \square \)

We end this subsection with the following generalization of one of the main results of [HM04]. The noncommutative \( \epsilon \)-neighborhood of 0 is the graded set \( \mathcal{U} = \bigcup_{n \geq 1} \mathcal{U}_n \) with
\[
    \mathcal{U}_n = \{(X, v) : X \in (\mathbb{R}^{n \times n}_{\text{sym}})^g, \ |X| < \epsilon, \ v \in \mathbb{R}^n\}
\]
(where \( |\cdot| \) is any norm on \((\mathbb{R}^{n \times n}_{\text{sym}})^g\)).

Given a (graded) subset \( S = \bigcup_{n \geq 1} S_n \) of \( \bigcup_{n \geq 1} ((\mathbb{R}^{n \times n}_{\text{sym}})^g \times \mathbb{R}^n) \), we say that the relaxed Hessian is positive on \( S \) if it is positive at each \((X, v) \in S\).
Corollary 1.6. Let $p$ be a given symmetric nc polynomial of degree $d$ in $g$ symmetric variables. If there is an $\epsilon > 0$ so that the relaxed Hessian is positive on a noncommutative $\epsilon$-neighborhood of 0, then the degree of $p$ is at most two and $p$ is convex.

Proof. The corollary is proved in subsection 6.3. □

1.3.3. Partial positivity of the Hessian. In this section we consider variants and corollaries of Theorem 1.5 and its consequence Theorem 1.4 pertaining to the nonnegativity hypotheses on the Hessian of the symmetric polynomial $p$ which imply bounds on the negative signature $\sigma_{-}(p')$.

Given a subspace $H \subset (\mathbb{R}^{n \times n})^g$, we say that the Hessian of $p$ is positive relative to $H$ at $(X,v)$ if for each $H \in H$,

$$0 \leq \langle p''(X)[H]v, v \rangle.$$

Theorem 1.7. Let $p$ be a symmetric nc polynomial of degree $d$ in $g$ symmetric variables and let $k$ be a given positive integer. Suppose $X \in (\mathbb{R}^{n \times n})^g$, $v \in \mathbb{R}^n$, and $H$ is a subspace of $(\mathbb{R}^{n \times n})^g$. If the set of vectors $\{V(X)[H]v : H \in H\}$ has codimension at most $kn - 1$ in $\mathbb{R}^{ng^2}$ and if the Hessian of $p$ is positive relative to $H$ at $(X,v)$, then

$$\sigma_{\min}(p'') < k.$$

The next theorem follows from Theorem 1.7 in much the same way that Theorem 1.4 follows from Theorem 1.5.

Theorem 1.8. Let $p$ be a symmetric nc polynomial of degree $d$ in $g$ symmetric variables and fix a positive integer $k$. Suppose

$$n > g \frac{\nu(\nu - 1)}{2}$$

$X \in (\mathbb{R}_{sym}^{n \times n})^g$, $v \in \mathbb{R}^n$, and there does not exist a nonzero polynomial $q$ (not necessarily symmetric) of degree less than or equal to $d - 2$ such that $q(X)v = 0$. If $H$ is a subspace of $(\mathbb{R}_{sym}^{n \times n})^g$ of codimension at most $kn$ and if the Hessian of $p$ is positive relative to $H$ at $(X,v)$, then

$$\sigma_{\min}(p'') < k + 1.$$

Proof. The last two theorems are proved in subsection 6.4. □

Remark 1.9. There are versions of Theorem 1.7 and Theorem 1.8 with the relaxed Hessian in place of the ordinary Hessian.

1.3.4. The modified Hessian. The modified Hessian of a symmetric nc polynomial $p$ is the expression,

$$p''_\lambda(x)[h] = p''(x)[h] + \lambda p'(x)[h]^tp'(x)[h].$$
The **modified Hessian is positive at** \((X,v)\) if there is a \(\lambda > 0\) so that

\[
0 \leq \langle p''(X)[H]v,v \rangle \quad \text{for every } H \in (\mathbb{R}_{sym}^{n \times n})^g.
\]

Note that if the modified Hessian of \(p\) is positive at \((X,v)\), then so is the relaxed Hessian. Indeed, in that case a single \(\lambda\) suffices for all \(\delta\).

The modified Hessian will play an important intermediate role between the analysis of the Hessian and relaxed Hessian.

### 1.4. The middle matrix

The proofs and refinements of Theorems [1.3] and [1.4] rest on a careful analysis of a certain representation of \(p''\) which we now introduce.

In [DHM07a] it was shown that if \(p(x)\) is a symmetric nc polynomial of degree \(d \geq 2\) in \(g\) noncommuting variables, then the Hessian \(p''(x)[h]\) admits a representation of the form

\[
p''(x)[h] = V(x)[h]^T Z(x) V(x)[h]
\]

\[
= [V_0^T, V_1^T, \ldots, V_\ell^T]
\]

\[
\begin{bmatrix}
Z_{00} & Z_{01} & \cdots & Z_{0,\ell-1} & Z_{0\ell} \\
Z_{10} & Z_{11} & \cdots & Z_{1,\ell-1} & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
Z_{\ell 0} & 0 & \cdots & 0 & 0
\end{bmatrix}
\begin{bmatrix}
V_0 \\
V_1 \\
\vdots \\
V_\ell
\end{bmatrix},
\]

in which \(\ell = d - 2\), \(V(x)[h]\) is the border vector with vector components \(V_j(x)[h]\) of height \(g^{j+1}\), and \(Z(x) = [Z_{ij}(x)], i,j = 0, \ldots, d - 2\), the **middle matrix**, is a symmetric matrix polynomial with matrix polynomial entries \(Z_{ij}(x)\) of size \(g^{i+1} \times g^{j+1}\) and degree no more than \((d-2)-(i+j)\) for \(i+j \leq d-2\) with \(Z_{ij}(x) = 0\) for \(i+j > d-2\). Since \(p\) is symmetric \(Z_{ij} = Z_{ji}\) and, since \(p\) has degree \(d\), \(Z_{ij}\) is constant when \(i+j = d-2\).

Let \(Z = Z(0)\) for \(0 \in \mathbb{R}^g\). We call \(Z\) the **scalar middle matrix**.

The main conclusions from [DHM07a] that are relevant to this paper are:

1. \(Z(x)\) is polynomially congruent to the scalar middle matrix \(Z = Z(0)\), i.e., there exists a matrix polynomial \(B(x)\) with an inverse \(B(x)^{-1}\) that is again a matrix polynomial such that

\[
Z(x) = B(x)^T Z(0) B(x) = B(x)^T B(x).
\]

2. \(\mu_\pm(Z) = \sigma^g_\pm(p''(x)[h])\).

3. If \(X \in (\mathbb{R}_{sym}^{n \times n})^g\), then

\[
\mu_\pm(Z(X)) = n \mu_\pm(Z).
\]
(4) The degree $d$ of $p(x)$ is subject to the bound
\[ d \leq 2\mu_\pm(Z) + 2. \]

(5) If $p(x) = p_0(x) + \cdots + p_d(x)$ is expressed as a sum of homogeneous polynomials $p_j(x)$ of degree $j$ for $j = 0, \ldots, d$, then the homogeneous components $p_k$ of $p$ with $k \geq 2$ can be recovered from $Z$ by the formula
\[ p_{i+j+2}(x) = \frac{1}{2}[x_1, \ldots, x_g]_{i+1}Z_{ij}([x_1, \ldots, x_g]_{j+1})^T \]
for $i + j \leq d - 2$,

where $[x_1, \ldots, x_g]_j$ denotes the $j$-fold Kronecker product that is defined in (2.2); it is a row vector with $g^j$ entries consisting of all the monomials (in $x$) of degree $j$.

If $\mu_-(Z) \leq 1$, then the bound (1.15) implies that $d \leq 4$ and consequently that $Z$ will be a matrix of the form
\[ Z = \begin{bmatrix} Z_{00} & Z_{01} & Z_{02} \\ Z_{10} & Z_{11} & 0 \\ Z_{20} & 0 & 0 \end{bmatrix}. \]

The next theorem, which describes the structure of $Z$ when $\mu_-(Z) \leq 1$, requires some additional notation. Viewing a $g \times g^2$ matrix such as $Z_{01}$, as a $g \times g$ block matrix $C = [c_{ij}]$ with entries $c_{ij} \in \mathbb{R}^{1 \times g}$, the structured transpose $C^{sT}$ of $C$ is the $g \times g$ block matrix with $ij$ entry equal to $c_{ji}$. Thus, for example, if $g = 2$ and
\[ C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}, \]
then $C^{sT} = \begin{bmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{bmatrix}$ and $C^T = \begin{bmatrix} c_{11}^T & c_{21}^T \\ c_{12}^T & c_{22}^T \end{bmatrix}$.

A number of identities involving this definition are presented in Lemmas 3.3–3.5.

Given a $g \times n$ matrix $A$, expressed in terms of its columns as
\[ A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \]
let $\text{vec}(A) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$,

and, conversely, given a vector $w \in \mathbb{R}^g$ with entries $a_1, \ldots, a_n \in \mathbb{R}^g$,
\[ w = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \]
let $\text{mat}_g(w) = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$. 
Thus, if $A$ and $w$ are as above, then
\[
\text{mat}_g \text{vec } A = A \quad \text{and} \quad \text{vec } \text{mat}_g w = w.
\]

**Theorem 1.10.** Let $Z$ denote the scalar middle matrix of the Hessian of a symmetric nc polynomial $p(x)$.

(I) If $\text{degree } p \geq 3$, then $\mu_-(Z) = 1$ if and only if
\[
Z = \begin{bmatrix}
Z_{00} & Z_{01} & Z_{02} \\
Z_{10} & Z_{11} & 0 \\
Z_{20} & 0 & 0
\end{bmatrix}
\]
where
(1) $Z_{01} = u(u^T \otimes y^T) + uv^T + (uv^T)^sT = Z_{01}^{sT}$ and $Z_{10} = Z_{10}^{T}$;
(2) $Z_{11} = (u \otimes I_g)A(u \otimes I_g)^T$; and
(3) $Z_{02} = u(u^T \otimes w_A^T)$; and $Z_{20} = Z_{20}^{T}$,
in which
(i) $u, y \in \mathbb{R}^g$ and $u$ is a unit vector;
(ii) $v \in \mathbb{R}^{g^2}$ and $(\text{mat}_g v)u = 0$.
(iii) $A = A^T \in \mathbb{R}^{g \times g}$ and $w_A = \text{vec}(A) \in \mathbb{R}^{g^2}$.
(iv) The matrix
\[
E_1 = \begin{bmatrix}
PZ_{00}P & (\text{mat}_g v)^T \\
\text{mat}_g v
\end{bmatrix}
\]
(with $P = I - uu^T$ and $A = U^T Z_{11} U$)
is positive semidefinite.

Moreover, if $\text{degree } p = 3$, then in addition to the conditions stated above, $A = 0$, $w_A = 0$, and $v = 0$.

(II) If $\text{degree } p = 2$, then $\mu_-(Z) = 1$ if and only if $Z_{ij} = 0$ for $i + j \geq 1$ and there exists a vector $u \in \mathbb{R}^g$ with $u^T u = 1$ such that $Z_{00}u = \lambda u$ with $\lambda < 0$ and $PZ_{00}P \succeq 0$ for $P = I_g - uu^T$.

(III) $\mu_-(Z) = 0$ if and only if $Z_{ij} = 0$ for $i + j \geq 1$ and $Z_{00} \succeq 0$.

**Proof.** See subsection 4.3. □

Theorem 1.3 will follow from Theorem 1.10 and the results in section 4.

The next theorem gives the explicit correspondence between Theorems 1.3 and 1.10.

**Theorem 1.11.** The entries of $Z$ in Theorem 1.10 correspond to terms of $p$ in Theorem 1.3 as follows

1. $\varphi(x) = \lfloor x_1 \cdots x_g \rfloor u$
2. $q(x) = \frac{1}{4}(u \otimes y + 2v)^T([x_1 \cdots x_g]_2)^T$
3. $f_0(x) = \frac{1}{2}x^TAx$
Proof. See subsection 4.5. □

Remark 1.12. If \( p_4(x) = \frac{1}{2} \varphi(x)[x_1 \cdots x_g]A[x_1 \cdots x_g]^T \varphi(x) \), then, since
\[
p_4(x) = \frac{1}{2}[x_1 \cdots x_g]^2 \mathcal{Z}_{11}([x_1 \cdots x_g]^2) = \frac{1}{2}[x_1 \cdots x_g] \mathcal{Z}_{02}([x_1 \cdots x_g]^3)^T,
\]
the formulas in Theorem 1.10 imply that
\[
\mathcal{Z}_{11} = UAU^T \quad \text{and} \quad \mathcal{Z}_{02} = u(u^T \otimes w^T_A) \quad \text{with} \quad w_a = \vec{A}.
\]
Thus, rank \( \mathcal{Z}_{11} \) can vary between 1 and \( g \), when rank \( \mathcal{Z}_{02} = 1 \).

It is not difficult to see that if \( q(x)[h] \) is a symmetric nc polynomial in the variables \( x \) and \( h \) and if \( q \) is homogeneous of degree two in \( h \), then it has a representation like that in equation (1.12). In particular, the modified Hessian can be expressed as
\[
(1.18) \quad p''_{\lambda}(x)[h] = \tilde{V}(x)[h]^T Z_{\lambda}(x) \tilde{V}(x)[h].
\]
Note here that the vector \( \tilde{V}(x)[h] \) is now of degree \( d - 1 \) in \( x \) and homogeneous of degree one in \( h \).

The proof of Theorem 1.13 relies on the following structure theorem for \( Z_{\lambda} \), the scalar middle matrix of the modified Hessian of \( p \).

Theorem 1.13. If \( p(x) \) is a symmetric polynomial of degree \( d \) in non-commutative symmetric variables \( x_1, \ldots, x_g \), then the middle matrix \( Z_{\lambda}(x) \) for the modified Hessian for \( p(x) \) is polynomially congruent to the constant matrix
\[
Z_{\lambda} = \begin{bmatrix} Z & 0 \\ 0 & \lambda \psi_{d-1}(0) \psi_{d-1}(0)^T \end{bmatrix}
\]
in which \( \psi_{d-1}(0) \) is a nonzero vector of size \( g^d \). Moreover, the congruence is independent of \( \lambda \) and consequently, if \( \lambda > 0 \), then
\begin{enumerate}
\item \( \mu_+(Z_{\lambda}) = \mu_+(Z) + 1 \) and \( \mu_-(Z_{\lambda}) = \mu_-(Z) \);
\item \( Z_{\lambda} \succeq 0 \iff Z \succeq 0 \);
\item \( Z_{\lambda} \succeq 0 \iff d \leq 2 \); and
\item \( d \leq 2 \sigma_{\min}^{-1}(p''_{\lambda}(x)[h]) + 2 \).
\end{enumerate}

Proof. This theorem is established in Section 5. □

1.5. Motivation. The theorems in this paper on the relaxed Hessian bear on a conjecture concerning the convexity of the positivity set
\[
\mathcal{D}_p := \bigcup_{n \geq 0} \mathcal{D}_p^n,
\]
of a symmetric nc polynomial \( p \), where \( \mathcal{D}_p^n \) is the connected component of
$\mathcal{P}_p^n := \{X = \{X_1, \ldots, X_g\} : X_j \in \mathbb{R}^{n \times n}_{\text{sym}} \text{ and } p(X) > 0\}$ containing 0. An example of a “convex” positivity domain $\mathcal{D}_p$, is a ball, namely a set of the form $\mathcal{D}_q$ where $q$ has the form

$$q(x) = c - \sum_j^w (a_j + L_j(x))^T(a_j + L_j(x))$$

where $a_j, c$ are real numbers and $L_j(x)$ are linear in $x_1, \ldots, x_g$, that is, $L(x) = b_1 x_1 + \cdots + b_g x_g$.

A (loosely stated) conjecture of Helton and McCullough is

**Conjecture 1.14.** (Helton and McCullough) If $p$ is “irreducible” and $\mathcal{D}_p$ is convex, then $\mathcal{D}_p$ is a ball.

1.5.1. *Noncommutative real algebraic geometry.* One might think of Conjecture [1.14] and its supporting results as the beginnings of a noncommutative real algebraic geometry in that this is like supposing “if the boundary $\partial \mathcal{D}_p$ of $\mathcal{D}_p$ has positive curvature” and concluding that “it has constant curvature.”

Since $(\mathbb{R}^{n \times n}_{\text{sym}})^g = \mathbb{R}^g$ when $n = 1$, it is reasonable to turn to functions $f$ on $\mathbb{R}^g$ (now considered with commuting variables) for guidance. In particular smooth quasiconvex functions $f$ on $\mathbb{R}^g$ are functions 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 $\partial \mathcal{C}_c$ and 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 same as the second fundamental form of differential geometry being nonnegative. In the terminology of this paper for $n = 1$ this condition for $X \in \mathbb{R}^g$ is that $p''(X)[H] \succeq 0$ for all $H \in \mathbb{R}^g$ satisfying $p'(X)[H] = 0$. In the scalar case ($n = 1$) $v$ is effectively 1, so we do not write it. An easy Lagrange multiplier argument shows that this is equivalent to there is a $\lambda > 0$ making the modified Hessian, $p''_\lambda(X)[H]$, nonnegative for all $H$. Thus a smooth function $f$ with nowhere vanishing gradient (except at its global minimum) is quasiconvex on a bounded domain $B \subset \mathbb{R}^g$ if and only if for $\lambda$ large enough $p''_\lambda(X)(X)[H]$ is nonnegative for all $X \in B \subset \mathbb{R}^g$ and all $H \in \mathbb{R}^g$.

The papers [DHM07b] and subsequent work represent progress on Conjecture [1.14] for the noncommutative case. We find that an appropriately restricted quadratic form based on the relaxed Hessian
plays the role of the second fundamental form very effectively and seems to be the key to understanding “noncommutative curvature.” Its power emanates to a large extent from the classification theorems established in this paper.

In a different direction we mention that many engineering problems connected with linear systems lead to constraints in which a polynomial or a matrix of polynomials with matrix variables is required to take a positive semidefinite value. Many of these problems are also “dimension free” in the sense that the form of the polynomial remains the same for matrices of all sizes. This is one very good reason to study noncommutative polynomials. If these polynomials are “convex” or “quasiconvex” in the unknown matrix variables, then numerical calculations are much more reliable. This motivated Conjecture 1.14 and our results suggest that matrix convexity in conjunction with a type of “system scalability” produces incredibly heavy constraints.

1.5.2. Noncommutative analysis. This article could also come under the general heading of “free analysis”, since the setting is a noncommutative algebra whose generators are “free” of relations. This is a bur- dgeoning area, of which free probability, invented by Voiculescu [Vo05] and [Vo06] is currently the largest component. The interested reader is referred to the web site [SV06] of the American Institute of Mathematics, in particular it gives the findings of the AIM workshop in 2006 on free analysis. Excellent work in other directions with free algebras has been carried out by Ball et al. [BGM06], Kaluzny-Verbotzky-Vinnikov, e.g. [KVV06], Popescu, e.g. [Po06], and by Voiculescu [Vo05] and [Vo06]. A fairly expository article describing noncommutative convexity, noncommutative semialgebraic geometry and relations to engineering is [HP07].

Acknowledgement The authors extend their thanks to the referee for his/her careful reading of the paper, and for catching a number of misprints and making useful suggestions for improving the clarity of the presentation.

2. Kronecker Product Notation

Recall the Kronecker product

\[
A \otimes B = \begin{bmatrix}
  a_{11}B & \cdots & a_{1q}B \\
  \vdots & \ddots & \vdots \\
  a_{p1}B & \cdots & a_{pq}B
\end{bmatrix}
\]

\[\text{\footnote{This section is taken from [DHM07a] for the convenience of the reader.}}\]
of a pair of matrices $A \in \mathbb{R}^{p \times q}$ and $B \in \mathbb{R}^{s \times t}$.

A number of formulas that occur in this paper are conveniently expressed in terms of $j$-fold iterated Kronecker products of row or column vectors of noncommuting variables. Accordingly we introduce the notations,

$$[x_1 \cdots x_g]_j = [x_1 \cdots x_g] \otimes \cdots \otimes [x_1 \cdots x_g] \quad (j \text{ times})$$

and

$$\begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}_j = \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \quad (j \text{ times}).$$

2.1. **Border vectors.** In this section we shall expand upon the representation of the Hessian $p''(x)[h]$ in terms of a middle matrix $Z(x)$ and the components $V_j = V_j(x)[h]$ of the border vector $V(x)[h]$ defined by the formulas

$$V_j^T(x)[h] = \begin{cases} [h_1 \cdots h_g] & \text{if } j = 0, \\
[x_1 \cdots x_g]_j \otimes [h_1 \cdots h_g] & \text{if } j = 1, 2, \ldots,
\end{cases}$$

Thus, for example, the components $[V_0^T V_1^T \cdots V_{d-2}^T]$ of the transpose $V^T$ of the border vector $V$ for polynomials of two noncommuting variables may be written

$$V_0^T(x)[h] = [h_1 \ h_2],$$

$$V_1^T(x)[h] = [x_1 \ x_2] \otimes [h_1 \ h_2],$$

$$V_2^T(x)[h] = [x_1 \ x_2] \otimes [x_1 \ x_2] \otimes [h_1 \ h_2],$$

$$\vdots$$

In this ordering,

$$V_1^T(x)[h] = [x_1 h_1 \ x_1 h_2 \ x_2 h_1 \ x_2 h_2]$$

and

$$V_2^T(x)[h] = [x_1 x_1 h_1 \ x_1 x_1 h_2 \ x_1 x_2 h_1 \ x_1 x_2 h_2 \ x_2 x_1 h_1 \ x_2 x_1 h_2 \ x_2 x_2 h_1 \ x_2 x_2 h_2].$$

However, since we are dealing with noncommuting variables, the rules for extracting $V_j$ from $V_j^T$ are a little more complicated than might be expected. Thus, for example, the given formula for $V_1^T$ implies that

$$V_1 = \begin{bmatrix} h_1 x_1 \\
 h_2 x_1 \\
 h_1 x_2 \\
 h_2 x_2 
\end{bmatrix} \neq \begin{bmatrix} h_1 \\
 h_2 
\end{bmatrix} \otimes \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$
as one might like. Nevertheless, the situation is not so bad, because the left hand side is just a permutation of the right hand side. This propagates:

\[(2.5)\]

\[
\begin{bmatrix}
h_1 \\
\vdots \\
h_g
\end{bmatrix} = V_0 \quad \text{and} \quad \begin{bmatrix}
h_1 \\
\vdots \\
h_g
\end{bmatrix} \otimes \begin{bmatrix}
x_1 \\
\vdots \\
x_g\end{bmatrix}_j = \Pi_j V_j \quad \text{for} \quad j = 1, 2, \ldots
\]

for suitably defined permutation matrices \( \Pi_j, \ j = 1, 2, \ldots \). The permutation matrix \( \Pi_j \) in formula (2.5) can also be characterized by the condition

\[(2.6)\]

\[
\begin{bmatrix}
x_1 \\
\vdots \\
x_g
\end{bmatrix}_{j+1} = \Pi_j \left( \begin{bmatrix}
x_1 & \cdots & x_g\end{bmatrix}_j \right)^T \quad \text{for} \quad j = 1, 2, \ldots
\]

The second set of formulas in (2.5) can also be written in the form

\[(2.7)\]

\[
\Pi_j V_j(x)[h] = \left( \begin{bmatrix}
h_1 \\
\vdots \\
h_g
\end{bmatrix} \otimes I_g \right) \left( \begin{bmatrix}
x_1 \\
\vdots \\
x_g\end{bmatrix}_j \right), \quad j = 1, 2, \ldots
\]

The general formula

\[(2.8)\]

\[
([u_1, \ldots, u_k] \otimes [v_1, \ldots, v_\ell]) a = ([u_1, \ldots, u_k]) A ([v_1, \ldots, v_\ell]^T)
\]

with

\[(2.9)\]

\[
a^T = [a_1, a_2, \ldots, a_{k\ell}] \quad \text{and} \quad A = \begin{bmatrix}
a_1 & \cdots & a_\ell \\
a_{\ell+1} & \cdots & a_{2\ell} \\
\vdots & \ddots & \vdots \\
a_{(k-1)\ell+1} & \cdots & a_{k\ell}
\end{bmatrix}
\]

is useful, as is the identity

\[(2.10)\]

\[
\begin{bmatrix}
u_1 \\
\vdots \\
u_k
\end{bmatrix} \otimes \begin{bmatrix}
v_1 \\
\vdots \\
v_\ell
\end{bmatrix} = \left( \begin{bmatrix}
u_1 \\
\vdots \\
u_k
\end{bmatrix} \otimes I_\ell \right) \begin{bmatrix}
v_1 \\
\vdots \\
v_\ell
\end{bmatrix}.
\]

2.2. Tracking derivatives. The Kronecker notation is useful for keeping track of the positions of the variables \( \{h_1, \ldots, h_g\} \) when computing the derivatives \( \frac{p'(x)[h]}{h} \) and \( \frac{p''(x)[h]}{h} \) of the polynomial \( p(x) \) in the direction \( h \). This is a consequence of two facts:

1. Every polynomial is a sum of homogeneous polynomials.
(2) Every homogeneous polynomial \( p = p(x) = p(x_1, \ldots, x_g) \) of degree \( d \) in \( g \) noncommuting variables can be expressed as a \( d \)-fold Kronecker product times a vector \( a \) with \( g^d \) real entries:

\[ p \text{ homogeneous of degree } d \implies p(x) = ([x_1 \cdots x_g]^d)a. \]

Thus, for example, if \( p(x) = ([x_1, \ldots, x_g]^3)a \) for some vector \( a \in \mathbb{R}^{g^3} \), then

\[ p'(x)[h] = ([h_1, \ldots, h_g] \otimes [x_1, \ldots, x_g]_2 + [x_1, \ldots, x_g] \otimes [h_1, \ldots, h_g] \otimes [x_1, \ldots, x_g] + [x_1, \ldots, x_g]_2 \otimes [h_1, \ldots, h_g])a. \]

Higher order directional derivatives can be tracked in a similar fashion.

In calculations of this type the identity

\[ ([x_1, \ldots, x_g]_k u)([y_1, \ldots, y_l]v) = ([x_1, \ldots, x_g]_k \otimes [y_1, \ldots, y_l])(u \otimes v), \]

for noncommuting (as well as commuting) variables \( x_1, \ldots, x_g, y_1, \ldots, y_l \) and vectors \( u \in \mathbb{R}^{g^k} \) and \( v \in \mathbb{R}^{\ell} \) is often useful.

3. Basic Identities

This section provides some basic identities needed for later sections. Some are easy, some take a bit of work. Most of the details are left to the reader.

3.1. Identities.

**Theorem 3.1.** Let \( u \in \mathbb{R}^g, w \in \mathbb{R}^{g^2}, A \in \mathbb{R}^{g \times g}, \varphi(x) = [x_1 \cdots x_g]u \) and let \( \Pi_j \) denote the permutations defined in (2.6). Then the following identities are in force:

1. \([x_1 \cdots x_g]_2(u \otimes A) = \varphi(x)[x_1 \cdots x_g]A.\)

2. \((u \otimes A)^T([x_1 \cdots x_g]_2)^T = A^T \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \varphi(x).\)

3. \([x_1 \cdots x_g](u \otimes A)^T = u^T \otimes ([x_1 \cdots x_g]A^T).\)

4. \([x_1 \cdots x_g]_2^T = (I_g \otimes [x_1 \cdots x_g]^T)[x_1 \cdots x_g]^T.\)

5. \(\begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_l \end{bmatrix} = (b^T \otimes a^T)([y_1 \cdots y_l] \otimes [x_1 \cdots x_g]^T).\)
Lemma 3.2. If \( \varphi^2 = (u^T \otimes u^T)([x_1 \cdots x_g]_2)^T, \varphi^3 = (u^T \otimes u^T \otimes u^T)([x_1 \cdots x_g]_3)^T, \) etc.

(7) \([x_1 \cdots x_g]uw^T([x_1 \cdots x_g]_2)^T = (u^T \otimes u^T)([x_1 \cdots x_g]_3)^T.\)

(8) \([x_1 \cdots x_g]_2wu^T = u^T \otimes ([\Pi_1^T w]^T([x_1 \cdots x_g]_2)^T).\)

(9) \([x_1 \cdots x_g]_2wu^T([x_1 \cdots x_g]_3)^T = (u^T (\Pi_1^T w)^T)([x_1 \cdots x_g]_3)^T.\)

(10) \(\begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \otimes I_g^k \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}_{k+1} = \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \) for \( k = 1, 2, \ldots.\)

(11) If \( w = \text{vec } A, \) then

\([x_1 \cdots x_g]A = w^T \left( I_g \otimes \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \right) \) and

\([x_1 \cdots x_g]_2w = [x_1 \cdots x_g]A^T([x_1 \cdots x_g])^T.\)

(12) \(A(u \otimes I_g)^T = u^T \otimes A.\)

(13) \(\left( I_{g^2} \otimes \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} \right)([x_1 \cdots x_g]_2)^T = ([x_1 \cdots x_g]_3)^T.\)

(14) If \( v \) a vector of size \( g^k \) and \( w \) a vector with \( g^l \) entries, then

\(\begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}_{k+l} = \left( v^T \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}_k \right) \left( w^T \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}_l \right).\)

(15) If \( a, c \in \mathbb{R}^k \) and \( b, d \in \mathbb{R}^l, \) then \( (a^T \otimes b^T)(c \otimes d) = (a^T c)(b^T d).\)

**Lemma 3.2.** If \( u \in R^g \) and \( w \in R^{g^2}, \) then

\([x_1 \cdots x_g] uw^T ([x_1 \cdots x_g]_2)^T + [x_1 \cdots x_g]_2wu^T([x_1 \cdots x_g])^T = (w^T \otimes u^T + u^T \otimes (\Pi_1^T w)^T)([x_1 \cdots x_g]_3)^T,\)

where \( \Pi_1 \) is the permutation defined in (2.6).

**Proof.** This is immediate from identities (7) and (9) in Theorem 3.1. \( \square \)
3.2. Structured transposes.

Lemma 3.3. If \( u \in \mathbb{R}^g \), \( U = u \otimes I_g \) and \( B \in \mathbb{R}^{g \times g^2} \), then

\[
(BUU^T)^{sT} = uu^T B^{sT} \quad \text{and} \quad (uu^T B)^{sT} = B^{sT} UU^T.
\]

Proof. Let \( B = [b_{ij}] \), with \( b_{ij} \in \mathbb{R}^{1 \times g} \) for \( i, j = 1, \ldots, g \). Then the entries \( c_{ij} \) of the matrix \( C = BUU^T = B(uu^T \otimes I_g) \) are

\[
c_{ij} = \sum_{s=1}^{g} b_{is}(u_s u_j I_g) = \sum_{s=1}^{g} u_s u_j b_{is} I_g
\]

\[
= \sum_{s=1}^{g} u_j u_s b_{is} = \sum_{s=1}^{g} (uu^T)_{js} b_{is},
\]

which coincides with the \( ji \) entry of \( uu^T B^{sT} \). Therefore, \( C^{sT} = uu^T B^{sT} \).

This justifies the first formula in (3.1). The second follows from the first upon replacing \( B \) by \( B^{sT} \), since \( (K^{sT})^{sT} = K \) for every \( K \in \mathbb{R}^{g \times g^2} \).

Lemma 3.4. If \( u \in \mathbb{R}^g \) and \( A \in \mathbb{R}^{g \times g} \), then

\[
(u^T \otimes A^T)^{sT} = u (\text{vec}A)^T.
\]

Proof. If \( A = [a_1 \cdots a_g] \) with \( a_j \in \mathbb{R}^g \) for \( j = 1, \ldots, g \), then the \( ij \) entry of \( u^T \otimes A^T \) is \( u_i a_j^T \). Therefore, the \( ij \) entry of \( (u^T \otimes A^T)^{sT} \) is \( u_i a_j^T \), which coincides with the \( ij \) entry of \( u(\text{vec}A)^T \).

Lemma 3.5. If \( u, v, w \) belong to \( \mathbb{R}^g \) and \( y \in \mathbb{R}^{g^2} \), then:

1. \( u(v^T \otimes w^T) = u \otimes (v^T \otimes w^T) \) and

\[
(u \otimes (v^T \otimes w^T))^{sT} = u^T \otimes (v \otimes w^T) = (v \otimes w^T)U^T = v(u^T \otimes w^T).
\]

2. \( (uy)^{sT} = (\text{mat}_g y)^T U^T. \)

Proof. This is a straightforward computation that is similar to the verification of the preceding two lemmas.

Lemma 3.6. The \( g \times g^2 \) matrix \( Z_{01} \) is block symmetric in the following sense:

\[
Z_{01}^{sT} = Z_{01}.
\]

Proof. This rests on the formula

\[
p_3 = \frac{1}{2}[x_1 \cdots x_g]Z_{01}[x_1 \cdots x_g]^T
\]

(see [1,16]) and the fact that \( p_3 = p_3^T \).
Remark 3.7. Lemma [3.6] is a special case of the more general fact that if $p$ is a symmetric nc polynomial of degree $d$ and $j \leq d - 2$ and if the $g \times g^{j+1}$ matrix $\mathcal{Z}_{0j}$ is written in block form as

$$\mathcal{Z}_{0j} = \begin{bmatrix} b_{11} & \cdots & b_{1g} \\ \vdots & \ddots & \vdots \\ b_{g1} & \cdots & b_{gg} \end{bmatrix} \text{ with } b_{st} \in \mathbb{R}^{1 \times g^j},$$

then $b_{st} = b_{ts}$.

3.3. A factoring lemma. The following lemma of Vladimir Berkovich [B] simplifies a number of calculations.

Lemma 3.8. Let $\varphi(x) = \sum_{j=1}^{g} a_j x_j$ and $\psi(x) = \sum_{j=1}^{g} b_j x_j$. Then the identity

$$\varphi(x)f_1(x) = f_2(x)\psi(x)$$

for nc polynomials $f_1(x) = f_1(x_1, \ldots, x_g)$ and $f_2(x) = f_2(x_1, \ldots, x_g)$ with $f_1(0) = f_2(0) = 0$ implies that there exist a polynomial $f_3(x) = f_3(x_1, \ldots, x_g)$ such that $f_1(x) = f_3(x)\psi(x)$ and $f_2(x) = \varphi(x)f_3(x)$.

Proof. Without loss of generality we may assume that $b_g \neq 0$ and introduce the nonsingular change of variables variables $y_1 = x_1, y_2 = x_2, \ldots, y_{g-1} = x_{g-1}$ and $y_g = \sum_{j=1}^{g} b_j x_j$. Then the given equality is of the form

$$(c_1 y_1 + \cdots + c_g y_g)f_1(y) = \tilde{f}_2(y)y_g$$

and hence, upon writing $\tilde{f}_1(y) = \sum_{i=1}^{k} m_i$ as a sum of monomials of degree at least one in the variables $y_1, \ldots, y_g$, obtain that each monomial $m_i$ is of the form $\tilde{m}_i = \tilde{m}_iy_g$, where $\tilde{m}_i$ is a monomial that is one degree lower than $m_i$. Therefore, $\tilde{f}_1(y) = \tilde{f}_3(y)y_g$. The identity $f_1(x) = f_3(x)\psi(x)$ follows upon rewriting the last formula in terms of $x_1, \ldots, x_g$. The identity $f_2(x) = \varphi(x)f_4(x)$ for some nc polynomial $f_4(x)$ is established in much the same way. But this in turn implies that $\varphi f_3 \psi = \varphi f_4 \psi$ and hence that $f_4 = f_3$.

4. Classifying $\mathcal{Z}$ with $\mu_-(\mathcal{Z}) = 1$

In this section we study the structure of the symmetric nc polynomial $p$ and the scalar middle matrix $\mathcal{Z} = \mathcal{Z}(0)$ ($0 \in \mathbb{R}^{g}$) associated with its Hessian $p''(x)[h]$ when $\mu_-(\mathcal{Z}) = 1$. In view of (1.13),

$$\mu_-(\mathcal{Z}) \leq 1 \implies d \leq 4 \implies \mathcal{Z} \text{ is of the form (1.17).}$$

Moreover,

$$\text{rank } \mathcal{Z}_{02} \leq 1,$$
thanks to the inequality
\[
\mu_{\pm}(\mathcal{Z}) \geq \mu_{\pm}(\mathcal{Z}_{11}) + \text{rank } \mathcal{Z}_{02};
\]
see e.g., [DHM07a] for a proof of the latter.

Thus, if \( \mu_{-}(\mathcal{Z}) \leq 1 \), then
\[
p = p_0 + p_1 + p_2 + p_3 + p_4,
\]
where \( p_j \) is either equal to a homogeneous polynomial of degree \( j \) or equal to zero. Moreover, the scalar middle matrices of \( p_4''(x)[h] \), \( p_3''(x)[h] \) and \( p_2''(x)[h] \) are
\[
\begin{bmatrix}
0 & 0 & \mathcal{Z}_{02} \\
0 & \mathcal{Z}_{11} & 0 \\
\mathcal{Z}_{20} & 0 & 0
\end{bmatrix},
\begin{bmatrix}
0 & \mathcal{Z}_{01} \\
\mathcal{Z}_{10} & 0
\end{bmatrix}
\]
respectively, and each of the polynomials \( p_4 \), \( p_3 \), \( p_2 \) can be recovered from the scalar middle matrix of its Hessian by formula (1.16).

If \( \mu_{-}(\mathcal{Z}) \leq 1 \) and \( p \) is not the zero polynomial, then there are four mutually exclusive possibilities:

(1) \( \mu_{-}(\mathcal{Z}) = 1 \) and \( \text{rank } \mathcal{Z}_{02} = 1 \).

(2) \( \mu_{-}(\mathcal{Z}) = 1 \), \( \mathcal{Z}_{02} = 0 \) and \( \text{rank } \mathcal{Z}_{01} = 1 \).

(3) \( \mu_{-}(\mathcal{Z}) = 1 \), \( \mathcal{Z}_{02} = 0 \), \( \mathcal{Z}_{01} = 0 \) and \( \mu_{-}(\mathcal{Z}_{00}) = 1 \).

(4) \( \mu_{-}(\mathcal{Z}) = 0 \), \( \mathcal{Z}_{02} = 0 \), \( \mathcal{Z}_{01} = 0 \) and \( \mu_{-}(\mathcal{Z}_{00}) = 0 \).

4.1. The degree four case. In this subsection we shall assume that the rank of \( \mathcal{Z}_{02} \) is one and hence that \( p \) has degree four and
\[
\mathcal{Z}_{02} = uw^T \quad \text{with } u \in \mathbb{R}^g, \ w \in \mathbb{R}^{g^3}, \ ||u|| = 1 \text{ and } w \neq 0.
\]
Then \( P = I_g - uu^T \) is the orthogonal projection of \( \mathbb{R}^g \) onto the orthogonal complement of the vector \( u \) in \( \mathbb{R}^g \).

To prove Theorems [1.10] and [1.3] it is convenient to first establish a number of lemmas.

\textbf{Lemma 4.1.} If \( \mathcal{Z} \) is of the form (1.17) and if (4.1) holds and
\[
E := \begin{bmatrix}
P\mathcal{Z}_{00}P & P\mathcal{Z}_{01} \\
P\mathcal{Z}_{10}P & P\mathcal{Z}_{11}
\end{bmatrix},
\]
then:

(1) \( \mu_{\pm}(\mathcal{Z}) = \mu_{\pm}(E) + 1 \).

(2) \( \mu_{-}(\mathcal{Z}) = 1 \iff E \succeq 0 \).
Proof. The identity
\[
G = \begin{bmatrix}
PZ_{00}P & PZ_{01} & uw^T \\
Z_{10}P & Z_{11} & 0 \\
wu^T & 0 & 0
\end{bmatrix} = K^T Z K
\]
with
\[
K = \begin{bmatrix}
I_g & 0 & 0 \\
0 & I_g^2 & 0 \\
-K_0 & -K_1 & I_g^3
\end{bmatrix},
\]
implies that \( Z \) is congruent to \( G \) and hence that \( \mu_{\pm}(Z) = \mu_{\pm}(G) \).
Moreover, since \( G = G_1 + G_2 \) with
\[
G_1 = \begin{bmatrix}
PZ_{00}P & PZ_{01} & 0 \\
Z_{10}P & Z_{11} & 0 \\
0 & 0 & 0
\end{bmatrix} \quad \text{and} \quad G_2 = \begin{bmatrix}
0 & 0 & uw^T \\
0 & 0 & 0 \\
wu^T & 0 & 0
\end{bmatrix}
\]
and the ranges of these two real symmetric matrices are orthogonal, it is readily checked that
\[
\mu_{\pm}(G) = \mu_{\pm}(G_1) + \mu_{\pm}(G_2) = \mu_{\pm}(G_1) + 1 = \mu_{\pm}(E) + 1,
\]
which justifies (1). Thus, as (2) is immediate from (1), the proof is complete. \( \square \)

The following lemma appears in [DHM07a]. The statement and proof are repeated here for the convenience of the reader.

Lemma 4.2. If \( Z \) is of the form (1.17) and (4.1) holds, then there exists a vector \( w_1 \in \mathbb{R}^{g^2} \) so that
\[
w^T = u^T \otimes w_1^T.
\]
Moreover, the homogeneous of degree four part of \( p \) can be written as
\[
p_4(x) = \varphi(x)f_0(x)\varphi(x)
\]
where \( \varphi(x) = [x_1, \ldots, x_g]u \) and \( f_0 \) is homogeneous of degree two in \( x \).

Proof. Recall that
\[
p_4 = \frac{1}{2}[x_1 \cdots x_g]Z_{02}([x_1 \cdots x_g]_3)^T
\]
and hence that
\[
p_4 = \frac{1}{2}[x_1 \cdots x_g]uw^T([x_1 \cdots x_g]_3)^T.
\]
Thus
\[ p_4 = \varphi f \text{ with } \varphi = [x_1 \cdots x_g]u \text{ and } f = \frac{1}{2} w^T ([x_1 \cdots x_g]_3)^T. \]

Since \( p_4 = p_4^T \), it follows that \( \varphi f = f^T \varphi \), and hence Lemma 3.8 implies that
\[ f = f_0 \varphi = \frac{1}{2} [x_1 \cdots x_g]_2 w_0 u^T ([x_1 \cdots x_g])^T, \]
for some degree 2 homogeneous polynomial
\[ f_0 = \frac{1}{2} [x_1 \cdots x_g]_2 w_0. \]

By identity (9) in Theorem 3.1 this can be written as
\[ f = \frac{1}{2} (u^T \otimes (\Pi_1^T w_0)^T) ([x_1 \cdots x_g]_3)^T, \]
where \( \Pi_i \) is the permutation that is defined in formula (2.6). Comparing the two formulas for \( f \), yields the result with \( w_1 = \Pi_1^T w_0 \).

Recall that
\[ P = I_g - uu^T \quad \text{and} \quad U = u \otimes I_g. \]

**Lemma 4.3.** If \( Z \) is of the form (1.17) and (4.1) holds, then:

1. \( Z_{02} = u(u^T \otimes w_1^T) \) for some vector \( w_1 \in \mathbb{R}^{g^2} \).
2. \( Z_{11} = UAU^T \), where \( A = \text{mat}_g(w_1) = A^T \).
3. \( Z_{11} = Z_{11} UU^T = UU^T Z_{11} \).

Moreover, if \( \mu_-(Z) = 1 \), then
\[ PZ_{01} = PZ_{01} UU^T \]
and there exist vectors \( y \in \mathbb{R}^g \) and \( v \in \mathbb{R}^{g^2} \) such that:

4. The entry \( Z_{01} \) in formula (4.10) can be expressed as
\[ Z_{01} = u(u^T \otimes y^T) + uv^T + (uv^T)^T = Z_{01}^T \]
\[ = u^T \otimes B^T + u(\text{vec}B)^T = B^T U^T + u(\text{vec}B)^T, \]
where
\[ B = \text{mat}_g \left( v + \frac{1}{2} u \otimes y \right) \quad \text{and} \quad (\text{mat}_g v)u = 0. \]

5. \( PB^T = (\text{mat}_g v)^T \).
If \( E \) and \( E_1 \) denote the matrices in (4.2) and (iv) of Theorem 1.10, respectively, and \( B \) is as in (4), then
\[
E_1 = \begin{bmatrix}
P Z_{00} P & P B^T \\
B P & U^T Z_{11} U
\end{bmatrix} = \begin{bmatrix}
P Z_{00} P & (\text{mat}_g v)^T \\
(mat_g v) & U^T Z_{11} U
\end{bmatrix}
\]
and
\[
E \succeq 0 \iff E_1 \succeq 0.
\]

Proof. Item (1) was established in Lemma 4.2. Item (2) rests on the interplay between the formulas
\[
p_4(x) = \frac{1}{2} [x_1 \cdots x_g] Z_{02}([x_1 \cdots x_g]_3)\]
and
\[
p_4(x) = \frac{1}{2} [x_1 \cdots x_g]_2 Z_{11}([x_1 \cdots x_g]_2)\]
In view of (1),
\[
p_4(x) = \frac{1}{2} \varphi(x)(u^T \otimes w_1)([x_1 \cdots x_g]_3)^T
\]
and
\[
\frac{1}{2} \varphi(x) (([x_1 \cdots x_g]_3)(u \otimes w_1))^T
\]
But, if the vector \( w_1 \in \mathbb{R}^{g^2} \) is expressed as
\[
w_1 = \text{vec} A = \text{vec} \begin{bmatrix} a_1 & \cdots & a_g \end{bmatrix} \quad \text{with} \quad a_i \in \mathbb{R}^g,
\]
then
\[
[x_1 \cdots x_g]_2 w_1 = [x_1 \cdots x_g] A^T ([x_1 \cdots x_g])^T
\]
and thus,
\[
[x_1 \cdots x_g] u [x_1 \cdots x_g] = [x_1 \cdots x_g]_2 U
\]
(by (1) of Theorem 3.1 with \( A = I_g \)), the last formula for \( p_4(x) \) can be rewritten as
\[
p_4(x) = \frac{1}{2} \varphi(x) [x_1 \cdots x_g] A \begin{bmatrix} x_1 \\
\vdots \\
x_g
\end{bmatrix}
\]
and
\[
= \frac{1}{2} [x_1 \cdots x_g]_2 U A U^T ([x_1 \cdots x_g]_2)^T,
\]
which, upon comparison with the formula for \( p_4 \) in terms of \( Z_{11} \), serves to justify (2).
Next, (3) is immediate from (2), since \( U^T U = I_g \).
The rest of the proof is carried out under the added assumption that \( \mu_-(Z) = 1 \). Then Lemma 4.3 implies that \( E \succeq 0 \) and hence, by a well
known argument (see e.g., Lemma 12.19 in [D07]), that $PZ_{01} = KZ_{11}$ for some $K \in \mathbb{R}^{g \times g^2}$. Therefore,
\[
PZ_{01}UU^T = KZ_{11}UU^T = KZ_{11} = PZ_{01},
\]
and, as $UU^T = uu^T \otimes I_g$,
\[
(4.8) \quad Z_{01} = PZ_{01}UU^T + uu^T Z_{01}(P \otimes I_g) + uu^T Z_{01}(uu^T \otimes I_g) = \alpha + \beta + \gamma.
\]

The next step is to analyze the three terms
\[
\alpha = PZ_{01}UU^T, \quad \beta = uu^T Z_{01}(P \otimes I_g) \quad \text{and} \quad \gamma = uu^T Z_{01}(uu^T \otimes I_g)
\]
in (4.8).

Let $y^T = u^T Z_{01}(u \otimes I_g)$. Then
\[
\gamma = uy^TU^T = u(u^T \otimes y^T),
\]
which has $ij$ block entry
\[
u_{ij}y^T \in \mathbb{R}^{1 \times g^2}.
\]
Therefore, $\gamma^{st} = \gamma$, since $u_iu_jy^T = u_ju_iy^T$.

Next, in view of Lemma 3.3, the structured transpose $\alpha^{st}$ of $\alpha$ is
\[
\alpha^{st} = (Z_{01}UU^T)^{st} - (uu^T Z_{01}UU^T)^{st} = uu^T Z_{01}^{st} - uu^T(uu^T Z_{01})^{st} = uu^T Z_{01}^{st}UU^T = uu^T Z_{01}(I_g^2 - UU^T) = uu^T Z_{01}(P \otimes I_g) = \beta,
\]
since $Z_{01}^{st} = Z_{01}$ by Lemma 3.3.

The first advertised form of $Z_{01}$ in (4) is obtained by setting $v^T = u^T Z_{01}(P \otimes I_g)$, because then $\beta = uv^T$ and $\alpha = \beta^{st}$. Moreover, since $v = (P \otimes I_g)w$ with $w = Z_{10}u$, the identity
\[
\text{mat}_g((P \otimes I_g)w) = (\text{mat}_g w)P^T
\]
implies that $(\text{mat}_g v)u = 0$, since $P^Tu = Pu = 0$. The second formula for $Z_{01}$ in (4) follows from the first formula and Lemma 3.3.

Next, since $(\text{mat}_g(u \otimes y))^T = uy^T$, it follows immediately from the last formula for $B$ in (4) that $PB^T = P(\text{mat}_g v)^T$. But this yields (5), since $(\text{mat}_g v)u = 0$.

Finally, since $Z_{01} = B^T U^T + u(\text{vec} B)^T$ by (4), and $Pu = 0$,
\[
E = \begin{bmatrix}
PZ_{00}P & PZ_{01} \\
Z_{10}P & Z_{11}
\end{bmatrix} = \begin{bmatrix}
PZ_{00}P & PB^TU^T \\
UBP & Z_{11}
\end{bmatrix}
\]
and hence, as $Z_{11} = UU^T Z_{11} UU^T$,
\[
E = \begin{bmatrix}
I_g & 0 \\
0 & U
\end{bmatrix} \begin{bmatrix}
PZ_{00}P & PB^TU^T \\
UBP & Z_{11}
\end{bmatrix} \begin{bmatrix}
I_g & 0 \\
0 & U^T
\end{bmatrix} = \begin{bmatrix}
I_g & 0 \\
0 & U
\end{bmatrix} E_1 \begin{bmatrix}
I_g & 0 \\
0 & U^T
\end{bmatrix}.
\]
Thus, $E_1 \succeq 0 \implies E \succeq 0$. On the other hand, the formula

$$E_1 = \begin{bmatrix} I_g & 0 \\ 0 & U^T \end{bmatrix} E \begin{bmatrix} I_g & 0 \\ 0 & U \end{bmatrix}$$

is also valid, since $U^T U = I_g$. Therefore, $E \succeq 0 \implies E_1 \succeq 0$ and consequently (6) holds.

**Lemma 4.4.** An nc polynomial $p$ is of the form $p = \varphi f \varphi$, where

(i) $\varphi = \sum u_j x_j$ and $u$ is a unit vector with entries $u_1, \ldots, u_g$;
(ii) $f = [x_1 \cdots x_g] A [x_1 \cdots x_g]^T$ with $A = A^T \in \mathbb{R}^{g \times g}$

if and only if the scalar middle matrix $Z$ of $p''$ is

$$Z = 2 \begin{bmatrix} 0 & 0 & u(u^T \otimes w_A) \\ 0 & UAU^T & 0 \\ (u \otimes w_A)u^T & 0 & 0 \end{bmatrix},$$

where $w_A = \text{vec}(A)$ and $U = u \otimes I_g$.

**Proof.** Suppose first that $p = \varphi f \varphi$ where $\varphi$ and $f$ satisfy conditions (i) and (ii), respectively. Then

$$p = [x_1 \cdots x_g] u [x_1 \cdots x_g] A ([x_1 \cdots x_g])^T u^T ([x_1 \cdots x_g])^T$$

$$= [x_1 \cdots x_g]_2 UAU^T ([x_1 \cdots x_g]_2)^T,$$

where the first equality comes from the hypotheses and the second from identity (1) (with $A = I_g$) in Theorem 3.1. Hence $Z_{11} = 2 UAU^T$.

Similarly, the formulas

$$\varphi(x) \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} A \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix} = \varphi(x) \begin{bmatrix} x_1 \\ \vdots \\ x_g \end{bmatrix}_2 w_A$$

$$= ([x_1 \cdots x_g]_3) (u \otimes w_A)$$

with $w_A = \text{vec}(A)$ imply that

$$p = ([x_1 \cdots x_g]_3) (u \otimes w_A) u^T ([x_1 \cdots x_g])^T,$$

and hence that $Z_{20} = 2 (u \otimes w_A) u^T$.

Conversely, if $Z$ is of the form (4.9), then the formulas

$$p = \frac{1}{2} [x_1, \ldots, x_g]_2 Z_{11} ([x_1, \ldots, x_g]_2)^T$$

and (1) of Theorem 3.1 (with $A = I_g$) lead easily to the conclusion that $p = \varphi f \varphi$, as claimed.

**Lemma 4.5.** An nc polynomial $p$ is of the form $p = \varphi q + q^T \varphi$, where

(i) $\varphi = \sum u_j x_j$ and $u$ is a unit vector with entries $u_1, \ldots, u_g$; and
if and only if the scalar middle matrix $Z$ of $p''$ is

$$Z = 2\begin{bmatrix} 0 & u^T \otimes C^T + u(\text{vec } C)^T \\ u \otimes C + (\text{vec } C)u^T & 0 \end{bmatrix}.$$  

Proof. If $p = \varphi q + q^T \varphi$ with $\varphi$ and $q$ as described in the first part of the lemma, then

$$p'' = 2\varphi'q' + \varphi q'' + 2(q')^T \varphi' + (q'')^T \varphi'$$

and the indicated form of $Z$ is easily obtained by direct computation with the aid of the identities in Theorem 3.1. Conversely, if $Z$ is of the form (4.10), then direct computation based on the formula

$$p = \frac{1}{2}[x_1, \ldots, x_g]Z_{01}([x_1, \ldots, x_g]_2)^T$$

serves to recover $p$. \hfill $\Box$

4.2. The case of degree three. In this subsection we assume that $Z_{02} = 0$ and that $Z_{01}$ is rank one.

Lemma 4.6. If $Z_{i+j} = 0$ in (1.14) for $i + j = 2$ and $Z_{01} = uw_1^T$ for some pair of vectors $u \in \mathbb{R}^g$ and $w_1 \in \mathbb{R}^{g^2}$ with $\|u\| = 1$ and $w_1 \neq 0$ and $P = I_g - uu^T$, then:

$$\mu_- (Z) = 1 \iff PZ_{00}P \succeq 0.$$  

Proof. The formula

$$\begin{bmatrix} Z_{00} & Z_{01} \\ Z_{10} & 0 \end{bmatrix} = \begin{bmatrix} I_g & K^T \\ 0 & I_{g^2} \end{bmatrix} \begin{bmatrix} PZ_{00}P & uw_1^T \\ w_1u^T & 0 \end{bmatrix} \begin{bmatrix} I_g & 0 \\ K & I_{g^2} \end{bmatrix},$$

with

$$K = w_1(w_1^Tw_1)^{-1}u^T Z_{00} \left\{ I_g - \frac{uw_1^T}{2} \right\}$$

implies that

$$\mu_- (Z) = \mu_- (PZ_{00}P) + \text{rank} (uw_1^T) = \mu_- (PZ_{00}P) + 1,$$

which justifies the claim. \hfill $\Box$

Lemma 4.7. An nc polynomial $q$ of degree three is of the form $q = \varphi \psi \varphi$, where

(i) $\varphi = [x_1 \cdots x_g]_u, u \in \mathbb{R}^g,$

(ii) $\psi = [x_1 \cdots x_g]_y$ and $y \in \mathbb{R}^g,$
if and only if the scalar middle matrix \( Z \) of \( q'' \) is of the form

\[
Z = 2 \begin{bmatrix} 0 & u(u^T \otimes y^T) \\ (u \otimes y)u^T & 0 \end{bmatrix}.
\]

Proof. If \( q = \varphi \psi \varphi \), with \( \varphi \) and \( \psi \) as in (i) and (ii), then

\[
q'' = 2 \varphi' \psi' \varphi + 2 \varphi' \psi \varphi' + 2 \varphi \psi' \varphi' .
\]

However, the only term that contributes to \( Z_{01} \) is

\[
2 \varphi' \psi' \varphi = 2 [h_1 \cdots h_g] u (\varphi \psi')^T.
\]

Therefore, since

\[
\varphi \psi' = [x_1 \cdots x_g] u [h_1 \cdots h_g] y
\]

it is readily seen that

\[
\psi' \varphi = (u^T \otimes y^T) ([x_1 \cdots x_g] \otimes [h_1 \cdots h_g])^T
\]

and hence that \( Z_{01} = 2 u (u^T \otimes y^T) \). Conversely, if \( Z_{01} = 2 u (u^T \otimes y^T) \), then the formula

\[
q = \frac{1}{2} [x_1 \cdots x_g] Z_{01} ([x_1 \cdots x_g]_2)^T
\]

implies that \( q = \varphi \psi \varphi \), as advertised. \( \square \)

**Lemma 4.8.** If \( Z_{02} = 0 \) and \( Z_{01} = u_1 w_1^T \) with \( u_1^T u_1 = 1 \), \( w_1 \in \mathbb{R}^g \) and \( w_1 \neq 0 \), then \( w_1 = u_1 \otimes y_1 \) for some nonzero vector \( y_1 \in \mathbb{R}^g \),

(4.11) \[ p_4(x) = 0 \quad \text{and} \quad p_3(x) = \varphi_1(x) f_1(x) \varphi_1(x) , \]

where

\[
f_1(x) = \frac{1}{2} [x_1 \cdots x_g] y_1 \quad \text{and} \quad \varphi_1(x) = [x_1 \cdots x_g] u_1.
\]

Proof. The formulas \( Z_{01} = u_1 w_1^T \) and (1.16) imply that

\[
p_3(x) = \frac{1}{2} [x_1 \cdots x_g] Z_{01} ([x_1 \cdots x_g]_2)^T = \varphi_1(x) f(x)^T ,
\]

where

\[
f(x) = \frac{1}{2} ([x_1 \cdots x_g]_2) w_1.
\]

The rest follows from the fact that \( p_3 = p_3^T \) and Lemma 3.8. \( \square \)
4.3. Proof of Theorem 1.10. If $\mu_-(Z) = 1$ and $\text{rank}Z_{02} = 1$, then Lemmas 4.1, 4.3 guarantee that $Z$ is of the form specified in part I of Theorem 1.10 and that $E_1 \succeq 0$. If $\mu_-(Z) = 1$, $Z_{02} = 0$ and $\text{rank}Z_{01} = 1$, then Lemma 4.6 guarantees that $Z$ is still of the form specified in part I of Theorem 1.10 but with $A = 0$ and $v = 0$ and that $PZ_{00}P \succeq 0$. Therefore,

$$E_1 = \begin{bmatrix} PZ_{00}P & 0 \\ 0 & 0 \end{bmatrix} \succeq 0.$$ 

If $\mu_-(Z) = 1$, $Z_{02} = 0$ and $Z_{01} = 0$, then there exists a unit vector $u \in \mathbb{R}^g$ such that $Z_{00}u = \lambda u$ with $\lambda < 0$ and, if $P = I_g - uu^T$ for this choice of $u$, then $PZ_{00}P \succeq 0$. If $\mu_-(Z) = 0$, then $Z_{02} = 0$, $Z_{01} = 0$ and $Z_{00} \succeq 0$.

Conversely, if $Z$ is of the form specified in Theorem 1.10 and $E_1 \succeq 0$, then, in view of Lemma 4.1 and (6) of Lemma 4.3, $\mu_-(Z) = 1$ if $w_A \neq 0$. If $w_A = 0$, then $A = 0$ and the constraint $E_1 \succeq 0$ implies that mat$_g v = 0$ (see, e.g., Lemma 12.19 in [D07]) and hence that $v = 0$. Thus, $Z_{01} = u(u^T \otimes y^T)$. If $y \neq 0$, then rank$Z_{01} = 1$ and $\mu_-(Z) = 1$. If $y = 0$, then $Z_{01} = 0$ and $\mu_-(Z) = \mu_-(Z) \leq 1$, since $PZ_{00}P \succeq 0$.

4.4. Proof of Theorem 1.3. Since $\mu_-(Z) = \sigma_{\min}(p'')$, the assumption $\sigma_{\min}(p'') \leq 1$ guarantees that the scalar middle matrix of the Hessian $p''(x)[h]$ of $p$ has the form indicated in Theorem 1.10 and that $E_1 \succeq 0$. Consequently, the homogeneous components of $p$ of degree $j$ with $j \geq 2$ may be computed from the entries in $Z$ and formula (1.16). Moreover, by (1) of Theorem 3.1,

$$p_4(x) = \frac{1}{2} [x_1 \cdots x_g]^T Z_{11}([x_1 \cdots x_g]^T)$$

$$= \frac{1}{2} [x_1 \cdots x_g]^T UAU^T([x_1 \cdots x_g]^T)$$

$$= \frac{1}{2} \varphi(x)[x_1 \cdots x_g] A([x_1 \cdots x_g])^T \varphi(x),$$

which implies that $A = 2Q(f_0)$. Similarly the formula

$$p_3(x) = \frac{1}{2} [x_1 \cdots x_g]^T Z_{01}([x_1 \cdots x_g]^T)$$

together with the formulas for $Z_{01}$ in (4) of Lemma 4.3 imply that

$$\varphi(x)q(x) = \frac{1}{2} [x_1 \cdots x_g]u(\text{vec}B)^T([x_1 \cdots x_g]^T).$$
and hence (with the help of the transpose of the second formula in (11) of Theorem 3.1) that
\[ q(x) = \frac{1}{2} (\text{vec} B)^T ([x_1 \cdots x_g]_2)^T = \frac{1}{2} [x_1 \cdots x_g] B ([x_1 \cdots x_g])^T. \]

Therefore, \( B = 2Q(q) \). Similarly, the formula
\[ p_2 = \frac{1}{2} [x_1 \cdots x_g] Z_{00} ([x_1 \cdots x_g])^T \]
implies that \( Z_{00} = 2Q(p_2) \). Thus, in view of (6) of Lemma 4.3, the matrices \( E_2 \) in Theorem 1.3 and \( E_1 \) in Theorem 1.10 are simply related:
\[ E_1 = 2E_2. \]

If \( p \) is of degree three, then \( v = 0 \) by Theorem 1.10 and hence the formula for \( B \) in (4) of Lemma 4.3 reduces to \( 2B = \text{mat}_g(u \otimes y) = yu^T \).

Therefore,
\[ q(x) = \frac{1}{4} [x_1 \cdots x_g] yu^T ([x_1 \cdots x_g])^T = f_1(x) \varphi(x) \]

Conversely, if \( p \) is the form specified in Theorem 1.3, then Lemmas 4.4, 4.5 and 4.7 guarantee that \( Z \) is of the form specified in Theorem 1.10 and hence that \( \mu_-(Z) = \sigma_{\text{min}}(p') \leq 1 \).

4.5. **Proof of Theorem 1.11**. Theorem 1.11 follows from the formulas
\[ p_4(x) = \frac{1}{2} [x_1 \cdots x_g]_2 Z_{11} ([x_1 \cdots x_g]_2)^T, \quad p_3(x) = \frac{1}{2} [x_1 \cdots x_g] Z_{01} ([x_1 \cdots x_g]_2)^T \]
and
\[ p_2(x) = \frac{1}{2} [x_1 \cdots x_g]_2 Z_{00} ([x_1 \cdots x_g])^T, \]
and appropriate choices of the identities in Theorem 3.1.

5. **Modified Hessians**

In this section, we introduce the modified Hessian
\[ p''_\lambda = p''_\lambda(x, h) = p''(x)[h] + \lambda \{p'(x)[h]\}^T \{p'(x)[h]\} \]
of a symmetric polynomial \( p(x) = p(x_1, \ldots, x_g) \) in \( g \) noncommuting variables. The first order of business is to establish a representation formula analogous to formula (1.12) for the new term \( \{p'(x)[h]\}^T \{p'(x)[h]\} \).

**Lemma 5.1.** Let \( p_k(x) \) be a symmetric nc polynomial that is homogeneous of degree \( k \) in the \( g \) symmetric variables \( x_1, \ldots, x_g \) and suppose that \( k \geq 1 \). Then \( p'_k(x)[h] \) can be expressed uniquely in the form
\[ p'_k(x)[h] = \sum_{j=0}^{k-1} \varphi_{kj}(x) V_j(x)[h], \]
where \( \varphi_{kj}(x) \) is a row polynomial of size \( 1 \times g^{j+1} \) in which the nonzero entries are homogeneous polynomials of degree \( k - 1 - j \).
Proof. The polynomial $p_k(x)$ can be expressed in the form

$$p_k(x) = u_k^T \begin{bmatrix} x_1 \\ \vdots \\ x_{g-k} \end{bmatrix}$$

for some vector $u_k \in \mathbb{R}^{g^k}$. Therefore,

$$p'_k(x)[h] = u_k^T \left\{ \begin{bmatrix} h_1 \\ \vdots \\ h_g \\ x_{g-k-1} \\ \vdots \\ x_{g-k} \end{bmatrix} \otimes \begin{bmatrix} x_1 \\ \vdots \\ x_{g-k-1} \\ \vdots \\ x_{g-k} \end{bmatrix} + \sum_{i=1}^{k-1} \begin{bmatrix} x_1 \\ \vdots \\ x_{g-i-1} \\ \vdots \\ x_{g-k} \end{bmatrix} \otimes \Pi_{k-1-i} V_{k-1-i} \right\},$$

where the $\Pi_j$ denote the permutations defined by formula (2.5) for $j = 1, \ldots, k-1$ and $\Pi_0 = I_g$. But the last formula for $p'_k(x)[h]$ can be rewritten in the form (5.1) by noting that

$$u_k^T \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_{g-k} \end{bmatrix} \otimes \Pi_{k-1-i} V_{k-1-i} \right\} = \begin{bmatrix} x_1, \ldots, x_g \end{bmatrix}_i A_i \Pi_{k-1-i} V_{k-1-i}$$

for a suitably defined matrix $A_i \in \mathbb{R}^{g \times g^{k-1}}$ and then setting

$$\varphi_{k-1-i}(x) = \begin{cases} u_k^T [x_1, \ldots, x_g]_i A_i \Pi_{k-1-i} & \text{for } i = 1, \ldots, k-1 \\ u_k^T \Pi_{k-1} & \text{for } i = 0. \end{cases}$$

Lemma 5.2. If $p(x)$ is a symmetric nc polynomial of degree $d$ in $g$ symmetric variables, then

$$\{ p'(x)[h] \}^T \{ p'(x)[h] \} = \begin{bmatrix} W_{00} & W_{01} & \cdots & W_{0k} \\ W_{10} & W_{11} & \cdots & W_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ W_{k0} & W_{k1} & \cdots & W_{kk} \end{bmatrix} \begin{bmatrix} V_0 \\ \vdots \\ V_k \end{bmatrix},$$

where:
(1) \( k = d - 1 \).
(2) The vectors \( V_j = V_j(x)[h] \) in formula (5.2) are given by formula (2.4) for \( j = 0, \ldots, d - 1 \).
(3) \( W_{ij} \) is a matrix of size \( g^{i+1} \times g^{j+1} \) and the entries in \( W_{ij} \) are polynomials in the noncommuting variables \( x_1, \ldots, x_g \) of degree \( \leq 2(d - 1) - (i + j) \).
(4) \( W_{ij} = W_{ji} \).
(5) The matrix \( W = W(x) \) admits the factorization
\[
\begin{bmatrix}
W_{00} & \cdots & W_{0k} \\
\vdots & & \vdots \\
W_{k0} & \cdots & W_{kk}
\end{bmatrix}
= \begin{bmatrix}
\psi_0 \\
\vdots \\
\psi_k
\end{bmatrix}
\begin{bmatrix}
\psi_0^T \\
\vdots \\
\psi_k^T
\end{bmatrix}.
\]

Proof. Items (1)–(4) are straightforward; (5) is discussed next.
If \( p(x) \) is a polynomial of degree \( d \) in \( g \) noncommuting variables, then
\[
p(x) = \sum_{j=0}^d p_j(x),
\]
where \( p_j \) is either equal to a homogeneous polynomial of degree \( j \) or to zero and \( p_d \) is not zero. Therefore,
\[
p'(x)[h] = \sum_{j=1}^d p_j'(x)[h] = \sum_{s=0}^{d-1} \psi_s(x)^T V_s(x)[h],
\]
follows by applying Lemma 5.1 to each of the terms \( p_j(x) \) in the sum. Thus, \( p'(x) = p'(x)[h] \) can be expressed in terms of the border vectors \( V_s = V_s(x)[h] \) and a unique choice of vector polynomials \( \psi_s = \psi_s(x) \) of size \( g^{s+1} \times 1 \) and degree \( d-1-s \) by the indicated formula. Consequently, the entries \( W_{ij} \) in the representation formula (5.2) can now be written in terms of the vector polynomials \( \psi_0, \ldots, \psi_k \) as
\[
W_{ij}(x) = \psi_i(x)\psi_j^T(x) \quad \text{for} \quad i, j \leq d - 1
\]
and hence the full matrix
\[
W(x) = \begin{bmatrix}
\psi_0(x) \\
\vdots \\
\psi_k(x)
\end{bmatrix}
\begin{bmatrix}
\psi_0^T(x), \ldots, \psi_k^T(x)
\end{bmatrix}, \quad \text{where} \quad k = d - 1.
\]

\[\square\]

The middle matrix for the modified Hessian of a symmetric polynomial of degree \( d \) is the \((d-1) \times (d-1)\) block matrix \( Z_\lambda \) (with polynomial entries)
\[
Z_\lambda = \begin{bmatrix}
Z & 0 \\
0 & 0
\end{bmatrix} + \lambda W,
\]
where \( Z \) is the middle matrix for \( p'' \). Thus, as \( \tilde{V}(x)[h] \) denotes the border vector of height \( g\tilde{\nu} \) (which includes monomials in \( x \) of degree
The scalar middle matrix of the modified Hessian is the matrix
\begin{equation}
Z_{\lambda} = \begin{bmatrix}
Z & 0 \\
0 & \lambda \psi_{d-1}(0) \psi_{d-1}(0)^T
\end{bmatrix}
\end{equation}

Note that this differs a bit from the earlier terminology since \(Z_{\lambda}(0) \neq Z_{\lambda}\). The key point is that \(Z_{\lambda}\) is constant and is polynomially congruent to \(Z_{\lambda}(x)\) via a congruence which does not depend upon \(\lambda\).

**Theorem 5.3.** Let \(p(x)\) be a symmetric nc polynomial of degree \(d \geq 1\) in \(g\) symmetric variables and let \(\psi_j(x)\) denote the coefficients of \(p'(x)[h]\) in formula (5.3). Then
\[\psi_{d-1}(x) = \psi_{d-1}(0) \neq 0\]
and there exists a block \(d \times d\) matrix-valued polynomial \(S\) with polynomial inverse so that
\[Z_{\lambda}(x) = S(x)^T Z_{\lambda} S(x)\]

In particular, \(S(X)\) is invertible when \(X \in (\mathbb{R}_{sym}^{n \times n})^g\), and
\[Z_{\lambda}(X) = S(X)^T (Z_{\lambda} \otimes I_n) S(X)\]

**Proof.** It is convenient to let \(y = \psi_{d-1}(x)\) and \(\tilde{\psi}^T := [\psi_0^T, \ldots, \psi_{d-2}^T]\). Then, as \(y \in \mathbb{R}^{g^d}\),
\begin{equation}
Z_{\lambda}(x) = \begin{bmatrix}
Z(x) & 0 \\
0 & 0
\end{bmatrix} + \lambda \begin{bmatrix}
\tilde{\psi}^T \\
\tilde{\psi}^T
\end{bmatrix} \begin{bmatrix}
\psi_0^T & \psi_1^T \\
\psi_1^T & \psi_2^T
\end{bmatrix} = \begin{bmatrix}
Z(x) + \lambda \tilde{\psi}^T \tilde{\psi} y^T & \lambda \tilde{\psi} y^T \\
\lambda y \tilde{\psi} y^T & \lambda y y^T
\end{bmatrix}
\end{equation}
\[= \begin{bmatrix}
I & \lambda \tilde{\psi} y^T (\lambda y y^T)^T \\
0 & I
\end{bmatrix} \begin{bmatrix}
Z(x) & 0 \\
0 & \lambda y y^T
\end{bmatrix} \begin{bmatrix}
I & (\lambda y y^T)^T \\
(\lambda y y^T)^T & I
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix},
\]
since the Moore-Penrose inverse \((\lambda y y^T)^T\) of \(\lambda y y^T\) is given by the formula
\[(\lambda y y^T)^T = \frac{y (y^T y)^{-2} y^T}{\lambda}\]
when \(\lambda \neq 0\), and
\[y^T (\lambda y y^T)^T (\lambda y y^T)^T = y^T (\lambda y y^T) (\lambda y y^T)^T = y^T.\]
Thus, upon setting $C(x) = (yy^T)^T y\widetilde{\psi}(x)$ and invoking formula (1.13), it follows that

$$Z_\lambda(x) = S(x)^T Z_\lambda S(x),$$

where

$$(5.8) \quad S(x) = \begin{bmatrix} B(x) & 0 \\ C(x) & I \end{bmatrix} \quad \text{and} \quad S(x)^{-1} = \begin{bmatrix} B(x)^{-1} & 0 \\ -C(x)B(x)^{-1} & I \end{bmatrix}$$

are both polynomial matrices. It is important to note that $S(x)$ does not depend on $\lambda$.

**Corollary 5.4.** If $\lambda > 0$, then, in the setting of Theorem 5.3

1. $\mu_+(Z_\lambda) = \mu_+(Z) + 1$ and $\mu_-(Z_\lambda) = \mu_-(Z)$.
2. $Z_\lambda \succeq 0 \iff Z \succeq 0$.
3. $Z_\lambda \succeq 0 \implies d \leq 2$ (thanks to (1.13)).

6. The Relaxed Hessian: Local Positivity

In this section we exploit the structure of the polynomial congruence of Theorem 1.13 to prove Theorems 1.5 and 1.4. We then establish Theorems 1.7 and 1.8.

6.1. Proof of Theorem 1.5. The starting point is the polynomial congruence formula (5.8) of Theorem 5.3. Again, we emphasize that, whereas $Z = Z(0)$ for $0 \in \mathbb{R}^g$, it is not the case that $Z_\lambda = Z(0)$.

It is convenient to let $F = (S^T)^{-1} S^{-1}$, and to bear in mind that $Z_\lambda(X) = Z_\lambda \otimes I_n$ for $X \in (\mathbb{R}^{n \times n})^g$, since $Z_\lambda$ is constant.

If $X \in (\mathbb{R}^{n \times n})^g$ and $\delta > 0$, then

$$Z_\lambda(X) + \delta I = S(X)^T (Z \otimes I_n + \delta F(X))S(X).$$

Thus, for $H \in (\mathbb{R}^{n \times n})^g$ and $v \in \mathbb{R}^n$, we have

$$(6.1) \quad \left\langle \left( p''_\lambda(X)[H] + \delta \widetilde{V}(X)[H]^T \widetilde{V}(X)[H] \right) v, v \right\rangle$$

$$= \langle (Z_\lambda \otimes I_n + \delta F(X))S(X)\widetilde{V}(X)[H]v, S(X)\widetilde{V}(X)[H]v \rangle.$$

The hypothesis that $\{ \widetilde{V}(X)[H]v : H \in (\mathbb{R}^{n \times n})^g \}$ has codimension at most $n - 1$ in $\mathbb{R}^{ng^2}$ and the invertibility of $S(X)$ implies that $\{ S(X)\widetilde{V}(X)[H]v : H \in (\mathbb{R}^{n \times n})^g \}$ also has codimension at most $n - 1$. Consequently, as the relaxed Hessian is positive at $(X, v)$ under the hypothesis of Theorem 1.5, we can select $\lambda$ such that the left side of formula (6.1) is nonnegative for each $H \in (\mathbb{R}^{n \times n})^g$, in which case $Z_\lambda \otimes I_n + \delta F(X)$ has at most $n - 1$ negative eigenvalues. It follows that $Z \otimes I_n + \delta \widetilde{F}(X)$, the upper left hand (block) corner of $Z_\lambda \otimes I_n + \delta F(X)$, has at most $n - 1$ negative eigenvalues for each $\delta > 0$. Since this upper
left hand corner does not depend upon $\lambda$ and $\delta > 0$ is arbitrary, it follows that $Z \otimes I_n$ has at most $n - 1$ negative eigenvalues. Therefore, by (1.14), $\mu_- (Z) = 0$. Thus, $p$ has degree at most two by (1.15).

6.2. Proof of Theorem 1.4. To see that Theorem 1.4 follows from Theorem 1.5 it suffices to show that, under the hypothesis on $(X, v)$ in Theorem 1.4, the subspace $\{ \tilde{V}(X)[H] v : H \in (\mathbb{R}_{sym}^{n \times n})^g \}$ has codimension at most $n - 1$ in $\mathbb{R}^{ng^2}$. For this, and subsequent proofs, we will invoke an estimate that is extracted from Lemmas 9.5 and 9.7 in [CHSY03], which are rephrased below as Lemma 6.1 for the convenience of the reader.

**Lemma 6.1 (CHSY Lemma).** Given a pair of positive integers $g$ and $r$, a matrix $X \in (\mathbb{R}_{sym}^{n \times n})^g$ and a vector $v \in \mathbb{R}^n$, let

$$\alpha_k = \sum_{j=0}^{k} g^j, \quad \mathcal{R}_k = \left\{ \begin{bmatrix} V_0(X)[H] v \\ \vdots \\ V_k(X)[H] v \end{bmatrix} : H \in (\mathbb{R}_{sym}^{n \times n})^g \right\},$$

and suppose that the set $\{ m(X)v : |m| \leq d - 1 \}$ is a linearly independent subset of $\mathbb{R}^n$. Then $\mathcal{R}_k$ is a subspace of $\mathbb{R}^{ng\alpha_k}$ and its codimension

$$\text{codim} \mathcal{R}_k \leq ng(\alpha_k - \alpha_r) + ga_r \frac{\alpha_r - 1}{2} \quad \text{if} \quad k \geq r,$$

with equality if $k = r$. (Recall that $|m|$ denotes the length of the monomial $m$.)

A key fact is that if $k = r$, then the bound on the codimension is independent of $n$.

The bound (6.2) follows easily from the following sequence of estimates, the first of which is based on Lemmas 9.5 and 9.7 in [CHSY03].

1. The codimension of $\mathcal{R}_r$ in $\mathbb{R}^{ng\alpha_r}$ is equal to $ga_r(\alpha_r - 1)/2$.
2. $\dim \mathcal{R}_r = ng\alpha_r - ga_r(\alpha_r - 1)/2$.
3. $\dim \mathcal{R}_k \geq ng\alpha_r - ga_r(\alpha_r - 1)/2$ if $k \geq r$.

To prove Theorem 1.4, note that the set $\{ m(X)v : |m| \leq d - 1 \}$ is linearly dependent if and only if there exists a set of nonzero numbers $q_m \in \mathbb{R}$ for $|m| \leq d - 1$ such that

$$q(X)v = \sum q_m m(X)v = 0.$$

Thus, $\{ m(X)v : |m| \leq d - 1 \}$ is linearly independent. By the CHSY Lemma, with $r = d - 1$ so that $\alpha_r = \tilde{v}$, the codimension of the set $\mathcal{R}_{d-1} = \{ \tilde{V}(X)[H] v : H \in (\mathbb{R}_{sym}^{n \times n})^g \}$ is less than $n$ (thanks to assumption (1.10)). The proof is completed by an application of Theorem 1.5.
6.3. **Proof of Corollary 1.6.** We begin with a definition and a preliminary lemma.

If \( \mathfrak{B}_n \subset (\mathbb{R}^{n \times n}_{\text{sym}})^g \times \mathbb{R}^n \) for \( n = 1, 2, \ldots \), then the graded set

\[
\mathfrak{B} = \bigcup_{n \geq 1} \mathfrak{B}_n
\]

is **closed with respect to direct sums** if whenever \((X^j, v^j) \in \mathfrak{B}_{n_j}\) for \( j = 1, 2, \ldots, k \), it follows that \((X, v) \in \mathfrak{B}_n\) for \( n = \sum n_j \), where \( X = \text{diag}\{X^1, \ldots, X^k\} \) and \( v = \text{vec}[v^1 \cdots v^k] \).

**Lemma 6.2.** Let \( k \) be a given positive integer and suppose \( \mathfrak{B} \) is closed with respect to direct sums. If for each \((X, v) \in \mathfrak{B}\) the set \( \{m(X)v : |m| \leq k\} \) is linearly dependent, then there exists a nonzero polynomial

\[
q = \sum_{|m| \leq k} q_m m(x)
\]

(which is not necessarily symmetric) of degree \( \leq k \) such that \( q(X)v = 0 \) for every \((X, v) \in \mathfrak{B}\).

**Proof.** See Lemma 4.1 of [DHM07b] and, for a cleaner proof, §6 of [HMV06]. \( \square \)

To prove Corollary 1.6, note that there does not exist a nonzero polynomial \( q \) of degree at most \( k \) such that \( q(X)v = 0 \) for every \( X \in (\mathbb{R}^{n \times n}_{\text{sym}})^g \) with \( \|X\| < \epsilon \) when \( v \) is a nonzero vector in \( \mathbb{R}^n \). Thus, Lemma 6.2 applied to the noncommutative \( \varepsilon \) neighborhood of zero (i.e., to the graded set \( \mathcal{U} \) that is defined just above the statement of Corollary 1.6) guarantees that there exists a pair \((X, v) \in \mathcal{U}_n\) for some choice of \( n \) such that \( \{m(X)v : |m| \leq k\} \) is linearly independent in \( \mathbb{R}^n \). This much is true for every positive integer \( k \). Now fix \( k = d - 1 \). Then \( n \geq \alpha d - 1 = \tilde{\nu} \). However, by considering \( \oplus_1(X, v) \) it may be assumed that \( n > g \tilde{\nu}(\tilde{\nu} - 1)/2 \), and hence that Theorem 1.4 is applicable.

6.4. **Proofs of Theorems 1.7 and 1.8.** For the proof of Theorem 1.7, the now familiar arguments show that the hypotheses imply that \( \mu_-(Z(X)) \leq kn - 1 \). On the other hand, \( \mu_-(Z(X)) = n \mu_-(Z) \). Hence \( \mu_-(Z) = \sigma^{\text{min}}(p'') < k \).

Theorem 1.8 follows from Theorem 1.7 in much the same way that Theorem 1.5 follows from Theorem 1.4. The main point is that, from the CHSY-Lemma with \( r = d - 2 \) so that \( \alpha_r = \nu \), the subspace \( \mathcal{R}_{d-2} \) has codimension less than \( n \) (thanks to assumption (1.11)). Thus, restricting \( H \) to the space \( \mathcal{H} \) of codimension at most \( nk \), it follows that

\[
\text{codim}\{V(X)[H]v : H \in \mathcal{H}\} \leq \text{codim} \mathcal{R}_{d-2} + \text{codim} \mathcal{H} \leq n - 1 + nk.
\]
In this appendix we present some useful connections between the polynomials $\psi_j(x)$ that are defined by formula (5.3) and the entries in the middle matrix $Z_{ij}(x)$ for the Hessian $p''(x)[h]$ for an arbitrary symmetric nc polynomial of degree $d$ in $g$ symmetric variables.

Theorem A.1. If $p(x)$ is a symmetric nc polynomial of degree $d$ in $g$ symmetric variables, then the coefficients $\psi_s$, $s = 0, \ldots, d - 1$, in formula (5.3) are related to the entries $Z_{0s}$ in the representation formula (1.12) for $p''(x)$ by the formula

\[\psi_s(x)^T = \frac{1}{2}[x_1 \cdots x_g]Z_{0s}(x) + \psi_s(0)^T \quad \text{for} \quad s = 0, \ldots, d - 1.\]

Proof. In view of formulas (1.16) and (2.6)

\[p(x) = \frac{1}{2} \sum_{j=0}^\ell [x_1 \cdots x_g]Z_{0j}([x_1 \cdots x_{g,j+1}]^T + q(x)\right\]

\[= \frac{1}{2} \sum_{j=0}^\ell [x_1 \cdots x_g]Z_{0j}\Pi_j^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_g \\ j+1 \end{bmatrix} + q(x),\]

where $\ell = d - 2$ and degree $q(x) \leq 1$. Therefore, since $Z_{0j} = Z_{0j}(0)$ is independent of $x$,

\[p'(x)[h] = I + II + q'(x)[h],\]

where

\[I = \frac{1}{2} \sum_{j=0}^\ell [h_1 \cdots h_g]Z_{0j}\Pi_j^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_g \\ j+1 \end{bmatrix} = \sum_{j=1}^{\ell+1} u_j^TV_j(x)[h]\]

for some choice of vectors $u_j \in \mathbb{R}^{g'+1}$,

\[q'(x)[h] = [h_1 \cdots h_g]u_0 = u_0^TV_0(x)[h] \quad \text{for some vector} \quad u_0 \in \mathbb{R}^g\]

and

\[II = \frac{1}{2} \sum_{j=0}^\ell [x_1 \cdots x_g]Z_{0j}\Pi_j^{-1} \begin{bmatrix} [h_1] \\ \vdots \\ [x_1] \\ \vdots \\ [h_g] \\ \vdots \\ [x_g]_j \end{bmatrix} \otimes \begin{bmatrix} [x_1] \\ \vdots \\ [h_1] \\ \vdots \\ [x_g]_j \end{bmatrix} + \cdots + \begin{bmatrix} [h_1] \\ \vdots \\ [x_1] \\ \vdots \\ [h_g] \\ \vdots \\ [x_g]_j \end{bmatrix} \otimes \begin{bmatrix} [x_1] \\ \vdots \\ [h_1] \\ \vdots \\ [x_g]_j \end{bmatrix}\]

\[= \frac{1}{2}[x_1 \cdots x_g] \sum_{j=0}^\ell Z_{0j}\Pi_j^{-1} \sum_{i=0}^j \theta_{j,j-i}\Pi_iV_i,\]
where, with the help of (2.5), the factor $\theta_{ji}$ can be written as

$$\theta_{ji} = \left[ \begin{array}{c} x_1 \\ \vdots \\ x_g \end{array} \right] \otimes I_{g^{j+1-i}} \quad \text{for } i = 1, \ldots, j \quad \text{and} \quad \theta_{j0} = I_{g^{j+1}}.$$

Thus, the double sum

$$\ell \sum_{i=0}^{j} Z_{0j} \Pi_j^{-1} \sum_{i=0}^{\ell} \theta_{j,j-i} \Pi_i V_i = \ell \sum_{i=0}^{j} \left( \sum_{j=i}^{\ell} Z_{0j} \Pi_j^{-1} \theta_{j,j-i} \Pi_i \right) V_i.$$

But the inner sum $\sum_{j=i}^{\ell} Z_{0j} \Pi_j^{-1} \theta_{j,j-i} \Pi_i$ can be reexpressed in terms of the matrix polynomials

$$K_j(x) = \Pi_j^{-1} \left( \left[ \begin{array}{c} x_1 \\ \vdots \\ x_g \end{array} \right] \otimes I_{g^{j+1}} \right) \Pi_j$$

and their products

$$K_{j+1}K_j = \Pi_{j+2}^{-1} \left( \left[ \begin{array}{c} x_1 \\ \vdots \\ x_g \end{array} \right] \otimes I_{g^{j+1}} \right) \Pi_j, \quad \cdots$$

as

(A.2) $Z_{0i} + Z_{0,i+1}K_i + Z_{0,i+2}K_{i+1}K_i + \cdots + Z_{0\ell}K_{\ell-1} \cdots K_i = Z_{0i}(x)$;

and the last identity follows from the formula $Z(x)A(x) = Z$ in Theorem 7.3 of [DHM07a]. Thus,

$$p'(x)[h] = \frac{1}{2} [x_1 \cdots x_g] \sum_{i=0}^{\ell} Z_{0i}(x)V_i(x)[h],$$

which, upon comparison with formula (5.3), implies that

$$\psi_i(x)^T = u_i^T + \frac{1}{2} [x_1 \cdots x_g] Z_{0i}(x) \quad \text{for } i = 0, \ldots, \ell + 1,$$

since $Z_{0,\ell+1}(x) = 0$. Therefore, $\psi_i(0)^T = u_i^T$ and the proof is complete. \hfill \Box

**Corollary A.2.** If $\psi_j(0) = 0$ for $j = 0, \ldots, \ell$ in the setting of Theorem A.1, then

(A.3) $W_{ij} = Z_{i0}QZ_{0j}$ for $i, j = 0, \ldots, \ell$,\hfill
where \( \ell = d - 2 \) and

\[
(A.4) \quad Q = \frac{1}{4} \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} [x_1 \cdots x_d].
\]

**Theorem A.3.** Let \( p_k(x) = p(x_1, \ldots, x_g) \) be a homogeneous symmetric nc polynomial of degree \( k \geq 1 \) in \( g \) symmetric variables and let

\[
(A.5) \quad p'_k(x)[h] = \sum_{s=0}^{k-1} \psi_{ks}(x)^T V_s(x)[h].
\]

Then:

1. \( \psi_{ks}(x) \) is a homogeneous polynomial of degree \( k - 1 - s \) for \( s = 0, \ldots, k - 1 \).
2. \( \psi_{ks}(0) = 0 \) for \( s = 0, \ldots, k - 2 \) when \( k \geq 2 \).
3. Formulas (A.3) and (A.4) are in force for \( i, j = 0, \ldots, k - 2 \) when \( k \geq 2 \).

**Proof.** Assertion (1) is immediate from (A.5), since \( \psi_{ks}(x)^T V_s(x)[h] \) is a homogeneous polynomial of degree \( k - 1 \) in \( x \) and \( V_s \) is a homogeneous polynomial of degree \( s \) in \( x \). Assertions (2) and (3) then follow easily from (1) and the preceding corollary. \( \square \)

Theorem A.3 also yields conclusions for nonhomogeneous polynomials, subject to some restrictions.

**Corollary A.4.** Let \( p = p(x) = p(x_1, \ldots, x_g) \) be a symmetric nc polynomial of degree \( d \geq 2 \) in \( g \) symmetric variables such that there are no terms of degree one and no terms of degree \( d - 1 \) in \( p(x) \), let \( \ell = d - 2 \) and let \( \psi_{ks}(x)^T \) denote the row vector polynomials defined by formula (5.3). Then \( \psi_0(0) = 0 \) and \( \psi_\ell(0) = 0 \).

**Proof.** Let \( p(x) = \sum_{k=0}^{d} c_k p_k(x) \), where \( p_k(x) \) is a homogeneous polynomial of degree \( k \) and \( c_k = 0 \) or \( c_k = 1 \). Then, by formula (A.5),

\[
p'(x)[h] = \sum_{k=1}^{d} c_k p_k'(x)[h]
\]

\[
= \sum_{k=1}^{d} c_k \sum_{s=0}^{k-1} \psi_{ks}(x)^T V_s(x)[h]
\]

\[
= \sum_{s=1}^{d-1} \left\{ \sum_{k=s+1}^{d} c_k \psi_{ks}(x)^T \right\} V_s(x)[h].
\]
Thus, in formula (5.3),

\[
\psi_0(x)^T = 0 \quad \text{and} \quad \psi_s(x)^T = \sum_{k=s+1}^{d} c_k \psi_{ks}(x)^T \quad \text{for} \quad s = 1, \ldots, d - 1.
\]

In particular,

\[
\psi_\ell(x)^T = \sum_{k=\ell+1}^{d} c_k \psi_{k\ell}(x)^T = c_{\ell+1} \psi_{\ell+1,\ell}(x)^T + c_{\ell+2} \psi_{\ell+2,\ell}(x)^T.
\]

Consequently,

\[
\psi_\ell(0)^T = c_{\ell+1} \psi_{\ell+2,\ell}(0)^T = 0,
\]

since \( c_{\ell+1} = 0 \), by assumption and \( \psi_{\ell+2,\ell}(0)^T = 0 \) by Theorem A.3. \( \square \)

REFERENCES

[BGM06] Joseph A. Ball, Gilbert Groenewald and Tanit Malakorn, Conservative structured noncommutative multidimensional linear systems, in: The state space method generalizations and applications, Oper. Theory Adv. Appl., 161, Birkhäuser, Basel, 2006, pp. 179–223,

[B] Vladimir Berkovich, Private communication.

[CHSY03] Juan F. Camino, J. William Helton, Robert E. Skelton and Jieping Ye, Matrix inequalities: A Symbolic Procedure to Determine Convexity Automatically, Integral Equations Operator Theory, 46 (2003), no. 4, 399-454.

[D07] Harry Dym, Linear Algebra in Action, Amer. Math. Soc., Rhode Island, 2007.

[DHM07a] Harry Dym, J. William Helton and Scott A. McCullough, The Hessian of a noncommutative polynomial has numerous negative eigenvalues, J. d’Analyse, 102 (2007), 29–76.

[DHM07b] Harry Dym, J. William Helton and Scott A. McCullough, Irreducible noncommutative defining polynomials for convex sets have degree four or less, Indiana Univ. Math. J., 56 (2007), no. 3, 1189–1231.

[HM04] J. William Helton and Scott A. McCullough, Convex noncommutative polynomials have degree two or less, SIAM Journal of Matrix Analysis, 25(2004), no. 4, 1124-1139.

[HMS02] J. William Helton, Mauricio C. de Oliveira, Mark Stankus and Robert L. Miller, NCAAlgebra Package for Mathematica, 2006.

[HMV06] J. William Helton, Scott A. McCullough and Victor Vinnikov, Noncommutative convexity arises from linear matrix inequalities, J. Funct. Anal., 240 (2006), no. 1, 105–191.
[HMe98] J. William Helton and Orlando Merino, Sufficient conditions for the optimization of matrix functions, Proc. Conf. on Decision and Control, 1998, pp. 1–5.

[HP07] J. William Helton and Mihai Putinar, Positive Polynomials in Scalar and Matrix Variables, the Spectral Theorem and Optimization, in: Operator theory, structured matrices, and dilations, Theta Ser. Adv. Math., 7, Theta, Bucharest, 2007, pp. 229–306.

[KVV06] Dmitry S. Kalyuzhnyi-Verbovetski˘ı and Victor Vinnikov, Noncommutative positive kernels and their matrix evaluations, Proc. Amer. Math. Soc. 134 (2006), no. 3, 805–816.

[Po06] Gelu Popescu, Operator theory on noncommutative varieties. Indiana Univ. Math. J., 55 (2006), no. 2, 389–442.

[SV06] Dimitri Shlyakhtenko and Dan Voiculescu, Free analysis workshop summary: American Institute of Mathematics. [http://www.aimath.org/pastworkshops/freeanalysis.html](http://www.aimath.org/pastworkshops/freeanalysis.html) 2006.

[Vo05] Dan Voiculescu, Free probability and the von Neumann algebras of free groups, Rep. Math. Phys. 55 (2005), no.1, 127-133.

[Vo06] Dan Voiculescu, Symmetries arising from free probability theory, in: Frontiers in number theory, physics, and geometry. I (2006), 231-243 Springer Berlin.
# Contents

1. Introduction and Main Results
   1.1. The setting
   1.2. Previous results
   1.3. Main Results
   1.4. The middle matrix
   1.5. Motivation
2. Kronecker Product Notation
   2.1. Border vectors
   2.2. Tracking derivatives
3. Basic Identities
   3.1. Identities
   3.2. Structured transposes
   3.3. A factoring lemma
4. Classifying $Z$ with $\mu_- (Z) = 1$
   4.1. The degree four case
   4.2. The case of degree three
   4.3. Proof of Theorem 1.10
   4.4. Proof of Theorem 1.3
   4.5. Proof of Theorem 1.11
5. Modified Hessians
6. The Relaxed Hessian: Local Positivity
   6.1. Proof of Theorem 1.5
   6.2. Proof of Theorem 1.4
   6.3. Proof of Corollary 1.6
   6.4. Proofs of Theorems 1.7 and 1.8
7. Appendix A. A connection between $Z$ and $W$
8. References


Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
E-mail address: harry.dym@weizmann.ac.il

Mathematics Department, University of California at San Diego, La Jolla, CA 92093
E-mail address: j1greene@math.ucsd.edu

Mathematics Department, University of California at San Diego, La Jolla, CA 92093
E-mail address: helton@math.ucsd.edu

Department of Mathematics, University of Florida, Gainesville, FL 32611-8105
E-mail address: sam@math.ufl.edu